

**UNIVERSIDAD POLITÉCNICA SALESIANA
SEDE QUITO**

CARRERA: INGENIERÍA DE SISTEMAS

**Trabajo de titulación previo a la obtención del título de: INGENIERA E
INGENIERO DE SISTEMAS**

**TEMA:
ANALIZAR, DISEÑAR E IMPLEMENTAR UNA SOLUCIÓN DE
“INTELIGENCIA DE NEGOCIOS” PARA LA TOMA DE DECISIONES
GERENCIALES APLICADAS EN EL ÁREA DE CONTROL FINANCIERO
PARA EL BANCO INTERNACIONAL**

**AUTORES:
ERIKA SUSANA ERAZO MEJÍA
PAÚL LEONARDO ORTIZ VALDEZ**

**DIRECTOR:
WALTER FERNANDO GAIBOR NARANJO**

Quito, mayo de 2015

**DECLARATORIA DE RESPONSABILIDAD Y AUTORIZACIÓN DE USO DEL
TRABAJO DE TITULACIÓN**

Nosotros, autorizamos a la Universidad Politécnica Salesiana la publicación total o parcial de este trabajo de titulación y su reproducción sin fines de lucro.

Además, declaramos que los conceptos, análisis desarrollados y las conclusiones del presente trabajo son de exclusiva responsabilidad de los autores.

Quito, 05 de mayo del 2015

Erika Susana Erazo Mejía

C.C. 171645297-2

Paúl Leonardo Ortiz Valdez

C.C. 172100977-5

DEDICATORIA

En primer lugar agradecer a Dios por llenarnos de sabiduría y habernos dado unos padres preocupados por nuestro futuro, que de una u otra forma nos han apoyado en cada momento de nuestras vidas, a ellos como hijos el eterno agradecimiento por haber forjado en nosotros valores y virtudes, que nos han ayudado tanto en lo personal como profesional.

AGRADECIMIENTO

Agradecemos a la Universidad “Politécnica Salesiana” y sus docentes, cuyas cualidades inmejorables, tanto personal como profesional, que nos ayudaron a forjar en nuestras mentes y corazón el deseo de superación.

Al Ingeniero Walter Gaibor por su invaluable apoyo durante el desarrollo del presente trabajo, que siempre estuvo ampliando nuestro conocimiento, solventando preguntas e inquietudes encontradas en el camino, con su gran apoyo logramos alcanzar la meta propuesta.

ÍNDICE

INTRODUCCIÓN	1
CAPÍTULO 1	2
PLANTEAMIENTO DEL PROBLEMA	2
1.1 Antecedentes	2
1.1.1 Identificación de indicadores del problema	4
1.1.2 Efectos que generan	4
1.2 Justificación.....	5
1.3 Alcance	6
1.4 Objetivos	9
1.4.1 General	9
1.4.1.1 Específicos.....	9
CAPÍTULO 2	10
MARCO TEÓRICO	10
2.1 Banco Internacional.....	10
2.1.1 Introducción	10
2.1.2 Historia.....	10
2.1.3 Misión	10
2.1.4 Visión	11
2.2 Antecedentes de la investigación.....	11
2.3 Inteligencia de negocios	11
2.3.1 Introducción	11
2.3.2 Conceptos de inteligencia de negocios	12
2.3.3 Importancia de la inteligencia de negocios	13

2.3.4	Arquitectura inteligencia de negocios	13
2.4	Metodologías de desarrollo BI	14
2.4.1	Definición de la metodología	14
2.4.2	Metodología Bill Inmon	14
2.4.3	Metodología de Ralph Kimball	16
2.5	Comparación de las metodologías	17
2.6	Elección de la metodología	18
2.7	Base de datos relacional	19
2.7.1	Tipos de relaciones en un modelo entidad – relación	19
2.7.1.1	Ventajas	19
2.7.1.2	Desventajas	19
2.7.2	Arquitectura de base de datos	20
2.7.3	Almacén de datos	20
2.8	Data warehouse	20
2.8.1.1	Función de un data warehouse	21
2.8.1.2	Elementos de un data warehouse	22
2.8.1.3	Elementos de una arquitectura data warehouse	25
2.8.1.4	Arquitectura de un data warehouse	26
2.8.2	Data mart	29
2.8.2.1	Definición data mart	29
2.8.2.2	Tipos de data mart	29
2.8.2.3	Carga de datos en data mart	30
2.8.3	Cubos OLAP	31
2.8.3.1	Formas de almacenamiento de los cubos	31

CAPÍTULO 3	32
METODOLOGÍA DE DESARROLLO	32
3.1. Metodología de Ralph Kimball	32
3.1.1. Planificación del proyecto.....	33
3.1.1.1. Preparación de la organización.....	33
3.1.1.2. Medir la Demanda	37
3.1.1.3. Determinación del alcance preliminar.....	38
3.1.1.4. Justificación del proyecto.....	39
3.1.1.5. Personal encargado del proyecto	39
3.1.1.6. Desarrollar el plan del proyecto.....	41
3.1.2. Definición de los requerimientos del negocio.....	42
3.1.2.1. Análisis preliminar de auditoria de los datos.....	45
3.1.2.2. Criterios de éxito.....	45
3.1.3. Diseño de la arquitectura técnica	45
2.8.3.2 Arquitectura back room.....	54
2.8.3.3 Arquitectura front room.....	56
3.1.4. Selección e instalación del producto	57
3.1.4.1. Posibles atajos para la evaluación.....	58
3.1.5. Modelo dimensional.....	60
3.1.5.1. Identificación de fuentes de datos.....	61
3.1.5.2. Diseño de arquitectura del bus del data warehouse	61
3.1.5.3. Proceso de modelado dimensional.....	62
3.1.5.4 Modelo lógico.....	69
3.1.6. Diseño físico	71
3.1.6.1. Estándares	71

3.1.6.2.	Diseño físico de tablas	73
3.1.6.3.	Diseño multidimensional	76
3.1.6.4.	Homologación de datos	78
3.1.6.5.	Integración de los datos ETL (extracción, transformación y carga)...	81
3.1.6.6.	Diseño del análisis y reporte de la información.....	98
CAPÍTULO 4		105
IMPLEMENTACION DE REPORTES		105
4.1	Desarrollo del análisis y reporte de la información.....	105
4.2	Implementación de reportes	105
4.2.1	Desarrollo tablero de control web	105
4.2.2	Reportería WEB	105
4.2.3	Reporte personalizados	107
4.2.4	Estrategia de capacitación.....	109
4.2.5	Pruebas de operación.....	109
4.2.5.1	Carga data mart captaciones y colocaciones	109
4.2.5.2	Carga data warehouse.....	115
CONCLUSIONES		116
RECOMENDACIONES		118
LISTA DE REFERENCIA		119
ANEXOS		121
GLOSARIO		132

ÍNDICE DE FIGURAS

<i>Figura 1.</i> Componentes de la arquitectura BI.....	13
<i>Figura 2.</i> Modelo Bill Inmon.....	15
<i>Figura 3.</i> Modelo Ralph Kimball	16
<i>Figura 4.</i> Esquema tipo estrella y tabla de hechos.....	28
<i>Figura 5.</i> Esquema Copo de nieve y tablas jerárquicas	28
<i>Figura 6.</i> Ejemplo data marts dependientes.....	29
<i>Figura 7.</i> Ejemplo Data mart Independientes	30
<i>Figura 8.</i> Business Dimensional Lifecycle (BDL) por Ralph Kimball	32
<i>Figura 9.</i> Dependencia Planificación Del Proyecto y Definición de los Requerimientos del negocio Elaborado: Paúl Ortiz, Erika Erazo	33
<i>Figura 10.</i> Análisis de resultados obtenidos de la encuesta.....	37
<i>Figura 11.</i> Dependencia de los requerimientos en todas las etapas.....	43
<i>Figura 12.</i> Componentes de la arquitectura técnica Back Room.....	55
<i>Figura 13.</i> Componentes arquitectura Front Room	57
<i>Figura 14.</i> Resultados de la Encuesta Evaluación de Software.....	58
<i>Figura 15.</i> Ejemplo de Granularidad de una Institución Bancaria	63
<i>Figura 16.</i> Detalle dimensión fecha.....	64
<i>Figura 17.</i> Diagrama de dimensión Institución financiera	64
<i>Figura 18.</i> Diagrama de dimensión Producto	65
<i>Figura 19.</i> Diagrama de dimensión Plaza.....	65
<i>Figura 20.</i> Diagrama de dimensión fecha RANKING BANCARIO.	66
<i>Figura 21.</i> Diagrama de dimensión Institución financiera RANKING.....	66
<i>Figura 22.</i> Diagrama de dimensión Producto RANKING.....	67
<i>Figura 23.</i> Diagrama de la tabla de hechos saldos, cuentas contables.	68
<i>Figura 24.</i> Diagrama de la tabla de hechos saldos, cuentas contables.	69
<i>Figura 25.</i> Data mart Captaciones Colocaciones. Fuente. Datamart Capcol Banco Internacional	77
<i>Figura 26.</i> Data mart Instituciones Financieras. Fuente. Data mart Ranking Banco Internacional	78
<i>Figura 27.</i> Control de flujo de la carga del proceso de homologación.....	79

<i>Figura 28.</i> Flujo de datos “Obtener nombres no validos”:	79
<i>Figura 29.</i> Flujo de datos “Obtener nombres por ingresar”.	80
<i>Figura 30.</i> Flujo de datos “Homologar Nombres”	80
<i>Figura 31.</i> Control de flujo de la carga del proceso de homologación agregado instituciones	81
<i>Figura 32.</i> Flujo de Tareas Dimensión Fecha	83
<i>Figura 33.</i> Flujo de datos “Dimensión Fecha”	84
<i>Figura 34.</i> Flujo de tareas “Dimensión Plaza”	84
<i>Figura 35.</i> Flujo de datos “Carga Tablas Cantones Provincias Regiones”	85
<i>Figura 36.</i> Flujo de datos “Dimensión Plaza”	85
<i>Figura 37.</i> Flujo de tareas “Dimensión Institución”	86
<i>Figura 38.</i> Flujo de Datos “Carga Tablas de Institución”	86
<i>Figura 39.</i> Flujo de Datos “Carga Tablas de Institución”	87
<i>Figura 40.</i> Flujo de tareas “Dimensión Producto”	87
<i>Figura 41.</i> Flujo de datos “Carga Tablas Productos”	88
<i>Figura 42.</i> Flujo de Datos “Dimensión Producto”	88
<i>Figura 43.</i> Flujo de Tareas “Tabla de hechos Capcol”	89
<i>Figura 44.</i> Flujo de tareas “Colocaciones”	91
<i>Figura 45.</i> Flujo de tareas “Captaciones”.	91
<i>Figura 46.</i> Flujo de datos "Dimensión Fecha"	92
<i>Figura 47.</i> Flujo de tareas "Dimensión Fecha"	93
<i>Figura 48.</i> Flujo de datos "Dimensión Institución"	93
<i>Figura 49.</i> Flujo de tareas "Dimensión Institución"	94
<i>Figura 50.</i> Flujo de datos "Dimensión Cuenta Contable"	94
<i>Figura 51.</i> Flujo de tareas "Dimensión Cuenta Contable"	95
<i>Figura 52.</i> Flujo de datos "Dimensión Plaza"	95
<i>Figura 53.</i> Flujo de tareas "Inserta data regiones"	96
<i>Figura 54.</i> Flujo de tareas "Procesa dimensión plaza"	96
<i>Figura 55.</i> Flujo de tareas "Carga tabla de hechos"	97
<i>Figura 56.</i> Creación de las fuentes de datos	99
<i>Figura 57.</i> Selección de dimensiones y tablas de hechos	100

<i>Figura 58.</i> Vista de Datos DWH	101
<i>Figura 59.</i> Método de selección construcción de un cubo.....	102
<i>Figura 60.</i> Selección de dimensión y tablas de hechos.	102
<i>Figura 61.</i> Estructura dimensión tiempo	103
<i>Figura 62.</i> Selección de medidas	103
<i>Figura 63.</i> Resumen de dimensiones y tabla de hechos	104
<i>Figura 64.</i> Comparativa de saldos	106
<i>Figura 65.</i> Reporte principales bancos.	106
<i>Figura 66.</i> Conexión SQL Server Analysis Services	107
<i>Figura 67.</i> Cubo OLAP MISDWH.....	107
<i>Figura 68.</i> Ejemplo 1 de consulta personalizadas.	108
<i>Figura 69.</i> Ejemplo 2 de consulta personalizadas	108
<i>Figura 70.</i> Paquete "Carga Mensual"	110
<i>Figura 71.</i> Flujo de control "Dimensión Fecha"	110
<i>Figura 72.</i> Flujo de tareas "Dimensión Fecha"	111
<i>Figura 73.</i> Validación de la tabla RCC_Fecha	111
<i>Figura 74.</i> Flujo de tareas "Dimension Institución"	112
<i>Figura 75.</i> Flujo de control "Dimensión cuenta contable"	112
<i>Figura 76.</i> Flujo de datos "dimensión cuenta contable"	113
<i>Figura 77.</i> Carga tabla de hechos	113
<i>Figura 78.</i> Carga tabla de hechos	114
<i>Figura 79.</i> Flujo de carga mensual para el data mart ranking	114
<i>Figura 80.</i> Carga DWHMIS	115

ÍNDICE DE TABLAS

Tabla 1. <i>Estudio comparativo de las metodologías Kimball vs Inmon</i>	18
Tabla 2. <i>Test de Litmus</i>	35
Tabla 3. <i>Asignación de roles y responsabilidades</i>	40
Tabla 4. <i>Plan de comunicación Banco Internacional</i>	41
Tabla 5. <i>Gestión de riesgos Banco Internacional</i>	42
Tabla 6. <i>Definición de requerimientos del negocio.</i>	44
Tabla 7. <i>Equipo de trabajo diseño de la arquitectura</i>	46
Tabla 8. <i>Implicaciones arquitectónicas Banco Internacional</i>	47
Tabla 9. <i>Modelo arquitectónico de alto nivel Banco Internacional</i>	49
Tabla 10. <i>Requisitos de la infraestructura</i>	50
Tabla 11. <i>Despliegue de la arquitectura técnica</i>	53
Tabla 12. <i>Definición de fuentes de datos.</i>	61
Tabla 13. <i>Arquitectura de bus del data warehouse</i>	61
Tabla 14. <i>Descripción de la tabla de hechos saldo y cuentas, captaciones y colocaciones</i>	68
Tabla 15. <i>Atributos dimensión fecha</i>	69
Tabla 16. <i>Atributos dimensión Institución</i>	70
Tabla 17. <i>Atributos de la dimensión cuenta contable</i>	70
Tabla 18. <i>Descripción tabla RKD_FECHA</i>	73
Tabla 19. <i>Descripción tabla RKD_Institucion</i>	73
Tabla 20. <i>Descripción tabla RKD_Plaza</i>	74
Tabla 21. <i>Descripción tabla RKD_producto</i>	74
Tabla 22. <i>Descripción tabla RKH_Saldos_Cuentas</i>	74
Tabla 23. <i>Descripción tabla RKBD_Fecha</i>	75
Tabla 24. <i>Descripción tabla RKBD_Plaza</i>	75
Tabla 25. <i>Descripción tabla RKDB_Institucion</i>	75
Tabla 26. <i>Descripción tabla RKBD_CuentaContable</i>	76
Tabla 27. <i>Descripción tabla RKBK_Saldos_Cuentas</i>	76
Tabla 28. <i>Descripción del archivo "Formatos.xls" por hoja</i>	88

ÍNDICE ANEXOS

Anexo 1. Reunión determinación del alcance	121
Anexo 2. Reunión determinación de la arquitectura data warehouse	122
Anexo 3. Revisión de la metodología de desarrollo	123
Anexo 4. Estudio de la metodología a implementar	124
Anexo 5. Encuesta de para medir la demanda de una aplicación BI	125
Anexo 6. Encuesta para medir la experiencia del equipo	128

RESUMEN

El presente proyecto de titulación consiste en analizar, diseñar e implementar una solución de “Inteligencia de negocios” en el área de control financiero para el Banco Internacional. Fue elaborado con la finalidad de brindar soporte en la toma de decisiones gerenciales para los objetivos estratégicos de la institución financiera.

El contenido de este documento consta de cuatro capítulos cada uno con su temática relacionada al desarrollo de la solución propuesta, que se detalla a continuación:

El capítulo uno describe la problemática actual de los métodos que lleva a cabo el área de control financiero, con el fin de obtener información para el soporte del proceso de toma de decisiones gerenciales. Menciona la información que se va analizar así como también la solución planteada en base a los problemas detectados.

El capítulo dos contiene toda la sustentación científica y teórica para cubrir toda la terminología utilizada en el desarrollo del documento.

El capítulo tres abarca la metodología de desarrollo y el despliegue de la solución de Inteligencia de negocios. Según el estudio realizado se eligió la metodología de Ralph Kimball debido a su gran enfoque en el desarrollo de aplicaciones de Inteligencia de negocios y su gran manejo de almacenes de datos.

El capítulo cuatro presenta la implementación de los reportes considerando los métodos de visualización y/o estándares definidos por la institución, además se expone las distintas maneras que los usuarios pueden acceder a la información.

ABSTRACT

This titling project is to analyze, design and implement a solution of "Business Intelligence" in the area of financial control for the International Bank. It was developed in order to provide support in making management decisions for the strategic objectives of the financial institution.

The contents of this document consists of four chapters each with its related to the development of the proposed solution subject, as detailed below:

Chapter one describes the current problems of the methods that performs the area of financial control in order to obtain information to support the process of management decision making. He mentions the information to be analyzed as well as the proposed solution based on the detected problems.

Chapter two contains all scientific and theoretical to cover all the terminology used in the development of the document support.

Chapter three covers methodology development and deployment of Business Intelligence solution. According to the study's methodology Ralph Kimball was chosen because of its strong focus on application development business intelligence and great handling data stores.

Chapter four presents the implementation considering reports display methods and / or standards defined by the institution, and the ways that users can access information exposed.

INTRODUCCIÓN

Hoy en día el análisis de información se ha convertido en un instrumento clave para obtener una ventaja competitiva en los negocios, los datos constituyen un elemento primordial para las organizaciones que poseen grandes volúmenes de información relevante, para mejorar el desempeño e innovar los procesos de tal manera que pueda transformarse en una empresa exitosa.

Justamente de ahí nace la necesidad de implementar nuevas herramientas IT (Information Technology) que permitan estar un paso adelante y anticiparse a los acontecimientos para la adecuada y oportuna toma de decisiones. Con ello se busca una posición dentro del mercado, por lo cual se necesita identificar oportunidades y así obtener ventaja sobre ella de manera rápida y efectiva.

La Inteligencia de negocios (BI) ayuda a las empresas a comprender, supervisar, analizar e incluso prever lo que va a pasar en el entorno general y en la empresa. Su uso permite convertir datos en información útil y relevante para distribuirla a las personas que la necesiten en el momento oportuno, a fin de que puedan tomar decisiones adecuadas en el momento oportuno y sobre todo bien fundamentadas.

La institución financiera Banco Internacional está enfocada en la calidad de sus productos y servicios, debido a las exigencias de sus clientes y a la gran cantidad de ofertas en todos los ámbitos del mercado; permitiendo que el cliente tenga más opciones para escoger y satisfacer sus necesidades.

CAPÍTULO 1

PLANTEAMIENTO DEL PROBLEMA

1.1 Antecedentes

La Inteligencia de negocios o Business Intelligence (BI) con el Data Warehouse (DWH), como componentes de alto nivel de los Sistemas de Información, tienen una serie de ventajas y beneficios para toda organización, los más relevantes se basan en estudiar cantidades de información y obtener análisis de ellas permitiendo un mejor desempeño de la organización. Con información más precisa y conocimiento, se puede mejorar el manejo operativo de la empresa, mejorar la toma de decisiones estratégicas en áreas críticas, que presenta un alto impacto en beneficios para el cliente.

Para realizar el presente proyecto se estudió a fondo la empresa desde sus orígenes, su evolución, el presente y hacia donde está enfocada la situación actual del Banco Internacional, de tal manera que permitió conocer desde varios enfoques la estructura organizacional y así poder identificar los problemas al momento de obtener información.

Se realizó todo el proceso para conocer sus necesidades con el fin de identificar el problema y dar una propuesta de solución que permitiera al área de Control Financiero mejorar el flujo de información de forma consistente y eficaz para la toma de decisiones gerenciales.

Las organizaciones financieras para generar competitividad en el mercado realizan estudios sobre: “Instituciones Financieras, Captaciones y Colocaciones” (SBS, 2014), provista por el Sistema Financiero Nacional, regulada por la Superintendencia de Bancos y Seguros.

Instituciones financieras: es la recopilación de la información financiera con respecto a los activos y pasivos de cada entidad, dentro de estos indicadores se analiza los saldos de todas las instituciones financieras y su principal competencia entre los bancos privados, así mismo el porcentaje de participación en el mercado económico que es reflejado en Sistema

Financiero Nacional, el cual permite conocer el top de los seis mejores bancos a nivel nacional. (SBS, 2014)

Captaciones y colocaciones: realiza un estudio económico comparativo por plaza, producto o medida financiera entre las instituciones financieras y el Banco Internacional basándose en la información de captaciones y colocaciones de los clientes, emitida por la Superintendencia de Bancos y Seguros. (SBS, 2014)

Captaciones: “constituyen todos los recursos del público que han sido depositados en una institución financiera a través de depósitos a la vista y depósitos a plazo u otros mecanismos.”

Colocación o crédito: “es un préstamo de dinero que un Banco otorga a su cliente, con el compromiso de que en el futuro, el cliente devolverá dicho préstamo en forma gradual, mediante el pago de cuotas, o en un solo pago y con un interés adicional que compensa al acreedor por el período que no tuvo ese dinero.”

Banco Internacional actualmente no dispone de una solución de inteligencia de negocios vinculada a indicadores gerenciales de: instituciones financieras, captaciones y colocaciones. Esto genera desventajas al utilizar procesos manuales para la recopilación de información de varias bases de datos operacionales.

Por lo tanto los usuarios que toman decisiones en función de las mismas experimentan inconvenientes en:

- Capacidad de extraer, depurar y agregar datos de varios sistemas de información.
- No contar con información inmediata, consolidada que permitan la entrega de información resumida y examinada al detalle, rápidamente.

Vistas personalizadas, capacidad de consulta, reportes y análisis relevantes que podrían ayudar a la toma de decisiones de una forma eficiente y objetiva.

1.1.1 Identificación de indicadores del problema

Uno de los problemas principales es la elaboración de reportes, generados de una forma manual. El mismo que conlleva tiempo de espera para cada consulta realizada a la base de datos por el área de sistemas para la extracción de los mismos, el cual provoca una gran cantidad de tráfico en la base de datos de producción. Los datos son extraídos y entregados a los usuarios en hojas de cálculo.

Cada uno de los usuarios se encarga de depurar, ordenar, organizar y clasificar la información, con el fin de realizar el proceso de análisis, el mismo que se requiere para la toma de decisiones estratégicas de la organización.

Al realizar este procedimiento, se presentan varios problemas que se detallan a continuación:

- Cada usuario aplica su criterio para realizar el procedimiento de depurar, ordenar, organizar y clasificar los datos que obtienen. Dado que no está estandarizado a nivel de toda la organización, al realizar este procedimiento de forma manual, los usuarios invierten mucho tiempo y los resultados obtenidos no son exactos. Al manipular una gran cantidad de datos, el beneficiario esta propenso a cometer una serie de errores.
- Cada uno de los usuarios obtiene hojas de cálculo creadas para cada proceso como resultado del análisis que realizan para obtener sus reportes, de modo que no es posible contar con una información histórica.

1.1.2 Efectos que generan

- La información se trabaja mediante archivos digitales aislados y procesados manualmente, ocasionando posibles pérdidas en el flujo de información al no tener un repositorio centralizado, por lo que se puede identificar la debilidad al no presentar información óptima y consolidada.

- El tiempo de ejecución que demanda el procesamiento de importantes volúmenes de datos es muy extenso debido a la cantidad de información.

1.2 Justificación

El Banco Internacional emprende un proceso de automatización basado en una solución de Inteligencia de negocios, apoyada por el modelamiento multidimensional de datos, con el fin de cubrir los requerimientos analíticos de los usuarios, que sea lo suficientemente robusta para alcanzar las necesidades futuras. Convirtiéndose en el repositorio central y la única fuente de datos consolidados, consultable para el análisis de la información de forma eficaz de alta disponibilidad.

La solución que se plantea para el análisis de información está basada en la elaboración de una aplicación de Inteligencia de negocios (BI) para el Área de Control Financiero, que estará conformado por dos data mart:

- Instituciones financieras
- Información de captaciones y colocaciones

Aplicando este sistema de inteligencia de negocios (BI), se pretende reducir los costos y optimizar los tiempos en lo que se refiere al tratamiento de la información se facilitará la administración de la misma personalizándola y adaptándola a las necesidades lo que permitirá la integración y depuración de los datos.

De acuerdo a este análisis, se pretende realizar el modelo de datos y obtener un conjunto de reportes mensuales (instituciones financieras), trimestrales (captaciones y colocaciones) que cumplan con las necesidades de información de los usuarios con los criterios obtenidos del área de control financiero y así poder tomar decisiones gerenciales que mejoren el rendimiento de banco.

1.3 Alcance

La elaboración de una aplicación de inteligencia de negocios (BI) para el área de control financiero.

La solución planteada estará conformada por dos data mart:

- Instituciones financieras
- Información de captaciones y colocaciones

La combinación de los dos data marts conformará el data warehouse (DWH), el cual brindará un soporte para optimizar el proceso de toma de decisiones gerenciales.

Esta información es provista por la Superintendencia de Bancos y Seguros (SBS) de forma periódica de la siguiente manera:

- **Instituciones financieras:** “es la recopilación de información con respecto a los activos y pasivos de cada entidad a nivel de bancos de forma mensual” (SBS, 2014).
- **Información de captaciones y colocaciones:** “es un estudio económico comparativo por plaza, producto o medida financiera de forma trimestral” (SBS, 2014).

Además contará con un front end que permita visualizar los principales indicadores financieros y realizará las siguientes acciones:

- Modificar los indicadores principales para la generación de reportes.
- Ingresar indicador, crea un nuevo indicador con su respectiva fórmula para generar la información.

Esta funcionalidad interactúa con el usuario final y la base de datos multidimensional, la cual ayuda a mitigar procesos internos con el área de sistemas. Por lo tanto la

información será presentada y procesada en tiempo real para la evaluación de los resultados.

La solución consistirá en brindar una gama de reportes de análisis gerenciales que permitan determinar oportunidades estratégicas a continuación el detalle de los mismos:

1. Principales bancos: presentará un reporte que permitirá observar la posición del banco frente a las demás instituciones financieras, que cuentan con los mayores volúmenes de cartera o depósitos en un determinado momento.

2. Competencia bancaria: mostrará información relevante sobre la diferencia de saldos y cuotas entre el banco y las demás instituciones financieras.

3. Variación de saldo: indicará la diferencia de saldos, mensual y anual, con el porcentaje individual mensual de cada producto, que están clasificados por:

- Depósito monto sin interés
- Depósito monto con interés
- Depósito de ahorro
- Depósito de plazo restringido
- Otros depósitos

4. Captaciones trimestrales: analizará trimestralmente la evolución de cada producto respecto a las captaciones totales de la entidad financiera.

5. Morosidad: contará con 2 reportes que son:

- **Cartera ampliada:** indicará el grado en el que un cliente no cumplió con una obligación pasada y consecuentemente podría no cumplir con sus obligaciones futuras.
- **Cobertura ampliada:** mostrará las provisiones de cartera que mantiene una entidad para cubrir eventuales pérdidas por no recuperación de créditos.

6. Concentración 100 mayores depositantes: visualizará los recursos captados del público que pertenecen a los 100 mayores depositantes.

7. Cobertura 100 mayores depositantes: que son los fondos altamente líquidos que dispone el banco para cubrir los depósitos de los 100 mayores clientes de captaciones, considerando un tiempo de 36 meses.

8. Distribución por plaza: permitirá obtener información por producto, subproducto región de captaciones y colocación.

9. Cuota de mercado: permitirá conocer la parte del mercado que consume los productos o servicios.

10. Solvencia: indicará el nivel de patrimonio que posee una entidad para cubrir, sus activos y contingentes ponderados por grado de riesgo. La ponderación de riesgo está definida por el órgano de control en función de la dificultad de convertir tales activos en recursos altamente líquidos; es decir, determina el grado en el que el patrimonio respalda el riesgo asumido en las distintas formas de inversión realizadas por el banco.

Aplicando el sistema inteligencia de negocios (BI), se pretende reducir recursos y optimizar el tiempo en el tratamiento de la información, por consiguiente la solución planteada está basada en tres capas: BD relacional, BD multidimensional y herramienta de visualización. Este proceso facilita la administración personalizada y adaptándola a las necesidades del negocio. Lo que permite la integración y depuración de los datos, haciendo que las personas encargadas de las decisiones de la institución sean más productivos en los siguientes ámbitos:

- Minimizar el tiempo de recolección de la información importante para el negocio.
- Proporcionar herramientas de análisis para hacer comparaciones y tomar decisiones inteligentes.

- Capacidad de extraer, depurar y agregar datos de múltiples sistemas de información en data mart.
- Almacenar datos en esquemas multidimensionales que permitan la entrega de información resumida y examinada al detalle, rápidamente.
- Capacidades de consultas, reportes y análisis relevantes que podrían ayudar a la toma de decisiones de una forma eficiente y objetiva.

1.4 Objetivos

1.4.1 General

Desarrollar una solución de Inteligencia de negocios que permita organizar y optimizar la información de instituciones financieras, captaciones y colocaciones provistas por el Sistema Financiero Nacional, ayudará a la toma de decisiones gerenciales, con base en requerimientos específicos del área de Control Financiero del Banco Internacional.

1.4.1.1 Específicos

- Desarrollar una base de datos multidimensional que permita obtener la información de una manera estándar, sencilla y sobre todo intuitiva para los usuarios.
- Generar cubos de datos multidimensionales para la extracción de información a través de herramientas OLAP (On-Line Analytical Processing), que permita crear reportes para la toma de decisiones gerenciales de la institución.
- Generar métodos o procesos: extracción, transformación y carga, que alimentarán a los data marts.
- Generar reportes globales o por secciones que reflejen el estado del Banco Internacional versus la competencia actual.
- Definir perfiles de acceso mediante políticas de seguridad de acuerdo a necesidades de los usuarios.

CAPÍTULO 2

MARCO TEÓRICO

2.1 Banco Internacional

2.1.1 Introducción

“Banco Internacional cuenta con una estratégica y una renovadora estructura organizacional, garantizando los recursos, tecnología y la infraestructura necesaria para satisfacer las necesidades de una sociedad cada vez más demandante y de un mercado global más competitivo.” (Banco Internacional, 2014)

“Actualmente cuenta con más de 400.000 clientes, con 87 puntos de atención a nivel nacional y 380 cajeros automáticos.” (Banco Internacional, 2014)

2.1.2 Historia

Banco Internacional fue creado el 29 de Mayo de 1973, la oficina matriz inicio sus operaciones con 19 empleados en Quito. En 1976 instala el primer sistema de computación.

“Banco Internacional es el primer Banco nacional en incorporarse a la red mundial Swift” (Banco Internacional, 2014) en el año de 1981. Se consagra por ser el Primer Banco en tener presencia en la región Oriente, aperturando su oficina en el Lago Agrio, dando soporte a economías agrícolas, ganaderas, comerciales y contribuyendo así a las empresas petroleras y compañías de servicios.

2.1.3 Misión

“Gestionar recursos y servicios financieros con alto grado de calidad, a fin de obtener la satisfacción de nuestros clientes, aportar al desarrollo nacional y crear valor para nuestro capital humano y nuestros accionistas.” (Banco Internacional, 2014)

2.1.4 Visión

Ser el mejor Banco del Ecuador:

- Contaremos con el mejor talento humano.
- Brindaremos servicios de la más alta calidad.
- Alcanzaremos altos estándares de eficiencia y productividad.
- Incrementaremos las capacidades sostenida y diversificadamente.
- Gestionaremos con prudencia la cartera de crédito.

2.2 Antecedentes de la investigación

Para realizar este proyecto se han revisado fuentes de información como libros, tesis, revistas, etc., que otorguen las bases teóricas para usar las prácticas eficientes en el desarrollo de sistemas de Inteligencia de negocios. Bajo el mismo enfoque la relación de la metodología a emplear son las que mejor se adaptan al desarrollo del presente proyecto.

2.3 Inteligencia de negocios

2.3.1 Introducción

Las organizaciones actualmente se desenvuelven en un mundo de negocios acelerado, esto obliga a las altas gerencias y directivos a buscar soluciones y estrategias que generen ventajas competitivas sobre sus rivales del mercado. La globalización es el principal impulsor de esta dinámica donde los mercados cambian rápidamente y las corporaciones necesitan disponer de más información pero menos tiempo para analizarla y tomar las decisiones correctas. Es aquí donde las tecnologías de información impulsan los avances en esta área y permiten que las empresas puedan tomar acción frente a sus problemas de manera más fácil utilizando la inteligencia de negocios.

Inteligencia de negocios (BI) es una aglomeración de sistemas y herramientas que permiten el análisis de los datos importantes para el proceso de proyección estratégica de la corporación. Estos sistemas permiten a una organización revelar indicadores de un

área específica de negocio, transformando la información en conocimiento, permitiendo tomar decisiones más rápidas y precisas.

La inteligencia de negocios combina un amplio conjunto de aplicaciones de análisis de datos, incluyendo el análisis de consultas, informes empresariales, procesamiento analítico en línea. La tecnología BI también incluye el software de visualización de datos para el diseño de tablas y otros de infografía, así como herramientas para la construcción de cuadros de mando que muestran los datos visualizados con las métricas del negocio e indicadores claves de rendimiento de una manera fácil de captar.

2.3.2 Conceptos de inteligencia de negocios

“Se entiende por Business Intelligence (BI) al conjunto de metodologías, aplicaciones, prácticas y capacidades enfocadas a la creación y administración de información que permite tomar mejores decisiones a los usuarios de una organización” (Curto Díaz, 2012, pág. 18)

Inteligencia de negocios es un término que hace referencia a sistemas de soporte de decisiones (Decision Support Systems), de información ejecutiva y gerencial” (Kimball, The Data Warehouse Toolkit, 2008). “Combina conceptos de compartimiento y almacenamiento de datos y conocimientos en gestión con herramientas de análisis, la presentación de información compleja y competitiva hacia los usuarios encargados de la toma de decisiones y planeamiento estratégico dentro de una empresa. (Kimball, The Data Warehouse Lifecycle Toolkit, 2008)

Una aplicación de BI puede ser tan simple como una herramienta de consultas o tan complejo como una minería de datos sofisticada o una aplicación de modelado. Las herramientas de consulta pueden ser utilizadas eficazmente por un pequeño porcentaje de la población que usa el sistema DWH/BI. La mayoría de los usuarios de negocios es probable que accedan a los datos a través de aplicaciones basadas en parámetros pre-construidos y plantillas que no requieren intervención de los usuarios

para construir directamente las consultas. Algunas de las aplicaciones más sofisticadas, pueden cargar los resultados de nuevo en los sistemas de código de operación, sistema de ETL, o área de presentación. (Kimball, The Data Warehouse Toolkit, 2008)

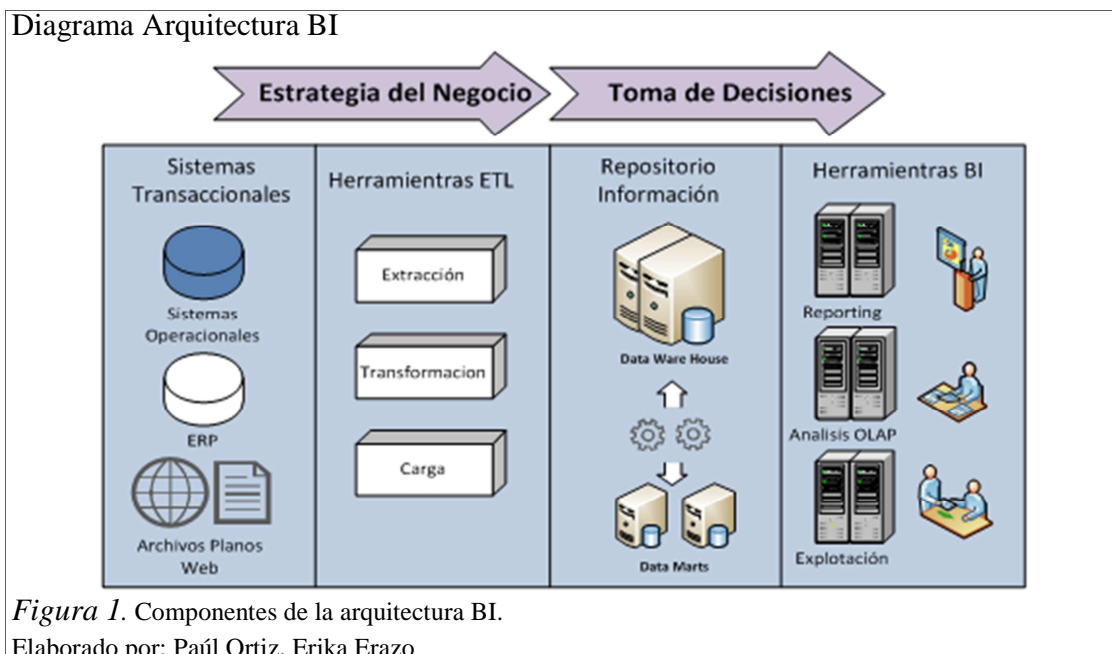
2.3.3 Importancia de la inteligencia de negocios

BI sirve para detectar eventos significativos e identificar, supervisar las tendencias de negocio, las aplicaciones BI buscan mejorar los procesos de toma de decisiones en todos los niveles de gestión y mejorar sus procesos tácticos y estratégicos de gestión.

Es importante que las organizaciones entiendan que es posible obtener beneficios de las nuevas tecnologías que se ofrecen para sistemas BI y que muy pronto se presentan como un activo básico y primordial para el funcionamiento de cualquier organización que quiera obtener ventaja competitiva sobre sus rivales del negocio.

2.3.4 Arquitectura inteligencia de negocios

Un sistema de inteligencia de negocios, se alimenta de varias fuentes de información, la arquitectura se divide en dos flujos: Estrategia del negocio y Toma de decisiones, como se indica en la siguiente ilustración.



La información se obtiene de: sistemas operacionales y transaccionales estructurados o no estructurados, que incluyen aplicaciones desarrolladas a la medida para la organización; fuentes de información externa (páginas web, estudios de mercado, estadísticas, etc.). Esta información es considerada estructurada pero ahora se puede ingresar información semiestructurada o desestructurada (imágenes, vídeos, correos electrónicos, etc.) gracias a nuevas tecnologías y programación avanzada de los sistemas.

Cabe recalcar que la información debe ser correcta y de gran calidad de manera que no se propaguen errores en presentación de la información a la organización y que luego será dificultoso localizarlos.

2.4 Metodologías de desarrollo BI

2.4.1 Definición de la metodología

Es necesario realizar un estudio comparativo de las dos metodologías más importantes, que son la de Ralph Kimball, y la de Bill Inmon, con enfoques distintos para la construcción de data warehouse (DWH), debido a la naturaleza del proyecto se analizará desde el punto de vista de la construcción de los data marts que forma parte de la solución propuesta.

2.4.2 Metodología Bill Inmon

Según Bill Inmon, “un almacén de datos es una colección de datos orientados por tema, integrados variables en el tiempo y no volátiles que se emplea como apoyo a la toma de decisiones gerenciales” (Inmon W. H., 2010).

Bill Inmon presenta una arquitectura de múltiples niveles en el cual el DWH recopila datos de diferentes fuentes, luego estos son integrados y finalmente distribuidos a subconjuntos denominados data marts (enfoque “Top – Down” “de arriba abajo”). En la ilustración se observa la distribución del DWH a los data marts.

Diagrama Metodología Bill Inmon

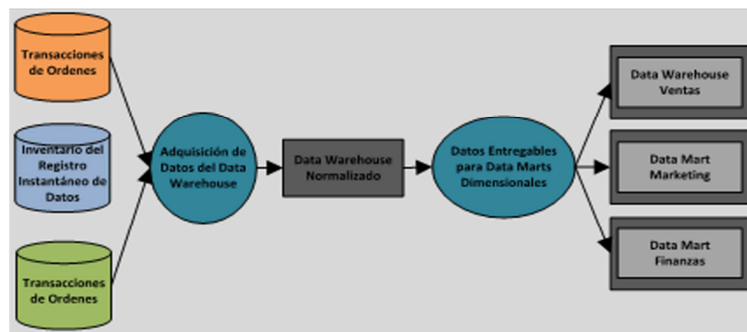


Figura 2. Modelo Bill Inmon

Elaborado: Paúl Ortiz, Erika Erazo

Además ve la necesidad de trasladar los datos obtenidos de diferentes sistemas transaccionales de las organizaciones a un lugar centralizado donde la información pueda ser consumida para el análisis en el CIF (Corporate Information Factory).

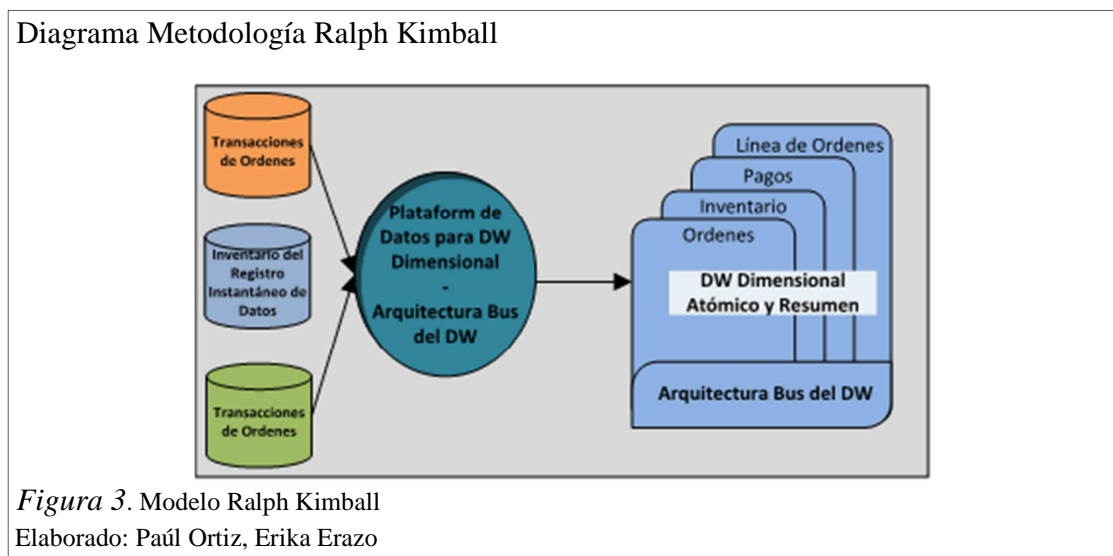
La información debe estar al mayor detalle permitirá el nivel más bajo de granularidad. La data marts departamentales son tratados como subconjuntos de este DWH corporativo, son construidas para cubrir necesidades individuales de análisis de cada departamento, y siempre consumen la información de un DWH centralizado.

El data warehouse es cargado con procesos ETL (extracción, transformación y carga), con datos obtenidos de los sistemas operacionales, previamente dichos datos deben pasar por un complejo proceso de conversión, donde son validados y consolidados en el DWH corporativo, una vez realizado este proceso comienza la actualización de los data marts departamentales, en ellos se organizan los datos de estructuras particularmente requeridas por cada uno de ellos.

El DWH corporativo presentado por Bill Inmon principalmente debe ayudar a las necesidades de todos los usuarios en la organización y no solo a un grupo en particular. Se trata de un método que minimiza los problemas de integración pero es costoso y de gran esfuerzo debido a que se trabaja con una gran cantidad de datos. Esta metodología es utilizada para proyectos de gran envergadura y para grandes corporaciones.

2.4.3 Metodología de Ralph Kimball

El modelo presentado por Ralph Kimball deja de lado la necesidad de un data warehouse centralizado debido a que la mayoría de usuarios desea obtener información más detallada. Además argumenta que la mejor práctica es almacenar los datos en data marts independientes y utilizando dimensiones lógicamente conectadas. Kimball propone el modelo de datos tipo estrella, lo cual permitirá la optimización de consultas y facilidad de uso de los data marts.



Esta metodología también es conocida como “Bottom-up” “de abajo hacia arriba”, y se basa en la construcción de prototipos que posteriormente son data marts independientes, conformados de una manera común a través de la estructura de bus, la unión de los mismos diseñan el DWH corporativo. Con esta característica se hace más flexible y sencilla de implementar, ya que se puede construir una data marts como primer elemento del sistema de inteligencia de negocios, y luego ir añadiendo consecutivamente otros de acuerdo a los requerimientos de cada área.

En este sistema, los procesos ETL extraen información de los sistemas operacionales y los procesan, cargando los data marts de una forma individual, teniendo en cuenta la estandarización de las dimensiones.

Kimball propone 4 fases para la construcción del DWH que son:

- **Elegir el proceso de negocio.**

El primer paso es elegir el área a modelar. Esta es una decisión del patrocinador del proyecto, y depende fundamentalmente del análisis de requerimientos y de los problemas analíticos obtenidos de los usuarios.

- **Establecer el nivel de granularidad.**

La granularidad se trata de definir el nivel más bajo de detalle. La elección de la granularidad depende de los requerimientos del negocio y lo que pueda ser útil de los datos actuales.

- **Elegir las dimensiones.**

Las dimensiones surgen de los requerimientos del negocio discutidas por el equipo de trabajo, también son facilitadas por el nivel de granularidad previamente establecido.

- **Identificar medidas y las tablas de hechos.**

Este paso consiste en identificar las medidas que nacen de los flujos del negocio. Una medida usa criterios definidos en las dimensiones, es un atributo de una tabla que se desea analizar, sumalizando o agrupando sus datos.

2.5 Comparación de las metodologías

En base a lo expuesto sobre las metodologías más relevantes en el mercado para la construcción de sistemas de inteligencia de negocios, se realizó un estudio comparativo, el cual ayuda en la elección de la metodología a utilizar en el presente proyecto.

El estudio comparativo se basó en los siguientes aspectos:

- Metodología y Arquitectura, Modelado de Datos, Dimensiones, Filosofía

Tabla 1. Estudio comparativo de las metodologías Kimball vs Inmon

Metodología y Arquitectura		
	Inmon	Kimball
Enfoque General	Top-down (Arriba - Abajo)	Bottom-up (Abajo – Arriba)
Estructura Arquitectónica	Data warehouse Centralizado alimentado por data marts departamentales	Data marts de acuerdo al modelo del negocio, se conforman de acuerdo a las dimensiones
Complejidad	Muy complejo	Muy simple
Modelado de Datos		
	Inmon	Kimball
Orientación de datos	Impulsada por los datos	Orientado por procesos
Herramientas	Tradicional (Diagrama entidad - relación)	Tradicional (Diagrama entidad - relación) Modelo dimensional
Accesibilidad del usuario final	Baja	Alta
Dimensiones		
	Inmon	Kimball
Periodo de tiempo	Continuo y discreto	Lentamente Continuo
Métodos	Banderas de tiempo	Dimensiones y llaves
Filosofía		
	Inmon	Kimball
Audiencia primaria	IT	Usuarios finales
Objetivo	Entregar la solución en base a métodos aprobados.	Ofrecer una solución transparente para el usuario final.

Nota. La tabla describe los aspectos principales que se debe cubrir en las metodologías
Elaborado por: Paul Ortiz y Erika Erazo

2.6 Elección de la metodología

De acuerdo a los resultados obtenidos en el estudio comparativo de las metodologías se elige la de Ralph Kimball para la gestión del proyecto, ya que proporciona una dirección muy versátil y una serie de herramientas prácticas que ayudan a la implementación de un sistema gerencial para la toma de decisiones en un menor tiempo.

Se acopla a las necesidades porque se pueden implementar data marts independientes en áreas específicas (sistema financieros, captaciones y colocaciones), con pocos recursos y

paulatinamente irlos integrándolos en un gran almacén de datos con un enfoque de menor a mayor.

2.7 Base de datos relacional

“Las bases de datos relacionales está conformada por tablas, las cuales representan a los campos y atributos en una estructura rectangular formada por filas y columnas.”. (Sánchez, 2010, pág. 12)

Cabe destacar que existe un proceso llamado normalización que permite manejar las bases de datos de una forma efectiva.

2.7.1 Tipos de relaciones en un modelo entidad – relación

Una relación es la unión de tablas en ambos sentidos mediante atributos individuales, por ejemplo: una persona puede tener varios vehículos, esto se considera una relación de uno a varios. Se puede diferenciar tres tipos de relaciones:

- Uno a Uno
- Uno a Varios
- Varios a Varios

2.7.1.1 Ventajas

- Garantizan la duplicidad de registros, a través de campos claves o llaves.
- Garantizan la integridad referencial.

2.7.1.2 Desventajas

- Un data warehouse realiza la reestructuración de los sistemas operacionales, lo que implica altos costos económicos y esfuerzo en tiempo.
- Utiliza aplicaciones, sistemas y almacenamientos específicos.

2.7.2 Arquitectura de base de datos

El rol de la arquitectura de base de datos es el de separar los programas de aplicación de la base de datos física. El esquema de una base de datos se define en tres niveles de abstracción distintos:

- Nivel interno
- Nivel conceptual
- Nivel externo

2.7.3 Almacén de datos

2.8 Data warehouse

La definición de data warehouse (DWH) ha sido planteada por varios autores, estas definiciones enfatizan el uso de la información dentro de una empresa para apoyar a los objetivos estratégicos con la toma de decisiones gerenciales.

Existen dos grandes autores con respecto al tema data warehouse: Bill Inmon y Ralph Kimball.

“El data warehouse es una copia de las transacciones de datos específicamente estructuradas para la consulta y el análisis; es la unión de todos los data marts de una entidad.” (Kimball, The Data Warehouse Toolkit, 2008)

“Un data warehouse o depósito de datos es una colección de datos orientados a temas, integrado, no volátil, de tiempo variante, que usa para el soporte del proceso de la toma de decisiones gerenciales.” (Inmon W. B., 2008)

Además se destaca un concepto muy concreto sobre la definición de un DWH:

“Los DWH son sistemas que almacenan datos históricos para ser utilizados por los DSS (Decision Support System) para toma de decisiones estratégicas. Dichos sistemas son eminentemente de consulta y están enfocados a extraer conocimiento de los datos históricos”. (Mazón , Pardillo, & Trujillo, 2010, pág. 6)

Se deduce que un data warehouse, puede verse como un repositorio de datos, para lo cual se utilizan herramientas para extraer, transformar y cargar la información en el almacén de datos.

2.8.1.1 Función de un data warehouse

“Los almacenes de datos no solo ofrecen información mejorada sino que también les facilita obtenerla a los responsables de tomar las decisiones. Incluso contiene la capacidad de modelar y remodelar los datos.” (Laudon & Laudon, 2010, pág. 238)

La función del DWH es obtener datos desde varias bases de datos de fuentes internas y externas, que almacenan información necesaria para grandes instituciones, proveen un ambiente para que las corporaciones realicen una mejor evaluación de la información que está siendo administrada por diversas aplicaciones operacionales.

Los sistemas para el soporte de decisiones gerenciales basadas en un DWH, pueden ser más prácticas con la explotación de datos, con el fin de ayudar en el estudio de la eficacia del negocio.

Las principales características que tiene un data warehouse son:

- **Orientado al tema**

Permite clasificar la información en base a los objetivos estratégicos planteados por la empresa.

- **Integrado**

En un ambiente data warehouse la información encontrada al interior está siempre integrada y se muestra de muchas maneras.

- **Tiempo variante**

Registra los cambios que se producen a lo largo del tiempo, reflejando los datos originales en cualquier momento.

- **No volátil**

La información es solo de lectura no se puede eliminar ni modificar una vez almacenada, siempre y cuando se encuentre estable.

2.8.1.2 Elementos de un data warehouse

2.8.1.2.1 Tipos de tablas de hechos

Una tabla de hechos es la representación de un proceso de negocio que permite guardar dos tipos de atributos diferenciados:

- Medidas del proceso
- Claves foráneas hacia registros de una tabla de dimensión.

Existen diferentes tipos de tablas de hechos:

- Transaction fact table constituye eventos que ocurre en un determinado tiempo, permite estudiar los datos con el máximo detalle.
- Factless fact table /coverage son tablas que no tienen medidas.
- Periodic snapshot fact table son tablas usadas para almacenar información de forma periódica con intervalos de tiempos regulares.
- Accumulating snapshot fact table representa un ciclo de vida que debe tener un principio y un final.

2.8.1.2.2 Tipos de dimensiones

Las dimensiones recogen los puntos de análisis de un hecho, estos se pueden clasificar como, SCD que son siglas de Slowly Changing Dimension, y se refiere a la política de actualización de datos de una dimensión.

- **SCD Tipo 0**

Nunca se cambia la información, ni se reescribe.

- **SCD Tipo 1**

Simplemente sustituye el valor de atributo en la fila dimensión, reemplazándolo con el valor actual. De este modo, el atributo siempre refleja la asignación más reciente.

Ventajas:

- Es apropiado si el cambio de atributo es una corrección.
- También puede ser apropiado si no hay ningún valor en la descripción.
- Es fácil de implementar.
- Este tipo es fácil de mantener y a menudo se utiliza para los datos de los cambios que son causadas por las correcciones de procesamiento.

Desventajas:

- No mantiene un historial de valores de los atributos anteriores.
- Se pierde toda la historia de los cambios de atributos.

- **SCD Tipo 2**

Es la principal técnica para el seguimiento de precisión cambiando lentamente atributos de dimensión, se recomienda tener claves subrogadas que es un identificador único que se asigna a cada registro de una tabla de dimensión.

Ventajas:

- Cambios de atributos históricos crean nuevos registros en lugar de actualizar las existentes.
- Es poder analizar todos los cambios de dimensión según se requiera.
- Crea una nueva fila de dimensión con una nueva clave principal de una sola columna para identificar unívocamente el nuevo perfil de producto.

Desventaja:

- Acelera el crecimiento tabla de dimensiones. Por lo tanto, puede ser una técnica inadecuada para las tablas de dimensiones que ya superan un millón de filas.

- **SCD Tipo 3**

Cambios de atributos fijos indican que el valor de la columna no debe cambiar, agrega una columna Dimensión.

- No agrega una nueva fila dimensión, sino que se añade una nueva columna para capturar el cambio de atributo.
- Permite ver nuevos e históricos datos de hechos por cualquiera de los nuevos valores de los atributos o antes.

- **SCD Tipo 4**

Se le conoce habitualmente como tablas históricas

- **SCD Tipo 6**

Es la combinación de los tipos 1,2 y 3.

2.8.1.2.3 Tipos de métricas

Una métrica son medidas que se utilizan para medir el avance de un proceso, existen dos tipos de métricas:

- **Métricas de realización de actividad (leading):** Permite medir la realización de la actividad. Por ejemplo, la participación de una institución financiera en el Sistema Financiero Nacional.

- **Métrica de resultados de actividad (lagging):** recogen los resultados de una actividad. Por ejemplo, el porcentaje de participación de una institución financiera en el Sistema Financiero Nacional.

2.8.1.3 Elementos de una arquitectura data warehouse

Los elementos de la arquitectura DWH, representan la estructura total de datos, y tienen una relación con los siguientes componentes: procesamiento, comunicación y presentación, que interactúan con los usuarios finales. A continuación se describe cada uno de los elementos:

- **Base de datos operacional**

Tienen información histórica que proveen una estructura de procesamiento eficiente, donde se almacenan todas las transacciones de la organización.

- **Nivel de acceso a la información**

Es el nivel donde incluye el hardware y software que permite al usuario final interactuar con las herramientas y aplicaciones desarrolladas para emitir reportes en hojas de cálculos y gráficos que normalmente usa para sus labores diarias. Por ejemplo: Excel, power pivot, crystal reports, etc

- **Nivel de acceso a datos**

El nivel de acceso a los datos se encarga de definir una interface entre las herramientas de acceso a la información y las bases de datos operacionales de forma universal. De tal manera que permite a los usuarios finales acceder a cualquier información en la empresa para su trabajo diario.

- **Nivel de directorio de datos (metadata)**

Permite almacenar todas las definiciones de los meta datos, es decir contiene información cuya finalidad es ayudar a la comprensión y navegación de los usuarios a través del DWH.

- **Nivel de gestión de procesos**

Permite realizar diversas tareas para construir y mantener el DWH actualizado.

- **Nivel de mensajes de aplicación**

El nivel de mensajes de aplicación permite circular la información a toda la red de la institución.

- **Nivel data warehouse (físico)**

Es un repositorio central de información donde se almacenan de forma que sea fácil de acceder.

- **Nivel de organización de datos**

Es el componente final de un DWH, incluye los procesos replica necesarios para seleccionar, editar y cargar datos en un DWH para acceder a los datos desde una base de datos operacional.

2.8.1.4 Arquitectura de un data warehouse

Existen principalmente tres enfoques en la arquitectura de un data warehouse:

Data warehouse virtual federado (Enterprise bus architecture)

También conocido como MD (Arquitectura multidimensional) consiste en una arquitectura basada en data mart independientes federados que puede hacer uso de una herramienta EII (Enterprise Information Integration).

EII (Enterprise Information Integration). es una tecnología de integración de datos que proporciona una visión común y consistente. Este tipo de técnicas permite a las

aplicaciones el acceso de datos dispersos, los cuales recogen resultados individuales, los combinan adecuadamente y devuelven el resultado de la consulta.

Características:

- Transparencia al acceso de datos.
- Integración de datos de diferentes fuentes.
- Rendimiento.
- Posibilidad de federar cualquier fuente de datos.

Corporate information factory (Enterprise data warehouse)

Consiste en una arquitectura en la que existe un data warehouse corporativos y unos data marts dependientes del mismo, que permite describir y categorizar los almacenes de información utilizados para operar y administrar infraestructuras robustas de BI.

Enterprise data warehouse

Permite separar la información por la edad de la misma y se clasifica por su uso, se caracteriza por incluir información estructurada como no estructurada y poder satisfacer las necesidades del usuario.

Esquemas para estructurar los datos

- **Esquema en estrella**

Se caracteriza por tener una tabla de hechos en el centro de su estructura, a la cual llegan las tablas de dimensiones mediante claves primarias simples de cada dimensión. La tabla de hechos tiene múltiples relaciones con otras tablas mediante (foreign keys), las tablas de dimensión se encuentran desnormalizadas (la data relacionada a una dimensión se guarda en la misma tabla). A continuación se muestra en la figura un diagrama tipo estrella.

Diagrama Tipo Estrella

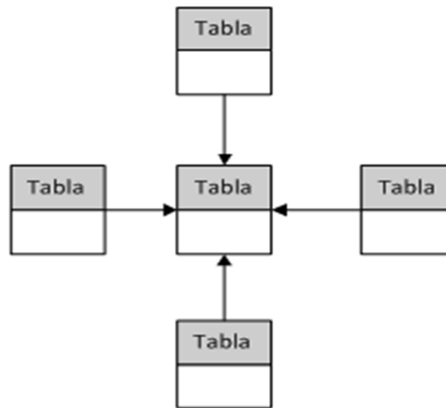


Figura 4. Esquema tipo estrella y tabla de hechos

Elaborado: Paúl Ortiz, Erika Erazo

- **Esquema copo de nieve**

Es un esquema basado en el tipo estrella, las tablas de dimensiones se normalizan en varias tablas, permitiendo tener varios caminos para llegar a los datos influyendo directamente sobre el rendimiento de las consultas. Por lo tanto la tabla de hechos deja de ser la única tabla del esquema que se relaciona con varias tablas, donde se crea una tabla por cada nivel de la jerarquía.

Diagrama Copo de Nieve

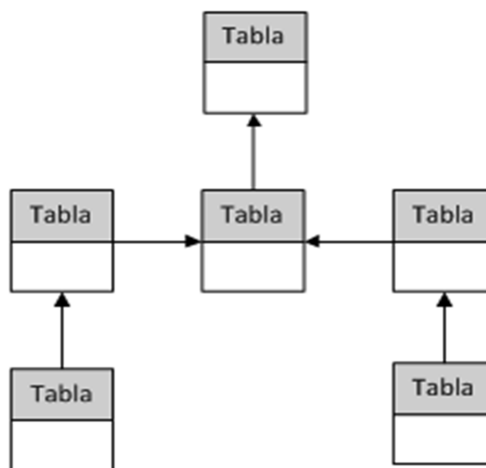


Figura 5. Esquema Copo de nieve y tablas jerárquicas

Elaborado: Paúl Ortiz, Erika Erazo

2.8.2 Data mart

2.8.2.1 Definición data mart

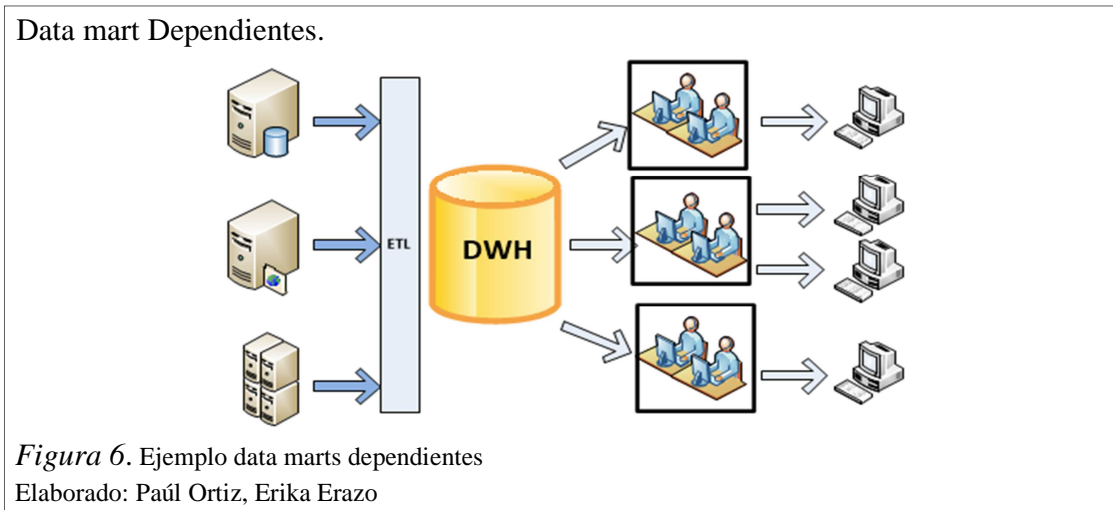
Los data marts son base de datos departamentales es decir subconjuntos de datos que ayuda a dar soporte a un departamento o unidad de negocio en particular, se caracteriza por tener una estructura óptima de los datos que puede ser llenado desde una base de datos transaccional y los que formaran parte de un data warehouse.

2.8.2.2 Tipos de data mart

Existen dos tipos de data mart que se detallan a continuación:

- **Dependientes**

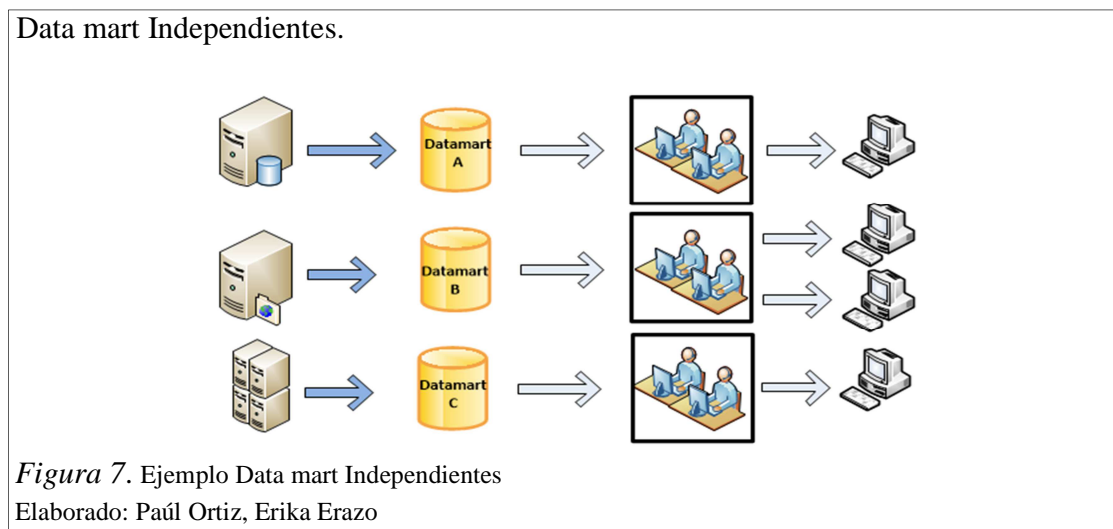
Son los que se constituyen a partir de un data warehouse central, es decir reciben sus datos de un repositorio.



- **Independientes**

Son aquellos data mart que no dependen de un data warehouse central, ya que pueden recibir los datos directamente sea mediante procesos internos de la fuente de datos o

mediante almacén de datos operacionales (ODS) es sistema que recopila información existente en los sistemas operacionales y ofrece a los usuarios finales.



2.8.2.3 Carga de datos en data mart

Para la carga de información de un data mart se puede utilizar técnicas de carga para la herramienta OLAP, una de las principales es Microsoft SQL Server mediante DTS (Data Transformation Services), con los cuales se puede realizar la carga con solo escoger la fuente y el destino.

Ventajas y desventajas

Ventajas

- Optimiza el tiempo de respuesta para obtener más rápido los datos solicitados por el usuario.
- El costo que implica la construcción de un data mart son menores a los de implementación de un data warehouse.
- Permite generar un data mart por áreas específicas.
- Permite realizar consultas sencillas.
- Permite llevar un historial de la información.

Desventajas

- No se adapta a altos volúmenes de información.

2.8.3 Cubos OLAP

Los cubos OLAP (**Online Analytical Progressing**) se crean en función a bases de datos multidimensionales, que permiten procesar grandes volúmenes de información.

2.8.3.1 Formas de almacenamiento de los cubos

- **MOLAP (Multidimensional Online Analytical Processing)**

La información del cubo son almacenados junto con sus agregaciones, en una estructura multidimensional que provee alto rendimiento y comprensión de datos.

- **ROLAP (Relational Online Analytical Processing)**

Toda la información del cubo, sus datos, agregación son almacenados en una base de datos relacional.

- **HOLAP (Hybrid Online Analytical Processing)**

Este tipo de almacenamiento es un híbrido de Molap y Rolap, la agregación de los datos es almacenada en una estructura multidimensional usada por Molap.

CAPÍTULO 3

METODOLOGÍA DE DESARROLLO

3.1. Metodología de Ralph Kimball

Debido a su gran enfoque para el desarrollo del proyecto la metodología establecida por Ralph Kimball establece ciertos procesos para llevar al éxito la construcción de la solución BI, formada por un DWH. Para su desarrollo se incluyen varias tareas que pueden ser realizadas en paralelo o en forma secuencial.

El correcto desarrollo de cada una de las fases planteadas en esa metodología garantiza la calidad y el correcto proceso de construcción del proyecto.

Los proyectos BI tienen un único objetivo, que es el negocio, por lo tanto los puntos importantes a considerar y ser más relevantes son las necesidades del negocio las cuales guiarán a lo largo de todo el proyecto. En general se contempla como base el ciclo de vida dimensional presentado por Ralph Kimball con el nombre Business Dimensional Lifecycle (BDL) (Kimball, The Data Warehouse Lifecycle Toolkit, 2008), este enfoque de implementación es ilustrado en la siguiente figura.

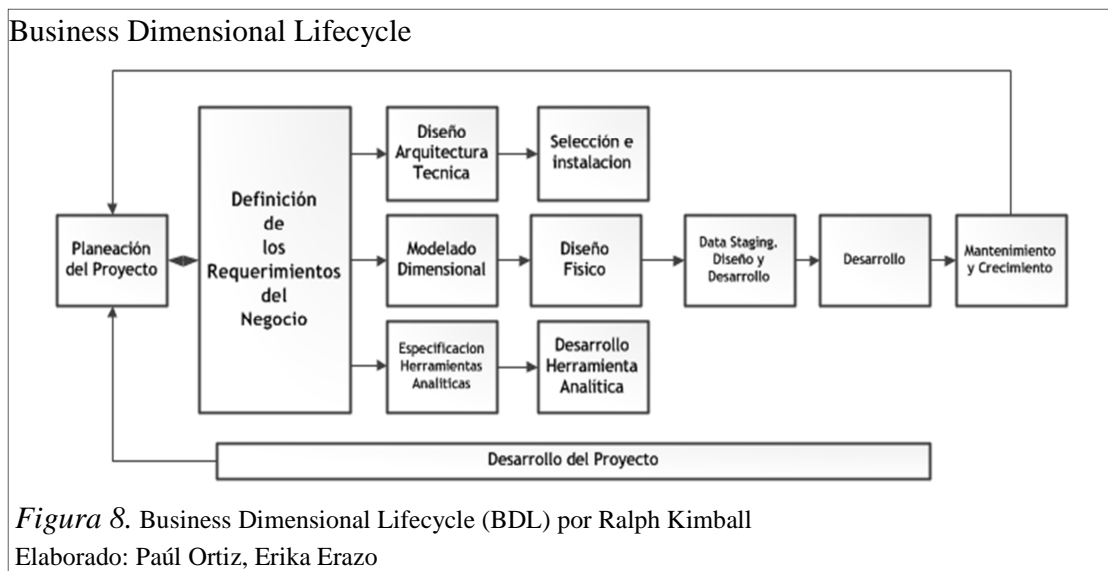


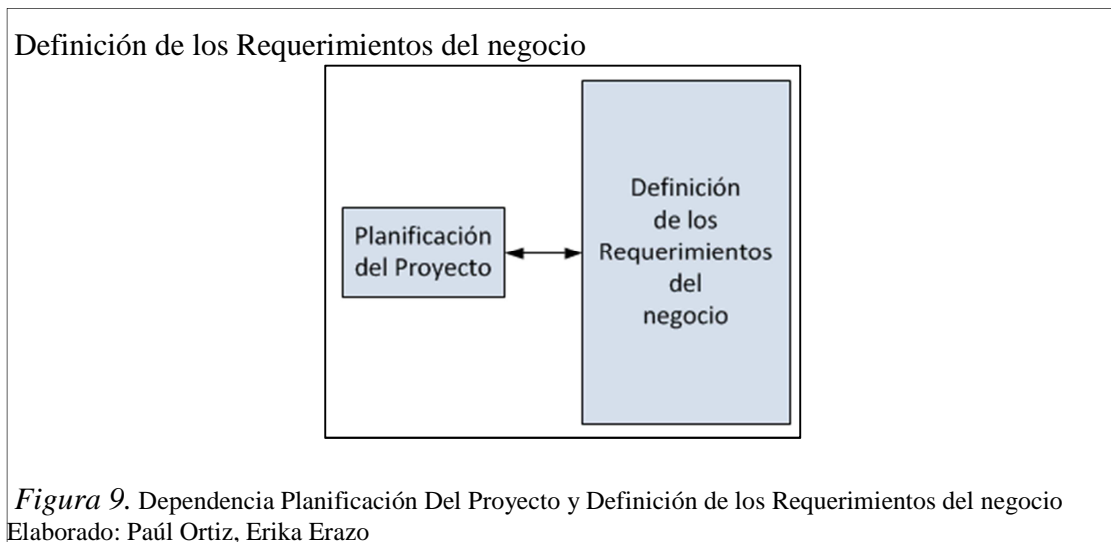
Figura 8. Business Dimensional Lifecycle (BDL) por Ralph Kimball

Elaborado: Paúl Ortiz, Erika Erazo

El diagrama muestra el encadenamiento de tareas de alto nivel requeridas para el diseño efectivo, desarrollo e implementación del proyecto BI. Es importante aclarar que el enfoque Kimball sobre BDL (Business Dimensional Lifecycle) se focaliza en secuencialidad y concurrencia y no intenta reflejar un proyecto en término de tiempos y plazos.

3.1.1. Planificación del proyecto

La planificación del proyecto es la primera etapa del ciclo de vida dimensional que aborda la definición y el alcance del proyecto, se centra en las necesidades del personal incluyendo la evaluación y justificación del negocio con la asignación de tareas, la duración y secuencia que podrían tener.



3.1.1.1. Preparación de la organización.

En la actualidad implementar un proyecto BI representa importantes gastos económicos, por lo tanto para el banco Internacional se evaluó la preparación de la organización a través del test de Litmus, que permite identificar cinco factores para realizar el análisis actual de la institución.

Además indica que no es necesario tener una alta calificación en todos los factores para avanzar con el proyecto, pero cualquier déficit representa riesgos o vulnerabilidades.

A continuación se describen los factores en orden de rango de importancia.

1. Fuerte promotor del negocio

Es el factor más crítico, los patrocinadores de negocios de la institución bancaria tienen una visión y es capaz de convencer al personal de la institución del impacto potencial del proyecto sobre la organización.

2. Motivación comercial

Un proyecto de inteligencia de negocios necesita resolver problemas críticos que se detectó con el análisis de requerimientos con el fin de reunir los recursos necesarios para un informe exitoso. A continuación se detalla los requerimientos levantados para la elaboración del presente proyecto.

- La información se trabaja mediante archivos digitales y procesados manualmente, ocasionando posibles pérdidas en el flujo de información al no tener un repositorio centralizado, por lo que se puede identificar la debilidad al no presentar información óptima y consolidada.
- El tiempo de ejecución que demanda el procesamiento de importantes volúmenes de datos es muy extenso debido a la cantidad de información.

3. Factibilidad de datos

La factibilidad de datos es la más crucial para realizar una solución BI, se debe conocer las fuentes de información y la facilidad de acceso para las mismas.

Para el desarrollo de la solución de Inteligencia de negocios se va a tomar la información de instituciones financieras, captaciones y colocaciones de las siguientes fuentes:

- Página web de la Superintendencia de Bancos
Permite visualizar datos reales de forma mensual de las instituciones financieras accediendo a la siguiente página:
http://www.sbs.gob.ec/practg/sbs_index?vp_art_id=&vp_tip=6&vp_buscr=/practg/pk_cons_bdd.p_bal_entdd_finnc

- SRVDWH03 Servidor donde registra los valores de cartera y captaciones del Banco Internacional.

4. Relación área tecnológica IT y área de negocios

Existe una fuerte relación entre TI y el área de negocios por que se trabaja conjuntamente para tener la información actualizada y soportar todas las necesidades de los diferentes productos y servicios a través de estas gerencias.

5. Cultura analítica

Para medir la preparación de la institución Banco Internacional se aplicó el test de litmus, que cuantifica los factores de preparación que intervienen en el desarrollo e implementación de la solución de inteligencia de negocios. El test de litmus califica la disposición de la institución en cuanto a la solución de inteligencia de negocios, de acuerdo a los siguientes ítems clasificados por factor de preparación:

Tabla 2. *Test de Litmus*

Factor de preparación	Baja preparación	↔	Alta preparación
Fuerte promotor del negocio 55/60 (1 más bajo, 10 más alto)			
10	No muy respetado	↔	Alta influencia en la organización
9	Puede pasar varias semanas el equipo sin reunirse con el	↔	Altamente disponible al equipo
8	“Me pondré en contacto con usted acerca de eso”	↔	Decisiones rápidas sobre problemas
9	Espera que “tu hagas las cosas”	↔	Activo, apoyo verbal y visible, predispuesto a enfrentar problemas
9	“¿Puedes incluir a 150 usuarios para mañana, verdad?”	↔	Expectativas realistas
10	“¿Un data que?”	↔	Conoce de data warehouse
Motivación comercial 17/21 (1 más bajo, 3 más alto)			
2	“¿Y el punto es?”	↔	La supervivencia de la compañía depende del data warehouse
3	El financiamiento es un gran inconveniente	↔	El costo no es un problema – es inevitable realizar esto
3	Visión poco clara	↔	Visión claramente definida
3	10 visiones diferentes de la solución	↔	Visión firme de la solución
1	Problema táctico	↔	Problema estratégico
3	Oportunidad de bajar costos	↔	Oportunidad de aumentar ganancias

2	Imposibilidad de cuantificar el retorno	↔	Gran retorno
Factibilidad de datos 6/9 (1 más bajo, 3 más alto)			
2	Un data warehouse requeriría la compra de abundante tecnología	↔	La organización posee infraestructura tecnológica fuerte
2	Todo el mundo está embargado en otro proyecto de gran tamaño	↔	Existe gente experimentada disponible
2	No existirá datos confiable hasta después de la implementación del ERP	↔	Existen datos con la calidad suficiente disponibles.
Relación área tecnológica IT y área de negocios 4/5 (0 más bajo, 1 más alto)			
1	El área de negocios busca consultoría sin el conocimiento del área de TI	↔	Las áreas de TI y de negocios trabajan hombro a hombro
1	El área de negocios crea un equipo alterno de tecnología para construir el data warehouse	↔	El área TI está muy compenetrada con el área de negocios.
0	“No podemos confiar en los números de sus sistemas”	↔	Fuerte confianza en los reportes existentes
1	Toma años obtener un reporte nuevo	↔	Rápida respuesta del área de tecnología a los requerimientos de información
1	Los usuarios ya ni siquiera envían requerimientos	↔	Lista de espera corta para los requerimientos
Cultura analítica 5/5 (0 más bajo, 1 más alto)			
1	Las decisiones se toman en base a instintos	↔	Las decisiones se toman en base a datos e información
1	Los usuarios no solicitan información	↔	Los usuarios del área de negocios ruegan por acceso a los datos; “Solo denme los datos y yo me las arreglo”
1	A los usuarios no les gusta los reportes actuales	↔	Los reportes actuales son constantemente usados en hojas de cálculo para análisis de tendencias históricas
1	Los usuarios necesitan que sus secretarías impriman sus correos	↔	Los usuarios tienen un conocimiento computacional razonable
1	El área de finanzas es muy posesivo acerca de la información	↔	La información es compartida libremente a través de la organización
TOTAL 87/100			

Nota. El test se basa en un total de 100 puntos. Fuente: (Kimball, The Data Warehouse Lifecycle Toolkit, 2008).

Elaborado: Paúl Ortiz, Erika Erazo

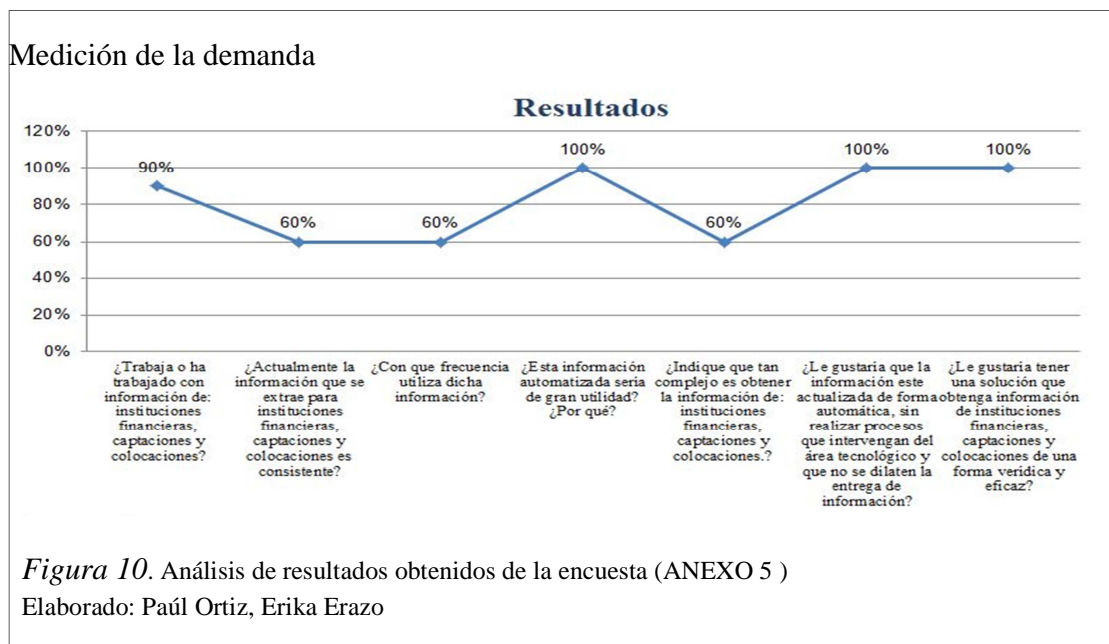
Los resultados del “test de Litmus” (Kimball, The Data Warehouse Lifecycle Toolkit, 2008) arrojó un resultado del 87/100, en cual indica la buena predisposición de la organización para el desarrollo e implementación de una solución de inteligencia de negocios.

3.1.1.2. Medir la Demanda

Para identificar en el Banco Internacional si existe demanda para aplicar una solución de inteligencia de negocios se analiza los tres casos que Kimball propone y son los siguientes:

- **En busca de demanda:** se presenta cuando la demanda no existe, para la institución financiera existe más de un fanático de los negocios que estaría interesado en la elaboración de una solución BI.
- **Demanda por parte de un fanático de los negocios en solitario:** Banco Internacional realiza una serie de análisis y reportes para generar información relevante, por lo tanto existe más de un fanático interesado en la solución propuesta.
- **Gran demanda:** para la institución financiera existe más de un solo fanático, puesto que en el área de control financiero está compuesto por seis personas las cuales son los responsables de generar análisis y reportes para la toma de decisiones de alta gerencia.

Para medir la demanda de información de: instituciones financieras, captaciones y colocaciones se realizó unas series de encuestas a un determinado grupo de usuarios en el área de control financiero.



Los resultados obtenidos desplegó que existe una gran necesidad para los usuarios del área de control financiero obtener información de calidad y de fácil acceso para realizar sus diferentes análisis y poder tomar decisiones oportunas para la institución.

3.1.1.3. Determinación del alcance preliminar.

El alcance acordado con la Institución Banco Internacional sobre el proyecto de inteligencia de negocios, tiene la siguiente definición.

Desarrollar e Implementar una solución de “Inteligencia de negocios” basada en información que otorga la Superintendencia de Bancos sobre ranking de instituciones financieras e información de captaciones y colocaciones por distribución de plaza.

La aplicación se enfoca en los distintos reportes gerenciales que se pueda obtener del ranking de instituciones financieras versus Banco Internacional, además de captaciones y colocaciones distribuidos por plaza, toda la información obtenida será motivo de análisis mensualmente.

Aplicando la solución de inteligencia de negocios se pretende reducir recursos y optimizar el tiempo en el tratamiento de la información por consiguiente, la solución planteada está basada en tres capas que se describe a continuación.

Capa 1 Relacional: se utiliza métodos de Extracción, Transformación y Carga, los mismo que alimentan los data mart conformados por base de datos relacionales de esta forma se obtiene información consistente y oportuna.

Capa 2 Multidimensional: los usuarios disponen de cubos de información OLAP (Online Analytical Progressing) para el análisis de consultas personalizadas que no sean previsto en el desarrollo de la solución.

Capa 3 Visualización: la solución cuenta con una serie de reportes mensuales predefinidos planeada para el apoyo de los ejecutivos facilitando el acceso a datos de calidad para sustentar la toma de decisiones.

3.1.1.4. Justificación del proyecto.

Debido a la gran cantidad de recursos que necesitan este tipo de proyectos se realiza una estimación de los costos y beneficios que evidencien el mejoramiento de la organización si se los implementa. La justificación puede tomar 2 enfoques de análisis que son:

- **Determinar las inversiones financieras y costos:** se realiza un estudio económico sobre las ganancias que genera el proyecto a la organización en caso de implementarlo.
- **Describir las necesidades insatisfechas que van a ser cubiertas:** de acuerdo a la definición de los requerimientos del negocio que nacen de los problemas de la organización y a su vez de las necesidades de usuario.

Hoy en día se vive una época en que la información es clave para tener una ventaja competitiva en el mundo de los negocios. Por lo tanto los usuarios que son encargados de desarrollar estrategias deben contar con información de fácil acceso y de alta calidad para la toma de decisiones.

Banco Internacional utiliza información de: instituciones financieras, captaciones y colocaciones, las mismas que es fundamental para desarrollar estrategias organizacionales a largo plazo que generen competitividad en el mercado.

Un factor determinante para realizar una solución de inteligencia de negocios es la automatización de reportes para reducir tiempos y evitar errores causados por manipulación manual de los datos. El mismo que es de gran utilidad para analizar la información obtenida, que ayuda para tomar las medidas necesarias dentro de la institución financiera e incrementar el número de captaciones y colocaciones.

3.1.1.5. Personal encargado del proyecto

Para el desarrollo del presente proyecto se determinó los siguientes roles y tareas, cabe recalcar que por ser un proyecto de titulación los tesistas asumen distintos roles en el proyecto:

Tabla 3. *Asignación de roles y responsabilidades*

Cargo	Área	Responsable	Funciones
Patrocinador de negocios	Control Financiero	Santiago Rodríguez	Cliente final del negocio
	Tecnología	Henry Pozo	
Conductor de negocios	Control Financiero	Santiago Rodríguez	Responsable de realizar levantamiento de necesidades para el desarrollo de la solución.
Líder de negocios	Tecnología	Henry Pozo	Personas con alto conocimiento del negocios e involucradas en el proyecto.
		Erika Erazo	
		Paúl Ortiz	
Usuario de negocios	Control Financiero	Karla Játiva	Usuarios finales que utilizan la solución BI.
		Jaime Aguilar	
		Iván Rodríguez	
		Edison Anchapaxi	
		Francisco Hidalgo	
		Santiago Rodríguez	
Analista de sistemas de negocio	Tecnología	Henry Pozo	Responsable de determinar las necesidades de la institución.
Desarrollador de aplicaciones analíticas	Tecnología	Erika Erazo, Paúl Ortiz	Responsable de diseñar y desarrollar el conjunto de plantillas analíticas.
Especialista Data Warehouse			Responsable de analizar los datos que deben tener las aplicaciones.
Director de proyecto			Persona con alto conocimiento del negocio y procesos internos de la institución.
Arquitecto técnico			Responsable de la arquitectura técnica y la seguridad en general.
Modelador de datos			Responsable de realizar el análisis de datos detallada y desarrollar el modelo de datos dimensional.
Administrador de base de datos			Responsable de la funcionalidad e integridad de la información en la base de datos
Coordinador de metadatos			Responsable de asegurar de que todos los metadatos se recogen, cogestione y difundan.

Nota: Las funciones pueden variar en base a las destrezas de los responsables.

Elaborado: Paúl Ortiz, Erika Erazo

3.1.1.6. Desarrollar el plan del proyecto

El Desarrollo del plan consiste en la identificación de todas las tareas y las estimaciones de esfuerzo necesarias para implementar la solución. El plan del proyecto debe identificar un punto de control de aceptación de los usuarios después de cada hito y entregables para asegurar que el proyecto está todavía en camino y que el negocio sigue siendo involucrado íntimamente. La metodología Kimball recomienda para un desarrollo del proyecto las siguientes etapas:

Organización y planificación

En esta etapa se define las actividades, esfuerzo y tiempo para el desarrollo del proyecto. Por la naturaleza del proyecto se asignado 2 recursos, se cuenta con disponibilidad de 20 horas/semana y desempeñan diferentes roles a lo largo del proyecto.

Plan de comunicación

El objetivo de esta etapa es considerar las reuniones del equipo, monitoreo del avance del proyecto, enfoque y estrategias de comunicación.

Con el fin de mantener informadas a las partes involucradas de los avances del proyecto y asegurar la intervención oportuna en caso que se requiera, se definió el siguiente plan de comunicación.

Tabla 4. *Plan de comunicación Banco Internacional*

Miembros	Frecuencia	Método	Información
Patrocinador	Mensual	Reunión informativa, presentación ejecutiva	Presentación de informes, entregables parciales
Área de Negocios	Semanal	Presentación ejecutiva	Implicaciones, expectativas de la comunidad, fechas críticas
Equipo del proyecto	Semanal	Reuniones de estado	Seguimiento de tareas
Colaboradores IT	Diaria	Reunión del personal IT	Gestión de las expectativas, necesidades de los recursos

Nota: La frecuencia es variante ya que los patrocinadores tienen actividades propias de su cargo.
Elaborado: Paúl Ortiz, Erika Erazo

Gestión de riesgos

Durante la planificación del proyecto se consideraron posibles riesgos, puesto que la solución está planeada para el área de control financiero y es la encargada de la planificación estratégica del banco, por lo tanto se fijaron los planes de respuesta específicos según probabilidad e impacto que se detallan a continuación.

Tabla 5. *Gestión de riesgos Banco Internacional*

Descripción del riesgo	Probabilidad	Impacto	Plan de respuesta al riesgo	
			Estrategia	Acción de respuesta
Indisponibilidad de recursos intermitente	ALTA	ALTO	Mitigar	Extender jornada laboral
Información inconsistente	MEDIA	ALTO	Revisar	Comparar con los archivos de origen las cifras totales
Integridad de datos	BAJA	ALTO	Verificar	Revisar la coherencia de los datos con respecto a las tendencias marcadas
Cambio requerimientos Cliente	MEDIA	MEDIO	Transferir	Re planificación de fechas críticas y tareas
Entorno y Plataforma de Desarrollo	ALTA	MEDIO	Evitar	Negociar el cambio del entorno y plataforma a tecnologías ya conocidas

Nota: Se pueden utilizar más de una estrategia para mitigar el riesgo
Elaborado: Paúl Ortiz, Erika Erazo

3.1.2. Definición de los requerimientos del negocio

El objetivo de ésta fase es definir todas las necesidades funcionales del negocio que puedan requerir los usuarios finales del DWH.

Los requerimientos identificados en ésta fase impactan casi en todos los aspectos de un proyecto BI. El resto de etapas que forman parte del proceso se nutren de la información definida en ésta etapa, cuánto más claros sean los requerimientos menos dudas surgen en el proceso de diseño e implementación a la hora de tomar decisiones sobre cómo dar solución a las necesidades requeridas por los usuarios. Por ello el enfoque de Kimball de la importancia de la definición de los requerimientos del negocio que se indica en la figura.

Importancia de los requerimientos del negocio



Figura 11. Dependencia de los requerimientos en todas las etapas

Fuente: (Kimball, *The Data Warehouse Lifecycle Toolkit*, 2008)

Elaborado: Paúl Ortiz, Erika Erazo

Para entender mejor los requerimientos se debe empezar por hablar con los usuarios del negocio que están relacionados con las funciones y la forma en la que los usuarios tomen decisiones, para ello se necesita convocar una serie de reuniones y entrevistas, lo cual permite comparar los requerimientos obtenidos con las fuentes de datos disponibles para lograr el soporte de estos requerimientos.

Para obtener la información de los requerimientos la metodología Kimball propone utilizar los siguientes métodos: selección de los entrevistados, entrevistas y análisis de las entrevistas.

Los requerimientos de un negocio es identificar las necesidades específicas que se presentan durante la entrevista al personal del área de control financiero al momento de requerir información. Los requerimientos consisten en brindar una gama de reportes de análisis gerenciales que permitan determinar oportunidades estratégicas.

En la siguiente tabla se detalle los requerimientos del negocio:

Tabla 6. *Definición de requerimientos del negocio.*

Tema analítico	Análisis o requerimiento	Proceso de negocio	Resultados
Principales Bancos	Posición del banco frente a las demás instituciones financieras.	Obtener información de mayor volumen de cartera o depósitos en un determinado momento	Información mensual
Competencia Bancaria	Información relevante sobre la diferencia de saldos y cuotas entre el banco y las demás instituciones financieras.	Obtener la variación de saldos, mensual y anual.	Depósitos monto sin interés. Depósitos monto con interés. Depósitos de ahorro. Depósitos de plazo restringido. Otros depósitos.
Captaciones Trimestrales	Analizar trimestralmente de captaciones.	Evolución de cada producto con respecto a las captaciones totales de la entidad financiera.	Información trimestralmente.
Morosidad	Analizar la cartera ampliada y la cobertura ampliada.	Indicar el grado en que el cliente no cumplió con una obligación. Mostrar las provisiones de cartera que mantiene una entidad.	Morosidad constara de dos reportes que son: Cartera Ampliada Cobertura Ampliada
Distribución por Plaza	Analizar la información por producto.	Obtener información de captaciones y colocación por producto, subproducto y región.	Información Trimestralmente.
Cuota de Mercado	Analizar los productos o servicios.	Conocer la parte del mercado que consume los productos o servicios.	Información mensual.
Solvencia	Analizar el nivel de patrimonio que posee una entidad para cubrir, sus activos y contingentes.	Determinar el grado en el que el patrimonio respalda el riesgo asumido en las distintas formas de inversión realizadas por el banco.	Información mensual.

Nota: Los requerimientos del negocio fueron obtenidos en base a entrevistas y reuniones
Elaborado: Paúl Ortiz, Erika Erazo.

3.1.2.1. Análisis preliminar de auditoría de los datos

Las instituciones financieras reportan sus balances generales hasta el décimo día laborable de cada mes a la Superintendencia de Bancos y Seguros, la misma presenta a la ciudadanía información transparente y actualizada, accediendo a su página web.

Por lo tanto la auditoría de los datos de: instituciones financieras e información de captaciones y colocaciones viene de una fuente confiable y con alta disponibilidad.

3.1.2.2. Criterios de éxito

- Disponer información histórica desde el 2010.
- Contar con un alto rendimiento de respuesta de consultas menores a 10 segundos.
- Permitir acceder a cubos de información para que los usuarios puedan realizar sus análisis personalizados para la toma de decisiones oportunas.

3.1.3. Diseño de la arquitectura técnica

En los sistemas de información el diseño de la arquitectura permite hacer un desarrollo más confiable y eficiente. Con la definición de la arquitectura se mejora la comunicación entre las diferentes áreas del proyecto, el planeamiento del proyecto, la flexibilidad y el mantenimiento del mismo.

Algunos equipos de trabajo no establecen la arquitectura y consideran que es una distracción y un obstáculo para el progreso de la construcción del DWH, y tienen la sensación de que las tareas son demasiado opacas, por lo que optan en omitir esta fase. Sin embargo, hay equipos de trabajo que dedica demasiado tiempo para el diseño arquitectónico.

Para mitigar los problemas que genera esta fase, Kimball en su metodología expone un proceso de 8 pasos en los cuales indica el correcto diseño arquitectónico sin extenderse demasiado en el tiempo.

Paso 1. Establecer un grupo de trabajo de diseño de la arquitectura

Es de gran utilidad establecer un pequeño grupo de dos a tres personas liderado por el arquitecto técnico encargado de trabajar con los datos del diseño. Es importante que todo el grupo se comprometa con las tareas que conlleva el diseño con el fin que vayan en la misma dirección.

Para el diseño de la arquitectura técnica se forma el siguiente equipo de trabajo:

Tabla 7. Equipo de trabajo diseño de la arquitectura

Rol	Nombre	Actividades
Director del proyecto	Erika Erazo, Paul Ortiz	Ingeniería de procesos, gestionar tareas, cronogramas, recolectar requerimientos del negocio
Arquitecto técnico	Erika Erazo, Paul Ortiz	Diseñar arquitectura SW y HW bajo los requerimientos obtenidos
Especialista DWH	Erika Erazo, Paul Ortiz	Estudiar infraestructura y tecnología existente
Desarrollador de aplicaciones analíticas	Erika Erazo, Paul Ortiz	Programación de herramientas de visualización

Nota: Las actividades son producto de los requerimientos del negocio
Elaborado: Paúl Ortiz, Erika Erazo

Paso 2. Requisitos relacionados con la arquitectura

La arquitectura se crea para apoyar las necesidades del negocio, el plan no es adquirir más productos que conlleven gastos económicos, sino crear un sistema que se acople completamente a los requerimientos del negocio.

Con el fin de descubrir las implicaciones arquitectónicas que llevan los requerimientos del negocio establecidos para el área de control financiero del Banco Internacional es necesario recolectar información provista por el área de TI y control financiero, en la siguiente tabla se evidencian los resultados obtenidos.

Tabla 8. *Implicaciones arquitectónicas Banco Internacional*

Requerimiento del Negocio	Técnica	Implicación arquitectónica	
		Back Room	Front Room
Información instituciones bancarias	Entrevista Checklist	Repositorio de fuentes externas Repositorio de fuentes internas Procesamiento ETL	Servidor de procesamiento OLAP Servidor de visualización reportería
Información captación y colocaciones	Entrevista Checklist	Repositorio de fuentes externas Repositorio de fuentes internas Procesamiento ETL	Servidor de procesamiento OLAP servidor de visualización reportería

Nota: Las implicaciones arquitectónicas están separas para cada requerimiento del negocio
Elaborado: Paúl Ortiz, Erika Erazo

Para el levantamiento de información se llevaron a cabo reuniones dentro de la organización para comprender la normativa vigente dentro del marco tecnológico, instrucciones técnicas previstas y los límites no negociables.

Paso 3. Documento de requisitos arquitectónicos.

Establecidos los requerimientos de negocio y llevado a cabo las entrevistas en este paso es necesario documentar las conclusiones.

Realizada las implicaciones arquitectónicas de los requerimientos del negocio para el proyecto BI a desarrollar en la institución se detectó el impacto que tiene estos en la arquitectura actual bajo los siguientes puntos:

- La información es provista por la SBS en archivos HTML descargada de la web del organismo, ya que se trata de una fuente externa es necesario contar con un repositorio en el cual se trabaja con información filtrada por el Banco.
- Se identificó una fuente interna de información que es provista por una base de datos relacional, que es utilizada por el área de control financiero para extraer datos históricos.
- Los procesos ETL, se construyen en base a las fuentes internas y externas identificadas, con el fin de limpiar la información que es almacenada en el DWH.

- El DWH es un repositorio centralizado con información de instituciones financieras, captaciones y colocaciones, previamente procesada por ETL's.
- El análisis de la información será provista por un servicio OLAP, el cual cuenta con una base de datos dimensional.
- Los cubos OLAP están disponibles para servicios de reportes previamente establecidos, o consultas dinámicas por usuario.

Paso 4. Desarrollo de un modelo arquitectónico de alto nivel

Definidos los requisitos de la arquitectura es hora de empezar a formular modelos para apoyar las necesidades identificadas.

Tabla 9. *Modelo arquitectónico de alto nivel Banco Internacional*

NIVEL DE DETALLE	Acceso a Datos (Qué)	Técnico (Cómo)		Infraestructura (Dónde)
		Back Room	Front Room	
Requerimientos del negocio y revisión	Se necesita Información de instituciones bancarias, captaciones y colocaciones comparadas con el Banco Internacional para mejorar el proceso de toma de decisiones	Los datos obtenidos de fuentes internas y externas procesados en un DWH centralizado, actualmente el proceso se realiza manualmente por los usuarios	El tiempo y recursos que toma realizar los reportes es muy extenso, retrasando la toma de decisiones gerenciales que contribuyen al desarrollo de la institución	El sistema requiere velocidad de procesamiento y menos intervención manual, el hw debe considerar disponibilidad y crecimiento del volumen de datos. Actualmente se cuenta con una infraestructura de hw flexible, y compatibilidad de plataformas
Arquitectura, Modelos y Documentos (Modelo Dimensional)	(Hechos y Dimensiones) Las entidades que componen la información son: fecha, institución producto y plaza, la relación que tienen es por: cuentas y saldos.	Los procesos ETL, tiene que estar certificados por el área de control financiero, realizando un validación periódica de la información. La información se almacena en 2 data mart, que conforman el DWH	Los usuarios cuentan con Cubos OLAP, para ver la información disponible por consultas personalizadas o reportes definidos primordiales que son: <ul style="list-style-type: none"> • Principales Bancos • Competencia Bancaria • Captaciones Trimestrales • Morosidad • Distribución por Plaza • Cuota de Mercado • Solvencia 	El hardware para para la solución requiere capacidad de procesamiento y aproximadamente un almacenamiento de 20 GB con un crecimiento mensual de 1GB. Los recursos de Infraestructura están disponibles y son provistos por el Área de TI del Banco Internacional.
Modelos detallados y especificaciones (Modelo lógico y físico)	Se cuenta con 2 fuentes: externa archivos HTML descargados de la página web de la SBS, interna servidor SRVDWH03 con datos históricos desde el 2010	En el mercado existen varios productos y estándares que cubren las capacidades necesitadas entre los más importantes: Microsoft Inteligencia de negocios, Oracle Inteligencia de negocios, Pentaho. Obligatoriamente se debe seguir el marco tecnológico y limitantes impuestas por el área de TI del Banco	Las plantilla de reportes esta provista por el área de control financiero, el acceso a los reportes es otorgado por perfiles de usuario con distintos permisos de acuerdo a la jerarquía del organigrama de la organización	Establecer un sistema siempre disponible con herramientas intuitivas que permitan al usuario con conocimientos básicos realizar sus reportes de forma sencilla y que demande menos tiempo
Implementación	La creación de bases de datos, índices, store procedimientos, etc. Se llevara a cabo con la nomenclatura que se utiliza actualmente en la institución	Analizar los mejores prácticas en la construcción de los procesos, extracción y carga de tal manera que no afecte al desempeño del sistema	Implementar los ambientes de reportaría y análisis, construir el grupo inicial de reportes y entrenar a los usuarios.	Instalar y probar los nuevos componentes de la infraestructura.

Nota: El modelo de alto nivel es establecido por el nivel de detalle.

Elaborado: Paúl Ortiz, Erika Erazo

Establecer reuniones de seguimiento e integración sobre el progreso del diseño ayudo en la recolección de información para satisfacer las necesidades del usuario ya establecidas en los requerimientos del negocio.

Paso 5. Diseño y especificaciones de los subsistemas

En este paso es fundamental hacer un diseño detallado de las necesidades de seguridad, así como la infraestructura física y sus configuraciones. Por otro lado las opciones de infraestructura, tales como el hardware del servidor y el software de base de datos, están predeterminados por la institución Banco Internacional y a procesos internos manejados por el área de TI.

Tabla 10. *Requisitos de la infraestructura*

Infraestructura	Técnicas	Análisis	Roles	Resultados
Back Room	Entrevista Checklist	Número de usuarios	Coordinador del proyecto arquitecto técnico	Usuarios que consumen la aplicación
		Compatibilidad de software		Compatibilidad de SW y HW
		Números de procesos de negocio		Consultas personalizadas, uso de recursos de infraestructura
		Recursos Económicos		Recurso monetario en caso de adquirir infraestructura
Front Room	Entrevista Checklist	Procesamiento del Sistema	Coordinador del proyecto arquitecto técnico	Comportamiento adecuado del sistema
		Crecimiento de datos		Proyección de crecimiento
		Cuellos de botella		Detectar procesos que lleven a cuellos de botella
		Conexiones simultaneas		Número de usuarios que puedan usar el sistema

Nota: Los resultados son en base a los análisis para cada infraestructura.
Elaborado: Paúl Ortiz, Erika Erazo

La tabla indica cuales son los requisitos mínimos que debe cumplir la infraestructura back-room y front-room además tienen que estar relacionados con las características propias del negocio. A continuación el estudio que se realizó para obtener los resultados.

Número de usuarios

El número de usuarios que usaran al aplicación esta predefinido por el área de control financiero, la solución BI que se plantea no tiene limitantes de acceso, esto depende de los procesos internos del área de seguridad para establecer los perfiles de acceso. El sistema contara con 3 roles principales a los cuales se pueden atar cualquier usuario del dominio.

- Rol Gerencial.
- Rol Jefatura
- Rol Control Financiero

Compatibilidad de software

La solución no tiene restricciones de compatibilidad de software, está construida en una plataforma convergente para los usuarios finales.

Números de procesos de negocio

Los procesos de negocio básicamente son los reportes establecidos en los requerimientos del negocio que son:

- Principales Bancos
- Competencia Bancaria
- Captaciones Trimestrales
- Morosidad
- Distribución por Plaza
- Cuota de Mercado
- Solvencia

Recursos económicos

Banco Internacional ha dispuesto que no se adquiera infraestructura software SW y/o hardware HW, ya que se cuenta con los recursos disponibles para implementar la solución BI.

Procesamiento del sistema

En condiciones normales las soluciones BI son capaces de soportar una carga masiva de trabajo que demande velocidad de procesamiento y tiempo de respuesta transparentes para los usuarios finales.

El estudio realizado en el Banco Internacional dio como resultado que la infraestructura HW provista por el área de TI, soportará la demanda que tenga la aplicación ya que se cuenta con servidores robustos y de alto rendimiento.

Crecimiento de datos

Hoy en día el crecimiento de los datos y donde almacenarlos es un problema que toda organización lo tiene, y más aún cuando se trata de información bancaria como es el caso del Banco Internacional.

En la actualidad el área de control financiero consume datos del servidor SRVDWH03, el tamaño aproximado de los datos es de 15 GB y aumentando. Ya que se ha establecido que el proyecto BI cuenta con información histórica desde el 2010, se contempla un tamaño total de 20 GB y un promedio de crecimiento de 1GB mensual, adicionalmente la solución BI ofrecen una estrategia eficiente, sostenible y a largo plazo para resolver los problemas del crecimiento masivo de datos.

Cuellos de botella

Se detectó los siguientes procesos que con llevan a un cuello de botella.

- La información entregada a gerencia es procesada manualmente por varios empleados.
- Recursos y tiempo que demanda generar la información para construir los reportes.
- Reproceso de la información por técnicas ineficientes de transformación y carga (ETL's).

Implementada la solución de BI, los cuellos de botella que se presentan actualmente se eliminan, otorgando fiabilidad a la información con menos tiempo de elaboración.

Conexiones simultáneas

La aplicación BI es altamente fiable por lo tanto los usuarios gerenciales que necesiten la información están conectados sin límites de sesiones.

Paso 6. Determinar las fases de aplicación de la arquitectura.

En ciertos casos no se puede poner en práctica todos los aspectos de la arquitectura técnica a la vez. Algunos no son negociables. Mientras que otros se pueden aplazar a fechas posteriores, con esto los requisitos de negocios para establecer las prioridades de la arquitectura.

Para el despliegue de la arquitectura técnica en el banco se ha establecido las siguientes fases:

Tabla 11. Despliegue de la arquitectura técnica

Fase	Actividad	Responsable
Requerimientos del negocio y revisión	Recopilar toda la información de los requerimientos del negocio que tenga impacto sobre la arquitectura actual	Director del proyecto, arquitecto técnico
Modelo Dimensional	Establecer hechos y dimensiones	Director del proyecto, arquitecto técnico, especialista DWH
Modelo lógico y físico	Identificar las fuentes que intervienen en los procesos de carga	Director del proyecto, arquitecto técnico especialista DWH
Implementación	Implementar la solución BI, manuales de usuario y documentación	Director del proyecto, arquitecto técnico, desarrollador de aplicaciones analíticas

Nota: Las fases establecidas son secuenciales y depende una de otra.
Elaborado: Paúl Ortiz, Erika Erazo

Paso 7. Documento de la arquitectura técnica.

El documento de arquitectura incluye información adecuada y útil de manera que los profesionales puedan proceder con la construcción del sistema.

2.8.3.2 Arquitectura back room

En el entorno de data warehousing, se conoce como back room al nivel que permite realizar procesos de data staging que consisten en el proceso de extracción, transformación y carga (ETL) desde la fuente de origen hacia el DWH.

Su principal objetivo es resolver cualquier tipo de problema que impida el traslado de los datos desde sus repositorios de origen hasta su destino (consultas).

Almacenes de datos del back room

Permite obtener repositorios de información permanente durante el proceso de extracción, transformación y carga.

La arquitectura back room del Banco Internacional está conformada por tres tipos de almacén de datos:

- **Sistema de origen:** es una de las principales fuentes de datos de la empresa llamados sistemas transaccionales que se encuentra en el servidor SRVDWH03 y archivos en Excel emitidos por la Superintendencia de Bancos.
- **Catálogo de metadatos:** es un repositorio que permite obtener información descriptiva del data warehouse.
- **Servidores de presentación:** son plataformas que se utiliza para almacenar los datos transformados que permite tener las consultas disponibles.

Arquitectura técnica back room

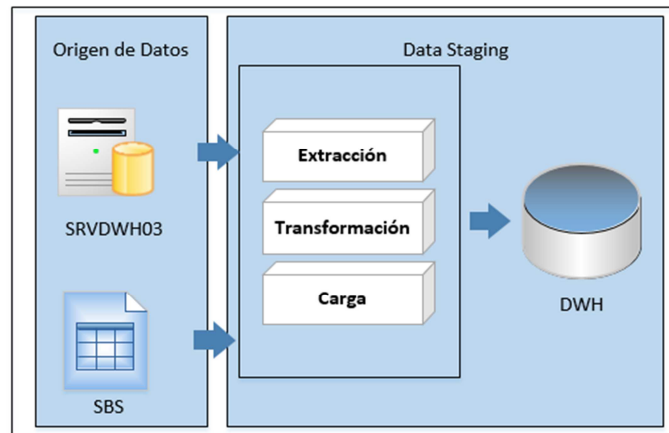


Figura 12. Componentes de la arquitectura técnica back room

Elaborado: Paúl Ortiz, Erika Erazo

Servicios del back room

Son herramientas y técnicas que se usan en los procesos del área temporal de datos.

Existen cuatro tipos de servicios principales:

- **Servicios de extracción de datos:** permite obtener los datos de sistemas fuentes.
- **Servicios de transformación de datos:** permite realizar cambios que los datos puedan requerir que pueden ser: generación de claves subrogadas, desnormalización, limpieza de datos, conversión de tipo de datos, etc.
- **Servicios de carga:** Permite realizar el proceso de carga para almacenar los datos en los sistemas destino para ser consumidos por los usuarios.
- **Servicios de control de tareas:** Ayuda a controlar los errores y permite incluir servicios de control de tareas como:
 - **Programación de tareas:** permite programar a una fecha u hora determinada los procesos de extracción y carga de información.

- **Monitoreo y registro:** permite monitorear la hora de inicio y finalización

2.8.3.3 Arquitectura front room

Permite explorar al máximo todas las funcionalidades y características que puede ofrecer el sistema en general, representa la capa de datos hacia el usuario. Entre las principales funcionalidades de este tipo de entorno son:

- **Servicios de navegación:** permite al usuario obtener de una forma ágil la información que necesita.
- **Servicios de acceso y seguridad:** permite asignar permisos a usuarios dependiendo de sus necesidades a través de autenticación o autorización de acceso.
- **Servicios de monitoreo de actividad:** permite realizar mejoras en los datos de la actividad del data warehouse a través del DBA (Data base administrator).
- **Servicios de manejo de consultas:** permite transformar las necesidades del usuario en consultas disponibles para el manejo de información, los siguientes son los más recomendables:
 - **Simplificación de contenidos:** permite generar consultas predefinidas e información de fácil entendimiento, limitando la visibilidad de campos y tablas.
 - **Reformulación de consultas:** permite analizar las peticiones de información por parte del usuario y estructurar de manera eficiente y de alto rendimiento.
 - **Redireccionamiento de consultas:** permite analizar las peticiones realizadas por el usuario y buscar en el árbol de metadata la ubicación de la información solicitada.
 - **Regulación de consultas:** permite limitar el tiempo máximo requerido de respuestas por cada consulta solicitada por el usuario.

Con el levantamiento de información obtenido en los distintos pasos, se ha establecido la arquitectura back room que se describe en la figura 6.

Arquitectura front room

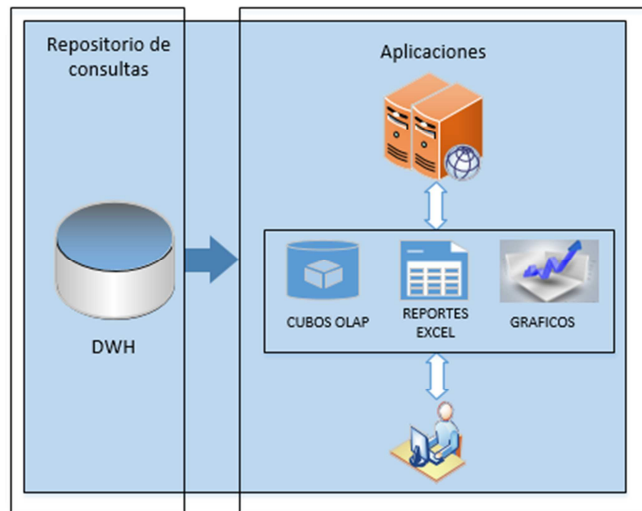


Figura 13. Componentes arquitectura front room

Elaborado: Paúl Ortiz, Erika Erazo

Paso 8. Revisar y finalizar la arquitectura técnica.

Terminado el plan de la arquitectura se debe notificar los participantes con detalle: equipo del proyecto, auspiciante y director. Tras la revisión, la documentación debe ser actualizada y utilizada inmediatamente en el proceso de selección del producto.

3.1.4. Selección e instalación del producto

Los proyectos de BI, que poseen entornos de DWH requiere del uso de productos y herramientas SW adicionales, esta fase permite al equipo de trabajo seleccionar la más óptima y su capacitación al personal para su manipulación.

Para la selección de los productos se debe hacer un riguroso estudio de cada una de ellas, mirar sus ventajas y desventajas y saber el alcance. Para la selección de productos la metodología Kimball recomienda elaborar una matriz de evaluación donde se menciona las características del producto de acuerdo a los requerimientos del negocio.

Esta fase puede verse reducida debido a que las empresas que desean implementar un sistema DWH ya tienen ciertos productos y herramientas predefinidos para usar, por lo cual el equipo de trabajo tiene que adaptarse al SW y HW provisto por la institución.

3.1.4.1. Posibles atajos para la evaluación

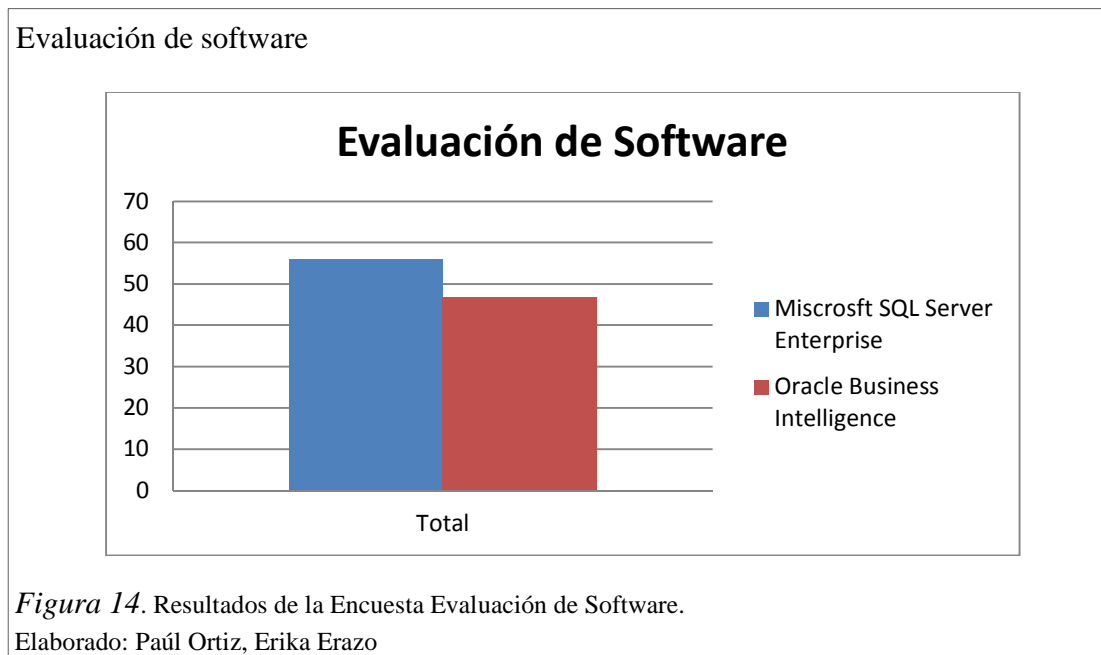
Consiste en determinar factores importantes que inclinen la balanza para la elección de un determinado producto, a continuación se detallan algunos.

Marco tecnológico: se basa en estándares de SW y HW de marcas específicas que impone la institución.

Recursos disponibles: utilizar las licencias que posee la organización como parte del proyecto, si se necesita la adquisición de un nuevo producto tiene que estar debidamente justificado

Experiencia del equipo: el equipo de trabajo ya cuenta con la experiencia de un determinado producto, esto evita capacitaciones sobre nuevos productos que lleven más tiempo en el desarrollo del proyecto.

Partiendo del marco tecnológico y los atajos expuestos se realizó una encuesta a los participantes del proyecto con el fin de tener una evaluación de los productos lo más realista posible a la situación de la organización, las preguntas son en base a los productos con mayor consumo en el mercado.



Los resultados de la encuesta sobre la selección de producto indican que el mejor producto para soportar el desarrollo de la solución BI y tomando en cuenta la experiencia del equipo es Microsoft SQL Server Enterprise.

La selección está basada en el diseño de la arquitectura técnica. De tal manera que la selección de productos y herramientas está definido de la siguiente manera:

- Plataforma HW
- SGBR: sistema gestor de base de datos
- Herramienta ETL
- Herramienta de consultas OLAP
- Herramienta de reporte

A continuación se describe los productos para cada herramienta:

Plataforma HW

La plataforma de hardware que dispone actualmente el Banco Internacional se caracteriza por escalabilidad, rendimiento y capacidad. Con estas primicias y en función de la disponibilidad de hardware se ha establecido la siguiente plataforma HW:

Marca: IBM XSeries 3550

CPU: 8 Core, Intel Xeon 2.60 GHz,

Memoria: 8GB

Sistema Operativo: Microsoft Windows Server 2003

Nombre: SRVDWH03

IP: 172.16.65.170

Sistema gestor de base de datos (SGBR)

Debido a los resultados de la evaluación del producto la suite Microsoft provee una plataforma SGBR, que por temas de licenciamiento de la organización se ha establecido SQL Server 2005.

Herramienta ETL

Tomando en cuenta la recomendación del proveedor de software (Microsoft), se decidió usar Integration Services 2005 por aspectos de compatibilidad que puedan afectar al desempeño de la solución BI.

Herramienta de consultas OLAP

Aprovechando el licenciamiento que posee la organización sobre la suite Microsoft SQL Server 2005, se utilizó el producto Analysis Services 2005 para la construcción de cubos y consultas OLAP (On-Line Analytical Processing).

Herramienta de reportes

Los cubos de información son accedidos desde Microsoft Excel ya que los usuarios están familiarizados con el uso del programa, además que la organización posee licencias para el producto.

Una vez, establecida la selección de los productos y licencias, se procede con la instalación de los mismos de acuerdo a las necesidades ya definidas en la arquitectura tecnológica documentando como se ha llevado a cabo y de acuerdo a procedimientos establecidos por el área de TI, de ser necesario justificar el por qué se difiere de instalaciones estándar.

Por último hacer énfasis en la verificación de la instalación sea correcta y que no exista problemas en los productos y herramientas para asegurar la correcta integración extremo a extremo del sistema DWH (permisos, usuarios, accesos, direcciones, licencias, etc.).

3.1.5. Modelo dimensional

El inicio de este modelo comienza con una matriz en donde se detalla la dimensionalidad de cada indicador y luego se especifican los diferentes grados de detalle que se acoplan al concepto del negocio, así como la granularidad de cada indicador y las diferentes jerarquías todo este conjunto forma el modelo dimensional del negocio o mapa dimensional.

3.1.5.1. Identificación de fuentes de datos

El equipo de trabajo es encargado de conocer cada data de información, entender el proceso y determinar la disponibilidad de cada uno con su respectivo encargado.

La metodología de Kimball recomienda tener las fuentes de información de forma ordenada, para ello recomienda usar el documento de definición de fuentes de datos.

Tabla 12. *Definición de fuentes de datos.*

Fuentes	Localización	Descripción
SBS	Página web Superintendencia de bancos	Información de instituciones financieras que reportan los bancos privados con las siguientes estructuras: <ul style="list-style-type: none"> • B11 Estados financieros. • B13 Balances diarios de instituciones financieras.
SBS	Página web Superintendencia de bancos	Información de captaciones y colocaciones dividida por plaza que reportan los bancos privados con la siguiente estructura: <ul style="list-style-type: none"> • B12 Detalle de captaciones y colocaciones.

Nota: SBS (Superintendencia de Bancos y Seguros), Fuente: (SBS, 2014)

Elaborado: Paúl Ortiz, Erika Erazo

La siguiente tabla detalla la definición de fuentes de datos y su respectiva ubicación de donde se obtiene la información para el proyecto.

3.1.5.2. Diseño de arquitectura del bus del data warehouse

Las dimensiones y los hechos conformadas permite la integración entre diferentes data mart, lo cual permite que no queden aislados entre sí y pueden ser usadas durante todo el proyecto, esto se le conoce como el bus del data warehouse.

Tabla 13. *Arquitectura de bus del data warehouse*

Dimensiones				
Data mart	Fecha	Producto	Institución	Plaza
Información de instituciones financieras	x	x	x	x
Información de captaciones y colocaciones	x	x	x	x

Nota: Arquitectura bus DWH para el banco internacional

Elaborado: Paúl Ortiz, Erika Erazo

La siguiente tabla detalla la arquitectura de bus del Data Warehouse, que tiene dimensiones conformadas y permite la integración entre los dos data mart que son: información de instituciones financieras, información de captaciones y colocaciones.

3.1.5.3. Proceso de modelado dimensional

De acuerdo a la metodología de Ralph Kimball se deben seguir los siguientes pasos, estos son secuenciales y el uno depende del siguiente. Los criterios a tomar en cuenta son:

1. El data mart
2. La granularidad de la tabla de hechos
3. Las dimensiones
4. Los hechos

A continuación se describe los pasos del proceso de modelado dimensional

3.1.5.3.1 Selección del data mart.

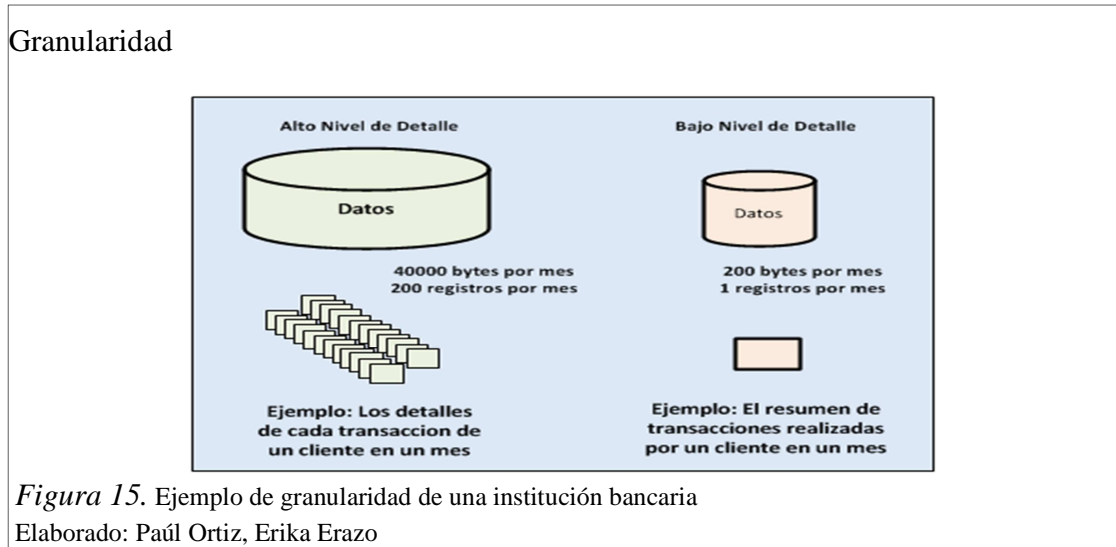
La solución BI para el banco Internacional consta de dos data mart que son:

- **Instituciones financieras:**
Permite conocer información real de cada institución del sistema financiero, que consiste en captar fondos del público e invertirlos en activos financieros, brindando una visión real de las organizaciones.
- **Información captaciones y colocaciones:**
Permite conocer información de captaciones y colocaciones de cada institución, la misma que se clasifica por: región, provincia y cantón.

3.1.5.3.2 Declaración de granularidad

La elección de la granularidad para el banco dependió de los requerimientos del negocio, se describe a continuación el grado de granularidad para cada data mart:

- Información de instituciones financieras: saldo por institución financiera, saldo por producto, saldo por mes.
- Información de captaciones y colocaciones: saldo por institución financiera, saldo por producto, saldo por mes y saldo por cantón.



3.1.5.3.3 Selección de dimensiones

Una forma de identificar las tablas de dimensiones es que sus atributos son posibles candidatos para ser encabezados en los informes, tablas pívot, cubos, o cualquier forma de visualización, unidimensional o multidimensional.

No se encontró tablas relevantes para ser utilizadas como tablas de dimensión, dichas tablas están construidas manualmente con la asesoría de la gente de negocios con el fin de establecer la información más relevante.

Información de captaciones y colocaciones CAPCOL:

- **Fecha:** consiste en la fecha que fue reportada la información a la Superintendencia de Bancos del Ecuador por las distintas instituciones financieras.

- **Institución financiera:** contiene las instituciones financieras que pertenecen al Sistema financiero nacional, agrupadas en bancos privados, públicos, cooperativas, mutualistas y sociedades financieras.
- **Plaza:** lugar geográfico en el que el valor del producto fue reportado, para esta dimensión el cantón es el nivel más bajo de granularidad seguido de las provincias y regiones.
- **Producto:** en esta dimensión se guardarán los diferentes productos que ofrece la entidad bancaria en captaciones y colocaciones.

Dimensión Fecha



Figura 16. Detalle dimensión fecha.

Elaborado: Paúl Ortiz, Erika Erazo

Dimensión Institución financiera

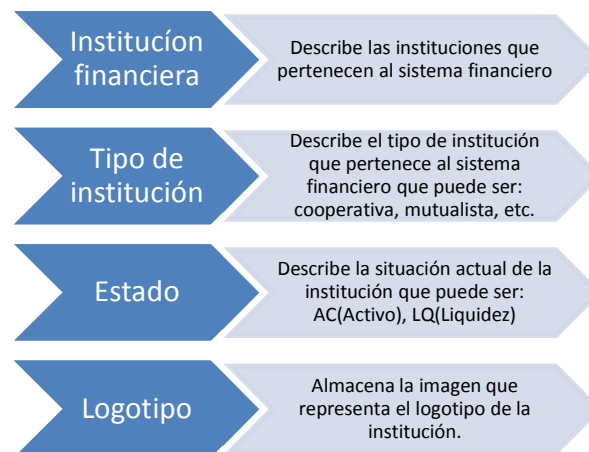
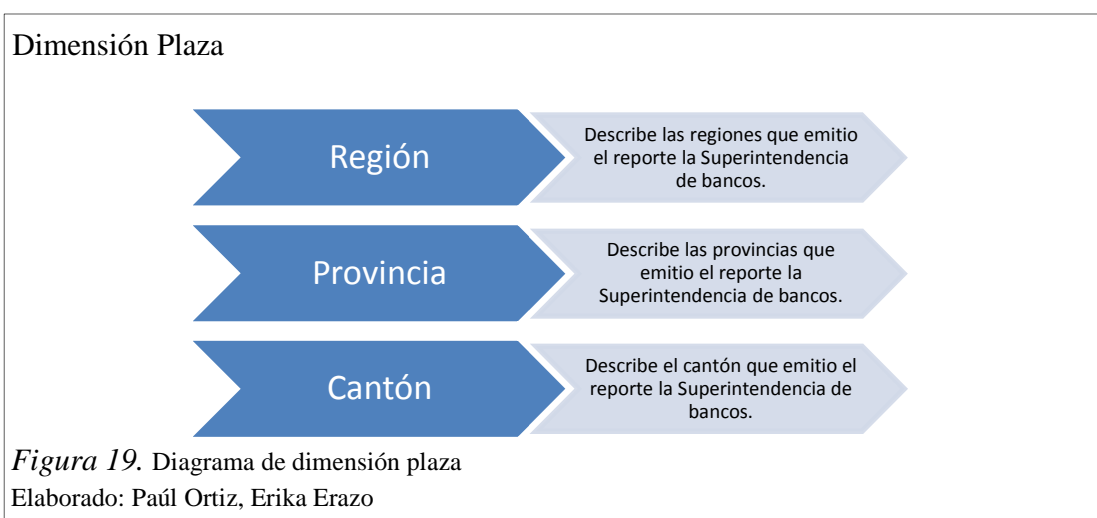
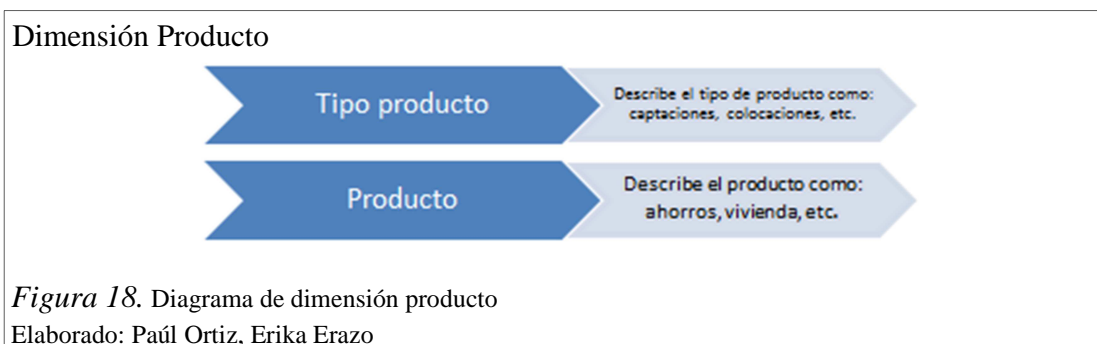


Figura 17. Diagrama de dimensión institución financiera

Elaborado: Paúl Ortiz, Erika Erazo



Información de instituciones financieras ranking

Las dimensiones para el ranking de instituciones financieras son las mismas con la diferencia que no se establece la dimensión plaza ya que los requerimientos del negocio no establecen necesidades de sitio geográfico, el fin es cubrir las necesidades de los reportes.

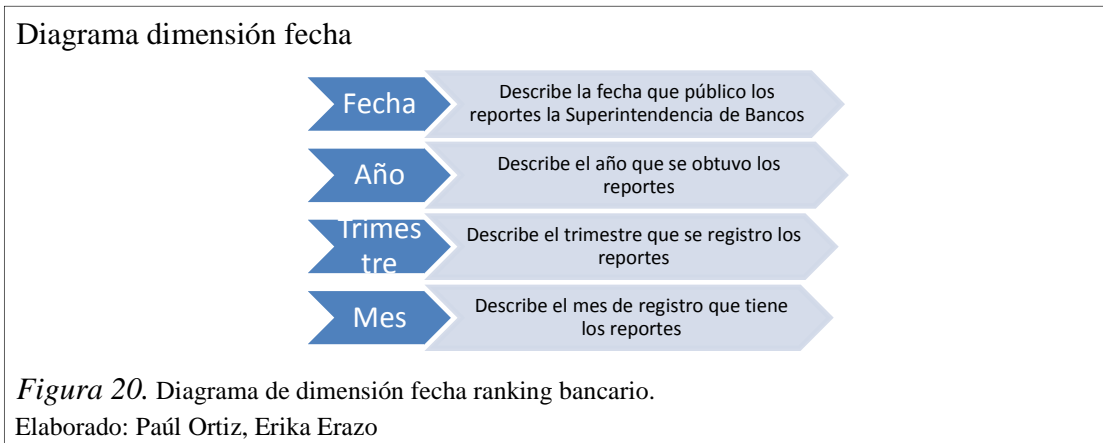
A continuación su descripción:

- **Fecha:** consiste en la fecha que fue reportada la información a la Superintendencia de Bancos del Ecuador por las distintas instituciones financieras.
- **Institución financiera:** contiene las instituciones financieras que pertenecen al Sistema financiero nacional, agrupadas en bancos privados, públicos, cooperativas, mutualistas y sociedades financieras.

- **Plaza:** lugar geográfico en el que el valor del producto fue reportado, para esta dimensión el cantón es el nivel más bajo de granularidad seguido de las provincias y regiones.

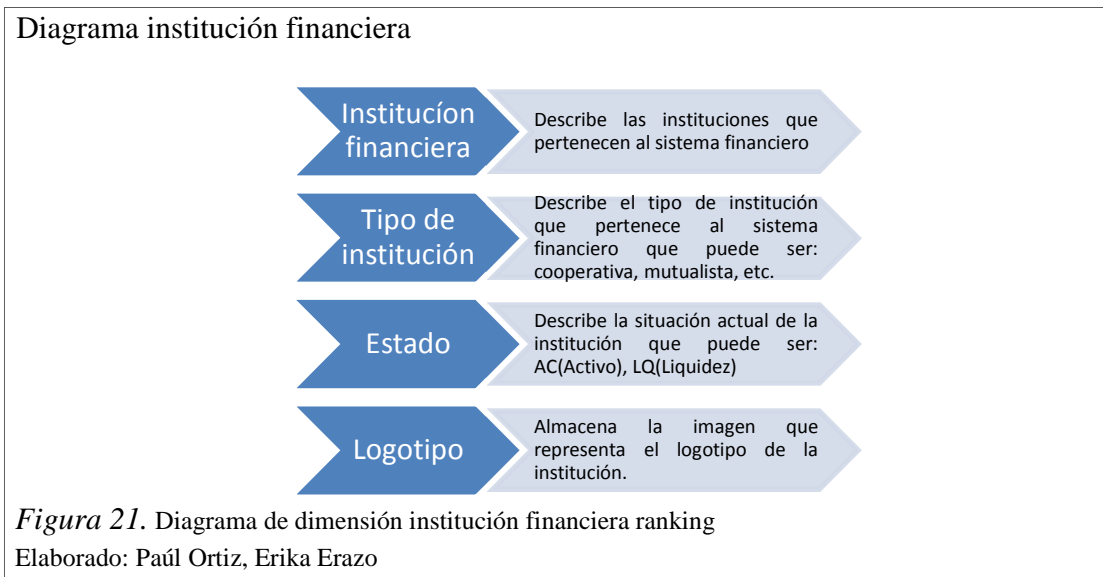
Dimensión fecha

La estructura de la dimensión fecha tiene el mismo concepto que se definió en el data mart de captaciones y colocaciones, como se describe a continuación.



Dimensión institución financiera

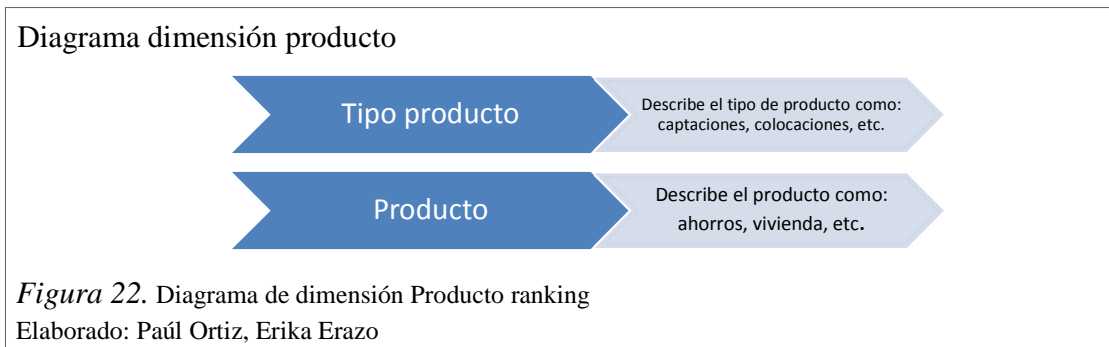
Bajo el mismo concepto de igualdad se estable la dimensión institución financiera, continuación al detalle:



Bajo el mismo concepto de igualdad se establece la dimensión institución financiera, continuación al detalle:

Dimensión producto

Para la dimensión producto se estableció el siguiente diagrama con las descripciones.



3.1.5.3.4 Selección de los hechos

Para la institución financiera cada tabla de hechos cuenta con dimensiones que la rodean con atributos o medidas del proceso organizacional de acuerdo a los requerimientos del negocio. Un registro contiene una medida expresada en números, por ejemplo: saldo por producto, saldo por cuentas contables, etc., sobre la cual se desea realizar una operación de agregación (promedio, conteo, suma, etc.).

Una vez establecida las dimensiones se construye las tablas de hechos que son independientes y de distinta estructura para cada data mart, esto debido a los diferentes requerimientos del negocio se estableció previamente en la arquitectura técnica.

El siguiente diagrama indica la tabla de hechos con sus dimensiones, que están directamente atados a las exigencias del usuario final.

Diagrama tabla de hechos captaciones colocaciones

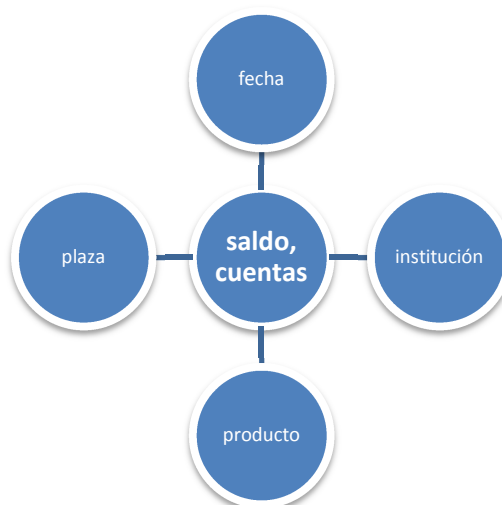


Figura 23. Diagrama de la tabla de hechos saldos, cuentas contables.

Elaborado: Paúl Ortiz, Erika Erazo

Tabla 14. Descripción de la tabla de hechos saldo y cuentas, captaciones y colocaciones

Nombre del Hecho	Descripción	Regla de Agregación
Saldo vigente	Cartera vigente	Sum
Saldo dificultades	Cartera QNDI y vencida	Sum
Saldo cartera	Cartera total	Sum
Saldo captación	Saldo captaciones	Sum
Saldo cuentas	Numero de cuentas	Sum

Nota: La tabla de hechos debe tener unas medidas que permitan funciones de agregación

Elaborado: Paúl Ortiz, Erika Erazo.

Para las instituciones financieras la tabla de hechos se basa en el saldo, ya que es necesario conocer el saldo total que generan cada una de las instituciones financieras.

Tabla de hechos instituciones financieras



Figura 24. Diagrama de la tabla de hechos saldos, cuentas contables.

Elaborado: Paúl Ortiz, Erika Erazo

3.1.5.4 Modelo lógico

Para Kimball, el diseño lógico es de gran ayuda al momento de construir el diseño físico, en el mismo se establece; los atributos y breves descripciones de las dimensiones y tabla de hechos que se utiliza en el proyecto BI.

3.1.5.4.1 Captaciones y colocaciones

Descripción de los atributos de la dimensión fecha

Tabla 15. Atributos dimensión fecha

Nombre Atributo	Descripción	Ejemplo
Fecha	Fecha emisión del reporte de la Superintendencia de Bancos	20140531 20140430
Fecha carga	Fecha en la que la información fue cargada al data mart	20140614 20140516
Anio	Año con el cual fueron registrados los valores	2014 2013
Semestre	Semestre que fue ingresada la información, para la institución Banco Internacional está dividido de la siguiente manera: Semestre 1: Enero, Febrero, Marzo, Abril, Mayo, Junio. Semestre 2: Julio, Agosto, Septiembre, Octubre, Noviembre, Diciembre	1 2
Trimestre	Trimestre que fue ingresada la información, para la institución Banco Internacional está dividido de la siguiente manera: 1 Trimestre: Enero, Febrero, Marzo	1 2 3 4

	2 Trimestre: Abril, Mayo, Junio 3 Trimestre: Julio, Agosto, Septiembre 4 Trimestre: Octubre, Noviembre, Diciembre	
Mes	Mes que fue ingresada los valores, para cada mes se le asigna un número. 1 Enero 2 Febrero	1 2

Nota: los atributos para la dimensión fecha se basan en periodos regulares
Elaborado: Paúl Ortiz, Erika Erazo

Descripción de los atributos de la dimensión Institución

Tabla 16. Atributos dimensión Institución

Nombre Atributo	Descripción	Ejemplo
Institución financiera	Identifica a Institución registrada en el sistema financiero nacional	Banco Internacional CNF
Tipo Institución	Identifica el tipo de institución que se trata	BANCOS PRIVADOS COOPERATIVAS MUTUALISTAS SOC. FINANCIERAS INST. PUBLICAS
Estado	Identifica el estado actual de la institución	AC (Activo) IA (Liquidación)
Color	Es un identificativo personalizado por color en hexadecimal para cada institución financiera.	#0000FF (Azul) #F7FE2E (Amarillo)
Logotipo	Guarda el nombre del archivo imagen del logotipo correspondiente a cada institución	imgBI.png imgBP.png

Nota: Los atributos color y logotipo son referencias para las herramientas de visualización
Elaborado: Paúl Ortiz, Erika Erazo

Descripción de los atributos de la dimensión cuenta contable

Los atributos para la dimensión cuenta contable son los descritos en la siguiente tabla.

Tabla 17. Atributos de la dimensión cuenta contable

Nombre Atributo	Descripción	Ejemplo
Cuenta Contable	Identifica a la cuenta contable que está inscrita según la superintendencia de bancos	1 3 11
Detalle	Detalla a que corresponde la cuenta contable	Caja Cartera de créditos comercial por vencer Intereses por cobrar inversiones
Producto	Describe el producto que pertenece la cuenta contable	Fondos disponibles Carteras de Créditos

Nota: el atributo cuenta contable viene dado por la SBS
Elaborado: Paúl Ortiz, Erika Erazo

3.1.6. Diseño físico

El diseño físico de la base de datos o arquitectura de datos se focaliza sobre la selección de las estructuras necesarias para soportar el diseño lógico. Esta etapa se basa en la definición de métodos que garanticen un acceso eficiente a los datos.

Los elementos principales de este proceso son la definición de estándares de nombres específicos del ambiente de la base de datos. Es muy probable que el modelo dimensional no se respete del todo puesto que en el modelo físico se contempla ciertas estrategias de implementación de tal manera que no sea transparente, esto debido a temas de agregación, índices, particionamiento, etc.

Por lo tanto, en el plan de indexación los índices son estructuras opcionales optimizadas y orientadas a conjuntos de operaciones.

Kimball dice, “las tablas de dimensión deben tener un único índice sobre las claves primarias y sería recomendable que el índice estuviera compuesto de un único atributo”. También recomienda el uso de índices tipo árbol-B en atributos de alta cardinalidad y aplicar índices de mapas de bits en atributos de cardinalidad media o baja.

3.1.6.1. Estándares

Para tener un buen manejo de los elementos que conforman los data marts, es necesario definir estándares de nombramiento para los objetos, con el fin de generar una nomenclatura que facilitara la integración de nuevos proyectos BI.

3.1.6.1.1 Bases de datos

Para las bases de datos se ha establecido las siglas MIS que significa Management Information System:

- Captaciones y colocaciones: MIS
- Instituciones financieras: MISRANKING

3.1.6.1.2 Tablas

Se ha dividido en 2 grupos:

Back room:

RCC: siglas ranking captaciones colocaciones

RKB: siglas ranking instituciones financieras

- <Identificador Principal>: descripción de la tabla
- [Identificador Secundario]: descripción a detalle y/o característica especial

RCC_<Identificador Principal>_[Identificador Secundario]

RKB_<Identificador Principal>_[Identificador Secundario]

Front Room:

- <Tipo>: identificador que indica si es un tabla de dimensión (D), tabla de hechos (H), u otro tipo (O)
- <Identificador Principal>: descripción de la tabla
- [Identificador Secundario]: descripción a detalle y/o característica especial

RK<Tipo>_<Identificador Principal>_[Identificador Secundario]

RKB<Tipo>_<Identificador Principal>_[Identificador Secundario]

3.1.6.1.3 Atributos

<Tabla>_<Identificador Principal>_[Identificador Secundario]

- <Tabla>: Se identifica con 3 caracteres que indican la tabla a la que pertenece el campo.
- <Identificador Principal>: descripción de la tabla
- [Identificador Secundario]: descripción a detalle y/o característica especial

3.1.6.2. Diseño físico de tablas

Con el fin de mejorar el diseño físico de las tablas Kimball recomienda:

- Mantener los nombres del modelo lógico
- Encontrar un balance entre nombres específicos y vagos

3.1.6.1.4 Captaciones y colocaciones tablas de dimensiones

RKD_Fecha

Tabla 18. Descripción tabla RKD_FECHA

Columna	Tipo Dato	Longitud	Nulos	PK	FK
fec_fecha	numeric	8	no	X	
fec_fecha_carga	numeric	8	no		
fec_anio_mes	int		no		
fec_anio	int		no		
fec_semestre	int		no		
fec_trimestre	int		no		
fec_mes	int		no		

Nota: Data mart captaciones y colocaciones
Elaborado: Paúl Ortiz, Erika Erazo

RKD_Institucion

Tabla 19. Descripción tabla RKD_Institucion

Columna	Tipo Dato	Longitud	Nulos	PK	FK
ifi_institucion	int		no	X	
ifi_institucion_nombre	varchar	50	no		
tif_tipo_institucion	int		no		
tif_tipo_institucion_nombre	varchar	50	no		
ifi_estado	char	2	no		
ifi_color_barra	char	7	no		
ifi_color_pie	char	7	no		
ifi_color_borde	char	7	no		
ifi_logotipo	varchar	50	no		

Nota: Data mart captaciones y colocaciones
Elaborado: Paúl Ortiz, Erika Erazo

RKD_Plaza

Tabla 20. Descripción tabla RKD_Plaza

Columna	Tipo Dato	Longitud	Nulos	PK	FK
plz_canton	int		no	X	
plz_canton_nombre	varchar	50	no		
plz_provincia	int		no		
plz_provincia_nombre	varchar	50	no		
plz_region	int		no		
plz_region_nombre	varchar	50	no		

Nota: Data mart captaciones y colocaciones

Elaborado: Paúl Ortiz, Erika Erazo

RKD_Producto

Tabla 21. Descripción tabla RKD_producto

Columna	Tipo Dato	Longitud	Nulos	PK	FK
prd_producto	int		no	X	
prd_producto_nombre	varchar	50	no		
prd_producto_descripcion	varchar	200	no		
prd_color_barras	char	7	no		
prd_color_pie	char	7	no		
tpr_tipo_producto	int		no		
tpr_tipo_producto_nombre	varchar	50	no		

Nota: Data mart captaciones y colocaciones

Elaborado: Paúl Ortiz, Erika Erazo

Captaciones y colocaciones tabla de hechos

RKH_Saldos_Cuentas

Tabla 22. Descripción tabla RKH_Saldos_Cuentas

Columna	Tipo Dato	Longitud	Nulos	PK	FK
fec_fecha	numeric	8	no	X	X
ifi_institucion	int		no	X	X
plz_canton	int		no	X	X
prd_producto	int		no	X	X
scu_saldo_vigente	float		no		
scu_saldo_dificultades	float		no		
scu_saldo_cartera	float		no		
scu_saldo_captacion	float		no		
scu_cuentas	int		no		

Nota: Data mart captaciones y colocaciones

Elaborado: Paúl Ortiz, Erika Erazo

3.1.6.1.5 Instituciones financieras Tablas de dimensiones

RKBD_Fecha

Tabla 23. Descripción tabla RKBD_Fecha

Columna	Tipo Dato	Longitud	Nulos	PK	FK
fec_fecha	nchar	10	no	X	
fec_FechaCarga	nchar	10	no		
fec_FechaAnioMes	nchar	10	no		
fec_FechaAnio	nchar	10	no		
fec_FechaSemestre	nchar	10	no		
fec_FechaTrimestre	nchar	10	no		
fec_FechaMes	nchar	10	no		

Nota: Data mart Instituciones financieras
Elaborado: Paúl Ortiz, Erika Erazo

RKBD_Plaza

Tabla 24. Descripción tabla RKBD_Plaza

Columna	Tipo Dato	Longitud	Nulos	PK	FK
plz_canton	int		no	X	
plz_canton_nombre	varchar	50	no		
plz_provincia	int		no		
plz_provincia_nombre	varchar	50	no		
plz_region	int		no		
plz_region_nombre	varchar	50	no		

Nota: Data mart Instituciones financieras
Elaborado: Paúl Ortiz, Erika Erazo

RKBD_Institucion

Tabla 25. Descripción tabla RKDB_Institucion

Columna	Tipo Dato	Longitud	Nulos	PK	FK
ifi_InstitucionCodigo	int		no	X	
ifi_InstitucionNombre	varchar	500	no		
tif_InstitucionTipo	int		no		
tif_InstitucionNombreTipo	varchar	50	no		
ifi_Institucionestado	char	2	no		
ifi_InstitucionColorBarra	char	7	no		
ifi_InstitucionColorPie	char	7	no		

Nota: Datamart Instituciones financieras.
Elaborado: Paúl Ortiz, Erika Erazo.

RKBD_CuentaContable

Tabla 26. Descripción tabla RKBD_CuentaContable

Columna	Tipo Dato	Longitud	Nulos	PK	FK
cco_CuentaContable	int		no	X	
cco_DetalleCuentaContable	varchar	500	no		
cco_ProductoDescripcion	varchar	1000	no		

Nota: Datamart Instituciones financieras.

Elaborado: Paúl Ortiz, Erika Erazo

Instituciones financieras tabla de hechos

RKBH_Saldos_Cuentas

Tabla 27. Descripción tabla RKBK_Saldos_Cuentas

Columna	Tipo Dato	Longitud	Nulos	PK	FK
fec_Fecha	nchar	10	No	X	X
ifi_InstitucionCodigo	int		No	X	X
cco_CuentaContable	int		No	X	X
scu_SALDO	int		No		

Nota: Datamart Instituciones financieras.

Elaborado: Paúl Ortiz, Erika Erazo

3.1.6.3. Diseño multidimensional

El diseño multidimensional a diferencia del modelo relacional, permite analizar datos mediante cubos OLAP. Además su gran escalabilidad ya que se pueden agregar hechos y dimensiones conforme se requiera y de acuerdo a los requerimientos del negocio.

Se estableció dos data marts que a su vez conforma el data warehouse del cual se consume la información para el apoyo de toma de decisiones en el área de control Financiero en la institución Banco Internacional.

Datamart captaciones colocaciones (CAPCOL)

Se construyó el data mart en base a los requerimientos del negocio, modelo dimensional y estándares a continuación se describen las dimensiones y tabla de hechos

- Dimensiones
 - RKD_Fecha, RKD_Producto, RKD_Institucion, RKD_plaza

- Tabla de hechos
 - RKH_Saldos_Cuentas

Data mart captaciones colocaciones

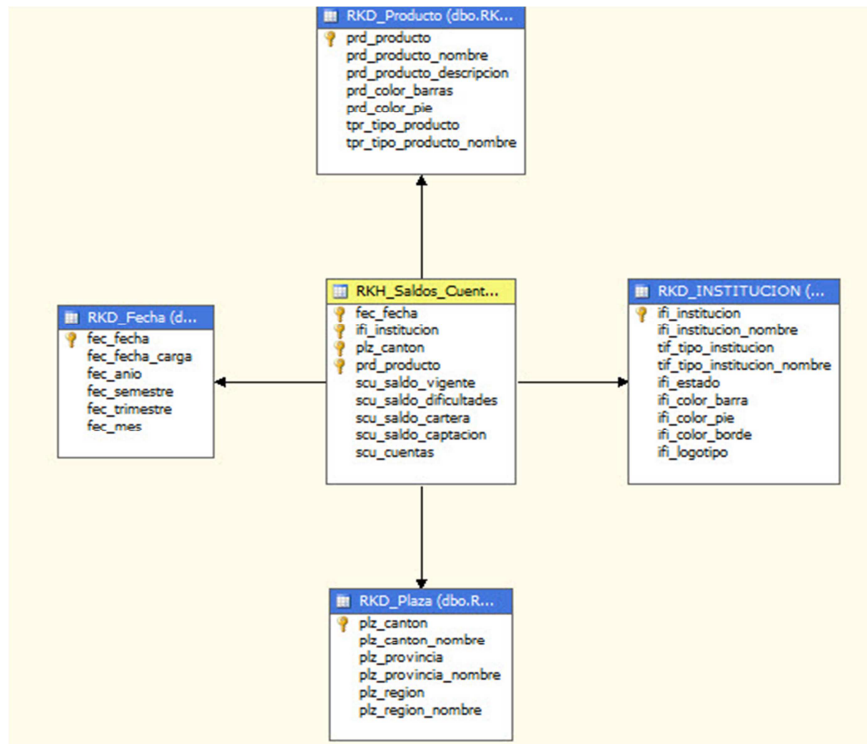
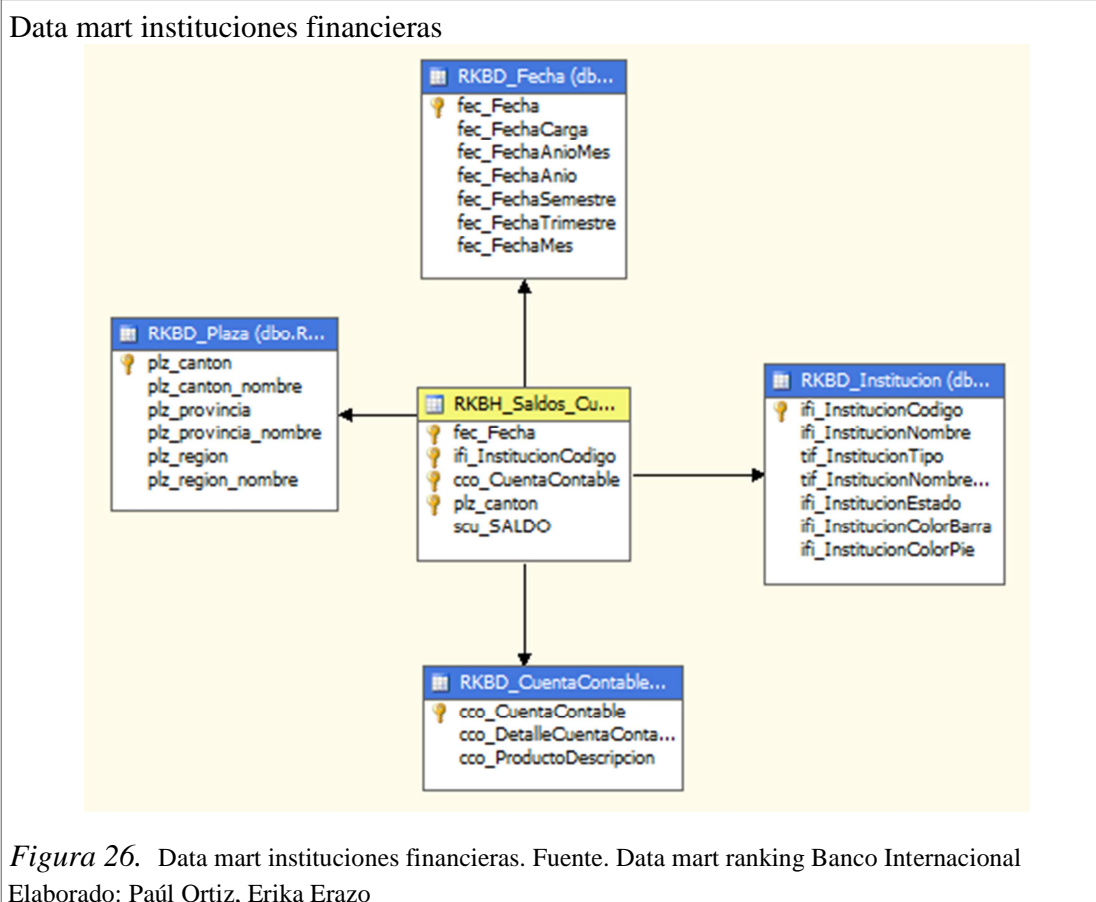


Figura 25. Data mart captaciones colocaciones. Fuente. Data mart capcol Banco Internacional
Elaborado: Paúl Ortiz, Erika Erazo

Data mart instituciones financieras (ranking bancario)

Se construyó el data mart en base a los requerimientos del negocio, modelo dimensional y estándares a continuación se describen las dimensiones y tabla de hechos

- **Dimensiones**
 - RKBD_Fecha, RKBD_Institucion, RKBD_CuentaContable, RKBD_Plaza
- **Tabla de hechos**
 - RKBH_Saldos_Cuentas



Como se observa en las ilustraciones la construcción de los data marts se utilizó un esquema tipo estrella.

3.1.6.4. Homologación de datos

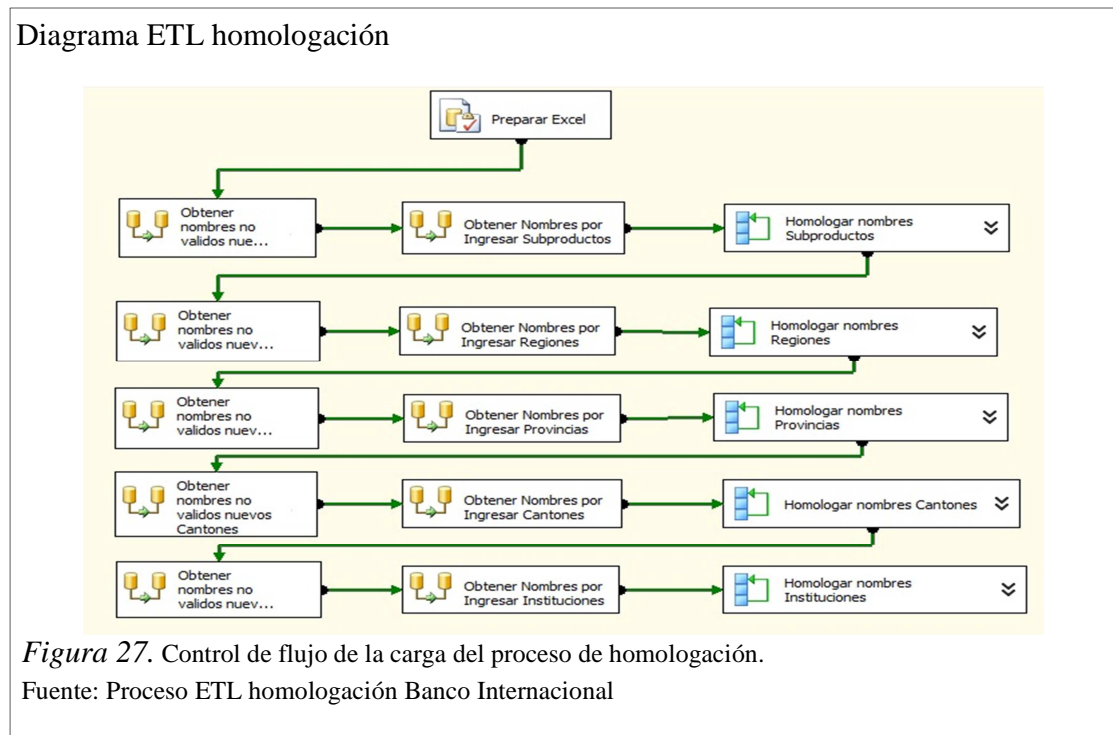
Los datos entregados por la Superintendencia de Bancos no siempre se encuentran homogéneos y estandarizados. Esto se debe a la falta de uniformidad de la información del reporte de la SBS que publica el nombre de los cantones, provincias, regiones, productos y especialmente las instituciones. Para esto se requiere que los datos ingresen en un proceso ETL (Extracción, transformación y carga) de limpieza.

“Los procesos ETL (Extraction, Transformation and Loading, extracción, transformación y carga) son los encargados de transportar los datos que contienen las fuentes de datos al

almacén de datos asegurando la calidad de los datos que serán luego analizados mediante las herramientas de explotación.” (Mazón , Pardillo, & Trujillo, 2010)

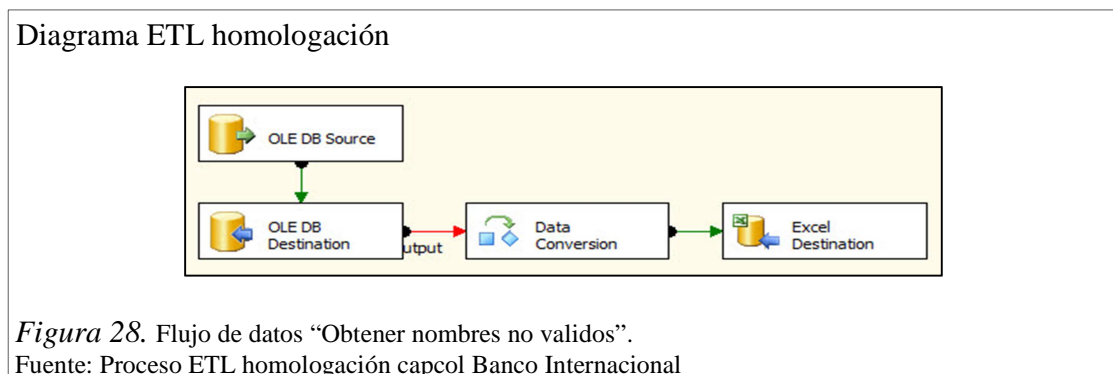
Homologación captaciones y colocaciones

En la siguiente figura se visualiza el proceso de limpieza para Captaciones y Colocaciones.



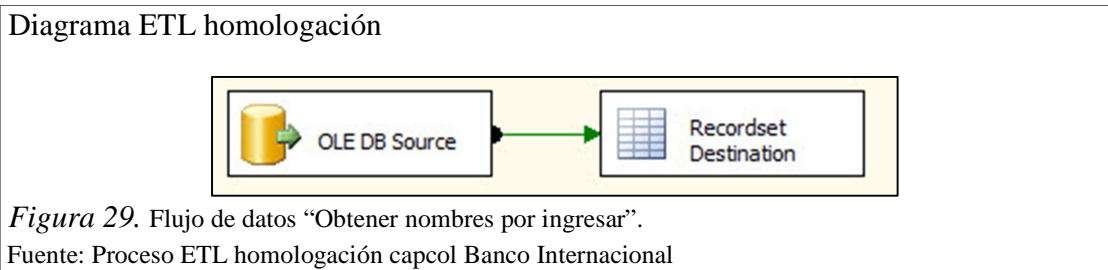
El proceso consiste en 3 flujos que se detallan a continuación:

1. Obtener nombres no validos nuevos



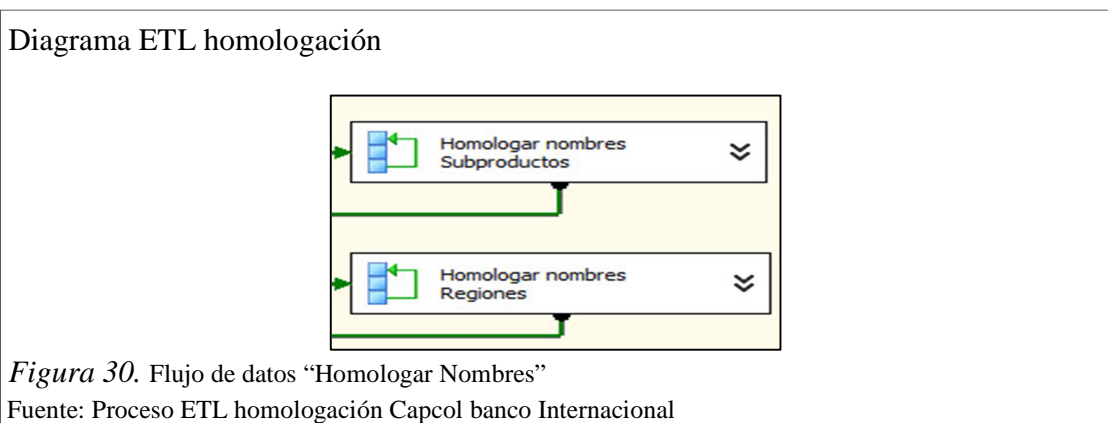
- Realiza una consulta en las tablas RCC_InformacionSuperCaptaciones, RCC_InformacionSuperColocaciones para obtener los nuevos registros de los reportes de la SBS.
- Si existen registros se inserta en la tablas RCC_Homologacion
- Si algún registro no pasa a la tabla RCC_Homologacion se realiza una conversión de datos.
- Los registros que no ingresan a la tabla RCC_Homologacion se insertan en un archivo Excel llamado Excepciones.

2. Obtener nombres por ingresar.



- Consulta los nuevos registros que se insertó en la tabla RCC_Homologacion
- Inserta los registros en un conjunto de memoria, en este caso la variable NombresPorIngresar.

3. Homologar nombres



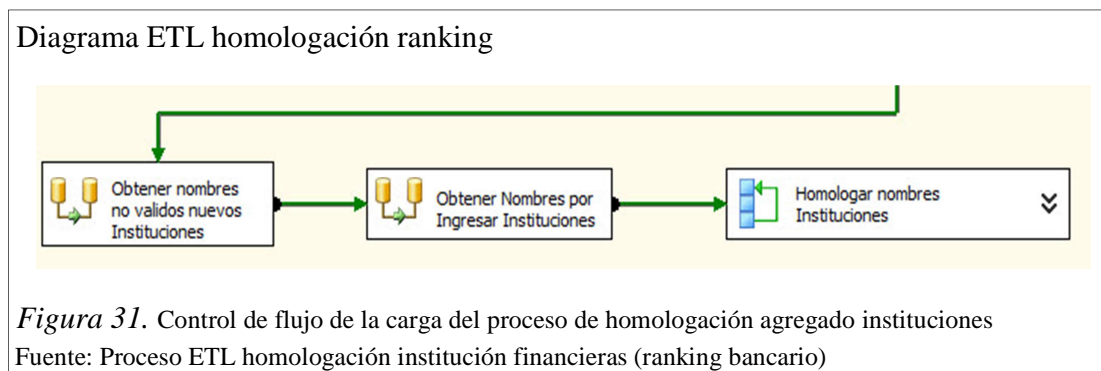
- Registra la variable NombresPorIngresar y le pasa a la variable Nombres

Todo este proceso se cumple para cada registro nuevo del reporte emitido por la SBS, subproductos, regiones, provincias, cantones. Esto con el propósito de tener información homogénea de las descripciones de los reportes para la toma de decisiones.

Antes de iniciar la carga es necesario limpiar el archivo de excepciones que es un archivo Excel.

Homologación instituciones financieras (Ranking)

Debido a los requerimientos del negocio se replica el proceso de homologación de captaciones y colocaciones, con la variante que se agrega el flujo del proceso para la homologación de instituciones, como se indica en la figura.



3.1.6.5. Integración de los datos ETL (extracción, transformación y carga)

Las principales actividades de esta fase del ciclo de vida son:

- Extracción de datos
- Transformación de los datos origen de tal manera que sea información
- Carga de datos

Proceso de Extracción: son aquellos requeridos para adquirir datos que permita la carga al modelo físico previamente diseñado.

Proceso de Transformación: permiten convertir o recodificar los datos que son previamente extraídos de las fuentes para la carga eficiente del modelo físico.

Proceso de Carga: es el que se encarga de poblar el DWH bajo procesos o métodos requeridos.

Esta fase no debe ser subestimada ya que las tareas son altamente críticas debido a que se está trabajando con la materia prima del DWH: los datos y la información. Al presentar información inconsistente en el DWH, nace la desconfianza y pérdida de credibilidad en el usuario causando efectos inmediatos e inevitables en el desarrollo del DWH. Es en esta etapa donde deben depurar todos los inconvenientes relacionados con la calidad de los datos fuentes.

Bajo esta definición se estableció procesos ETL independientes que permiten la extracción, transformación y carga para cada uno de los data mart de la solución de inteligencia de negocios, por lo tanto el proceso se divide en dos partes:

1. Captaciones y colocaciones (Capcol).
2. Instituciones financieras (Ranking).

La herramienta que se utilizó para la creación de los ETL es Microsoft SQL Server Integration Services 2005, que previamente se definió en la fase de selección e instalación.

“Microsoft SQL Server Integration Services (SSIS) es una plataforma para la creación de soluciones empresariales de transformaciones de datos e integración de datos. Integration Services sirve para resolver complejos problemas empresariales mediante la copia o descarga de archivos, el envío de mensajes de correo electrónico como respuesta a eventos, la actualización de almacenamientos de datos, la limpieza y minería de datos, y la administración de objetos y datos de SQL Server” (msdn.microsoft.com, 2014)

Arquitectura SSIS (SQL Server Integration Services).

Motor de flujo de tareas

El motor de flujo de tareas orientado a las operaciones es el que proporciona la compatibilidad operativa y de recurso en tiempo de ejecución para el motor de flujo de datos.

Motor de flujo de datos

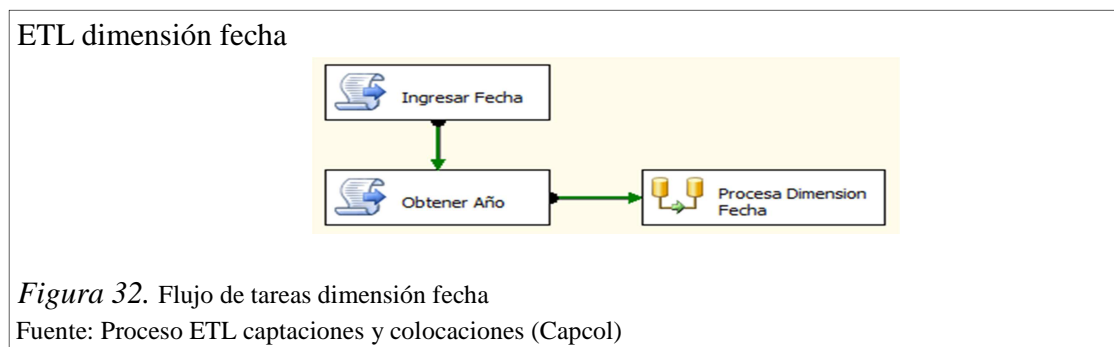
“El flujo de datos proporciona tres componentes: orígenes, transformaciones y destinos. Los orígenes extraen datos de almacenes de datos tales como tablas y vistas en bases de datos relacionales, archivos y bases de datos de Analysis Services. Las transformaciones modifican, resumen y limpian datos. Los destinos cargan datos en almacenes de datos o crean conjuntos de datos almacenados en la memoria.” (Microsoft, 2014)

3.1.6.5.1 ETL captaciones y colocaciones (Capcol)

El proceso consiste en cargar información en las tablas de dimensiones y posteriormente en la tabla de hechos. A continuación se describe el proceso para cada carga del ETL.

Dimensión fecha.

El proceso de la dimensión fecha consiste en cargar los datos en la tabla RKD_Fecha, básicamente obtiene la fecha ingresada con el formato (AAAAMMDD) mediante el cuadro de dialogo, seguidamente pasa al flujo de datos como indica la figura.



El flujo de datos como primer paso realiza la conversión de los datos que obtiene desde la fuente, seguidamente pasa al SCD (Slowly Changing Dimension) que consiste en transformar los datos en una medida mediante los atributos de entrada.

ETL Dimensión fecha

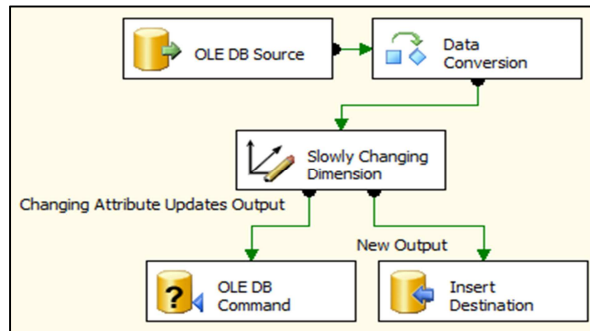


Figura 33. Flujo de datos “Dimensión Fecha”

Fuente: Proceso ETL captaciones y colocaciones (Capcol)

Dimensión plaza

El proceso de carga de la dimensión plaza comienza eliminando los datos de la tabla (), esto permite en cada ejecución homologar las regiones, provincias, cantones. Seguidamente pasa al flujo de datos donde carga la información a la tabla y posteriormente procesar los datos de la dimensión.

ETL dimensión plaza

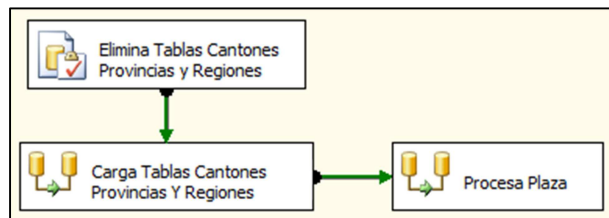


Figura 34. Flujo de tareas “Dimensión Plaza”

Fuente: Proceso ETL captaciones y colocaciones (Capcol)

En el flujo de datos, utiliza como fuente de datos la plantilla de Excel FormatosRan.xls” la cual contiene las regiones, provincias y cantones atados a un código.

ETL dimensión plaza

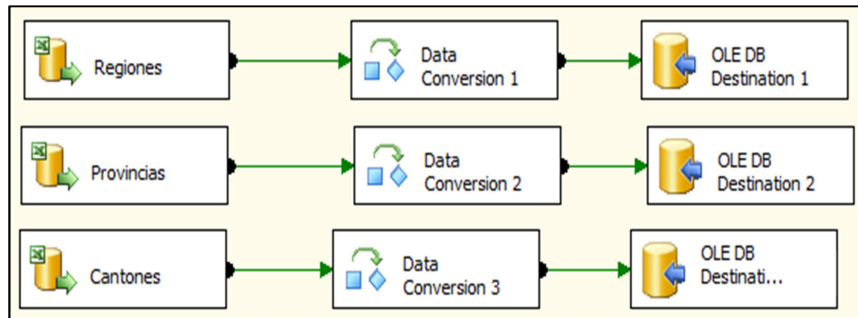


Figura 35. Flujo de datos “Carga Tablas Cantones Provincias Regiones”

Fuente: Proceso ETL captaciones y colocaciones (Capcol)

Finalmente en el flujo de datos procesa plaza, convierte la tabla ya en su totalidad cargada en dimensión la cual permite unir con la tabla de hechos, el atributo que permite la unión es código del cantón que es el nivel más bajo de granularidad para la dimensión plaza.

ETL dimensión plaza

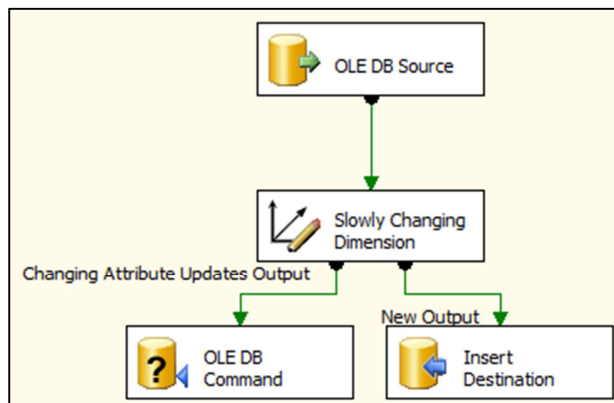


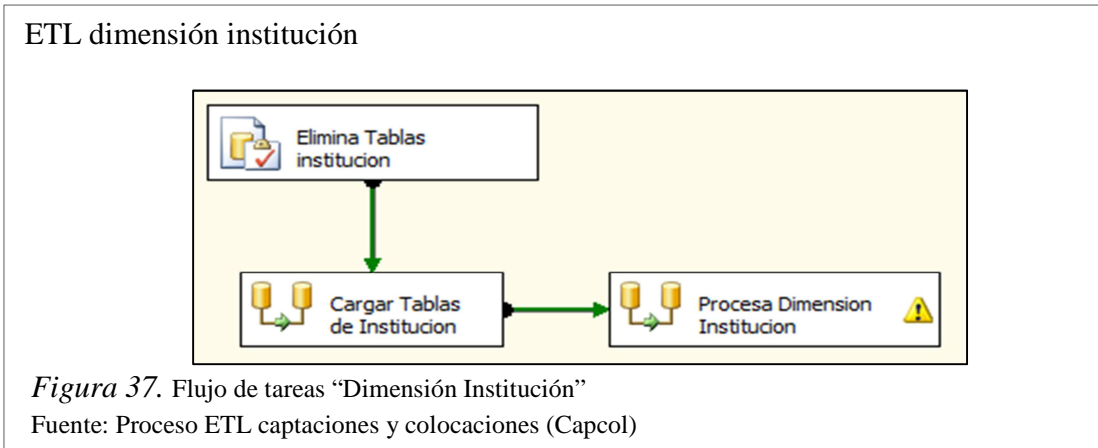
Figura 36. Flujo de datos “Dimensión Plaza”

Fuente: Proceso ETL captaciones y colocaciones (Capcol)

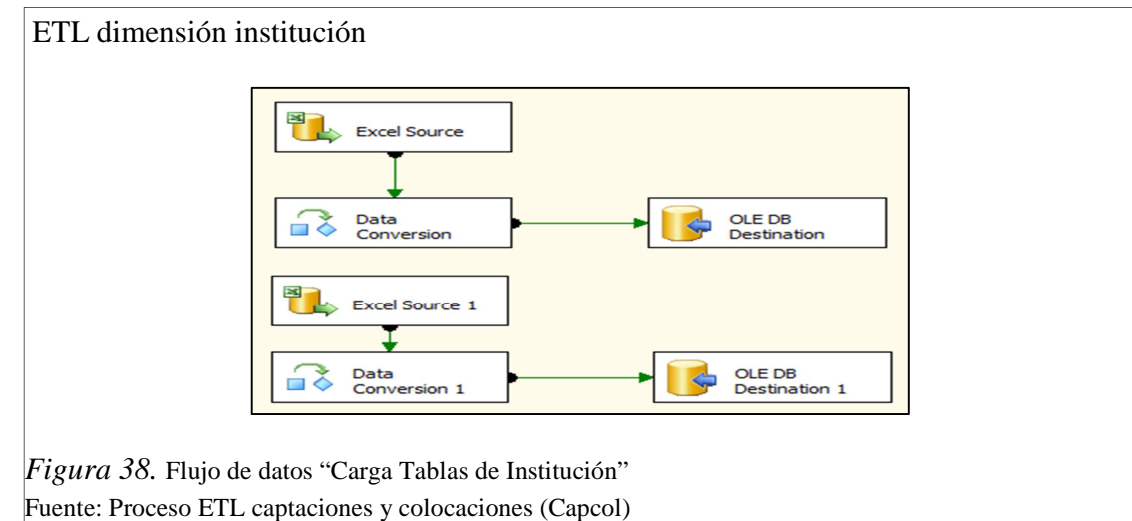
Dimensión institución

Al igual que el proceso de la dimensión plaza, el proceso de la dimensión institución comienza con la eliminación de los datos de tabla RCC_Instituciones_Financieras y

RCC_Tipo_Institucion. De esta manera se puede continuar con los flujos de datos: carga tablas de iInstitución y procesa dimensión institución.



En el flujo de datos “carga tablas de institución” atreves de un plantilla de Excel “Formatos.xls” como fuente de datos, vuelve a cargar en las tablas RCC_Instituciones_Financieras y RCC_Tipo_Institucion las cuales contienen toda información de las instituciones financiera como se indica en la ilustración 38.



Una vez concluida la carga de las tablas es necesario realizar un join por la columna que identifica el tipo de institución de esta manera se obtiene los atributos de entrada para SCD lo que permite cargar datos de las instituciones financieras en la dimensión RKD_Institucion.

ETL dimensión institución

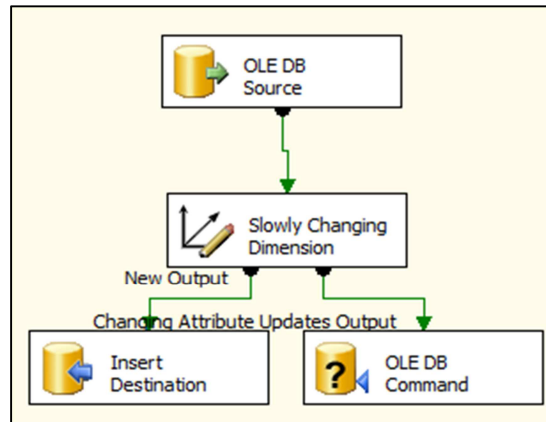


Figura 39. Flujo de datos “Carga Tablas de Institución”

Fuente: Proceso ETL captaciones y colocaciones (Capcol)

Dimensión producto

Siempre es necesario comenzar con la eliminación de la data que existe en las tablas, para la dimensión producto es así como comienza el flujo de tareas. Esto facilita a la presentación de los datos. Seguidamente empieza el flujo de datos de carga de tablas producto y después pasa a procesar los datos de tal manera que la tabla sea una tabla de dimensión.

ETL dimensión producto

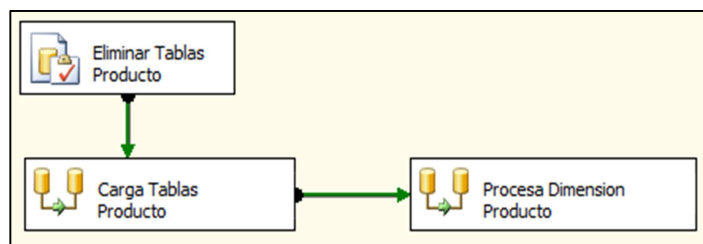


Figura 40. Flujo de tareas “Dimensión Producto”

Fuente: Proceso ETL captaciones y colocaciones (Capcol)

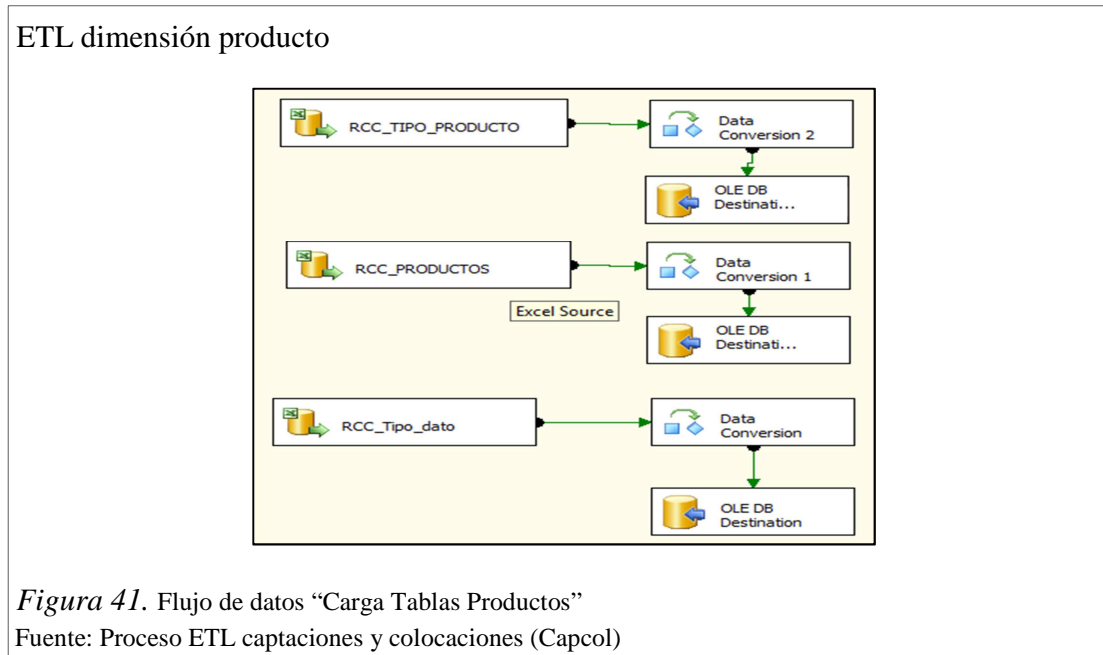
Para el flujo de datos “Carga Tablas Producto” se obtiene de la plantilla “Formatos.xlsx” con la diferencia que cada fuente se direcciona a una hoja del mismo archivo Excel, como se detalla a continuación:

Tabla 28. Descripción del archivo "Formatos.xls" por hoja

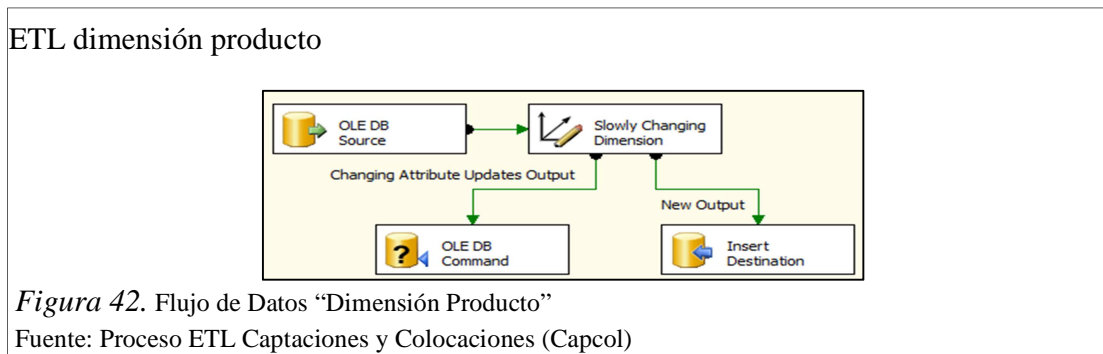
Descripción	Hoja
RCC_TIPO_PRODUCTO	RCC_TIPO_PRODUCTO
RCC_PRODUCTOS	RCC_PRODUCTOS
RCC_TIPO_DATO	RCC_Tipo_dato

Nota: el archivo Excel tiene que estar con las respectivas hojas descritas
 Elaborado por: Paul Ortiz, Erika Erazo

En la siguiente ilustración indica el detalle del proceso "Carga Tablas Producto".



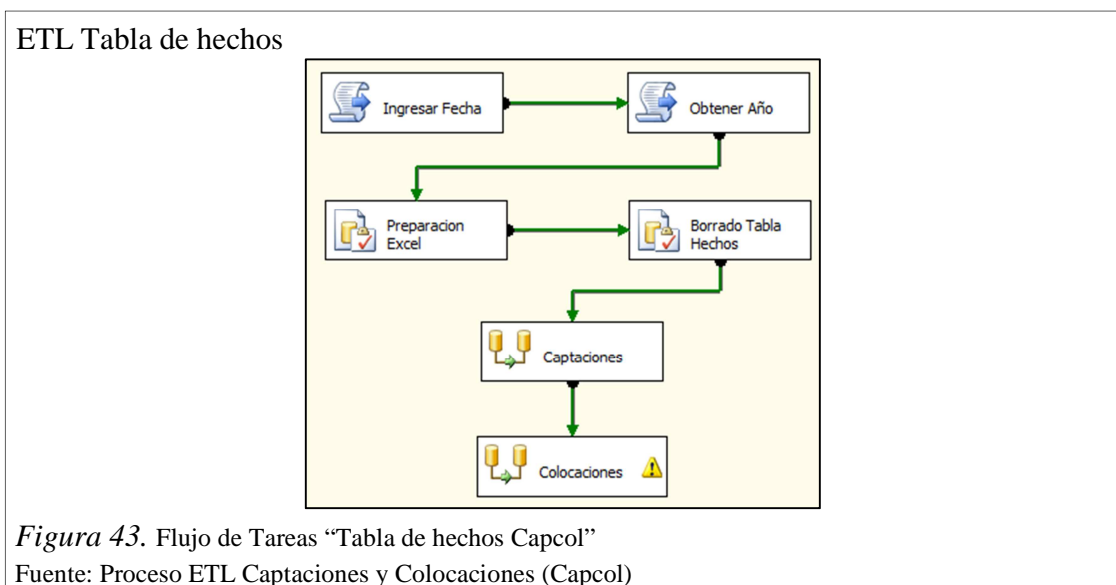
Para el flujo de datos "Dimensión Producto", se obtiene los atributos de entrada para el SCD desde las tablas ya cargadas con la información de los productos, esto permite cargar a la tabla RKD_productos que ya puede ser considerada como dimensión.



Carga tabla de hechos

La tabla de hechos permite la evaluación principal de los indicadores de captaciones y colocaciones de acuerdo al nivel de granularidad.

El proceso empieza con recrear la tablas que se van utilizar para realizar la carga en cada uno de los flujos tanto para Captaciones y Colocaciones. También realiza un borrado de los datos que existe en la tabla de hechos resultado de procesos pasados. Como se indica en la siguiente figura.



El flujo de tareas se dividió en dos flujos de datos tanto para captaciones como colocaciones, esto con el fin de realizar un proceso más efectivo al momento de la carga y garantizando la integridad de los datos.

Para cada uno de los flujos de datos “Captaciones y Colocaciones” se usó el componente “Lookup” provista por SSIS, la transformación búsqueda exacta (lookup) consiste en unir los datos de las columnas de entrada en un conjunto referencia de datos, ya que se trata de una tabla de hechos los conjuntos de referencia son las tablas de dimensiones que conforman el DWH. Con este componente se logra tener una tabla de hechos sin redundancia de datos e información consistente.

Proceso captaciones

A continuación se detalla el flujo de tareas que carga la información de captaciones en la tabla de hechos RKH_Saldos_Cuentas.

- Se obtiene los datos desde la tabla RCC_InformacionSuperCaptaciones
- Se agrega una columna con la información del mes en el siguiente formato 201404
- Lookup Fecha: une los datos de la tabla referencia RKD_Fecha y los datos de entrada de fecha atreves de la búsqueda exacta, si existe errores son guardados en el archivo de excel Expeciones.xls
- Lookup Producto: une los datos de la tabla referencia RKD_Producto y los datos de entrada de producto atreves de la búsqueda exacta, si existe errores son guardados en el archivo de excel Expeciones.xls
- Lookup Plaza: une los datos de la tabla referencia RKD_Plaza y los datos de entrada de plaza atreves de la búsqueda exacta, si existe errores son guardados en el archivo de excel Expeciones.xls
- Lookup Institución: une los datos de la tabla referencia RKD_Institucion y los datos de entrada de instituciones atreves de la búsqueda exacta, si existe errores son guardados en el archivo de excel Expeciones.xls
- Se convierte el campo fecha
- Se realiza la función de agregación que permite obtener datos para agrupados.

En la siguiente ilustración se observa todo flujo del proceso.

ETL captaciones

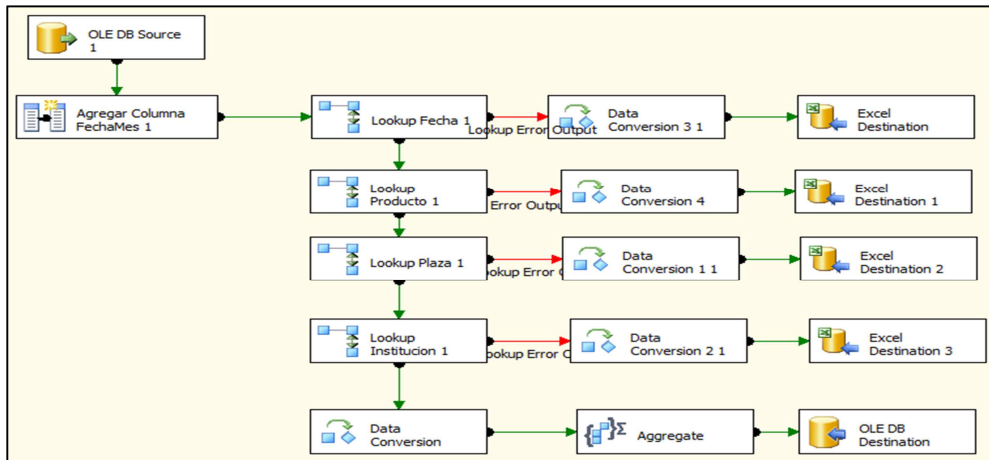


Figura 44. Flujo de tareas “Colocaciones”

Fuente: Proceso ETL Captaciones y Colocaciones (Capcol)

Proceso colocaciones

A continuación se detalla el flujo de tareas que carga la información de colocaciones en la tabla de hechos RKH_Saldos_Cuentas. Básicamente consiste en el mismo proceso de captaciones con la única diferencia que la fuente de datos es la tabla RCC_InformacionSuperColocaciones.

En la siguiente ilustración se observa todo flujo del proceso.

ETL colocaciones

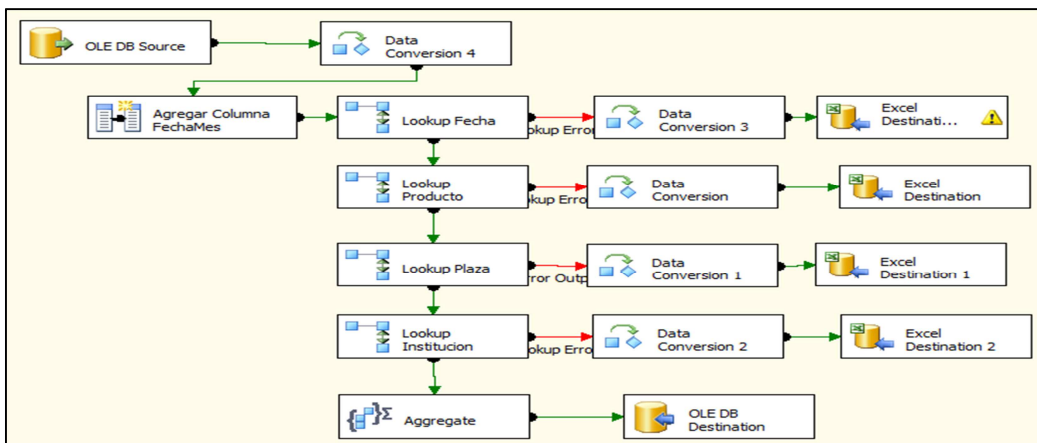


Figura 45. Flujo de tareas “Captaciones”.

Fuente: Proceso ETL Captaciones y Colocaciones (Capcol)

Con la finalización de la carga de la tabla de hechos RKH_Saldos_Cuentas, se da por culminado el ETL captaciones colocaciones.

3.1.6.5.2 ETL instituciones financieras.

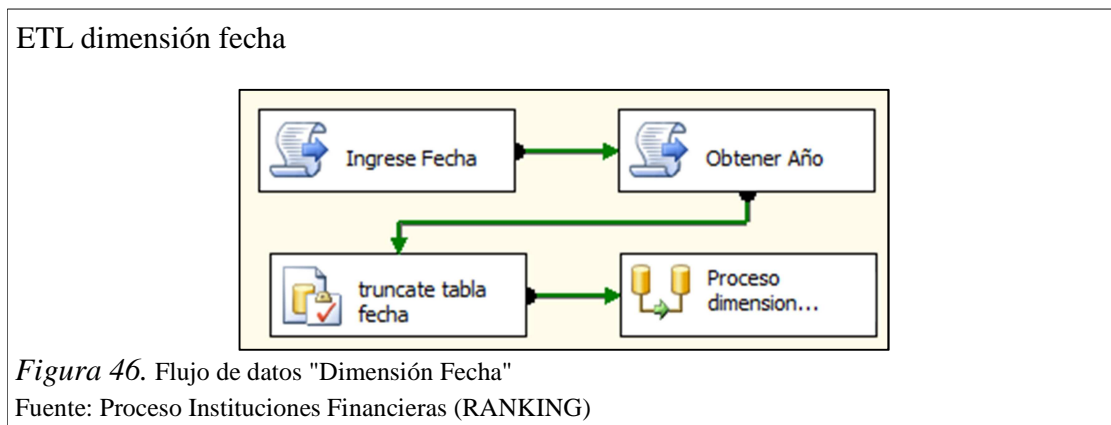
Como se tratan de data marts independientes es necesario establecer un proceso para la extracción, transformación y carga que maneje los datos de las instituciones financieras, básicamente el proceso tiene la misma lógica que maneja el ETL captaciones y colocaciones.

A continuación se detalla el proceso de carga para las dimensiones y la tabla de hechos respectivamente.

Dimensión Fecha

Flujo de datos.

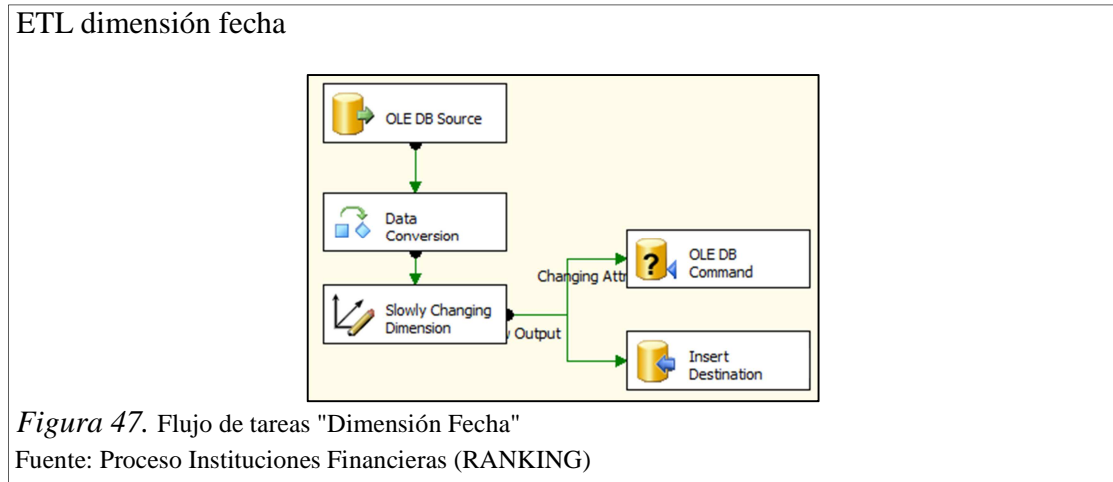
El proceso comienza con la dimensión fecha, en el flujo de datos el primer paso es ingresar la fecha, el siguiente flujo obtiene el año de la fecha ingresada, seguidamente borra los registros de cargas antiguas en el paso “truncate tabla fecha” y continúa con el flujo de tareas, como se visualiza en la siguiente figura.



Flujo de tareas

Como primer paso del flujo de tareas es obtener los datos de la fuente, en el siguiente paso es convertir el tipo de dato para el campo fecha, al igual que en los procesos de

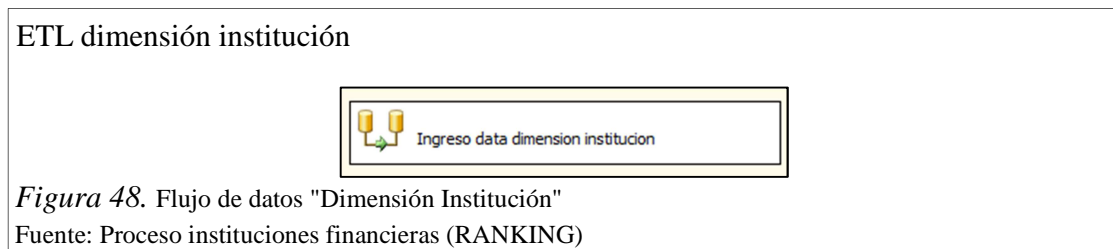
captaciones y colocaciones se utiliza el componente SCD que permite cargar los datos de fecha como tabla de dimensión. En la figura adjunta se indica los pasos del flujo.



Dimensión institución

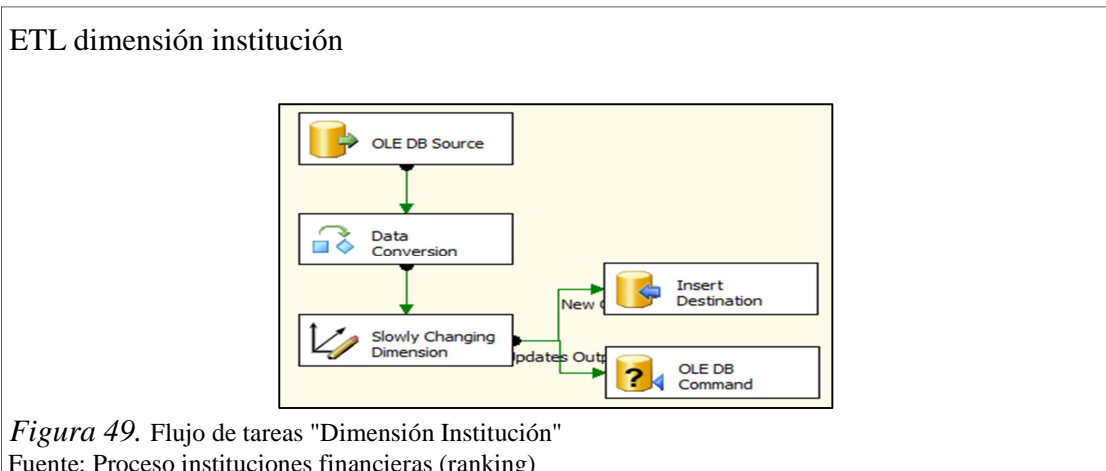
Flujo de datos

Para la dimensión institución, se tiene un único flujo de datos, como se indica en la figura.



Flujo de tareas

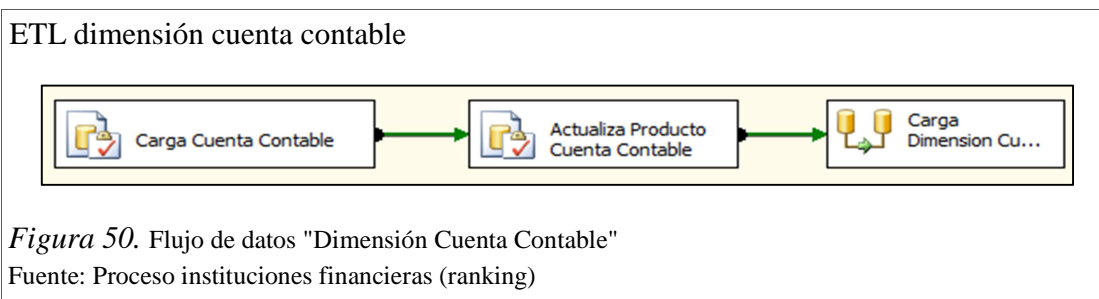
El flujo de tareas empieza por obtener los datos de la fuente, después realiza la conversión del atributo en el paso “Data Conversion”, seguidamente ingresa al paso SCD donde transforma los datos, los mismos que son parte de la dimensión RKD_Institucion. El detalle de los pasos se describe en la siguiente figura.



Dimensión cuenta contable

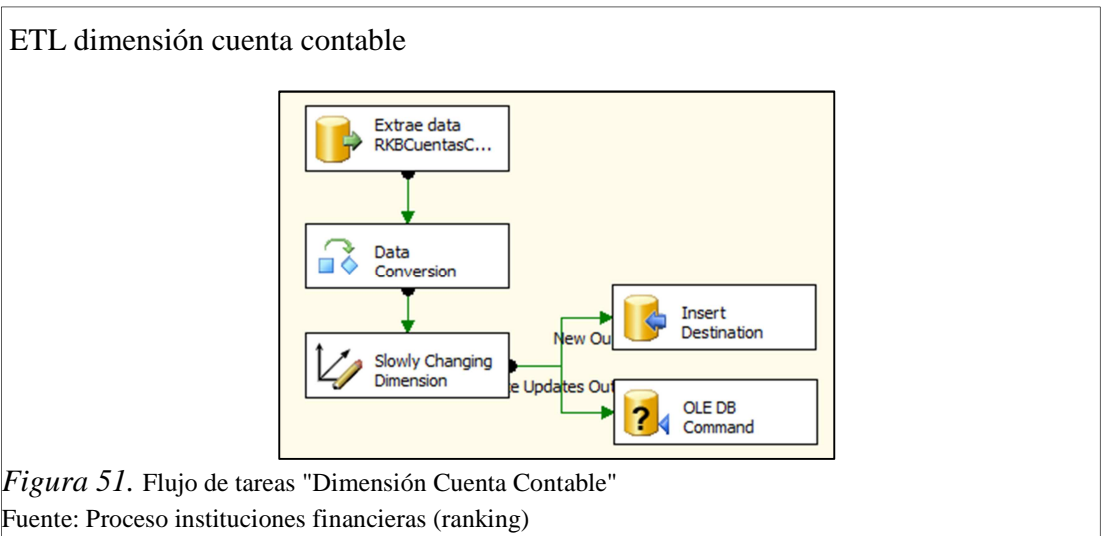
Flujo de datos

El flujo de datos consta de tres pasos, el primer paso es carga las cuentas contables que se encuentran en la tabla, el segundo paso realizar un cruce de los datos si existe diferencias realiza la actualización, el siguiente paso es el flujo de tareas, en la figura se visualiza los tres pasos del flujo de datos.



Flujo de tareas

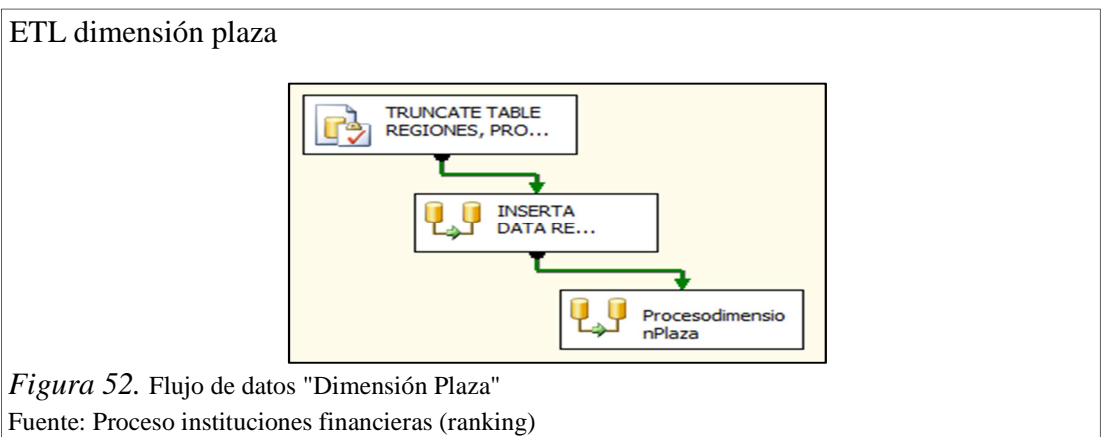
Se obtiene datos referentes a las cuentas contables desde la tabla, se convierte los datos de tal manera que tenga el formato adecuado, los datos ingresan como parámetros al SCD de tal manera que se conviertan en datos sobre una tabla de dimensión, en la siguiente ilustración se observa todo el flujo de tareas descrito.



Dimensión plaza

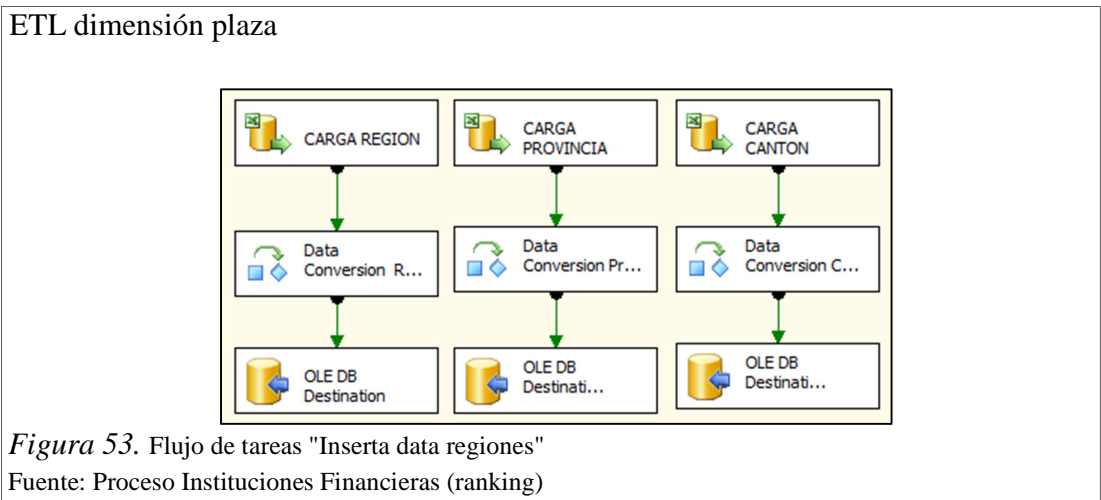
Flujo de datos

En el flujo de datos de la dimensión plaza comienza con el borrado de datos de cargas anteriores, seguido por el flujo de tareas “Inserta data regiones”, y al finalizar el flujo de tareas “Procesamiento Plaza”, como se observa en la siguiente figura.



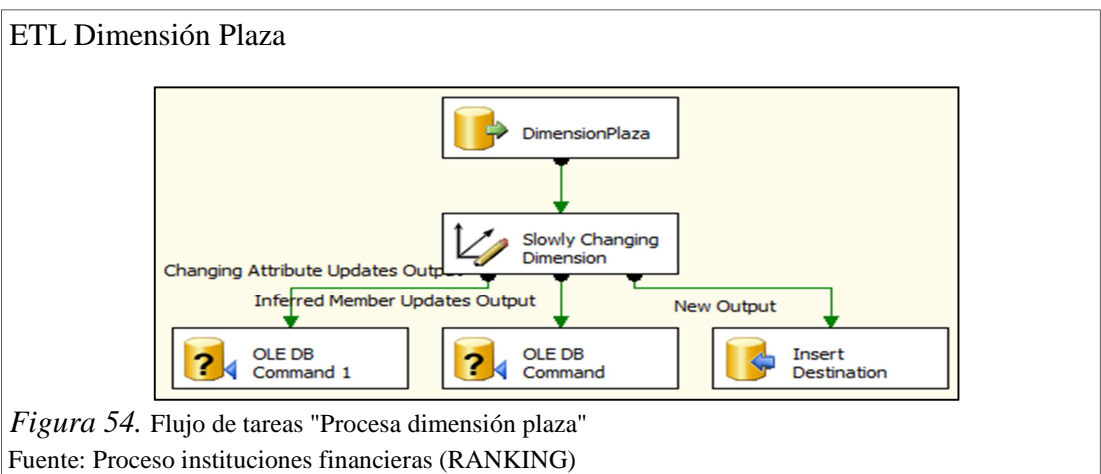
Flujo de tareas

En el primer flujo de tareas, realiza la carga de regiones, provincia y cantón, estos datos se obtiene del proceso de homologación previamente explicado, en el siguiente paso se procede con la conversión y carga en el destino.



Flujo de tareas

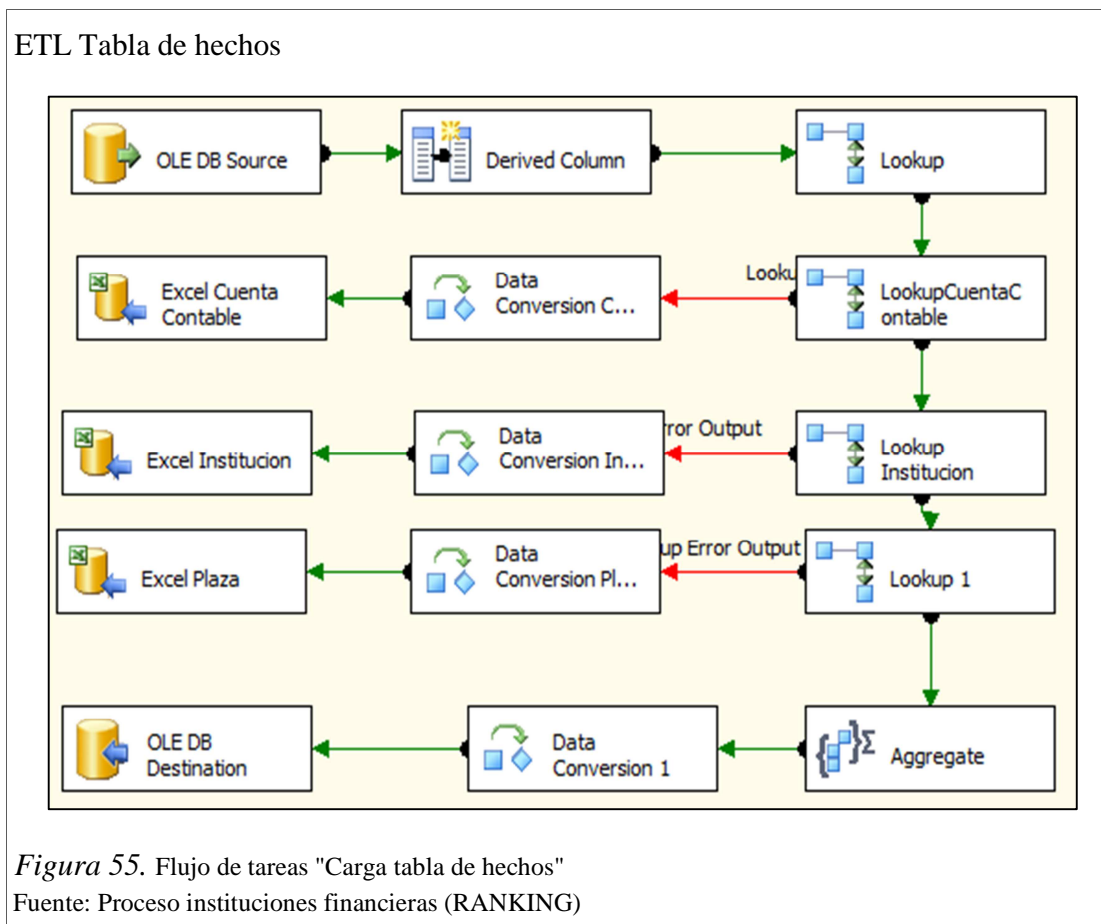
Para finalizar con el procesamiento de la dimensión plaza se obtiene los datos cargados del flujo de tareas anterior y se los transforma en tabla de dimensiones mediante el SCD.



Con la finalización de la carga de las tablas de dimensiones que conforman el datamart de instituciones financieras, se procede con la carga de la tabla de hechos, que tiene la misma lógica que se utilizó en el proceso de captaciones y colocaciones esto con el fin de garantizar la calidad de los datos.

Tabla de hechos instituciones financieras

El flujo de tareas de la tabla de hechos, como primer paso obtiene la información de instituciones financieras que previamente pasó por el proceso de homologación de datos, en el paso siguiente se añade una columna perteneciente a la fecha, la misma tiene un formato (YYYYMM), al igual que el proceso ETL de captaciones y colocaciones se utilizó el componente “Lookup” que permite la unión de los registros de búsqueda exacta con el conjunto de referencia que se basa en consultas de las tablas de hechos. Finalmente se realiza la función de agregación lo cual permite agrupar los datos. Se observa en la figura siguiente todos los pasos para la carga de la tabla de hechos para el datamart instituciones financieras.



Terminada la carga de la tabla de hechos, concluye el proceso ETL Instituciones financieras (Ranking).

3.1.6.6. Diseño del análisis y reporte de la información

Representa al conjunto de aplicaciones que consumen la información que reside en el modelo dimensional donde se consulta y se analiza los datos.

La funcionalidad viene dado por las aplicaciones de BI que aportan valor al negocio del DWH ya que la meta es dotar de nuevas capacidades o posibilidades al negocio para soportar y mejorar la toma de decisiones.

Esta fase tiene un nivel de complejidad avanzado debido que se construye la forma visible y tangible que se presenta al usuario final mostrándole la potencia de la que se dota al sistema de inteligencia de negocios.

Los reportes tanto de CAPCOL y RANKING FINCANCIERO estan centralizados en un solo DWH que permite a los usuarios del área de control financiero realizar el análisis de la información con el fin de mejorar la toma de decisiones en beneficio de la institución.

3.1.6.6.1 Data warehouse Control Financiero

La construcción del DWH Control Financiero está dada por la unión de los data marts Captaciones y Colocaciones e Instituciones Financieras, siendo el entorno front room para el usuario final. Por lo tanto se diseñó un cubo dimensional en base al DWH.

Se identificó tres dimensiones que son recurrentes en los data marts, estas dimensiones permiten combinar cada una de las tablas de hechos.

A continuación se describe el proceso que se desarrolló para diseñar el DWH Control Financiero.

1. Creación de la base de datos **MISDWH**: En esta base se alojara las tablas de dimensiones y tablas de hechos de los data marts Captaciones y colocaciones e Instituciones Financieras.

2. Identificación de las dimensiones:

- RDK_Fecha
- RKD_Institucion
- RKD_Plaza
- RKD_Producto
- RKDB_CuentaContable

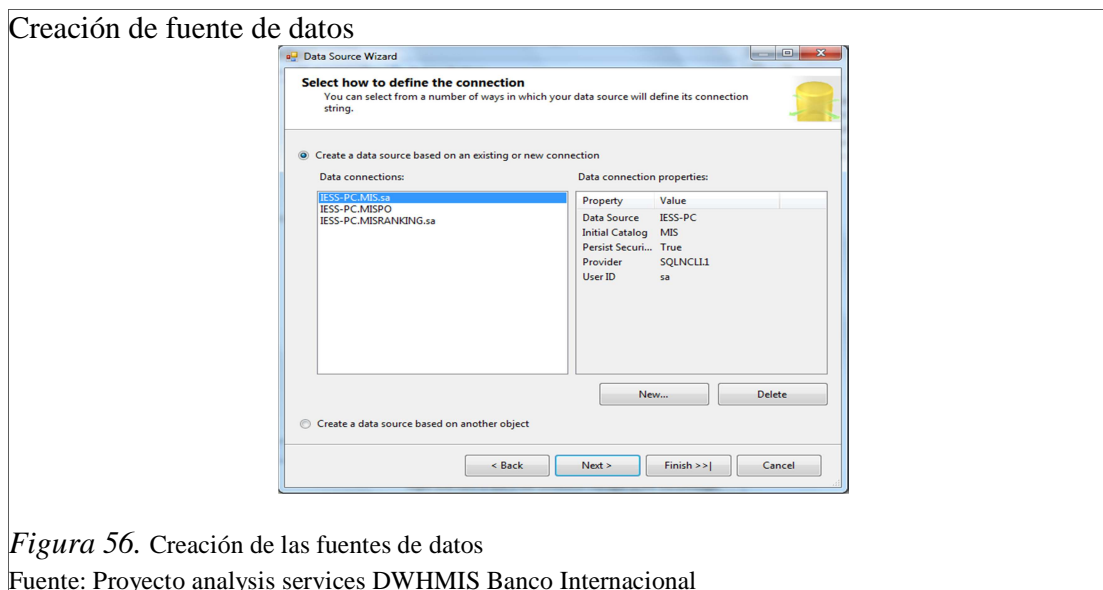
3. Tabla de hechos

- RKH_Saldos_Cuentas (Datamart Capcol)
- RKBH_SaldosCuentas (Datamart Ranking)

El proceso de creación del cubo MISDWH se desarrolló con la herramienta SQL Server Analysis Services, que permite diseñar, crear y administrar estructuras multidimensionales que contienen datos agregados desde otros orígenes de datos, como bases de datos relacionales.

Los cubos de datos necesitan fuente de datos de donde consumir la información, para el cubo MISDWH se crearon dos conexiones para las bases de datos:

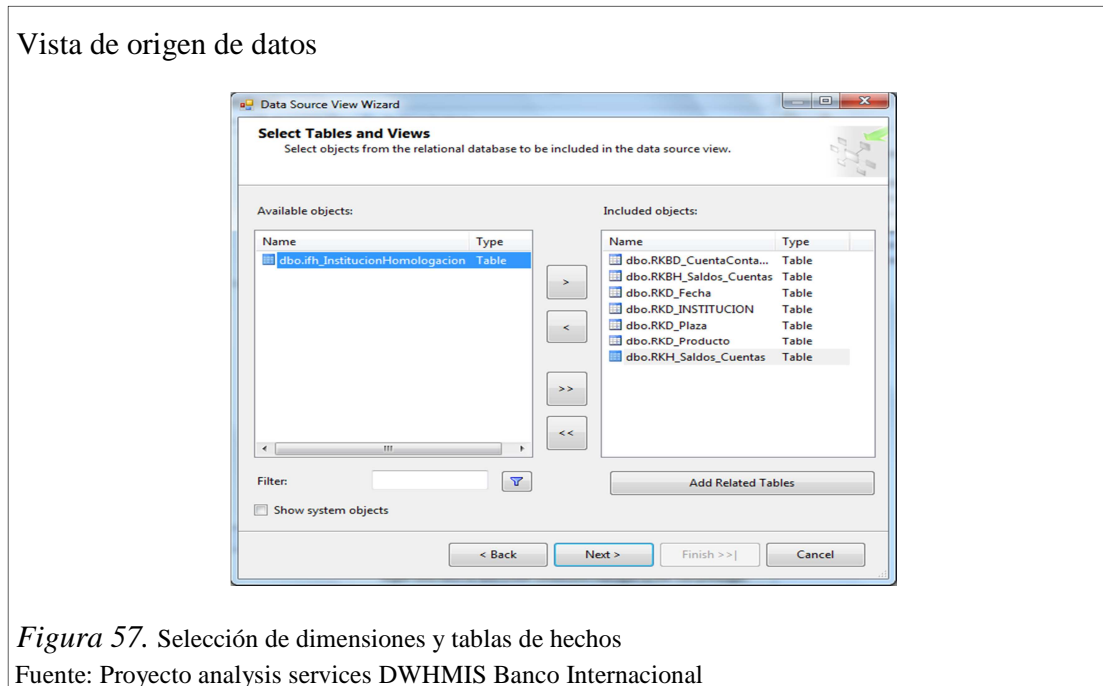
- MIS: Captaciones y Colocaciones
- MISRANKING: Instituciones Financieras



Con las conexiones ya creadas, se procede a crear las vistas de orígenes de datos.

En la vista de orígenes de datos el asistente de creación indica los objetos disponibles de las fuentes de datos, es necesario seleccionar las tablas de hechos y tablas de dimensiones de cada data mart de CAPCOL y RANKING.

En la siguiente ilustración se indica los objetos que tiene la vista de origen de datos



Es estrictamente necesario que el diseño lógico de las tablas este bien conformado por sus llaves primarias y foráneas, esto determinara la correcta creación de la vista de datos y posteriormente del cubo.

Además se tiene que tener en cuenta las relaciones que tienen que tener las tablas de hechos con el fin de no tener redundancia de datos, esto permite la integración de los datos de los dos data mart en el data warehouse del área de control financiero, consolidando la información en un solo repositorio centralizado.

Las tablas de hechos de CAPCOL y RANKING se pueden unir por las tablas de dimensiones y sus claves primarias:

- RKD_Fecha (fec_fecha)
- RKD_Plaza (plz_canton)
- RKD_Instituacion (ifi_institucion)

Esto se puede visualizar en la siguiente ilustración que indica la vista de origen de datos ya creada.

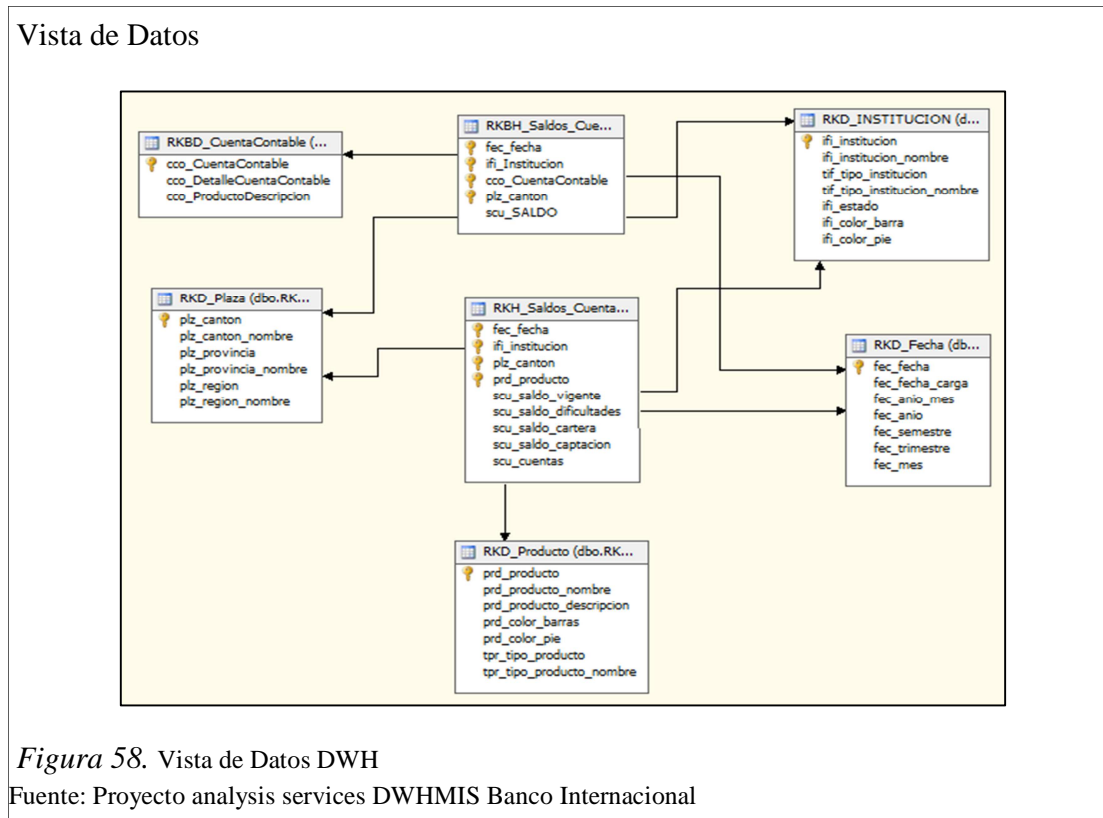


Figura 58. Vista de Datos DWH

Fuente: Proyecto analysis services DWHMIS Banco Internacional

Conformada la vista de origen de datos, se continua con la creación del cubo multidimensional, tomando como guía el asistente de creación del cubo, como se observa en la siguiente figura.

Asistente de creación del cubo

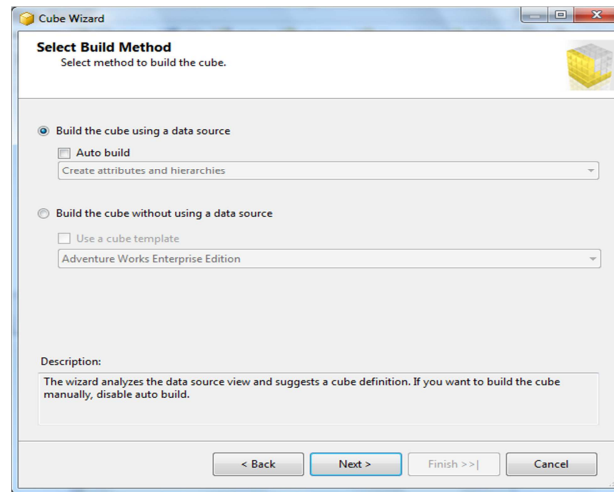


Figura 59. Método de selección construcción de un cubo

Fuente: Proyecto analysis services DWHMIS Banco Internacional

El siguiente paso es seleccionar la tabla de dimensión tiempo, adicionalmente se indica las tablas de dimensiones y las tablas de hechos del cubo, como se muestra en la ilustración.

Asistente de creación del cubo

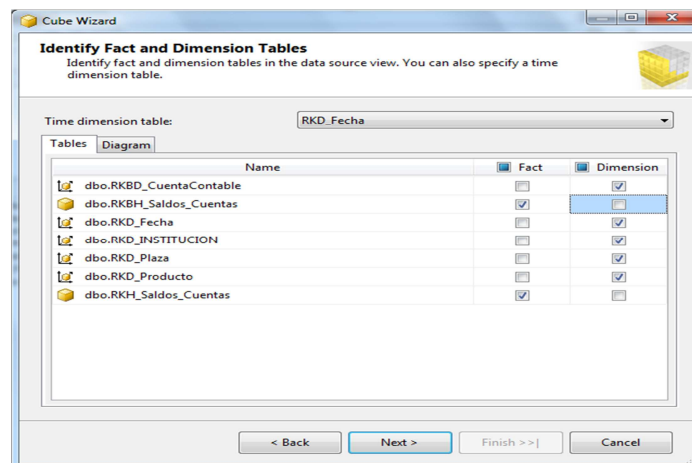
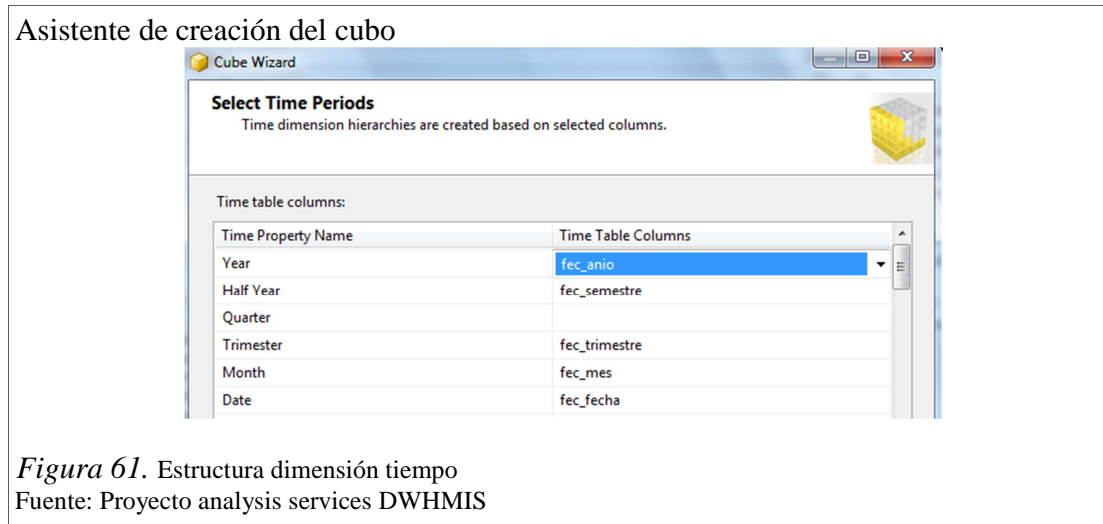


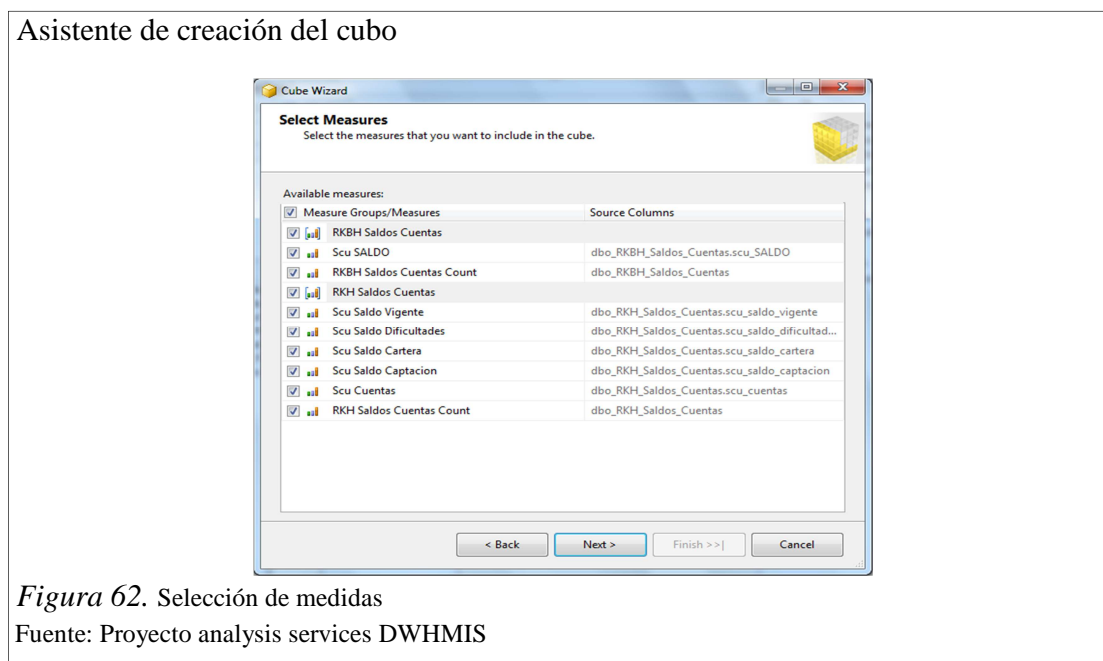
Figura 60. Selección de dimensión y tablas de hechos.

Fuente: Proyecto analysis services DWHMIS

Se procede a establecer la estructura de la dimensión tiempo indicando: año, semestre, trimestre, mes y fecha actual. Estos campos tienen que estar definidos previamente en la tabla de dimensión RKD_Fecha.



El siguiente paso del asistente de creación del cubo es la selección de las medidas, las mismas son tomadas de las tablas de hechos antes seleccionadas y que pertenecen a los data mart de captaciones y colocaciones y/o instituciones financieras.



El asistente de creación de cubos, una vez seleccionada las medidas, indica un resumen de las medidas que se van a añadir a la tabla de hechos, como se indica en la siguiente figura.

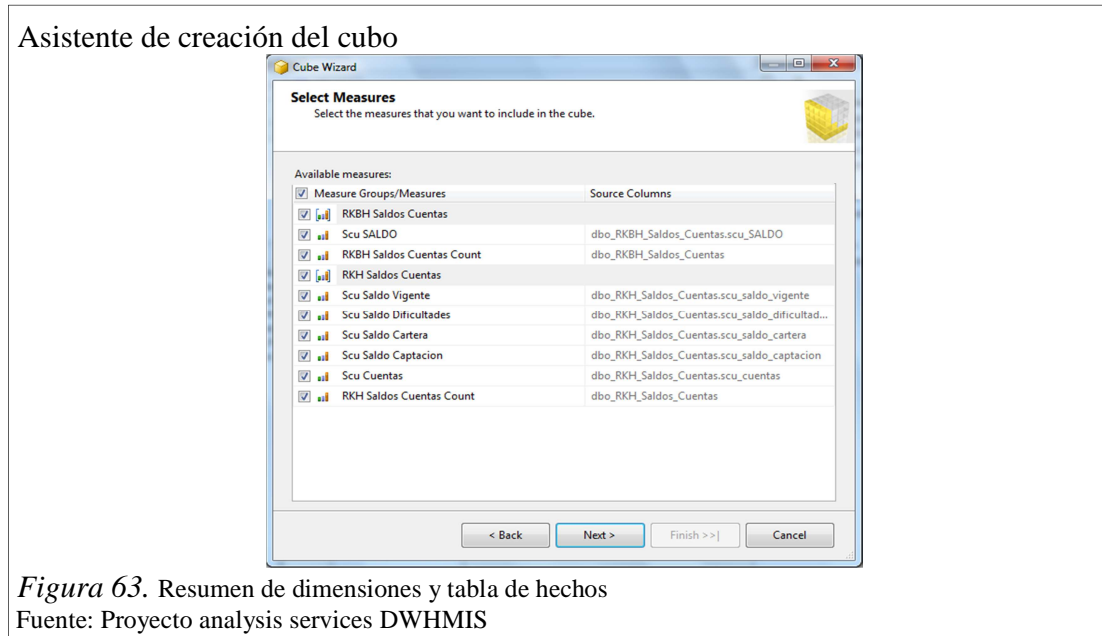


Figura 63. Resumen de dimensiones y tabla de hechos
Fuente: Proyecto analysis services DWHMIS

Con este último paso ya se tiene creado el cubo DWH con las dimensiones y tablas de hechos correspondientes a la unión de los data marts CAPCOL y RANKING.

En el siguiente capítulo se muestra los tipos de visualización que pueden consumir la información del cubo.

CAPÍTULO 4

IMPLEMENTACION DE REPORTE

4.1 Desarrollo del análisis y reporte de la información

Una vez realizado el diseño multidimensional se procede con el desarrollo del entorno front-end que se trata de la configuración de los metadatos del negocio y la infraestructura que utilizara las herramientas de BI. Además de la construcción y validación de aplicaciones BI analíticas y operacionales, es necesario establecer reuniones con el auspiciante del proyecto con el fin de presentar prototipos de la presentación de datos.

4.2 Implementación de reportes

4.2.1 Desarrollo tablero de control web

“Un tablero de control (dashboard) de inteligencia de negocios es una herramienta de visualización de datos que muestra el estado actual de métricas e indicadores clave de rendimiento para una empresa.” (Méndez del Río, 2012)

Con el fin de tener un interfaz para el usuario final, se desarrolló un tablero de control mediante la web, en el cual se indican los reportes ya predefinidos por los requerimientos del negocio. Con esto se mitiga los archivos Excel que se manejan de forma manual, para observar los indicadores del área, que tienen un alto impacto en los objetivos estratégicos de la organización.

El tablero de control se construyó en base a la herramienta Microsoft Visual Studio 2010, esto debido a que la institución cuenta con el licenciamiento oficial del proveedor, además de no interferir con las normas y estándares que tiene el área de IT, para el desarrollo de nuevas aplicaciones para usuarios finales.

4.2.2 Reportería WEB

El levantamiento de información entregada al inicio del proyecto, sirvió para detectar los reportes que están definidos en la web, lo cual ayuda al proceso de toma de decisiones.

El usuario que ingrese al sitio web, como primer gráfico encontrara un indicador del Banco Internacional con los saldos de todos los productos, realiza una comparativa de saldos del mes en curso con el mes del año anterior y de hace 2 años. Teniendo como resultado el siguiente gráfico.

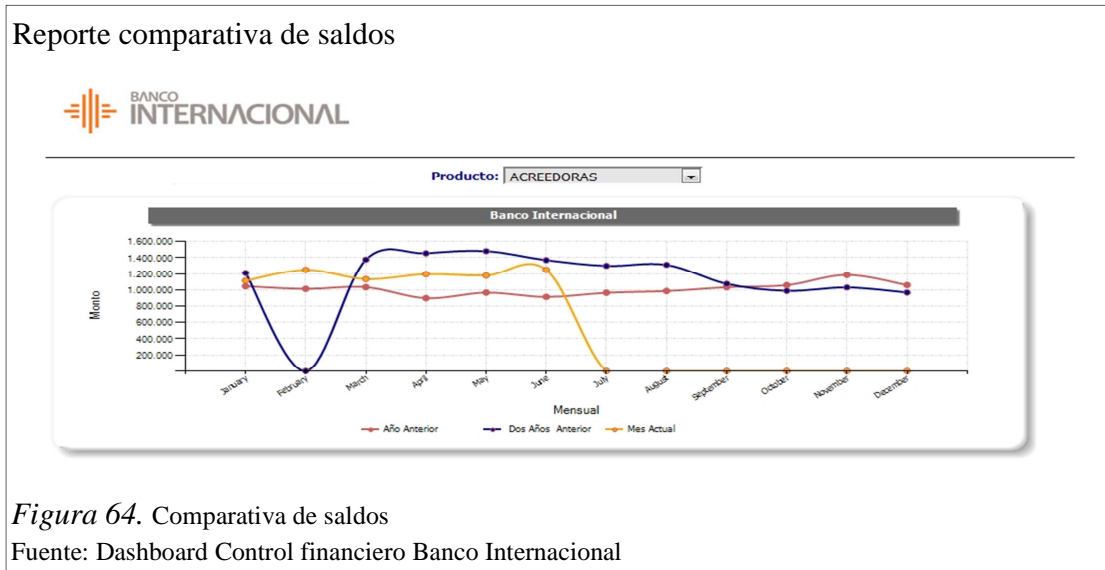


Figura 64. Comparativa de saldos

Fuente: Dashboard Control financiero Banco Internacional

A continuación se presenta un ejemplo de uno de los reportes que presenta el tablero de control (dashboard).

Principales Bancos: permite observar la posición del banco frente a las demás instituciones financieras, que cuentan con los mayores volúmenes de cartera o depósitos en un determinado momento.

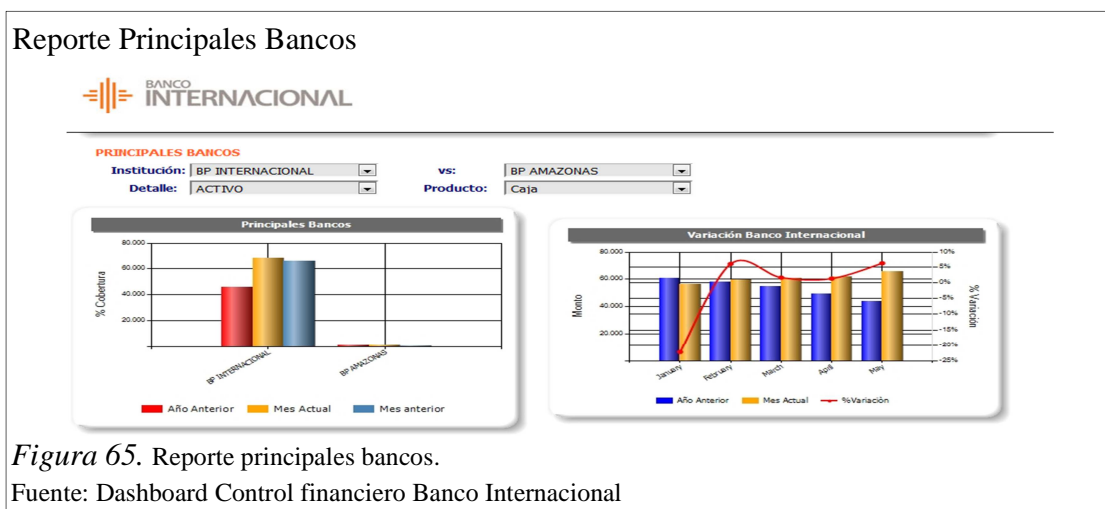


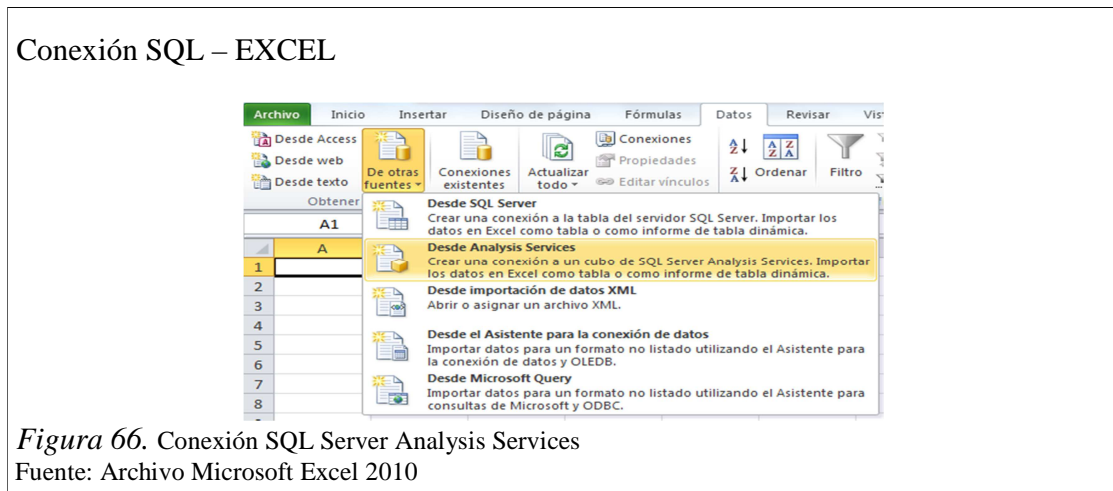
Figura 65. Reporte principales bancos.

Fuente: Dashboard Control financiero Banco Internacional

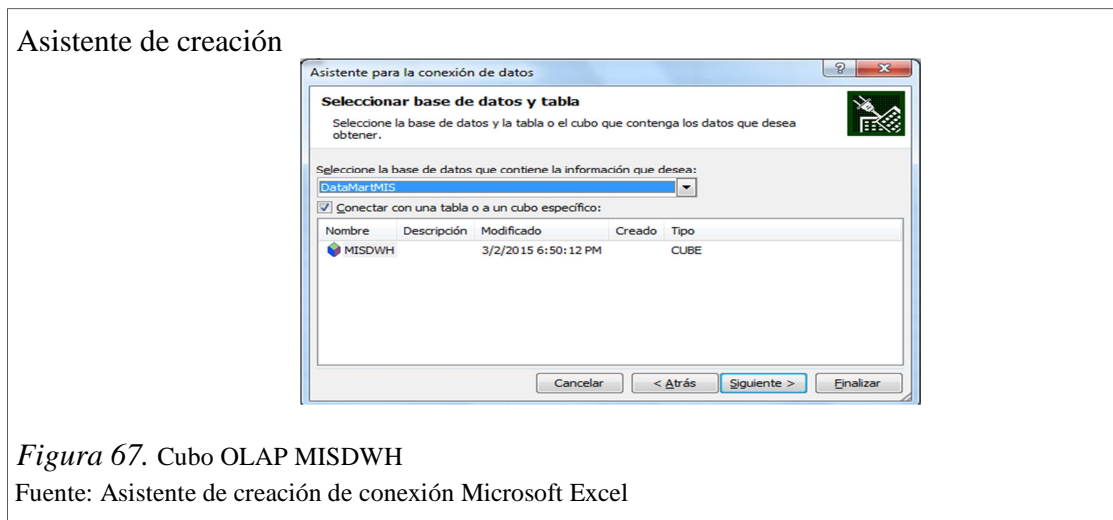
4.2.3 Reporte personalizados

La facilidad de contar con un repositorio centralizado, permite a los usuarios expertos poder armar sus reportes en base a consultas personalizadas desde la herramienta Microsoft Excel.

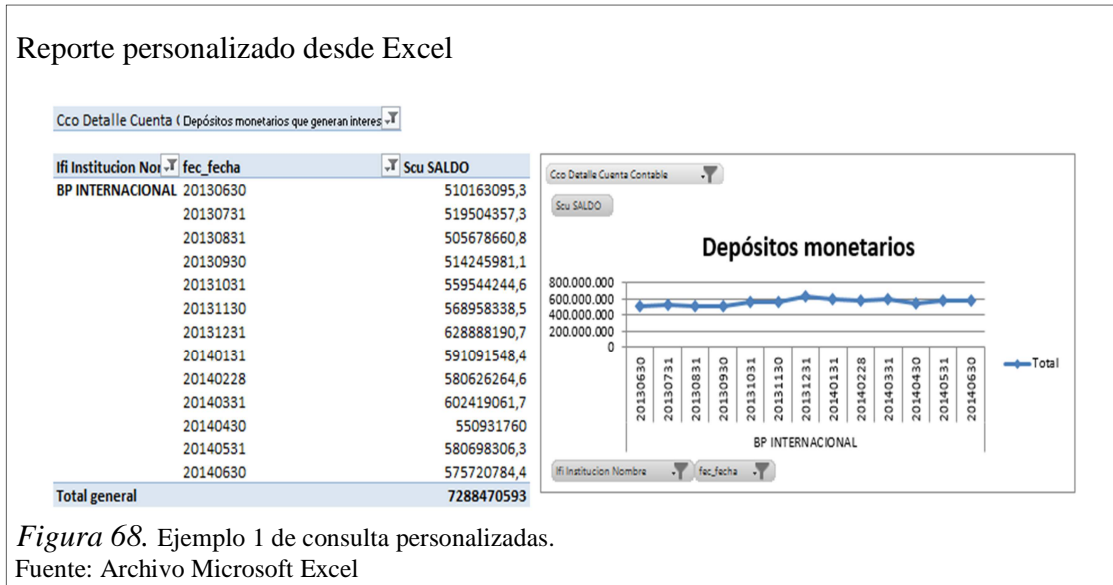
Los usuarios tienen que conectarse mediante el asistente de creación de conexión “Analysis Services”, como se detalla en la siguiente figura.



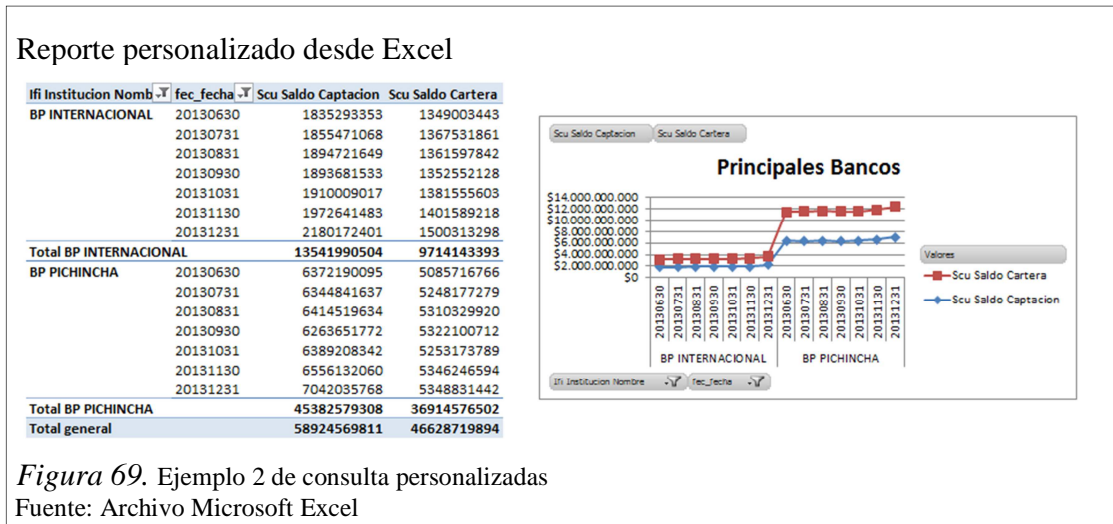
Después de ingresar el nombre del servidor y las credenciales de acceso el asistente indica el cubo que se encuentra alojado en el servidor. Como se visualiza en la siguiente figura.



Ya establecida la conexión el usuario experto puede generar los reportes de acuerdo a sus requerimientos. A continuación dos ejemplos generados por conexión directa al cubo multidimensional.



En el primer ejemplo se puede observar un reporte que indica los depósitos monetarios de una institución financiera, gracias a las tablas dinámicas que provee Excel es mucho más sencillo para el usuario obtener distintos reportes.



En ejemplo dos genera el reporte de principales bancos de acuerdo a una fecha determinada, esto se puede lograr estableciendo filtros por fecha.

4.2.4 Estrategia de capacitación

Las capacitaciones sobre el uso de herramienta web se las realizan el en área de control financiero, las mismas tienen como finalidad instruir al usuario sobre la información que arroja cada gráfico. También se capacitó a los usuarios sobre la utilización del asistente de conexión desde archivos Excel que permiten construir consultas avanzadas para un mejor detalle de la información.

4.2.5 Pruebas de operación

A continuación se describe el funcionamiento de los procesos ETL que se implementaron en el data staging, esto con el objetivo de cargar los data marts, y posteriormente estos sean consumidos por data warehouse del área de control financiero.

4.2.5.1 Carga data mart captaciones y colocaciones

Se construyó un paquete centralizado llamado “CargaMensual.dtsx” el cual hace referencia a los diferentes flujos para proporcionar la carga al data mart. Como se indica en la figura el paquete cuenta con cinco llamados que se detallan a continuación:

1. Ingresar Fecha: se debe ingresar la fecha de carga o corte en formato AAAAMMDD
2. Carga Dimensión Fecha
3. Carga Dimensión Institución
4. Carga Dimensión Cuentas
5. Carga Tabla de hechos

Los paquetes ETL unificados en un solo proceso centralizado ayudan a controlar de mejor manera las ejecuciones permitiendo a los operadores llevar un óptimo monitoreo de los flujos completados correctamente.

ETL Carga Mensual

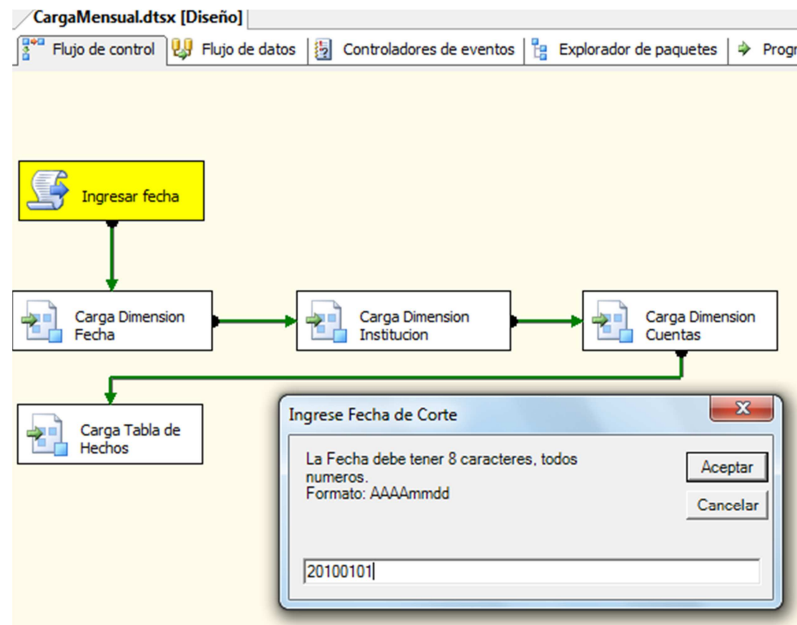


Figura 70. Paquete "Carga Mensual"

Fuente: Data mart captaciones y colocaciones

Carga Dimensión Fecha

Ingresa la fecha el flujo hace el llamado del paquete "CargadimensionFecha.dtsx". También se tiene que ingresar la fecha de carga en el mismo formato AAAAMMDD.

ETL Dimensión Fecha

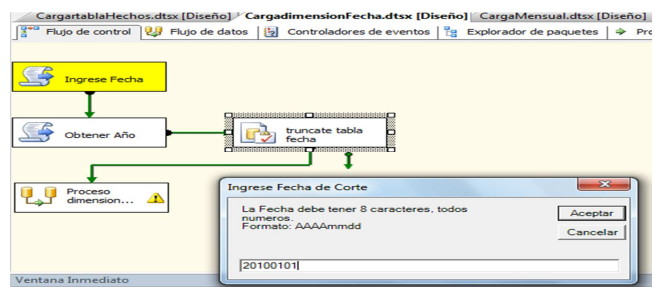
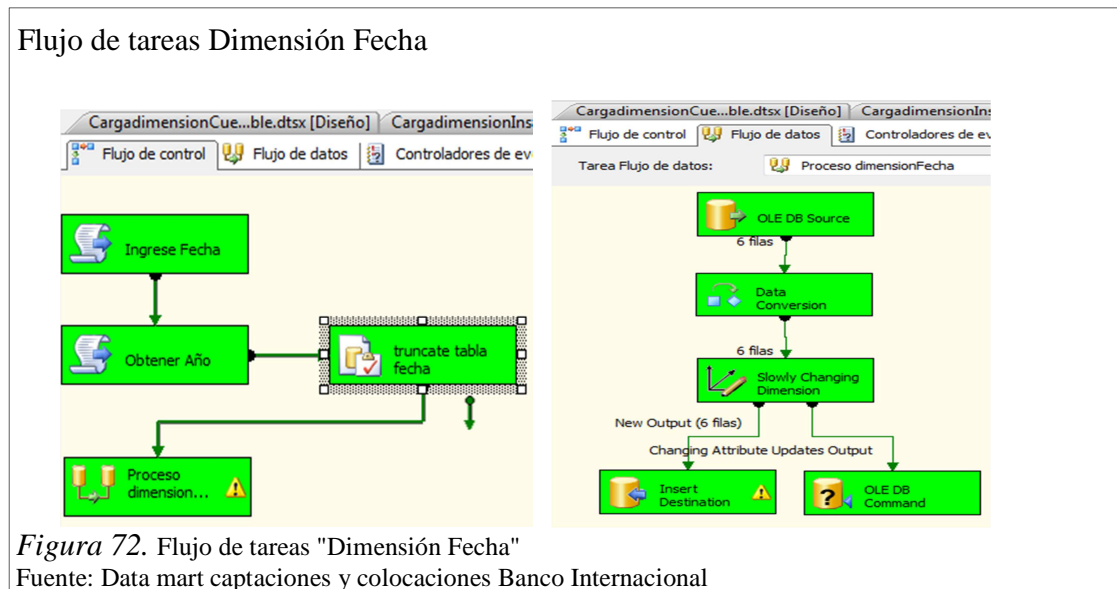


Figura 71. Flujo de control "Dimensión Fecha"

Fuente: Data mart captaciones y colocaciones

El proceso ETL dimensión fecha es el único que necesita una intervención de los operadores ya que es necesario que se ingrese la fecha de carga. Dicha fecha es el punto de partida para las cargas posteriores.

En la siguiente figura se puede observar que el proceso ETL es secuencial y depende de la ejecución correcta del anterior flujo. De esta manera se puede detectar los posibles errores en la ejecución.



El color verde en el flujo de control de la dimensión fecha, indica la carga satisfactoria; por lo tanto la prueba es exitosa.

En la siguiente figura se puede observar que al realizar una consulta a la tabla dimensión fecha se obtiene como resultado las fechas de la carga mensual.

Consulta de resultados dimensión fecha

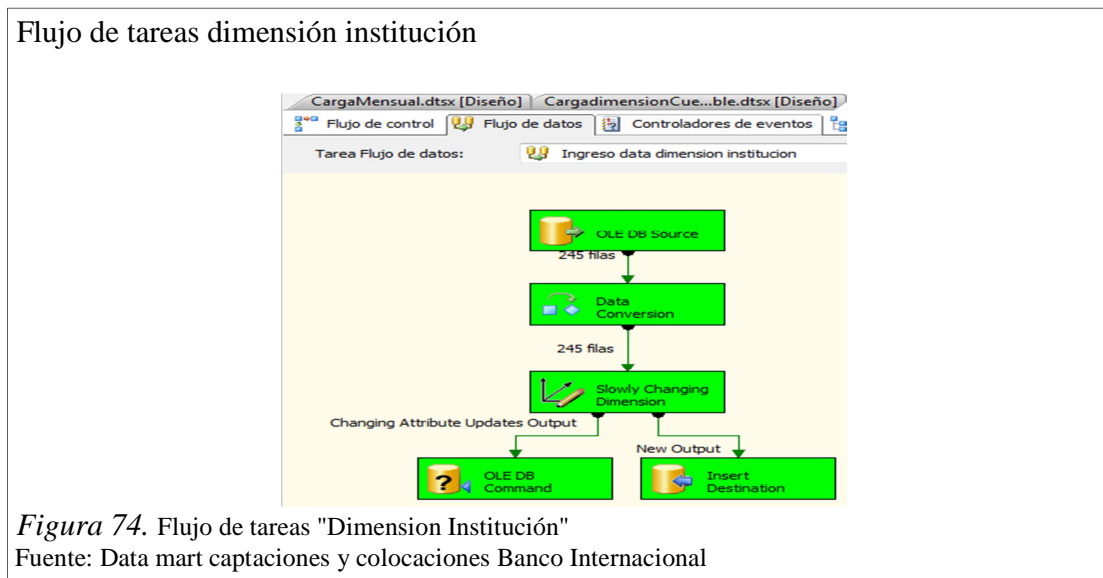
fec_Fecha	fec_FechaCarga	fec_FechaAnioMes	fec_FechaAnio	fec_FechaSemestre	fec_FechaTrimestre	fec_FechaMes
20070131	20130701	200701	2007	1	1	1
20070228	20130701	200702	2007	1	1	2
20070331	20130701	200703	2007	1	1	3
20070430	20130701	200704	2007	1	2	4
20070531	20130701	200705	2007	1	2	5

Figura 73. Validación de la tabla RCC_Fecha
 Fuente: Data mart captaciones y colocaciones Banco Internacional

Como se puede observar el campo fec_FechaCarga es donde se guarda la fecha que se ingresó manualmente al comenzar el proceso de carga.

Carga dimensión institución

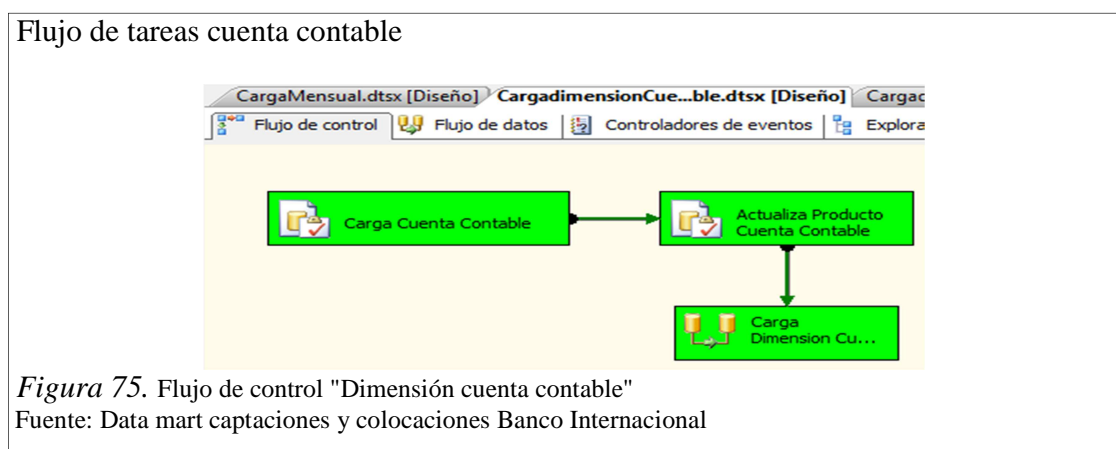
El siguiente flujo es la carga de la dimensión Institución tiene el nombre “CargadimensionInstitucion.dtsx” este flujo de tareas ingresa toda las descripciones de las instituciones financieras.



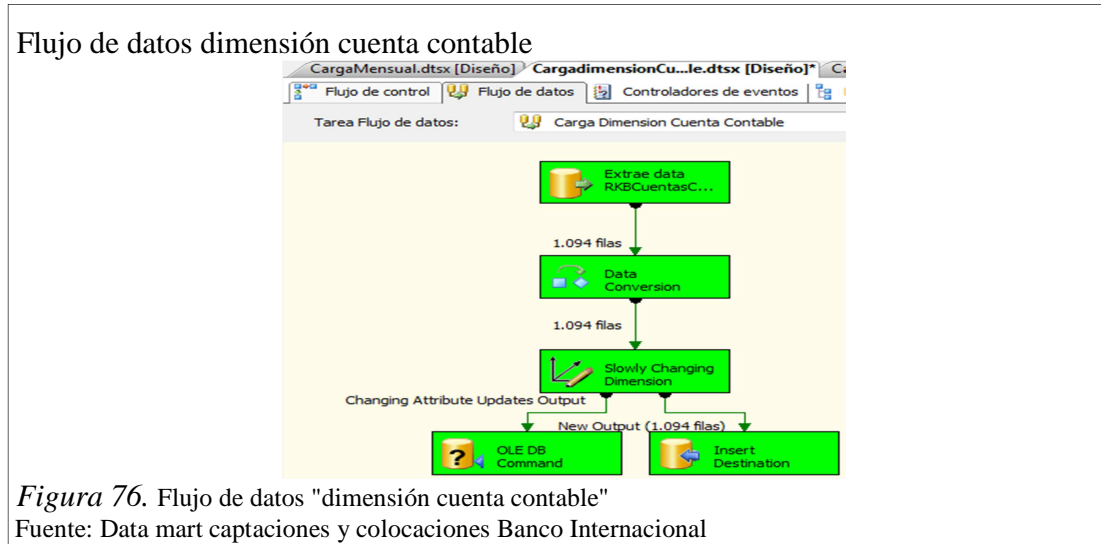
El color verde en el flujo de control de la dimensión fecha, indica la carga satisfactoria; por lo tanto la prueba es exitosa.

Carga dimensión cuenta contable

En el flujo de control de la carga contable se indica tres flujos de datos, como se indica a continuación:



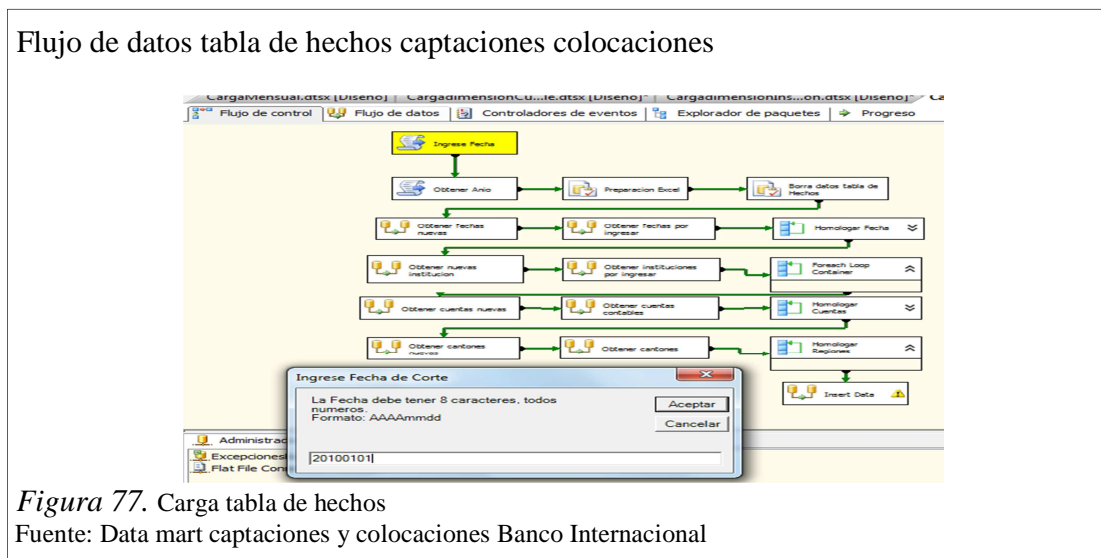
En proceso continuo con el flujo de tareas que crea la dimensión cuenta contable como se puede observar en la figura.



El color verde en el flujo de control de la dimensión fecha, indica la carga satisfactoria; por lo tanto la prueba es exitosa.

Carga de tabla de hechos

Terminado el proceso de carga de las dimensiones, el paquete principal por ultimo hace el llamado del paquete de carga de la tabla de hechos, como se indica en la siguiente figura.



En el proceso de carga de la tabla de hechos, realiza la llamada de todas las tablas de dimensiones que intervienen en el data mart ranking,

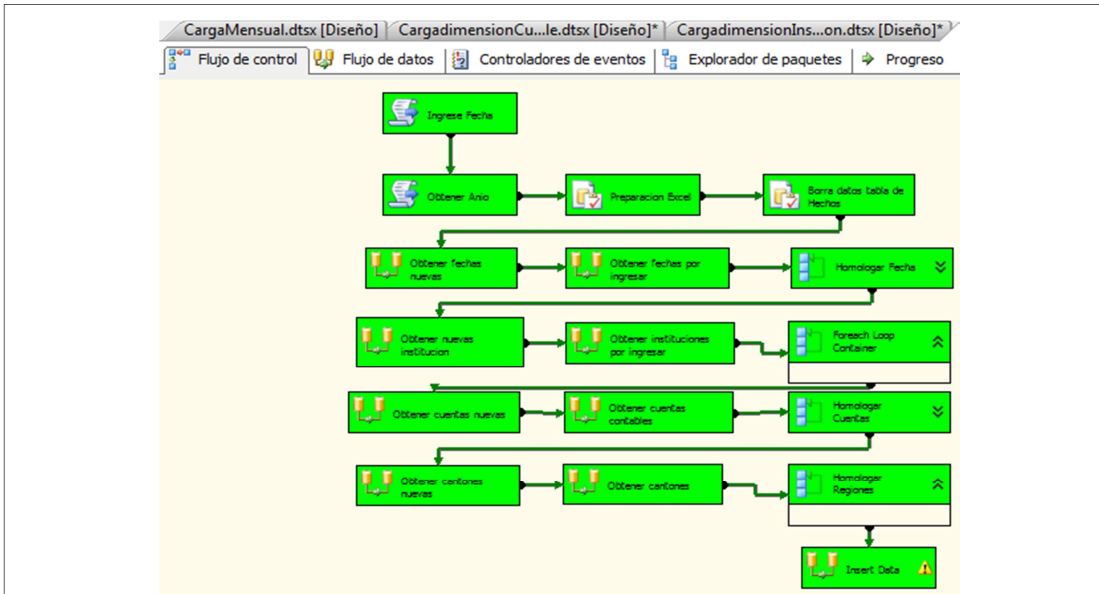


Figura 78. Carga tabla de hechos
Fuente: Data mart captaciones y colocaciones Banco Internacional

En la siguiente figura se indica todo el proceso de carga del data mart ranking financiero completado satisfactoriamente.

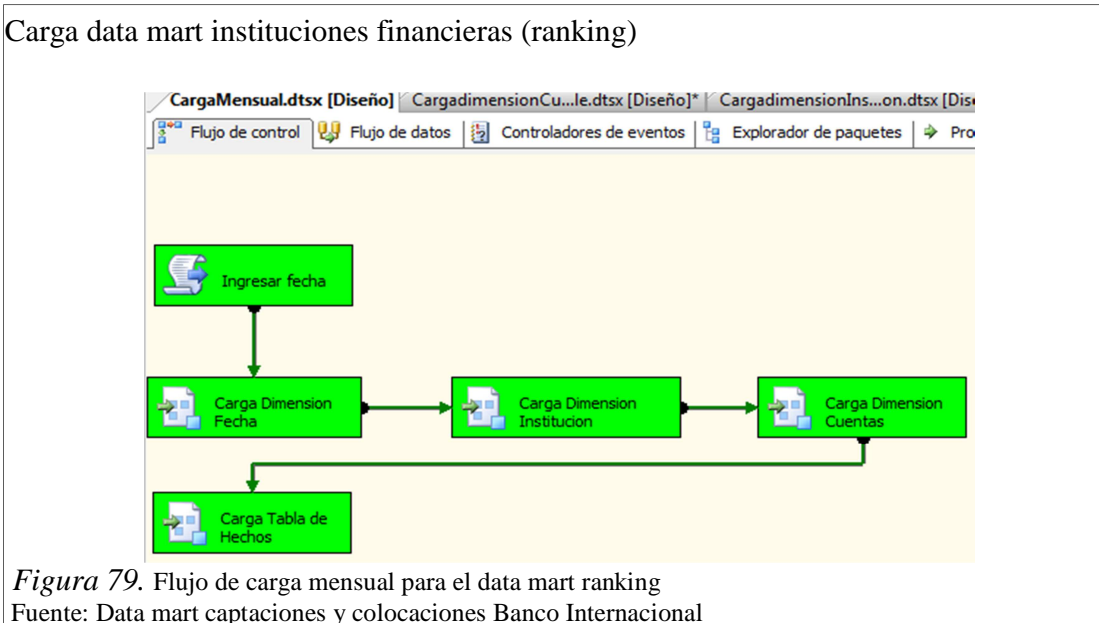


Figura 79. Flujo de carga mensual para el data mart ranking
Fuente: Data mart captaciones y colocaciones Banco Internacional

El proceso en color verde indica que el proceso fue generado satisfactoriamente.

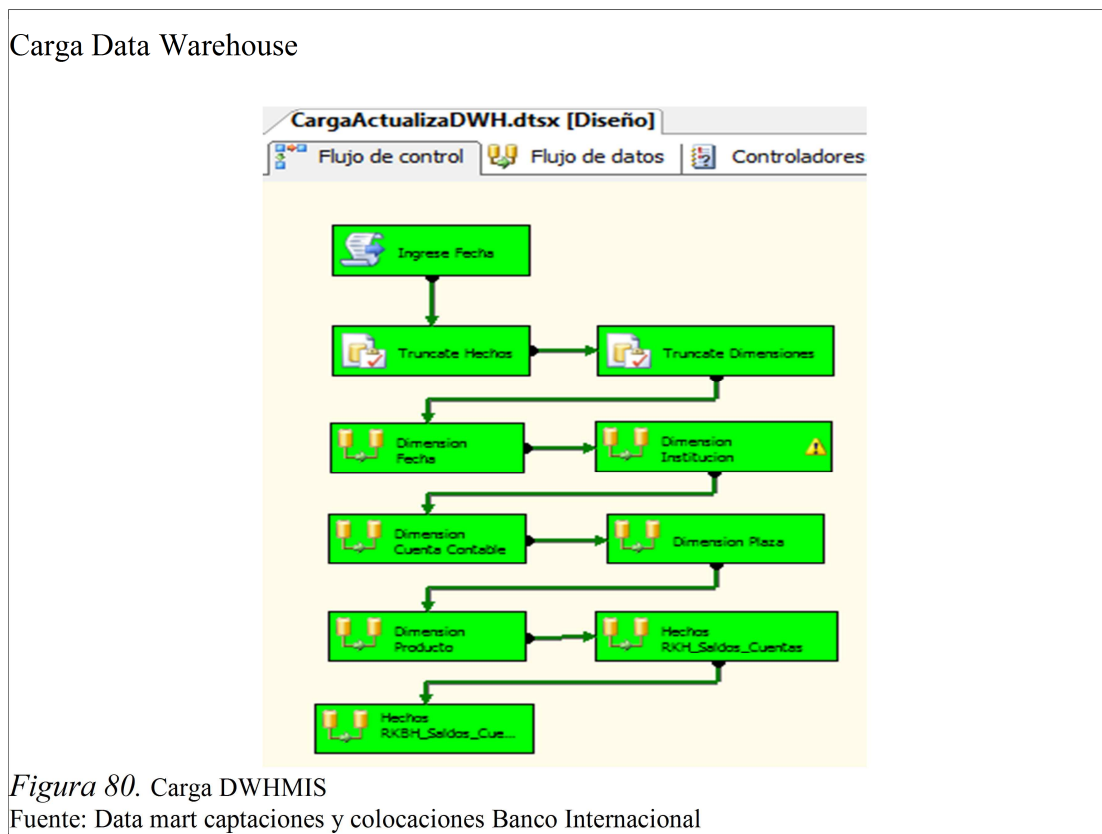
4.2.5.2 Carga data warehouse

Una vez terminado el proceso de carga de los data marts, se comienza la carga del DWH que es el repositorio centralizado de consultas.

El proceso de carga del DWH comienza con los siguientes pasos:

1. Ingresar la fecha
2. Borrar la tabla de hechos y las tablas de dimensiones
3. Carga dimensiones: Fecha, Institución, Cuenta Contable, Plaza, Producto
4. Carga tabla de hechos: Saldo Cuentas

En la siguiente figura se indica todo el flujo de carga.



El proceso en color verde indica que el proceso fue generado satisfactoriamente. Y se encuentra listo para realizar las consultas OLAP.

CONCLUSIONES

- Con el desarrollo del proyecto Inteligencia de negocios se ha logrado dar una solución completa y rentable a la principal problemática del área de control financiero del Banco Internacional. Inicialmente se tiene una área sin las herramientas necesarias para tomar decisiones óptimas sobre la “Información de Instituciones Bancarias e Información de Captaciones y Colocaciones “, debido principalmente a la dispersión de datos entre las diferentes fuentes de información, falta de control, manipulación manual de archivos y organización en los procesos que no son apoyados por ningún sistema de información que pueda automatizar la creación de reportes gerenciales.
- La integración de la Inteligencia de negocios con los almacenes de datos proporciona repositorios de información que representan la plataforma para emitir los análisis de datos y explotación de conocimiento a cargo de los procesos especializados como OLAP. Esta constituye como la herramienta que produce los elementos de información necesarios para la toma de decisiones.
- Es indispensable el uso de herramientas BI en las instituciones financieras ya que por su escalabilidad permite una gran adaptación a los procesos, facilitan la consolidación y obtención de los datos del sistema financiero, que permite destinar menos tiempo en su recolección y extracción, para enfocarse más en el estudio y análisis de la información.
- La metodología propuesta por Ralph Kimball para implementar la solución de Inteligencia de negocios, ayudó en la gestión del proyecto de manera ordenada, así mismo como sus etapas fueron fáciles de aplicar en la construcción de los procesos ETL(Extracción, Transformación y Carga) y el modelado dimensional.

- Se usó herramientas de la suite de Microsoft SQL Server 2005, para la construcción de los procesos ETL se utilizó SQL Sever Integration Services, y SQL Server Analysis Services para el modelado del cubo OLAP, debido a la gran documentación que poseen dichas herramientas y el conocimiento del equipo de trabajo no se presentaron inconvenientes en el desarrollo de la solución BI.
- El sistema como está diseñado permite implementar nuevas funcionalidades en los reportes de acuerdo a los requerimientos del negocio. La herramienta y la metodología desarrollada, son fáciles de migrar a realidades de otras instituciones financieras.

RECOMENDACIONES

- Estudiar de manera conjunta con el área técnica los requerimientos del negocio levantados, de tal manera que se pueda determinar si existen las fuentes de datos para dar soporte a los requerimientos establecidos, esto con el fin de no ofrecer reportes que no se puedan generar.
- Es importante destacar que los métodos de extracción, transformación y carga son parte fundamental de la inteligencias de negocios, deben estar bien definidos y estructurados, de tal manera que soporte inmensos volúmenes de información, datos “sucios” e incoherentes.
- Se recomienda el uso de una metodología para el desarrollo de proyectos de inteligencia de negocios y/o almacenes de datos, sus etapas evolutivas e incrementales, así como sus técnicas y documentación se adaptan a cualquier escenario propuesto.

LISTA DE REFERENCIA

- Banco Internacional. (5 de Agosto de 2014). *Banco Internacional*. Obtenido de Banco Internacional: <http://www.bancointernacional.com.ec/>
- Brian Knigh, Erik Veerman . (2007). *SQL Server 2005 Integration Services*. Indianapolis: Wiley.
- Curto Díaz, J. (2012). *Introducción al Business Intelligence*. Santiago de Chile: Editorial UOC.
- Date, C. J. (2012). *Introducción a los sistemas de bases de datos*. Barcelona: Pearson .
- Inmon, W. B. (2008). *The Architecture for the Next Generation of Data Warehousing*. Burlington: Elsevier Inc.
- Inmon, W. H. (2010). *Building the Data Warehouse*. California: Wiley.
- Internacional, B. (01 de 10 de 2014). <http://www.bancointernacional.com.ec/>. Obtenido de <http://www.bancointernacional.com.ec/>.
- John C. Hancock, R. T. (2006). *Practical Business Intelligence with SQL Server 2005*. California: Pearson Education.
- Kimball, R. (2008). *The Data Warehouse Lifecycle Toolkit*. New York: Wiley Inc.
- Kimball, R. (2008). *The Data Warehouse Toolkit*. New York: John Wiley & Sons, Inc.
- Laudon, K. C., & Laudon, J. P. (2010). *Sistemas de información gerencial*. Mexico: Pearson.
- Leonard, A., Masson, M., & Mitchell, T. (2012). *SQL Server 2012 Integration Services Design Patterns*. San Francisco: Apress.
- Mazón , N., Pardillo, J., & Trujillo, J. C. (2010). *Diseño y explotación de almacenes de datos*. Alicante: Editorial Club Universitario.

- Méndez del Río, L. (2012). *Más allá del Business Intelligence*. Barcelona: Ediciones Gestion 2000.
- Microsoft. (2014). *Microsoft TechNet*. Obtenido de Microsoft TechNet.
- Mike Gunderloy, T. S. (2006). *SQL Server's Developer's Guide to OLAP with Analysis Services*. San Francisco: John Wiley & Sons.
- msdn.microsoft.com. (2014). *MSDN*. Obtenido de MSDN.
- PRABHU, C. (2012). *DATA WAREHOUSING: Concepts, Techniques, Products and Applications*. New Delhi: PHI Learning.
- Sánchez, J. (2010). *Principios sobre Bases de Datos Relacionales*.
- SBS, S. d. (20 de 05 de 2014). *Superintendencia de Bancos y Seguros*. Recuperado el 01 de 10 de 2014, de SBS: www.sbs.gob.ec
- Scheps, S. (2011). *Business Intelligence For Dummies*. Indianapolis: John Wiley & Sons.

ANEXOS

Anexo 1. Reunión determinación del alcance



Minuta de reunión

Datos de la reunión

Fecha: Quito 7 de Febrero del 2014

Lugar: Banco Internacional

Objetivo/s: Determinar el alcance del proyecto.

Reunión convocada por: Erika Erazo, Paúl Ortiz

Participantes

Nombre y apellido	Cargo	Referencia
Erika Erazo	Programador	Tesista
Henry Pozo	Analista de información gerencial	Supervisor
Paul Ortiz	Tesista	Tesista

Temas tratados

1. Tema 1

Detalle del tema:

- Elaboración de una aplicación de Inteligencia de Negocios (BI) para el Área de control financiero.

2. Tema 2

Detalle del tema:

- Alcance del proyecto estará con formado por dos data mart independientes los mismo que conformaran un Data Warehouse, el cual brindará información relevante para la toma de decisiones.

Compromisos asumidos

Descripción	Responsable
<ul style="list-style-type: none">• Realizar el análisis de información para la elaboración de un Data Warehouse.• Realizar dos data mart que conformaran un Data Warehouse.	Erika Erazo Paul Ortiz

Temas pendientes:

1. Tema 1

Detalle del tema: Arquitectura de desarrollo.

Próxima Reunión: 12 de Abril 2014

FIRMA

Anexo 2. Reunión determinación de la arquitectura data warehouse .



Minuta de reunión

Datos de la reunión

Fecha: Quito 14 de Abril del 2014

Lugar: Banco Internacional

Objetivo/s: Arquitectura de un Data Warehouse.

Reunión convocada por: Erika Erazo, Paul Ortiz

Participantes

Nombre y apellido	Cargo	Referencia
Erika Erazo	Programador	Tesista
Henry Pozo	Analista de información gerencial	Supervisor
Paul Ortiz	Tesista	Tesista

Temas tratados

1. Tema 1

Detalle del tema:

- Escoger la arquitectura de un Data Warehouse.
- La arquitectura a utilizar es MD (Arquitectura Multidimensional) la misma que permite tener dos data mart independientes.

Compromisos asumidos

Descripción	Responsable
Realizar la arquitectura Data Warehouse como MD (Arquitectura Multidimensional). El mismo que estará conformado por dos data mart independientes.	Erika Erazo Paul Ortiz

Temas pendientes:

1. Tema 1

Detalle del tema: Metodología de desarrollo.

Próxima Reunión: 16 de Mayo 2014

FIRMA

Anexo 3. Revisión de la metodología de desarrollo



Minuta de reunión

Datos de la reunión

Fecha: Quito 16 de Mayo del 2014

Lugar: Banco Internacional

Objetivo/s: Metodología de desarrollo

Reunión convocada por: Erika Erazo, Paul Ortiz

Participantes

Nombre y apellido	Cargo	Referencia
Erika Erazo	Programador	Tesista
Henry Pozo	Analista de información gerencial	Supervisor
Paul Ortiz	Tesista	Tesista

Temas tratados

1. Tema 1

Detalle del tema:

- Analizar la metodología con la cual se va a desarrollar la solución BI.

Compromisos asumidos

Descripción	Responsable
Realizar el análisis de la metodología a utilizar para el desarrollo de la solución BI. Comparación de ventajas y desventajas entre la metodología de Ralph kimball, Inmon	Erika Erazo Paul Ortiz

Temas pendientes:

1. Tema 1

Detalle del tema: Escoger la metodología de desarrollo.

FIRMA

Próxima Reunión: 20 de Junio 2014

Anexo 4. Estudio de la metodología a implementar



Minuta de reunión

Datos de la reunión

Fecha: Quito 20 de Junio del 2014

Lugar: Banco Internacional

Objetivo/s: Escoger la metodología de desarrollo.

Reunión convocada por: Erika Erazo, Paul Ortiz

Participantes

Nombre y apellido	Cargo	Referencia
Erika Erazo	Programador	Tesista
Henry Pozo	Analista de información gerencial	Supervisor
Paul Ortiz	Tesista	Tesista

Temas tratados

1. Tema 1

Detalle del tema:

- Se escoge la metodología de Ralph Kimball para la gestión del proyecto, ya que proporciona una dirección versátil y unas series de herramientas prácticas que ayudan para la implementación de la solución BI.

Compromisos asumidos

Descripción	Responsable
Desarrollar la solución de inteligencia de negocios utilizando la metodología de Ralph Kimball.	Erika Erazo Paul Ortiz

Temas pendientes:

1. Tema 1

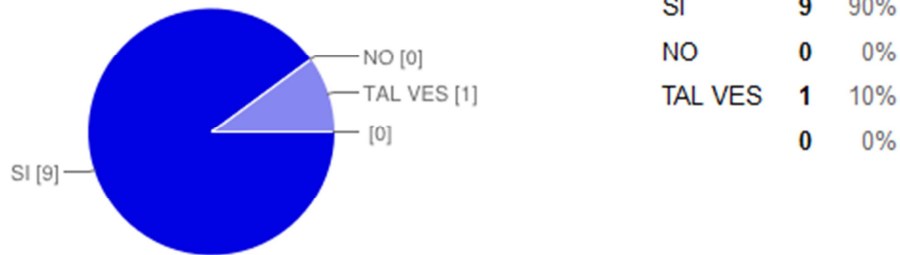
Detalle del tema: Escoger la metodología de desarrollo.

Próxima Reunión:

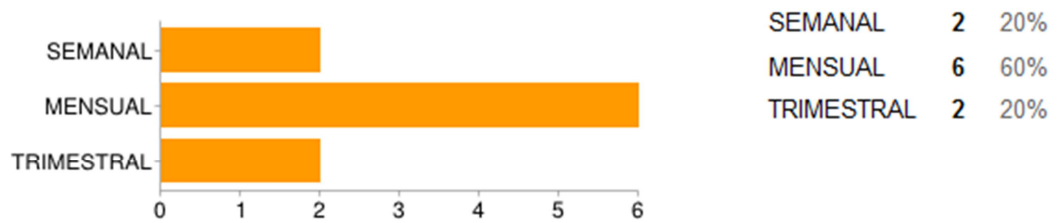
FIRMA

Anexo 5. Encuesta de para medir la demanda de una aplicación BI

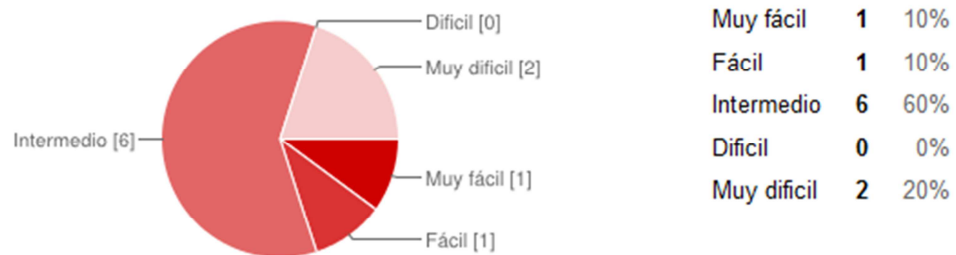
1. ¿Trabaja o ha trabajado con información de: instituciones financieras, captaciones y colocaciones?



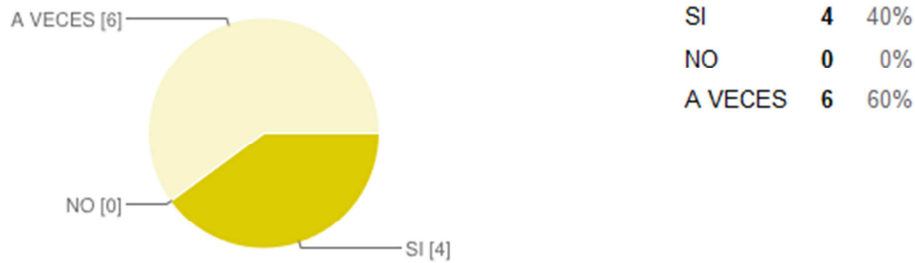
2. ¿Con que frecuencia utiliza dicha información?



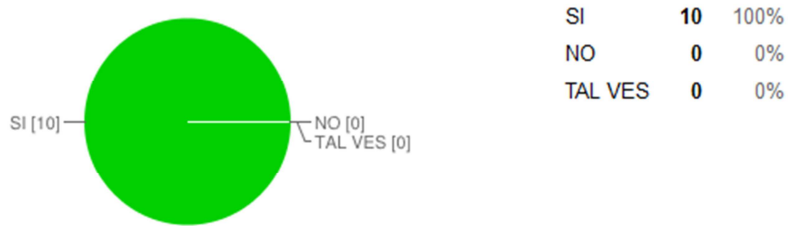
3. ¿Indique que tan complejo es obtener la información de: instituciones financieras, captaciones y colocaciones? siendo:



4. ¿Actualmente la información que se extrae para instituciones financieras, captaciones y colocaciones es consistente?



5. ¿Le gustaría que la información este actualizada de forma automática, sin realizar procesos que intervengan del área tecnológico y que no se dilaten la entrega de información?



6. ¿Esta información automatizada sería de gran utilidad? ¿Por qué?

- Por qué de esta manera se tiene una data consistente y de alta disponibilidad sin tener que depender de un proceso del área de tecnología y así poder cumplir con los procesos dependientes del área de financiero.
- Sí, facilitaría muchos procesos que están de por medio para la consulta de información que se requiera.
- Seria de mucha utilidad tanto para el área de sistemas en disminuir la carga operativa y para control financiero de muy utilidad al obtener la información de alta calidad y disponibilidad.

- Permite obtener resultados rápidamente para medirnos como estamos con respecto a la competencia de una forma rápida y mitigando los errores humano en manipulación de información.
 - Porque se puede tomar decisiones a tiempo y generar ventaja sobre la competencia.
 - Por qué ayuda al banco a tomar decisiones más acertadas sobre ciertos productos que ofrece.
 - Para mejorar la toma de decisiones.
 - Sería de utilidad ya que medio de un análisis de esta información obtenida se pudiera tomar las medidas necesarias dentro de la institución financiera para incrementar el número de captaciones y colocaciones.
 - Para mejoramiento de proceso y toma de decisiones.
 - Liberaría mucho tiempo que se puede dedicar al análisis. Hoy en día mucho de ese tiempo se dedica a la tarea operativa de procesamiento de información.
7. ¿Le gustaría tener una solución que obtenga información de instituciones financieras, captaciones y colocaciones de una forma verídica y eficaz?



Anexo 6. Encuesta para medir la experiencia del equipo

Encuesta para medir la experiencia del equipo sobre las herramientas BI actuales

1. ¿Escoja el software que permite la factibilidad de uso para los usuarios finales?



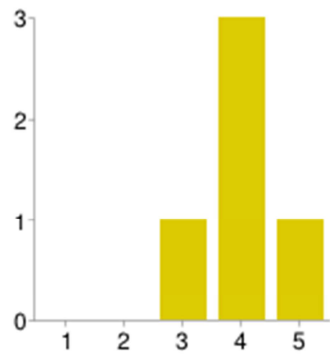
2. ¿Cuál del siguiente software permite la factibilidad de uso para los usuarios desarrolladores?



3. ¿Microsoft SQL Server Enterprise permite escalabilidad de volumen de datos?

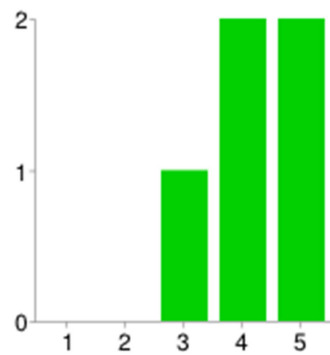


4. ¿Oracle Business Intelligence permite escalabilidad de volumen de datos?



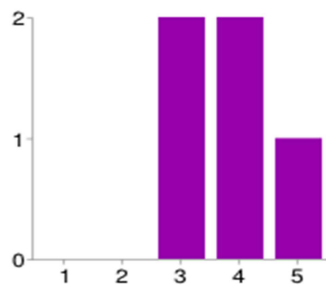
1	0	0%
2	0	0%
3	1	20%
4	3	60%
5	1	20%

5. ¿Microsoft SQL Server Enterprise permite escalabilidad de usuarios?



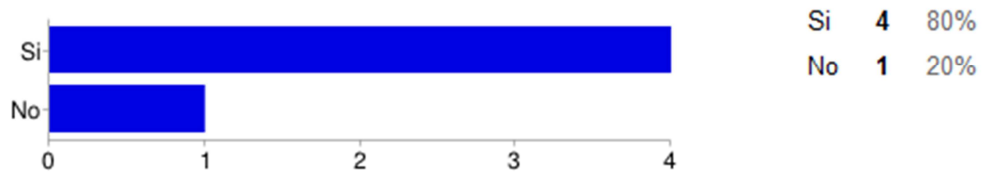
1	0	0%
2	0	0%
3	1	20%
4	2	40%
5	2	40%

6. ¿Oracle Business Intelligence permite escalabilidad de volumen de datos?

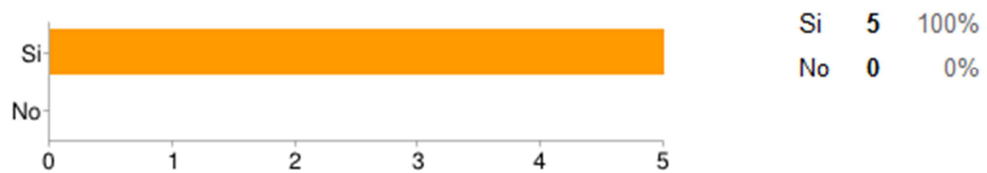


1	0	0%
2	0	0%
3	2	40%
4	2	40%
5	1	20%

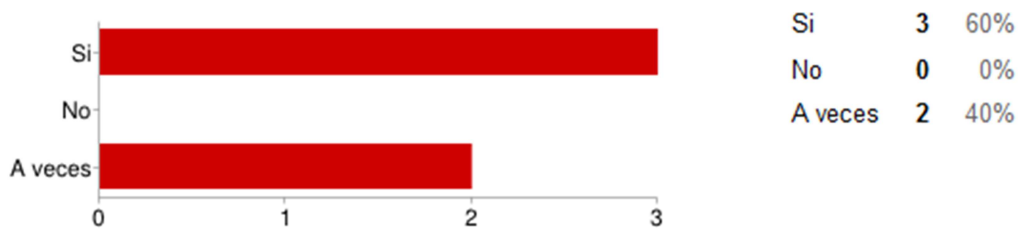
7. Microsoft SQL Server Enterprise ¿Cuál de los siguientes software permite la integración con distintas plataformas?



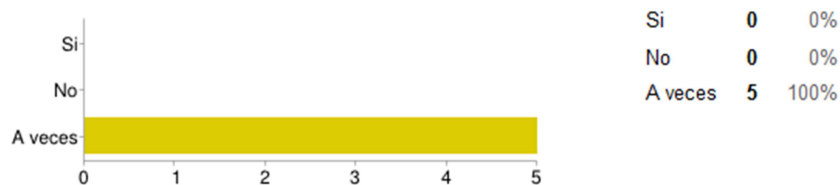
8. Oracle Business Intelligence ¿Cuál de los siguientes software permite la integración con distintas plataformas?



9. Microsoft SQL Server Enterprise ¿Cuál de los siguientes software cuenta con soporte técnico y la ayuda es inmediata?



10. Oracle Business Intelligence ¿Cuál de los siguientes software cuenta con soporte técnico y la ayuda es inmediata?



11. ¿Qué motor de bases de datos maneja mejor?, ¿por qué?

- SQL por soporte técnico y documentación.
- SQL Server, puesto a que es un producto de Microsoft con muy buena documentación, escalabilidad, rendimiento y totalmente compatible con ambientes Windows tanto en servidor como cliente.
- Microsoft Sql he trabajado en este motor ya varios años creo que permite optimizar procesos, la única desventaja que le veo es las licencias son muy costosas.
- Microsoft SQL Server Enterprise.
- Microsoft SQL Server, porque es el motor de base de datos que actualmente utiliza en la institución que presto mis servicios.

GLOSARIO

BI (BUSINESS INTELLIGENT)

BD (BASE DE DATOS)

BDL (BUSINESS INTELLIGENT LIFECYCLE)

CIF (CORPORATE INFORMATION FACTORY)

DBA (DATABASE ADMINISTRATOR)

DTS (DATA TRANSFORMATION SERVICES)

DWH (DATA WAREHOUSE)

EII (ENTERPRISE INFORMATION INTEGRATION)

ETL (EXTRECT, TRANSFORM AND LOAD)

FK (FOREIGN KEY)

HOLAP (HYBRID ONLINE ANALYTICAL PROGRESSING)

HTML (HYPERTEXT MARKUP LANGUAGE)

HW (HADWHARE)

MD (MODEL DIMENSIONAL)

MOLAP (MULTIDIMENSIONAL ONLINE ANALYTICAL PROCESSING)

ROLAP (RELATIONAL ONLINE ANALYTICAL PROGRESSING)

ODS (OPERATIONAL DATA SOURCE)

OLAP (ON-LINE ANALYTICAL PROGRESSING)

PK (PRIMARY KEY)

TI (Information technology)

SBS (SUPERINTENDENCIA DE BANCOS Y SEGUROS)

SCD (SLOWLY CHANGING DIMENSION)

SGBR (SISTEMA DE GESTOR DE BASE DE DATOS)

SSIS (SQL SERVER INTEGRATION SERVICES)

SQL (STRUCTURED QUERY LANGUAGE)

SW (SOWFTARE)