



**“Análisis, diseño e implementación de un modelo de Business Intelligence,
enfocado a la toma de decisiones de la gestión de investigación del
Vicerrectorado de Investigación, Innovación y Transferencia de Tecnología de la
UFA-ESPE”**

Acosta Córdova, Diego Hernán y Cajamarca Tobay, Jhonny Orlando

Departamento de Ciencias de la Computación

Carrera de Ingeniería de Sistemas e Informática

Trabajo de titulación, previo a la obtención del título de Ingeniero en Sistemas e
Informática

Ing. Reyes Chicango, Rolando Patricio Ph. D

22 de junio del 2020



Urkund Analysis Result

Analysed Document: Acosta Córdova Diego_Cajamarca Tobay Jhonny.docx (D75843077)
Submitted: 6/30/2020 6:49:00 PM
Submitted By: mgutierrez@difusion.com.mx
Significance: 0 %

Sources included in the report:

Instances where selected sources appear:

0



Firmado electrónicamente por:
**ROLANDO PATRICIO
REYES CHICANGO**

Ing. Rolando Patricio, Reyes Chicango Ph.D
Director de Proyecto



DEPARTAMENTO DE CIENCIAS DE LA COMPUTACIÓN

CARRERA DE TECNOLOGÍAS DE LA INFORMACIÓN

CERTIFICACIÓN

Certifico que el trabajo de titulación, “Análisis, diseño e implementación de un modelo de Business Intelligence, enfocado a la toma de decisiones de la gestión de investigación del Vicerrectorado de Investigación, Innovación y Transferencia de Tecnología de la UFA-ESPE” fue realizado por los señores *Acosta Córdova, Diego Hernán y Cajamarca Tobay, Jhonny Orlando*; el mismo que ha sido revisado en su totalidad y analizado por la herramienta de verificación de similitud de contenido; por lo tanto cumple con los requisitos teóricos legales, científicos, técnicos y metodológicos establecidos por la Universidad de Fuerzas Armadas ESPE, razón por la cual me permito acreditar y autorizar para que lo sustenten públicamente.

Sangolquí, 22 junio del 2020

Firma:

Ing. Reyes Chicango, Rolando Patricio Ph.D

CC: 1713270328



ESPE
UNIVERSIDAD DE LAS FUERZAS ARMADAS
INNOVACIÓN PARA LA EXCELENCIA

**DEPARTAMENTO DE CIENCIAS DE LA COMPUTACIÓN
CARRERA DE TECNOLOGÍAS DE LA INFORMACIÓN**

RESPONSABILIDAD DE AUTORÍA

Nosotros, *Acosta Córdova, Diego Hernán* con cédula de ciudadanía: 1715882328 y *Cajamarca Tobay, Jhonny Orlando* con cédula de ciudadanía: 1723644488, declaramos que el contenido, ideas y criterios del trabajo de titulación: **“Análisis, diseño e implementación de un modelo de Business Intelligence, enfocado a la toma de decisiones de la gestión de investigación del Vicerrectorado de Investigación, Innovación y Transferencia de Tecnología de la UFA-ESPE”** es de nuestra autoría y responsabilidad, cumpliendo con los requisitos legales, teóricos, científicos, técnicos y metodológicos establecidos por la Universidad de Fuerzas Armadas ESPE, respetando los derechos intelectuales de terceros y referenciando las citas bibliográficas.

Consecuentemente el contenido de la investigación mencionada es veraz.

Sangolquí, 22 junio del 2020.

Firmas:

Acosta Córdova, Diego Hernán

CC: 1715882328

Cajamarca Tobay, Jhonny Orlando

CC: 1723644488



ESPE

UNIVERSIDAD DE LAS FUERZAS ARMADAS
INNOVACIÓN PARA LA EXCELENCIA

DEPARTAMENTO DE CIENCIAS DE LA COMPUTACIÓN

CARRERA DE TECNOLOGÍAS DE LA INFORMACIÓN

AUTORIZACIÓN DE PUBLICACIÓN

Nosotros, *Acosta Córdova, Diego Hernán* con cédula de ciudadanía: 1715882328 y *Cajamarca Tobay, Jhonny Orlando* con cédula de ciudadanía: 1723644488, autorizamos a la Universidad de las Fuerzas Armadas ESPE, publicar el trabajo de titulación: "Análisis, diseño e implementación de un modelo de Business Intelligence, enfocado a la toma de decisiones de la gestión de investigación del Vicerrectorado de Investigación, Innovación y Transferencia de Tecnología de la UFA-ESPE" en el Repositorio Institucional, cuyo contenido, ideas y criterios son de nuestra responsabilidad.

Sangolquí, 22 junio del 2020.

Firmas:

Acosta Córdova, Diego Hernán

CC: 1715882328

Cajamarca Tobay, Jhonny Orlando

CC: 1723644488

DEDICATORIA

El presente trabajo de titulación está dedicado a Dios, por permitirme finalizar mi carrera profesional y estar rodeado de una hermosa familia y de buenos amigos. A mi esposa María Albán por su amor incondicional, quién me ayudado siempre a finalizar las metas trazadas. A mis padres Euler Acosta y Varsovia Córdova, quienes me guiaron por el buen camino y me dieron el impulso para seguir estudiando. A mis hermanos José y Deisy por sus consejos y una grata niñez. A mi abuelito, tías y primos con quienes he compartido buenos momentos durante mi vida y siempre han estado prestos a ayudarme.

Diego Acosta

El presente trabajo de titulación está dedicado principalmente a Dios y a mis padres: Jorge Cajamarca y Rosa Tobay; por brindarme su amor, esfuerzo, trabajo y sacrificio durante todos estos años. De igual forma a mi familia y amigos del círculo cercano, gracias a su apoyo y cariño incondicional logré alcanzar una meta más trazada en mi vida, fortaleciéndome de esta manera en el aspecto humano y profesional. A todos aquellos que me brindaron una mano amiga y no permitieron rendirme en el camino es dedicado el presente trabajo.

Jhonny Cajamarca.

AGRADECIMIENTO

Al finalizar este trabajo de titulación lleno de retos y expectativas, quiero agradecer a todas las personas que ayudaron de forma directa o indirectamente a concluir este trabajo. Un agradecimiento especial a mi familia que siempre estuvo presente en cada momento de mi formación como Ingeniero en Sistemas e Informática. Un agradecimiento especial a los docentes, quienes basados en su experiencia y profesionalismo nos ayudaron encaminándonos por el mejor camino a llegar a los objetivos propuestos. También quiero agradecer a la Universidad de las Fuerzas Armadas – ESPE, institución de calidad que me abrió las puertas para mi formación profesional.

Diego Acosta

El presente trabajo de titulación es fruto de la dedicación y esfuerzo de mis padres y familiares cercanos, que fueron mi motivación para alcanzar un objetivo más. Gracias por ser las personas que me motivaron a convertir mis sueños en realidad, por acompañarme en cada agotadora jornada de estudio y por confiar en mis destrezas y habilidades. Gracias a Dios por sus bendiciones e infinito amor y también por permitirme estar rodeado de excelentes personas, que con su ejemplo supieron encaminarme por el camino correcto brindándome siempre ese apoyo incondicional y de esta manera recorrer este largo camino lleno de obstáculos y sacrificios; donde sin duda alguna todos fueron directa e indirectamente parte de este logro alcanzado. Por esto y muchas más razones que quizá las palabras sean cortas para expresarlo quedo muy agradecido y orgulloso del trabajo alcanzado.

Jhonny Cajamarca.

ÍNDICE DE CONTENIDOS

PORTADA	1
CERTIFICACIÓN	3
RESPONSABILIDAD DE AUTORÍA	4
AUTORIZACIÓN DE PUBLICACIÓN	5
DEDICATORIA	6
AGRADECIMIENTO	7
ÍNDICE DE CONTENIDOS	8
ÍNDICE DE TABLAS	10
ÍNDICE DE FIGURAS	11
RESUMEN	13
ABSTRACT	14
CAPÍTULO I	15
INTRODUCCIÓN	15
Antecedentes	15
Planteamiento del Problema	17
Justificación	18
Objetivos	20
Objetivo General	20
Objetivos Específicos	20
Alcance	21
CAPÍTULO II	23
ESTADO DEL ARTE	23
Marco Teórico	23
Inteligencia de negocios	23
Metodologías de uso para BI	25
Herramientas de software de BI	28
Procesos ETL	32
Bases de Datos	34
Cubos de Información	38
Indicadores (KPI)	40
Dashboard	41
Revisión Sistemática de Literatura Ligera	42
CAPÍTULO III	56
METODOLOGÍA DE LA INVESTIGACIÓN	56
Diseño de la investigación	56
CAPÍTULO IV	64
MAPEO DE DATOS	64
Establecimiento	64
Implementación	84
Construcción del datamart con la metodología Kimball	84
Procesos ETL	105
Construcción de los dashboard	111
CAPÍTULO V	121

	9
PRESENTACIÓN DE LOS RESULTADOS _____	121
Validación de datos _____	121
Resultados de la validación de datos _____	122
CONCLUSIONES Y RECOMENDACIONES _____	127
Conclusiones _____	127
Recomendaciones _____	128
ANEXOS _____	130
BIBLIOGRAFÍA _____	141

ÍNDICE DE TABLAS

Tabla 1. <i>PREGUNTAS DE INVESTIGACIÓN</i>	21
Tabla 2. <i>HERRAMIENTAS DE SOFTWARE DE BI</i>	29
Tabla 3. <i>GRUPO DE CONTROL DEFINIDO (GC)</i>	45
Tabla 4. <i>CADENA DE BÚSQUEDA</i>	47
Tabla 5. <i>ACTIVIDADES DE SUSTENTACIÓN DEL PROYECTO</i>	56
Tabla 6. <i>CRITERIOS PARA LA TRANSFORMACIÓN DE DATOS</i>	61
Tabla 7. <i>ENTREVISTA REALIZADA AL DIRECTOR DE LA UGI</i>	67
Tabla 8. <i>ENTREVISTA DEL ANALISTA DE GESTIÓN DE PROYECTOS</i>	68
Tabla 9. <i>ENTREVISTA DEL ANALISTA FINANCIERO DE INVESTIGACIÓN</i> ..	69
Tabla 10. <i>ENTREVISTA DE LA ANALISTA DE DIFUSIÓN DE CIENCIA</i>	70
Tabla 11. <i>ENTREVISTA REALIZADA A LA ANALISTA DE MOVILIDAD</i>	71
Tabla 12. <i>INDICADOR DE PROYECTOS DE INVESTIGACIÓN I</i>	72
Tabla 13. <i>INDICADOR DE PROYECTOS DE INVESTIGACIÓN II</i>	73
Tabla 14. <i>INDICADOR DE PROYECTOS DE INVESTIGACIÓN III</i>	73
Tabla 15. <i>INDICADOR DE GRUPOS DE INVESTIGACIÓN IV</i>	74
Tabla 16. <i>INDICADOR GRUPOS DE INVESTIGACIÓN V</i>	74
Tabla 17. <i>INDICADOR DE DIFUSIÓN I</i>	75
Tabla 18. <i>INDICADOR DE DIFUSIÓN II</i>	75
Tabla 19. <i>INDICADOR DE DIFUSIÓN III</i>	76
Tabla 20. <i>INDICADOR DE DIFUSIÓN IV</i>	76
Tabla 21. <i>INDICADOR DE DIFUSIÓN V</i>	77
Tabla 22. <i>INDICADOR DE MOVILIDAD I</i>	77
Tabla 23. <i>INDICADOR DE MOVILIDAD II</i>	78
Tabla 24. <i>INDICADOR DE MOVILIDAD III</i>	78
Tabla 25. <i>INDICADOR DE MOVILIDAD IV</i>	79
Tabla 26. <i>INDICADOR DE MOVILIDAD V</i>	79
Tabla 27. <i>INDICADOR FINANCIERO I</i>	80
Tabla 28. <i>INDICADOR FINANCIERO II</i>	81
Tabla 29. <i>INDICADOR FINANCIERO III</i>	81
Tabla 30. <i>INDICADOR FINANCIERO IV</i>	82
Tabla 31. <i>INDICADOR FINANCIERO V</i>	82
Tabla 32. <i>FUENTES DE DATOS DE LOS INDICADORES DE RESULTADOS</i> 83	83
Tabla 33. <i>PLANIFICACIÓN DE LA CONSTRUCCIÓN DEL DATAMART</i>	84
Tabla 34. <i>DEFINICIÓN DE REQUERIMIENTOS</i>	85
Tabla 35. <i>HERRAMIENTAS DE SOFTWARE</i>	88
Tabla 36. <i>RESULTADOS DE LAS ENCUESTAS DEL MODELO DE BI IMPLEMENTADO</i>	123

ÍNDICE DE FIGURAS

Figura 1. <i>RETOS EN LA IMPLEMENTACIÓN DE BI</i>	24
Figura 2. <i>PROCESO DE BUSINESS INTELLIGENCE</i>	25
Figura 3. <i>CICLO DE VIDA DIMENSIONAL DEL NEGOCIO</i>	26
Figura 4. <i>ARQUITECTURA DE KIMBALL</i>	27
Figura 5. <i>ARQUITECTURA DE INMON</i>	28
Figura 6. <i>CUADRO MÁGICO DE GARTNER</i>	31
Figura 7. <i>PROCESO ETL</i>	32
Figura 8. <i>ESQUEMA EN FORMA DE ESTRELLA</i>	36
Figura 9. <i>ESQUEMA COPO DE NIEVE</i>	37
Figura 10. <i>ARQUITECTURA DE BI APLICADA EN LA UTM</i>	52
Figura 11. <i>DISEÑO DE LA INVESTIGACIÓN</i>	58
Figura 12. <i>REPRESENTACIONES ESTADÍSTICAS GRÁFICAS -TABLEAU</i> ...	62
Figura 13. <i>ORGANIGRAMA DE LA UFA-ESPE</i>	65
Figura 14. <i>DISEÑO DE ARQUITECTURA TÉCNICA</i>	87
Figura 15. <i>INSTALACIÓN DE ORACLE 10 G</i>	88
Figura 16. <i>ORACLE SQL DEVELOPER</i>	89
Figura 17. <i>INSTALACIÓN DEL POWER DESIGNER</i>	90
Figura 18. <i>SOFTWARE POWER DESIGNER INSTALADO</i>	90
Figura 19. <i>SOFTWARE PENTAHO DATA INTEGRATION</i>	91
Figura 20. <i>INSTALACIÓN DE TABLEAU DESKTOP</i>	92
Figura 21. <i>CREACIÓN TABLA BASE ESIGEF</i>	93
Figura 22. <i>CREACIÓN TABLA PAGOS</i>	93
Figura 23. <i>CREACIÓN TABLAS TEMPORALES FINANCIERAS</i>	94
Figura 24. <i>CREACIÓN TABLA TEMPORAL DE PROYECTOS DE INVESTIGACIÓN</i>	95
Figura 25. <i>CREACIÓN TABLA TEMPORAL DE MOVILIDAD</i>	95
Figura 26. <i>CREACIÓN DE LAS TABLAS TEMPORALES DE DIFUSIÓN</i>	96
Figura 27. <i>GENERACIÓN DE SCRIPTS</i>	97
Figura 28. <i>MODELO MULTIDIMENSIONAL FINANCIERO</i>	98
Figura 29. <i>MODELO MULTIDIMENSIONAL DE PROYECTOS DE INVESTIGACIÓN</i>	99
Figura 30. <i>MODELO MULTIDIMENSIONAL DE MOVILIDAD</i>	99
Figura 31. <i>MODELO MULTIDIMENSIONAL DE DIFUSIÓN</i>	100
Figura 32. <i>CONEXIÓN DE BASE DE DATOS</i>	101
Figura 33. <i>CREACIÓN DE LOS TABLESPACE</i>	101
Figura 34. <i>CREACIÓN DE USUARIOS</i>	102
Figura 35. <i>PRIVILEGIOS DE USUARIO</i>	102
Figura 36. <i>CONEXIÓN AL DSA BU UGI ESPE</i>	103
Figura 37. <i>SCHEMAS DSA</i>	104
Figura 38. <i>CREACIÓN DE TABLAS FINANCIERAS EN EL ESQUEMA DSA</i> . 104	104
Figura 39. <i>CREACIÓN DE TABLAS FINANCIERAS EN EL ESQUEMA DWH</i> 105	105
Figura 40. <i>TRANSFORMACIÓN DE DATOS</i>	106
Figura 41. <i>CONEXIÓN A ORACLE</i>	106

	12
Figura 42. OBJETO PARA OBTENER DATOS DE ARCHIVOS EXCEL.....	107
Figura 43. CONFIGURACIÓN DEL OBJETO MICROSOFT EXCEL INPUT...	108
Figura 44. NOMBRE DE LA HOJA DEL ARCHIVO EXCEL.....	108
Figura 45. CONFIGURACIÓN OBJETO TABLE OUTPUT.....	109
Figura 46. CONFIGURACIÓN TABLA DESTINO.....	109
Figura 47. UNIÓN DE OBJETOS.....	110
Figura 48. PROCESOS DE LIMPIEZA DE DATOS.....	110
Figura 49. EJECUCIÓN DE STORE PROCEDURES.....	111
Figura 50. CONEXIÓN DE TABLEAU A LA BASE DE DATOS DE ORACLE.	112
Figura 51. ACCESO A LA INFORMACIÓN DE LA BASE DE DATOS DE ORACLE.....	113
Figura 52. CREACIÓN DE DASHBOARD FINANCIEROS.....	113
Figura 53. SCRIPT SQL PERSONALIZADO.....	114
Figura 54. NUEVA PERSONALIZACIÓN DE SQL.....	114
Figura 55. EXTRACCIÓN DE DATOS DE TABLEAU.....	115
Figura 56. COPIA DE DATOS EN ARCHIVO TIPO HYPER.....	115
Figura 57. HERRAMIENTAS DE TABLEAU.....	116
Figura 58. RESULTADOS DE LOS INDICADORES FINANCIEROS.....	117
Figura 59. RESULTADOS DE LOS PROYECTOS DE INVESTIGACIÓN.....	118
Figura 60. RESULTADOS DE LOS INDICADORES DE MOVILIDAD.....	119
Figura 61. RESULTADOS DE LOS INDICADORES DE DIFUSIÓN.....	120
Figura 62. RESULTADOS DE LA ENCUESTA DEL MODELO DE BI IMPLEMENTADO.....	124
Figura 63. GRADO DE ACEPTACIÓN DEL SISTEMA DE BI.....	124
Figura 64. USABILIDAD DEL MODELO DE BI IMPLEMENTADO.....	125

RESUMEN

La investigación y producción científica son uno de los pilares fundamentales en las Instituciones de educación superior, por tal motivo, la Universidad de las Fuerzas Armadas "ESPE" posee en su organización estructural al Vicerrectorado de Investigación, Innovación y Transferencia de Tecnología (VIITT) que cumple con las actividades de docencia, investigación científica y vinculación con la sociedad. Dentro de este vicerrectorado se encuentra la Unidad de Gestión de la Investigación (UGI), encargada de las actividades de investigación de la universidad.

La UGI posee bastante información esparcida en varias fuentes de datos, por tal motivo existe muchas de las veces duplicidad de datos que dificulta el análisis de la información y la toma de decisiones de forma eficiente. Para dar una solución al problema encontrado se propone diseñar un modelo de Business Intelligence enfocado a la toma de decisiones de la gestión de Investigación del VIITT. Al final de este trabajo de titulación, se cuenta con modelos de datamarts que almacenan los datos más relevantes de la UGI. Además, la información depurada se presenta en dashboards elaborados en base a indicadores de resultados apegados a los parámetros de evaluación del Consejo de Aseguramiento de la Calidad de la Educación Superior (CACES), que permite conocer el avance del cumplimiento de las actividades de investigación dando una concepción más amplia para una mejor toma de decisiones.

PALABRAS CLAVE:

- **ACTIVIDADES DE INVESTIGACIÓN.**
- **INDICADORES.**
- **DATAMART.**
- **BUSINESS INTELLIGENCE.**
- **TOMA DE DECISIONES.**

ABSTRACT

Scientific research and production are one of the fundamental pillars in higher education institutions, for this reason, the University of the Armed Forces "ESPE" has in its structural organization the Vice-Rector's Office for Research, Innovation and Technology Transfer (VIITT) that complies with teaching, scientific research, and links with society. Within this vice-chancellor's office is the Research Management Unit (UGI), in charge of the university's research activities.

The UGI has a lot of information scattered in several data sources, for this reason there is often duplication of data that makes it difficult to analyze the information and make decisions efficiently. In order to provide a solution to the problem encountered, it is proposed to design a Business Intelligence model focused on decision-making in VIITT Research management. At the end of this degree work, there are models of datamarts that store the most relevant UGI data. In addition, the refined information is presented in dashboards prepared based on results indicators attached to the evaluation parameters of the Higher Education Quality Assurance Council (CACES), which allows us to know the progress of compliance with research activities giving a broader conception for better decision making.

KEY WORDS:

- Research Activities.
- Indicators.
- Datamart.
- Business Intelligence.
- Decision making.

CAPÍTULO I

INTRODUCCIÓN

Antecedentes

De acuerdo con la ley orgánica de educación superior (LOES) que entró en vigor el 12 de octubre del 2010, consideró a la calidad académica como un pilar importante en las instituciones de educación superior. Por ello, ha iniciado un proceso de mejora continua, la cual, está basada en las directrices desarrollados por los órganos rectores de la educación superior como son: la Secretaria de Educación Superior, Ciencia, Tecnología e Innovación (SENESCYT), Consejo de Educación Superior (CES) y Consejo de Aseguramiento de la Calidad de la Educación Superior (CACES). Estas entidades, fomentan principalmente los pilares de investigación, ciencia, tecnología e innovación para el desarrollo nacional.

Por otra parte, se encuentra la LOES en su Art. 8, literales d y k respectivamente, donde consagra que las instituciones de educación superior deben: “aportar al desarrollo del pensamiento universal, al despliegue de la producción científica y a la promoción de las transferencias e innovaciones tecnológicas”, además de “fomentar y ejecutar programas de investigación de carácter científico tecnológico y pedagógico que coadyuven al mejoramiento y protección del ambiente y promuevan al desarrollo sustentable nacional.” (LOES, 2010). Es por esta razón, se crea el Art. 10 del Estatuto de la UFA-ESPE, referente a los órganos colegiados en donde se determina la existencia del Vicerrectorado de Investigación, Innovación y transferencia de Tecnología (VIITT). El cual, debe alinearse a las disposiciones de investigación, innovación y transferencia de tecnología de la LOES. Así mismo, mediante la Orden de

Rectorado ESPE-HCUP-OR-2015-001, se aprueba el Reglamento Orgánico de Gestión Organizacional por procesos de la UFA-ESPE, en donde en su Art 54 se establece claramente la gestión de la investigación y mediante la resolución ESPE-HCUP-ES-2015-075 se aprueba el definitivo Reglamento de Investigación de la UFA-ESPE (HCU, 2015) que articula las actividades de investigación, desarrollo, innovación y vinculación de manera sistemática para promover el desarrollo de la investigación y así lograr la solución de problemas de las Fuerzas Armadas y de la sociedad ecuatoriana.

Ahora bien, respecto al proceso de investigación de la ESPE, este se inicia con la definición de las líneas de investigación, y en base a ellas, se realiza las diferentes convocatorias para presentar proyectos de investigación científica y tecnológica tanto civil como militar. Los proyectos de investigación presentados son evaluados por pares ciegos, luego se los prioriza a fin de formar parte del portafolio de proyectos de investigación de la ESPE (UGI, 2018). Como parte resultante de los proyectos de investigación existen las publicaciones indexadas, las cuales han permitido que la UFA-ESPE obtenga el primer lugar en los años: 2016, 2017 y el tercer lugar en el año: 2018 de los artículos publicados en revistas indexadas Scopus¹, de esta manera se ubica en el puesto 218 de Iberoamérica con 495 publicaciones, según el ranking iberoamericano de instituciones de educación superior (SIR Iber)² en el año 2018.

Para que el VIITT pueda gestionar todas las actividades de investigación referentes a los procesos de movilidad, publicaciones indexadas y proyectos de

¹ SCOPUS es una base de datos bibliográfica de resúmenes y citas de artículos de revistas científicas, libros y conferencias.

² SIR IBER es un informe anual sobre el comportamiento científico de las instituciones de Educación Superior Iberoamericanas publicado por el Grupo de Investigación Scimago.

investigación necesitan indicadores de cumplimiento que estén enfocados en la acreditación del CACES usados mediante la Unidad de Gestión de la Investigación (UGI). Sin embargo, la UGI no puede evidenciar los indicadores debido a que no posee una herramienta automatizada para el control de las actividades de investigación de la UFA-ESPE.

Planteamiento del Problema

Para llevar el control de las actividades de investigación y que estas se encuentren alineadas con los objetivos de investigación de la UFA-ESPE, se debe realizar un análisis de los procesos y resultados de investigación como una oportunidad para mejorar. Basados en esta premisa, se ha identificado que la UGI presenta problemas de seguimiento y control de las actividades de investigación. Estos problemas tienen varias causas, la principal, aparentemente es que sus analistas manejan gran cantidad de información que se encuentra esparcida en varios subsistemas de investigación, los mismos que están en distintas fuentes de información (Ejemplo: Excel) y con diferentes formatos, lo que afecta a la obtención y consolidación de información para la toma de decisiones.

Otra causa identificada es la duplicidad de datos que existe en las diferentes fuentes de información. Prueba de ello, es el memorando N.-ESPE-DCCO-2019-0313-M perteneciente a la UGI, en donde (Fonseca, 2019) manifiesta que el subsistema de publicaciones es carente de interoperabilidad debido a que no accede a la información existente, es decir, existe duplicidad de datos, lo que proporciona como resultados información redundante.

Además, se puede identificar la baja producción de reportes científicos por parte de la UFA-ESPE, situación que es informada por la dirección de

postgrados, tal y como se refleja en el memorándum N.- ESPE-CPOS-2019-1206-M. Es por esa razón que esta situación afecta a obtención de las metas y políticas de investigación de la universidad.

Con respecto a los indicadores, estos son incipientes en referencia al seguimiento y control de actividades de investigación. Lo cual, dificulta el análisis de los resultados de investigación de los investigadores, como es la realización de proyectos o simplemente analizar sus actividades científicas. Afectando de esta manera a la evaluación y acreditación por parte del CACES, institución que en su modelo de evaluación institucional de universidades y escuelas politécnicas, pone en manifiesto en el criterio INVESTIGACIÓN, en lo relacionado a la producción científica “ indicadores, que miden el desarrollo de artículos, libros, capítulos de libros y ponencias de los miembros de las Instituciones de Educación Superior” (Sanchez, Coronel, Labanda, Chamaba, & Guamán, 2018).

Justificación

La investigación y producción de contenido científico es vital en la educación superior, porque es la mejor manera de aportar al estudiante contenidos de calidad que eleven el nivel académico (Vélez & Dávila, 2018). Ante esta realidad, el VIITT de la UFA-ESPE tendrá a su disposición bases de datos multidimensionales actualizadas como herramienta de soporte de las actividades de investigación más importantes como son: proyectos, publicaciones y movilidad que se realiza en la UGI. Esta información disminuirá los tiempos de espera en la generación de los reportes de los resultados de investigación.

Por otra parte, según (Méndez del Río, 2006) menciona: “La información es uno de los activos potencialmente más valiosos en una empresa” y la ausencia de información depurada o incompleta hacen que las organizaciones tomen decisiones a través de la intuición (Akinci & Smith, 2012). Por eso, a través de la propuesta del uso de procesos ETL se busca obtener un mejor manejo de los datos de las actividades de investigación, lo cual permite contar con información relevante de los procesos de la UGI. Datos que sirven para mejorar la toma de decisiones del VIITT de la UFA-ESPE.

Con respecto a los indicadores debemos tomar en cuenta que son procesos de control que identifican cambios en una organización (ONUMJERES, 2010). Por ello se vuelve necesario analizar los datos de la UGI y evaluarlos a través de indicadores cuantitativos que están relacionados con el tiempo y cantidad. Con estos indicadores se puede llevar a cabo el control de las actividades de investigación enfocados en la acreditación del CACES, y sobre todo buscar estrategias que permitan el cumplimiento de los objetivos de investigación de la UFA-ESPE. Para cumplir con estos objetivos, es necesario que los indicadores sean automatizados mediante el uso de herramientas de BI.

Estas herramientas permiten combinar los datos resultantes de las actividades de investigación para tener una perspectiva histórica y actual del estado de la UGI. La integración y contextualización de estos datos genera conocimiento que sirve para escoger las alternativas más convenientes para la mejora continua y el éxito de la UFA-ESPE.

Por último, tenemos la validación de los resultados obtenidos en los reportes de acuerdo con los indicadores de resultados que deben estar acordes a la información de cada proceso de la UGI para que la información sea fiable.

De acuerdo con las justificaciones escritas en los párrafos anteriores, es preciso poner en práctica los conocimientos adquiridos en la Carrera de Ingeniería e Informática, para contribuir a la solución del problema expuesto en el planteamiento del problema, mediante la propuesta de un diseño correcto de BI que sirva para el análisis de la información más relevante de los subsistemas de la UGI. Asimismo, la necesidad del VIITT de contar con información oportuna y adecuada para el análisis y evaluación de resultados, nos permite aplicar indicadores de control de resultados con el objetivo de promover una administración exitosa.

Objetivos

Objetivo General

Desarrollar un modelo de Business Intelligence enfocado a la toma de decisiones de la gestión de investigación del Vicerrectorado de Investigación, Innovación y Transferencia de Tecnología de la UFA-ESPE.

Objetivos Específicos

- i. Diseñar modelos multidimensionales de bases de datos (datamart) a fin de almacenar los datos de las actividades de investigación.
- ii. Diseñar indicadores de control de actividades de investigación enfocados en la acreditación del CACES.
- iii. Aplicar los procesos de extracción, transformación y carga (ETL) a fin de obtener los datos más relevantes de las actividades de investigación.
- iv. Validar con el usuario final los reportes generados mediante una herramienta de análisis de BI a fin de verificar el cumplimiento de los indicadores.

Alcance

Para definir el alcance de nuestra investigación se establecen las siguientes preguntas asociadas a los objetivos específicos, tal como se muestra en la *Tabla 1*.

Tabla 1.

PREGUNTAS DE INVESTIGACIÓN

Objetivos específicos	Pregunta de investigación
i. Diseñar modelos multidimensionales de base de datos (datamarts) a fin de almacenar los datos de las actividades de investigación.	RQ1: ¿Es viable diseñar datamarts para almacenar los datos de investigación e innovación tecnológica?
ii. Diseñar indicadores de control de actividades de investigación enfocados en la acreditación del CACES.	RQ2: ¿Es posible diseñar indicadores de control de actividades de investigación enfocados en la acreditación del CACES?
iii. Aplicar los procesos de extracción, transformación y carga (ETL) a fin de obtener los datos más relevantes de las actividades de investigación.	RQ3: ¿Es posible extraer los datos más relevantes de las actividades de investigación, transformarlos y cargarlos en un nuevo almacén de datos?
iv. Validar con el usuario final los reportes generados mediante una herramienta	RQ4: ¿Es posible validar los resultados de los indicadores de control a partir de un dashboard y el usuario final?

Objetivos específicos	Pregunta de investigación
de análisis de BI a fin de verificar el cumplimiento de los indicadores.	

Nota: Esta tabla muestra los objetivos específicos del presente trabajo de investigación y las preguntas que se van a investigar.

El alcance previsto en nuestra investigación cubre desde:

- Objetivo I: Levantamiento de requerimientos de usuarios y fuentes de información que intervienen en el proceso de registro y control de actividades de investigación a fin de diseñar modelos multidimensionales de bases de datos.
- Objetivo II: Análisis y diseño de aproximadamente cinco indicadores para el proceso de proyectos de investigación, cinco indicadores para el proceso de publicaciones de investigación, cinco indicadores para el proceso de movilidad y cinco indicadores para el proceso financiero, dando un total de veinte indicadores que serán implementados en la UGI, los cuales serán los más representativos para el control de actividades de investigación del VIITT y estarán enfocados en la acreditación del CACES.
- Objetivo III: Desarrollo e implementación de procesos de extracción, transformación y carga de datos de las actividades de investigación a fin de consolidar los datos históricos e incrementales en un repositorio centralizado.
- Objetivo IV: Diseño de Dashboards con el apoyo de herramientas de BI enfocados en los indicadores de resultados de actividades de investigación y validados a través de un estudio de caso práctico en el VIITT de la UFA-ESPE con los usuarios finales.

CAPÍTULO II

ESTADO DEL ARTE

Marco Teórico

El marco teórico que se desarrolla a continuación nos muestra los conceptos más relevantes y necesarios para sustentar el desarrollo del presente proyecto de graduación. Los principales conceptos utilizados se detallan a continuación.

Inteligencia de negocios

➤ **Definición:**

Conocida también como Business Intelligence o BI en inglés, se define como “Un conjunto de estrategias, acciones y herramientas enfocadas a la administración y creación de conocimientos mediante el análisis de datos existentes en una organización” (Tello & Perusquia, 2016). Estas estrategias permiten recabar información importante para la toma de decisiones acertadas y posteriores acciones que las empresas pueden emprender, para generar una ventaja competitiva que responda a los problemas del negocio (Azita, 2011).

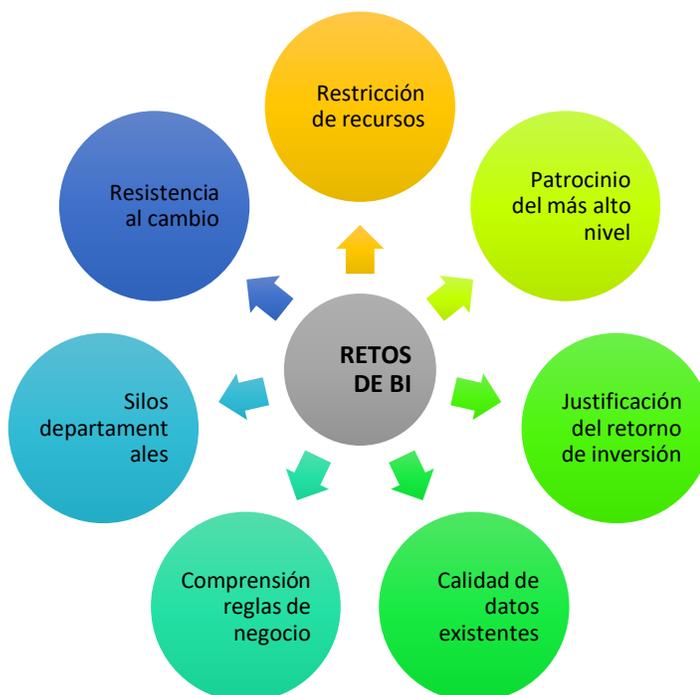
➤ **Importancia:**

Los sistemas de BI en la actualidad son de vital importancia debido a que permiten a las empresas tener el control total de su información. Esta información la empresa puede utilizar como base para llevar a cabo cambios fundamentales de acuerdo con la razón social de la organización, es decir, establecer nuevas estrategias, formar una nueva cooperación, adquirir nuevos clientes, crear nuevos mercados y ofrecer nuevos servicios o productos a sus clientes. (Chaudhary, 2014). Sin embargo, el uso de BI no es sencillo, se debe

enfrentar diferentes situaciones que son mostradas en la *Figura 1*. En esta figura se puede observar los problemas más comunes que se presentan al implementar la tecnología de BI; dentro de los cuales, tenemos a dos grandes limitantes como son: la “restricción de recursos” y la “resistencia al cambio”. Luego tenemos presente al “patrocinio de alto nivel” y la “justificación del retorno de inversión” en donde deben estar involucrados todo el personal de la empresa para que se pueda dar viabilidad al proyecto; y como último poseemos a la “comprensión de las reglas de negocio”, la “calidad de datos existentes” y los “silos departamentales” que se refieren a los datos y procesos internos propios de la organización que deben estar disponibles para su análisis (Solano, 2016).

Figura 1.

RETOS EN LA IMPLEMENTACIÓN DE BI

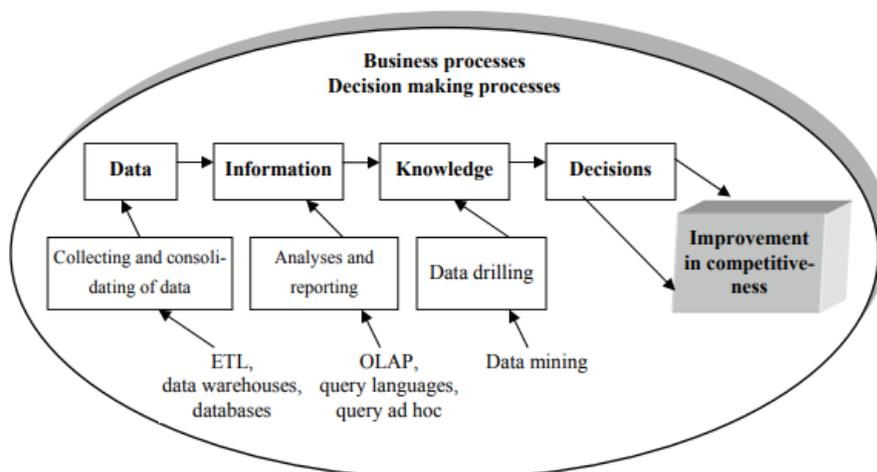


Nota: El gráfico representa los retos que enfrenta el BI en la actualidad. (Solano, 2016)

Luego de los retos mencionados en la implementación de BI dentro de la organización, se debe iniciar con el proceso de BI como se muestra en la *Figura 2*. En esta figura se incluye 4 pasos fundamentales: el primer paso es la “recolección inteligente de los datos” de las diversas fuentes de datos, el segundo paso es “integrar y analizar la información”, el tercer paso es “obtener el conocimiento de los datos” y el paso final del proceso de BI es “tomar decisiones que mejore la competitividad empresarial” (Olszak, 2007).

Figura 2.

PROCESO DE BUSINESS INTELLIGENCE



Nota: El gráfico representa el proceso de BI en la toma de decisiones. (Olszak, 2007).

Metodologías de uso para BI:

En la actualidad existen varias metodologías (ejemplo: Data-Driven Approach, Value-Chain Data Approach, Process-Driven Approach, Event-Driven Approach, etc.) Esta diversidad depende a que cada fabricante de software de BI busque imponer su producto en el mercado. A pesar de ello, a estas metodologías se las puede clasificar en dos grandes grupos: las primeras

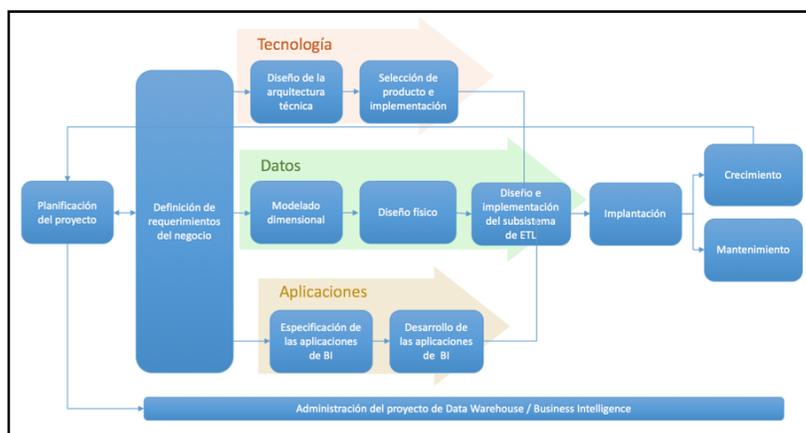
metodologías denominadas “Bottom-up” o en forma ascendente representado por Ralph Kimball y las segundas metodologías llamadas “Top-down” o descendente en la cual su principal actor es Bill Inmon. Para explicar de mejor forma estas metodologías se las detalla a continuación.

➤ **La Metodología Kimball (Bottom-up)**

Kimball propone una metodología que consta de un modelo dimensional conocido como ciclo de vida dimensional del negocio, la cual se puede visualizar en la *Figura 3*. Esta figura tiene el objetivo de indicar las tareas que se realizan al momento de definir los requerimientos del negocio que están enfocados en 3 áreas diferentes (tecnología, datos y aplicaciones), caminos que se combinan cuando se implementa el proceso de ETL³. Además, tiene una relación directa con la planificación del proyecto y una relación indirecta con la administración del proyecto de BI.

Figura 3.

CICLO DE VIDA DIMENSIONAL DEL NEGOCIO



Nota: El gráfico representa el ciclo de vida de la metodología de Kimball. (Calvo, 2017).

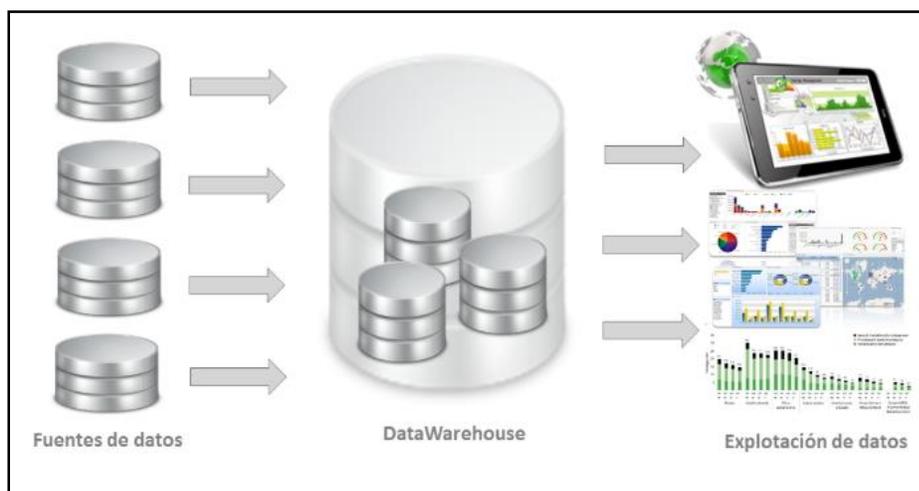
³ ETL, es un proceso de extracción de datos a partir de un número ilimitado de fuentes, transformación de formatos de los datos y carga de datos un único repositorio centralizado para su análisis.

Esta metodología empleada para la construcción de un almacén de datos. Se basa en cuatro principios básicos que son: centrarse en el negocio, construir una infraestructura de información adecuada, realizar entregas en incrementos significativos y como último principio es ofrecer una solución completa en documentación, capacitación y soporte (Rivadeneira, 2010).

En base a los principios de Ralph Kimball, se puede identificar que el esquema de arquitectura de esta metodología se encuentra conformada por datamarts organizados en modelos dimensionales de datos que conforman el DataWarehouse. Estos modelos pueden ser esquemas en forma de estrella o copo de nieve diseñados, de tal forma que faciliten la consulta y generación de reportes al momento de explotar los datos, tal y como se observa en la *Figura 4*, (Dertiano, 2015).

Figura 4.

ARQUITECTURA DE KIMBALL



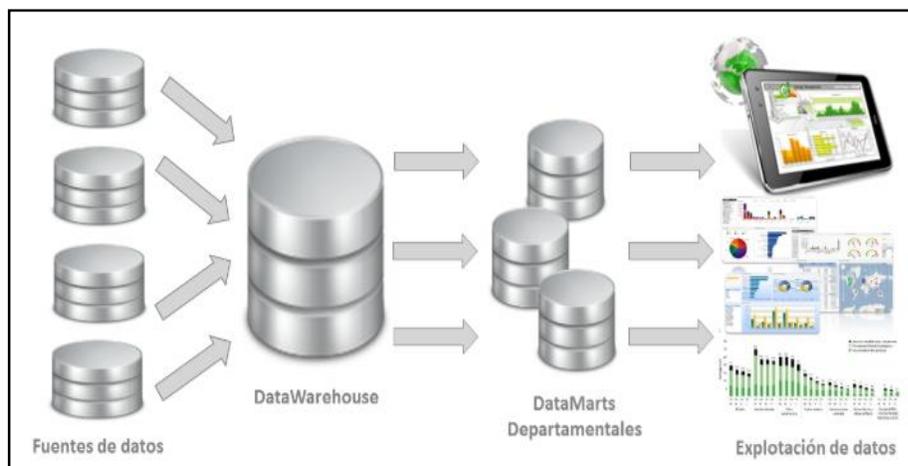
Nota: El gráfico representa como está formado la arquitectura de Kimball. (Dertiano, 2015)

➤ **La Metodología de Inmon (Top-down)**

Es una metodología utilizada para almacenar grandes volúmenes de datos, a esta metodología se llama INMON, debido a que es un modelo centralizado donde integra a todos los datos de una organización con el fin de normalizar y evitar la redundancia de datos. Su arquitectura posee varios niveles donde existen varios datamarts temáticos dependientes del DataWarehouse con datos específicos de un departamento de la empresa y son los puntos de acceso para las herramientas de reportería (Inmon, 2002), tal como se observa en la *Figura 5*, (Dertiano, 2015).

Figura 5.

ARQUITECTURA DE INMON



Nota: El gráfico representa la arquitectura de Inmon. (Dertiano, 2015)

Herramientas de software de BI

Las herramientas de software de BI ayudan a los ejecutivos en la automatización del análisis de datos empresariales con el fin de proporcionar información de calidad que permita mejorar la toma de decisiones gerenciales en la organización. (Antonio, Ordoñez, & Ortega, 2014). Estas herramientas están

enfocadas en la administración y creación de conocimiento mediante el análisis de datos existentes en una organización, por tal motivo en la actualidad las empresas de vanguardia utilizan herramientas de BI como apoyo en la toma de decisiones tanto en instituciones públicas como privadas. También, esta tecnología permite visualizar mediante gráficos estadísticos las acciones históricas, el comportamiento y el estado real de una organización, obteniendo ventaja de su información. (Medina, 2014).

En la *Tabla 2*, se muestran algunas herramientas que se incluyen en las soluciones de BI y que permiten la extracción, integración, representación y análisis de datos. A estas herramientas las hemos clasificado por tipo, función y ejemplos representativos de cada uno.

Tabla 2.

HERRAMIENTAS DE SOFTWARE DE BI.

TIPO	FUNCIÓN	EJEMPLOS
Bases de Datos	Herramienta tecnológica que permite almacenar de forma organizada datos de un sistema de información para su posterior utilización (Raffino, 2019).	Oracle MySql
Integración de datos	Software que utiliza procesos ETL para extraer, transformar y cargar los datos de las diferentes fuentes de información a un nuevo DataWarehouse. (Espinosa, 2010)	Clover ETL Pentaho Data Integration Oracle Data Integrator

TIPO	FUNCIÓN	EJEMPLOS
Modelado de datos	Software que permite construir el modelo de Inteligencia Empresarial, que es la abstracción del modelo físico.	Power Designer PowerArchitect
Reportería	Software para generación de reportes e informes utilizando los indicadores y dimensiones de acuerdo con la organización.	AgataReport Pentaho Reporting iReport Tableau
Cubos de información	Herramienta tecnológica que ayuda en el procesamiento analítico en línea y para el análisis multidimensional, lo cual nos permite tener una visión más rápida de los datos. (ECURed, 2019)	Pentaho Schema Workbeanch
Cuadros de Mando	Permite visualizar de forma rápida los indicadores más importantes de una organización (QAEC, 2019)	Balance Score Card Microsoft Power BI
Minería de Datos	Análisis de la información de interés para la predicción de tendencias, comportamiento e identificación de patrones ocultos.	Rapid Miner SAS Analytics

Nota: Esta tabla muestra el tipo, función y ejemplos de las herramientas de BI.

La herramienta de software de BI más utilizada en la actualidad y que reúne todos los requisitos necesarios para la solución de problemas de

inteligencia empresarial es Tableau⁴. De acuerdo con el cuadro mágico de Gartner⁵ se puede identificar a esta herramienta de BI líder por sexto año consecutivo. Gracias a sus prestaciones es considerado como el modelo de excelencia para análisis visuales, intuitivos e interactivos que ayudan a la toma de decisiones empresariales (Ajenstat, 2018). Tal como se observa en la *Figura 6*.

Figura 6.

CUADRO MÁGICO DE GARTNER



Nota: El gráfico representa la posición de la herramienta Tableau dentro del cuadro mágico de Gartner. (Howson, Richardson, & Kronz, 2019)

⁴ Tableau, es una plataforma de análisis integral más eficaz, segura y flexible de datos.

⁵ Cuadro Mágico de Gartner, es una consultora de investigación de mercado que realiza una representación gráfica de la situación del mercado de un producto en un momento determinado.

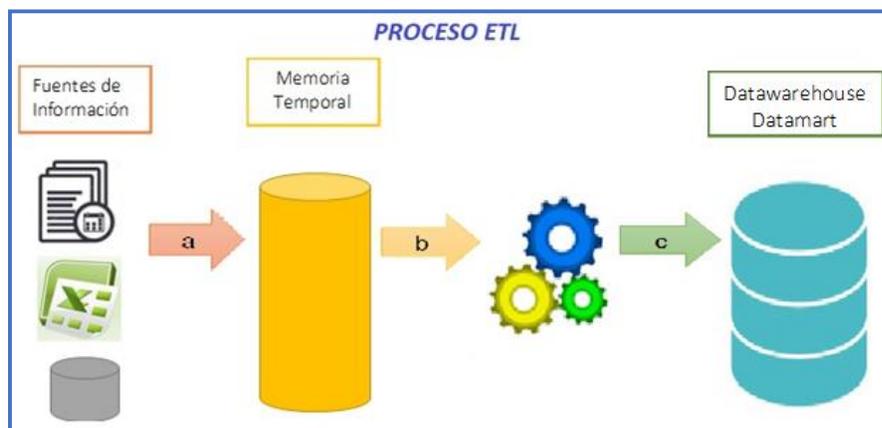
Es por ello, que Tableau es una plataforma que provee un análisis integral seguro y flexible que se adapta a todo tipo de empresa y transforma los datos en información útil, mediante representaciones gráficas a partir de la información limpia que se obtiene de los procesos ETL (ClarKAt, 2019).

Procesos ETL.

El proceso ETL que en el idioma inglés significa “Extract, Transform and Load” permite extraer datos desde múltiples fuentes heterogéneas para luego limpiarlos, normalizarlos y cargarlos en un nuevo almacén de datos (Trujillo & Lujan, 2008). Estos almacenes pueden ser un DataMart o un DataWarehouse que son contenedores en donde se va a explotar la información mediante herramientas de reportería para obtener una ventaja competitiva. Estos nuevos datos depurados que resultan luego del proceso ETL, se los puede cargar en otros sistemas de la organización para colaborar con otros procesos de negocio de la empresa (Trujillo & Lujan, 2008). Los procesos ETL: extraer, transformar y cargar se puede observar en la *Figura 7* y se explica a continuación.

Figura 7.

PROCESO ETL



Nota: El gráfico representa los procesos ETL que se van a realizar en la investigación.

a) Extraer

Es la primera parte del proceso ETL que consiste en extraer los datos de los diferentes sistemas de origen que pueden ser: bases de datos relacionales, bases de datos no relacionales, ficheros planos y otras estructuras diferentes, entre otros. Los datos pueden encontrarse en diferentes formatos (Ejemplo: Excel, archivos planos) por lo que la extracción convierte a los datos en un nuevo formato preparado para iniciar el proceso de transformación (Valero, Bigeek. , 2016).

b) Transformar

La fase de transformación se aplica una serie de cálculos necesarios, reglas de negocio y funciones sobre los datos extraídos para convertirlos en datos que cumplan con estándares de calidad para luego ser cargados en un almacén de datos. En este proceso de transformar datos, se logra mediante la unión de algunas fuentes de datos donde puede ser necesario una pequeña manipulación de su estructura, lo que se deriva en nuevos datos que son necesarios para que los usuarios finales puedan analizarlos, explotarlos y tomar las decisiones estratégicas de acuerdo con los objetivos de la institución. (PowerData, 2013). Debemos tener en cuenta que estas transformaciones pueden ser complejas y riesgosas por lo que deben tomarse todo el tiempo posible para realizarlos de la mejor manera minimizando totalmente los errores.

c) Cargar

Esta parte del proceso inicia luego de la fase de transformación de los datos, cuando estos se encuentran depurados y listos para ser cargados en el sistema de destino (PowerData, 2013). En esta parte del proceso los atributos que se obtuvieron de las fuentes de datos pueden ser mapeados con las

correspondientes fuentes de destino, de esta forma se pueden volcar los datos limpios al nuevo almacén de datos que tiene nuevas características de acuerdo con los objetivos empresariales. Para lo cual, se debe asegurar la consistencia de los datos y se debe priorizar la calidad de carga ante cualquier eventualidad, considerando parámetros de seguridad para evitar errores durante este proceso (Clavero & Carlos, 2016).

Bases de Datos

Una base de datos es un “almacén” que nos permite guardar grandes cantidades de información de un mismo contexto y en forma organizada (Pérez, 2007). Surgieron por la década de 1970 como bases de datos relacionales y están orientados a guardar los datos de una organización para luego utilizarlas de una forma rápida y ágil mediante sistemas informáticos (Leonard & Castro, 2013). Las bases de datos relacionales cumplen con un modelo relacional. Esto quiere decir que poseen relaciones predefinidas entre ellos y se organizan en un conjunto de tablas con columnas y filas para guardar la información. Además, utilizan un lenguaje SQL para consulta y mantenimiento de los datos (Amazon, 2019).

Asimismo, a finales del siglo pasado aparecen las bases de datos no relacionales o NoSql que no requieren estructuras fijas como tablas. Estas bases de datos tienen esquemas más flexibles, escalables, de alto rendimiento y altamente funcionales con una variedad de modelos de datos que se adapta a las aplicaciones modernas como juegos, redes sociales y la web. (Amazon, 2019).

➤ **Data Waterhouse**

La definición más conocida fue propuesta por *Inmon*, quien manifiesta que "Un Data Waterhouse es una colección de datos orientados a temas, integrados, no-volátiles y variante en el tiempo, organizados para soportar necesidades empresariales" (Inmon W. H., 1992). Esta colección de datos reúne toda la información de interés tanto interna como externa de una organización; lo que proporciona una visión global de la misma a los usuarios. Los datos de esta colección pueden colocarse dentro de un intervalo de tiempo, almacenando datos históricos que permiten pronosticar tendencias y patrones de datos. Estas tendencias pueden ser analizadas por diferentes perspectivas y a grandes velocidades de respuesta, gracias a herramientas tecnológicas de reportería que benefician a las empresas a ser más competitivas, eficientes y sobre todo a los usuarios finales para que puedan mejorar la toma de decisiones.

Se debe tener en cuenta que los elementos de un DataWaterhouse son: los "hechos" que representan los procesos del negocio, las "dimensiones" que simbolizan las distintas vistas del proceso del negocio y las métricas que miden el proceso del negocio mediante indicadores.

➤ **Datamarts**

El datamart es un subconjunto dependiente o independiente de un DataWarehouse. Posee la misma funcionalidad y complejidad, pero de tamaño menor con una población de datos específicos. (Caraguay, 2018). Está orientado al análisis, almacenamiento e integración de los diferentes tipos de datos de un área de la organización. Para modelar un datamart se debe tomar en cuenta ¿Cuál de los esquemas (estrella o copo de nieve) más se apega a nuestras

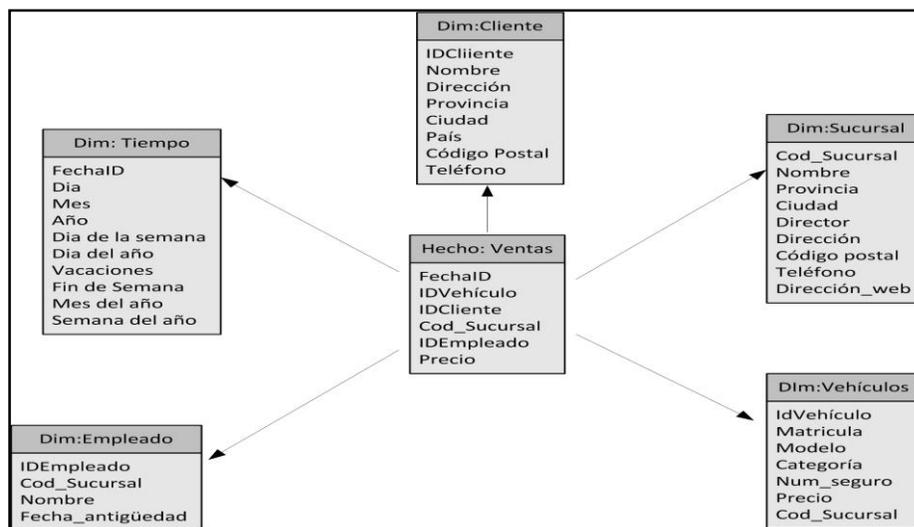
necesidades para obtener los resultados esperados? (Salvador, 2017). A continuación, se detallan los diferentes tipos de datamart.

- **Datamart con Esquema en Estrella:**

También llamada como “star schema” en inglés, es un modelo de datos relacional simple que tiene una tabla central de hechos que contiene datos cuantitativos medibles que sirven para el análisis del negocio. Esta tabla “hechos” está rodeada de tablas “dimensiones” que son los atributos de los hechos. En la *Figura 8* podemos observar este esquema.

Figura 8.

ESQUEMA EN FORMA DE ESTRELLA.



Nota: El gráfico representa la forma de cómo está compuesto un esquema de datamart en forma de estrella. (Rochina, 2017).

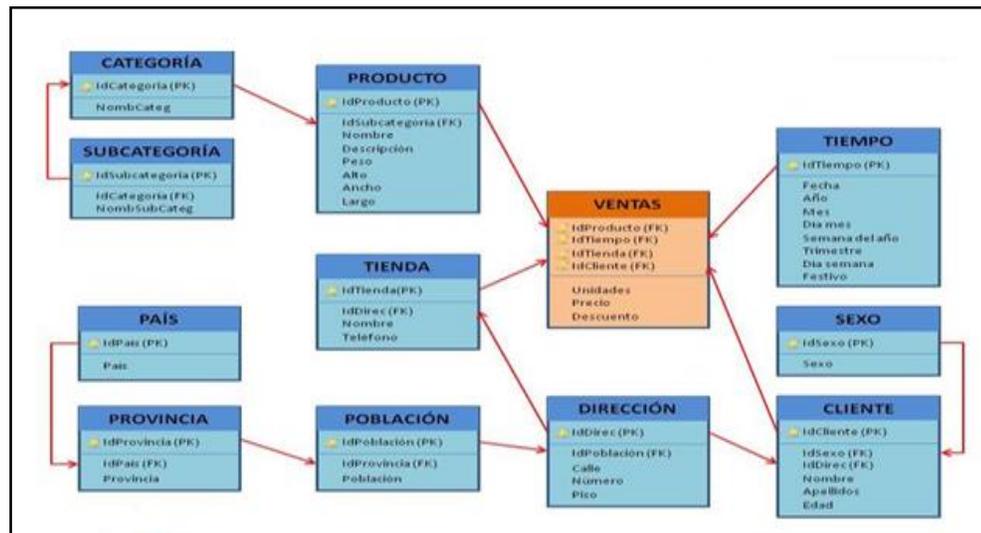
- **Datamart con Esquema Copo de Nieve.**

Este esquema conocido también como *snowflake schema* en inglés, es un modelo que posee una estructura más compleja que el esquema anterior,

debido a que algunas “dimensiones” no están relacionadas directamente con la tabla de “hechos”, sino con otras “dimensiones” (Valero, Bigeek Modelo Dimensional, 2017). En la *Figura 9*, se puede identificar este esquema.

Figura 9.

ESQUEMA COPO DE NIEVE.



Nota: El gráfico representa la forma de cómo está compuesto un esquema de datamart en forma de copo de nieve (WIKIDOT, 2015)

Según (Moody, 2010), los datamarts suelen adaptarse a las necesidades de un grupo de usuarios, debido a que colocan la información importante de un área funcional de la organización en manos de quien realiza la toma de decisiones. También son muy útiles para trabajar con cubos de información como las OLAP (Online Analytical Processing).

Cubos de Información

Los cubos de información son estructuras de datos organizados y jerarquizados dentro de una estructura multidimensional de acuerdo con los factores de negocio de la empresa u organización (ECURed, 2019). Dentro de las principales características tenemos: la facilidad de uso y la rapidez de respuesta. Los cubos de información se clasifican de la siguiente manera:

a) OLAP

En inglés (On-Line Analytic Processing) que significa procesamiento analítico en línea, es un método ágil, flexible y rápido para organizar grandes cantidades de datos de acuerdo con su jerarquía u organización multidimensional, donde su objetivo es almacenar la información en una base de datos que permita realizar consultas históricas, complejas, analíticas e iterativas de forma más efectiva para generar reportes e informes que soportan decisiones estratégicas (Conesa, 2010).

b) OLTP

En inglés (On-Line Transaction Processing) que significa procesamiento de transacciones en línea, es un gestor transaccional que facilita y administra aplicaciones de bases de datos donde se puede insertar, actualizar y eliminar información. En las OLTP se usan instrucciones SQL para generar un proceso atómico, por ejemplo, se puede usar un commit para validar las modificaciones realizadas o un rollback para que sea invalidado y guardado como permanentes los cambios ejecutados en las bases de datos operacionales (Sinnexus, 2018).

c) MOLAP

En inglés (Multidimensional On-Line Analytic Processing) que significa procesamiento analítico en línea multidimensional, este cubo almacena los datos

en las diferentes dimensiones existentes para potenciar su visualización y análisis (Conesa, 2010). El sistema MOLAP utiliza una arquitectura de dos niveles: el primer nivel es la base de datos multidimensional encargada del manejo, acceso y obtención de los datos y el segundo nivel es el motor analítico que realiza las operaciones en el menor tiempo de respuesta optimizando las consultas para que los usuarios finales visualicen los análisis OLAP.

d) ROLAP

En inglés (Relational OLAP) que significa procesamiento analítico en línea relacional, trabaja directamente sobre las bases de datos relacionales, este cubo posee una estructura de 3 niveles: el primer nivel es una base de datos relacional donde se almacenan los datos multidimensionales en tablas organizadas, el segundo nivel es el motor OLAP donde se realiza las consultas y el análisis de los datos y el tercer nivel es una herramienta especializada para la representación de la información (Sinnexus, 2018).

Los cubos ROLAP son escalables y se utilizan para tener acceso a grandes conjuntos de datos, que se consulta con poca frecuencia, por ejemplo, datos históricos (Laiseca, 2019).

e) HOLAP

En inglés (Hybrid On Line Analytic Processing) que significa procesamiento analítico en línea híbrido. Consiste en una combinación de las arquitecturas de los cubos MOLAP y ROLAP para brindar una mejor solución en el procesamiento de datos. Los datos agregados y pre-calculados se almacenan en estructuras multidimensionales y los datos de menor nivel de detalle se almacenan en estructuras relacionales (Bernabeu, 2009). “El almacenamiento

HOLAP suele ser adecuado para cubos que requieren tiempos cortos de respuesta” (Laiseca, 2019).

Indicadores (KPI)

En inglés, “Key Performance Indicator” es conocido también como indicadores claves de resultados. Estos son valores tangibles que permiten medir y cuantificar resultados de los procesos organizacionales para verificar si una organización está en el camino correcto en la consecución de los objetivos empresariales (CulturaCrm, 2016). Esta variable dimensional se correlaciona con otras variables presentes en los procesos de la organización y se caracteriza por informar al usuario el estado actual del sistema evaluado en un tiempo determinado (COLPATRIA, s.f.). Con los KPI una empresa se permite tomar decisiones preventivas o correctivas al comparar el valor obtenido del sistema con el valor esperado. Existen dos tipos de indicadores:

➤ **Indicadores Cualitativos**

Los indicadores cualitativos son subjetivos y están relacionados con el nivel de calidad, donde la concepción está proporcionado al tipo de servicio o producto que se ofrece a los clientes. La subjetividad del indicador cualitativo permite tener una mayor riqueza informativa (DeGregorio, 2008). Su valoración se toma en cuenta con el grado de eficiencia y nivel de productividad (Oliveras, 2018).

➤ **Indicadores Cuantitativos**

Se refiere a los indicadores que poseen una fórmula de cálculo y funciones matemáticas. Su representación numérica está relacionada con el tiempo que se utiliza para llevar a cabo una tarea y la cantidad representativa del trabajo realizado. Además, este tipo de indicador está orientado a obtener un

mayor grado de objetividad en la información obtenida (DeGregorio, 2008).

Existen algunos tipos de indicadores cuantitativos como:

a) Indicadores de Gestión

Los indicadores de gestión se utilizan para realizar el monitoreo, seguimiento y evaluación del rendimiento de los procesos y actividades de la gestión de una organización, además permiten evaluar el logro de los objetivos empresariales propuestos (Colmayor, 2017).

b) Indicadores Financieros

Estos indicadores son utilizados para relacionar estados financieros y compararlos con estados financieros de años anteriores o con los estados de otras organizaciones del mismo sector, para evaluar el grado de eficiencia del consumo de los recursos. Asimismo, estos indicadores nos ayudan a obtener conclusiones sobre la situación financiera real de la empresa.

(Fontalvo, De la Hoz, & Vergara, 2012)

c) Indicadores de Acreditación

Son indicadores formulados para la evaluación del desempeño de las instituciones de educación superior para verificar si cumple o no, con las características y estándares de calidad académicas de las universidades. El CACES, establece indicadores que miden el desarrollo de artículos, libros, capítulos de libros y ponencias de los miembros de las Instituciones de Educación Superior en el Ecuador.

Dashboard

“Los dashboard son mecanismos de representación visual utilizados en un sistema de medición operativa de resultados, que mide el desempeño contra objetivos y umbrales usando datos de tiempo adecuado” (Kerzner, 2013). Es por

ello que el dashboard permiten diagnosticar de una forma visual el estado actual de una empresa, agilizando en la toma de decisiones gerenciales a través del seguimiento y evaluación periódica mediante el uso de indicadores de resultados. Existen 3 tipos de dashboard (operativo, táctico y estratégico), tal y como se describe a continuación:

➤ **Dashboard Operativo:**

Los dashboard operativos enfatizan el monitoreo más que el análisis y la administración. (Kerzner, 2013).

➤ **Dashboard Táctico:**

Un dashboard táctico rastrea procesos y proyectos departamentales que son de interés para un segmento de la organización. (Kerzner, 2013).

➤ **Dashboard Estratégico:**

Los dashboard estratégicos alinean la organización en torno a los objetivos estratégicos y lidera para que todos los departamentos o áreas de una empresa avancen en la misma dirección (Martínez, 2017).

Revisión Sistemática de Literatura Ligera

En esta parte se describen varios artículos de investigaciones que se encuentran relacionados con la problemática identificada dentro del VIITT de la UFA-ESPE. La Revisión Sistemática de Literatura (SLR de las siglas del inglés Systematic Literature Review) es una estrategia metodológica que permite obtener de manera rigurosa estudios fiables, para enfocar la presente investigación en un contexto adecuado. (Zambrano, Reyes, Castro, & Fonseca, 2018). Para tal efecto, nos hemos planteado la siguiente pregunta de investigación (RQ de las siglas del inglés research question).

RQ: ¿Es posible que mediante el diseño de datamarts enfocados en BI ayuden a la toma de decisiones de la gestión de investigación del VIITT, y permita el control de las actividades de investigación en la UFA-ESPE?

Para dar respuesta a la RQ planteada, como se mencionó anteriormente, se realizó una SRL basado en las guías propuestas por (Kitchenham & Charters, 2007). Este protocolo reestructurado de revisión de literatura propone la definición de: (1) Planteamiento del objetivo de búsqueda, (2) Conformación del grupo de control (GC), (3) Construcción de la cadena de búsqueda, (4) Selección de estudios primarios.

1. Planteamiento del objetivo de búsqueda

El objetivo de la búsqueda es encontrar varios estudios relacionados con: Datamarts, procesos ETL, actividades de investigación, gestión de control e indicadores de calidad de la educación superior para obtener una vista panorámica de problemas, causas y soluciones afines a nuestra investigación para estructurar nuestra búsqueda como alcanzable se definieron los siguientes criterios de inclusión y exclusión:

Criterios de inclusión:

- Artículos relacionados con la gestión de control, procesos ETL e indicadores de calidad de la educación superior.
- Investigaciones publicadas en revistas, congresos o en libros que sean indexados por Scopus, IEEEExplore o ACM Digital Library.
- Artículos en inglés y español.

Criterios de exclusión.

- Artículos publicados después del año 2010.
- Artículos cortos, tutoriales y resúmenes de presentaciones o conferencias.

- Artículos no accesibles en texto completo.
- Artículos en otros idiomas.

2. Conformación del grupo de control (GC)

Según Zhang & otros, (GC) el grupo de control es un conjunto de investigaciones relevantes que permiten establecer estudios similares bajo criterios explícitos establecidos en la RQ (Zhang, Alibabar, & Tell, 2010). Estos estudios son la fuente de información de donde se obtienen los términos o palabras claves relacionadas con el contexto de la temática investigada que servirán como base para la conformación de la cadena de búsqueda.

Para conformar el GC es necesario los siguientes pasos:

- Identificación de los investigadores.
- Búsqueda de artículos candidatos.
- Validación.
- Conformación del grupo de control.

Para nuestra investigación se contó con la participación de dos investigadores, los cuales, propusieron estudios relevantes de bases digitales académicas prestigiosas como Scopus y IEEEExplore, como candidatos para conformar el GC. Luego se realizó una validación cruzada entre investigadores con el propósito de verificar que las investigaciones tengan una relación directa con la RQ propuesta anteriormente, para finalmente seleccionar 5 artículos que conformen el GC.

En la *Tabla 3*, se muestran los estudios que integran el grupo de control, así como, las palabras clave identificadas en cada estudio y el número de veces que se repite esta palabra.

Tabla 3.*GRUPO DE CONTROL DEFINIDO (GC).*

N.	TÍTULO	PALABRAS CLAVE						Total	Porcentaje
		ETL (BI)	Datamart	Dashboard (Cuadros de mando)	Indicators (Indicador)	Toma de decisiones			
1	Computer system for the management and scientific divulgation of the “Universidad Nacional de Loja” Design and realization of an ETL method in	2	12	2	2	0	18	6,4%	
2	business intelligence project. Data Migration from a Product	47	14	2	0	7	70	24,9%	
3	to a Data Warehouse	13	22	0	0	0	35	12,5%	

N.	PALABRAS CLAVE							Total	Porcentaje
	TÍTULO	ETL (BI)	Datamart	Dashboard (Cuadros de mando)	Indicators (Indicador)	Toma de decisiones			
	Using ETL Tool. Business intelligence solutions for processing huge data to the business user's using dashboards.	10	5	11	0	5	31	11.0%	
4	Significance of data integration and ETL in business intelligence framework for higher education.	84	25	0	0	18	127	45,2%	
5									
	Total	156	78	15	2	30	281		
	PORCENTAJE	55.5%	27.8%	5.3%	0.7%	10.7%		100%	

Nota: Esta tabla muestra los estudios del grupo de control y los porcentajes de las palabras claves que se repiten.

3. Construcción de la cadena de búsqueda

La construcción de la cadena de búsqueda inicia con el análisis detallado de los artículos de la *Tabla 3*. De esta tabla, se obtienen las palabras claves que son las que aparecen con mayor frecuencia y están relacionados con el tema de estudio. Por ello que para conformar esta cadena de búsqueda se combinaron estas palabras claves a través de operadores lógicos: OR para sinónimos y AND para combinar palabras de otros contextos.

En primera instancia se conformó la cadena de búsqueda inicial que corresponde a la N.º 1, de la *Tabla 4*, en donde se muestran las cadenas de búsqueda realizadas que se probaron en la base IEEEExplore. De esta cadena se obtuvieron 450 estudios de las cuales, solamente 2 estudios pertenecían al GC; por lo cual fue necesario realizar 3 iteraciones adicionales con diferentes combinaciones de palabras claves, hasta que se encontró la cadena de búsqueda ideal que es la iteración N.º 4. Se denomina búsqueda ideal porque retorna un mayor número de estudios que se encuentran dentro del GC.

Tabla 4.

CADENA DE BÚSQUEDA

Iteración	Cadena de búsqueda	Resultados	Grupo Control
N.º 1	ALL: (((ETL) OR (Datamart) OR (dashboard)) AND ((indicators) OR (toma de decisiones)))	450	2
N.º 2	ALL: (((Datamart) AND etl) OR panel de control) AND toma de decisiones)	385	1

Iteración	Cadena de búsqueda	Resultados	Grupo Control
N.º 3	ALL: (((((Datamart) OR etl) AND dashboard) OR toma de decisiones) OR indicators)	250	2
N.º 4	ALL: (((((Datamart) AND etl) AND BI) OR indicators)	26	3

Nota: Esta tabla muestra la cadena de búsqueda con las palabras claves y los resultados de estudios encontrados en las bases científicas.

4. Proceso de Selección de Estudios Primarios y Extracción de Datos

A partir de la conformación de la cadena de búsqueda N.º 4 se obtuvieron 26 estudios en la base digital IEEEExplore, entre los cuales aparecen 3 estudios que conforman el GC. Luego para estudiar la factibilidad del objetivo principal, en la búsqueda fueron aplicados filtros de inclusión y exclusión antes establecidos en el primer punto de esta sección.

Como resultado de esta actividad se seleccionaron 15 estudios los cuales fueron revisados y finalmente se escogieron 8 investigaciones, que fueron denominados estudios primarios (EP).

5. Resultados de la Investigación Sistemática de Literatura Ligera.

- **EP1** (Muñoz, Mazon, & Trujillo, 2011) **ETL Process Modeling Conceptual for Data Warehouses: A Systematic Mapping Study**

La presente investigación presenta un estudio de mapeo sistemático de literatura sobre los procesos ETL. Dichos procesos son los factores clave en el éxito de los proyectos de almacén de datos a pesar de que son complejos,

propenso a errores y sobre todo costosos. Los autores sugieren, que es necesario que se tome un tiempo adecuado para dichos procesos y afirman que existe una visión clara del modelado de procesos ETL, pero que no cubren lo suficiente en estos enfoques por lo que es necesario más esfuerzo para garantizar mejores resultados.

- **EP2** (Kabiri & Chiadmi, A method for modelling and organizing ETL processes, 2013): **A method for modelling and organizing ETL processes.**

Este documento presenta un estudio motivado por la escasa existencia de métodos de modelos ETL, y tiene el objetivo de presentar un método de modelaje y organización de procesos ETL para las tareas de la extracción de datos, limpieza, conformidad y carga. Para lo cual, los autores proponen un modelo conceptual de procesos ETL ayudados de un framework para la gestión de procesos ETL denominado KANTARA. Los autores obtienen como resultado una modularización de los procesos ETL y llegan a la conclusión que este modelo es lo más apegado a la experiencia del mundo real reduciendo su complejidad y costo. Dándonos una pauta de cómo organizar la información de las actividades de investigación de la UFA-ESPE.

- **EP3** (Suneetha & Krishna, 2017): **Business intelligence solutions for processing huge data to the business user's using dashboards**

En este trabajo de investigación, los autores son motivados por los problemas que se presentan en el sistema de gestión de las bases de datos relacionales (RDBMS) Por esta razón, se plantearon el objetivo de resolver estos problemas. Para lo cual, los autores proponen la aplicación de BI para el proceso de extracción, transformación y carga de los datos iniciales, proporcionando una herramienta en la cual el usuario final pueda monitorear los

resultados en un dashboard. Los autores llegan a la conclusión que los clientes pueden visualizar fácilmente su negocio a través de representaciones estadísticas y gráficas de los indicadores de la empresa. Por lo tanto, este estudio brinda un primer paso para fundamentar la gestión de control a través de herramientas de BI, las cuales pueden ser aplicadas en las actividades de los investigadores de la UFA-ESPE.

- **Ep4** (Sevilla, Reinoso, Ordoñez, & Ortega, 2014) **SOLUCIÓN BUSINESS INTELLIGENCE (BI) CORPORATIVA: Transformando Datos en Conocimiento.**

En esta investigación los autores plantean una solución para el aumento de datos dentro de la gestión de la información de una organización basado en la inteligencia de negocios. Los autores reúnen estrategias y herramientas enfocadas a la administración y creación de conocimiento mediante el análisis de datos existentes en la empresa. Asimismo, utilizan la metodología de Kimball para el desarrollo de una solución de BI para el servicio público de empleo en Madrid. La idea del estudio es relacionar a las personas con un perfil profesional. Los autores usan un modelo dimensional en forma de estrella, donde la tabla de “hechos” va ubicada en el centro y las tablas de “dimensiones” alrededor de esta. Las tablas de “hechos” tienen columnas de datos llamadas métricas y las de “dimensiones” tienen columnas con atributos y niveles de jerarquías. Las métricas son indicadores que permiten cuantificar los hechos. Toda esta información es muy importante dentro de nuestra investigación porque se debe evaluar las actividades de investigación del VIITT a través de indicadores de calidad.

- **EP5 (Keerin, 2016): Development of business intelligence solution for personnel administration.**

El autor plantea su artículo motivado en el tema de la administración del personal de una Institución. El objetivo es buscar el éxito de una organización evaluando las habilidades, competencias y méritos relativos a cada empleado. Para lo cual, propone una solución de BI para la planificación del personal con un estudio de caso en una universidad tailandesa de Rajabhat. Se aplica un módulo de ETL para convertir los datos transaccionales en un DataWarehouse que sirve para presentar los datos resumidos. Este experimento da como resultado datos de manera más rápida y oportuna para la toma de decisiones de los ejecutivos. La conclusión que proporciona el autor se refiere a que la eficiencia de una organización depende de la exactitud de la información obtenida.

Este estudio nos sirve para nuestro tema de investigación, porque se busca apoyar a la toma de decisiones del VIITT mediante la extracción de información resumida y precisa de todos los datos existentes en los subsistemas de la UGI.

- **EP6 (Hamizah, Shahizan, & Mi, 2015). Significance of data integration and ETL in business intelligence framework for higher education.**

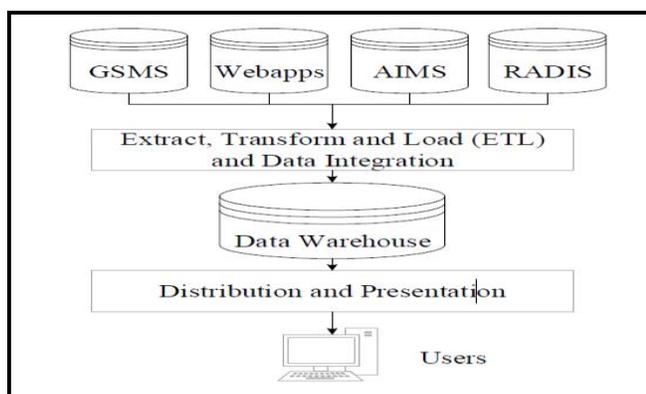
Los autores manifiestan en este artículo que la educación superior requiere una gran cantidad de decisiones, pronósticos y análisis. Por lo tanto, un componente básico e importante para el BI en una institución de educación superior es la integración de datos antes de su almacenaje. Para lo cual, utiliza como caso de estudio a los estudiantes y docentes de la Universidad Tecnológica de Malasia (UTM) para comprender mejor la relación de los datos

en el proceso ETL. Se identifican dos formas de integración de datos: la primera en forma manual y la segunda mediante la integración ayudada por herramientas tecnológicas, en donde la integración de datos es capaz de eliminar errores, corregir datos faltantes y sobre todo es crucial para copiar solo datos de interés empresarial que podrían contribuir a la toma de decisiones.

Los autores plantean la arquitectura de BI que se debe aplicar en la UTM para una correcta integración de grandes cantidades de datos tal y como se observa en la *Figura 10*. Información muy útil en nuestra investigación debido a la información producida dentro del VIITT de la Ufa-ESPE relacionada con sus pilares fundamentales que son: La investigación, la docencia y vinculación con la sociedad.

Figura 10.

ARQUITECTURA DE BI APLICADA EN LA UTM



Nota: Este gráfico muestra la arquitectura utilizada en la Universidad Tecnológica de Malasia. (Hamizah, Shahizan, & Mi, 2015).

- **EP7** (Sanchez, Coronel, Labanda, Chamaba, & Guamán, 2018): **Computer system for the management and scientific divulgation of the “Universidad Nacional de Loja”.**

En esta investigación los autores buscan la manera de facilitar a los miembros de las instituciones de educación superior del Ecuador (Universidad Nacional de Loja), un sistema de gestión y publicación de producción científica de los grupos de investigación con el objetivo de cumplir con los requerimientos del Consejo de Evaluación, Acreditación y Aseguramiento de la Calidad de la Educación Superior (CEAACES) que son solicitados en los procesos de acreditación de la educación superior en el Ecuador. Para lo cual, se implementa una herramienta informática de apoyo a la investigación científica llamada SIPAC, dando como resultados la gestión de las producciones científicas, principalmente los definidos por el modelo genérico de evaluación del CEAACES. Por esta razón, los autores concluyen que esta herramienta informática genera beneficios a los investigadores en la educación superior porque podrán conocer si se encuentran dentro de los valores mínimos de los indicadores del CEAACES.

Este artículo, nos ayuda en la investigación porque refleja la alta rigurosidad como los indicadores del CEAACES evalúan la calidad de la producción científica de las instituciones de educación superior.

- **EP8 (Krishma & Suneetha, 2016) Business Intelligence solutions for processing huge data to the Business user's using dashboards**

En esta investigación los autores nos manifiestan que la inteligencia empresarial ha comenzado a resolver problemas de las bases de datos relacionales aplicando procesos ETL, que extraen y transforman los datos al área de los objetivos propuestos por las organizaciones, para al final representar estos datos mediante tableros y cuadros estadísticos según la necesidad de los usuarios empresariales, lo cual sirve para mejorar el rendimiento del negocio.

También los autores emplearon herramientas tecnológicas como PowerCenter para realizar los workflows de la información de acuerdo a los objetivos buscados, además usaron un Data Warehouse Administrative Console que es un planificador que sirve para programar los flujos de trabajo, es decir donde se organiza la información seleccionada que van a utilizar y mediante diferentes tipos de mapeo de datos realizados con la aplicación PowerCenter, extraen, transforman y cargan los datos al nuevo almacén y como último paso utilizan una interfaz del software de Oracle Business Intelligence 10g para obtener los paneles y cuadros de información requeridos.

Esta investigación nos plantea la forma de cómo podemos realizar nuestro proceso ETL, para extraer, transformar y cargar los datos de las actividades de investigación del VIITT de la UFA-ESPE.

6. Caracterización de los Estudios Primarios

De acuerdo con la RQ planteada al inicio de la revisión sistemática de literatura ligera (Pág. 30) y a la revisión literaria podemos afirmar que, si es posible mediante el diseño de datamarts enfocados en BI ayuden a la toma de decisiones de la gestión de la gestión de investigación del VIITT, y permita el control de las actividades de investigación en la UFA-ESPE. Debido a que, se identificó dentro de la literatura que los autores consideran, que los sistemas de BI basados en información de procesos ETL, pueden ayudar a los gerentes a obtener los datos de manera más rápida y oportuna sobre su empresa, agilizando la toma de decisiones y el control de esta. Además, los autores manifiestan que los data warehouse y datamarts pueden ser la mejor herramienta para conseguir información de calidad de una institución.

Por otra parte, los enfoques de investigación empleados en los estudios primarios aportan modelos y herramientas que pueden servir a las instituciones de educación superior, pero ninguna de estas contribuciones ha sido probada lo suficiente como para constituirse en una solución definitiva o un modelo estándar que se pueda utilizar para representar los escenarios ETL dentro de las actividades de investigación de la UFA-ESPE. Esto se debe a que cada institución es un mundo diferente debido a su razón social, sus productos, sus servicios o el medio en donde se desenvuelve. Por tal motivo la metodología que se va a implementar en el siguiente capítulo es propia con el fin de apegar a los objetivos propuestos en nuestro proyecto de titulación.

CAPÍTULO III

METODOLOGÍA DE LA INVESTIGACIÓN

Este capítulo describe la investigación realizada. Por medio de esta metodología se logra alcanzar los objetivos planteados con resultados que permitan sustentar el presente trabajo de investigación.

En este proyecto de graduación se utilizó una metodología de investigación cuantitativa debido al uso de indicadores y técnicas estadísticas para medir aspectos o variables de interés de la UFA-ESPE. (Hueso & Cascant, 2012).

Diseño de la investigación

El diseño de investigación es el marco que ha sido creado mediante métodos y técnicas elegidos por el investigador con el fin de encontrar respuestas a las preguntas de investigación planteados. En la *Tabla 5*, se identifican las fases y las actividades que se van a realizar para sustentar las preguntas planteadas y los objetivos específicos del presente proyecto de titulación.

Tabla 5.

ACTIVIDADES DE SUSTENTACIÓN DEL PROYECTO.

Objetivos específicos	Pregunta de investigación	Fase	Actividad
i. Diseñar modelos multidimensionales de base de datos (Datamarts) a fin de almacenar los datos e investigar.	RQ1: ¿Es viable diseñar para almacenar los datos de investigación e innovación tecnológica?	I Identificación	Revisión Sistemática de Literatura Ligera.

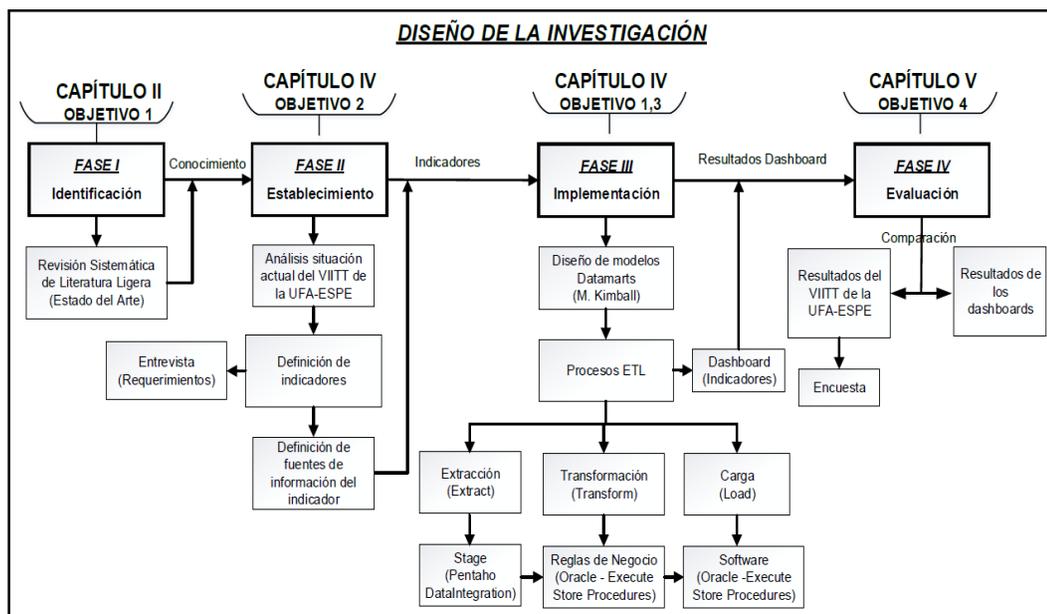
Objetivos específicos	Pregunta de investigación	Fase	Actividad
ii. Diseñar indicadores de control de actividades de investigación enfocados en la acreditación del CACES.	RQ2: ¿Es posible diseñar indicadores de control de actividades de investigación enfocados en la acreditación del CACES?	II Establecimiento	Análisis de la situación actual del VIIT. Definición de indicadores (Entrevista) Definición de fuentes de información.
iii. Aplicar los procesos de extracción, transformación y carga (ETL) a fin de obtener los datos más relevantes de las actividades de investigación.	RQ3: ¿Es posible extraer los datos más relevantes de las actividades de investigación, transformarlos y cargarlos en un nuevo almacén de datos?	III Implementación	Diseños de Datamarts Procesos ETL. Indicadores de resultados. Dashboards.
iv. Validar con el usuario final los reportes generados mediante una herramienta de análisis de BI a fin de verificar el cumplimiento de los indicadores.	RQ4: ¿Es posible validar los resultados de los indicadores de control a partir de un dashboard y el usuario final?	IV Evaluación	Comparación de resultados del VIITT vs Resultados del dashboard (Encuesta).

Nota: Esta tabla muestra las fases y las actividades que se van a realizar en la investigación.

Luego de definir las fases y actividades, se procede a elaborar el diseño de investigación que se encuentra plasmada en la *Figura 11*, la misma que consta de un diagrama de flujo que une las actividades con cada fase mediante entradas y salidas de procesos realizados en forma ordenada.

Figura 11.

DISEÑO DE LA INVESTIGACIÓN.



Nota: Esta gráfica muestra los pasos a seguir para cumplir con los objetivos del trabajo de titulación.

A continuación, se explica detalladamente cada una de estas fases.

a) Primera Fase (Identificación):

Esta fase denominada "Identificación" está relacionada con el objetivo N.º 1 y está enfocada en realizar una revisión sistemática de literatura, etapa fundamental y necesaria para obtener de manera fiable estudios relacionados con el tema de investigación. Para esta parte se utiliza la investigación científica documental. Esta investigación es necesaria porque su estudio se realiza a partir

de un SLR de diferentes fuentes bibliográficas que pueden ser: artículos, libros y capítulos de libros pertenecientes a bases de datos científicas como: SCOPUS y IEEEXplore que sirven para profundizar con información de calidad los conocimientos relacionados con BI que ayuden a la gestión de investigación del VIITT de la UFA-ESPE.

b) Segunda Fase (Establecimiento):

Esta fase se denomina “establecimiento” está relacionada con el objetivo específico N.º 2, que corresponde al diseño de indicadores de control. Para esta parte es necesario realizar tres actividades importantes: la primera actividad es el análisis de la situación actual del VIITT, la segunda actividad es el diseño de los indicadores y la tercera actividad es el establecimiento de las fuentes de información de los indicadores de resultados.

La primera actividad es necesaria para conocer como está conformada el VIITT, los procesos que realiza y los problemas que tiene en la actualidad. Esta información es indispensable para continuar con la segunda actividad, que es el diseño de aproximadamente veinte indicadores de resultados que corresponden a los requerimientos del VIITT de la UFA-ESPE y representan el alcance de nuestra investigación. En la segunda actividad es necesario utilizar una técnica de recolección de datos muy conocida como es la entrevista. La entrevista será aplicada a las personas encargadas de cada proceso de las actividades de investigación a fin de obtener respuestas verbales a los interrogantes planteados sobre el tema propuesto.

Los indicadores se van a diseñar en base a los requerimientos obtenidos en las entrevistas y apegados con el modelo de evaluación Institucional de Universidades y Escuelas Politécnicas del CACES. Luego de construir los

indicadores, viene la tercera actividad que es establecer las fuentes de datos de donde proviene la información.

c) Tercera Fase (Implementación):

Esta fase pertenece a la implementación de la investigación y está relacionado con los objetivos específicos N.º 1 y N.º 3.

El objetivo específico N.º 1 corresponde al diseño de los datamarts, para esta fase es necesario conocer los objetos y procesos de quienes interactúan en el VIITT e identificar las relaciones que existen entre las variables. También, es necesario la ayuda de una herramienta de diseño de bases de datos que permita representar los tres niveles de abstracción que son: conceptual, lógico y físico.

En nuestro caso, utilizaremos el software llamado "Power Designer", que permite elaborar diagramas UML para cubrir la especificación de conceptos, atributos y funciones de las actividades de investigación del VIITT de la UFA-ESPE.

El modelo dimensional que se va a utilizar es el esquema en forma de estrella, debido a que estructura los datos de una forma comprensible para el usuario y sobre todo genera un alto rendimiento en las búsquedas y reportería de la información. Asimismo, la metodología de BI que se va a emplear para la construcción del nuevo almacén de datos es la Metodología de Kimball debido a su versatilidad que simplifica la complejidad de la construcción de una solución BI a corto plazo y esto ayuda a la implementación de Datamarts de una forma más rápida.

El objetivo específico N.º 3 se refiere a los procesos ETL para depurar la información. En estos procesos lo primero que se realiza, es la extracción de los

datos de las diferentes fuentes hacia un Stage⁶. Para esto se va a utilizar un software llamado Pentaho Data Integration.

Una vez que los datos iniciales están cargados en el stage, comienza la segunda parte del proceso que es la transformación de los datos de acuerdo con las reglas del negocio. Para la transformación de los datos se va a seguir los criterios de evaluación expuestos en la *Tabla 6*. Estos criterios de evaluación van a hacer aplicados mediante Store Procedures en Oracle hasta lograr depurar por completo los datos del Stage.

Tabla 6.

CRITERIOS PARA LA TRANSFORMACIÓN DE DATOS.

N.-	Criterios de evaluación	Ejemplo	Corrección
1	Palabras repetidas	Diego Diego Acosta	Diego Acosta
2	Palabras en orden	Orlando Jhonny	Jhonny Orlando
3	Mayúsculas/minúsculas	INGENIERO	Ingeniero
4	Caracteres faltantes	Departameto	Departamento
5	Errores tipográficos	Másterr	Máster
6	Espacios en blancoSistemas	Sistemas
7	Fechas	12-01/2019	12/01/2019

Nota: Esta tabla muestra ejemplos en los criterios de transformación de datos.

⁶ Stage, es un área temporal de almacenamiento de datos utilizada para realizar procesos ETL.

Para finalizar el proceso ETL, tenemos que realizar la carga de los datos depurados a los diferentes modelos de DataMarts diseñados anteriormente.

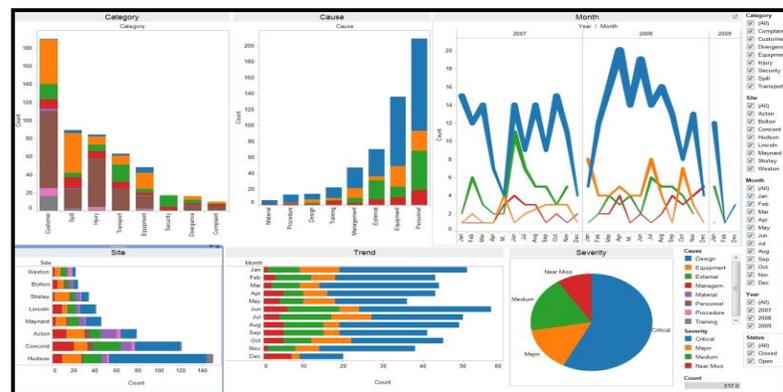
Para esta actividad se va a utilizar la ejecución de Store Procedures en Oracle.

La última parte de esta fase corresponde a la ejecución de los indicadores de resultados de las actividades de investigación del VIITT, mediante el desarrollo de dashboards en base a los indicadores realizados en la fase anterior. Para esta parte, se va a utilizar un software de BI llamado "Tableau". Esta herramienta permite presentar la información importante del estado actual del VIITT de la UFA-ESPE. Esta información aparecerá en forma de representaciones gráficas como se visualiza en la *Figura 12*. Estas representaciones son:

- Diagramas de barras.
- Diagramas circulares.
- Diagramas de tendencia.
- Mapas.

Figura 12.

REPRESENTACIONES ESTADÍSTICAS GRÁFICAS -TABLEAU



Nota: Esta gráfica muestra ejemplos de representaciones gráficas del software Tableau.

(BIS, 2017)

d) Cuarta Fase (Evaluación):

La evaluación es la última fase de nuestra metodología de investigación y está relacionada con el objetivo específico N.º 4 de la presente investigación.

Esta fase inicia con los resultados obtenidos en los dashboards, los cuales van a ser evaluados y validados con las personas responsables de cada área del VIITT. Para esto, se va a utilizar una técnica para recuperar información de los resultados como es la encuesta. La encuesta se va a aplicar a los responsables de cada área del VIITT para corroborar la información de los dashboard con la información que ellos poseen.

El resultado final servirá para conocer si el modelo de BI aplicado es eficiente dentro del VIITT de la UFA-ESPE.

CAPÍTULO IV

MAPEO DE DATOS

En el presente capítulo, se va a desarrollar dos componentes importantes del diseño de investigación planteado en el capítulo anterior. El primer componente es la fase II denominado “Establecimiento” y comprende tres actividades: (la primera actividad es el análisis de la situación actual del VIITT de la UFA-ESPE, la segunda actividad es la definición de indicadores y la tercera actividad es la definición de las fuentes de información de los indicadores).

El segundo componente es la fase III denominado “Implementación”. En esta fase se va a realizar el mapeo de datos⁷ y abarca también tres actividades: (la primera actividad es el diseño del DataMart, la segunda actividad son los procesos ETL y la tercera actividad es la ejecución de los dashboard).

Establecimiento.

Análisis de la Situación Actual del VITT.

El VIITT se encuentra en el nivel operativo dentro de la estructura organizacional de la UFA-ESPE como se identifica en la *Figura 13* y está conformada por ocho componentes:

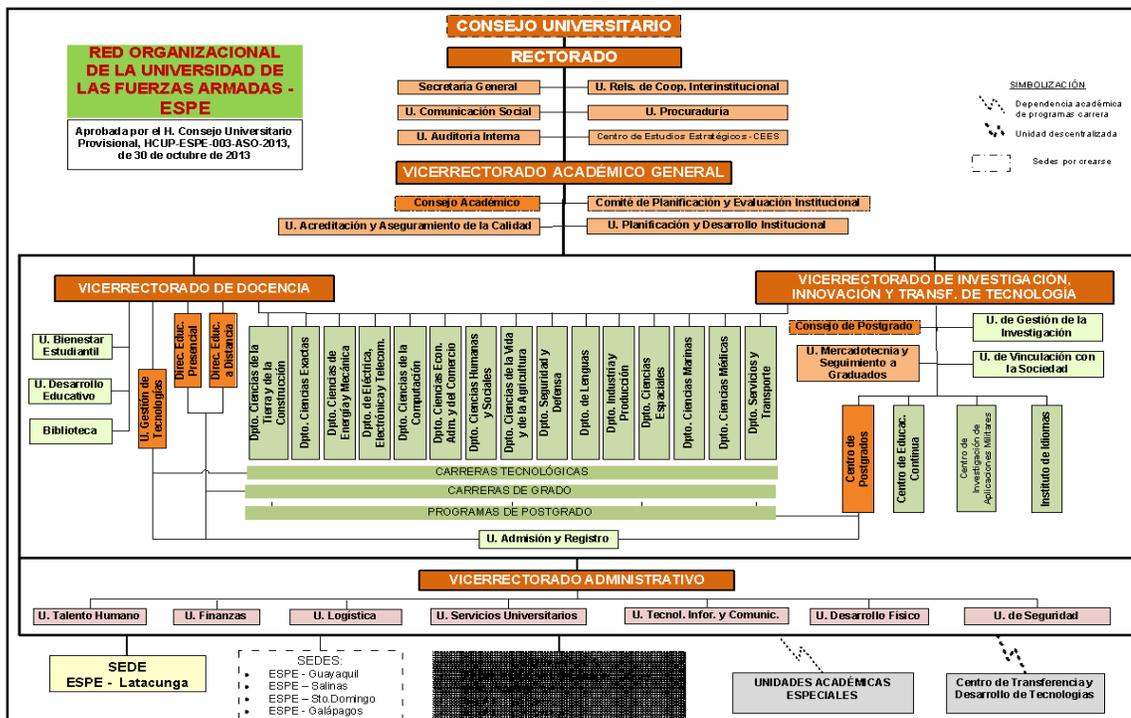
- Consejo de Posgrados
- Centro de Posgrados
- Centro de educación continua.
- Centro de investigaciones de aplicaciones militares
- Instituto de idiomas

⁷ Mapeo de Datos, consiste en asignar los datos de un archivo origen a sus campos destinatarios relacionados.

- La unidad de gestión de la investigación (UGI).
- La unidad de vinculación con la sociedad.
- La unidad de mercadotecnia y seguimiento de graduados.

Figura 13.

ORGANIGRAMA DE LA UFA-ESPE.



Nota: Esta gráfica muestra la red organizacional de la UFA-ESPE. (UGI, 2019).

El primer componente es la UGI y su misión es: “Generar y proteger el conocimiento, apoyar a la formación de investigadores de excelencia, creativos y humanistas, así como facilitar el apoyo técnico a la gestión de la investigación, mediante la aplicación de métodos y procedimientos ágiles, dinámicos y sencillos, a fin de contribuir a la implementación de alternativas de solución a los problemas del país, acordes con el Plan Nacional de Desarrollo”.

La UGI tiene tres macroprocesos: el primer macroproceso pertenece a los proyectos de investigación, el segundo a la movilidad académica y el tercero a la difusión de la ciencia. Los problemas encontrados dentro de la UGI son los siguientes: gran cantidad de información esparcida en varias fuentes de datos, información con diferentes formatos, duplicidad de datos, datos incompletos en las fuentes de información, indicadores incipientes de control de actividades de investigación y baja producción de resultados científicos durante el año 2018. Es por este motivo que el presente trabajo de titulación está enfocado en resolver los problemas antes mencionados en la gestión de investigación.

Definición de Indicadores.

En esta parte se obtiene los indicadores de resultados de las actividades de investigación. Los indicadores son identificados luego de usar una técnica de investigación como es la “entrevista”. La entrevista se empleó como un medio natural de comunicación con el fin de obtener información relevante de las 5 personas responsables de las actividades de investigación del VIITT. Además, fue estructurada con 6 preguntas previamente elaboradas para acoger sugerencias y opiniones para el desarrollo del proyecto de titulación.

Las preguntas y respuestas de los responsables de este proceso se detallan en las *Tablas: 7,8, 9, 10 y 11* que se detallan a continuación:

Tabla 7.*ENTREVISTA REALIZADA AL DIRECTOR DE LA UGI.*

N.-	Preguntas	Respuestas
1	¿Qué función desempeña dentro del departamento del VIITT?	Director de la Unidad de Gestión de la Investigación del VIITT de la UFA-ESPE.
2	¿Qué procesos realiza en el área laboral?	Solicita información a los responsables de las actividades de investigación para conocer el estado actual de la investigación de la UFA-ESPE.
3	¿Cuáles son los productos finales de cada proceso?	Informe de gestión de docencia. Informe de gestión de Investigación. Informe de gestión de vinculación con la sociedad.
4	¿Las fuentes de datos que utiliza?	Docencia. Investigación. Vinculación con la sociedad.
5	¿Las bases de datos que alimenta?	UGI
6	¿Qué indicador de resultados utiliza?	Porcentaje de cumplimiento de los grupos de investigación. Porcentaje de publicaciones realizadas en revistas de impacto.

Nota: Esta tabla muestra las preguntas de la entrevista realizada al director de la UGI.

Tabla 8.*ENTREVISTA DEL ANALISTA DE GESTIÓN DE PROYECTOS.*

N.-	Preguntas	Respuestas
1	¿Qué función desempeña dentro del departamento del VIITT?	Analista de gestión de proyectos de investigación.
2	¿Qué procesos realiza en el área laboral?	Dentro de los procesos que realiza esta la gestión de Líneas de investigación, la gestión de Grupos de Investigación, la gestión de participantes de los grupos de investigación y lleva el control del presupuesto asignado a los grupos de investigación.
3	¿Cuáles son los productos finales de cada proceso?	Informe de gestión de proyectos de investigación.
4	¿Las fuentes de datos que utiliza?	Líneas de investigación. Grupos de investigación. Proyectos de investigación.
5	¿Las bases de datos que alimenta?	Matriz de Proyectos de Investigación (Excel).
6	¿Qué indicador de resultados utiliza?	Presupuesto asignado por años. Presupuesto ejecutado por años.

Nota: Esta tabla muestra las preguntas de la entrevista realizada al analista de gestión de proyectos del VIITT.

Tabla 9.*ENTREVISTA DEL ANALISTA FINANCIERO DE INVESTIGACIÓN.*

N.-	Preguntas	Respuestas
1	¿Qué función desempeña dentro del departamento del VIITT?	Analista financiero del VIITT.
2	¿Qué procesos realiza en el área laboral?	POA, PAC. Presupuesto UGI.
3	¿Cuáles son los productos finales de cada proceso?	Liquidación de ejecución presupuestaria por proyecto de investigación.
4	¿Las fuentes de datos que utiliza?	Estado financiero ESIGEF.
5	¿Las bases de datos que alimenta?	ESIGEF.
6	¿Qué indicador de resultados utiliza?	Porcentaje de ejecución presupuesto por año. Porcentaje de ejecución presupuesto por proyecto. Porcentaje de ejecución presupuesto total por años. Liquidación de proyectos formato institucional.

Nota: Esta tabla muestra las preguntas de la entrevista realizada al analista financiero de investigación.

Tabla 10.*ENTREVISTA DE LA ANALISTA DE DIFUSIÓN DE CIENCIA*

N.-	Preguntas	Respuestas
1	¿Qué función desempeña dentro del departamento del VIITT?	Analista de difusión de la ciencia.
2	¿Qué procesos realiza en el área laboral?	Organización de eventos de investigación de la UFA-ESPE. Divulgación de revistas científicas de la UFA-ESPE. Difusión de Congresos, Libros, Capítulos de libros y Publicaciones científicas.
3	¿Cuáles son los productos finales de cada proceso?	Informe de gestión de la difusión de publicaciones científicas.
4	¿Las fuentes de datos que utiliza?	Grupos de investigación, Proyectos de investigación. Publicaciones científicas y Movilidad.
5	¿Las bases de datos que alimenta?	Matriz de publicaciones científicas (Excel)
6	¿Qué indicador de resultados utiliza?	Porcentaje de publicaciones en revistas indexadas. Porcentaje

Nota: Esta tabla muestra las preguntas de la entrevista realizada al analista de difusión.

Tabla 11.*ENTREVISTA REALIZADA A LA ANALISTA DE MOVILIDAD.*

N.-	Preguntas	Respuestas
1	¿Qué función desempeña dentro del departamento del VIITT?	Analista de movilidad de proyectos de investigación.
2	¿Qué procesos realiza en el área laboral?	Recepción y trámite de las solicitudes de movilidad y publicaciones indexadas.
3	¿Cuáles son los productos finales de cada proceso?	Resoluciones del Comité de Movilidad Plan de Movilidad Institucional
4	¿Las fuentes de datos que utiliza?	Líneas de investigación Reglamento para la Movilidad del Profesor e Investigador de la Universidad de las Fuerzas Armadas ESPE. Proyectos de investigación.
5	¿Las bases de datos que alimenta?	Base de datos de Artículos y Congresos (Excel).
6	¿Qué indicador de resultados utiliza?	Número de solicitudes aprobadas/ Número de solicitudes presentadas. Número de estancias año actual/ Número de estancias año anterior. Número de salidas por movilidad año actual/ Número de salidas por movilidad año anterior Profesores invitados año actual/ año anterior

Nota: Esta tabla muestra las preguntas de la entrevista realizada al analista de movilidad.

Luego de analizar la información proporcionada en las entrevistas y de revisar los indicadores que utiliza el CACES para evaluar a las instituciones de educación superior, se procede a definir los 20 indicadores de resultados de las actividades de Investigación del VIITT.

Los indicadores serán descritos de acuerdo con los siguientes parámetros: nombre del indicador, descripción, periodo de consulta, fuente de información, destino, fórmula y resultado. A continuación, se detallan 20 indicadores de resultados que comprenden desde la *Tabla: 12 a la Tabla 31*.

Tabla 12.

INDICADOR DE PROYECTOS DE INVESTIGACIÓN I.

INDICADOR N.- 1	
Nombre del indicador:	Total de líneas de investigación (TOLI)
Descripción:	Representa el total de líneas de investigación existentes en la UFA-ESPE.
Periodo de consulta:	Mensual
Fuente de información:	Matriz Líneas de investigación.
Destino:	Unidad de Gestión de la Investigación
Fórmula:	$TOLI = \sum \text{Líneas de Investigación}$
Resultado:	Se despliega en forma de barras verticales.



Nota: Esta tabla muestra el total de líneas de investigación de la ESPE.

Tabla 13.

INDICADOR DE PROYECTOS DE INVESTIGACIÓN II.

INDICADOR N.- 2	
Nombre del indicador:	Total de grupos de investigación (TOGI).
Descripción:	Representa el total de grupos de investigación en la UFA-ESPE.
Periodo de consulta:	Mensual.
Fuente de información:	Matriz proyectos.
Destino:	Unidad de Gestión de la Investigación.
Fórmula:	$TOGI = \sum \text{Grupos de Investigación.}$
Resultado:	Se despliega en forma de barras verticales.



Nota: Esta tabla muestra el total de grupos de investigación de la ESPE.

Tabla 14.

INDICADOR DE PROYECTOS DE INVESTIGACIÓN III.

INDICADOR N.- 3	
Nombre del indicador:	Total de fuentes de financiamiento. (TFFI)
Descripción:	Representa el porcentaje de las fuentes de financiamiento asignados para los grupos de investigación.
Periodo de consulta:	Mensual.
Fuente de información:	Matriz proyectos.
Destino:	Unidad de Gestión de la Investigación.
Fórmula:	$TFFI = \frac{\text{Fuente de financiamiento} * 100}{\sum \text{Total de fuentes de financiamiento}}$
Resultado:	Se despliega en forma de anillo.



Nota: Esta tabla muestra las fuentes de financiamiento de la ESPE.

Tabla 15.

INDICADOR DE GRUPOS DE INVESTIGACIÓN IV.

INDICADOR N.- 4	
Nombre del indicador:	Tasa proyectos de investigación por departamento. (TPID)
Descripción:	Representa el total de proyectos de investigación existentes en cada departamento de la UFA-ESPE.
Periodo de consulta:	Mensual
Fuente de información:	Matriz proyectos
Destino:	Unidad de Gestión de la Investigación
Fórmula:	$TPID = \sum \text{Proyectos de investigación por departamento.}$
Resultado:	Se despliega en forma de barras verticales.



Nota: Esta tabla muestra el total de grupos de investigación de la ESPE.

Tabla 16.

INDICADOR GRUPOS DE INVESTIGACIÓN V.

INDICADOR N.- 5	
Nombre del indicador:	Tasa de presupuesto planificado por proyecto de investigación (TPPI).
Descripción:	Representa el presupuesto planificado para cada grupo de investigación.
Periodo de consulta:	Mensual
Fuente de información:	Matriz proyectos
Destino:	Unidad de Gestión de la Investigación
Fórmula:	$TPPI = \text{Presupuesto planificado por proyecto de investigación.}$
Resultado:	Se despliega en forma de barras verticales.



Nota: Esta tabla muestra el presupuesto planificado de los proyectos de investigación.

Tabla 17.

INDICADOR DE DIFUSIÓN I.

INDICADOR N.- 6	
Nombre del indicador:	Resumen de difusión (REDI).
Descripción:	Representa el total de capítulos de libros, congresos, libros y publicaciones indexadas.
Periodo de consulta:	Mensual.
Fuente de información:	Matriz de difusión.
Destino:	Unidad de Gestión de la Investigación.
Fórmula:	$REDI = \sum \text{capítulos de libros, } \sum \text{congresos, } \sum \text{libros y } \sum \text{publicaciones indexadas.}$
Resultado:	Se despliega en forma de gráficos independientes. <div style="display: flex; justify-content: space-around; align-items: center; margin-top: 10px;"> <div style="text-align: center;"> <p>Resumen Difusión</p> <p>Número Capítulos</p>  <p>1</p> </div> <div style="text-align: center;"> <p>Número Congresos</p>  <p>20</p> </div> <div style="text-align: center;"> <p>Número Libros</p>  <p>21</p> </div> </div>

Nota: Esta tabla muestra los libros y publicaciones indexadas.

Tabla 18.

INDICADOR DE DIFUSIÓN II.

INDICADOR N.- 7	
Nombre del indicador:	Total de congresos realizados en el exterior (TCEX).
Descripción:	Representa el total de congresos realizados en países alrededor del mundo.
Periodo de consulta:	Mensual
Fuente de información:	Matriz proyectos
Destino:	Unidad de Gestión de la Investigación
Fórmula:	$TCEX = \sum \text{Congresos realizados en el exterior}$
Resultado:	Se despliega en forma de mapamundi. <div style="display: flex; justify-content: center; align-items: center; margin-top: 10px;">  </div>

Nota: Esta tabla muestra el total de congresos realizados en el exterior.

Tabla 19.

INDICADOR DE DIFUSIÓN III.

INDICADOR N.- 8	
Nombre del indicador:	Tasa de producción de revistas de impacto (TPRI).
Descripción:	Representa el porcentaje de publicaciones realizadas en revistas de impacto mundial y regional.
Periodo de consulta:	Mensual
Fuente de información:	Matriz publicaciones
Destino:	Unidad de Gestión de la Investigación
Fórmula:	$TPRI = \frac{\text{Revista de impacto}}{\text{Total revistas de impacto}} * 100$
Resultado:	Se despliega en forma circular



Nota: Esta tabla muestra el total de grupos de investigación de la ESPE.

Tabla 20.

INDICADOR DE DIFUSIÓN IV.

INDICADOR N.- 9	
Nombre del indicador:	Tasa de publicaciones en bases indexadas (TPBD).
Descripción:	Representa el porcentaje de publicaciones realizadas en bases de datos indexadas.
Periodo de consulta:	Mensual
Fuente de información:	Matriz publicaciones
Destino:	Unidad de Gestión de la Investigación
Fórmula:	$TPBD = \frac{\sum \text{Publicaciones en bases indexadas}}{\text{Total}} * 100$
Resultado:	Se despliega en forma de columnas verticales.

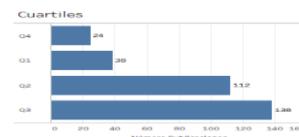


Nota: Esta tabla muestra el total de publicaciones en bases de datos indexadas.

Tabla 21.

INDICADOR DE DIFUSIÓN V.

INDICADOR N.- 10	
Nombre del indicador:	Tasa de publicaciones según el cuartil (TPSQ)
Descripción:	Representa el total de publicaciones ubicadas en grupos de cuartiles según su importancia.
Periodo de consulta:	Mensual
Fuente de información:	Base de movilidad.
Destino:	Unidad de Gestión de la Investigación
Dimensiones:	Movilidad
Fórmula:	$TPSQ = \sum \text{Publicaciones segun su cuartil e importancia.}$
Resultado:	Se despliega en forma de barras horizontales.



Nota: Esta tabla muestra el total de publicaciones ubicadas en cuartiles.

Tabla 22.

INDICADOR DE MOVILIDAD I.

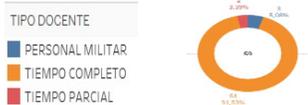
INDICADOR N.- 11	
Nombre del indicador:	Tasa de movilidad de congresos anuales (TMCA).
Descripción:	Este indicador está relacionado con los estándares y elementos proyectivos de evaluación del CACES.
Periodo de consulta:	Anual
Fuente de información:	Base de movilidad.
Destino:	Unidad de Gestión de la Investigación
Fórmula:	$TMCA = \sum \text{Congresos} * \text{Países}$
Resultado:	Se despliega en forma de mapamundi.



Nota: Esta tabla representa los congresos realizados en los diferentes países en el año.

Tabla 23.

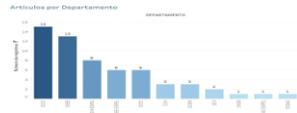
INDICADOR DE MOVILIDAD II.

INDICADOR N.- 12	
Nombre del indicador:	Tasa de publicación de artículos por docente (TPAD).
Descripción:	Este indicador está relacionado con la función sustantiva de investigación del CACES, en la dimensión de resultados de fuentes de información de ponencias presentadas en congresos.
Periodo de consulta:	Mensual
Fuente de información:	Base de movilidad.
Destino:	Unidad de Gestión de la Investigación
Fórmula:	$TPAD = \frac{\text{Dedicación docentes}}{\text{Total docentes}} * 100$ <p>Se despliega en forma de anillo las publicaciones por tipo de docente.</p>
Resultado:	

Nota: Esta tabla muestra los artículos presentados en congresos científicos por docente.

Tabla 24.

INDICADOR DE MOVILIDAD III.

INDICADOR N.- 13	
Nombre del indicador:	Número de artículos por departamento (NAPD)
Descripción:	Este indicador sirve para llevar el control por parte de la UGI y está relacionado con el estándar 10 del CACES.
Periodo de consulta:	Mensual
Fuente de información:	Base de movilidad.
Destino:	Unidad de Gestión de la Investigación
Fórmula:	$NAPD = \sum \text{articulos} * \text{departamento}$ <p>Se despliega en forma de barras verticales.</p>
Resultado:	

Nota: Esta tabla muestra los artículos publicados por departamentos de la ESPE.

Tabla 25.

INDICADOR DE MOVILIDAD IV.

INDICADOR N.- 14	
Nombre del indicador:	Tasa de estancias por departamento (TEPD).
Descripción:	Este indicador sirve para el control de la UGI y está relacionado con el estándar 8 “Planificación de los procesos de investigación” del CACES.
Periodo de consulta:	Mensual
Fuente de información:	Base de movilidad.
Destino:	Unidad de Gestión de la Investigación
Fórmula:	$TIBC = \frac{\sum \text{Publicaciones en bases indexadas}}{\sum \text{Total de Publicaciones indexadas}} * 100$
Resultado:	Se despliega en forma circular las bases indexadas y los cuartiles.



Nota: Esta tabla muestra las estancias fuera del país por departamentos.

Tabla 26.

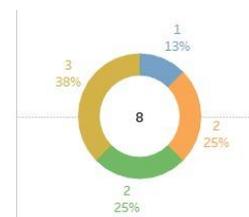
INDICADOR DE MOVILIDAD V.

INDICADOR N.- 15	
Nombre del indicador:	Tasa de docentes invitados por departamento (TDID)
Descripción:	Representa la cantidad de docentes extranjeros que han sido invitados a congresos realizados por la UFA-ESPE. Está relacionado con el estándar proyectivo B “uso social del conocimiento” del CACES.
Periodo de consulta:	Anual.
Fuente de información:	de Matriz de publicaciones
Destino:	Unidad de Gestión de la Investigación

Fórmula:
$$TDID = \frac{\sum \text{docentes invitados}}{\text{Total de docentes invitados}} * \text{departamento}$$

Resultado:

Se despliega en forma de anillo, en el centro está el total de docentes invitados en el año.



Nota: Esta tabla muestra los docentes extranjeros invitados por departamentos.

Tabla 27.

INDICADOR FINANCIERO I.

INDICADOR N.- 16

Nombre del indicador: Valor total codificado y devengado de investigación (VCDI).

Descripción: Muestra los valores asignados por el estado y los valores gastados mensualmente por parte de la UFA-ESPE en lo relacionado a la Investigación.

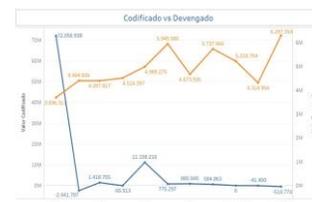
Periodo de consulta: Mensual.

Fuente de información: Datos ESIGEF.

Destino: Unidad de Gestión de la Investigación – Financiero de Investigación.

Fórmula: $\sum \text{valor codificado} * \text{mes vs } \sum \text{valor devengado} * \text{mes.}$

Resultado: Se despliega en forma de gráfico de tendencia.



Nota: Esta tabla muestra los valores financieros codificados y devengados asignados a la ESPE.

Tabla 28.

INDICADOR FINANCIERO II.

INDICADOR N.- 17	
Nombre del indicador:	Valor asignado a proyectos de investigación (VAPY)
Descripción:	Representa el valor asignado por el estado para el desarrollo de cada proyecto de investigación.
Periodo de consulta:	Mensual.
Fuente de información:	Datos ESIGEF.
Destino:	Unidad Gestión Investigación – Financiero de Investigación
Fórmula:	\sum valor codificado (proyectos)
Resultado:	Se despliega en forma de barras verticales



Nota: Esta tabla muestra el valor asignado a los proyectos de investigación ESPE.

Tabla 29.

INDICADOR FINANCIERO III.

INDICADOR N.- 18	
Nombre del indicador:	Valor asignado a programas de investigación (VAPG).
Descripción:	Representa el valor asignado por el estado para el desarrollo de cada programa de investigación.
Periodo de consulta:	Mensual
Fuente de información:	Datos ESIGEF
Destino:	Unidad de Gestión de la Investigación – Financiero de Investigación
Fórmula:	\sum valor codificado (programas)
Resultado:	Se despliega en forma de anillo.



Nota: Esta tabla muestra el valor asignado para cada programa de investigación.

Tabla 30.

INDICADOR FINANCIERO IV.

INDICADOR N.- 19	
Nombre del indicador:	Valor asignado a actividades de investigación (VAAI).
Descripción:	Representa el valor asignado por el estado para el desarrollo de las actividades de investigación.
Periodo de consulta:	Mensual
Fuente de información:	Datos ESIGEF
Destino:	Unidad de Gestión de la Investigación – Financiero de Investigación
Fórmula:	\sum valor codificado (actividades)
Resultado:	Se despliega en forma de barras verticales



Nota: Esta tabla muestra el valor asignado por el estado para las actividades de inv.

Tabla 31.

INDICADOR FINANCIERO V.

INDICADOR N.- 20	
Nombre del indicador:	Valor gastado y valor de saldo de pagos (VGSE).
Descripción:	Muestra el valor gastado y el saldo correspondiente de la estructura de pagos por meses.
Periodo de consulta:	Mensual
Fuente de información:	Datos ESIGEF
Destino:	Unidad de Gestión de la Investigación – Financiero de Investigación
Fórmula:	\sum Valor gastado vs Valor pagado
Resultado:	Se despliega en forma de gráfico de barras combinado.



Nota: Esta tabla muestra los valores gastados y saldos existentes de investigación.

Definición de las fuentes de información del indicador.

Después de concluir con la obtención de indicadores de resultados, vamos a continuar con el análisis de las fuentes de datos. En la *Tabla 32*, se determina de que fuentes de datos vamos a obtener la información para procesar los datos con los indicadores de resultados.

Tabla 32.

FUENTES DE DATOS DE LOS INDICADORES DE RESULTADOS

Indicador	Tipo de indicador	Nombre del Indicador	Fuente de datos	Tipo de datos
1	Grupo de Investigación I	TOLI	Matriz líneas de Inv.	Excel.
2	Grupo de Investigación II	TOGI	Matriz proyectos.	Excel.
3	Grupo de Investigación III	TFFI	Matriz proyectos.	Excel.
4	Grupo de Investigación IV	TPID	Matriz proyectos	Excel.
5	Grupo de Investigación V	TPPI	Matriz proyectos	Excel.
6	Difusión I	REDI	Matriz publicaciones	Excel.
7	Difusión II	TCEX	Matriz publicaciones	Excel.
8	Difusión III	TPRI	Matriz publicaciones	Excel.
9	Difusión IV	TPBD	Matriz publicaciones	Excel.
10	Difusión V	TPSQ	Matriz publicaciones	Excel.
11	Movilidad I	TMCA	Base de movilidad.	Excel.
12	Movilidad II	TPAD	Base de movilidad.	Excel.
13	Movilidad III	NAPD	Base de movilidad.	Excel.
14	Movilidad IV	TEPD	Base de movilidad.	Excel.
15	Movilidad V	TDID	Base de movilidad.	Excel.

16	Financiero I	VCDI	Datos ESIGEF	Excel.
17	Financiero II	VAPY	Datos ESIGEF	Excel.
18	Financiero III	VAPG	Datos ESIGEF	Excel.
19	Financiero IV	VAAI	Datos ESIGEF	Excel.
20	Financiero V	VGSE	Datos ESIGEF	Excel.

Nota: Esta tabla muestra las fuentes de datos necesarias para elaborar los indicadores de resultados.

Implementación.

Construcción del datamart con la metodología Kimball

Para esta parte, la metodología Kimball contempla las siguientes etapas:

a) Planificación

En esta etapa se encuentran detalladas las actividades que se van a cumplir para la creación de los datamart. En la *Tabla 33*, se identifica cada una de las actividades y el tiempo aproximado que se demora en cada etapa.

Tabla 33.

PLANIFICACIÓN DE LA CONSTRUCCIÓN DEL DATAMART.

Orden	Etapa	Actividad	Tiempo
1	Requerimientos	Definición de indicadores.	10 días
2	Análisis de arquitectura	Arquitectura técnica.	5 días
3	Herramientas	Instalación de software a utilizar.	3 días

Orden	Etapa	Actividad	Tiempo
4	Diseño	Modelo dimensional.	5 días
5	Construcción	Datamart.	10 días

Nota: Esta tabla muestra la planificación necesaria para la construcción de los Datamart.

b) Requerimientos

En esta etapa, se definen los requerimientos para el diseño de los datamarts, en este caso es la definición de los indicadores de resultados de investigación que se realizó de forma detallada en la actividad 4.1.2 de este proyecto de titulación. Sin embargo, en la *Tabla 34*, se materializa estos indicadores en forma general.

Tabla 34.

DEFINICIÓN DE REQUERIMIENTOS.

Indicador	Contexto	Nombre del Indicador	Resultado
1	Proyecto de Investigación	TOLI	Gráfico forma de barras verticales
2	Proyecto de Investigación	TOGI	Gráfico forma de barras verticales
3	Proyecto de Investigación	TFFI	Gráfico circular en forma de anillo

Nota: Esta tabla muestra los requerimientos de los usuarios en forma generalizada.

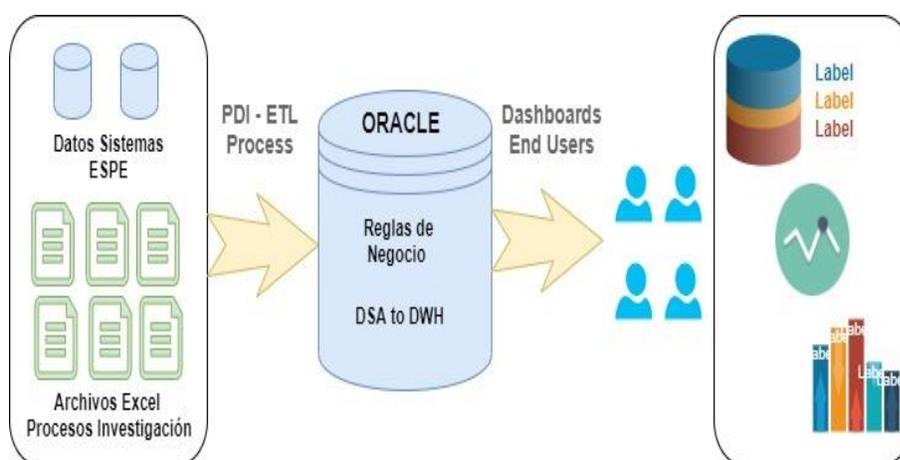
4	Proyecto de Investigación	TPID	Gráfico forma de barras verticales
5	Proyecto de Investigación	TPPI	Gráfico forma de barras verticales
6	Difusión	REDI	Gráficos representativos.
7	Difusión	TCEX	Gráfico mapamundi
8	Difusión	TPRI	Gráfico circular en forma de anillo.
9	Difusión	TPBD	Gráfico forma de barras verticales
10	Difusión	TPSQ	Gráfico forma barras horizontales
11	Movilidad	TMCA	Gráfico en forma de mapamundi.
12	Movilidad	TPAD	Gráfico en forma de anillo.
13	Movilidad	NAPD	Gráfico barras verticales.
14	Movilidad	TEPD	Gráfico en forma de anillo.
15	Movilidad	TDID	Gráfico en forma de anillo.
16	Financiero	VCDI	Gráfico en forma de tendencia.
17	Financiero	VAPY	Gráfico barras verticales.
18	Financiero	VAPG	Gráfico en forma de anillo.
19	Financiero	VAAI	Gráfico barras verticales.
20	Financiero	VGSE	Gráfico barras combinadas.

c) Análisis de arquitectura.

En esta etapa, la actividad a cumplir es el diseño de la arquitectura técnica de nuestro proyecto de titulación, para esto se ha considerado un diseño sencillo como se puede identificar en la *Figura 14*. Este diseño de arquitectura técnica inicia con la depuración de los datos de las diferentes actividades de investigación, continua con el diseño de los datamarts en donde se va a almacenar la información limpia, sigue con los reportes de los dashboard en base a los indicadores de resultados y termina con la toma de decisiones por parte del VIITT de la UFA-ESPE.

Figura 14.

DISEÑO DE ARQUITECTURA TÉCNICA.



Nota: El gráfico representa el diseño de arquitectura técnica que se va a realizar para obtener los indicadores de resultados de investigación. (Elaboración Propia).

d) Herramientas.

Esta etapa comprende la instalación de los softwares necesarios para el desarrollo del proyecto de titulación. En la *Tabla 35*, se enumera las herramientas seleccionadas.

Tabla 35.**HERRAMIENTAS DE SOFTWARE.**

Orden	Nombre de la herramienta	Función
1	Oracle	Base de Datos
2	SQL Developer	Cliente de base de datos
3	Power Designer	Modelamiento de datos
4	Pentaho Data Integration	ETL
5	Tableau Desktop	Dashboard, Business Intelligence

Nota: Esta tabla muestra las herramientas de software que vamos a instalar.

- **Instalación de la base de datos.**

El gestor de base de datos que se procede a instalar es Oracle 10g, por ser una base de datos líder en gestores de bases de datos según el cuadrante de Gartner, definiendo el uso de esta herramienta. La instalación es sencilla gracias al asistente de instalación que viene por defecto como se puede observar en la *Figura 15*.

Figura 15.**INSTALACIÓN DE ORACLE 10 G**

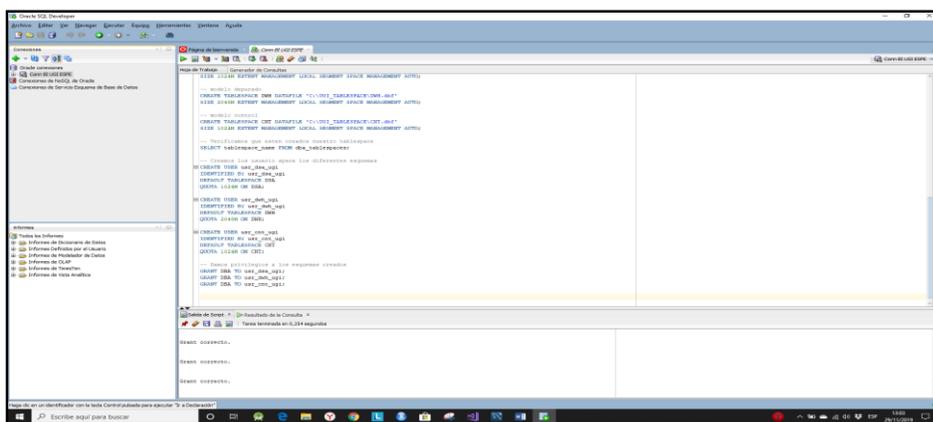
Nota: El gráfico representa la instalación del software Oracle 10G.

- **Instalación del cliente de la base de datos.**

En esta actividad vamos a utilizar el software sql developer porque permite interactuar y desarrollar bases de datos en Oracle. La interacción lo realiza como un cliente para conectarse al gestor de bases de datos de Oracle. Esta herramienta gráfica, portable, de fácil uso y administración se puede observar en la *Figura 16*.

Figura 16.

ORACLE SQL DEVELOPER



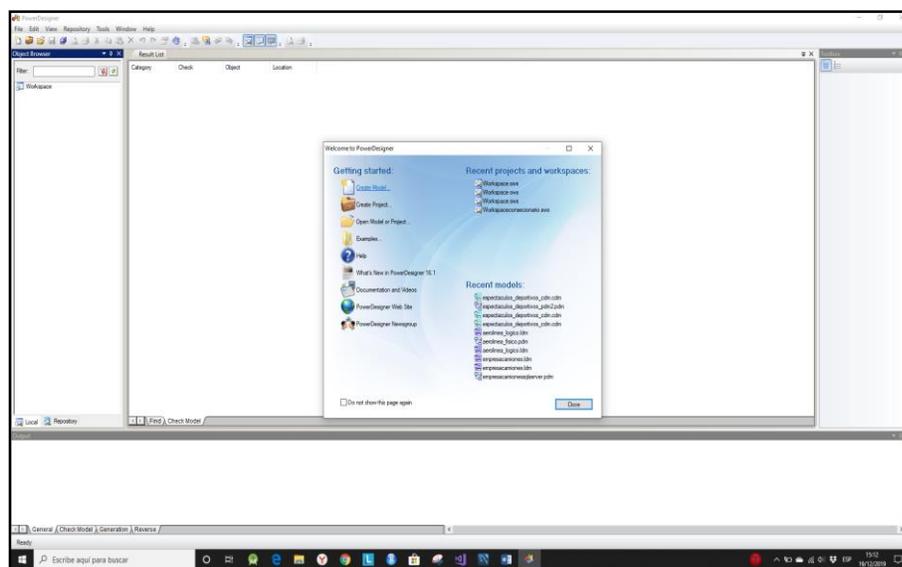
Nota: El gráfico representa la instalación del software Oracle SQL Developer.

- **Instalación de Power Designer.**

El software de modelamiento empresarial que se va a instalar es el power designer, herramienta aprendida en clases y de fácil empleo que permite el análisis, diseño y construcción de una base de datos orientado a modelos de datos a nivel físico y conceptual. Al ejecutar el instalador del programa aparece el asistente de instalación por defecto que facilita este proceso. En la *Figura 17* podemos identificar el asistente de instalación y en la *Figura 18* el software instalado.

Figura 17.**INSTALACIÓN DEL POWER DESIGNER**

Nota: El gráfico representa la instalación del software Power designer.

Figura 18.**SOFTWARE POWER DESIGNER INSTALADO.**

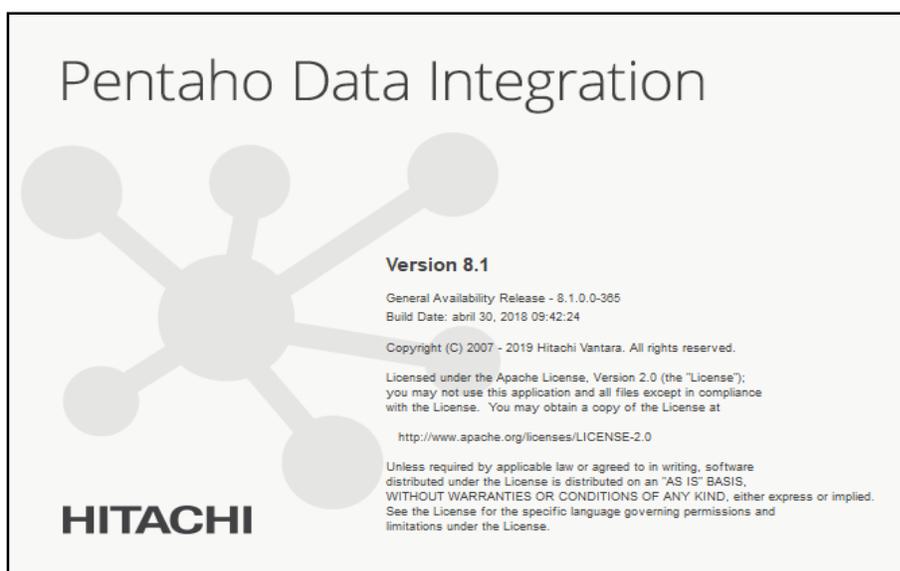
Nota: El gráfico representa el software power designer instalado.

- **Instalación de Pentaho Data Integration**

La herramienta que se va a instalar para los procedimientos ETL es “Pentaho Data Integration” por qué tiene una versión libre que permite utilizar la mayoría de las características del software. Además, es una herramienta portable que facilita su empleo. En la *Figura 19* se muestra la herramienta lista para funcionar.

Figura 19.

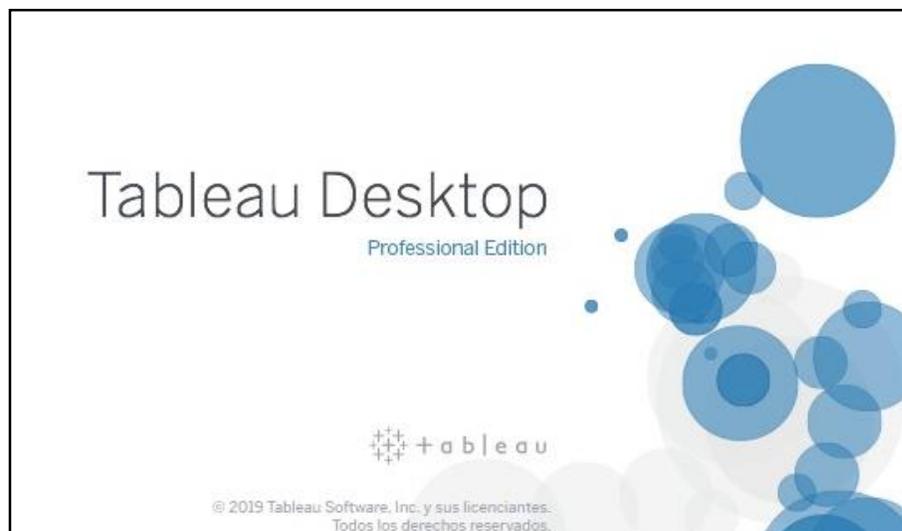
SOFTWARE PENTAHO DATA INTEGRATION.



Nota: El gráfico representa la instalación del software Pentaho Data Integration

- **Instalación de Tableau**

El software que se va a instalar para realizar los dashboard es Tableau Desktop, porque es una herramienta líder en análisis de datos e inteligencia Empresarial según el cuadrante de Gartner. En la *Figura 20* se puede observar la pantalla inicial de Tableau Desktop.

Figura 20.**INSTALACIÓN DE TABLEAU DESKTOP**

Nota: El gráfico representa la instalación del software Tableau Desktop.

e) Diseño

Esta etapa corresponde al diseño de los modelos de datos.

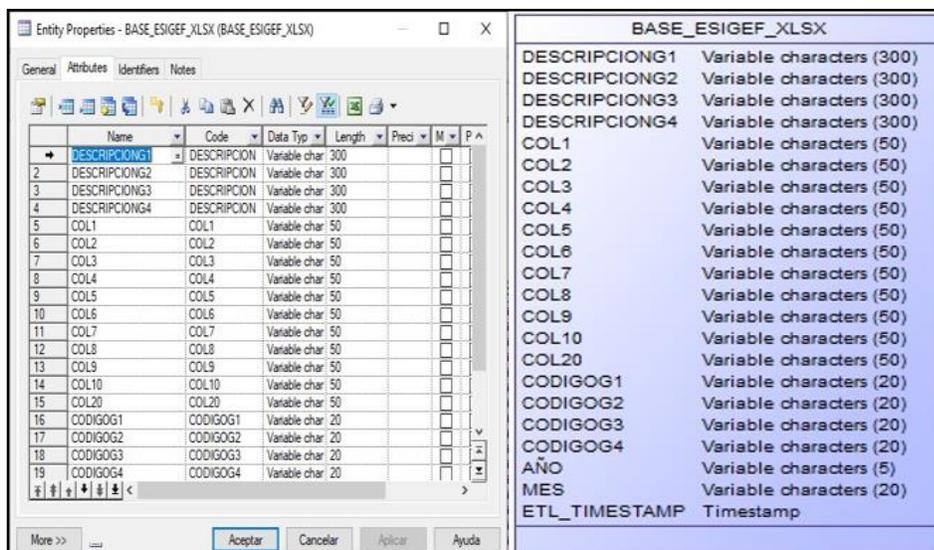
- Modelos de Datos (STAGE)

A partir de las fuentes de información identificadas en la *Tabla 32*, se procede a realizar los modelos de datos lógicos que almacenarán en el Data Stage (DSA) la carga de datos temporales desde las fuentes de datos. Estos modelos de datos son realizados con ayuda del software power designer.

A continuación, se va a iniciar con la creación del stage financiero que consta de dos tablas principales: BASE_ESIGEF y PAGOS que contienen los atributos de las fuentes de información. En la *Figuras 21 y 22* se pueden identificar la creación de estas tablas.

Figura 21.

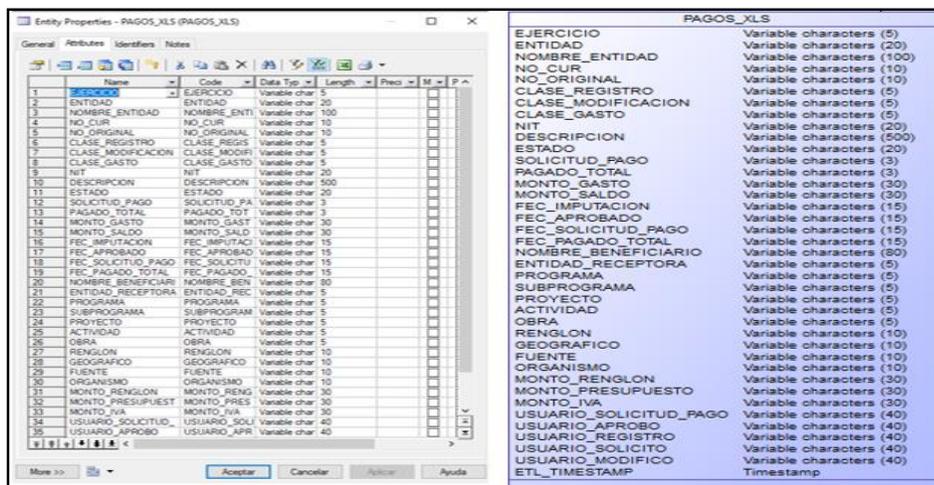
CREACIÓN TABLA BASE ESIGEF



Nota: El gráfico representa la creación de la tabla Base ESIGEF donde se van a almacenar los datos financieros.

Figura 22.

CREACIÓN TABLA PAGOS



Nota: El gráfico representa la creación de la tabla PAGOS donde se van a almacenar los valores pagados.

Luego de crear las tablas de carga iniciales se procede a elaborar un par de tablas nuevas con los nombres anteriores más los caracteres “TMP 1” al final de los nombres. Estas tablas sirven para la limpieza de los datos mediante stored procedures⁸ desde la carga inicial de los archivos excel hasta el mapeo de los campos del negocio. En la *Figura 23*, podemos identificar las tablas temporales financieras creadas.

Figura 23.

CREACIÓN TABLAS TEMPORALES FINANCIERAS.

BASE ESQUEF XLS	PAGOS XLS	PAGOS TMP1
DESCRIPCION31 Variable characters (200)	EJERCICIO Variable characters (5)	EJERCICIO Integer
DESCRIPCION32 Variable characters (100)	ENTIDAD Variable characters (200)	ENTIDAD Variable characters (20)
DESCRIPCION33 Variable characters (200)	NOMBRE ENTIDAD Variable characters (100)	NOMBRE ENTIDAD Variable characters (100)
DESCRIPCION34 Variable characters (200)	NO CUR Variable characters (10)	NO CUR Variable characters (10)
COL1 Variable characters (50)	NO ORIGINAL Variable characters (10)	NO ORIGINAL Variable characters (10)
COL2 Variable characters (50)	CLASE REGISTRO Variable characters (5)	CLASE REGISTRO Variable characters (5)
COL3 Variable characters (50)	CLASE MODIFICACION Variable characters (5)	CLASE MODIFICACION Variable characters (5)
COL4 Variable characters (50)	CLASE GASTO Variable characters (5)	CLASE GASTO Variable characters (5)
COL5 Variable characters (50)	NIT Variable characters (20)	NIT Variable characters (20)
COL6 Variable characters (50)	DESCRIPCION Variable characters (500)	DESCRIPCION Variable characters (500)
COL7 Variable characters (50)	ESTADO Variable characters (20)	ESTADO Variable characters (20)
COL8 Variable characters (50)	SOLICITUD PAGO Variable characters (3)	SOLICITUD PAGO Variable characters (3)
COL9 Variable characters (50)	PAGADO TOTAL Variable characters (30)	PAGADO TOTAL Variable characters (3)
COL10 Variable characters (50)	MONTO GASTO Variable characters (30)	MONTO GASTO Variable characters (3)
COL11 Variable characters (50)	MONTO SALDO Variable characters (30)	MONTO SALDO Number (19,2)
COL12 Variable characters (50)	FEC IMPUTACION Variable characters (15)	FEC IMPUTACION Date
COL13 Variable characters (50)	FEC APROBADO Variable characters (15)	FEC APROBADO Date
COL14 Variable characters (50)	FEC SOLICITUD PAGO Variable characters (15)	FEC SOLICITUD PAGO Date
COL15 Variable characters (50)	FEC PAGADO TOTAL Variable characters (15)	FEC PAGADO TOTAL Date
COL16 Variable characters (50)	NOMBRE BENEFICIARIO Variable characters (80)	NOMBRE BENEFICIARIO Variable characters (80)
COL17 Variable characters (50)	ENTIDAD RECEPTORA Variable characters (5)	ENTIDAD RECEPTORA Variable characters (5)
COL18 Variable characters (50)	PROGRAMA Variable characters (5)	PROGRAMA Variable characters (5)
COL19 Variable characters (50)	SUBPROGRAMA Variable characters (5)	SUBPROGRAMA Variable characters (5)
COL20 Variable characters (50)	PROYECTO Variable characters (5)	PROYECTO Variable characters (5)
COL21 Variable characters (50)	ACTIVIDAD Variable characters (5)	ACTIVIDAD Variable characters (5)
COL22 Variable characters (50)	OBRA Variable characters (5)	OBRA Variable characters (5)
COL23 Variable characters (50)	RENGLON Variable characters (10)	RENGLON Variable characters (10)
COL24 Variable characters (50)	GEOGRAFICO Variable characters (10)	GEOGRAFICO Variable characters (10)
COL25 Variable characters (50)	FUENTE Variable characters (10)	FUENTE Variable characters (10)
COL26 Variable characters (50)	ORGANISMO Variable characters (10)	ORGANISMO Number (19,2)
COL27 Variable characters (50)	MONTO RENGLON Variable characters (30)	MONTO RENGLON Number (19,2)
COL28 Variable characters (50)	MONTO PRESUPUESTO Variable characters (30)	MONTO PRESUPUESTO Number (19,2)
COL29 Variable characters (50)	MONTO IVA Variable characters (30)	MONTO IVA Variable characters (30)
COL30 Variable characters (50)	USUARIO SOLICITUD PAGO Variable characters (40)	USUARIO SOLICITUD PAGO Variable characters (40)
COL31 Variable characters (50)	USUARIO APROBO Variable characters (40)	USUARIO APROBO Variable characters (40)
COL32 Variable characters (50)	USUARIO REGISTRO Variable characters (40)	USUARIO REGISTRO Variable characters (40)
COL33 Variable characters (50)	USUARIO SOLICITO Variable characters (40)	USUARIO SOLICITO Variable characters (40)
COL34 Variable characters (50)	USUARIO MODIFICO Variable characters (40)	USUARIO MODIFICO Variable characters (40)
COL35 Variable characters (50)	ETL_TIMESTAMP Timestamp	ETL_TIMESTAMP Timestamp

Nota: El gráfico representa la creación de las tablas financieras temporales.

A continuación, se realiza el mismo procedimiento para crear los siguientes stages:

- Stage de proyectos de investigación como se identifica en la Figura 24.
- Stage de movilidad como se muestra en la Figura 25.
- Stage de difusión como se observa en la Figura 26.

⁸ Stored procedures, son procedimientos almacenados dentro de una base de datos que son ejecutados mediante peticiones.

Figura 24.

CREACIÓN TABLA TEMPORAL DE PROYECTOS DE INVESTIGACIÓN.

MATRIZ_PROYECTOS_XLS		MATRIZ_PROYECTOS_TMP1	
CODIGO	Variable characters (30)	CODIGO	Variable characters (30)
PROYECTO	Variable characters (500)	PROYECTO	Variable characters (500)
TIPO	Variable characters (100)	TIPO	Variable characters (100)
OBJETIVO	Variable characters (500)	OBJETIVO	Variable characters (500)
NOMBRE_PROGRAMA	Variable characters (300)	NOMBRE_PROGRAMA	Variable characters (300)
LINEA_INVESTIGACION	Variable characters (100)	LINEA_INVESTIGACION	Variable characters (100)
GRUPO_INVESTIGACION	Variable characters (100)	GRUPO_INVESTIGACION	Variable characters (100)
ESTADO	Variable characters (20)	ESTADO	Variable characters (20)
FACULTAD	Variable characters (100)	FACULTAD	Variable characters (100)
FECHA_INICIO	Variable characters (100)	FECHA_INICIO	Date
FECHA_FIN_PLANIFICADA	Variable characters (100)	FECHA_FIN_PLANIFICADA	Date
FECHA_FIN_REAL	Variable characters (100)	FECHA_FIN_REAL	Date
FUENTE_FINANCIAMIENTO	Variable characters (100)	FUENTE_FINANCIAMIENTO	Variable characters (100)
DIRECTOR	Variable characters (100)	DIRECTOR	Variable characters (100)
CODIRECTOR	Variable characters (100)	CODIRECTOR	Variable characters (100)
INVESTIGADOR1	Variable characters (100)	INVESTIGADOR1	Variable characters (100)
INVESTIGADOR2	Variable characters (100)	INVESTIGADOR2	Variable characters (100)
INVESTIGADOR3	Variable characters (100)	INVESTIGADOR3	Variable characters (100)
ESTUDIANTE1	Variable characters (100)	ESTUDIANTE1	Variable characters (100)
ESTUDIANTE2	Variable characters (100)	ESTUDIANTE2	Variable characters (100)
PRESUPUESTO_PLANIFICADO	Variable characters (30)	PRESUPUESTO_PLANIFICADO	Number (19,2)
PRESUPUESTO_EJECUTADO	Variable characters (30)	PRESUPUESTO_EJECUTADO	Number (19,2)
PORCENTAJE_CUMPLIMIENTO	Variable characters (10)	PORCENTAJE_CUMPLIMIENTO	Number (19,2)

Nota: El gráfico representa la creación de las tablas temporales de proyectos de investigación.

Figura 25.

CREACIÓN TABLA TEMPORAL DE MOVILIDAD.

ARTICULOS_CONGRESO_XLS		ARTICULOS_CONGRESO_TMP1	
NUMERO	Variable characters (20)	NUMERO	Variable characters (20)
NOMBRE	Variable characters (100)	NOMBRE	Variable characters (100)
CEDULA	Variable characters (15)	CEDULA	Variable characters (15)
AUTOR	Variable characters (100)	AUTOR	Variable characters (100)
DEPARTAMENTO	Variable characters (100)	DEPARTAMENTO	Variable characters (100)
TIPO_DOCENTE	Variable characters (30)	TIPO_DOCENTE	Variable characters (30)
NOMBRE_ARTICULO	Variable characters (1000)	NOMBRE_ARTICULO	Variable characters (1000)
NOMBRE_REVISTA	Variable characters (1000)	NOMBRE_REVISTA	Variable characters (1000)
CONGRESO_PROC_SCOPUS	Variable characters (60)	CONGRESO_PROC_SCOPUS	Variable characters (60)
CONGRESO_INDX_SCOPUS	Variable characters (60)	CONGRESO_INDX_SCOPUS	Variable characters (60)
REVISTA_SCOPUS	Variable characters (20)	REVISTA_SCOPUS	Variable characters (20)
REVISTA_LATINDEX	Variable characters (20)	REVISTA_LATINDEX	Variable characters (20)
ISI_WEB	Variable characters (20)	ISI_WEB	Variable characters (20)
EBSCO	Variable characters (20)	EBSCO	Variable characters (20)
LIBRO_MEMS_ISBN	Variable characters (20)	LIBRO_MEMS_ISBN	Variable characters (20)
PAIS	Variable characters (50)	PAIS	Variable characters (50)
CIUDAD	Variable characters (50)	CIUDAD	Variable characters (50)
FECHA_INICIO	Variable characters (50)	FECHA_INICIO	Variable characters (50)
FECHA_FIN	Variable characters (50)	FECHA_FIN	Variable characters (50)
VALOR_PASAJE	Variable characters (50)	VALOR_PASAJE	Variable characters (50)
VALOR_VIATICOS	Variable characters (50)	VALOR_VIATICOS	Variable characters (50)
VALOR_INSCRIPCION	Variable characters (50)	VALOR_INSCRIPCION	Variable characters (50)
OBSERVACION	Variable characters (500)	OBSERVACION	Variable characters (500)
ANIO	Variable characters (10)	ANIO	Variable characters (10)
ETL_TIMESTAMP	Timestamp	ETL_TIMESTAMP	Timestamp

Nota: El gráfico representa la creación de las tablas temporales de movilidad.

Figura 26.

CREACIÓN DE LAS TABLAS TEMPORALES DE DIFUSIÓN.

CAPITULOS_LIBROS_XLS		CAPITULOS_LIBROS_TMP1	
CODIGO_IES	Variable characters (20)	CODIGO_IES	Variable characters (20)
TIPO_PUBLICACION	Variable characters (100)	TIPO_PUBLICACION	Variable characters (100)
CODIGO_PUBLICACION	Variable characters (200)	CODIGO_PUBLICACION	Variable characters (200)
TITULO_CAPITULO	Variable characters (300)	TITULO_CAPITULO	Variable characters (300)
TITULO_LIBRO	Variable characters (300)	TITULO_LIBRO	Variable characters (300)
CODIGO_ISBN	Variable characters (30)	CODIGO_ISBN	Variable characters (30)
EDITOR_COMPILADOR	Variable characters (500)	FECHA_PUBLICACION	Date
PAGINAS	Variable characters (50)	IDENTIFICACION_PARTICIPANTE	Variable characters (30)
FECHA_PUBLICACION	Variable characters (30)	PARTICIPACION	Variable characters (50)
CAMPO_DETALLADO	Variable characters (50)	NOMBRES	Variable characters (100)
FILIACION	Variable characters (30)	DEPARTAMENTO	Variable characters (100)
IDENTIFICACION_PARTICIPANTE	Variable characters (30)	SEDE	Variable characters (100)
PARTICIPACION	Variable characters (50)	OBSERVACIONES	Variable characters (300)
NOMBRES	Variable characters (100)	ANIO	Integer
DEPARTAMENTO	Variable characters (100)	ETL_TIMESTAMP	Timestamp
SEDE	Variable characters (100)		
OBSERVACIONES	Variable characters (300)		
ANIO	Variable characters (5)		
ETL_TIMESTAMP	Timestamp		

LIBROS_XLS		LIBROS_TMP1	
CODIGO_IES	Variable characters (20)	CODIGO_IES	Variable characters (20)
TIPO_PUBLICACION	Variable characters (100)	TIPO_PUBLICACION	Variable characters (100)
CODIGO_PUBLICACION	Variable characters (200)	CODIGO_PUBLICACION	Variable characters (200)
TITULO_LIBRO	Variable characters (300)	TITULO_LIBRO	Variable characters (300)
CODIGO_ISBN	Variable characters (30)	CODIGO_ISBN	Variable characters (30)
FECHA_PUBLICACION	Variable characters (30)	FECHA_PUBLICACION	Variable characters (30)
CAMPO_DETALLADO	Variable characters (50)	IDENTIFICACION_PARTICIPANTE	Variable characters (30)
REVISADO_PARES	Variable characters (30)	PARTICIPACION	Variable characters (50)
FILIACION	Variable characters (30)	NOMBRES	Variable characters (100)
IDENTIFICACION_PARTICIPANTE	Variable characters (30)	DEPARTAMENTO	Variable characters (100)
PARTICIPACION	Variable characters (50)	SEDE	Variable characters (100)
NOMBRES	Variable characters (100)	OBSERVACIONES	Variable characters (300)
DEPARTAMENTO	Variable characters (100)	ANIO	Variable characters (5)
SEDE	Variable characters (100)	ETL_TIMESTAMP	Timestamp
OBSERVACIONES	Variable characters (300)		
ANIO	Variable characters (5)		
ETL_TIMESTAMP	Timestamp		

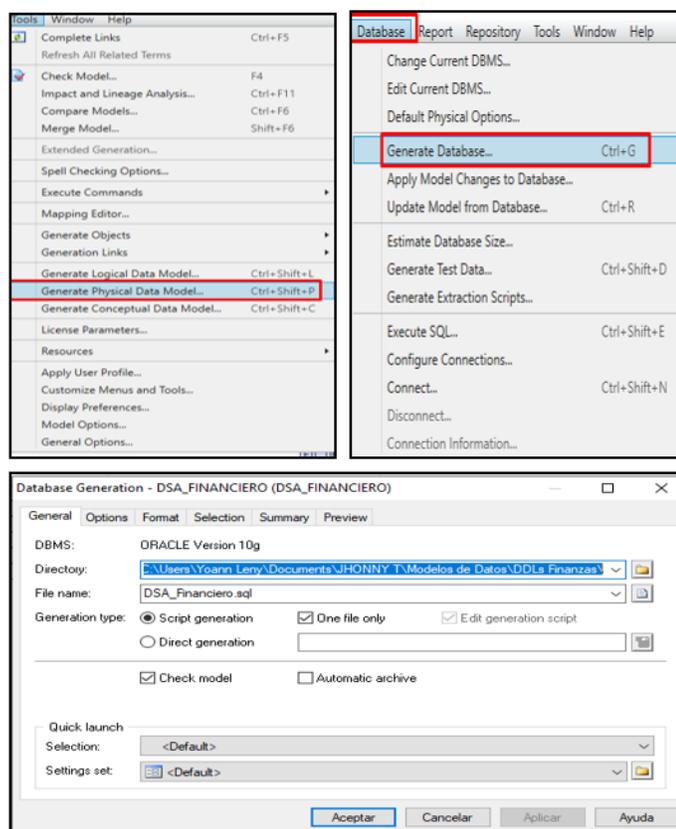
PUBLICACIONES_XLS		PUBLICACIONES_TMP1	
CODIGO_IES	Variable characters (20)	CODIGO_IES	Variable characters (20)
TIPO_PUBLICACION	Variable characters (100)	TIPO_PUBLICACION	Variable characters (100)
TIPO_ARTICULO	Variable characters (50)	TIPO_ARTICULO	Variable characters (50)
CODIGO_PUBLICACION	Variable characters (200)	CODIGO_PUBLICACION	Variable characters (200)
TITULO_PUBLICACION	Variable characters (500)	TITULO_PUBLICACION	Variable characters (500)
BASE_DATOS_INDEXADA	Variable characters (50)	BASE_DATOS_INDEXADA	Variable characters (50)
CODIGO_ISSN	Variable characters (30)	CODIGO_ISSN	Variable characters (30)
NOMBRE_REVISTA	Variable characters (300)	NOMBRE_REVISTA	Variable characters (300)
CUARTIL	Variable characters (5)	CUARTIL	Variable characters (5)
NUMERO_REVISTA	Variable characters (100)	FECHA_PUBLICACION	Date
FECHA_PUBLICACION	Variable characters (30)	SJR	Number (19,2)
SJR	Variable characters (30)	LINK_PUBLICACION	Variable characters (500)
CAMPO_DETALLADO	Variable characters (50)	LINK_REVISTA	Variable characters (500)
ESTADO	Variable characters (50)	FILIACION	Variable characters (30)
LINK_PUBLICACION	Variable characters (500)	IDENTIFICACION_PARTICIPANTE	Variable characters (30)
LINK_REVISTA	Variable characters (500)	PARTICIPACION	Variable characters (50)
FILIACION	Variable characters (30)	NOMBRES	Variable characters (100)
IDENTIFICACION_PARTICIPANTE	Variable characters (30)	DEPARTAMENTO	Variable characters (100)
PARTICIPACION	Variable characters (50)	SEDE	Variable characters (100)
NOMBRES	Variable characters (100)	OBSERVACIONES	Variable characters (300)
DEPARTAMENTO	Variable characters (100)	ANIO	Variable characters (5)
SEDE	Variable characters (100)	ETL_TIMESTAMP	Timestamp
OBSERVACIONES	Variable characters (300)		
ANIO	Variable characters (5)		
ETL_TIMESTAMP	Timestamp		

Nota: El gráfico representa la creación de las tablas temporales de difusión.

A partir del modelo lógico, se crea el modelo de datos físico para obtener los scripts⁹ de creación de las tablas para la base de datos. Para esta actividad en la opción “Tools” escogemos la opción Generate Physical Data Model. Una vez creado el modelo físico se habilita en la barra de herramientas una opción que se llama “Database” y damos clic en Generate Database, a continuación, se despliega una ventana que indica la dirección en donde se va a crear el script. Este procedimiento se muestra en la *Figura 27* y se repite para cada uno de los stages.

Figura 27.

GENERACIÓN DE SCRIPTS.



Nota: El gráfico representa la generación de scripts que van a crear las tablas de datos.

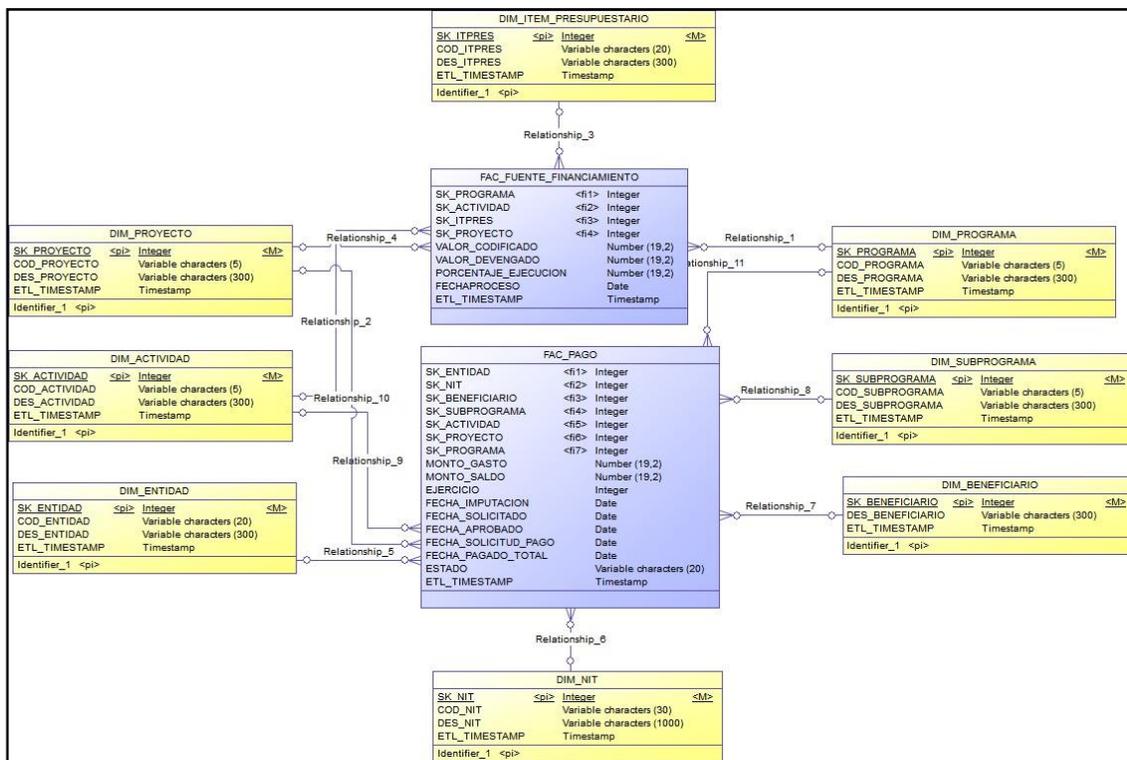
⁹ Script es un archivo que contiene un conjunto de comandos a ejecutarse en un motor de base de datos.

- **Modelos de Datos (Multidimensionales propuestos).**

A continuación, se muestran los modelos multidimensionales propuestos para almacenar la información depurada de las actividades de investigación: proyectos de investigación, movilidad, difusión y financiero. Estos modelos tienen dos colores (morado y amarillo), con color morado tenemos a los “hechos” y con color amarillo a las “dimensiones”. En la *Figura 28*, se puede observar el modelo multidimensional financiero, en la *Figura 29*, se muestra el modelo multidimensional de proyectos de investigación, en la *Figura 30*, encontramos al modelo multidimensional de movilidad y por último en la *Figura 31*, se detalla al modelo multidimensional de difusión.

Figura 28.

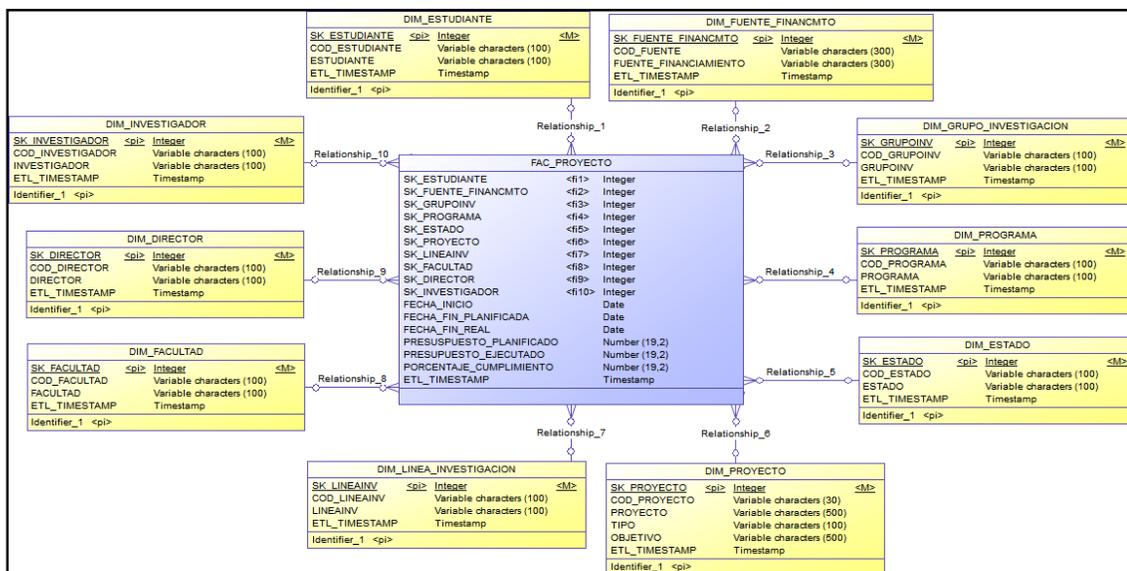
MODELO MULTIDIMENSIONAL FINANCIERO



Nota: El gráfico representa las dimensiones del modelo financiero.

Figura 29.

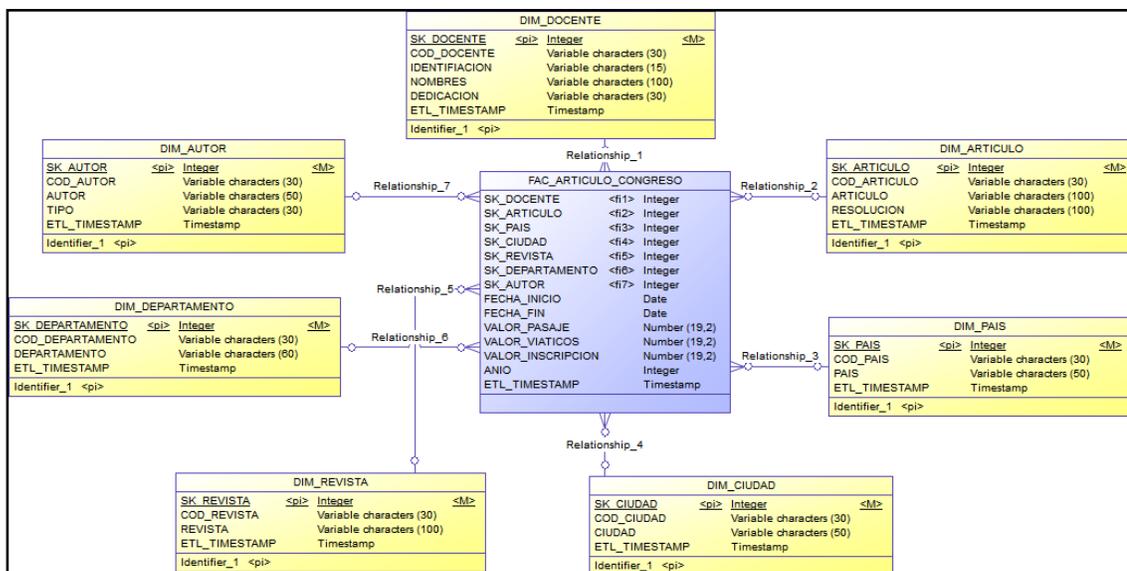
MODELO MULTIDIMENSIONAL DE PROYECTOS DE INVESTIGACIÓN



Nota: El gráfico representa las dimensiones del modelo de proyectos de investigación

Figura 30.

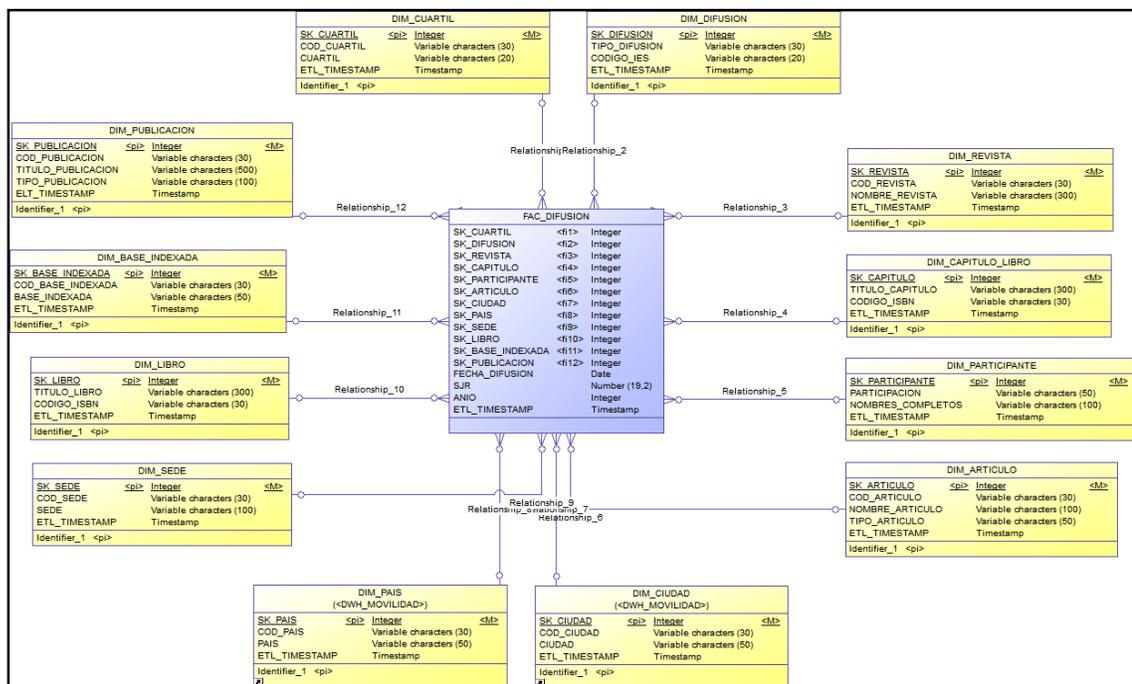
MODELO MULTIDIMENSIONAL DE MOVILIDAD



Nota: El gráfico representa las dimensiones del modelo de datos de movilidad.

Figura 31.

MODELO MULTIDIMENSIONAL DE DIFUSIÓN.



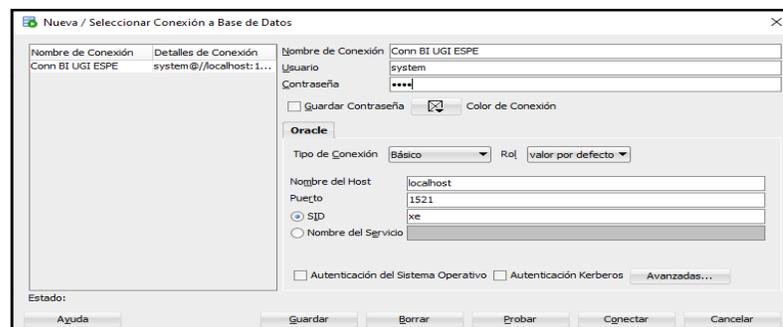
Nota: El gráfico representa las dimensiones del modelo de datos de difusión.

f) Construcción

Esta etapa corresponde a la construcción de los datamart. El datamart es el repositorio donde se va a almacenar los datos depurados más relevantes de las actividades del VIITT de la UFA-ESPE. Para la construcción del datamart vamos a realizar las siguientes actividades:

- Creación de esquemas y usuarios

Luego de ejecutar la herramienta SQL Developer, se procede a crear una nueva conexión con el nombre "Conn BI UGI ESPE". En esta nueva conexión se va a crear la arquitectura que vamos a utilizar en los Datamarts, como se identifica en la Figura 32.

Figura 32.**CONEXIÓN DE BASE DE DATOS**

Nota: El gráfico representa el proceso de conexión a la base de datos.

Después, de estar conectados a la base Conn BI UGI ESPE, se comienza a crear los tablespace para almacenar los diferentes esquemas de bases de datos. En la *Figura 33*, se muestra los comandos para crear 3 tipos de Tablespace:

- Tablespace DSA: para el modelo temporal.
- Tablespace DWH: para el modelo depurado.
- Tablespace CNT: para el modelo control.

Figura 33.**CREACIÓN DE LOS TABLESPACE.**

```

-- CREACIÓN DE BASES DE DATOS
-- modelo temporal
CREATE TABLESPACE DSA DATAFILE 'C:\UGI_TABLESPACE\DSA.dbf'
SIZE 1024M EXTENT MANAGEMENT LOCAL SEGMENT SPACE MANAGEMENT AUTO;
-- modelo depurado
CREATE TABLESPACE DWH DATAFILE 'C:\UGI_TABLESPACE\DWH.dbf'
SIZE 2048M EXTENT MANAGEMENT LOCAL SEGMENT SPACE MANAGEMENT AUTO;
-- modelo control
CREATE TABLESPACE CNT DATAFILE 'C:\UGI_TABLESPACE\CNT.dbf'
SIZE 1024M EXTENT MANAGEMENT LOCAL SEGMENT SPACE MANAGEMENT AUTO;

```

Salida de Script x Resultado de la Consulta x
Tarea terminada en 0,254 segundos

```

TABLESPACE DSA creado.
TABLESPACE DWH creado.
TABLESPACE CNT creado.

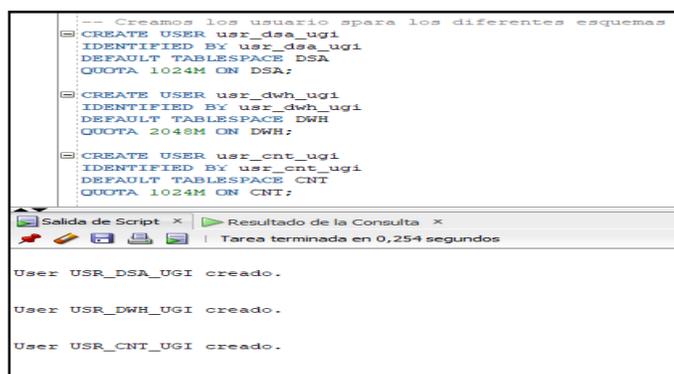
```

Nota: El gráfico representa la creación de los tablespace en la base de datos.

Después, de haber creado los tablespaces debemos crear los usuarios respectivos que podrán acceder a los objetos de estos, para lo cual se ejecutó los comandos que se encuentran en la *Figura 34*.

Figura 34.

CREACIÓN DE USUARIOS.



```
-- Creamos los usuario para los diferentes esquemas
CREATE USER usr_dsa_ugi
IDENTIFIED BY usr_dsa_ugi
DEFAULT TABLESPACE DSA
QUOTA 1024M ON DSA;

CREATE USER usr_dwh_ugi
IDENTIFIED BY usr_dwh_ugi
DEFAULT TABLESPACE DWH
QUOTA 2048M ON DWH;

CREATE USER usr_cnt_ugi
IDENTIFIED BY usr_cnt_ugi
DEFAULT TABLESPACE CNT
QUOTA 1024M ON CNT;
```

Salida de Script x Resultado de la Consulta x
Tarea terminada en 0,254 segundos

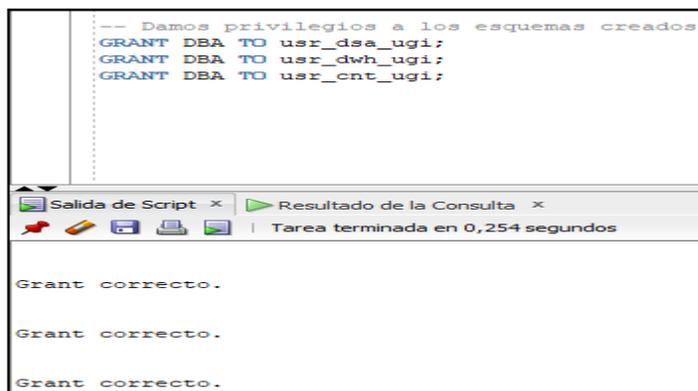
User USR_DSA_UGI creado.
User USR_DWH_UGI creado.
User USR_CNT_UGI creado.

Nota: El gráfico representa la creación de los usuarios que usaran los tablespaces.

Por último, se ejecuta los comandos de la *Figura 35*, que proporcionan los privilegios necesarios para que los usuarios tengan acceso a las bases de datos y a los objetos correspondientes.

Figura 35.

PRIVILEGIOS DE USUARIO



```
-- Damos privilegios a los esquemas creados
GRANT DBA TO usr_dsa_ugi;
GRANT DBA TO usr_dwh_ugi;
GRANT DBA TO usr_cnt_ugi;
```

Salida de Script x Resultado de la Consulta x
Tarea terminada en 0,254 segundos

Grant correcto.
Grant correcto.
Grant correcto.

Nota: El gráfico representa los privilegios de usuario.

- **Conexión a la base de datos.**

Al ejecutar el software sql developer se despliega una ventana y dentro de esta se procede a seleccionar la opción que dice “nueva conexión a base de datos” y en la ventana que aparece completamos con los siguientes datos:

Nombre de la conexión: Conn DSA BI UGI ESPE

Usuario: usr_dsa_ugi

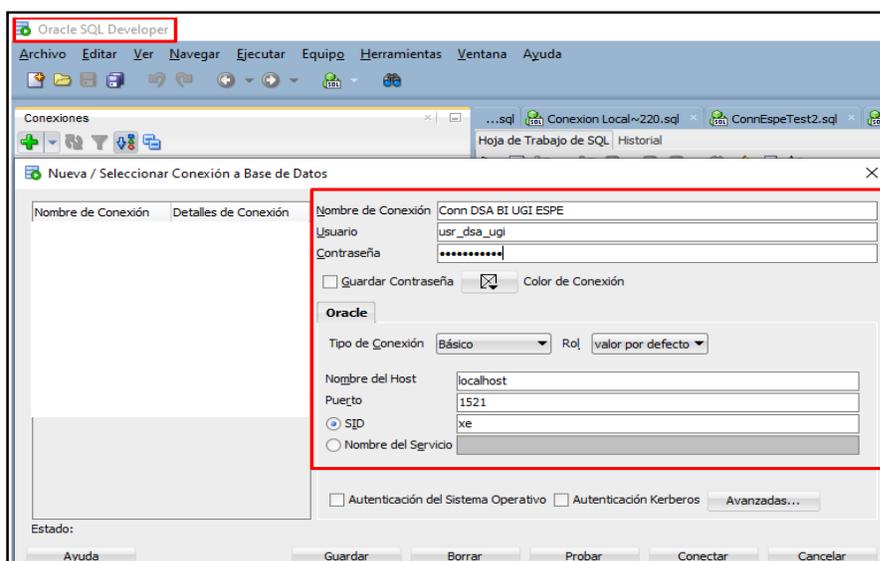
Password:

Host: localhost (temporalmente).

En la *Figura 36*, podemos observar el procedimiento realizado.

Figura 36.

CONEXIÓN AL DSA BU UGI ESPE.

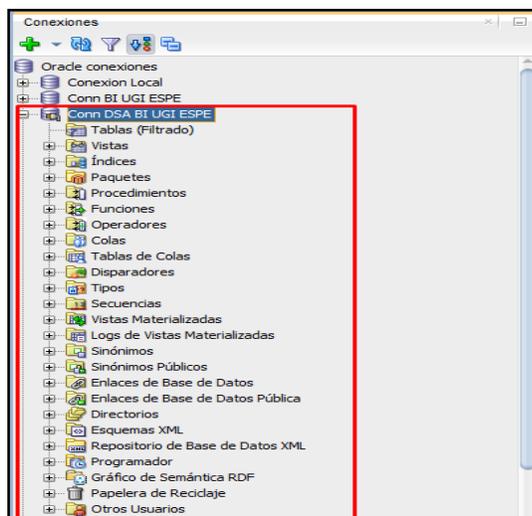


Nota: El gráfico representa la conexión al DSA BU UGI ESPE

Una vez creada la conexión se despliega los objetos en el schema DSA, como se puede identificar en la *Figura 37*.

Figura 37.

SCHEMAS DSA

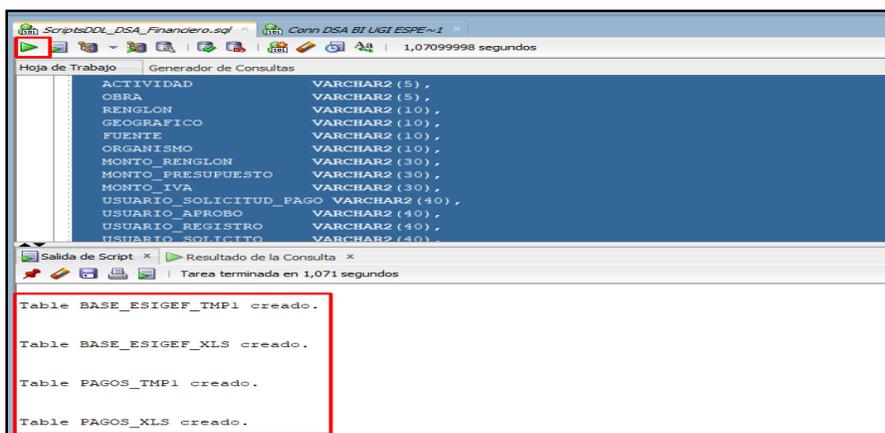


Nota: El gráfico representa los schemas creados.

Luego vamos a ejecutar el script creado anteriormente en la hoja de trabajo SQL. En la *Figura 38*, se puede identificar la creación de las tablas para el stage financiero.

Figura 38.

CREACIÓN DE TABLAS FINANCIERAS EN EL ESQUEMA DSA.



Nota: El gráfico representa la creación de las tablas financieras en el esquema DSA.

Por último, vamos a ejecutar el script creado en la hoja de trabajo SQL. En la *Figura 39*, se puede observar la creación de las tablas para el DWH financiero.

Figura 39.

CREACIÓN DE TABLAS FINANCIERAS EN EL ESQUEMA DWH

```

/*-----*/
create table FAC_FUENTE_FINANCIAMIENTO (
  SK_PROGRAMA          INTEGER,
  SK_ACTIVIDAD         INTEGER,
  SK_TIPORES           INTEGER,
  SK_PROYECTO         INTEGER,
  VALOR_CODIFICADO     NUMBER(19,2),
  VALOR_DEVENGADO     NUMBER(19,2),
  PORCENTAJE_EJECUCION NUMBER(19,2),
  FECHAPROCESO        DATE,
  ETL_TIMESTAMP        TIMESTAMP DEFAULT CURRENT_TIMESTAMP
);
/*-----*/
/* Table: FAC_PAGO */
/*-----*/
create table FAC_PAGO (
  SK_ENTIDAD           INTEGER
);

```

Salida de Script x

Tarea terminada en 0,562 segundos

Table FAC_FUENTE_FINANCIAMIENTO creado.

Table FAC_PAGO creado.

Nota: El gráfico representa la creación de las tablas financieras en el esquema DWH

Luego realizamos el mismo procedimiento para los otros procesos tanto para los Stage como para los DWH.

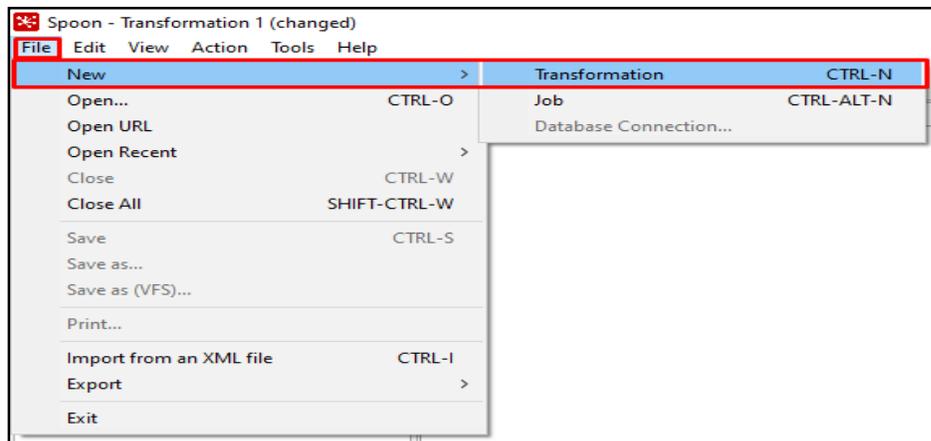
Procesos ETL

Esta etapa corresponde a la descarga de datos de las actividades del VIITT, que permite integrar los datos más importantes para los procesos operativos y analíticos de los dashboard. Esta etapa inicia luego de definir la estructura de datos en el stage y termina con la carga de los datos depurados en el DataMart.

Para esta actividad, vamos a ejecutar el software Pentaho Data Integration para crear una nueva transformación que permita realizar las cargas de los archivos Excel al Stage como se puede identificar en la *Figura 40*.

Figura 40.

TRANSFORMACIÓN DE DATOS

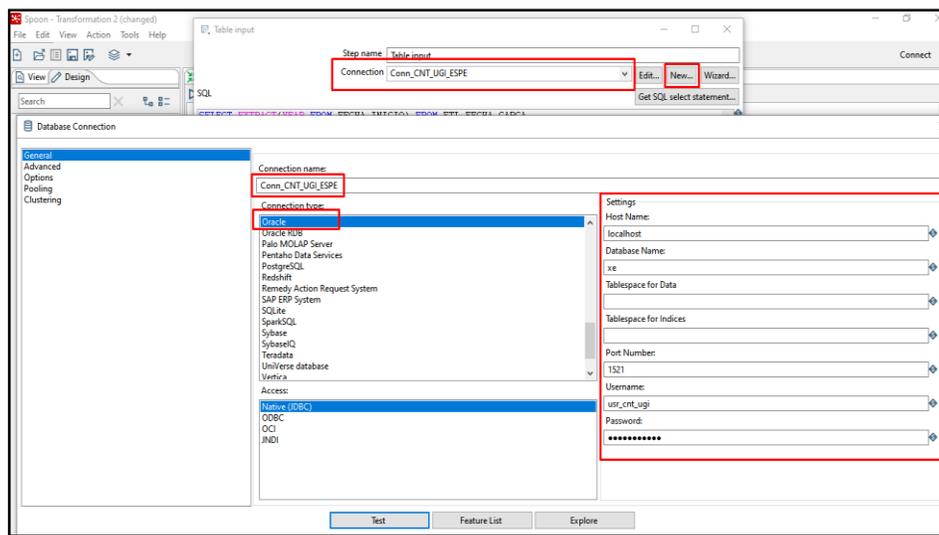


Nota: El gráfico representa la transformación de datos ayudados de Pentaho.

Posterior a esto, aparece una nueva ventana que permite generar una conexión a un schema de Oracle como se observa en la *Figura 41*, en este paso se debe configurar la conexión, el puerto, el host, el username, la contraseña y el nombre de la database.

Figura 41.

CONEXIÓN A ORACLE.

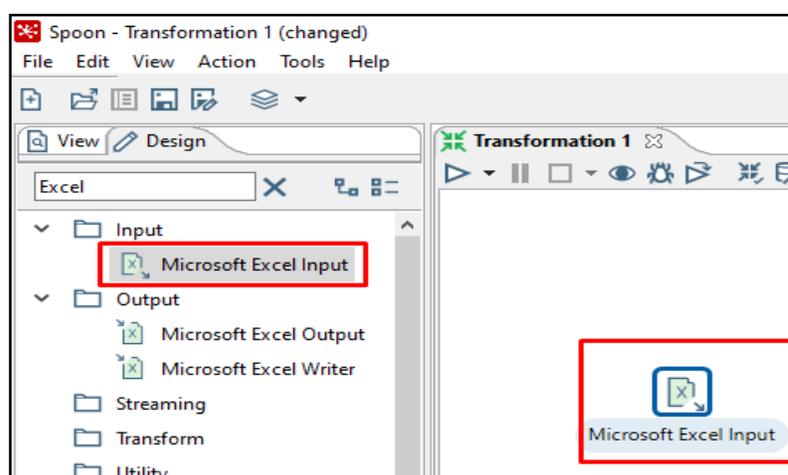


Nota: El gráfico representa la conexión a la herramienta de Oracle.

Después, se abre una nueva ventana en donde se elige el objeto “table input” y dentro de este escogemos la opción “Microsoft Excel Input” para acceder a los datos de los archivos Excel de las actividades de investigación. En la *Figura 42*, podemos identificar este paso.

Figura 42.

OBJETO PARA OBTENER DATOS DE ARCHIVOS EXCEL.

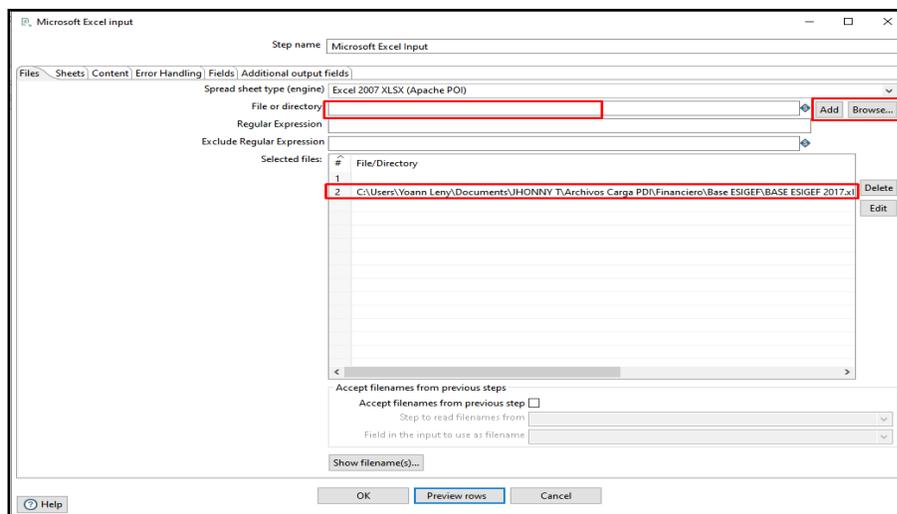


Nota: El gráfico representa la forma de extraer los datos de los archivos fuentes.

Luego, configuramos el objeto seleccionado para apuntar hacia el archivo de Excel correspondiente a cada actividad de investigación como se muestra en la *Figura 43*, para esto se debe tener en cuenta que la herramienta de Pentaho va a necesitar la dirección del archivo excel descargado del sistema financiero “ESIGEF”. Este archivo descargado debe tener en la pestaña el nombre “BASE” Como se identifica en la *Figura 44*.

Figura 43.

CONFIGURACIÓN DEL OBJETO MICROSOFT EXCEL INPUT



Nota: El gráfico representa la configuración de la herramienta según el tipo de dato.

Figura 44.

NOMBRE DE LA HOJA DEL ARCHIVO EXCEL

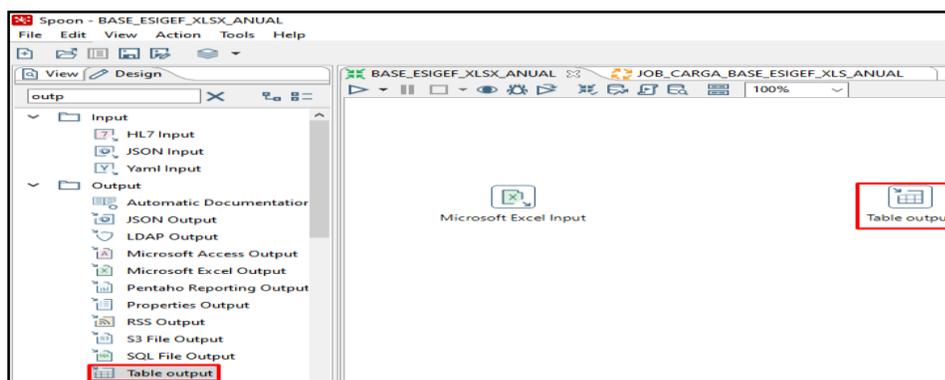
	A	B	C	D	E	F	G	H	I	J	K	L	M
1	DESCRIPCION1	DESCRIPCION1	DESCRIPCION1	DESCRIPCION1	DESCRIPCION1	COL1	COL2	COL3	COL4	COL5	COL6	COL7	COL8
2	PROVISION	FORMACION SIN PROYEC	Servicios Personales por			2200005	2676716,6	4876721,6	0	4614231,33	4614231,33	4614231,33	262490,27
3	DESARROLL	ADMINISTR/ SIN PROYEC	Materiales de Aseo			6836	15411,71	22247,71	5048,54	17199,17	4677,57	4677,57	5048,54
4	GESTION AD	ADMINISTR/ SIN PROYEC	Viaticos y Subsistencias			100563	-63835,15	36727,85	12262,06	24390,35	24390,35	24390,35	12337,5
5	DESARROLL	ADMINISTR/ SIN PROYEC	Maquinarias y Equipos (I			30000	463,39	30463,39	146,05	30317,34	19634,02	19634,02	146,05
6	ADMINISTR	FORMACION SIN PROYEC	Pasajes al Interior			3000	-3000	0	0	0	0	0	0
7	SERVICIOS T	ADMINISTR/ SIN PROYEC	Maquinarias y Equipos (I			17785	60727	78512	1845	76445,6	0	0	2066,4
8	DESARROLL	ADMINISTR/ SIN PROYEC	Insumos- Bienes- Materi			34355	61854,95	96209,95	18116,6	78093,34	29037,34	29037,34	18116,61
9	GESTION	DE FORMACION SIN PROYEC	Energia Electrica			6500	-6500	0	0	0	0	0	0
10	DESARROLL	ADMINISTR/ SIN PROYEC	Edificios- Locales- Reside			2856	-2856	0	0	0	0	0	0
11	GESTION DE	GESTION DE SIN PROYEC	Pasajes al Exterior			0	13000	13000	750,18	11955,42	11955,42	11955,42	1044,58
12	SERVICIO DE	ADMINISTR/ SIN PROYEC	Combustibles - Lubricant			0	0	0	0	0	0	0	0
13	GESTION DI	ADMINISTR/ SIN PROYEC	Mobiliarios (Bienes Mue			0	0	0	0	0	0	0	0
14	GESTION DI	ADMINISTR/ SIN PROYEC	Remediacion - Restaura			0	0	0	0	0	0	0	0
15	DESARROLL	ADMINISTR/ SIN PROYEC	Manejo de Cocina-de Ho			0	0	0	0	0	0	0	0
16	DESARROLL	ADMINISTR/ SIN PROYEC	Herramientas (Bienes Mi			0	466,94	466,94	178,43	288,51	288,51	288,51	178,43
17	PROCURADL	ADMINISTR/ SIN PROYEC	Fiscalizacion e Inspeccio			0	1048,32	1048,32	0	1048,32	1048,32	1048,32	0
18	PROCURADL	ADMINISTR/ SIN PROYEC	Edificios- Locales y Resid			0	27,36	27,36	0	24	24	24	3,36
19	PROVISION	FORMACION SIN PROYEC	A Jubilados Patronales			0	2268,53	2268,53	0	0	0	0	2268,53
20	FASE 2017, C	FORMACION FORTALECIN	Aporte Patri FORTALECIN			0	6092,3	6092,3	0	6092,3	6092,3	6092,3	0
21	CURSO DE N	GESTION DE PROYECTO F	Decimocuar PROYECTO F			0	7458,76	7458,76	0	0	0	0	7458,76
22	INDEMNIZA	ADMINISTR/ INDEMNIZA	Decimocuar INDEMNIZA			0	0	0	0	0	0	0	0
23	PROYECTO F	GESTION DE PROYECTO F	Honorarios PROYECTO F			2759491	-2759491	0	0	0	0	0	0
24	PROYECTO F	GESTION DE PROYECTO F	Infraestruct PROYECTO F			0	0	0	0	0	0	0	0

Nota: El gráfico representa como se obtiene los datos de excel apuntando al nombre del archivo.

Después, de la configuración del objeto de Excel, se selecciona el objeto table output como se muestra en la *Figura 45*. Para almacenar los datos en las tablas Stage.

Figura 45.

CONFIGURACIÓN OBJETO TABLE OUTPUT

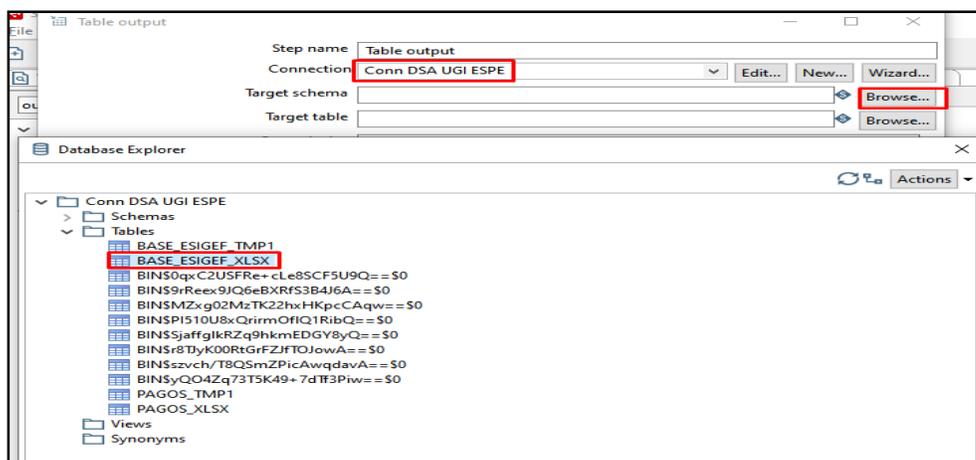


Nota: El gráfico representa la configuración de Pentaho para acceder a los table output.

Dentro del “Table output” se configura los parámetros de conexión y se selecciona la tabla destino de los datos. En la *Figura 46*, se puede identificar este paso.

Figura 46.

CONFIGURACIÓN TABLA DESTINO

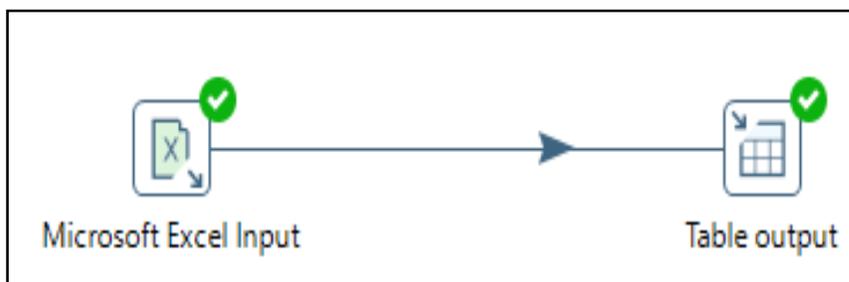


Nota: El gráfico representa la configuración de la tabla destino en Pentaho.

A continuación, se unen los objetos configurados anteriormente y se obtiene la secuencia de descarga de datos de Excel a las tablas Stage. En la *Figura 47*, podemos observar este paso.

Figura 47.

UNIÓN DE OBJETOS



Nota: El gráfico representa la unión de los objetos en Pentaho Data Integration.

Una vez obtenidos los datos en el Stage, se ejecuta las funciones y procedimientos de limpieza de datos mediante sentencias SQL para obtener las columnas de datos depuradas como se observa en la *Figura 48*.

Figura 48.

PROCESOS DE LIMPIEZA DE DATOS.

```

CONVERTIR_CADENA_NUMBER(COL10),
CONVERTIR_CADENA_NUMBER(COL20)/100,
CODIGOG1 ESTRUCTURA, -- ESTRUCTURA
CASE
WHEN LENGTH(CAST(CONVERTIR_CADENA_ENTERO(CODIGOG2) AS VARCHAR(5)))=1
THEN CONCAT('0',CAST(CONVERTIR_CADENA_ENTERO(CODIGOG2) AS VARCHAR(5)))
WHEN LENGTH(CAST(CONVERTIR_CADENA_ENTERO(CODIGOG2) AS VARCHAR(5)))=2
THEN CAST(CONVERTIR_CADENA_ENTERO(CODIGOG2) AS VARCHAR(5))
ELSE 'XX'
END CODIGO_PROGRAMA,
SUBSTR(LIMPIAR_CADENA(CODIGOG3),-3) CODIGO_PROYECTO,
SUBSTR(LIMPIAR_CADENA(CODIGOG1),-3) CODIGO_ACTIVIDAD,
LIMPIAR_CADENA(CODIGOG4) ITEM_PRESUPUESTARIO,
TO_DATE( CONCAT( CONCAT( '01' ,

```

Nota: El gráfico representa el proceso de limpieza de datos con sentencias SQL.

Con la información depurada se ejecutan los “store procedures” que permiten llevar los datos depurados del Stage hacia tablas del schema DWH (modelos dimensionales propuestos) para el módulo de finanzas. En la *Figura 49*, se puede identificar este proceso.

Figura 49.

EJECUCIÓN DE STORE PROCEDURES

```
-- CARGA DE LA TABLA TEMPORAL
CALL CRG_DSA_BASE_ESIGEF_TMP1('DSA');
-- CARGA DE LAS DIMENSIONES
CALL CRG_DWH_DIM_ACTIVIDAD('DWH');
CALL CRG_DWH_DIM_ITEM_PRESUP('DWH');
CALL CRG_DWH_DIM_PROGRAMA('DWH');
CALL CRG_DWH_DIM_PROYECTO('DWH');
-- CARGA DE LA TABLA CENTRAL: FAC_FUENTE_FINANCIAMIENTO
CALL CRG_DWH_FAC_FUENTE_FINANCIAMTO('DWH');
```

Nota: El gráfico representa la limpieza de datos mediante store procedures.

Construcción de los dashboard

Para este segundo componente de nuestro proyecto de titulación vamos a utilizar el software Tableau Desktop instalado con anterioridad y vamos a realizar las siguientes actividades.

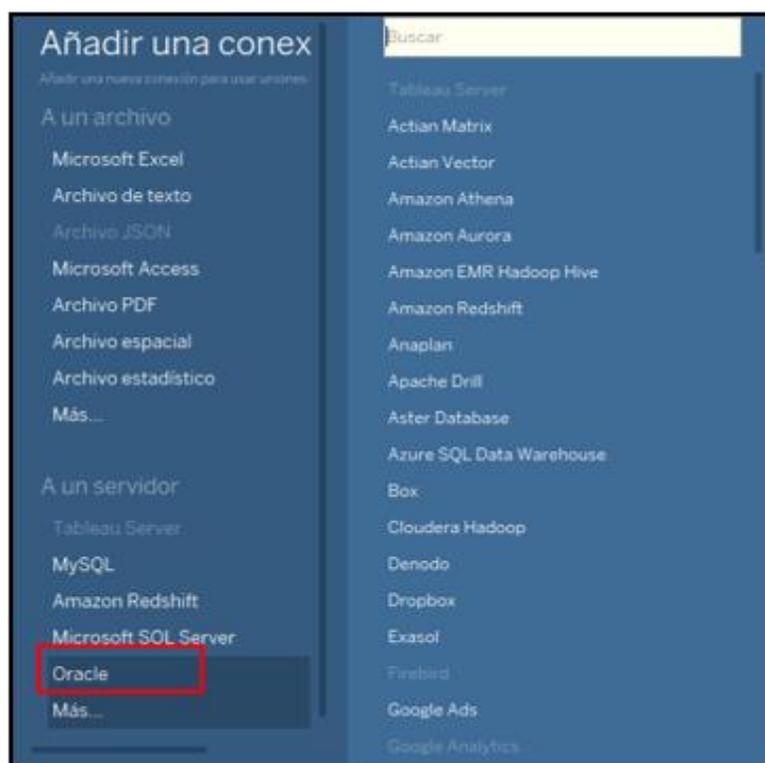
- Creación de Dashboards en Tableau Desktop.

Con Tableau Desktop existe dos formas de realizar la conexión a los datos para generar los dashboard: la primera es mediante una forma predeterminada aprovechando las bondades del software y la segunda es mediante un script personalizado que va a ejecutar lo que nosotros queremos mostrar.

Primera forma: al momento de ejecutar el programa Tableau Desktop se abre una ventana como se muestra en la *Figura 50*, en la cual se procede a añadir una conexión a los datos, en este caso escogemos la opción de conectarnos a un servidor de Oracle.

Figura 50.

CONEXIÓN DE TABLEAU A LA BASE DE DATOS DE ORACLE.

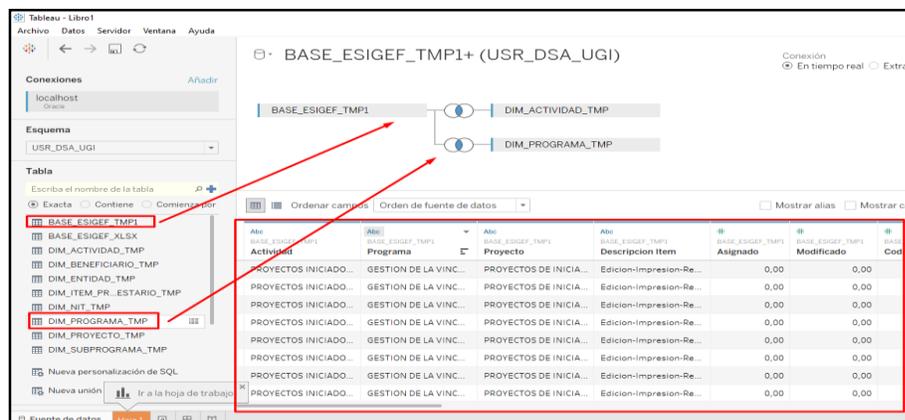


Nota: El gráfico representa la conexión a la herramienta Tableau Desktop.

Después, se abre otra ventana como se muestra en la *Figura 51*, en donde debemos arrastrar las tablas de la base de datos conectada y el software comienza a generar los joins de forma automática. En esta parte se debe tomar en cuenta el criterio del join al momento de unir las dimensiones con los hechos.

Figura 51.

ACCESO A LA INFORMACIÓN DE LA BASE DE DATOS DE ORACLE.

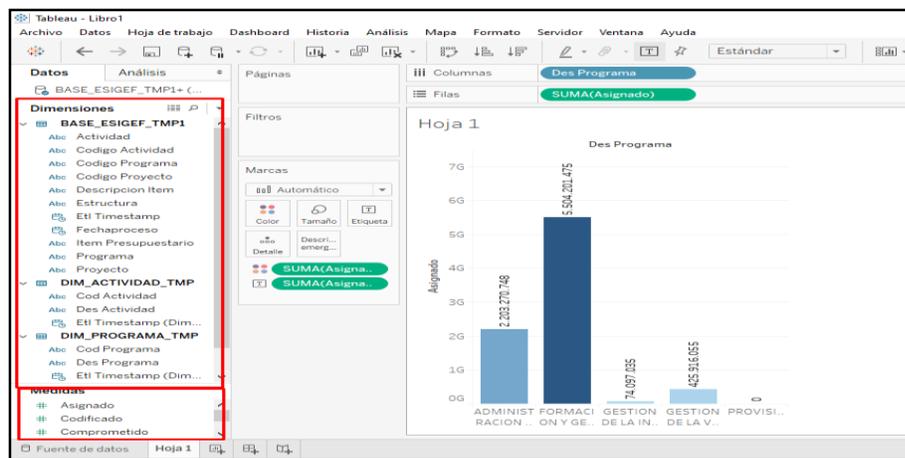


Nota: El gráfico representa el acceso a la información de la base de datos de Oracle

Luego, que la herramienta Tableau tenga acceso a la información financiera de la base de datos de Oracle, se despliega una hoja de trabajo en la cual están disponibles las dimensiones y medidas para realizar los cálculos estadísticos de los indicadores de resultados financieros. En la Figura 52, podemos observar la generación de estos dashboards.

Figura 52.

CREACIÓN DE DASHBOARD FINANCIEROS.

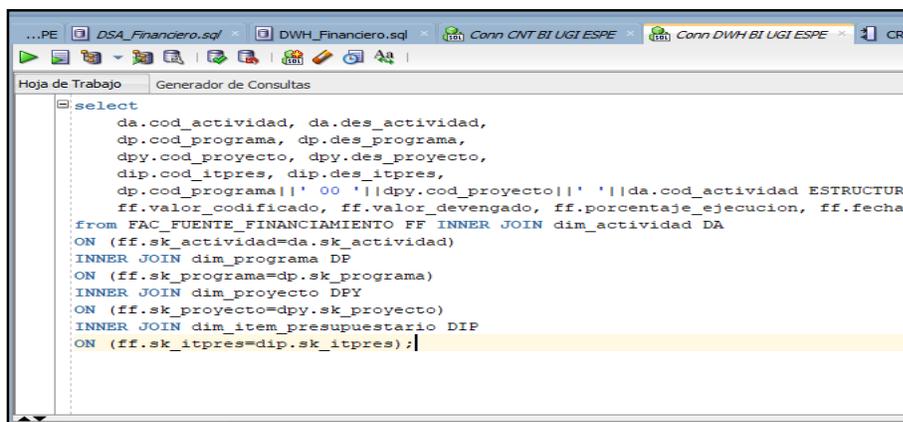


Nota: El gráfico representa la creación de los dashboard financieros en Tableau.

Segunda forma: es la forma más adecuada porque se va a mapear los datos mediante un script sql personalizado como se muestra en la *Figura 53*.

Figura 53.

SCRIPT SQL PERSONALIZADO



```

select
da.cod_actividad, da.des_actividad,
dp.cod_programa, dp.des_programa,
dpy.cod_proyecto, dpy.des_proyecto,
dip.cod_itpres, dip.des_itpres,
dp.cod_programa||' 00 '||dpy.cod_proyecto||' '||da.cod_actividad ESTRUCTURA
ff.valor_codificado, ff.valor_devengado, ff.porcentaje_ejecucion, ff.fechaap
from FAC_FUENTE_FINANCIAMIENTO FF INNER JOIN dim_actividad DA
ON (ff.sk_actividad=da.sk_actividad)
INNER JOIN dim_programa DP
ON (ff.sk_programa=dp.sk_programa)
INNER JOIN dim_proyecto DPY
ON (ff.sk_proyecto=dpy.sk_proyecto)
INNER JOIN dim_item_presupuestario DIP
ON (ff.sk_itpres=dip.sk_itpres);

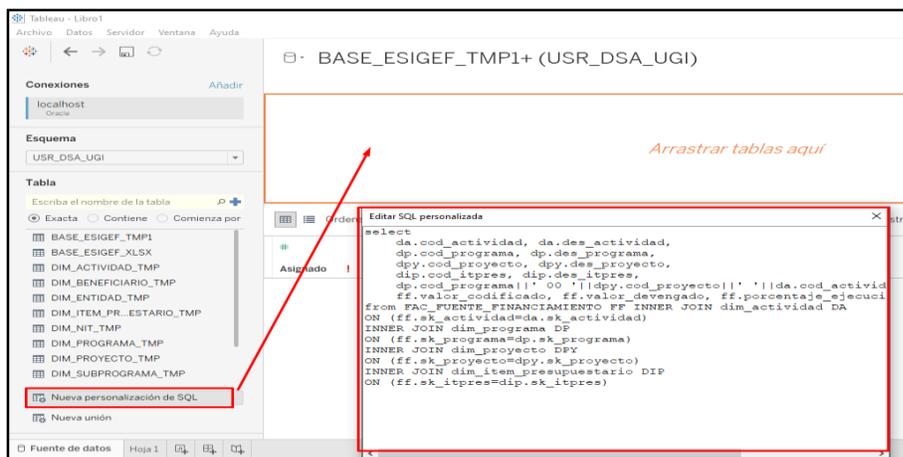
```

Nota: El gráfico representa el acceso a los datos mediante sentencias SQL.

Luego de conectarnos a la base de datos de Oracle se abre una ventana en donde se va a elegir la opción “Nueva personalización de SQL”, y vamos a pegar y ejecutar el script como se muestra en la *Figura 54*.

Figura 54.

NUEVA PERSONALIZACIÓN DE SQL.

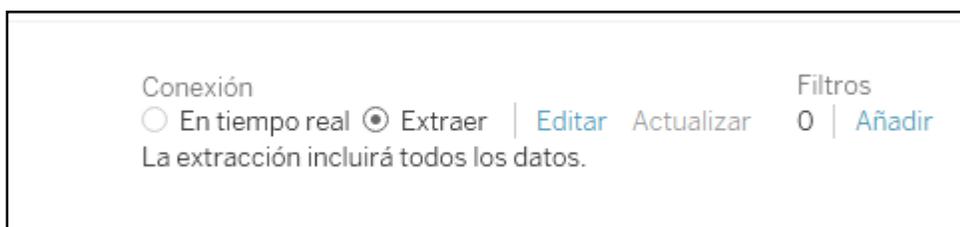


Nota: El gráfico representa la personalización de datos mediante SQL.

Se recomienda que la conexión sea de tipo extracto para no saturar con consultas demasiado pesadas a la base de datos como se observa en la *Figura 55*. Esto quiere decir que Tableau realiza de forma automática una consulta a la base de datos de toda la información y guarda una copia en un archivo propio de Tableau con extensión tipo hyper. De donde va a tomar los datos para presentar en los dashboard, como se muestra en la *Figura 56*.

Figura 55.

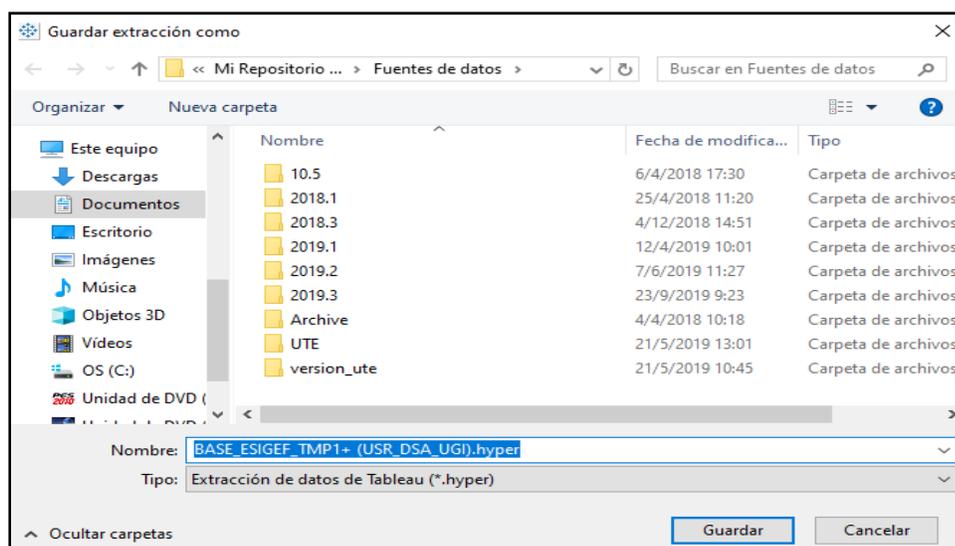
EXTRACCIÓN DE DATOS DE TABLEAU



Nota: El gráfico representa el proceso de extracción de datos con Tableau.

Figura 56.

COPIA DE DATOS EN ARCHIVO TIPO HYPER.

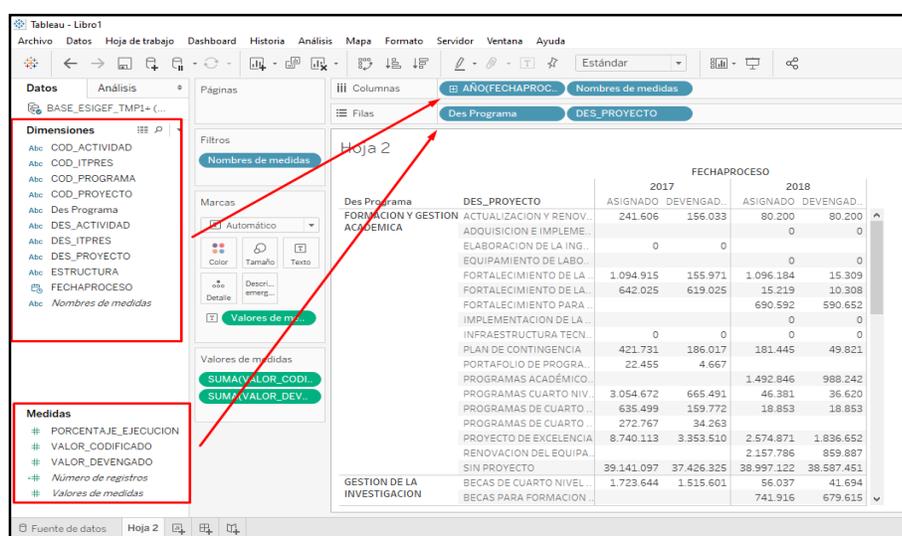


Nota: El gráfico representa como Tableau guarda una copia de las fuentes de datos.

Luego de la extracción de datos, se abre una ventana principal que despliega las herramientas disponibles de Tableau para el análisis de los datos de las actividades de Investigación de acuerdo con los indicadores de resultados planteados anteriormente. En la *Figura 57*, se puede identificar estas herramientas.

Figura 57.

HERRAMIENTAS DE TABLEAU.



Nota: El gráfico muestra los recursos de Tableau Desktop disponibles

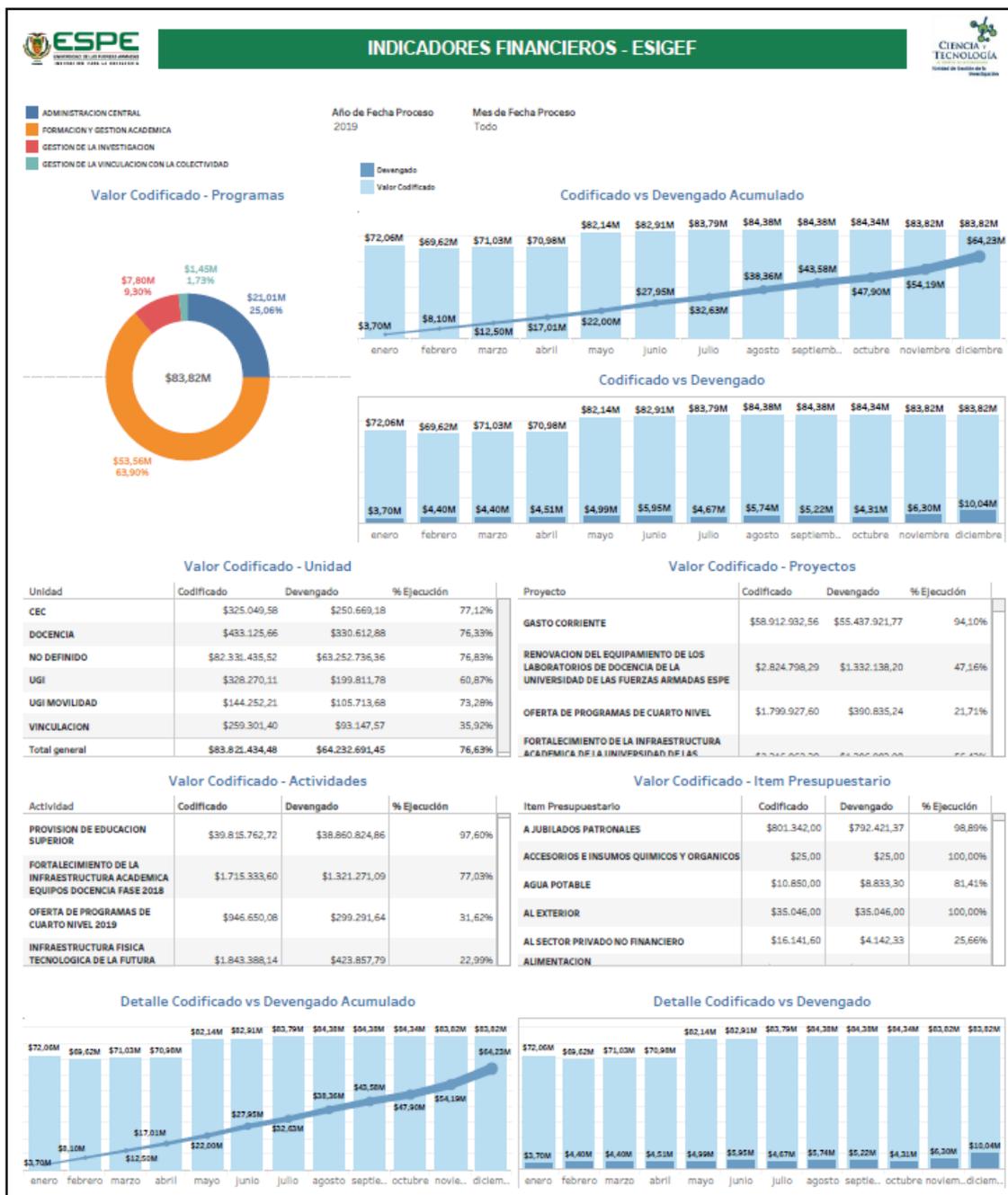
- Resultados de los dashboard.

A continuación, se muestra los resultados de los indicadores obtenidos en los dashboard. En la *Figura 58*, tenemos el resultado de los indicadores financieros. En la *Figura 59*, se muestra los resultados de los indicadores de proyectos de investigación. En la *Figura 60*, se observa los resultados de los indicadores de movilidad y por último en la *Figura 61*, se encuentran los resultados de los indicadores de difusión.

- Resultados de las actividades financieras.

Figura 58.

RESULTADOS DE LOS INDICADORES FINANCIEROS

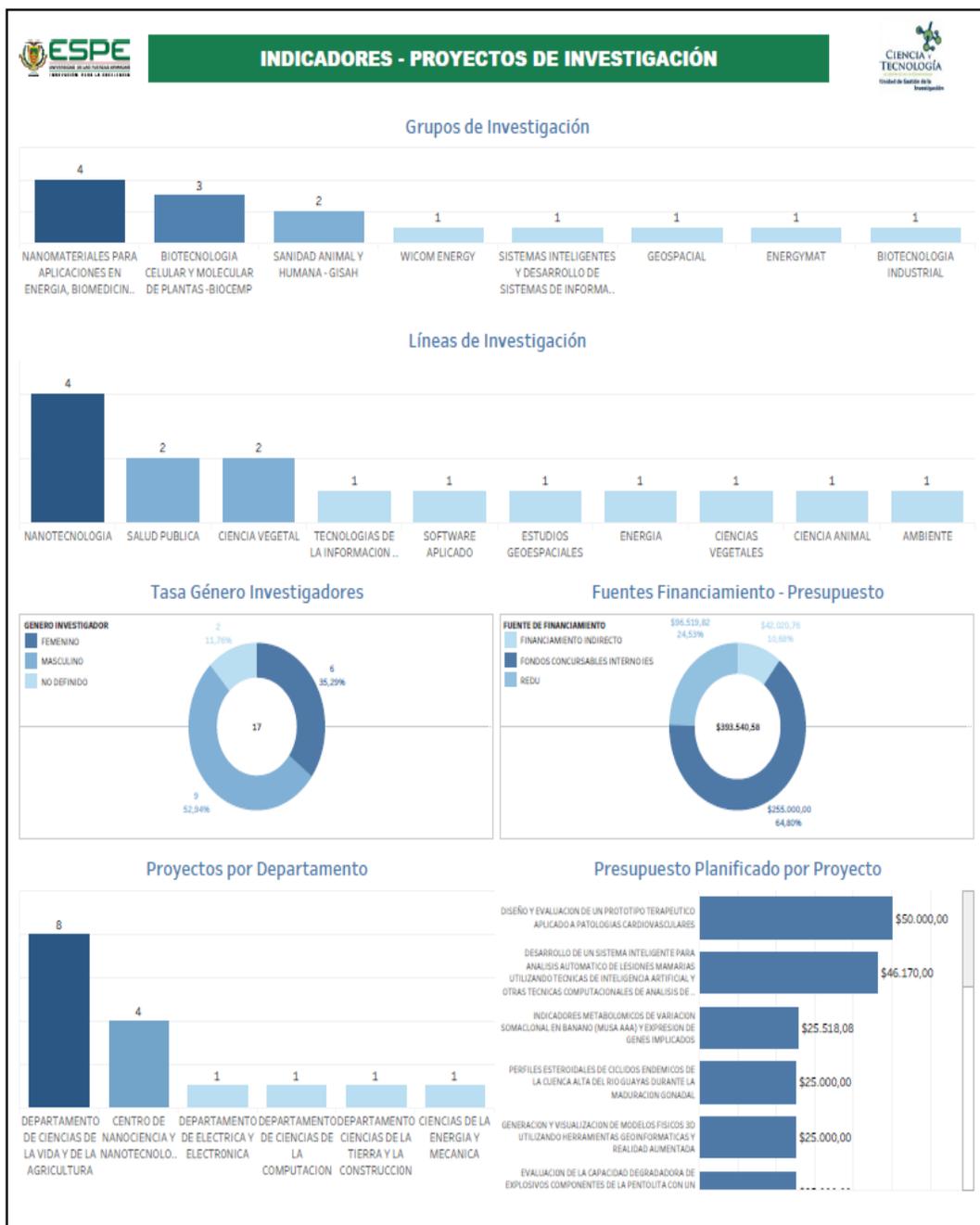


Nota: El gráfico representa los dashboard de resultados de los indicadores financieros.

- Resultados de las actividades de los proyectos de investigación.

Figura 59.

RESULTADOS DE LOS PROYECTOS DE INVESTIGACIÓN

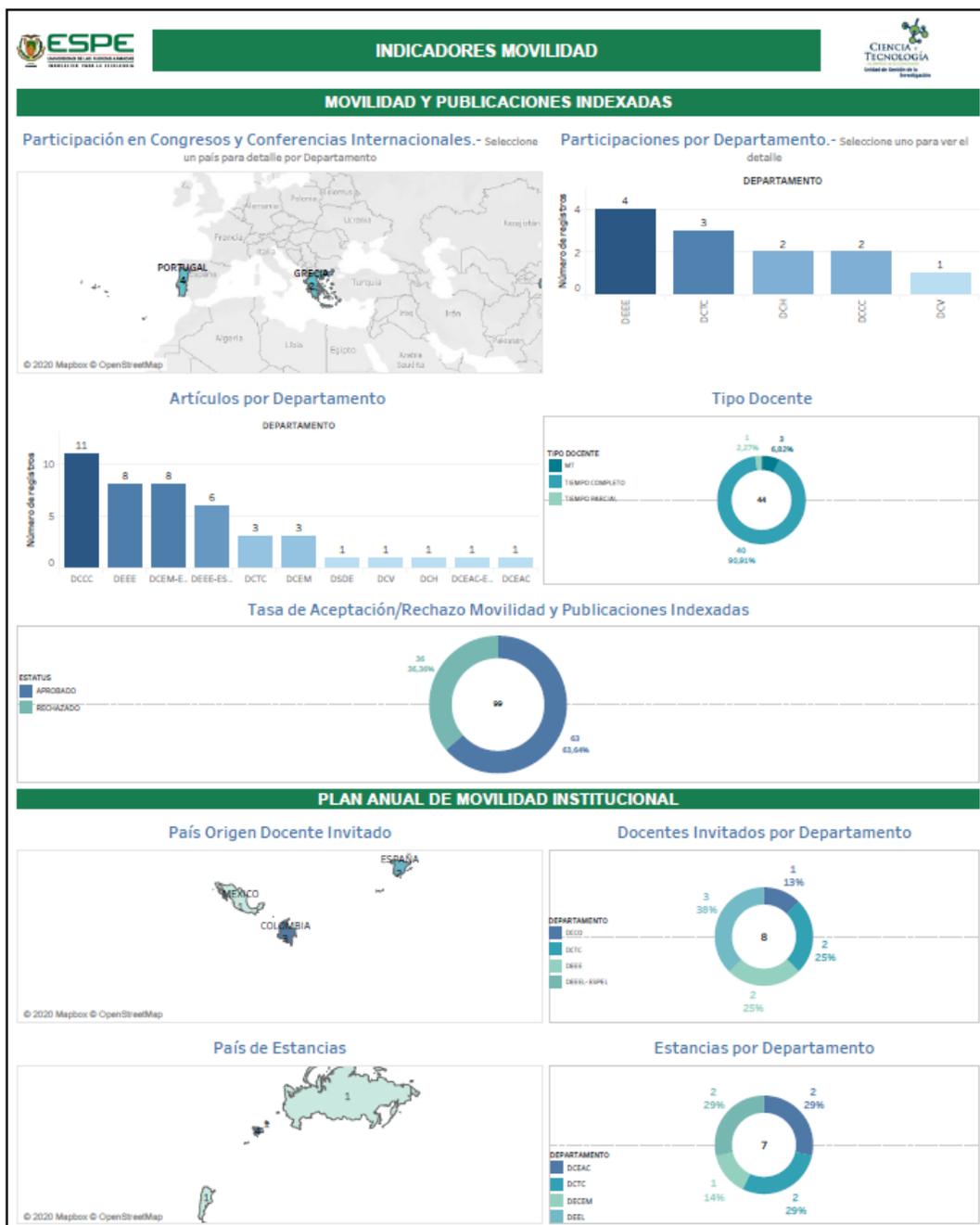


Nota: El gráfico representa los dashboard de resultados de los indicadores de proyectos de investigación

- Resultados de las actividades de movilidad.

Figura 60.

RESULTADOS DE LOS INDICADORES DE MOVILIDAD

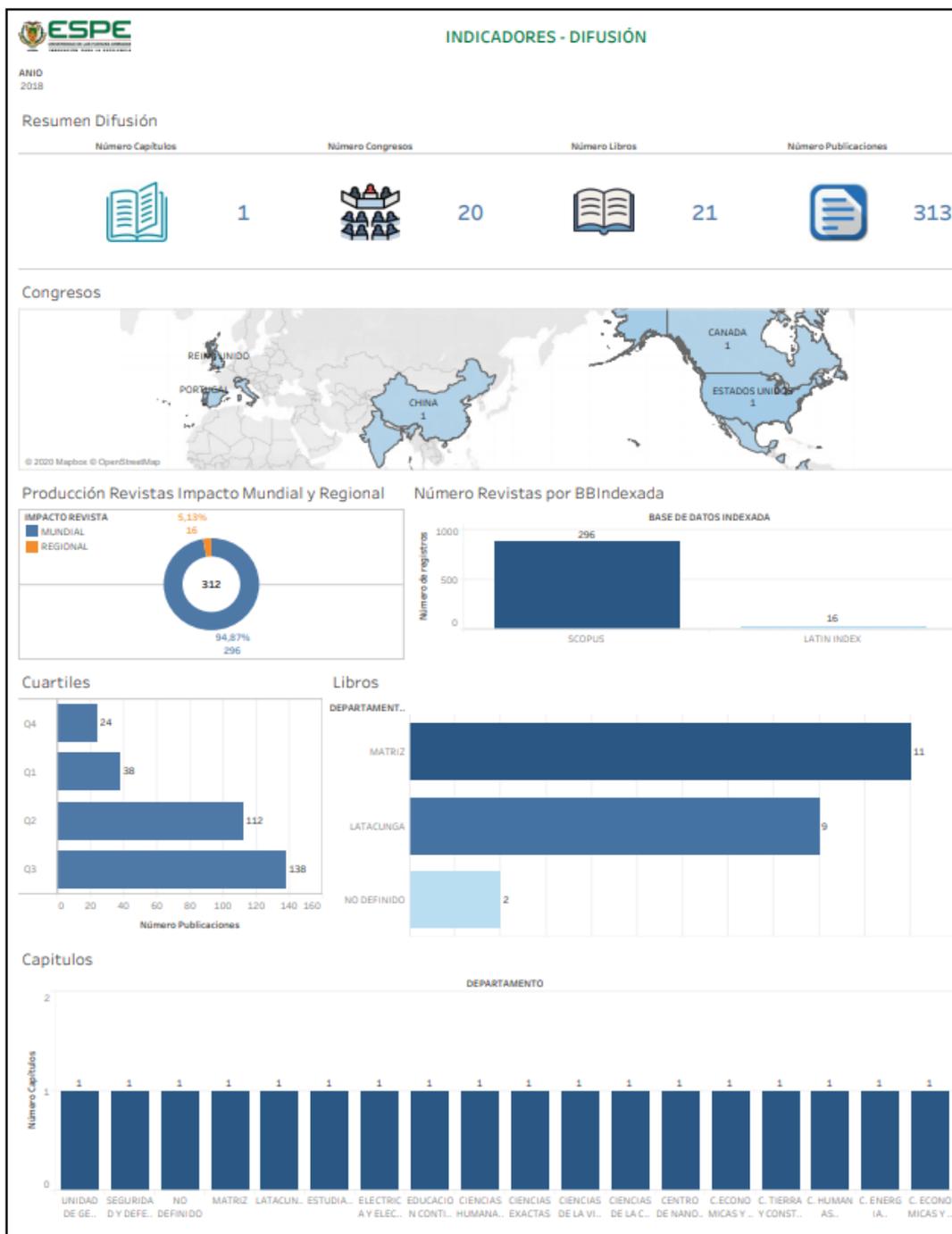


Nota: El gráfico representa los dashboard de resultados de los indicadores de movilidad.

- Resultados de las actividades de difusión.

Figura 61.

RESULTADOS DE LOS INDICADORES DE DIFUSIÓN.



Nota: El gráfico representa los dashboard de resultados de los indicadores de difusión.

CAPÍTULO V

PRESENTACIÓN DE LOS RESULTADOS

En el presente capítulo, se va a realizar una presentación de los resultados generados en los dashboard. Estos resultados se evalúan en forma conjunta con cada uno de los responsables de las actividades de investigación de la UGI.

Validación de datos

Esta parte comprende la verificación y validación de los resultados de los dashboard, por tal motivo se va a utilizar un procedimiento de recolección de datos como es la encuesta. En la encuesta se ha planteado preguntas cerradas para que sean completadas por los encargados de las actividades de investigación de la UGI y usuarios generales, de esta forma obtendremos resultados fiables.



ENCUESTA DE SATISFACCIÓN DEL CLIENTE

La presente encuesta tiene la finalidad de conocer el criterio acerca del uso de Tableros de control (Dashboard) y la importancia de utilizar Business Intelligence en el VIITT de la UFA-ESPE. En estos tableros están definidos los indicadores de las actividades de investigación en forma gráfica de acuerdo con cada área. Favor complete la siguiente encuesta sobre los resultados de los dashboard. Coloque una "X" en el casillero de su preferencia.

ORD.	PREGUNTAS	SI	NO
1	¿El sistema implementado es amigable con el usuario?		

2	¿Los colores implementados en el dashboard captan la atención de los usuarios?		
3	¿Los dashboards implementados están relacionados con la acreditación del CACES?		
4	¿Es clara y entendible la manera en que muestran la información los dashboards?		
5	¿La información que brinda el dashboard está acorde a los procesos y resultados de la UGI?		
6	¿Los gráficos de los dashboard son los adecuados para expresar los resultados de los indicadores implementados?		
7	¿Los resultados de los dashboard proporcionan información útil para la toma de decisiones?		
8	¿Los dashboard son dinámicos?		
9	¿Es necesario contar con una base de datos centralizada para almacenar la información de la UGI?		
10	¿La unidad de gestión de la investigación cuenta con un área de Business Intelligence?		

Gracias por su colaboración.

Resultados de la validación de datos

Después de aplicar las encuestas a los responsables de las actividades de investigación de la UGI y a otros usuarios relacionados con el tema investigado, se recuperó las respuestas de las encuestas como se puede observar en la *Tabla 36*.

Tabla 36.*RESULTADOS DE LAS ENCUESTAS DEL MODELO DE BI IMPLEMENTADO.*

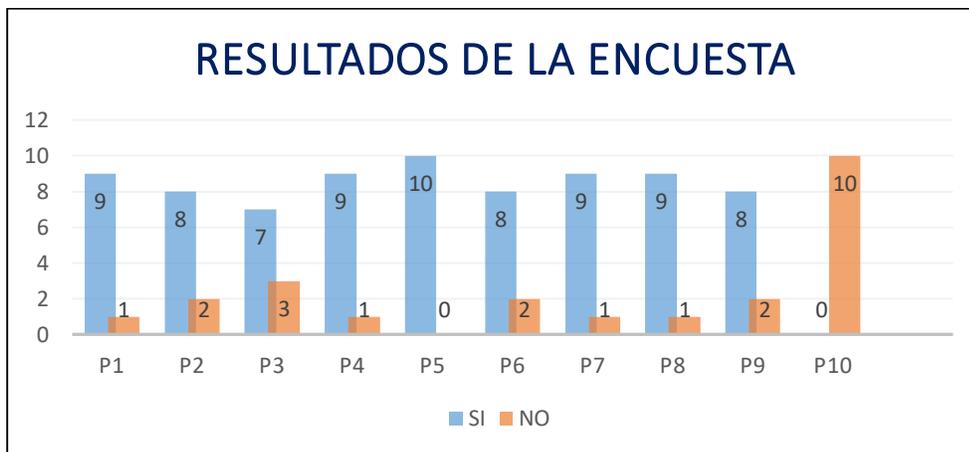
PREGUNTAS	SI	NO
Pregunta 1	9	1
Pregunta 2	8	2
Pregunta 3	7	3
Pregunta 4	9	1
Pregunta 5	8	2
Pregunta 6	8	2
Pregunta 7	9	1
Pregunta 8	9	1
Pregunta 9	8	2
Pregunta 10	0	10

Nota: Esta tabla muestra los resultados de las encuestas realizadas a los responsables de las actividades de Investigación y a los usuarios finales.

Luego se procesó estos datos mediante un esquema gráfico y se realizó un análisis de las preguntas evaluadas. En la *Figura 62*, se identifica el resultado sobre el uso de los dashboards.

Figura 62.

RESULTADOS DE LA ENCUESTA DEL MODELO DE BI IMPLEMENTADO.



Nota: El gráfico representa los resultados de la encuesta realizada a los usuarios.

Los resultados obtenidos de las encuestas indican que desde la pregunta 1 a la 9 existe una gran aceptación por parte de los usuarios como se identifica en la *Figura 63*, lo que representa un 86 % de aceptación del sistema de BI implementado, por lo tanto se puede ir ajustando el sistema de acuerdo con las recomendaciones por parte de los usuarios finales.

Figura 63.

GRADO DE ACEPTACIÓN DEL SISTEMA DE BI



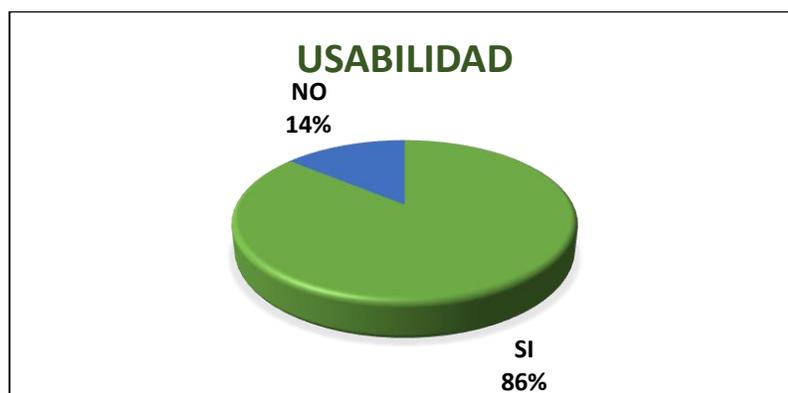
Nota: El gráfico representa el resultado del grado de aceptación del usuario.

En cambio en la pregunta 10 la totalidad de los usuarios encuestados afirman que la UGI no cuenta con un área de Business Intelligence. Esta afirmación nos permite concluir que es necesario la implementación de un área de BI en el VIITT, por la gran importancia que representa para el control y análisis de las actividades de investigación.

Las preguntas 1, 2, 4, 6 y 8 están relacionadas con el funcionamiento y la usabilidad de los dashboards en donde la aceptación es de un 86% por parte de los usuarios como se puede identificar en la *Figura 64*. Por lo tanto, se puede considerar que es de gran utilidad la implementación de un modelo de Business Intelligence, enfocado a la toma de decisiones de la gestión de investigación del vicerrectorado de investigación, innovación y transferencia de tecnología de la UFA-ESPE debido a la satisfacción del cliente al usar este sistema.

Figura 64.

USABILIDAD DEL MODELO DE BI IMPLEMENTADO



Nota: El gráfico representa la usabilidad del sistema ante los usuarios finales.

Por otro lado, la pregunta 3 está relacionado a una normativa vigente como es el CACES y enfocado en la acreditación de la universidad. En esta pregunta tenemos que un 70 % de los usuarios están de acuerdo que los indicadores de los dashboard están

apegados a la acreditación y un 30% que no están de acuerdo. Esto se debe a que los indicadores financieros son más informativos sobre el valor asignado y gastado, por lo tanto, se puede considerar aspectos para que los indicadores financieros se apeguen más hacia la acreditación.

Los resultados de la pregunta 5, están relacionados a ciertos procesos del VIITT y tienen un 100 % de aceptación de los usuarios. Esto es de vital importancia porque demuestra que el sistema de BI implementado cuenta con los mismos resultados de los procesos de las actividades de investigación que llevan a cargo cierto usuarios. Por esta razón podemos estar seguros de que los valores reflejados en los dashboard son los correctos.

Otra pregunta muy importante es la número 7, porque está relacionado con los resultados proporcionados por los dashboard. Estos resultados son producto del análisis de datos automáticos que realiza la herramienta de BI, que está configurado a las necesidades del VIITT. Esta pregunta tiene una aprobación del 90% por parte de los encuestados, lo que permite concluir que la información de los reportes es confiable y facilita una mejor toma de decisiones por la rapidez de los resultados. Por otro lado el 10 % que no está seguro en confiar en los resultados de los dashboard, es porque desconfían de las herramientas tecnológicas por desconocimiento de su funcionamiento.

Por último, la pregunta 9 tiene una aceptación del 80% por parte de los encuestados, lo que nos permite interpretar que es necesario tener una base de datos centralizada para almacenar la información de la UGI, por las ventajas que esto proporciona al momento de analizar los datos. En cambio el 20 % restante prefiere tener su propia información almacenada de forma independiente.

CONCLUSIONES Y RECOMENDACIONES

Conclusiones

- De acuerdo con los resultados obtenidos, se puede concluir que se lograron alcanzar los objetivos propuestos del presente proyecto de titulación por las siguientes razones: se implementó los datamarts para almacenar los datos depurados de las actividades de investigación de la UGI. Y en base a estos datos depurados se pudo ejecutar los diferentes dashboards que son los indicadores de resultados enfocados en la acreditación del CACES y por último con estos reportes se logró evaluar con el usuario final los resultados obtenidos
- El uso de la metodología de Kimball es la adecuada para la implementación de almacenes de datos pequeños y medianos como son los datamarts dentro de proyectos a corto plazo. Su metodología reduce la complejidad del presente trabajo de titulación al dividir en etapas su desarrollo. Esto permite tener una estructura simple de los datamarts para que admitan consultas sencillas de los datos almacenados.
- La centralización de datos en un solo repositorio depurado permite tener una única verdad de la información pese a que se manejen diferentes fuentes de datos permite analizar la misma de forma objetiva.
- Los indicadores de resultados enfocados en el CACES y plasmados en los dashboards, permiten medir la calidad de la investigación de la UFA-ESPE. Estos reportes gráficos de los resultados de las actividades de investigación ayudan en la toma de decisiones y en el control por parte del VIITT.
- En el presente proyecto de grado, se demuestra la factibilidad de aplicar herramientas de BI en la administración integral de las actividades de investigación de la UGI. Por el motivo que en el VIITT no existe un área de Business Intelligence

que lleve los reportes automáticos y actualizados constantemente para medir el progreso de las actividades de investigación de una forma eficiente y sobre todo por la gran aceptación y grado de satisfacción de los usuarios finales.

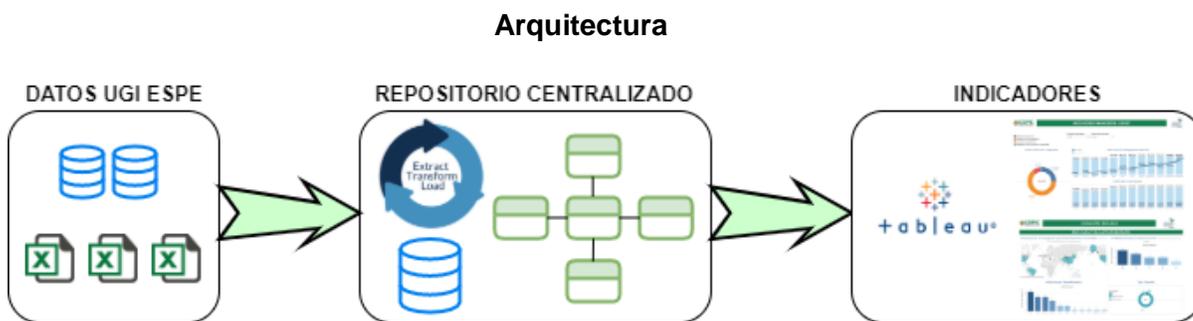
- Los beneficios del análisis, diseño e implementación de un modelo de Business Intelligence, enfocado a la toma de decisiones de la gestión de investigación del vicerrectorado de investigación, innovación y transferencia de tecnología de la UFA-ESPE, se obtendrán a corto plazo, reflejando ventajas competitivas, identificando oportunidades para mejorar de forma constante y generando valor agregado dentro de la UGI.

Recomendaciones

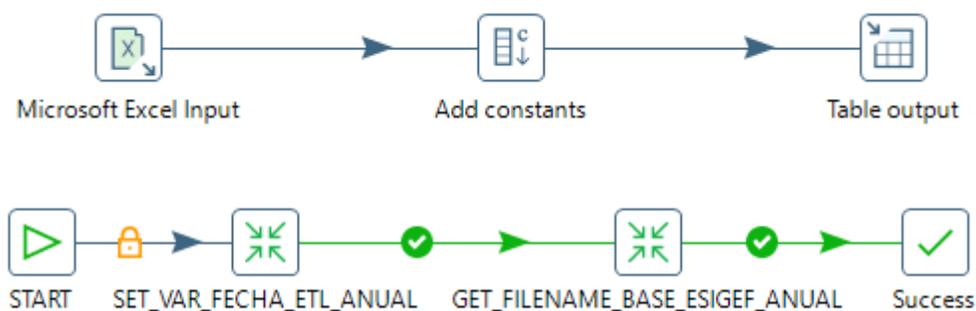
- Se recomienda el uso de herramientas de Business Intelligence que se ajusten a las necesidades de la UGI para aprovechar las bondades de estas herramientas tecnológicas, en busca de generar conocimiento y estrategias de calidad para la mejora continua del área de investigación del VIITT de la UFA-ESPE.
- Es recomendable tener dashboards concisos y precisos a fin de optimizar tiempo en el análisis de la información de las actividades de investigación siendo de gran ayuda en la toma de decisiones del VIITT.
- Los procesos ETL de las diferentes fuentes de datos deben integrarse de tal manera que se pueda contar con información de primera mano actualizada, depurada y almacenada correctamente en los datamarts cuando se necesite realizar los diferentes tipos de análisis.
- Es necesario que los indicadores enfocados en la acreditación del CACES se mejoren de forma constante de acuerdo con la realidad presente para que puedan cumplir con su objetivo.

- Se recomienda que la UFA-ESPE tenga un área propia de Business Intelligence con personal capacitado en el tema para que proporcione la información actualizada y depurada para la toma de decisiones y sobre todo para que el mismo sea el responsable de implementar nuevas reglas de negocio a lo largo del tiempo.

ANEXOS

MANUAL DE USUARIO MODELO DE BUSINESS INTELIGENCE**Ingesta de Datos hacia el repositorio centralizado**

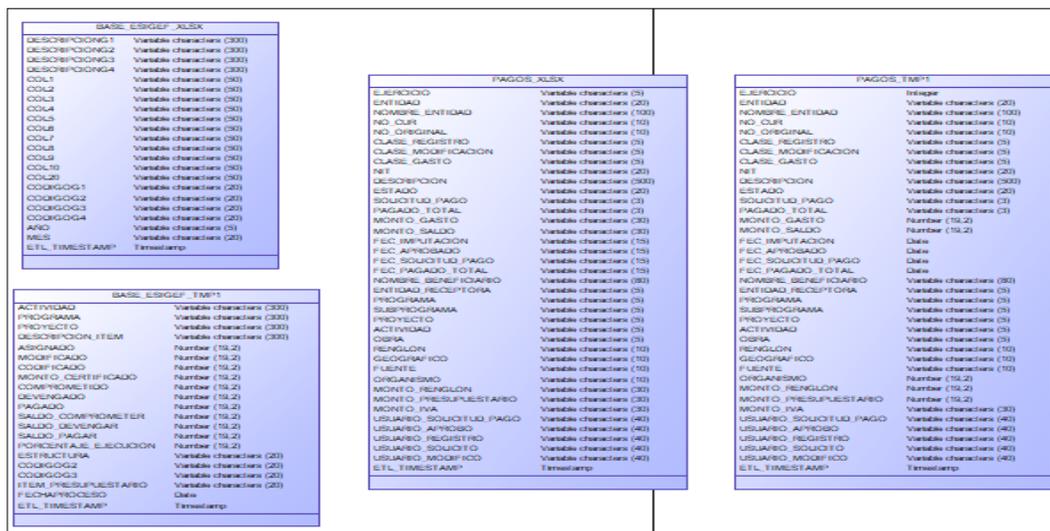
La ingesta de datos de los diferentes archivos que se manejan a nivel interno en la UGI ESPE se realiza mediante procesos ETL realizados en la herramienta Pentaho Data Integration, a continuación se presenta una serie de imágenes donde se puede observar la forma en que se cargaran los datos de los diferentes archivos finales de los usuarios.



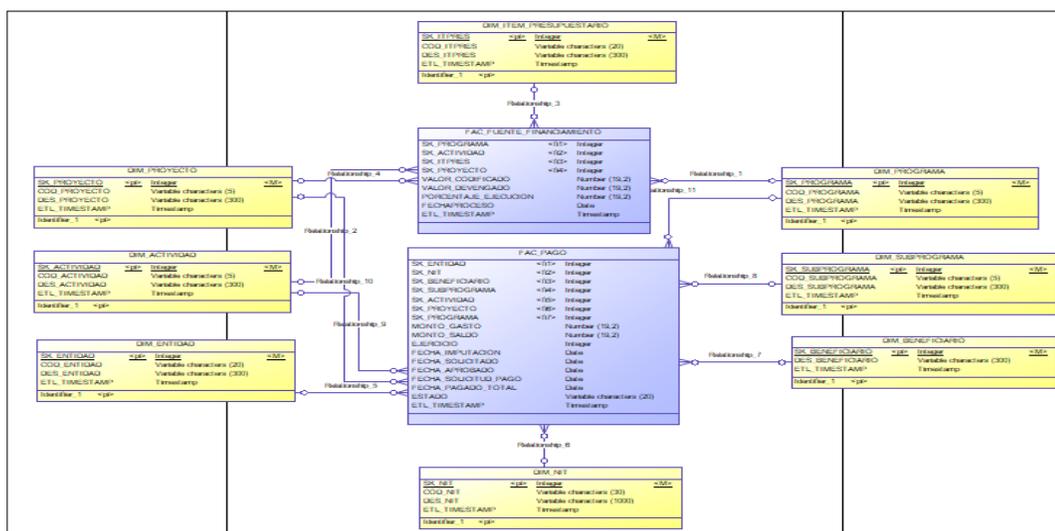
La ejecución de transformaciones y jobs se lo realiza mediante procesos calendarizados y bajo demanda de tal manera que los datos se encuentran cargados y disponibles en el

repositorio centralizado (Base de Datos Oracle), en la cual se definió una serie de modelos dimensionales que almacenan los datos en modelos estrella que facilita el análisis de datos históricos, que a su vez previamente se depuran mediante código SQL con Procedimientos Almacenados desde un schema DSA hacia un DWH para tener datos íntegros y de calidad para el análisis visual en Tableau.

Modelo de Datos en el Schema DSA



Modelo de Datos en el Schema DWH

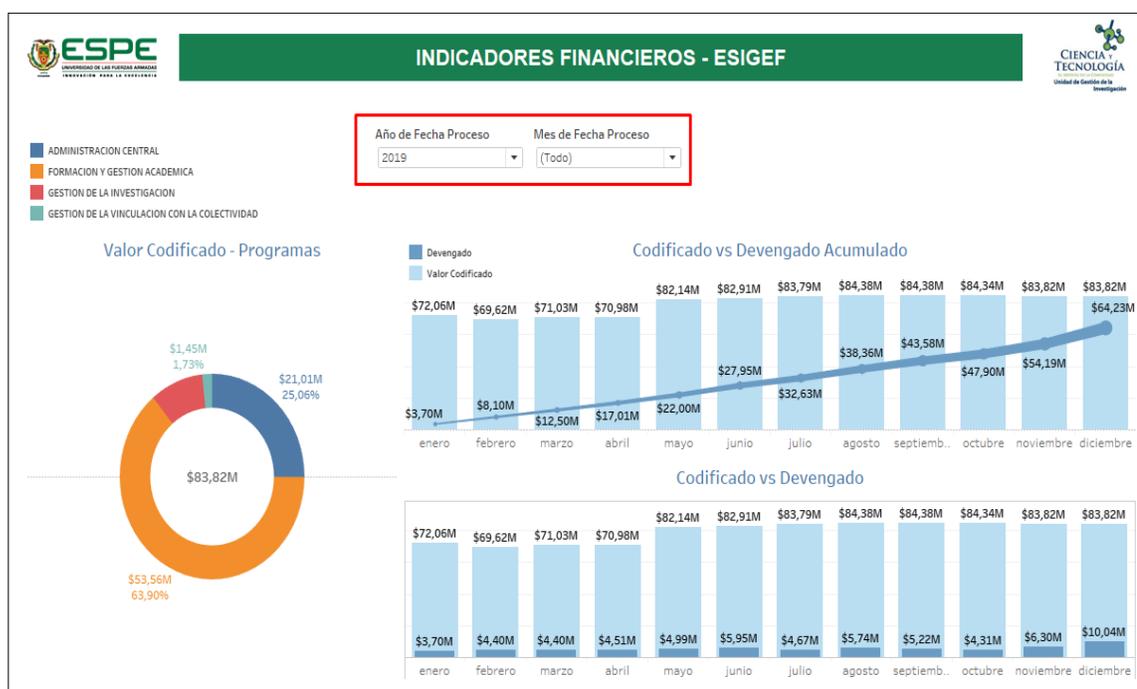


Visualización de Indicadores con Tableau

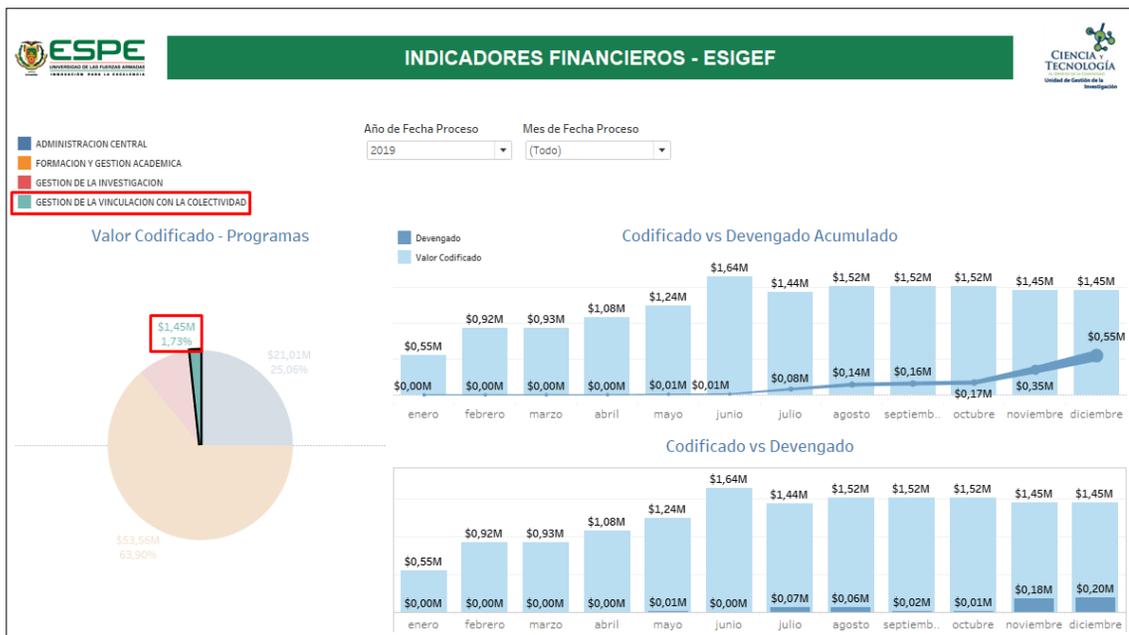
Para la visualización de indicadores se lo realizará con la herramienta de Business Intelligence Tableau Software, misma que facilita a los usuarios finales el descubrimiento de contenido visual mediante representaciones fáciles de interpretar por su facilidad de uso. Para esto se tiene cuatro módulos en los cuales contienen al menos 5 indicadores, a continuación se adjunta imágenes de los dashboards:

Módulo Finanzas

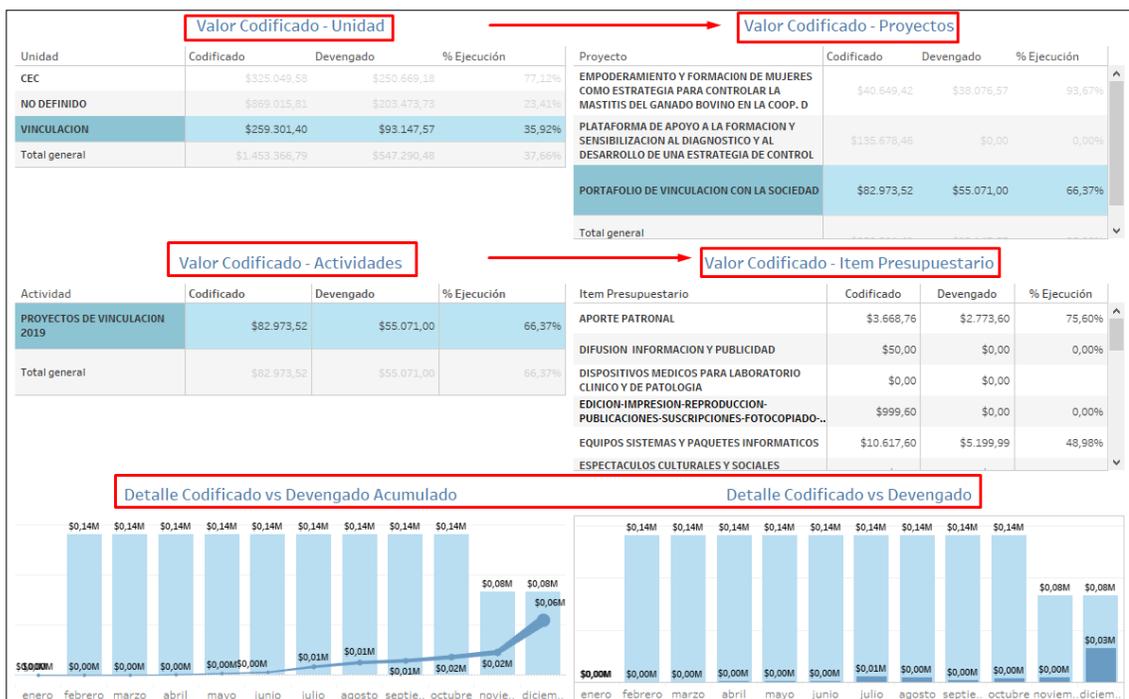
En el módulo de finanzas se toman los datos que tienen como base el sistema ESIGEF presentando los datos de la siguiente manera en el dashboard a continuación:



El módulo financiero cuenta con filtros de año y mes que facilitan el análisis de los datos por cortes de fechas, también se cuenta con la opción de interactuar con el Dashboard pinchando sobre un Programa que se tenga interés en analizar, con lo cual cambian las cifras de acuerdo con el seleccionado.

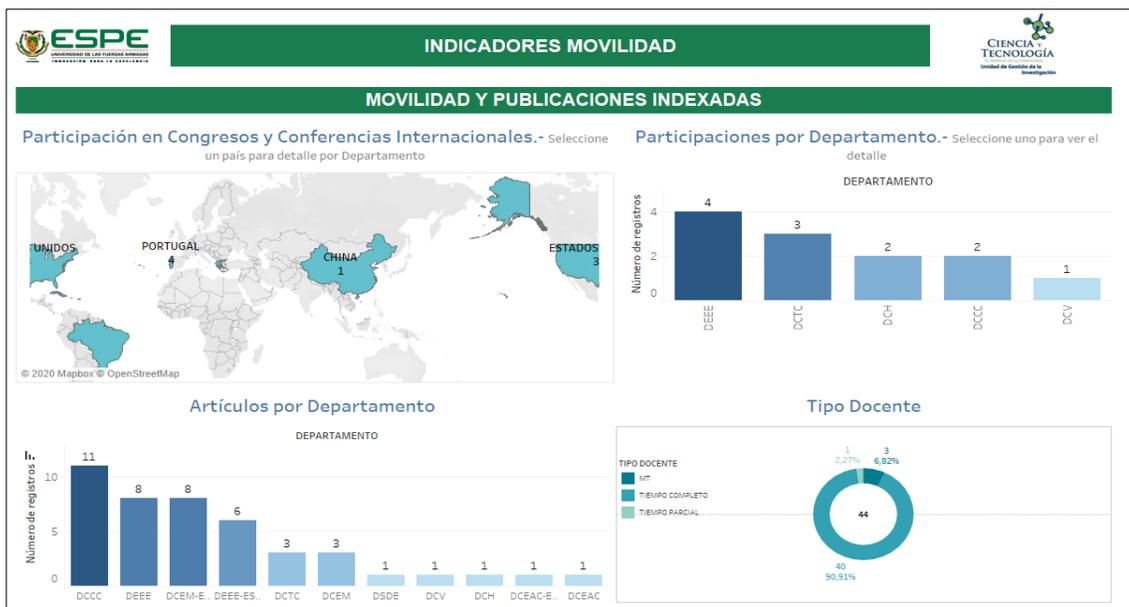


De la misma manera se puede interactuar con el Dashboard de tal manera que es posible revisar los datos de forma escalonada en los diferentes niveles de jerarquía que se presenta a continuación.

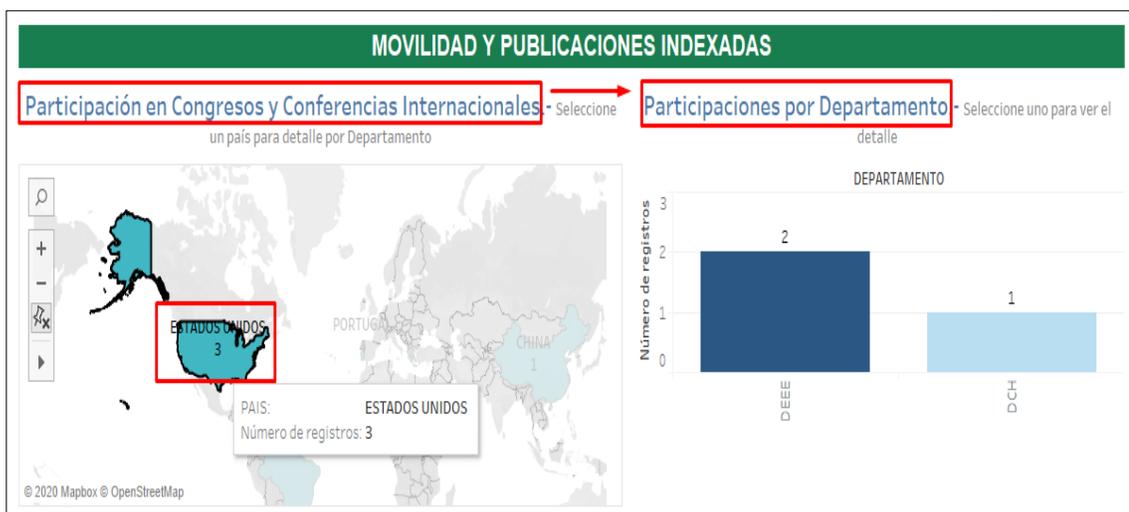


Módulo Movilidad

Para el módulo de movilidad se toma como base el archivo Excel que maneja el usuario de negocio, el mismo se adjunta a continuación una captura de pantalla.



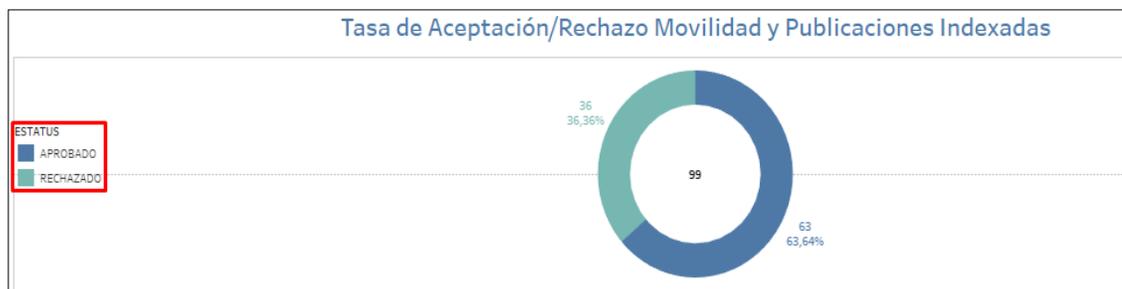
El Dashboard se divide en 2 grandes bloques, el primero trata acerca de “Movilidad y Publicaciones Indexadas” donde se tienen la interacción con el Mapa en Tableau para tener el dinamismo con los siguientes gráficos como se presenta a continuación:



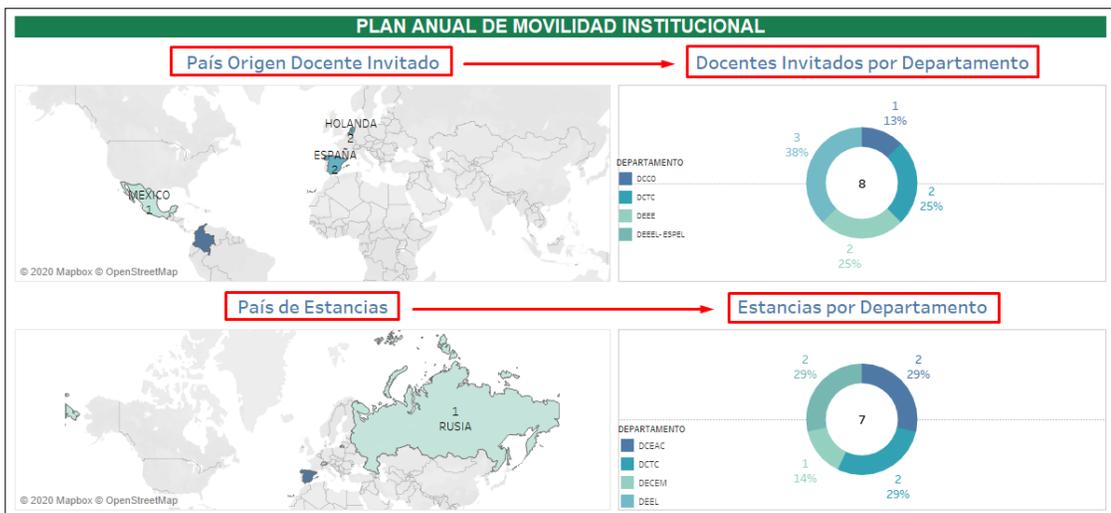
Se presentan también los artículos por Departamento ordenados de forma descendente para conocer cuál es que más aportado científicamente y a su vez se cuenta con un gráfico donde se puede ver la composición de los Tipos de Docente.



Y la tasa de aceptación - rechazo, lo cual es importante ya que permite tener conocimiento acerca del porcentaje de publicaciones que se aceptan o rechazan del total de solicitudes.



El segundo bloque del Dashboard trata acerca del "Plan Anual de Movilidad Institucional", donde se tiene el análisis de los Docentes Invitados de forma que se puede ver los datos de forma geográfica mediante un mapa y la composición por Departamento, así como también, de las Estancias de igual manera de forma geográfica y por Departamento.

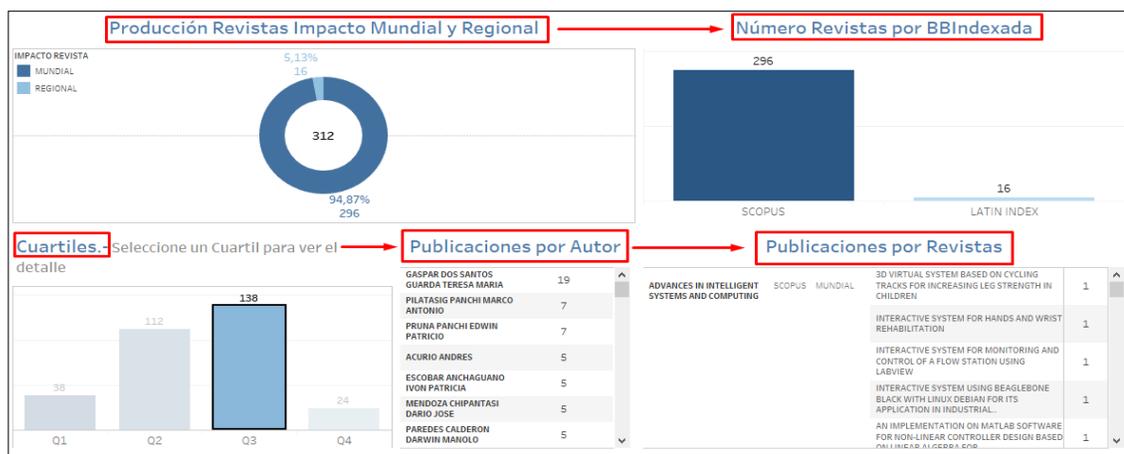


Módulo Difusión

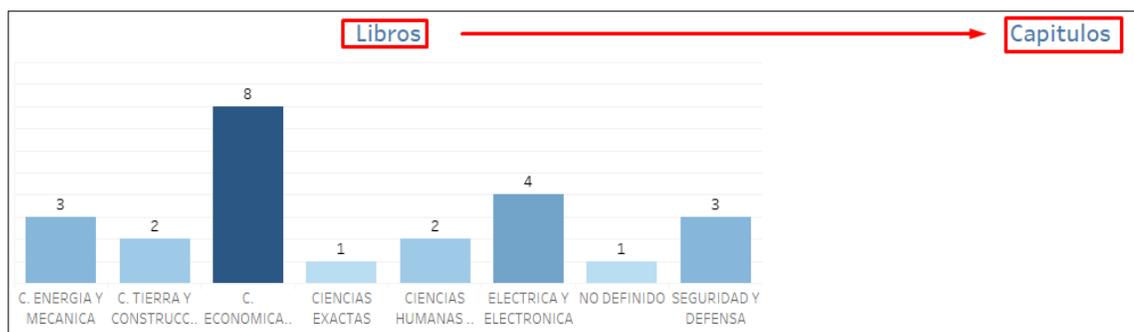
Para el módulo de difusión se toma como base el archivo Excel que maneja el usuario de negocio, donde se tiene la interacción directa con el dashboard mediante un filtro de año para poder analizar la información por corte de fechas. El mismo se adjunta a continuación una captura de pantalla:



Dentro de este módulo se maneja datos de publicaciones, congresos, libros y capítulos; mismos que se encuentran distribuidos de la siguiente manera: en primera instancia tenemos los congresos los cuales se pueden ver los datos de forma geográfica en un mapa como se mostraba anteriormente. En la interacción directa con el Dashboard se puede realizar el análisis de acuerdo con la Producción de Revistas en impacto Mundial y Regional y por Base Indexada y a su vez interactuar con los cuartiles registrados para conocer los autores asociados al cuartil seleccionado y las publicaciones en las diferentes revistas como se muestra a continuación.

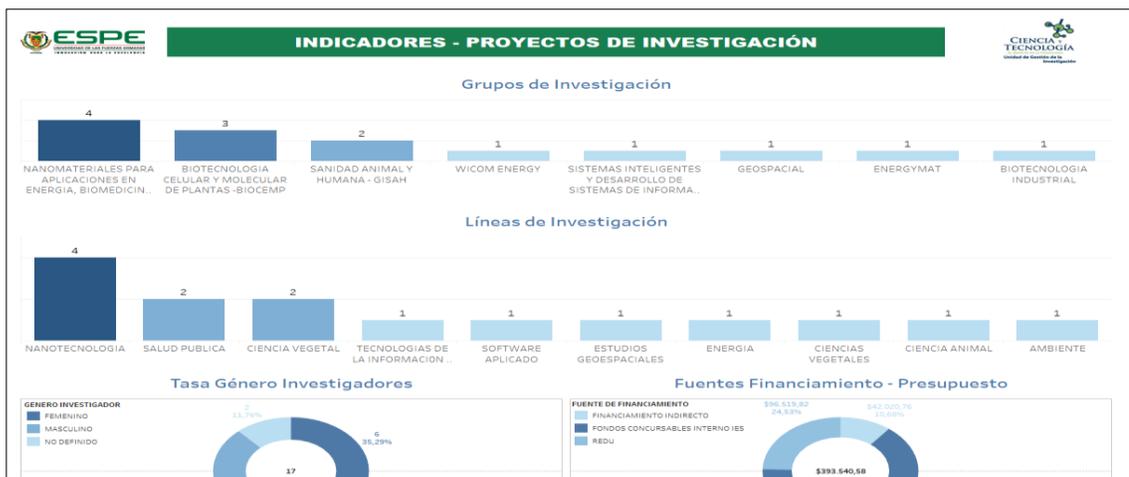


Finalmente se muestra el aporte de Libros y Capítulos por Departamento para conocer el aporte que han contribuido cada uno para estos dos indicadores como se muestra a continuación.



Módulo Proyectos

Para el módulo de proyectos se toma como base el archivo Excel que maneja el usuario de negocio donde tenemos el dashboard que se presenta a continuación.



Para los indicadores que se refieren a los grupos y líneas de investigación se tiene 2 gráficos de barras donde se indica el aporte que han contribuido cada departamento de forma descendente de la siguiente manera:



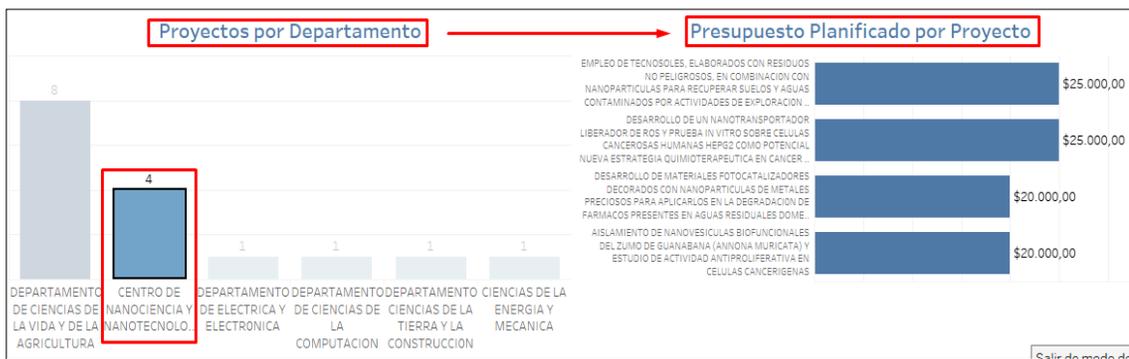
Por otra parte se dispone del análisis mediante los indicadores de composición para el género de Investigadores y las diferentes Fuentes de Financiamiento como se muestra a continuación:



Con respecto a los proyectos que se manejan por departamento, se dispone de un gráfico de barras verticales para conocer dentro de cada departamento cual maneja más presupuesto en sus proyectos como se muestra a continuación.



La interacción directa con el dashboard se da cuando se selecciona un departamento para mostrar los diferentes valores de presupuesto por proyecto como se muestra a continuación:



BIBLIOGRAFÍA

- Ajenstat, F. (27 de febrero de 2018). *TABLEAU*. Obtenido de TABLEAU: <https://www.tableau.com/es-es/about/blog/2018/2/tableau-named-leader-gartner-magic-quadrant-six-years-row-82534>
- Akinci, C., & Smith, S. (2012). INTUITION IN MANAGEMENT RESEARCH. *International Journal of Management Reviews*, 104-122.
- Amazon. (2019). AWS. Obtenido de <https://aws.amazon.com/es/relational-database/>
- Antonio, R., Ordoñez, J., & Ortega, J. (2014). Solución de BI Corporativa, Transformando datos en conocimiento. *Revista de Ciencia, Tecnología y Medio Ambiente*, Volumen XII.
- Azita, S. (2011). An approach to building and implementation of business intelligence system in exchange stock companies. *Australian Journal of Basic and Applied Sciences*, 1491-1495.
- Bernabeu, D. (6 de Mayo de 2009). *DataPRIX*. Obtenido de <https://www.dataprix.com/es/data-warehousing-y-metodologia-hefesto/arquitectura-del-data-warehouse/34-datawarehouse-manager#x1-490003.4.5>
- BIS, S. (31 de julio de 2017). *BIS Tecnologías de la información en México y el mundo*. Obtenido de <https://bismx.wordpress.com/2017/07/31/que-es-tableau-software/>
- Calvo, D. (21 de Noviembre de 2017). *Inteligencia de Negocios Diego Calvo*. Obtenido de <http://www.diegocalvo.es/inteligencia-de-negocio/ciclo-de-vida-kimball/>
- Caraguay, J. (2018). *ANÁLISIS DE DATOS DEL CONSUMO ELÉCTRICO PARA MEJORAR LA TOMA DE DECISIONES UTILIZANDO INTELIGENCIA DE NEGOCIOS*". Ibarra: UTN. Obtenido de <http://reader.digitalbooks.pro/content/preview/books/43005/book/OEBPS/chapter02.xhtml>
- Chaudhary, S. (2014). Management factors for strategic BI success. *IGI Global, Disseminator of Knowledge*, 16.
- ClarKAt. (2019). *ClarKAt SOLUCIONES*. Obtenido de <https://www.clarcat.com/tableau-software/>
- Clavero, A., & Carlos, C. (2016). En *Salvaguarda y seguridad de los datos*. Madrid: Paraninfo.
- Colmayor. (09 de 05 de 2017). *Institución Universitaria Colegio Mayor de Antioquia*. Obtenido de [http://www.colmayor.edu.co/uploaded_files/images/archivos/normograma/manuales/Manualindicadoresversion20\(may8\)10\(1\).pdf](http://www.colmayor.edu.co/uploaded_files/images/archivos/normograma/manuales/Manualindicadoresversion20(may8)10(1).pdf)
- COLPATRIA, A. (s.f.). *Objetivos e indicadores de gestión*. Obtenido de https://asesoriavirtualaxacolpatria.co/axafiles/gestor_contenidos/Objetivos_e_Indicadores_de_gestion_SG-SST_680_2016_08_01_14_54_01.pdf
- Conesa, J. C. (2010). *Introducción al Business Intelligence*. México: Editorial UOC.

- CulturaCrm. (26 de 12 de 2016). *CULTURA CRM*. Obtenido de <http://culturacrm.com/business-intelligence/kpi-business-intelligence/>
- DeGregorio, A. (2008). *Indicadores cuantitativos versus indicadores cualitativos*. Barcelona: Ajuntament de Barcelona.
- Dertiano, V. (06 de abril de 2015). *bigEEK ARQUITECTURAS BI*. Obtenido de <https://blog.big-geek.com/arquitectura-el-enfoque-de-ralph-kimball/>
- ECURed. (3 de agosto de 2019). *EcuRed*. Obtenido de https://www.ecured.cu/Cubos_OLAP
- Espinosa, R. (25 de febrero de 2010). *DATA Prix*. Obtenido de https://www.dataprix.com/blogs/respinosamilla/herramientas-etl-que-son-para-que-valen-productos-mas-conocidos-etl-s-open-sour?utm_campaign=procesos-etl&utm_source=hs_automation&utm_medium=email&utm_content=9063552&_hsenc=p2ANqtz-9hXdnZo04IUnFIgu4Ks-qu55TW
- Fonseca, E. (2019). *Informe sobre el sistema de Publicaciones de la UGI*. Sanglquí.
- Fontalvo, T., De la Hoz, E., & Vergara, J. (2012). Aplicación de análisis discriminante para evaluar el mejoramiento de los indicadores financieros en las empresa de Barranquilla. *Revista Chilena de Ingenieros*, 320-330.
- Hamizah, A., Shahizan, M., & Mi, L. (2015). Significance of Data Integration and ETL in Business Intelligence Framework for Higher Education. *2015 International Conference on Science in Information Technology*. Indonesia: IEEE.
- HCU. (16 de junio de 2015). ESPE-HCU. *Resolución*. Sangolquí, Pichincha, Ecuador.
- Howson, C., Richardson, J., & Kronz, A. (13 de febrero de 2019). *Gartner peer insights*. Obtenido de <https://www.gartner.com/doc/reprints?id=1-68720FP&ct=190213&st=sb>
- Hueso, A., & Cascant, J. (2012). *Metodología y Técnicas*. Valencia: Universitat Politècnica de Valencia.
- Inmon. (2002). *Building the Data Warehouse (Third Edition)*. New York: Jhon Wiley & Sons.
- Inmon, W. H. (24 de 09 de 1992). El padre del almacenamiento de datos. California: Computer World. Obtenido de Data Warehouse: http://www.comprendamos.org/alephzero/45/data_warehouse_analisis_de_datos_no_estructurados.html
- Kabiri, A., & Chiadmi, D. (2013). A method for modelling and organazing ETL processes. *Second International Conference on the Innovative Computing Technology (INTECH 2012)*, 23, pág. 5. Casablanca, Morocco.
- Keerin, P. (2016). Development of business intelligence solution for personnel administration. *IEEE. 2016 Second Asian Conference on Defence Technology (ACDT)*, 22, 1-8.
- Kerzner, H. (2013). *Project Management Metrics, KPI's and dashboard*. New York: Jhon Wiley & sons.

- Kitchenham, B., & Charters, S. (2007). *Guidelines for performing Systematic Literature Reviews in Software Engineering*.
- Krishma, N., & Suneetha, M. (2016). Business Intelligence solutions for processing huge data to the Business user's using dashboards. *International conference on Signal Processing, Communication, Power and Embedded System*.
- Laiseca, J. (2019). *Consultor*. Obtenido de jesuslaiseca.com:
<https://sites.google.com/site/jesuslaiseca/olap/molap-rolap-y-holap>
- Leonard, E., & Castro, Y. (2013). Metodologías para desarrollar Almacén de datos. *Revista de Arquitectura e Ingeniería*, 1-12.
- LOES. (12 de Octubre de 2010). Fines de la Educación Superior. Quito, Pichincha, Ecuador.
- Martínez, D. (24 de Julio de 2017). *UNIR. Metodología para el diseño de dashboards orientado hacia el registro de evidencias en el proceso de evaluaciones institucionales*. Obtenido de
<https://reunir.unir.net/bitstream/handle/123456789/6171/MARTINEZ%20ROBALINO%2C%20DANIEL%20ANDRES.pdf?sequence=1&isAllowed=y>
- Medina, E. (2014). Business Intelligence: La información como arma competitiva. *Sinergia e innovación Universidad Peruana de Ciencias Aplicadas UPC*. Obtenido de
<https://repositorioacademico.upc.edu.pe/handle/10757/333779>.
- Méndez del Río, L. (2006). Más allá de Business Intelligence. En *Experiencias de éxito* (pág. 219). Barcelona.
- Moody, D. (2010). From Enterprise Models to Dimensional Models: A Methodology for Data Warehouse and Data Mart Design. *Australian Data Management Conference* (pág. Volumen 28). Parkville, Australia: DMDW 2000.
- Muñoz, L., Mazon, N., & Trujillo, J. (2011). ETL Process Modeling Conceptual for Data Warehouses. *IEEE LATIN AMERICA TRANSACTIONS*, Volumen 9.
- Oliveras, E. (2018). *Grupo P&A*. Obtenido de Blog Indicadores cuantitativos y cualitativos:
<https://blog.grupo-pya.com/indicadores-de-evaluacion-cualitativos-y-cuantitativos/>
- Olszak, C. Z. (2007). Approach to Building and Implementing Business Intelligence Systems. *Interdisciplinary Journal of Information, Knowledge and Management*, volumen 2.
- ONUMUJERES. (31 de octubre de 2010). *ONUMUJERES*. Obtenido de
<http://www.endvawnow.org/es/articles/336-indicadores.html>
- Pérez, D. (26 de octubre de 2007). *PLATZI, Maestros del Web Bases de Datos*. Obtenido de
<http://www.maestrosdelweb.com/que-son-las-bases-de-datos/>
- PowerData. (13 de junio de 2013). *Procesos ETL: Extracción, Transformación, Carga*. Obtenido de <https://blog.powerdata.es/el-valor-de-la-gestion-de-datos/bid/288859/procesos-etl-extracci-n-transformaci-n-carga>

- QAEC. (Julio de 2019). *Asociación Española para la calidad*. Obtenido de <https://www.aec.es/web/guest/centro-conocimiento/cuadro-de-mando>
- Raffino, M. (30 de octubre de 2019). *Sistemas de Información*. Obtenido de <https://concepto.de/base-de-datos/>
- Rivadeneira, G. (2010). *Universidad Católica de Salta*. Obtenido de <https://www.ucasal.edu.ar/hm/ingenieria/cuadernos/archivos/5-p56-rivadere-formateado.pdf>
- Rochina, P. (10 de Febrero de 2017). *Guía para construir un DataWarehouse. Caso práctico*. Obtenido de <https://revistadigital.inesem.es/informatica-y-tics/guia-construir-datawarehouse/>
- Salvador, R. (21 de julio de 2017). *Dara Analytics*. Obtenido de <https://blogs.solidq.com/es/business-analytics/data-warehouse-y-data-marts-esquema-en-estrella-11/>
- Sanchez, M., Coronel, E., Labanda, M., Chamaba, L., & Guamán, R. (13-16 de june de 2018). Computer system for the management and scientific divulgation of the “Universidad Nacional de Loja”. Loja, Loja, Ecuador.
- Sevilla, N., Reinoso, A., Ordoñez, J., & Ortega, J. (2014). SOLUCIÓN BUSINESS INTELLIGENCE (BI) CORPORATIVA; Transformando datos en conocimiento. *Revista de Ciencia, Tecnología y Medio Ambiente SEPARATA*, Volumen XII.
- Sinnexus. (01 de 01 de 2018). *Business Intelligence* . Obtenido de https://www.sinnexus.com/business_intelligence/olap_vs_oltp.aspx
- Solano, H. (01 de Julio de 2016). *INCAE Business School*. Obtenido de <https://www.incae.edu/es/blog/2016/07/01/retos-y-beneficios-del-business-intelligence.html>
- Suneetha, M., & Krishna, N. (2017). Business intelligence solutions for processing huge data to the business user's using dashboards. *Conferencia Internacional sobre Procesamiento de Señal, Comunicación, Energía y Sistemas Integrados (SCOPEs)*. Paralakhemundi, India.
- Tello, E., & Perusquia, J. (2016). Inteligencia de negocios: estrategia para el desarrollo de competitividad en empresas de base tecnológica. *Contaduría y Administración*, 127-158.
- Trujillo, J., & Lujan, S. (2008). A UML Based Approach for Modeling ETL Process in Data Warehouse. *International Conference on Conceptual Modeling* (págs. 307-320). Chicago: Universidad Alicante.
- UGI. (2018). *Sistema de Investigación de la ESPE*. Obtenido de Unidad de gestión de Investigación ESPE: <http://ugi.espe.edu.ec/ugi/wp-content/uploads/2013/06/SISTEMA-DE-INVESTIGACI%C3%93N-DE-LA-ESPE.pdf>
- UGI. (2019). Sangolqui., Pichincha, Ecuador.
- Valero, M. (16 de Mayo de 2016). *Bigeeek* . Obtenido de Business Intelligence: <https://blog.bi-geek.com/que-es-una-etl/>

Valero, M. (9 de octubre de 2017). *Bigeeek Modelo Dimensional*. Obtenido de <https://blog.bi-geek.com/modelo-dimENSIONAL/>

Vélez, M., & Dávila, M. (2018). La calidad de la educación superior.

WIKIDOT. (8 de diciembre de 2015). *WIKIDOT Implementación de un sistema business intelligence*. Obtenido de <http://cuboolapbd.wikidot.com/>

Zambrano, E., Reyes, R., Castro, J., & Fonseca, E. (2018). Métricas que podrían usarse en el desarrollo e palicaciones móviles para personas con deficiencias visuales. Una revisión de literatura preliminar. . *RISTI*, 985-999.

Zhang, H., Alibabar, M., & Tell, P. (28 de diciembre de 2010). Identifying relevant studies in software engineering. Australia, Australia.