



**Implementación de un sistema de business intelligence para la toma de
decisiones de una agencia de courier**

Jaramillo Cañadas, Miguel Leopoldo

Vicerrectorado de Investigación, Innovación y Transferencia de Tecnología

Centro de Posgrados

Maestría en Gestión de Sistemas de Información e Inteligencia de Negocios

Trabajo de titulación, previo a la obtención del título de Magíster en Gestión de
Sistemas de Información e Inteligencia de Negocios

Msc. Gualotuña Alvarez, Tatiana Marisol

23 de diciembre del 2021

COPYLEAKS

TRABAJO_TITULACION_JARAMILLO_v3.pdf

Scanned on: 17:55 March 2, 2022 UTC



Overall Similarity Score



Results Found



Total Words in Text

Identical Words	1030
Words with Minor Changes	206
Paraphrased Words	1121
Omitted Words	1964



Firmado electrónicamente por:
**TATIANA MARISOL
GUALOTUNA
ALVAREZ**



**VICERRECTORADO DE INVESTIGACIÓN, INNOVACIÓN Y TRANSFERENCIA DE
TECNOLOGÍA**

CENTRO DE POSGRADOS

CERTIFICACIÓN

Certifico que el trabajo de titulación, “**Implementación de un sistema de *business intelligence* para la toma de decisiones de una agencia de *courier***” fue realizado por el señor **Miguel Leopoldo Jaramillo Cañadas** el cual ha sido revisado y analizado en su totalidad por la herramienta de verificación de similitud de contenido; por lo tanto cumple con los requisitos legales, teóricos, científicos, técnicos y metodológicos establecidos por la Universidad de las Fuerzas Armadas ESPE, razón por la cual me permito acreditar y autorizar para que lo sustente públicamente.

Sangolquí, 15 de diciembre 2021



Firmado electrónicamente por:
**TATIANA MARISOL
GUALOTUNA
ALVAREZ**

Gualotuña Álvarez, Tatiana Marisol

C.C.: 1711498418



ESPE
UNIVERSIDAD DE LAS FUERZAS ARMADAS
INNOVACIÓN PARA LA EXCELENCIA

VICERRECTORADO DE INVESTIGACIÓN, INNOVACIÓN Y
TRANSFERENCIA DE TECNOLOGÍA

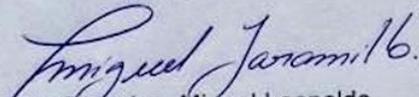
CENTRO DE POSGRADOS

RESPONSABILIDAD DE AUTORÍA

Yo **Jaramillo Cañadas, Miguel Leopoldo** con cédula de ciudadanía No. 2100160544, declaro que el contenido, ideas y criterios del trabajo de titulación: **Implementación de un sistema de business intelligence para la toma de decisiones de una agencia de courier** es de mi autoría y responsabilidad, cumpliendo con los requisitos legales, teóricos, científicos, técnicos y metodológicos establecidos por la Universidad de las Fuerzas Armadas ESPE, respetando los derechos intelectuales de terceros y referenciando las citas bibliográficas.

Sangolquí, 15 de diciembre 2021

Firma


Jaramillo Cañadas, Miguel Leopoldo

C.C.: 2100160544



VICERRECTORADO DE INVESTIGACIÓN, INNOVACIÓN Y
TRANSFERENCIA DE TECNOLOGÍA

CENTRO DE POSGRADOS

AUTORIZACIÓN DE PUBLICACIÓN

Yo **Jaramillo Cañadas, Miguel Leopoldo** con cédula de ciudadanía No.2100160544, autorizo a la Universidad de las Fuerzas Armadas ESPE publicar el trabajo de titulación: **Implementación de un sistema de business intelligence para la toma de decisiones de una agencia de courier** en el Repositorio Institucional, cuyo contenido, ideas y criterios son de mi responsabilidad.

Sangolquí, 15 de diciembre 2021

Firma

Jaramillo Cañadas, Miguel Leopoldo

C.C.: 2100160544

Agradecimientos

Agradezco a Dios y a Jesús porque me sostienen en bendiciones día tras día.

Mi agradecimiento a mi novia Maritza Badillo por su apoyo y empuje incondicional desde antes de empezar esta Maestría. Gratitud de corazón a mis padres por su ejemplo de principios y fortaleza. A mi hermana Evita y Santiago por estar conmigo siempre con cariño, a Felipe y Tomas por brindarme esas alegrías y sus enseñanzas de almas puras.

A todos ellos por darme una nueva oportunidad de vida y poder vivir esta bendición junto con ellos.

Dedicatoria

Dedico este logro profesional a Dios y a Jesús por su amor incondicional, por su divina presencia, por ser mi luz que me guía día a día y por brindarme su bendición.

A mis abuelitas que son mis ángeles en el cielo y que sigo su ejemplo de amor.

A mi novia Maritza Badillo y a mi familia Luis, Ruth, Evita, Santiago, Felipe y Tomas, por su apoyo y amor incondicional.

Índice de contenido

Certificado del director	2
Responsabilidad de autoría.....	3
Autorización de publicac	4
Agradecimientos	5
Dedicatoria.....	6
Resumen	14
Abstract	15
Capítulo I	16
Introducción	16
Antecedentes	17
Justificación e importancia.....	18
Planteamiento del problema	19
Objetivo general	22
Objetivos específicos.....	22
Preguntas de investigación.....	23
Capítulo II	24
Estado del arte y marco teórico.....	24
Estado del arte	24
Criterios de inclusión	24
Criterios de exclusión	24
Cadena de búsqueda.....	26
Conclusiones de trabajos relacionados	30
Categorización de las variables de investigación	31
Categorización de la variable independiente.....	32
Inteligencia de Negocios	32

Datos, información y conocimiento.....	35
Almacenes de datos.....	36
Procesos de soporte para toma de decisiones.....	37
Metodologías, técnicas y herramientas para el desarrollo de soluciones de Inteligencia de Negocios.....	41
Decisiones efectivas y oportunas.....	42
El rol de los modelos matemáticos.....	44
Arquitectura de un sistema de Inteligencia de Negocios.....	46
Ciclo de la Inteligencia de Negocios.....	49
Desarrollo de un sistema de Inteligencia de Negocios.....	52
Data Warehouse (DW) y On Line Analytical Proccesing (OLAP).....	54
Data mart	56
Pentaho	56
Microsoft Power BI.....	58
Sistema de Gestión de Base de Datos.....	59
PostgreSQL para Inteligencia de Negocios.....	59
Metodología de Bill Inmon.....	60
Metodología de Ralph Kimball.....	63
Diseño y desarrollo del sistema de BI.....	67
Data Marts y modelado dimensional.....	67
Conexión de los Data Marts a la arquitectura del Bus de DW.....	69
Puntos clave para el diseño de soluciones DW/BI.....	74
Pautas para el diseño de un DW de bajo riesgo.....	75
Selección de la metodología de trabajo.....	78
Categorización de la variable dependiente.....	82
Agencia de Courier.....	82
Atención al cliente.....	83
Proceso de atención al cliente.....	83

Eficiencia en la atención al cliente.....	84
Capítulo III	85
Análisis de estado actual.....	85
Capítulo IV.....	90
Marco metodológico.....	90
Implementación del sistema de Business Intelligence.....	90
Planificación del proyecto.....	91
Alcance del proyecto.....	91
Factibilidad del proyecto.....	91
Definición de los requerimientos	103
Diseño de la arquitectura técnica	107
Infraestructura existente.....	108
Productos y herramientas para el desarrollo de la solución.....	108
Diagramas de alto nivel de la arquitectura técnica	109
Modelado dimensional	110
Dimensiones identificadas.....	111
Hechos identificados	116
Diseño físico	120
Diseño y desarrollo ETL.....	129
Diseño y desarrollo de las aplicaciones.....	136
Creación del cubo y generación de la visualización de datos	136
Implementación.....	139
Capítulo V.....	141
Análisis e interpretación de resultados	141
Capítulo VI.....	152
Conclusiones y recomendaciones.....	153
Conclusiones.....	153
Recomendaciones.....	154

Bibliografía.....	156
Anexos.....	159

Índice de tablas

Tabla 1 Grupos de Control	25
Tabla 2 Cadena de búsqueda	26
Tabla 3 Tipos de datos.....	34
Tabla 4 Criterios e indicadores para comparación de metodologías	78
Tabla 5 Escala de Likert.....	80
Tabla 6 Ponderación de indicadores de las metodologías revisadas	81
Tabla 7 Plan del proyecto.....	92
Tabla 8 Resumen de problemas detectados	104
Tabla 9 Lista de requerimientos identificados	105
Tabla 10 Reportes identificados	106
Tabla 11 Indicador identificado.....	107
Tabla 12 Dimensiones identificadas	111
Tabla 13 Atributos de la dimensión fecha	112
Tabla 14. Atributos de la dimensión envío.	114
Tabla 15 Atributos de la dimensión gasto.	115
Tabla 16 Tabla de Hechos identificados.....	117
Tabla 17 Atributos de la tabla de hechos ventas.	117
Tabla 18 Atributos de la tabla de hechos gastos_operativos.	119
Tabla 19 Base de datos.	121
Tabla 20 Mapeo de datos fuente a destino.....	130
Tabla 21 Resumen de resultados de las pruebas funcionales	140
Tabla 22 Promedio índice de crecimiento 2019 - 2020.	152

Índice de figuras

Figura 1 Descripción del problema usando diagrama de Ishikawa.....	21
Figura 2 Categorización de las variables	31
Figura 3 Proceso ETL	38
Figura 4 Esquema del proceso OLAP	39
Figura 5 Proceso de minería de datos.....	40
Figura 6 Proceso de reportería (reporting).....	40
Figura 7 Beneficios de un sistema de Inteligencia de Negocio	43
Figura 8 Arquitectura típica de un sistema de Inteligencia de Negocios.....	47
Figura 9 Componentes de un sistema de Inteligencia de Negocios	48
Figura 10 Ciclo de la Inteligencia de Negocios	50
Figura 11 Fases del desarrollo de un sistema de Inteligencia de Negocios	52
Figura 12 Pentaho BI. Capas y componentes	57
Figura 13 Ciclo de vida metodología de Ralph Kimball	64
Figura 14 Fases de la metodología.	66
Figura 15 Esquema horizontal de un entorno de Data Warehouse.	68
Figura 16 Arquitectura de Bus de Data Warehouse.....	70
Figura 17 Priorización de impacto versus factibilidad.....	75
Figura 18 Organigrama de la estructura organizacional.....	86
Figura 19 Esquema de trabajo para la solución del problema.....	90
Figura 20 Modelo lógico dimensión fecha.....	114
Figura 21 Modelo lógico dimensión envío.....	115
Figura 22 Modelo lógico dimensión gasto.....	116
Figura 24 Modelo lógico tabla de hechos ventas.	118
Figura 24 Modelo lógico tabla de hechos gastos_ operativos.	119
Figura 25 Modelo dimensional del almacén de datos de la agencia de courier	120

Figura 26 Proceso de creación de la base de datos.	123
Figura 27 Proceso de creación de la dimensión Fecha: dim_fecha	124
Figura 28 Proceso de creación de la dimensión Envío: dim_envio.....	124
Figura 29 Proceso de creación de la dimensión Gasto: dim_gasto.....	125
Figura 30 Proceso de creación del hecho Ventas: fact_ventas.....	126
Figura 31 Proceso de creación de hecho Gastos operativos: fact_gastos_operativos.	127
Figura 32 Diagrama general del proceso de ETL.	129
Figura 33 Proceso de ETL.	133
Figura 34 Construcción de las tablas de dimensiones.	134
Figura 35 Proceso de limpieza inicial de las tablas de dimensiones.	134
Figura 36 Proceso de limpieza inicial de las tablas de hechos.....	135
Figura 37 Proceso de construcción de las tablas de hechos.....	135
Figura 38 Modelo del cubo de datos BI_OLAP1	136
Figura 39 Corrida del cubo BI_OLAP1: años 2019 y 2020 acumulados por año. ...	137
Figura 40 Corrida del cubo BI_OLAP1: año 2019 acumulado por producto.	138
Figura 41 Corrida del cubo BI_OLAP1: años 2019 y 2020 acumulado por producto.	138
Figura 42 Servicio más demandado por año.	142
Figura 43 Cantidad de servicios gestionados en 2019.....	142
Figura 44 Cantidad de servicios gestionados en 2020.....	143
Figura 45 Día de la semana con mayor y menor número de servicios (2019).	144
Figura 46 Día de la semana con mayor y menor número de servicios (2020).	145
Figura 47 Mes con mayor y menor número de ventas (2019 – 2020)	146
Figura 48 Año con mayor y menor número de ventas acumuladas.....	147
Figura 49 Comportamiento de la rentabilidad en un período.....	148
Figura 50 Comparativo de Ventas mayo 2019 vs. Ventas mayo 2020.	149
Figura 51 Comparativo índice de crecimiento Servientrega - DHL.....	150

Resumen

Ante el éxito de las grandes organizaciones, los intermediarios y las empresas de courier independientes comienzan a crecer en masa. Estas compañías independientes han hecho asequible que las pequeñas empresas locales y los vendedores en línea puedan ofrecer entregas competitivas. En este sentido las agencias de Courier han sido un puente importante entre los vendedores con sus clientes, en esta comunicación se genera gran cantidad de información debido a seguidas transacciones que se realizan diariamente. Sin embargo, existen problemas como la falta de actualización en nuevas tecnologías que aporten información sobre las actividades que impactan directamente a la toma de decisiones en la gestión del negocio de los Courier, este puede verse afectado en la calidad de prestación del servicio de la agencia. Por lo tanto, el objetivo general de esta investigación es implementar una solución de BI que, mejore la eficiencia en la toma de decisiones.. Como solución a esta problemática se implementó un sistema de Business Intelligence para la agencia de Courier basado en una investigación de Estado del Arte en la cual se determinó que la metodología de Ralph Kimball era la que más se adaptaba a esta solución y mediante la integración de herramientas de análisis de datos se obtuvieron, satisfactoriamente, las respuestas a las necesidades de información que apoyan la toma de decisiones en la agencia de Courier, en todos los niveles, es decir, niveles operacionales, tácticos y estratégicos, así como también, esta solución está contribuyendo a mejorar la calidad del dato, debido a que a través de los procesos de extracción y transformación se identifican los datos que no cumplen con los aspectos de calidad establecidos.

Palabras Clave

- **BUSINESS INTELLIGENCE**
- **TOMA DE DECISIONES**
- **ANÁLISIS DE DATOS**

Abstract

Faced with the success of large organizations, intermediaries and independent courier companies begin to grow en masse. These independent companies have made it affordable for local small businesses and online sellers to offer competitive deliveries. In this sense, Courier agencies have solved the connection of organizations with their clients, through direct contact with clients, generating a large amount of information due to the number of transactions carried out daily. However, there are problems such as the lack of updating in new technologies that provide information on the activities that directly impact decision-making in the management of the Courier business, this can be affected in the quality of the service provided by the agency. Therefore, the general objective of this research is to implement a BI solution that improves the efficiency in decision making.

As a solution to this problem, a Business Intelligence system was implemented for the Courier agency based on a State of the Art investigation in which it was determined that Ralph Kimball's methodology was the one that best suited this solution and through the integration of Data analysis tools were satisfactorily obtained responses to the information needs that support decision-making in the Courier agency, at all levels, that is, operational, tactical and strategic levels, as well as this solution It is helping to improve the quality of the data, because through the extraction and transformation processes the data that do not comply with the established quality aspects are identified.

Key words

- **BUSINESS INTELLIGENCE**
- **DECISION MAKING**
- **DATA ANALYSIS**

Capítulo I

Introducción

Un negocio moderno y adaptado a los cambios tecnológicos actuales es, un negocio centrado en el cliente, esto significa que los procesos existentes deben transformarse completamente con ayuda de herramientas que apoyen las operaciones principales de cada organización. Este proceso en sí, es complejo debido a que se requiere de la migración y modernización de sistemas que cumplan con las expectativas a cumplir en la evolución de los mercados y el crecimiento financiero sostenido.

Para la transformación del negocio se necesita modificar procesos que se enfoquen en los clientes, con el propósito de crear nuevos canales de ingresos, acortando los tiempos de respuesta mediante un enfoque ágil. Esto permitirá una mejor experiencia para el cliente, como es el caso de un procesamiento de pedidos más rápido y eficiente.

La Inteligencia de Negocios juega un papel importante en esta transformación, debido a que aporta las herramientas necesarias para el logro de los objetivos y la toma de decisiones oportunas que benefician a la organización, mejorando la productividad en general. Las tareas rutinarias requieren mucho tiempo; en contraste con esto, la utilización de las tecnologías adecuadas, permite que la empresa se enfoque en lo que más importa, sus clientes y el ahorro de sus recursos.

Con la Inteligencia de Negocios, las empresas y las organizaciones pueden exponer y dar respuesta a preguntas con solo el uso estructurado de los datos que generan en sus actividades diarias. Con los datos históricos y actuales se puede obtener una visión prospectiva del comportamiento del negocio en distintas áreas, proporcionando así beneficios, al optimizar el rendimiento gracias al uso inteligente de

la información. Todo esto conduce a la reducción de costos, resultados más confiables, nuevas oportunidades de negocios, anticipación a las necesidades de los clientes e incremento en la eficiencia.

Antecedentes

En la actualidad las personas necesitan comprar cosas en línea con el fin de enviarlas a otro destino. Para tal fin se ha apoyado con los automóviles, los ferrocarriles y las carreteras principales, por lo que la industria del correo despegó rápidamente. Ahora, existe una red integral de transportistas y mensajeros que se extiende por todo el mundo. La evolución del transporte moderno hizo posible enviar un paquete de un extremo del mundo a otro en solo un día, e Internet abrió los servicios de entrega para todos.

Ante el éxito de las grandes organizaciones, los intermediarios y las empresas de Courier independientes comienzan a crecer en masa. Estas compañías independientes han hecho asequible que las pequeñas empresas locales y los vendedores en línea puedan ofrecer entregas competitivas. Ahora, los clientes pueden obtener cotizaciones rápidamente, realizar un seguimiento de las entregas y evitar complejos trámites para obtener un servicio más rápido y sencillo. (Secure Media Services, 2018)

El primer uso del término “Inteligencia de Negocios” se lo imputa a Richard Miller, en su libro *Cyclopedia of Commercial and Business Anecdotes*. Lo utilizó para describir cómo Sir Henry Furnese, un banquero exitoso, se benefició a partir de la información reuniéndola activamente y actuando sobre ella antes de su competencia.

En 1970, Edgar Codd publicó un artículo citado “Un modelo relacional de datos para grandes bancos de datos compartidos”. Así comenzó el camino en la creación de las bases de datos relacionales, lo que permitió una capacidad mucho más amplia para almacenar y manipular datos. (Limp, 2020)

Ahora bien, con respecto a la actividad objeto de estudio, se puede afirmar que el Courier es uno de los servicios más antiguos del mundo. Este sistema ha evolucionado constantemente alineado a las nuevas necesidades y circunstancias del mundo hasta la época actual, caracterizada por nuevos retos basados en la transformación digital. Como consecuencia se ha observado la brecha entre la tasa de crecimiento anual de los envíos Courier (2%) y del comercio electrónico del país (8% promedio), pese a que ambos deberían ir de la mano. (Landa, Núñez, & Valdivia, 2017)

Como se evidencia, el crecimiento de envíos por Courier en comparación al comercio electrónico tiene un lento crecimiento debido a varias razones, una de las cuales es la falta de inversión en sistemas tecnológicos en las agencias de recepción de paquetes que permitan a estas ser más competitivas. El poder de posicionamiento que puede tener un negocio se basa en el correcto aprovechamiento de los datos para la toma de decisiones, a través de la implementación de Sistema de Inteligencia de Negocios. (Rosado & Rico, 2010)

Justificación e importancia

En la actualidad los procesos de análisis de datos de la agencia en estudio son llevados en gran parte de forma manual, lo que determina en muchos casos que la gerencia tome decisiones mal fundamentadas y de forma empírica, generando pérdidas económicas y de posicionamiento en el mercado. A pesar de que cuenta con un sistema informático para la gestión operativa del negocio (logística), se hace imprescindible colocarse a la vanguardia de los nuevos conjuntos de técnicas de información como lo es la Inteligencia de Negocios para ir acorde con el crecimiento, tanto tecnológico como económico, ya que conlleva a la mejor toma de decisiones que potencien nuevas oportunidades de ampliación en los ámbitos nombrados.

El manejo de datos con herramientas o sistemas inteligentes, proporciona una ventaja competitiva con respecto a los demás, lo que implica garantizar la supervivencia de la empresa en el tiempo.

El servicio de Courier ha evolucionado constantemente, a la par de las nuevas necesidades y circunstancias del mundo, hasta la época actual en la que existen nuevos retos basados en la transformación digital. Un punto de gran importancia dentro de los componentes que forman el Courier, son las agencias de recepción de paquetes, las cuales tienen contacto directo con los clientes, generándose gran cantidad de información debido al número de transacciones que se realizan diariamente. Todos estos datos pueden ser utilizados de manera inteligente, procesándolos a través de sistemas tecnológicos para obtener resultados de gran valor del éxito en la gestión y control del negocio. Dichos sistemas tecnológicos pueden:

- Dar una visibilidad más amplia y precisa del negocio.
- Centralizar la información en un solo repositorio de datos.
- Reducir el tiempo requerido para el procesamiento de la información.
- Mejorar la toma de decisiones reduciendo pérdidas económicas.

El fin de la implementación de un sistema de Business Intelligence es catapultar a la agencia de Courier hacia la era de la transformación digital. Al estar a la vanguardia tecnológica, podrá ser más competitiva frente a otras empresas del ramo, por lo que da garantía de aumento en la calidad y eficiencia del servicio.

Planteamiento del problema

Las agencias de Courier son sucursales distribuidas en sitios estratégicos a nivel nacional e internacional, en donde se reciben los paquetes para el posterior envío hacia sus respectivos destinos. La agencia seleccionada en el presente estudio actualmente consolida la información del giro del negocio sin un proceso

automatizado, en forma manual, debido a que la información se maneja tanto en archivos físicos como digitales, lo que provoca que se realice un análisis de los distintos procesos sin exactitud e invirtiendo gran cantidad de tiempo con información no precisa. Lo que puede derivar en errores significativos en la toma de decisiones.

No existen herramientas de gestión y control de negocio que permitan realizar análisis de importancia para la gestión, asimismo, el costo de implementar un Sistema de Planificación de Recursos Empresariales es muy elevado en comparación con las ganancias generadas por las actividades realizadas, lo cual dificulta formalizar un estudio real del margen de ganancias y pérdidas, ventas, rotación del personal y posibles inversiones. La toma de decisiones no es realizada consultando alguna base de datos histórica, sino a partir de experiencias propias de la gerencia, lo que ha causado desventaja a la hora de conseguir nuevas oportunidades para generar un mayor incremento en las ganancias.

Los resultados obtenidos de la toma de decisiones no orientadas adecuadamente, son el impacto directo de la ausencia de un plan de marketing y estudios de nichos de mercado, estas prácticas no se han desarrollado o tomado en cuenta desde el inicio de las operaciones de la agencia, causando que los clientes potenciales y los que ya han utilizado el servicio desconozcan los beneficios, promociones y costos de los distintos servicios de envío que se ofrecen.

Adicionalmente, al no detallar un panorama claro de ventas, centrado en las necesidades de los clientes, así como el desconocimiento de procedimientos automatizados apoyados en herramientas tecnológicas, se ha provocado que la toma de decisiones sea ineficiente, en la actualidad no se conoce a ciencia cierta cuanto personal de atención al cliente especializado se requiere para cubrir las horas o días en los que se recibe una mayor afluencia de clientes, por lo que no se ha podido plantear estrategias sostenibles para competir con otras agencias que realizan las mismas actividades; por ejemplo, el desconocimiento de los horarios de funcionamiento de los competidores ha traído como consecuencia que muchos de

los clientes al requerir el servicio no lo contraten, debido a que no esté disponible al público en momentos o días claves.

Por los problemas detallados y la falta de actualización en nuevas tecnologías que aporten información sobre las actividades que impactan directamente a la toma de decisiones en la gestión del negocio, este puede verse afectado en la calidad de prestación del servicio de la empresa. Esto representa una desventaja con relación a los competidores, retrasa el crecimiento económico y la ampliación de operaciones, el posicionamiento en otras zonas y la expansión de los servicios, planificación de estrategias que aporten al incremento de la eficacia de una asistencia de calidad y efectividad al momento de alta demanda de transacciones y oportunidades comerciales.

En la Figura 1 se puede observar la descripción del problema usando diagrama de Ishikawa.

Figura 1

Descripción del problema usando diagrama de Ishikawa



Objetivo general

Implementar un sistema de Business Intelligence que, mediante la integración de herramientas de análisis de datos, mejore la eficiencia en la toma de decisiones por parte del propietario de la agencia de courier.

Objetivos específicos

OE1: Identificar herramientas, técnicas y metodologías de inteligencia de negocios más utilizados en el negocio de Courier, para la implementación de un sistema de Business Intelligence.

OE2: Identificar el comportamiento del estado actual de los datos de una de las agencias de Courier en lo concerniente a la toma de decisiones.

OE3: Implementar un sistema de Business Intelligence en la agencia de Courier, como producto del análisis de literatura, identificando al que mejor se ajuste a las necesidades de una agencia de Courier, y que facilite la toma de decisiones teniendo una visión a corto plazo de las utilidades del negocio.

OE4: Evaluar los resultados del sistema Business Intelligence implementado, respecto de la toma de decisiones.

Preguntas de investigación

OE1-RQ1: ¿Cuáles son las estrategias planteadas en las agencias de Courier para fortalecer la toma de decisiones?

OE1-RQ2: ¿Qué tipos de soluciones tecnológicas son utilizadas para perfeccionar la toma de decisiones?

OE2-RQ1: ¿Cuál es la situación actual de la agencia de Courier en torno a las ganancias?

OE2-RQ2: ¿Cuál es la situación actual de la agencia de Courier frente a otras agencias del mismo ramo y que están cercanas respecto a las ganancias?

OE3-RQ1: ¿Cuál es la metodología de implementación de sistema de Business Intelligence que más se ajusta a las necesidades del problema bajo estudio?

OE3-RQ2: ¿La implementación del sistema de Business Intelligence se realizó bajo los criterios idóneos del problema bajo estudio?

OE4-RQ1: ¿La información resultante de la implementación del sistema de Business Intelligence es relevante para la toma de decisiones respecto al negocio?

OE4-RQ2: ¿La información resultante de la implementación del sistema de Business Intelligence mejoró la eficiencia en la toma de decisiones respecto al negocio?

Capítulo II

Estado del arte y marco teórico

Estado del arte

Para evaluar los trabajos relacionados se ha realizado un estudio del arte con el objetivo de responder las preguntas de investigación propuestas en el presente documento.

Criterios de inclusión

Los criterios de inclusión en los cuales se basó el presente análisis son:

- Artículos desarrollados en español o inglés.
- Artículos contentivos de conceptos y procedimientos relacionados con la implementación de sistemas de Business Intelligence.
- Artículos relacionados con agencias de Courier.

Criterios de exclusión

Como criterios de exclusión se tienen:

- Artículos que no tengan relación con la implementación de sistemas de Inteligencia de Negocios.
- Artículos referentes a organización y manejo de datos que no estén relacionados con el tema de investigación.

Se realizó un análisis de literatura para identificar herramientas, técnicas y metodologías de Inteligencia de Negocios más utilizados en el negocio de Courier,

escogiendo los textos que a criterio propio aportaron más información relevante a la implementación de una solución de Business Intelligence.

La literatura fue revisada y seleccionada de acuerdo con los criterios de inclusión y exclusión establecidos para tal fin, se utilizaron palabras claves, contenidos y resúmenes de los cuales se pudo elaborar como resultado los grupos de control relacionados con el tema.

La conformación de estos grupos de control que sirven de base para la investigación se muestra en la Tabla 1.

Tabla 1

Grupos de Control

Grupo	Título	Palabras clave
EC1	A Comprehensive Method for Data Warehouse Design	data warehouse, multidimensional modeling, design methods, UML.
EC2	Implementación de un Datamart como una solución de Inteligencia de Negocios para el área de logística de T-Impulso.	Datamart, Datawarehouse, Inteligencia de Negocios, logística, compras, proveedores.
EC3	Design and Implementation of an Enterprise Data Warehouse	Data warehousing, transactional databases, Extract, Transform and Load (ETL).

EC4	Business Intelligence: Strategy for competitiveness development in technology-based firms	Business Intelligence, Business competitiveness, Knowledge management, Business innovation, Decision making process.
EC5	El impacto de las herramientas de inteligencia de negocios en la toma de decisiones de los ejecutivos	Inteligencia de negocios, tecnologías de información, decisiones, información, ciencias de la computación.

Cadena de búsqueda

La cadena de búsqueda fue realizada para identificar las palabras clave que más se repiten o que tiene mayor frecuencia en los artículos analizados en cada uno de los grupos de control, arrojando como resultado lo mostrado en la Tabla 2.

Tabla 2

Cadena de búsqueda

Cadena de búsqueda	EC1	EC2	EC3	EC4	EC5	Frecuencia
("data" or "information") and ("analysis" or "investigation") and ("data warehouse") and		X	X	X		37

De acuerdo con los resultados obtenidos luego de aplicar los criterios de la cadena de búsqueda, se consiguen seleccionar ocho artículos que contienen información relevante para la investigación. A continuación, se presenta un resumen analítico de cada uno de ellos con la finalidad de verificar que dichos conceptos aportan una explicación cónsona y orientada al estudio.

Implementación de un datamart como solución de inteligencia de negocios bajo la metodología de Ralph Kimball para la empresa Corporación Corrales SAC –Lima; 2018, realizado por Villegas (2018). El autor propone que al implementar un datamart bajo la metodología de Ralph Kimball mejora el flujo de la información para la toma de decisiones a nivel operativo y estratégico, asegurando la automatización de los procesos de unificación, transformación, carga y análisis de datos para generar reportes de información y conocimiento de gran utilidad.

Implementación de un Data mart para mejorar la toma de decisiones del área de logística de Sedachimbote S.A.", realizado por Ramos (2018). El autor propone el desarrollo e implementación de un Datamart, a través de la teoría de inteligencia de negocios bajo la implementación de la metodología de desarrollo Ralph Kimball. Asegurando la mejora en la toma de decisiones en el área de logística de la empresa mejorando los tiempos para obtener información correspondiente a reportes de costos de aprovisionamiento, consulta de gasto presupuestal y los niveles de satisfacción por parte de la gerencia, administración y finanzas al obtener datos más fidedignos y en menor tiempo al momento de aplicar el ETL.

Design and Implementation of an Enterprise Data Warehouse realizado por Gulzar (2016). El autor menciona que para solucionar los problemas existentes en las empresas que manejan diferentes bases de datos donde dedican gran esfuerzo para la búsqueda eficaz de información existe la necesidad real de tener una ubicación única de almacenamiento e intercambio de datos donde los usuarios puedan utilizar fácilmente la información y tomar decisiones comerciales aplicando ETL, de esta manera ofrece una guía técnica que facilite el manejo de esta tecnología explicando como los usuarios pueden consumir los datos en el almacén de datos de la empresa a través de informes y otra inteligencia empresarial.

Intelligent dashboard to support for decision making in business Courier realizado por Pinto et al. (2016). Los autores proponen un estudio sobre un dashboard inteligente para apoyar la toma de decisiones en empresas de Courier con bases en las técnicas de IA, utilizando una red neuronal de tipo Perceptrón Multicapa entrenada por el algoritmo de error back-propagation para verificar y estudiar las previsiones y absentismo, asimismo incluyeron otro tipo de red neuronal artificial el mapa auto organizado de Kohonen para agrupamientos y así mejorar la visualización del dashboard.

Design and implementation of data warehouse with data model using survey-based services data realizado por Boon Keon y Nor Ezam (2014). Los autores proponen el diseño del almacén de datos para la creación de múltiples repositorios de datos y el diseño de un proceso de extracción, transformación y limpieza de los mismos, conocido como proceso ETL por sus siglas, con el cual se podrá alimentar dichos repositorios a partir de la fuente seleccionada. Asegurando que el desarrollo de múltiples bancos de datos habilitará una generación de informes mucho más sencilla al identificar una dimensión común entre los mismos.

Application of Data Warehouse and Data Mining in the Steel Enterprise Information Integration System realizado por Shenglei y Guoqing (2014). Proponen el uso de SQL Server para implementar la base de datos en una empresa de acero, para construir modelos de sistemas de datos de información.

Towards goal-driven access to process warehouse: Integrating goals with process warehouse for business process analysis realizado por Khurram y Jelena (2011). Esta investigación propone un enfoque en la integración que abarca dos niveles, conceptual y de implementación; el primero define los términos que se utilizan para relacionar los objetivos y procesar el almacén; el otro, se utiliza para implementar los conceptos y las relaciones con la intención de vincular los datos del almacén con los registros en las bases de datos.

Development of the data warehouse model for public authority's accounts in Croatia realizado por Sretenović, Kovačić y Jovanović (2016). Los investigadores describen el proceso de creación y organización del almacén de datos y el proceso ETL de extracción, limpieza y transformación de datos en un modelo de almacenamiento multidimensional, que se basa en el modelo de metadatos.

Conclusiones de trabajos relacionados

Al efectuar el análisis y estudio de los conceptos e investigaciones seleccionados, se evidencia la necesidad de centralizar todos los datos de los procesos que realiza continuamente la agencia de Courier bajo estudio, con la finalidad de poder aplicar y generar estrategias de Inteligencia de Negocios que contribuyan al crecimiento de la empresa.

Los criterios para la selección de herramientas de la implementación se basaron principalmente en el costo de licenciamiento, accesibilidad, complejidad en la configuración, facilidad en la administración de la herramienta, visualización de resultados y documentación disponible en el internet, de esta manera se seleccionó a Pentaho, Microsoft BI y Postgress.

La metodología de la solución selecciona fue Ralph Kimball la cual fue selecciona a través de un análisis detallado en el numeral 2.2.1.21

Categorización de las variables de investigación

La implementación de un sistema de Business Intelligence influye favorablemente en la toma de decisiones en la agencia de courier objeto del estudio. Se realizó un análisis de las variables de investigación, dando como resultado lo siguiente:

- Variable independiente: sistema de Business Intelligence.
- Variable dependiente: eficiencia.

Figura 2

Categorización de las variables



Categorización de la variable independiente

Inteligencia de Negocios

En la actualidad se le da un peso importante a la información como el principal activo de la organización, pero no basta con detentar la información; pues algo peor que no tenerla disponible es compilar en demasía y no saber qué hacer con ella, cómo organizarla y manipularla para beneficio en la empresa.

Business Intelligence (BI) es una metodología de gestión aplicadas por herramientas de software con el fin de generar ganancias en los procesos de toma de decisiones.

Los softwares de BI y otras tecnologías se usan para generar conocimiento que contribuyen para alcanzar el punto óptimo de las decisiones cuando estas lo ameritan. (Puerta, 2015)

La Inteligencia de Negocios proporciona valor a la empresa, abarca todas las funciones y todos los sectores. Esto, concierne a la totalidad de la organización y, más allá, a los clientes y proveedores. Sin embargo, la Inteligencia de Negocios sólo puede aportar valor al negocio cuando es utilizada eficazmente. Existe una correlación entre uso eficaz de la Inteligencia de Negocios y rendimiento de la empresa. No obstante, tener un mejor acceso a los datos no afecta al rendimiento de la empresa; la diferencia está en lo que hacen con los datos. (Howson, 2008)

La Inteligencia de Negocios permite el uso de datos de una manera eficiente y práctica, para evidenciar oportunidades en la toma de decisiones pertinentes. El objetivo final de las metodologías de BI es presentar una guía y conjunto de buenas prácticas en todo el entorno empresarial.

En su sentido más básico, la Inteligencia de Negocio proporciona a los directivos información sobre lo que sucede en la empresa. Sin esta herramienta tecnológica, los gerentes pueden estar operando sin ninguna visión hasta que se publican los números financieros. Con la Inteligencia de Negocios, la información es accesible de forma oportuna y flexible para proporcionar, el panorama de algunas actividades como:

- Cuál es el estado de las ventas en varias regiones y por varias líneas de productos.
- Si los gastos se ajustan al plan o se salen del presupuesto.
- Si las capacidades de los almacenes están en niveles óptimos.
- Si los canales de venta están donde deben estar.

Cuando una métrica en particular no está donde debería, la Inteligencia de Negocios permite a los usuarios explorar los detalles subyacentes para determinar el porqué de esa circunstancia y tomar medidas para mejorar la situación. En el pasado, los gerentes realizaban análisis de forma manual y no tenían flexibilidad para explorar por qué el negocio operaba de cierta manera.

Por ejemplo, muchas empresas usan BI para monitorear los gastos y asegurarse de que los costos no excedan los presupuestos. En lugar de esperar hasta el cierre del trimestre para descubrir que los gastos excesivos han reducido la rentabilidad, el acceso oportuno a estos datos permite a los gerentes identificar qué unidad de negocios está por encima del presupuesto y tomar medidas inmediatas para reducir costos. (Howson, 2008)

Se puede destacar que la Inteligencia de Negocios busca garantizar el acceso a la información de manera independiente, es decir, los datos a ser analizados y procesados pueden provenir de fuentes internas y externas, pudiendo los usuarios manipular la información a través de las distintas herramientas de análisis para el apoyo de toma de decisiones oportunas.

Datos

En un contexto amplio se puede decir que un dato es la información emitida por un dispositivo que debe ser procesada para que sea significativa. (Merriam-Webster, 2020)

Tabla 3

Tipos de datos.

Tipo De Dato	Descripción	Ejemplo
Estructurados	La mayoría de las fuentes de datos tradicionales son datos estructurados, datos con formato o esquema fijo que poseen campos fijos. Este tipo de datos se dividen en estáticos y dinámicos. Los datos estructurados son aquellos de mayor facilidad para acceder, gracias a su estructura bien especificada. Al tener un formato definido facilitan el trabajo.	Array, cadena de caracteres y registros, listas, pilas, colas, árboles, archivos.
No Estructurados	Son aquellos que no están organizados bajo algún patrón; tampoco son almacenados de manera	Imágenes según su resolución de píxeles, audio, video, fotografías, documentos impresos,

	relacional, o con base jerárquica de datos. Estos datos son los más difíciles de dominar por cualquier analista, pero debido al gran crecimiento han nacido herramientas que facilitan su manipulación.	imágenes digitales, mensajes de correo electrónico, mensajes instantáneos.
Semi Estructurados	Estos datos no tienen un formato definido, sino etiquetas que facilitan separarlos entre sí. La lectura de este tipo de datos requiere el uso de reglas complejas que determinan cómo proceder después de la lectura de cada pieza de información.	Archivos comprimidos, ejecutables binarios, paquetes TCP/IP, XML y otros lenguajes de marcado.

Fuente: (Joyanes, 2013)

Datos, información y conocimiento

Los datos pueden ser de carácter administrativo, logístico, comercial o científico, y deben ser procesados de forma sistemática y analítica para su aprovechamiento; al mismo tiempo, estos procesos deben ser capaces de transformar dichos datos en información y conocimiento valorables para la toma de decisiones.

Entre los datos, información y conocimiento se puede ver las siguientes diferencias:

- Datos: por lo general, los datos representan una codificación estructurada de entidades primarias únicas, así como de transacciones en las que intervienen dos o más entidades primarias. Por ejemplo, para un minorista los datos se refieren a entidades primarias como clientes, puntos de venta y artículos, mientras que los recibos de caja representan las transacciones comerciales.
- Información: Es el resultado de las acciones de extracción y elaboración de datos, y parece tener sentido en quienes la reciben en un ámbito específico.
- Conocimiento: Cuando la información se utiliza para tomar decisiones y desarrollar las acciones correspondientes se convierte en conocimiento. El conocimiento extraído de esta manera conducirá eventualmente a acciones dirigidas a resolver problemas complejos detectados. (Vercellis, 2009)

Proveer toda esta información a los encargados de toma de decisiones contribuye grandemente en que estas sean las más idóneas y las que generen mayor beneficio para la empresa.

Almacenes de datos

La tecnología de base de datos estándar es utilizada para recopilar, almacenar y procesar grandes cantidades de datos operativos que permiten ver más de cerca los almacenes de datos actuales e históricos como fuentes de información y de toma de las mejores decisiones empresariales.

Esta arquitectura opera de la siguiente manera: los datos se extraen, los datos son integrados, transformados y limpiados, para luego ser cargados en el almacén de datos, la información se estructura en cubos multidimensionales para responder a consultas dinámicas, logrando una buena presentación de los mismos con facilidad de interpretación.

Los almacenes de datos o Data Warehouse (DW por sus siglas en inglés), son conocidos como una gran base de datos analíticos que deriva su información para una variedad de sistemas de producción y está estructurado para la consulta, la presentación de informes y el análisis. Los almacenes de datos, que pueden incluir tanto datos de transacción como de no transacción, se utilizan como la base para la toma de decisiones, teniendo por objeto satisfacer los requisitos de una gran comunidad de usuarios empresariales. La información de un almacén de datos puede tener acceso a una variedad de herramientas fáciles de utilizar para el manejo de los datos de primera mano. (Khan, 2003)

Procesos de soporte para toma de decisiones

El análisis de la información facilita a las empresas la comprensión de sus negocios y a tomar decisiones empresariales a tiempo. (Mamani, 2018)

Las tecnologías de información se enfocan en la resolución de las necesidades del cliente y la manera de responder apropiadamente al cambiante ambiente de negocio. Las organizaciones deben ser conscientes de la importancia de sus datos e información; es necesario que la organización, a los fines de ser competitiva en el mercado realice una adecuada y oportuna gestión de la información interna y externa.

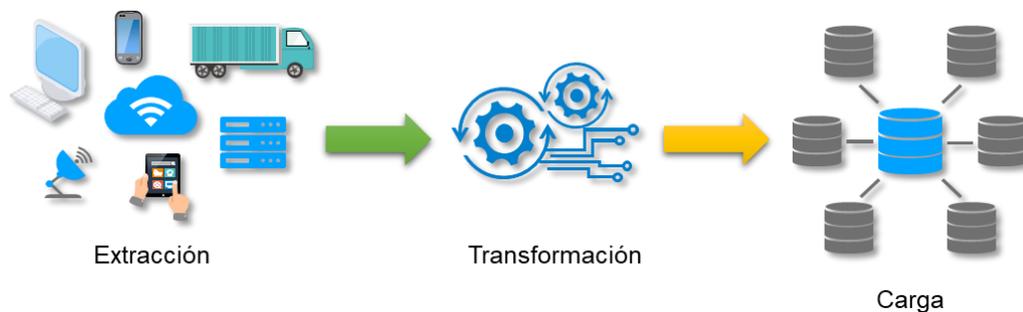
En el mismo orden de ideas, (Evelson, 2008), define el Business Intelligence como “la tarea que permite a las empresas combinar y analizar datos procedentes de fuentes diversas y obtener una visión integrada, completa y actualizada, con el fin de pronosticar lo que sucederá en su contexto y en la misma organización”.

Entre los procesos comprendidos en una solución BI están los siguientes:

- Proceso de extracción, transformación y carga de datos (ETL Process): es un proceso que se permite extraer datos, transformarlos según nuestras necesidades y cargarlos en los entornos destino. Los entornos origen y destino son usualmente bases de datos y/o ficheros, que pueden ser colas de mensajes, así como ficheros u otras fuentes de datos estructurados, semiestructurados o no estructurados (Curto, 2012). Este proceso se puede detallar en la Figura 3.

Figura 3

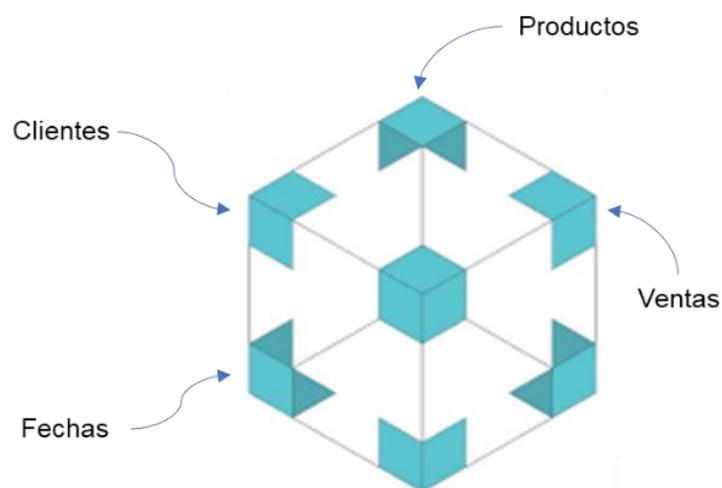
Proceso ETL



- Proceso analítico en línea (OLAP Process): es un sistema gestor de base de datos multidimensional cuyo proceso funciona de forma dedicada a la toma de decisiones, posee varias dimensiones visibles, jerarquizadas en varias granularidades y sigue un modelo lógico multidimensional (Puerta, 2015). La Figura 4 muestra un esquema del proceso.

Figura 4

Esquema del proceso OLAP



- Proceso de minería de datos (Data Mining): se refiere al proceso de utilización de los recursos de BI en la obtención de datos y generación de informes para ayudar en la toma de decisiones. Se hace la búsqueda de los datos relacionados con un determinado asunto y su cruce haciendo que las informaciones importantes sean identificadas en el posterior análisis de datos (Puerta, 2015). Esto se muestra en la Figura 5.

Figura 5

Proceso de minería de datos



- Proceso de reportería (Reporting): esta fase de difusión, consiste en entregar reportes. Esto implica el uso de herramientas BI para que los usuarios tengan la capacidad de manejar los datos de manera sencilla (Canales Inocencio, 2014), tal como se detalla en la Figura 6.

Figura 6

Proceso de reportería (reporting)



Metodologías, técnicas y herramientas para el desarrollo de soluciones de Inteligencia de Negocios

En la actualidad la sistematización de los procesos y la globalización de las comunicaciones obliga a las empresas a estar a la altura de las exigencias del entorno. Se debe tomar en cuenta que, la ejecución de proyectos de este tipo debe operarse gradualmente, en lo que se refiere a la Inteligencia de Negocios, con la finalidad de establecer prioridades y administración de recursos, físicos y económicos para ir de la mano con el crecimiento que se pretende obtener en la empresa.

Esos datos incluyen los aspectos comerciales, financieros y transacciones administrativas, así como rutas de navegación en la web, correos electrónicos, textos e hipertextos, por nombrar sólo algunos ejemplos. Su accesibilidad abre oportunidades prometedoras (Vercellis, 2009).

En ese mismo orden de ideas, se puede establecer una serie de pasos para la recolección y manejo de los datos generados por las operaciones dentro de la empresa de courier, los cuales darán origen

a información de vital importancia en la oportuna toma de decisiones con respecto a las ganancias y otros aspectos de interés para la organización.

Decisiones efectivas y oportunas

De acuerdo con Vercellis (2009), es de suma importancia saber que:

En las organizaciones complejas, públicas o privadas, las decisiones se toman de manera continua. Esas decisiones pueden ser más o menos críticas, tener efectos a largo o a corto plazo e implicar a personas y funciones de diversos niveles jerárquicos. La capacidad de estos trabajadores del conocimiento para tomar decisiones, tanto a nivel individual como comunitario, es uno de los principales factores que influyen en el rendimiento y la fuerza competitiva de una organización determinada.

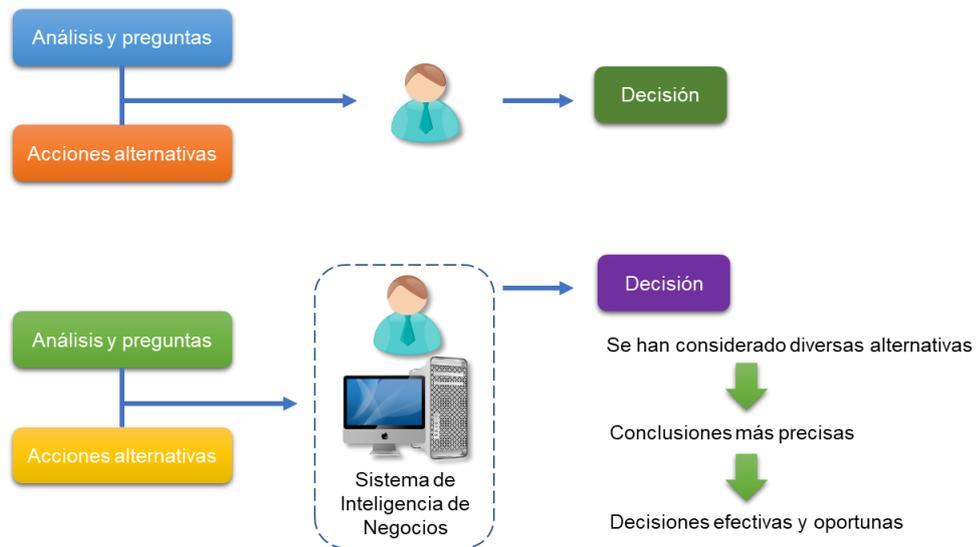
Comprender lo que significa la toma oportuna de decisiones conlleva a estar a la par de los competidores y poder superarlos en cuanto a la calidad de servicio, a la par que ofrecer mejores servicios de alta eficacia y efectividad, esto impulsa a la empresa a superar y solucionar problemas en su gestión, con lo que se minimizarán las posibles pérdidas, es decir, se aumentando las ganancias de forma óptima, que es el objetivo principal que se persigue. Se tiene que tomar en cuenta lo siguiente:

- Decisiones efectivas. La aplicación de métodos analíticos rigurosos permite a los encargados de adoptar decisiones basándose en información y conocimientos más fiables. Un examen a profundidad y comprensión de los métodos de trabajo conduce a un juicio más amplio de la lógica del proceso.
- Decisiones oportunas. las empresas operan en entornos económicos caracterizados por niveles de competencia cada vez mayores y de un gran dinamismo. En consecuencia, la capacidad de reaccionar rápidamente a las acciones de los competidores y a las nuevas condiciones del mercado es un factor crítico para el éxito o incluso la supervivencia de una empresa. (Vercellis, 2009)

En la siguiente figura se representan los beneficios que aporta una solución de BI.

Figura 7

Beneficios de un sistema de Inteligencia de Negocio



Fuente: (Vercellis, 2009)

En la medida en que se puedan disponer de herramientas que proporcionen información importante para la toma de decisiones oportunas, se tendrá ventaja considerable sobre los competidores, dando lugar a un crecimiento sostenible en el tiempo, garantizando un servicio óptimo y de calidad.

Se puede confiar en que los sistemas de Inteligencia de Negocios contribuyen considerablemente en las decisiones de posicionamiento de mercado, creación de estrategias de crecimiento organizacional y reducción de riesgos económicos debido a que con este tipo de sistemas se pueden evaluar y analizar un gran número de alternativas y distinguir las que son favorables y las que no lo son.

El rol de los modelos matemáticos

Los sistemas de Inteligencia de Negocios proporcionan a los encargados de toma de decisiones información y conocimiento a partir

de los datos extraídos a través de la aplicación de modelos y algoritmos matemáticos.

Esta forma de extracción de información y conocimiento reduce en gran parte el cálculo de totales y porcentajes que son de interés en la organización, los cuales pueden ser representados a través de histogramas sencillos de interpretar, al mismo tiempo, se dispone de otros tipos de gráficos para realizar análisis más complejos de acuerdo con las necesidades o problemas de estudio.

En términos generales, adoptar un sistema de Inteligencia de Negocios, contribuye a tener un acercamiento científico y racional para el manejo de soluciones en las empresas por muy complejas que estas sean.

En relación a la idea anterior, se puede señalar que, un típico análisis aplicando Inteligencia de Negocios tiene las siguientes características:

- Primero, se identifican los objetivos del análisis y el rendimiento de los indicadores que serán usados para evaluar alternativas.
- Los modelos matemáticos son desarrollados con la finalidad de cruzar gran cantidad de variables de control y determinar sus relaciones, parámetros y métricas de evaluación.
- Finalmente, el sistema de análisis de los datos evalúa los efectos en el rendimiento determinado por la variación de las variables de control y los cambios de parámetros.

Como objetivo principal, se pretende incrementar la efectividad en la toma de decisiones, ya que, los modelos matemáticos permiten a los decisores estar más enfocados en los aspectos más relevantes a ser analizados, induciendo a un entendimiento profundo de las

situaciones presentadas. El conocimiento adquirido, al momento de la aplicación de estos modelos, conlleva a que puedan ser implementados en organizaciones con las mismas características, dando resultados muchos más precisos a los producidos por la toma de decisiones de forma empírica (Vercellis, 2009).

Arquitectura de un sistema de Inteligencia de Negocios

La arquitectura de un sistema de Inteligencia de Negocios, la cual se observa en la Figura 8, se conforma principalmente por tres componentes que se describen a continuación:

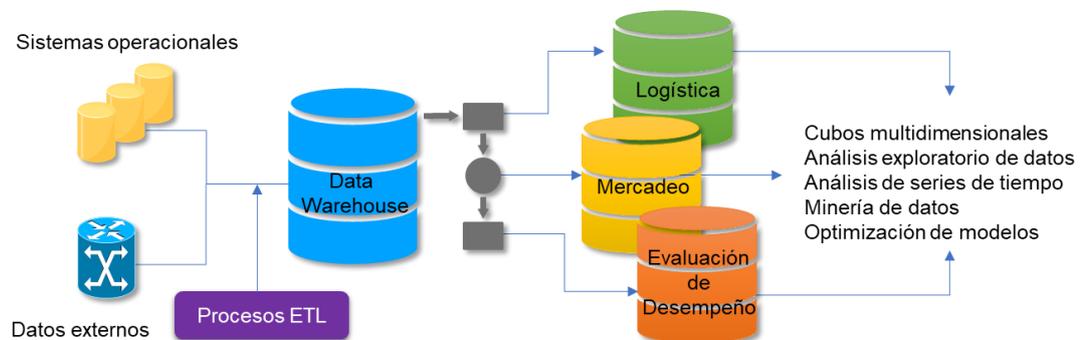
- Fuentes de datos: en una primera etapa, es necesario reunir e integrar los repositorios de datos y clasificarlos en primarios y secundarios. Las fuentes consisten en la mayor parte de los datos pertenecientes a los sistemas y operaciones de la empresa, deben ser incluidos datos no estructurados como documentos, emails y al mismo tiempo datos de fuentes externas. Generalmente hablando, se requiere de un gran esfuerzo para lograr la unificación e integración de las diferentes fuentes de datos.
- Data Warehouse y Data Marts: utilizando herramientas y métodos de extracción y transformación de datos conocida como extracción, transformación y carga, los datos originados desde las distintas fuentes son almacenados en repositorios de datos orientados y construidos para el análisis con herramientas de Inteligencia de Negocios, estos repositorios son conocidos como Data Warehouse. En las próximas secciones se describen con más detalle.
- Metodologías de la Inteligencia de Negocios: finalmente, los datos son extraídos y usados para alimentar modelos matemáticos y

metodologías de análisis que sirven de apoyo a los encargados de las tomas de decisiones. En un sistema de Inteligencia de Negocios son muchas las aplicaciones que deben ser implementadas para abarcar los procesos dentro de la empresa. Entre ellas se tienen:

- Cubos multidimensionales.
- Análisis exploratorio de los datos.
- Análisis de series de tiempo.
- Minería de datos.
- Optimización de modelos.

Figura 8

Arquitectura típica de un sistema de Inteligencia de Negocios



Fuente: (Vercellis, 2009)

A continuación, se presenta en la Figura 9, un esquema piramidal de los componentes principales de un sistema de Inteligencia de Negocios. Anteriormente, se explicaron los dos primeros niveles de dicha pirámide, por consiguiente, se detallan los bloques sucesivos a estos con lo que se complementa el concepto de dichos sistemas.

Figura 9

Componentes de un sistema de Inteligencia de Negocios



Fuente: (Vercellis, 2009)

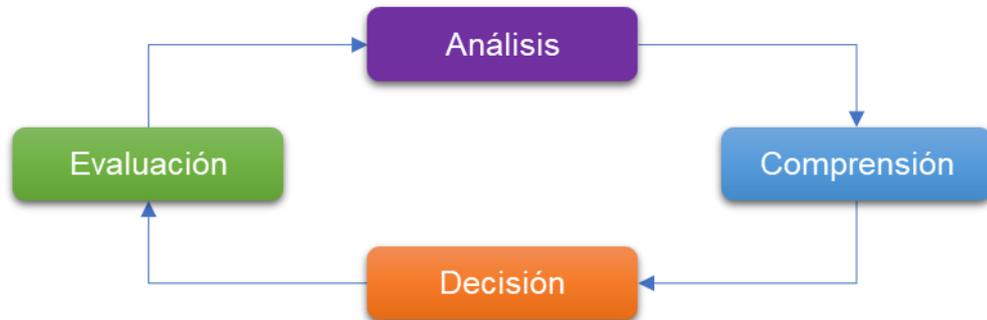
- Exploración de datos: en este tercer nivel de la pirámide están las herramientas para realizar un análisis de Inteligencia de Negocios pasiva, que consiste en los modos de consulta y sistemas de presentación de informes. Se denominan metodologías pasivas porque se pide a los responsables de la toma de decisiones que generen hipótesis o definir los criterios de extracción de datos, y luego utilizar las herramientas de análisis para encontrar respuestas y confirmar su visión original.
- Minería de datos. el cuarto nivel incluye metodologías activas de Inteligencia de Negocios, cuyo propósito es la extracción de información y conocimientos a partir de los datos.

- Optimización: Permiten determinar la mejor solución de un conjunto de acciones alternativas, bastante extenso e incluso infinito.
- Decisiones: por último, la parte superior de la pirámide corresponde a la elección y la adopción efectiva de una decisión específica.

A medida que se avanza hacia la cima de la pirámide, los sistemas de Inteligencia de Negocios ofrecen herramientas de apoyo más avanzadas de tipo activo. En la parte inferior, las competencias requeridas son proporcionadas a los administradores de bases de datos. Analistas y expertos en modelos estadísticos son responsables de las fases intermedias. Por último, las actividades de los responsables de la toma de decisiones aparecen en la parte superior. (Vercellis, 2009)

Ciclo de la Inteligencia de Negocios

Cada análisis de Inteligencia de Negocios sigue su propio camino según el campo de aplicación, la actitud personal de los responsables de la toma de decisiones y las metodologías analíticas disponibles. Sin embargo, es posible identificar una trayectoria cíclica ideal que caracteriza la evolución de un análisis típico de Inteligencia de Negocios, como se muestra en la Figura 12.

Figura 10*Ciclo de la Inteligencia de Negocios*

Fuente: (Vercellis, 2009)

A continuación, se describen cada uno de las etapas del ciclo de la Inteligencia de Negocios:

- **Análisis:** durante esta fase, es necesario reconocer y explicar con precisión el problema en cuestión. Los encargados de la toma de decisiones deben crear una representación mental del mismo, identificando los factores críticos que se perciben como más relevantes. La disponibilidad de metodologías de Inteligencia de Negocios puede ayudar en esta etapa, al permitir adoptar decisiones y desarrollar rápidamente diversas vías de investigación. La exploración de los cubos de datos en un análisis multidimensional permite modificar hipótesis de manera flexible y rápida, hasta llegar a un esquema de interpretación que se considere satisfactorio.
- **Comprensión:** la segunda fase permite a los responsables de la toma de decisiones comprender más profundamente el problema, a menudo a nivel causal. La información obtenida en la fase de análisis se transforma en conocimiento durante la fase de

comprensión. Por una parte, la extracción del conocimiento puede ocurrir debido a la intuición de los encargados de toma de decisiones y, por lo tanto, basarse en su experiencia y posiblemente en información no estructurada que disponen. Por otra parte, los modelos de aprendizaje inductivo, pueden resultar muy útiles durante esta fase, en específico cuando se aplican a datos estructurados.

- **Decisión:** durante esta tercera fase, los conocimientos obtenidos como resultado de la comprensión se convierten en decisiones y acciones. La disponibilidad de metodologías de Inteligencia de Negocios permite que las fases de análisis y comprensión se ejecuten con mayor rapidez, de modo que se puedan adoptar decisiones más eficaces y oportunas, útiles a las prioridades estratégicas de una organización determinada. Esto conduce a una reducción general del tiempo de ejecución del ciclo de análisis – decisión – acción – revisión y a un proceso de adopción de decisiones de mejor calidad.
- **Evaluación:** por último, la cuarta fase del ciclo de Inteligencia de Negocios implica la medición y evaluación del rendimiento. A continuación, deben elaborarse distintas métricas que no se limiten exclusivamente a los aspectos financieros, sino que tengan en cuenta los principales indicadores de rendimiento definidos para los diferentes departamentos de la empresa. (Vercellis, 2009)

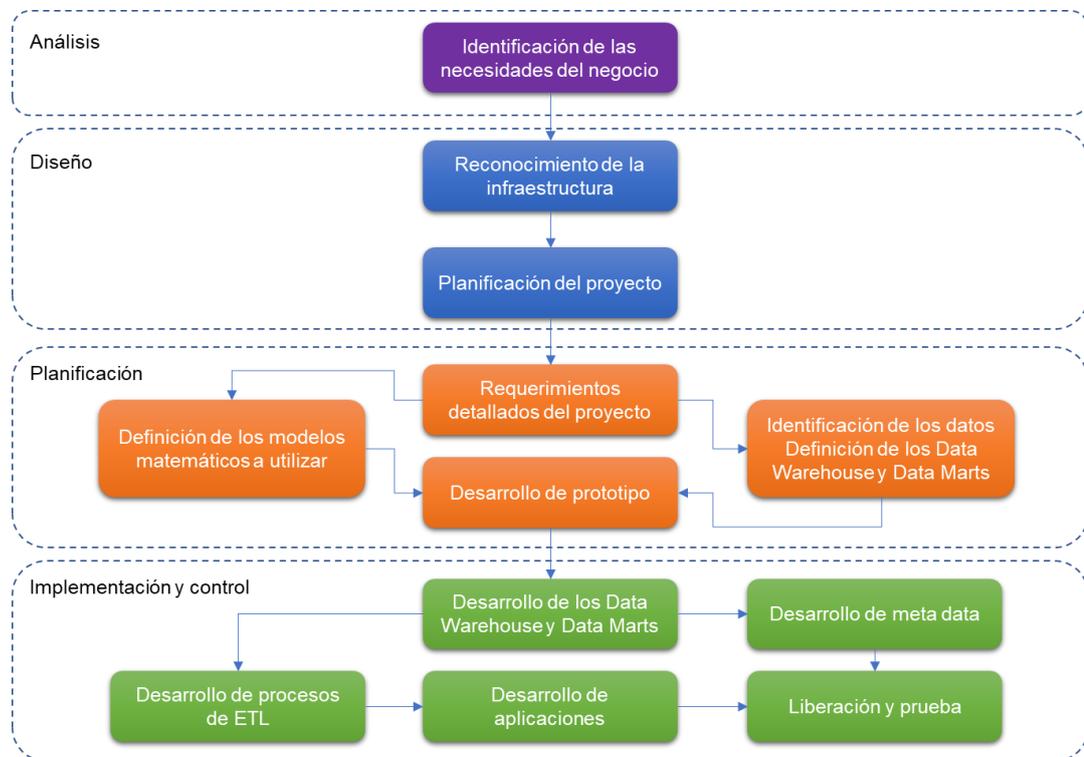
Desarrollo de un sistema de Inteligencia de Negocios

El desarrollo de un sistema de Inteligencia de Negocios puede asociarse a un proyecto, con un objetivo final específico, los tiempos y costos de desarrollo previstos, y la utilización y coordinación de los recursos necesarios para llevar a cabo las actividades planificadas.

La Figura 11 muestra el ciclo de desarrollo típico de una arquitectura de Inteligencia de Negocios. Es evidente que, el camino específico seguido por cada organización puede diferir del que se indica en la figura, es decir, se deben considerar adaptaciones según los procesos específicos.

Figura 11

Fases del desarrollo de un sistema de Inteligencia de Negocios



Fuente: (Vercellis, 2009)

En este orden de ideas, se describe cada una de las fases que sigue el desarrollo de un sistema de Inteligencia de Negocios:

- **Análisis:** se deben identificar cuidadosamente las necesidades de la organización en relación con el desarrollo de un sistema de Inteligencia de Negocios. Esta fase preliminar se lleva a cabo generalmente mediante una serie de entrevistas a los trabajadores de la organización. Es necesario describir claramente los objetivos y prioridades generales del proyecto, así como establecer los costos y beneficios derivados del desarrollo del sistema que se pretende implementar.
- **Diseño:** incluye dos procesos y tiene como objetivo derivar un plan provisional de la arquitectura general, teniendo en cuenta cualquier desarrollo en un futuro próximo y la evolución del sistema en poco tiempo. En primer lugar, es necesario hacer una evaluación de la información existente de la infraestructura. Se utiliza la clásica metodología de gestión de proyectos, que permite establecer un plan al efecto, identificando las fases de desarrollo, las prioridades, los plazos de ejecución previstos y los costos, junto con las funciones y los recursos necesarios.
- **Planificación:** en esta etapa se incluye un proceso en el que se definen y describen con mayor detalle las funciones del sistema a implementar. Posteriormente, se evalúan los datos existentes, así como otros que podrían recuperarse externamente. Esto permite diseñar las estructuras de información. Por último, es conveniente crear un prototipo de sistema, de bajo costo y con capacidades limitadas, a fin de descubrir de antemano cualquier discrepancia entre las necesidades reales y las especificaciones del proyecto.

- Implementación y control: esta última fase consta de cinco procesos principales. En primer lugar, se desarrolla el Data Warehouse y cada Data Mart específico. Estos representan las infraestructuras de información que alimentarán al Sistema de Inteligencia de Negocios. A fin de explicar el significado de los datos contenidos en el Data Warehouse y las transformaciones aplicadas anteriormente a los datos primarios, se debe crear un archivo de metadatos, igualmente, se establecen procedimientos para extraer y transformar los datos existentes en las fuentes, cargándolas en el Data Warehouse y los Data Mart, es decir, los procesos de ETL. El siguiente paso tiene por objeto desarrollar las aplicaciones básicas de Inteligencia de Negocios que permitan los análisis previstos que se llevarán a cabo. Finalmente, el sistema es liberado para ser probado y usado. (Vercellis, 2009)

Data Warehouse (DW) y On Line Analytical Proccesing (OLAP)

Una de los más importantes métodos de integración de fuentes de datos está basado en la arquitectura de Data Warehouse. Dentro de este modelo, los datos provienen de diferentes fuentes, son extraídos, filtrados, combinados y guardados en un repositorio de datos central llamado Data Warehouse. Este repositorio es incluso enriquecido por datos históricos y resúmenes de informes.

Un Data Warehouse es analizado por el procedimiento llamado On Line Analytical Proccesing, OLAP por sus siglas en inglés, cuyo propósito es descubrir tendencias, patrones de comportamiento y anomalías escondidas entre los datos. Los resultados obtenidos de estos análisis

forman la base para realizar la toma de decisiones. Las aplicaciones OLAP analizan los datos mediante complejas consultas que van desde unas pocas a varias operaciones de unión, filtrado, agrupación y agregación de los mismos. Como estas consultas son muy complejas y a menudo leen terabytes de datos, su ejecución puede llevar una gran cantidad de minutos, horas o incluso días. Por lo tanto, una cuestión clave es la eficiencia del DW.

Las soluciones bien desarrolladas para este problema se basan en vistas materializadas y en la reescritura de consultas, así como en estructuras avanzadas de índices. (Wrembel & Koncilia, 2007)

Las siguientes características siguen distinguiendo a las herramientas de OLAP de las herramientas de consulta e información de negocios:

- **Multidimensional:** los usuarios analizan valores numéricos de diferentes dimensiones como producto, tiempo y geografía. Un informe, por otro lado, puede ser unidimensional, como la lista de precios de productos en un momento dado.
- **Permanentemente rápido:** a medida que los usuarios navegan en diferentes dimensiones se puede desglosar de año a trimestre, esperando, por ejemplo, 24 horas, 24 minutos, o incluso 24 segundos para una obtener una respuesta, lo cual no es aceptable.
- **Altamente interactivo:** la exploración es la forma en que los usuarios interactúan con los datos de OLAP. Permite a estos ver la información desde diferentes perspectivas. De igual manera, faculta a los usuarios para filtrar los datos dentro de estas dimensiones como las ventas en alguna ciudad específica y luego para otra diferente.

- Niveles variables de integración: para asegurar tiempos de consulta predecibles, los procesos OLAP agregan los datos de diferentes maneras.
- Cálculos multidimensionales: con las múltiples dimensiones vienen cálculos más complejos. Estos cálculos a menudo deben realizarse en un orden específico e implican números de entrada que los usuarios podrían no ver nunca. Por otro lado, los informes detallados suelen basarse en simples subtotales o cálculos de valores que se muestran en el propio informe. (Howson, 2008)

Data mart

Un Data mart es un subconjunto de los datos procedentes de un Data Warehouse. Un Data mart puede atender las necesidades de una unidad comercial, función, proceso o aplicación. Debido a que está alineado con un requisito empresarial determinado, algunas empresas pueden querer saltarse el DW y construir un Data mart independiente.

Pentaho

El verdadero valor de los sistemas de BI reside en su uso para tomar decisiones que conduzcan a una mayor rentabilidad, menores costos, mayor eficiencia, crecimiento de la participación en el mercado, mayor satisfacción de los empleados o cualesquiera otros que sean los objetivos de su organización. El beneficio añadido de utilizar la solución de código abierto para lograr esto es la gran relación calidad – precio y flexibilidad del software. Esto permite a cualquier organización, con o sin fines de lucro, grande o pequeña, implementar y utilizar este tipo de herramientas con el fin de tomar mejores decisiones.

Pentaho es una suite de Inteligencia de Negocios en lugar de un producto único; se compone de una colección de programas que trabajan juntos para crear y entregar soluciones de Inteligencia de Negocios. Algunos de estos componentes proporcionan funcionalidades muy básicas, como la autenticación de usuarios o la gestión de la conexión a bases de datos. Otros componentes ofrecen funcionalidades a un nivel más alto, como la visualización de datos mediante tablas y gráficos.

A menudo, pero no siempre, los componentes que ofrecen una funcionalidad de alto nivel se basan en otros que ofrecen una funcionalidad de bajo nivel. Como tal, la colección de programas que forma el conjunto completo puede verse como una serie de componentes.

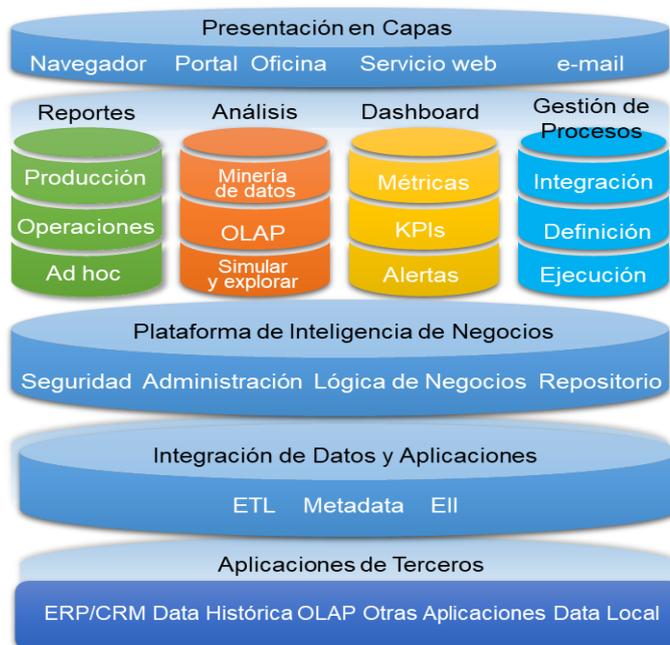
Las principales áreas funcionales de BI, información, análisis, cuadros de mando y gestión de procesos, constituyen la capa intermedia de estos componentes, mientras que la plataforma de BI en sí misma ofrece características básicas de seguridad y administración.

La integración de los datos completa el grupo de módulos y es necesaria para obtener datos de varios sistemas de origen en un entorno de DW compartido. (Bouman & Van Dongen, 2009)

En la Figura 12 se puede observar las capas y componentes de la suite de Pentaho.

Figura 12

Pentaho BI. Capas y componentes



Fuente: (Bouman & Van Dongen, 2009)

Microsoft Power BI

Si bien el Power BI suele clasificarse como una herramienta de visualización, la realidad es que no es una herramienta única, sino más bien un conjunto de herramientas y servicios interrelacionados que forman un completo ecosistema propio de Inteligencia de Negocios. Abarca desde las fuentes de datos hasta el modelado, el análisis y la visualización. Igualmente, este ecosistema incluye componentes que son propios de Power BI, así como otras tecnologías de Microsoft que interoperan con él, e integraciones de terceros.

Es evidente que existen herramientas de apoyo a la Inteligencia de Negocios de gran amplitud y alcance, las cuales van a contribuir en la generación de reportes e indicadores que conlleven a una mejor y oportuna toma de decisiones por parte de los encargados para esto. Los aportes tecnológicos de estas herramientas son incalculables en la

gestión de la organización y se deben aprovechar al máximo para potenciar a la empresa en este creciente y competitivo mercado.

Sistema de Gestión de Base de Datos

PostgreSQL (www.postgresql.org) es un sistema de gestión de bases de datos objeto – relacional (RDBMS por sus siglas en inglés), distribuido bajo licencia BSD.

PostgreSQL para Inteligencia de Negocios

Entre las características más relevantes de este motor de datos están las siguientes:

- Se integra perfectamente con herramientas externas y aplicaciones de Minería de Datos, análisis OLAP y de generación de reportes.
- Maneja procedimientos almacenados agregando flexibilidad al DW (se puede analizar y transformar los datos dentro del manejador).
- Utiliza el concepto de Tablespace (contenedor de objetos como tablas e índices), pueden crearse o modificarse en cualquier momento permitiendo mejorar en escalabilidad y performance. A su vez, los objetos se organizan en grupos lógicos denominados esquemas.
- Soporta particionamiento horizontal a través de herencia, la tabla total es abstracta y heredan las particiones conteniendo los datos.
- Pueden implementarse vistas con, por ejemplo, agregación sobre las tablas de hechos (fact).

- Existen herramientas externas para ETL comerciales y de código libre, por ejemplo, Kettle que es parte de Pentaho Data Integration.
- Utiliza estándares ODBC y/o JDBC

Metodología de Bill Inmon

La metodología de Bill Inmon presenta un diseño descendente o top-down: El Data Warehouse es una parte de toda la solución de Inteligencia de Negocios. Una empresa tiene un Data Warehouse, y los Data Marts, tienen como fuente de información al Data Warehouse. La información en el Data Warehouse se almacena en tercera forma normal.

Estructura: la construcción de un Data Warehouse tiene dos componentes principales; el diseño de la interfaz de los sistemas operativos y el diseño del almacén de datos en sí mismo. Sin embargo, el término “diseño” no es del todo preciso porque sugiere que los elementos pueden planificarse de antemano. Los requisitos del almacén de datos no pueden conocerse hasta que esté parcialmente cargado y en uso, y los enfoques de diseño que han funcionado anteriormente no necesariamente serán suficientes para los siguientes almacenes de datos, ya que estos se construyen de forma heurística, en la que una fase de desarrollo depende totalmente de los resultados obtenidos en la fase anterior. En primer lugar, se carga una parte de los datos, a continuación, se utiliza y se examina por el analista del sistema de soporte de decisión. Luego, basándose en los comentarios del usuario final los datos se modifican y/o se añaden otros datos para después construir otra parte del almacén de datos, y así sucesivamente. Este proceso de retroalimentación continúa durante todo el ciclo de vida del Data Warehouse.

Modelado de datos: existen tres niveles de modelado de datos, los cuales son, el modelado de alto nivel, llamado diagrama de relación de entidades o ERD por sus siglas en inglés, el modelado de nivel medio o conjunto de elementos de datos o DIS por sus siglas en inglés y el modelado de bajo nivel, llamado modelo físico.

- La dirección y el número de las flechas indican la cardinalidad de la relación. De este modo, se minimizan las dependencias transitivas. Las entidades que se muestran en el nivel de la ERD están en el nivel más alto de abstracción, definiendo cuáles entidades pertenecen al ámbito del modelo y cuáles entidades no están determinadas, esto se denomina ámbito de integración, el cual a su vez define los límites del modelo de datos y debe definirse antes de que comience el proceso de modelado.
- El modelo de datos de nivel medio es establecido luego de creado el modelo de datos de alto nivel. Dentro del modelo de nivel medio se encuentran cuatro elementos básicos:
 1. Una agrupación primaria de datos: la cual existe una vez y sólo una vez, para cada entidad.
 2. Una agrupación secundaria de datos: esta contiene atributos de datos que pueden existir varias veces para cada entidad.
 3. Un conector: este indica las relaciones de los datos entre las principales entidades. El conector relaciona los datos de una agrupación con otra. Una relación identificada en el nivel ERD da lugar a un registro en el nivel DIS.
 4. Tipos de datos: los tipos de datos identifican los supertipos de datos y a los subtipos de datos.

Estos cuatro elementos de modelado de datos se utilizan para identificar los atributos de los datos en un modelo de datos y la relación

entre esos atributos. Cuando se identifica una relación en el nivel ERD, se manifiesta mediante un par de relaciones conectoras en el nivel DIS.

- El modelo de datos físico: Se incluye la optimización de las características de rendimiento. Con el almacén de datos, el primer paso del diseño es decidir la granularidad y la partición de los datos. Este paso es crucial y por supuesto, la estructura clave se modificada para añadir el elemento del tiempo, para el que cada unidad de datos es relevante.

El modelo de datos y el desarrollo iterativo: en todos los casos, el almacén de datos se construye mejor de forma iterativa. Esto significa que primero se construye una parte del almacén de datos, y luego se construye otra parte del mismo y así sucesivamente. A continuación, se exponen algunas razones por las que el desarrollo iterativo es importante:

- El historial de éxitos de la industria así lo sugiere.
- El usuario final es incapaz de articular muchos requisitos hasta que se realiza la primera iteración.
- La dirección no se comprometerá plenamente hasta que al menos algunos resultados reales sean tangibles y evidentes.
- Los resultados visibles deben verse rápidamente.

Normalización y desnormalización: en el nivel más bajo del diseño de la base de datos, puede producirse una ligera desnormalización cuando toda la organización mira los datos de la misma manera. Algunas técnicas de desnormalización ligera de los datos incluyen la creación de matrices de datos, la creación de datos redundantes con criterio y la elaboración de índices convencionales.

La estructura básica del registro del Data Warehouse es la que contiene una marca de tiempo, una clave, datos directos y datos secundarios. El diseño de bases de datos de los almacenes de datos, de

una forma u otra, siguen este sencillo patrón. Las tablas de referencia deben colocarse en el almacén de datos y gestionarse en función del tiempo, como cualquier otro dato.

Esquema de datos dimensional: el esquema estrella es una técnica de diseño de bases de datos que a veces se aplica erróneamente al entorno del Data Warehouse. El enfoque multidimensional de este, es un enfoque en el que el diseño de la base de datos se basa en las ocurrencias de datos dentro de una entidad y en cómo se accederá a esos datos. Se determina, que el esquema de estrella corresponde como base del diseño del Data Mart. (Inmon, 2005)

Metodología de Ralph Kimball

La metodología de Ralph Kimball presenta un diseño ascendente o bottom – up: el Data Warehouse es el conglomerado de todos los Data Marts en la empresa; la información siempre se almacena en el modelo dimensional.

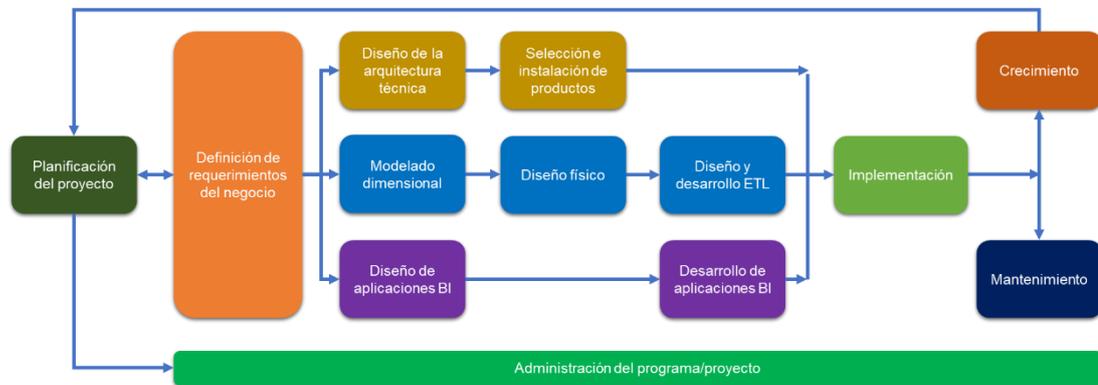
No indica la existencia física de un Data Warehouse, sino que argumenta que es la unión de los Data Marts que han seguido el proceso de conformación de dimensiones.

Kimball manifiesta que los datos primitivos pueden transformarse en información con sentido.

Esta metodología es de las más utilizadas para el desarrollo de soluciones de BI y DW y se fundamenta en los aspectos que se presentan a continuación.

Figura 13

Ciclo de vida metodología de Ralph Kimball



Fuente: (Kimball & Ross, 2016)

- Planificación y gestión del programa/proyecto: se desarrolla la hoja de ruta para la puesta en marcha del programa/proyecto, incluyendo el alcance, la justificación y la dotación de personal.
- Requerimientos del negocio: esta etapa es muy importante en el ciclo de vida de la metodología propuesta por Kimball, ya que estos datos impulsan la mayoría de las decisiones previas y posteriores. Se identifican las principales oportunidades en toda la empresa, se priorizan en función del valor empresarial y la viabilidad.
- Fase tecnológica: La vía tecnológica comienza con el diseño de la arquitectura del sistema para establecer una lista de capacidades necesarias, seguido de la selección e instalación de productos que satisfagan esas necesidades.

- Fase de datos: Es el diseño de un modelo dimensional para abordar los requisitos de la empresa, teniendo en cuenta las realidades de los datos subyacentes. Se realiza un modelado dimensional de los datos. Los modelos dimensionales pueden ser instanciados en bases de datos relacionales. La matriz de bus es muy importante porque sirve simultáneamente de guía técnica, gestión y de foro de comunicación con los ejecutivos. El ciclo de vida describe el proceso ETL agrupados en cuatro operaciones principales: extraer los datos de la fuente, realizar transformaciones y conformar la entrega de los datos a la capa de presentación y la gestión del entorno y los procesos de ETL.
- Fase de la Inteligencia de Negocios: en esta etapa algunos miembros del proyecto están sumidos en la tecnología y los datos, otros se centran en identificar y construir aplicaciones de BI.
- Implementación, mantenimiento y crecimiento: las tres vías del ciclo de vida de la metodología convergen en el despliegue, reuniendo la tecnología, los datos y las aplicaciones de BI.

Por último, es imprescindible tener en cuenta que el objetivo general debe ser la aceptación por parte del negocio, los resultados de la solución DW/BI para apoyar la toma de decisiones. Este objetivo debe permanecer en mente durante todo el ciclo de vida de diseño, desarrollo y despliegue de cualquier sistema DW/BI. (Kimball & Ross, 2016)

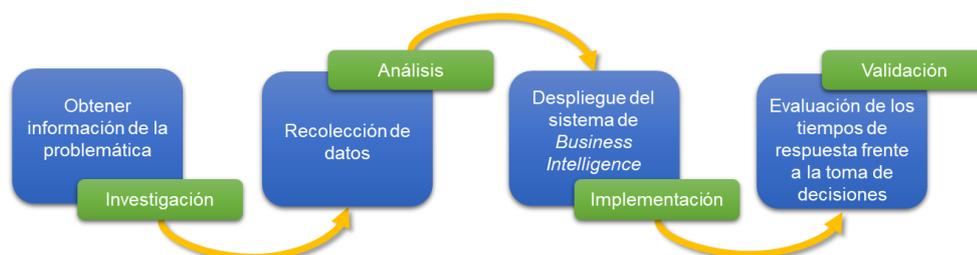
En el mismo orden de ideas, a continuación, se describen las etapas complementarias para el desarrollo del presente proyecto, de acuerdo al problema planteado:

- Fase 1: análisis inicial de literatura con el propósito de identificar herramientas, técnicas y metodologías de Inteligencia de Negocios. En esta fase se pretende seleccionar aquellas que se ajusten mejor a las necesidades actuales de la empresa en estudio, con lo cual se obtendrán mejores resultados de acuerdo con los objetivos planteados.
- Fase 2: exploración del estado actual de los datos y herramientas utilizadas para la toma de decisiones en la agencia de courier para identificar el problema o dificultad principal, con la finalidad de poner en práctica todo lo encontrado en la primera fase de análisis.
- Fase 3: implementar un sistema de Business Intelligence identificando la mejor solución que optimice o influya favorablemente en la toma de decisiones.
- Fase 4: validación de los resultados de la implementación del sistema de Business Intelligence, en la cual se evalúa y certifica que se logren los resultados deseados, tomando en cuenta todo lo utilizado en las fases previas, con la finalidad de obtener una visión objetiva del negocio.

Cada una de las cuatro fases descritas se pueden observar en la Figura 8.

Figura 14

Fases de la metodología.



Diseño y desarrollo del sistema de BI

De acuerdo con la metodología escogida para el diseño, desarrollo e implementación de un adecuado sistema de BI que se ajuste a las necesidades del negocio, se tomó en cuenta considerar cómo hacer frente a los nuevos requerimientos de las empresas.

En primer lugar, para el diseño del sistema de BI se deben tomar en cuenta los siguientes aspectos:

Data Marts y modelado dimensional

Existen un número de criterios empresariales a tomar en cuenta para la implementación de un sistema de BI que sirva de guía para lograr la mejor gestión. Están íntimamente relacionados con un adecuado diseño de un Data Warehouse acorde a las necesidades, los Data Marts y el modelado dimensional. Se consideran estos dos temas conjuntamente donde todos los criterios se pueden abordar simultáneamente.

Se toman en cuenta dos principios clave para guiar los diseños de DW:

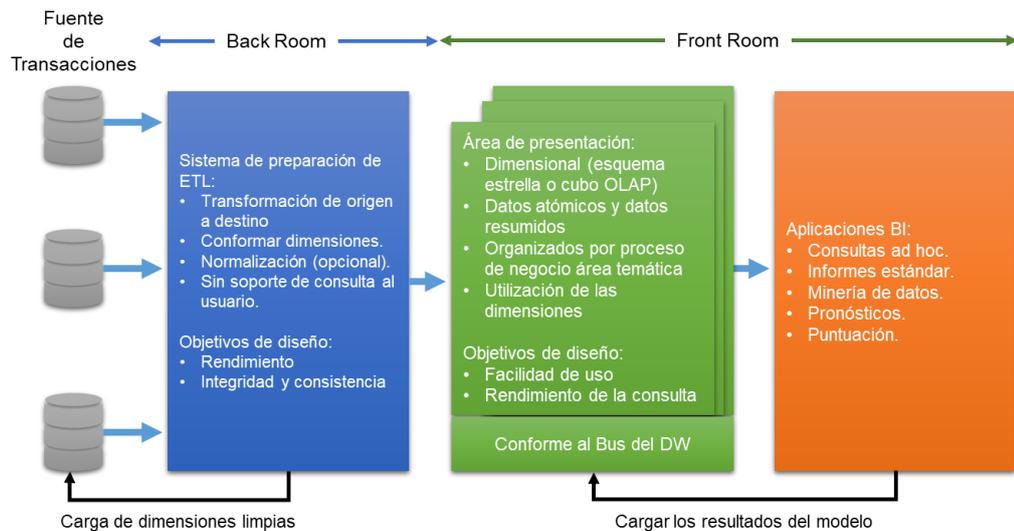
- Separar sus arquitecturas, donde se traza una línea muy clara entre el Back Room y el Front Room. En el Back Room es donde se extraen, transforman y cargan los datos, y en el Front Room es donde se ponen a disposición los datos para su presentación.
- Construir los Data Marts orientados a la presentación en torno a modelos dimensionales, no a modelos de entidad relación (ER) normalizados.

La Figura 15 muestra un esquema horizontal típico de un entorno de Data Warehouse. En el extremo izquierdo se ven los sistemas tradicionales orientados a las transacciones. La responsabilidad del DW

comienza cuando se extraen los datos de los sistemas heredados y se llevan a la zona de preparación de ETL del Back Room.

Figura 15

Esquema horizontal de un entorno de Data Warehouse.



Fuente: (Kimball & Ross, 2016)

El área de preparación de datos es la operación completa del Back Room del DW, en el que se limpian, depuran, combinan, clasifican, buscan, se añaden claves, eliminan duplicados, se reúnen los datos, se archivan y exportan. Los datos que llegan al área de preparación suelen no estar limpios, mal formados y en un formato de archivo plano. Si se tiene suerte, los datos llegan en una tercera forma normal prístina, pero eso no es común. Cuando se haya terminado de limpiar y reestructurar los datos, se pueden dejar en formato plano o almacenarlos en tercera forma normal. Las representaciones relacionales en tercera forma normal son lo ideal, pero son el resultado final de un trabajo arduo en formatos no relacionales.

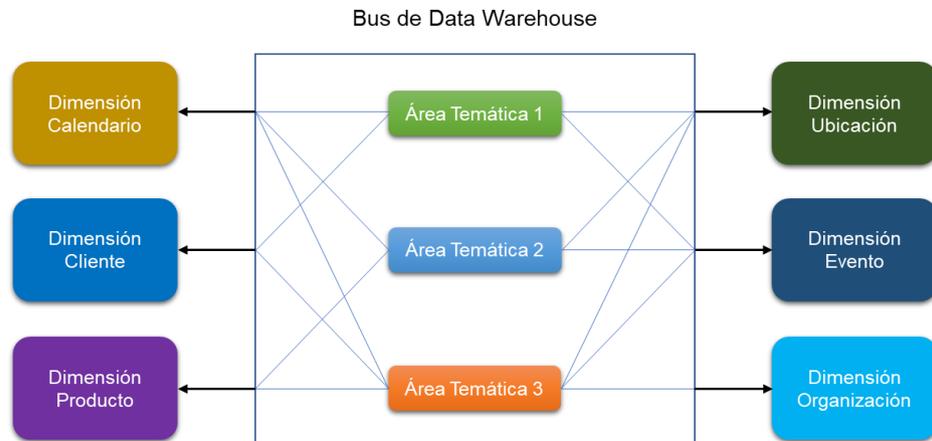
El requisito arquitectónico clave para el área de almacenamiento de datos está fuera de los límites de los usuarios de la empresa. al igual que el área de preparación de datos y la zona de almacenamiento de datos del DW. No deben existir distracciones al momento de proporcionar disponibilidad de datos, índices, agregaciones, series temporales, integración sincrónica entre áreas temáticas y especialmente en la seguridad a nivel de usuario.

Conexión de los Data Marts a la arquitectura del Bus de DW

La Figura 16 muestra cómo conectar varios Data Marts de procesos de negocio (etiquetados como “áreas temáticas”) junto con dimensiones conformadas, que suelen incluir elementos, como el calendario (tiempo), cliente, producto, ubicación, evento y la organización. Se deben considerar los hechos conformados, que implican cualquier medida que exista en más de un Data Mart. Para conformar varias instancias de ingresos, se debe tomar en cuenta que las definiciones técnicas de cada una de ellas sean las mismas, de modo que se puedan comparar y sumar los ingresos por separado. Si no se pueden conformar dos versiones de ingresos, deben etiquetarse de forma diferente para que los usuarios no comparen ni añadan variantes.

Figura 16

Arquitectura de Bus de Data Warehouse.



Fuente: (Kimball & Ross, 2016)

La arquitectura del bus del DW combina los dos conceptos clave mencionados anteriormente: los Data Marts y los modelos dimensionales. Se separa la arquitectura general del Data Warehouse de forma limpia para que el área de preparación de ETL y el área de presentación realicen funciones claramente diferentes. La arquitectura de bus, que define el área de presentación de datos, se basa en la similitud predecible de los modelos dimensionales para que la empresa pueda conectar todos los Data Marts. Este enfoque es un marco para abordar todos los nuevos requerimientos del DW. (Kimball & Ross, 2016)

Teniendo en cuenta la arquitectura de los DW, las organizaciones pueden enfrentar los nuevos retos para optimizar la toma de decisiones por lo que deben tomar en cuenta los siguientes criterios para su implementación:

- Desarrollo descentralizado e incremental: la disciplina de definir las dimensiones y los hechos conformados antes de iniciar los proyectos de Data Marts separados es la clave principal en este punto.
- Anticipación al cambio continuo a medida que evolucionan las necesidades empresariales y las fuentes de datos disponibles: la similitud de todos los esquemas dimensionales permite anticipar los efectos de los cambios inesperados. Cuando los usuarios, por ejemplo, comienzan a hacer preguntas centradas en la geografía y no en la categoría del producto, no cambia mucho el Data Mart dimensional. La ubicación y el producto son simplemente dos dimensiones conectadas simétrica y equivalentemente a la mayoría de las mismas tablas de hechos. Se pueden añadir nuevos elementos de datos y nuevos atributos dimensionales, nuevos hechos adicionales y dimensiones completamente nuevas.
- Despliegue rápido: una vez que se han establecido las dimensiones y los hechos conformados, los equipos de Data Mart separados pueden proceder de forma independiente entre sí. En muchos casos, tiene sentido construir las primeras tablas en cada Data Mart como imágenes dimensionales de fuentes subyacentes únicas.
- Desglose sin problemas hasta los datos más bajos posibles: desglosar no es más que agregar un encabezado de fila a un informe existente. En una arquitectura de bus de dimensiones conformada, se sabe que estos encabezados de fila están disponibles en las dimensiones y tendrán un significado consistente a medida que se desciende de tablas de hechos más globales a las que lo son menos.

- Las partes (Data Marts) se suman al todo (Data Warehouse): Esto es una consecuencia directa de la definición de los Data Marts como subconjuntos lógicos completos del DW. Se añade más significado a esta definición mostrando la estructura de cada Data Mart (dimensional) y cómo conectarlos entre sí (la arquitectura del bus).
- Data Marts implementados en tecnologías diversas e incompatibles: se puede dejar que los Data Marts sean incompatibles en los niveles más bajos de hardware y software porque no insistimos en hacer que estos últimos se comuniquen directamente. Al realizar consultas separadas a cada Data Mart utilizando el llamado SQL Multipass simplemente se combinan los conjuntos de respuestas en una capa de aplicaciones de nivel superior. Este enfoque tiene el beneficio significativo de que las consultas separadas evitan una serie de problemas lógicos complejos asociados con el intento de unir tablas de hechos con diferentes números de elementos.
- Disponibilidad 24 x 7: Al pensar claramente en la separación del Back Room del Front Room, se tiene que el requisito de disponibilidad se refiere al área de presentación de datos. El primer paso para lograr esta disponibilidad en su totalidad es implementar el área de almacenamiento de datos en una máquina separada o en un proceso separado del área de presentación de datos. La salida final del área de preparación es un conjunto de archivos de carga para el área de presentación. Sin embargo, cargar e indexar archivos en la base de datos de presentación final puede ser un proceso largo que puede dejarla fuera de línea. La carga de la base de datos de cada día se convierte en un archivo “temporal” que

comienza como una copia completa de la base de datos de presentación normal.

- Publicar los resultados del DW en todas partes, preferiblemente a través de Internet: los proveedores de herramientas de BI se han enfocado en proporcionar interfaces de usuario habilitadas para la web. En cierto sentido, la mayoría de los propietarios de Data Warehouse se encontrarán presentando sus resultados a través de intranets o incluso en Internet, ya sea que lo planeen o no.
- Asegurar los resultados del almacén de datos en todas partes, especialmente a través de Internet: la desventaja de utilizar Internet es la exposición a problemas de seguridad. El propietario del Data Warehouse es especialmente vulnerable debido a la sensibilidad de gran parte de los datos subyacentes e irónicamente, al éxito del almacén en la publicación de datos de manera efectiva para todos los usuarios comerciales.
- Respuesta casi instantánea a todas las solicitudes: mejorar los tiempos de respuesta para las consultas de los usuarios implica ser disciplinados para usar estructuras de bases de datos simples y predecibles, aumentar el uso de índices de bases de datos con agregaciones de bases de datos y usar SQL Multipass en lugar de SQL complejo. Todos estos enfoques utilizan en gran medida el enfoque del modelo dimensional, y existe un creciente número de experiencia y tecnología en estas áreas basado en supuestos dimensionales.
- Facilidad de uso, especialmente para los no entusiastas de la informática: Este último requisito completa el círculo. La motivación del diseño original para las tablas de hechos y tablas de dimensiones en el enfoque dimensional ante todo deben

representar los datos de una manera comprensible para el beneficio de los usuarios finales mediante la unión de fuentes de datos separadas utilizando el enfoque de “dimensiones ajustadas”. (Kimball & Ross, 2016)

Se debe tomar muy en cuenta que estos criterios para la implementación de un DW traen consigo ventajas y desventajas. Entre las desventajas se tiene que son exigentes y no son tan fáciles de ajustar a las técnicas de diseño y enfoque de gestión que se usan tradicionalmente. Por otro lado, una de las principales ventajas es que con los Data Marts enfocados a un diseño dimensional puede lograr cubrir los requerimientos de los usuarios finales.

Puntos clave para el diseño de soluciones DW/BI

Las soluciones más eficientes de BI implican un adecuado diseño de DW que ofrezcan como resultado un conjunto de buenas prácticas desde el punto de vista empresarial para el manejo y distribución de la información generada por los datos que son conformados en este tipo de soluciones.

Se deben tener en cuenta algunos puntos importantes que servirán de base para obtener los mejores resultados posibles al momento de la implementación, a saber, se debe tener presente lo siguiente:

- Adoptar un enfoque empresarial.
- Practicar la Inteligencia de Negocios.
- Diseñar esquemas dimensionales.
- Utilizar dimensiones adaptadas para la integración.
- Planificar cuidadosamente la arquitectura ETL.

Este conjunto de buenas prácticas son las recomendadas para los diseñadores y desarrolladores de soluciones de Inteligencia de Negocios,

de acuerdo con la metodología escogida. Se pretende crear un pensamiento crítico al momento de elección de este enfoque. (Kimball & Ross, 2016)

Pautas para el diseño de un DW de bajo riesgo

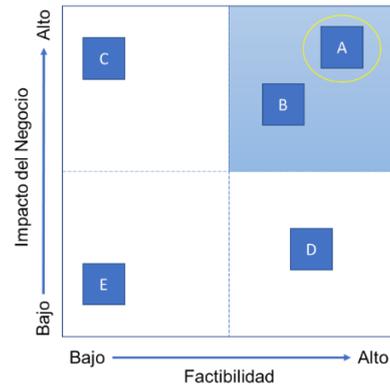
En el clima económico actual, los proyectos de DW/BI enfrentan dos presiones poderosas y conflictivas. Por un lado, los usuarios empresariales quieren una visión más centrada de sus herramientas de BI en la satisfacción y rentabilidad del cliente. Por otro lado, estos mismos usuarios están bajo una enorme presión para controlar los costos y reducir los riesgos. La explosión de nuevas fuentes de datos y nuevos modos de entrega disponibles para BI realmente agudiza este dilema.

Para evitar este conflicto presentado, es recomendable seguir las siguientes pautas para minimizar los posibles problemas que se presenten:

- Trabajar en lo correcto: se procede a realizar una lista de todos los proyectos potenciales de DW/BI y colocarlos en una cuadrícula simple de 2 x 2 (Figura 18).
- Dar control a los usuarios: la transferencia de control significa tener usuarios directamente involucrados y responsables de cada proyecto de DW/BI. Obviamente, estos usuarios tienen que aprender a trabajar con TI para poder hacer demandas razonables. La cuadrícula de impacto – factibilidad que se muestra en la Figura 17 es un buen punto de partida.

Figura 17

Priorización de impacto versus factibilidad.



Fuente: (Kimball & Ross, 2016)

- Proceder de forma incremental: muchos proyectos de DW/BI están gravitando hacia lo que podría llamarse un enfoque “ágil” que enfatiza los lanzamientos frecuentes y las correcciones a mitad de camino. Este tipo de enfoque se puede adaptar con éxito a proyectos de toda la empresa, como la gestión de datos maestros y la integración empresarial.
- Comenzar con una gobernanza enfocada y liviana: la gobernanza consiste en reconocer el valor de sus activos de datos y gestionarlos de forma responsable, es parte de una cultura más amplia que reconoce el valor de sus activos de datos y está respaldada e impulsada por altos ejecutivos. A nivel de un proyecto individual, la gobernanza implica identificar, catalogar, valorar, asignar responsabilidades, asegurar, proteger, cumplir, controlar, mejorar, establecer prácticas consistentes, integrarse en todas las áreas temáticas, planificar el crecimiento, planificar para cosechar valor.
- Construir una plataforma universal simple: es cierto en el espacio de BI que: no se puede predecir la naturaleza de las herramientas de BI de cara al usuario. En el futuro, ¿qué será más importante?:

análisis predictivo de minería de datos, entrega a dispositivos móviles, informes por lotes, alertas en tiempo real o algo en lo que aún no se ha pensado. Afortunadamente, se tiene una buena respuesta a esta pregunta; se debe reconocer que el DW es la plataforma única para todas las formas de BI. Este punto de vista permite notar que la interfaz del DW para todas las formas de BI debe ser simple y universal. El modelado dimensional cumple estos objetivos como interfaz para todas las formas de BI.

- Integrar usando dimensiones conformadas: la integración en toda la empresa se ha elevado a la cima de la lista de controladores técnicos de DW/BI junto con la calidad y latencia de los datos. El modelado dimensional proporciona un conjunto simple de procedimientos para lograr la integración que las herramientas de BI pueden utilizar de manera eficaz.
- Gestionar la calidad de pocas pantallas a la vez: las pantallas de calidad de datos se pueden implementar una a la vez, lo que permite que el desarrollo del sistema de calidad de datos crezca de manera incremental.
- Utilizar claves sustitutas en todo momento: una recomendación aparentemente pequeña para reducir el riesgo de desarrollo de DW/BI es asegurarse de construir todas las dimensiones (incluso las dimensiones de tipo 1) con claves primarias sustitutas. Esto lo aísla de las sorpresas posteriores cuando adquiere una nueva división que tiene sus propias ideas sobre las claves. Asimismo, todas las bases de datos se ejecutarán más rápido con este tipo de claves. (Kimball & Ross, 2016)

Selección de la metodología de trabajo

Para obtener una solución de BI que sirva como herramienta tecnológica para el alcance de los objetivos de la empresa, es necesario evaluar las metodologías más conocidas y seleccionar la que más se ajuste a los propósitos que se persiguen; en este caso y de acuerdo con la literatura revisada, la evaluación y selección se realizará entre la metodología de Ralph Kimball y la metodología propuesta por Bill Inmon, anteriormente descritas.

Para la evaluación de la metodología más adecuada se definen algunos criterios e indicadores que se toman como parámetros para su comparación, estos se muestran en la siguiente tabla:

Tabla 4

Criterios e indicadores para comparación de metodologías

Criterio	Indicadores	Inmon	Kimball
	Enfoque general	De arriba a abajo	De abajo a arriba
			Los Data marts
Metodología y arquitectura	Estructura arquitectónica	Almacén de datos de toda la empresa (atómico), alimenta las bases de datos departamentales.	modelan un único proceso de negocio; la consistencia de la empresa se consigue a través de un bus de datos y

		dimensiones conformes
	Complejidad del método	Complejo Sencillo
	Comparación con metodologías de desarrollo establecidas	Derivado de la metodología en espiral Proceso de cuatro pasos; se aleja de los métodos RDBMS
	Discusión del diseño físico	Completo Ligero
	Orientación de los datos	Basados en la materia o en los datos Orientado a los procesos
Modelado de datos	Herramientas	Modelado dimensional; se aleja del modelado relacional
	Accesibilidad para el usuario final	Baja Alta
	Usuario principal	Profesionales de la informática Usuarios finales
Filosofía	Lugar en la organización	Parte integrante de la Fábrica de Transformador y conservador de datos operativos

	Información Corporativa	
Objetivo	Ofrecer una solución técnica sólida basada en métodos y tecnologías de bases de datos de probada eficacia	Ofrecer una solución que facilite a los usuarios finales la consulta directa de los datos y obtener tiempos de respuesta razonables

Fuente: (Breslin, 2004)

Para seleccionar la metodología más idónea para el presente caso de estudio, se procedió a crear una escala de Likert con la finalidad de asociar una valoración a cada uno de los indicadores de las dos metodologías consideradas como aplicables, por lo cual se creó una escala que determina el grado de cada indicador: bajo, medio, alto, a cada una de estas características se le asigna un valor numérico que determina el peso, en el que se tiene: uno (1) para bajo, dos (2) para medio y tres (3) para alto.

A continuación, se presenta una comparación en la Tabla 4 definiendo la escala y valoración para los indicadores a ser utilizados para la medición de cada una de las metodologías.

Tabla 5

Escala de Likert

ESCALA	VALORACIÓN
Bajo	1
Medio	2
Alto	3

Del mismo modo, se muestra en la Tabla 5 la ponderación de los indicadores a ser tomados en cuenta para la selección de la metodología más adecuada:

Tabla 6*Ponderación de indicadores de las metodologías revisadas*

Indicador	Metodología	
	Kimball	Inmon
Enfoque general	3	3
Estructura arquitectónica	3	2
Complejidad del método	3	2
Comparación con metodologías de desarrollo establecidas	3	2
Discusión del diseño físico	2	1
Orientación de los datos	3	2
Herramientas	2	3
Accesibilidad para el usuario final	3	1
Usuario principal	3	1

Lugar en la organización	2	2
Objetivo	3	2
Resultados	30	21

De acuerdo con los resultados obtenidos luego de la aplicación del método de escogencia desarrollado, se selecciona para el presente trabajo de estudio la metodología de Ralph Kimball para la implementación de la solución de BI.

Resulta beneficioso para el proyecto seguir dicha metodología, debido a los criterios e indicadores ya evaluados, asimismo su filosofía de trabajo es conveniente, ya que se basa en el desarrollo incremental en ciclos rápidos: en cada iteración generar entregables que se irán refinando en cada ciclo.

Dicha metodología será ajustada a las necesidades y características propias del proyecto.

Categorización de la variable dependiente

Agencia de Courier

En relación con la agencia de Courier, (Soriano, 2000, citando en Cárdenas, 2012) la define como "...una empresa destinada a la importación y exportación de mercaderías tales como paquetes y documentos alrededor del mundo".

En lo que respecta al desarrollo del servicio de Courier en Ecuador, con el incremento de remesas del exterior desde el 2002 al 2005 las empresas Courier despiertan el correo internacional y nacional, especialmente al primero donde se observa un interesante nicho de mercado que promete resultados halagadores en la economía nacional y

particular de aquellos que toman el riesgo de hacer negocio en esta área. (Cárdenas, 2012)

Hoy en día, las entregas se realizan sin parar en todo el mundo y son las más rápidas y eficientes que se han hecho nunca. Los sistemas automatizados de reserva, pedido y procesamiento permiten organizar más envíos, mientras que los modernos aviones y vehículos de logística permiten que los pedidos más grandes puedan llegar a todo el mundo gracias a las redes de transporte disponibles. (TNT-Fedex, 2015)

Atención al cliente

Es el servicio que prestan y proporcionan las empresas a sus clientes para comunicarse directamente con ellos.

En caso de que estos necesiten manifestar reclamos, sugerencias, los clientes de una empresa deberán comunicarse de manera expedita con la misma. (Ucha, 2009)

Proceso de atención al cliente

En referencia con los pasos o procesos de atención al cliente, se puede nombrar lo siguiente:

- **Contacto:** El usuario acudirá con el fin de obtener una respuesta rápida, resolver una duda o problema, entre otras.
- **Obtención de la información:** la empresa debe esforzarse en comprender correctamente sus demandas.
- **Resolución:** bien se trate de dar respuesta a un problema técnico o aclarar una duda acerca del producto o servicio, esta es la fase decisiva del proceso de atención al cliente.

- Finalización: en esta etapa final, es necesario confirmar que la demanda del cliente ha sido satisfecha y despedirlo conforme al protocolo establecido. Estas encuestas deben ser breves y claras para que puedan ser un instrumento realmente útil para la empresa. (Navarro, 2020)

Eficiencia en la atención al cliente

Entre las distintas herramientas de medición de la satisfacción del cliente se encuentran:

- La encuesta: su propósito es obtener información de los clientes, por medio de preguntas concretas sobre temas como las expectativas previas del producto que va a adquirir y contrastarlas con su percepción después de la compra. Estas encuestas pueden realizarse de forma directa, por correo electrónico o por vía telefónica, aunque en cualquiera de las formas, existe la dificultad que se genera respecto al tema que se desea abordar para obtener la tabulación de los datos.
- Buzón de sugerencias: es el método más sencillo y de menor costo, pues simplemente se requiere de formatos donde los usuarios o clientes puedan dejar por escrito sus dudas, quejas o reclamos, en buzones ubicados de forma visible en puntos de alta confluencia de personas.
- Para llevar a cabo esto, las empresas localizan a estos por medio de sus bases de datos, en donde se pueda determinar cuáles de ellos no han estado activos en un largo tiempo. (Zonalogística, 2019)

Capítulo III

Análisis de estado actual

Servientrega Ecuador S.A. es reconocida como líder en el servicio de Courier nacional, ha logrado el posicionamiento de la marca gracias al compromiso con Ecuador, generando empleo, con personal capacitado para convertirse en un soporte logístico, tecnológico y administrativo, buscando siempre la satisfacción de sus clientes. (Servientrega Ecuador S.A., 2015)

La agencia de Courier seleccionada para la implementación del Sistema de Inteligencia de Negocios es una oficina franquiciada de Servientrega Ecuador S.A., se encuentra ubicada en la capital Quito.

La agencia comparte la visión, misión, objetivos, y toda la filosofía de atención al cliente y prestación de servicios de la franquicia ya mencionada.

Ha incorporado, recientemente, la prestación de otros servicios adicionales que buscan aumentar los ingresos del negocio, estos son: la venta de papelería y recarga de servicios (agua, electricidad, telefonía móvil celular y tv por cable).

Su estructura organizacional está compuesta por tres unidades: gerencia general (propietario), administración (una persona) y operaciones (una persona).

Esta agencia funciona de forma independiente de toda la estructura principal de la empresa de Courier por ser una franquicia de la misma; solo se adquieren los derechos de utilización de logos, prestación de los servicios de Courier y acceso al sistema central para el registro de las transacciones de

envíos y despachos. La Figura 18 representa la estructura organizacional de la agencia.

Figura 18

Organigrama de la estructura organizacional.



Gerencia general: contratación de personal, gestión de compras, control y seguimiento de la administración y operaciones del negocio; búsqueda de nuevos negocios que incrementen los ingresos de la empresa.

Administración: contabilidad, facturación, gestión de pagos y cuentas por cobrar.

Operaciones: recepción, despacho y registro de paquetes y documentos; atención a la venta de la papelería y recarga de los servicios disponibles.

Los servicios de Courier: comprenden varias etapas, que van desde las agencias que reciben los paquetes y documentos a ser enviados, seguido de la logística de transporte, que en principio se hace por vía terrestre y dependiendo de la distancia, por vía aérea. Finalmente, es destinada a otra oficina desde donde se hará el despacho vía terrestre o el retiro por el destinatario que recibe el envío. Los servicios de envío prestados son los siguientes:

- Paquetería liviana (carga): Cajas o paquetes desde 2 hasta 30 kilogramos por unidad de empaque.
- Documentos: Los documentos deben ser de hasta dos kilos máximo de peso por unidad de empaque. (Servientrega Ecuador S.A., 2015)

Las transacciones relacionadas con los procesos de Courier son registradas en el sistema corporativo de Servientrega Ecuador S.A.; al que se tiene un acceso restringido al perfil de franquiciado: registro de paquetes y documentos, modificación sobre el registro ingresado, eliminación de registro ingresado; emisión de etiquetas, guías y notas de entrega; declaración de valor, estatus de los envíos, emisión de reportes de operaciones (diarias, semanales y mensuales).

Los reportes de operaciones de la agencia son emitidos en formato .pdf, estos se descargan localmente y después, son convertidos a hojas de cálculo Excel; las que posteriormente, son utilizadas por la gerencia y la administración, con fines de análisis y control. Se resalta que solo se accede a la información de esta agencia.

Dichos reportes son prefijados por el sistema corporativo, por consiguiente, rígidos; no presentan la posibilidad de visualizar la información cuándo y cómo se desee, lo que no permite hacer análisis precisos que ayuden a conocer y prever la salud financiera del negocio.

Las operaciones referidas a las ventas de los servicios adicionales de la agencia son registradas de forma manual en una hoja de cálculo sin mucha rigurosidad. No se cuenta con alguna aplicación o herramienta tecnológica adecuada que apoye la gestión empresarial de la agencia: contabilidad, ventas, compras, pagos, inventario, entre otros procesos.

Dichas hojas son clasificadas de acuerdo al proceso administrativo:

- Ventas diarias que después se consolidan por mes.

- Inventario, compra y venta del material de papelería.
- Pago a proveedores y nómina.
- Cuentas contables.

La revisión y análisis de los datos de los archivos fuentes arrojó problemas en la calidad de los mismos, tales como: productos sin información de envío, campos nulos, campos con ceros, productos enviados sin peso registrado, inexactitud en los registros e inconsistencia; todo esto representa un panorama poco confiable en cuanto a información se refiere.

Toda esta situación genera retrasos al momento de obtener información, se invierte mucho tiempo en su procesamiento y el resultado de dicho procesamiento no es preciso y está lleno de inconsistencias; lo que conduce a no tener una visión clara de la situación financiera de la agencia.

En resumen, los problemas encontrados pueden ser agrupados en:

- Tecnológicos: ausencia de soluciones y herramientas tecnológicas de apoyo para la gestión de negocios.
- Marketing: carencia de estudios de comportamiento de mercado y clientes.
- Administrativos: se ejecutan análisis tardíos y sin exactitud, la información que se genera de las operaciones no es precisa, gran parte de la información se ingresa de forma manual.
- Recursos humanos: no se tienen estrategias para la asignación de personal en momentos de mayor tráfico de clientes, tampoco para la capacitación de dicho personal; lo que los limita en la comprensión y gestión de los procesos de la agencia.
- Estratégicos: la agencia de Courier no tiene definidas su misión y visión de negocio, así como sus objetivos estratégicos, situación que no permite fijar el rumbo correcto de la empresa.

El comportamiento del estado actual de los datos evidencia que la toma de decisiones en la agencia de Courier no está basada en hechos, sino en intuición, es decir, no se sustenta en indicadores de negocio que puedan guiar a la empresa a establecer estrategias que favorezcan la fijación de metas de crecimiento.

Capítulo IV

Marco metodológico

Implementación del sistema de Business Intelligence

Para la implementación de la solución de BI se ha seguido el esquema de trabajo mostrado en la Figura 19, que es el resultado de la adaptación de la Metodología de Kimball ya explicada con anterioridad.

Figura 19

Esquema de trabajo para la solución del problema.



En relación con el Ciclo de vida metodología de Ralph Kimball es importante detallar que:

- En el ítem de Diseño de la Arquitectura (de la Figura 19) se desarrolla:
 - Diseño de la Arquitectura técnica.
 - Selección e instalación de productos.
- En el ítem Diseño y desarrollo de aplicaciones (de la Figura 19) se desarrolla:
 - Diseño de aplicaciones BI
 - Desarrollo de aplicaciones BI

Planificación del proyecto

Constituye la etapa inicial del proyecto, donde se determinó el alcance del proyecto y la factibilidad del mismo; se elaboró el plan de trabajo, identificando las fases de desarrollo, los productos y los tiempos de ejecución, conjuntamente con las funciones y los recursos necesarios.

Alcance del proyecto

Este proyecto propone la implementación de una solución de Inteligencia de Negocios para la toma de decisiones en una agencia de Courier, dicha implementación consiste en la primera iteración del sistema y abarca:

- La construcción del almacén de datos (data mart)
- Creación de análisis multidimensional
- La elaboración de las aplicaciones de consulta de información (reportes y consultas ad hoc)

La solución está circunscrita a un desarrollo local (stand alone) cuyo crecimiento está previsto con la ejecución de un futuro proyecto, supeditado a una inversión de capital mayor.

Este desarrollo se ha llevado a cabo utilizando los recursos tecnológicos disponibles actualmente en la agencia, recordando que se trata de una pequeña oficina franquiciada de Servientrega Ecuador S.A. que desea mejorar su toma de decisiones.

Factibilidad del proyecto

El proyecto es económica y técnicamente factible, porque la empresa cuenta con el hardware y el software necesarios para sustentar la solución de BI planteada. No requiere adquirir licencias adicionales, y la suite de

base y derivadas																		
Identificar la correspondencia de data desde la fuente al destino	•Tesisista.							X										
Diseño físico																		
Diseñar tablas físicas y columnas	•Tesisista.								X									
Elaborar la correspondencia origen – destino	•Tesisista.								X									
Estimar el tamaño de la base de datos	•Tesisista.								X									
Diseñar la base de datos	•Tesisista.									X								
Diseñar la base de datos OLAP	•Tesisista.									X								

Instalar y configurar el RDBMS	•Tesista.									X							
Construir la estructura de almacenamiento físico	•Tesista.									X							
Crear tablas e índices	•Tesista.									X							
Crear base de datos OLAP	•Tesista.									X							
Aceptación del usuario	•Responsable del negocio										X						
Desarrollo de los ETL																	
Diseño del proceso ETL	•Tesista.										X						
Desarrollar estrategias predeterminadas para la extracción de datos	•Tesista.										X						

Desarrollar estrategias predeterminadas para archivar datos extraídos	•Tesisista.										X						
Desarrollar estrategias predeterminadas para vigilar la calidad de los datos	•Tesisista.										X						
Desarrollar estrategias predeterminadas para la gestión de cambios de las dimensiones	•Tesisista.										X						
Diseñar planes de ETL para cada tabla	•Tesisista.										X						

Establecer la secuencia de Jobs	•Tesisista.										X						
Cargar la tabla de fechas	•Tesisista.											X					
Construir cargas históricas e incrementales para tablas de dimensiones	•Tesisista.											X					
Construir cargas históricas e incrementales de tablas de hechos	•Tesisista.											X					
Diseñar, construir y probar la automatización del sistema ETL	•Tesisista.											X					
Aceptación del usuario	•Responsable												X				

	del negocio																		
Diseño y desarrollo de las aplicaciones																			
Probar las herramientas de BI	•Tesisista.												X						
Configurar la seguridad del usuario	•Tesisista.												X						
Desarrollar aplicaciones de BI	•Tesisista.												X	X					
Validar los datos y el modelo de datos	•Tesisista.													X					
Configurar la planificación de ejecución de reportes	•Tesisista.														X				
Probar las aplicaciones de BI y	•Tesisista.														X				

	negocio . •Tesisista.																				
Aceptación del usuario	•Respon sable del negocio .																			X	
Analizar y corregir errores	•Tesisista.																			X	
Documentar	•Tesisista.																			X	X
Liberación	•Tesisista.																				X
Aceptación del usuario	•Respon sable del negocio .																				X

Definición de los requerimientos

La identificación y valoración de las necesidades de la agencia de Courier, se llevó a cabo a través de una entrevista, usando como instrumento de recolección de datos un cuestionario preparado para tal fin, como se muestra en el Anexo 1. Estos requerimientos establecen la base para diseñar y desarrollar el DW y las aplicaciones.

Productos generados: necesidades del negocio identificadas, fuentes de datos de origen identificadas y analizadas; identificación de reportes, como se muestra en el Anexo 2; identificación de consultas e indicadores, como se muestra en el Anexo 3.

En la tabla 8 se puede apreciar un resumen de los problemas detectados con el levantamiento de información, la tabla 9 presenta un listado de los requerimientos de información identificados, la tabla 10 recoge una lista de reportes necesarios y finalmente, la tabla 11 muestra el indicador identificado.

Tabla 8

Resumen de problemas detectados

Número	Descripción	Tipo	Prioridad
1	Registro manual de todas las operaciones	Administrativo	Alta
2	Ausencia de soluciones y herramientas tecnológicas de apoyo para la gestión de negocios	Tecnológico	Alta
4	Baja calidad de los datos registrados manualmente: inexactitud e inconsistencia en los registros, palabras truncadas, campos nulos o en blanco.	Administrativo	Alta
3	Análisis tardíos e imprecisos basados en datos de poca calidad	Estratégico	Alta
4	Ausencia de fijación de metas claras	Estratégico	Alta

5	Inexistencia de capacitación y desarrollo del talento humano	RRHH	Alta
6	Inexistencia de estudios de mercado	Estratégico	Alta
7	Ausencia de misión, visión y objetivos de la agencia	Estratégico	Alta

Tabla 9*Lista de requerimientos identificados*

Código	Requerimiento
REQ01	Visualizar el servicio más demandado por año
REQ02	Visualizar la cantidad de servicios gestionados por semana, mes y año
REQ03	Conocer qué día de la semana se vende más y menos
REQ04	Conocer cuál es el mes con mayor y menor cantidad de ventas
REQ05	Conocer el año con mayor y menor cantidad de ventas
REQ06	Visualizar el comportamiento de la rentabilidad en un período determinado
REQ07	Comparar las ventas de un período específico vs. períodos anteriores
REQ08	Generar consultas personalizadas
REQ09	Visualizar las ventas de un período específico

Tabla 10*Reportes identificados*

Reporte	Área	Definición
Servicio más demandado por año: envío de documentos, envío de paquetes.	Ventas	Visualizar los servicios de envío de documentos y paquetes realizados en un año.
Cantidad de servicios gestionados por semana, mes y año.	Ventas	Analizar y comparar los servicios más demandados en un año específico.
Día de la semana con mayor y menor número de servicios por año.	Ventas	Determinar los días de la semana con mayor y menor cantidad de servicios enviados en un año específico.
Mes con mayor y menor número de ventas.	Ventas	Determinar los meses con mayor y menor ingreso por ventas en un año específico.
Año con mayor y menor número de ventas acumuladas.	Ventas	Determinar el mayor y menor ingreso por ventas en un año específico.
Comparativo de las ventas del mes de un año vs. Las ventas del mismo mes de otro año.	Ventas	Comparar las ventas netas de un mes específico de un año respecto a las ventas netas del mismo mes en otro año.
Comportamiento de la rentabilidad en un período determinado.	Ventas	Visualizar la rentabilidad del negocio en un período específico.

Tabla 11*Indicador identificado*

Indicador	Área	Descripción
Índice de crecimiento de ventas	Finanzas	Medir la tasa a la que la empresa puede aumentar sus ingresos por ventas durante un periodo fijo y compararlos con alguna otra empresa de la competencia.

Diseño de la arquitectura técnica

Con el diseño de la arquitectura se estableció el marco de trabajo general, es decir, los componentes específicos de la solución (motor de la base de datos, la herramienta de ETL, herramientas de acceso y despliegue de datos, plataforma de hardware).

En mismo orden de ideas, se cuenta con el apoyo de las herramientas escogidas, en el presente caso Pentaho para todo el modelado de la solución y Microsoft Power BI para la visualización de los resultados obtenidos. Estas herramientas fueron escogidas principalmente por tener una accesibilidad a la licencia, entre otros criterios mencionados en el Capítulo 2: Licencia Libre en Pentaho y Licencia free (limitada) en Microsoft Power BI.

La arquitectura técnica está basada en los siguientes tipos de componentes:

- **Servicios:** son las funciones necesarias para cumplir las tareas requeridas del almacén de datos: servicios de extracción, transformación y carga de datos, servicios de control de tareas,

servicios de consulta (visualización, gestión de consultas, acceso, seguridad, entre otros).

- Repositorios de datos: lugar de aterrizaje de los datos.
- Otros componentes detallados más adelante.

Infraestructura existente

- Hardware:

Microprocesador: Intel Core i7-6700HQ 2.50GHz

Memoria RAM: 16 GB

Disco Duro: 1 TB

Pantalla: 15"

- Software:

Sistema operativo Windows 7 Enterprise

Office 365 con Power BI Desktop

Mozilla Firefox

Productos y herramientas para el desarrollo de la solución

- Motor de base de datos PostgreSQL.
- Pentaho versión comunitaria (código libre).

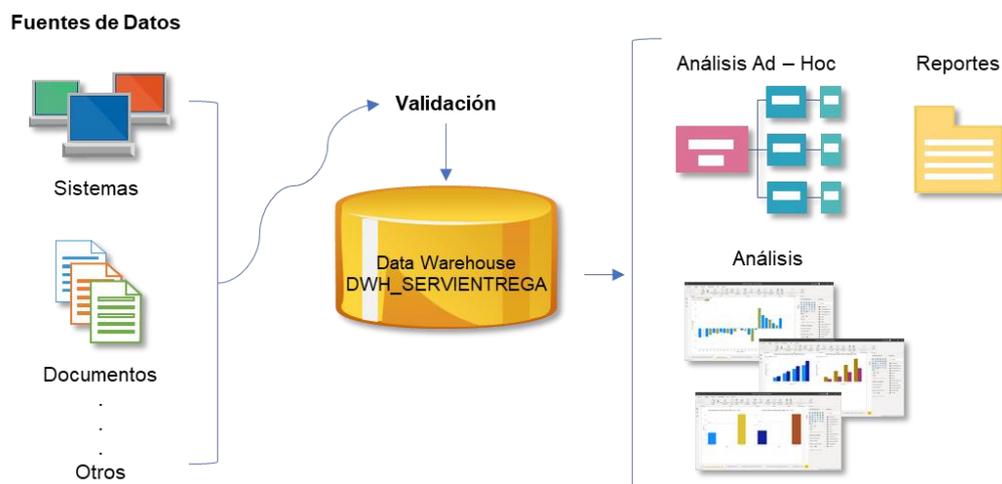
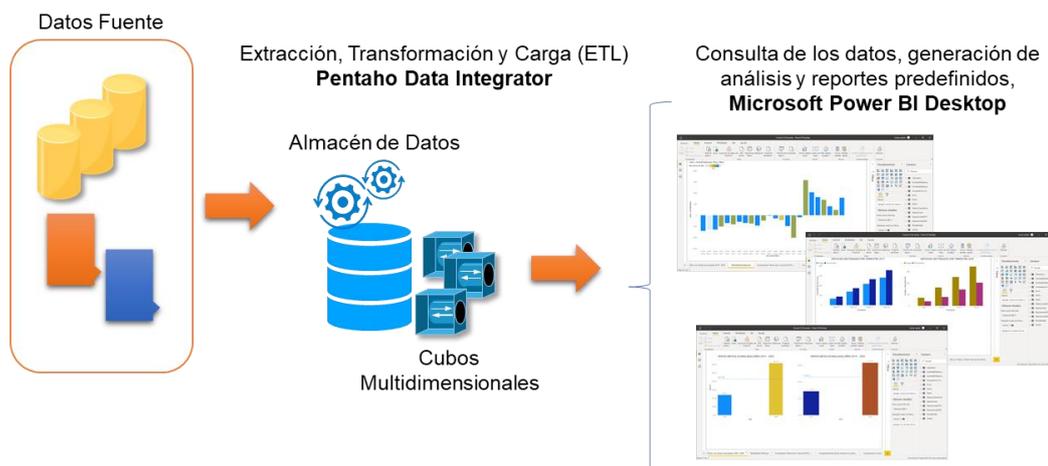
Pentaho Integración de Datos - PDI (Kettle): para los procesos ETL.

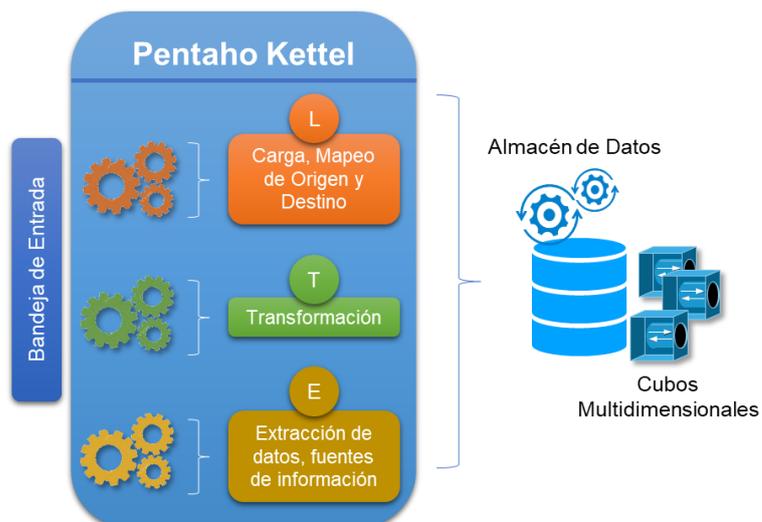
Gestión de procesos: para administrar y configurar las reglas, logs, y programar las tareas.

Consola de administración Pentaho: para crear el usuario administrador, gestionar usuarios, roles, y crear tareas programadas (Jobs) de PDI.

- Power BI Desktop: para la generación de reportes, consultas ad hoc y análisis OLAP.

Diagramas de alto nivel de la arquitectura técnica





Modelado dimensional

El proceso de modelado dimensional se diseñó en tres pasos básicos:

1. Elección del proceso de negocio.
2. Identificación de las dimensiones
3. Identificación de los hechos

Conforme a las necesidades del cliente, se determinó la creación de un Data mart como almacén de datos para la solución, por lo cual hubo que identificar previamente las dimensiones y hechos para proceder al modelado dimensional; modelo que posteriormente se convirtió en la base de datos física.

La primera tarea del modelado dimensional se concentró en la identificación de las dimensiones y hechos, los cuales se presentan a continuación:

Dimensiones identificadas

Las dimensiones representan los conceptos de negocio y es a través de sus atributos que se realizan las consultas requeridas. Las dimensiones identificadas y atributos de la dimensión fecha, constan en las Tabla 6 y Tabla 7, respectivamente. En la Figura 20 se muestra el modelo lógico dimensión fecha. Del mismo modo, la dimensión envío y gastos se visualizan en las Tabla 8, Figura 21, Tabla 9 y Figura 22, respectivamente.

Tabla 12

Dimensiones identificadas

Dimensión	Descripción
Fecha	Almacena el detalle mínimo de una fecha, lo que permite navegarla en plazos determinados. Es de uso común en el almacén de datos.
Envío	Contiene los servicios y productos de la agencia de Courier.
Gasto	Almacena los gastos operativos de la agencia de Courier.

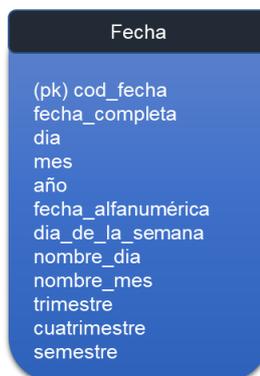
Tabla 13*Atributos de la dimensión fecha*

Atributo	Descripción
Cod fecha	Es la clave primaria de la dimensión (PK), es generada automáticamente y sirve para asociar los datos de la dimensión con la tabla de hechos.
Fecha completa	Fecha completa numérica.
día	Día del mes. Valores permitidos: del 1 al 31.
mes	Mes del año. Valores permitidos: del 1 al 12.
año	Año calendario.
Fecha alfanumérica	Fecha alfanumérica.
Día de la semana	Día de la semana. Valores del 1 al 7.

Nombre día	Nombre del día de la semana. Los valores permitidos son: lunes, martes, miércoles, jueves, viernes, sábado y domingo.
Nombre mes	Nombre del día de la semana. Los valores permitidos son: desde enero hasta diciembre.
trimestre	Trimestre correspondiente a la fecha.
cuatrimestre	Cuatrimestre correspondiente a la fecha.
semestre	Semestre correspondiente a la fecha.

Figura 20

Modelo lógico dimensión fecha.

**Tabla 14.**

Atributos de la dimensión envío.

Atributo	Descripción
Cod envío	Es la clave primaria de la dimensión (PK), es generada automáticamente y sirve para asociar los datos de la dimensión con la tabla de hechos.
producto	Se refiere a los productos que son objeto de envío.
Tipo de producto	

	Se refiere a los subproductos que son objeto de envío. Valores permitidos: documento y carga.
peso	Es el peso que tiene el subproducto enviado.

Figura 21

Modelo lógico dimensión envío.

**Tabla 15**

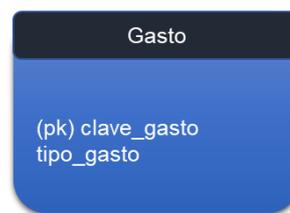
Atributos de la dimensión gasto.

Atributo	Descripción
Clave gasto	Es la clave primaria de la dimensión (PK), es generada automáticamente y sirve para asociar los datos de la dimensión con la tabla de hechos.
Tipo gasto	

Es la categorización del gasto.

Figura 22

Modelo lógico dimensión gasto.



Hechos identificados

Las tablas de Hechos contienen las métricas o medidas que refieren a los indicadores usados por el negocio para evaluar sus procesos. Cada métrica se relaciona con las filas correspondientes en cada una de las tablas de dimensiones a través de claves únicas (claves foráneas).

Se identificaron dos tablas de hechos que cubren las necesidades del negocio. Estas se muestran en la tabla 10, de igual manera se muestran sus atributos en las tablas 11 y 12 respectivamente.

Tabla 16*Tabla de Hechos identificados.*

Hecho	Descripción
ventas	Almacena las ventas diarias de los productos de la agencia de Courier.
Gastos operativos	Almacena los gastos operativos mensuales de la agencia de Courier.

Tabla 17*Atributos de la tabla de hechos ventas.*

Medida	Descripción
Cod fecha (fk)	Es la clave foránea de la dimensión Fecha, la misma permite relacionar cada medida con su fecha correspondiente.
Cod envío (fk)	Es la clave foránea de la dimensión Envío, la misma permite relacionar cada medida con sus atributos correspondientes.
Cantidad de despacho	

	Es la cantidad de productos despachados/enviados.
Monto venta bruta	Es la venta bruta asociada al envío de productos.
Monto venta neta	Es un porcentaje de la venta bruta y representa el margen que recibe la agencia de courier por cada venta realizada. Ese porcentaje siempre es de 30%, salvo en el año 2020 que fue de 25%.

Figura 23

Modelo lógico tabla de hechos ventas.

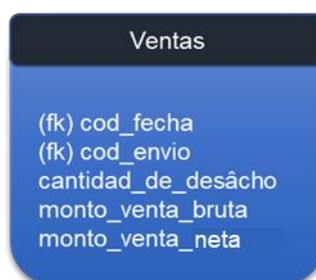


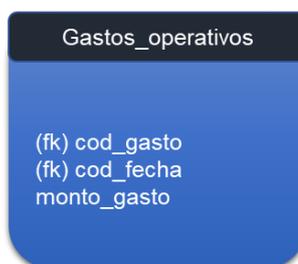
Tabla 18

Atributos de la tabla de hechos gastos_operativos.

Medida	Descripción
cod_fecha (fk)	Es la clave foránea de la dimensión Fecha, la misma permite relacionar cada medida con su fecha correspondiente.
cod_gasto (fk)	Es la clave foránea de la dimensión Gasto, la misma permite relacionar cada medida con sus atributos correspondientes.
monto_gasto	Se refiere al monto del gasto.

Figura 24

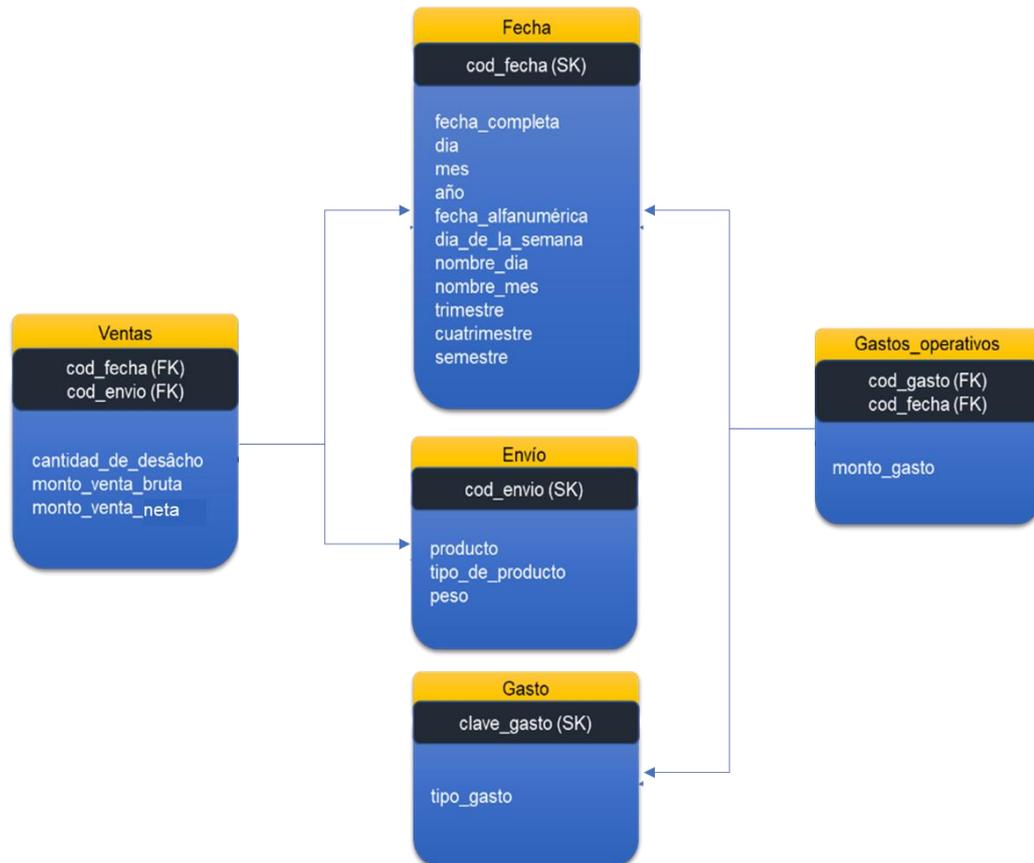
Modelo lógico tabla de hechos gastos_operativos.



Producto generado: modelo dimensional mostrado en Figura 25.

Figura 25

Modelo dimensional del almacén de datos de la agencia de courier



Diseño físico

El diseño físico se centró en: creación de la base de datos y construcción del Data mart.

Posterior al modelado dimensional se establecieron los tipos de datos de las tablas, como se puede apreciar en la tabla 13.

Como se comentó anteriormente, el gestor de base de datos seleccionado para trabajar es PostgreSQL y para la creación de las estructuras de datos se utilizaron las herramientas que él mismo proporciona para la tarea.

Tabla 19*Base de datos.*

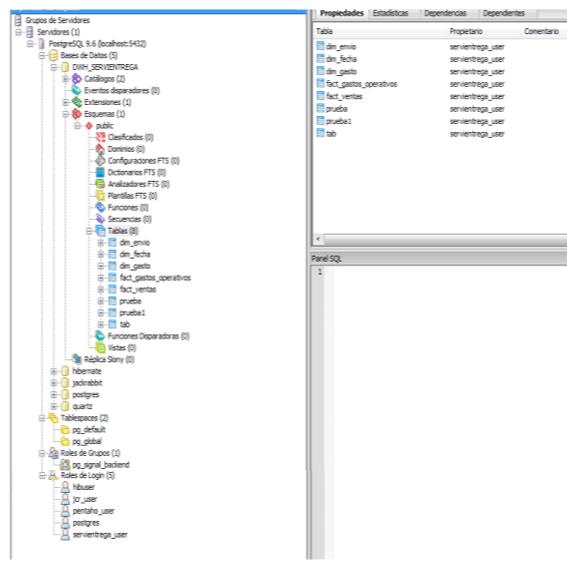
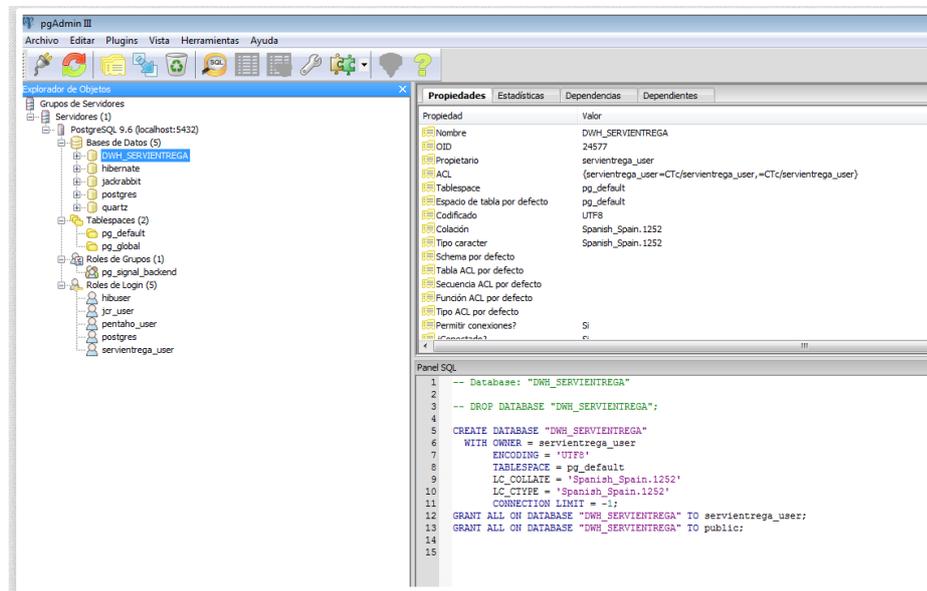
Nombre de tabla/columna	Tipo de dato	¿Permite nulos?	Clave primaria	Comentarios
Fecha.cod_fecha	Integer	No	Si	Autogenerada
Fecha.fecha_completa	Date	No	No	
Fecha.día	Integer	No	No	
Fecha.mes	Integer	No	No	
Fecha.año	Integer	No	No	
Fecha.fecha_alfanumerica	Varchar	No	No	
Fecha.día_de_la_semana	Varchar	No	No	
Fecha.nombre_dia	Varchar	No	No	
Fecha.nombre_mes	Varchar	No	No	
Fecha.trimestre	Integer	No	No	
Fecha.cuatrimestre	Integer	No	No	
Fecha.semestre	Integer	No	No	
Envio.cod_envio	Integer	No	Si	Autogenerada
Envio.producto	Varchar	No	No	

Envio.tipo_de_producto	Varcha r	No	No	
Envio.peso	Decim al	No	No	
Gasto.clave_gasto	Integer	No	Si	Autogenerad a
Gasto.tipo_gasto	Integer	No	No	
Ventas.cod_fecha	Integer	No	Si	
Ventas.cod_envio	Integer	No	Si	
Ventas.cantidad_de_despach o	Integer	No	No	
Ventas.monto_venta_bruta	Decim al	No	No	
Ventas.monto_venta_neta	Decim al	No	No	
Gastos_operativos.cod_fecha	Integer	No	Si	
Gastos_operativos.cod_gasto	Integer	No	Si	
Gastos_operativos.monto_ga sto	Decim al	No	No	

La Figura 26 muestra el proceso de creación de la base de datos de la solución: DWH_SERVIENTREGA.

Figura 26

Proceso de creación de la base de datos.

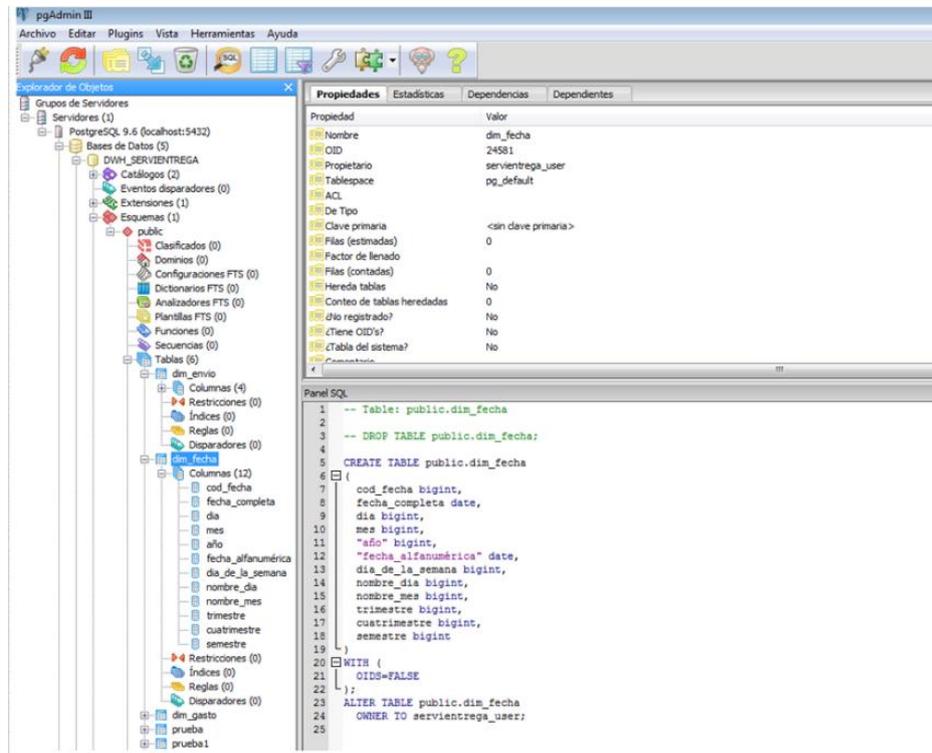


A continuación, a través de las figuras presentadas, se muestra el proceso de creación de las tablas de dimensiones y hechos.

La Figura 27 muestra la creación de la tabla de dimensión Fecha: `dim_fecha`.

Figura 27

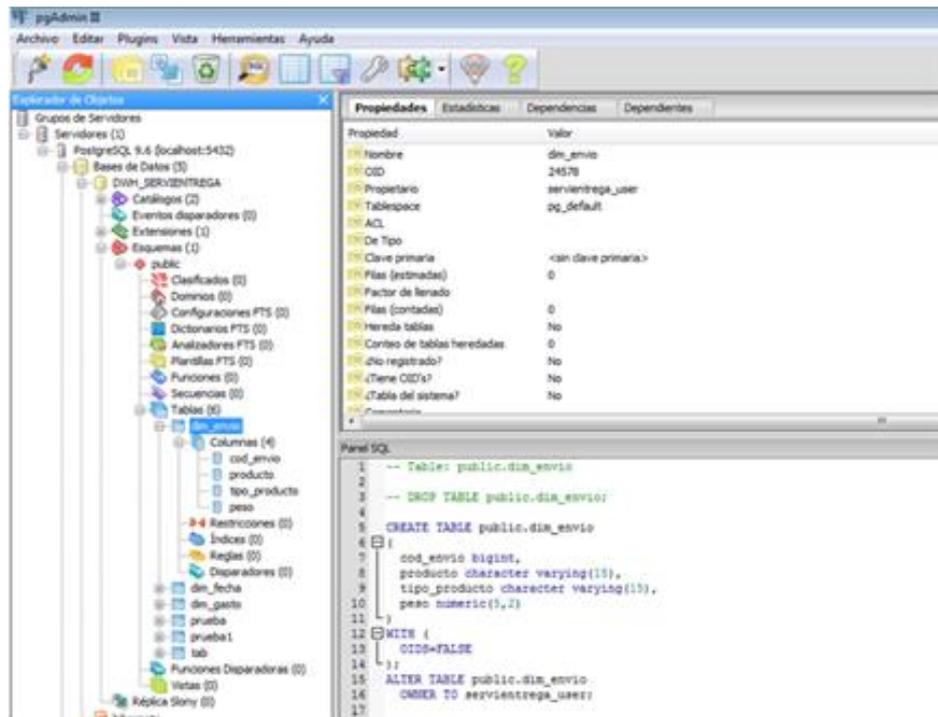
Proceso de creación de la dimensión Fecha: dim_fecha



En la Figura 28 se muestra el proceso de creación de la tabla de dimensión Envío: dim_envio.

Figura 28

Proceso de creación de la dimensión Envío: dim_envio.

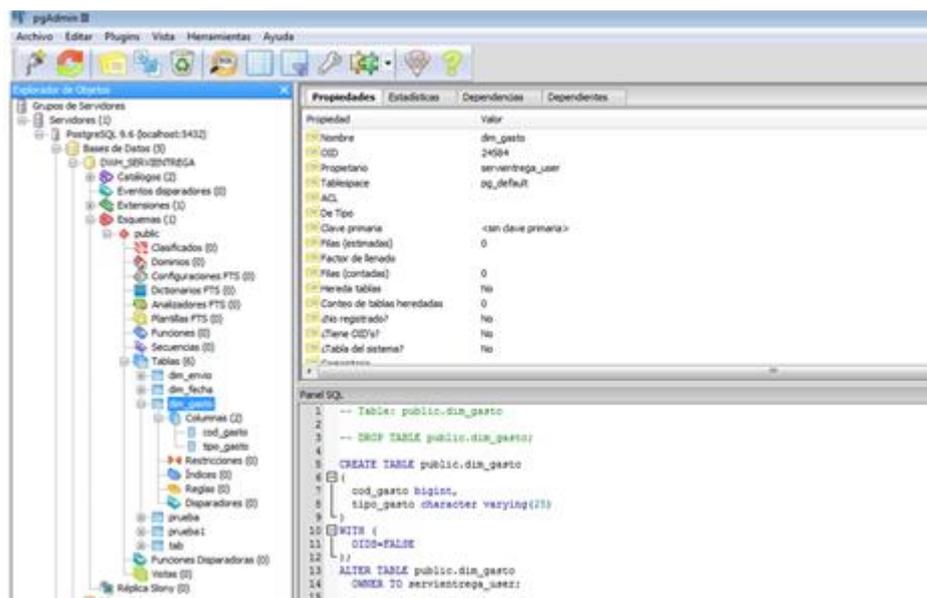


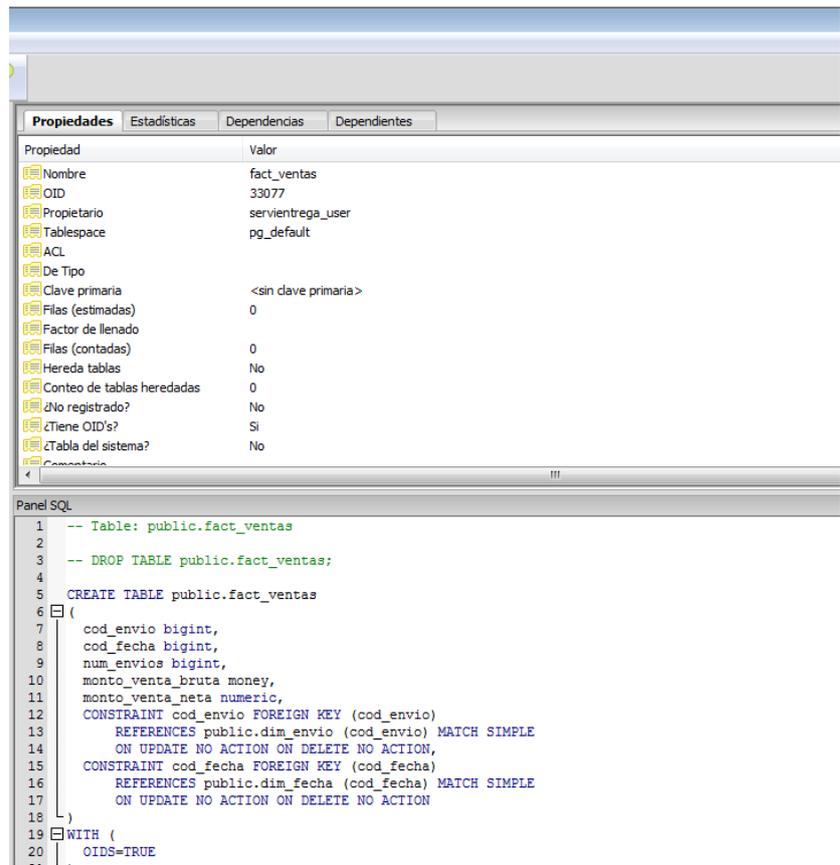
En la Figura 29 se puede apreciar la creación de la tabla de dimensión

Gasto: dim_gasto.

Figura 29

Proceso de creación de la dimensión Gasto: dim_gasto.

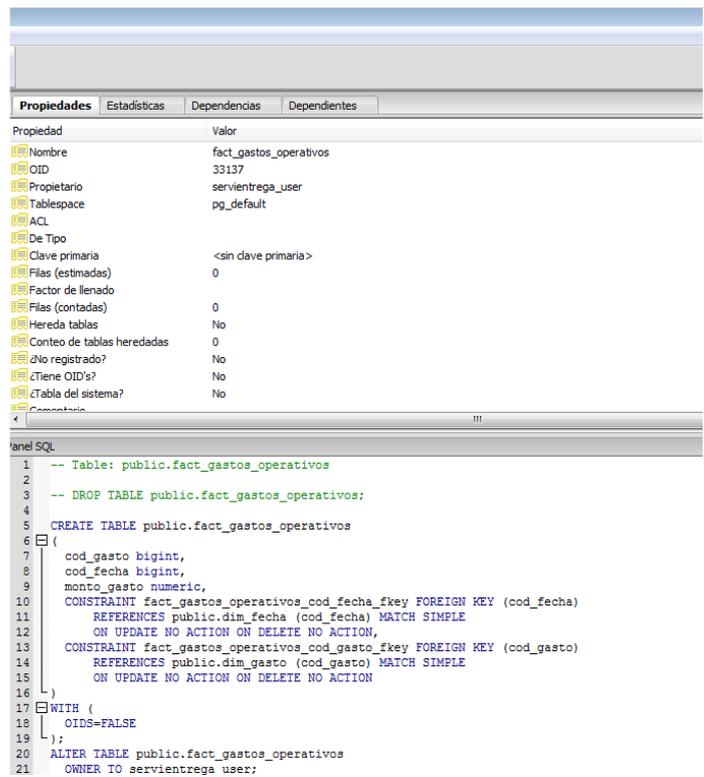




La Figura 31 presenta la creación de la tabla de hechos Gastos operativos: fact_gastos_operativos.

Figura 31

Proceso de creación de hecho Gastos operativos: fact_gastos_operativos.



El anexo 5 ofrece una vista del proceso de creación de la base de datos y el data mart y sus elementos: espacio, índices, tablas de dimensiones y hechos, roles, usuarios.

Diseño y desarrollo ETL

Esta fase se centró en tres subprocessos fundamentales: extracción (E), transformación (T) y carga (L); necesarios para poblar la base de datos ya creada (Data mart).

Los datos fueron extraídos de los archivos fuente, sometidos a un proceso de transformación (limpieza, formateo y reglas de negocio) para adecuarlos a las necesidades (técnicas y de negocios), y finalmente fueron cargados en el Data Mart DWH_SERVIENTREGA.

La herramienta utilizada para llevar a cabo los procesos de ETL es Pentaho Integración de Datos (PDI), la cual contiene una serie de componentes diseñados para ejecutar tareas específicas.

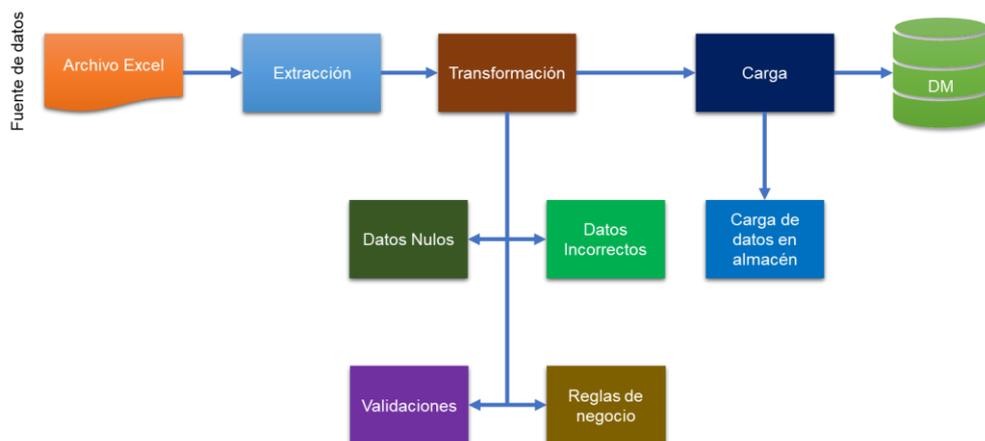
Para diseñar las transformaciones y trabajos se utilizó Spoon, la interfaz gráfica de usuario (GUI) del PDI. A través de esta se crearon las conexiones a los datos y las transformaciones directas e indirectas.

El producto generado en esta fase es el diseño y ejecución del proceso de ETL sobre la base de datos dimensional.

En la Figura 32 se muestra un diagrama general del proceso de ETL.

Figura 32

Diagrama general del proceso de ETL.



Los archivos fuentes son hojas de Excel generadas por el sistema transaccional corporativo, los mismos fueron descargados por el usuario responsable de acuerdo con las instrucciones emitidas.

Antes de dar inicio al desarrollo de los ETL se elaboró un mapeo de datos fuente a destino para cada tabla. Esta información es mostrada en la Tabla 20.

Tabla 20

Mapeo de datos fuente a destino.

Tabla/ Archivo fuente	Columna/campo fuente	Tabla destino	Columna destino	Tipo de dato	Longit ud	Notas de transformaci ón de datos
Dimensiones						
		Fecha	Cod_fecha	Integer		
Operacione s	Fecha_completa	Fecha	Fecha_compl eta	Date		La fecha completa se descompus o en los

diferentes atributos que representan el nivel de granularidad de una fecha

Fecha	Dia	Integer	2
-------	-----	---------	---

Fecha	Mes	Integer	2
-------	-----	---------	---

Fecha	Año	Integer	4
-------	-----	---------	---

Fecha	Fecha_alfanu mérica	Varcha r	20
-------	------------------------	-------------	----

Fecha	Dia_de_la_se mana	Varcha r	10
-------	----------------------	-------------	----

Fecha	Nombre_dia	Varcha r	10
-------	------------	-------------	----

Fecha	Nombre_mes	Varcha r	10
-------	------------	-------------	----

Fecha	Trimestre	Integer	1
-------	-----------	---------	---

Fecha	Cuatrimestre	Integer	1
-------	--------------	---------	---

Fecha	Semestre	Integer	1
-------	----------	---------	---

Envio	Cod_envio	Integer	
-------	-----------	---------	--

	Producto	Envio	Producto	Varcha r	15
	Tipo_de_produc to	Envio	Tipo_de_pro ducto	Varcha r	15
	Peso	Envio	Peso	Decima l	3,2
		Gasto	Cod_gasto	Integer	
Gastos_ operativos	S/n	Gasto	Tipo_gasto	Varcha r	25

Hechos

		Venta	Cod_fecha	Integer	
		Venta	Cod_envio	Integer	
	Num_envios	Venta	Cantidad_de _despacho	Integer	3
	Ventas	Venta	Monto_venta _bruta	Decima l	10,2
		Venta	Monto_venta _neta	Decima l	10,2
					Campo calculado, representa el 30% de las ventas brutas, salvo

					para el año 2020 que es el 25%
	Gastos _operat ivos	Cod_gasto	Integer		
	Gastos _operat ivos	Cod_fecha	Integer		
Gasto	Gastos _operat ivos	Gastos_oper ativos.monto _gasto	Decima I	10,2	

La Figura 33 muestra el proceso de ETL. Asimismo, se detalla la construcción de las tablas de dimensiones en el proceso ETL, procesos de limpieza inicial de las tablas de dimensiones, proceso de limpieza inicial de las tablas de hechos en la Figura 34, Figura 35, Figura 36 y Figura 37.

Figura 33

Proceso de ETL.

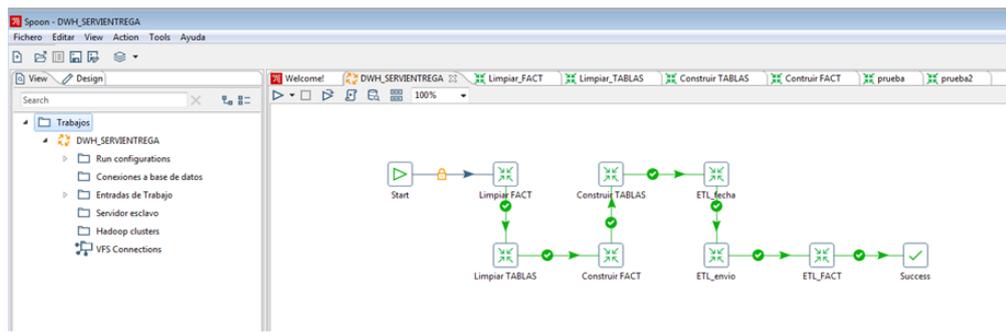


Figura 34

Construcción de las tablas de dimensiones.

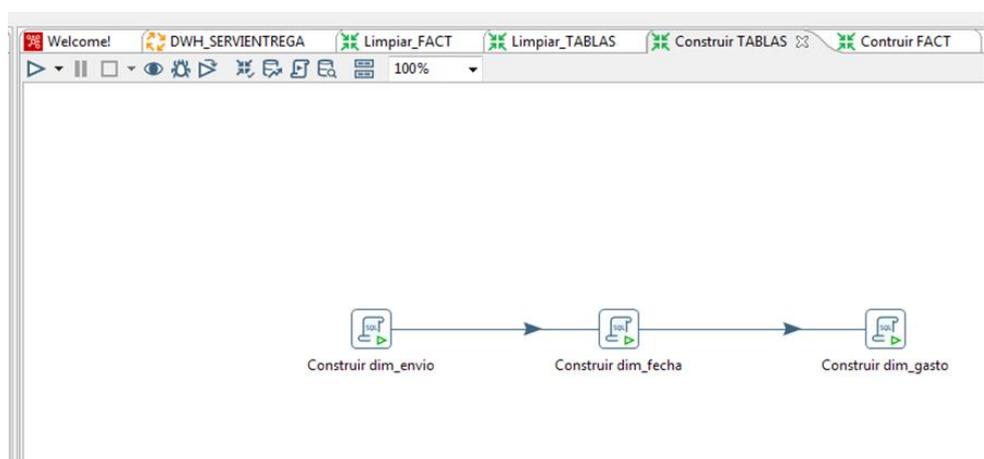


Figura 35

Proceso de limpieza inicial de las tablas de dimensiones.

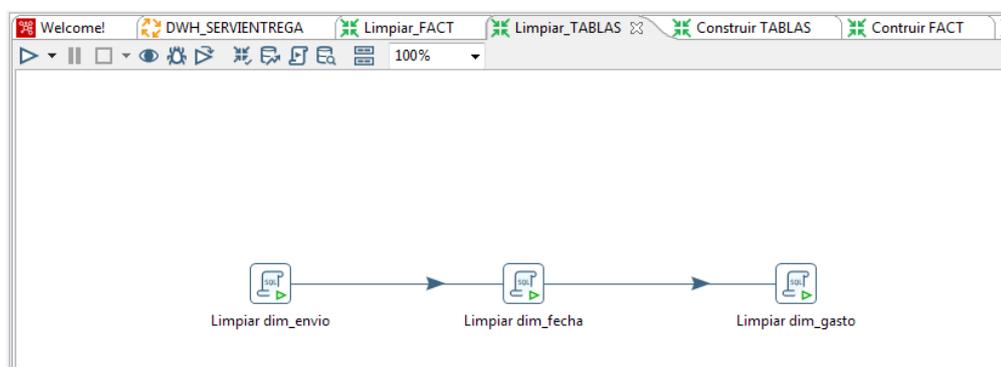
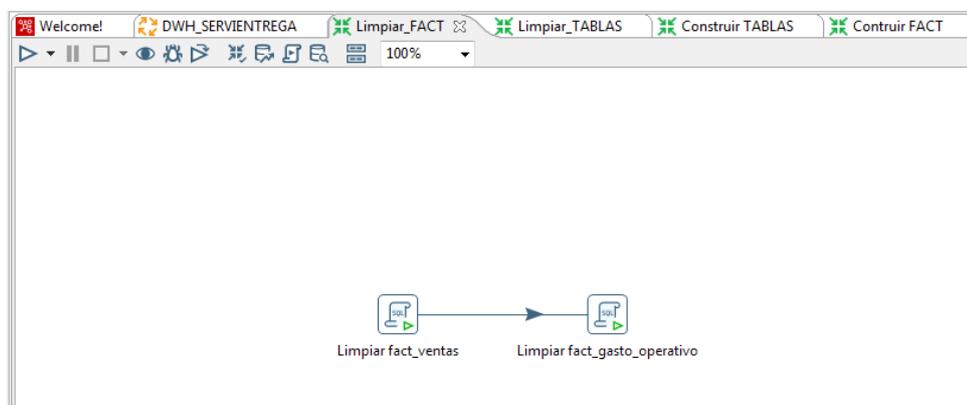
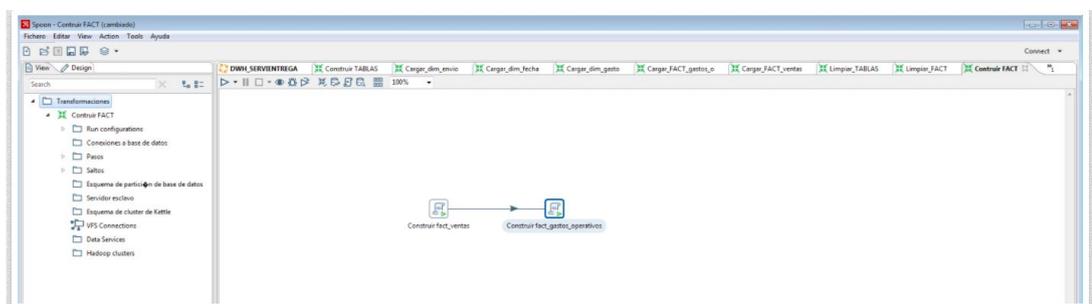


Figura 36

Proceso de limpieza inicial de las tablas de hechos

**Figura 37**

Proceso de construcción de las tablas de hechos



El Anexo 7 muestra una vista de las diferentes tareas ejecutadas dentro del proceso de extracción, transformación y carga de datos para construir en su totalidad el Data Mart DWH_SERVIENTREGA, dicho repositorio de datos representa el corazón de la solución de Inteligencia de Negocios planteada y es la fuente de información única para realizar las consultas y análisis de información requeridos por el usuario.

Diseño y desarrollo de las aplicaciones

Estas actividades fueron abordadas después de poblar el almacén de datos.

Se desarrollaron los aplicativos de acceso a los datos del Data mart: reportes, consultas ad hoc y análisis OLAP; por lo cual se utilizó la herramienta Microsoft Power BI.

Esta herramienta es muy sencilla de usar, maneja muy bien la autenticación de usuarios y su seguridad; y se puede integrar rápidamente con diferentes fuentes de datos; para efectos del proyecto la única fuente de datos a la que se conecta es el Data Mart.

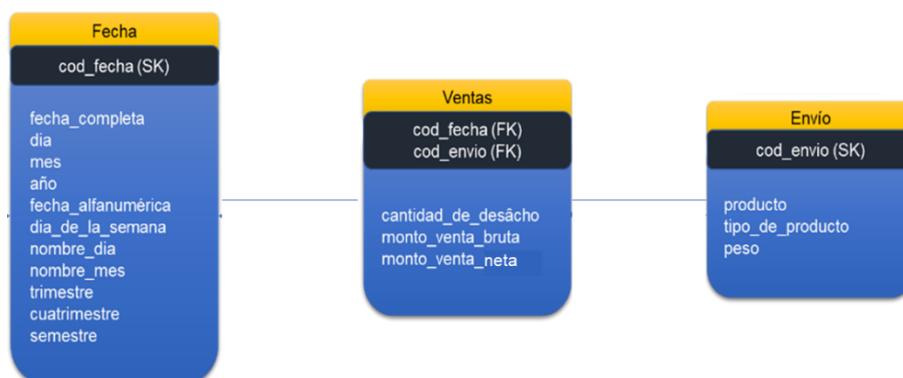
Los productos de esta fase son: el cubo y las visualizaciones de datos.

Creación del cubo y generación de la visualización de datos

El modelo del cubo es muy sencillo, conforme a las necesidades manifestadas por del usuario. Se muestra a continuación en la figura 38.

Figura 38

Modelo del cubo de datos BI_OLAP1



El cubo da la facilidad de hacer cálculos tan complejos o sencillos como los requiera el usuario, por ejemplo, se puede analizar la información correspondiente a un tipo de servicio (producto), contrastando sus detalles de ventas por fecha.

Las figuras 39, 40 y 41 muestran varias corridas del cubo para los años 2019 y 2020.

Figura 39

Corrida del cubo BI_OLAP1: años 2019 y 2020 acumulados por año.

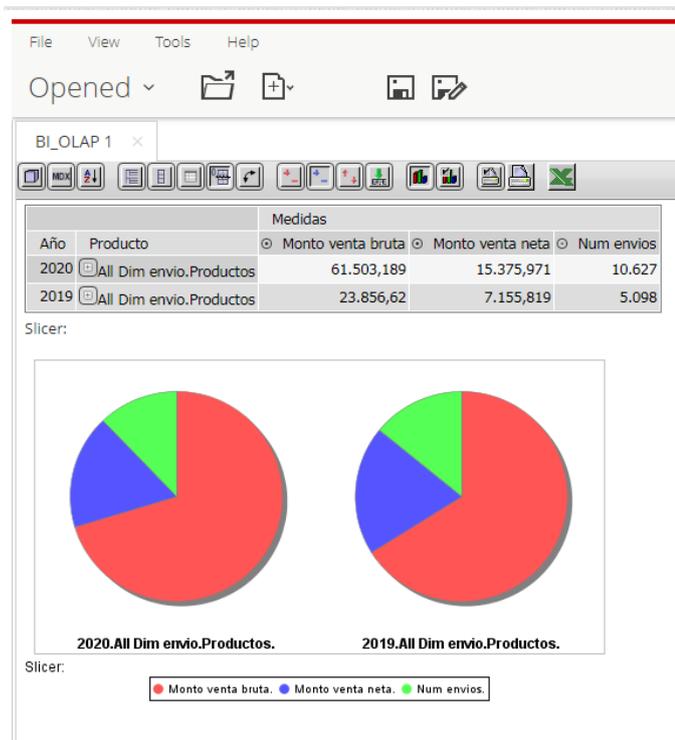


Figura 40

Corrida del cubo BI_OLAP1: año 2019 acumulado por producto.

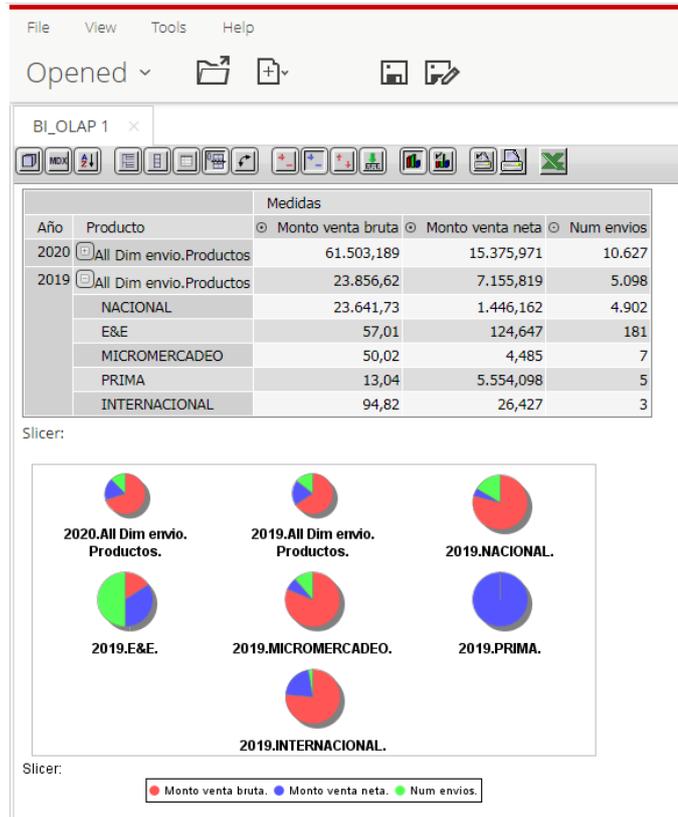
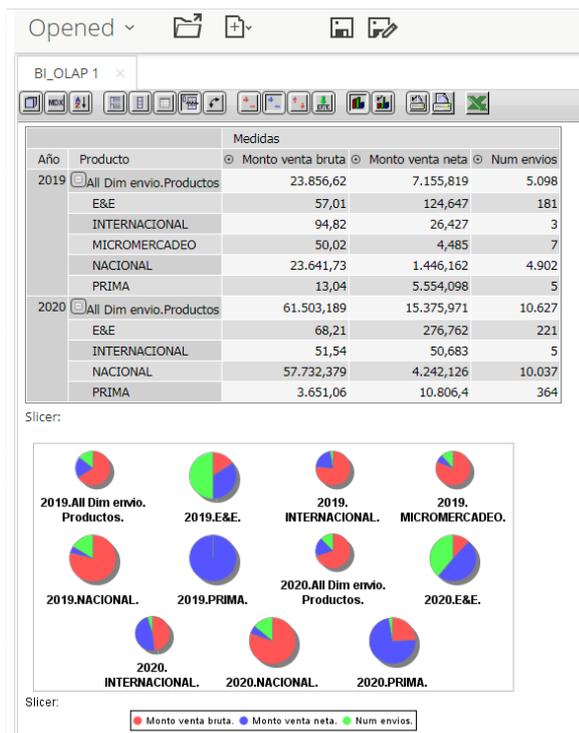


Figura 41

Corrida del cubo BI_OLAP1: años 2019 y 2020 acumulado por producto.



Implementación

Esta fase marca el inicio del pase a producción de la solución desarrollada, etapa que contempló cuatro actividades fundamentales a saber:

- Pruebas (técnicas y funcionales)
- Capacitación del usuario
- Liberación
- Aprobación del usuario

Al tratarse de un sistema poco complejo, con herramientas de uso intuitivo, la capacitación del usuario resultó bastante satisfactoria.

Las pruebas técnicas fueron realizadas sobre cada producto generado por fase y finalmente, el proceso completo.

Las pruebas con el usuario se realizaron siguiendo el script de prueba preparado para cada reporte; puede ser consultado en el anexo 4.

Tabla 21*Resumen de resultados de las pruebas funcionales*

Función	Resultado	Observación
Login al sistema	Satisfactorio	Desde la consola del usuario
Ejecutar el proceso de ETL	Satisfactorio	Haciendo clic al icono de Pentaho desde la consola del usuario.
Ejecutar Power BI	Satisfactorio	Haciendo clic al icono desde la consola del usuario
Ejecutar reportes	Satisfactorio	Haciendo clic al icono del reporte dentro de Power BI
Ejecutar consultas ad hoc	Satisfactorio	Haciendo clic al icono de consulta dentro de Power BI. Ingresando los parámetros de filtro.

Esta implementación se encuentra en fase de Producción dentro de la agencia de Courier en estudio.

Capítulo V

Análisis e interpretación de resultados

El análisis de la información ha arrojado unos resultados interesantes, que de no haber sido implementado el sistema de Inteligencia de Negocios, se hubiesen mantenido ocultos, sin aportar beneficio alguno al propietario de la agencia en cuanto a la toma de decisiones. Este problema se venía dando dentro de la agencia previo a esta implementación, y se lo pudo notar cuando se recolecto los datos (ANEXO 6), los mismos que no tenían ningún tratamiento y se encontraban incompletos. Por otra parte, era de carácter urgente realizar un análisis de datos para incorporar una estrategia debido a la creciente competencia en el sector. Con los resultados de este análisis, el propietario de la agencia aplicó inmediatamente estrategias de publicidad y personal para incrementar sus ventas basado en el tratamiento de datos.

Sobre la base de los datos suministrados para la realización del presente trabajo, se identificaron los siguientes indicadores y consultas asociadas a los procesos de la agencia, que son de gran utilidad y que serán graficados con la herramienta mencionada.

El anexo 2 muestra la descripción de los indicadores identificados en el levantamiento de requerimientos; mientras que en el anexo 3 se presenta la descripción de los reportes identificados.

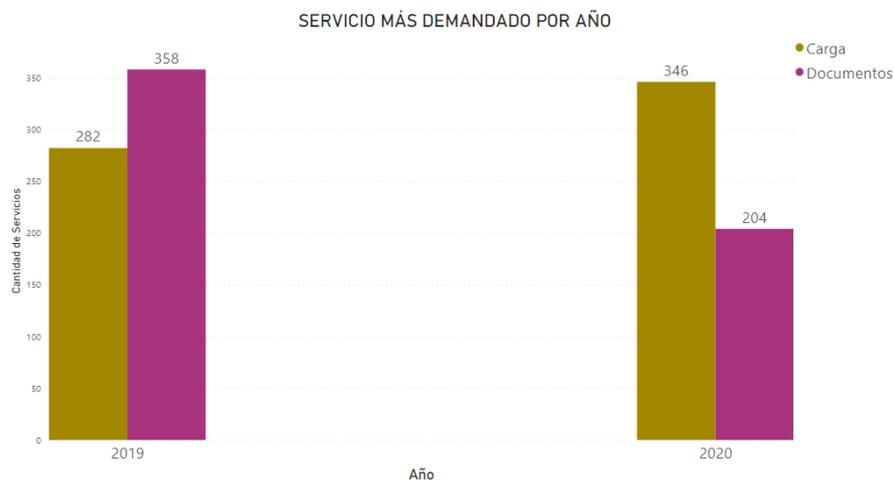
Adicionalmente, ya se ha mencionado con anterioridad, que la solución cuenta con la facilidad de realizar consultas ad hoc.

A continuación, se muestran los resultados de la ejecución de las consultas requeridas por el usuario.

Gráfico resultante de la ejecución de la consulta de Servicio más demandado: envío de documentos, envío de paquetes.

Figura 42

Servicio más demandado por año.



En la Figura 42 se analizan y comparan los servicios más demandados en los años 2019 y 2020 donde se puede apreciar las cantidades de envíos de paquetes (carga) y documentos, se observa que para el año 2020 hubo un incremento de servicios de envío de carga de un 22% al mismo tiempo que se evidencia una disminución de los servicios de envío de documentos en un 56,98% comparado con el año 2019.

Gráfico resultante de la ejecución de la consulta de Cantidad de servicios gestionados por mes en un año específico (2019)

Figura 43

Cantidad de servicios gestionados en 2019.

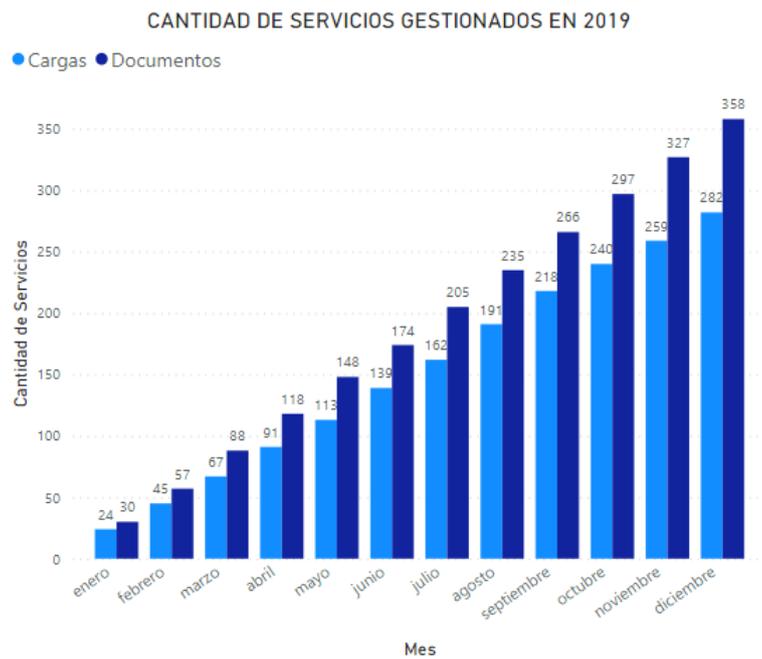
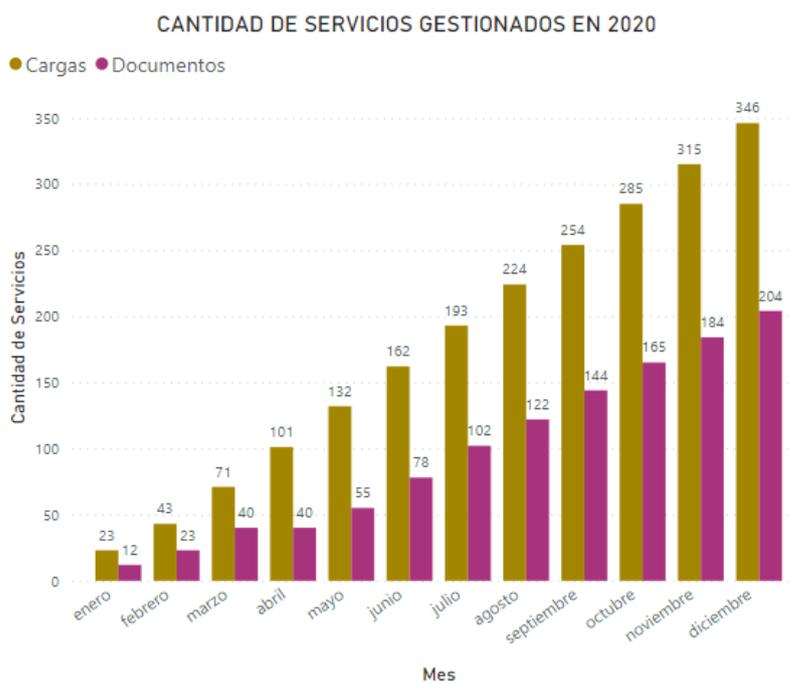


Gráfico resultante de la ejecución de la consulta de Cantidad de servicios gestionados por mes en un año específico (2020)

Figura 44

Cantidad de servicios gestionados en 2020.

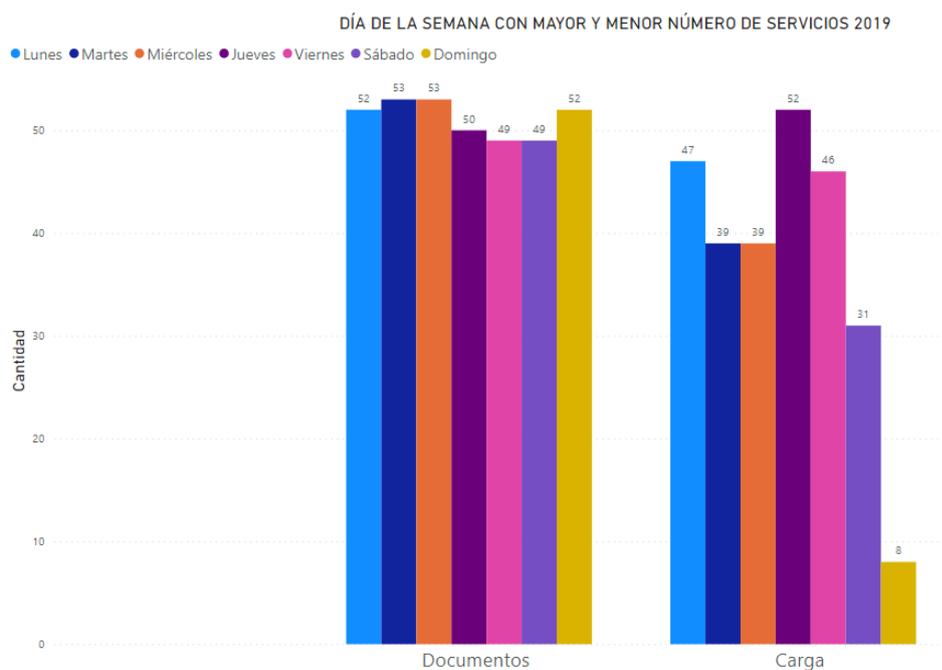


Las Figuras 43 y 44 muestran la cantidad de servicios gestionados por mes en un año específico, en el caso los años 2019 y 2020, donde se puede ver un comportamiento de aumento mes a mes de cada uno de los servicios prestados de envío (carga y documentos), se observa igualmente, que si bien hubo aumento continuo en la cantidad de ambos servicios, los documentos evidenciaron una caída de demanda en el año 2020 con respecto al 2019, mientras que aumentaron los servicios de envío de carga para el año 2020.

Gráfico resultante de la ejecución de la consulta de Día de la semana con mayor y menor número de servicios (2019)

Figura 45

Día de la semana con mayor y menor número de servicios (2019).

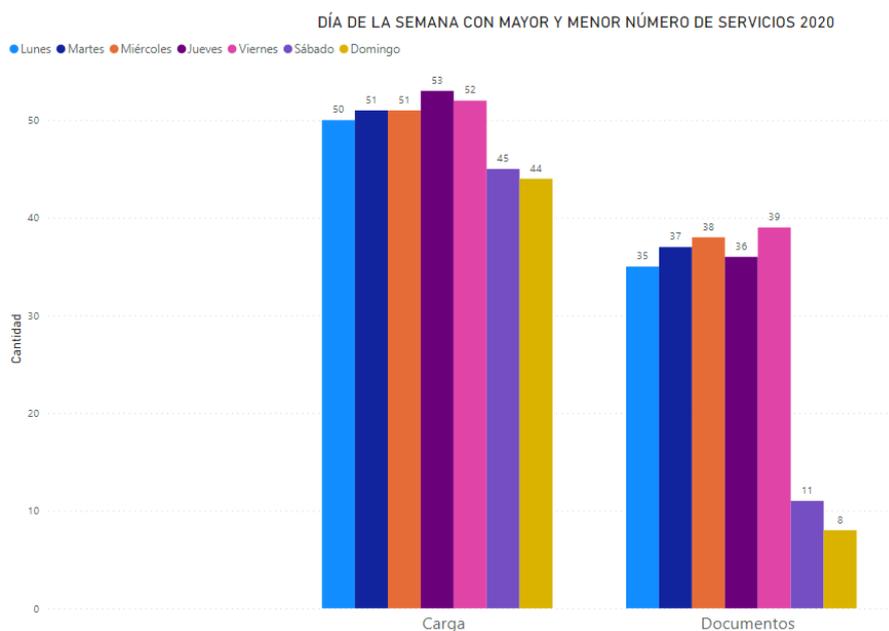


Se puede observar en la Figura 45, correspondiente al año 2019, que los días con mayor número de servicios de envíos de documentos son los días martes y miércoles, mientras que para el despacho de cargas es el día jueves donde se gestionan más servicios de este tipo. Al mismo tiempo, se puede determinar que el día con menos servicios de envío de documentos son los días viernes y sábado, mientras que para el servicio de envío de carga es el día domingo.

Gráfico resultante de la ejecución de la consulta de Día de la semana con mayor y menor número de servicios (2020)

Figura 46

Día de la semana con mayor y menor número de servicios (2020).



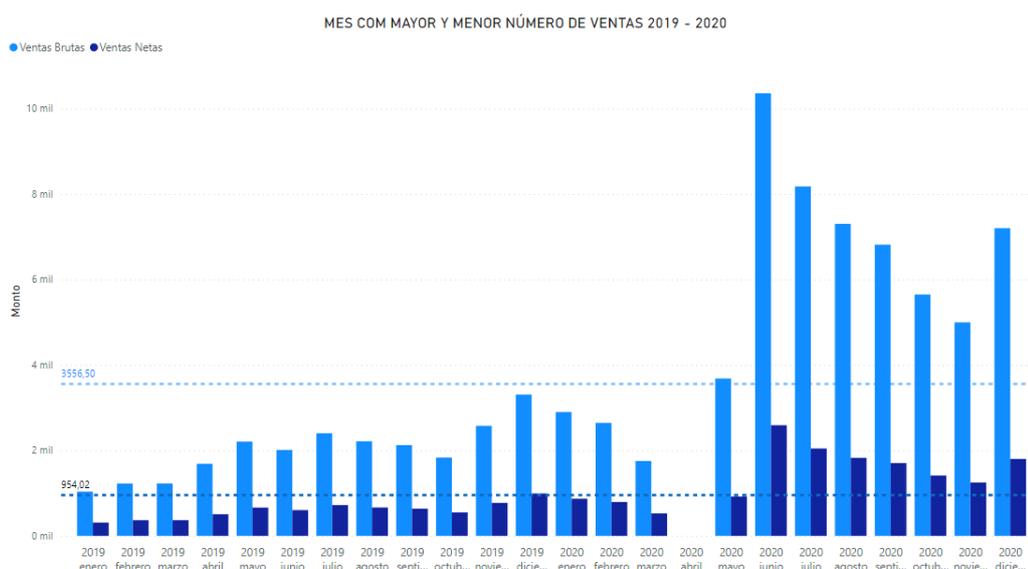
En la Figura 46, correspondiente al año 2020, el día con mayor número de servicios de envíos de documentos es el día jueves, mientras que para el envío de

cargas es el día viernes donde se gestionan más servicios de este tipo. Al mismo tiempo, se puede determinar que el día con menos servicios de envío de documentos y de carga es el día domingo.

Gráfico resultante de la ejecución de la consulta de Mes con mayor y menor número de ventas (2019 – 2020)

Figura 47

Mes con mayor y menor número de ventas (2019 – 2020)



La Figura 47, muestra el monto de las ventas netas y brutas discriminadas por mes en el período 2019 – 2020 donde se refleja que para el año 2019 el mes con mayor venta bruta registrada es diciembre con un monto total de 3.308,89, mientras que para el año 2020 es el mes de junio con un monto de 10.361,44. Por otra parte, el mes con menores ventas brutas registradas en el año 2019 es el mes de enero con un monto de 1.037,16 y para el año 2020 es el mes de marzo con un monto de 1.753,72.

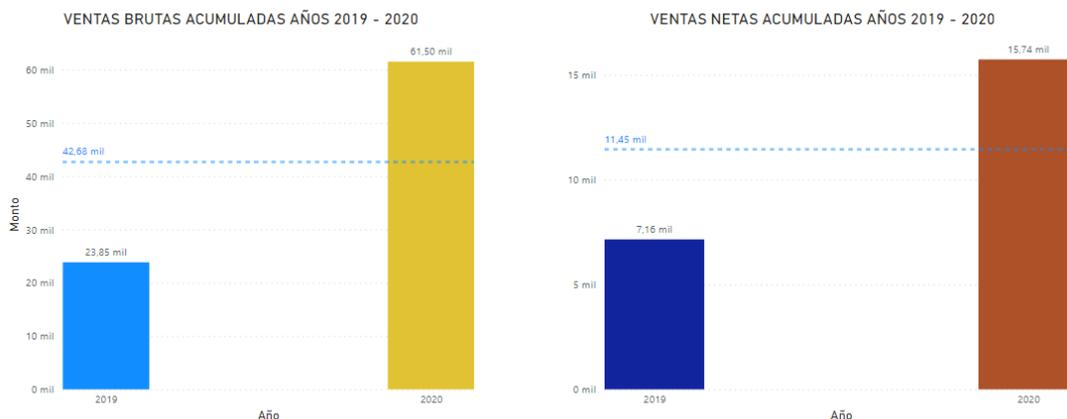
Para las ventas netas en el período 2019 – 2020, el mes con mayores ventas netas registradas para 2019 es el mes de diciembre con un monto de 992,67, mientras que para el año 2020 es el mes de junio con un monto registrado de 2590,36. El menor monto de ventas netas para el año 2019 se registró en el mes de enero por un monto de 311,15 mientras que para el año 2020 fue en el mes de mayo por un monto de 921,14. Hay que destacar que para el mes de abril de 2020 no se registraron transacciones de ventas debido a las restricciones impuestas por el gobierno para la protección ciudadana con miras a disminuir la propagación del virus Covid19.

Se debe hacer mención, que las comisiones recibidas por ventas brutas para el año 2019 era de un 30% por transacción, mientras que para el año 2020 este porcentaje fue ajustado al 25% por transacción, a pesar de la disminución de la alícuota de la comisión, se registró un incremento en las ventas en general. El monto promedio de ventas brutas en el período es de 3.556,50 y para las ventas netas es de 954,02.

Gráfico resultante de la ejecución de la consulta de Año con mayor y menor número de ventas acumuladas

Figura 48

Año con mayor y menor número de ventas acumuladas.

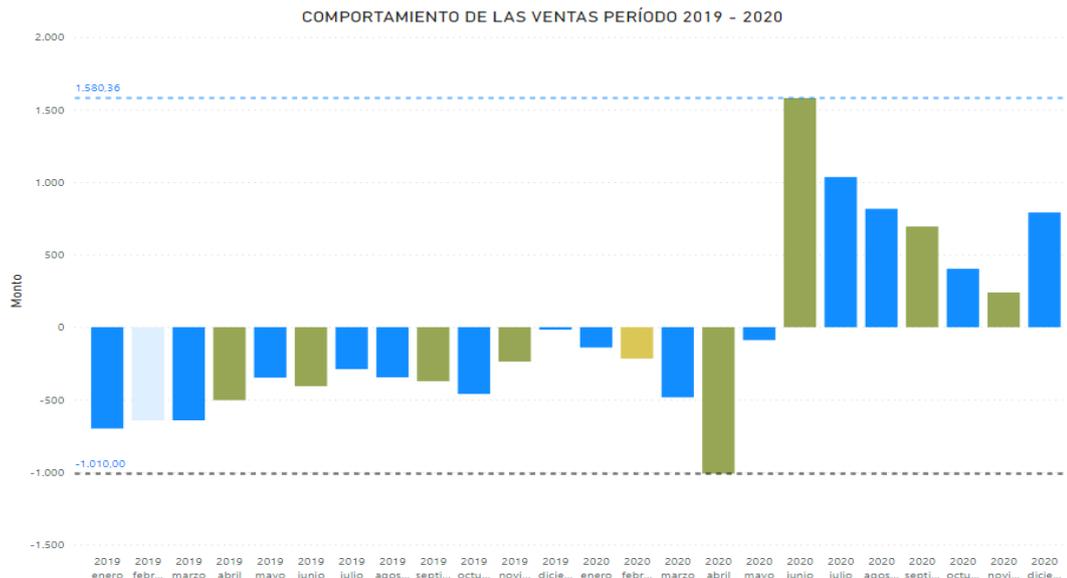


En la Figura 48 se muestran los montos de las ventas netas y brutas acumuladas por año en el período 2019 – 2020 donde se puede ver que para el año 2019 el monto de las ventas brutas acumuladas es de 23.850, mientras que para el año 2020 es de 61.500 arrojando como promedio un monto de 42.675. Para las ventas netas acumuladas en el mismo período se tiene que, para el año 2019 registra un monto de 7.160 y para el año 2020 un monto de 15.740 dando como promedio un monto de 11.450.

Gráfico resultante de la ejecución de la consulta de Comportamiento de la rentabilidad en un período

Figura 49

Comportamiento de la rentabilidad en un período.

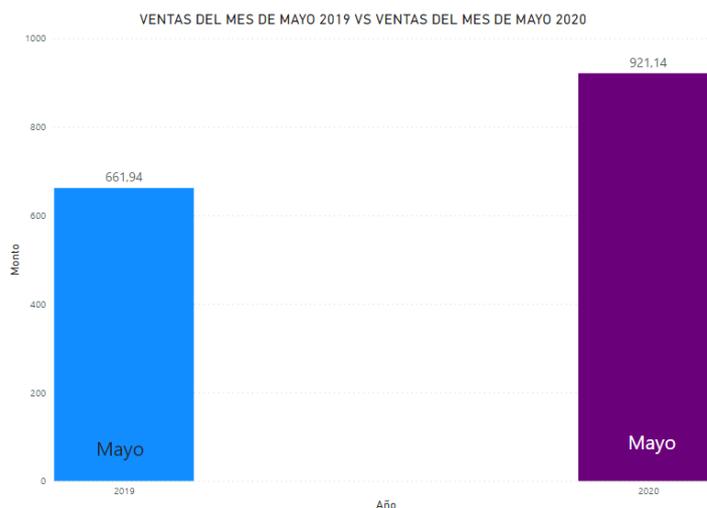


La Figura 49 representa el comportamiento de la rentabilidad en un período determinado, en el caso es 2019 – 2020, donde se puede observar que el mayor monto registrado fue en el mes de junio del año 2020 por 1.580,36 mientras que el monto más bajo registrado fue en el mes de abril del mismo año debido a que para ese mes se aplicaron medidas de seguridad para evitar la propagación del virus Covid19.

Gráfico resultante de la ejecución de la consulta de Comparativo de Ventas de mes y año anterior vs. Ventas de mes y año siguiente

Figura 50

Comparativo de Ventas mayo 2019 vs. Ventas mayo 2020.

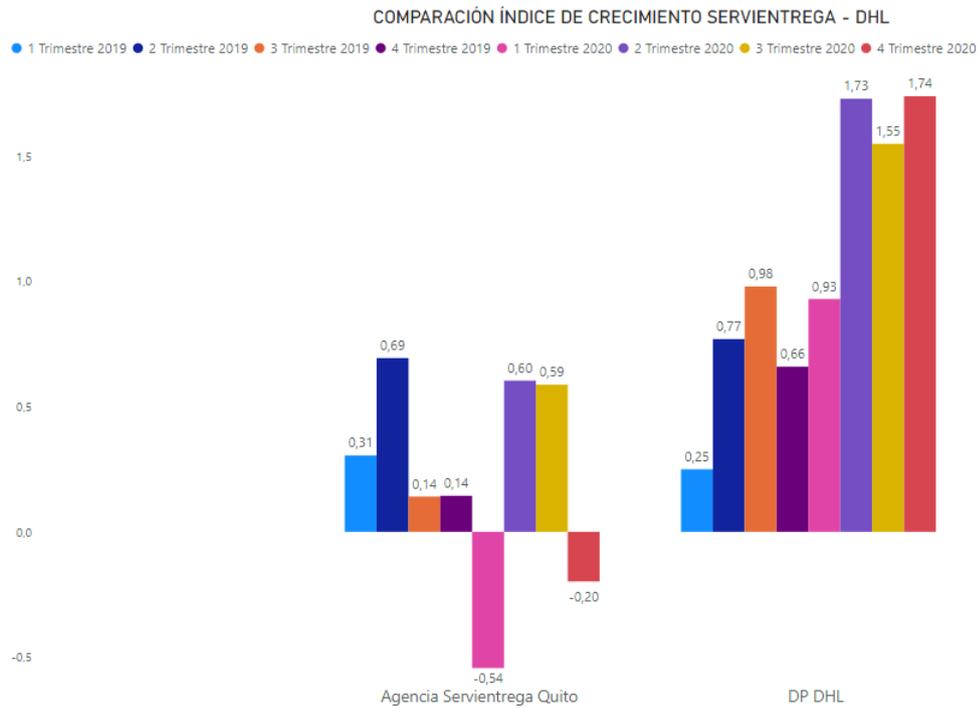


En la Figura 50 se muestra el gráfico comparativo de las ventas netas del mes de mayo del año 2019 versus el mes de mayo de 2020, donde se puede observar que hubo un incremento del 39,33% del monto del año 2020 con respecto al año 2019 del mes seleccionado.

Gráfico resultante de la ejecución de la consulta de Comparativo índice de crecimiento Servientrega vs. Competencia

Figura 51

Comparativo índice de crecimiento Servientrega - DHL.



La Figura 51 muestra la comparación del índice de crecimiento durante los cuatro trimestres del año 2019 y 2020 respectivamente entre la compañía internacional DP DHL y Servientrega S.A, (oficina en estudio), ambas presentes en el mercado de courier en la ciudad de Quito, Ecuador. El índice de crecimiento de las ventas se calcula con la siguiente fórmula:

$$ICV = \frac{\text{ventas del período actual} - \text{ventas del período anterior}}{(\text{ventas del período anterior}) * 100}$$

Se puede ver que el valor de este índice para los dos primeros trimestres del año 2019 para ambas empresas, es muy similar, sin embargo, la agencia en estudio, en los dos últimos trimestres del mismo año, estos valores presentan una disminución considerable en comparación con los dos anteriores, mientras que para la otra empresa hay un aumento significativo en el tercer trimestre, pero presentando una disminución a tomar en cuenta en el cuarto trimestre del año.

Para los trimestres del año 2020, Servientrega S.A, presenta una fuerte caída en su primer trimestre, sin embargo, para los trimestres dos y tres, se compensa y sostiene el crecimiento respectivamente, presentando de nuevo para el cuarto trimestre una caída, menos fuerte en comparación con el primero, relevante para ser tomada en cuenta por el gerente para realizar ajustes en las operaciones.

En el mismo orden de ideas, la competencia presenta para el mismo año un crecimiento sostenido y en aumento, a excepción del tercer trimestre que presenta una leve baja del índice.

El promedio del índice de crecimiento de cada una de las empresas en el período 2019 – 2020 se presenta en la Tabla 15:

Tabla 22

Promedio índice de crecimiento 2019 - 2020.

Valor del índice de crecimiento (%)										
Compañía	Trimestres 2019				Promedio	Trimestres 2020				Promedio
	1	2	3	4		1	2	3	4	
Servientrega	0,3	0,6	0,1	0,1	0,32	-	0,6	0,5	-	0,11
DP DHL	1	9	4	4		0,5	0	9	0,2	
						4			0	
	0,2	0,7	0,9	0,6	0,67	0,9	1,7	1,5	1,7	1,49
	5	7	8	6		3	3	5	4	

Conclusiones y recomendaciones

Conclusiones

- Al disponer de una visión general del negocio con respecto a su estructura organizacional y principalmente de su flujo operacional, a través de la implementación de una solución de BI se pudieron establecer estrategias que favorecen a las metas de crecimiento planteadas.
- La identificación y evaluación de las herramientas, técnicas y metodologías de Inteligencia de Negocios más utilizados, así como el estudio de la literatura, permitió seleccionar las que más se ajustaron a las necesidades (técnicas, funcionales y económicas) de la agencia de courier y usarlas en el desarrollo e implementación de la solución planteada; habilitando, al mismo tiempo, la posibilidad de crecimiento del sistema para perfeccionar sus funciones. La metodología escogida fue de Ralph Kimball y se utilizó herramientas de Pentaho y Microsoft Power BI.
- La revisión y análisis de los datos de los archivos fuentes arrojó problemas en la calidad de los mismos como: productos sin información de envío, campos nulos, campos con ceros, productos enviados sin peso registrado, inexactitud en los registros e inconsistencia; todo esto representa un panorama poco confiable en cuanto a información se refiere.
- El conocimiento del estado de los datos representa una oportunidad de mejora continua que se verá reflejada en una toma de decisiones asertiva y oportuna.
- La solución de BI implementada favoreció todo el proceso de la agencia de courier respecto a:
 - La oportunidad y temporalidad de los datos.
 - La obtención de indicadores y métricas claves del negocio cuándo y dónde se necesiten; lo que conlleva a un mejor funcionamiento actual del negocio.

- La reducción sustancial del tiempo de obtención de indicadores y métricas, y de igual manera mejorar la capacidad de respuesta ante necesidades de información.
- Permitir la toma de decisiones basadas en hechos.

Recomendaciones

- Se recomienda eliminar el registro de las transacciones de forma manual y hacer uso de los sistemas dispuestos para tal fin, de esa manera se puede asegurar una mejor calidad de los datos. La automatización de los procesos mejora la productividad de los empleados.
- Se recomienda la adquisición de un software de Gestión Empresarial (ERP) especial para PYMES, que le permita gestionar de manera integral los diferentes procesos de su negocio. La automatización de los procesos mejora la productividad de los empleados y asegura la calidad del dato.
- Se hace necesaria la capacitación continua del personal encargado del manejo de las herramientas de BI implementadas, para obtener un mayor beneficio, por los resultados del estudio del negocio; esto con el fin de poder realizar los análisis más convenientes y aprovechar al máximo el potencial de este tipo de tecnología.
- El software de Gestión Empresarial (ERP) y el sistema de Inteligencia de Negocios implementado se pueden integrar fácilmente entre sí, asegurando un buen control de los procesos de negocio, registro adecuado de las operaciones diarias y un buen proceso de toma de decisiones.

- Se debe revisar con frecuencia la infraestructura tecnológica con la finalidad de mejorar gradualmente la generación y obtención de los datos a través programas informáticos actualizados y asequibles para la empresa.

Bibliografía

- Boon Keon, S., & Nor Ezam, S. (2014). Desing and implematation of data warehouse with data model using survey-based services data. Luton, UK: Fourth edition of the International Conference on the Innovative Computing Technology (INTECH 2014).
- Bouman, R., & Van Dongen, J. (2009). Pentaho Solutions: Business Intelligence and DataWarehousing with Pentaho and MySQL. Indianapolis, IN 46256: Wiley Publishing, Inc.
- Breslin, M. (2004). www.semanticscholar.org. Obtenido de DW MODELS Data Warehousing Battle of the Giants: Comparing the Basics of the Kimball and Inmon Models.: <https://www.semanticscholar.org/paper/DW-MODELS-Data-Warehousing-Battle-of-the-Giants-%3A-Breslin/c80f8aaea5bf58846b0125b460401fed8230c2d2>
- Canales Inocencio, B. (14 de Julio de 2014). Procesos de la inteligencia de negocios. Obtenido de <https://www.gestiopolis.com/procesos-inteligencia-negocios>
- Cárdenas, J. M. (2012). Creación de una empresa de servicios de courier ubicada en la parroquia Eloy Alfaro del Cantón Quito. Quito.
- Curto, J. (2012). Introducción al Business Intelligence. Editorial UOC.
- Deckler, G. (2019). Learn Power BI. Birmingham, UK: Packt Publishing Ltd.
- Evelson, B. (2008). Topic Overview: Business Intelligence. Cambridge: Forrester Research.
- Howson, C. (2008). Successful Business Intelligence. Secrets to Making BI a Killer App. New York: McGraw-Hill.

- Inmon, W. H. (2005). Building the Data Warehouse. Indianapolis: Wiley Publishing Inc.
- Joyanes, L. (2013). Big Data: Análisis de grandes volúmenes de datos en organizaciones. México: Alfaomega Grupo Editor, S.A. de C.V.
- Khan, H. A. (2003). Data Warehousing 101: Concepts and Implementation. Lincoln: iUniverse.
- Khurram, S., & Jelena, Z. (2011). Towards goal-driven access to process warehouse: Integrating goals with process warehouse for business process analysis. Gosier: Fifth international conference on research challenges in information science.
- Kimball, R., & Ross, M. (2016). The Kimball Group Reader: Relentlessly Practical Tools for Data Warehousing and Business Intelligence. Indianapolis, IN 46256: John Wiley & Sons, Inc.
- Landa, G., Núñez, D., & Valdivia, M. (2017). Impacto de la cadena de suministros courier en el comercio internacional del país. Obtenido de <https://repositorio.up.edu.pe/handle/11354/1971>
- Limp, P. (2020). Exploring the History of Business Intelligence. Obtenido de <https://www.toptal.com/project-managers/it/history-of-business-intelligence>
- Mamani, Y. (2018). Business Intelligence: herramientas para la toma de decisiones en procesos de negocio. Universidad Nacional Micaela Bastidas de Apurímac.
- Merriam-Webster. (28 de noviembre de 2020). Data. In Merriam-Webster.com dictionary. Obtenido de <https://www.merriam-webster.com/dictionary/data>
- Navarro, G. (21 de agosto de 2020). ¿Cuáles son las fases del proceso de atención al cliente? Obtenido de Fases del proceso de atención al cliente: puntos de

dolor del cliente: <https://www.icr-evolution.com/blog/fases-del-proceso-de-atencion-al-cliente/>

Puerta, A. (2015). Business Intelligence y la tecnología de la información. IT Campus Academy.

Rosado, A. A., & Rico, D. W. (30 de abril de 2010). Inteligencia de Negocios: Estado del arte. Obtenido de *Cientia Et Technica*, 1(44), 321-326:
<https://revistas.utp.edu.co/index.php/revistaciencia/article/view/1803>

Secure Media Services. (09 de Julio de 2018). A Brief History of the Courier Industry. Obtenido de Secure Media Services:
<https://www.securemediaservices.co.uk/uncategorised/a-brief-history-of-the-courier-industry/>

Servientrega Ecuador S.A. (2015). Servientrega. Obtenido de Quienes Somos:
<http://www.servientrega.com.ec/index.html>

Shenglei, P., & Guoqing, J. (2014). Application of Data Warehouse and Data Mining in the Steel Enterprise Information Integration System. Hangzhou, China: Sixth International Conference on Intelligent Human-Machine Systems and Cybernetics.

Sretenović, M., Kovačić, B., & Jovanović, V. (2016). Development of the data warehouse model for public authorities accounts in Croatia. Opatija: 39th International Convention on Information and Communication Technology, Electronics and Microelectronics (MIPRO).

Tableau Software, LLC. (2021). www.tableau.com. Obtenido de www.tableau.com/why-tableau: www.tableau.com

TNT-Fedex. (09 de Junio de 2015). History of courier services. Obtenido de <https://direct.tnt.co.uk/blog/historyofcourierservices>

- Trujillo, J. C., Mazón, J. N., & Pardillo, J. (2011). Diseño y explotación de almacenes de datos. Conceptos básicos de modelado multidimensional. Alicante: ECU - Editorial Club Universitario.
- Ucha, F. (noviembre de 2009). Definición de Atención al cliente. Obtenido de Definición ABC: <https://www.definicionabc.com/economia/atencion-al-cliente.php>
- Vercellis, C. (2009). Business Intelligence: Data Mining and Optimization for Decision Making. Chichester, West Sussex: John Wiley & Sons, Ltd.
- Wrembel, R., & Koncilia, C. (2007). Data Warehouses and OLAP: Concepts, Architectures and Solutions. Hershey, PA: IRM Press (an imprint of Idea Group Inc.).
- Zonalogística. (04 de diciembre de 2019). ¿Cómo medir el servicio al cliente? Obtenido de Zonalogística Conocimiento de Expertos: <https://zonalogistica.com/como-medir-el-servicio-al-cliente/>

