



ESPE

UNIVERSIDAD DE LAS FUERZAS ARMADAS
INNOVACIÓN PARA LA EXCELENCIA

DEPARTAMENTO DE CIENCIAS DE LA COMPUTACIÓN

CARRERA DE INGENIERÍA EN SISTEMAS

TESIS PREVIO A LA OBTENCIÓN DEL TÍTULO

DE INGENIERO EN SISTEMAS

AUTORES: ALVAREZ CALVOPIÑA CRISTIAN DARIO,

ESPARZA MONTES DIEGO ARMANDO

TEMA: ANÁLISIS, DISEÑO E IMPLEMENTACIÓN DE UN DATAMART

UTILIZANDO HERRAMIENTAS OPEN SOURCE PARA LA UNIDAD

ADMINISTRATIVA Y FINANCIERA DE LA ESPE.

DIRECTORA: ING. LORENA DUQUE

CODIRECTORA: ING. DORYS QUIROZ

SANGOLQUÍ, JULIO 2014

CERTIFICACIÓN

Certifico que el presente trabajo fue realizado en su totalidad por los señores Diego Armando Esparza Montes y Cristian Darío Alvarez Calvopiña como requerimiento parcial a la obtención del título de INGENIERO EN SISTEMAS E INFORMÁTICA.

Sangolquí, Julio 2014

Ing. Lorena Duque

DIRECTORA DE TESIS

Ing. Dorys Quiroz

CODIRECTORA DE TESIS

AUTORÍA DE RESPONSABILIDAD

Yo Cristian Darío Alvarez Calvopiña, yo Diego Armando Esparza Montes, declaramos que el trabajo aquí descrito es de nuestra autoría; que no ha sido presentado previamente para ningún grado o calificación profesional; y, que hemos consultado las referencias bibliográficas que se incluyen en el presente documento.

La Universidad de las Fuerzas Armadas ESPE, puede hacer uso de los derechos correspondientes a este trabajo, según lo establecido por Ley de Propiedad Intelectual, por su reglamento y por la normativa institucional vigente.

Sangolquí, Julio 2014

Cristian Darío Alvarez Calvopiña

Diego Armando Esparza Montes

AUTORIZACIÓN

Nosotros:

Cristian Darío Alvarez Calvopiña

Y

Diego Armando Esparza Montes

Autorizamos a la Universidad de las Fuerzas Armadas ESPE la publicación, en la biblioteca virtual de la Institución del trabajo **ANÁLISIS, DISEÑO E IMPLEMENTACIÓN DE UN DATAMART UTILIZANDO HERRAMIENTAS OPEN SOURCE PARA LA UNIDAD ADMINISTRATIVA Y FINANCIERA DE LA ESPE**, cuyo contenido, ideas y criterios son de nuestra exclusiva responsabilidad y autoría.

Sangolquí, Julio 2014

Cristian Darío Alvarez Calvopiña

Diego Armando Esparza Montes

DEDICATORIAS

Este proyecto de tesis va dedicado a mis padres Napoleon Esparza y Gloria Montes por la ayuda, la motivación, el amor que me dieron para cumplir un objetivo más en mi vida. También a mis dos queridos hermanos Cristian Esparza y Marcel Esparza por brindarme su apoyo, sus consejos para ser mejor en lo personal y lo laboral.

DIEGO ARMANDO ESPARZA MONTES

El presente proyecto de tesis va dedicado principalmente a mis padres Guillermo Alvarez y Nelly Calvopiña por el apoyo incondicional brindado a lo largo de toda mi vida estudiantil. También a mis hermanos, familiares cercanos y amigos en general que han estado presentes apoyándome y ayudándome a seguir adelante en los momentos difíciles.

CRISTIAN DARÍO ALVAREZ CALVOPÍÑA

AGRADECIMIENTOS

Agradezco primeramente a Dios que me da la fuerza para seguir adelante con mis metas, a mis padres y hermanos por su apoyo incondicional y el cariño que me demuestran a diario, a la Universidad de las Fuerzas Armadas ESPE y sus docentes por brindarme sus conocimientos que a futuro los pondré a prueba y a mis amigos de la universidad y del colegio por haberme demostrado una amistad sincera.

DIEGO ARMANDO ESPARZA MONTES

Agradezco principalmente a todo el personal que conforma las unidades administrativa y financiera de la Universidad de las Fuerzas Armadas ESPE por el apoyo y conocimiento brindado, a la directora de tesis Ingeniera Lorena Duque por su guía a lo largo de todo el proyecto, a mis padres y familiares por su apoyo y cariño incondicional demostrado a lo largo de toda mi vida estudiantil, a mi enamorada Ana María Yanez por su amor, comprensión y buenos deseos durante el transcurso del proyecto, finalmente a mis amigos por su apoyo y amistad brindada.

CRISTIAN DARÍO ALVAREZ CALVOPIÑA

RESUMEN

La Universidad de las Fuerzas Armadas ESPE almacena información de la Dirección Financiera y Talento Humano en diferentes bases de datos y en archivos planos, lo que hace difícil su procesamiento para generar conocimiento de acuerdo a las necesidades del usuario y presenta inconvenientes al realizar reportes de forma manual que emplean mucho tiempo en obtener información para análisis gerencial. El proyecto propuesto fue orientado al análisis, diseño e implementación de un conjunto de data marts (Data Warehouse) que servirá para simplificar procesos, mejorar el manejo y la consolidación de la información gestionada por la Dirección Financiera y Talento Humano. El proyecto se realizó empleando la metodología de Ralph Kimball y como herramienta de Business Intelligence open source se empleó Pentaho Community Edition. Los resultados mostraron que mediante la consolidación y depuración de información, se pudo reducir el tiempo de respuesta en la generación de informes, logrando un análisis eficiente y simplificado de la información, disminuyendo el tráfico generado en las consultas hacia las bases de datos, liberando recursos dentro del servidor. En conclusión este proyecto permite la simplificación de procesos y un análisis dinámico e intuitivo para la toma de decisiones.

PALABRAS CLAVES: DATA MARTS, DATA WAREHOUSE, OPEN SOURCE, PENTAHO COMMUNITY EDITION, RALPH KIMBALL.

ABSTRACT

The University of the Armed Forces ESPE stores information of Financial and Human Resource in different databases and flat files, making it difficult to generate knowledge processing according to user needs and drawbacks to make such reports manual that spend much time in management information for analysis. The proposed project was oriented analysis, design and implementation of a set of data marts (Data Warehouse) that will simplify processes, improve the management and consolidation of the information managed by the Finance and Human Resource. The project was made using the Ralph Kimball methodology and a tool for open source Business Intelligence Pentaho Community Edition was used. The results showed that by consolidating and debugging information could reduce the response time in reporting, achieving an efficient and simplified data analysis, reducing the traffic generated queries to databases, freeing resources inside the server. In conclusion, this project allows the simplification of processes and a dynamic and intuitive analysis for decision making.

KEY WORDS: DATA MARTS, DATA WAREHOUSE, OPEN SOURCE, PENTAHO COMMUNITY EDITION, RALPH KIMBALL.

CONTENIDO

CAPÍTULO 1.....	1
ESTRUCTURA ORGANIZACIONAL DE LA UNIVERSIDAD DE LAS FUERZAS ARMADAS ESPE	1
1.1 ESTRUCTURA ORGANIZACIONAL DE LA UNIVERSIDAD DE LAS FUERZAS ARMADAS ESPE.....	1
1.1.1 Antecedentes	1
1.1.2 Macrosistema de la Universidad de las Fuerzas Armadas ESPE.....	2
1.1.3 Sistemas de Dirección, Valor y Soporte	3
1.1.4 Red de Procesos de la Universidad de las Fuerzas Armadas ESPE.....	3
1.1.5 Cadena de valor de la Universidad de las Fuerzas Armadas ESPE	4
1.1.6 Red Institucional de la Universidad de las Fuerzas Armadas ESPE.....	5
CAPÍTULO 2.....	7
INTRODUCCIÓN	7
2.1 TEMA	7
2.2 INTRODUCCIÓN	7
2.3 PLANTEAMIENTO DEL PROBLEMA	8
2.4 JUSTIFICACIÓN E IMPORTANCIA	8
2.5 ALCANCE.....	8
2.6 OBJETIVOS	9
2.6.1 Objetivo General	9
2.6.2 Objetivos Específicos.....	9
2.7 METODOLOGÍA	10
2.8 HERRAMIENTAS	12
2.8.1 Motor de base de datos.....	12
2.8.1.1 Oracle.....	12
2.8.1.2 Sybase	13
2.8.1.3 Microsoft access.....	13
2.9 HERRAMIENTAS BI	13
2.9.1 Pentaho community.....	13
2.10 FACTIBILIDAD.....	14
2.10.1 Factibilidad Técnica.....	14

2.10.1.1	Requisitos de Hardware	15
2.10.1.2	Requisitos de Software.....	15
2.10.2	Factibilidad Económica.....	15
2.10.2.1	Operativa.....	16
2.10.2.2	Recursos de personal Técnico.....	16
CAPÍTULO 3.....		17
MARCO TEÓRICO.....		17
3.1	BUSINESS INTELLIGENCE.....	17
3.1.1	Introducción	17
3.1.2	Conceptos y definiciones	18
3.1.3	Aplicaciones.....	18
3.1.4	Arquitectura	19
3.2	DATA WAREHOUSE	20
3.2.1	Introducción	20
3.2.2	Definición	21
3.2.3	Características	21
3.2.3.1	Orientada al negocio	21
3.2.3.2	Integrada	23
3.2.3.3	Variante en el tiempo	24
3.2.3.4	No volátil	25
3.2.4	Cualidades.....	26
3.2.5	Ventajas.....	27
3.2.6	Desventajas	28
3.2.7	Redundancia.....	29
3.2.8	Estructura	30
3.2.8.1	Tablas de hecho agregadas y pre agregadas.....	32
3.2.9	Flujo de Datos	33
3.3	ARQUITECTURA DEL DATA WAREHOUSING	34
3.3.1	Introducción	34
3.3.2	OLTP.....	35
3.3.3	Load Manager	36
3.3.3.1	Extracción	36

3.3.3.2	Transformación	37
3.3.3.2.1	Codificación.....	38
3.3.3.2.2	Medida de atributos.....	38
3.3.3.2.3	Convenciones de nombramiento.....	39
3.3.3.2.4	Fuentes múltiples	40
3.3.3.2.5	Limpieza de datos	40
3.3.3.3	Carga.....	41
3.3.3.4	Proceso ETL.....	41
3.3.3.5	Tareas del ETL.....	42
3.3.4	Data Warehouse Manager	43
3.3.4.1	Base de datos multidimensional.....	44
3.3.4.2	Tablas de dimensiones	45
3.3.4.2.1	Tablas de dimensión tiempo	46
3.3.4.3	Tablas de hechos	47
3.3.4.4	Cubo multidimensional: introducción.....	50
3.3.4.4.1	Indicadores.....	51
3.3.4.4.2	Atributos	51
3.3.4.4.3	Jerarquías	52
3.3.4.4.4	Relación	53
3.3.4.4.5	Granularidad	54
3.3.4.5	Tipos de modelamiento de un Data Warehouse.....	54
3.3.4.5.1	Esquema en estrella.....	54
3.3.4.5.2	Esquema copo de nieve.....	55
3.3.4.5.3	Esquema constelación.....	57
3.3.4.6	Cubo Multidimensional: creación y ejemplificación	58
3.3.5	Data Mart	62
3.3.6	Metodologías de desarrollo de un proyecto de BI	64
3.3.6.1	Elección de la metodología para el desarrollo del proyecto.....	65
3.3.6.2	Metodología de Ralph Kimball.....	65
3.3.6.2.1	Planificación y Administración del Proyecto.....	66
3.3.6.2.2	Definición de los requerimientos del negocio.....	66
3.3.6.2.3	Diseño técnico de la arquitectura	66

3.3.6.2.4	Selección de Productos e Instalación	67
3.3.6.2.5	Modelado Dimensional	67
3.3.6.2.6	Diseño Físico	67
3.3.6.2.7	Diseño y Desarrollo de la Presentación de Datos	68
3.3.6.2.8	Especificación de Aplicaciones para Usuarios Finales.....	68
3.3.6.2.9	Desarrollo de Aplicaciones para Usuarios Finales	68
3.3.6.2.10	Implementación	68
3.3.6.2.11	Mantenimiento y Crecimiento	68
3.3.6.2.12	Gestión del Proyecto.....	68
3.3.7	Herramientas	69
3.3.7.1	Plataforma Pentaho Open Source Business Intelligence.....	69
3.3.7.2	Módulos de la Plataforma Pentaho BI	69
3.3.7.2.1	Reporting	69
3.3.7.2.2	Análisis	69
3.3.7.2.3	Mondrian.....	70
3.3.7.2.4	Dashboards.....	70
3.3.7.2.5	Integración de Datos	70
3.3.7.2.6	Diseño de Cubo.....	71
3.3.7.2.7	Base de Datos.....	71
CAPÍTULO 4.....		72
DESARROLLO DEL PROYECTO		72
4.1	DESARROLLO DEL BI PARA LA UNIDAD ADMINISTRATIVA Y FINANCIERA DE LA UNIVERSIDAD DE LAS FUERZAS ARMADAS ESPE.....	72
4.2	PLANEACIÓN Y ADMINISTRACIÓN DEL PROYECTO.....	72
4.2.1	Definición del proyecto.....	72
4.2.2	Preparación para un proyecto de BI.....	73
4.2.3	Alcance	73
4.2.4	Justificación	73
4.2.5	Planificación del proyecto.....	74
4.2.6	Administración del proyecto.....	75
4.3	DEFINICIÓN DE LOS REQUERIMIENTOS DEL NEGOCIO	75
4.3.1	Requerimientos módulos de las unidades administrativa y financiera.	75

4.3.2	Análisis y requerimientos.....	80
4.4	DISEÑO TÉCNICO DE LA ARQUITECTURA	80
4.4.1	Estándares	80
4.4.1.1	Almacenamiento de las carpetas del desarrollo de los datamarts.	81
4.4.1.1.1	Carpetas general	81
4.4.1.1.2	Subcarpetas	81
4.4.1.1.3	Nombre de los módulos	83
4.4.1.1.4	Nombre de los submódulos.....	84
4.4.1.2	Estándares para modelamiento.....	84
4.4.1.3	Estándar para base de datos stage	85
4.4.1.3.1	Nombre de objetos	85
4.4.1.3.2	Nombre de la base de datos stage	85
4.4.1.3.3	Nombre del esquema stage	86
4.4.1.4	Estándar para nombre y esquema de base de datos del datamart.....	86
4.4.1.4.1	Nombre de la base de datos datamart	86
4.4.1.4.2	Nombre del esquema datamart.....	86
4.4.1.5	Estándar para dimensiones y tabla de hechos	87
4.4.1.6	Estándares para la herramienta PENTAHO	92
4.4.1.6.1	Estándar pentaho data integration (PDI).....	93
4.4.1.6.2	Estándar pentaho report designer (PRD)	94
4.4.1.6.3	Estándar pentaho schema workbench (PSW)	94
4.4.1.6.4	Pentaho user console.....	95
4.4.1.7	Directorio de las publicaciones de reportes y cubos en pentaho.....	98
4.4.2	Esquema de base de datos en la fuente y el repositorio temporal	99
4.4.2.1	Diagrama General Presupuestos	100
4.4.2.2	Mapeo de Valores de la fuente al repositorio temporal módulo presupuestos.....	100
4.4.2.3	Diagrama de Archivos Planos usados para el módulo de presupuestos...	101
4.4.2.4	Diagrama de base de datos entidad relación en el repositorio temporal para el cubo de presupuestos.....	102
4.4.2.5	Diagrama general Activo Fijos	103
4.4.2.6	Mapeo de valores de la fuente al repositorio temporal módulo activos fijos.....	103

4.4.2.7	Diagrama de base de datos de la fuente para el módulo de activos fijos.....	104
4.4.2.8	Diagrama de base de datos entidad relación en el repositorio temporal para el cubo de activos fijos.....	105
4.4.2.9	Diagrama general inventarios	106
4.4.2.10	Mapeo de valores de la fuente al repositorio temporal módulo inventarios.....	106
4.4.2.11	Diagrama de base de datos de la Fuente para el módulo de Inventarios.....	107
4.4.2.1	Diagrama de base de datos entidad relación en el repositorio temporal para el cubo de inventarios.....	108
4.4.2.2	Diagrama general roles de pago - shifre	109
4.4.2.3	Mapeo de valores de la fuente al repositorio temporal módulo roles de pago – shifre....	110
4.4.2.4	Diagrama de base de datos de la fuente para el módulo roles de pago – shifre.....	111
4.4.2.5	Diagrama de base de datos entidad relación en el repositorio temporal para el cubo de roles de pago – shifre.....	112
4.4.2.6	Diagrama general matriculación	113
4.4.2.7	Mapeo de valores de la fuente al repositorio temporal módulo de matriculación.....	113
4.4.2.8	Diagrama de base de datos de la Fuente para el módulo de matriculación.....	114
4.4.2.9	Diagrama de base de datos entidad relación en el repositorio temporal para el cubo de matrículas.....	115
4.4.2.10	Diagrama general roles de pago - banner.....	116
4.4.2.11	Mapeo de valores de la fuente al repositorio temporal módulo roles de pago - banner	116
4.4.2.12	Diagrama de base de datos de la fuente del módulo de roles de pago – banner.....	117
4.4.2.13	Diagrama de base de datos entidad relación en el repositorio temporal para los cubos de roles de pago - banner	118
4.4.2.14	Diagrama general cuentas por cobrar.....	119
4.4.2.15	Mapeo de valores de la fuente al repositorio temporal módulo cuentas por cobrar.....	119
4.4.2.16	Diagrama de base de datos de la fuente de cuentas por cobrar	120

4.4.2.17	Diagrama de base de datos entidad relación en el repositorio temporal para los cubos de cuentas por cobrar.	121
4.4.3	Entorno Back Room.....	122
4.4.4	Entorno Front Room	122
4.5	SELECCIÓN DE PRODUCTOS E INSTALACIÓN	123
4.6	MODELADO DIMENSIONAL	124
4.6.1	Bus Matrix Data Warehouse	124
4.6.2	Modelado Dimensional	125
4.6.2.1	Diagrama general para el módulo de presupuestos.....	125
4.6.2.2	Mapeo de datos presupuestos.....	126
4.6.2.3	Modelo Dimensional Presupuestos	127
4.6.2.4	Mapeo de datos para tablas de dimensiones de Presupuestos.....	128
4.6.2.5	Mapeo de datos para tablas de hechos de Presupuestos.....	138
4.6.2.6	Diagrama general para el módulo de activos fijos.....	140
4.6.2.7	Mapeo de datos activos fijos	140
4.6.2.8	Modelo Dimensional Activos Fijos	142
4.6.2.9	Mapeo de datos para tablas de dimensiones de Activos Fijos.	143
4.6.2.10	Mapeo de datos para tablas de hechos de Activos Fijos.	154
4.6.2.11	Diagrama general para el módulo de inventarios.....	156
4.6.2.12	Mapeo de datos inventarios.....	156
4.6.2.13	Modelo Dimensional Inventarios.....	158
4.6.2.14	Mapeo de datos para tablas de dimensiones de Inventarios.....	159
4.6.2.15	Mapeo de datos para tablas de hechos de Inventarios.....	162
4.6.2.16	Diagrama general para el módulo de roles de pago (shifre)	163
4.6.2.17	Mapeo de datos roles de pago (shifre)	163
4.6.2.18	Modelo dimensional roles de pago (shifre).....	165
4.6.2.19	Mapeo de datos para tablas de dimensiones de Roles de pago (Shifre).....	166
4.6.2.20	Mapeo de datos para tablas de hechos de Roles de pago (Shifre).....	173
4.6.2.21	Diagrama general para el módulo de matrículas.....	174
4.6.2.22	Mapeo de datos de matrículas.	175
4.6.2.23	Modelo dimensional de matrículas.	176
4.6.2.24	Mapeo de datos para tablas de dimensiones de matrículas.	177

4.6.2.25	Mapeo de datos para tablas de hechos de matrículas.....	181
4.6.2.26	Diagrama general para el módulo de roles de pago - banner.....	183
4.6.2.27	Mapeo de datos roles de pago - banner.....	184
4.6.2.28	Modelo dimensional de roles de pago - banner.....	185
4.6.2.29	Mapeo de datos para dimensiones y hechos de roles de pago - banner.....	186
4.6.2.30	Diagrama general para el módulo de cuentas por cobrar.....	192
4.6.2.31	Mapeo de datos cuentas por cobrar.....	193
4.6.2.32	Modelo dimensional de cuentas por cobrar.....	194
4.6.2.33	Mapeo de datos para dimensiones y hechos de cuentas por cobrar.	195
4.7	MANUAL DE USUARIO SPOON.....	198
4.7.1	Proceso de extracción transformación y carga.....	198
4.7.2	Diseño del etl para el caso de nuestro proyecto.....	200
4.7.2.1	Entrada de tabla.....	200
4.7.2.2	Selecciona / renombra valores.....	204
4.7.2.3	Mapeo de valores.....	206
4.7.2.4	Añadir secuencia.....	207
4.7.2.5	Información del sistema.....	209
4.7.2.6	Insertar / actualizar.....	210
4.7.2.7	Añadir secuencia de auditoria.....	213
4.7.2.8	Salida de tabla de auditoria.....	215
4.7.3	Ejecutar el ETL.....	218
4.8	MANUAL CARGA DE ARCHIVOS CSV. AL SERVIDOR DE PENTAHO ..	221
4.9	MANUAL CARGA DE ARCHIVO DE EJECUCIÓN PRESUPUESTARIA ...	226
4.10	MANUAL DE USUARIO SCHEMA WORKBENCH.....	228
4.10.1	Introducción.....	228
4.10.2	Creación conexión.....	229
4.10.3	Creación de cubo.....	230
4.10.4	Creación dimensiones.....	233
4.10.5	Creación medidas.....	237
4.10.6	Guardar cubos.....	239
4.10.7	Publicación de cubos.....	240

4.10.7.1	Configuración.....	240
4.10.7.2	Publicación.....	240
4.10.8	Uso de un cubo en la consola de usuario	244
4.11	MANUAL DE USUARIO DE REPORT DESIGNER.....	251
4.11.1	Introducción	251
4.11.2	Proceso de creación de un reporte.....	251
4.11.3	Creación de parámetros.....	257
4.11.4	Diseño de reporte	258
4.11.4.1	Report Header	258
4.11.4.2	Details	259
4.11.4.3	Report Footer	260
4.11.5	Diseño de parámetros.....	261
4.11.6	Nombre del reporte	262
4.11.7	Previsualizar el reporte.....	263
4.11.8	Guardar y publicar el reporte	264
4.11.9	Crear gráfico dentro de un reporte (OPCIONAL)	268
4.11.9.1	Parámetros para el subreporte.....	271
4.11.10	Uso del reporte en la consola de usuario.....	273
4.12	MANUAL DE USUARIO DE DASHBOARD.....	277
4.12.1	INTRODUCCIÓN	277
4.12.2	PROCESO DE CREACIÓN DE UN DASHBOARD	277
4.12.3	Crear parámetros	283
4.12.4	Crear MDX Queries	284
4.12.5	Crear SQL Queries.....	285
4.12.6	Crear componentes.....	286
4.12.6.1	Advanced properties	287
CAPÍTULO 5.....		291
5.1	Conclusiones	291
5.2	Recomendaciones	292
BIBLIOGRAFIA		293

LISTADO DE FIGURAS

Figura 1: Macrosistema Universidad de las Fuerzas Armadas ESPE	2
Figura 2: Red de procesos en el macrosistema de la Universidad de las Fuerzas Armadas ESPE	4
Figura 3: Cadena de valor de la Universidad de las Fuerzas Armadas ESPE.....	4
Figura 4: Estructura organizacional de la Universidad de las Fuerzas Armadas ESPE.....	5
Figura 5: Estructura de la metodología para BI.....	11
Figura 6: Etapas de metodología de Ralph Kimball	12
Figura 7: Arquitectura pentaho BI.....	20
Figura 8: Función del datawarehouse.....	25
Figura 9: Estructura de datawarehouse.....	30
Figura 10: Flujo de datos del datawarehouse.....	33
Figura 11: Arquitectura del datawarehousing.....	34
Figura 12: Transformación de un OLTP a un DW	38
Figura 13: Transformación de medidas de OLTP a DW.....	39
Figura 14: Convención del nombramiento	39
Figura 15: Múltiples fuentes.....	40
Figura 16: Proceso de un ETL	42
Figura 17: Tablas de dimensiones	45
Figura 18: Tabla de hechos con dimensiones.....	47
Figura 19: Tabla de hechos	49
Figura 20: Cubo multidimensional	50
Figura 21: Jerarquía de dimensión fecha.....	52
Figura 22: Esquema estrella.....	55
Figura 23: Esquema copo de nieve.....	56
Figura 24: Esquema constelación	57
Figura 25: Estructura de un cubo multidimensional.....	58
Figura 26: Ejemplo de un esquema estrella	59
Figura 27: Cubo de ventas con dos dimensiones.....	59
Figura 28: Cubo ventas con dimensiones, indicadores	60
Figura 29: Matriz de las dos dimensiones.....	60
Figura 30: Cubo ventas con 3 dimensiones y 1 indicador.....	61
Figura 31: Estructura del cubo de ventas.....	61
Figura 32: Cubo ventas 4 dimensiones y 1 indicador	61
Figura 33: Cubo de ventas agrupadas por años.....	62
Figura 34: Arquitectura TOP-DOWN.....	63

Figura 35: Arquitectura BOTTOM-UP	63
Figura 36: Estructura carpetas.....	83
Figura 37: Ejemplo estándar nombre tabla.....	88
Figura 38: Ejemplo 3 iniciales nombre tabla.....	89
Figura 39: Ejemplo nombre atributo.....	89
Figura 40: Ejemplo usuario repetido.....	93
Figura 41: Ejemplo formato reporte.....	97
Figura 42: Pantalla filtros	98
Figura 43: Diagrama general presupuestos.....	100
Figura 44: Diagrama de archivos planos para presupuestos.....	101
Figura 45: Modelo entidad relación presupuesto.....	102
Figura 46: Diagrama general activos fijos.....	103
Figura 47: Diagrama de la fuente de activos fijos.....	104
Figura 48: Modelo entidad relación activos fijos	105
Figura 49: Diagrama general inventarios	106
Figura 50: Diagrama de la fuente de inventarios	107
Figura 51: Modelo entidad relación inventario.....	108
Figura 52: Diagrama general roles de pago - shifre	109
Figura 53: Diagrama de la fuente de roles de pago - shifre.....	111
Figura 54: Modelo entidad relación roles de pago - Shifre.....	112
Figura 55: Diagrama general matriculación	113
Figura 56: Diagrama de la fuente de matriculación.....	114
Figura 57: Modelo entidad relación matrículas.....	115
Figura 58: Diagrama general roles de pago - banner	116
Figura 59: Diagrama de la fuente de roles de pago - banner	117
Figura 60: Modelo entidad relación roles de pago - banner.....	118
Figura 61: Diagrama general cuentas por cobrar	119
Figura 62: Diagrama de la fuente de cuentas por cobrar	120
Figura 63: Modelo entidad relación cuentas por cobrar.....	121
Figura 64: Back Room	122
Figura 65: Front Room	123
Figura 66: Matriz datawarehouse	124
Figura 67: Diagrama general de presupuestos	125
Figura 68: Modelo dimensional de presupuestos	127
Figura 69: Diagrama general activos fijos.....	140
Figura 70: Modelo dimensional activos fijos	142
Figura 71: Diagrama general inventarios	156
Figura 72: Modelo dimensional inventarios.....	158
Figura 73: Diagrama general roles de pago – shifre	163

Figura 74: Modelo dimensional roles de pago - shifre	165
Figura 75: Diagrama general matrículas	174
Figura 76: Modelo dimensional matrículas	176
Figura 77: Diagrama general roles de pago - banner	183
Figura 78: Modelo dimensional roles de pago – banner.....	185
Figura 79: Diagrama general cuentas por cobrar	192
Figura 80: Modelo dimensional cuentas por cobrar.....	194
Figura 81: Carga de Spoon	198
Figura 82: Pantalla de entrada de spoon	199
Figura 83: Pantalla de menú spoon	199
Figura 84: Pantalla de interfaz de transformación	200
Figura 85: Pantalla de elemento entrada de tabla.....	201
Figura 86: Pantalla de entrada de tabla.....	201
Figura 87: Pantalla para crear conexión a una bd.....	202
Figura 88: Pantalla para probar conexión de bd	203
Figura 89: Pantalla de entrada de tabla.....	203
Figura 90: Pantalla para previsualizar	204
Figura 91: Pantalla spoon con dos elementos	204
Figura 92: Proceso de unión entre dos pasos.....	205
Figura 93: Pantalla 'selección/renombrar valores'.....	205
Figura 94: Pantalla 'selección/renombrar valores' con valores.....	206
Figura 95: Pantalla con mapeo de valores.....	207
Figura 96: Mapeo de valores con valores	207
Figura 97: Pantalla con secuencia 1	208
Figura 98: Añadir secuencia 1.....	208
Figura 99: Pantalla con información sistema.....	209
Figura 100: Selección información del sistema.....	209
Figura 101: Selección tipo información del sistema	210
Figura 102: Selección información del sistema.....	210
Figura 103: Pantalla con insertar/actualizar.....	211
Figura 104: Interfaz insertar / actualizar	211
Figura 105: Campos 'insertar / actualizar'	212
Figura 106: Ejecución script de tabla	213
Figura 107: Resultado de tabla creada.....	213
Figura 108: Pantalla con paso secuencia 2	214
Figura 109: Añadir secuencia 2	215
Figura 110: Paso de error	215
Figura 111: Crear salida de tabla	216
Figura 112: Datos de la salida de tabla	216

Figura 113: Ejecución script tabla auditoria	217
Figura 114: Resultado tabla auditoria creada.....	218
Figura 115: Botón para ejecutar ETL.....	218
Figura 116: Ejecución ETL.....	219
Figura 117: Guardar ETL	219
Figura 118: Ruta para guardar ETL	220
Figura 119: ETL Ejecutado	220
Figura 120: Pestaña datos en Excel 2013	221
Figura 121: Venta importar archivo de texto.....	222
Figura 122: Asistente para importar texto paso 1	222
Figura 123: Asistente para importar texto paso 2	223
Figura 124: Asistente para importar texto paso 3	223
Figura 125: Archivo de texto exportado.....	224
Figura 126: Modificación de archivo de texto	224
Figura 127: Ventana guardar como.....	225
Figura 128: Archivo de ejecución presupuestaria.....	226
Figura 129: Ventana guardar como.....	227
Figura 130: Interfaz Schema Workbench.....	228
Figura 131: Creación conexión	229
Figura 132: Configuración conexión.....	229
Figura 133: Conexión correcta.....	230
Figura 134: Crear nuevo esquema.....	231
Figura 135: Crear nuevo cubo.....	231
Figura 136: Creación tabla en cubo	232
Figura 137: Atributos de tabla del cubo	232
Figura 138: Crear dimensión	233
Figura 139: Crear tabla de dimensión.....	233
Figura 140: Atributos de tabla de dimensión	234
Figura 141: Crear nivel	234
Figura 142: Atributos del nivel.....	235
Figura 143: Dimensión con dos niveles.....	235
Figura 144: Atributos de la jerarquía	236
Figura 145: Atributos de la dimensión	237
Figura 146: Crear medidas	237
Figura 147: Atributos de las medidas.....	238
Figura 148: Cubo con dimensiones y medidas	238
Figura 149: Guardar cubo.....	239
Figura 150: Ruta y nombre para guardar el cubo	239
Figura 151: Publicación de cubo.....	241

Figura 152: Conexión al repositorio	242
Figura 153: Ruta de publicación del cubo	243
Figura 154: Conexión en consola de administración.....	243
Figura 155: Cubo publicado satisfactoriamente	244
Figura 156: URL de la consola de usuario.....	244
Figura 157: Interfaz de login de la consola de usuario	245
Figura 158: Interfaz consola de usuario	245
Figura 159: Nuevo saiku analytics.....	246
Figura 160: Ventana de saiku	246
Figura 161: Listado de cubos publicados.....	247
Figura 162: Area de trabajo para un cubo.....	247
Figura 163: Manipulación del cubo.....	248
Figura 164: Filtro del campo mes	249
Figura 165: Filtro del campo actividad	249
Figura 166: Manejo de gráficos en saiku.....	250
Figura 167: Iconos de formatos de exportación	250
Figura 168: Carga Report Designer	252
Figura 169: Interfaz report designer	252
Figura 170: Menú para abrir una plantilla.....	253
Figura 171: Escoger y abrir la plantilla.....	253
Figura 172: Crear una fuente de datos JDBC	254
Figura 173: Añadir nueva conexión.....	254
Figura 174: Configuración de la conexión.....	255
Figura 175: Mensaje de conexión satisfactoria	255
Figura 176: Crear queries para el reporte	256
Figura 177: Previsualizar información de un query	256
Figura 178: Añadir parámetros en el reporte.....	257
Figura 179: Configuración de un parámetro	257
Figura 180: Interfaz de parámetros creados.....	258
Figura 181: Añadir etiquetas en report header.....	259
Figura 182: Añadir campos a details.....	260
Figura 183: Añadir funciones en report footer	261
Figura 184: Añadir etiquetas en page header	262
Figura 185: Añadir parámetros en page header	262
Figura 186: Poner nombre al reporte	263
Figura 187: Reporte pre visualizado	264
Figura 188: Menú para publicar el reporte	264
Figura 189: Mensaje para guardar reporte	265
Figura 190: Guardar reporte.....	265

Figura 191: Conexión al repositorio	266
Figura 192: Configuración para publicar el reporte	267
Figura 193: Mensaje de reporte publicado.....	267
Figura 194: Añadir subreport.....	268
Figura 195: Añadir gráfico al reporte.....	269
Figura 196: Seleccionar query para el gráfico.....	269
Figura 197: Selección tipo de gráfico.....	270
Figura 198: Añadir campos por serie	270
Figura 199: Configuración del gráfico	271
Figura 200: Añadir parámetros al subreport	272
Figura 201: Importar parámetros	272
Figura 202: Gráfico pre visualizado.....	273
Figura 203: URL de la consola de usuario.....	274
Figura 204: Interfaz de login de la consola de usuario	274
Figura 205: Búsqueda del reporte publicado.....	275
Figura 206: Reporte abierto.....	275
Figura 207: Tipo de salida del reporte	276
Figura 208: Ejecución del reporte.....	276
Figura 209: URL de la consola de usuario.....	278
Figura 210: Interfaz de login de la consola de usuario	278
Figura 211: Abrir CDE	279
Figura 212: Guardar dashboard.....	279
Figura 213: Opciones de Layout.....	280
Figura 214: Opciones de plantillas	280
Figura 215: Opciones de plantillas personalizadas	281
Figura 216: Mensaje para aplicar plantilla.....	281
Figura 217: Diseño de la plantilla	282
Figura 218: Previsualizar plantilla.....	282
Figura 219: Crear parámetro	283
Figura 220: Crear componente para el parámetro.....	283
Figura 221: Crear MDX queries	284
Figura 222: 3 MDX queries creados.....	285
Figura 223: Crear SQL queries	286
Figura 224: Creación gráfico de barras	287
Figura 225: Propiedades avanzadas	288
Figura 226: Propiedad de valores visibles	288
Figura 227: Propiedad de orientación.....	288
Figura 228: Previsualizar dashboard con un gráfico de barras	289
Figura 229: 4 Componentes de gráfico en barras	290

Figura 230: Dashboard completo	290
--------------------------------------	-----

LISTADO DE CUADROS

Cuadro 1 Factibilidad económica	15
Cuadro 2 Toma Requerimientos.....	75
Cuadro 3 Estándar carpeta general	81
Cuadro 4 Estándar subcarpetas	82
Cuadro 5 Estándar módulos.....	84
Cuadro 6 Estándar submódulos	84
Cuadro 7 Estándar modelamiento.....	85
Cuadro 8 Estándar de nombre de base de datos stage.....	85
Cuadro 9 Estándar de nombre de esquema stage.....	86
Cuadro 10 Estándar de nombre de base de datos datamart.....	86
Cuadro 11 Estándar de nombre de esquema datamart.....	87
Cuadro 12 Estándar tablas de hechos y dimensiones.....	87
Cuadro 13 Estándar campos de dimensiones.....	88
Cuadro 14 Definición de dimensión tiempo	90
Cuadro 15 Estándar campos para tabla de hechos	91
Cuadro 16 Estándar usuarios y roles.....	92
Cuadro 17 Estándar nombre de conexiones	93
Cuadro 18 Estándar nombre de transformaciones.....	93
Cuadro 19 Estándar nombre de jobs	93
Cuadro 20 Estándar nombre de reportes.....	94
Cuadro 21 Estándar nombre de cubos.....	94
Cuadro 22 Estándar nombre de dashboard.....	95
Cuadro 23 Estándar nombre de vistas	95
Cuadro 24 Estructura estándar nombre de reportes ad-hoc.....	96
Cuadro 25 Ejemplo nombre reporte ad-hoc.....	96
Cuadro 26 Formato reportes	97
Cuadro 27 Formato filtros	98
Cuadro 28 Mapeo de valores de la fuente al repositorio temporal de presupuestos.....	100
Cuadro 29 Mapeo de valores de la fuente al repositorio temporal de activos fijos.....	103
Cuadro 30 Mapeo de valores de la fuente al repositorio temporal de inventarios	106

Cuadro 31 Mapeo de valores de la fuente al repositorio temporal de roles de pago - shifre	110
Cuadro 32 Mapeo de valores de la fuente al repositorio temporal de matriculación	113
Cuadro 33 Mapeo de valores de la fuente al repositorio temporal de roles de pago - banner	116
Cuadro 34 Mapeo de valores de la fuente al repositorio temporal de cuentas por cobrar	119
Cuadro 35 Selección de productos e instalación	123
Cuadro 36 Mapeo de datos de presupuesto	126
Cuadro 37 Estructura dimensión tiempo	130
Cuadro 38 Estructura dimensión proyecto	131
Cuadro 39 Estructura dimensión programa	132
Cuadro 40 Estructura dimensión actividad	133
Cuadro 41 Estructura dimensión subprograma	134
Cuadro 42 Estructura dimensión ítem	135
Cuadro 43 Estructura dimensión geográfica	136
Cuadro 44 Estructura dimensión fuente	137
Cuadro 45 Estructura tabla de hechos presupuesto	138
Cuadro 46 Mapeo de datos activos fijos	140
Cuadro 47 Estructura dimensión tipo	143
Cuadro 48 Estructura dimensión subtipo	144
Cuadro 49 Estructura dimensión subsección	145
Cuadro 50 Estructura dimensión sección	146
Cuadro 51 Estructura dimensión dirección	147
Cuadro 52 Estructura dimensión origen	148
Cuadro 53 Estructura dimensión área	149
Cuadro 54 Estructura dimensión custodio	150
Cuadro 55 Estructura dimensión clase	151
Cuadro 56 Estructura dimensión código del bien	152
Cuadro 57 Estructura dimensión dependencia	153
Cuadro 58 Estructura tabla de hechos activos fijos	154
Cuadro 59 Mapeo de datos inventarios	157
Cuadro 60 Estructura dimensión tipo de transacción	159
Cuadro 61 Estructura dimensión rubro	160
Cuadro 62 Estructura dimensión bodega	161
Cuadro 63 Estructura tabla de hechos inventario	162
Cuadro 64 Mapeo de datos roles de pago - shifre	164
Cuadro 65 Estructura dimensión cargo	166

Cuadro 66 Estructura dimensión tipo de empleado.....	167
Cuadro 67 Estructura dimensión empleado de la Universidad de las Fuerzas Armadas ESPE	168
Cuadro 68 Estructura dimensión ciudadano.....	169
Cuadro 69 Estructura dimensión ingresos y egresos.....	170
Cuadro 70 Estructura dimensión departamento.....	171
Cuadro 71: Estructura dimensión docente.....	172
Cuadro 72 Estructura tabla de hechos roles de pago - shifre.....	173
Cuadro 73 Mapeo de datos matrículas.....	175
Cuadro 74 Estructura dimensión campus banner.....	177
Cuadro 75 Estructura dimensión alumno banner.....	178
Cuadro 76 Estructura dimensión arancel banner.....	179
Cuadro 77 Estructura dimensión periodo banner	180
Cuadro 78 Estructura tabla de hechos matrículas por periodo	181
Cuadro 79 Estructura tabla de hechos matrículas por tiempo.....	182
Cuadro 80 Mapeo de datos roles de pago - banner.....	184
Cuadro 81 Estructura dimensión empleados banner	186
Cuadro 82 Estructura dimensión disposición banner	188
Cuadro 83 Estructura dimensión tipo empleado banner.....	189
Cuadro 84 Estructura dimensión rubro banner	190
Cuadro 85 Estructura tabla de hechos roles de pago banner	191
Cuadro 86 Mapeo de datos cuentas por cobrar.....	193
Cuadro 87 Estructura dimensión cliente.....	195
Cuadro 88 Estructura dimensión factura	196
Cuadro 89 Estructura tabla de hechos cuentas por cobrar	197

CAPÍTULO 1

ESTRUCTURA ORGANIZACIONAL DE LA UNIVERSIDAD DE LAS FUERZAS ARMADAS ESPE

1.1 ESTRUCTURA ORGANIZACIONAL DE LA UNIVERSIDAD DE LAS FUERZAS ARMADAS ESPE

1.1.1 Antecedentes

El 16 de junio de 1922 la “Escuela de Oficiales de Ingenieros” inicia sus actividades académicas con miembros de las Fuerzas Armadas, ante la necesidad de tecnificar los mandos cambia su nombre por Escuela de Artillería e Ingenieros en 1936. Amplía su pensum académico para nivelarlo con las demás universidades ecuatorianas, en 1948 denominándola Escuela Técnica de Ingenieros; luego en 1972 abre sus puertas a estudiantes civiles; y en 1977 se convierte en Universidad de las Fuerzas Armadas ESPE.

Este crecimiento institucional ha sido significativo para la Universidad de las Fuerzas Armadas ESPE, de tal forma que, en la actualidad cuenta con aproximadamente 19 mil estudiantes en sus 17 facultades, 3 institutos en la modalidad presencial; 3 facultades, 1 instituto en la modalidad a distancia; 28 programas de maestrías, 14 diplomados y 4 programas de especialización.

La estructura organizacional de la Universidad de las Fuerzas Armadas ESPE ha sido modificada en dos ocasiones (1978 y 1994) durante su evolución y desarrollo.

En el año de 1994 la Universidad de las Fuerzas Armadas ESPE modificó su estructura al crear tres vicerrectorados, con el propósito de generar una administración ágil y eficiente, y principalmente para impulsar la investigación científica y tecnológica. Sin embargo, dicha estructura mantiene como unidades

básicas de trabajo a las Facultades, las que se dedican por sobre todo a la docencia, sin un significativo desarrollo de la investigación y de la extensión.

La Constitución Política de la República del Ecuador y la Ley de Educación Superior contemplan como funciones fundamentales de las instituciones de Educación Superior las de **Docencia, Investigación y Extensión**. La Universidad de las Fuerzas Armadas como parte del Sistema Nacional de Educación Superior y en cumplimiento de dichas premisas, las integra como la esencia de su **Misión Institucional**.

Con referencia al régimen académico, el Estatuto de la Universidad de las Fuerzas Armadas ESPE (Art. 31) establece que “Cada uno de los programas carrera contará con una dirección; serán asistidos en la docencia, por los departamentos y centros según las áreas de formación profesional, y la estructura administrativa que exija el cumplimiento de sus objetivos”.

El estatuto de la Universidad de las Fuerzas Armadas ESPE (Art. 38) determina que “Los programas y proyectos de investigación y de extensión serán ejecutados en los departamentos, centros u otros organismos que se crearen para el efecto.”

1.1.2 Macrosistema de la Universidad de las Fuerzas Armadas ESPE

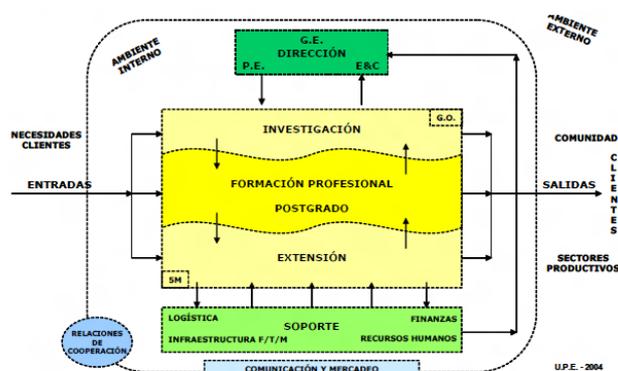


Figura 1: Macrosistema Universidad de las Fuerzas Armadas ESPE, (ESPE, 2010)

El esquema ilustra el enfoque de la Universidad de las Fuerzas Armadas ESPE como un Macrosistema abierto que interactúa con el ambiente externo. El núcleo de dicho Macrosistema está constituido por los Sistemas de Investigación, Docencia y Extensión, que son los que a su vez constituyen la Misión de la Institución.

Cabe indicar que el Macrosistema de la Universidad de las Fuerzas Armadas ESPE consta en el Plan Estratégico Institucional aprobado por el H. Consejo Politécnico con fecha 28 de Enero de 2005.

1.1.3 Sistemas de Dirección, Valor y Soporte

El Macrosistema de la Universidad de las Fuerzas Armadas ESPE está a su vez constituido por los siguientes Sistemas o Macro procesos:

De Dirección: Responsables del gobierno de la institución y por consiguiente de la Gestión de su desarrollo. Orientan a la Universidad de las Fuerzas Armadas ESPE al logro de sus metas de largo plazo mediante Directrices derivadas de la Planificación Estratégica.

De Valor: Ejecutores de la misión institucional bajo las directrices de la alta administración y por consiguiente responsables de las tres actividades básicas universitarias: Docencia, Investigación y Extensión.

De Soporte: Responsables de brindar servicios de soporte y habilitación a todas las unidades de la institución para su eficaz funcionamiento.

1.1.4 Red de Procesos de la Universidad de las Fuerzas Armadas ESPE

Los procesos que integran el Macrosistema de la Universidad de las Fuerzas Armadas ESPE están en el siguiente gráfico, así como las principales relaciones internas y externas.

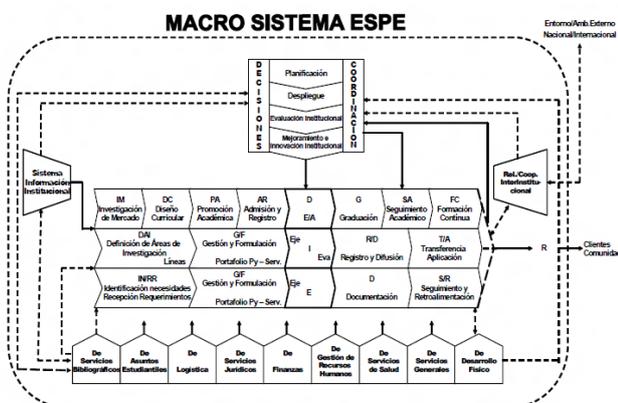


Figura 2: Red de procesos en el macrosistema de la Universidad de las Fuerzas Armadas ESPE, (ESPE, 2010)

1.1.5 Cadena de valor de la Universidad de las Fuerzas Armadas ESPE

La cadena de valor de la Universidad de las Fuerzas Armadas ESPE, está conformada por todas las actividades generadoras de valor agregado, como son: Docencia, Investigación y Extensión que se encuentran representadas en el siguiente gráfico:



Figura 3: Cadena de valor de la Universidad de las Fuerzas Armadas ESPE, (ESPE, 2010)

1.1.6 Red Institucional de la Universidad de las Fuerzas Armadas

ESPE

De conformidad con el Artículo 6 del Estatuto de la Universidad de las Fuerzas Armadas, aprobado mediante Decreto Ejecutivo No. 1585 de fecha 13 de junio de 2001, la estructura organizacional de la Universidad de las Fuerzas Armadas ESPE, es la siguiente:

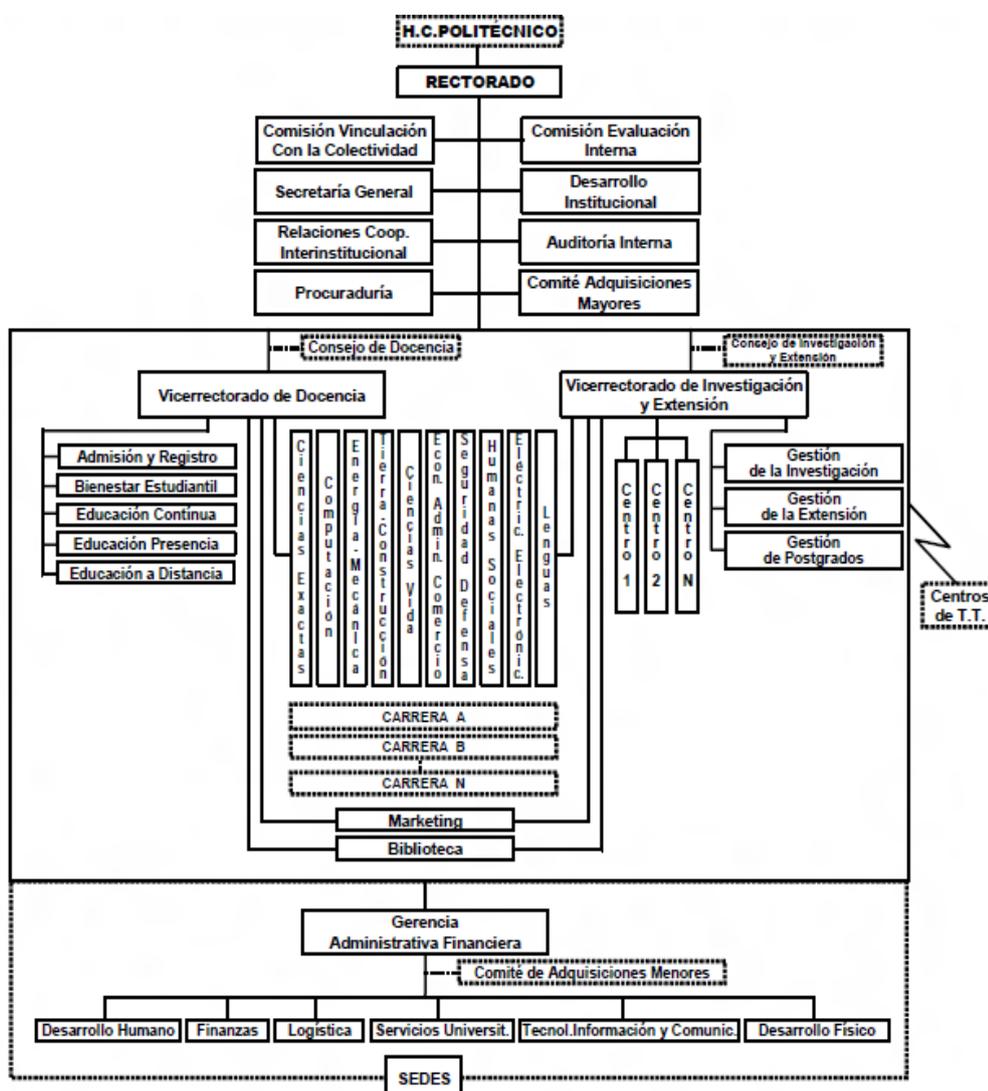


Figura 4: Estructura organizacional de la Universidad de las Fuerzas Armadas ESPE, (ESPE, 2010)

Las características principales de la Red Organizacional, son las siguientes:

- Cumplimiento cabal de las tres funciones principales de una universidad: docencia, investigación y extensión.
- Genera cambios en las actitudes de todos los miembros de la comunidad politécnica, para propiciar mayores logros puntuales en los campos de la investigación, extensión, así como en la docencia.
- Permite transformar una politécnica estrictamente docente a un instituto de educación superior que integra la investigación y el servicio a la comunidad como ejes fundamentales de su desarrollo.
- Los proyectos que integran los campos de docencia, investigación y extensión progresivamente serán más objetivos, propiciando el mejoramiento de la calidad institucional.
- Permite una mayor flexibilidad para que las carreras profesionales puedan crearse, modificarse o eliminarse de acuerdo a las exigencias del entorno.
- Evita la duplicidad de materias, facilita la estandarización en contenidos, así como la docencia especializada.
- Agrupa las áreas del conocimiento en departamentos, lo que refuerza la especialización y el trabajo en equipo.
- Se generan proyectos interdisciplinarios para satisfacer requerimientos específicos.
- Se diseñan planes curriculares más flexibles y en diferentes modalidades de estudio, de tal manera que facilitan al alumno optimizar su tiempo y sus actividades.
- La organización departamental funciona sobre la base de los principios de autogestión y autofinanciamiento, de manera que los participantes se apropien de los diferentes proyectos y puedan actuar basándose en resultados.
- La organización facilita la colaboración entre profesores y alumnos de las diferentes áreas de conocimiento para la revisión, generación, actualización y estandarización de los programas de estudio de una misma asignatura.

CAPÍTULO 2

INTRODUCCIÓN

2.1 TEMA

Análisis, Diseño e Implementación de un Data Mart Utilizando Herramientas Open Source Para Las Unidades Administrativa y Financiera de la Universidad de las Fuerzas Armadas ESPE.

2.2 INTRODUCCIÓN

La Dirección de Talento Humano, busca un cambio de la cultura organizacional hacia la administración del talento humano presente y del futuro, preparando y potenciando las competencias de los Servidores Públicos, Docentes, Militares y Trabajadores u Obreros, acordes con el desarrollo tecnológico y los requerimientos de la Institución; para lo cual el Reclutamiento, Selección y Contratación de personal se lo hace tomando en cuenta instrumentos de gestión para la vinculación de profesionales idóneos en el puesto de trabajo; la capacitación está focalizada para alcanzar los objetivos institucionales (planificación, ejecución y evaluación de los planes de capacitación); la Evaluación del Desempeño se encuentra encaminada a evaluar el rendimiento óptimo del desempeño para conseguir resultados con eficiencia y eficacia para la Organización, Acciones de Bienestar Social a fin de levantar los niveles de motivación que se reflejarán en el desempeño de las actividades, la administración de las remuneraciones; y la Salud Ocupacional para determinar la aptitud laboral, historia clínica ocupacional de ingresos y salidas de la Institución, así como programas de vigilancia de la salud.

Este cambio va de la mano con la aplicación de acciones de mejora continua en los procesos y el trabajo conjunto y coordinado con todas las dependencias, en el fomento de una cultura de calidad dentro de la Institución.

2.3 PLANTEAMIENTO DEL PROBLEMA

La Dirección financiera y recursos humanos de la Universidad de las Fuerzas Armadas ESPE, maneja una cantidad considerable de información utilizando varias bases de datos (Sybase, Oracle y Access), como un repositorio muy completo de los datos. Esta información no se encuentra procesada de acuerdo a las necesidades institucionales o de los órganos de control universitario como el CES, SENESCYT, CEACES; de tal manera que cuando se requieren reportes, informes estadísticos o tabulares es necesario realizarlos de forma manual.

Existen muchas relaciones entre los datos que no han podido utilizarse para efectos de planificación, organización, logística, etc. Siendo por tanto indispensable el procesamiento, adecuando de estos datos para cubrir estas necesidades.

2.4 JUSTIFICACIÓN E IMPORTANCIA

El proyecto propuesto busca procesar y reorganizar la información de la base de datos de la dirección financiera y recursos humanos, logrando con esto la simplificación de procesos tales como la elaboración de reportes, consultas determinadas y en general la toma de decisiones.

La solución que se plantea es el diseño e implementación de un Data Mart, el mismo que estará conformado por la recopilación de información obtenida de las distintas bases de datos ya sean Oracle, Sybase o Access, logrando con esto un análisis eficiente y simplificado de la información, disminuyendo considerablemente el tráfico generado en las consultas hacia la base de datos y por consiguiente liberando recursos dentro del servidor.

2.5 ALCANCE

Se realizará el diseño e implementación de un Data Mart para la automatización y depuración de la información contenida en la base de datos de las Unidades Administrativa y Financiera de la Universidad de las Fuerzas Armadas ESPE. Esto se realizará con la metodología de Ralph Kimball la misma que ayudará

en la planificación, toma de requerimientos, análisis, construcción y en general en cada una de las fases que componen el ciclo de vida del proyecto.

La aplicación contempla procesos de consolidación de la información, para que esta pueda ser utilizada en reportes variados que permitan considerar la información para planes de mejora o políticas institucionales. Además de que puede recogerse la información para realizar minería de datos y efectuar correlaciones entre éstos, se realizarán reportes automatizados con información precisa, en base a los requerimientos del área antes mencionada.

Una vez implementado el Data Mart para las Unidades Administrativas y Financieras, podrá ser integrado con Data Marts que se realizaran posteriormente y que conformarán el Data Warehouse de la Universidad de las Fuerzas Armadas ESPE, entre estos se encuentra por ejemplo el Data Mart Académico.

2.6 OBJETIVOS

2.6.1 Objetivo General

Diseñar e implementar un Data Mart mediante el uso de la metodología de Ralph Kimbal y con la utilización de herramientas open source para las Unidades Administrativa y Financiera de la Universidad de las Fuerzas Armadas ESPE.

2.6.2 Objetivos Específicos

- Realizar la toma de requerimientos en colaboración con los encargados de la dirección financiera para el diseño e implementación del Data Mart.
- Utilizar la metodología de Ralph Kimbal para el desarrollo del proyecto.
- Identificar las necesidades de la unidad administrativa y financiera de la Universidad de las Fuerzas Armadas ESPE para realizar un análisis y visualización de su información.

- Especificar un esquema de Data Mart y las consultas de análisis asociadas.
- Construir el Data Mart cumpliendo con los reportes solicitados por la Unidad Administrativa y Financiera de la Universidad de las Fuerzas Armadas ESPE.
- Automatizar la información de la Unidad Administrativa y Financiera de la Universidad de las Fuerzas Armadas ESPE mediante la herramientas Pentaho Dashboard editor, Pentaho Data Integration, Pentaho Schema Workbench, Report Designer, la cual ayudará a la construcción del datamart para obtener informes detallados, gráficos interactivos, creación de cubos, dashboards para la visualización del usuario acorde a los requerimientos.

2.7 METODOLOGÍA

El Data Warehouse es un conglomerado de todos los Data Marts dentro de una empresa, siendo una copia de los datos transaccionales estructurados de una forma especial para el análisis, de acuerdo al Modelo Dimensional (no normalizado), que incluye, como se vió, las dimensiones de análisis y sus atributos, su organización jerárquica, así como los diferentes hechos de negocio que se quieren analizar. Por un lado se tiene tablas para las representar las dimensiones y por otro lado tablas para los hechos (las facts tables). Los diferentes Data Marts están conectados entre sí por la llamada bus structure, que contiene los elementos anteriormente citados a través de las dimensiones conformadas (que permiten que los usuarios puedan realizar queries conjuntos sobre los diferentes data marts, pues este bus contiene los elementos en común que los comunican). Una dimensión conformada puede ser, por ejemplo, la dimensión cliente, que incluye todos los atributos o elementos de análisis referentes a los clientes y que puede ser compartida por diferentes data marts (ventas, pedidos, gestión de cobros, etc.).

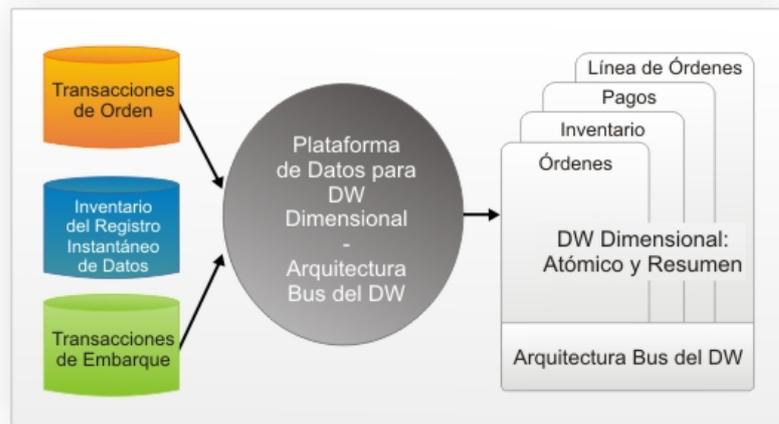


Figura 5: Estructura de la metodología para BI, (Espinoza, 2011)

Esta metodología también se referencia como Bottom-up, pues al final el Data warehouse Corporativo no es más que la unión de los diferentes data marts, que están estructurados de una forma común a través de la estructura de bus. Esta característica le hace más flexible y sencilla de implementar, pues se puede construir un Data Mart como primer elemento del sistema de análisis, y luego ir añadiendo otros que comparten las dimensiones ya definidas o incluyen otras nuevas. En este sistema, los procesos ETL extraen la información de los sistemas operacionales y los procesan igualmente en el área stage, realizando posteriormente el llenado de cada uno de los Data Mart de una forma individual, aunque siempre respetando la estandarización de las dimensiones (dimensiones conformadas).

La metodología para la construcción del Data warehouse incluye las 4 fases que son:

- Selección del proceso de negocio.
- Definición de la granularidad de la información.
- Elección de las dimensiones de análisis.
- Identificación de los hechos o métricas. Tratamiento de los cambios, Dimensiones Lentamente Cambiantes (SCD).

La Metodología de kimball realiza los siguientes pasos para la construcción de los data marts

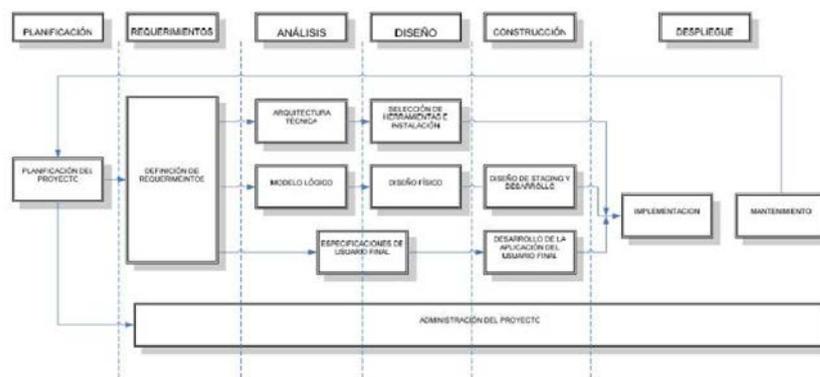


Figura 6: Etapas de metodología de Ralph Kimball, (Lopez, 2008)

Etapas de la Metodología de Kimball.

- **Planificación**, Planificación del Proyecto.
- **Requerimientos**, Definición de requerimientos.
- **Análisis**, Arquitectura técnica.
- **Diseño**, Selección de herramientas, Diseño físico.
- **Construcción**, Diseño de la puesta en escena (Staging) y desarrollo.
- **Despliegue**, Implantación y Mantenimiento.

2.8 HERRAMIENTAS

2.8.1 Motor de base de datos

2.8.1.1 Oracle

Oracle soporta todas las funciones que se esperan de un servidor "serio": un lenguaje de diseño de base de datos muy completo (PL/SQL) que permite implementar diseños "activos", con triggers y procedimientos almacenados, con una integridad referencial declarativa bastante potente

Oracle es el motor de base de datos relacional más usado a nivel mundial y se lo destaca por su:

- Soporte de transacciones.
- Estabilidad.

- Escalabilidad.
- Soporte multiplataforma.

2.8.1.2 Sybase

Es una plataforma para el manejo de datos que provee servicios de tiempo real para aplicaciones con gran volumen de transacciones.

La mayoría de aplicaciones de la Universidad de las Fuerzas Armadas ESPE de Intranet e Internet están desarrolladas para trabajar con esta Base de Datos.

2.8.1.3 Microsoft access

Microsoft Access, también conocido como **Microsoft Office Access**, es un sistema de gestión de base de datos de Microsoft que combina la relacional Microsoft Jet Database Engine con una interfaz gráfica de usuario y herramientas de desarrollo de software. Es miembro de la Microsoft Office suite de aplicaciones, incluido en las ediciones Professional y superior o se vende por separado. El 12 de mayo de 2010, la versión actual de Microsoft Access 2010 fue lanzado por Microsoft en Office 2010, Microsoft Office Access 2007 es la versión anterior.

MS Access almacena los datos en un formato propio basado en el motor de base de datos Access Jet. También puede importar o vincular directamente a los datos almacenados en otras aplicaciones y bases de datos.

2.9 HERRAMIENTAS BI

2.9.1 Pentaho community

Es una herramienta open source, es decir, de fuente totalmente abierta que incluye el análisis de los ETL, metadatos y capacidades de presentación de informes. Pentaho ofrece varios productos como:

Pentaho Analysis Services: cuyo nombre código Mondrian es una servidor OLAP (procesamiento analítico en línea) escrito en Java. Es compatible con el MDX

(expresiones multidimensionales) y el lenguaje de consulta XML para el Análisis y especificaciones de la interfaz olap4j.

Pentaho Reporting: Consiste en un motor de presentación, capaz de generación de informes programáticos sobre la base de un archivo de definición XML. Sobre esta solución se han desarrollado muchas herramientas, por ejemplo informes, diseñadores de interfaz gráfica de usuario, y asistentes tipo wizard. Un uso notable de esta herramienta es el Generador de informes para OpenOffice.org

Pentaho Data Mining: Es una envoltura alrededor del proyecto Weka. Es una suite de software que usa estrategias de aprendizaje de máquina, aprendizaje automático y minería de datos. Cuenta con series de clasificación, regresión, reglas de asociación, y algoritmos de clustering, así apoyar las tareas de análisis predictivo.

Pentaho Dashboard: Es una plataforma integrada para proporcionar información sobre sus datos, donde se pueden ver informes, gráficos interactivos y los cubos creados con las herramientas Pentaho Schema Workbench.

Pentaho para Apache Hadoop: Es un conector de bajo nivel para facilitar el acceso a MUY grandes volúmenes manejados en el proyecto Apache Hadoop, la Suite de Pentaho BI para Hadoop permite abordar los mayores desafíos que experimentan los usuarios de Hadoop - , sobre su empinada curva de aprendizaje técnico, la falta de personal técnico cualificado y la falta de disponibilidad de las aplicaciones de desarrollo y despliegue para llevar a cabo la integración de datos e inteligencia de negocios con Hadoop.

2.10 FACTIBILIDAD

2.10.1 Factibilidad Técnica

La Universidad de las Fuerzas Armadas ESPE cuenta con el área de UTICS, la misma que proveerá de los siguientes requerimientos:

2.10.1.1 Requisitos de Hardware

- Servidor de Base de Datos
- Servidor de Aplicaciones
- Computador para uso de los desarrolladores

2.10.1.2 Requisitos de Software

- Sistema operativo indistinto (Windows o Linux)
- Motor de Base de Datos
- Herramientas de Business Intelligence

2.10.2 Factibilidad Económica

En la parte económica se necesitará lo siguiente:

Cuadro 1
Factibilidad económica

Detalle	Cantidad	Valor Unitario	Total
Desarrolladores	2	800 x mes	1600
Oracle	1	0	0
Sybase	1	0	0
Pentaho Community	1	0	0
Internet	1	30	30
Computador	1	800	800
Impresora	1	150	150
Licencias	1	0	0
TOTAL			2580

NOTA: La Universidad de las Fuerzas Armadas ESPE cuenta con las licencias para las distintas bases de datos que ocupa que son Oracle, Sybase y Access por lo tanto no tiene costo y la herramienta BI es gratuita.

2.10.2.1 Operativa

Debido a que este proyecto será realizado para las direcciones financieras y administrativas de la Universidad de las Fuerzas Armadas ESPE, se cuenta con su total colaboración. Explícitamente el director/a ha dado el visto bueno para el diseño e implementación del presente proyecto que a su vez será presentado como tesis de grado.

2.10.2.2 Recursos de personal Técnico

Dentro del proyecto existe la necesidad de contar con gente especializada a nivel del área de sistemas es decir en el ámbito técnico, desde este punto de vista en el proyecto intervendrán:

- Director de carrera – Ingeniero de Sistemas
- Director de tesis – Ingeniero de Sistemas
- Coodirector de tesis – Ingeniero de Sistemas
- Informantes (para elaboración de plan de tesis) – Ingenieros de Sistemas
- Tesistas – Egresados de la carrera de Ingeniería de Sistemas

CAPÍTULO 3

MARCO TEÓRICO

3.1 BUSINESS INTELLIGENCE

3.1.1 Introducción

Históricamente, la tecnología de Business Intelligence ha encontrado lugar en dos niveles primarios: entre los altos ejecutivos quienes necesitan obtener información estratégica y entre los administradores de la línea de negocios que son responsables del análisis táctico. Estas tradicionales actividades de soporte a la decisión son importantes, pero ellos solamente muestran superficialmente el potencial de la inteligencia de negocios dentro de la empresa., involucrando quizá el 5% de los usuarios y el 10% de los datos disponibles (Information Builders, 2005).

Desde principios de los 90's, las aplicaciones de BI han evolucionado dramáticamente en muchas direcciones, debido al crecimiento exponencial de la información. Desde reportes operacionales generados por mainframes, modelación estadística de campañas publicitarias, ambientes OLAP multidimensionales para analistas así como dashboards y scorecards para ejecutivos. Las compañías empiezan a demandar más formas de analizar y realizar reportes de datos.

Las inversiones en aplicaciones empresariales, tales como planeación de recursos (ERP) y administración de la relación con el cliente (CRM), han resultando en una enorme cantidad de datos dentro de las organizaciones. Estas organizaciones ahora quieren apalancar estas inversiones y usar la información para ayudarles a tomar mejores decisiones, se más ágiles con organización y tener una mayor comprensión de cómo correr sus negocios.

Por ellos mucha pequeña y mediana empresa está adoptando BI para ayudarles a poner en marcha sus negocios.

El corazón de Business Intelligence es la habilidad de una organización para acceder y analizar la información, y entonces explotar su ventaja competitiva. En la

era digital, las capacidades que ofrece Business Intelligence será la diferencia entre el éxito y el fracaso.

3.1.2 Conceptos y definiciones

Las aplicaciones de Business Intelligence (BI) son herramientas de soporte de decisiones que permiten en tiempo real, acceso interactivo, análisis y manipulación de información crítica para la empresa. Estas aplicaciones proporcionan a los usuarios un mayor entendimiento que les permite identificar las oportunidades y los problemas de los negocios. Los usuarios son capaces de acceder y apalancar una vasta cantidad de información y analizar sus relaciones y entender las tendencias que últimamente están apoyando las decisiones de los negocios. Estas herramientas previenen una potencial pérdida de conocimiento dentro de la empresa que resulta de una acumulación masiva reinformación que no es fácil de leer o de usar.

3.1.3 Aplicaciones

Durante el periodo formativo, las compañías han descubierto activamente nuevas maneras de usar sus datos para apoyar la toma de decisiones, realizar una optimización de procesos y realizar reportes operacionales. Y durante esta era de invenciones, los vendedores de tecnología de BI han construido nichos de software para implementar cada nuevo patrón de aplicaciones que las compañías inventan. Estos patrones de aplicación resultan en productos de software centrados exclusivamente en cinco estilos de BI (Microstrategy, 2002), tales como:

- a. Reporte empresarial.** Los reportes escritos son usados para generar reportes estáticos altamente formateados destinados para ampliar su distribución con mucha gente.
- b. Cubos de análisis.** Los cubos basados en herramientas de BI son usados para proveer capacidades analíticas a los administradores de negocios.

c. Vistas Ad Hoc Query y análisis. Herramientas OLAP relacionales son usadas para permitir a los expertos visualizar la base de datos y ver cualquier respuesta y convertirla en información transaccional de bajo nivel.

d. Data mining y análisis estadísticos. Son herramientas usadas para desempeñar modelado predictivo o para descubrir la relación causa efecto entre dos métricas.

e. Entrega de reportes y alertas. Los motores de distribución de reportes son usados para enviar reportes completos o avisos a un gran número de usuarios, dichos reportes se basan en suscripciones, calendarios, etc

Hasta este punto, las grandes empresas han tenido que comprar diferentes conjuntos de herramientas de BI a distintos vendedores, con cada herramienta dirigida a una nueva aplicación de BI y cada una de ellas dando al usuario funcionalidad en solo uno de los estilos de BI.

3.1.4 Arquitectura

La solución Business Intelligence OpenSource Pentaho pretende ser una alternativa a las soluciones propietarias tradicionales más completas: Business Objects, Cognos, Microstrategy, Microsoft, etc... por lo que incluye todos aquellos componentes que se puede encontrar en las soluciones BI propietarias más avanzadas:

- Reporting
- Analysis
- Dashboards
- Workflow
- Data Mining
- ETL
- Single Sign-On. Ldap

- Auditoría de uso y rendimiento
- Planificador
- Notificador
- Seguridad. Perfiles

El siguiente esquema muestra la arquitectura estructurada de las diferentes componentes que forman parte de pentaho:

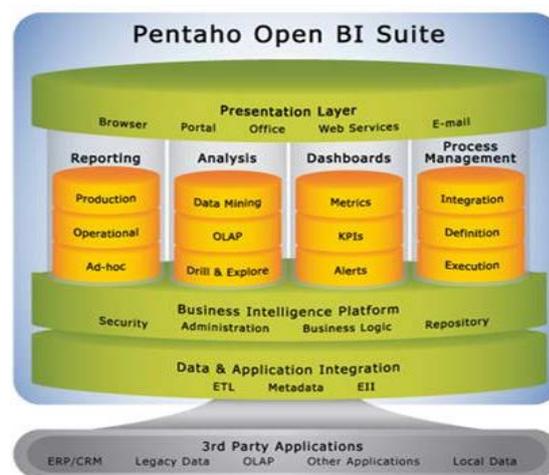


Figura 7: Arquitectura pentaho BI, (Espinoza, 2011)

3.2 DATA WAREHOUSE

3.2.1 Introducción

Debido a que para llevar a cabo BI, es necesario gestionar datos guardados en diversos formatos, fuentes y tipos, para luego depurarlos e integrarlos, además de almacenarlos en un solo destino o base de datos que permita su posterior análisis y exploración, es imperativo y de vital importancia contar con un proceso que satisfaga todas estas necesidades. Este proceso se denomina Data Warehousing.

El Data Warehousing, es el encargado de extraer, transformar, consolidar, integrar y centralizar los datos que la empresa genera en todos los ámbitos de su actividad diaria de negocios (compras, ventas, producción, etc.) y/o información externa relacionada. Permitiendo de esta manera el acceso y exploración de la información requerida, a través de una amplia gama de posibilidades de análisis

multivariantes, con el objetivo final de dar soporte al proceso de toma de decisiones estratégico y táctico.

3.2.2 Definición

El Data Warehousing posibilita la extracción de datos de sistemas operacionales y fuentes externas, permite la integración y homogeneización de los datos de toda la empresa, provee información que ha sido transformada y resumida, para que ayude en el proceso de toma de decisiones estratégicas y tácticas.

El Data Warehousing, convertirá entonces los datos operacionales de la empresa en una herramienta competitiva, debido a que pondrá a disposición de los usuarios indicados la información pertinente, correcta e integrada, en el momento que se necesita.

Pero para que el Data Warehousing pueda cumplir con sus objetivos, es necesario que la información que se extrae, transforma y consolida, sea almacenada de manera centralizada en una base de datos con estructura multidimensional denominada Data Warehouse (DW).

Una de las definiciones más famosas sobre DW, es la de William Harvey Inmon, quien define: “Un Data Warehouse es una colección de datos orientada al negocio, integrada, variante en el tiempo y no volátil para el soporte del proceso de toma de decisiones de la gerencia”

3.2.3 Características

3.2.3.1 Orientada al negocio

La primera característica del DW, es que la información se clasifica en base a los aspectos que son de interés para la organización. Esta clasificación afecta el diseño y la implementación de los datos encontrados en el almacén de datos, debido

a que la estructura del mismo difiere considerablemente a la de los clásicos procesos operacionales orientados a las aplicaciones.

A continuación, y con el fin de obtener una mejor comprensión de las diferencias existentes entre estos dos tipos de orientación, se realizará un análisis comparativo:

- Con respecto al nivel de detalle de los datos, el DW excluye la información que no será utilizada exclusivamente en el proceso de toma de decisiones; mientras que en los procesos orientados a las aplicaciones, se incluyen todos aquellos datos que son necesarios para satisfacer de manera inmediata los requerimientos funcionales de la actividad que soporten. Por ejemplo, los datos comunes referidos al cliente, como su dirección de correo electrónico, fax, teléfono, D.N.I., código postal, etc., que son tan importantes de almacenar en cualquier sistema operacional, no son tenidos en cuenta en el depósito de datos por carecer de valor para la toma de decisiones, pero sí lo serán aquellos que indiquen el tipo de cliente, su clasificación, ubicación geográfica, sexo, edad, etc.
- En lo que concierne a la interacción de la información, los datos operacionales mantienen una relación continua entre dos o más tablas, basadas en alguna regla comercial vigente; en cambio las relaciones encontradas en los datos residentes del DW son muchas, debido a que por lo general cada tabla del mismo estará conformada por la integración de varias tablas u otras fuentes del ambiente operacional, cada una con sus propias reglas de negocio inherentes.

El origen de este contraste es totalmente lógico, ya que el ambiente operacional se diseña alrededor de las aplicaciones u programas que necesite la organización para llevar a cabo sus actividades diarias y funciones específicas. Por ejemplo, una aplicación de una empresa minorista manejará: stock, lista de precios, cuentas corrientes, pagos diferidos, impuestos, retenciones, ventas, notas de crédito, compras, etc. De esta manera, la base de datos combinará estos elementos en una estructura que se adapte a sus necesidades.

En contraposición, por ejemplo, para un fabricante el ambiente DW se organizará alrededor de entidades de alto nivel tales como: clientes, productos, rubros, proveedores, vendedores, zonas, etc. Que son precisamente aquellos sujetos mediante los cuales se desea analizar la información. Esto se debe a que el depósito de datos se diseña para realizar consultas e investigaciones sobre las actividades de la organización y no para soportar los procesos que se realizan en ella.

En síntesis, la ventaja de contar con procesos orientados a la aplicación, está fundamentada en la alta accesibilidad de los datos, lo que implica un elevado desempeño y velocidad en la ejecución de consultas, ya que las mismas están predeterminadas; mientras que en el DW para satisfacer esta ventaja se requiere que la información este des normalizada, es decir, con redundancia y que la misma esté dimensionada, para evitar tener que recorrer toda la base de datos cuando se necesite realizar algún análisis determinado, sino que simplemente la consulta sea enfocada por variables de análisis que permitan localizar los datos de manera rápida y eficaz, para poder de esta manera satisfacer una alta demanda de complejos exámenes en un mínimo tiempo de respuesta.

3.2.3.2 Integrada

La integración implica que todos los datos de diversas fuentes que son producidos por distintos departamentos, secciones y aplicaciones, tanto internas como externas, deben ser consolidados en una instancia antes de ser agregados al DW. A este proceso se lo conoce como Extracción, Transformación y Carga de Datos (Extraction, Transformation and Load - ETL).

La integración de datos, resuelve diferentes tipos de problemas relacionados con las convenciones de nombres, unidades de medidas, codificaciones, fuentes múltiples, etc., cada uno de los cuales será correctamente detallado y ejemplificado más adelante.

La causa de dichos problemas, se debe principalmente a que a través de los años los diseñadores y programadores no se han basado en ningún estándar concreto para definir nombres de variables, tipos de datos, etc., ya sea por carecer de ellos o

por no creer que sean necesarios. Por lo cual, cada uno por su parte ha dejado en cada aplicación, módulo, tabla, etc., su propio estilo personalizado, confluyendo de esta manera en la creación de modelos muy inconsistentes e incompatibles entre sí.

Los puntos de integración afectan casi todos los aspectos de diseño, y cualquiera sea su forma, el resultado es el mismo, ya que la información será almacenada en el DW en un modelo globalmente aceptable y singular, aun cuando los sistemas operacionales y demás fuentes almacenen los datos de maneras disímiles, para que de esta manera el usuario final este enfocado en la utilización de los datos del depósito y no deba cuestionarse sobre la confiabilidad o solidez de los mismos.

3.2.3.3 Variante en el tiempo

Debido al gran volumen de información que se manejará en el DW, cuando se le realiza una consulta, los resultados deseados demorarán en originarse. Este espacio de tiempo que se produce desde la búsqueda de datos hasta su consecución es del todo normal en este ambiente y es, precisamente por ello, que la información que se encuentra dentro del depósito de datos se denomina de tiempo variable.

Esta característica básica, es muy diferente de la información encontrada en el ambiente operacional, en el cual, los datos se requieren en el momento de acceder, es decir, que se espera que los valores procurados se obtengan a partir del momento mismo de acceso.

Esto contribuye a una de las principales ventajas del almacén de datos: los datos son almacenados junto a sus respectivos históricos. Esta cualidad que no se encuentra en fuentes de datos operacionales, garantiza poder desarrollar análisis de la dinámica de la información, pues ella es procesada como una serie de instantáneas, cada una representando un periodo de tiempo. Es decir, que gracias al sello de tiempo se podrá tener acceso a diferentes versiones de la misma información.

Es importante tener en cuenta la granularidad de los datos, así como también la intensidad de cambio natural del comportamiento de los fenómenos de la actividad

que se desarrolle, para evitar crecimientos incontrolables y desbordamientos de la base de datos.

El intervalo de tiempo y periodicidad de los datos debe definirse de acuerdo a la necesidad y requisitos de los usuarios.

Es elemental aclarar, que el almacenamiento de datos históricos, es lo que permite al DW desarrollar pronósticos y análisis de tendencias y patrones, a partir de una base estadística de información.

3.2.3.4 No volátil

La información es útil para el análisis y la toma de decisiones solo cuando es estable. Los datos operacionales varían momento a momento, en cambio, los datos una vez que entran en el DW no cambian.

La actualización, o sea, insertar, eliminar y modificar, se hace de forma muy habitual en el ambiente operacional sobre una base, registro por registro, en cambio en el depósito de datos la manipulación básica de los datos es mucho más simple, debido a que solo existen dos tipos de operaciones: la carga de datos y el acceso a los mismos.

Por esta razón es que en el DW no se requieren mecanismos de control de la concurrencia y recuperación.

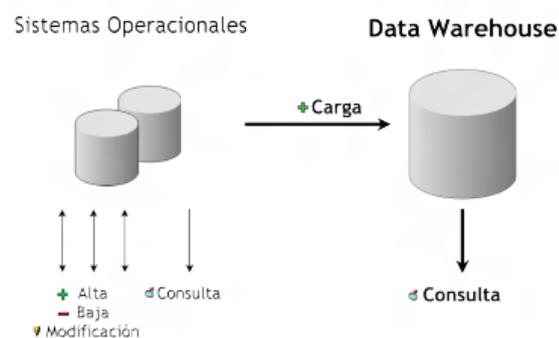


Figura 8: Función del datawarehouse, (Bernabeu, 2009)

3.2.4 Cualidades

Una de las primeras cualidades que se puede mencionar del DW, es que maneja un gran volumen de datos, debido a que consolida en su estructura la información recolectada durante años, proveniente de diversas fuentes y áreas, en un solo lugar centralizado. Es por esta razón que el depósito puede ser soportado y mantenido sobre diversos medios de almacenamiento.

Además, como ya se ha mencionado, el almacén de datos presenta la información sumariada y agregada desde múltiples versiones, y maneja información histórica.

Organiza y almacena los datos que se necesitan para realizar consultas y procesos analíticos, con el propósito de responder a preguntas complejas y brindarles a los usuarios finales la posibilidad de que mediante una interface amigable, intuitiva y fácil de utilizar, puedan tomar decisiones sobre los datos sin tener que poseer demasiados conocimientos informáticos. El DW permite un acceso más directo, es decir, la información gira en torno al negocio, y es por ello que también los usuarios pueden sentirse cómodos al explorar los datos y encontrar relaciones complejas entre los mismos.

Cabe aclarar que el Data Warehousing no se compone solo de datos, ni tampoco solo se trata de un depósito de datos aislado. El Data Warehousing hace referencia a un conjunto de herramientas para consultar, analizar y presentar información, que permiten obtener o realizar análisis, reporting, extracción y explotación de los datos, con alta performance, para transformar dichos datos en información valiosa para la organización.

Con respecto a las tecnologías que son empleadas, se pueden encontrar las siguientes:

Arquitectura cliente/servidor.

- Técnicas avanzadas para replicar, refrescar y actualizar datos.
- Software front-end, para acceso y análisis de datos.

- Herramientas para extraer, transformar y cargar datos en el depósito, desde múltiples fuentes muy heterogéneas.
- Sistema de Gestión de Base de Datos (SGBD).

Todas las cualidades expuestas anteriormente, son imposibles de saldar en un típico ambiente operacional, y esto es una de las razones de ser del Data Warehousing.

3.2.5 Ventajas

A continuación se enumerarán algunas de las ventajas más sobresalientes que trae aparejada la implementación de un Data Warehousing y que ejemplifican de mejor modo sus características y cualidades:

- Transforma datos orientados a las aplicaciones en información orientada a la toma de decisiones.
- Integra y consolida diferentes fuentes de datos (internas y/o externas) y departamentos empresariales, que anteriormente formaban islas, en una única plataforma sólida y centralizada.
- Provee la capacidad de analizar y explotar las diferentes áreas de trabajo y de realizar un análisis inmediato de las mismas.
- Permite reaccionar rápidamente a los cambios del mercado.
- Aumenta la competitividad en el mercado.
- Elimina la producción y el procesamiento de datos que no son utilizados ni necesarios, producto de aplicaciones mal diseñadas o ya no utilizadas.
- Mejora la entrega de información, es decir, información completa, correcta, consistente, oportuna y accesible. Información que los usuarios necesitan, en el momento adecuado y en el formato apropiado.
- Logra un impacto positivo sobre los procesos de toma de decisiones. Cuando los usuarios tienen acceso a una mejor calidad de información, la empresa puede lograr por sí misma: aprovechar el enorme valor potencial de sus

recursos de información y transformarlo en valor verdadero; eliminar los retardos de los procesos que resultan de información incorrecta, inconsistente y/o inexistente; integrar y optimizar procesos a través del uso compartido e integrado de las fuentes de información; permitir al usuario adquirir mayor confianza acerca de sus propias decisiones y de las del resto, y lograr así, un mayor entendimiento de los impactos ocasionados.

- Aumento de la eficiencia de los encargados de tomar decisiones.
- Los usuarios pueden acceder directamente a la información en línea, lo que contribuye a su capacidad para operar con mayor efectividad en las tareas rutinarias o no. Además, pueden tener a su disposición una gran cantidad de valiosa información multidimensional, presentada coherentemente como fuente única, confiable y disponible en sus estaciones de trabajo. Así mismo, los usuarios tienen la facilidad de contar con herramientas que les son familiares para manipular y evaluar la información obtenida en el DW, tales como: hojas de cálculo, procesadores de texto, software de análisis de datos, software de análisis estadístico, reportes, tableros, etc.
- Permite la toma de decisiones estratégicas y tácticas.

3.2.6 Desventajas

A continuación se enumerarán algunas de las desventajas más comunes que se pueden presentar en la implementación de un Data Warehousing:

- Requiere una gran inversión, debido a que su correcta construcción no es tarea sencilla y consume muchos recursos, además, su misma implementación implica desde la adquisición de herramientas de consulta y análisis, hasta la capacitación de los usuarios.
- Existe resistencia al cambio por parte de los usuarios.
- Los beneficios del almacén de datos son apreciados en el mediano y largo plazo. Este punto deriva del anterior, y básicamente se refiere a que no todos los usuarios confiarán en el DW en una primera instancia, pero sí lo harán

una vez que comprueben su efectividad y ventajas. Además, su correcta utilización surge de la propia experiencia.

- Si se incluyen datos propios y confidenciales de clientes, proveedores, etc, el depósito de datos atentará contra la privacidad de los mismos, ya que cualquier usuario podrá tener acceso a ellos.
- Infravaloración de los recursos necesarios para la captura, carga y almacenamiento de los datos.
- Infravaloración del esfuerzo necesario para su diseño y creación.
- Incremento continuo de los requerimientos del usuario.
- Subestimación de las capacidades que puede brindar la correcta utilización del DW y de las herramientas de BI en general.

3.2.7 Redundancia

Debido a que el DW recibe información histórica de diferentes fuentes, sencillamente se podría suponer que existe una repetición de datos masiva entre el ambiente DW y el operacional. Por supuesto, este razonamiento es superficial y erróneo, de hecho, hay una mínima redundancia de datos entre ambos ambientes.

Para entender claramente lo antes expuesto, se debe considerar lo siguiente:

- Los datos del ambiente operacional se filtran antes de pertenecer al DW. Existen muchos datos que nunca ingresarán, ya que no conforman información necesaria o suficientemente relevante para la toma de decisiones.
- El horizonte de tiempo es muy diferente entre los dos ambientes.
- El almacén de datos contiene un resumen de la información que no se encuentra en el ambiente operacional.
- Los datos experimentan una considerable transformación, antes de ser cargados al DW. La mayor parte de los datos se alteran significativamente al ser seleccionados, consolidados y movidos al depósito.

En vista de estos factores, se puede afirmar que, la redundancia encontrada al cotejar los datos de ambos ambientes es mínima, ya que generalmente resulta en un porcentaje menor del 1 %.

3.2.8 Estructura

Los DW estructuran los datos de manera muy particular y existen diferentes niveles de esquematización y detalle que los delimitan

En la siguiente figura se puede apreciar mejor su respectiva estructura.

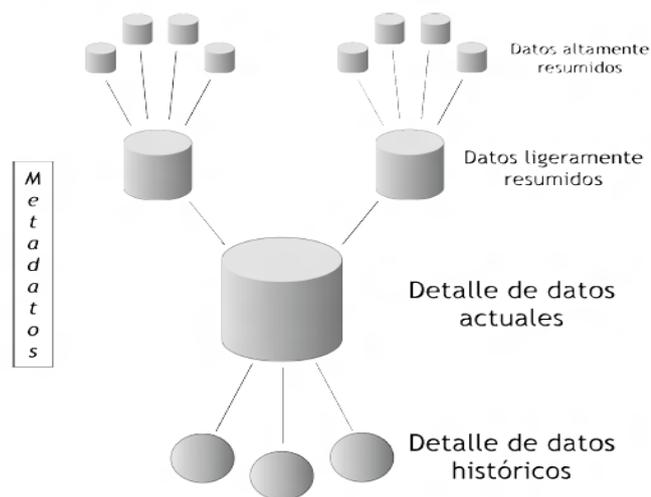


Figura 9: Estructura de datawarehouse, (Bernabeu, 2009)

Como se puede observar, los almacenes de datos están compuestos por diversos tipos de datos, que se organizan y dividen de acuerdo al nivel de detalle o granularidad que posean.

A continuación se explicarán cada uno de estos tipos de datos:

- **Detalle de datos actuales:** son aquellos que reflejan las ocurrencias más recientes. Generalmente se almacenan en disco, aunque su administración sea costosa y compleja, con el fin de conseguir que el acceso a la información sea sencillo y veloz, ya que son bastante voluminosos. Su gran tamaño se debe a que los datos residentes poseen el más bajo nivel de granularidad, o sea, se

almacenan a nivel de detalle. Por ejemplo, aquí es donde se guardaría el detalle de una venta realizada en tal fecha.

- **Detalle de datos históricos:** representan aquellos datos antiguos, que no son frecuentemente consultados. También se almacenan a nivel de detalle, normalmente sobre alguna forma de almacenamiento externa, ya que son muy pesados y en adición a esto, no son requeridos con mucha periodicidad. Este tipo de datos son consistentes con los de Detalle de datos actuales. Por ejemplo, en este nivel, al igual que en el anterior, se encontraría el detalle de una venta realizada en tal fecha, pero con la particularidad de que el día en que se registró la venta debe ser lo suficientemente antigua, para que se considere como histórica.
- **Datos ligeramente resumidos:** son los que provienen desde un bajo nivel de detalle y suman o agrupan los datos bajo algún criterio o condición de análisis. Habitualmente son almacenados en disco. Por ejemplo, en este caso se almacenaría la sumación del detalle de las ventas realizadas en cada mes.
- **Datos altamente resumidos:** son aquellos que compactan aún más a los datos ligeramente resumidos. Se guardan en disco y son muy fáciles de acceder. Por ejemplo, aquí se encontraría la sumación de las ventas realizadas en cada año.
- **Metadatos:** representan la información acerca de los datos. De muchas maneras se sitúa en una dimensión diferente al de otros datos del DW, ya que su contenido no es tomado directamente desde el ambiente operacional.

Estos diferentes niveles de detalle o granularidad, se obtienen a través de tablas de hechos agregados y/o pre agregadas.

3.2.8.1 Tablas de hecho agregadas y pre agregadas

Las tablas de hechos agregadas y pre agregadas se utilizan para almacenar un resumen de los datos, es decir, se guardan los datos en niveles de granularidad superior a los que inicialmente fueron obtenidos y/o gestionados.

Para obtener tablas agregadas o pre agregadas, es necesario establecer un criterio por el cual realizar el resumen. Por ejemplo, esto ocurre cuando se desea obtener información de ventas sumariadas por mes.

Cada vez que se requiere que los datos en una consulta se presenten en un nivel de granularidad superior al que se encuentran alojados en el Data Warehouse, se debe llevar a cabo un proceso de agregación.

El objetivo general de las tablas de hechos agregadas y pre agregadas es similar, pero cada una de ellas tiene una manera de operar diferente:

- Tablas de hechos agregadas: se generan luego de que se procesa la consulta correspondiente a la tabla de hechos que se resumirá. En general, la agregación se produce dinámicamente a través de una instrucción SQL que incluya sumalizaciones.
- Tablas de hechos pre agregadas: se generan antes de que se procese la consulta correspondiente a la tabla de hechos que se resumirá. De esta manera, la consulta se realiza contra una tabla que ya fue previamente sumariada. Habitualmente, estas sumalizaciones se calculan a través de procesos ETL.

Las tablas de hechos pre agregadas cuentan con los siguientes beneficios:

- Reduce la utilización de recursos de hardware que normalmente son incurridos en el cálculo de las sumalizaciones.
- Reduce el número de registros que serán analizados por el usuario.
- Reduce el tiempo utilizado en la generación de consultas por parte del usuario.

Las tablas de hechos pre agregadas son muy útiles en los siguientes casos generales:

- Cuando los datos a nivel detalle (menor nivel granular) son innecesarios y/o no son requeridos.
- Cuando una consulta sumariada a determinado nivel de granularidad es solicitado con mucha frecuencia.
- Cuando los datos son muy abundantes, y las consultas demoran en ser procesadas demasiado tiempo.

Como contrapartida, las tablas de hechos pre agregadas presentan una serie de desventajas:

- Requieren que se mantengan y gestionen nuevos procesos ETL.
- Demandan espacio de almacenamiento extra en el depósito de datos.

3.2.9 Flujo de Datos

El DW posee un flujo de datos estándar y generalizado, el cual puede apreciarse mejor en la siguiente figura.

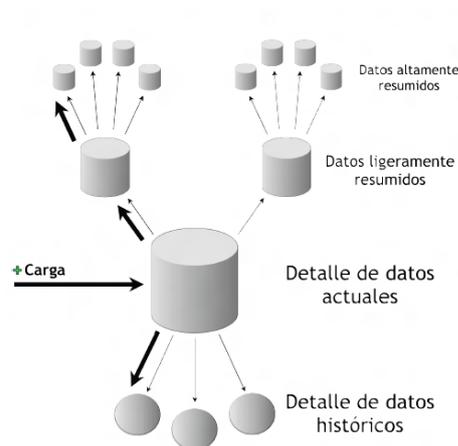


Figura 10: Flujo de datos del datawarehouse, (Bernabeu, 2009)

Cuando la información ingresa al depósito de datos se almacena a nivel de Detalle de datos actuales. Los datos permanecerán allí hasta que ocurra alguno de los tres eventos siguientes:

- Sean borrados del depósito de datos.
- Sean resumidos, ya sea a nivel de Datos ligeramente resumidos o a nivel de Datos altamente resumidos.
- Sean archivados a nivel de Detalle de datos históricos.

3.3 ARQUITECTURA DEL DATA WAREHOUSING

3.3.1 Introducción

En este punto y teniendo en cuenta que ya se han detallado claramente las características generales del Data Warehousing, se definirán y describirán todos los componentes que intervienen en su arquitectura o ambiente.

A través del siguiente gráfico se explicitará la estructura del Data Warehousing:

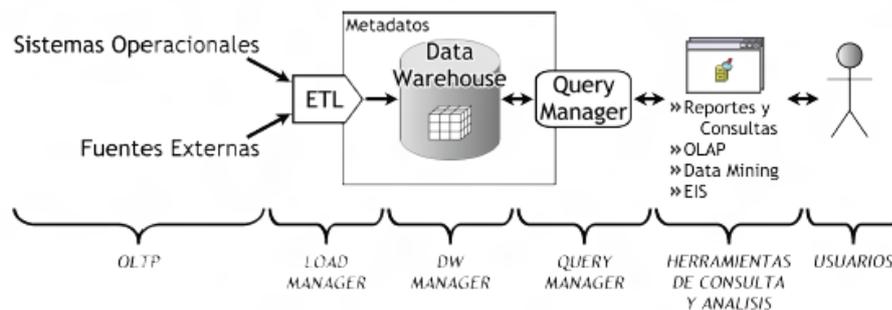


Figura 11: Arquitectura del datawarehousing, (Bernabeu, 2009)

Tal y como se puede apreciar, el ambiente está formado por diversos elementos que interactúan entre sí y que cumplen una función específica dentro del sistema. Por ello es que al abordar la exposición de cada elemento se lo hará en forma ordenada y teniendo en cuenta su relación con las demás partes.

Básicamente, la forma de operar del esquema superior se resume de la siguiente manera:

- Los datos son extraídos desde aplicaciones, bases de datos, archivos, etc. Esta información generalmente reside en diferentes tipos de sistemas, orígenes y arquitecturas y tienen formatos muy variados.
- Los datos son integrados, transformados y limpiados, para luego ser cargados en el DW.
- La información del DW se estructura en cubos multidimensionales, los cuales preparan esta información para responder a consultas dinámicas con una buena performance.
- Los usuarios acceden a los cubos multidimensionales del DW utilizando diversas herramientas de consulta, exploración, análisis, reportes, etc.

A continuación se detallará cada uno de los componentes de la arquitectura del Data Warehousing, teniendo como referencia siempre el gráfico antes expuesto, pero resaltando el tema que se tratará.

3.3.2 OLTP

OLTP (On Line Transaction Processing), representa toda aquella información transaccional que genera la empresa en su accionar diario, además, de las fuentes externas con las que puede llegar a disponer.

Como ya se ha mencionado, estas fuentes de información, son de características muy disímiles entre sí, en formato, procedencia, función, etc.

Entre los OLTP más habituales que pueden existir en cualquier organización se encuentran:

- Archivos de textos.
- Hipertextos.
- Hojas de cálculos.
- Informes semanales, mensuales, anuales, etc.
- Bases de datos transaccionales.

3.3.3 Load Manager

Para poder extraer los datos desde los OLTP, para luego manipularlos, integrarlos y transformarlos, para posteriormente cargar los resultados obtenidos en el DW, es necesario contar con algún sistema que se encargue de ello. Precisamente, los ETL (Extracción, Transformación y Carga) son los que cumplirán con tal fin.

Tal y como sus siglas lo indican, los ETL, extraen datos de las diversas fuentes que se requieran, los transforman para resolver posibles problemas de inconsistencias entre los mismos y finalmente, después de haberlos depurado se procede a su carga en el depósito de datos.

En síntesis, las funciones específicas de los ETL son tres:

- Extracción.
- Transformación.
- Carga.

A continuación, se detallará cada una de estas funciones, se expondrá cuál es el proceso que llevan a cabo los ETL y se enumerarán cuáles son sus principales tareas.

3.3.3.1 Extracción

Es aquí, en donde, basándose en las necesidades y requisitos del usuario, se exploran las diversas fuentes OLTP que se tengan a disposición, y se extrae la información que se considere relevante al caso.

Si los datos operacionales residen en un SGBD Relacional, el proceso de extracción se puede reducir a, por ejemplo, consultas en SQL o rutinas programadas. En cambio, si se encuentran en un sistema no convencional o fuentes externas, ya sean textuales, hipertextuales, hojas de cálculos, etc, la obtención de los mismos puede ser un tanto más dificultoso, debido a que, por ejemplo, se tendrán que realizar cambios de formato y/o volcado de información a partir de alguna herramienta específica.

Una vez que los datos son seleccionados y extraídos, se guardan en un almacenamiento intermedio, lo cual permite, entre otras ventajas:

- Manipular los datos sin interrumpir ni paralizar los OLTP, ni tampoco el DW.
- Almacenar y gestionar los metadatos que se generarán en los procesos ETL.
- Facilitar la integración de las diversas fuentes, internas y externas.

El almacenamiento intermedio constituye en la mayoría de los casos una base de datos en donde la información puede ser almacenada por ejemplo en tablas auxiliares, tablas temporales, etc. Los datos de estas tablas serán los que finalmente (luego de su correspondiente transformación) poblarán el DW.

3.3.3.2 Transformación

Esta función es la encargada de convertir aquellos datos inconsistentes en un conjunto de datos compatibles y congruentes, para que puedan ser cargados en el DW. Estas acciones se llevan a cabo, debido a que pueden existir diferentes fuentes de información, y es vital conciliar un formato y forma única, definiendo estándares, para que todos los datos que ingresarán al DW estén integrados.

Los casos más comunes en los que se deberá realizar integración, son los siguientes:

- Codificación.
- Medida de atributos.
- Convenciones de nombramiento.
- Fuentes múltiples.

Además de lo antes mencionado, esta función se encarga de realizar el proceso de Limpieza de Datos (Data Cleansing).

3.3.3.2.1 Codificación

Una inconsistencia muy típica que se encuentra al intentar integrar varias fuentes de datos, es la de contar con más de una forma de codificar un atributo en común. Por ejemplo, en el campo “estado”, algunos diseñadores completan su valor con “0” y “1”, otros con “Apagado” y “Encendido”, otros con “off” y “on”, etc. Lo que se debe realizar en estos casos, es seleccionar o recodificar estos atributos, para que cuando la información llegue al DW, esté integrada de manera uniforme.

En la siguiente figura, se puede apreciar que de varias formas de codificar se escoge una, entonces cuando surge una codificación diferente a la seleccionada, se procede a su transformación.

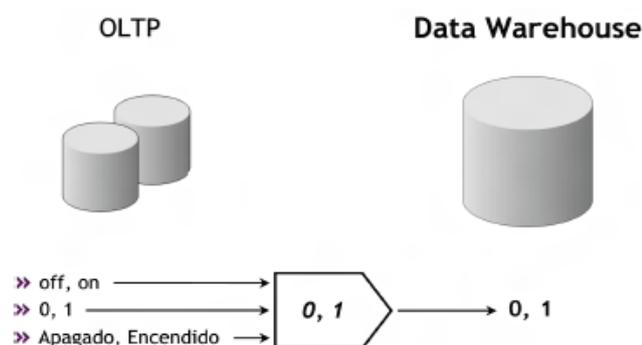


Figura 12: Transformación de un OLTP a un DW, (Bernabeu, 2009)

3.3.3.2.2 Medida de atributos

Los tipos de unidades de medidas utilizados para representar los atributos de una entidad, varían considerablemente entre sí, a través de los diferentes OLTP. Por ejemplo, al registrar la longitud de un producto determinado, de acuerdo a la aplicación que se emplee para tal fin, las unidades de medidas pueden ser explicitadas en centímetros, metros, pulgadas, etc.

En esta ocasión, se deberán estandarizar las unidades de medidas de los atributos, para que todas las fuentes de datos expresen sus valores de igual manera. Los algoritmos que resuelven estas inconsistencias son generalmente los más complejos.

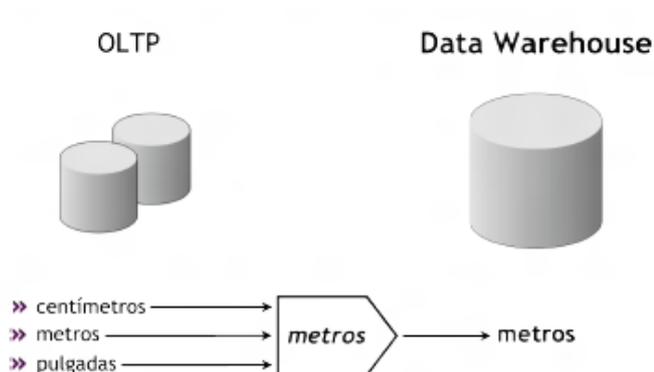


Figura 13: Transformación de medidas de OLTP a DW, (Bernabeu, 2009)

3.3.3.2.3 Convenciones de nombramiento

Usualmente, un mismo atributo es nombrado de diversas maneras en los diferentes OLTP. Por ejemplo, al referirse al nombre del proveedor, puede hacerse como “nombre”, “razón_social”, “proveedor”, etc. Aquí, se debe utilizar la convención de nombramiento que para el usuario sea más comprensible.



Figura 14: Convención del nombramiento, (Bernabeu, 2009)

3.3.3.2.4 Fuentes múltiples

Un mismo elemento puede derivarse desde varias fuentes. En este caso, se debe elegir aquella fuente que se considere más fiable y apropiada.

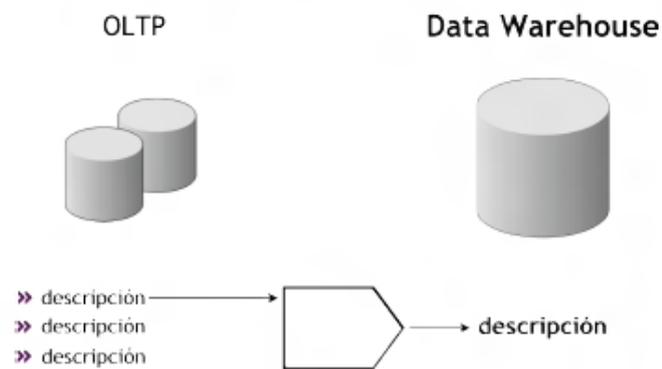


Figura 15: Múltiples fuentes, (Bernabeu, 2009)

3.3.3.2.5 Limpieza de datos

Su objetivo principal es el de realizar distintos tipos de acciones contra el mayor número de datos erróneos, inconsistentes e irrelevantes.

- Las acciones más típicas que se pueden llevar a cabo al encontrarse con Datos Anómalos (Outliers) son:
 - Ignorarlos.
 - Eliminar la columna.
 - Filtrar la columna.
 - Filtrar la fila errónea, ya que a veces su origen, se debe a casos especiales.
 - Reemplazar el valor.
 - Discretizar los valores de las columnas. Por ejemplo de 1 a 2, poner “bajo”; de 3 a 7, “óptimo”; de 8 a 10, “alto”. Para que los outliers caigan en “bajo” o en “alto” sin mayores problemas.

- Las acciones que suelen efectuarse contra Datos Faltantes (Missing Values) son:
 - Ignorarlos.

- Eliminar la columna.
- Filtrar la columna.
- Filtrar la fila errónea, ya que a veces su origen, se debe a casos especiales.
- Reemplazar el valor.
- Esperar hasta que los datos faltantes estén disponibles.

Un punto muy importante que se debe tener en cuenta al elegir alguna acción, es el de identificar el porqué de la anomalía, para luego actuar en consecuencia, con el fin de evitar que se repitan, agregándole de esta manera más valor a los datos de la organización. Se puede dar que en algunos casos, los valores faltantes sean inexistentes, ya que por ejemplo, un nuevo asociado o cliente, no poseerá consumo medio del último año.

3.3.3.3 Carga

Este proceso es el responsable de cargar la estructura de datos del DW con:

- Aquellos datos que han sido transformados y que residen en el almacenamiento intermedio.
- Aquellos datos de los OLTP que tienen correspondencia directa con el depósito de datos.

Se debe tener en cuenta, que los datos antes de moverse al almacén de datos, deben ser analizados con el propósito de asegurar su calidad, ya que este es un factor clave, que no debe dejarse de lado.

3.3.3.4 Proceso ETL

A continuación, se explicará en síntesis el accionar del proceso ETL, y cuál es la relación existente entre sus diversas funciones. En la siguiente figura se puede apreciar mejor lo antes descrito:

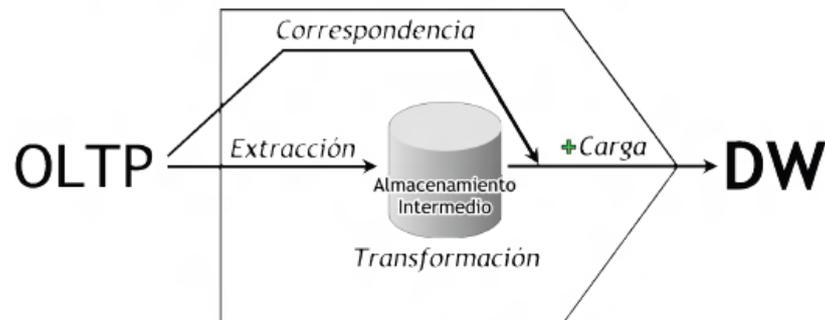


Figura 16: Proceso de un ETL, (Bernabeu, 2009)

Los pasos que se siguen son:

- Se extraen los datos relevantes desde los OLTP y se depositan en un almacenamiento intermedio.
- Se integran y transforman los datos, para evitar inconsistencias.
- Se cargan los datos desde el almacenamiento intermedio hasta el DW. Si existiesen correspondencias directas entre datos de los OLTP y el DW, se procede también a su respectiva carga.

3.3.3.5 Tareas del ETL

Los ETL, son los encargados de realizar dos tareas bien definidas:

- Carga Inicial (Initial Load).
- Actualización, mantenimiento o refresco periódico (siempre teniendo en cuenta un intervalo de tiempo predefinido para tal operación).

La carga inicial, se refiere precisamente a la primera carga de datos que se le realizará al DW. Por lo general, esta tarea consume un tiempo bastante considerable, ya que se deben insertar registros que han sido generados aproximadamente, y en casos ideales, durante más de cinco años.

Los mantenimientos periódicos mueven pequeños volúmenes de datos, y su frecuencia está dada en función del gránulo del DW y los requerimientos del usuario. El objetivo de esta tarea es añadir al depósito aquellos datos nuevos que se fueron generando desde el último refresco.

Antes de realizar una nueva actualización, es necesario identificar si se han producido cambios en las fuentes originales de los datos recogidos, desde la fecha del último mantenimiento, a fin de no atentar contra la consistencia del DW.

Para efectuar esta operación, se pueden realizar las siguientes acciones:

- Cotejar las instancias de los OLTP involucrados.
- Utilizar disparadores en los OLTP.
- Recurrir a Marcas de Tiempo (Time Stamp), en los registros de los OLTP.
- Comparar los datos existentes en los dos ambientes (OLTP y DW).
- Hacer uso de técnicas mixtas.

Si este control consume demasiado tiempo y esfuerzo, o simplemente no puede llevarse a cabo por algún motivo en particular, existe la posibilidad de cargar el DW desde cero, este proceso se denomina Carga Total (Full Load).

3.3.4 Data Warehouse Manager

El DW Manager presenta las siguientes características y funciones:

- Transforma e integra los datos fuentes y del almacenamiento intermedio en un modelo adecuado para la toma de decisiones.
- Gestiona el depósito de datos a través de tablas de hechos y tablas de dimensiones, y lo organiza en torno a una base de datos multidimensional. Esto permite que se puedan crear cubos multidimensionales.
- Permite realizar todas las funciones de definición y manipulación del depósito de datos, para poder soportar todos los procesos de gestión del mismo.

- Es el encargado de ejecutar y definir las políticas de particionamiento. El objetivo de realizar esto, es conseguir una mayor eficiencia y performance en las consultas al no tener que manejar todo el grueso de los datos. Esta política debe aplicarse sobre la tabla de hechos que, como se explicará más adelante, es en la que se almacena toda la información que será analizada.
- Realiza copias de resguardo incrementales o totales de los datos del DW.
- Se constituye típicamente al combinar un SGBD con software y aplicaciones dedicadas.
- Posee un repositorio de datos propio.
- Gestiona y mantiene metadatos.

3.3.4.1 Base de datos multidimensional

Las bases de datos multidimensionales, proveen una estructura que permite, a través de la creación y consulta a un cubo multidimensional, tener acceso flexible a los datos, para explorar y analizar sus relaciones, y consiguientes resultados.

Las bases de datos multidimensionales implican tres variantes posibles de modelamiento, que permiten realizar consultas de soporte de decisión:

- Esquema en estrella (Star Scheme).
- Esquema copo de nieve (Snowflake Scheme).
- Esquema constelación o copo de estrellas (Starflake Scheme).

Los mencionados esquemas pueden ser implementados de diversas maneras, que, independientemente al tipo de arquitectura, requieren que toda la estructura de datos este desnormalizada o semi desnormalizada, para evitar desarrollar uniones

(Join) complejas para acceder a la información, con el fin de agilizar la ejecución de consultas. Los diferentes tipos de implementación son los siguientes:

- Relacional – ROLAP.
- Multidimensional – MOLAP.
- Híbrido – HOLAP.

3.3.4.2 Tablas de dimensiones

Las tablas de dimensiones definen como están los datos organizados lógicamente y proveen el medio para analizar el contexto del negocio.

Representan los aspectos de interés, mediante los cuales el usuario podrá filtrar y manipular la información almacenada en la tabla de hechos.

En la siguiente figura se pueden apreciar algunos ejemplos:

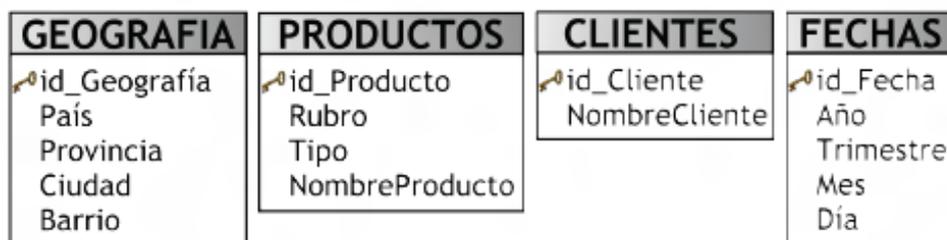


Figura 17: Tablas de dimensiones, (Bernabeu, 2009)

Como se puede observar, cada tabla posee un identificador único y al menos un campo o dato de referencia que describe los criterios de análisis relevantes para la organización, estos son por lo general de tipo texto.

Los datos dentro de estas tablas, que proveen información del negocio o que describen alguna de sus características, son llamados datos de referencia.

Más detalladamente, cada tabla de dimensión podrá contener los siguientes campos:

- Clave principal o identificador único.
- Clave foráneas.

- Datos de referencia primarios: datos que identifican la dimensión. Por ejemplo: nombre del cliente.
- Datos de referencia secundarios: datos que complementan la descripción de la dimensión. Por ejemplo: e-mail del cliente, fax del cliente, etc.

Usualmente la cantidad de tablas de dimensiones, aplicadas a un tema de interés en particular, varían entre tres y quince.

Debe tenerse en cuenta, que no siempre la clave primaria del OLTP, se corresponde con la clave primaria de la tabla de dimensión relacionada. Es recomendable manejar un sistema de claves en el DW totalmente diferente al de los OLTP, ya que si estos últimos son recodificados, el almacén quedaría inconsistente y debería ser poblado nuevamente en su totalidad.

3.3.4.2.1 Tablas de dimensión tiempo

En un DW, la creación y el mantenimiento de una tabla de dimensión Tiempo es obligatoria y la definición de granularidad y estructuración de la misma depende de la dinámica del negocio que se esté analizando. Toda la información dentro del depósito, como ya se ha explicado, posee su propio sello de tiempo que determina la ocurrencia de un hecho específico, representando de esta manera diferentes versiones de una misma situación.

Es importante tener en cuenta que el tiempo no es solo una secuencia cronológica representada de forma numérica, sino que posee fechas especiales que inciden notablemente en las actividades de la organización. Esto se debe a que los usuarios podrán por ejemplo analizar las ventas realizadas teniendo en cuenta el día de la semana en que se produjeron, quincena, mes, trimestre, semestre, año, estación, etc.

Existen muchas maneras de diseñar esta tabla, y en adición a ello no es una tarea sencilla de llevar a cabo. Por estas razones se considera una buena práctica

evaluar con cuidado la temporalidad de los datos, la forma en que trabaja la organización, los resultados que se esperan obtener del almacén de datos relacionados con una unidad de tiempo y la flexibilidad que se desea obtener de dicha tabla.

Si bien, el lenguaje SQL ofrece funciones del tipo DATE, en la tabla de dimensión Tiempo, se modelan y presentan datos temporales que no pueden calcularse a través de consultas SQL, lo cual le añade una ventaja más.

3.3.4.3 Tablas de hechos

Las tablas de hechos contienen, precisamente, los hechos que serán utilizados por los analistas de negocio para apoyar el proceso de toma de decisiones.

Los hechos son datos instantáneos en el tiempo, que son filtrados, agrupados y explorados a través de condiciones definidas en las tablas de dimensiones.

Los datos presentes en las tablas de hechos constituyen el volumen de la bodega, y pueden estar compuestos por millones de registros dependiendo de su granularidad y antigüedad de la organización. Los más importantes son los de tipo numérico.

El registro del hecho posee una clave primaria que está compuesta por las claves primarias de las tablas de dimensiones relacionadas a este.

En la siguiente imagen se puede apreciar un ejemplo de lo antes mencionado:

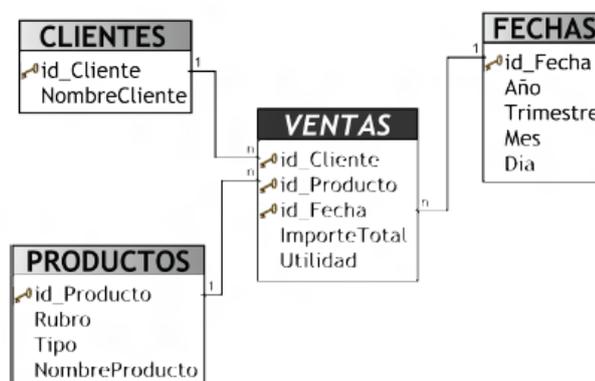


Figura 18: Tabla de hechos con dimensiones, (Bernabeu, 2009)

Como se muestra en la figura anterior, la tabla de hechos “VENTAS” se ubica en el centro, e irradiando de ella se encuentran las tablas de dimensiones “CLIENTES”, “PRODUCTOS” y “FECHAS”, que están conectadas mediante sus claves primarias. Es por ello que la clave primaria de la tabla de hechos es la combinación de las claves primarias de sus dimensiones. Los hechos en este caso son “ImporteTotal” y “Utilidad”.

A continuación, se entrará más en detalle sobre la definición de un hecho, también llamado dato agregado:

- Los hechos son aquellos datos que residen en una tabla de hechos y que son utilizados para crear indicadores, a través de sumalizaciones preestablecidas al momento de crear un cubo multidimensional. Debido a que una tabla de hechos se encuentra interrelacionada con sus respectivas tablas de dimensiones, permite que los hechos puedan ser accedidos, filtrados y explorados por los valores de los campos de estas tablas de dimensiones, obteniendo de este modo una gran capacidad analítica.

Las sumalizaciones no están referidas solo a sumas, sino también a promedios, mínimos, máximos, totales por sector, porcentajes, fórmulas predefinidas, etc, dependiendo de los requerimientos de información del negocio.

Para ejemplificar este nuevo concepto de hechos, se enumerarán algunos que son muy típicos y fáciles de comprender:

- $\text{ImporteTotal} = \text{precioProducto} * \text{cantidadVendida}$
- $\text{Rentabilidad} = \text{utilidad} / \text{PN}$
- $\text{CantidadVentas} = \text{cantidad}$
- $\text{PromedioGeneral} = \text{AVG}(\text{notasFinales})$

A la izquierda de la igualdad se encuentran los hechos; a la derecha los campos de los OLTP que son utilizados para representarlos. En el último ejemplo se realiza un pre cálculo para establecer el hecho.

Existen dos tipos de hechos, los básicos y los derivados, a continuación se detallarán cada uno de ellos, teniendo en cuenta para su ejemplificación la siguiente tabla de hechos:

HECHOS	
id_Dimensión1	
id_Dimensión2	
id_DimensiónN	
precio	
cantidad	
total	

Figura 19: Tabla de hechos, (Bernabeu, 2009)

- Hechos básicos: son los que se encuentran representados por un campo de una tabla de hechos. Los campos “precio” y “cantidad” de la tabla anterior son hechos básicos.
- Hechos derivados: son los que se forman al combinar uno o más hechos con alguna operación matemática o lógica y que también residen en una tabla de hechos. Estos poseen la ventaja de almacenarse previamente calculados, por lo cual pueden ser accedidos a través de consultas SQL sencillas y devolver resultados rápidamente, pero requieren más espacio físico en el DW, además de necesitar más tiempo de proceso en los ETL que los calculan. El campo “total” de la tabla anterior en un hecho derivado, ya que se conforma de la siguiente manera:
 - $\text{total} = \text{precio} * \text{cantidad}$

Los campos “precio” y “cantidad”, también pertenecen a la tabla “HECHOS”. Cabe resaltar, que no es necesario que los hechos derivados se compongan únicamente con hechos pertenecientes a una misma tabla.

Los hechos son gestionados con el principal objetivo de que se creen indicadores basados en ellos, para posteriormente incluir estos últimos en algún cubo multidimensional.

3.3.4.4 Cubo multidimensional: introducción

Un cubo multidimensional o hipercubo, representa o convierte los datos planos que se encuentran en filas y columnas, en una matriz de N dimensiones.

Los objetos más importantes que se pueden incluir en un cubo multidimensional, son los siguientes:

- **Indicadores:** sumalizaciones que se efectúan sobre algún hecho, perteneciente a una tabla de hechos.
- **Atributos:** campos o criterios de análisis, pertenecientes a tablas de dimensiones.
- **Jerarquías:** representa una relación lógica entre dos o más atributos.

De esta manera en un cubo multidimensional, los atributos existen a lo largo de varios ejes o dimensiones, y la intersección de las mismas representa el valor que tomará el indicador que se está evaluando.

En la siguiente representación matricial se puede ver más claramente lo que se acaba de decir.

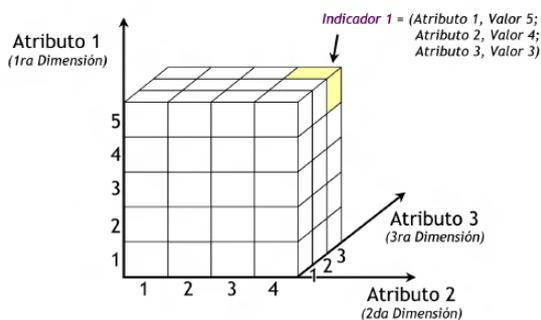


Figura 20: Cubo multidimensional, (Bernabeu, 2009)

Para la creación del cubo de la figura anterior, se definieron tres Atributos (“Atributo 1”, “Atributo 2” y “Atributo 3”) y se definió un Indicador (“Indicador 1”). Entonces el cubo quedo compuesto por 3 dimensiones o ejes (una por cada Atributo),

cada una con sus respectivos valores asociados. También, se ha seleccionado una intersección al azar para demostrar la correspondencia con los valores de las Atributos. En este caso, el indicador “Indicador 1”, representa el cruce del Valor “5” de “Atributo 1”, con el Valor “4” de “Atributo 2” y con el Valor “3” de “Atributo 3”.

Se puede observar, que el resultado del análisis está dado por los cruces matriciales de acuerdo a los valores de las dimensiones seleccionadas.

Más específicamente, para acceder a los datos del DW, se deben ejecutar las consultas sobre algún cubo multidimensional previamente definido. Dicho cubo debe incluir entre otros objetos: indicadores, atributos, jerarquías, etc, basados en los campos de las tablas de dimensiones y de hechos, que se deseen analizar. De esta manera, las consultas son respondidas con gran performance, minimizando al máximo el tiempo que se hubiese incurrido en realizar dicha consulta sobre una base de datos transaccional.

3.3.4.4.1 Indicadores

Los indicadores son sumalizaciones efectuadas sobre algún hecho, que serán incluidos en algún cubo multidimensional, con el fin de analizar los datos almacenados en el DW. El valor que estos adopten estará condicionado por los atributos/jerarquías que se utilicen para analizarlos.

Los indicadores, además de hechos, pueden estar compuestos por otros indicadores, pero no ambos simultáneamente. Pueden utilizarse para su creación funciones de sumalización (suma, conteo, promedio, etc), funciones matemáticas, estadísticas, operadores matemáticos y lógicos.

3.3.4.4.2 Atributos

Los atributos constituyen los criterios de análisis que se utilizarán para analizar los indicadores dentro de un cubo multidimensional. Los mismos se basan, en su gran mayoría, en los campos o datos de referencia de las tablas de dimensiones.

Dentro de un cubo multidimensional, los atributos son los ejes del mismo.

3.3.4.4.3 Jerarquías

Una jerarquía representa una relación lógica entre dos o más atributos pertenecientes a un cubo multidimensional; siempre y cuando posean su correspondiente relación “padre-hijo”.

Las jerarquías poseen las siguientes características:

- Pueden existir varias en un mismo cubo.
- Están compuestas por dos o más niveles.
- Se tiene una relación “1-n” o “padre-hijo” entre atributos consecutivos de un nivel superior y uno inferior.

Por lo general, las jerarquías pueden identificarse fácilmente, debido a que existen relaciones “1-n” o “padre-hijo” entre los propios atributos de un cubo.

La principal ventaja de manejar jerarquías, reside en poder analizar los datos desde su nivel más general al más detallado y viceversa, al desplazarse por los diferentes niveles.

La siguiente figura muestra un pequeño ejemplo de lo recién expuesto:

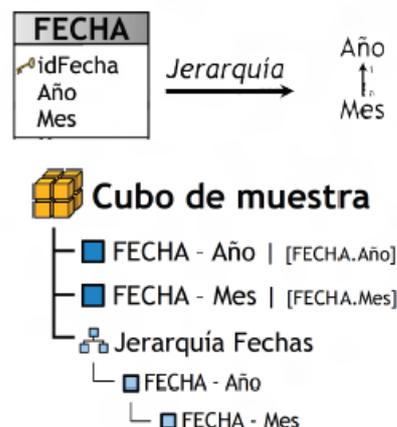


Figura 21: Jerarquía de dimensión fecha, (Bernabeu, 2009)

Aquí, se puede apreciar claramente cómo se construye una jerarquía:

- 1) Se crearon dos atributos, “FECHA - Año” y “FECHA - Mes”, los cuales están constituidos de la siguiente manera:
 - “FECHA - Año” = FECHA.Año
 - “FECHA - Mes” = FECHA.Mes

A la izquierda de la igualdad se encuentra el nombre del atributo; a la derecha el nombre del campo de la tabla de dimensión “FECHA”.

- 2) Se creó la jerarquía llamada “Jerarquía Fechas”, en la cual se colocó el atributo más general en la cabecera y se comenzó a disgregar los niveles hacia abajo. En este caso, la figura se explica como sigue:
 - Un mes del año pertenece solo a un año. Un año puede poseer uno o más meses del año.

3.3.4.4 Relación

Una relación representa la forma en que dos atributos interactúan dentro de una jerarquía. Existen básicamente dos tipos de relaciones:

- Explícitas: son las más comunes y se pueden modelar a partir de atributos directos y están en línea continua de una jerarquía, por ejemplo, un país posee una o más provincias y una provincia pertenece solo a un país.
- Implícitas: son las que ocurren en la vida real, pero su relación no es de vista directa, por ejemplo, una provincia tiene uno o más ríos, pero un río pertenece a una o más provincias. Como se puede observar, este caso se trata de una relación muchos a muchos.

3.3.4.4.5 Granularidad

La granularidad representa el nivel de detalle al que se desea almacenar la información sobre el negocio que se esté analizando. Por ejemplo, los datos referentes a ventas o compras realizadas por una empresa, pueden registrarse día a día, en cambio, los datos pertinentes a pagos de sueldos o cuotas de socios, podrán almacenarse a nivel de mes.

Mientras mayor sea el nivel de detalle de los datos, se tendrán mayores posibilidades analíticas, ya que los mismos podrán ser resumidos o sumariados. Es decir, los datos que posean granularidad fina (nivel de detalle) podrán ser resumidos hasta obtener una granularidad media o gruesa. No sucede lo mismo en sentido contrario, ya que por ejemplo, los datos almacenados con granularidad media podrán resumirse, pero no tendrán la facultad de ser analizados a nivel de detalle. O sea, si la granularidad con que se guardan los registros es a nivel de día, estos datos podrán sumariarse por semana, mes, semestre y año, en cambio, si estos registros se almacenan a nivel de mes, podrán sumariarse por semestre y año, pero no lo podrán hacer por día y semana.

3.3.4.5 Tipos de modelamiento de un Data Warehouse

3.3.4.5.1 Esquema en estrella

El esquema en estrella, consta de una tabla de hechos central y de varias tablas de dimensiones relacionadas a esta, a través de sus respectivas claves. En la siguiente figura se puede apreciar un esquema en estrella estándar:

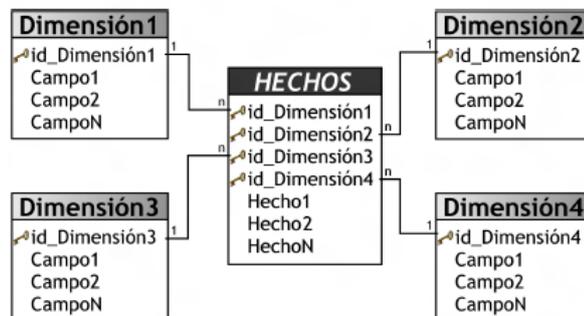


Figura 22: Esquema estrella, (Bernabeu, 2009)

El esquema en estrella es el más simple de interpretar y optimiza los tiempos de respuesta ante las consultas de los usuarios. Este modelo es soportado por casi todas las herramientas de consulta y análisis, y los metadatos son fáciles de documentar y mantener, sin embargo es el menos robusto para la carga y es el más lento de construir.

A continuación se destacarán algunas características de este modelo, que ayudarán a comprender mejor el porqué de sus ventajas:

- Posee los mejores tiempos de respuesta.
- Su diseño es fácilmente modificable.
- Existe paralelismo entre su diseño y la forma en que los usuarios visualizan y manipulan los datos.
- Simplifica el análisis.
- Facilita la interacción con herramientas de consulta y análisis.

3.3.4.5.2 Esquema copo de nieve

Este esquema representa una extensión del modelo en estrella cuando las tablas de dimensiones se organizan en jerarquías de dimensiones.

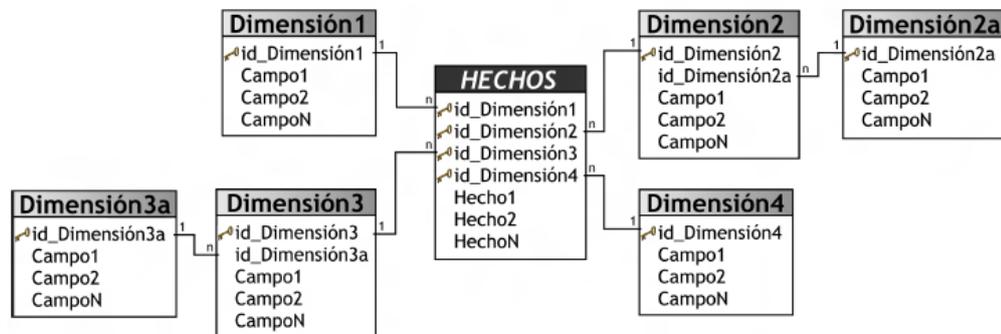


Figura 23: Esquema copo de nieve, (Bernabeu, 2009)

Como se puede apreciar en la figura anterior, existe una tabla de hechos central que está relacionada con una o más tablas de dimensiones, quienes a su vez pueden estar relacionadas o no con una o más tablas de dimensiones.

Este modelo es más cercano a un modelo de entidad relación, que al modelo en estrella, debido a que sus tablas de dimensiones están normalizadas.

Una de los motivos principales de utilizar este tipo de modelo, es la posibilidad de segregarse los datos de las tablas de dimensiones y proveer un esquema que sustente los requerimientos de diseño. Otra razón es que es muy flexible y puede implementarse después de que se haya desarrollado un esquema en estrella.

Se pueden definir las siguientes características de este tipo de modelo:

- Posee mayor complejidad en su estructura.
- Hace una mejor utilización del espacio.
- Es muy útil en tablas de dimensiones de muchas tuplas.
- Las tablas de dimensiones están normalizadas, por lo que requiere menos esfuerzo de diseño.
- Puede desarrollar clases de jerarquías fuera de las tablas de dimensiones, que permiten realizar análisis de lo general a lo detallado y viceversa.

A pesar de todas las características y ventajas que trae aparejada la implementación del esquema copo de nieve, existen dos grandes inconvenientes de ello:

- Si se poseen múltiples tablas de dimensiones, cada una de ellas con varias jerarquías, se creará un número de tablas bastante considerable, que pueden llegar al punto de ser inmanejables.
- Al existir muchas uniones y relaciones entre tablas, el desempeño puede verse reducido.

3.3.4.5.3 Esquema constelación

Este modelo está compuesto por una serie de esquemas en estrella, y tal como se puede apreciar en la siguiente figura, está formado por una tabla de hechos principal (“HECHOS_A”) y por una o más tablas de hechos auxiliares (“HECHOS_B”), las cuales pueden ser sumalizaciones de la principal. Dichas tablas yacen en el centro del modelo y están relacionadas con sus respectivas tablas de dimensiones.

No es necesario que las diferentes tablas de hechos compartan las mismas tablas de dimensiones, ya que, las tablas de hechos auxiliares pueden vincularse con solo algunas de las tablas de dimensiones asignadas a la tabla de hechos principal, y también pueden hacerlo con nuevas tablas de dimensiones.

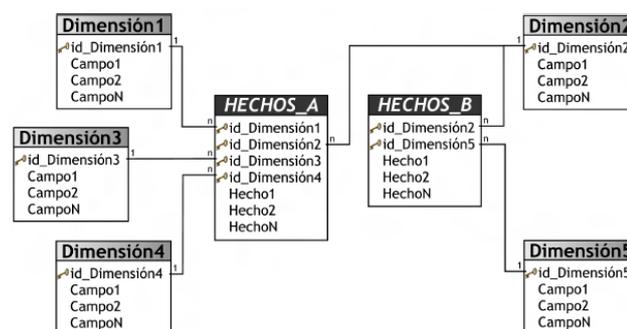


Figura 24: Esquema constelación, (Bernabeu, 2009)

Su diseño y cualidades son muy similares a las del esquema en estrella, pero posee una serie de diferencias con el mismo, que son precisamente las que lo destacan y caracterizan. Entre ellas se pueden mencionar:

- Permite tener más de una tabla de hechos, por lo cual se podrán analizar más aspectos claves del negocio con un mínimo esfuerzo adicional de diseño.
- Contribuye a la reutilización de las tablas de dimensiones, ya que una misma tabla de dimensión puede utilizarse para varias tablas de hechos.
- No es soportado por todas las herramientas de consulta y análisis.

3.3.4.6 Cubo Multidimensional: creación y ejemplificación

Ahora que ya se tiene una visión general de los tipos de modelamiento e implementación de un DW, se volverá a abordar el tema de los cubos multidimensionales, pero esta vez se hará énfasis en su construcción y se ejemplificará cada paso, a fin de que se puedan visualizar mejor los resultados de cada acción.

La forma que se utilizará para graficar el cubo que se creará, será la siguiente:

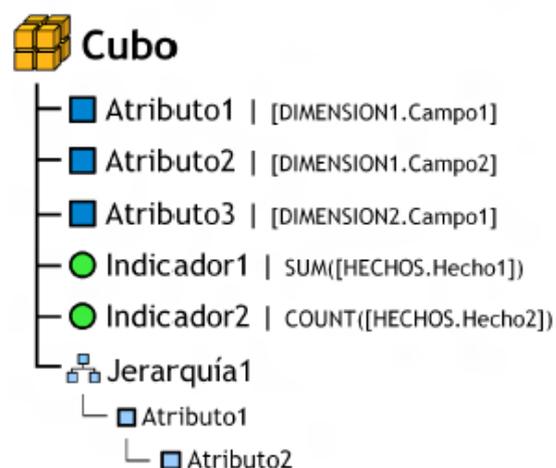


Figura 25: Estructura de un cubo multidimensional, (Bernabeu, 2009)

Tal y como se puede observar, el gráfico toma una estructura de árbol, en la cual en la raíz figura el cubo en cuestión y dependiendo de este sus diferentes objetos relacionados. En el caso de las jerarquías, los atributos que la componen, también

deben estructurarse en forma de árbol, teniendo en cuenta su respectiva relación padre-hijo.

Se tomará como base para la realización de los ejemplos, el siguiente esquema en estrella:



Figura 26: Ejemplo de un esquema estrella, (Bernabeu, 2009)

Como primer paso se creará un cubo multidimensional llamado “Cubo de Ventas”.

Luego se crearán dos atributos:

- De la tabla de dimensión “PRODUCTOS”, se tomará el campo “Producto” para la creación del atributo denominado:
 - “PRODUCTOS - Producto”.
- De la tabla dimensión “MARCAS”, se tomará el campo “Marca” para la creación del atributo denominado:
 - “MARCAS - Marca”.



Figura 27: Cubo de ventas con dos dimensiones, (Bernabeu, 2009)

También se creará un indicador:

- De la tabla de hechos “VENTAS”, se sumará el hecho “Venta” para crear el indicador denominado:
 - “VENTAS - Venta”.

La fórmula utilizada para crear este indicador es la siguiente:

- “VENTAS - Venta” = SUM(VENTAS.Venta).

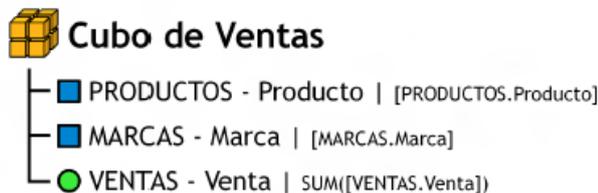


Figura 28: Cubo ventas con dimensiones, indicadores, (Bernabeu, 2009)

En este momento, se tiene un cubo multidimensional de dos dimensiones, cuya representación matricial sería la siguiente:

PRODUCTOS - Producto	P1	40	25	60
	P2	21	55	45
	P3	13	32	43
		M1	M2	M3
		MARCAS - Marca		

Figura 29: Matriz de las dos dimensiones, (Bernabeu, 2009)

Este cubo posee dos ejes o dimensiones, “PRODUCTOS - Producto” y “MARCAS - arca”. La intersección de los ejes representa las ventas de cada producto con su respectiva marca (indicador “VENTAS - Venta”).

Por ejemplo:

- Las ventas asociadas al producto “P1” y a la marca “M1” son 40.
- Las ventas asociadas al producto “P1” y a la marca “M2” son 25.
- Las ventas asociadas al producto “P1” y a la marca “M3” son 60.

Ahora, al cubo planteado se le agregará un nuevo atributo:

- De la tabla de dimensión “CLIENTES”, se tomará el campo “Cliente” para la creación del atributo denominado:
 - “CLIENTES - Cliente”.

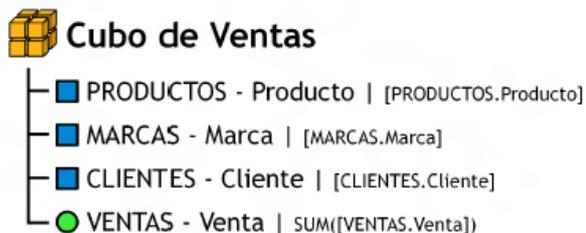


Figura 30: Cubo ventas con 3 dimensiones y 1 indicador, (Bernabeu, 2009)

De esta manera, ahora se tiene un cubo multidimensional de tres dimensiones, cuya representación matricial sería la siguiente:

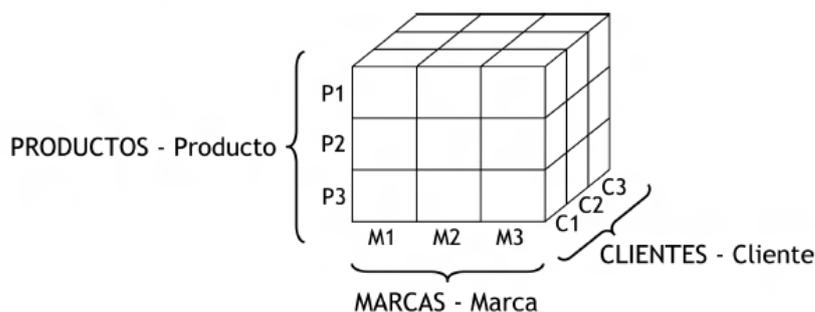


Figura 31: Estructura del cubo de ventas, (Bernabeu, 2009)

En este caso los valores del indicador “VENTAS - Venta” están dados de acuerdo a las ventas de cada producto, de cada marca, a cada cliente.

Para finalizar, se añadirá un cuarto atributo al cubo:

- De la tabla de dimensión “TIEMPO”, se tomará el campo “Año” para la creación del atributo denominado:
 - “TIEMPO - Año”.

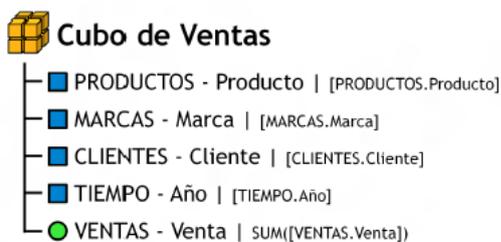


Figura 32: Cubo ventas 4 dimensiones y 1 indicador, (Bernabeu, 2009)

Entonces, la representación matricial del cubo multidimensional resultante sería la siguiente:

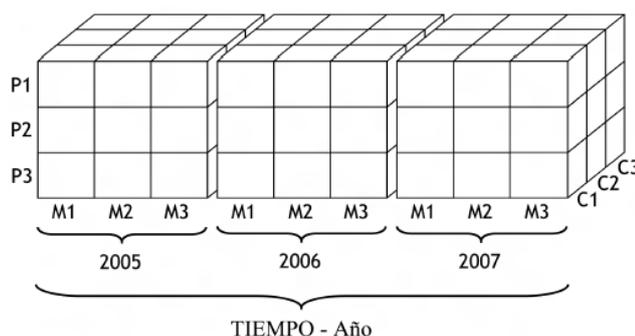


Figura 33: Cubo de ventas agrupadas por años, (Bernabeu, 2009)

Los valores del indicador “VENTAS - Venta” en este momento, estarán condicionados por las ventas de cada producto, de cada marca, a cada cliente, en cada año.

Esta última imagen, demuestra claramente los conceptos expuestos de la tabla de dimensión tiempo, donde se decía que pueden existir diferentes versiones de la situación del negocio.

Cabe aclarar que pueden crearse tantos cubos se deseen y que los mismos pueden coexistir sin ningún inconveniente.

3.3.5 Data Mart

En adición al DW principal pueden existir varios Data Mart (DM) departamentales. Un DM es la implementación de un DW con alcance restringido a un área funcional, problema en particular, departamento, tema o grupo de necesidades.

Muchos depósitos de datos comienzan siendo Data Mart, para, entre otros motivos, minimizar riesgos y producir una primera entrega en tiempos razonables. Pero, una vez que estos se han implementado exitosamente, su ámbito se irá ampliando paulatinamente.

De acuerdo a las operaciones que se deseen o requieran desarrollar, los DM pueden adoptar las siguientes arquitecturas:

- Top-Down: primero se define el DW y luego se desarrollan, construyen y cargan los DM a partir del mismo. En la siguiente figura se encuentra detallada esta arquitectura:

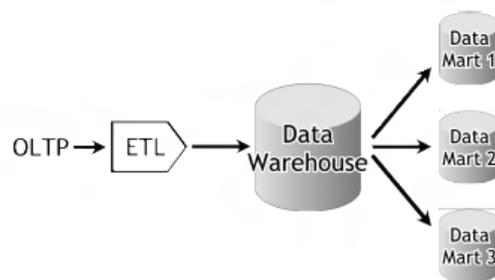


Figura 34: Arquitectura TOP-DOWN, (Bernabeu, 2009)

Como se puede apreciar, el DW es cargado a través de procesos ETL y luego este alimenta a los diferentes DM, cada uno de los cuales recibirá los datos que correspondan al tema o departamento que traten. Esta forma de implementación cuenta con la ventaja de no tener que incurrir en complicadas sincronizaciones de hechos, pero requiere una gran inversión y una gran cantidad de tiempo de construcción.

- Bottom-Up: en esta arquitectura, se definen previamente los DM y luego se integran en un DW centralizado. La siguiente figura presenta esta implementación:

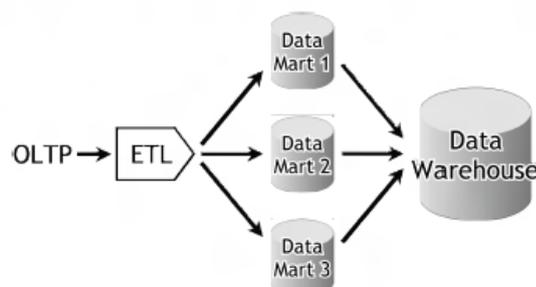


Figura 35: Arquitectura BOTTOM-UP, (Bernabeu, 2009)

Los DM se cargan a través de procesos ETL, los cuales suministrarán la información adecuada a cada uno de ellos. En muchas ocasiones, los DM son implementados sin que exista el DW, ya que tienen sus mismas características pero con la particularidad de que están enfocados en un tema específico. Luego de que hayan sido creados y cargados todos los DM, se procederá a su integración con el depósito. La ventaja que trae aparejada este modelo es que cada DM se crea y pone en funcionamiento en un corto lapso de tiempo y se puede tener una pequeña solución a un costo no tan elevado. Luego que todos los DM estén puestos en marcha, se puede decidir si construir el DW o no. El mayor inconveniente está dado en tener que sincronizar los hechos al momento de la consolidación en el depósito.

Dentro de las ventajas de aplicar un Data Mart a un negocio, se han seleccionado las siguientes:

- Son simples de implementar.
- Conllevan poco tiempo de construcción y puesta en marcha.
- Permiten manejar información confidencial.
- Reflejan rápidamente sus beneficios y cualidades.
- Reducen la demanda del depósito de datos.

3.3.6 Metodologías de desarrollo de un proyecto de BI

La metodología de Business Intelligence da mayor grado de certidumbre que nuestro proyecto cumplirá los objetivos trazados y en el tiempo convenido desde el punto de vista de un proyecto de Tecnología de Información, pero cabe aclarar que los proyectos de BI tienen 2 perspectivas, la de negocios y la de TI.

El uso de la metodología puede orientar a definir correctamente los objetivos de negocio y definir las metas del proyecto, pero no garantiza un proyecto exitoso debido a los factores directos e indirectos que influyen en este tema, como lo son las herramientas de BI seleccionadas, la cultura organizacional, perfiles técnicos y de negocio de los consultores, presupuesto del proyecto, entre otros.

3.3.6.1 Elección de la metodología para el desarrollo del proyecto

Aunque no es una regla, tanto en proyectos en general de TI así como proyectos de BI, la metodología se liga de cierta manera a la tecnología seleccionada, en el caso de BI, la plataforma seleccionada puede definir puntos esenciales relacionados a características propias de la tecnología, mejores prácticas, documentación de procesos y escenarios, definición de componentes, perfiles involucrados, etc.

Dentro de los rubros de la metodología que no se consideran afectados por la plataforma tecnológica se puede considerar los documentos de toma de requerimientos de negocio deberían ser similares para cualquier metodología.

La metodología con la que será desarrollado el proyecto es parte medular del mismo, por lo que deberá ser elegida con mucha precaución y buscando la que sea más afín a la tecnología seleccionada, tomando en cuenta estas apreciaciones la metodología con la que se desarrollará el presente proyecto es la metodología de Ralph Kimball.

3.3.6.2 Metodología de Ralph Kimball

Según (Kimball, 1998), la metodología de Ralph Kimball, llamada Modelo, se basa en lo que se denomina Ciclo de Vida Dimensional del Negocio. Esta metodología es considerada una de las técnicas favoritas a la hora de construir un Data Warehouse.

En el Modelo Dimensional se constituyen modelos de tablas y relaciones con el propósito de optimizar la toma de decisiones, con base en las consultas hechas en una base de datos relacional que están ligadas con la medición o un conjunto de mediciones de los resultados de los procesos de negocio.

El Modelo Dimensional es una técnica de diseño lógico que tiene como objetivo presentar los datos dentro de un marco de trabajo estándar e intuitivo, para permitir su acceso con un alto rendimiento. Cada Modelo Dimensional está compuesta por una tabla con una llave combinada, llamada tabla de hechos, y con un conjunto de tablas más pequeñas llamadas tablas de dimensiones.

3.3.6.2.1 Planificación y Administración del Proyecto

Busca identificar la definición y el alcance que tiene el proyecto de DWH. Esta etapa se concentra sobre la definición del proyecto, donde, a nivel de planificación, se establece la identidad del mismo, el personal, desarrollo del plan de proyecto, el seguimiento y la monitorización.

3.3.6.2.2 Definición de los requerimientos del negocio

Es un factor determinante en el éxito de un proceso de DWH. Los diseñadores de los Data Warehouse deben tener en claro cuáles son los factores claves que guían el negocio para determinar efectivamente los requerimientos y traducirlos en consideraciones de diseño apropiadas.

3.3.6.2.3 Diseño técnico de la arquitectura

En esta fase se deben tener en cuenta tres factores: los requerimientos de negocio, los actuales entornos técnicos, y las directrices técnicas y estratégicas futuras planificadas por la compañía, lo que permitirá establecer el diseño de la arquitectura técnica del entorno del Data Warehouse.

El proceso de diseño de la arquitectura técnica está compuesto de 8 pasos:

- Establecer un grupo de trabajo de arquitectura
- Requisitos relacionados con la arquitectura

- Documento de requisitos arquitectónicos
- Desarrollo de un modelo arquitectónico de alto nivel
- Diseño y especificación de los subsistemas
- Determinar las fases de aplicación de la arquitectura
- Documento de la arquitectura técnica
- Revisar y finalizar la arquitectura técnica

3.3.6.2.4 Selección de Productos e Instalación

Se evalúa y selecciona cuales son los componentes necesarios específicos de la arquitectura (plataforma de hardware, motor de la BD, herramienta de ETL, etc).

Luego de realizar la instalación de los componentes previamente evaluados y seleccionados, se recomiendan las siguientes premisas:

- Comprender el proceso de compras corporativas.
- Elaborar una matriz de evaluación del producto.
- Realizar la investigación de mercados.
- Filtrar opciones y realizar evaluaciones más detalladas.
- Manejo de un prototipo.
- Selección del producto, instalación y negociación.

3.3.6.2.5 Modelado Dimensional

Se comienza con una matriz donde se determina la dimensionalidad de cada indicador para luego especificar los diferentes grados de detalle dentro de cada concepto del negocio.

3.3.6.2.6 Diseño Físico

Se centra en la selección de las estructuras necesarias para soportar el diseño lógico. Un elemento principal de este proceso es la definición de estándares del

entorno de la base de datos. La indexación y las estrategias de particionamiento se determinan en esta etapa.

3.3.6.2.7 Diseño y Desarrollo de la Presentación de Datos

Tiene como principales actividades la extracción, transformación y carga (ETL). Estas actividades son altamente críticas ya que tienen que ver con la materia prima del Data Warehouse que son los datos.

3.3.6.2.8 Especificación de Aplicaciones para Usuarios Finales

Se identifican los roles o perfiles de usuarios para los diferentes tipos de aplicaciones necesarias en base al alcance de los perfiles detectados.

3.3.6.2.9 Desarrollo de Aplicaciones para Usuarios Finales

Involucra configuraciones de los metadatos y construcción de reportes específicos.

3.3.6.2.10 Implementación

Representa el correcto funcionamiento de la tecnología, los datos y las aplicaciones de usuarios finales accesibles para el usuario del negocio.

3.3.6.2.11 Mantenimiento y Crecimiento

Se basa en la necesidad de continuar con las actualizaciones de forma constante para así lograr la evolución de las metas por conseguir.

3.3.6.2.12 Gestión del Proyecto

Asegura que todas las actividades del ciclo de vida se lleven a cabo de manera sincronizada.

3.3.7 Herramientas

3.3.7.1 Plataforma Pentaho Open Source Business Intelligence

Pentaho es una plataforma de BI “orientada a la solución” y “centrada en procesos”, incluye todos los principales componentes requeridos para implementar soluciones basadas en procesos y ha sido concebido desde sus inicios para ser una solución basada en procesos.

Pentaho es una infraestructura de herramientas de análisis e informes integrados con un motor workflow de procesos de negocio. Lo que permitirá ejecutar las reglas de negocio necesarias, expresadas en forma de procesos y actividades para presentar y entregar información adecuada.

3.3.7.2 Módulos de la Plataforma Pentaho BI

3.3.7.2.1 Reporting

Un módulo de los informes ofrece la solución adecuada a las necesidades de los usuarios. Pentaho Reporting es una solución basada en el proyecto JFreeReport y permite generar informes ágiles y de gran capacidad. Pentaho Reporting permite la distribución de los resultados del análisis en múltiples formatos - todos los informes incluyen la opción de imprimir o exportar a formato PDF, XLS, HTML y texto. Los reportes Pentaho permiten también programación de tareas y ejecución automática de informes con una determinada periodicidad.

3.3.7.2.2 Análisis

Pentaho Análisis suministra a los usuarios un sistema avanzado de análisis de información. Con uso de las tablas dinámicas (pivot tables, crosstabs), generadas por Mondrian y JPivot, el usuario puede navegar por los datos, ajustando la visión de los

datos, los filtros de visualización, añadiendo o quitando los campos de agregación. Los datos pueden ser representados en una forma de SVG o Flash, los dashboards widgets, o también integrados con los sistemas de minería de datos y los portales web (portlets). Además, con el Microsoft Excel Analysis Services, se puede analizar los datos dinámicos en Microsoft Excel (usando la conexión a OLAP server Mondrian).

3.3.7.2.3 Mondrian

Mondrian utiliza MDX como lenguaje de consulta, que no tendría nada de raro si no fuera porque fue un lenguaje propuesto por Microsoft.

Mondrian funciona sobre las bases de datos estándar del mercado: Oracle, DB2, SQL-Server, MySQL, lo cual habilita y facilita el desarrollo de negocio basado en la plataforma Pentaho.

JPivot está considerado con un “proyecto hermano” de Mondrian. Al incorporarse Mondrian también lo ha hecho JPivot.

3.3.7.2.4 Dashboards

Todos los componentes del módulo Pentaho Reporting y Pentaho Análisis pueden formar parte de un Dashboard. En Pentaho Dashboards es muy fácil incorporar una gran variedad en tipos de gráficos, tablas y velocímetros (Dashboard widgets) e integrarlos con los Portlets JSP, en donde podrá visualizar informes, gráficos y análisis OLAP.

3.3.7.2.5 Integración de Datos

Se realiza con una herramienta Kettle ETL (Pentaho Data Integration) que permite implementar los procesos ETL. Últimamente Pentaho lanzó una nueva versión que marcó un gran paso adelante en OSBI ETL y que hizo de Pentaho Data Integration una alternativa interesante para las herramientas comerciales.

3.3.7.2.6 Diseño de Cubo

Para el diseño de cubos se usará la herramienta OLAP Mondrian, que permite diseñar y crear cubos de información para análisis multidimensional.

Dichos cubos se componen de archivos XML y en ellos se definen las Dimensiones y las conexiones de los datos. Los archivos XML por lo general son complejos de realizar manualmente por lo que es común utilizar herramientas gráficas para realizar la edición de estos. Como ejemplo de estas herramientas Open Source Pentaho se tiene Cube Designer para la Creación de cubos y el Workbench para la edición de los mismos.

3.3.7.2.7 Base de Datos

Una conexión de base de datos describe el método mediante el cual Kettle se conectará a una base de datos. Se pueden crear conexiones específicas para un Trabajo o Transformación o guardarlas en el catálogo de Kettle para reutilizarlas en múltiples Transformaciones o Trabajos.

CAPÍTULO 4

DESARROLLO DEL PROYECTO

4.1 DESARROLLO DEL BI PARA LA UNIDAD ADMINISTRATIVA Y FINANCIERA DE LA UNIVERSIDAD DE LAS FUERZAS ARMADAS ESPE.

Para el desarrollo del Data Mart se utilizará la metodología de Ralph Kimball lo que permitirá dar un seguimiento adecuado al desarrollo del proyecto.

4.2 PLANEACIÓN Y ADMINISTRACIÓN DEL PROYECTO

4.2.1 Definición del proyecto

La Dirección financiera y recursos humanos de la Universidad de las Fuerzas Armadas ESPE, maneja una cantidad considerable de información utilizando varias bases de datos (Sybase, Oracle y Access), como un repositorio muy completo de los datos. Esta información no se encuentra procesada de acuerdo a las necesidades institucionales o de los órganos de control universitario como el CES, SENESCYT, CEACES; de tal manera que cuando se requieren reportes, informes estadísticos o tabulares es necesario realizarlos de forma manual.

Existen muchas relaciones entre los datos que no han podido utilizarse para efectos de planificación, organización, logística, etc. Siendo por tanto indispensable el procesamiento, adecuando de estos datos para cubrir estas necesidades.

4.2.2 Preparación para un proyecto de BI

La Universidad de las Fuerzas Armadas ESPE, muestra interés en el desarrollo del proyecto, la Unidad de Tecnologías de la Información UTICS, fue la encargada de presentar este proyecto como soporte para el análisis de información y toma de decisiones para los encargados de procesar dicha información en las áreas involucradas.

4.2.3 Alcance

Se realizará el diseño e implementación de un Data Mart para la automatización y depuración de la información contenida en la base de datos de las Unidades Administrativa y Financiera de la Universidad de las Fuerzas Armadas ESPE. Esto se realizará con la metodología de Ralph Kimball la misma que ayudará en la planificación, toma de requerimientos, análisis, construcción y en general en cada una de las fases que componen el ciclo de vida del proyecto.

La aplicación contempla procesos de consolidación de la información, para que esta pueda ser utilizada en reportes variados que permitan considerar la información para planes de mejora o políticas institucionales. Se realizarán reportes automatizados con información precisa, en base a los requerimientos del área antes mencionada.

Una vez implementado el Data Mart para las Unidades Administrativas y Financieras, podrá ser integrado con Data Marts que se realizaran posteriormente y que conformarán el Data Warehouse de la Universidad de las Fuerzas Armadas ESPE, entre estos se encuentra por ejemplo el Data Mart Académico.

4.2.4 Justificación

El proyecto propuesto busca procesar y reorganizar la información de la base de datos de la dirección financiera y recursos humanos, logrando con esto la

simplificación de procesos tales como la elaboración de reportes, consultas determinadas y en general la toma de decisiones.

La solución que se plantea es el diseño e implementación de un Data Mart, el mismo que estará conformado por la recopilación de información obtenida de las distintas bases de datos ya sean Oracle, Sybase o Access, logrando con esto un análisis eficiente y simplificado de la información, disminuyendo considerablemente el tráfico generado en las consultas hacia la base de datos y por consiguiente liberando recursos dentro del servidor.

4.2.5 Planificación del proyecto

Se asigna roles a las personas involucradas en el ciclo de vida del desarrollo del proyecto de BI, con el objetivo de tener responsables de cada fase del ciclo de vida del proyecto. Tomando en consideración las necesidades e intereses de la Universidad de las Fuerzas Armadas ESPE en el desarrollo del proyecto se asignaron los siguientes responsables.

- ***Patrocinador del Proyecto:*** Universidad de las Fuerzas Armadas ESPE.
- ***Administrador de Base de Datos:*** Ing. Lorena Duque
- ***Director del Proyecto:*** Ing. Lorena Duque
- ***Desarrollador:*** Estudiantes Egresados (Cristian Alvarez, Diego Esparza)
- ***Arquitecto de Datos:*** Estudiantes Egresados (Cristian Alvarez, Diego Esparza)
- ***Diseñador:*** Estudiantes Egresados (Cristian Alvarez, Diego Esparza)
- ***Personal involucrado en el proyecto:*** empleados del área administrativa y financiera de la Universidad de las Fuerzas Armadas ESPE.

4.2.6 Administración del proyecto

El proyecto se desarrolla en las instalaciones de las UTICS por lo que se contará con el apoyo y monitoreo de todo el personal involucrado en desarrollo del BI, realizando revisiones periódicas del avance y evolución del mismo.

4.3 DEFINICIÓN DE LOS REQUERIMIENTOS DEL NEGOCIO

La toma de requerimientos se lo realizó con los usuarios encargados de procesar la información tanto de la unidad administrativa como financiera de la Universidad de las Fuerzas Armadas ESPE, cada requerimiento fue descrito en el formato establecido previamente en la metodología.

A continuación se describen los requerimientos divididos de acuerdo a los módulos existentes dentro de cada unidad.

4.3.1 Requerimientos módulos de las unidades administrativa y financiera.

Cuadro 2
Toma Requerimientos

Módulo	Identificador	Nombre Requerimiento	Descripción
Presupuesto	PR001	ITEMS POR ACTIVIDAD	El sistema debe permitir generar un reporte, el cual presente la información consolidada de ítems y actividad en una tabla cruzada.
	PR002	VALOR PAGADO POR ITEM	El sistema debe permitir generar un reporte, en el cual se presenten el valor pagado para cada uno de los ítems, agrupados por actividad.
	PR003	CUENTAS SEPARADAS	El sistema debe permitir conocer la inversión y el gasto para cada uno de los ítems, agrupados por actividades.
	PR004	VALOR ASIGNADO POR ITEM	El sistema debe permitir generar un reporte, en el cual se presenten el valor asignado para cada uno de los ítems, agrupados por actividad.

Activos Fijos	PR005	VALOR ASIGNADO POR MESES	El sistema debe permitir generar un reporte, en el cual se presente el valor asignado a cada ítem o actividad para cada uno de los meses.
	PR006	VALOR PAGADO POR MESES	El sistema debe permitir generar un reporte, en el cual se presente el valor pagado a cada ítem o actividad para cada uno de los meses.
	PR007	VALOR ASIGNADO POR PROGRAMA	El sistema debe permitir generar un reporte, en el cual se presente el valor asignado a cada programa para cada uno de los meses, o semestres.
	PR008	VALOR PAGADO POR PROGRAMA	El sistema debe permitir generar un reporte, en el cual se presente el valor pagado a cada programa para cada uno de los meses, o semestres.
	PR009	VALOR ASIGNADO POR PROYECTO	El sistema debe permitir generar un reporte, en el cual se presente el valor asignado a cada proyecto para cada uno de los meses, o semestres.
	PR010	VALOR PAGADO POR PROYECTO	El sistema debe permitir generar un reporte, en el cual se presente el valor pagado a cada proyecto para cada uno de los meses, o semestres.
	AF001	VALOR POR TIPOS	El sistema debe permitir conocer el valor de los bienes de corta o larga duración, agrupados por tipos.
	AF002	BIENES POR SECCIÓN	El sistema debe permitir conocer el valor de los tipos de bienes existentes en cada una de las secciones de la Universidad de las Fuerzas Armadas ESPE
	AF003	BIENES POR DIRECCIÓN	El sistema debe permitir conocer el valor de los tipos de bienes existentes en cada uno de los campus de la Universidad de las Fuerzas Armadas ESPE.
	AF004	BIENES POR ÁREA	El sistema debe permitir conocer el valor de los tipos de bienes existentes en cada una de las áreas de la Universidad de las Fuerzas Armadas ESPE.
	AF005	BIENES FECHA ADQUISICIÓN	El sistema debe permitir conocer el valor de los tipos de bienes según la fecha en la que hayan sido adquiridos.
	AF006	BIENES POR CLASE	El sistema debe permitir conocer el valor de los tipos de bienes clasificados de acuerdo a su clase.
	AF007	VALOR POR SUBTIPOS	El sistema debe permitir conocer el valor de los bienes de corta o larga duración, agrupados por tipos y subtipos.
	AF008	VALOR POR CUSTODIO	El sistema debe permitir conocer el valor de los bienes de corta o larga

Inventarios	IN001	INGRESOS Y EGRESOS POR BODEGA	duración, agrupados por custodio al cual pertenezcan dichos bienes. El sistema debe permitir conocer los ingresos y egresos para cada una de las bodegas.
	IN002	INGRESOS Y EGRESOS POR UNIDAD	El sistema debe permitir conocer los ingresos y egresos para cada una de las unidades administrativas.
	IN003	INGRESOS Y EGRESOS POR TRANSACCIÓN	El sistema debe permitir conocer los ingresos y egresos por tipo de transacción.
	IN004	INGRESOS Y EGRESOS POR RUBRO	El sistema debe permitir conocer los ingresos y egresos por tipo de rubro.
Roles de Pago (SHIFRE)	RS001	SUELDO MENSUAL POR EMPLEADO	El sistema debe permitir conocer los sueldos mensuales de cada uno de los empleados.
	RS002	SUELDO MENSUAL POR TIPO DE EMPLEADO	El sistema debe permitir conocer los sueldos mensuales y el tipo de empleado (docente, administrativo, etc.) de cada uno de los empleados.
	RS003	SUELDO MENSUAL POR RUBRO	El sistema debe permitir conocer los sueldos mensuales, divididos en rubros.
	RS004	SUELDO MENSUAL POR TIPO DE RUBRO	El sistema debe permitir conocer los sueldos mensuales, divididos en ingresos, egresos o patrono.
	RS005	SUELDO MENSUAL POR TIPO DE DOCENTE	El sistema debe permitir conocer los sueldos mensuales, por tipo de docente.
	RS006	SUELDO MENSUAL POR DEPARTAMENTO	El sistema debe permitir conocer los sueldos mensuales, de los empleados pertenecientes a un departamento determinado.
	RS007	SUELDO MENSUAL POR CARGO	El sistema debe permitir conocer los sueldos mensuales, de los empleados agrupados por cargo.
	RS008	SUELDO MENSUAL POR CATEGORÍA	El sistema debe permitir conocer los sueldos mensuales, de los empleados y la categoría a la cual pertenecen.
	RS009	SUELDO MENSUAL DISCAPACIDAD	El sistema debe permitir conocer los sueldos mensuales, de los empleados con algún tipo de discapacidad.
	RS010	SUELDO MENSUAL TERCERA EDAD	El sistema debe permitir conocer los sueldos mensuales, de los empleados pertenecientes a la tercera edad.
	RS011	SUELDO MENSUAL POSTGRADO	El sistema debe permitir conocer los sueldos mensuales, de los empleados con título de cuarto nivel.
	RS012	SUELDO MENSUAL POR TIPO DE PROFESOR	El sistema debe permitir conocer los sueldos mensuales, por tipo de profesor ej.: docente antiguo, ocasional, etc.
	RS013	VALOR LIQUIDO	El sistema debe permitir conocer el

		POR EMPLEADO	valor líquido recibido por mes para cada empleado.
	RS014	SUELDO POR JORNADA DE TRABAJO	El sistema debe permitir conocer los sueldos mensuales de los empleados por jornada de trabajo ej.: horario militar, horario normal diurno, etc.
	RS015	SUELDO POR FORMA DE TRABAJO	El sistema debe permitir conocer los sueldos mensuales de los empleados por forma de trabajo ej.: contrato servidores públicos, contrato docentes, etc.
	RS016	CANTIDAD DOCENTES POR DEPARTAMENTO	El sistema debe permitir conocer la cantidad de docentes que existen en cada dedicación en un determinado tiempo y por departamento ej: ESPE MATRIZ - DEP.. LENGUAS - FORT. ACADÉMICO, ESPE MATRIZ - EDUCACIÓN FÍSICA - FORT. ACADÉMICO, etc.
	RS017	VALOR HORA CLASE POR DEPARTAMENTO	El sistema debe permitir conocer el valor de la hora clase de docentes que existen en cada dedicación en un determinado tiempo y por departamento ej: ESPE MATRIZ - DEP.. LENGUAS - FORT. ACADÉMICO, ESPE MATRIZ - EDUCACIÓN FÍSICA - FORT. ACADÉMICO, etc.
Roles de Pago (BANNER)	RB001	SUELDO POR RUBROS	El sistema debe permitir conocer los sueldos de cada empleado, desglosado por rubros y tipos de rubros.
	RB002	SUELDO POR TIPO DE EMPLEADO	El sistema debe permitir conocer los sueldos de acuerdo al tipo de empleado.
	RB003	SUELDO POR AÑO Y MES	El sistema debe permitir conocer los sueldos de cada empleado para un año y mes determinado.
	RB004	SUELDO POR DISPOSICIÓN	El sistema debe permitir conocer los sueldos de acuerdo a la disposición de cada empleado.
	RB005	SUELDO POR ESTADO CIVIL	El sistema debe permitir conocer los sueldos de acuerdo al estado civil de cada empleado.
	RB006	EMPLEADO POR CAMPUS	El sistema debe permitir conocer los campus a los que pertenece cada empleado.
	RB007	TÍTULOS POR EMPLEADO	El sistema debe permitir conocer por cada empleado, el nivel del título, el título, la fecha de obtención, la universidad, el país, registro senescyt y la sub área del título.
	RB008	TIEMPO DEDICACIÓN	El sistema debe permitir conocer por cada empleado, el tiempo dedicación (tiempo completo, tiempo parcial).

	RB009	ASIGNATURAS POR DOCENTE	El sistema debe permitir conocer por cada docente, el nombre de la asignatura, y las horas que se dicta.
	RB010	PUBLICACIONES POR EMPLEADO	El sistema debe permitir conocer el nombre de la publicación, el nombre del publicador y la fecha de la publicación.
	RB011	DISCAPACITADOS	El sistema debe permitir conocer los sueldos de acuerdo a la disposición de cada empleado.
Matriculación	M001	ARANCEL POR PERIODO	El sistema debe permitir conocer los valores de cada uno de los aranceles, para un período determinado.
	M002	VALOR RECAUDADO POR PERIODO	El sistema debe permitir conocer el valor total recaudado, es decir la sumatoria de todos los aranceles de varios tipos.
	M003	ARANCEL POR TIPO DE ALUMNO	El sistema debe permitir conocer el valor por aranceles desglosados o totalizados, para cada tipo de alumno existente.
	M004	ARANCEL POR CARRERA	El sistema debe permitir conocer el valor por aranceles desglosados o totalizados, para cada carrera existente.
	M005	ARANCEL POR CAMPUS	El sistema debe permitir conocer el valor por aranceles desglosados o totalizados, para cada uno de los campus.
	M006	ARANCEL POR AÑOS	El sistema debe permitir conocer el valor por aranceles desglosados o totalizados, por cada año.
		ARANCEL POR MESES	El sistema debe permitir conocer el valor por aranceles desglosados o totalizados, para cada mes.
	M008	ARANCEL POR SEMANA	El sistema debe permitir conocer el valor por aranceles desglosados o totalizados, para cada semana.
Cuentas por cobrar	C001	SALDO POR AÑOS	El sistema debe permitir conocer los saldos de facturación desde el año 2007 hasta el actual.
	C002	SALDO POR DETALLE	El sistema debe permitir conocer el saldo de facturación por detalle de factura.
	C003	SALDO POR CLIENTE	El sistema debe permitir conocer el saldo de facturación por cliente.
	C004	REPORTE FACTURACIÓN	El reporte debe permitir conocer el saldo de facturación desde el año 2007 hasta el actual.
	C005	REPORTE CARTERA Y PROYECTO	El reporte debe permitir conocer el saldo de cartera y proyecto desde el año 2004 hasta el 2008.

4.3.2 Análisis y requerimientos

Los distintos OLTP necesarios para el desarrollo del proyecto son variados, entre estos destacan el servidor de base de datos Oracle y archivos planos como csv y xls.

El objetivo de implementar un BI para las unidades administrativa y financiera de la Universidad de las Fuerzas Armadas ESPE, consiste en poder analizar la información de manera consolidada y precisa, sin depender de los aplicativos con los que trabajan estas unidades como por ejemplo: Olympo, Esigef o Shifre. Los usuarios podrán acceder a la información a través de la consola de usuario de la herramienta BI.

4.4 DISEÑO TÉCNICO DE LA ARQUITECTURA

4.4.1 Estándares

Los estándares son las reglas o medidas que se van a utilizar en el proyecto como un requisito para obtener una presentación de la información estableciendo una gran calidad de orden. Los estándares se los van a definir mediante un esquema que tendrá varias columnas y se van a especificar a continuación:

- **Alcance:** Se indica el nombre o atributos del objeto al que se va a estandarizar.
- **Prefijo:** Se almacena el nombre del prefijo que va a tener el objeto.
- **Número de caracteres:** Se indica el número de caracteres que tendrá el nombre del objeto.
- **Descripción:** Se indica el estándar del nombre completo que debe tener cada uno de los objetos.
- **Ejemplos:** Se coloca un nombre del resultado final del estándar.
- **Nombres:** Se coloca el nombre con el que va a quedar definido el objeto.

4.4.1.1 Almacenamiento de las carpetas del desarrollo de los datamarts.

Estas carpetas estarán ubicadas en el servidor de pentaho en el directorio:
/opt/pentaho/biserver-ce/pentaho-solutions/

4.4.1.1.1 Carpeta general

A continuación se muestra el esquema que va a tener las carpetas generales que servirán para almacenar toda la información acerca de un área de trabajo de la Universidad de las Fuerzas Armadas ESPE, para este caso se tomará los nombres referente a nuestro tema de tesis.

Cuadro 3
Estándar carpeta general

Alcance	Prefijo	Número Caracteres	Descripción	Ejemplos
CARPETA GENERAL	DATAMART_	PREFIJO + MÁXIMO 15	DATAMART_ + NOMBRE	DATAMART_FINANZAS_ RRHH

En donde:

DATAMART_: Es el prefijo que se va a establecer para las carpetas generales.

NOMBRE: Es un nombre cualquiera con el cual el usuario pueda diferenciar cada uno de los módulos.

4.4.1.1.2 Subcarpetas

Las subcarpetas creadas tendrán como objetivo organizar la información de manera específica, en nuestro caso se creará las subcarpetas de acuerdo a nuestro tema de tesis y tendrán el siguiente esquema:

Cuadro 4
Estándar subcarpetas

Alcance	Número Caracteres	Descripción	Ejemplos
SUBCARPETAS	MÁXIMO 15	NOMBRE	FINANZAS, PERSONAL

En donde:

NOMBRE: Es un nombre cualquiera que pondrá el usuario para diferenciar sus áreas de trabajo.

Dentro de cada subcarpeta se tendrá varias carpetas las cuales tendrán el siguiente nombre:

- **ETL:** Dentro de esta carpeta estarán 3 carpetas:
 - **TRANSFORMACIÓN:** En esta carpeta se guardarán carpetas que tendrán el nombre del submódulo y en ellas se va alojar los procesos ETL en extensión .ktr los cuales se los realiza en la herramienta Pentaho Data Integration (SPOON).
 - **INPUT:** Se guardarán todos los archivos planos (csv,xls,txt,etc) que se vaya a utilizar para la carga de datos en la base datos stage.
 - **JOB:** Se guardarán todos los Jobs en extensión kjb que se realicen en la herramienta spoon de Pentaho Data Integration.
- **REPORT:** Se guardarán los archivos de extensión prpt, que son los reportes que se realizan en la herramienta pentaho report designer
- **WORKBENCH:** Se guardarán los cubos en extensión xml los cuales se los realiza en la herramienta pentaho schema workbench
- **DASHBOARD:** Se guardarán los dashboard en extensión xml los cuales se los realiza en la consola de usuario (Pentaho user console).
- **VIEW:** Se guardarán las vistas que se crean en la consola de usuario (Pentaho user console) por el usuario, estas vistas se las obtiene de una vista de análisis del jpivot o del plugin sayku analytics, se manipula al cubo como se requiera y se manda a guardar con información específica de un cubo determinado.

La estructura queda de la siguiente forma:

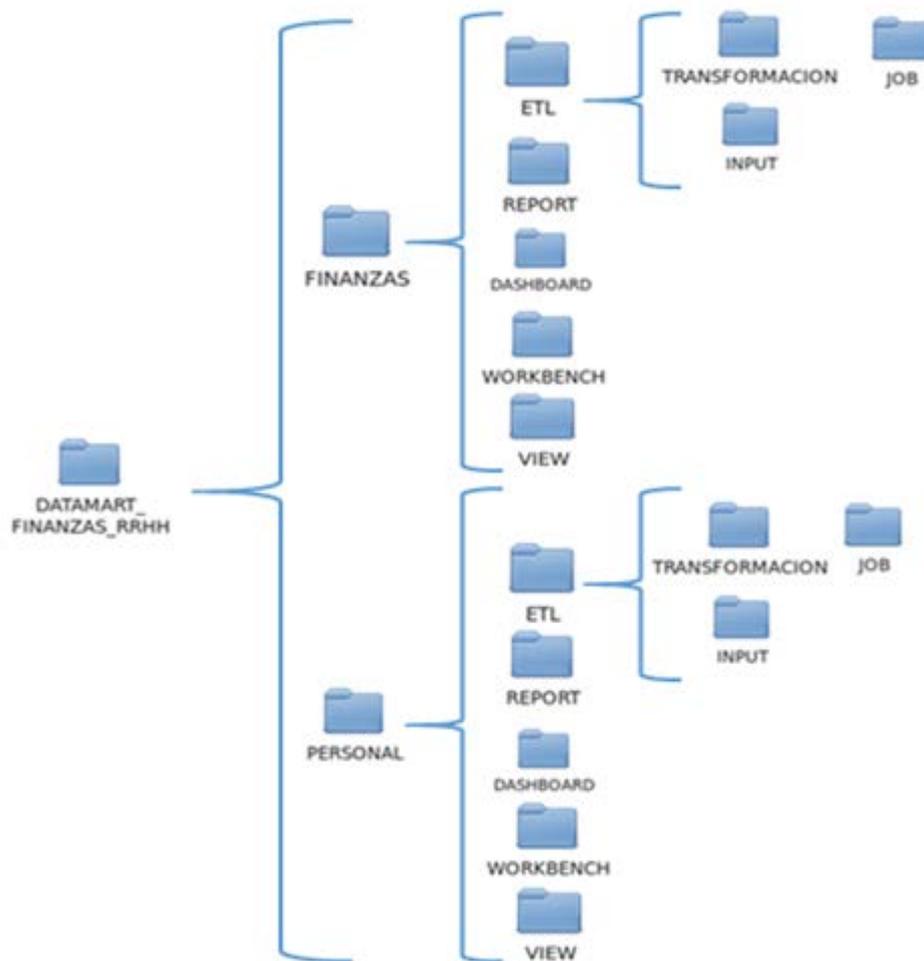


Figura 36: Estructura carpetas

A continuación se va a definir la estructura que tendrán los estándares de los módulos y submódulos. Para este caso se indicará los nombres referente al proyecto de tesis.

4.4.1.1.3 Nombre de los módulos

Los nombres de los módulos se refieren a los departamentos de la Universidad de las Fuerzas Armadas ESPE y tendrán la siguiente estructura:

Cuadro 5
Estándar módulos

Alcance	Número Caracteres	Descripción	Nombre
MÓDULOS	MÁXIMO 15	NOMBRE	FINANZAS, PERSONAL

4.4.1.1.4 Nombre de los submódulos

Los nombres de los submódulos se refieren a las distintas áreas de trabajo que existen en los departamentos de la Universidad de las Fuerzas Armadas ESPE y tendrán la siguiente estructura:

Cuadro 6
Estándar submódulos

Alcance	Número Caracteres	Descripción	Nombre
MÓDULOS	MÁXIMO 20	NOMBRE	PRESUPUESTO ACTIVOS CUENTASXCOBRAR ROLES MATRÍCULAS

En donde:

NOMBRE: Es un nombre cualquiera con el que se pueda identificar un departamento, área o unidad de trabajo de la Universidad de las Fuerzas Armadas ESPE en caso de ser un módulo, si es submódulo será un nombre de un área de trabajo referente al módulo.

4.4.1.2 Estándares para modelamiento

El estándar para los nombres de los archivos de los modelos entidad relación y dimensional tendrán la siguiente estructura.

Cuadro 7
Estándar modelamiento

Alcance	Prefijo	Número Caracteres	Descripción	Ejemplo
Entidad relación	MOD_ER_	MÁXIMO 20	MOD_ER_ + 3 PRIMERAS LETRAS NOMBRE MODULO + CÓDIGO NUMÉRICO 4 DÍGITOS	MOD_ER_FIN0001
Entidad Dimensional	MOD_DIM_	MÁXIMO 20	MOD_DIM_ + 3 PRIMERAS LETRAS NOMBRE MODULO + CÓDIGO NUMÉRICO 4 DÍGITOS	MOD_DIM_FIN0001

4.4.1.3 Estándar para base de datos stage

4.4.1.3.1 Nombre de objetos

Para la creación de objetos en la base de datos stage, se tomara como referencia los estándares que están definidos en la Universidad de las Fuerzas Armadas ESPE.

4.4.1.3.2 Nombre de la base de datos stage

El estándar para nombres de las bases de datos stage tendrá el siguiente formato:

Cuadro 8
Estándar de nombre de base de datos stage

Alcance	Número Caracteres	Descripción	Nombre
Base de datos stage	MÁXIMO 8	Es un nombre cualquiera que identifique el propósito de la base de datos.	ESPE

4.4.1.3.3 Nombre del esquema stage

El estándar para nombres de esquemas stage tendrá el siguiente formato:

Cuadro 9
Estándar de nombre de esquema stage

Alcance	Número Caracteres	Descripción	Nombre
Esquema stage	MÁXIMO 8	Es un nombre cualquiera que identifique el propósito del esquema.	STGESPE

4.4.1.4 Estándar para nombre y esquema de base de datos del datamart.

4.4.1.4.1 Nombre de la base de datos datamart

El estándar para nombres de las bases de datos del datamart tendrá el siguiente formato:

Cuadro 10
Estándar de nombre de base de datos datamart

Alcance	Número Caracteres	Descripción	Nombre
Base de datos datamart	MÁXIMO 8	Es un nombre cualquiera que identifique el propósito de la base de datos.	ESPE

4.4.1.4.2 Nombre del esquema datamart

El estándar para nombres de esquemas del datamart tendrá el siguiente formato:

Cuadro 11
Estándar de nombre de esquema datamart

Alcance	Número Caracteres	Descripción	Nombre
Esquema datamart	MÁXIMO 8	Es un nombre cualquiera que identifique el propósito del esquema.	DWESPE

Ejemplo:

/root/SCRIPTS/TABLAS/FINANZAS/PRESUPUESTOS

4.4.1.5 Estándar para dimensiones y tabla de hechos

Para que sea más claro los ejemplos de los siguientes estándares se tomarán de referencia la dimensión **docente**, el módulo de **finanzas** y el submódulo **roles**.

Los nombres para las tablas de dimensión y de hechos en el datamart tienen la siguiente estructura:

Cuadro 12
Estándar tablas de hechos y dimensiones

Alcance	Prefijo	Número Caracteres	Descripción	Ejemplo
tablas de dimensiones	DIM_	MÁXIMO 20	DIM_NOMBRE TABLA	DIM_DOCENTE
tablas de hechos	FAC_	MÁXIMO 20	FAC_NOMBRE SUBMODULO	FAC_ROLES

Donde:

- **NOMBRE TABLA:** será un nombre cualquiera pero que haga referencia al nombre de la tabla de la base de datos stage.

Ejemplo:



Figura 37: Ejemplo estándar nombre tabla

- **NOMBRE SUBMÓDULO:** ya están definidos anteriormente en el literal 4.4.1.1.4.

NOTA: En el caso de que se repita el nombre de la tabla de hechos, el usuario pondrá un nombre cualquiera que haga referencia al submódulo.

El estándar de nombres para campos en tablas de dimensiones tendrá la siguiente estructura:

Cuadro 13
Estándar campos de dimensiones

Alcance	Prefijo	Número Caracteres	Descripción	Ejemplo
Claves primarias	SK_	MÁXIMO 15	SK_NOMBRE TABLA	SK_DOCENT
Campos	D	MÁXIMO 25	D + 3 INICIALES NOMBRE TABLA_NOMBRE ATRIBUTO	DDOC_CIU_CEDULA

Donde:

Las 3 iniciales de la tabla se refieren a las 3 iniciales contando desde el quinto carácter de la tabla de la base de datos stage.

En caso de que se llegaran a repetir las 3 iniciales con otra tabla se tomará las iniciales necesarias contando desde el quinto carácter de la tabla de la base de datos stage para poderse diferenciar de las otras tablas.

El nombre del atributo se lo define haciendo referencia a los campos de la tabla de la base de datos stage.

Ejemplo de las 3 iniciales del nombre de la tabla para la tabla del datamart:

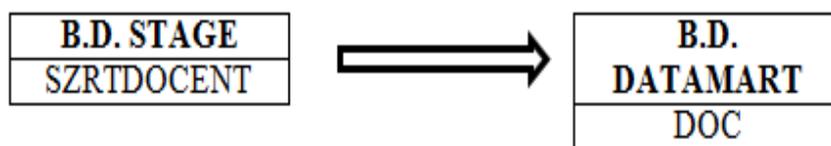


Figura 38: Ejemplo 3 iniciales nombre tabla

Ejemplo del nombre del atributo para la tabla del datamart

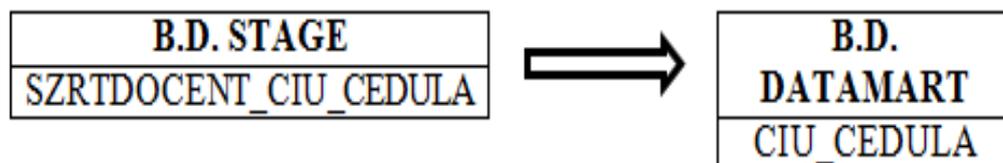


Figura 39: Ejemplo nombre atributo

DIMENSIÓN TIEMPO

La dimensión tiempo va a ser definida de la siguiente forma:

Cuadro 14
Definición de dimensión tiempo

Alcance	Número Caracteres	Descripción	Nombre
Clave primaria de las dimensión tiempo	MÁXIMO 15	SK_FECHA	SK_FECHA
		FORMATO DD/MM/YY	FECHA
		FORMATO YYYY	AÑO
		NOMBRE DEL SEMESTRE	NSEMESTRE
		DÍGITOS DEL 1 AL 4	TRIMESTRE
		DÍGITOS DEL 1 AL 12	MES
		DÍGITOS DEL 1 AL 53	SEMANA
		EL DÍA EN DÍGITOS	DÍA
		EL DÍA DE LA SEMANA EN DÍGITOS (1 AL 7)	DIASEMANA
		T+NUMERO DEL TRIMESTRE+/+EL AÑO (YY)	NTRIMESTRE
		NOMBRE DEL MES	NMES
		3 PRIMERAS LETRAS DEL MES	NMES3L
		SEM+NUMERO DE LA SEMANA+/+AÑO (YY)	NSEMANA
Campos de la dimensión tiempo	MÁXIMO 15	NUMERO DEL DÍA + 3 PRIMERAS LETRAS DEL MES	NDIA
		NOMBRE DEL DÍA DE LA SEMANA	NDIASEMANA

El estándar de nombres para campos en tablas de hechos tendrá la siguiente estructura:

Cuadro 15
Estándar campos para tabla de hechos

Alcance	Prefijo	Número Caracteres	Descripción	Ejemplo
Claves primarias	SK_	MÁXIMO 15	SK_NOMBRE SUBMODULO	SK_ROLES
Claves secundarias		MÁXIMO 15	CLAVE PRIMARIA DE LA DIMENSIÓN	SK_DOCENT
Campos	F	MÁXIMO 25	F + 3 INICIALES NOMBRE SUBMODULO_NOMBRE ATRIBUTO	FROL_VALOR

NOTA: Las dimensiones y las tablas de hechos tendrán definidos 3 campos adicionales que se crean automáticamente con los procesos ETL y serán de ayuda para una tabla de auditoria. Los 3 campos se explican a continuación:

- **FECHA_ETL:** Se almacena la fecha en la que se cargó la transformación.
- **NOMBRE_ETL:** Se almacena el nombre con el que se guarda la transformación.
- **IP_EQUIPO:** Se almacena la dirección ip del equipo.

Nombre de la tabla de auditoria en el datamart.

Esta tabla será la encargada de cargar toda la información que tenga algún tipo de error. Se la va a definir con el siguiente nombre:

- TB_AUDITORIA

Esta tabla auditoria será estática en cuestión a campos y tendrá los 3 campos mencionados anteriormente en la nota, adicionalmente tendrá 4 campos más y el esquema de la tabla auditoria será de la siguiente forma:

CAMPOS:

- **FECHA_ETL:** Se almacena la fecha en la que se cargó la transformación.
- **NOMBRE_ETL:** Se almacena el nombre con el que se guarda la transformación.
- **IP_EQUIPO:** Se almacena la dirección ip del equipo.
- **Cod_error:** Se almacena un código de error.
- **Campo_error:** Se almacena el campo donde existe un tipo de error.
- **Descr_error:** Se almacena una descripción del error.
- **Num_error:** Se almacena el número de errores

4.4.1.6 Estándares para la herramienta PENTAHO**Pentaho administration console**

El estándar para nombres de usuarios y roles de la administración de consola de pentaho tendrán la siguiente estructura:

Cuadro 16
Estándar usuarios y roles

Alcance	Número Caracteres	Descripción	Ejemplo
Usuarios	MÁXIMO 20	Primera letra del primer nombre + el primer apellido y un número secuencial que inicia en 1.	calvarez1
Roles	MÁXIMO 15	Nombre de roles establecidos por pentaho.	ceo, is, dev, etc

En caso de que el usuario se llegara a repetir, se incrementa el número del final.

Ejemplo de usuarios repetidos:



Figura 40: Ejemplo usuario repetido

El estándar para nombres de las conexiones de las bases de datos que se crean en pentaho tendrá el siguiente formato:

Cuadro 17
Estándar nombre de conexiones

Alcance	Número Caracteres	Descripción	EJEMPLO
Database connections	MÁXIMO 8	Es el nombre de la base de datos a la que el usuario requiera conectarse.	ESPE

4.4.1.6.1 Estándar pentaho data integration (PDI)

El Estándar de nombres de las transformaciones tendrá la siguiente estructura:

Cuadro 18
Estándar nombre de transformaciones

Alcance	Prefijo	Número Caracteres	Descripción	Ejemplo
Transformación	ETL_	MÁXIMO 25	ETL_NOMBRE DIMENSIÓN O TABLA DE HECHOS	ETL_DIM_TIPO / ETL_FAC_PRESUPUESTO
Transformacion (Stage)	ETL_CARGA -	MÁXIMO 25	ETL_CARGA_NOMBRE TABLA	ETL_CARGA_TIPO

El estándar de nombres de los job tendrá la siguiente estructura:

Cuadro 19
Estándar nombre de jobs

Alcance	Prefijo	Número Caracteres	Descripción	Ejemplo
Job	JOB_	MÁXIMO 20	JOB_NOMBRE SUBMODULO	JOB_PRESUPUESTO

NOTA: En el caso que existan varios job para un submódulo, se pone cualquier nombre que haga referencia al submódulo.

DIRECTORIO DE LOS LOGS

Los logs son el detalle de la ejecución del job de cada módulo y se encuentra en la siguiente estructura: root/JOBS_LOGS/

4.4.1.6.2 Estándar pentaho report designer (PRD)

El estándar de nombres de los reportes tendrá la siguiente estructura:

Cuadro 20
Estándar nombre de reportes

Alcance	Prefijo	Número Caracteres	Descripción	Ejemplo
Reportes	REP_	15	REP_3 PRIMERAS LETRAS NOMBRE MODULO_3 PRIMERAS LETRAS NOMBRE SUBMODULO + CÓDIGO NUMÉRICO SECUENCIAL DE 4 DÍGITOS	REP_FIN_PRE0001

4.4.1.6.3 Estándar pentaho schema workbench (PSW)

El estándar de nombres de los cubos tendrá la siguiente estructura:

Cuadro 21
Estándar nombre de cubos

Alcance	Prefijo	Número Caracteres	Descripción	Ejemplo
Cubos	CUB_	15	CUB_3 PRIMERAS LETRAS NOMBRE MODULO_3 PRIMERAS LETRAS NOMBRE SUBMODULO + CÓDIGO NUMÉRICO SECUENCIAL DE 4 DÍGITOS	CUB_FIN_PRE0001

4.4.1.6.4 Pentaho user console

Community dashboard editor (CDE)

El estándar de nombres de los dashboard tendrá la siguiente estructura:

Cuadro 22

Estándar nombre de dashboard

Alcance	Prefijo	Número Caracteres	Descripción	Ejemplo
Dashboard	DAS_	15	DAS_3 PRIMERAS LETRAS NOMBRE MODULO_3 PRIMERAS LETRAS NOMBRE SUBMODULO + CÓDIGO NUMÉRICO SECUENCIAL DE 4 DÍGITOS	DAS_FIN_PRE0001

El estándar de nombres de las vistas tendrá la siguiente estructura:

Cuadro 23

Estándar nombre de vistas

Alcance	Prefijo	Número Caracteres	Descripción	Ejemplo
VISTAS	VIEW_	MÁXIMO 15	VIEW_3 PRIMERAS LETRAS NOMBRE SUBMODULO_NOMBRE	VIEW_PRE_ACTIVIDAD

En donde:

NOMBRE: Es un nombre cualquiera que identifique a la vista que el usuario realice.

ESTÁNDARES PARA REPORTES AD-HOC

NOMBRE DEL REPORTE

Los nombres de los reportes adhoc que se realicen ya sean con jpivot o con el plugin saiku reporting deben ser de 8 caracteres y van de acuerdo al siguiente esquema:

Cuadro 24
Estructura estándar nombre de reportes ad-hoc

1	2	3	4	5	6	7	8
P	Z	Inicial módulo	Inicial submódulo	Código numérico del Reporte Identificativo			

En donde:

- P: Identificativo del archivo que denomina a un Reporte.
- Z: Identificativo que representa a que los Reportes son de propiedad de la Universidad de las Fuerzas Armadas ESPE.

A continuación se muestra un ejemplo tomando en cuenta que el módulo es finanzas y el submódulo es roles y el esquema quedaría de la siguiente manera:

Cuadro 25
Ejemplo nombre reporte ad-hoc

1	2	3	4	5	6	7	8
P	Z	F	R	0001			

Formato del reporte en pentaho report designer

UNIVERSIDAD DE LAS FUERZAS ARMADAS ESPE.

Debe salir de una tabla denominada GUBINST del campo GUBINST_NAME y Filtrar por GUBINST_KEY =INS

Cuadro 26 Formato reportes

ORDEN FILA	NOMBRE	ORIGEN	NOMBRE TABLA	TAMAÑO LETRA	TIPO LETRA
1	Nombre de la Institución	TABLA	GUBINST	14	COMIC SANS - NEGRITA
2	Nombre del Reporte			10	COMIC SANS
3	Parámetros-Título			10	COMIC SANS - NEGRITA
4	Parámetros - Dato			10	COMIC SANS
5	Fecha lado superior derecho			8	COMIC SANS
6	Línea color negra			1	PTO
7	Títulos de Columnas			8	ARIAL - NEGRITA
8	Línea color negra			1	PTO
9	Datos o Contenido del Reporte			8	ARIAL
10	Pie de Firma Nombre			7	ARIAL - NEGRITA
11	Pie de Firma Cargo			7	ARIAL - NEGRITA
12	Página lado inferior derecha			7	ARIAL - NEGRITA

Ejemplo reporte

ESCUELA POLITECNICA DEL EJERCITO

EDIFICIO AULAS Y CAPACIDAD

CAMPUS: ESPE MATRIZ SANGOLQUI

Jueves Marzo 19 2011 14:00 PM

CEDULA	NOMBRE	ID	PAGOS MATRICULAS
1715141235	MENCIAS JARAMILLO VICTORIA	L000000123	451.25
1715141235	MENCIAS JARAMILLO VICTORIA	L000000123	451.25
1715141235	MENCIAS JARAMILLO VICTORIA	L000000123	451.25
1715141235	MENCIAS JARAMILLO VICTORIA	L000000123	451.25

EDISON SOSA

JEFE DE BIENES

Pág. 1 de 1

Figura 41: Ejemplo formato reporte

PANTALLA DE FILTROS

Se debe borrar las leyendas de Parámetros para el informe e introduzca parámetros para el informe, solo queda el título del reporte y los filtros.

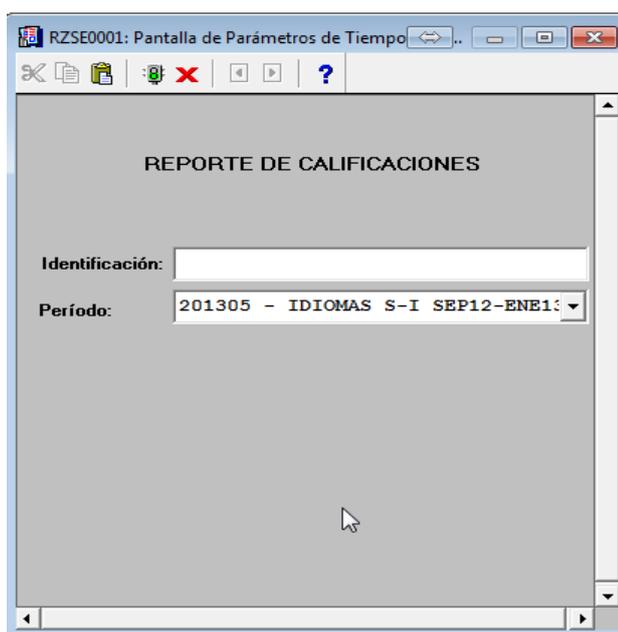


Figura 42: Pantalla filtros

Cuadro 27
Formato filtros

ORDEN FILAS	NOMBRE	ORIGEN	NOMBRE TABLA	TAMAÑO LETRA	TIPO LETRA
1	Título del reporte			10	COMIC SANS - NEGRITA
2	Título de los parámetros			10	COMIC SANS - NEGRITA

4.4.1.7 Directorio de las publicaciones de reportes y cubos en pentaho

Las rutas de los cubos y reportes van a variar en el caso de que existan diferentes carpetas generales y subcarpetas, pero para el caso nuestro los reportes y cubos serán publicados en el siguiente directorio:

FINANZAS

- **REPORTES:**/opt/pentaho/biserver-ce/pentaho-solutions/DATAMART_FINANZAS_RRHH/FINANZAS/REPORT
- **CUBOS:**/opt/pentaho/biserver-ce/pentaho-solutions/DATAMART_FINANZAS_RRHH/FINANZAS/WORKBENCH

PERSONAL

- **REPORTES:**/opt/pentaho/biserver-ce/pentaho-solutions/DATAMART_FINANZAS_RRHH/PERSONAL/REPORT
- **CUBOS:**/opt/pentaho/biserver-ce/pentaho-solutions/DATAMART_FINANZAS_RRHH/PERSONAL/WORKBENCH

Contraseña para publicaciones de cubos y reportes

La contraseña para la publicación de los cubos y los reportes quedo establecida de la siguiente manera: **p4ssw0rd**

Nombres de las publicaciones de reporte y cubos en pentaho

Los nombres de publicación para los reportes será el mismo nombre con el que se guarda el reporte y lo mismo se aplicará para los cubos

4.4.2 Esquema de base de datos en la fuente y el repositorio temporal

Después del análisis de los requerimientos se diseñó un esquema de base de datos para cada uno de los módulos involucrados en el proyecto, estos diagramas fueron diseñados a partir de las tablas de las base de datos relacionales y a partir de archivos planos.

A continuación se muestran cada uno de los diagramas diseñados para cada uno de los módulos dentro de las unidades administrativa y financiera, estos diagramas

servirán para generar la base de datos temporal la cual a su vez poblará los datos de nuestro datamart.

4.4.2.1 Diagrama General Presupuestos

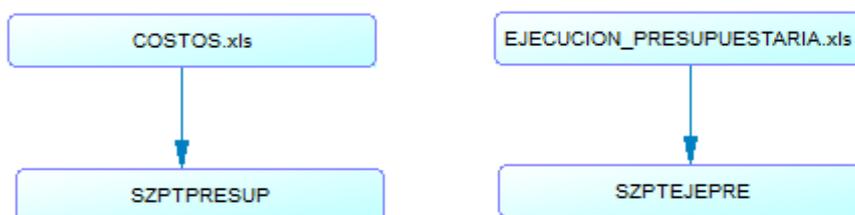


Figura 43: Diagrama general presupuestos

4.4.2.2 Mapeo de Valores de la fuente al repositorio temporal módulo presupuestos

Cuadro 28

Mapeo de valores de la fuente al repositorio temporal de presupuestos

FUENTE DE DATOS ORIGINAL (archivos .xls)	DESCRIPCIÓN	STAGE
COSTOS.xls	Tabla que contiene todos los valores de los distintos programas	szptpresup
EJECUCION_PRESUPUESTARIA.xls	Tabla que contiene los códigos y los nombres de los programas, actividades etc.	szptejepre

4.4.2.3 Diagrama de Archivos Planos usados para el módulo de presupuestos

COSTOS.xls	
FECHA	Date
DESCRIPCIONG1	String
DESCRIPCIONG2	String
COL1	Number
COL2	Number
COL3	Number
COL4	Number
COL5	Number
COL6	Number
COL7	Number
COL8	Number
COL9	Number
COL10	Number
COL20	Number
CODIGOG1	String
CODIGOG2	String

EJECUCIÓN_PRESUPUESTARIA.xls	
EJERCICIO	String
SECTOR	String
NOMBRE	String
SECTORIAL	String
NOMBRE	String
ENTIDAD	String
UNIDAD_EJECUTORA	String
NOMBRE	String
PROGRAMA	String
NOM_ESTRUCTURA	String
SUBPROGRAMA	String
NOM_ESTRUCTURA	String
PROYECTO	String
NOM_ESTRUCTURA	String
ACTIVIDAD	String
NOM_ESTRUCTURA	String
GRUPO	String
ITEM	String
NOMBRE	String
GEOGRAFICO	String
FUENTE	String
ORGANISMO	String
CORRELATIVO	String
FUNCION	String
INICIAL	String
CODIFICADO	String
PRECOMPROMISO	String
COMPROMISO	String
DEVENGADO	String
PAGADO	String

Figura 44: Diagrama de archivos planos para presupuestos

4.4.2.4 Diagrama de base de datos entidad relación en el repositorio temporal para el cubo de presupuestos

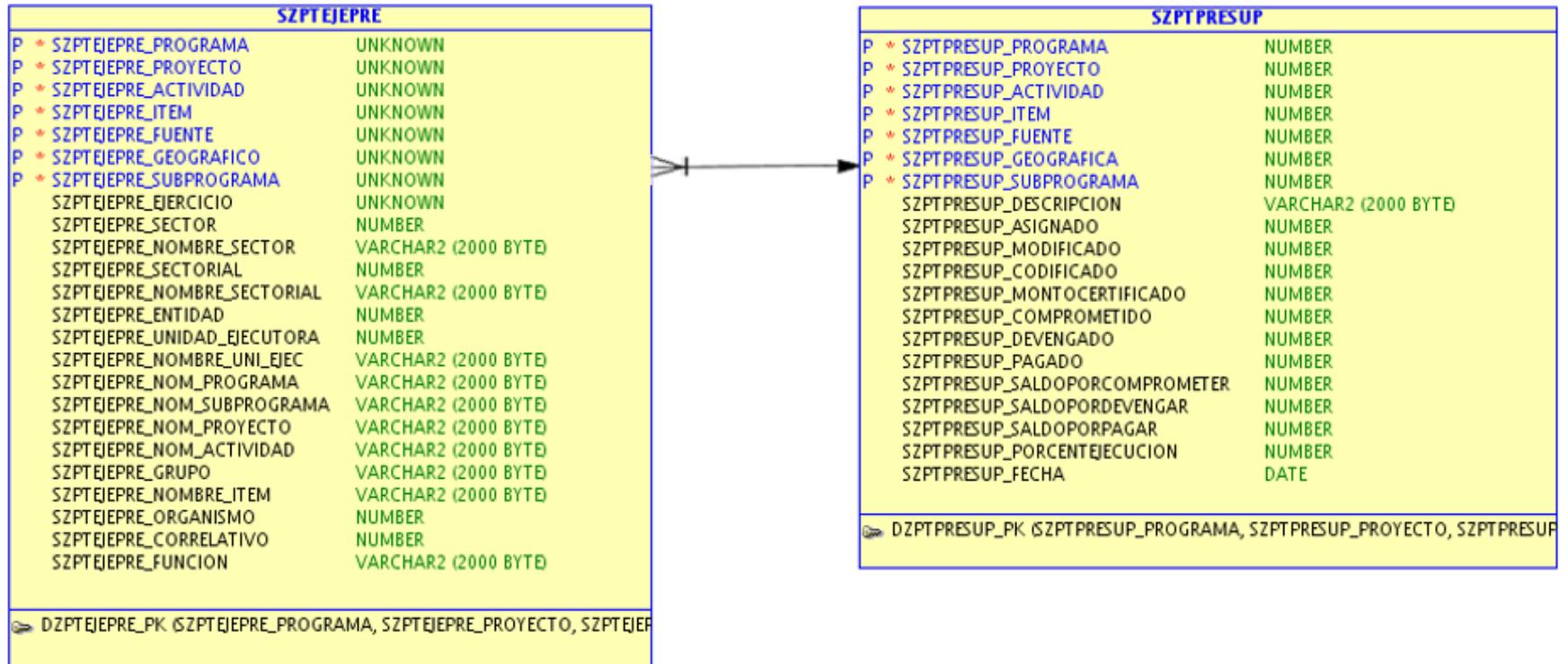


Figura 45: Modelo entidad relación presupuesto

4.4.2.5 Diagrama general Activo Fijos

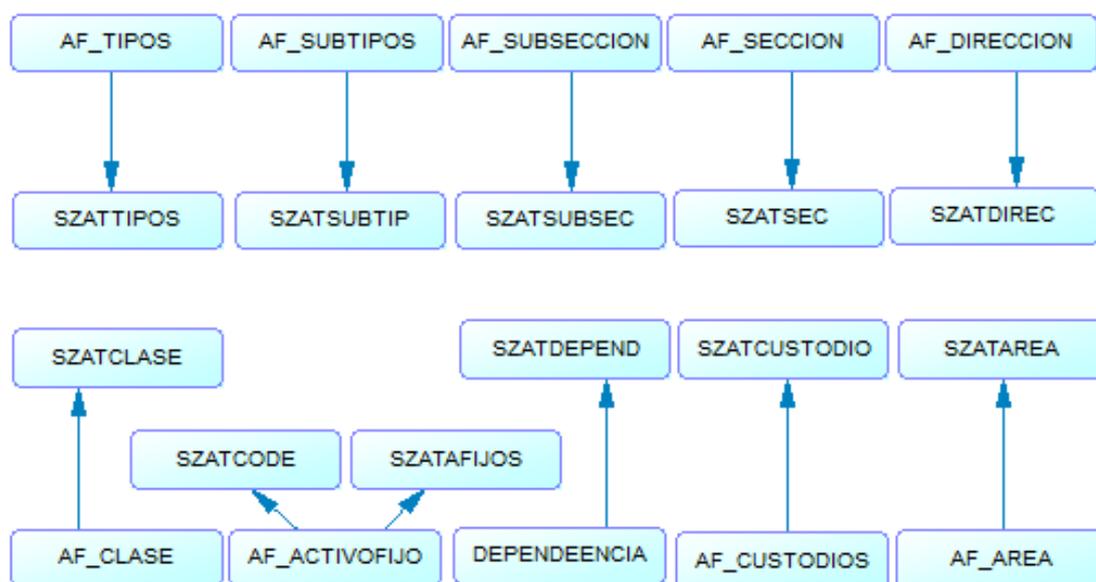


Figura 46: Diagrama general activos fijos

4.4.2.6 Mapeo de valores de la fuente al repositorio temporal módulo activos fijos

Cuadro 29
Mapeo de valores de la fuente al repositorio temporal de activos fijos

FUENTE DE DATOS ORIGINAL	DESCRIPCIÓN	STAGE
af_tipos	Tabla que contiene los códigos y las descripciones de los tipos.	szattipos
af_subtipos	Tabla que contiene los códigos y las descripciones de los subtipos.	szatsubtip
af_subseccion	Tabla que contiene los códigos y las descripciones de las subsecciones.	szatsubsec
af_seccion	Tabla que contiene los códigos y las descripciones de las secciones.	szatsec
af_direccion	Tabla que contiene los códigos y las descripciones de las direcciones.	szatdirec
af_area	Tabla que contiene los códigos y las descripciones de las áreas.	szatarea
af_clase	Tabla que contiene los códigos y las descripciones de las clases.	szatclase
af_activofijo	Tabla que contiene los códigos para el código del bien.	szatcode
dependencia	Tabla que contiene los códigos de la empresa, el nombre de la institución y la localidad.	szatdepend
af_custodios	Tabla que contiene los bienes y custodios.	szatcustodio
af_activofijo	Tabla que contiene las medidas para los activos.	szatafijos

4.4.2.8 Diagrama de base de datos entidad relación en el repositorio temporal para el cubo de activos fijos

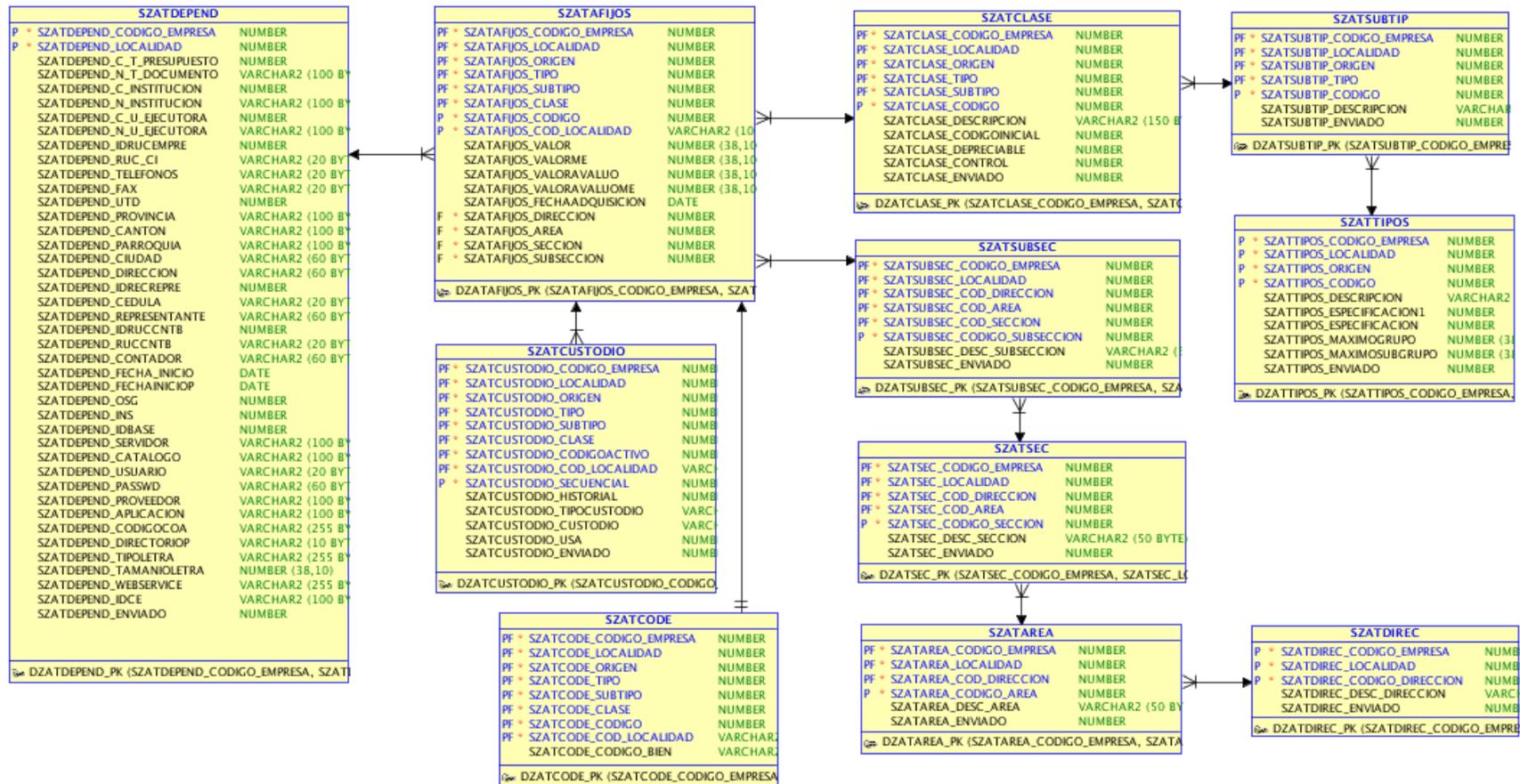


Figura 48: Modelo entidad relación activos fijos

4.4.2.9 Diagrama general inventarios

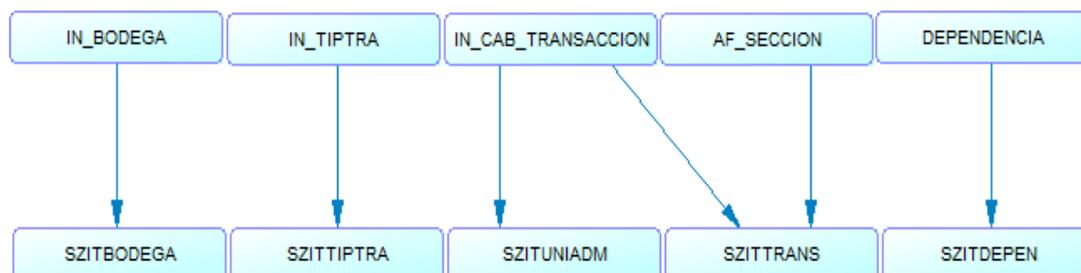


Figura 49: Diagrama general inventarios

4.4.2.10 Mapeo de valores de la fuente al repositorio temporal módulo inventarios

Cuadro 30

Mapeo de valores de la fuente al repositorio temporal de inventarios

FUENTE DE DATOS ORIGINAL	DESCRIPCIÓN	STAGE
in_bodega	Tabla que contiene los códigos y las descripciones de las bodegas.	szitbodega
in_tiptra	Tabla que contiene los tipos de transacción del inventario.	szittiptra
in_cab_transaccion	Tabla que contiene las unidades operativas de inventario.	szituniadm
af_seccion	Tabla que contiene los rubros (ingresos y egresos) para los inventarios.	szittrans
dependencia	Tabla que contiene los códigos de la empresa, el nombre de la institución y la localidad.	szitdepend
in_cab_transaccion	Tabla que contiene las medidas para inventario.	szittrans

4.4.2.11 Diagrama de base de datos de la Fuente para el módulo de Inventarios

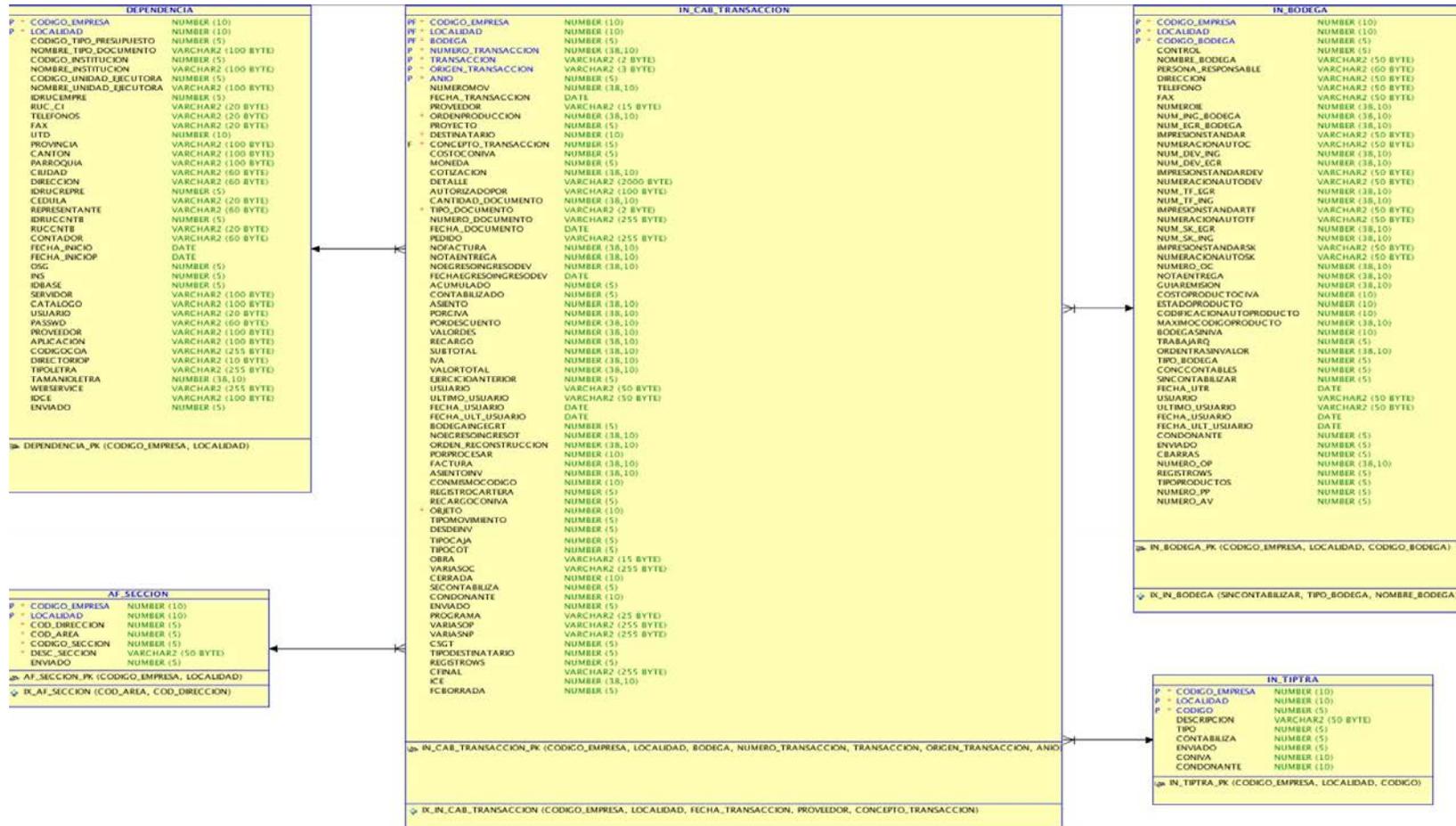


Figura 50: Diagrama de la fuente de inventarios

4.4.2.1 Diagrama de base de datos entidad relación en el repositorio temporal para el cubo de inventarios

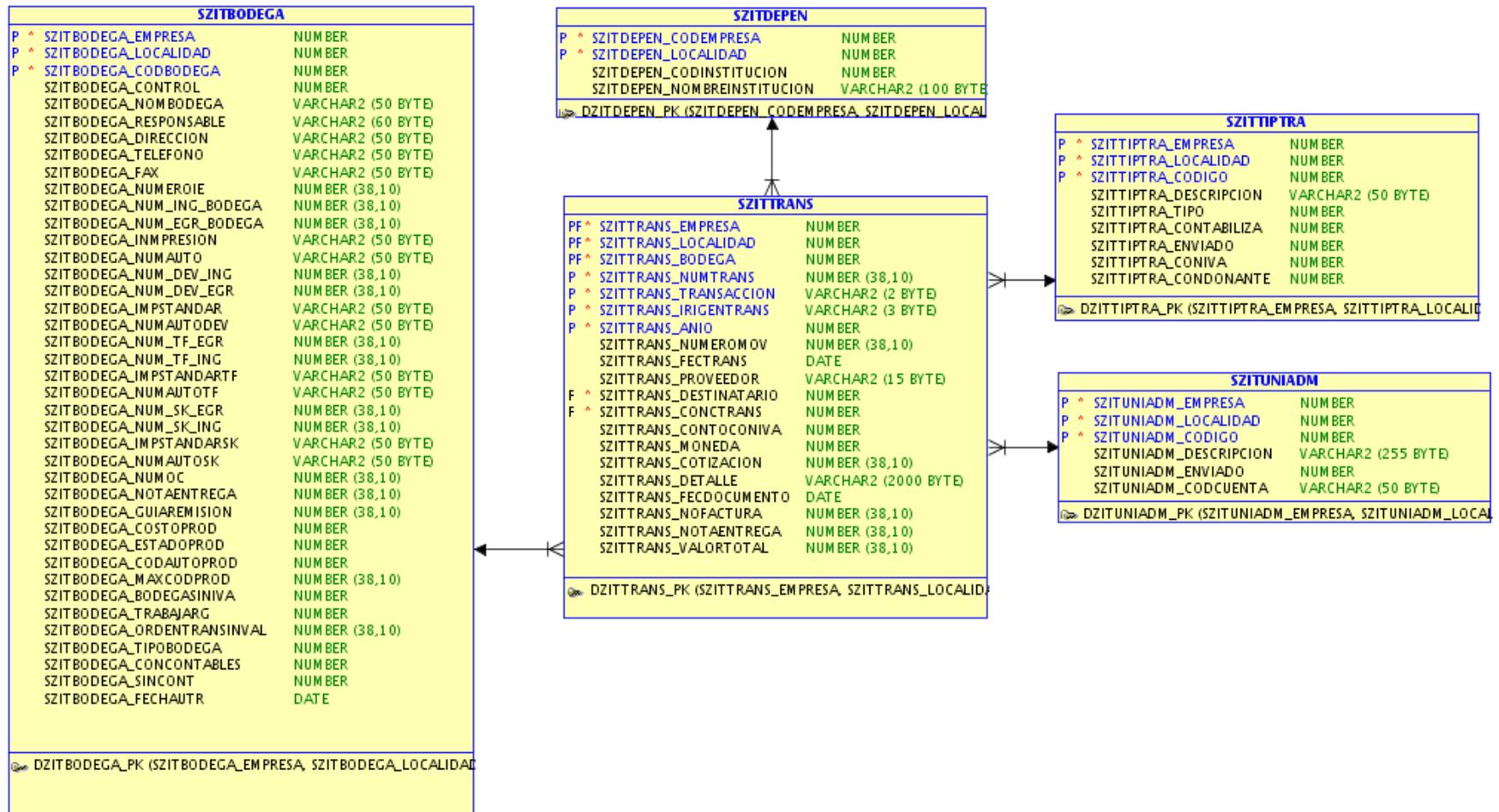


Figura 51: Modelo entidad relación inventario

4.4.2.2 Diagrama general roles de pago - shifre

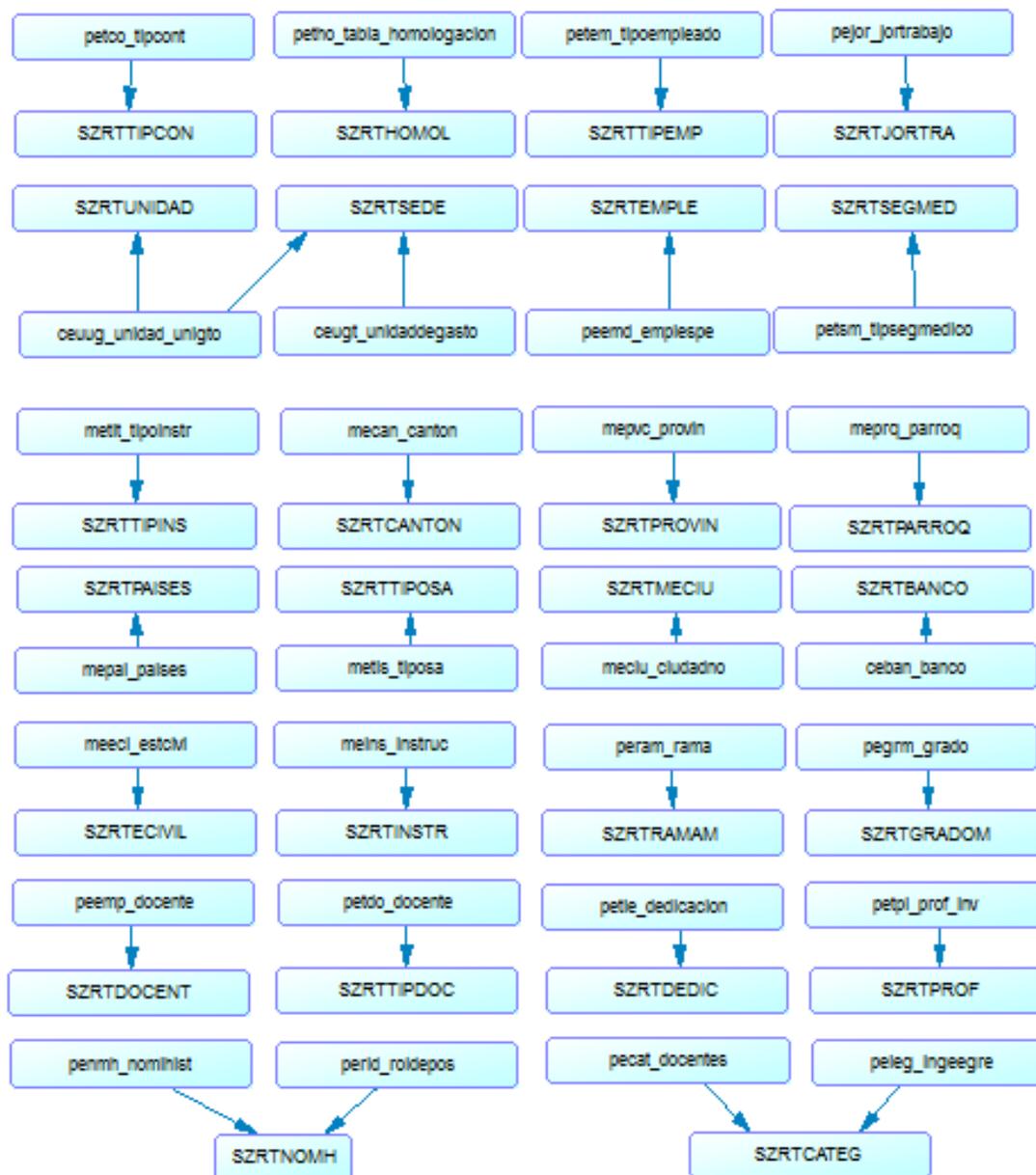


Figura 52: Diagrama general roles de pago - shifre

4.4.2.3 Mapeo de valores de la fuente al repositorio temporal módulo roles de pago - shifre

Cuadro 31
Mapeo de valores de la fuente al repositorio temporal de roles de pago - shifre

FUENTE DE DATOS ORIGINAL	DESCRIPCIÓN	STAGE
peemd_emplespe petho_tabla_homologacion petem_tipoempleado pejor_jortrabajo petco_tipcont petsm_tipsegmedico	Tabla que contiene los datos del empleado de la Universidad de las Fuerzas Armadas ESPE.	szrtemple szrttipemp szrthomol szrtjortra szrttipcon szrtsegmed szrtunidad szrtsede
ceuug_unidad_unigto	Tabla que contiene los cargos que tiene el empleado de la Universidad de las Fuerzas Armadas ESPE.	szrtunidad
petem_tipoempleado	Tabla que contiene el tipo de empleado que tiene cada persona.	szrttipemp
metit_tipoinstr mecan_canton mepvc_provin meprq_parroq mepai_paises metis_tiposa meciu_ciudadno ceban_banco meeci_estcivl meins_instruc peram_rama pegrm_grado	Tabla que contiene los datos personales de cada persona.	szrtmeciu szrttiposa szrtprovin szrtcanton szrtparroq szrtpaises szrtcecivil szrtinstr szrtgradom szrtbanco szrttramam szrttipins szrtdocent szrttipdoc szrtdedic szrtprof szrtcateg szrtunidad szrttipcon szrtsede szrtemple
peemp_docente petdo_docente petie_dedicacion petpi_prof_inv pecat_docentes	Tabla que contiene los datos acerca de los docentes de la Universidad de las Fuerzas Armadas ESPE.	szrtdocent szrttipdoc szrtdedic szrtprof szrtcateg szrtunidad szrttipcon szrtsede szrtemple
ceuug_unidad_unigto	Tabla que contiene los departamentos que existen en la Universidad de las Fuerzas Armadas ESPE.	szrtunidad
peieg_ingeege	Tabla que contiene los ingresos y los egresos que tiene cada empleado de la Universidad de las Fuerzas Armadas ESPE.	szrtingegr
ceugt_unidaddegasto ceuug_unidad_unigto	Tabla que contiene las sedes de la Universidad de las Fuerzas Armadas ESPE.	szrtsede
penmh_nomihist perld_roldepos	Tabla que tiene los sueldos de los empleados de la Universidad de las Fuerzas Armadas ESPE.	szrtnomih

4.4.2.5 Diagrama de base de datos entidad relación en el repositorio temporal para el cubo de roles de pago – shifre

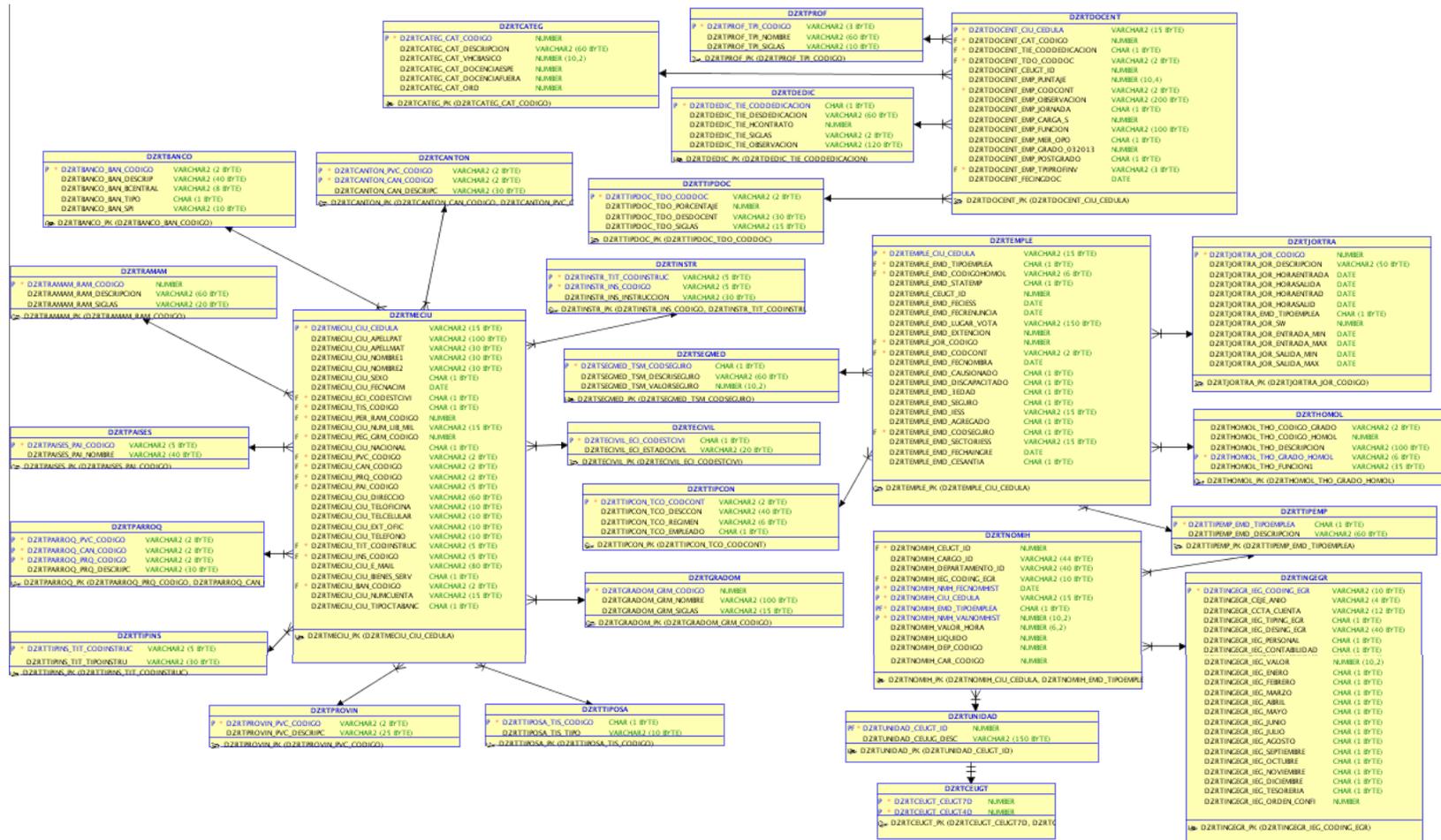


Figura 54: Modelo entidad relación roles de pago - Shifre

4.4.2.6 Diagrama general matriculación

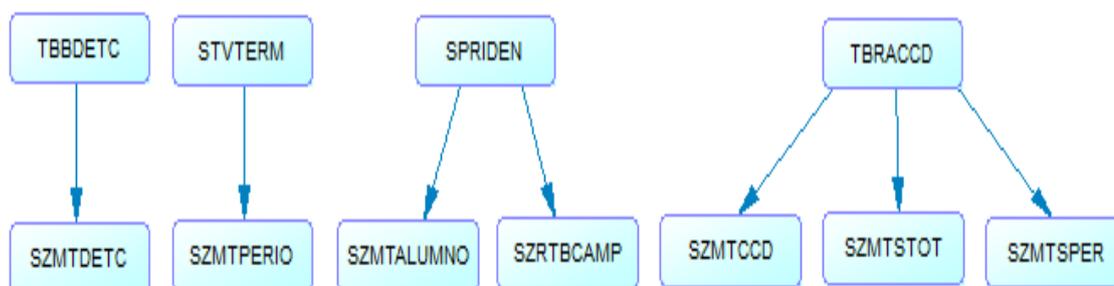


Figura 55: Diagrama general matriculación

4.4.2.7 Mapeo de valores de la fuente al repositorio temporal módulo de matriculación

Cuadro 32

Mapeo de valores de la fuente al repositorio temporal de matriculación

FUENTE DE DATOS ORIGINAL	DESCRIPCIÓN	STAGE
tbbdetc	Tabla que contiene los aranceles agrupados y desglosados de la Universidad de las Fuerzas Armadas ESPE.	szmtdetc
stvterm	Tabla que contiene los periodos del sistema banner de la Universidad de las Fuerzas Armadas ESPE.	szrtbperio
spriden, tbraccd	Tabla que contiene los nombres y tipos de alumno.	szmtalumno
spriden	Tabla que contiene los campus de la Universidad de las Fuerzas Armadas ESPE.	szrtbcamp
tbraccd	Tabla que contiene los pagos y deudas de los alumnos por periodo.	szmtalumno szmtccd
tbraccd	Tabla que contiene los pagos y deudas de los alumnos por tiempo.	szmtalumno szmtccd
tbraccd	Tabla que contiene los pagos y deudas de los alumnos por tiempo.	szmtstot szmtasper

4.4.2.8 Diagrama de base de datos de la Fuente para el módulo de matriculación

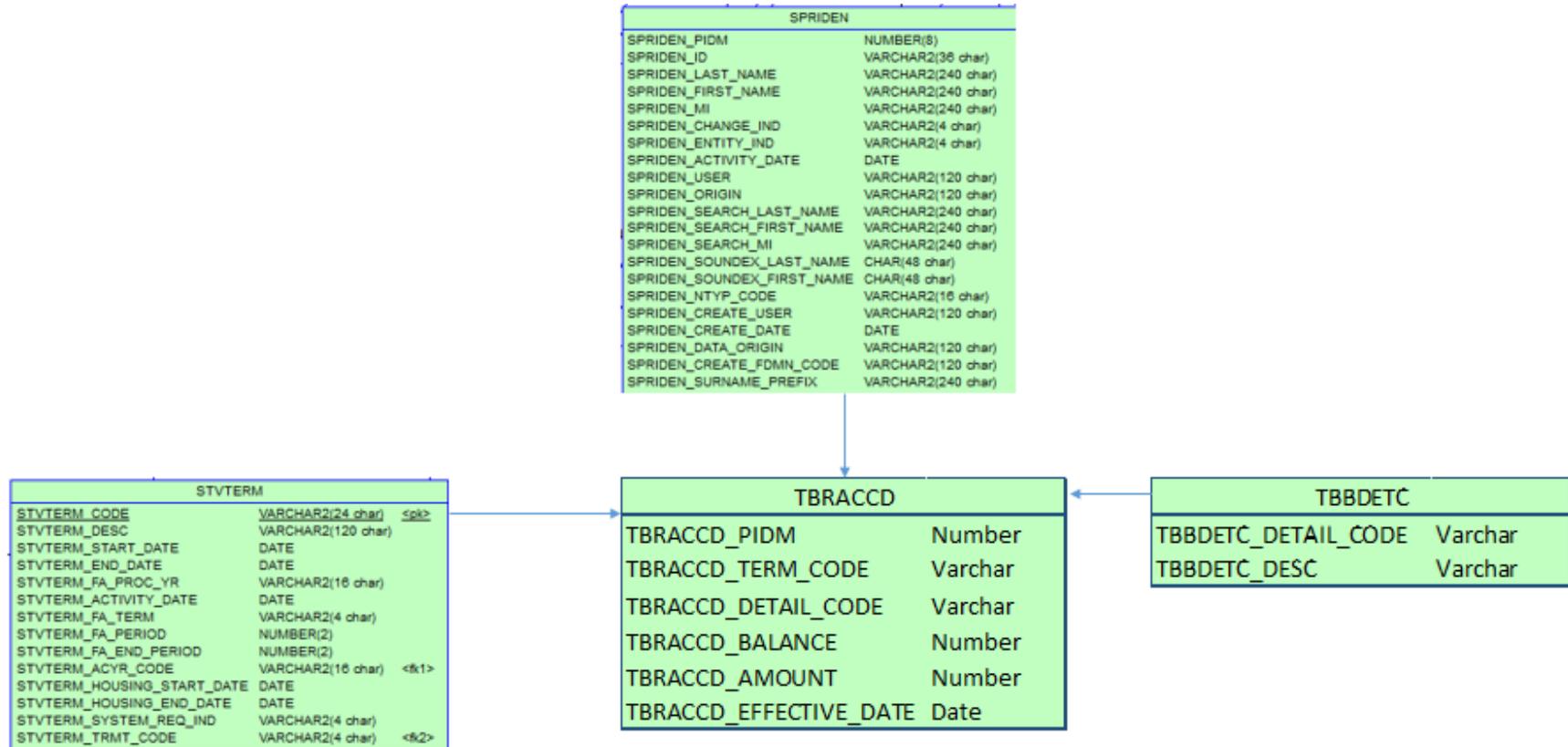


Figura 56: Diagrama de la fuente de matriculación

4.4.2.9 Diagrama de base de datos entidad relación en el repositorio temporal para el cubo de matrículas

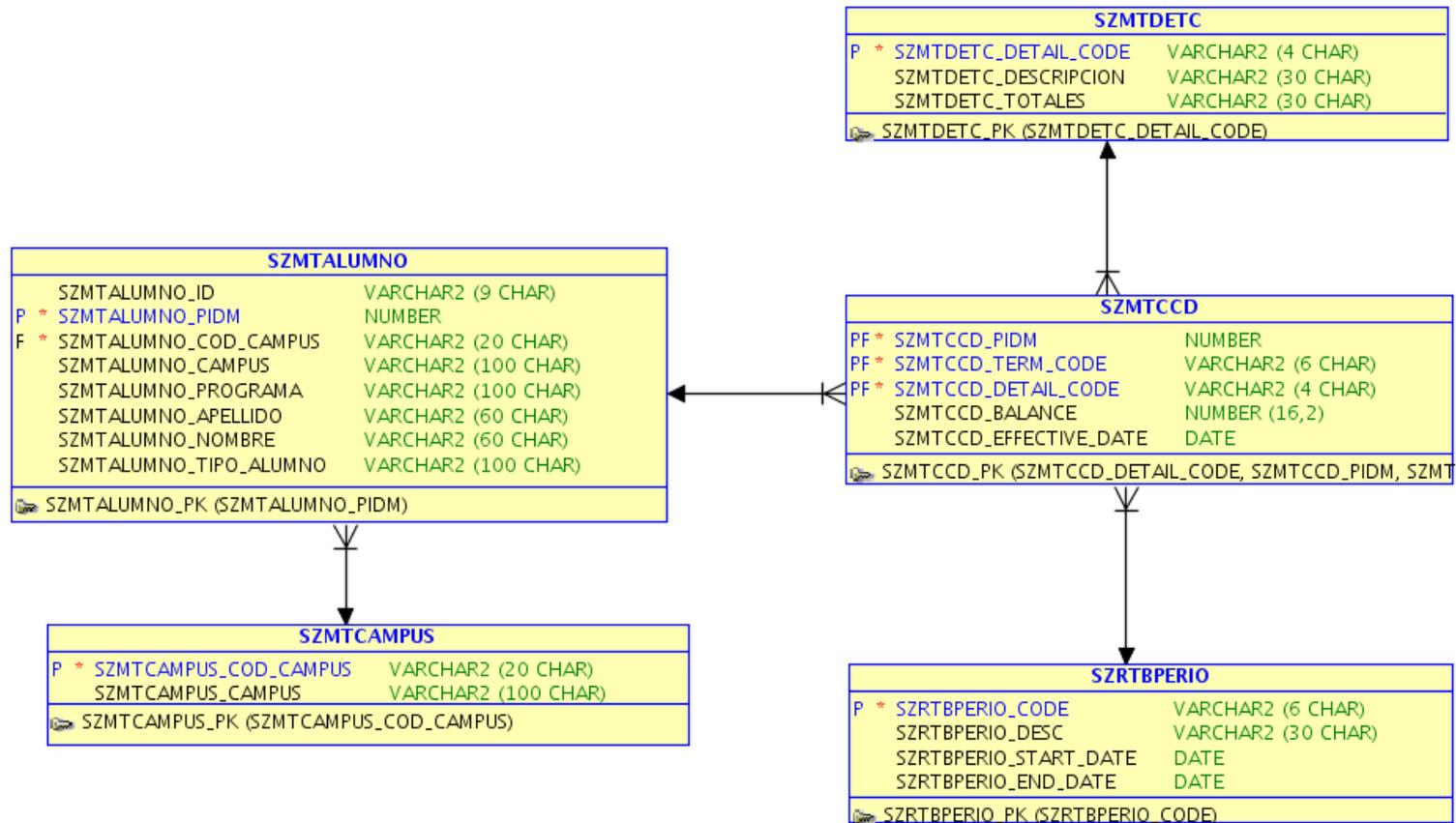


Figura 57: Modelo entidad relación matrículas

4.4.2.10 Diagrama general roles de pago - banner

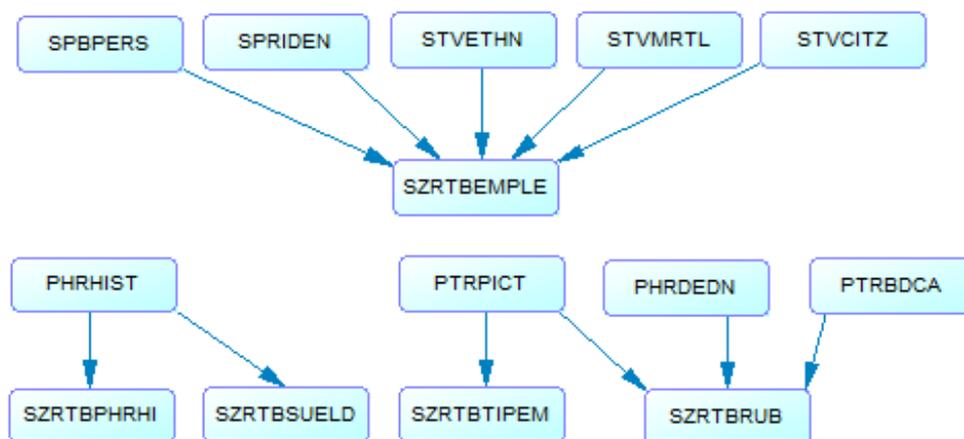


Figura 58: Diagrama general roles de pago - banner

4.4.2.11 Mapeo de valores de la fuente al repositorio temporal módulo roles de pago - banner

Cuadro 33

Mapeo de valores de la fuente al repositorio temporal de roles de pago - banner

FUENTE DE DATOS ORIGINAL	DESCRIPCIÓN	STAGE
spbpers spriden stvethn stvmrtl stvcitz	Tablas que contienen todos los datos personales de los empleados administrativos, docentes y alumnos.	szrtbemple
phrhist	Tablas que contienen las disposiciones de la Universidad de las Fuerzas Armadas ESPE.	szrtbphrhi
ptrpict	Tablas que contienen los distintos tipos de empleados de la Universidad de las Fuerzas Armadas ESPE.	szrtbtipem
phrdedn ptrbdca ptrpict	Tablas que contienen los rubros y los tipos de rubros.	szrtbrub
phrhist	Tabla que contiene los sueldos de los empleados.	szrtbsueld szrtbphrhi

4.4.2.12 Diagrama de base de datos de la fuente del módulo de roles de pago – banner

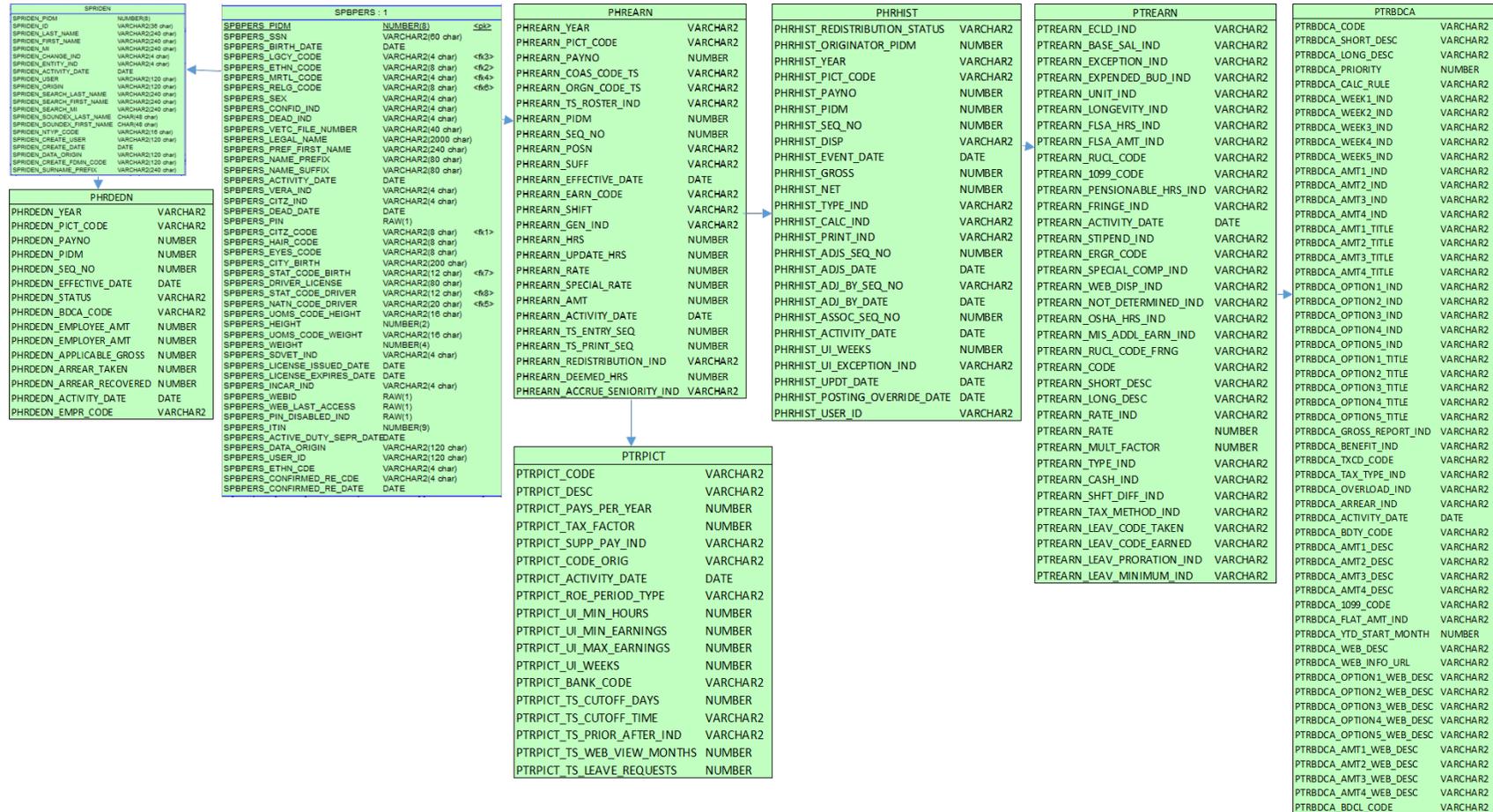


Figura 59: Diagrama de la fuente de roles de pago – banner

4.4.2.13 Diagrama de base de datos entidad relación en el repositorio temporal para los cubos de roles de pago - banner

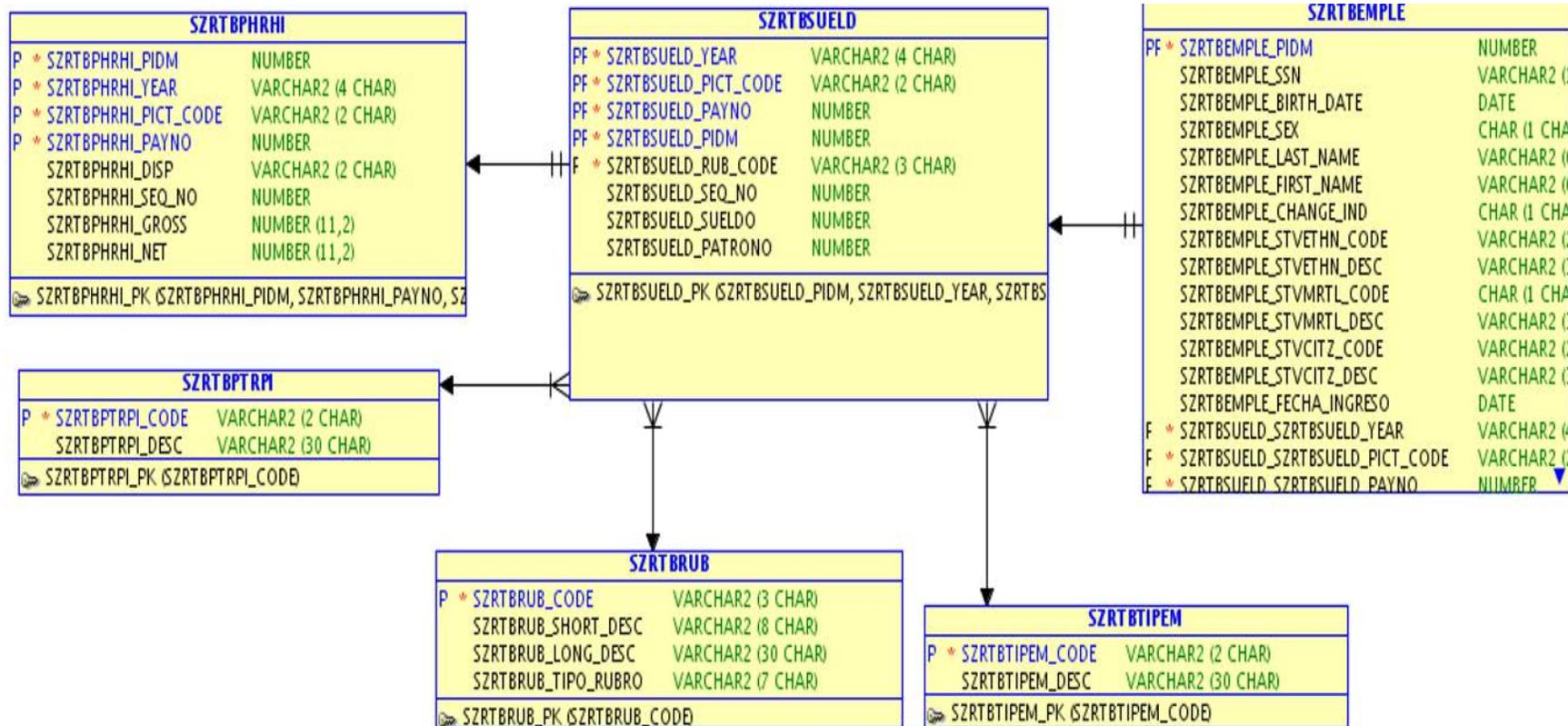


Figura 60: Modelo entidad relación roles de pago - banner

4.4.2.14 Diagrama general cuentas por cobrar

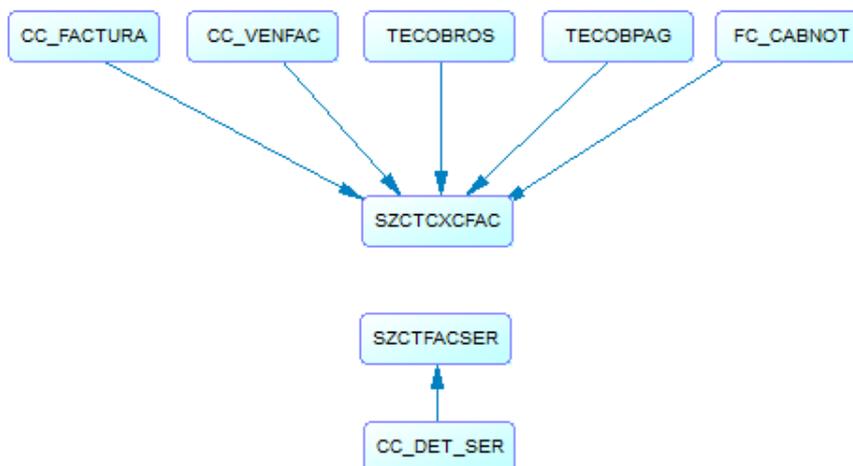


Figura 61: Diagrama general cuentas por cobrar

4.4.2.15 Mapeo de valores de la fuente al repositorio temporal módulo cuentas por cobrar

Cuadro 34

Mapeo de valores de la fuente al repositorio temporal de cuentas por cobrar

FUENTE DE DATOS ORIGINAL	DESCRIPCIÓN	STAGE
cc_factura cc_venfac tecobros tecobpag fc_cabnot	Tabla que contiene las cédulas y los nombres de los clientes.	szctcxcfac
cc_det_ser	Tabla que contiene los números de facturas con su detalle.	szctfacser
cc_factura cc_venfac tecobros tecobpag fc_cabnot	Tabla que contiene la medida de saldo para cuentas.	szctcxcfac

4.4.2.17 Diagrama de base de datos entidad relación en el repositorio temporal para los cubos de cuentas por cobrar.

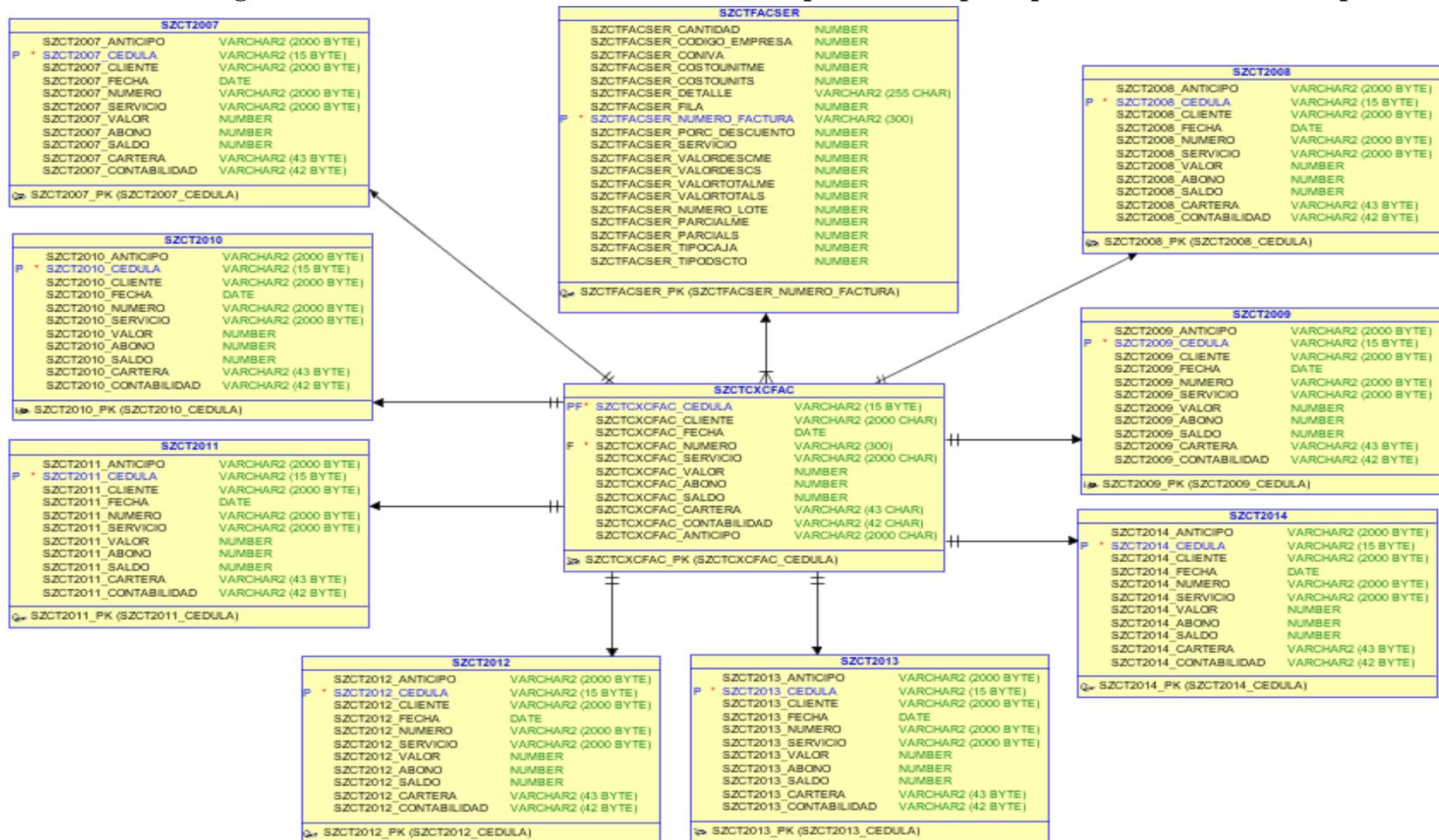


Figura 63: Modelo entidad relación cuentas por cobrar

4.4.3 Entorno Back Room

Los datos para la población del datamart, serán extraídos en parte del servidor de base de datos Oracle de la Universidad de las Fuerzas Armadas ESPE y también de archivos xls. El proceso ETL se lo realizará mediante la herramienta Pentaho Data Integration (PDI), de la suite de Pentaho y el datamart será alojado un servidor de base de datos Oracle 10g.

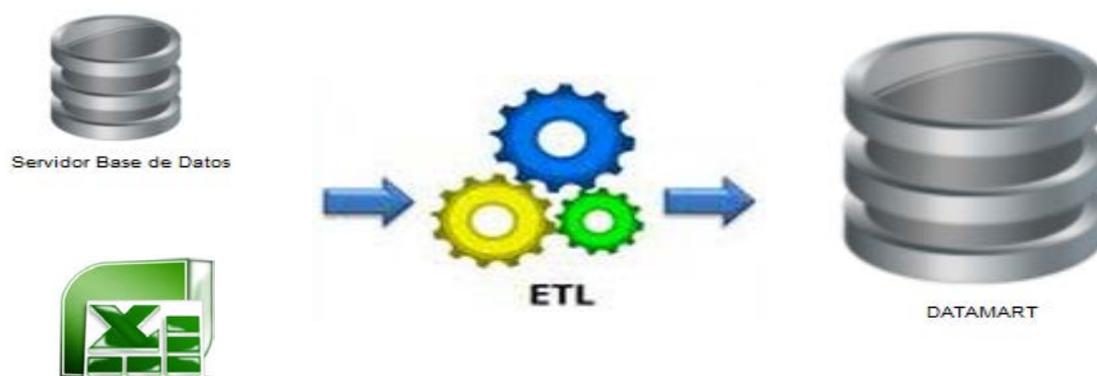


Figura 64: Back Room

4.4.4 Entorno Front Room

Una vez poblado el datamart, la información podrá ser visualizada y analizada mediante la consola de usuario de Pentaho en diversas formas como son: vistas de análisis, reportes ad hoc y dashboards.

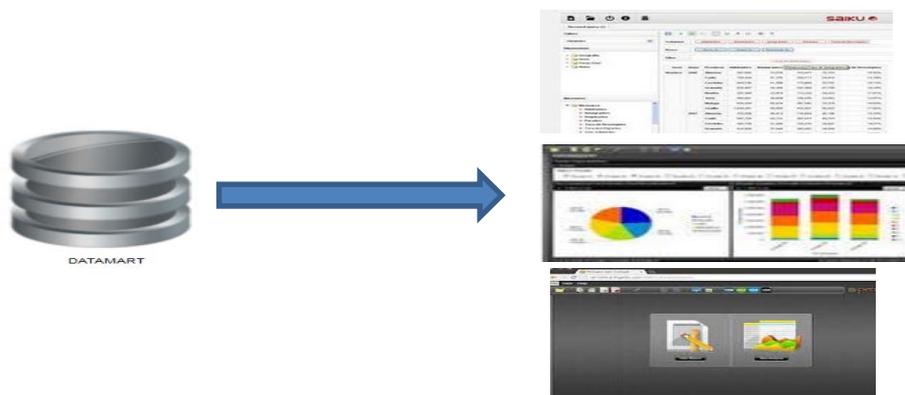


Figura 65: Front Room

4.5 SELECCIÓN DE PRODUCTOS E INSTALACIÓN

Para el desarrollo del proyecto de BI, se trabajó con el conjunto de herramientas de Pentaho Community, detalladas en el siguiente cuadro.

Cuadro 35
Selección de productos e instalación

Producto	Característica	Uso	Hardware
Pentaho BI Server 4.8	Servidor OLAP	Administración de roles y usuarios. Conexiones JNDI. Publicación de reportes. Publicación de Cubos. Elaboración de reportes ad hoc Elaboración de dashboards Creación de carpetas y asignación de permisos	Centos 6 12 GB RAM Disco duro 1TB QuadCore
Pentaho Data Integration Stable Release - 4.4.0	Diseñador gráfico de transformaciones y trabajos del sistema de ETL's	Diseño y elaboración del proceso del proceso de extracción transformación y carga (ETL). Diseño y elaboración de Jobs (modo batch), para ejecutarlos automáticamente en intervalos regulares.	Centos 6 12 GB RAM Disco duro 1TB QuadCore
Pentaho Schema Workbench 3.2.1.13885	Herramienta empleada para el diseño y publicación de cubos OLAP	Permite la construcción de los esquemas del mondrian. Permite publicar esquemas y cubos en el servidor BI, para ser utilizados en los análisis por los usuarios de la plataforma. Se define las relaciones entre dimensiones y medidas de un cubo, con las tablas y los campos de la base de datos (dimensional)	Hardware de acuerdo a equipo de usuario
Pentaho Report Designer 3.9.0-GA	Herramienta para construcción de reportes	Permite el diseño, elaboración y publicación de reportes dentro del servidor BI, para su posterior ejecución y análisis.	Hardware de acuerdo a equipo de usuario
CDE 13.06.05	Community Dashboards editor	Diseño y construcción de tableros de control.	Centos 6 12 GB RAM Disco duro 1TB QuadCore
Saiku Analytics 2.5	Complemento para vistas de análisis	Permite manipular los cubos publicados en el servidor BI, presentado una interfaz amigable y sencilla para el usuario.	12 GB RAM Disco duro 1TB QuadCore
Saiku Reporting 1.0	Complemento para reportes ad-hoc	Permite elaborar reportes ad-hoc de las fuentes de datos existentes en la consola de usuario del servidor BI.	12 GB RAM Disco duro 1TB QuadCore

4.6 MODELADO DIMENSIONAL

4.6.1 Bus Matrix Data Warehouse

MODULOS	Finanzas						Personal			
	Presupuesto	Activos fijos	Inventario	Matriculas		Cuentas por cobrar	Roles Shifre	Roles Banner		
TABLA DE HECHOS	fac_presupuesto	fac_activos	fac_inventario	fac_matriculast	fac_matriculasp	fac_cuentas	fac_rolas	fac_rolas_b_emple	fac_rolas_b_docen	fac_rolas_sueldo_b
dim_programa	x									
dim_subprograma	x									
dim_proyecto	x									
dim_actividad	x									
dim_item	x									
dim_geografico	x									
dim_fuente	x									
dim_tipo		x								
dim_subtipo		x								
dim_clase		x								
dim_area		x								
dim_seccion		x								
dim_subseccion		x								
dim_origen		x								
dim_direccion		x								
dim_codigobien		x								
dim_dependencia		x	x							
dim_tiempo	x	x	x	x		x	x			x
dim_bodega			x							
dim_tiptra			x							
dim_uniadm			x							
dim_rubro			x							
dim_emplespe							x			
dim_cargo							x			
dim_tipoempleado							x			
dim_ciudadano							x			
dim_ingegr							x			
dim_departamento							x			
dim_docente							x			
dim_campus				x	x					
dim_periodo					x			x	x	
dim_arancel				x	x					
dim_alumno				x	x					
dim_direccion_b								x		
dim_fecsalida_b								x		
dim_nombramiento_b								x		
dim_campus_b								x	x	
dim_departamento_b								x	x	
dim_carrera_b									x	
dim_asignatura_b									x	
dim_diruni_b								x		
dim_email_b								x		
dim_nbrjobs_b								x		
dim_publicaciones_b								x		
dim_discapacitado_b								x		
dim_titulos_b								x		
dim_emple_b								x	x	x
dim_rubro_b										x
dim_tipoempleado_b								x		
dim_tipodocente_b									x	
dim_cliente						x				
dim_factura						x				

Figura 66: Matriz datawarehouse

4.6.2 Modelado Dimensional

El modelo lógico para el área administrativa y financiera de la Universidad de las Fuerzas Armadas ESPE, fue realizado siguiendo el esquema de estrella que optimiza el tiempo de respuesta en consultas complejas, siguiendo los pasos del modelo dimensional de Ralph Kimball, en nuestro caso se han creado modelos para cada uno de los módulos que intervienen en las áreas mencionadas, entre estos constan los siguientes: módulo de presupuestos, módulo de activos fijos, módulo de inventarios, módulo de roles de pago y módulo cuentas por cobrar. Se creó dimensiones que contienen las características de las entidades del negocio y tablas de hechos que contienen las medidas o cantidades numéricas para el análisis del negocio. A continuación se detalla el proceso de modelado para cada uno de los módulos mencionados.

4.6.2.1 Diagrama general para el módulo de presupuestos

El diagrama muestra las dimensiones y la relación que tienen con las tablas de la base de datos de origen.

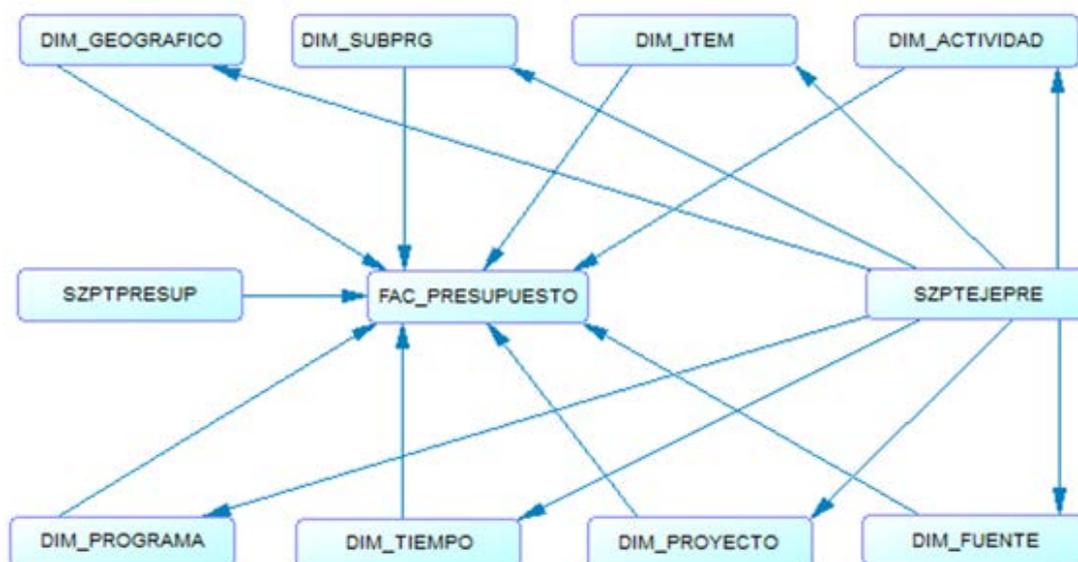


Figura 67: Diagrama general de presupuestos

4.6.2.2 Mapeo de datos presupuestos

En la siguiente tabla se describe el origen de datos desde el servidor de base de datos o archivos planos hacia las dimensiones y tabla de hechos del datamart.

Cuadro 36
Mapeo de datos de presupuesto

FUENTE DE DATOS ORIGINAL	DESCRIPCIÓN	STAGE	DIMENSIÓN (DIM)	HECHOS (FAC)
(archivo plano)	Tabla que contiene los códigos y los nombres de los proyectos.	szptejepre	DIM_PROYECTO	
(archivo plano)	Tabla que contiene los códigos y los nombres de los programas.	szptejepre	DIM_PROGRAMA	
(archivo plano)	Tabla que contiene los códigos y los nombres de las actividades.	szptejepre	DIM_ACTIVIDAD	
(archivo plano)	Tabla que contiene los códigos y los nombres de los subprogramas.	szptejepre	DIM_SUBPROGRAMA	
(archivo plano)	Tabla que contiene los códigos y los nombres de los items.	szptejepre	DIM_ITEM	
(archivo plano)	Tabla que contiene los códigos de la ubicación geográfica.	szptejepre	DIM_GEOGRAFICO	
(archivo plano)	Tabla que contiene los códigos de las fuentes.	szptejepre	DIM_FUENTE	
(archivo plano)	Tabla que contiene las medidas para presupuestos.	szptpresup		FAC_PRESUPUESTO

4.6.2.3 Modelo Dimensional Presupuestos

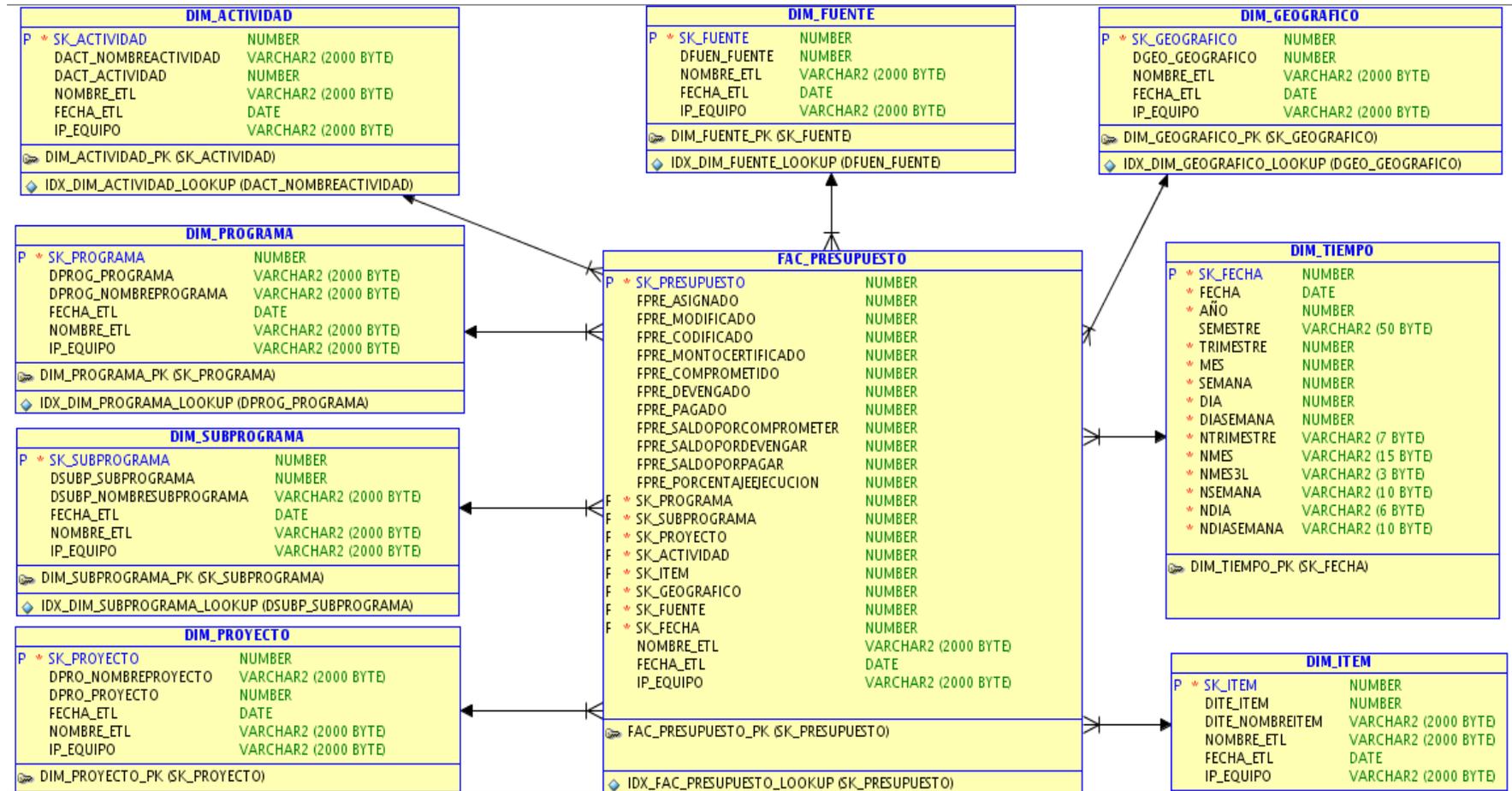


Figura 68: Modelo dimensional de presupuestos

4.6.2.4 Mapeo de datos para tablas de dimensiones de Presupuestos

NOTA: El esquema para el mapeo de datos de tablas de dimensiones y tablas de hechos de todos los módulos del proyecto, tendrá un formato que se especifica a continuación:

- **TABLA:** Nombre de la tabla de dimensión o tabla de hechos.
- **TIPO DE TABLA:** Dimensional o hechos.
- **DESCRIPCIÓN:** Descripción de la información que contiene la tabla.
- **ESQUEMA:** Nombre del usuario donde se encuentra la tabla de dimensión o tabla de hechos.

DESTINO

- **Columna:** Se almacena los campos de la tabla de dimensiones o tabla de hechos.
- **Descripción:** Es una descripción de la información que contiene cada campo de la tabla de dimensiones o tabla de hechos.
- **Tipo de dato:** Es el tipo de dato que tiene cada campo con sus respectivas longitudes.
- **Clave:** Se almacena las claves primarias con la palabra **pk** y claves secundarias con la palabra **fk** de la tabla.
- **Fk para:** Se coloca el nombre de la tabla de hechos en la fila donde se encuentra la clave primaria porque la clave primaria de la dimensión se va a convertir en clave secundaria en la tabla de hechos.
- **Null:** Se coloca la palabra **Y** a los campos que pueden tener datos nulos y **N** a los campos que no pueden tener datos nulos.

ORIGEN

- **Esquema:** Es el nombre del usuario. En el caso que se coloque la palabra **sistema**, quiere decir que la información se extrae de un proceso ETL en la herramienta spoon.

- **Tabla:** Es el nombre de la tabla de donde se extrae la información. En el caso que se coloque la palabra **sistema**, quiere decir que la información se extrae de un proceso ETL en la herramienta spoon.
- **Campo:** El nombre del campo que se utilizan para la extracción de la información. En el caso que se coloque la palabra **sistema**, quiere decir que la información se extrae de un proceso ETL en la herramienta spoon.
- **Tipo de dato:** Se coloca el tipo de dato para cada campo con sus respectivas longitudes.

Las siguientes tablas muestran el destino y origen de los datos para dimensiones y tabla de hechos del módulo presupuestos.

Cuadro 37
Estructura dimensión tiempo

TABLA	DIM_TIEMPO								
TIPO DE TABLA	DIMENSIONAL								
DESCRIPCIÓN	TABLA CON REGISTROS DE TIEMPO								
ESQUEMA	DWESPE								
	DESTINO					ORIGEN			
Columna	Descripción	Tipo de dato	Clave	Fk para	Null	Esquema	Tabla	Campo	Tipo de dato
SK_FECHA	Clave subrogada de la dimensión tiempo	Number(38)	pk	FAC_ACTIVOS	N				
FECHA	Fecha (DD/MM/YY)	Date			N				
AÑO	Año de la fecha	Number(5)			N				
NSEMESTRE	Número romano y año del semestre	Varchar2(50)			N				
TRIMESTRE	Número del trimestre	Number(5)			N				
MES	Número del mes	Number(5)			N				
SEMANA	Número de la semana	Number(5)			N				
DÍA	Número del día	Number(5)			N				
DIASEMANA	Número del día de la semana	Number(5)			N				
NTRIMESTRE	Nombre del trimestre y el año	Varchar2(7)			N				
NMES	Nombre del mes	Varchar2(15)			N				
NMES3L	Tres iniciales del mes	Varchar2(3)			N				
NSEMANA	Abreviatura de semana (Sem), el número de la semana y el año	Varchar2(10)			N				
NDIA	Número del día y el nombre del mes con tres iniciales	Varchar2(6)			N				
NDIASEMANA	Nombre del día	Varchar2(10)			N				

Cuadro 38
Estructura dimensión proyecto

TABLA		DIM_PROYECTO							
TIPO DE TABLA		DIMENSIONAL							
DESCRIPCIÓN		TABLA CON REGISTROS DE LOS PROYECTOS DE PRESUPUESTO.							
ESQUEMA		DWESPE							
		DESTINO				ORIGEN			
Columna	Descripción	Tipo de dato	Clave	Fk para	Null	Esquema	Tabla	Campo	Tipo de dato
sk_proyecto	Clave subrogada de la dimensión proyecto	Number(38)	pk	FAC_PRESUPUESTO	N	Sistema	Sistema	Sistema	Number(38)
dpro_nombre proyecto	Nombre del proyecto de presupuesto	Varchar2(150)			Y	STGESPE	szptejepre	szptejepre_nom_proyecto	Varchar(150)
dpro_proyecto	Código del proyecto de presupuesto	Number(38)			Y	STGESPE	szptejepre	szptejepre_proyecto	Number(38)
nombre_etl	Nombre de la transformación	Varchar2(2000)			Y	Sistema	Sistema	Sistema	Varchar2(2000)
fecha_etl	Fecha en la que se realiza la transformación	Date			Y	Sistema	Sistema	Sistema	Date
ip_equipo	Dirección ip del equipo	Varchar2(2000)			Y	Sistema	Sistema	Sistema	Varchar2(2000)

Cuadro 39
Estructura dimensión programa

TABLA		DIM_PROGRAMA							
TIPO DE TABLA	DIMENSIONAL								
DESCRIPCIÓN	TABLA CON REGISTROS DE LOS PROGRAMAS DE PRESUPUESTO.								
ESQUEMA	DWESPE								
Columna	Descripción	DESTINO	Clave	Fk para	Null	Esquema	Tabla	ORIGEN	Tipo de dato
		Tipo de dato						Campo	
sk_programa	Clave subrogada de la dimensión programa	Number(38)	pk	FAC_PRESU PUESTO	N	Sistema	Sistema	Sistema	Number(38)
dprog_programa	Código del programa de presupuesto	Number(38)			Y	STGESPE	szptejepre	szptejepre_programa	Number(38)
dprog_nombreprograma	Nombre del programa de presupuesto	Varchar2(150)			Y	STGESPE	szptejepre	szptejepre_nombreprograma	Varchar2(150)
nombre_etl	Nombre de la transformación	Varchar2(2000)			Y	Sistema	Sistema	Sistema	Varchar2(2000)
fecha_etl	Fecha en la que se realiza la transformación	Date			Y	Sistema	Sistema	Sistema	Date
ip_equipo	Dirección ip del equipo	Varchar2(2000)			Y	Sistema	Sistema	Sistema	Varchar2(2000)

Cuadro 40
Estructura dimensión actividad

TABLA	DIM_ACTIVIDAD								
TIPO DE TABLA	DIMENSIONAL								
DESCRIPCIÓN	TABLA CON REGISTROS DE LAS ACTIVIDADES DE PRESUPUESTO.								
ESQUEMA	DWESPE								
Columna	Descripción	DESTINO				ORIGEN			
		Tipo de dato	Clave	Fk para	Null	Esquema	Tabla	Campo	Tipo de dato
sk_actividad	Clave subrogada de la dimensión actividad	Number(38)	pk	FAC_PRESUPUESTO	N	Sistema	Sistema	Sistema	Number(38)
dact_nombreactividad	Nombre de la actividad de presupuesto	Varchar2(150)			Y	STGESPE	szptejepre	szptejepre_nombreactividad	Varchar2(150)
dact_actividad	Código de la actividad de presupuesto	Number(38)			Y	STGESPE	szptejepre	szptejepre_actividad	Number(38)
nombre_etl	Nombre de la transformación	Varchar2(2000)			Y	Sistema	Sistema	Sistema	Varchar2(2000)
fecha_etl	Fecha en la que se realiza la transformación	Date			Y	Sistema	Sistema	Sistema	Date
ip_equipo	Dirección ip del equipo	Varchar2(2000)			Y	Sistema	Sistema	Sistema	Varchar2(2000)

Cuadro 41
Estructura dimensión subprograma

TABLA		DIM_SUBPROGRAMA							
TIPO DE TABLA	DIMENSIONAL								
DESCRIPCIÓN	TABLA CON REGISTROS DE LOS SUBPROGRAMAS DE PRESUPUESTO.								
ESQUEMA	DWESPE								
Columna	Descripción	DESTINO				ORIGEN			
		Tipo de dato	Clave	Fk para	Null	Esquema	Tabla	Campo	Tipo de dato
sk_subprograma	Clave subrogada de la dimensión subprograma	Number(38)	pk	FAC_PRESUPUESTO	N	Sistema	Sistema	Sistema	Number(38)
dsubp_subprograma	Código del subprograma de presupuesto	Number(38)			Y	STGESPE	szptejepre	szptejepre_subprograma	Number(38)
dsubp_nombresubprograma	Nombre del subprograma de presupuesto	Varchar2(150)			Y	STGESPE	szptejepre	szptejepre_nom_subprograma	Varchar2(150)
nombre_etl	Nombre de la transformación	Varchar2(2000)			Y	Sistema	Sistema	Sistema	Varchar2(2000)
fecha_etl	Fecha en la que se realiza la transformación	Date			Y	Sistema	Sistema	Sistema	Date
ip_equipo	Dirección ip del equipo	Varchar2(2000)			Y	Sistema	Sistema	Sistema	Varchar2(2000)

Cuadro 42
Estructura dimensión ítem

TABLA		DIM_ITEM							
TIPO DE TABLA	DIMENSIONAL								
DESCRIPCIÓN	TABLA CON REGISTROS DE LOS ITEMS DE PRESUPUESTO.								
ESQUEMA	DWESPE								
		DESTINO				ORIGEN			
Columna	Descripción	Tipo de dato	Clave	Fk para	Null	Esquema	Tabla	Campo	Tipo de dato
sk_item	Clave subrogada de la dimensión ítem	Number(38)	pk	FAC_PRES UPUESTO	N	Sistema	Sistema	Sistema	Number(38)
dite_item	Código del ítem de presupuesto	Number(38)			Y	STGESPE	szptejepre	szptejepre_item	Number(38)
dite_nombreitem	Nombre del ítem de presupuesto	Varchar2(150)			Y	STGESPE	szptejepre	szptejepre_nombre_item	Varchar2(150)
nombre_etl	Nombre de la transformación	Varchar2(2000)			Y	Sistema	Sistema	Sistema	Varchar2(2000)
fecha_etl	Fecha en la que se realiza la transformación	Date			Y	Sistema	Sistema	Sistema	Date
ip_equipo	Dirección ip del equipo	Varchar2(2000)			Y	Sistema	Sistema	Sistema	Varchar2(2000)

Cuadro 43
Estructura dimensión geográfica

TABLA	DIM_GEOGRAFICO								
TIPO DE TABLA	DIMENSIONAL								
DESCRIPCIÓN	TABLA CON REGISTROS DE LA UBICACIÓN GEOGRÁFICA DE PRESUPUESTO.								
ESQUEMA	DWESPE								
		DESTINO					ORIGEN		
Columna	Descripción	Tipo de dato	Clave	Fk para	Null	Esquema	Tabla	Campo	Tipo de dato
sk_geografico	Clave subrogada de la dimensión geográfico	Number(38)	pk	FAC_PRES UPUESTO	N	Sistema	Sistema	Sistema	Number(38)
dgeo_geografico	Código de la ubicación geográfica de presupuesto	Number(38)			Y	STGESPE	szptejepre	szptejepre_geografico	Number(38)
nombre_etl	Nombre de la transformación	Varchar2(2000)			Y	Sistema	Sistema	Sistema	Varchar2(2000)
fecha_etl	Fecha en la que se realiza la transformación	Date			Y	Sistema	Sistema	Sistema	Date
ip_equipo	Dirección ip del equipo	Varchar2(2000)			Y	Sistema	Sistema	Sistema	Varchar2(2000)

Cuadro 44
Estructura dimensión fuente

TABLA		DIM FUENTE							
TIPO DE TABLA	DIMENSIONAL								
DESCRIPCIÓN	TABLA CON REGISTROS DE LAS FUENTES DE PRESUPUESTO.								
ESQUEMA	DWESPE								
Columna	Descripción	DESTINO				ORIGEN			
		Tipo de dato	Clave	Fk para	Null	Esquema	Tabla	Campo	Tipo de dato
sk_fuente	Clave subrogada de la dimensión proyecto	Number(38)	pk	FAC_PRES UPUESTO	N	Sistema	Sistema	Sistema	Number(38)
dfuen_fuente	Código de la fuente de presupuesto	Number(38)			Y	STGESPE	szptejepre	szptejepre_fuente	Number(38)
nombre_etl	Nombre de la transformación	Varchar2(2000)			Y	Sistema	Sistema	Sistema	Varchar2(2000)
fecha_etl	Fecha en la que se realiza la transformación	Date			Y	Sistema	Sistema	Sistema	Date
ip_equipo	Dirección ip del equipo	Varchar2(2000)			Y	Sistema	Sistema	Sistema	Varchar2(2000)

4.6.2.5 Mapeo de datos para tablas de hechos de Presupuestos

Cuadro 45
Estructura tabla de hechos presupuesto

TABLA		FAC PRESUPUESTO								
TIPO DE TABLA	HECHOS									
DESCRIPCIÓN	TABLA DE HECHOS PARA EL DATAMART DE PRESUPUESTO									
ESQUEMA	DWESPE								ORIGEN	
Columna	Descripción	Tipo de dato	Clave	Fk para	Null	Esquema	Tabla	Campo	Tipo de dato	
sk_presupuesto	Clave subrogada de la tabla de hechos de presupuesto	Number(38)	pk		N	sistema	sistema	sistema	Number(38)	
sk_programa	Clave subrogada de la dimensión programa	Number(38)	fk		N	DWESPE	dim_programa	sk_programa	Number(38)	
sk_subprograma	Clave subrogada de la dimensión subprograma	Number(38)	fk		N	DWESPE	dim_subprograma	sk_subprograma	Number(38)	
sk_proyecto	Clave subrogada de la dimensión proyecto	Number(38)	fk		N	DWESPE	dim_proyecto	sk_proyecto	Number(38)	
sk_actividad	Clave subrogada de la dimensión actividad	Number(38)	fk		N	DWESPE	dim_actividad	sk_actividad	Number(38)	
sk_item	Clave subrogada de la dimensión ítem	Number(38)	fk		N	DWESPE	dim_item	sk_item	Number(38)	
sk_geografico	Clave subrogada de la dimensión geográfico	Number(38)	fk		N	DWESPE	dim_geografico	sk_geografico	Number(38)	
sk_fuente	Clave subrogada de la dimensión fuente	Number(38)	fk		N	DWESPE	dim_fuente	sk_fuente	Number(38)	
sk_fecha	Clave subrogada de la dimensión tiempo	Number(38)	fk		N	DWESPE	dim_tiempo	sk_fecha	Number(38)	
fpre_asignado	Medida con el valor asignado de presupuesto	Number(38)			Y	STGESPE	szptpresup	szptpresup_asignado	Number(38)	
fpre_modificado	Medida con el valor modificado de presupuesto	Number(38)			Y	STGESPE	szptpresup	szptpresup_modificado	Number(38)	
fpre_codificado	Medida con el valor codificado de presupuesto	Number(38)			Y	STGESPE	szptpresup	szptpresup_codificado	Number(38)	

fpre_montocertificado	Medida con el valor montocertificado de presupuesto	Number(38)	Y	STGESPE	szptpresup	szptpresup_monto_certificado	Number(38)
fpre_comprometido	Medida con el valor comprometido de presupuesto	Number(38)	Y	STGESPE	szptpresup	szptpresup_comprometido	Number(38)
fpre_devengado	Medida con el valor devengado de presupuesto	Number(38)	Y	STGESPE	szptpresup	szptpresup_devengado	Number(38)
fpre_pagado	Medida con el valor pagado de presupuesto	Number(38)	Y	STGESPE	szptpresup	szptpresup_pagado	Number(38)
fpre_saldoporcomprometer	Medida con el valor saldoporcomprometer de presupuesto	Number(38)	Y	STGESPE	szptpresup	szptpresup_saldo_por_comprometer	Number(38)
fpre_saldopordevengar	Medida con el valor saldopordevengar de presupuesto	Number(38)	Y	STGESPE	szptpresup	szptpresup_saldo_por_devengar	Number(38)
fpre_saldoporpagar	Medida con el valor saldoporpagar de presupuesto	Number(38)	Y	STGESPE	szptpresup	szptpresup_saldo_por_pagar	Number(38)
fpre_porcentajeejecucion	Medida con el valor porcentajeejecucion de presupuesto	Number(38)	Y	STGESPE	szptpresup	szptpresup_porcent_ejecucion	Number(38)
nombre_etl	Nombre de la transformación	Varchar2(2000)	Y	Sistema	Sistema	Sistema	Varchar2(2000)
fecha_etl	Fecha en la que se realiza la transformación	Date	Y	Sistema	Sistema	Sistema	Date
ip_equipo	Dirección ip del equipo	Varchar2(2000)	Y	Sistema	Sistema	Sistema	Varchar2(2000)

4.6.2.6 Diagrama general para el módulo de activos fijos

El diagrama muestra las dimensiones y la relación que tienen con las tablas de la base de datos de origen.

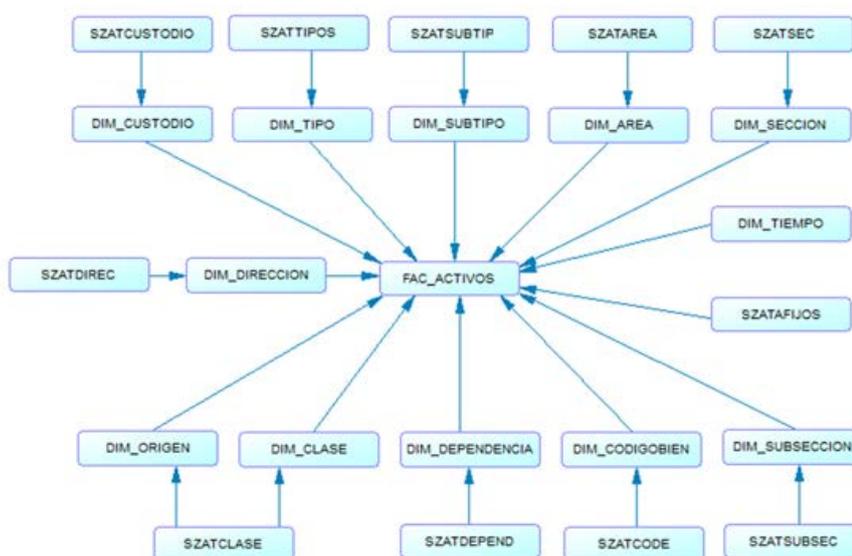


Figura 69: Diagrama general activos fijos

4.6.2.7 Mapeo de datos activos fijos

En la siguiente tabla se describe el origen de datos desde el servidor de base de datos o archivos planos hacia las dimensiones y tablas de hechos del datamart.

Cuadro 46
Mapeo de datos activos fijos

FUENTE DE DATOS ORIGINAL	DESCRIPCIÓN	STAGE	DIMENSIÓN (DIM)	HECHOS (FAC)
af_tipos	Tabla que contiene los códigos y las descripciones de los tipos.	szattipos	DIM_TIPO	
af_subtipos	Tabla que contiene los códigos y las descripciones de los subtipos.	szatsubtip	DIM_SUBTIPO	
af_subseccion	Tabla que contiene los códigos y las	szatsubsec	DIM_SUBSECCION	

	descripciones de las subsecciones.			
af_seccion	Tabla que contiene los códigos y las descripciones de las secciones.	szatsec	DIM_SECCION	
af_direccion	Tabla que contiene los códigos y las descripciones de las direcciones.	szatdirec	DIM_DIRECCION	
af_tipos	Tabla que contiene los códigos del origen.	szattipos	DIM_ORIGEN	
af_area	Tabla que contiene los códigos y las descripciones de las áreas.	szatarea	DIM_AREA	
af_clase	Tabla que contiene los códigos y las descripciones de las clases.	szatclase	DIM_CLASE	
af_activofijo	Tabla que contiene los códigos para el código del bien.	szatcode	DIM_CODIGOBIEN	
dependencia	Tabla que contiene los códigos de la empresa, el nombre de la institución y la localidad.	szatdepend	DIM_DEPENDENCIA	
af_custodios	Tabla que contiene los bienes y custodios.	szatcustodio	DIM_CUSTODIO	
af_activofijo	Tabla que contiene las medidas para los activos.	szatafijos		FAC_ACTIVOS

4.6.2.8 Modelo Dimensional Activos Fijos

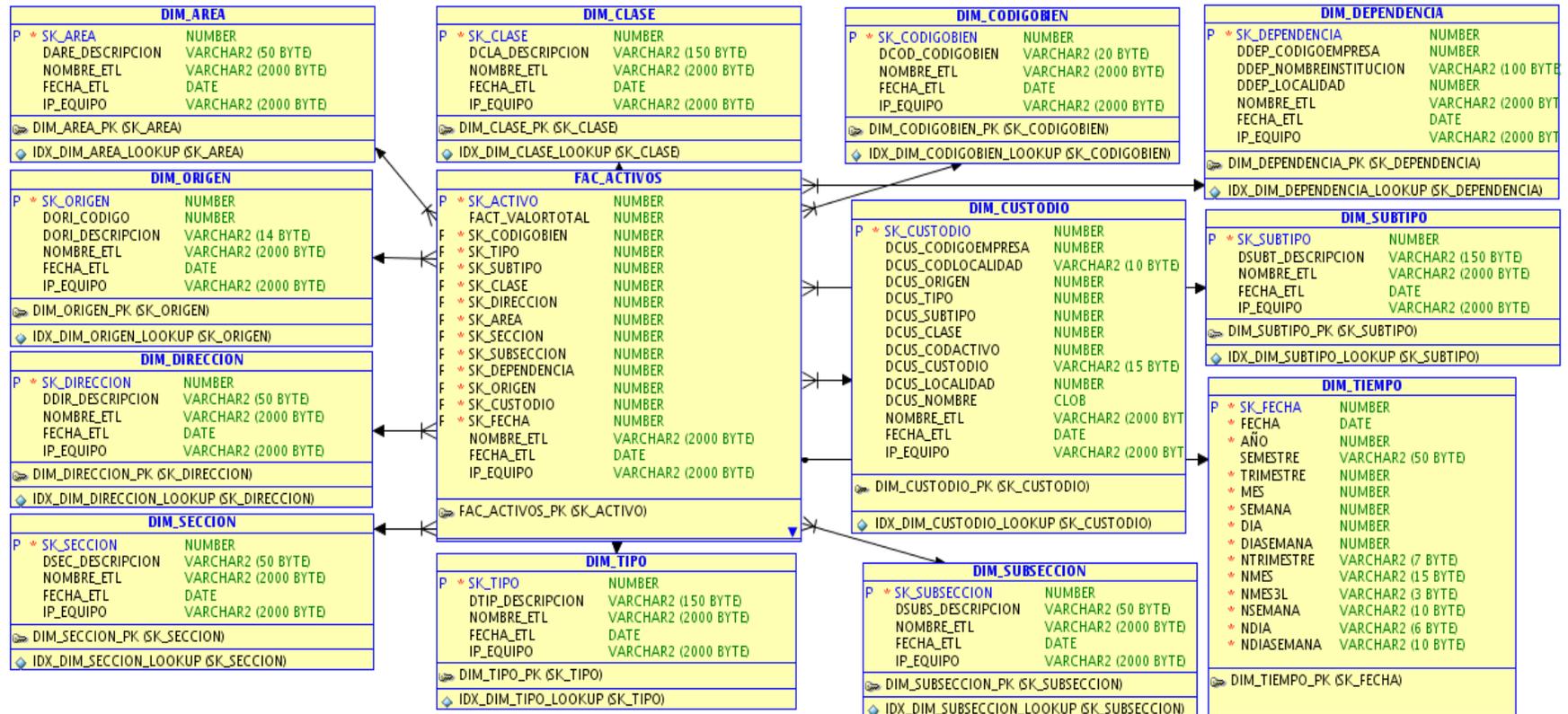


Figura 70: Modelo dimensional activos fijos

4.6.2.9 Mapeo de datos para tablas de dimensiones de Activos Fijos.

Las siguientes tablas muestran el destino y origen de los datos para dimensiones del módulo activos fijos, se detalla nombres de los campos de destino y origen, los tipos de datos y sus esquemas respectivos.

Cuadro 47
Estructura dimensión tipo

TABLA	DIM_TIPO								
TIPO DE TABLA	DIMENSIONAL								
DESCRIPCIÓN	TABLA CON REGISTROS DE LOS TIPOS DE ACTIVOS								
ESQUEMA	DWESPE								
		DESTINO				ORIGEN			
Columna	Descripción	Tipo de dato	Clave	Fk para	Null	Esquema	Tabla	Campo	Tipo de dato
sk_tipo	Clave subrogada de la dimensión tipos	Number(38)	pk	FAC_ACTIVOS	N	Sistema	Sistema	Sistema	Number(38)
dtip_descripcion	Nombre del tipo de activo	Varchar2(150)			Y	STGESPE	szattipos	szattipos_descripcion	Varchar2(150)
nombre_etl	Nombre de la transformación	Varchar2(2000)			Y	Sistema	Sistema	Sistema	Varchar2(2000)
fecha_etl	Fecha en la que se realiza la transformación	Date			Y	Sistema	Sistema	Sistema	Date
ip_equipo	Dirección ip del equipo	Varchar2(2000)			Y	Sistema	Sistema	Sistema	Varchar2(2000)

Cuadro 48
Estructura dimensión subtipo

TABLA		DIM_SUBTIPO							
TIPO DE TABLA		DIMENSIONAL							
DESCRIPCIÓN		TABLA CON REGISTROS DE LOS SUBTIPOS DE ACTIVOS							
ESQUEMA		DWESPE							
		DESTINO				ORIGEN			
Columna	Descripción	Tipo de dato	Clave	Fk para	Null	Esquema	Tabla	Campo	Tipo de dato
sk_subtipo	Clave subrogada de la dimensión subtipo	Number(38)	pk	FAC_ACTIVOS	N	Sistema	Sistema	Sistema	Number(38)
dsubt_descripcion	Nombre del subtipo de activo	Varchar2(150)			Y	STGESPE	szatsubtip	szatsubtip_descripcion	Varchar2(150)
nombre_etl	Nombre de la transformación	Varchar2(2000)			Y	Sistema	Sistema	Sistema	Varchar2(2000)
fecha_etl	Fecha en la que se realiza la transformación	Date			Y	Sistema	Sistema	Sistema	Date
ip_equipo	Dirección ip del equipo	Varchar2(2000)			Y	Sistema	Sistema	Sistema	Varchar2(2000)

Cuadro 49
Estructura dimensión subsección

TABLA		DIM_SUBSECCION							
TIPO DE TABLA		DIMENSIONAL							
DESCRIPCIÓN		TABLA CON REGISTROS DE LAS SUBSECCIONES DE ACTIVOS							
ESQUEMA		DWESPE							
		DESTINO				ORIGEN			
Columna	Descripción	Tipo de dato	Clave	Fk para	Null	Esquema	Tabla	Campo	Tipo de dato
sk_subseccion	Clave subrogada de la dimensión subsección	Number(38)	pk	FAC_ACTIVOS	N	Sistema	Sistema	Sistema	Number(38)
dsubs_descripcion	Nombre de la subsección del activo	Varchar2(150)			Y	STGESPE	szatsubsec	szatsubsec_desc_subseccion	Varchar2(50)
nombre_etl	Nombre de la transformación	Varchar2(2000)			Y	Sistema	Sistema	Sistema	Varchar2(2000)
fecha_etl	Fecha en la que se realiza la transformación	Date			Y	Sistema	Sistema	Sistema	Date
ip_equipo	Dirección ip del equipo	Varchar2(2000)			Y	Sistema	Sistema	Sistema	Varchar2(2000)

Cuadro 50
Estructura dimensión sección

TABLA		DIM_SECCION							
TIPO DE TABLA		DIMENSIONAL							
DESCRIPCIÓN		TABLA CON REGISTROS DE LAS SECCIONES DE ACTIVOS							
ESQUEMA		DWESPE							
		DESTINO				ORIGEN			
Columna	Descripción	Tipo de dato	Clave	Fk para	Null	Esquema	Tabla	Campo	Tipo de dato
sk_sección	Clave subrogada de la dimensión sección	Number(38)	pk	FAC_ACTIVOS	N	Sistema	Sistema	Sistema	Number(38)
dsec_descripcion	Nombre de la sección de activo	Varchar2(150)			Y	STGESPE	szatsec	szatsec_desc_seccion	Varchar2(150)
nombre_etl	Nombre de la transformación	Varchar2(2000)			Y	Sistema	Sistema	Sistema	Varchar2(2000)
fecha_etl	Fecha en la que se realiza la transformación	Date			Y	Sistema	Sistema	Sistema	Date
ip_equipo	Dirección ip del equipo	Varchar2(2000)			Y	Sistema	Sistema	Sistema	Varchar2(2000)

Cuadro 51
Estructura dimensión dirección

TABLA	DIM_DIRECCION									
TIPO DE TABLA	DIMENSIONAL									
DESCRIPCIÓN	TABLA CON REGISTROS DE LAS DIRECCIONES DE ACTIVOS									
ESQUEMA	DWESPE									
		DESTINO						ORIGEN		
Columna	Descripción	Tipo de dato	Clave	Fk para	Null	Esquema	Tabla	Campo	Tipo de dato	
sk_direccion	Clave subrogada de la dimensión dirección	Number(38)	pk	FAC_ACTIVOS	N	Sistema	Sistema	Sistema	Number(38)	
ddir_descripcion	Nombre de la dirección de activo	Varchar2(150)			Y	STGESPE	szatdirec	szatdirec_desc_direccion	Varchar2(150)	
nombre_etl	Nombre de la transformación	Varchar2(2000)			Y	Sistema	Sistema	Sistema	Varchar2(2000)	
fecha_etl	Fecha en la que se realiza la transformación	Date			Y	Sistema	Sistema	Sistema	Date	
ip_equipo	Dirección ip del equipo	Varchar2(2000)			Y	Sistema	Sistema	Sistema	Varchar2(2000)	

Cuadro 52
Estructura dimensión origen

TABLA		DIM_ORIGEN							
TIPO DE TABLA		DIMENSIONAL							
DESCRIPCIÓN		TABLA CON REGISTROS DEL ORIGEN DE ACTIVOS							
ESQUEMA		DWESPE							
		DESTINO				ORIGEN			
Columna	Descripción	Tipo de dato	Clave	Fk para	Null	Esquema	Tabla	Campo	Tipo de dato
sk_origen	Clave subrogada de la dimensión origen	Number(38)	pk	FAC_ACTIVOS	N	Sistema	Sistema	Sistema	Number(38)
dori_codigo	Código del origen de activo	Number(10)			Y	STGESPE	szattipos	szattipos_origen	Number(10)
dori_descripcion	Nombre del origen de activo	Varchar2(150)			Y	sistema	sistema	sistema	Varchar2(150)
nombre_etl	Nombre de la transformación	Varchar2(2000)			Y	Sistema	Sistema	Sistema	Varchar2(2000)
fecha_etl	Fecha en la que se realiza la transformación	Date			Y	Sistema	Sistema	Sistema	Date
ip_equipo	Dirección ip del equipo	Varchar2(2000)			Y	Sistema	Sistema	Sistema	Varchar2(2000)

Cuadro 53
Estructura dimensión área

TABLA		DIM_AREA							
TIPO DE TABLA		DIMENSIONAL							
DESCRIPCIÓN		TABLA CON REGISTROS DE LAS ÁREAS DE ACTIVOS							
ESQUEMA		DWESPE							
		DESTINO				ORIGEN			
Columna	Descripción	Tipo de dato	Clave	Fk para	Null	Esquema	Tabla	Campo	Tipo de dato
sk_area	Clave subrogada de la dimensión área	Number(38)	pk	FAC_ACTIVOS	N	Sistema	Sistema	Sistema	Number(38)
dare_descripcion	Nombre del área del activo	Varchar2(150)			Y	STGESPE	szatarea	szatarea_desc_area	Varchar2(150)
nombre_etl	Nombre de la transformación	Varchar2(2000)			Y	Sistema	Sistema	Sistema	Varchar2(2000)
fecha_etl	Fecha en la que se realiza la transformación	Date			Y	Sistema	Sistema	Sistema	Date
ip_equipo	Dirección ip del equipo	Varchar2(2000)			Y	Sistema	Sistema	Sistema	Varchar2(2000)

Cuadro 54 Estructura dimensión custodio

TABLA		DIM_CUSTODIO							
TIPO DE TABLA		DIMENSIONAL							
DESCRIPCIÓN		TABLA CON REGISTROS DE LOS CUSTODIOS DE ACTIVOS							
ESQUEMA		DWESPE							
		DESTINO				ORIGEN			
Columna	Descripción	Tipo de dato	Clave	Fk para	Null	Esquema	Tabla	Campo	Tipo de dato
sk_custodio	Clave subrogada de la dimensión custodio	Number(38)	pk	FAC_ACTIVOS	N	Sistema	Sistema	Sistema	Number(38)
dcus_codigoempresa	Código de empresa del activo	Number(38)			Y	STGESPE	szatcustodio	szatcustodio_codigo_empresa	Number(38)
dcus_codlocalidad	Código de la localidad del activo	Varchar2(10)			Y	STGESPE	szatcustodio	szatcustodio_cod_localidad	Varchar2(10)
dcus_origen	Código del origen del activo	Number(38)			Y	STGESPE	szatcustodio	szatcustodio_origen	Number(38)
dcus_tipo	Código del tipo del activo	Number(38)			Y	STGESPE	szatcustodio	szatcustodio_tipo	Number(38)
dcus_subtipo	Código del subtipo del activo	Number(38)			Y	STGESPE	szatcustodio	szatcustodio_subtipo	Number(38)
dcus_clase	Código de la clase del activo	Number(38)			Y	STGESPE	szatcustodio	szatcustodio_clase	Number(38)
dcus_codactivo	Código del activo	Number(38)			Y	STGESPE	szatcustodio	szatcustodio_codigoactivo	Number(38)
dcus_custodio	Descripción del custodio del activo	Varchar2(15)			Y	STGESPE	szatcustodio	szatcustodio_custodio	Varchar2(15)
dcus_localidad	Localidad del activo	Number(38)			Y	STGESPE	szatcustodio	szatcustodio_localidad	Number(38)
nombre_etl	Nombre de la transformación	Varchar2(2000)			Y	Sistema	Sistema	Sistema	Varchar2(2000)
fecha_etl	Fecha en la que se realiza la transformación	Date			Y	Sistema	Sistema	Sistema	Date
ip_equipo	Dirección ip del equipo	Varchar2(2000)			Y	Sistema	Sistema	Sistema	Varchar2(2000)

Cuadro 55
Estructura dimensión clase

TABLA	DIM_CLASE								
TIPO DE TABLA	DIMENSIONAL								
DESCRIPCIÓN	TABLA CON REGISTROS DE LAS CLASES DE ACTIVOS								
ESQUEMA	DWESPE								
		DESTINO					ORIGEN		
Columna	Descripción	Tipo de dato	Clave	Fk para	Null	Esquema	Tabla	Campo	Tipo de dato
sk_clase	Clave subrogada de la dimensión clase	Number(38)	pk	FAC_ACTIVOS	N	Sistema	Sistema	Sistema	Number(38)
dcla_descripcion	Nombre de la clase del activo	Varchar2(150)			Y	STGESPE	szatclase	szatclase_descripcion	Varchar2(150)
nombre_etl	Nombre de la transformación	Varchar2(2000)			Y	Sistema	Sistema	Sistema	Varchar2(2000)
fecha_etl	Fecha en la que se realiza la transformación	Date			Y	Sistema	Sistema	Sistema	Date
ip_equipo	Dirección ip del equipo	Varchar2(2000)			Y	Sistema	Sistema	Sistema	Varchar2(2000)

Cuadro 56
Estructura dimensión código del bien

TABLA	DIM_CODIGOBIEN									
TIPO DE TABLA	DIMENSIONAL									
DESCRIPCIÓN	TABLA CON REGISTROS DE LOS CÓDIGOS DEL BIEN DE CADA ACTIVO									
ESQUEMA	DWESPE									
		DESTINO						ORIGEN		
Columna	Descripción	Tipo de dato	Clave	Fk para	Null	Esquema	Tabla	Campo	Tipo de dato	
sk_codigobien	Clave subrogada de la dimensión codigobien	Number(38)	pk	FAC_ACTI VOS	N	Sistema	Sistema	Sistema	Number(38)	
dcod_codigobien	Código del bien del activo	Varchar2(20)			Y	STGESPE	szatcode	szatcode_codigo_bien	Varchar2(20)	
nombre_etl	Nombre de la transformación	Varchar2(2000)			Y	Sistema	Sistema	Sistema	Varchar2(2000)	
fecha_etl	Fecha en la que se realiza la transformación	Date			Y	Sistema	Sistema	Sistema	Date	
ip_equipo	Dirección ip del equipo	Varchar2(2000)			Y	Sistema	Sistema	Sistema	Varchar2(2000)	

Cuadro 57
Estructura dimensión dependencia

TABLA	DIM_DEPENDENCIA								
TIPO DE TABLA	DIMENSIONAL								
DESCRIPCIÓN	TABLA CON REGISTROS DE DEPENDENCIA DE ACTIVOS								
ESQUEMA	DWESPE								
	DESTINO					ORIGEN			
Columna	Descripción	Tipo de dato	Clave	Fk para	Null	Esquema	Tabla	Campo	Tipo de dato
sk_dependencia	Clave subrogada de la dimensión dependencia	Number(38)	pk	FAC_ACTIVOS	N	Sistema	Sistema	Sistema	Number(38)
ddep_codigoempresa	Código de la empresa del activo	Number(38)			Y	STGESPE	szatdepend	szatdepend_codigo_empresa	Number(38)
ddep_nombreinstitucion	Nombre de la institución del activo	Varchar2(150)			Y	STGESPE	szatdepend	szatdepend_n_institucion	Varchar2(150)
ddep_localidad	Nombre de la localidad del activo	Number(38)			Y	STGESPE	szatdepend	szatdepend_localidad	Number(38)
nombre_etl	Nombre de la transformación	Varchar2(2000)			Y	Sistema	Sistema	Sistema	Varchar2(2000)
fecha_etl	Fecha en la que se realiza la transformación	Date			Y	Sistema	Sistema	Sistema	Date
ip_equipo	Dirección ip del equipo	Varchar2(20)			Y	Sistema	Sistema	Sistema	Varchar2(20)

4.6.2.10 Mapeo de datos para tablas de hechos de Activos Fijos.

Cuadro 58
Estructura tabla de hechos activos fijos

TABLA		FAC_ACTIVOS							
TIPO DE TABLA		HECHOS							
DESCRIPCIÓN		TABLA DE HECHOS PARA EL DATAMART DE ACTIVOS FIJOS							
ESQUEMA		DWESPE							
		DESTINO				ORIGEN			
Columna	Descripción	Tipo de dato	Clave	Fk para	Null	Esquema	Tabla	Campo	Tipo de dato
sk_activos	Clave subrogada de la tabla de hechos de activos	Number(38)	pk		N	sistema	sistema	sistema	Number(38)
sk_codigobien	Clave subrogada de la dimensión codigobien	Number(38)	fk		N	DWESPE	dim_codigobien	sk_codigobien	Number(38)
sk_tipo	Clave subrogada de la dimensión tipo	Number(38)	fk		N	DWESPE	dim_tipo	sk_tipo	Number(38)
sk_subtipo	Clave subrogada de la dimensión subtipo	Number(38)	fk		N	DWESPE	dim_subtipo	sk_subtipo	Number(38)
sk_clase	Clave subrogada de la dimensión clase	Number(38)	fk		N	DWESPE	dim_clase	sk_clase	Number(38)
sk_direccion	Clave subrogada de la dimensión dirección	Number(38)	fk		N	DWESPE	dim_direccion	sk_direccion	Number(38)
sk_area	Clave subrogada de la dimensión	Number(38)	fk		N	DWESPE	dim_area	sk_area	Number(38)

área								
sk_seccion	Clave subrogada de la dimensión sección	Number(38)	fk	N	DWESPE	dim_seccion	sk_seccion	Number(38)
sk_subseccion	Clave subrogada de la dimensión subsección	Number(38)	fk	N	DWESPE	dim_subseccion	sk_subseccion	Number(38)
sk_dependencia	Clave subrogada de la dimensión dependencia	Number(38)	fk	N	DWESPE	dim_dependencia	sk_dependencia	Number(38)
sk_origen	Clave subrogada de la dimensión origen	Number(38)	fk	N	DWESPE	dim_origen	sk_origen	Number(38)
sk_fecha	Clave subrogada de la dimensión tiempo	Number(38)	fk	N	DWESPE	dim_tiempo	sk_fecha	Number(38)
fact_valortotal	Medida de la suma del valor avaluó + valor de activos	Number(38)		Y	STGESPE	szatafijos	szatafijos_valoravaluo , szatafijos_valor	Number(38)
nombre_etl	Nombre de la transformación	Varchar2(2000)		Y	Sistema	Sistema	Sistema	Varchar2(2000)
fecha_etl	Fecha en la que se realiza la transformación	Date		Y	Sistema	Sistema	Sistema	Date
ip_equipo	Dirección ip del equipo	Varchar2(2000)		Y	Sistema	Sistema	Sistema	Varchar2(2000)

4.6.2.11 Diagrama general para el módulo de inventarios

El diagrama muestra las dimensiones y la relación que tienen con las tablas de la base de datos de origen.

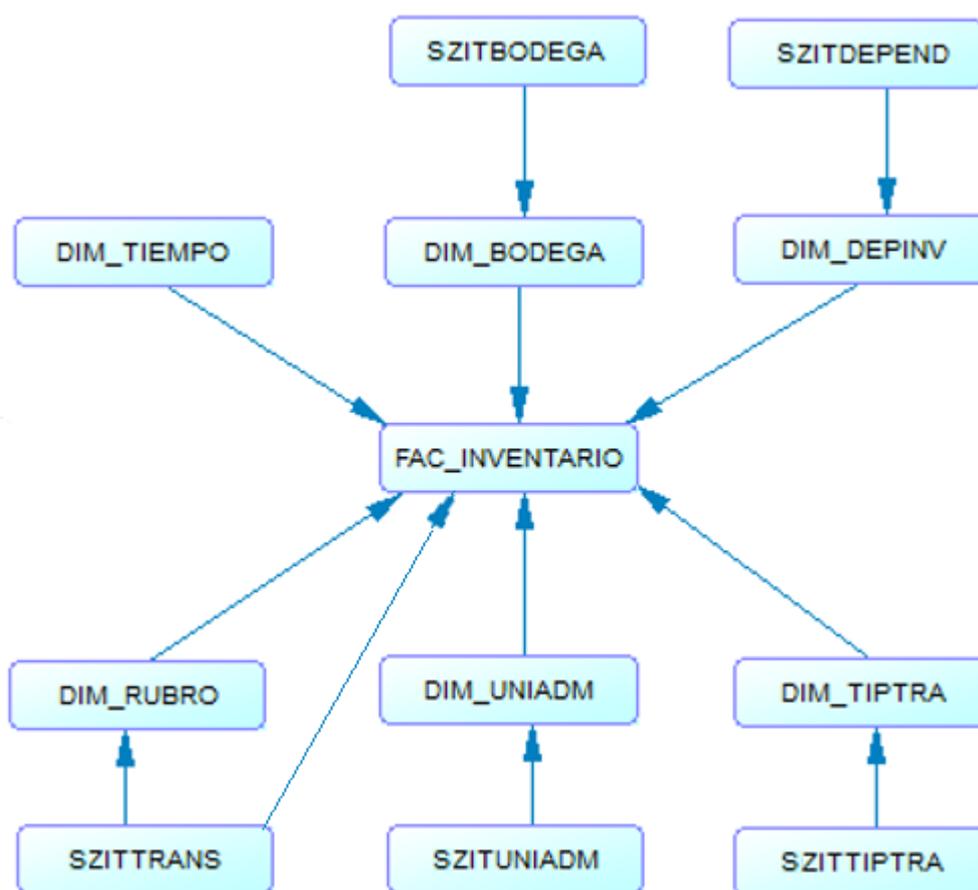


Figura 71: Diagrama general inventarios

4.6.2.12 Mapeo de datos inventarios

En la siguiente tabla se describe el origen de datos desde el servidor de base de datos o archivos planos hacia las dimensiones y tablas de hechos del datamart.

Cuadro 59
Mapeo de datos inventarios

FUENTE DE DATOS ORIGINAL	DESCRIPCIÓN	STAGE	DIMENSIÓN (DIM)	HECHOS (FAC)
in_bodega	Tabla que contiene los códigos y las descripciones de las bodegas.	szitbodega	DIM_BODEGA	
in_tiptra	Tabla que contiene los tipos de transacción del inventario.	szittiptra	DIM_TIPTRA	
in_cab_transaccion	Tabla que contiene las unidades operativas de inventario.	szituniadm	DIM_UNIADM	
af_seccion	Tabla que contiene los rubros (ingresos y egresos) para los inventarios.	szittrans	DIM_RUBRO	
dependencia	Tabla que contiene los códigos de la empresa, el nombre de la institución y la localidad.	szitdepend	DIM_DEPINV	
in_cab_transaccion	Tabla que contiene las medidas para inventario.	szittrans		FAC_INVENTARIO

4.6.2.13 Modelo Dimensional Inventarios

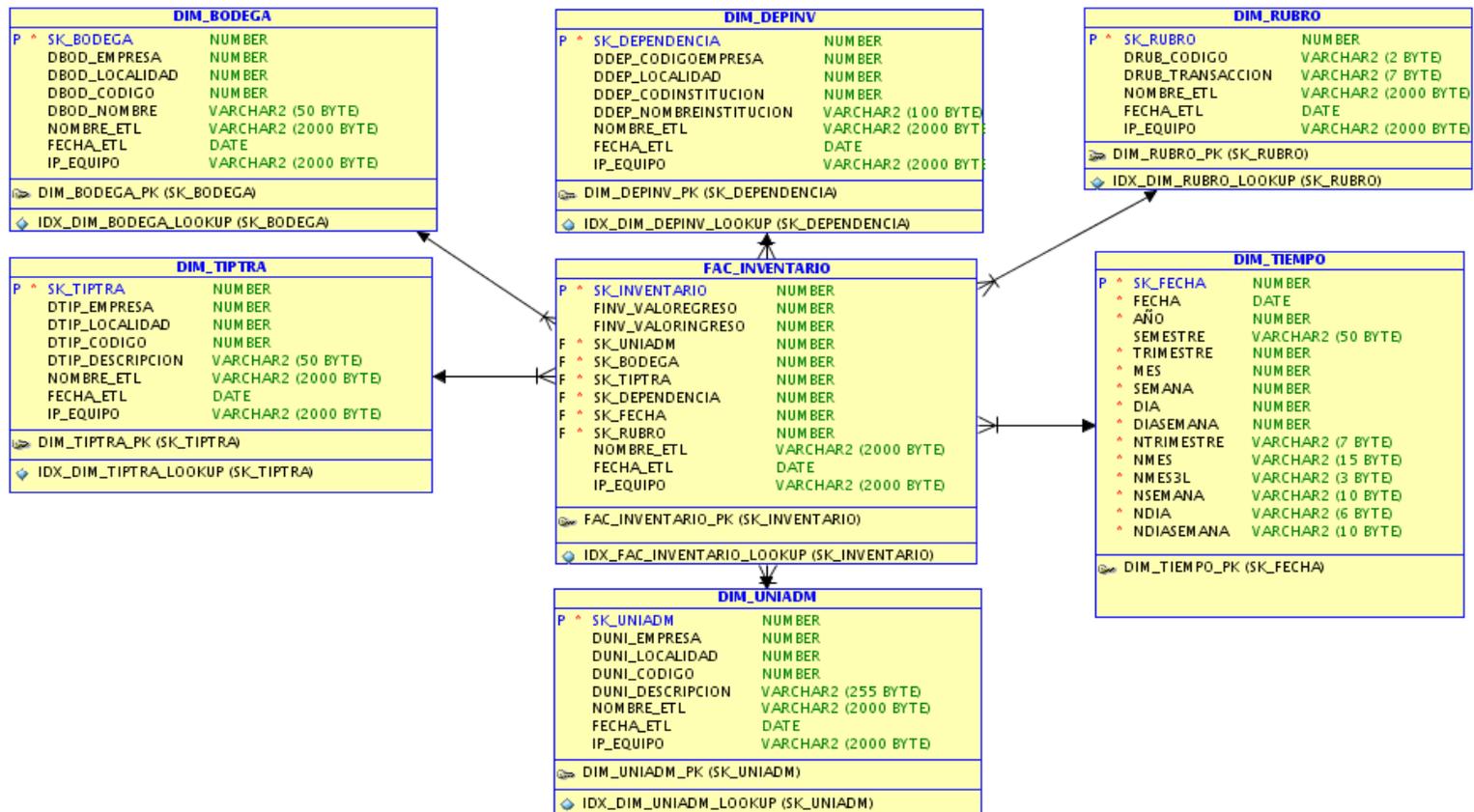


Figura 72: Modelo dimensional inventarios

4.6.2.14 Mapeo de datos para tablas de dimensiones de Inventarios.

Las siguientes tablas muestran el destino y origen de los datos para dimensiones del módulo inventarios, se detalla nombres de los campos de destino y origen, los tipos de datos y sus esquemas respectivos.

Cuadro 60
Estructura dimensión tipo de transacción

TABLA	DIM_TIPTRA								
TIPO DE TABLA	DIMENSIONAL								
DESCRIPCIÓN	TABLA CON REGISTROS DE TIPOS DE TRANSACCIÓN DE LOS INVENTARIOS								
ESQUEMA	DWESPE								
		DESTINO				ORIGEN			
Columna	Descripción	Tipo de dato	Clave	Fk para	Null	Esquema	Tabla	Campo	Tipo de dato
sk_tiptra	Clave subrogada de la dimensión tiptra.	Number(38)	pk	FAC_INVENTARIO	N	Sistema	Sistema	Sistema	Number(38)
dtip_empresa	Código de la empresa del inventario.	Number(38)			Y	STGESPE	szitiptra	szitiptra_empresa	Number(38)
dtip_localidad	Código de la localidad del inventario.	Number(38)			Y	STGESPE	szitiptra	szitiptra_localidad	Number(38)
dtip_codigo	Código del tipo de transacción del inventario.	Number(38)			Y	STGESPE	szitiptra	szitiptra_codigo	Number(38)
dtip_descripcion	Nombre del tipo de transacción del inventario.	Varchar2(50)			Y	STGESPE	szitiptra	szitiptra_descripcion	Varchar2(50)
nombre_etl	Nombre de la transformación	Varchar2(2000)			Y	Sistema	Sistema	Sistema	Varchar2(2000)
fecha_etl	Fecha en la que se realiza la transformación	Date			Y	Sistema	Sistema	Sistema	Date
ip_equipo	Dirección ip del equipo	Varchar2(2000)			Y	Sistema	Sistema	Sistema	Varchar2(2000)

Cuadro 61
Estructura dimensión rubro

TABLA	DIM RUBRO									
TIPO DE TABLA	DIMENSIONAL									
DESCRIPCIÓN	TABLA CON REGISTROS DE LOS INGRESOS O EGRESOS DE INVENTARIO.									
ESQUEMA	DWESPE									
Columna	Descripción	DESTINO				ORIGEN				Tipo de dato
		Tipo de dato	Clave	Fk para	Null	Esquema	Tabla	Campo		
sk_rubro	Clave subrogada de la dimensión rubro.	Number(38)	pk	FAC_INVEN TARIO	N	Sistema	Sistema	Sistema	Number(38)	
drub_codigo	Código de los rubros del inventario.	Varchar2(2)			Y	STGESPE	szittrans	szittrans_transaccion	Varchar2(2)	
drub_transaccion	Nombre del rubro del inventario.	Varchar2(7)			Y	STGESPE	szittrans	szittrans_transaccion	Varchar2(7)	
nombre_etl	Nombre de la transformación	Varchar2(2000)			Y	Sistema	Sistema	Sistema	Varchar2(2000)	
fecha_etl	Fecha en la que se realiza la transformación	Date			Y	Sistema	Sistema	Sistema	Date	
ip_equipo	Dirección ip del equipo	Varchar2(2000)			Y	Sistema	Sistema	Sistema	Varchar2(2000)	

Cuadro 62
Estructura dimensión bodega

TABLA	DIM_BODEGA								
TIPO DE TABLA	DIMENSIONAL								
DESCRIPCIÓN	TABLA CON REGISTROS DE BODEGAS DE INVENTARIO								
ESQUEMA	DWESPE								
	DESTINO					ORIGEN			
Columna	Descripción	Tipo de dato	Clave	Fk para	Null	Esquema	Tabla	Campo	Tipo de dato
sk_bodega	Clave subrogada de la dimensión bodega.	Number(38)	pk	FAC_INVENTARIO	N	Sistema	Sistema	Sistema	Number(38)
dbod_empresa	Código de la empresa.	Number(38)			Y	STGESPE	szitbodega	szitbodega_empresa	Number(38)
dbod_localidad	Código de la localidad.	Number(38)			Y	STGESPE	szitbodega	szitbodega_localidad	Number(38)
dbod_codigo	Código de la bodega.	Number(38)			Y	STGESPE	szitbodega	szitbodega_codbodega	Number(38)
dbod_nombre	Nombre de la bodega.	Varchar2(50)			Y	STGESPE	szitbodega	szitbodega_nombodega	Varchar2(50)
nombre_etl	Nombre de la transformación	Varchar2(2000)			Y	Sistema	Sistema	Sistema	Varchar2(2000)
fecha_etl	Fecha en la que se realiza la transformación	Date			Y	Sistema	Sistema	Sistema	Date
ip_equipo	Dirección ip del equipo	Varchar2(2000)			Y	Sistema	Sistema	Sistema	Varchar2(2000)

4.6.2.15 Mapeo de datos para tablas de hechos de Inventarios.

Cuadro 63
Estructura tabla de hechos inventario

TABLA	FAC_INVENTARIO								
TIPO DE TABLA	HECHOS								
DESCRIPCIÓN	TABLA DE HECHOS PARA EL DATAMART DE INVENTARIO								
ESQUEMA	DWESPE								
	DESTINO					ORIGEN			
Columna	Descripción	Tipo de dato	Clave	Fk para	Null	Esquema	Tabla	Campo	Tipo de dato
sk_inventario	Clave subrogada de la tabla de hechos de inventario.	Number(38)	pk		N	sistema	sistema	sistema	Number(38)
sk_uniadm	Clave subrogada de la dimensión uniadm.	Number(38)	fk		N	DWESPE	dim_uniadm	sk_uniadm	Number(38)
sk_bodega	Clave subrogada de la dimensión bodega.	Number(38)	fk		N	DWESPE	dim_bodega	sk_bodega	Number(38)
sk_tiptra	Clave subrogada de la dimensión tiptra.	Number(38)	fk		N	DWESPE	dim_tiptra	sk_tiptra	Number(38)
sk_dependencia	Clave subrogada de la dimensión dependencia.	Number(38)	fk		N	DWESPE	dim_dependencia	sk_dependencia	Number(38)
sk_fecha	Clave subrogada de la dimensión tiempo	Number(38)	fk		N	DWESPE	dim_tiempo	sk_fecha	Number(38)
sk_rubro	Clave subrogada de la dimensión rubro	Number(38)	fk		N	DWESPE	dim_rubro	sk_rubro	Number(38)
finv_valoregreso	Medida con el valor egreso.	Number(38)			N	STGESPE	szittrans	szittrans_valortotal	Number(38)
finv_valoringreso	Medida con el valor ingreso.	Number(38)			N	STGESPE	szittrans	szittrans_valortotal	Number(38)
nombre_etl	Nombre de la transformación	Varchar2(2000)			Y	Sistema	Sistema	Sistema	Varchar2(2000)
fecha_etl	Fecha en la que se realiza la transformación	Date			Y	Sistema	Sistema	Sistema	Date
ip_equipo	Dirección ip del equipo	Varchar2(2000)			Y	Sistema	Sistema	Sistema	Varchar2(2000)

4.6.2.16 Diagrama general para el módulo de roles de pago (shifre)

El diagrama muestra las dimensiones y la relación que tienen con las tablas de la base de datos de origen.

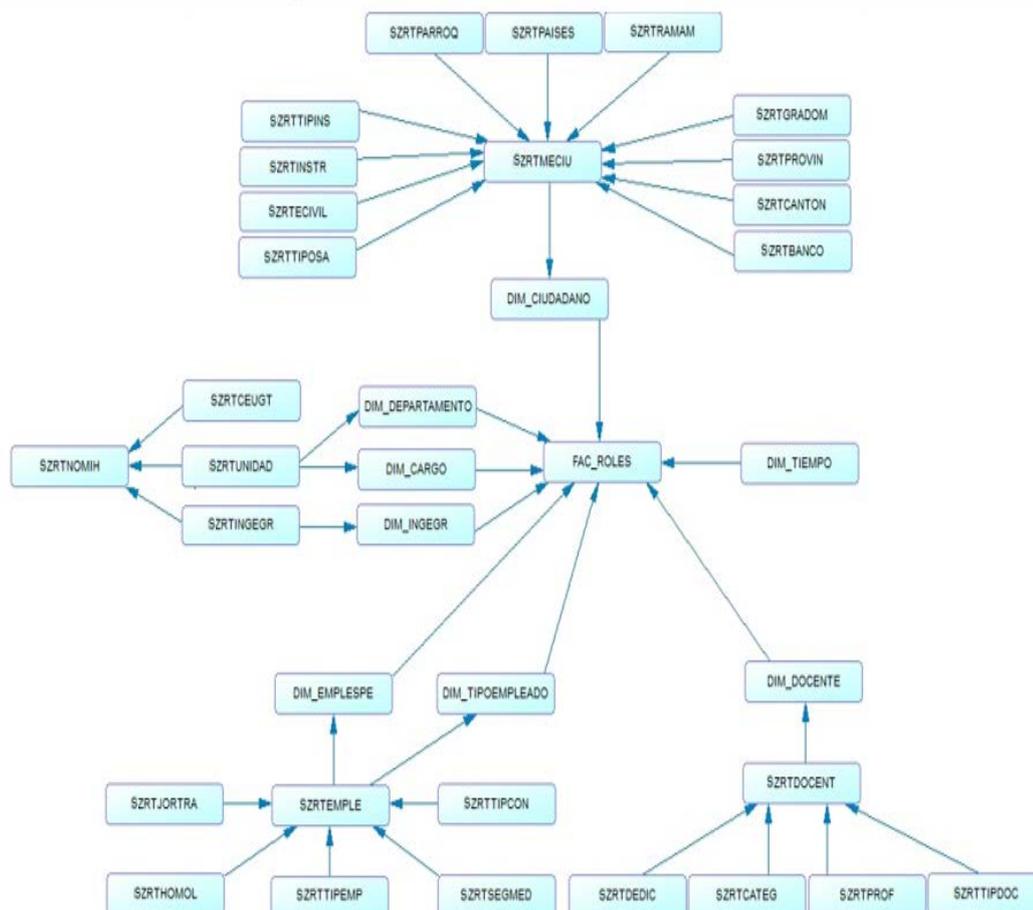


Figura 73: Diagrama general roles de pago – shifre

4.6.2.17 Mapeo de datos roles de pago (shifre)

En la siguiente tabla se describe el origen de datos desde el servidor de base de datos o archivos planos hacia las dimensiones y tablas de hechos del datamart.

Cuadro 64
Mapeo de datos roles de pago - shifre

FUENTE DE DATOS ORIGINAL	DESCRIPCIÓN	STAGE	DIMENSIÓN (DIM)	HECHOS (FAC)
peemd_emplepe petho_tabla_homologacion petem_tipoempleado pejor_jortrabajo petco_tipcont petsm_tipsegmedico	Tabla que contiene los datos del empleado de la Universidad de las Fuerzas Armadas ESPE.	szrtemple szrttipemp szrthomol szrtjotra szrttipcon szrtsegmed szrtunidad szrtsede	DIM_EMPLESPE	
ceugg_unidad_unigto	Tabla que contiene los cargos que tiene el empleado de la Universidad de las Fuerzas Armadas ESPE.	szrtunidad	DIM_CARGO	
petem_tipoempleado	Tabla que contiene el tipo de empleado que tiene cada persona.	szrttipemp	DIM_TIPOEMPLEADO	
metit_tipinstr mecan_canton mepvc_provin meprq_parroq mepai_paises metis_tiposa meciu_ciudadno ceban_banco meeci_estcivl meins_instruc peram_rama pegrm_grado	Tabla que contiene los datos personales de cada persona.	szrtmeciu szrttiposa szrtprovin szrtcanton szrtparroq szrtpais szrtcivil szrtinstr szrtgradom szrtbanco szrttramam szrttipins	DIM_CIUDADANO	
peemp_docente petdo_docente petie_dedicacion petpi_prof_inv pecat_docentes	Tabla que contiene los datos acerca de los docentes de la Universidad de las Fuerzas Armadas ESPE.	szrtdocent szrttipdoc szrtddedic szrtprof szrtcateg szrtunidad szrttipcon szrtsede szrtemple	DIM_DOCENTE	
ceugg_unidad_unigto	Tabla que contiene los departamentos que existen en la Universidad de las Fuerzas Armadas ESPE.	szrtunidad	DIM_DEPARTAMENTO	
peieg_ingeegre	Tabla que contiene los ingresos y los egresos que tiene cada empleado de la Universidad de las Fuerzas Armadas ESPE.	szrtingegr	DIM_INGEGR	
ceugt_unidad egasto ceugg_unidad_unigto	Tabla que contiene las sedes de la Universidad de las Fuerzas Armadas ESPE.	szrtsede	DIM_SEDE	
penmh_nomih perld_roldep os	Tabla que tiene los sueldos de los empleados de la Universidad de las Fuerzas Armadas ESPE.	szrtnomih		FAC_ROLES

4.6.2.18 Modelo dimensional roles de pago (shifre)

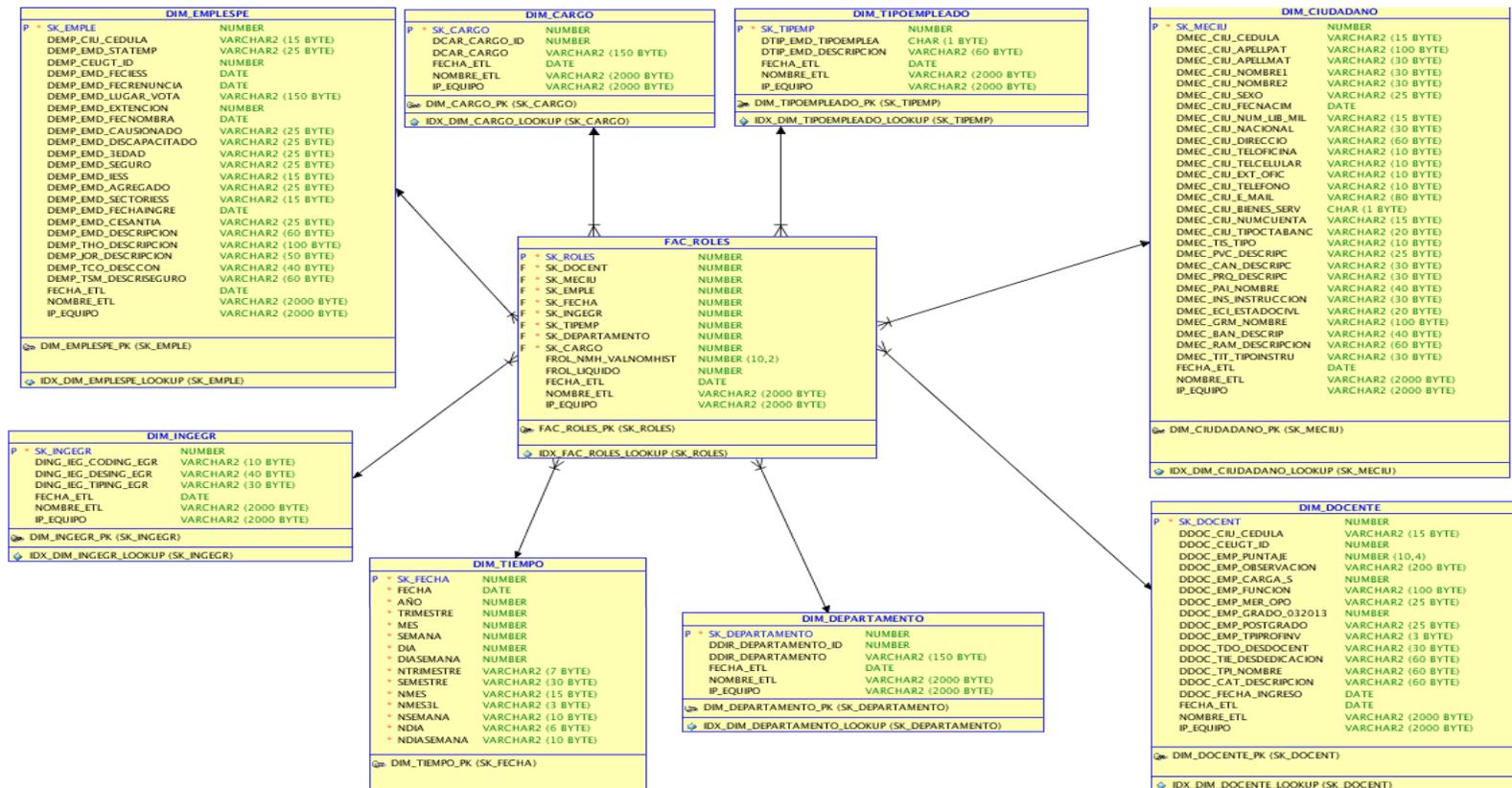


Figura 74: Modelo dimensional roles de pago - shifre

4.6.2.19 Mapeo de datos para tablas de dimensiones de Roles de pago (Shifre).

Las siguientes tablas muestran el destino y origen de los datos para dimensiones del módulo roles de pago, se detalla nombres de los campos de destino y origen, los tipos de datos y sus esquemas respectivos.

Cuadro 65
Estructura dimensión cargo

TABLA	DIM_CARGO								
TIPO DE TABLA	DIMENSIONAL								
DESCRIPCIÓN	TABLA CON REGISTROS DE LOS CARGOS DEL PERSONAL.								
ESQUEMA	DWESPE								
		DESTINO				ORIGEN			
Columna	Descripción	Tipo de dato	Clave	Fk para	Null	Esquema	Tabla	Campo	Tipo de dato
sk_cargo	Clave subrogada de la dimensión cargo.	Number(38)	pk	FAC_ROLES	N	Sistema	Sistema	Sistema	Number(38)
dcar_cargo_id	Código del cargo.	Number(38)			Y	STGESPE	szrtunidad	szrtunidad_ceugt_id	Number(38)
dcar_cargo	Nombre del cargo	Varchar2(150)			Y	STGESPE	szrtunidad	szrtunidad_ceuug_desc	Varchar2(150)
nombre_etl	Nombre de la transformación	Varchar2(2000)			Y	Sistema	Sistema	Sistema	Varchar2(2000)
fecha_etl	Fecha en la que se realiza la transformación	Date			Y	Sistema	Sistema	Sistema	Date
ip_equipo	Dirección ip del equipo	Varchar2(2000)			Y	Sistema	Sistema	Sistema	Varchar2(2000)

Cuadro 66
Estructura dimensión tipo de empleado

TABLA		DIM_TIPOEMPLEADO							
TIPO DE TABLA		DIMENSIONAL							
DESCRIPCIÓN		TABLA CON REGISTROS DE LOS TIPOS DE EMPLEADOS.							
ESQUEMA		DWESPE							
		DESTINO				ORIGEN			
Columna	Descripción	Tipo de dato	Clave	Fk para	Null	Esquema	Tabla	Campo	Tipo de dato
sk_tipemp	Clave subrogada de la dimensión cargo.	Number(38)	pk	FAC_ROLES	N	Sistema	Sistema	Sistema	Number(38)
dtip_emd_tipoemplea	Código del tipo de empleado.	char(1)			Y	STGESPE	szrttipemp	szrttipemp_emd_tipoemplea	char(1)
dtip_emd_descripcion	Descripción del tipo de empleado	Varchar2(60)			Y	STGESPE	szrttipemp	szrttipemp_emd_descripcion	Varchar2(60)
nombre_etl	Nombre de la transformación	Varchar2(2000)			Y	Sistema	Sistema	Sistema	Varchar2(2000)
fecha_etl	Fecha en la que se realiza la transformación	Date			Y	Sistema	Sistema	Sistema	Date
ip_equipo	Dirección ip del equipo	Varchar2(2000)			Y	Sistema	Sistema	Sistema	Varchar2(2000)

Cuadro 67
Estructura dimensión empleado de la Universidad de las Fuerzas Armadas ESPE

TABLA	DIM_EMPLESPE								
TIPO DE TABLA	DIMENSIONAL								
DESCRIPCIÓN	TABLA CON REGISTROS DE LOS DATOS DEL EMPLEADO DE LA UNIVERSIDAD DE LAS FUERZAS ARMADAS ESPE.								
ESQUEMA	DWESPE								
	DESTINO					ORIGEN			
Columna	Descripción	Tipo de dato	Clave	Fk para	Null	Esquema	Tabla	Campo	Tipo de dato
sk_emple	Clave subrogada de la dimensión cargo.	Number(38)	pk	FAC_ROLES	N	Sistema	Sistema	Sistema	Number(38)
demp_ciu_cedula	Cedula del empleado.	Varchar2(15)			Y	STGESPE	szrtemplate	szrtemplate_ciu_cedula	Varchar2(15)
demp_emd_statemp	Estado del empleado.	Varchar2(25)			Y	STGESPE	szrtemplate	szrtemplate_emd_statemp	Varchar2(25)
demp_ceugt_id	Código ceugt	Number(38)			Y	STGESPE	szrtemplate	szrtemplate_ceugt_id	Number(38)
demp_emd_feciess	Fecha de afiliado al iess.	Date			Y	STGESPE	szrtemplate	szrtemplate_emd_feciess	Date
demp_emd_fecrenuncia	Fecha de renuncia.	Date			Y	STGESPE	szrtemplate	szrtemplate_emd_fecrenuncia	Date
demp_emd_lugar_vota	Lugar de votación.	Varchar2(150)			Y	STGESPE	szrtemplate	szrtemplate_emd_lugar_vota	Varchar2(150)
demp_emd_extencion	Extensión del número de teléfono	Number(38)			Y	STGESPE	szrtemplate	szrtemplate_emd_extencion	Number(38)
demp_emd_fecnombra	Fecha del nombramiento.	Date			Y	STGESPE	szrtemplate	szrtemplate_emd_fecnombra	Date
demp_emd_causionado	Causionado (si-no)	Varchar2(25)			Y	STGESPE	szrtemplate	szrtemplate_emd_causionado	Varchar2(25)
demp_emd_discapacitado	Discapacitado (si-no)	Varchar2(25)			Y	STGESPE	szrtemplate	szrtemplate_emd_discapacitado	Varchar2(25)
demp_emd_3edad	Tercera edad (si-no)	Varchar2(25)			Y	STGESPE	szrtemplate	szrtemplate_emd_3edad	Varchar2(25)
demp_emd_seguro	Seguro (si-no)	Varchar2(25)			Y	STGESPE	szrtemplate	szrtemplate_emd_seguro	Varchar2(25)
demp_emd_iess	Numero de iess	Varchar2(25)			Y	STGESPE	szrtemplate	szrtemplate_emd_iess	Varchar2(25)
demp_emd_agregado	Agregado (si-no)	Varchar2(25)			Y	STGESPE	szrtemplate	szrtemplate_emd_agregado	Varchar2(25)
demp_emd_sectoriess	Sector iess	Varchar2(25)			Y	STGESPE	szrtemplate	szrtemplate_emd_sectoriess	Varchar2(25)
demp_emd_fechaingre	Fecha de ingreso a la Universidad de las Fuerzas Armadas ESPE.	Date			Y	STGESPE	szrtemplate	szrtemplate_emd_fechaingre	Date
demp_emd_cesantia	Cesantía (si-no)	Varchar2(25)			Y	STGESPE	szrtemplate	szrtemplate_emd_cesantia	Varchar2(25)
demp_emd_descripcion	Tipo de empleado.	Varchar2(60)			Y	STGESPE	szrttipemp	szrttipemp_emd_descripcion	Varchar2(60)
demp_tho_descripcion	Descripción de cargo	Varchar2(100)			Y	STGESPE	szrthomol	szrthomol_tho_descripcion	Varchar2(100)
demp_jor_descripcion	Descripción de jornadas	Varchar2(50)			Y	STGESPE	szrtjortra	szrtjortra_jor_descripcion	Varchar2(50)
demp_tco_desccon	Descripción de nombramiento	Varchar2(40)			Y	STGESPE	szrttipcon	szrttipcon_tco_desccon	Varchar2(40)
demp_tsm_descriseguro	Tipo de seguro.	Varchar2(60)			Y	STGESPE	szrtsegmed	szrtsegmed_tsm_descriseguro	Varchar2(60)
nombre_etl	Nombre de la transformación	Varchar2(2000)			Y	Sistema	Sistema	Sistema	Varchar2(2000)
fecha_etl	Fecha en la que se realiza la transformación	Date			Y	Sistema	Sistema	Sistema	Date
ip_equipo	Dirección ip del equipo	Varchar2(2000)			Y	Sistema	Sistema	Sistema	Varchar2(2000)

Cuadro 68

Estructura dimensión ciudadano

TABLA	DIM_CIUADANO								
TIPO DE TABLA	DIMENSIONAL								
DESCRIPCIÓN	TABLA CON REGISTROS DE LOS DATOS PERSONALES.								
ESQUEMA	DWESPE								
	DESTINO					ORIGEN			
Columna	Descripción	Tipo de dato	Clave	Fk para	Null	Esquema	Tabla	Campo	Tipo de dato
sk_meciu	Clave subrogada de la dimensión ciudadano.	Number(38)	pk	FAC_ROLES	N	Sistema	Sistema	Sistema	Number(38)
dmecciu_cedula	Número de cédula.	Varchar2(15)			Y	STGESPE	szrtmeciu	szrtmeciu_ciu_cedula	Varchar2(15)
dmecciu_apellpat	Apellido paterno.	Varchar2(100)			Y	STGESPE	szrtmeciu	szrtmeciu_ciu_apellpat	Varchar2(100)
dmecciu_apellmat	Apellido materno.	Varchar2(30)			Y	STGESPE	szrtmeciu	szrtmeciu_ciu_apellmat	Varchar2(30)
dmecciu_nombre1	Primero nombre.	Varchar2(30)			Y	STGESPE	szrtmeciu	szrtmeciu_ciu_nombre1	Varchar2(30)
dmecciu_nombre2	Segundo nombre.	Varchar2(30)			Y	STGESPE	szrtmeciu	szrtmeciu_ciu_nombre2	Varchar2(30)
dmecciu_sexo	Sexo del ciudadano.	Varchar2(25)			Y	STGESPE	szrtmeciu	szrtmeciu_ciu_sexo	Varchar2(25)
dmecciu_fecnacim	Fecha de nacimiento.	Date			Y	STGESPE	szrtmeciu	szrtmeciu_ciu_fecnacim	Date
dmecciu_num_lib_mil	Numero de libreta militar.	Varchar2(15)			Y	STGESPE	szrtmeciu	szrtmeciu_ciu_num_lib_mil	Varchar2(15)
dmecciu_nacional	Nacionalidad.	Varchar2(30)			Y	STGESPE	szrtmeciu	szrtmeciu_ciu_nacional	Varchar2(30)
dmecciu_direccion	Dirección del ciudadano.	Varchar2(60)			Y	STGESPE	szrtmeciu	szrtmeciu_ciu_direccio	Varchar2(60)
dmecciu_teloficina	Número de teléfono de la oficina.	Varchar2(10)			Y	STGESPE	szrtmeciu	szrtmeciu_ciu_teloficina	Varchar2(10)
dmecciu_telcelular	Número del teléfono celular.	Varchar2(10)			Y	STGESPE	szrtmeciu	szrtmeciu_ciu_telcelular	Varchar2(10)
dmecciu_ext_ofic	Número de extensión.	Varchar2(10)			Y	STGESPE	szrtmeciu	szrtmeciu_ciu_ext_ofic	Varchar2(10)
dmecciu_telefono	Número de teléfono del domicilio.	Varchar2(10)			Y	STGESPE	szrtmeciu	szrtmeciu_ciu_telefono	Varchar2(10)
dmecciu_e_mail	Correo electrónico.	Varchar2(80)			Y	STGESPE	szrtmeciu	szrtmeciu_ciu_e_mail	Varchar2(80)
dmecciu_bienes_serv	Bienes del servicio.	Char(1)			Y	STGESPE	szrtmeciu	szrtmeciu_ciu_bienes_serv	Char(1)
dmecciu_numcuenta	Número de la cuenta bancaria.	Varchar2(25)			Y	STGESPE	szrtmeciu	szrtmeciu_ciu_numcuenta	Varchar2(25)
dmecciu_tipoctabanc	Tipo de cuenta bancaria.	Varchar2(30)			Y	STGESPE	szrtmeciu	szrtmeciu_ciu_tipoctabanc	Varchar2(30)
dmecciu_tis_tipo	Tipo de sangre.	Varchar2(30)			Y	STGESPE	szrttiposa	szrttiposa_tis_tipo	Varchar2(30)
dmecciu_pvc_descripc	Provincia del domicilio.	Varchar2(40)			Y	STGESPE	szrtprovin	szrtprovin_pvc_descripc	Varchar2(40)
dmecciu_can_descripc	Cantón del domicilio.	Varchar2(30)			Y	STGESPE	szrtcanton	szrtcanton_can_descripc	Varchar2(30)
dmecciu_prq_descripc	Parroquia del domicilio.	Varchar2(20)			Y	STGESPE	szrtparroq	szrtparroq_prq_descripc	Varchar2(20)
dmecciu_pai_nombre	País del domicilio.	Varchar2(100)			Y	STGESPE	szrtpaíses	szrtpaíses_pai_nombre	Varchar2(100)
dmecciu_ins_instruccion	Nombre de instrucción.	Varchar2(40)			Y	STGESPE	szrtinstr	szrtinstr_ins_instruccion	Varchar2(40)
dmecciu_eci_estadocivil	Estado civil.	Varchar2(60)			Y	STGESPE	szrtcivil	szrtcivil_eci_estadocivil	Varchar2(60)
dmecciu_grm_nombre	Nombre del grado.	Varchar2(40)			Y	STGESPE	szrtgradom	szrtgradom_grm_nombre	Varchar2(40)
dmecciu_ban_descripc	Descripción del banco.	Varchar2(30)			Y	STGESPE	szrtbanco	szrtbanco_ban_descripc	Varchar2(30)
dmecciu_ram_descripc	Nombre de la rama militar.	Varchar2(40)			Y	STGESPE	szrttramam	szrttramam_ram_descripcion	Varchar2(40)
dmecciu_tit_tipoinstru	Nombre del tipo de instrucción	Varchar2(30)			Y	STGESPE	szrttipins	szrttipins_tit_tipoinstru	Varchar2(30)

Cuadro 69
Estructura dimensión ingresos y egresos

TABLA		DIM_INGEGR							
TIPO DE TABLA		DIMENSIONAL							
DESCRIPCIÓN		TABLA CON REGISTROS DE LOS RUBROS Y TIPOS DE RUBROS.							
ESQUEMA		DWESPE							
		DESTINO				ORIGEN			
Columna	Descripción	Tipo de dato	Clave	Fk para	Null	Esquema	Tabla	Campo	Tipo de dato
sk_ingegr	Clave subrogada de la dimensión ingegr.	Number(38)	pk	FAC_ROLES	N	Sistema	Sistema	Sistema	Number(38)
ding_ieg_coding_egr	Código del rubro.	Varchar2(10)			Y	STGESPE	szrtingegr	szrtingegr_ieg_coding_egr	Varchar2(10)
ding_ieg_desing_egr	Nombre del rubro	Varchar2(40)			Y	STGESPE	szrtingegr	szrtingegr_ieg_desing_egr	Varchar2(40)
ding_ieg_tiping_egr	Descripción del tipo de rubro	Varchar2(30)			Y	STGESPE	szrtingegr	szrtingegr_ieg_tiping_egr	Varchar2(30)
nombre_etl	Nombre de la transformación	Varchar2(2000)			Y	Sistema	Sistema	Sistema	Varchar2(2000)
fecha_etl	Fecha en la que se realiza la transformación	Date			Y	Sistema	Sistema	Sistema	Date
ip_equipo	Dirección ip del equipo	Varchar2(2000)			Y	Sistema	Sistema	Sistema	Varchar2(2000)

Cuadro 70
Estructura dimensión departamento

TABLA		DIM_DEPARTAMENTO							
TIPO DE TABLA		DIMENSIONAL							
DESCRIPCIÓN		TABLA CON REGISTROS DE LOS DEPARTAMENTOS DE LA UNIVERSIDAD DE LAS FUERZAS ARMADAS ESPE.							
ESQUEMA		DWESPE							
		DESTINO				ORIGEN			
Columna	Descripción	Tipo de dato	Clave	Fk para	Null	Esquema	Tabla	Campo	Tipo de dato
sk_departamento	Clave subrogada de la dimensión departamento.	Number(38)	pk	FAC_ROLES	N	Sistema	Sistema	Sistema	Number(38)
ddir_departamento_id	Código del departamento.	Number(38)			Y	STGESPE	szrtunidad	szrtunidad_ceugt_id	Number(38)
ddir_departamento	Nombre del departamento	Varchar2(150)			Y	STGESPE	szrtunidad	szrtunidad_ceuug_desc	Varchar2(150)
nombre_etl	Nombre de la transformación	Varchar2(2000)			Y	Sistema	Sistema	Sistema	Varchar2(2000)
fecha_etl	Fecha en la que se realiza la transformación	Date			Y	Sistema	Sistema	Sistema	Date
ip_equipo	Dirección ip del equipo	Varchar2(2000)			Y	Sistema	Sistema	Sistema	Varchar2(2000)

Cuadro 71: Estructura dimensión docente

TABLA		DIM_DOCENTE							
TIPO DE TABLA		DIMENSIONAL							
DESCRIPCIÓN		TABLA CON REGISTROS DE LOS DATOS DEL DOCENTE.							
ESQUEMA		DWESPE							
		DESTINO				ORIGEN			
Columna	Descripción	Tipo de dato	Clave	Fk para	Null	Esquema	Tabla	Campo	Tipo de dato
sk_docent	Clave subrogada de la dimensión cargo.	Number(38)	pk	FAC_ROLES	N	Sistema	Sistema	Sistema	Number(38)
ddoc_ciu_cedula	Número de cedula.	Varchar2(15)			Y	STGESPE	szrtdocent	szrtdocent_ciu_cedula	Varchar2(15)
ddoc_ceugt_id	Código ceugt	Number(38)			Y	STGESPE	szrtdocent	szrtdocent_ceugt_id	Number(38)
ddoc_emp_puntaje	Valor de puntaje.	Number(10,4)			Y	STGESPE	szrtdocent	szrtdocent_emp_puntaje	Number(10,4)
ddoc_emp_observacion	Observación del docente.	Varchar2(200)			Y	STGESPE	szrtdocent	szrtdocent_emp_observacion	Varchar2(200)
ddoc_emp_carga_s	Carga semanal.	Number(38)			Y	STGESPE	szrtdocent	szrtdocent_emp_carga_s	Number(38)
ddoc_emp_funcion	Función del docente.	Varchar2(100)			Y	STGESPE	szrtdocent	szrtdocent_emp_funcion	Varchar2(100)
ddoc_emp_mer_opo	Méritos y oposición.	Varchar2(25)			Y	STGESPE	szrtdocent	szrtdocent_emp_mer_opo	Varchar2(25)
ddoc_emp_grado_032013	Grado RMU2013	Number(38)			Y	STGESPE	szrtdocent	szrtdocent_emp_grado_032013	Number(38)
ddoc_emp_postgrado	Descripción de postgrado.	Varchar2(25)			Y	STGESPE	szrtdocent	szrtdocent_emp_postgrado	Varchar2(25)
ddoc_emp_tpiprofilv	Código de profesión.	Varchar2(3)			Y	STGESPE	szrtdocent	szrtdocent_emp_tpiprofilv	Varchar2(3)
ddoc_tdo_desdocent	Título de docente.	Varchar2(30)			Y	STGESPE	szrttipdoc	szrttipdoc_tdo_desdocent	Varchar2(30)
ddoc_tie_desdedicacion	Descripción de dedicación.	Varchar2(60)			Y	STGESPE	szrtdedic	szrtdedic_tie_desdedicacion	Varchar2(60)
ddoc_tpi_nombre	Tipo de docente.	Varchar2(60)			Y	STGESPE	szrtprof	szrtprof_tpi_nombre	Varchar2(60)
ddoc_cat_descripcion	Descripción de categoría.	Varchar2(60)			Y	STGESPE	szrtcateg	szrtcateg_cat_descripcion	Varchar2(60)
ddoc_fecha_ingreso	Fecha de ingreso del docente.	Date			Y	STGESPE	szrtdocent	szrtdocent_fecingdoc	Date
nombre_etl	Nombre de la transformación	Varchar2(2000)			Y	Sistema	Sistema	Sistema	Varchar2(2000)
fecha_etl	Fecha en la que se realiza la transformación	Date			Y	Sistema	Sistema	Sistema	Date
ip_equipo	Dirección ip del equipo	Varchar2(2000)			Y	Sistema	Sistema	Sistema	Varchar2(2000)

4.6.2.20 Mapeo de datos para tablas de hechos de Roles de pago (Shifre).

Cuadro 72
Estructura tabla de hechos roles de pago - shifre

TABLA		FAC_ROLES							
TIPO DE TABLA		HECHOS							
DESCRIPCIÓN		TABLA DE HECHOS CON REGISTROS DE LOS CARGOS DEL PERSONAL.							
ESQUEMA		DESTINO				ORIGEN			
Columna	Descripción	Tipo de dato	Clave	Fk para	Null	Esquema	Tabla	Campo	Tipo de dato
sk_roles	Clave subrogada de la tabla de hechos de roles.	Number(38)	pk		N	Sistema	Sistema	Sistema	Number(38)
sk_docent	Clave subrogada de la dimensión docente.	Number(38)	Fk		N	DWESPE	dim_docente	sk_docent	Number(38)
sk_meciu	Clave subrogada de la dimensión ciudadano.	Number(38)	Fk		N	DWESPE	dim_ciudadano	sk_meciu	Number(38)
sk_emple	Clave subrogada de la dimensión emplespe	Number(38)	Fk		N	DWESPE	dim_emplespe	sk_emple	Number(38)
sk_fecha	Clave subrogada de la dimensión tiempo.	Number(38)	Fk		N	DWESPE	dim_tiempo	sk_fecha	Number(38)
sk_ingegr	Clave subrogada de la dimensión ingegr.	Number(38)	Fk		N	DWESPE	dim_ingegr	sk_ingegr	Number(38)
sk_tipemp	Clave subrogada de la dimensión tipoempleado	Number(38)	Fk		N	DWESPE	dim_tipoempleado	sk_tipemp	Number(38)
sk_departamento	Clave subrogada de la dimensión departamento.	Number(38)	Fk		N	DWESPE	dim_departamento	sk_departamento	Number(38)
sk_cargo	Clave subrogada de la dimensión cargo.	Number(38)	Fk		N	DWESPE	dim_cargo	sk_cargo	Number(38)
frol_nmh_valnomhist	Medida con el valor del sueldo	Number(38)			Y	STGESPE	szrtnomih	szrtnomih_nmh_valnomhist	Number(38)
frol_liquido	Medida con el valor del sueldo líquido.	Number(38)			Y	STGESPE	szrtnomih	szrtnomih_liquido	Number(38)
nombre_etl	Nombre de la transformación	Varchar2(2000)			Y	Sistema	Sistema	Sistema	Varchar2(2000)
fecha_etl	Fecha en la que se realiza la transformación	Date			Y	Sistema	Sistema	Sistema	Date
ip_equipo	Dirección ip del equipo	Varchar2(2000)			Y	Sistema	Sistema	Sistema	Varchar2(2000)

4.6.2.21 Diagrama general para el módulo de matrículas.

El diagrama muestra las dimensiones y la relación que tienen con las tablas de la base de datos de origen.

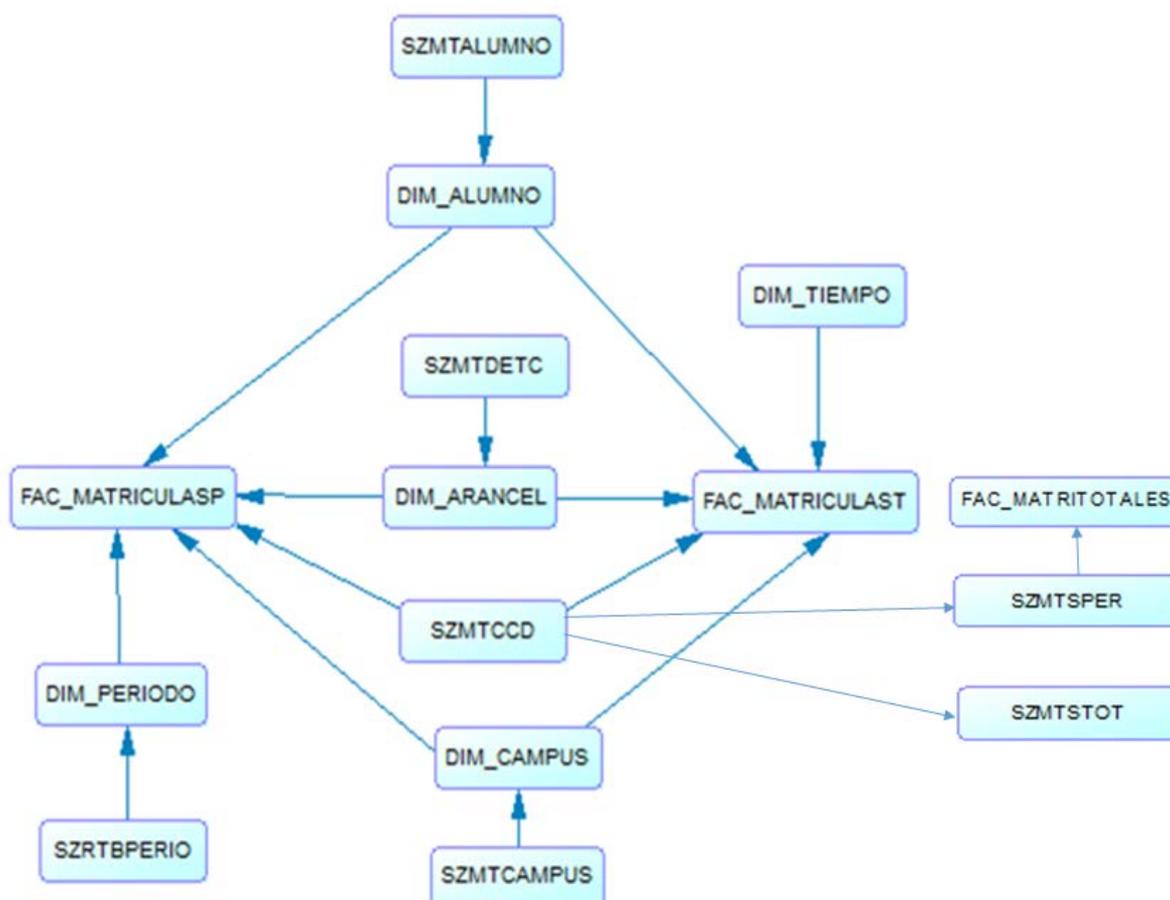


Figura 75: Diagrama general matrículas

4.6.2.22 Mapeo de datos de matrículas.

En la siguiente tabla se describe el origen de datos desde el servidor de base de datos o archivos planos hacia las dimensiones y tablas de hechos del datamart.

Cuadro 73
Mapeo de datos matrículas

FUENTE DE DATOS ORIGINAL	DESCRIPCIÓN	STAGE	DIMENSIÓN (DIM)	HECHOS (FAC)
tbbdetc	Tabla que contiene los aranceles agrupados y desglosados de la Universidad de las Fuerzas Armadas ESPE.	szmtdetc	DIM_ARANCEL	
stvterm	Tabla que contiene los periodos del sistema banner de la Universidad de las Fuerzas Armadas ESPE.	szrtbperio	DIM_PERIODO	
spriden, tbraccd	Tabla que contiene los nombres y tipos de alumno.	szmtalumno	DIM_ALUMNO	
spriden	Tabla que contiene los campus de la Universidad de las Fuerzas Armadas ESPE.	szrtbcamp	DIM_CAMPUS	
tbraccd	Tabla que contiene los pagos y deudas de los alumnos por periodo.	szmtalumno szmtccd		FAC_MATRICULASP
tbraccd	Tabla que contiene los pagos y deudas de los alumnos por tiempo.	szmtalumno szmtccd		FAC_MATRICULAST
tbraccd	Tabla que contiene los pagos y deudas de los alumnos por tiempo.	szmtstot szmtsper		FAC_MATRITOTALES

4.6.2.24 Mapeo de datos para tablas de dimensiones de matrículas.

Las siguientes tablas muestran el destino y origen de los datos para dimensiones del módulo de matrículas, se detalla nombres de los campos de destino y origen, los tipos de datos y sus esquemas respectivos.

Cuadro 74
Estructura dimensión campus banner

TABLA		DIM_CAMPUS_B							
TIPO DE TABLA		DIMENSIONAL							
DESCRIPCIÓN		TABLA CON REGISTROS DE LOS CAMPUS DE LA UNIVERSIDAD DE LAS FUERZAS ARMADAS ESPE.							
ESQUEMA		DWESPE							
		DESTINO				ORIGEN			
Columna	Descripción	Tipo de dato	Clave	Fk para	Null	Esquema	Tabla	Campo	Tipo de dato
sk_campus	Clave subrogada de la dimensión campus.	Number(38)	pk	FAC_MATRICULAS	N	Sistema	Sistema	Sistema	Number(38)
dcam_cod_campus	Código del campus.	Varchar2(100)			Y	STGESPE	szrtbcamp	szrtbcamp_cod_campus	Varchar2(100)
dcam_campus	Nombre del campus.	Varchar2(100)			Y	STGESPE	szrtbcamp	szrtbcamp_campus	Varchar2(100)
nombre_etl	Nombre de la transformación	Varchar2(2000)			Y	Sistema	Sistema	Sistema	Varchar2(2000)
fecha_etl	Fecha en la que se realiza la transformación	Date			Y	Sistema	Sistema	Sistema	Date
ip_equipo	Dirección ip del equipo	Varchar2(2000)			Y	Sistema	Sistema	Sistema	Varchar2(2000)

Cuadro 75
Estructura dimensión alumno banner

TABLA		DIM_ALUMNO							
TIPO DE TABLA		DIMENSIONAL							
DESCRIPCIÓN		TABLA CON REGISTROS DE LOS ALUMNOS DE LA UNIVERSIDAD DE LAS FUERZAS ARMADAS ESPE.							
ESQUEMA		DWESPE							
		DESTINO				ORIGEN			
Columna	Descripción	Tipo de dato	Clave	Fk para	Null	Esquema	Tabla	Campo	Tipo de dato
sk_alumno	Clave subrogada de la dimensión alumno.	Number(38)	pk	FAC_MATRICULAS	N	Sistema	Sistema	Sistema	Number(38)
dalu_id	Código id del alumno.	Varchar2(9)			Y	STGESPE	szmtalumno	szmtalumno_id	Varchar2(100)
dalu_pidm	Código pidm del alumno.	Number(38)			Y	STGESPE	szmtalumno	szmtalumno_pidm	Number(38)
dalu_programa	Programa del alumno.	Varchar2(100)			Y	STGESPE	szmtalumno	szmtalumno_programa	Varchar2(100)
dalu_apellido	Apellidos del alumno.	Varchar2(60)			Y	STGESPE	szmtalumno	szmtalumno_apellido	Varchar2(60)
dalu_nombre	Nombres del alumno.	Varchar2(260)			Y	STGESPE	szmtalumno	szmtalumno_nombre	Varchar2(60)
dalu_tipo_alumno	Tipo de alumno.	Varchar2(100)			Y	STGESPE	szmtalumno	szmtalumno_tipo_alumno	Varchar2(100)
nombre_etl	Nombre de la transformación	Varchar2(2000)			Y	Sistema	Sistema	Sistema	Varchar2(2000)
fecha_etl	Fecha en la que se realiza la transformación	Date			Y	Sistema	Sistema	Sistema	Date
ip_equipo	Dirección ip del equipo	Varchar2(2000)			Y	Sistema	Sistema	Sistema	Varchar2(2000)

Cuadro 76
Estructura dimensión arancel banner

TABLA		DIM_ARANCEL							
TIPO DE TABLA		DIMENSIONAL							
DESCRIPCIÓN		TABLA CON REGISTROS DE LOS ARANCELES DE LA UNIVERSIDAD DE LAS FUERZAS ARMADAS ESPE.							
ESQUEMA		DWESPE							
		DESTINO				ORIGEN			
Columna	Descripción	Tipo de dato	Clave	Fk para	Null	Esquema	Tabla	Campo	Tipo de dato
sk_arancel	Clave subrogada de la dimensión arancel.	Number(38)	pk	FAC_MATRICULAS	N	Sistema	Sistema	Sistema	Number(38)
dara_detail_code	Código del arancel.	Varchar2(4)			Y	STGESPE	szmtdetc	szmtdetc_detail_code	Varchar2(4)
dara_descripcion	Nombre del arancel.	Varchar2(30)			Y	STGESPE	szmtdetc	szmtdetc_descripcion	Varchar2(30)
dara_totales	Aranceles agrupados.	Varchar2(30)			Y	STGESPE	szmtdetc	szmtdetc_totales	Varchar2(30)
nombre_etl	Nombre de la transformación	Varchar2(2000)			Y	Sistema	Sistema	Sistema	Varchar2(2000)
fecha_etl	Fecha en la que se realiza la transformación	Date			Y	Sistema	Sistema	Sistema	Date
ip_equipo	Dirección ip del equipo	Varchar2(2000)			Y	Sistema	Sistema	Sistema	Varchar2(2000)

Cuadro 77
Estructura dimensión periodo banner

TABLA		DIM_PERIODO							
TIPO DE TABLA		DIMENSIONAL							
DESCRIPCIÓN		TABLA CON REGISTROS DE LOS PERIODOS DEL SISTEMA BANNER DE LA UNIVERSIDAD DE LAS FUERZAS ARMADAS ESPE.							
ESQUEMA		DWESPE							
		DESTINO				ORIGEN			
Columna	Descripción	Tipo de dato	Clave	Fk para	Null	Esquema	Tabla	Campo	Tipo de dato
sk_periodo	Clave subrogada de la dimensión periodo.	Number(38)	pk	FAC_MATRICULAS	N	Sistema	Sistema	Sistema	Number(38)
dper_code	Código del periodo.	Varchar2(6)			Y	STGESPE	szrtbperio	szrtbperio_code	Varchar2(6)
dper_desc	Nombre del periodo.	Varchar2(30)			Y	STGESPE	szrtbperio	szrtbperio_desc	Varchar2(30)
dper_start_date	Fecha inicial del periodo.	Date			Y	STGESPE	szrtbperio	szrtbperio_start_date	Date
dper_end_date	Fecha final del periodo.	Date			Y	STGESPE	szrtbperio	szrtbperio_end_date	Date
nombre_etl	Nombre de la transformación	Varchar2(2000)			Y	Sistema	Sistema	Sistema	Varchar2(2000)
fecha_etl	Fecha en la que se realiza la transformación	Date			Y	Sistema	Sistema	Sistema	Date
ip_equipo	Dirección ip del equipo	Varchar2(2000)			Y	Sistema	Sistema	Sistema	Varchar2(2000)

4.6.2.25 Mapeo de datos para tablas de hechos de matrículas.

Cuadro 78
Estructura tabla de hechos matrículas por periodo

TABLA		FAC_MATRICULASP							
TIPO DE TABLA		HECHOS							
DESCRIPCIÓN		TABLA DE HECHOS CON REGISTROS DE LAS MATRÍCULAS DE LOS ALUMNOS POR PERIODO.							
ESQUEMA		DESTINO				ORIGEN			
Columna	Descripción	Tipo de dato	Clave	Fk para	Null	Esquema	Tabla	Campo	Tipo de dato
sk_matriculasp	Clave subrogada de la tabla de hechos de matrículas.	Number(38)	pk		N	Sistema	Sistema	Sistema	Number(38)
sk_alumno	Clave subrogada de la dimensión alumno.	Number(38)	Fk		N	DWESPE	dim_alumno	sk_alumno	Number(38)
sk_periodo	Clave subrogada de la dimensión periodo.	Number(38)	Fk		N	DWESPE	dim_periodo	sk_periodo	Number(38)
sk_arancel	Clave subrogada de la dimensión arancel	Number(38)	Fk		N	DWESPE	dim_arancel	sk_arancel	Number(38)
sk_campus	Clave subrogada de la dimensión campus.	Number(38)	Fk		N	DWESPE	dim_campus_b	sk_campus	Number(38)
fmat_balance	Medida con el valor del pago de matrículas.	Number(16,2)			Y	STGESPE	szmtccd	szmtccd_balance	Number(16,2)
nombre_etl	Nombre de la transformación	Varchar2(2000)			Y	Sistema	Sistema	Sistema	Varchar2(2000)
fecha_etl	Fecha en la que se realiza la transformación	Date			Y	Sistema	Sistema	Sistema	Date
ip_equipo	Dirección ip del equipo	Varchar2(2000)			Y	Sistema	Sistema	Sistema	Varchar2(2000)

Cuadro 79
Estructura tabla de hechos matrículas por tiempo

TABLA		FAC_MATRICULAST							
TIPO DE TABLA		HECHOS							
DESCRIPCIÓN		TABLA DE HECHOS CON REGISTROS DE LAS MATRÍCULAS DE LOS ALUMNOS POR TIEMPO.							
ESQUEMA		DESTINO				ORIGEN			
Columna	Descripción	Tipo de dato	Clave	Fk para	Null	Esquema	Tabla	Campo	Tipo de dato
sk_matriculasp	Clave subrogada de la tabla de hechos de matrículas.	Number(38)	pk		N	Sistema	Sistema	Sistema	Number(38)
sk_alumno	Clave subrogada de la dimensión alumno.	Number(38)	Fk		N	DWESPE	dim_alumno	sk_alumno	Number(38)
sk_fecha	Clave subrogada de la dimensión tiempo.	Number(38)	Fk		N	DWESPE	dim_tiempo	sk_fecha	Number(38)
sk_arancel	Clave subrogada de la dimensión arancel	Number(38)	Fk		N	DWESPE	dim_arancel	sk_arancel	Number(38)
sk_campus	Clave subrogada de la dimensión campus.	Number(38)	Fk		N	DWESPE	dim_campus_b	sk_campus	Number(38)
fmat_balance	Medida con el valor del pago de matrículas.	Number(16,2)			Y	STGESPE	szmtccd	szmtccd_balance	Number(16,2)
nombre_etl	Nombre de la transformación	Varchar2(2000)			Y	Sistema	Sistema	Sistema	Varchar2(2000)
fecha_etl	Fecha en la que se realiza la transformación	Date			Y	Sistema	Sistema	Sistema	Date

4.6.2.26 Diagrama general para el módulo de roles de pago - banner.

El diagrama muestra las dimensiones y la relación que tienen con las tablas de la base de datos de origen.

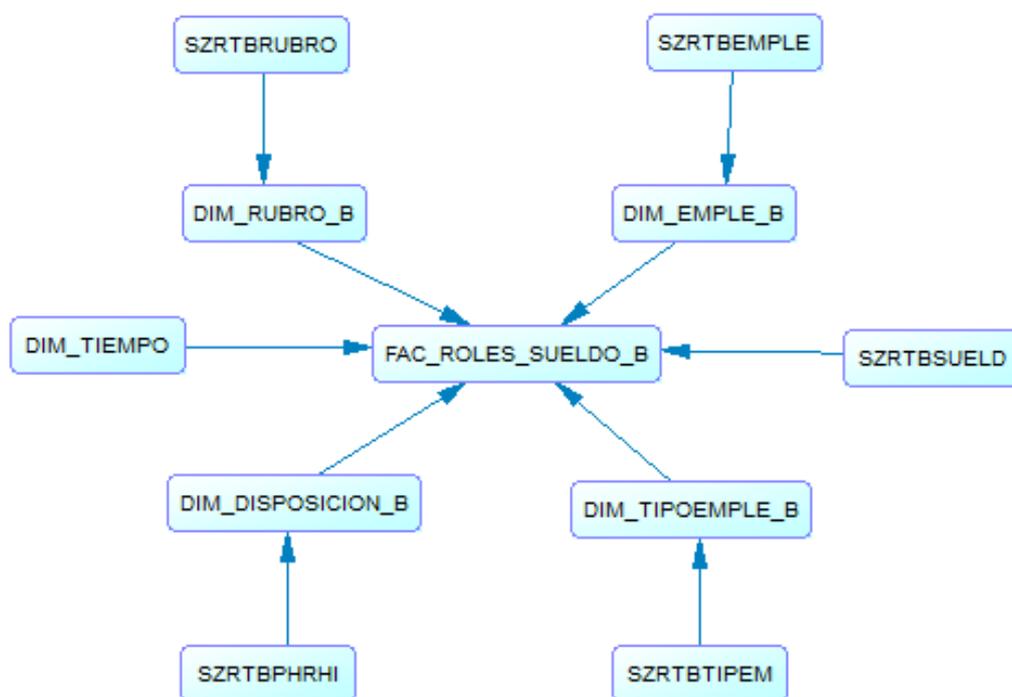


Figura 77: Diagrama general roles de pago - banner

4.6.2.27 Mapeo de datos roles de pago - banner.

En la siguiente tabla se describe el origen de datos desde el servidor de base de datos o archivos planos hacia las dimensiones y tablas de hechos del datamart.

Cuadro 80
Mapeo de datos roles de pago - banner

FUENTE DE DATOS ORIGINAL	DESCRIPCIÓN	STAGE	DIMENSIÓN (DIM)	HECHOS (FAC)
spbpers spriden stvethn stvmrtl stvcitz	Tablas que contienen todos los datos personales de los empleados administrativos, docentes y alumnos.	szrtbempl e	DIM_EMPLE_B	
phrhist	Tablas que contienen las disposiciones de la Universidad de las Fuerzas Armadas ESPE.	szrtbphrhi	DIM_DISPOSICION_B	
ptrpict	Tablas que contienen los distintos tipos de empleados de la Universidad de las Fuerzas Armadas ESPE.	szrtbtipem	DIM_TIPOEMPLE_B	
phrdedn ptrbdca ptrpict	Tablas que contienen los rubros y los tipos de rubros.	szrtbrub	DIM_RUBRO_B	
phrhist	Tabla que contiene los sueldos de los empleados.	szrtbsueld szrtbphrhi		FAC_ROLES_SUELDOS_B

4.6.2.28 Modelo dimensional de roles de pago - banner

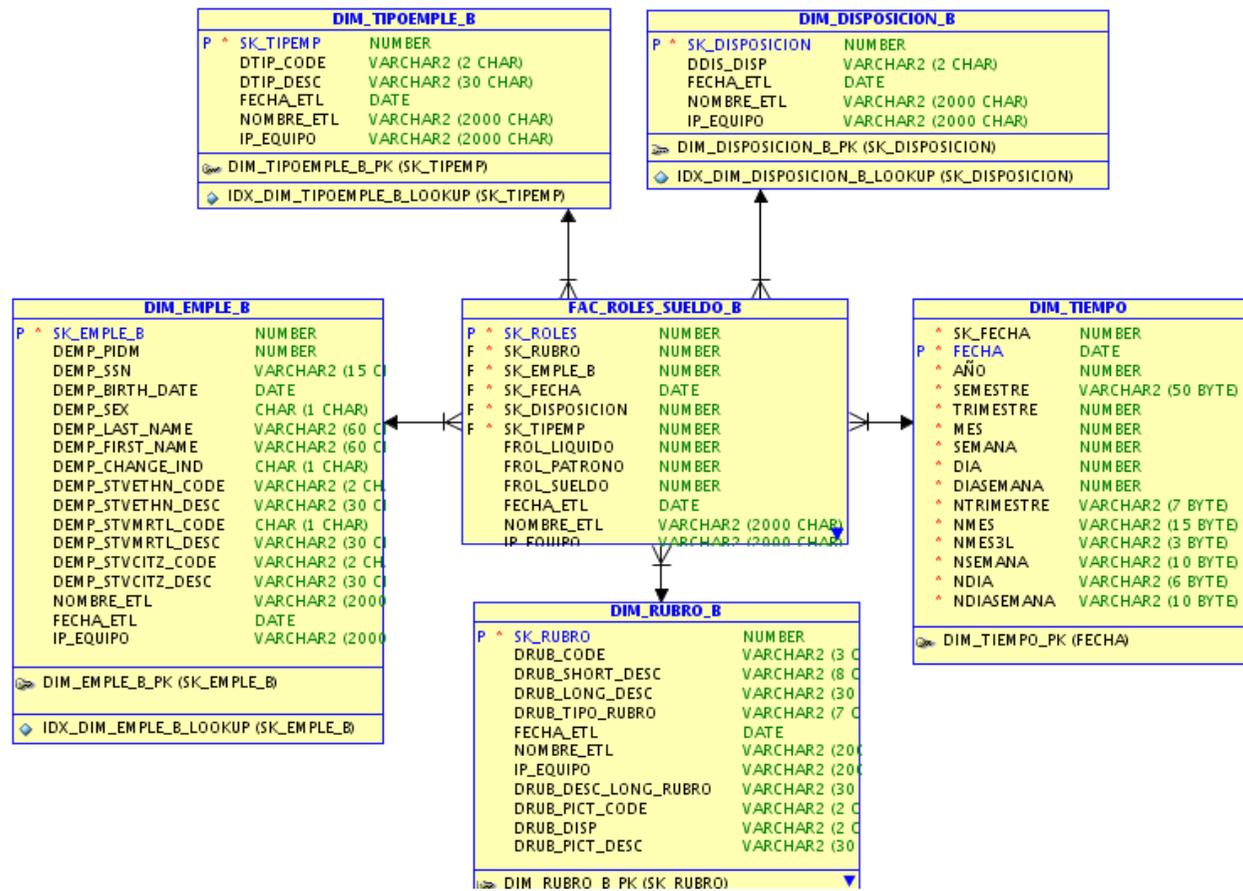


Figura 78: Modelo dimensional roles de pago – banne

4.6.2.29 Mapeo de datos para dimensiones y hechos de roles de pago - banner.

Las siguientes tablas muestran el destino y origen de los datos para dimensiones y tablas de hechos del módulo de roles de pago - banner, se detalla nombres de los campos de destino y origen, los tipos de datos y sus esquemas respectivos.

Cuadro 81
Estructura dimensión empleados banner

TABLA		DIM_EMPLE_B							
TIPO DE TABLA	DIMENSIONAL								
DESCRIPCIÓN	TABLA CON REGISTROS DE LOS EMPLEADOS DEL SISTEMA BANNER DE LA UNIVERSIDAD DE LAS FUERZAS ARMADAS ESPE.								
ESQUEMA	DWESPE								
		DESTINO				ORIGEN			
Columna	Descripción	Tipo de dato	Clave	Fk para	Null	Esquema	Tabla	Campo	Tipo de dato
sk_emple_b	Clave subrogada de la dimensión campus.	Number(38)	pk	FAC_ROLES_EMPLE_B,	N	Sistema	Sistema	Sistema	Number(38)
demp_pidm	Identificador del empleado docente o administrativo	Varchar2(2000)			Y	STGESPE	szrtbemple	szrtbemple_pidm	Varchar2(2000)
demp_ssn	Indica el número de cédula del empleado	Varchar2(2000)			Y	STGESPE	szrtbemple	szrtbemple_ssn	Varchar2(2000)
demp_birth_date	Muestra la fecha de nacimiento del empleado	Date			Y	STGESPE	szrtbemple	szrtbemple_birth_date	Date
demp_sex	Muestra el género del empleado	Varchar2(2000)			Y	STGESPE	szrtbemple	szrtbemple_sex	Varchar2(2000)
demp_last_name	Muestra los nombres del empleado	Varchar2(2000)			Y	STGESPE	szrtbemple	szrtbemple_last_name	Varchar2(2000)

demp_first_name	Muestra los apellidos del empleado	Varchar2(2000)	Y	STGESPE	szrtbemple	szrtbemple_first_name	Varchar2(2000)
demp_change_ind	Indica si es el registro más actual en la tabla	Varchar2(2000)	Y	STGESPE	szrtbemple	szrtbemple_change_ind	Varchar2(2000)
demp_stvethn_code	Muestra el código de la etnia del empleado	Varchar2(2000)	Y	STGESPE	szrtbemple	szrtbemple_stvethn_code	Varchar2(2000)
demp_stvethn_desc	Muestra la etnia del empleado	Varchar2(2000)	Y	STGESPE	szrtbemple	szrtbemple_stvethn_desc	Varchar2(2000)
demp_stvmrtl_code	Muestra el código de la libreta militar	Varchar2(2000)	Y	STGESPE	szrtbemple	szrtbemple_stvmrtl_code	Varchar2(2000)
demp_stvmrtl_desc	Muestra el número de la libreta militar	Varchar2(2000)	Y	STGESPE	szrtbemple	szrtbemple_stvmrtl_desc	Varchar2(2000)
demp_stvcitz_code	Muestra el código de la nacionalidad del empleado	Varchar2(2000)	Y	STGESPE	szrtbemple	szrtbemple_stvcitz_code	Varchar2(2000)
demp_stvcitz_desc	Muestra la nacionalidad del empleado	Varchar2(2000)	Y	STGESPE	szrtbemple	szrtbemple_stvcitz_desc	Varchar2(2000)
nombre_etl	Nombre de la transformación	Varchar2(2000)	Y	Sistema	Sistema	Sistema	Varchar2(2000)
fecha_etl	Fecha en la que se realiza la transformación	Date	Y	Sistema	Sistema	Sistema	Date
ip_equipo	Dirección ip del equipo	Varchar2(2000)	Y	Sistema	Sistema	Sistema	Varchar2(2000)

Cuadro 82
Estructura dimensión disposición banner

TABLA	DIM_DISPOSICION_B									
TIPO DE TABLA	DIMENSIONAL									
DESCRIPCIÓN	TABLA CON REGISTROS DE LAS DISPOSICIONES DEL SISTEMA BANNER DE LA UNIVERSIDAD DE LAS FUERZAS ARMADAS ESPE.									
ESQUEMA	DWESPE									
		DESTINO					ORIGEN			
Columna	Descripción	Tipo de dato	Clave	Fk para	Null	Esquema	Tabla	Campo	Tipo de dato	
sk_disposicion	Clave subrogada de la dimensión disposición.	Number(38)	pk	FAC_ROLES_SUELDO_B,	N	Sistema	Sistema	Sistema	Number(38)	
ddis_disp	Número de disposición.	Varchar2(2)			Y	STGESPE	szrtbphrhi	szrtbphrhi_disp	Varchar2(2)	
nombre_etl	Nombre de la transformación	Varchar2(2000)			Y	Sistema	Sistema	Sistema	Varchar2(2000)	
fecha_etl	Fecha en la que se realiza la transformación	Date			Y	Sistema	Sistema	Sistema	Date	
ip_equipo	Dirección ip del equipo	Varchar2(2000)			Y	Sistema	Sistema	Sistema	Varchar2(2000)	

Cuadro 83
Estructura dimensión tipo empleado banner

TABLA		DIM_TIPOEMPLE_B							
TIPO DE TABLA	DIMENSIONAL								
DESCRIPCIÓN	TABLA CON REGISTROS DE LOS TIPOS DE EMPLEADO DEL SISTEMA BANNER DE LA UNIVERSIDAD DE LAS FUERZAS ARMADAS ESPE.								
ESQUEMA	DWESPE								
	DESTINO					ORIGEN			
Columna	Descripción	Tipo de dato	Clave	Fk para	Null	Esquema	Tabla	Campo	Tipo de dato
sk_tipemp	Clave subrogada de la dimensión tipo empleado.	Number(38)	pk	FAC_ROLES_SUELDO_B,	N	Sistema	Sistema	Sistema	Number(38)
dtip_code	Código del tipo empleado.	Varchar2(3)			Y	STGESPE	szrtbtipem	szrtbtipem_code	Varchar2(3)
dtip_desc	Descripción del tipo empleado.	Varchar2(30)			Y	STGESPE	szrtbtipem	szrtbtipem_desc	Varchar2(30)
nombre_etl	Nombre de la transformación	Varchar2(2000)			Y	Sistema	Sistema	Sistema	Varchar2(2000)
fecha_etl	Fecha en la que se realiza la transformación	Date			Y	Sistema	Sistema	Sistema	Date
ip_equipo	Dirección ip del equipo	Varchar2(2000)			Y	Sistema	Sistema	Sistema	Varchar2(2000)

Cuadro 84
Estructura dimensión rubro banner

TABLA	DIM_RUBRO_B									
TIPO DE TABLA	DIMENSIONAL									
DESCRIPCIÓN	TABLA CON REGISTROS LOS RUBROS Y SUS TIPOS DEL SISTEMA BANNER DE LA UNIVERSIDAD DE LAS FUERZAS ARMADAS ESPE.									
ESQUEMA	DWESPE									
		DESTINO					ORIGEN			
Columna	Descripción	Tipo de dato	Clave	Fk para	Null	Esquema	Tabla	Campo	Tipo de dato	
sk_rubro	Clave subrogada de la dimensión rubro.	Number(38)	pk	FAC_ROLES_SUELDO_B,	N	Sistema	Sistema	Sistema	Number(38)	
drub_code_rubro	Código del rubro.	Varchar2(3)			Y	STGESPE	szrtbrub	szrtbrub_code_rubro	Varchar2(3)	
drub_desc_long_rubro	Descripción del rubro.	Varchar2(30)			Y	STGESPE	szrtbrub	szrtbrub_desc_long_rubro	Varchar2(30)	
drub_tipo_rubro	Tipo de rubro.	Varchar2(7)			Y	STGESPE	szrtbrub	szrtbrub_tipo_rubro	Varchar2(7)	
drub_pict_code	Código del tipo de empleado.	Varchar2(2)			Y	STGESPE	szrtbphrhi	szrtbphrhi_pict_code	Varchar2(2)	
drub_disp	Disposición.	Varchar2(2)			Y	STGESPE	szrtbphrhi	szrtbphrhi_disp	Varchar2(2)	
drub_pict_desc	Descripción del tipo de empleado,	Varchar2(30)			Y	STGESPE	szrtbptrpi	szrtbptrpi_pict_desc	Varchar2(30)	
nombre_etl	Nombre de la transformación	Varchar2(2000)			Y	Sistema	Sistema	Sistema	Varchar2(2000)	
fecha_etl	Fecha en la que se realiza la transformación	Date			Y	Sistema	Sistema	Sistema	Date	
ip_equipo	Dirección ip del equipo	Varchar2(2000)			Y	Sistema	Sistema	Sistema	Varchar2(2000)	

Cuadro 85
Estructura tabla de hechos roles de pago banner

TABLA	FAC_ROLES_SUELDO_B								
TIPO DE TABLA	HECHOS								
DESCRIPCIÓN	TABLA DE HECHOS CON LOS SUELDOS DE LOS EMPLEADOS DE LA UNIVERSIDAD DE LAS FUERZAS ARMADAS ESPE DEL SISTEMA BANNER.								
ESQUEMA	DWESPE								
	DESTINO					ORIGEN			
Columna	Descripción	Tipo de dato	Clave	Fk para	Null	Esquema	Tabla	Campo	Tipo de dato
sk_rols	Clave subrogada de la tabla de los sueldos de roles banner.	Number(38)	pk		N	Sistema	Sistema	Sistema	Number(38)
sk_emple_b	Clave subrogada de la dimensión empleado.	Number(38)	Fk		N	DWESPE	dim_emple_b	sk_emple_b	Number(38)
sk_rubro	Clave subrogada de la dimensión rubro.	Number(38)	Fk		N	DWESPE	dim_rubro_b	sk_rubro	Number(38)
sk_fecha	Clave subrogada de la dimensión tiempo.	Number(38)	Fk		N	DWESPE	dim_tiempo	sk_fecha	Number(38)
frol_valor	Medida que contiene los sueldos.	Number(12,2)			Y	STGESPE	szrtbsueld	szrtbsueld_valor	Number(12,2)
nombre_etl	Nombre de la transformación	Varchar2(2000)			Y	Sistema	Sistema	Sistema	Varchar2(2000)
fecha_etl	Fecha en la que se realiza la transformación	Date			Y	Sistema	Sistema	Sistema	Date
ip_equipo	Dirección ip del equipo	Varchar2(2000)			Y	Sistema	Sistema	Sistema	Varchar2(2000)

4.6.2.30 Diagrama general para el módulo de cuentas por cobrar.

El diagrama muestra las dimensiones y la relación que tienen con las tablas de la base de datos de origen.

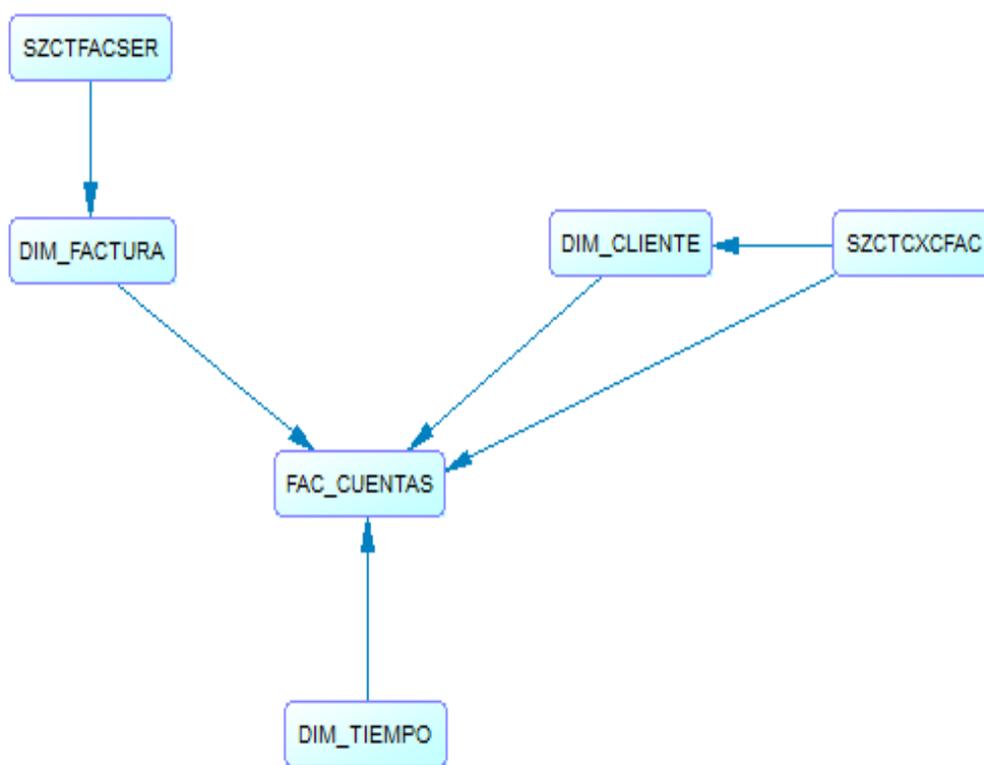


Figura 79: Diagrama general cuentas por cobrar

4.6.2.31 Mapeo de datos cuentas por cobrar.

En la siguiente tabla se describe el origen de datos desde el servidor de base de datos o archivos planos hacia las dimensiones y tablas de hechos del datamart.

Cuadro 86
Mapeo de datos cuentas por cobrar

FUENTE DE DATOS ORIGINAL	DESCRIPCIÓN	STAGE	DIMENSIÓN (DIM)	HECHOS (FAC)
cc_factura cc_venfac tecobros tecobpag fc_cabnot	Tabla que contiene las cédulas y los nombres de los clientes.	szctcxfac	DIM_CLIENTE	
cc_det_ser	Tabla que contiene los números de facturas con su detalle.	szctfacser	DIM_FACTURA	
cc_factura cc_venfac tecobros tecobpag fc_cabnot	Tabla que contiene la medida de saldo para cuentas.	szctcxfac		FAC_CUENTAS

4.6.2.32 Modelo dimensional de cuentas por cobrar

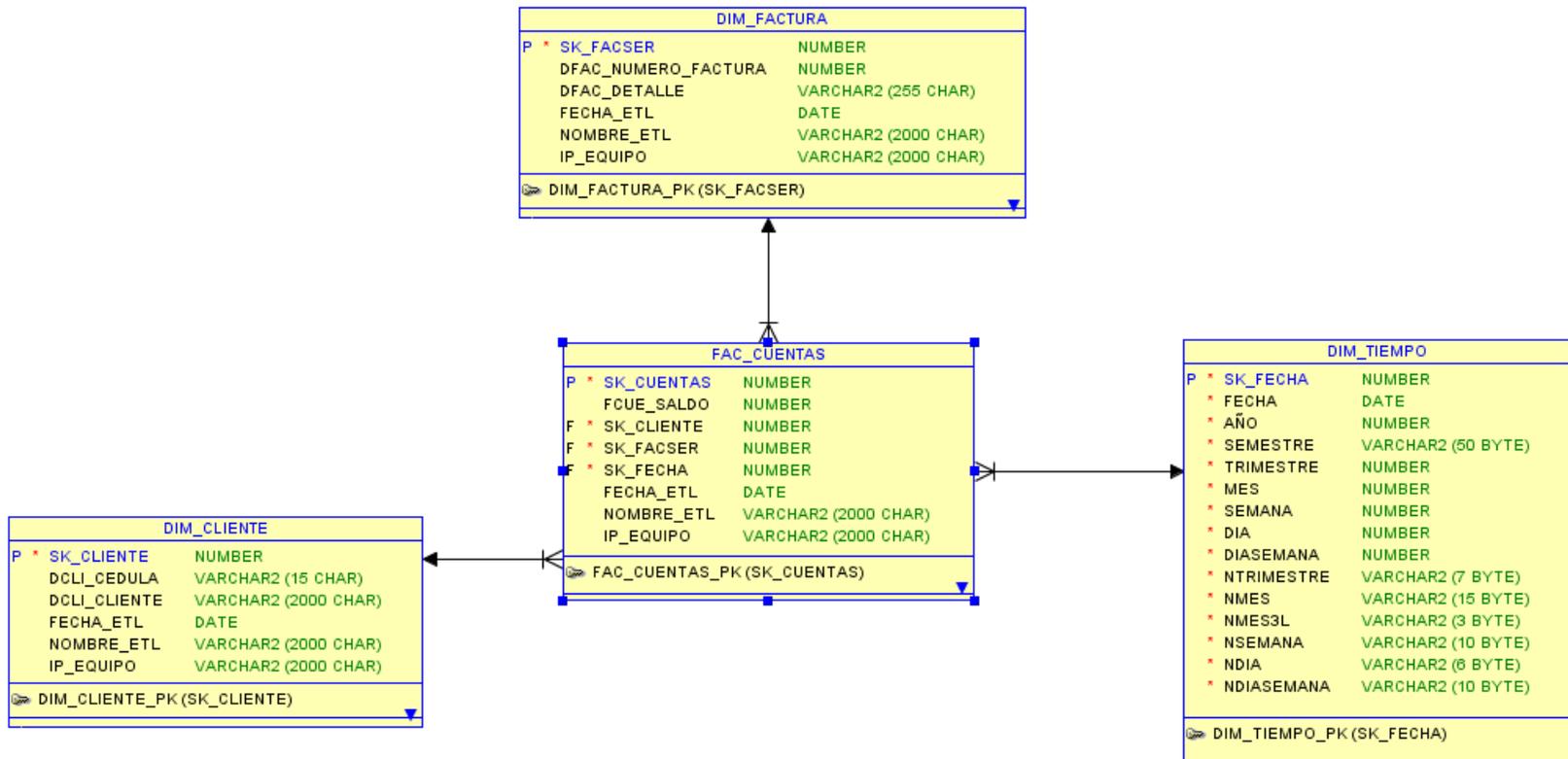


Figura 80: Modelo dimensional cuentas por cobrar

4.6.2.33 Mapeo de datos para dimensiones y hechos de cuentas por cobrar.

Las siguientes tablas muestran el destino y origen de los datos para dimensiones y tablas de hechos del módulo de cuentas por cobrar, se detalla nombres de los campos de destino y origen, los tipos de datos y sus esquemas respectivos.

Cuadro 87
Estructura dimensión cliente

TABLA		DIM_CLIENTE							
TIPO DE TABLA	DIMENSIONAL								
DESCRIPCIÓN	TABLA CON REGISTROS DE LOS CLIENTES DE CUENTAS POR COBRAR								
ESQUEMA	DWESPE								
		DESTINO				ORIGEN			
Columna	Descripción	Tipo de dato	Clave	Fk para	Null	Esquema	Tabla	Campo	Tipo de dato
sk_cliente	Clave subrogada de la dimensión cliente.	Number(38)	pk	FAC_CUENTAS	N	Sistema	Sistema	Sistema	Number(38)
dcli_cedula	Cedula o ruc del cliente.	Varchar2(15)			Y	STGESPE	szctxcfac	szctxcfac_cedula	Varchar2(15)
dcli_cliente	Nombre del cliente.	Varchar2(2000)			Y	STGESPE	szctxcfac	szctxcfac_cliente	Varchar2(2000)
nombre_etl	Nombre de la transformación	Varchar2(2000)			Y	Sistema	Sistema	Sistema	Varchar2(2000)
fecha_etl	Fecha en la que se realiza la transformación	Date			Y	Sistema	Sistema	Sistema	Date
ip_equipo	Dirección ip del equipo	Varchar2(2000)			Y	Sistema	Sistema	Sistema	Varchar2(2000)

Cuadro 88
Estructura dimensión factura

TABLA	DIM_FACTURA								
TIPO DE TABLA	DIMENSIONAL								
DESCRIPCIÓN	TABLA CON REGISTROS DE LAS FACTURAS CON SU DETALLE DE CUENTAS POR COBRAR								
ESQUEMA	DWESPE								
Columna	Descripción	DESTINO	Clave	Fk para	Null	Esquema	Tabla	ORIGEN	Tipo de dato
		Tipo de dato						Campo	
sk_facser	Clave subrogada de la dimensión factura.	Number(38)	pk	FAC_CUENTAS	N	Sistema	Sistema	Sistema	Number(38)
dfac_numero_factura	Número de la factura.	Number			Y	STGESPE	szctfacser	szctfacser_numero_factura	Number
dfac_detalle	Detalle de la factura.	Varchar2(255)			Y	STGESPE	szctfacser	szctfacser_detalle	Varchar2(255)
nombre_etl	Nombre de la transformación	Varchar2(2000)			Y	Sistema	Sistema	Sistema	Varchar2(2000)
fecha_etl	Fecha en la que se realiza la transformación	Date			Y	Sistema	Sistema	Sistema	Date
ip_equipo	Dirección ip del equipo	Varchar2(2000)			Y	Sistema	Sistema	Sistema	Varchar2(2000)

Cuadro 89
Estructura tabla de hechos cuentas por cobrar

TABLA		FAC_CUENTAS							
TIPO DE TABLA		HECHOS							
DESCRIPCIÓN		TABLA DE HECHOS PARA EL DATAMART DE CUENTAS POR COBRAR							
ESQUEMA		DWESPE							
		DESTINO				ORIGEN			
Columna	Descripción	Tipo de dato	Clave	Fk para	Null	Esquema	Tabla	Campo	Tipo de dato
sk_cuentas	Clave subrogada de la tabla de hechos de cuentas por cobrar	Number(38)	pk		N	sistema	sistema	sistema	Number(38)
sk_cliente	Clave subrogada de la dimensión cliente	Number(38)	fk		N	DWESPE	dim_cliente	sk_cliente	Number(38)
sk_facser	Clave subrogada de la dimensión factura	Number(38)	fk		N	DWESPE	dim_factura	sk_facser	Number(38)
sk_fecha	Clave subrogada de la dimensión tiempo	Number(38)	fk		N	DWESPE	dim_tiempo	sk_fecha	Number(38)
fcue_saldo	Medida del saldo	Number(38)			Y	STGESPE	szctxcfac	szctxcfac_saldo	Number(38)
nombre_etl	Nombre de la transformación	Varchar2(2000)			Y	Sistema	Sistema	Sistema	Varchar2(2000)
fecha_etl	Fecha en la que se realiza la transformación	Date			Y	Sistema	Sistema	Sistema	Date
ip_equipo	Dirección ip del equipo	Varchar2(2000)			Y	Sistema	Sistema	Sistema	Varchar2(2000)

4.7 MANUAL DE USUARIO SPOON

4.7.1 Proceso de extracción transformación y carga

El proceso ETL se lo realizó con la herramienta spoon que proviene del pentaho data integration, la cual permitirá en primer paso realizar una extracción de la información de fuentes de datos OLTP, seguido de una transformación de los datos y por ultimo cargar la nueva información en una base de datos destino de tipo Data Warehouse.

Descripción de la interfaz de usuario.

Primero se inicia el spoon con unos archivos que se encuentran en la carpeta que se descarga en el internet.

Para Windows se ejecuta el archivo spoon.bat para iniciar el spoon mientras que para Linux se ejecuta en un terminal el archivo spoon.sh, con esto carga y se abre la aplicación.

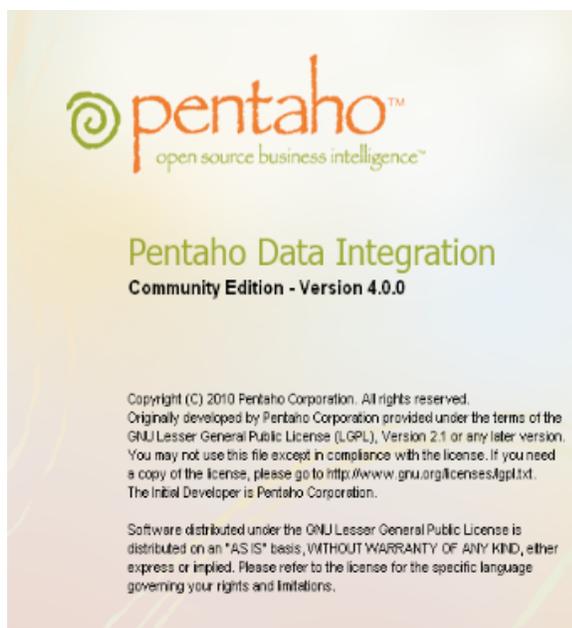


Figura 81: Carga de Spoon



Figura 82: Pantalla de entrada de spoon

En la siguiente imagen se puede observar que se puede crear varias cosas con la herramienta spoon para lo cual se va a explicar la creación de una transformación, se hace clic en transformación.

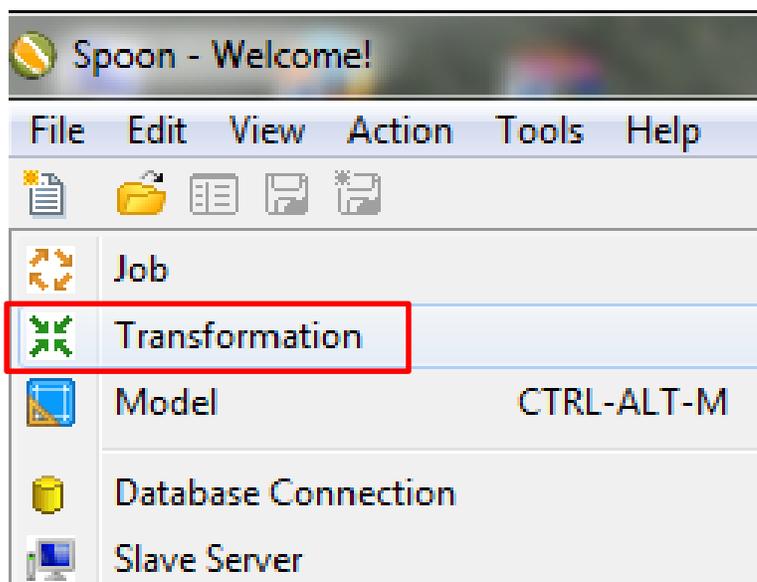


Figura 83: Pantalla de menú spoon

CREAR TRANSFORMACIÓN

Al momento que se crea una transformación, se abre una pestaña y todos los elementos que se pueden utilizar para realizar un proceso ETL como se puede ver en la siguiente imagen.

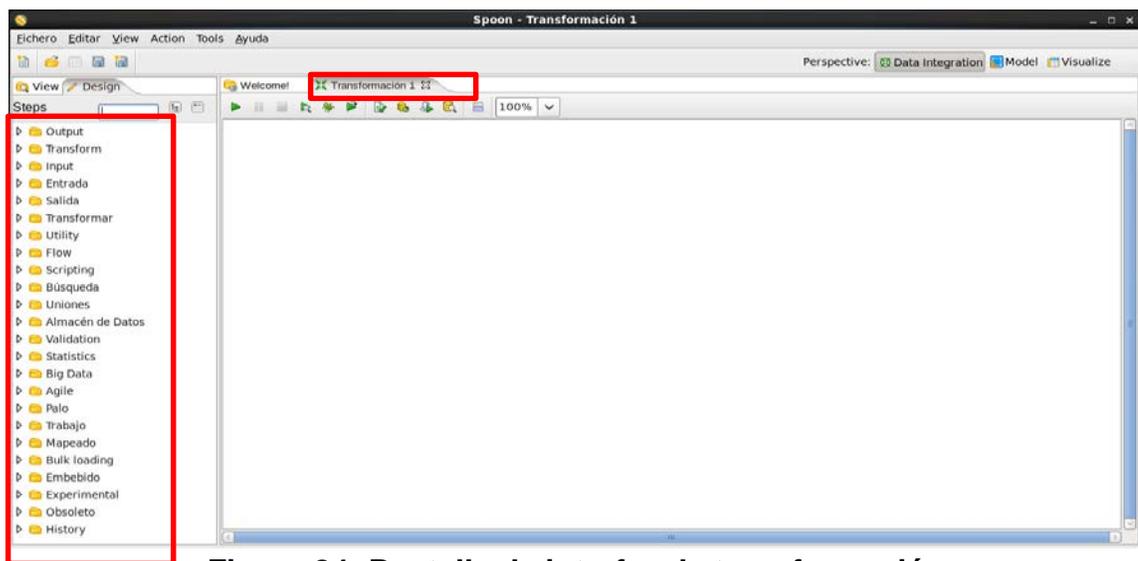


Figura 84: Pantalla de interfaz de transformación

4.7.2 Diseño del etl para el caso de nuestro proyecto

Para el caso de nuestro aplicativo se va a realizar un proceso ETL para la creación de una dimensión denominada dim_rubro y las demás dimensiones se las realiza de una forma muy similar.

4.7.2.1 Entrada de tabla

Como primer paso se abre la carpeta entrada que se encuentra en la pestaña design, luego se selecciona el elemento entrada de tabla que se encuentra en la carpeta entrada y se la arrastra a la parte derecha como se la puede ver en la siguiente imagen.

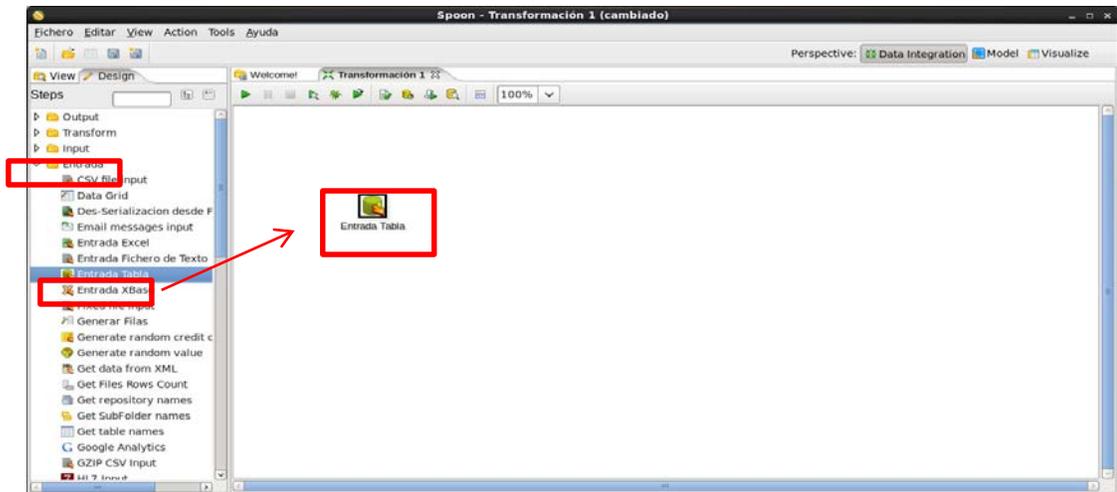


Figura 85: Pantalla de elemento entrada de tabla

Al hacer doble clic sobre el elemento “entrada de tabla”, se abrirá una pantalla como la siguiente donde se coloca un nombre al paso, se realiza la conexión a una base de datos.

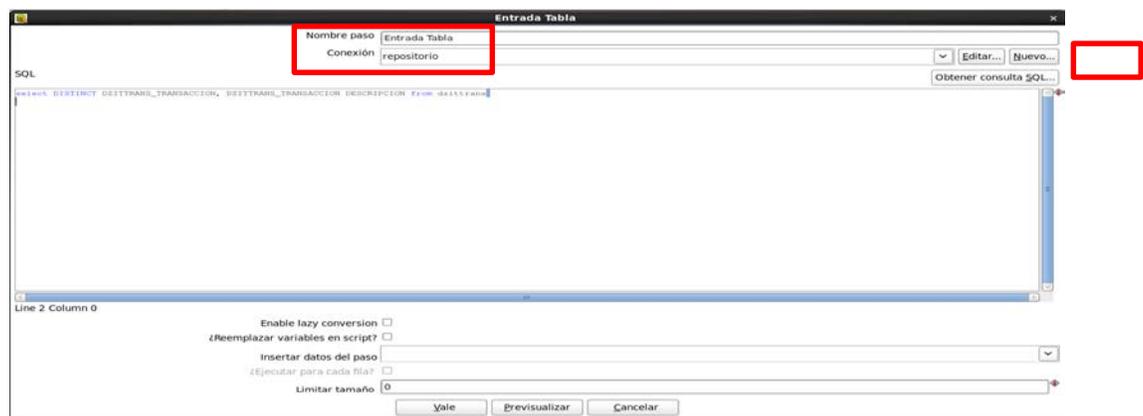


Figura 86: Pantalla de entrada de tabla

Al hacer clic sobre el botón “nuevo” en la imagen anterior se abre una ventana en la cual se realiza la conexión a una base de datos de donde se va a extraer la información como se muestra en la siguiente imagen.

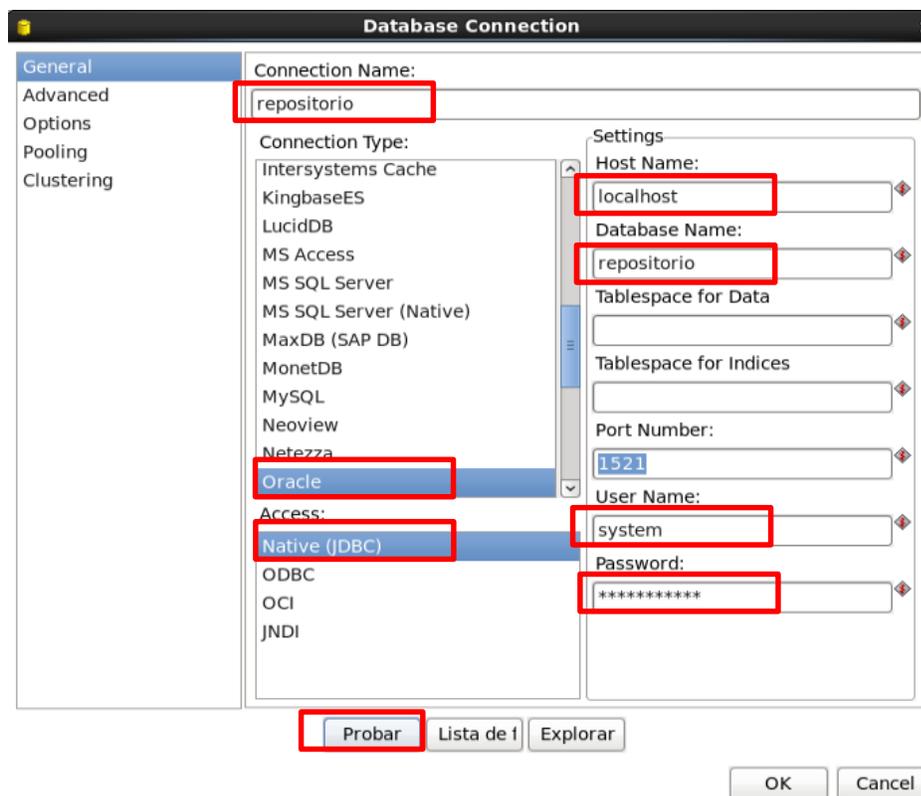


Figura 87: Pantalla para crear conexión a una bd

Como se pudo ver en la imagen anterior se pone un nombre cualquiera al nombre de la conexión, se escoge el tipo de conexión y acceso se llena los campos que sean necesarios para realizar una conexión a la base satisfactoriamente y para esto se comprueba haciendo clic en el botón probar y si la conexión es correcta aparecerá la una ventana como la siguiente imagen.

Si todo es correcto se hace clic en el botón vale y seguido de esto se hace clic en el botón ok para salir de las pantallas de conexión de base de datos.



Figura 88: Pantalla para probar conexión de bd

Una vez que se tiene la conexión, se coloca la sentencia sql que el usuario necesite para obtener la información como se puede ver en la siguiente imagen.

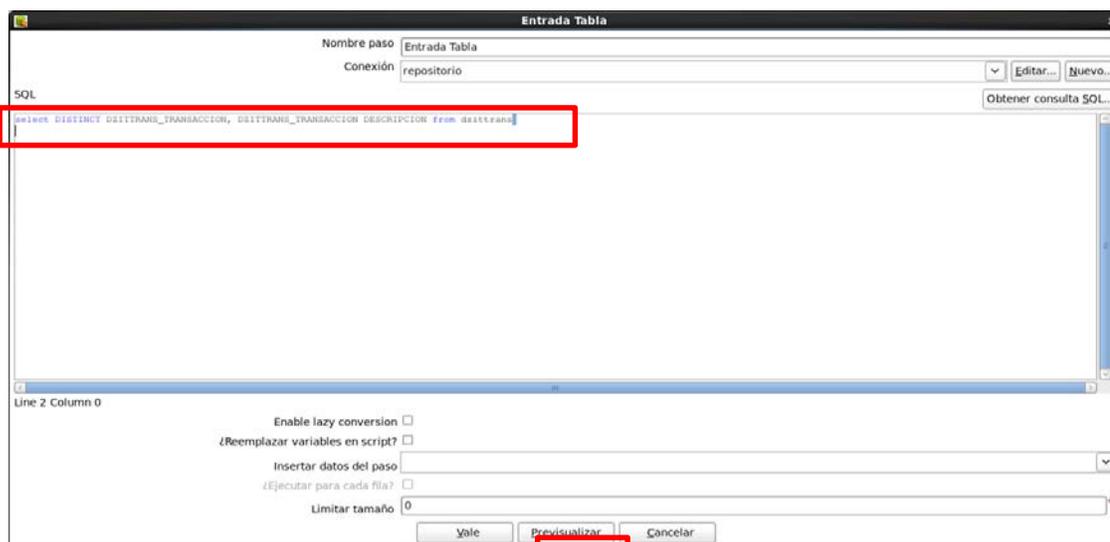


Figura 89: Pantalla de entrada de tabla

Luego de tener la sentencia sql, se hace clic en el botón previsualizar, esto se lo realiza únicamente para comprobar que el query se ejecute correctamente y se muestre la información necesaria como se muestra en la siguiente imagen.

Luego de comprobar que todo esté bien, se hace clic en el botón cerrar y después hacer clic en el botón vale y con esto estaría listo nuestra entrada de tabla de nuestro caso práctico.

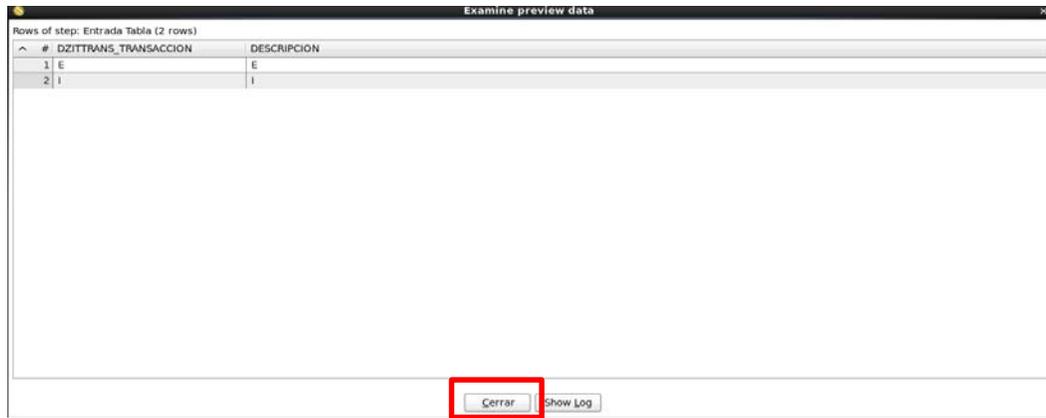


Figura 90: Pantalla para previsualizar

4.7.2.2 Selecciona / renombra valores

De la misma forma como se arrastró el elemento anterior, ahora se hará el mismo procedimiento pero escogiendo el elemento llamado “Selección/Renombrar valores” que se encuentra en la pestaña design, en la carpeta transformar.

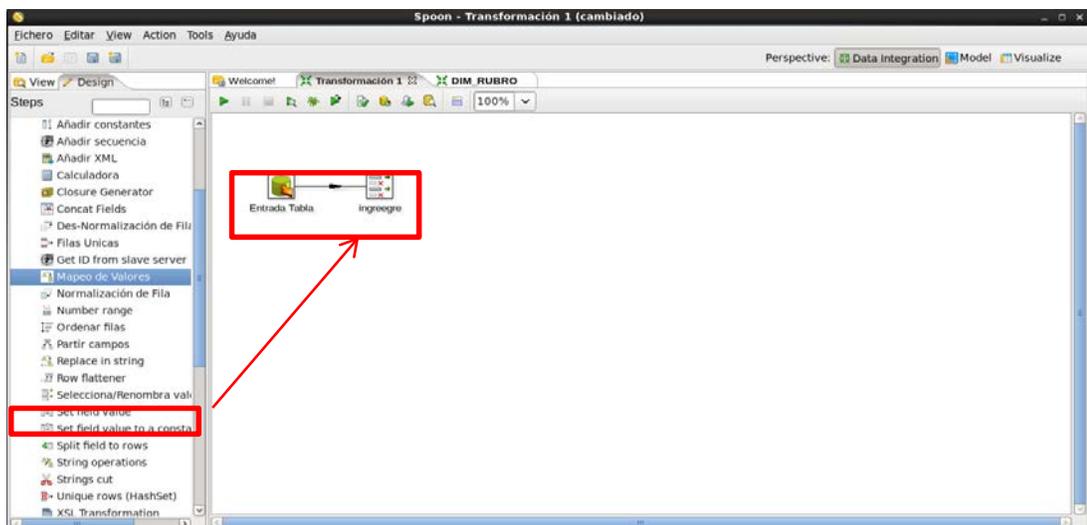


Figura 91: Pantalla spoon con dos elementos

Para realizar la unión de los dos elementos se debe colocar con el puntero sobre la “entrada de tabla” y aparecerán unas opciones que se muestran en la parte inferior del elemento como se podrá ver en la siguiente imagen. Luego se coloca con el puntero sobre la opción que representa una hoja con una flecha hacia la derecha y hacer clic sin soltarlo y se arrastra hacia el elemento que en nuestro caso se llama “ingreegre” y con esto se une dos elementos como se puede ver en la imagen anterior, esta representación gráfica o línea que une a dos pasos se lo denomina como **SALTO**.



Figura 92: Proceso de unión entre dos pasos

Al hacer doble clic sobre el nuevo elemento se podrá colocar un nombre al paso como se puede ver en la siguiente imagen.

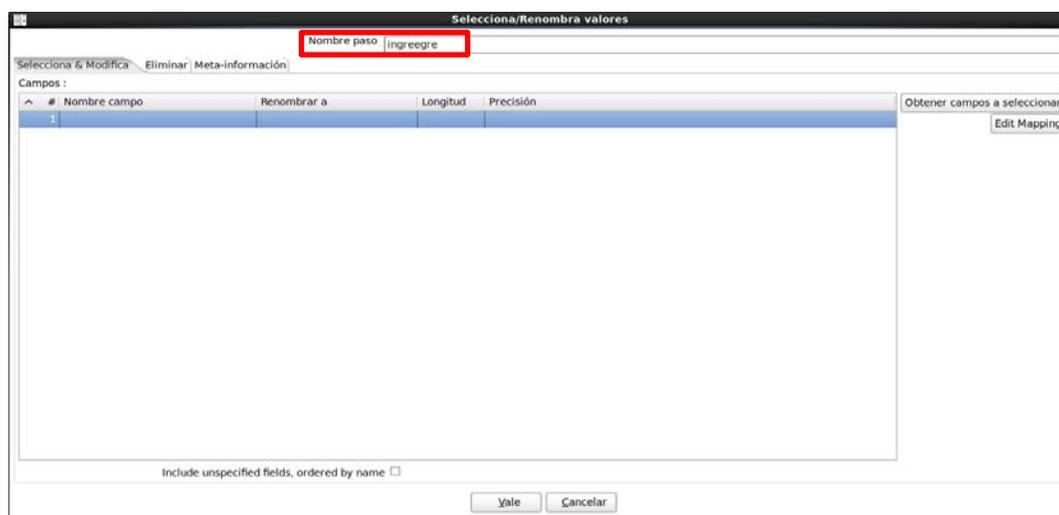


Figura 93: Pantalla 'selección/renombra valores'

Ahora se hace clic en el botón “**obtener campos a seleccionar**” para que se muestren todos los campos que se está trayendo desde la entrada de tabla y estos campos se mostraran en la columna que se llama “**nombre campo**” y el usuario puede renombrar los campos en la columna “**Renombrar a**” y hacer clic en vale si todo esta correcto como se puede ver en la siguiente imagen.

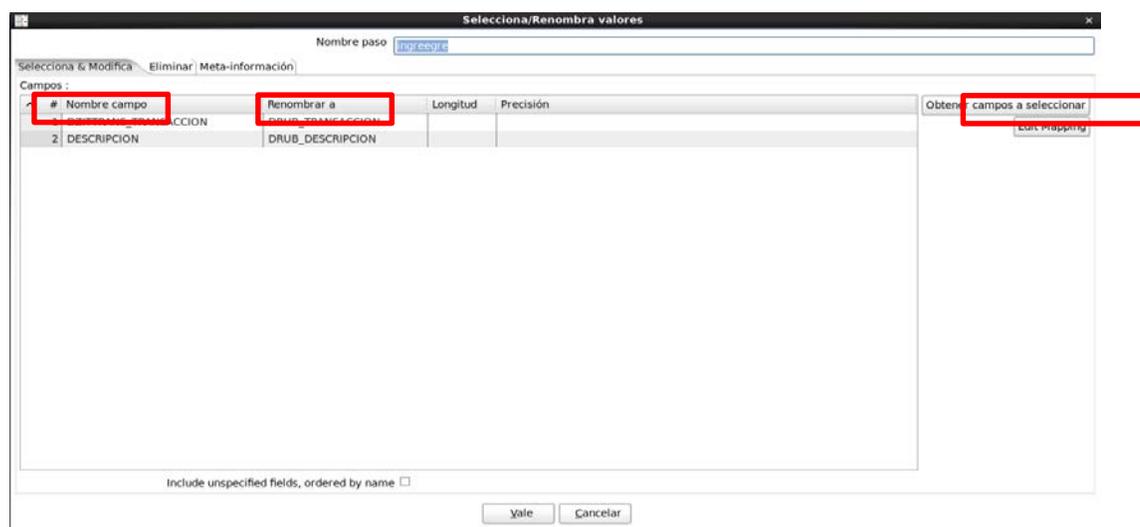


Figura 94: Pantalla 'selección/renombra valores' con valores

4.7.2.3 Mapeo de valores

Luego se realiza el mismo procedimiento de arrastrar otro elemento llamado “mapeo de valores” que se encuentra en la pestaña design, en la carpeta transformar y después se realiza el proceso de la unión de dos pasos explicado anteriormente entre el elemento llamado **ingeegre** y el mapeo de valores como se puede ver en la siguiente imagen.

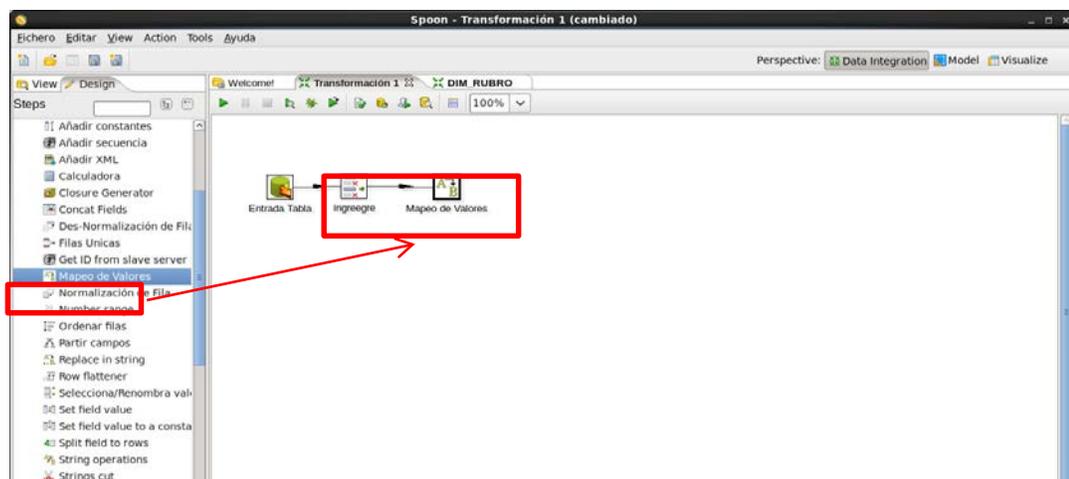


Figura 95: Pantalla con mapeo de valores

Al hacer doble clic sobre el mapeo de valores se pone un nombre a este paso, se escoge el campo al cual se le va a realizar un cambio a los valores y por último en la columna “**valor origen**” se coloca todos los valores que vienen de la base de datos y que sean necesarios hacerles un cambio. En la columna “**valor destino**” se escribe los nuevos valores que se obtendrán como se puede ver en la siguiente imagen y finalmente hacer clic en vale si todo está correcto.

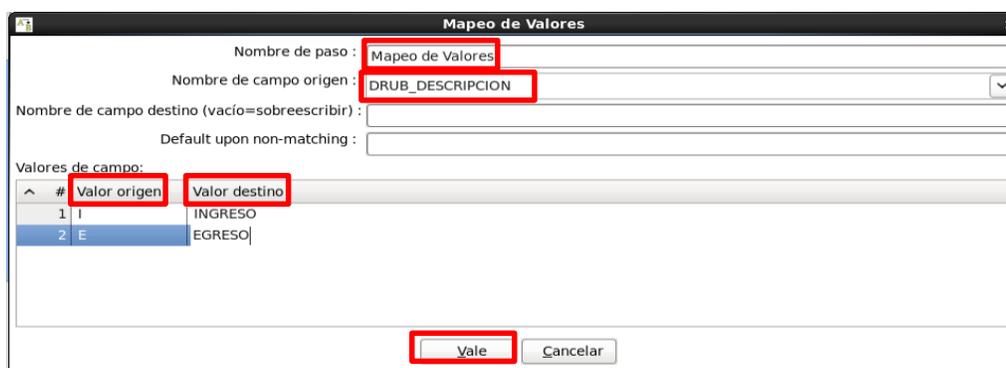


Figura 96: Mapeo de valores con valores

4.7.2.4 Añadir secuencia

Ahora se realiza el mismo procedimiento de arrastrar otro elemento llamado “añadir secuencia” que se encuentra en la pestaña design, en la carpeta transformar y después se realiza la unión de dos pasos entre el elemento mapeo de valores con añadir secuencia como se puede ver en la siguiente imagen.

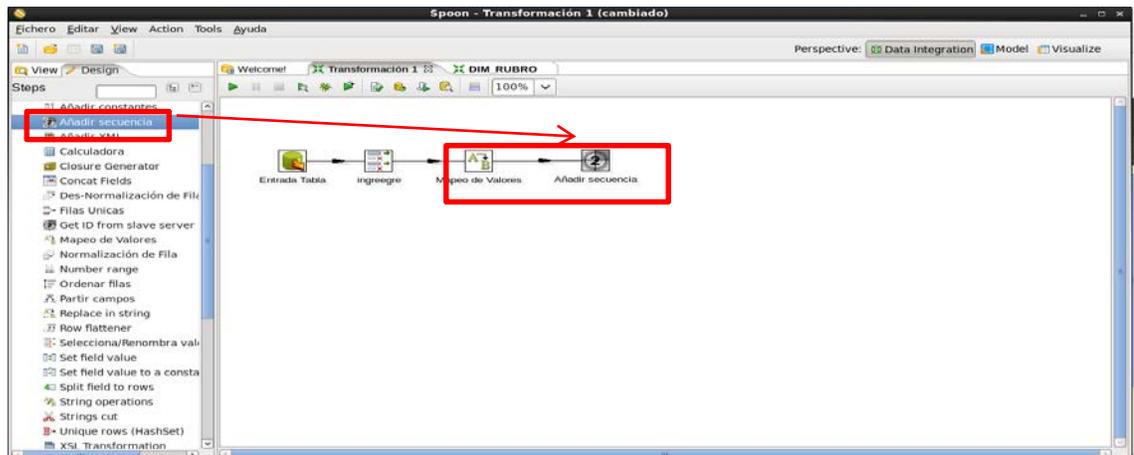


Figura 97: Pantalla con secuencia 1

Se hace doble clic en el elemento añadir secuencia y se pone un nombre al paso, el nombre de valor y opcionalmente colocar un nombre al contador en la parte posterior. Se verifica que todo este correcto y se hace clic en vale como se observa en la siguiente imagen.

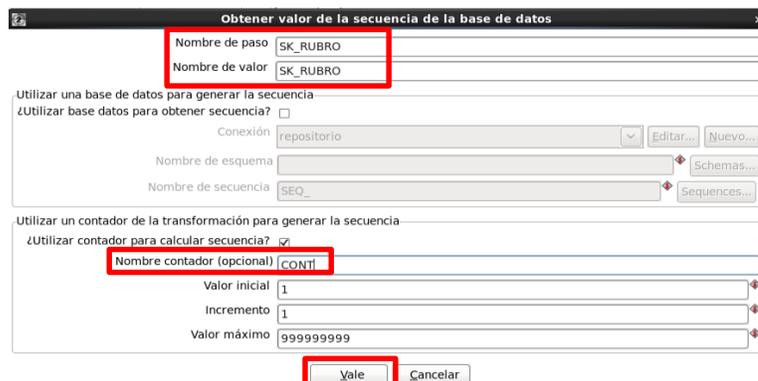


Figura 98: Añadir secuencia 1

4.7.2.5 Información del sistema

Ahora se realiza el mismo procedimiento de arrastrar otro elemento llamado “información de sistema” que se encuentra en la pestaña design, en la carpeta entrada y realizar la unión de dos pasos entre el elemento añadir secuencia con información del sistema como se puede ver en la siguiente imagen.

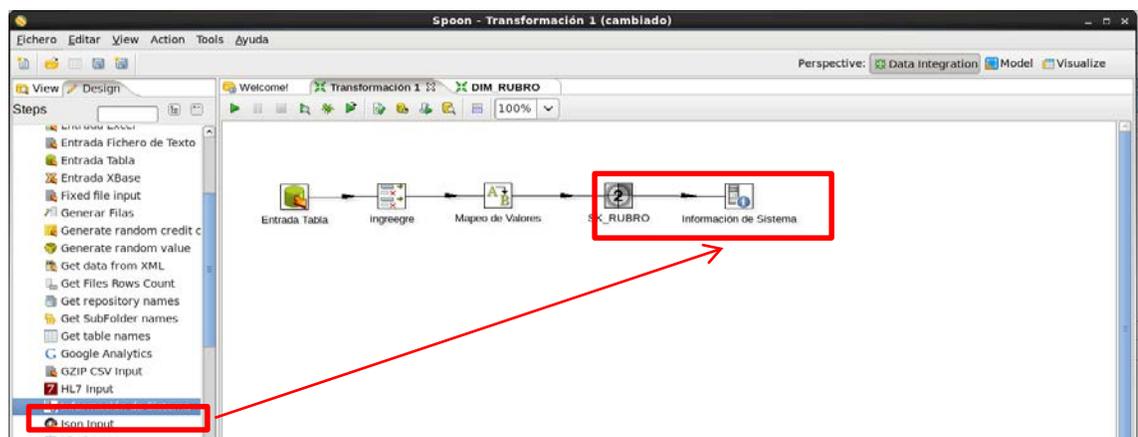


Figura 99: Pantalla con información sistema

Se hace doble clic en el elemento información del sistema, se pone un nombre al paso, en la columna “**nombre**” se van a colocar tres nombres los cuales van a ser: FECHA_ETL, NOMBRE_ETL, IP_EQUIPO.

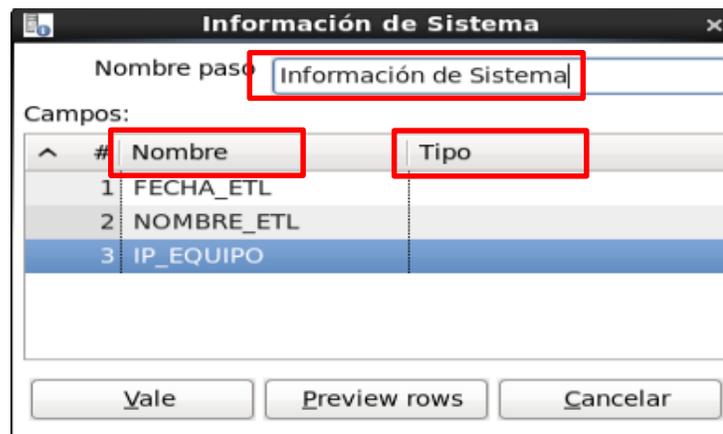


Figura 100: Selección información del sistema

Luego de tener los tres nombres al hacer clic en la columna tipo de la imagen anterior, se abrirá una ventana como se muestra en la siguiente imagen y se debe seleccionar los tipos de información del sistema.

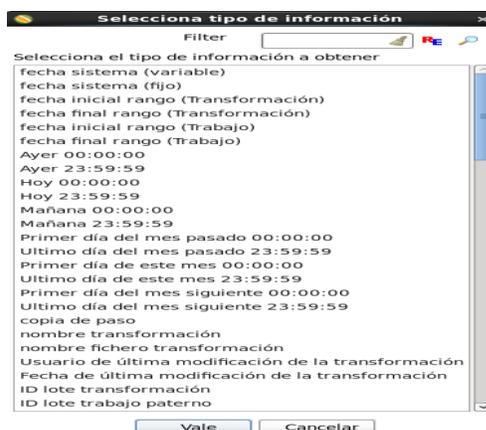


Figura 101: Selección tipo información del sistema

A continuación se muestran los tres tipos que se seleccionaron para cada nombre y luego de verificar que todo esté correcto se hace clic en vale.

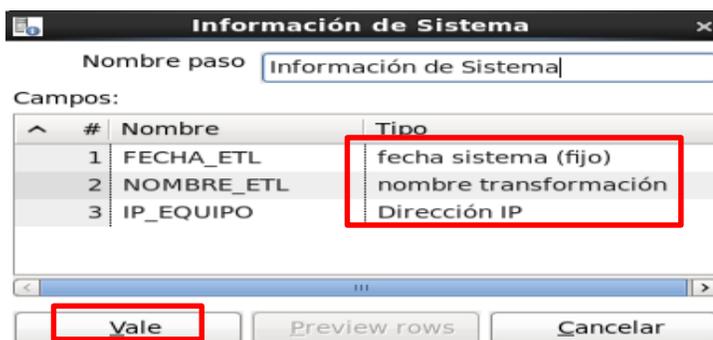


Figura 102: Selección información del sistema

4.7.2.6 Insertar / actualizar

Ahora se realiza el mismo procedimiento de arrastrar otro elemento llamado “insertar / actualizar” que se encuentra en la pestaña design, en la carpeta salida y realizar la unión de dos pasos entre el elemento información del sistema y el elemento insertar / actualizar como se puede ver en la siguiente imagen.

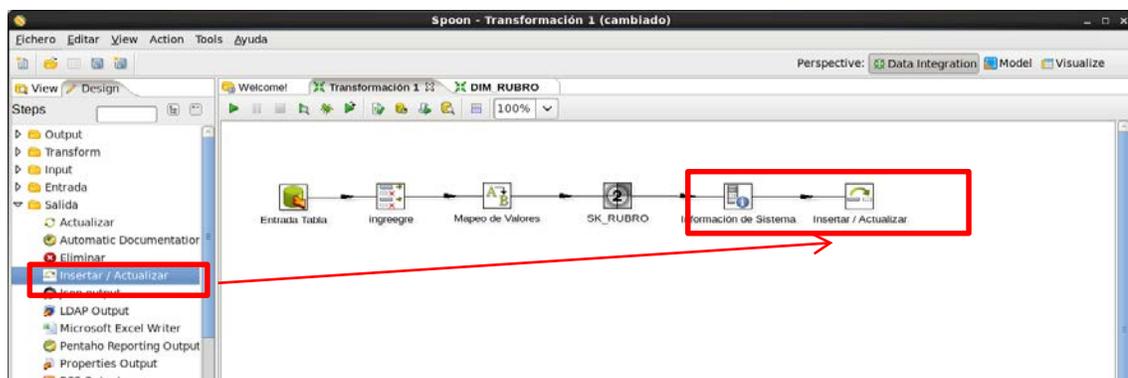


Figura 103: Pantalla con insertar/actualizar

Ahora al hacer doble clic en el botón insertar /actualizar se abrirá otra ventana en donde se pone primeramente el nombre del paso, luego se hace clic en el botón nuevo para crear una nueva conexión donde se va almacenar la nueva información.

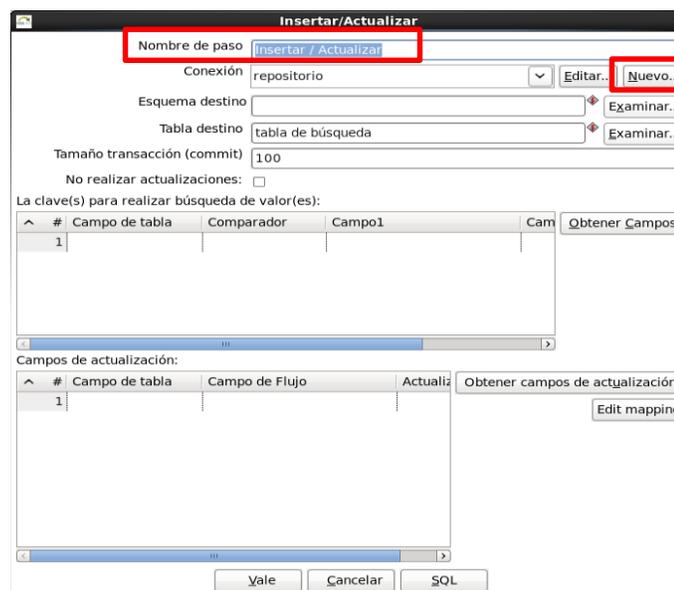


Figura 104: Interfaz insertar / actualizar

La nueva conexión es procedimiento que fue explicado al comienzo del manual, ahora se coloca un nombre cualquiera a la tabla destino, se hace clic en el botón “**obtener campos**” y se deja solo el campo sk_rubro que es el campo creado en el paso de añadir secuencia y en la parte inferior a eso, de igual forma se hace clic en “**obtener**

campos de actualización” y se pondrán todos los campos como se puede ver en la siguiente imagen.

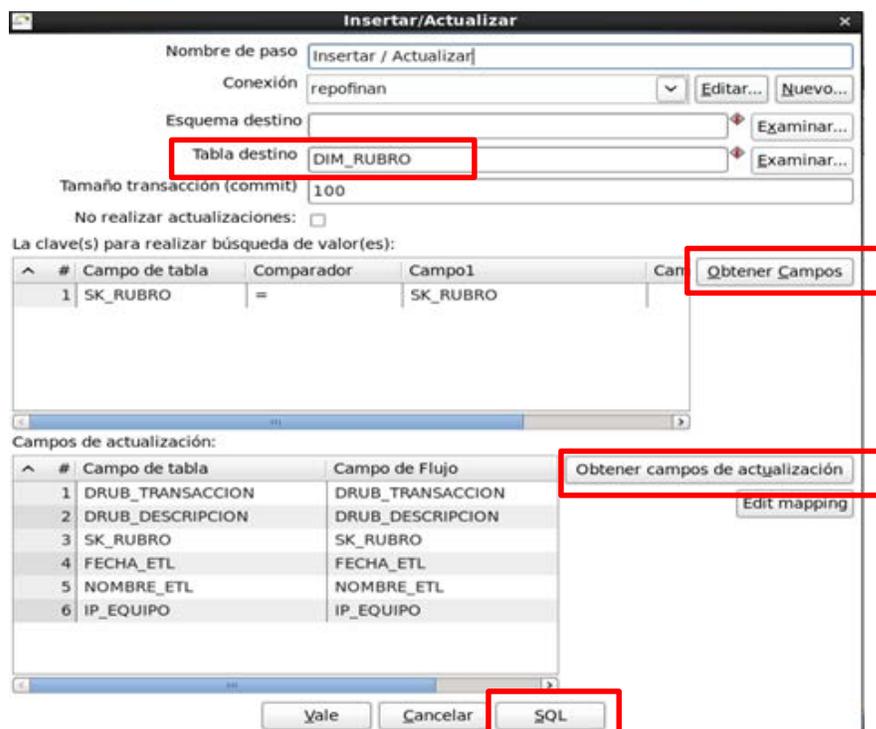


Figura 105: Campos 'insertar / actualizar'

Una vez que todo este correcto se hace clic en el botón sql y se abrirá una ventana con una sentencia sql en la cual se crea la tabla como se muestra en la siguiente imagen.

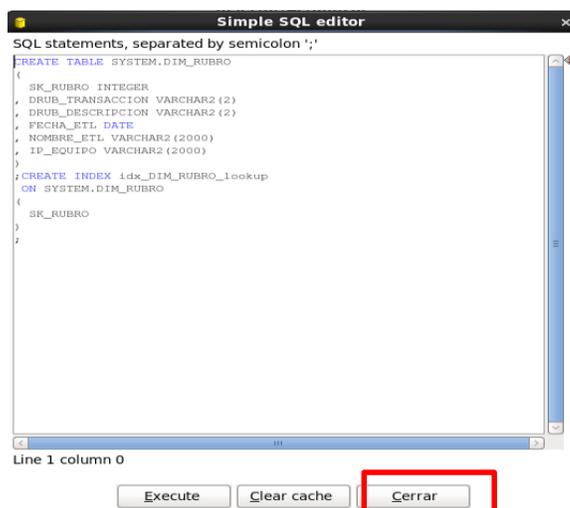


Figura 106: Ejecución script de tabla

Se verifica que la sentencia esta creada como se desea y luego se hace clic en el botón “**execute**” para que se ejecute el script y se cree la tabla como se muestra en la siguiente imagen. Después se hace clic en vale, seguido de eso se hace clic en cerrar y por último se hace clic en vale para cerrar todas las ventanas.

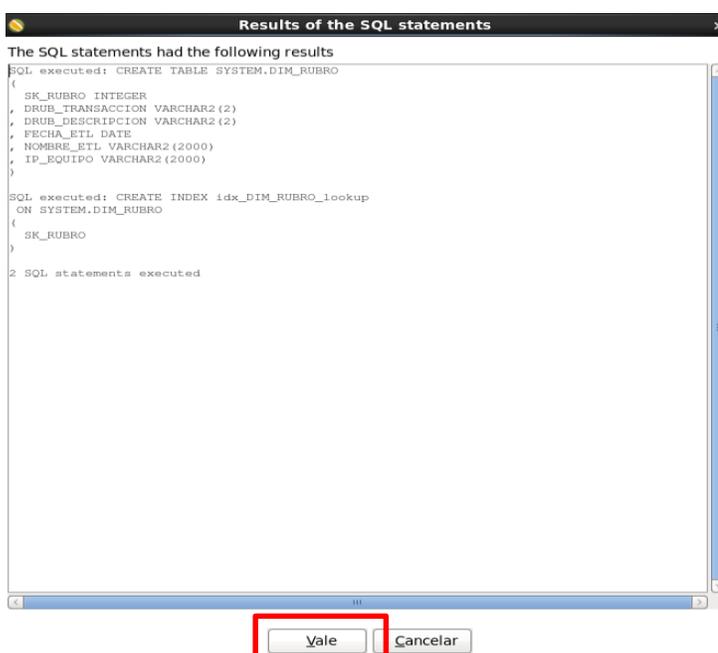


Figura 107: Resultado de tabla creada

4.7.2.7 Añadir secuencia de auditoria

De aquí los siguientes pasos son creados para realizar una auditoría de la nueva información creada.

Primero se arrastra otro elemento para añadir una secuencia que se lo puede encontrar en la pestaña design, en la carpeta transformar, luego se realiza la unión de los pasos insertar/actualizar y la nueva secuencia. Esta unión es diferente porque al momento de arrastrar la línea para unirlos a los dos pasos, aparecerán dos opciones:

1. Main output of step
2. Error handling of step

Se escoge la segunda opción y la línea se pondrá de color rojo y con una x como se puede ver a continuación.

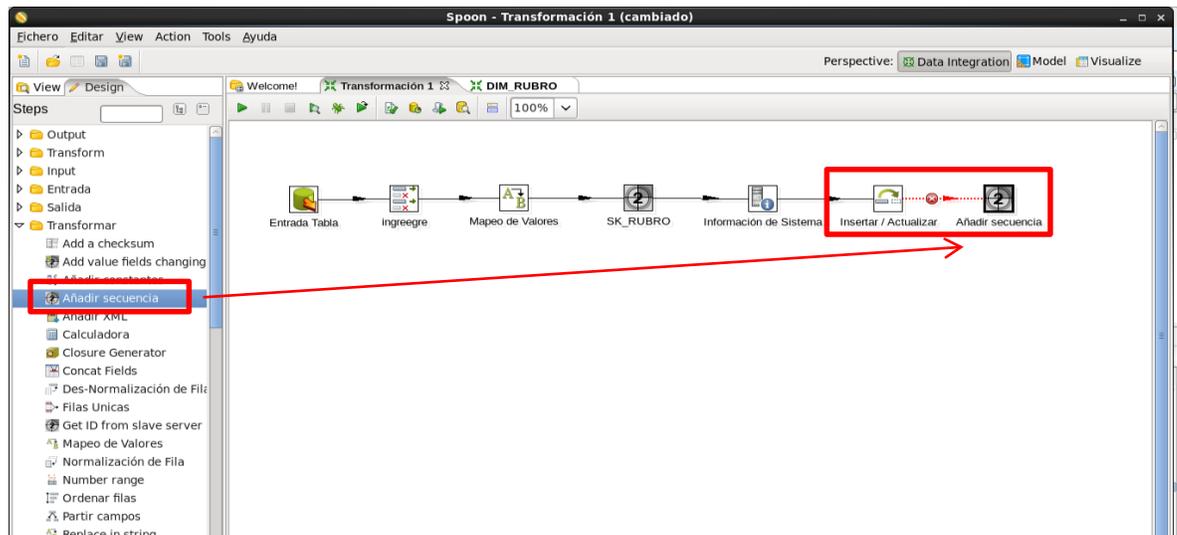


Figura 108: Pantalla con paso secuencia 2

Ahora se hace doble clic en el elemento de añadir secuencia y de la misma forma como se hizo anteriormente, se pone un nombre al paso, un nombre al valor y opcionalmente un nombre al contador como se puede ver en la siguiente imagen.

Obtener valor de la secuencia de la base de datos

Nombre de paso:

Nombre de valor:

Utilizar una base de datos para generar la secuencia

¿Utilizar base datos para obtener secuencia?

Conexión:

Nombre de esquema:

Nombre de secuencia:

Utilizar un contador de la transformación para generar la secuencia

¿Utilizar contador para calcular secuencia?

Nombre contador (opcional):

Valor inicial:

Incremento:

Valor máximo:

Figura 109: Añadir secuencia 2

Ahora se va hacer clic en la x del siguiente símbolo  , esto abre una ventana y se llena los cuatro campos como se muestra en la siguiente imagen y luego de eso hacer clic en ok.



Figura 110: Paso de error

4.7.2.8 Salida de tabla de auditoria

Como último proceso se va a explicar la salida de tabla de auditoria donde va a cargar información únicamente si el proceso de ETL tiene alguna falla al cargar la tabla de la dimensión.

Primero se arrastra el elemento para añadir la salida de tabla que se lo puede encontrar en la pestaña design, en la carpeta salida. Luego se realiza la unión de los pasos añadir secuencia y la salida de tabla.

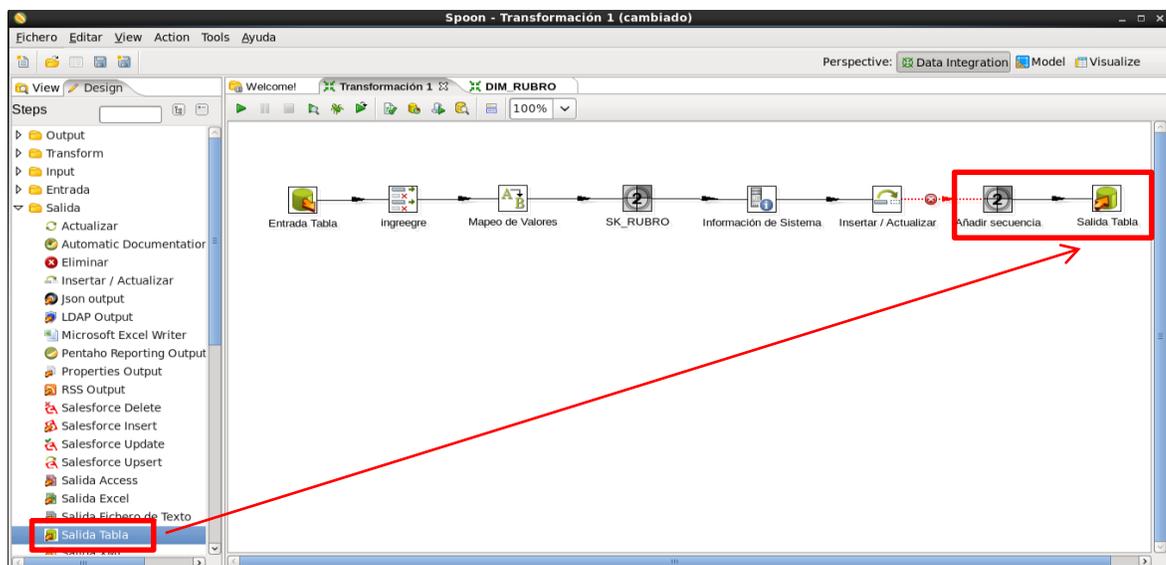


Figura 111: Crear salida de tabla

Al hacer doble clic en salida de tabla, aparece una nueva ventana donde se coloca un nombre al paso, se escoge una conexión se pone un nombre a la tabla destino y después se hace clic en la opción SQL.

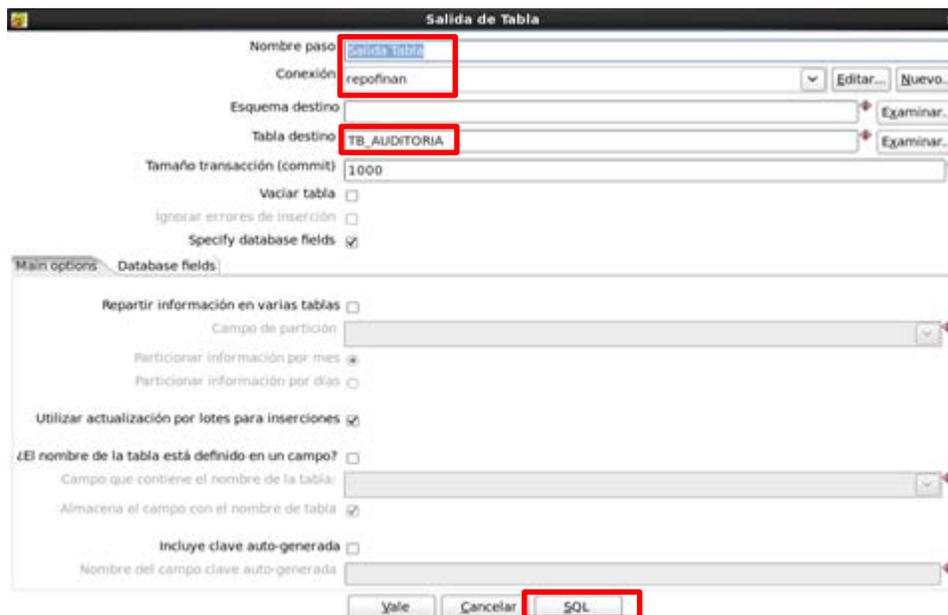


Figura 112: Datos de la salida de tabla

En la nueva ventana que aparece solo se debe hacer clic en la opción “**execute**” para que se ejecute el script y se genere la nueva tabla.

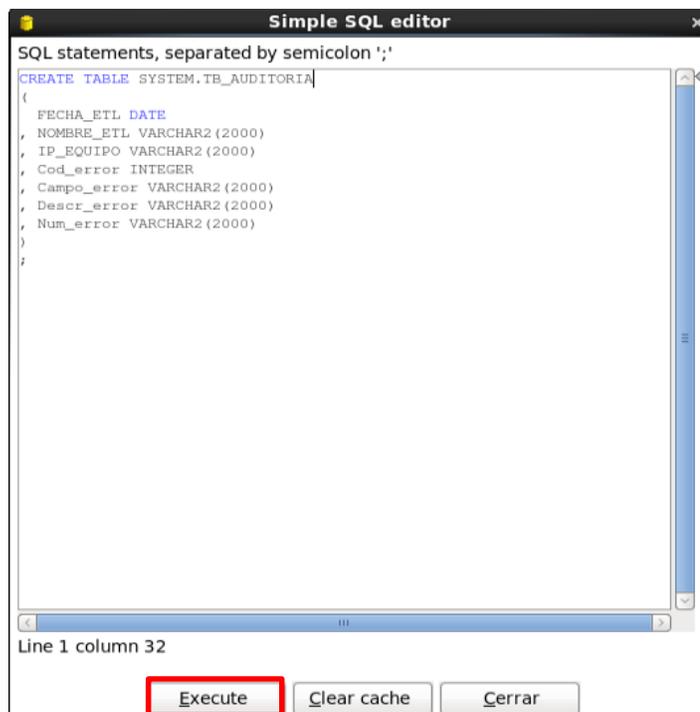


Figura 113: Ejecución script tabla auditoria

Finalmente en la siguiente imagen se hace clic en vale donde se está verificando que la nueva tabla ha sido creada.

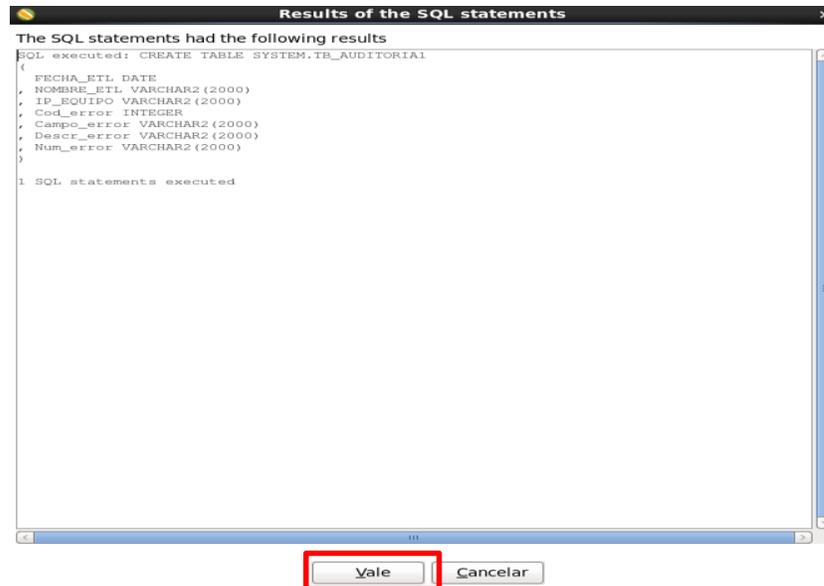


Figura 114: Resultado tabla auditoria creada

4.7.3 Ejecutar el ETL

Una vez creado todo el proceso ETL para la creación de la dimensión se ejecuta haciendo clic en el botón que está marcado con el recuadro en la siguiente imagen.

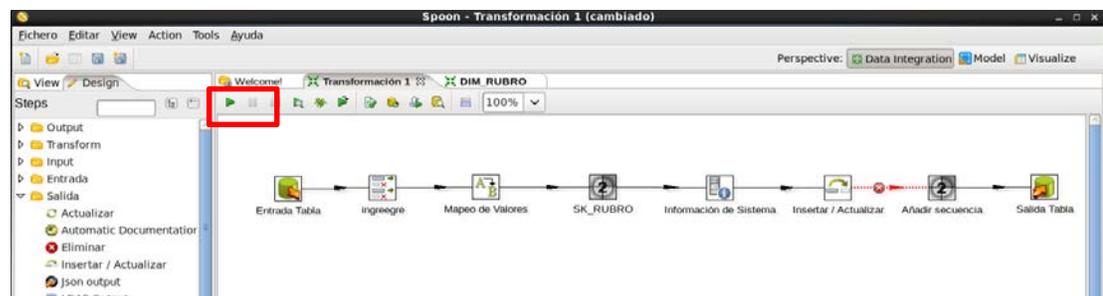


Figura 115: Botón para ejecutar ETL

Al hacer el paso anterior aparece una nueva pantalla y se hace clic en ejecutar para que se empiece a cargar la nueva dimensión.

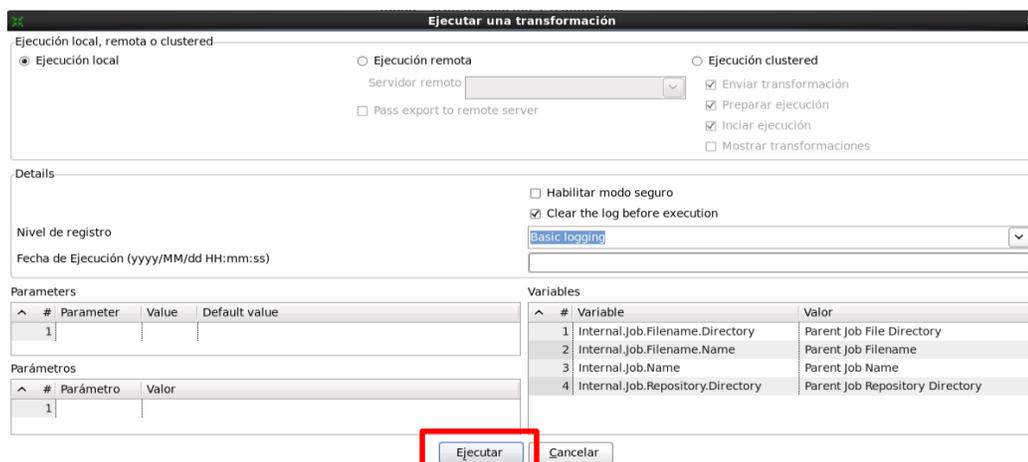


Figura 116: Ejecución ETL

Antes de que se ejecute el ETL la herramienta le va a pedir que guarde el archivo, es por eso que a continuación se hace clic en la opción si

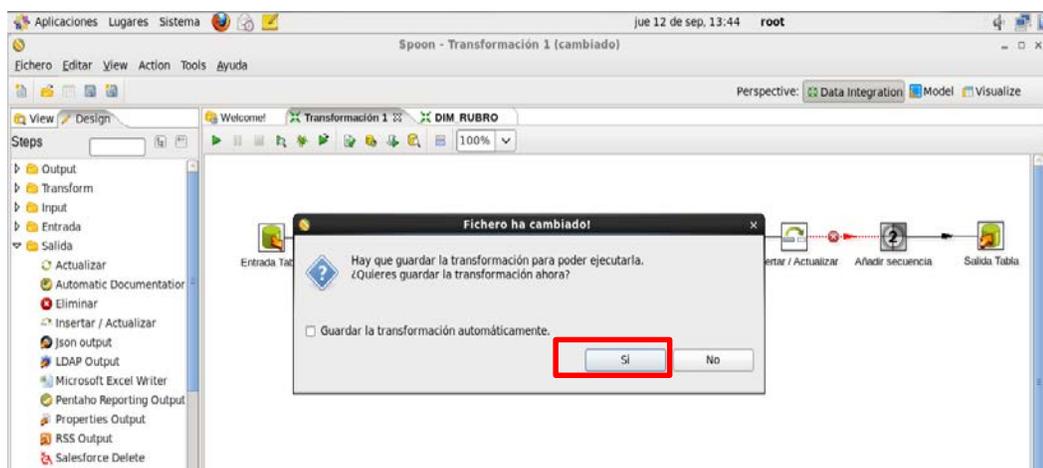


Figura 117: Guardar ETL

Ahora se abrirá un ventana en la cual se pone la ruta que el usuario desee y el nombre del archivo con el que le va a identificar, en este caso se está guardando en la ruta /root/Escritorio/ETL con el nombre DIM_RUBRO.

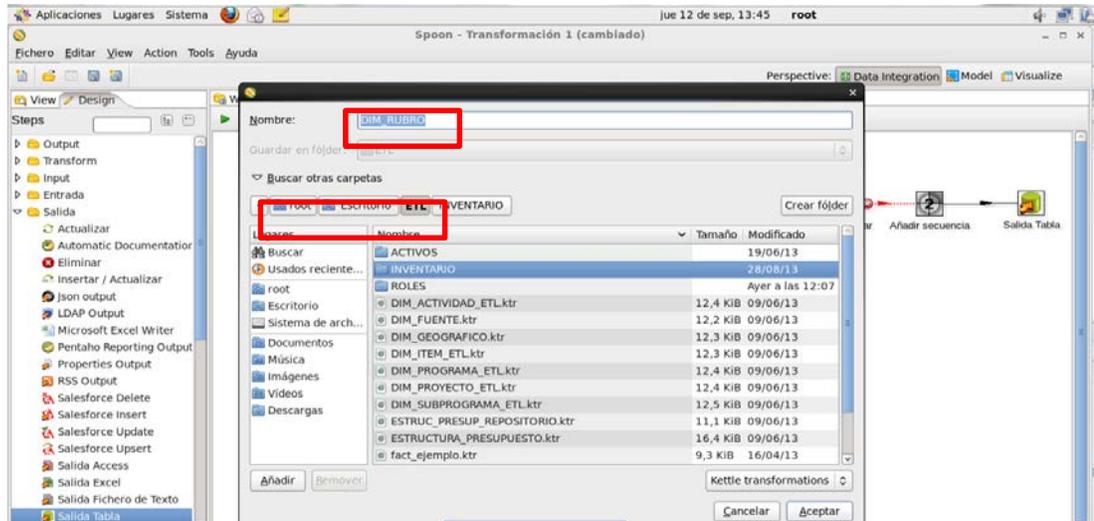


Figura 118: Ruta para guardar ETL

Por último se muestra la pantalla final de cómo quedaría un proceso ETL ejecutado en el cual en la parte inferior se detalla los registros que lee, escribe cada elemento y la salida que es la cantidad de registros que cargaron y sin errores.

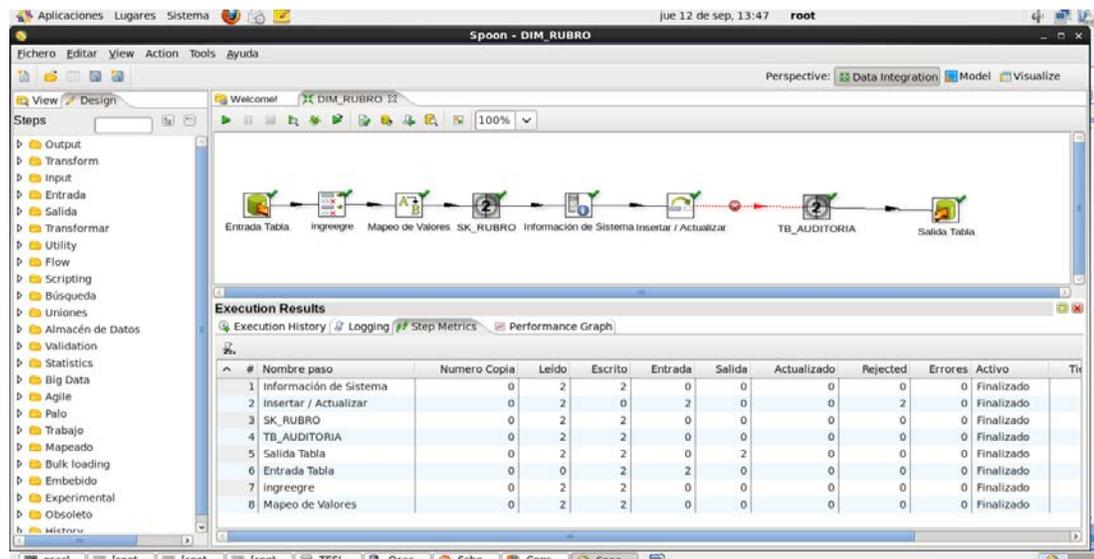


Figura 119: ETL Ejecutado

4.8 MANUAL CARGA DE ARCHIVOS CSV. AL SERVIDOR DE PENTAHO

1. El archivo de texto (csv), debe ser extraído del sistema ESIGEF, para el mes en curso.
2. Abrir una hoja de cálculo en Microsoft Excel y realizar los siguientes pasos:

a. Ir a la pestaña DATOS

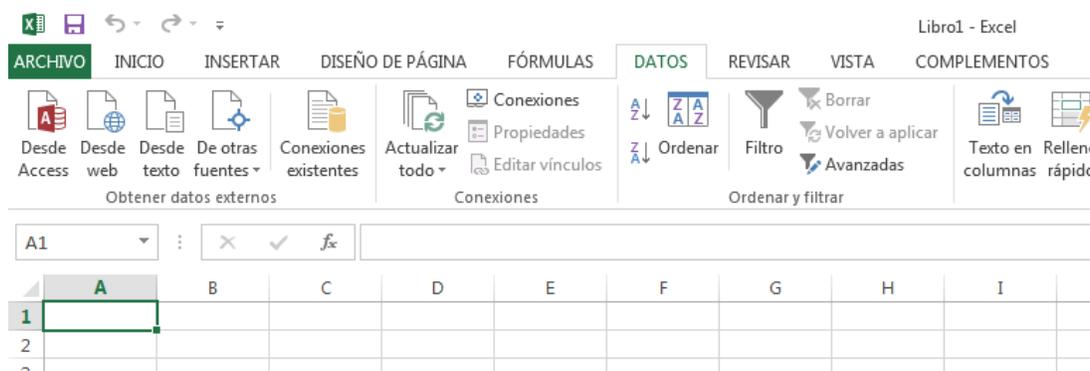


Figura 120: Pestaña datos en Excel 2013

- b. En la sección “Obtener datos externos”, hacer clic en “Desde texto” y a continuación, seleccionar el archivo csv y hacer clic en importar.

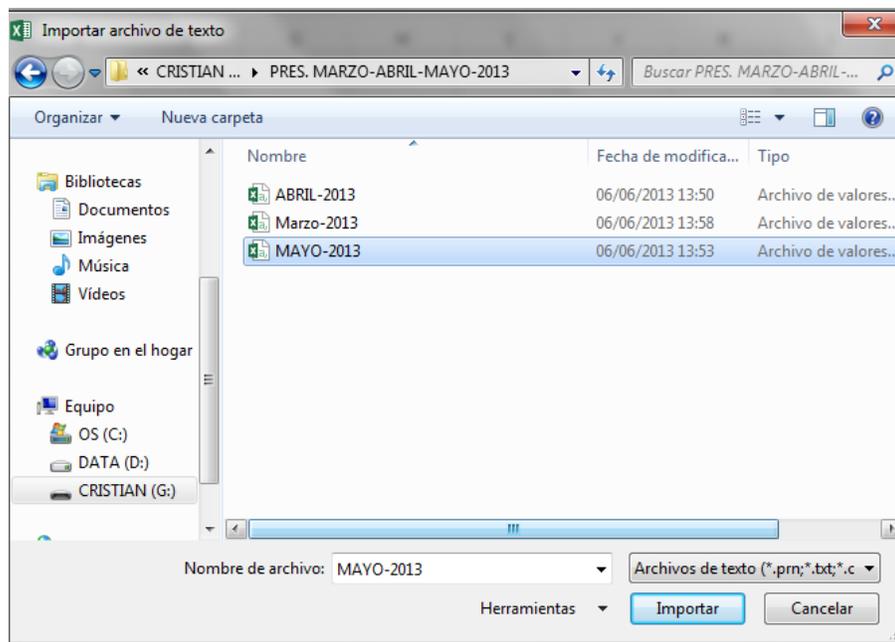


Figura 121: Venta importar archivo de texto

- c. Se desplegará la siguiente pantalla de configuración, en la que se selecciona “tipo de archivo”, la opción “Delimitados”. A continuación hacer clic en siguiente.

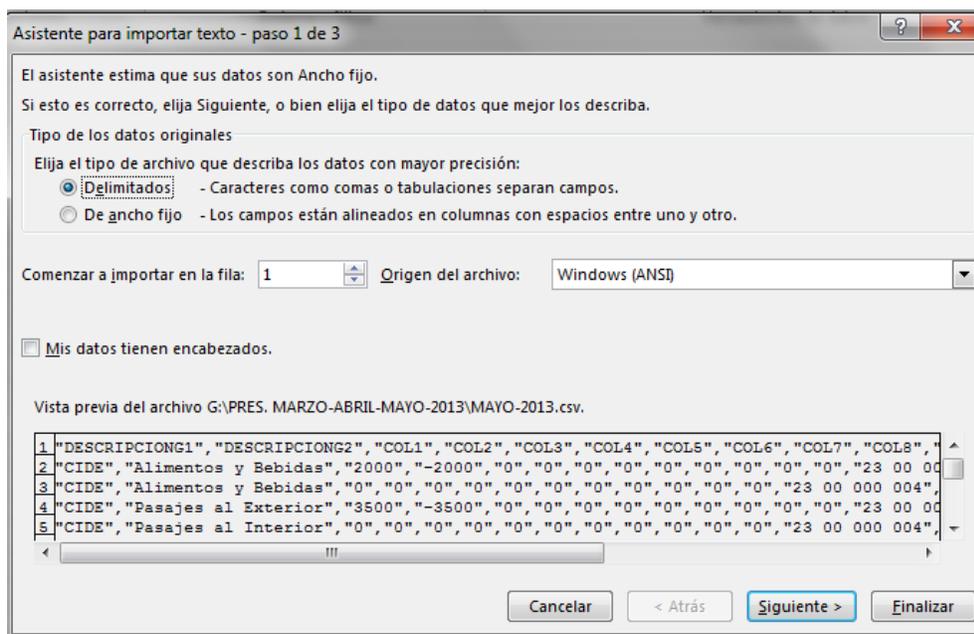


Figura 122: Asistente para importar texto paso 1

- d. Se desplegará la siguiente pantalla, en la que se selecciona “Coma” en la sección de “Separadores”, la caja de texto correspondiente a “Calificador de texto” debe quedar con el carácter de comilla. A continuación hacer clic en siguiente.

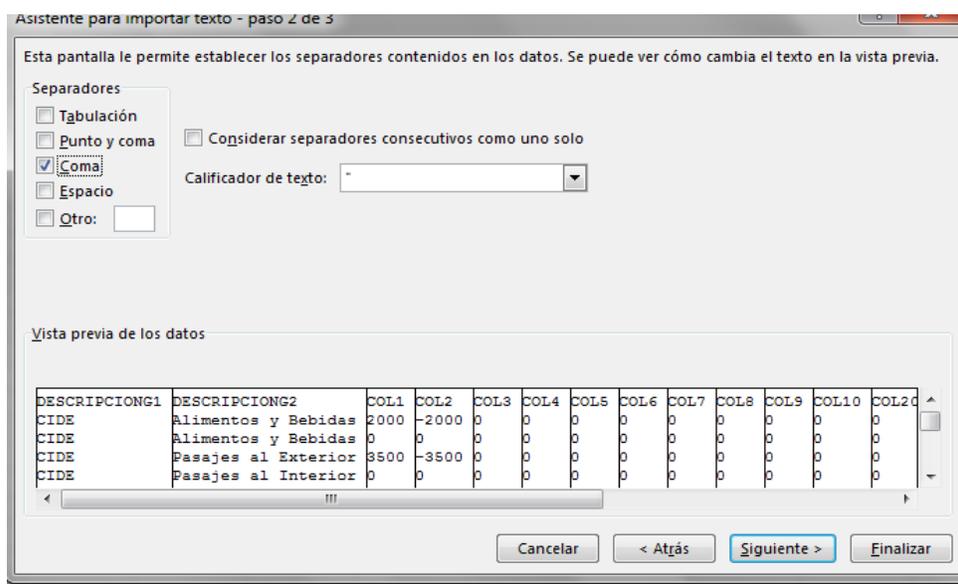


Figura 123: Asistente para importar texto paso 2

En la siguiente pantalla, se deja las opciones que se muestran por defecto y hacer clic en finalizar.

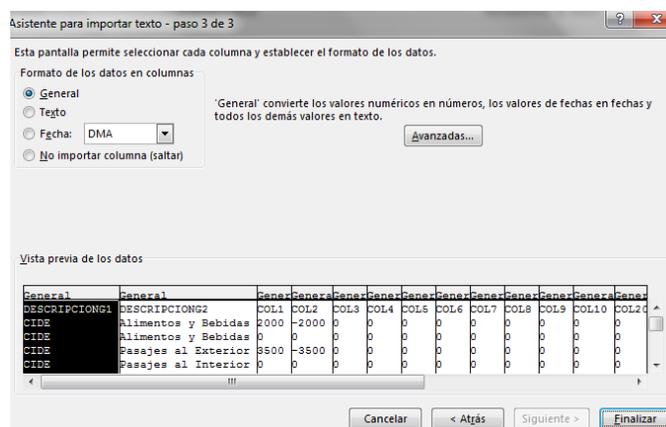
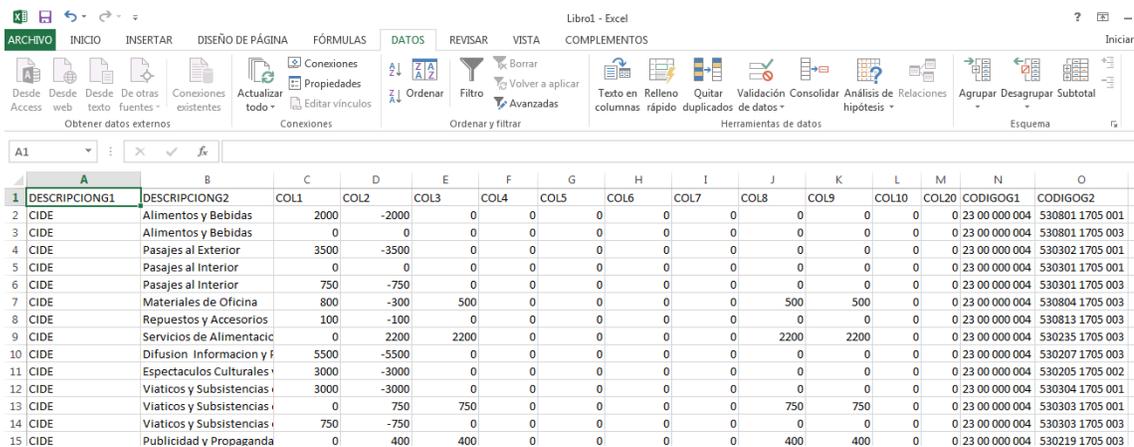


Figura 124: Asistente para importar texto paso 3

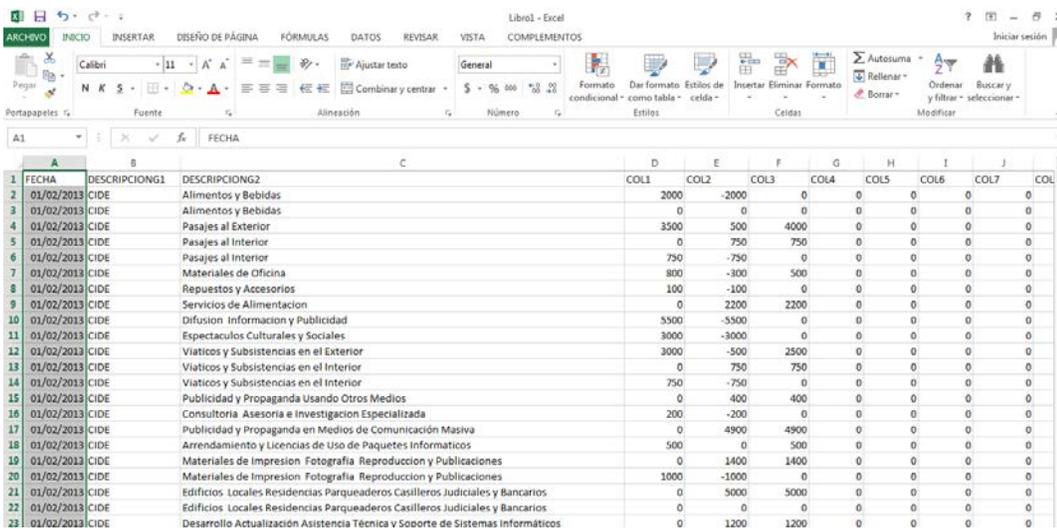
e. Se importará el archivo csv, a formato Excel, de la siguiente manera:



	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O
1	DESCRIPCIONG1	DESCRIPCIONG2	COL1	COL2	COL3	COL4	COL5	COL6	COL7	COL8	COL9	COL10	COL20	CODIGOG1	CODIGOG2
2	CIDE	Alimentos y Bebidas	2000	-2000	0	0	0	0	0	0	0	0	0	0 23 00 000 004	530801 1705 001
3	CIDE	Alimentos y Bebidas	0	0	0	0	0	0	0	0	0	0	0	0 23 00 000 004	530801 1705 003
4	CIDE	Pasajes al Exterior	3500	-3500	0	0	0	0	0	0	0	0	0	0 23 00 000 004	530302 1705 001
5	CIDE	Pasajes al Interior	0	0	0	0	0	0	0	0	0	0	0	0 23 00 000 004	530301 1705 001
6	CIDE	Pasajes al Interior	750	-750	0	0	0	0	0	0	0	0	0	0 23 00 000 004	530301 1705 003
7	CIDE	Materiales de Oficina	800	-300	500	0	0	0	0	500	500	0	0	0 23 00 000 004	530804 1705 003
8	CIDE	Repuestos y Accesorios	100	-100	0	0	0	0	0	0	0	0	0	0 23 00 000 004	530813 1705 003
9	CIDE	Servicios de Alimentación	0	2200	2200	0	0	0	0	2200	2200	0	0	0 23 00 000 004	530235 1705 003
10	CIDE	Difusión Información y f	5500	-5500	0	0	0	0	0	0	0	0	0	0 23 00 000 004	530207 1705 003
11	CIDE	Espectáculos Culturales y	3000	-3000	0	0	0	0	0	0	0	0	0	0 23 00 000 004	530205 1705 002
12	CIDE	Viaticos y Subsistencias i	3000	-3000	0	0	0	0	0	0	0	0	0	0 23 00 000 004	530304 1705 001
13	CIDE	Viaticos y Subsistencias i	0	750	750	0	0	0	0	750	750	0	0	0 23 00 000 004	530303 1705 001
14	CIDE	Viaticos y Subsistencias i	750	-750	0	0	0	0	0	0	0	0	0	0 23 00 000 004	530303 1705 003
15	CIDE	Publicidad y Propaganda	0	400	400	0	0	0	0	400	400	0	0	0 23 00 000 004	530219 1705 003

Figura 125: Archivo de texto exportado

f. A continuación, se debe insertar una nueva columna que llevara como cabecera el nombre “FECHA” y en todas sus filas se colocará la fecha correspondiente al mes del reporte extraído en el formato: DD/MM/ASA, como se muestra en la siguiente figura:



	A	B	C	D	E	F	G	H	I	J	K
1	FECHA	DESCRIPCIONG1	DESCRIPCIONG2	COL1	COL2	COL3	COL4	COL5	COL6	COL7	COL8
2	01/02/2013	CIDE	Alimentos y Bebidas	2000	-2000	0	0	0	0	0	0
3	01/02/2013	CIDE	Alimentos y Bebidas	0	0	0	0	0	0	0	0
4	01/02/2013	CIDE	Pasajes al Exterior	3500	500	4000	0	0	0	0	0
5	01/02/2013	CIDE	Pasajes al Interior	0	750	750	0	0	0	0	0
6	01/02/2013	CIDE	Pasajes al Interior	750	-750	0	0	0	0	0	0
7	01/02/2013	CIDE	Materiales de Oficina	800	-300	500	0	0	0	0	0
8	01/02/2013	CIDE	Repuestos y Accesorios	100	-100	0	0	0	0	0	0
9	01/02/2013	CIDE	Servicios de Alimentación	0	2200	2200	0	0	0	0	0
10	01/02/2013	CIDE	Difusión Información y Publicidad	5500	-5500	0	0	0	0	0	0
11	01/02/2013	CIDE	Espectáculos Culturales y Sociales	3000	-3000	0	0	0	0	0	0
12	01/02/2013	CIDE	Viaticos y Subsistencias en el Exterior	3000	-500	2500	0	0	0	0	0
13	01/02/2013	CIDE	Viaticos y Subsistencias en el Interior	0	750	750	0	0	0	0	0
14	01/02/2013	CIDE	Viaticos y Subsistencias en el Interior	750	-750	0	0	0	0	0	0
15	01/02/2013	CIDE	Publicidad y Propaganda Usando Otros Medios	0	400	400	0	0	0	0	0
16	01/02/2013	CIDE	Consultoría Asesoría e Investigación Especializada	200	-200	0	0	0	0	0	0
17	01/02/2013	CIDE	Publicidad y Propaganda en Medios de Comunicación Masiva	0	4900	4900	0	0	0	0	0
18	01/02/2013	CIDE	Arrendamiento y Licencias de Uso de Paquetes Informáticos	500	0	500	0	0	0	0	0
19	01/02/2013	CIDE	Materiales de Impresión Fotografía Reproducción y Publicaciones	0	1400	1400	0	0	0	0	0
20	01/02/2013	CIDE	Materiales de Impresión Fotografía Reproducción y Publicaciones	1000	-1000	0	0	0	0	0	0
21	01/02/2013	CIDE	Edificios Locales Residencias Parquederos Casilleros Judiciales y Bancarios	0	5000	5000	0	0	0	0	0
22	01/02/2013	CIDE	Edificios Locales Residencias Parquederos Casilleros Judiciales y Bancarios	0	0	0	0	0	0	0	0
23	01/02/2013	CIDE	Desarrollo Actualización Asistencia Técnica y Soporte de Sistemas Informáticos	0	1200	1200	0	0	0	0	0

Figura 126: Modificación de archivo de texto

3. A continuación se debe guardar el archivo generado en formato “**libro de Excel 97-2003 (*.xls)**” con el nombre: **COSTOS**

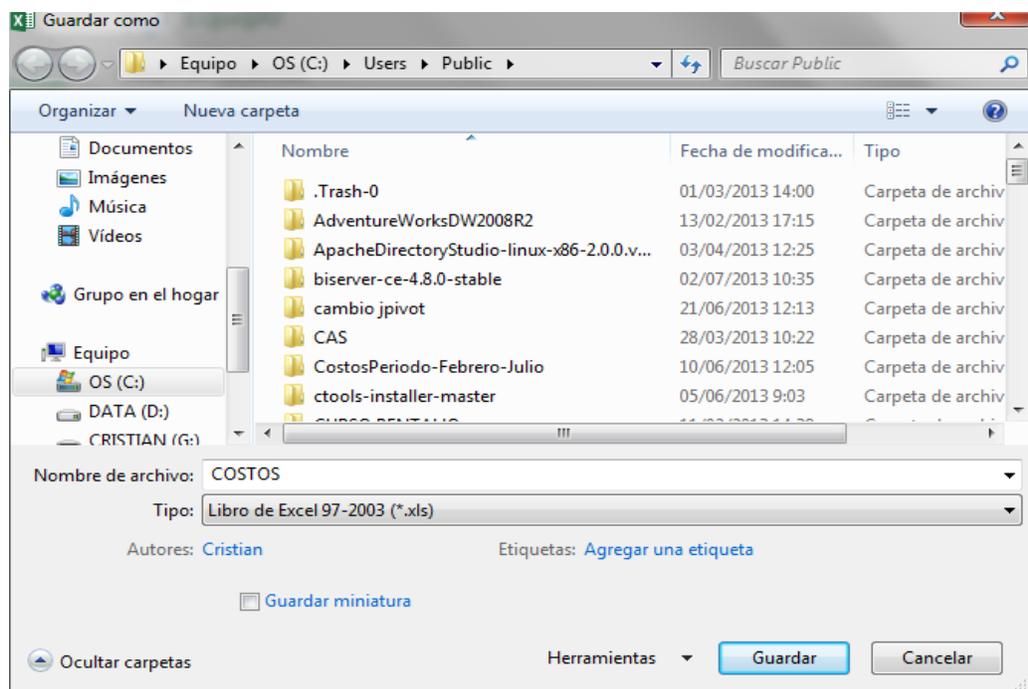


Figura 127: Ventana guardar como

4. Finalmente el archivo de Excel generado, debe ser copiado en el servidor de pentaho con ip: 10.1.1.35 en la siguiente ruta:

**/opt/pentaho/biserver-Ce/pentaho-
solutions/DATAMART_FINANZAS_RRHH/FINANZAS/ETL/INPUT**

4.9 MANUAL CARGA DE ARCHIVO DE EJECUCIÓN PRESUPUESTARIA

1. El formato que trae el archivo de ejecución presupuestaria no debe ser modificado:

UNIDAD	NOMBRE	PROGRAMA	NOM. ESTRUCTURA	PROYECTO	NOM DEL PRO	COIGO	CUP	NOMBRE	CUP	ACTIVIDAD	NOM ACTIVIDAD	GG	ITEM	NOMBRE	FUENTE ORGAN
1	CENTRO DE TRANSFERENCIA Y DESARROLLO TECNOLÓGICO		TRANSFERENCIA Y DES	0	SIN PROYECTO					1	TRANSFERENCIA Y 53		530101	Agua Potable	2
2	CENTRO DE TRANSFERENCIA Y DESARROLLO TECNOLÓGICO		TRANSFERENCIA Y DES	0	SIN PROYECTO					1	TRANSFERENCIA Y 53		530104	Energía Eléctrica	2
3	CENTRO DE TRANSFERENCIA Y DESARROLLO TECNOLÓGICO		TRANSFERENCIA Y DES	0	SIN PROYECTO					1	TRANSFERENCIA Y 53		530105	Telecomunicaciones	2
4	CENTRO DE TRANSFERENCIA Y DESARROLLO TECNOLÓGICO		TRANSFERENCIA Y DES	0	SIN PROYECTO					1	TRANSFERENCIA Y 53		530106	Servicio de Correo	2
5	CENTRO DE TRANSFERENCIA Y DESARROLLO TECNOLÓGICO		TRANSFERENCIA Y DES	0	SIN PROYECTO					1	TRANSFERENCIA Y 53		530201	Transporte de Pasajeros	2
6	CENTRO DE TRANSFERENCIA Y DESARROLLO TECNOLÓGICO		TRANSFERENCIA Y DES	0	SIN PROYECTO					1	TRANSFERENCIA Y 53		530204	Edición Impresión	2
7	CENTRO DE TRANSFERENCIA Y DESARROLLO TECNOLÓGICO		TRANSFERENCIA Y DES	0	SIN PROYECTO					1	TRANSFERENCIA Y 53		530205	Eventos Públicos y C	2
8	CENTRO DE TRANSFERENCIA Y DESARROLLO TECNOLÓGICO		TRANSFERENCIA Y DES	0	SIN PROYECTO					1	TRANSFERENCIA Y 53		530207	Difusión Informacion	2
9	CENTRO DE TRANSFERENCIA Y DESARROLLO TECNOLÓGICO		TRANSFERENCIA Y DES	0	SIN PROYECTO					1	TRANSFERENCIA Y 53		530218	Publicidad y Propaga	2
10	CENTRO DE TRANSFERENCIA Y DESARROLLO TECNOLÓGICO		TRANSFERENCIA Y DES	0	SIN PROYECTO					1	TRANSFERENCIA Y 53		530299	Otros Servicios Gene	998
11	CENTRO DE TRANSFERENCIA Y DESARROLLO TECNOLÓGICO		TRANSFERENCIA Y DES	0	SIN PROYECTO					1	TRANSFERENCIA Y 53		530299	Otros Servicios Gene	2
12	CENTRO DE TRANSFERENCIA Y DESARROLLO TECNOLÓGICO		TRANSFERENCIA Y DES	0	SIN PROYECTO					1	TRANSFERENCIA Y 53		530301	Pasajes al Interior	2
13	CENTRO DE TRANSFERENCIA Y DESARROLLO TECNOLÓGICO		TRANSFERENCIA Y DES	0	SIN PROYECTO					1	TRANSFERENCIA Y 53		530302	Pasajes al Exterior	2
14	CENTRO DE TRANSFERENCIA Y DESARROLLO TECNOLÓGICO		TRANSFERENCIA Y DES	0	SIN PROYECTO					1	TRANSFERENCIA Y 53		530303	Votivos y Substancie	2
15	CENTRO DE TRANSFERENCIA Y DESARROLLO TECNOLÓGICO		TRANSFERENCIA Y DES	0	SIN PROYECTO					1	TRANSFERENCIA Y 53		530304	Votivos y Substancie	2
16	CENTRO DE TRANSFERENCIA Y DESARROLLO TECNOLÓGICO		TRANSFERENCIA Y DES	0	SIN PROYECTO					1	TRANSFERENCIA Y 53		530402	Edificios Locales Res	2
17	CENTRO DE TRANSFERENCIA Y DESARROLLO TECNOLÓGICO		TRANSFERENCIA Y DES	0	SIN PROYECTO					1	TRANSFERENCIA Y 53		530403	Mobiliarios	2
18	CENTRO DE TRANSFERENCIA Y DESARROLLO TECNOLÓGICO		TRANSFERENCIA Y DES	0	SIN PROYECTO					1	TRANSFERENCIA Y 53		530404	Maquinarias y Equip	2
19	CENTRO DE TRANSFERENCIA Y DESARROLLO TECNOLÓGICO		TRANSFERENCIA Y DES	0	SIN PROYECTO					1	TRANSFERENCIA Y 53		530405	Vehiculos	2
20	CENTRO DE TRANSFERENCIA Y DESARROLLO TECNOLÓGICO		TRANSFERENCIA Y DES	0	SIN PROYECTO					1	TRANSFERENCIA Y 53		530409	Libros y Colecciones	2
21	CENTRO DE TRANSFERENCIA Y DESARROLLO TECNOLÓGICO		TRANSFERENCIA Y DES	0	SIN PROYECTO					1	TRANSFERENCIA Y 53		530499	Otras Instalaciones I	2
22	CENTRO DE TRANSFERENCIA Y DESARROLLO TECNOLÓGICO		TRANSFERENCIA Y DES	0	SIN PROYECTO					1	TRANSFERENCIA Y 53		530502	Edificios Locales Res	2
23	CENTRO DE TRANSFERENCIA Y DESARROLLO TECNOLÓGICO		TRANSFERENCIA Y DES	0	SIN PROYECTO					1	TRANSFERENCIA Y 53		530601	Consultoría Asesoria	2
24	CENTRO DE TRANSFERENCIA Y DESARROLLO TECNOLÓGICO		TRANSFERENCIA Y DES	0	SIN PROYECTO					1	TRANSFERENCIA Y 53		530601	Consultoría Asesoria	998
25	CENTRO DE TRANSFERENCIA Y DESARROLLO TECNOLÓGICO		TRANSFERENCIA Y DES	0	SIN PROYECTO					1	TRANSFERENCIA Y 53		530602	Servicio de Auditoría	2
26	CENTRO DE TRANSFERENCIA Y DESARROLLO TECNOLÓGICO		TRANSFERENCIA Y DES	0	SIN PROYECTO					1	TRANSFERENCIA Y 53		530603	Servicio de Capacita	2
27	CENTRO DE TRANSFERENCIA Y DESARROLLO TECNOLÓGICO		TRANSFERENCIA Y DES	0	SIN PROYECTO					1	TRANSFERENCIA Y 53		530702	Arrendamiento y Licen	2
28	CENTRO DE TRANSFERENCIA Y DESARROLLO TECNOLÓGICO		TRANSFERENCIA Y DES	0	SIN PROYECTO					1	TRANSFERENCIA Y 53		530703	Arrendamiento de Eq	2
29	CENTRO DE TRANSFERENCIA Y DESARROLLO TECNOLÓGICO		TRANSFERENCIA Y DES	0	SIN PROYECTO					1	TRANSFERENCIA Y 53				2

Figura 128: Archivo de ejecución presupuestaria

2. La pestaña que contiene la información a ser subida al servidor se llama 'GASTOS'
3. Se debe modificar el nombre del archivo por el siguiente: "EJECUCION_PRESUPUESTARIA" y el archivo debe ser guardado en formato "libro de Excel 97-2003 (*.xls)".

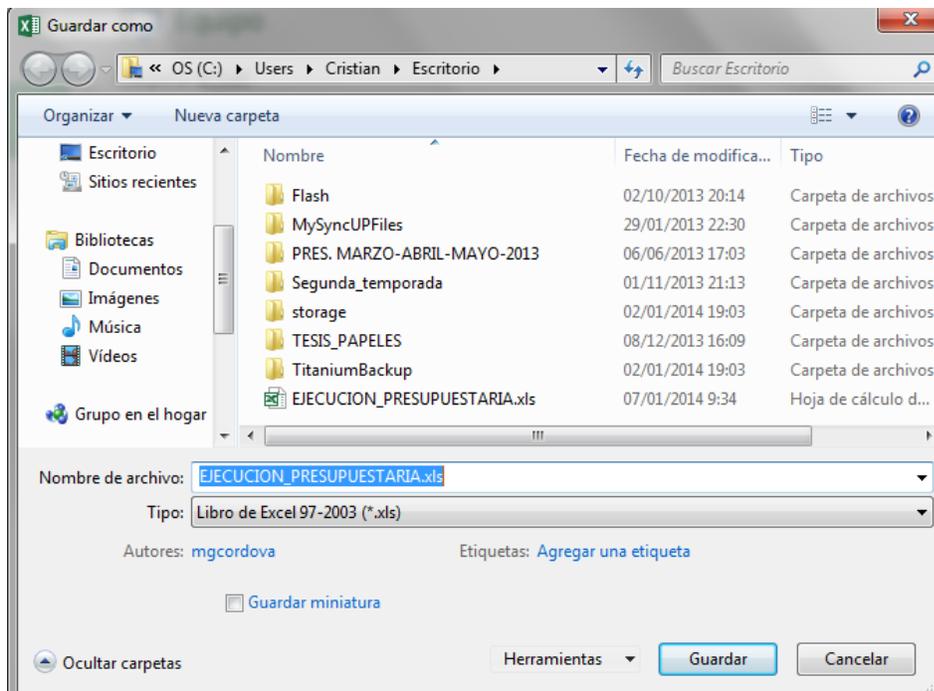


Figura 129: Ventana guardar como

4. Finalmente el archivo de Excel, debe ser copiado en el servidor de pentaho con ip: 10.1.1.35 en la siguiente ruta:

/opt/pentaho/biserver-Ce/pentaho-solutions/DATAMART_FINANZAS_RRHH/FINANZAS/ETL/INPUT

4.10 MANUAL DE USUARIO SCHEMA WORKBENCH

4.10.1 Introducción

Este manual de usuario está enfocado en todos los pasos técnicos que se debe realizar para una creación y una publicación de un cubo de información para un análisis OLAP.

Descripción de la interfaz de usuario

Primero se inicia el mondrian schema workbench con unos archivos que se encuentran en la carpeta que se descarga en el internet.

Para Windows se ejecuta el archivo workbench.bat para iniciar el schema workbench mientras que para Linux se ejecuta en un terminal el archivo workbench.sh, con esto carga y se abre la aplicación.

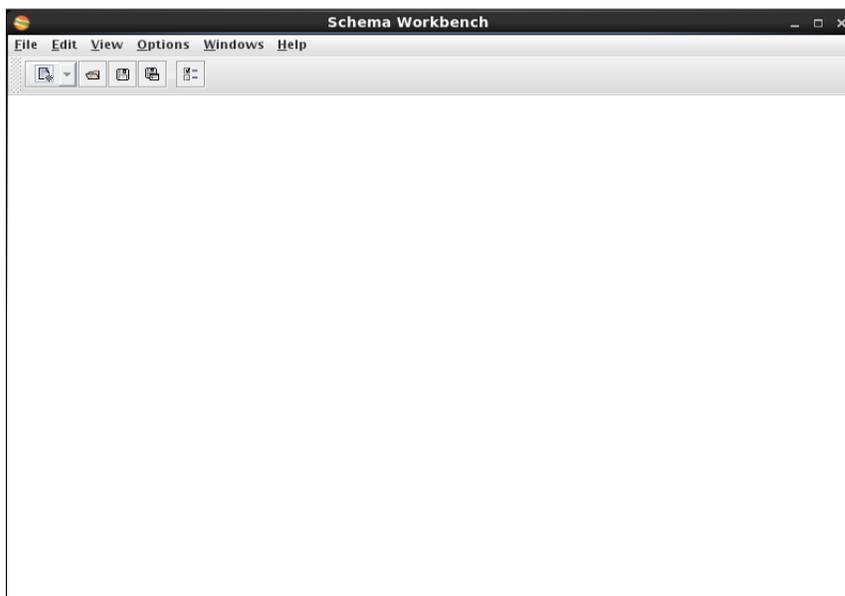


Figura 130: Interfaz Schema Workbench

4.10.2 Creación conexión

Antes de crear un cubo se debe crear la conexión a la base de datos de donde se va a extraer la información, para esto se dirige a la pestaña **options -> Connection**.

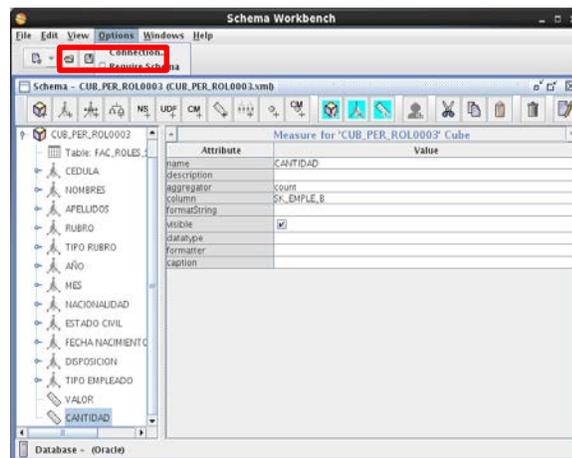


Figura 131: Creación conexión

Una vez que se abre la ventana para la conexión a la base de datos, se pone un nombre a la nueva conexión, se escoge tipo de conexión, la forma de acceso y las configuraciones necesarias para establecer una conexión, con esto hacer clic en el botón **Test**.

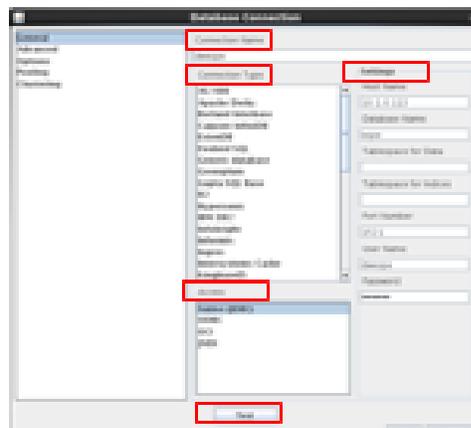


Figura 132: Configuración conexión

Si la conexión es correcta aparece un mensaje con la información de la conexión, hacer clic en el botón **ok**.

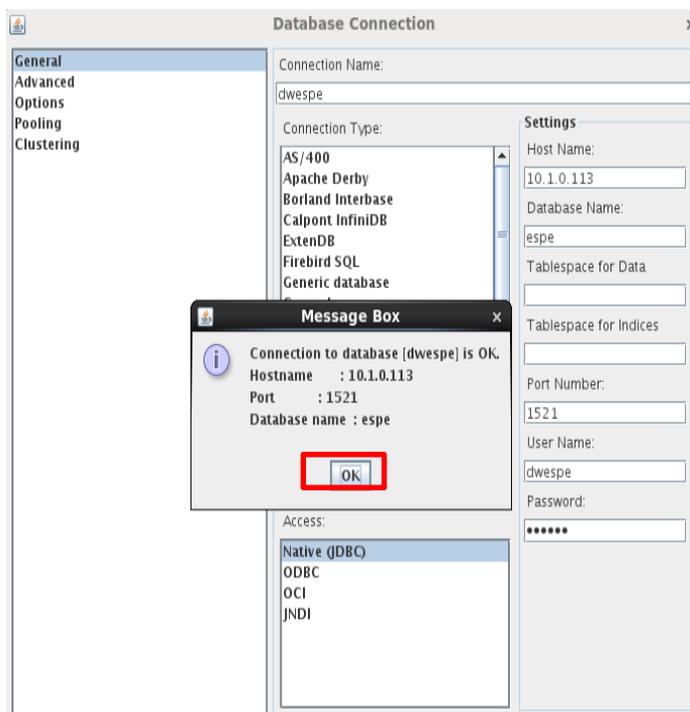


Figura 133: Conexión correcta

4.10.3 Creación de cubo

Para crear un cubo se debe primero crear un esquema, esto se crea en la opción: **file -> new -> schema**.

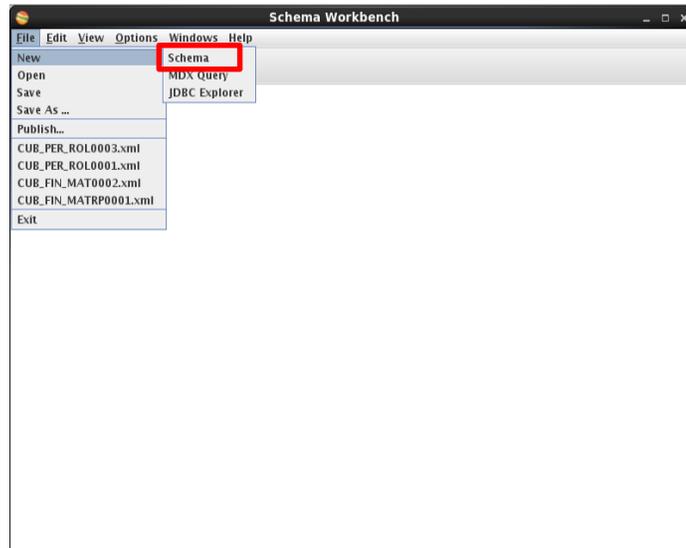


Figura 134: Crear nuevo esquema

Una vez creado el esquema se pone un nombre en la parte derecha donde se encuentran los atributos y luego se hace clic derecho sobre “**schema**”, aparece unas opciones y se escoge **add cube** para crear el cubo.

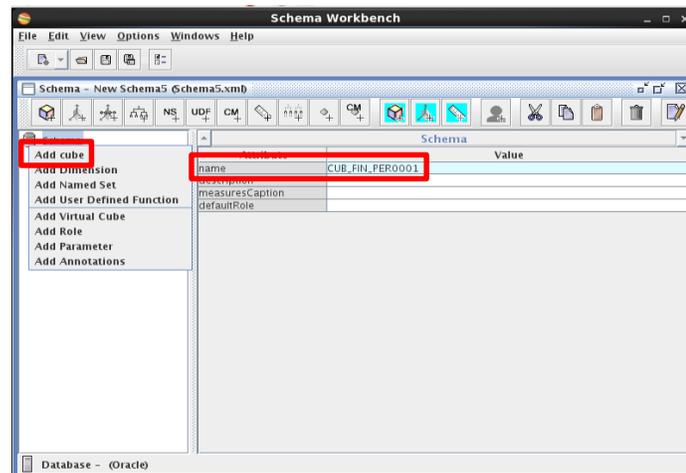


Figura 135: Crear nuevo cubo

Se pone un nombre al cubo en la parte de atributos que se encuentra en la parte derecha y hacer clic derecho sobre el cubo creado y escoger la opción **add table**.

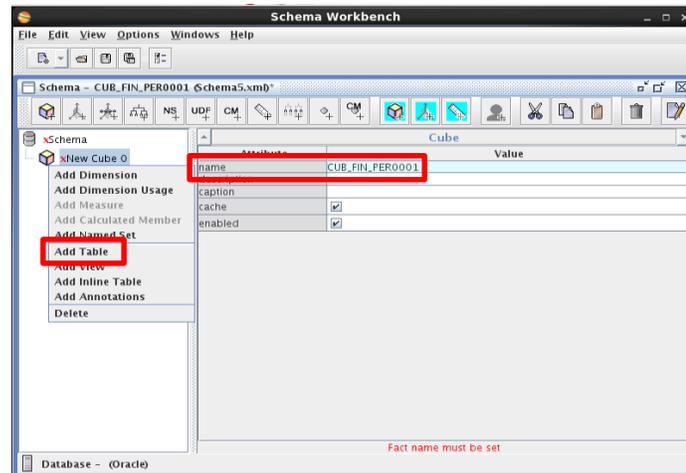


Figura 136: Creación tabla en cubo

Hacer clic en la tabla creada y llenar el atributo “**schema**”, aquí se pondrá el nombre del esquema donde se encuentran las dimensiones y tablas de hechos. En el atributo “**name**” se escoge la tabla de hechos del modelo multidimensional que se requiere hacer el cubo de información.

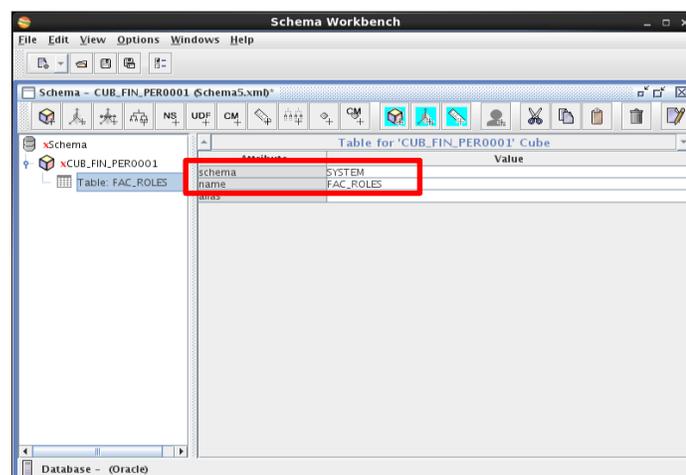


Figura 137: Atributos de tabla del cubo

4.10.4 Creación dimensiones

Hacer clic derecho sobre el cubo en la parte de la izquierda y escoger la opción **add dimensión** para crear la primera dimensión del cubo.

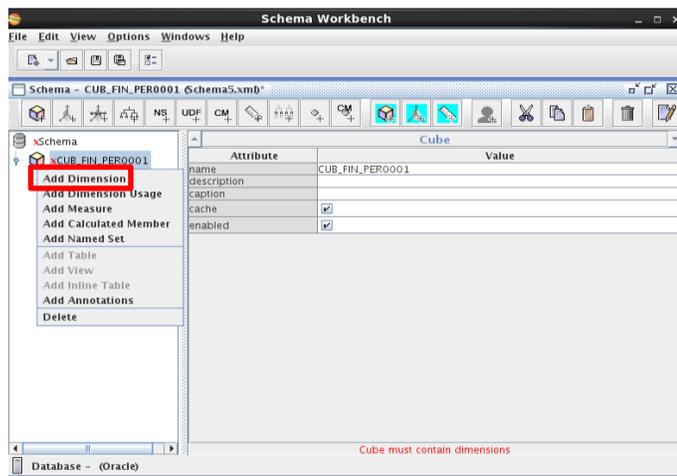


Figura 138: Crear dimensión

Al momento de crear la dimensión se genera una jerarquía automáticamente sobre la dimensión como se puede ver en la siguiente imagen. Ahora hacer clic derecho sobre la jerarquía y escoger la opción **add table**.

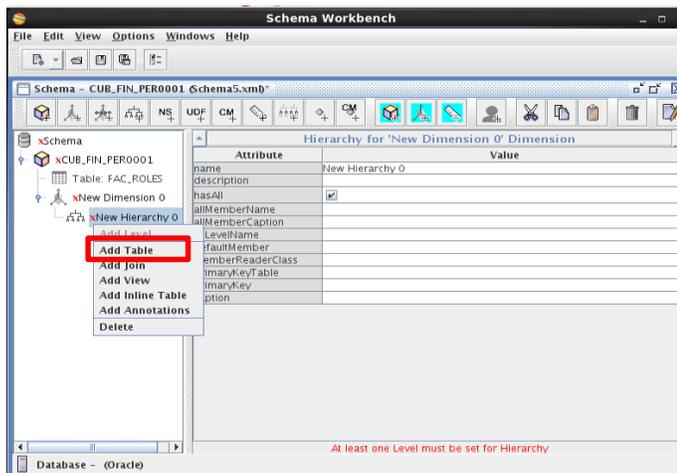


Figura 139: Crear tabla de dimensión

Se coloca sobre la tabla creada para la dimensión y se llena los atributos de “**schema**”, aquí se pone el mismo esquema que se colocó en la tabla anteriormente explicada y en “**name**” se escoge la dimensión del modelo multidimensional que se requiere hacer el cubo.

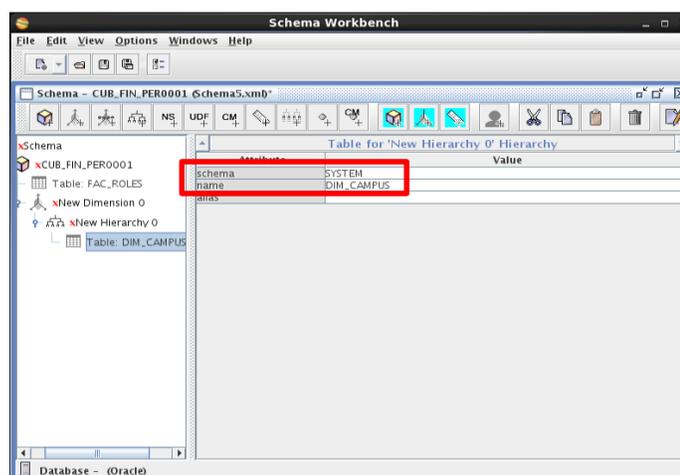


Figura 140: Atributos de tabla de dimensión

Hacer clic derecho sobre la jerarquía y escoger la opción **add leve**.

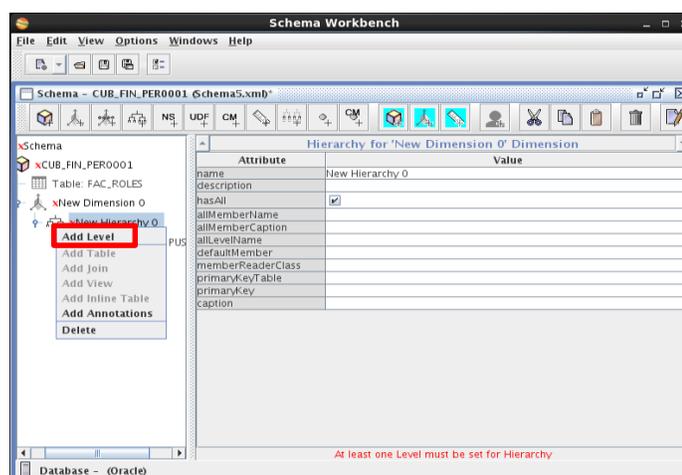


Figura 141: Crear nivel

Se coloca sobre el nivel creado y llenar los atributos:

- **Name:** Cualquier nombre que identifique al campo
- **Columna:** Se escoge el campo que requiera de la dimensión
- **Taype:** Se escoge el tipo de dato que tiene el campo de la dimensión.

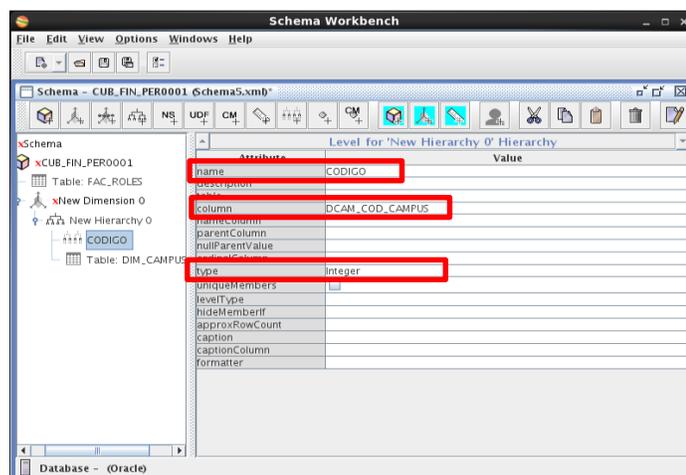


Figura 142: Atributos del nivel

La figura 143 muestra el resultado de dos niveles en el caso que el usuario requiera y el procedimiento sería el mismo que se explicó en el paso anterior.

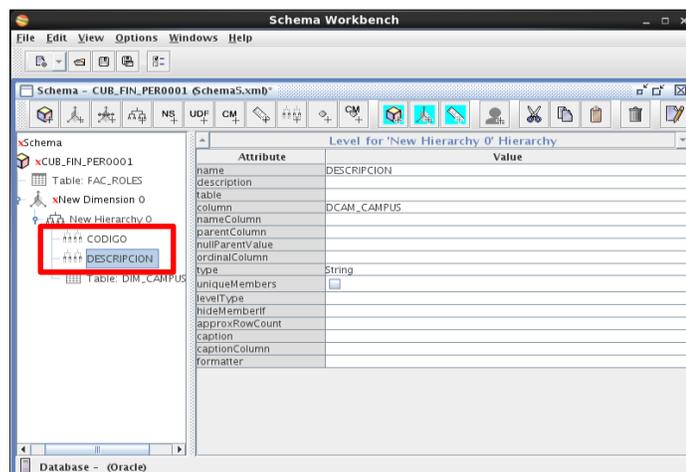


Figura 143: Dimensión con dos niveles

Una vez terminado los niveles, se coloca sobre la jerarquía y se llena los atributos:

- **Name:** Se pone un nombre referente a la dimensión o el nombre de la dimensión.

- **AllMemberName:** Se pone “all” o “todos”, esto será un nivel que genera automáticamente con este atributo y este nivel va a mostrar todos los registros que se encuentren en la dimensión.
- **primaryKey:** Se escoge la llave primaria de la dimensión.

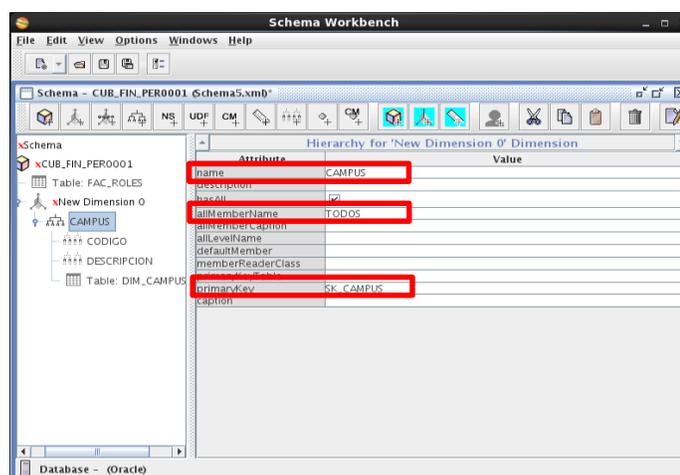


Figura 144: Atributos de la jerarquía

Con esto estarían completos los niveles y la jerarquía de la dimensión, ahora se coloca sobre la dimensión y se llena los atributos:

- **Name:** Se pone un nombre referente a la dimensión o el nombre de la dimensión.
- **ForeignKey:** Se escoge la llave secundaria o foránea de la tabla de hechos que hace referencia a la dimensión a la que se está utilizando.

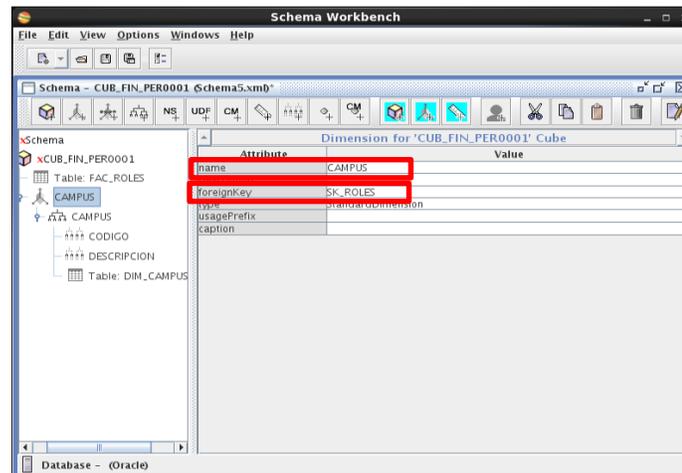


Figura 145: Atributos de la dimensión

4.10.5 Creación medidas

Hacer clic derecho sobre el cubo y escoger la opción **add measure** para crear una medida en el cubo de información.

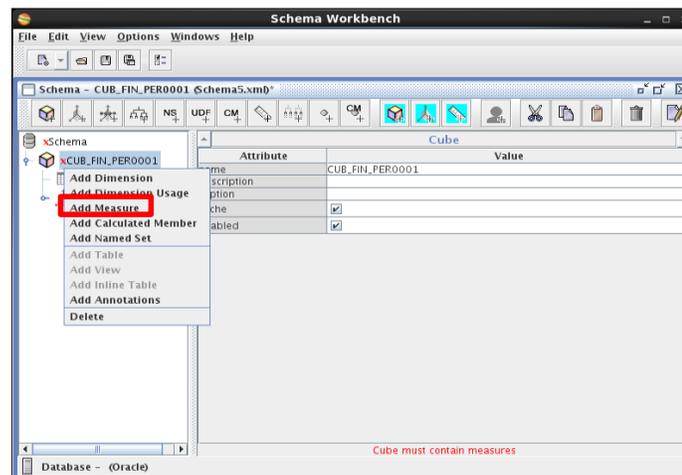


Figura 146: Crear medidas

Se coloca sobre la nueva medida y se llena los siguientes atributos:

- **Name:** Se pone un nombre a la medida.
- **Aggregator:** Se escoge la agregación que se requiera para la medida.
- **Columna:** Se escoge el campo que se va a utilizar como medida.

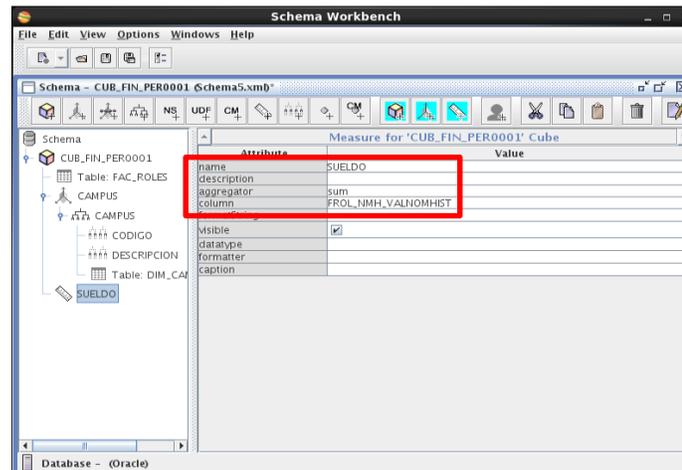


Figura 147: Atributos de las medidas

Al crear las dimensiones y medidas necesarias con un modelo multidimensional como se explicó anteriormente, el cubo se va a mostrar de la siguiente forma:

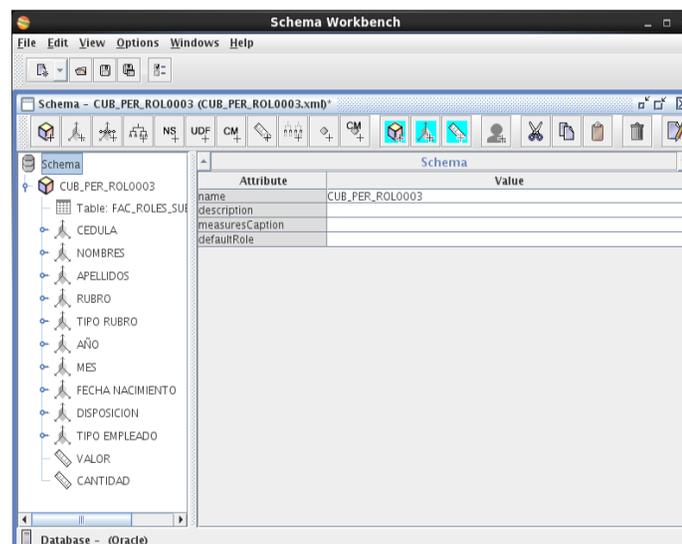


Figura 148: Cubo con dimensiones y medidas

4.10.6 Guardar cubos

Una vez que se tenga el cubo listo se dirige a guardarlo en la opción: **file -> Save As**

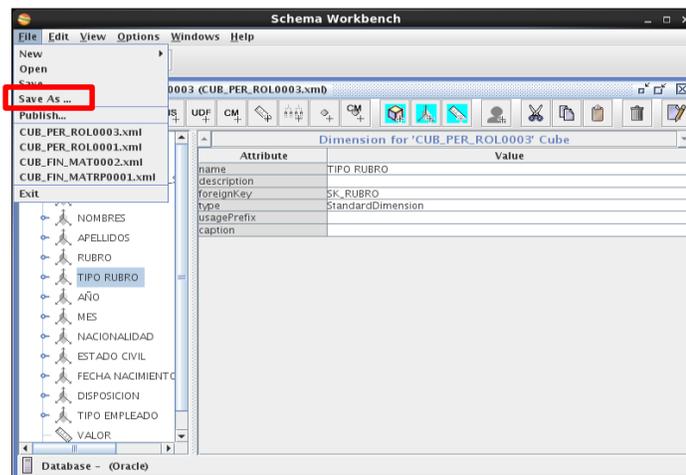


Figura 149: Guardar cubo

Se abrirá la ventana para poder guardar, escoger la ruta y poner un nombre con el que se va a guardar el cubo. En este caso se está guardando en la ruta **root** con el nombre **CUB_PER_ROL0003.xml**.

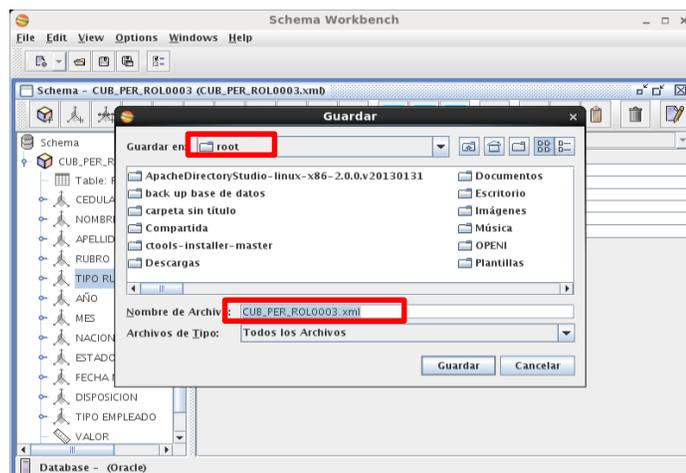


Figura 150: Ruta y nombre para guardar el cubo

4.10.7 Publicación de cubos

Para la publicación de cubos lo primero que se tiene en cuenta es que debe estar levantado el BI Server y en la consola de administración debe estar creado el usuario con el que va a publicar el cubo y también debe estar creada la conexión a la base de datos.

4.10.7.1 Configuración

Antes de publicar se debe configurar el archivo que se encuentra en la siguiente ruta:

/biserver-ce/pentaho-solutions/system/publisher_config.xml.

En este archivo se busca la siguiente parte:

```
<publisher-config>  
<publisher-password></publisher-password>  
</publisher-config>
```

Aquí se define el password que el usuario requiera para la publicación de cubos y reportes entre: <publisher-password></publisher-password>

Ejemplo: **<publisher-password>password</publisher-password>**

4.10.7.2 Publicación

Hacer clic en la opción: **file -> publish**

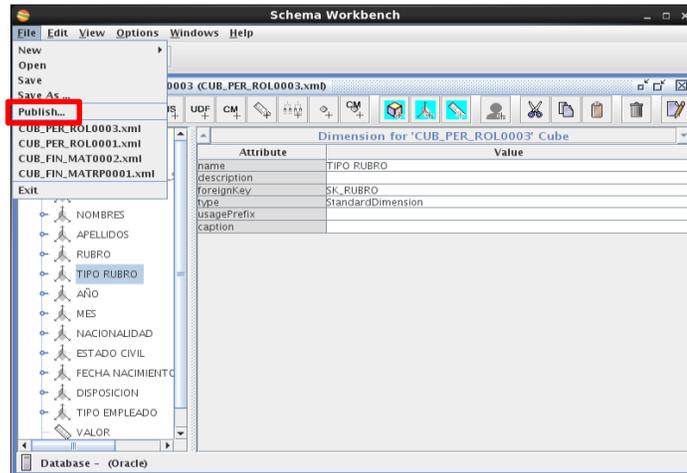


Figura 151: Publicación de cubo

Se va abrir una ventana donde se llena los siguientes campos como se puede ver en la siguiente imagen.

- **URL:** Se coloca la dirección del servidor.
- **PUBLISH PASSWORD:** Se pone la contraseña de la configuración que fue explicada anteriormente.
- **USER:** Se pone un usuario que este creado en la consola de administración de pentaho.
- **PASSWORD:** La contraseña del usuario.

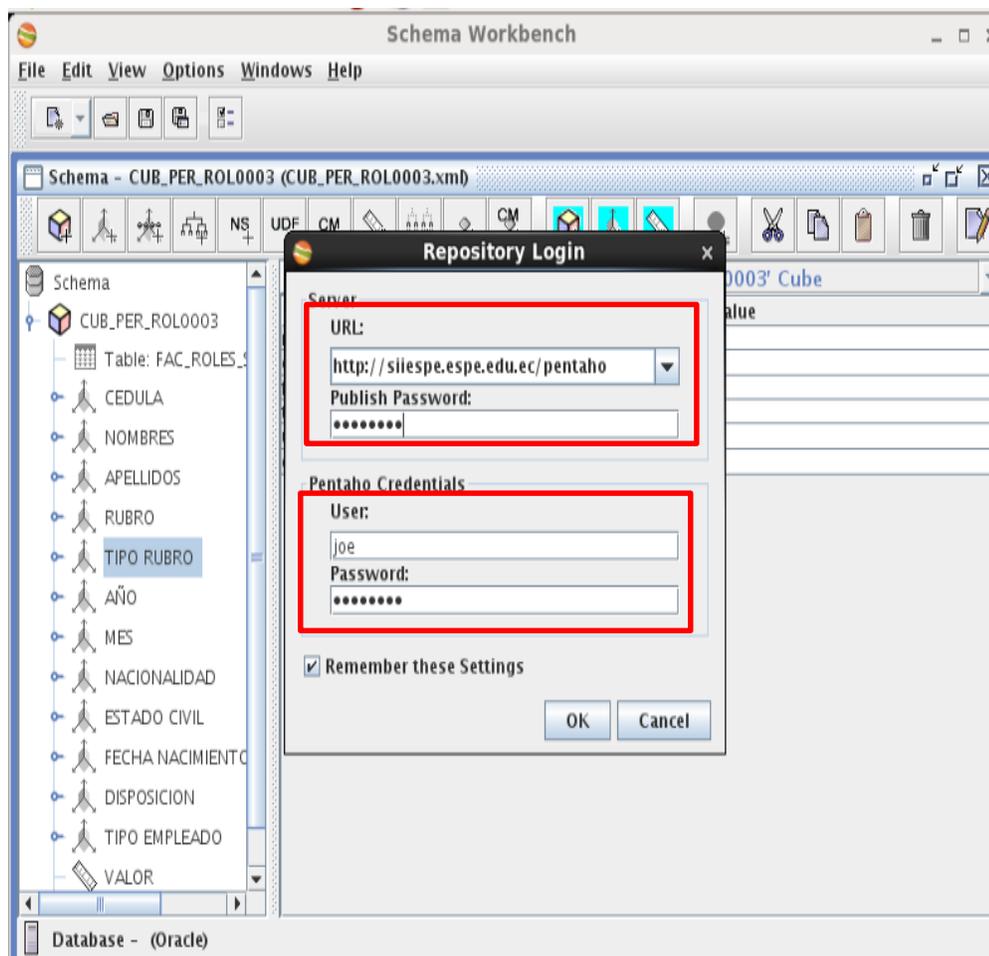


Figura 152: Conexión al repositorio

Con el paso anterior se conecta al repositorio del servidor y se abre una nueva ventana donde se va a definir la ruta de publicación del cubo y en la parte de “**pentaho or JNDI Data Source**” se pone el mismo nombre que se da a la conexión al crear en la consola de administración.

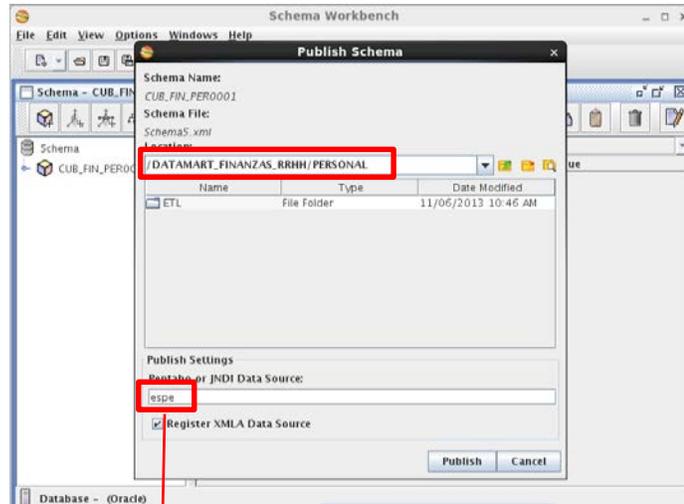


Figura 153: Ruta de publicación del cubo

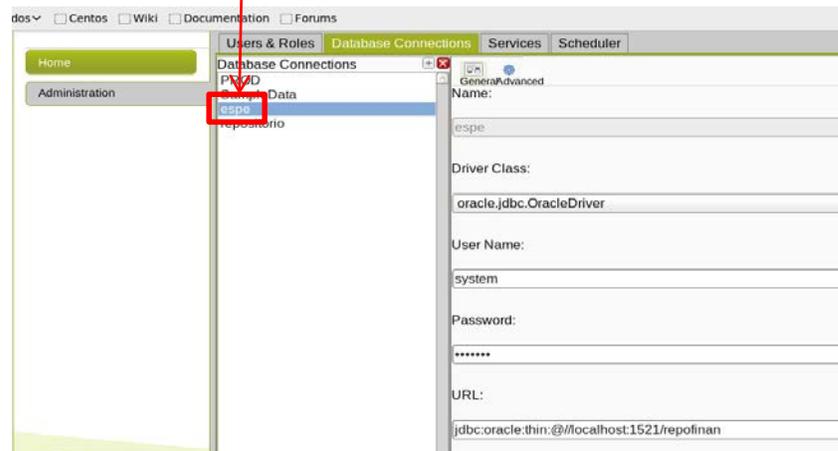


Figura 154: Conexión en consola de administración

Finalmente al publicarlo correctamente, se muestra un mensaje indicando que el cubo se ha publicado satisfactoriamente.

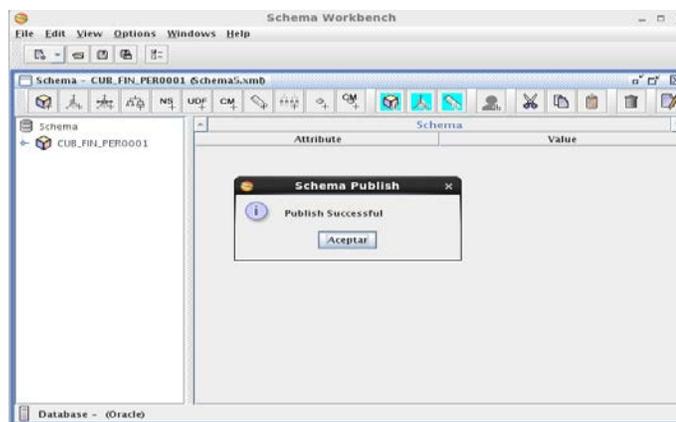


Figura 155: Cubo publicado satisfactoriamente

4.10.8 Uso de un cubo en la consola de usuario

Ingresar a la siguiente dirección web: **siiespe.espe.edu.ec**, preferiblemente usando Mozilla Firefox o Google Chrome. Se desplegará la siguiente pantalla:

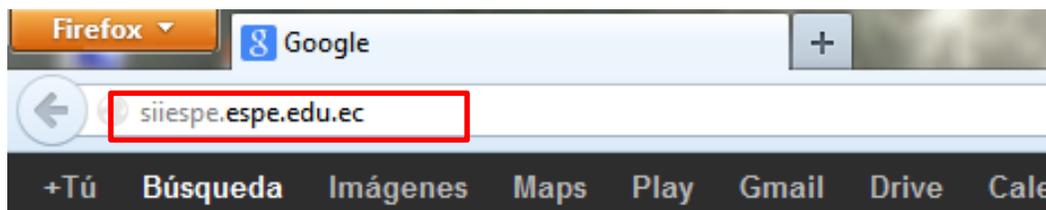


Figura 156: URL de la consola de usuario

Se abre la consola de usuario, se escribe el usuario y el password asignado a cada usuario por el administrador. Un ejemplo puede ser usuario: **joe** y password: *********



Figura 157: Interfaz de login de la consola de usuario

Finalmente se abre la consola de usuario conectada con un usuario para usarla.

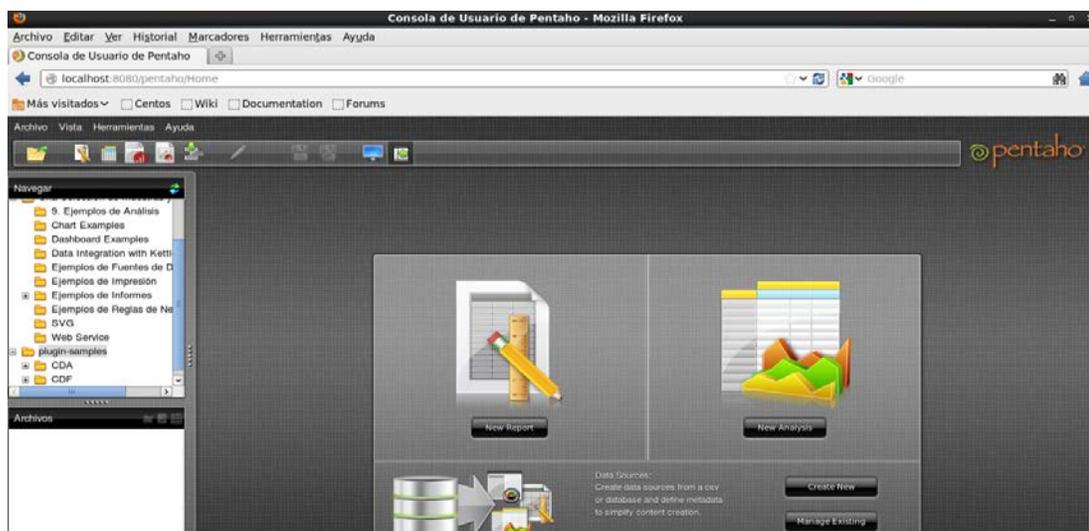


Figura 158: Interfaz consola de usuario

Hacer clic en el ícono que se encuentra en el recuadro de la figura 159, esto abrirá el área de trabajo de saiku analytics la cual permite manipular los cubos publicados.

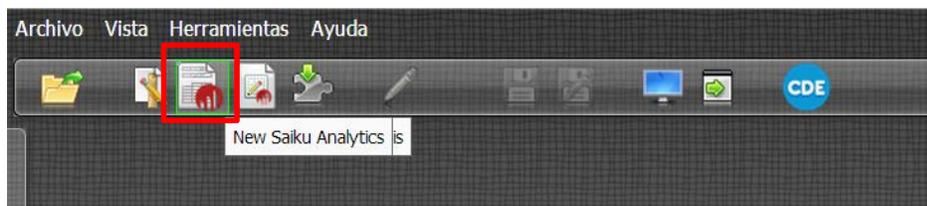


Figura 159: Nuevo saiku analytics

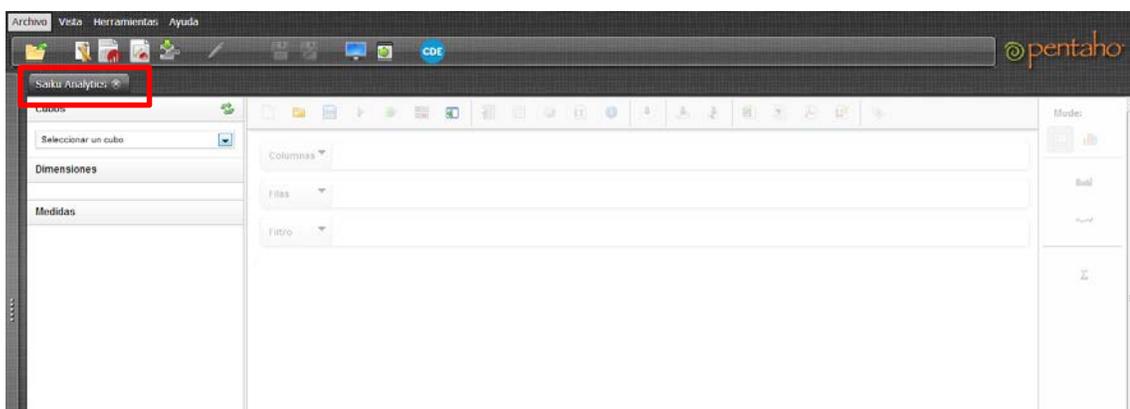


Figura 160: Ventana de saiku

Una vez abierto una ventana de saiku analytics, va a pedir que seleccione un cubo, así que, hacer clic en “Seleccionar un cubo” y escoger el cubo que requiera el usuario, para este caso se va a escoger el que se encuentra en el recuadro como se muestra en la figura 161.

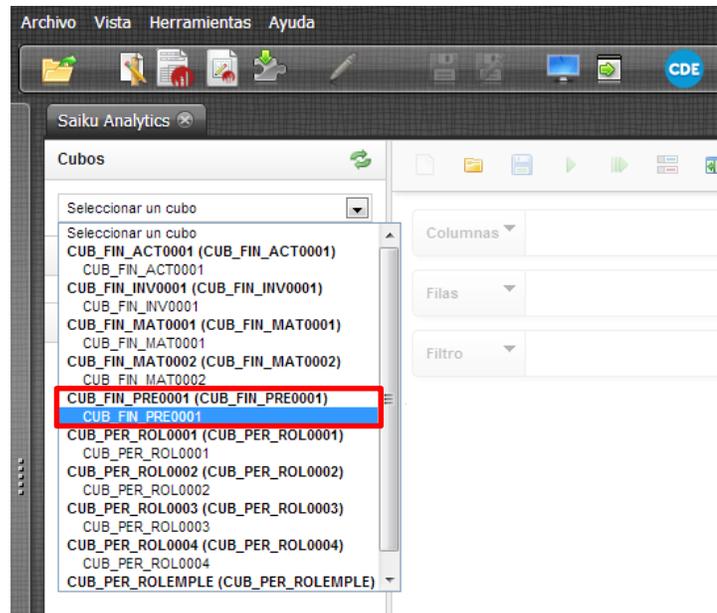


Figura 161: Listado de cubos publicados

Se desplegará la siguiente área de trabajo.

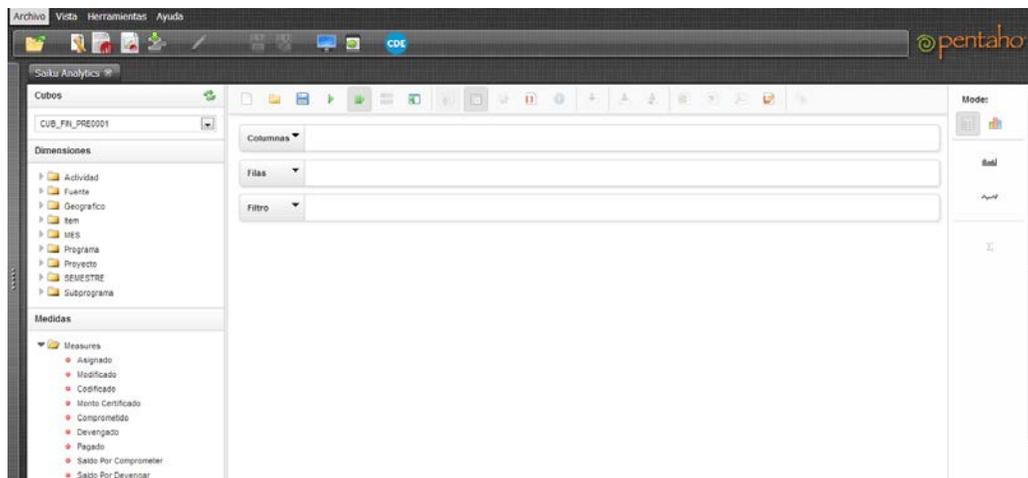


Figura 162: Area de trabajo para un cubo

Como se puede apreciar en la figura 162, al abrir un cubo se mostrarán las dimensiones y medidas existentes, las mismas que se puede ubicar como filas o columnas según las necesidades del usuario. Además se pueden realizar filtros de cualquier campo existente en las dimensiones como se observa en la figura 163.

The screenshot shows the Pentaho Saiku Analytics interface. The 'Filtro' section is set to 'MES Q1'. A red arrow points to the 'MES' dimension in the left sidebar. The main table displays data for various activities across different months.

Actividad	530104	530105	530106	530201	530204	530207	530209	530219	530299	530301	530302	530303	530304	530402	5304
CC. DE LA COMPUTACION		0			0		0	0	0	0	0	0	0	0	
CC. DE LA ENERGIA Y MECANICA		0		0	0	0			0	0	0	0	0		
CC. DE LA TIERRA		0	0	0	0				0	30,72	0	0	0		
CC. DE LA VIDA	12.230,78	650,38			0	0			0	68,16	0	80	0		
CC. ECONOMICAS Y ADMINISTRATIVAS		0			0	0			0	0	0	0	0	0	0
CC. EXACTAS		0			0	0			0	0	0	0	0		
CC. HUMANAS		49,44			0	0			0	26,00	0	1.200	0		

Figura 163: Manipulación del cubo

El reporte generado muestra como columnas los distintos ítems existentes y como filas las actividades correspondientes a los departamentos.

El reporte fue generado para los meses de: enero, febrero y marzo de 2013 y el filtro se realizó arrastrando el campo mes a la sección de filtro como se ilustra en la figura 163 y posteriormente se selecciona los meses requeridos como se muestra en la figura 164:

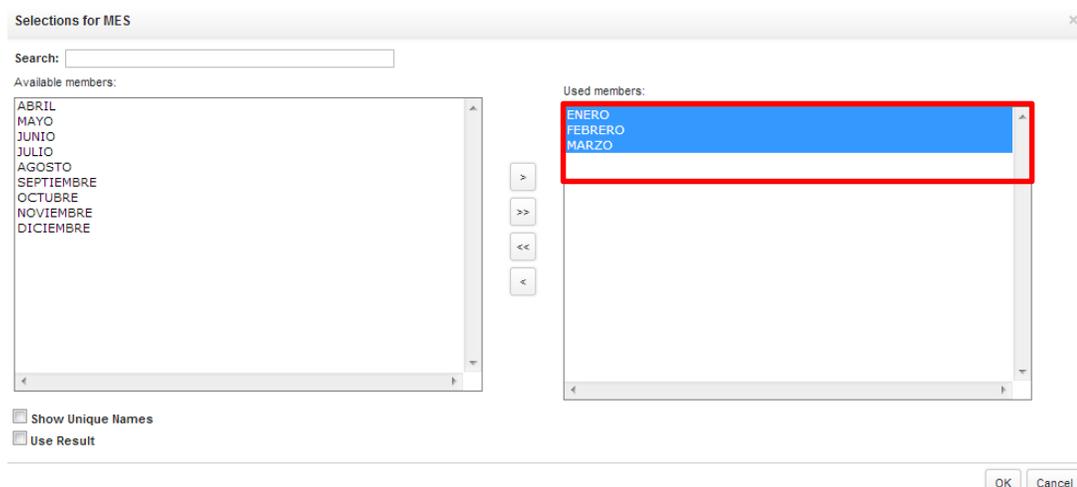


Figura 164: Filtro del campo mes

Para que las actividades que se muestran correspondan únicamente a los departamentos, simplemente hacer clic en la fila actividad y seleccionar las actividades requeridas como se muestra en la figura 165.

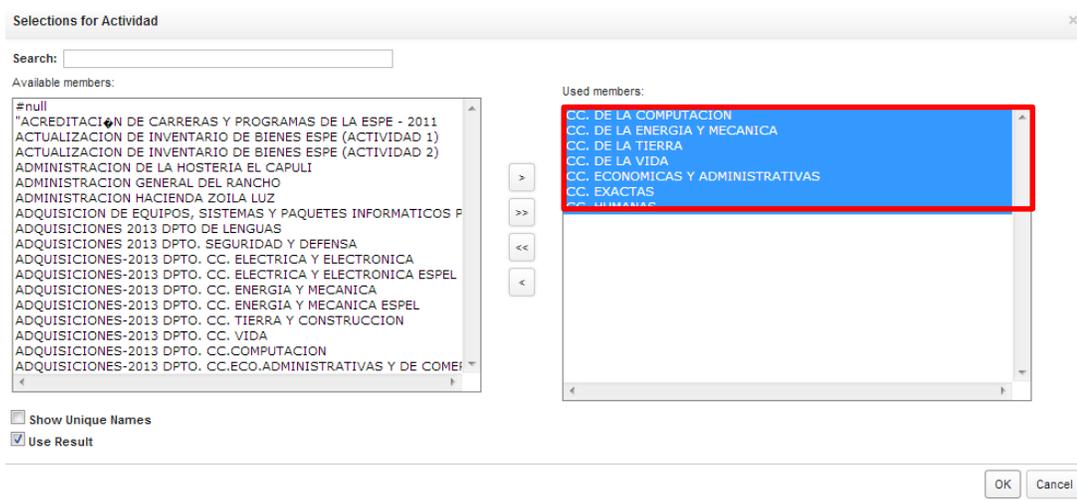


Figura 165: Filtro del campo actividad

Una vez que se visualizó un ejemplo de un cubo, también se puede en la parte derecha del área de trabajo optar por visualizar gráficamente los resultados obtenidos en el reporte. Se puede seleccionar distintos tipos de gráficos como se observa en la figura 166 y se escoge según las necesidades del usuario.

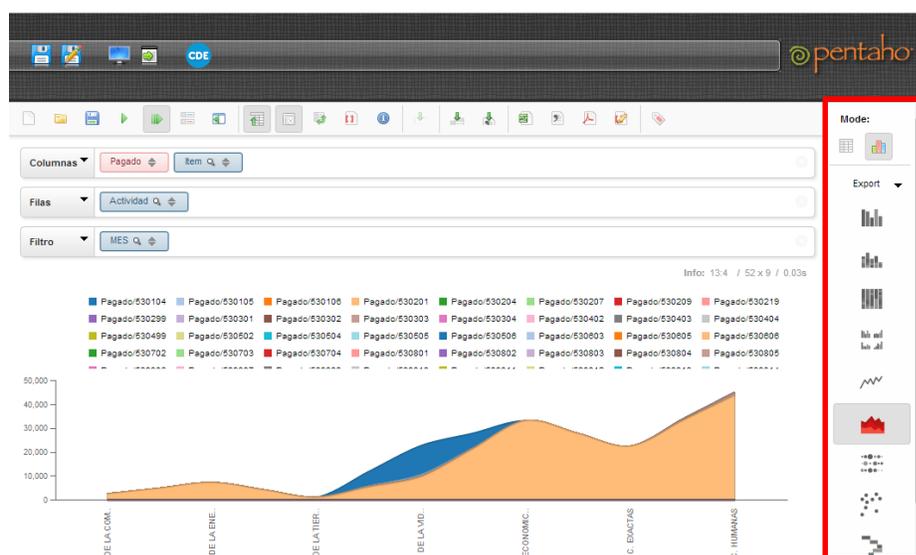


Figura 166: Manejo de gráficos en saiku

El reporte generado puede ser exportado a distintos formatos (XLS, CSV, PDF), los mismos que se muestra en la barra de tareas del área de trabajo.

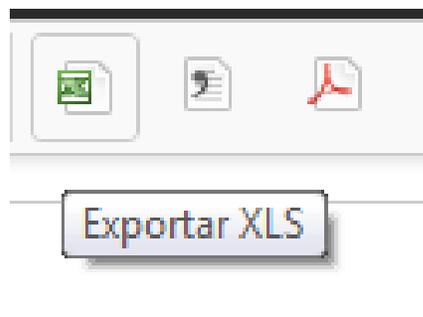


Figura 167: Iconos de formatos de exportación

Al hacer clic en cualquiera de los tres íconos, el reporte se exportará automáticamente en el formato seleccionado.

4.11 MANUAL DE USUARIO DE REPORT DESIGNER

4.11.1 Introducción

La herramienta de pentaho cuenta con una aplicación independiente que se encarga únicamente de la creación de reportes que son una parte indispensable para un análisis de negocio y esta aplicación se llama pentaho report designer (PRD). Esta aplicación tiene una interfaz gráfica que permite crear, dar un formato personalizado y publicar los reportes para la visualizar en la consola de usuario.

4.11.2 Proceso de creación de un reporte

Para realizar este manual se va a tomar como ejemplo un reporte que se creó de acuerdo a un requerimiento del usuario en donde se pide del módulo de personal, los roles de pago por rubros y además se da una breve explicación de la creación de un gráfico dentro de un reporte.

Primero se inicia el pentaho report designer con unos archivos que se encuentran en la carpeta que se descarga en el internet.

Para Windows se ejecuta el archivo report-designer.bat para iniciar el report designer mientras que para Linux se ejecuta en un terminal el archivo report-designer.sh, con esto carga y se abre la aplicación.

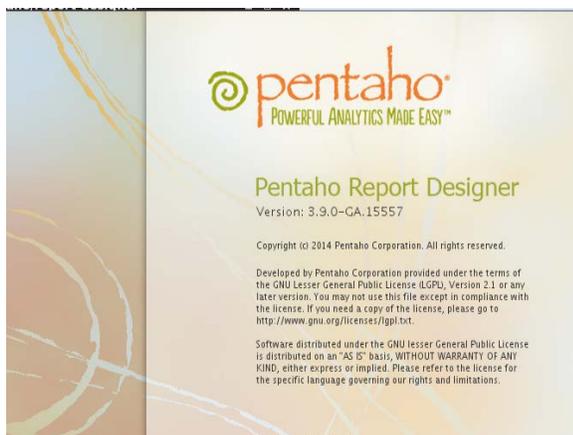


Figura 168: Carga Report Designer

Como requisito indispensable para la creación del reporte es que debe tener una serie de parámetros obligatorios. Para esto se va a utilizar una plantilla diseñada con los estándares establecidos por la Universidad de las Fuerzas Armadas ESPE.

Una vez abierto el report designer, hacer clic en **file**.



Figura 169: Interfaz report designer

Ahora se va a desplegar una serie de opciones en la cual se va hacer clic en **open**.

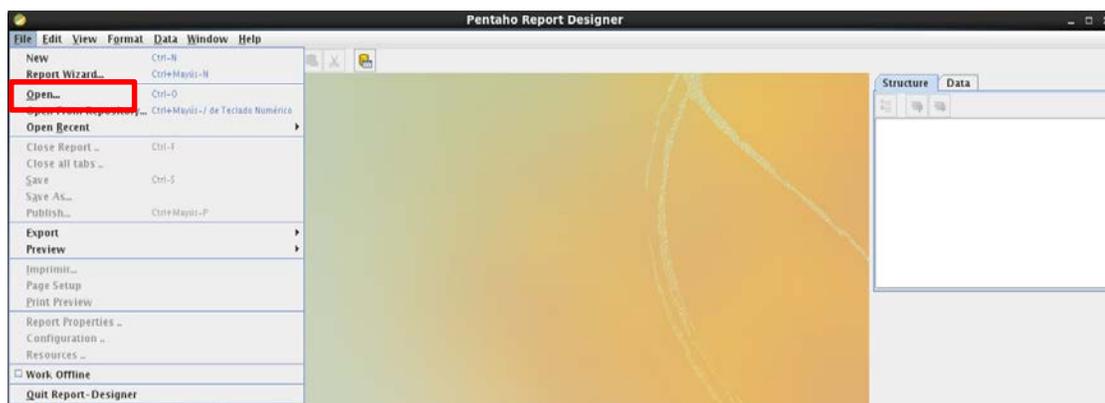


Figura 170: Menú para abrir una plantilla

Se va abrir una ventana en donde se va a escoger la plantilla ya definida en extensión .prpt y hacer clic en **abrir**.

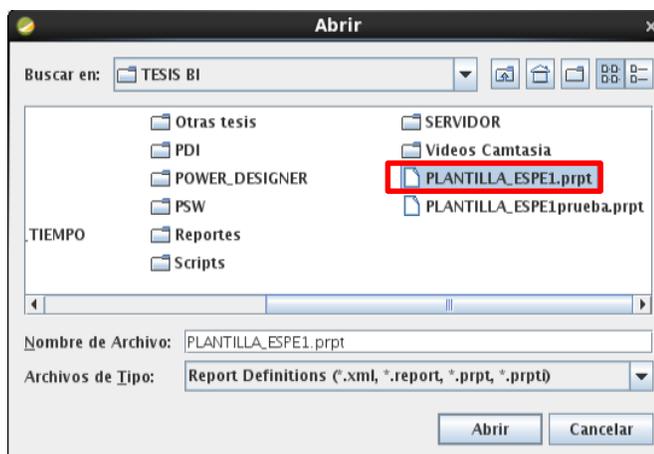


Figura 171: Escoger y abrir la plantilla

La plantilla se abre y se muestra como la figura 172, ahora en la parte derecha se debe ubicar en la pestaña **data**, hacer clic derecho sobre **data sets** y finalmente escoger la opción **JDBC**.

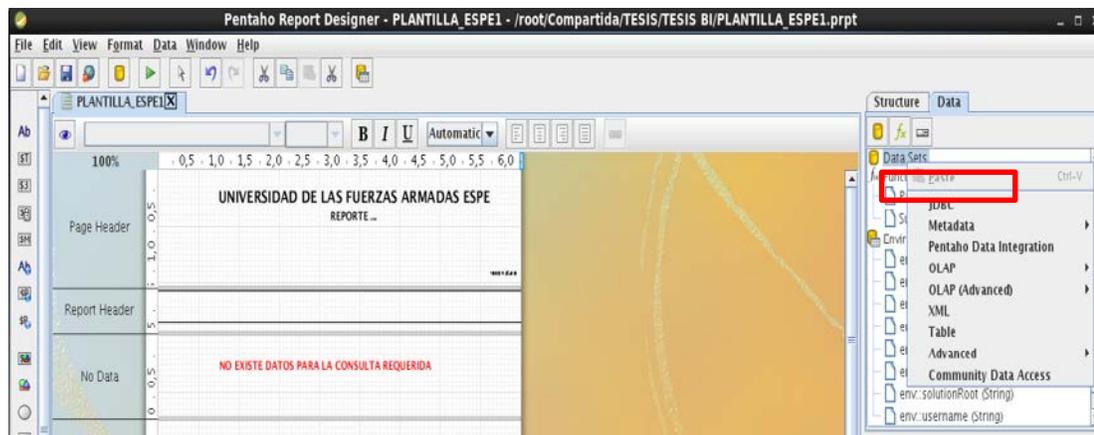


Figura 172: Crear una fuente de datos JDBC

Se abrirá una nueva ventana en donde se va a realizar primero una conexión JDBC, para esto se ubica en la parte superior izquierda y hacer clic en el símbolo que está en el recuadro.

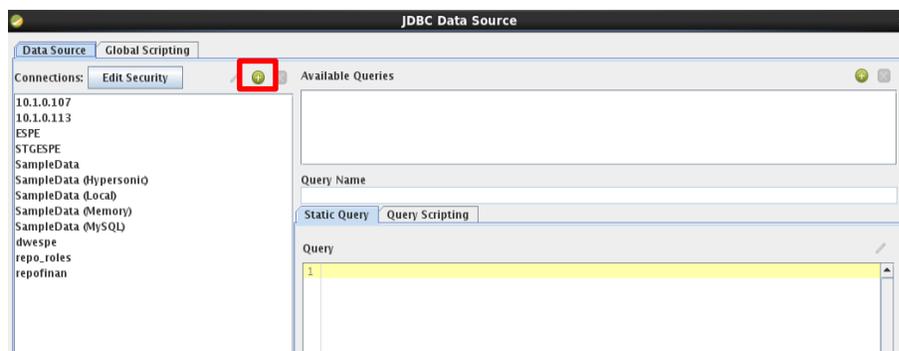


Figura 173: Añadir nueva conexión

En esta nueva ventana, para este caso, como connection type se va a escoger Oracle y en la parte inferior en Access se escoge native (JDBC) y con esto llenar todos los campos necesarios como se observa en la figura 174 para tener una conexión satisfactoria. Si necesita verificar que la conexión esta correcta, hacer clic en **test** y aparecerá una mensaje como se muestra en la figura 175 caso contrario corregir las configuraciones de la conexión y por ultimo hacer clic en **ok**.

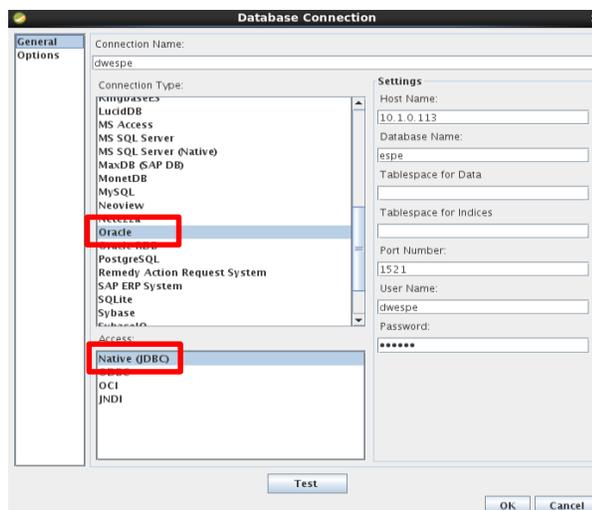


Figura 174: Configuración de la conexión

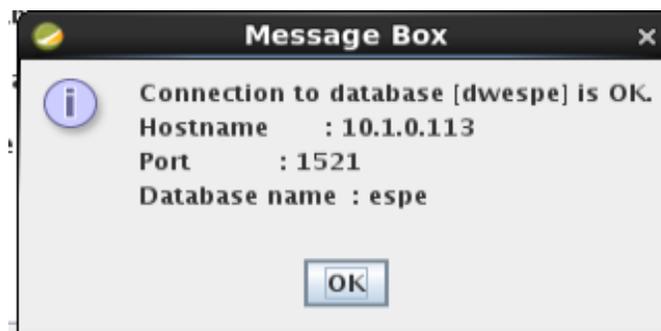


Figura 175: Mensaje de conexión satisfactoria

Una vez creado la conexión, ahora se va a ubicar en la parte superior derecha y hacer clic en el símbolo que se encuentra en el recuadro. Se va a crear un query en donde, primero se va a poner un nombre que lo identifique al query en el campo de **query name** y la sentencia sql con la información que necesita el usuario obtener para el requerimiento pedido en el campo de **query**. Para el caso del reporte que se va a explicar, se va a realizar el mismo procedimiento y se va a añadir 4 queries más que van a servir para los parámetros requeridos que son: año, mes, tipo de empleado y tipo de rubro.

En el caso de verificar si los queries se encuentran correctos, hacer clic en **preview** y se va a mostrar una ventana como la figura 177, caso contrario debe arreglar el query y al final hacer clic en **ok**.

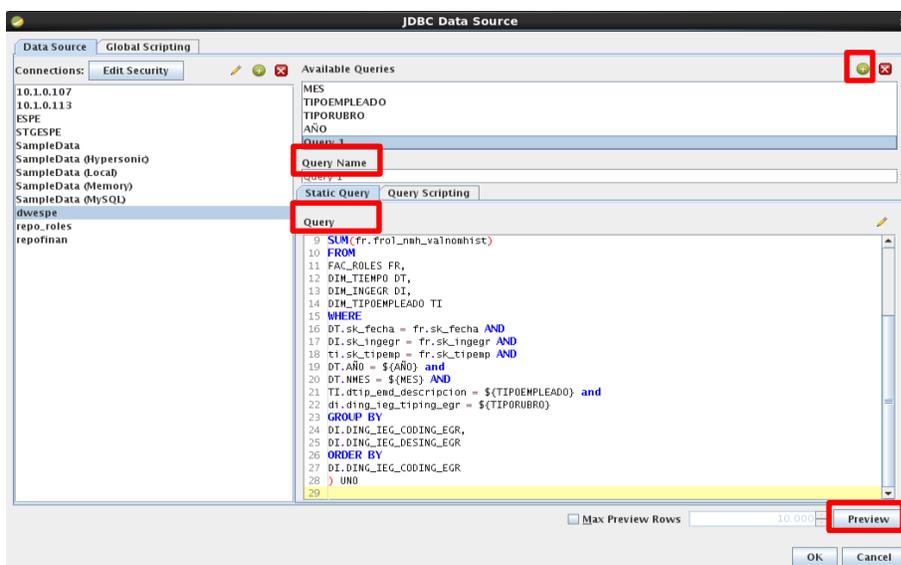


Figura 176: Crear queries para el reporte



Figura 177: Previsualizar información de un query

4.11.3 Creación de parámetros

Ahora se va a crear parámetros que son requisitos para el reporte, para esto se ubica en la parte derecha en la pestaña data, hacer clic derecho sobre parameters y escoger **add parameter**.

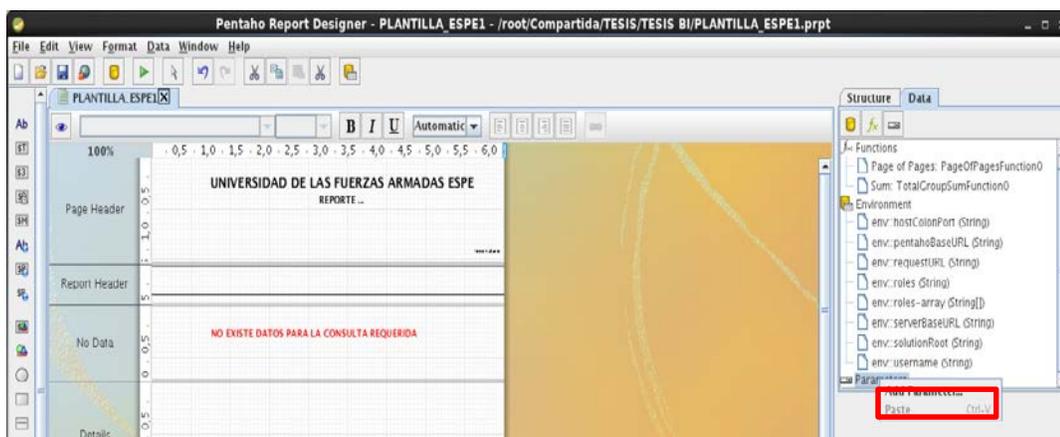


Figura 178: Añadir parámetros en el reporte

En la nueva ventana que se abre, en este caso, se llena los campos para el parámetro de año como se muestra en la figura 179 y hacer clic en **ok**.

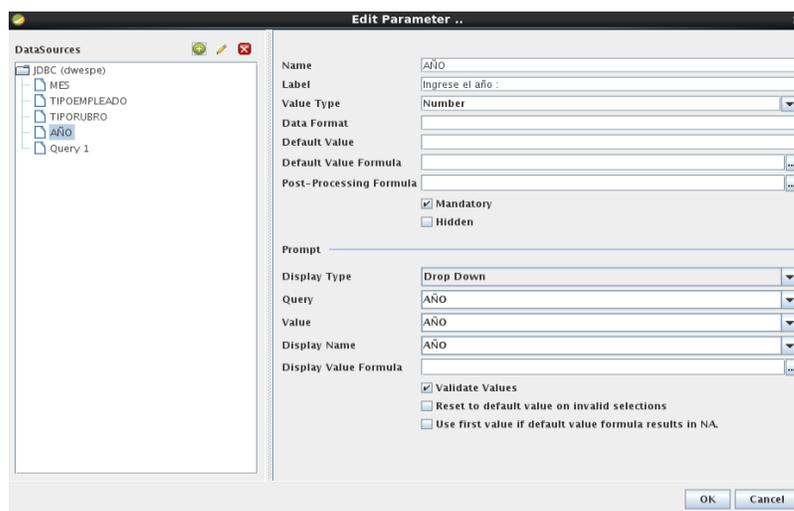


Figura 179: Configuración de un parámetro

Realizar el mismo procedimiento de los dos pasos anteriores para los parámetros requeridos que son los que se encuentran en el recuadro de la figura 180.

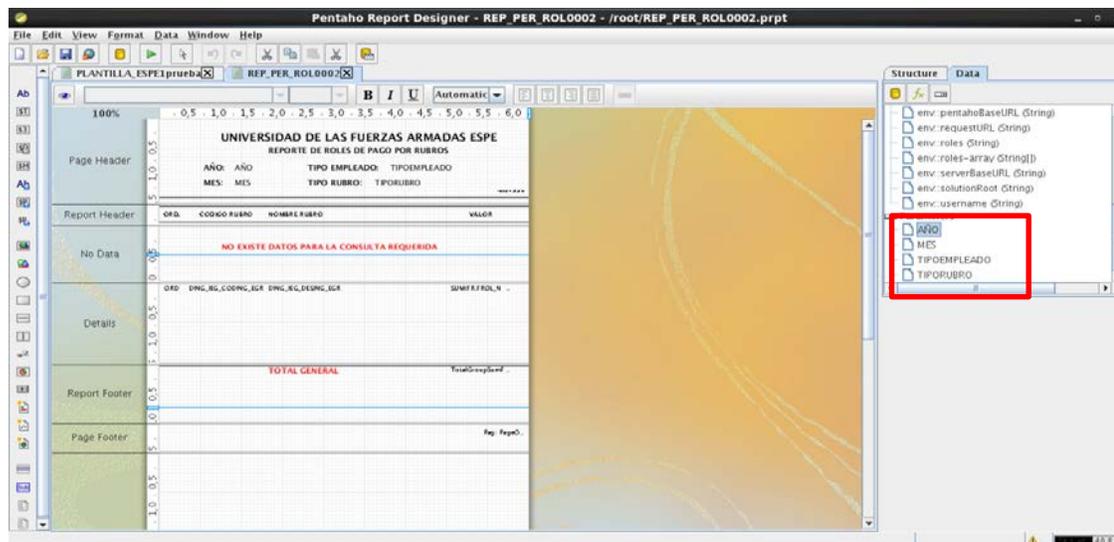


Figura 180: Interfaz de parámetros creados

4.11.4 Diseño de reporte

4.11.4.1 Report Header

El report header es el espacio donde se colocan las cabeceras, para esto se ubica en la parte superior izquierda, hacer clic sin soltar al símbolo **label** que está identificado con el recuadro en la figura 181 y se arrastra para la parte de report header de la estructura del reporte.

Una vez que el label se encuentra en la report header se le va a dar un nombre, para esto se selecciona al label, luego se ubica en la parte superior derecha en la pestaña structure, después en la parte inferior derecha se ubica en la pestaña attributes y en el atributo value se coloca el nombre, en este caso se puso ORD.

Si se requiere se puede hacer los cambios necesarios al label en la pestaña style y attributes.

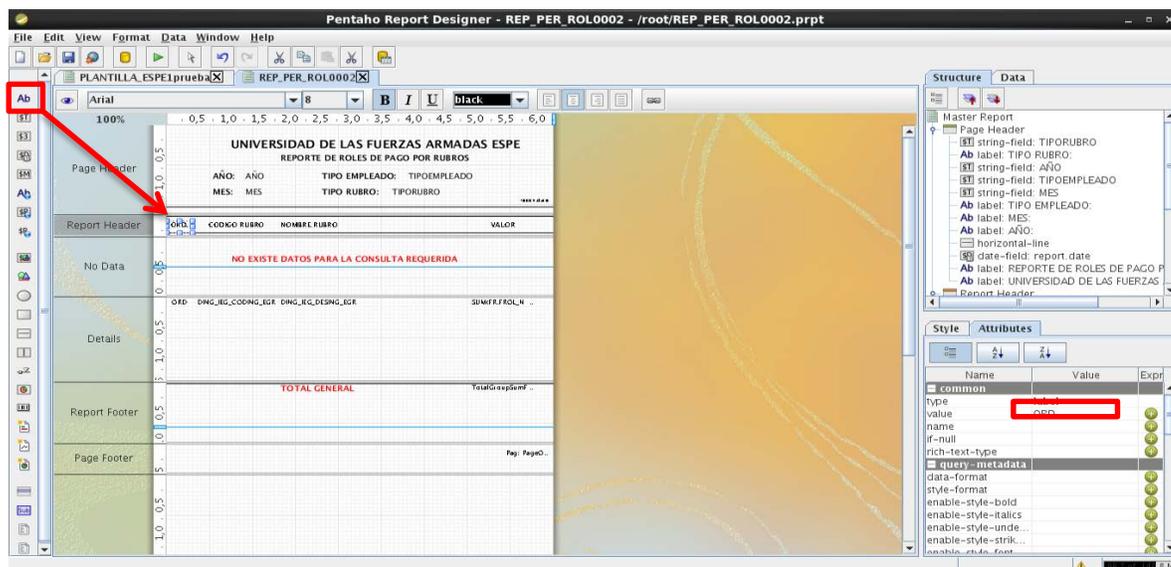


Figura 181: Añadir etiquetas en report header

4.11.4.2 Details

Details es el espacio donde se coloca los detalles que se requiere para el reporte, para esto se ubica en la parte superior derecha en la pestaña data, luego hacer clic sin soltar a los campos del query principal donde se obtiene la información para el reporte, estos se encuentran en el recuadro de la figura 182 y arrastrar hacia la parte details de la estructura del reporte.

Para realizar algún cambio de formato a los detalles, se ubica en la pestaña de structure, luego en la pestaña de style, de igual forma como se hace para las cabeceras.

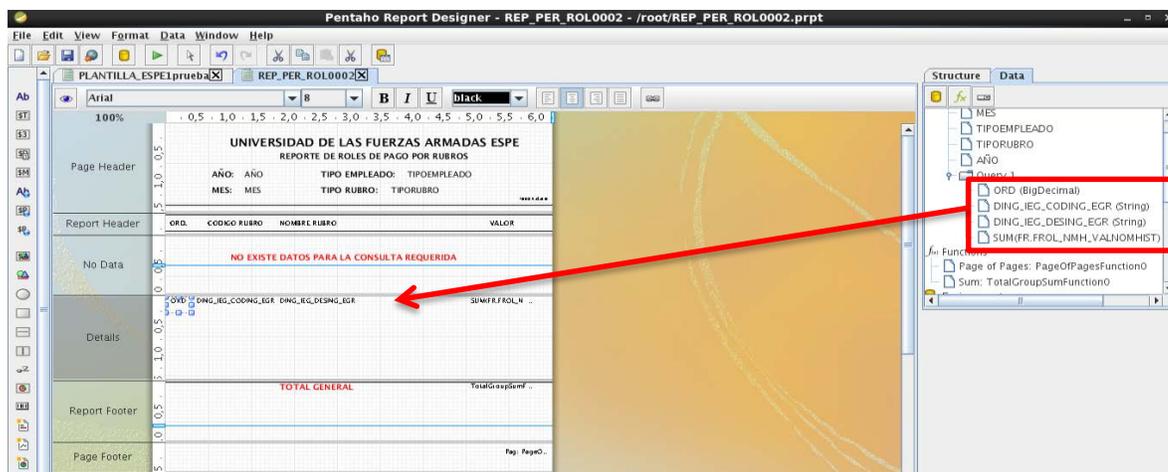


Figura 182: Añadir campos a details

4.11.4.3 Report Footer

El report footer es el espacio para el pie de página del reporte, en este caso se va a colocar un total general de todo el detalle. Para esto se ubica en la parte superior derecha en la pestaña data, ahí se puede obtener varias funciones que ofrece la herramienta y para este caso se tomara la función **sum** para obtener el total general. Entonces se hace clic en la **función sum** sin soltar y se le arrastra hacia el report footer de la estructura del reporte. Además se debe ubicar en la parte inferior derecha y en el valor del campo field name, se va a seleccionar el campo que necesita que sume como se muestra en el recuadro de la figura 183.

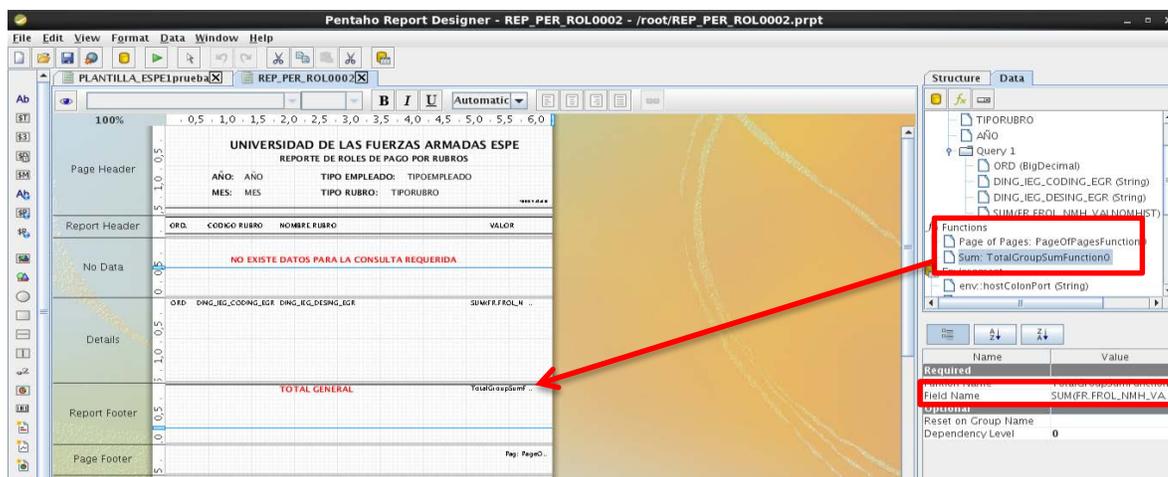


Figura 183: Añadir funciones en report footer

4.11.5 Diseño de parámetros

Ahora se va a colocar etiquetas para los parámetros en la cabecera de página para poder visualizar los filtros al momento de ejecución del reporte. Primero se arrastra un **label** hacia el page header de la estructura del reporte como se muestra en la figura 184, luego se coloca un nombre con el que se le identifique a la etiqueta del parámetro. El nombre se lo pone en la parte inferior derecha como se muestra en el recuadro, en este caso es **AÑO** y de la misma forma se hace para los parámetros restantes.

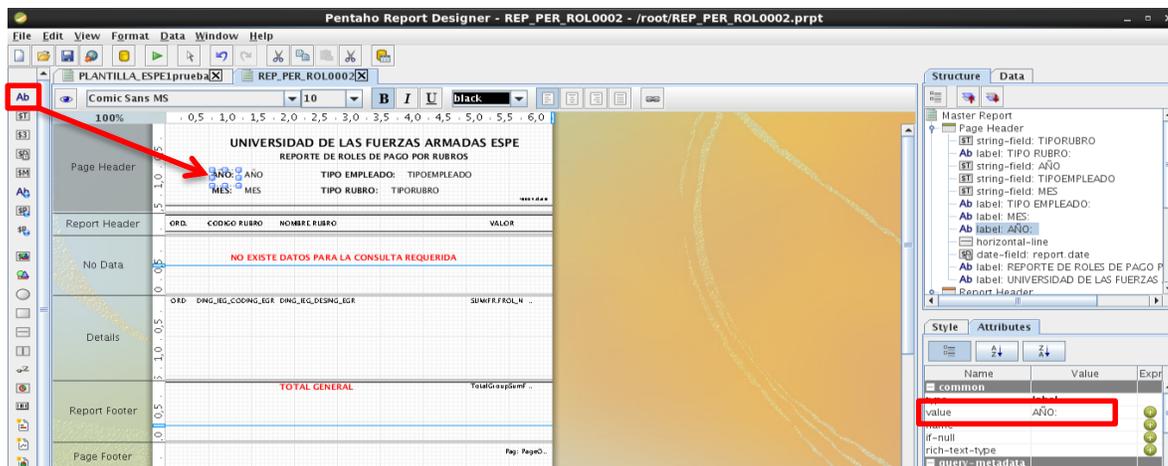


Figura 184: Añadir etiquetas en page header

Una vez puestas las etiquetas con sus nombres, se procede a colocar los valores de los parámetros en el page header. Primero se ubica en la parte derecha en la pestaña data y en la sección que dice parameters se encuentran los 4 parámetros que se utilizan para el reporte, se coge y se arrastra cada uno de ellos hacia el page header como se observa en la interpretación de la figura 185.

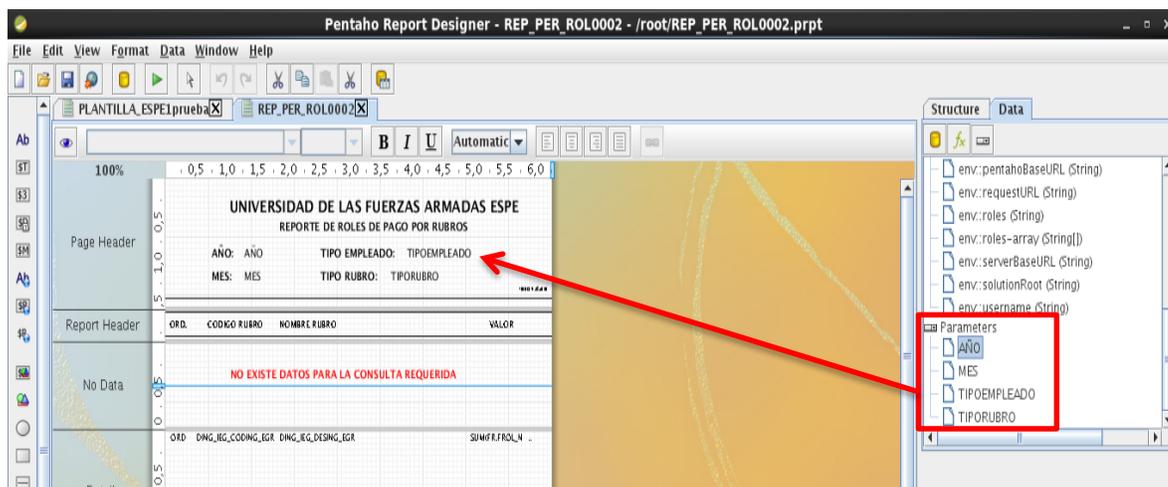


Figura 185: Añadir parámetros en page header

4.11.6 Nombre del reporte

Una vez realizado todo lo explicado, finalmente se coloca un nombre al reporte y con eso estaría completo. Para esto, seleccionar la etiqueta donde se coloca el nombre del reporte que viene definida en la plantilla y se encuentra en page header de la estructura del reporte, luego se ubica en la pestaña de structure seguido de la pestaña attributes que se encuentran en la parte derecha y en el atributo value colocar el nombre que requiera el usuario, para este caso es **reporte de roles de pago por rubros**.

El reporte está completo, ahora se ubica en la parte superior izquierda y hacer clic en el símbolo que se encuentra en el recuadro como se observa en la figura 186 para visualizar el reporte.

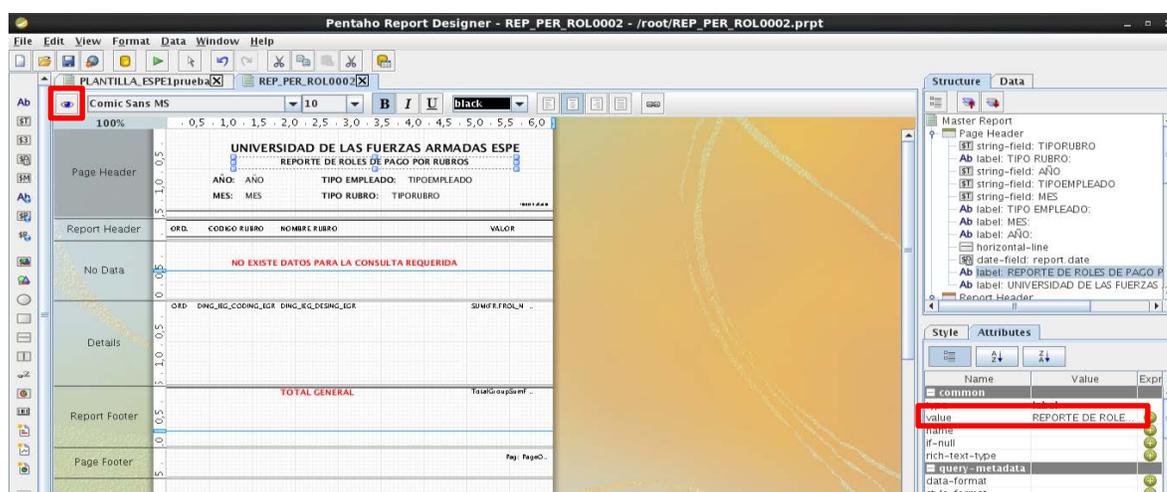


Figura 186: Poner nombre al reporte

4.11.7 Previsualizar el reporte

Finalmente la figura 187 muestra la visualización del reporte, donde se escoge, primero los 4 parámetros establecidos por los requerimientos del usuario y luego se va a ejecutar para que se despliegue el detalle del reporte.

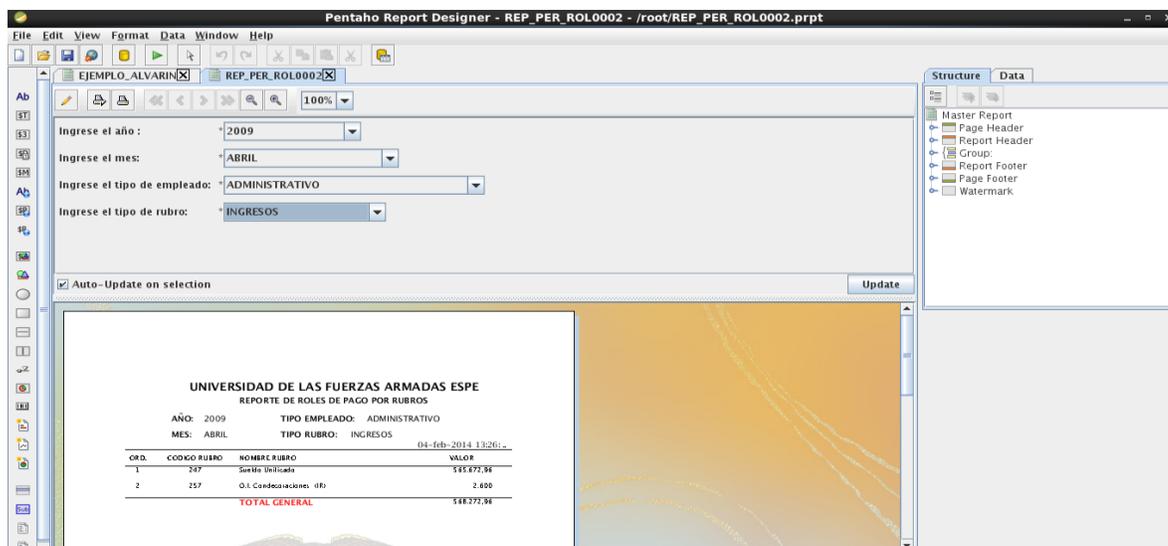


Figura 187: Reporte pre visualizado

4.11.8 Guardar y publicar el reporte

Una vez terminado el reporte y visualizado, ahora se va a publicar para que se pueda visualizar en la consola de usuario. Para esto se ubica en file y escoger la opción **publish**.

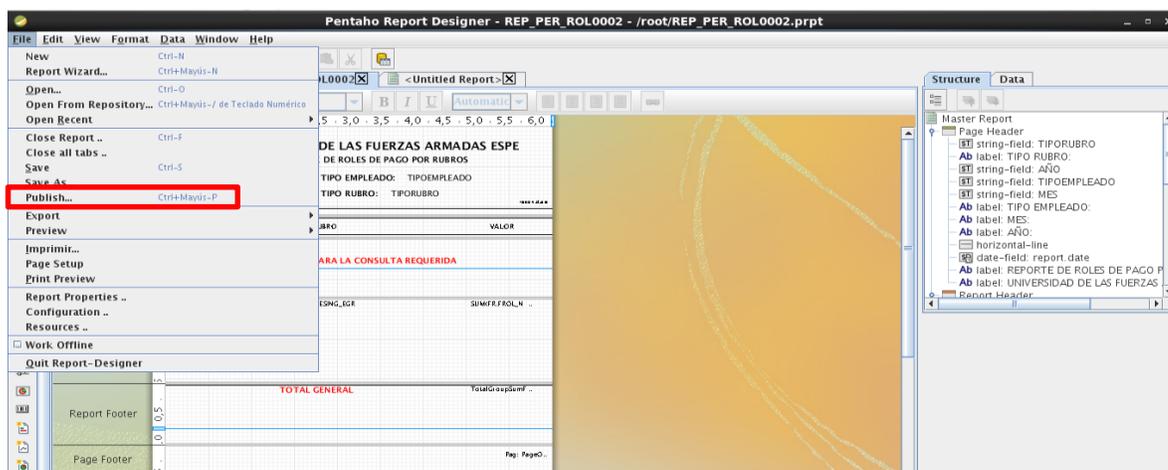


Figura 188: Menú para publicar el reporte

Al momento que se quiere publicar, aparecerá un mensaje para guardar los cambios que se han hecho en el reporte como se muestra en la figura 189, hacer clic en **sí**.

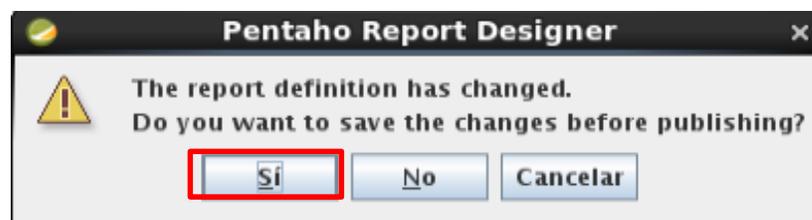


Figura 189: Mensaje para guardar reporte

Ahora se va a guardar el reporte, se lo puede guardar donde el usuario requiera y se debe guardar con un nombre que cumpla los estándares que ya están definidos, en este caso el reporte se llamará **REP_PER_ROL002.prpt**, luego hacer clic en **guardar**.



Figura 190: Guardar reporte

Se va a abrir una nueva ventana donde se llena los siguientes campos como se puede ver en la figura 191 y luego hacer clic en **ok**.

- **URL:** Se coloca la dirección del servidor.
- **USER:** Se pone un usuario que este creado en la consola de administración de pentaho.
- **PASSWORD:** La contraseña del usuario.



The image shows a 'Login' dialog box with the following fields and controls:

- Server:**
 - URI:
 - Timeout:
- Pentaho Credentials:**
 - User:
 - Password:
- Remember These Settings
- OK button
- Cancel button

Figura 191: Conexión al repositorio

En esta nueva ventana, pedirá que llene los siguientes campos para poder publicar el reporte a la consola de usuario.

- **FILE NAME:** Es el nombre con el que se guardó el archivo
- **TITTLE:** Se puede poner cualquier nombre con el que identifique el título del reporte
- **REPORT DESCRIPTION:** Se puede poner cualquier tipo de descripción acerca de lo que muestra el reporte.
- **PUBLISH PASSWORD:** La contraseña de las publicaciones.

Además de llenar esos campos, se debe colocar una dirección donde se va a guardar la publicación del reporte, esto ya está definido de igual forma por los estándares y finalmente hacer clic en **ok**.

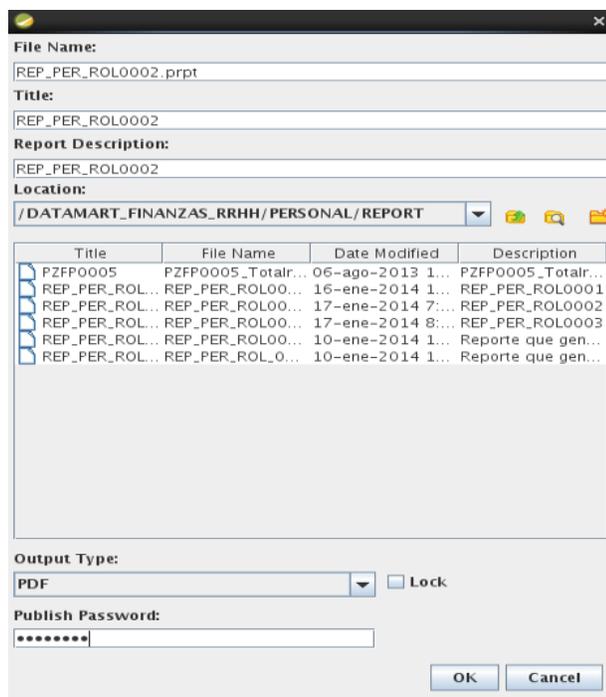


Figura 192: Configuración para publicar el reporte

La figura 193 muestra un mensaje que indica que el reporte fue publicado satisfactoriamente y ya se va a poder visualizar en la consola de usuario.

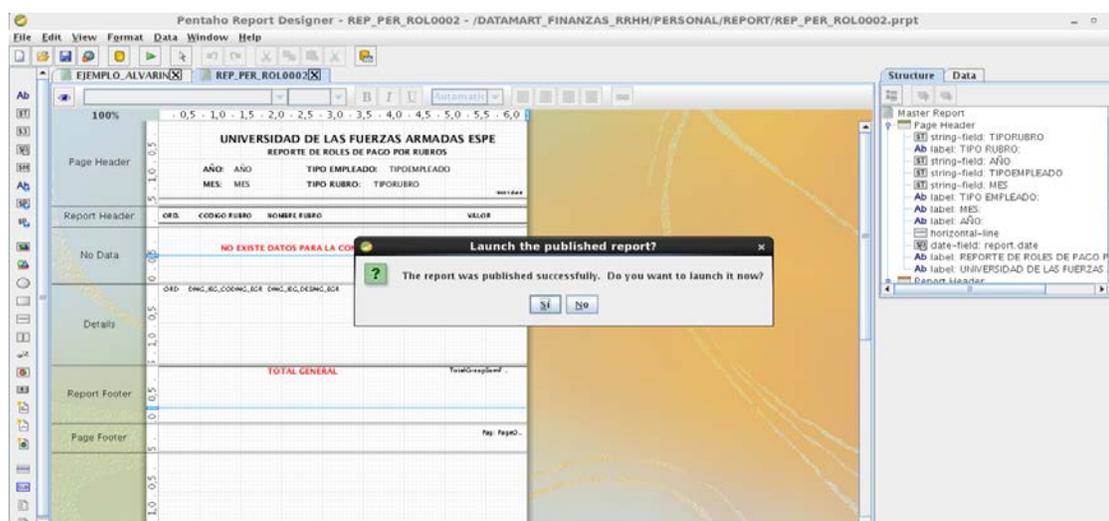


Figura 193: Mensaje de reporte publicado

4.11.9 Crear gráfico dentro de un reporte (OPCIONAL)

En caso de realizar un gráfico, aquí se hará una breve explicación para crear un gráfico dentro de un reporte. Como primer paso, se ubica en la parte izquierda donde se encuentra la barra de herramientas y hacer clic sin soltar en **subreport** que es el símbolo que está en el recuadro, luego arrastrar hacia el report footer de la estructura del reporte como se ilustra en la figura 194.

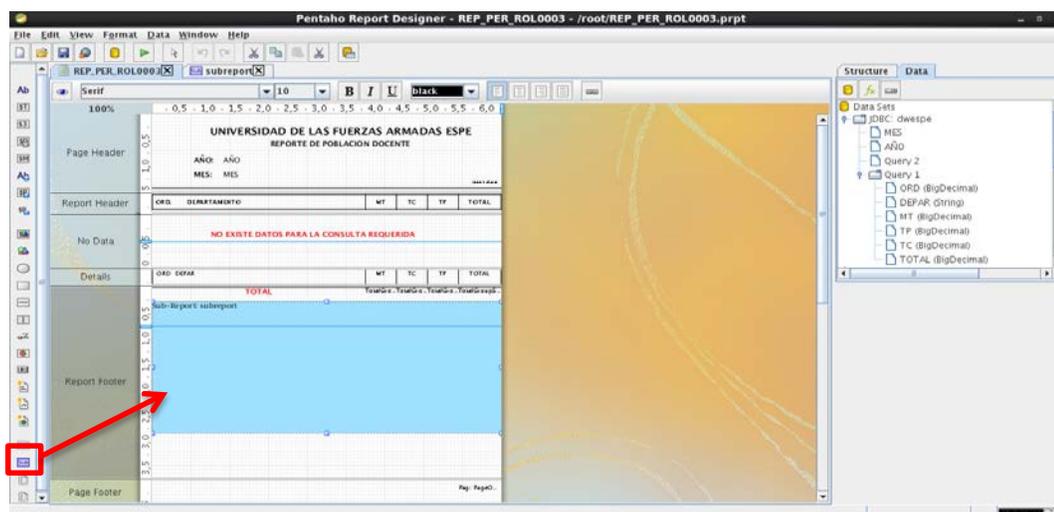


Figura 194: Añadir subreport

Al crearse el subreport se hace un cuadro en color celeste como se observa en la figura 194, ahora hacer doble clic sobre el espacio celeste y se abrirá una nueva ventana del subreport.

Una vez que se tiene la estructura del subreport, hacer clic en **chart**, que es el símbolo que se encuentra en el recuadro y arrastrar hacia el report footer como se ilustra en la figura 195.

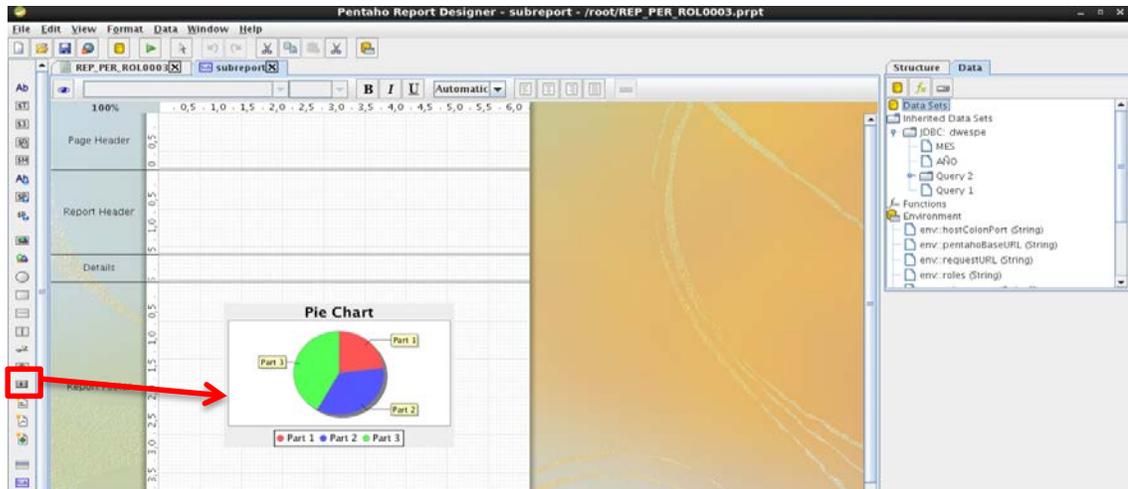


Figura 195: Añadir gráfico al reporte

Ahora se ubica en la parte derecha en la pestaña data, hacer clic derecho sobre el query que tenga la información que el usuario necesite para la creación del gráfico y escoger la opción **select query**.

Una vez que escogió el query que va a usar para el gráfico, hacer doble clic sobre el gráfico que está en el report footer de la estructura del subreporte.

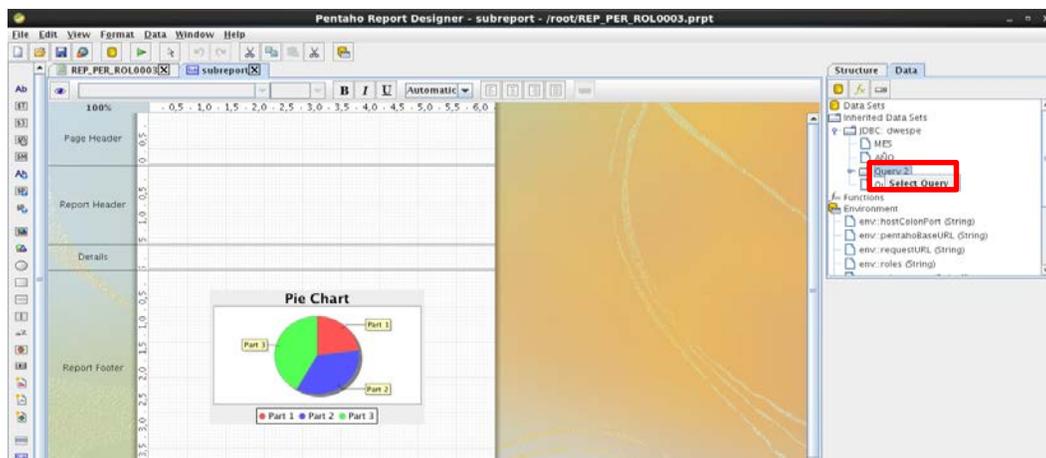


Figura 196: Seleccionar query para el gráfico

Se abre una nueva ventana, se escoge un tipo de gráfico, para este caso hacer clic en el **gráfico de pastel**, luego se ubica en la parte derecha y hacer clic en valor de la propiedad de **series-by-field** como se muestra en el recuadro de la figura 197.

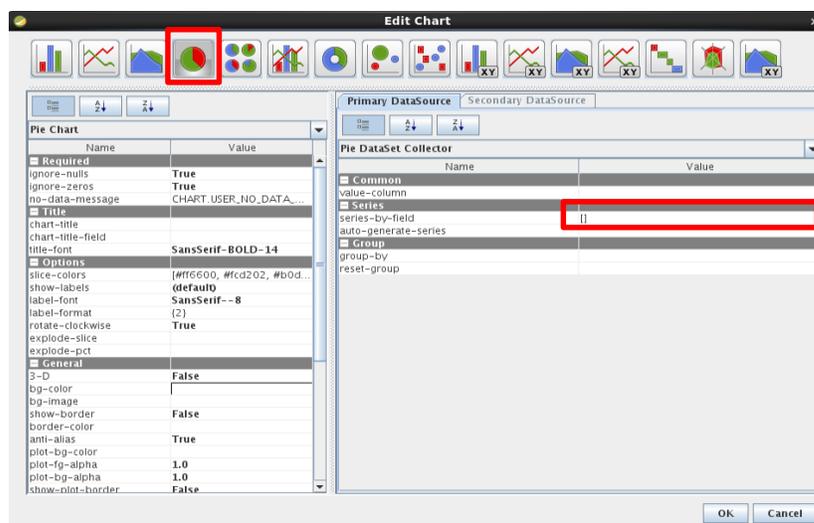


Figura 197: Selección tipo de gráfico

En esta nueva ventana, hacer clic en el símbolo que se encuentra en el recuadro de la figura 198 y en el nuevo espacio que aparecerá debajo de value, hacer clic sobre el espacio y escoger el campo que el usuario necesite que se muestre como el campo por cada serie, en este caso se escogió **DEDICACIÓN** y por último hacer clic en **ok**.



Figura 198: Añadir campos por serie

Ahora se ubica en la parte izquierda y se pone un título al gráfico, luego en la parte derecha se pone el campo que será el detalle para el gráfico como se observa en los recuadros de la figura 199 y por último hacer clic en **ok**.

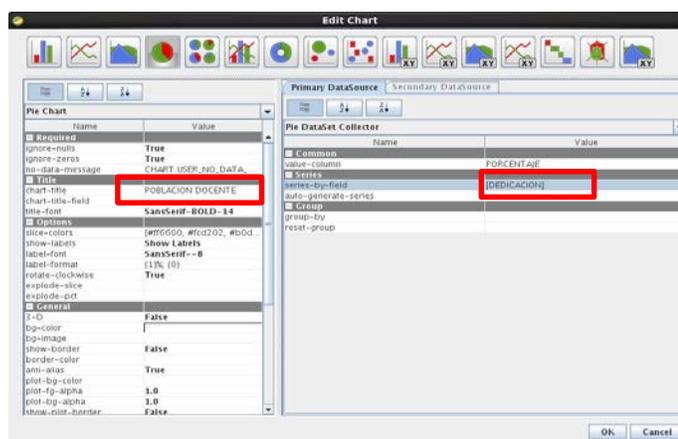


Figura 199: Configuración del gráfico

4.11.9.1 Parámetros para el subreport

Una vez realizado todos los ajustes para el gráfico, ahora se va a importar los parámetros que se toman en cuenta en el reporte general. Para esto se ubica en la parte derecha en la pestaña data y hacer clic en el símbolo que se encuentra en el recuadro de la figura 200.

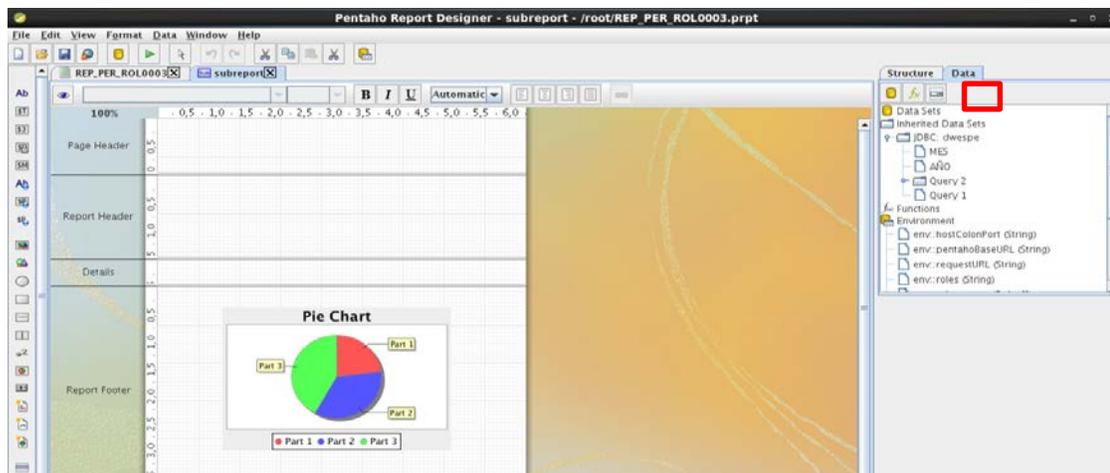


Figura 200: Añadir parámetros al subreport

Se abre una ventana nueva que es para los parámetros del subreporte en donde se va a ubicar en la parte izquierda en la sección de importar parámetros y hacer clic en el símbolo que se encuentra en el recuadro de la figura 201.

Ahora aparecerá un espacio en blanco debajo de inner name y outer name, en la columna de outer name, hacer clic y escoger el parámetro que necesita importar al subreporte mientras que en la columna de inner name, hacer clic y escoger el campo que se relacione con el parámetro del outer name, en este caso se importó dos parámetros que son: **año** y **mes**.

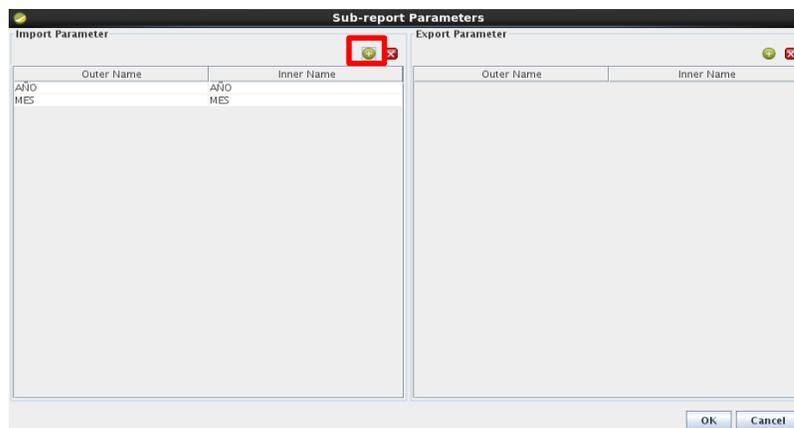


Figura 201: Importar parámetros

Una vez que se realice todo lo explicado referente a la creación del gráfico, se pre visualiza el reporte para comprobar que el gráfico se muestra como requiere el usuario, en este caso el gráfico se muestra como la figura 202. Con esto se finaliza la explicación de la creación de un gráfico dentro de un reporte.

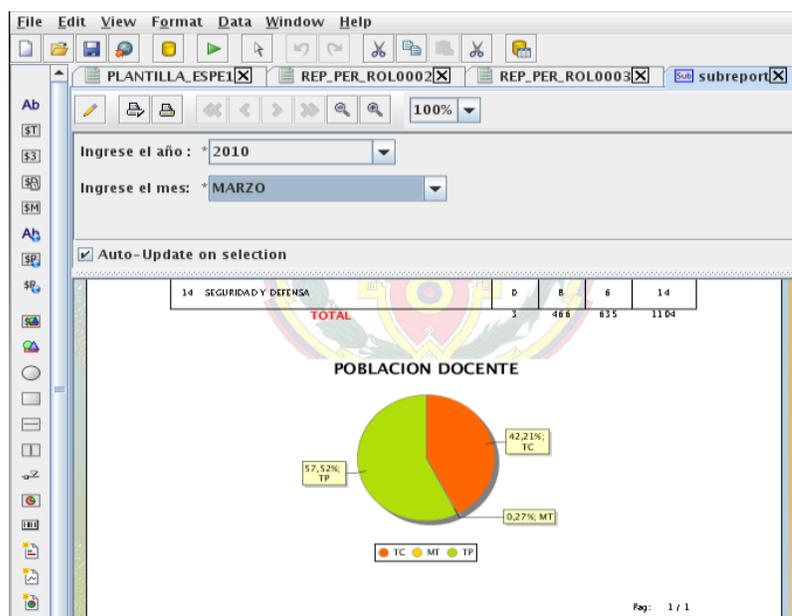


Figura 202: Gráfico pre visualizado

4.11.10 Uso del reporte en la consola de usuario

Para poder visualizar el reporte realizado como primer paso, ingresar a la siguiente dirección web: siespe.espe.edu.ec, preferiblemente usando Mozilla Firefox o Google Chrome. Se desplegará la siguiente pantalla:

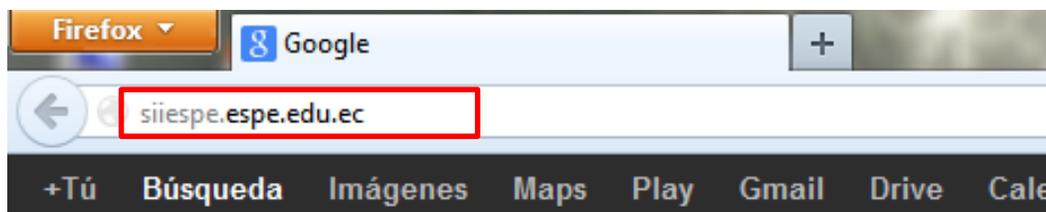


Figura 203: URL de la consola de usuario

Se abre la consola de usuario, se escribe el usuario y el password asignado a cada usuario por el administrador. Un ejemplo puede ser usuario: **joe** y password: *********



Figura 204: Interfaz de login de la consola de usuario

Una vez abierto la consola de usuario, se ubica en la parte izquierda donde se encuentran las carpetas para navegar, hacer clic en la carpeta **report** que está en el recuadro como se observa en la figura 205 y en la parte inferior van aparecer los reportes publicados, para visualizar el reporte realizado, hacer doble clic en el reporte que se encuentra en el recuadro.

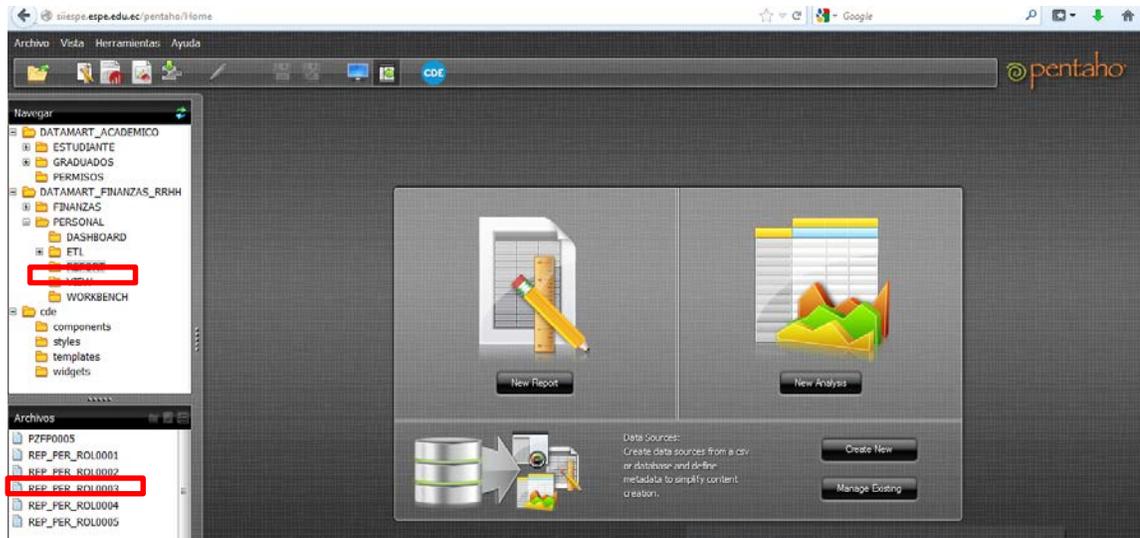


Figura 205: Búsqueda del reporte publicado

Se abre el reporte y se va a observar las 4 cajas de texto de los parámetros que tiene el reporte. Esto se necesita llenar para que el reporte se ejecute y muestre el detalle.

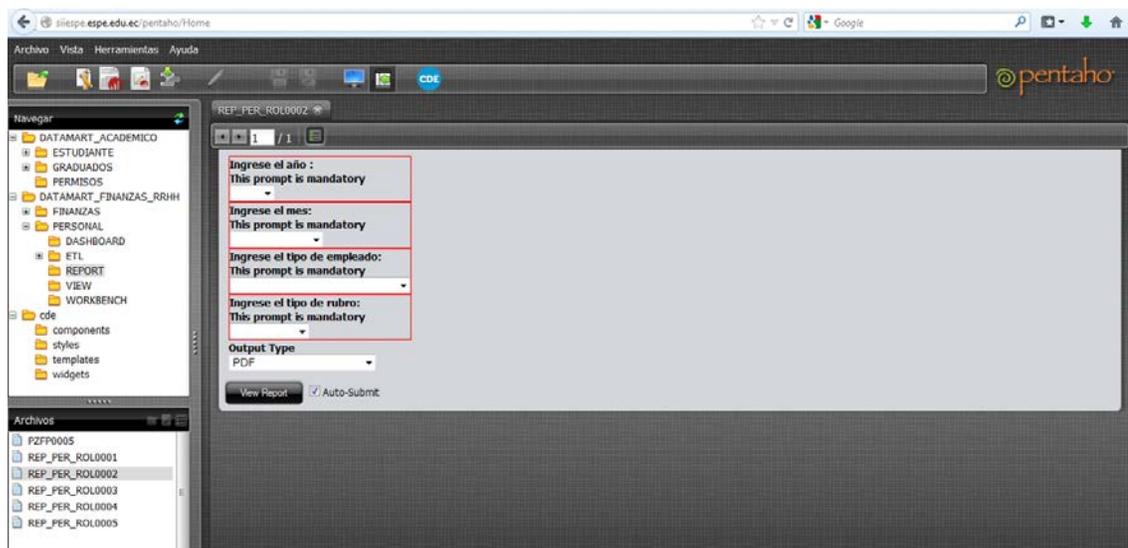


Figura 206: Reporte abierto

Se escoge una opción que el usuario requiera para cada parámetro y luego escoge el tipo de salida como se puede ver en el recuadro de la figura 207.

Output Type

- PDF
- TML (Paginated)
- TML (Single Page)
- Excel
- Excel 2007
- Comma Separated Value
- Rich-Text-Format
- Text

UNIVERSIDAD DE LAS FUERZAS ARMADAS ESPE
 REPORTE DE ROLES DE PAGO POR RUBROS

AÑO: 2012 TIPO EMPLEADO: ADMINISTRATIVO
 MES: NOVIEMBRE TIPO RUBRO: INGRESOS 04-Feb-2014 13:47:47

ORD.	CODIGO RUBRO	NOMBRE RUBRO	VALOR
1	200	Otros Ingresos	5.674,02
2	247	Sueldo Unificado	656.815,03
3	250	O.I. Sueldo Docente (IR)	52.544,04

Figura 207: Tipo de salida del reporte

En este caso se escogió pdf como tipo de salida del reporte y con todos los parámetros completos, en la parte inferior se despliega el detalle del reporte.

Output Type

PDF

View Report Auto-Submit

UNIVERSIDAD DE LAS FUERZAS ARMADAS ESPE
 REPORTE DE ROLES DE PAGO POR RUBROS

AÑO: 2012 TIPO EMPLEADO: ADMINISTRATIVO
 MES: NOVIEMBRE TIPO RUBRO: INGRESOS 04-Feb-2014 13:47:47

ORD.	CODIGO RUBRO	NOMBRE RUBRO	VALOR
1	200	Otros Ingresos	5.674,02
2	247	Sueldo Unificado	656.815,03
3	250	O.I. Sueldo Docente (IR)	52.544,04

Figura 208: Ejecución del reporte

4.12 MANUAL DE USUARIO DE DASHBOARD

4.12.1 INTRODUCCIÓN

La herramienta de pentaho tiene la opción de poder realizar cuadros de mando o también conocidos como la creación de los dashboard. Este gran aporte se lo puede realizar con la ayuda del community dashboard editor (CDE) que es un plugin que viene integrado en el marketplace de la consola de usuario y sirve para crear, editar los dashboard mediante una interfaz web gráfica.

4.12.2 PROCESO DE CREACIÓN DE UN DASHBOARD

Para la creación de los dashboard, se los crea en la consola de usuario de la herramienta de pentaho. Para esto se va a proceder a indicar todos los pasos que se realizó para la elaboración de un dashboard del submódulo matrículas con los requerimientos establecidos por el usuario final en donde el dashboard va a tener un nombre de título, un parámetro que será el periodo y tendrá 4 gráficos que son:

- Matrículas desglosadas
- Pagos desglosados
- Total Matrículas y pagos
- Total deudores

Para realizar el dashboard como primer paso, ingresar a la siguiente dirección web: **siiespe.espe.edu.ec**, preferiblemente usando Mozilla Firefox o Google Chrome. Se desplegará la siguiente pantalla:

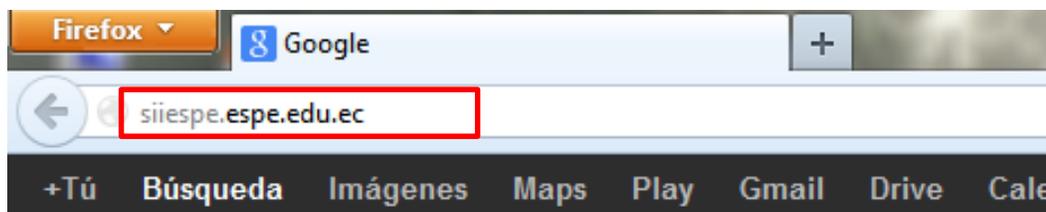


Figura 209: URL de la consola de usuario

Se abre la consola de usuario, se escribe el usuario y el password asignado a cada usuario por el administrador. Un ejemplo puede ser usuario: **joe** y password: *********



Figura 210: Interfaz de login de la consola de usuario

Se abre la consola de usuario conectada con un usuario, con esto hacer clic en la opción **CDE** para que se abra una nueva ventana en la que se crea el dashboard y hacer clic en la opción **save** para guardar el dashboard.

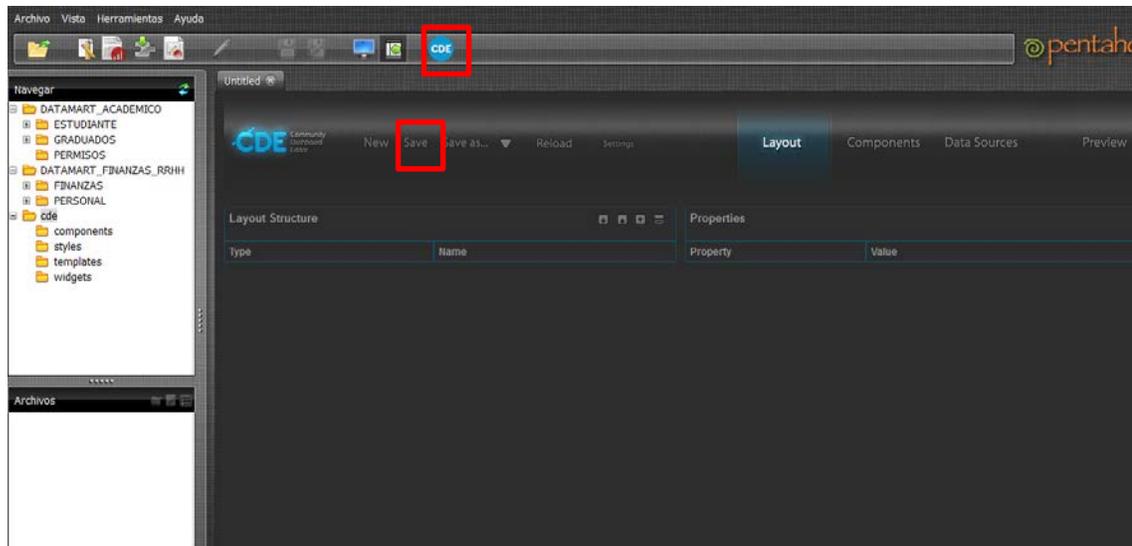


Figura 211: Abrir CDE

En la figura 212 se muestra como se guarda el archivo en la ruta que el usuario desee, se coloca el nombre al archivo y al título del dashboard y finalmente hacer clic en **ok** para que se guarde completamente.

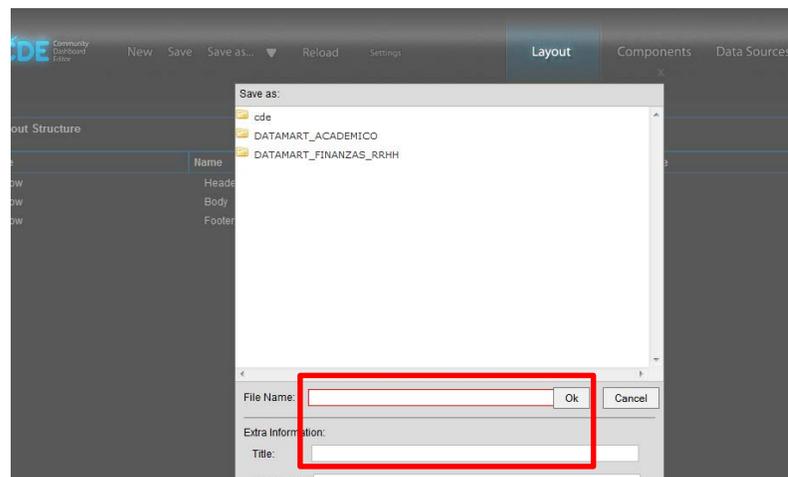


Figura 212: Guardar dashboard

Una vez guardado, se va a realizar el diseño del dashboard. Para esto el usuario se debe colocar en la opción de layout y hacer clic en el símbolo que se encuentra en el recuadro que sirve para seleccionar la plantilla.

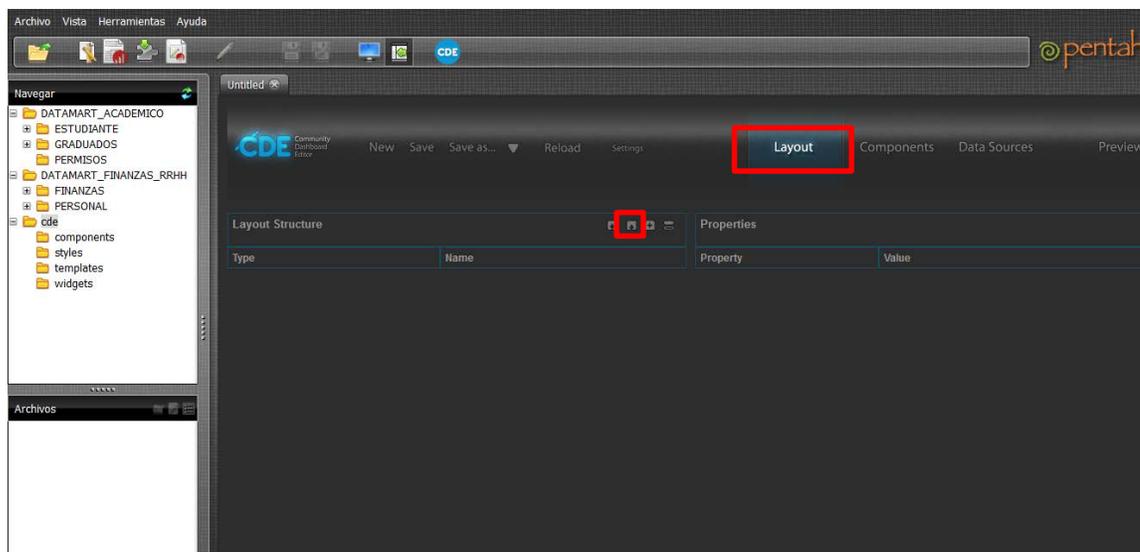


Figura 213: Opciones de Layout

En la siguiente pantalla se puede escoger la plantilla que desee el usuario, para este caso se va a usar una plantilla personalizada, por lo tanto hacer clic en **my templates**.

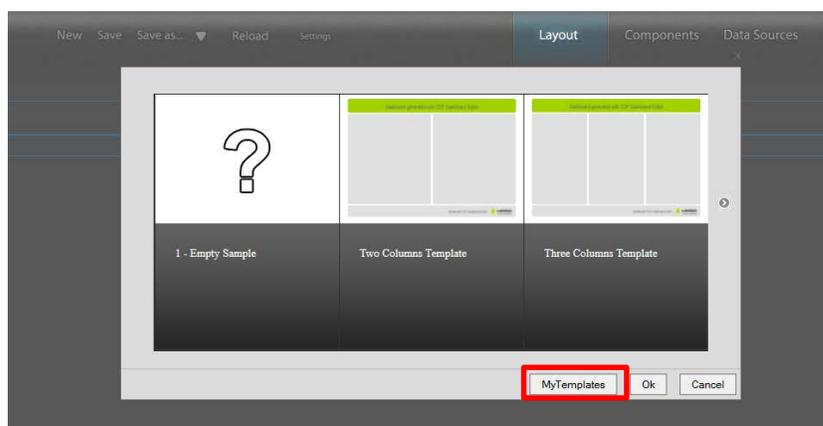


Figura 214: Opciones de plantillas

Aquí se va a mostrar una plantilla llamada **MATRICULACIÓN**, que es una plantilla elaborada con una estructura para los requerimientos del dashboard de matrículas. Seleccionar la plantilla y luego ok.

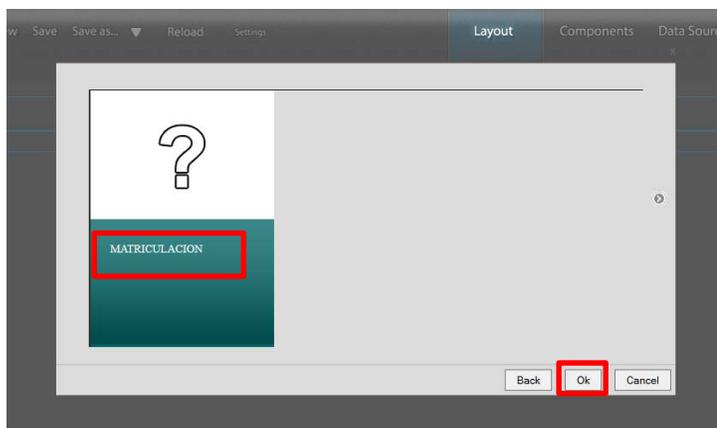


Figura 215: Opciones de plantillas personalizadas

Se va a mostrar un mensaje para cargar la plantilla. Hacer clic en **ok** para tener el dashboard con la plantilla cargada.

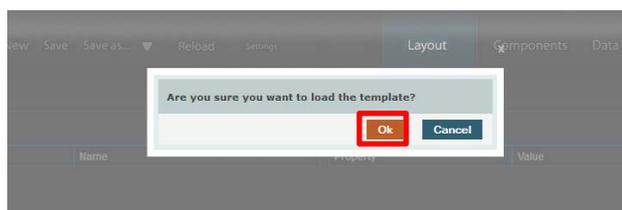


Figura 216: Mensaje para aplicar plantilla

El dashboard tendrá cargada la plantilla y a continuación se va a mostrar la estructura que tiene. Para visualizar en modo gráfico el diseño de la plantilla, hacer clic en **preview**.

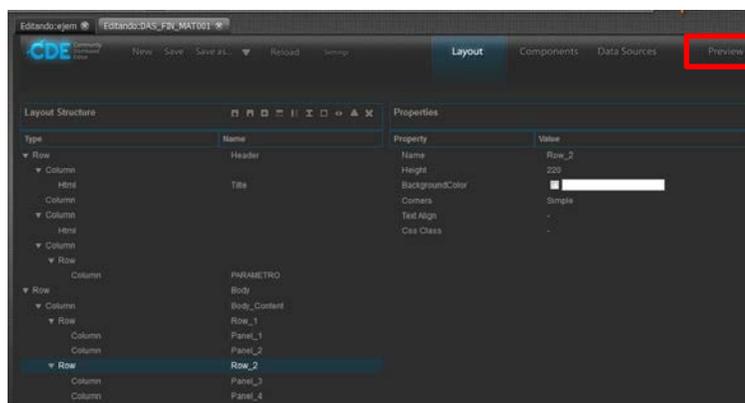


Figura 217: Diseño de la plantilla

A continuación se va a mostrar el diseño que tiene la plantilla de matriculación seleccionada para la elaboración del dashboard de matrículas en donde se observa un título, un espacio para el parámetro que va a ser periodo y por ultimo 4 espacios para los gráficos que se van a mostrar una vez que ya se termine de explicar este manual.

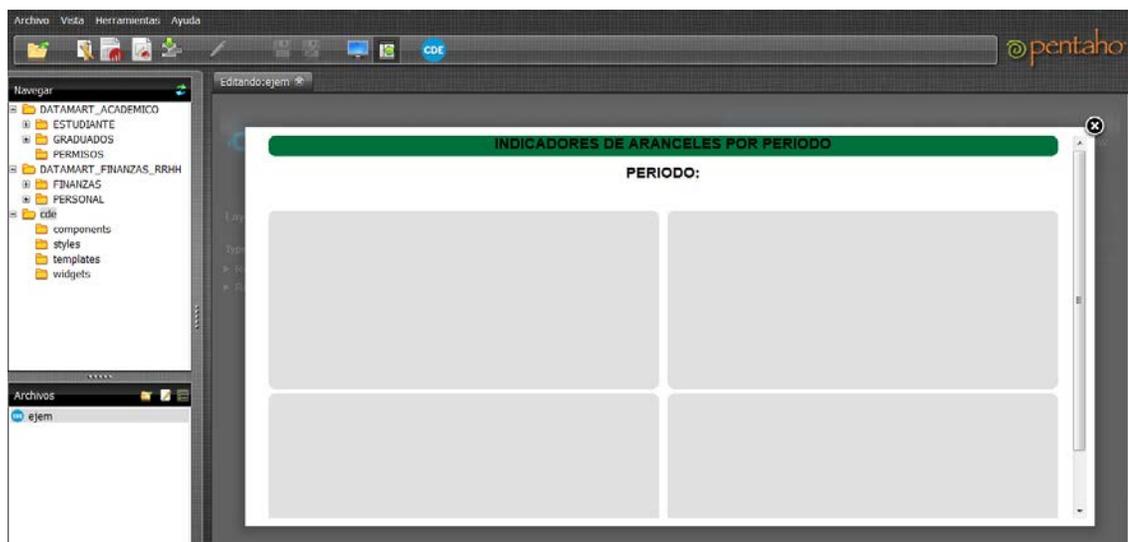


Figura 218: Previsualizar plantilla

4.12.3 Crear parámetros

Para crear un parámetro, se debe ubicar en components, luego en la parte de la izquierda hacer clic en **generic** para que se desplieguen las opciones en donde se va a seleccionar la opción **simple parameter**. Por ultimo llenar los valores del parámetro como el usuario requiera.

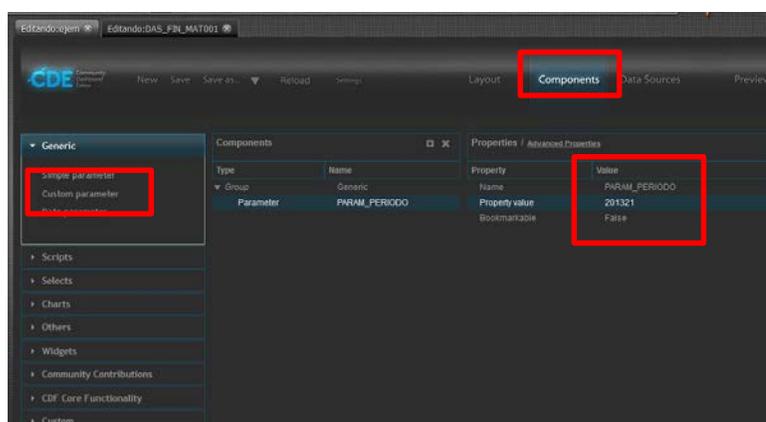


Figura 219: Crear parámetro

Nuevamente se ubica en la parte de la izquierda, hacer clic en **selects** y seleccionar la opción **select component**, con esto el usuario va a llenar los valores de la misma forma como se hizo en el paso anterior.

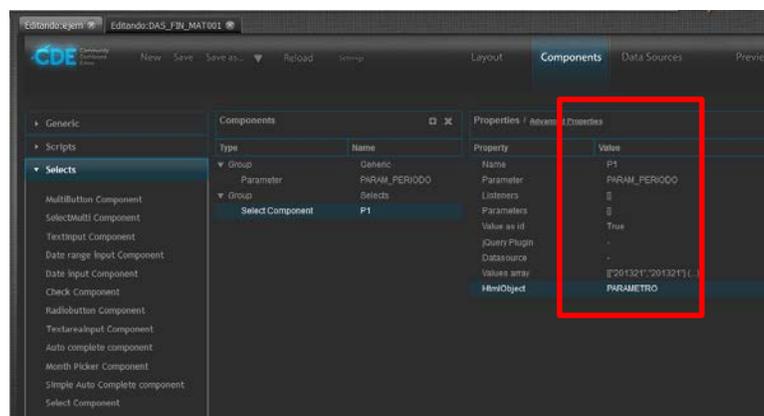


Figura 220: Crear componente para el parámetro

4.12.4 Crear MDX Queries

Ahora se va a crear los queries que van a servir para obtener la información que se va a desplegar en el dashboard. Para esto se ubica en Data Sources, luego en la parte izquierda hacer clic en **MDX queries** y seleccionar la opción **mdx over mondrianJndi**. Finalmente llenar los valores necesarios para obtener la información.

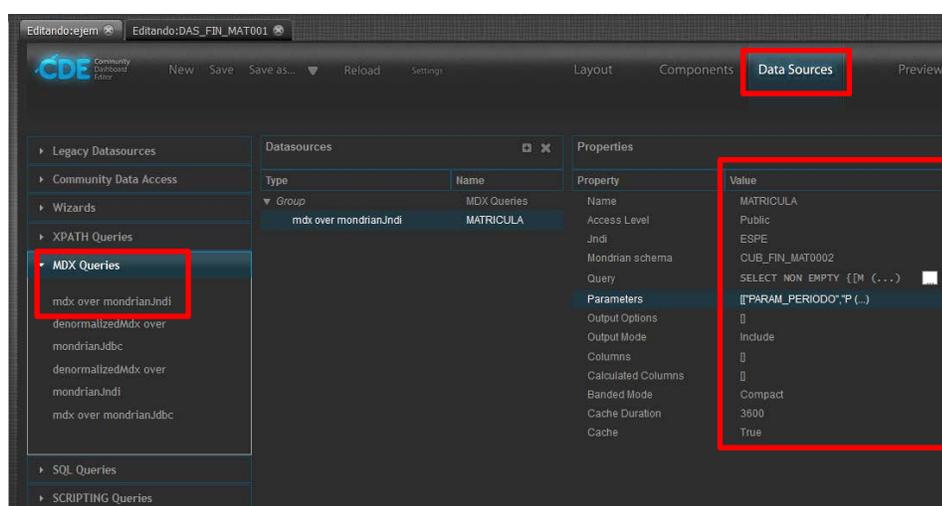


Figura 221: Crear MDX queries

Para el dashboard de matrículas se va a necesitar 3 queries MDX en total, por lo tanto se realiza el mismo procedimiento del paso anterior con los otros 2 queries requeridos y se va a obtener como resultado los 3 queries MDX que se encuentran en el recuadro de la figura 222.

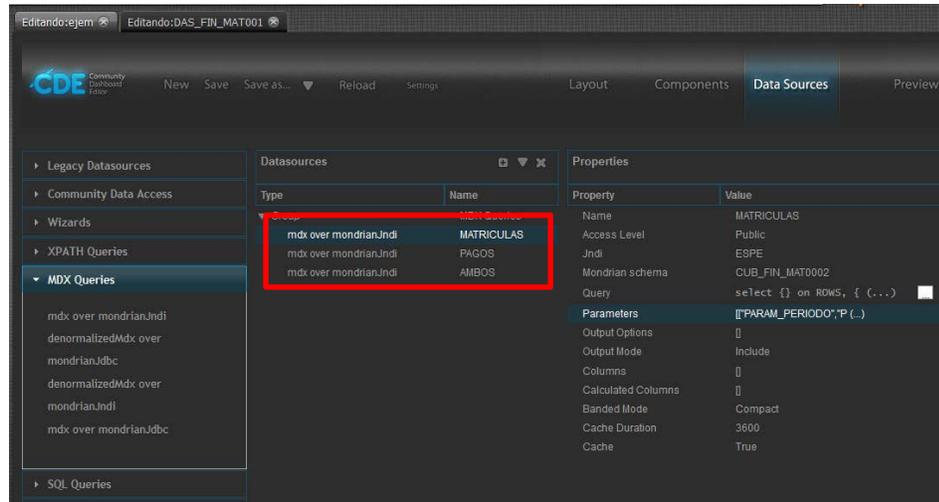


Figura 222: 3 MDX queries creados

4.12.5 Crear SQL Queries

Adicionalmente se requiere un query sql para el cuarto grafico del dashboard de matrículas, por lo tanto, en la parte izquierda, hacer clic en **SQL queries** y seleccionar la opción **sql over sqljndi**. Por ultimo llenar los valores necesarios para obtener la información del último gráfico.

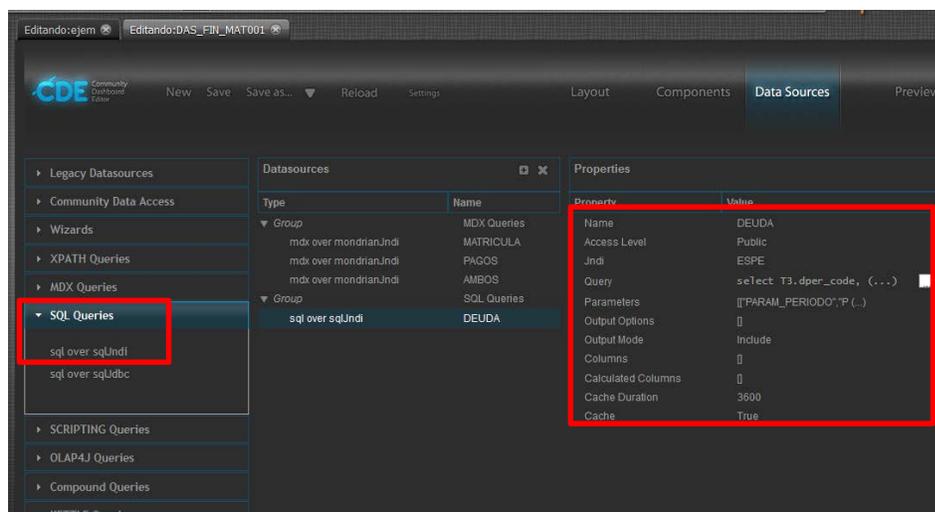


Figura 223: Crear SQL queries

4.12.6 Crear componentes

Ahora se va a escoger los tipos de gráficos que se van a utilizar en el dashboard que van a ser barras, por eso como primer paso, se debe ubicar en components, hacer clic en **charts** que se encuentra en la parte izquierda y en las opciones que se despliegan, seleccionar **CCC bar chart**. Finalmente llenar los valores en la parte derecha como se muestra en la siguiente imagen para que el procedimiento del primer grafico este completo.

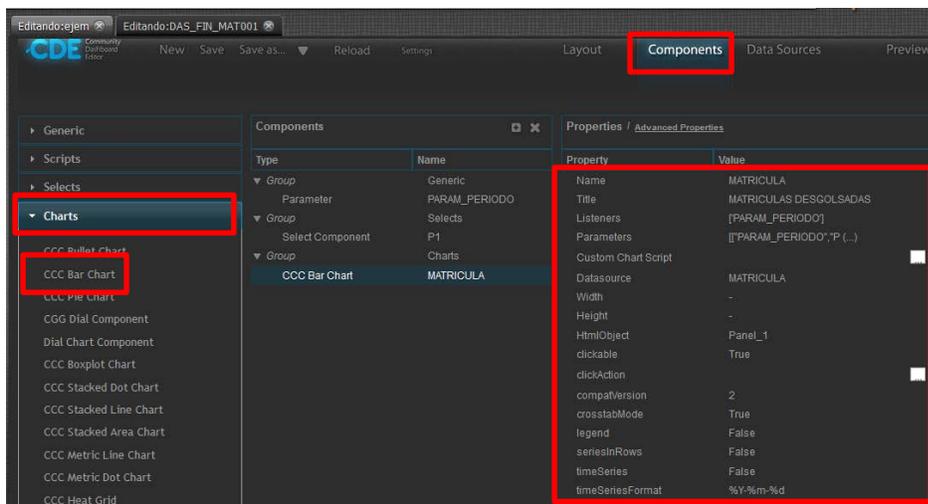


Figura 224: Creación gráfico de barras

4.12.6.1 Advanced properties

Una vez que se llena los valores del grafico de barras, en este caso se necesita modificar dos propiedades más pero que no se visualizan a primera vista. Para esto hacer clic en **Advanced Properties**, aquí se va a desplegar muchas más propiedades con sus valores, se buscara la propiedad `valuesvisible` para poner el valor **true** como se observa en la figura 226 y en la propiedad `orientation` para poner el valor **horizontal** como se observa en la figura 227. Con estos últimos cambios, el primer gráfico está listo y para visualizar lo que se ha hecho, hacer clic en **preview**.

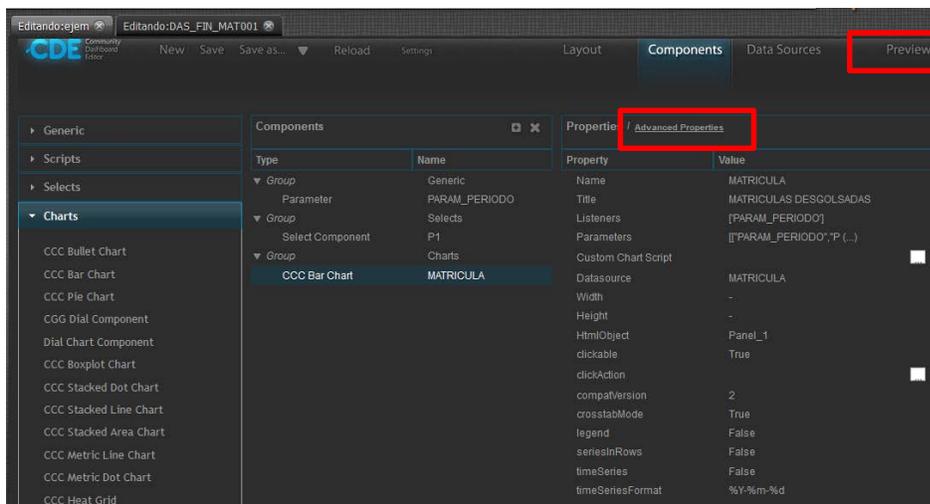


Figura 225: Propiedades avanzadas



Figura 226: Propiedad de valores visibles

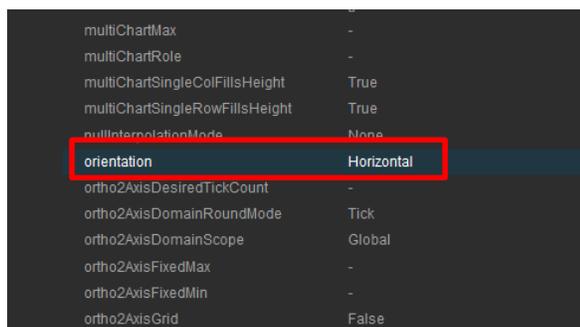


Figura 227: Propiedad de orientación

A continuación se muestra el pre visualización del dashboard de matrículas con la opción de escoger el parámetro de periodo en la parte superior y con el primer gráfico de matrículas desglosadas explicado en los pasos anteriores.

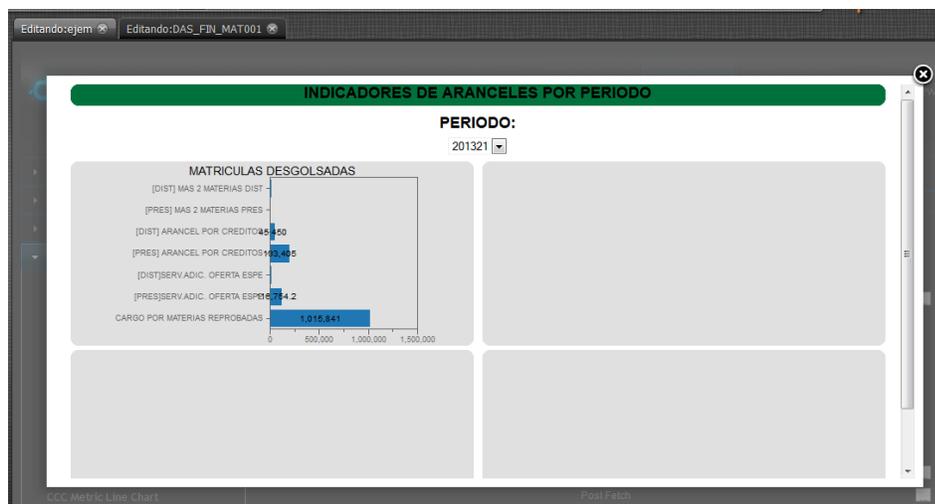


Figura 228: Previsualizar dashboard con un gráfico de barras

Para los 3 gráficos restantes, se realiza el mismo procedimiento de escoger el gráfico de barras y de completar los valores de las propiedades como se explicó anteriormente con el primer gráfico de matrículas desglosadas.

Una vez que ya esté listo el proceso de elaboración de los 4 gráficos, hacer clic en **save** para que se guarde el dashboard con todos los cambios que se ha hecho y por ultimo hacer clic en **preview** para mostrar el resultado final del dashboard de matrículas.

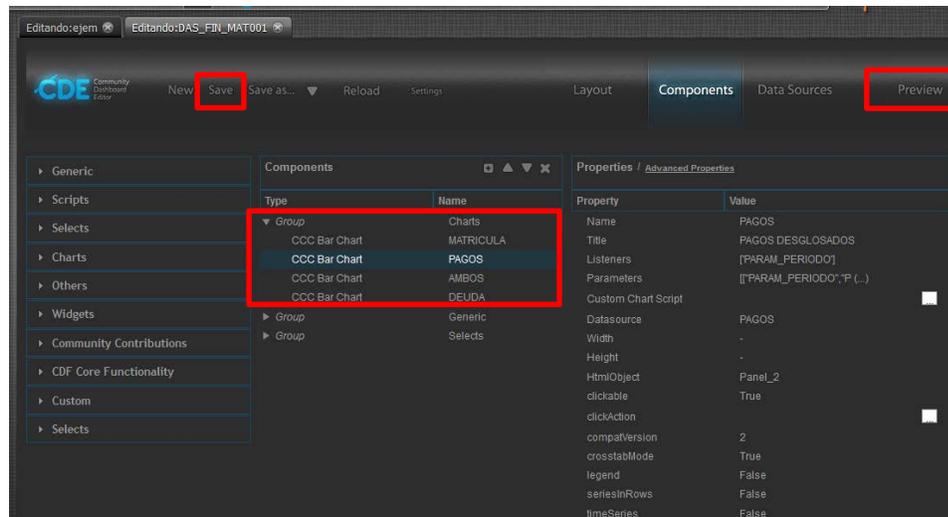


Figura 229: 4 Componentes de gráfico en barras

La figura 230 muestra el dashboard completo con su título, la opción de escoger el periodo como parámetro en la parte superior y los 4 gráficos de barras reflejando la información sobre el submódulo de matrículas.

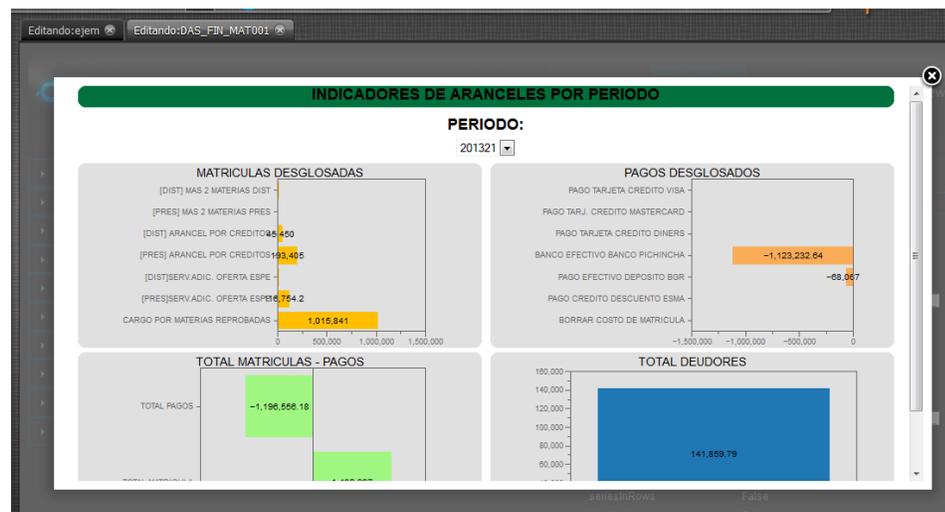


Figura 230: Dashboard completo

CAPÍTULO 5

5.1 Conclusiones

- Con la realización del proyecto se puede concluir que toda la documentación que pueda ofrecer la empresa al programador acerca del negocio es muy importante e indispensable en el proceso de creación de datamarts ya que la documentación brindada por la empresa es útil para establecer todos los parámetros para un datamart exitoso.
- Con el análisis que se empleó para realizar los datamarts de cada una de las áreas del financiero y de talento humano, se puede detectar que se carga únicamente la información útil que fue considerado en la toma de requerimientos y esto sirve para que los cubos de cada área de negocio brinden una respuesta muy rápida al momento de hacer una consulta y con esto se puede demostrar que al realizar una buena solución, se va a obtener mejoras y una optimización del análisis del negocio.
- Una vez finalizado el proyecto se puede concluir que la metodología de Ralph Kimball es la más adecuada gracias a la arquitectura que posee, pero en cuanto a documentos entregables no tiene estándares bien definidos por lo que se dificulta la elaboración de documentos finales que deben ser entregados al usuario.

5.2 Recomendaciones

- Antes de realizar la parte práctica, se recomienda pedir y leer detalladamente toda la documentación que le pueda brindar la empresa ya que es muy importante que se tenga conocimiento de lo que se va a realizar y esto podría ahorrar tiempo al momento de conocer lo que se tiene hecho y lo que falta por implementar.
- Para realizar un proyecto de este tipo, es muy recomendable identificar todos los recursos que brinda la empresa en los cuales serán de gran ayuda para un diseño e implementación de la solución y con esto saber si los recursos son aptos para dar capacidad a toda la información que se maneje y para el número de usuarios que lo vayan a dar uso.
- Se recomienda establecer estándares a todo lo que se vaya a implementar, ya sea en el área de base de datos como en el área de la herramienta, con esto los estándares reducen el tiempo en los procesos y además mejora la calidad del proyecto para prevenir almacenamiento de información innecesaria.

BIBLIOGRAFIA

Bernabeu, R. (2009). *DATA WAREHOUSING: Investigación y Sistematización de Conceptos*. Buenos Aires.

ESPE. (2010). *Sitio Web De La Universidad de las Fuerzas Armadas*. Obtenido de http://www.espe.edu.ec/portal/files/ley_transparencia/archivos/ESTRUCTURA_ESPE.pdf

Espinoza, R. (Marzo de 2011). *DESCUBRIENDO EL BUSINESS INTELLIGENCE*. Obtenido de <http://churriwifi.wordpress.com/>

Kimball, R. (1998). *The Data Warehouse Lifecycle Toolkit*. California: John Wiley.

Lopez, J. L. (2008). *Desarrollo de un Data Mart Para La Biblioteca de la Universidad Politécnica Nacional*. Quito.

GLOSARIO

- **ETL:** Extract Transformation Load (Extracción, Transformación y Carga)
- **BI:** Business Intelligence
- **OLAP:** On Line Analytical Processing
- **ROLAP:** On Line Analytical Processing Relational
- **MOLAP:** Multidimensional On Line Analytical Processing
- **HOLAP:** Hybrid On Line Analytical Processing
- **MDX:** Multidimensional Expression
- **JDNI:** Java Naming and Directory Interface
- **JDBC:** Java Database Connectivity
- **OLTP:** On Line Transaction Processing
- **PDI:** Pentaho Data Integration
- **PSW:** Pentaho Schema Workbench
- **PRD:** Pentaho Report Designer
- **CDE:** Community Dashboard editor
- **Arquitectura:** Diseño de una solución de business intelligence
- **XML:** eXtensible Markup Language
- **UTIC:** Unidad de Tecnología de Información y Comunicación
- **ERP:** Enterprise Resource Planning
- **CRM:** Customer Relationship Management
- **DASHBOARD:** Tablero de control

BIOGRAFÍA

Nombres y Apellidos:

Diego Armando Esparza Montes

Lugar y Fecha de Nacimiento:

Quito, 24 de Mayo de 1987

Educación Primaria:

Escuela “La Salle” – Quito

Educación Secundaria:

Escuela “La Salle” – Quito

Educación Superior:

Universidad de las Fuerzas Armadas ESPE – Sangolquí

Ingeniería de Sistemas e Informática

Títulos Obtenidos:

Suficiencia en el idioma Inglés

Nombres y Apellidos:

Cristian Dario Alvarez Calvopiña

Lugar y Fecha de Nacimiento:

Latacunga, 18 de Setiembre de 1988

Educación Primaria:

Escuela Franciscana “San Andrés” – Quito

Educación Secundaria:

Colegio Técnico Experimental Aeronáutico de Aviación Civil “COTAC” – Quito

Educación Superior:

Universidad de las Fuerzas Armadas ESPE – Sangolquí

Ingeniería de Sistemas e Informática

Títulos Obtenidos:

Suficiencia en el idioma Inglés

HOJA DE LEGALIZACIÓN DE FIRMAS

ELABORADO POR:

Cristian Darío Alvarez Calvopiña

Diego Armando Esparza Montes

DIRECTOR DE LA CARRERA

Ing. Mauricio Campaña MsC.

Sangolquí, Julio del 2014