

---

**Diseño de un Data Mart para el almacenamiento histórico de información de canales, oficinas y transacciones para el área operativa de canales y oficinas del Banco Agrario de Colombia.**

---

Autores  
José Luis Espinosa Unibio

Director  
Andres Julian Aristizabal Cardona

Co-Director  
Mauricio Garcés Restrepo

Cliente  
Banco Agrario de Colombia



**Universidad de Bogotá Jorge Tadeo Lozano**  
Facultad de Ciencias Naturales e Ingeniería  
*Especialización en Desarrollo de Bases de Datos*

Bogotá - Colombia, Mayo de 2025

# Índice

	<b>Página</b>
<b>Resumen</b>	<b>IV</b>
<b>Abstract</b>	<b>V</b>
<b>1. Introducción</b>	<b>1</b>
<b>2. Descripción del Problema</b>	<b>2</b>
<b>3. Objetivos</b>	<b>7</b>
3.1. Objetivo General . . . . .	7
3.2. Objetivos Específicos . . . . .	7
<b>4. Requerimientos</b>	<b>8</b>
4.1. Requerimientos del negocio . . . . .	8
4.2. Requerimientos Funcionales . . . . .	8
4.3. Requerimientos de calidad . . . . .	8
4.4. Requerimientos de Implementación . . . . .	9
<b>5. Estado del Arte</b>	<b>10</b>
<b>6. Marco Teórico</b>	<b>13</b>
<b>7. Solución propuesta</b>	<b>20</b>
7.1. Descripción general de la solución . . . . .	20
7.2. Modelo conceptual . . . . .	21
7.3. Estándares de la solución . . . . .	25
7.4. Condiciones de diseño, propuesta de implementación y evaluación . . . . .	27
<b>8. Planeación del Trabajo</b>	<b>30</b>
8.1. Descomposición de actividades WBS . . . . .	35
8.2. Diagrama de Gantt . . . . .	36
<b>9. Presupuesto</b>	<b>37</b>
<b>10. Conclusiones</b>	<b>38</b>
<b>Bibliografía</b>	<b>40</b>

## Índice de figuras

1.	Datos de Ingresos Netos. . . . .	4
2.	Resultados del modelo. . . . .	5
3.	Matriz de confusión. . . . .	5
4.	Arquitectura de un Data Warehouse. . . . .	14
5.	Ejemplo de Modelo Estrella. . . . .	15
6.	Ejemplo de Modelo Copo de Nieve (Snowflake). . . . .	16
7.	Modelo Metodología Top-Down . . . . .	16
8.	Modelo Metodología Botton-Up . . . . .	17
9.	Arquitectura propuesta . . . . .	20
10.	Tabla Histórica de Transacciones . . . . .	21
11.	Tabla dimensión canal . . . . .	22
12.	Tabla dimensión clientes . . . . .	22
13.	Tabla dimensión oficinas . . . . .	23
14.	Tabla dimensión tipo transacción . . . . .	23
15.	Tabla dimensión producto . . . . .	24
16.	Tabla dimensión tiempo . . . . .	24
17.	Tabla dimensión municipio . . . . .	25
18.	Tabla dimensión municipio . . . . .	25
19.	Modelo estrella de alto nivel . . . . .	27
20.	Figura n: Descomposición de Actividades. . . . .	35
21.	Figura 1: Diagrama - Parte 1 . . . . .	36
22.	Figura 1: Diagrama - Parte 2 . . . . .	36

## Índice de tablas

1. Presupuesto . . . . .	37
--------------------------	----

## Resumen

---

El presente trabajo investigativo tiene como objetivo dar a conocer una propuesta de diseño de un Data Mart para los usuarios de la gerencia operativa de canales y oficinas del Banco Agrario de Colombia que les permita acceder a la información que se encuentre allí almacenada y que sirva de apoyo en la toma de decisiones.

Actualmente el Banco Agrario de Colombia cuenta con una Bodega de Datos en donde se extra la información de diferentes fuentes de datos productivas del Banco (entre bases de datos productivas e insumos de archivos planos). Este modelo lo que pretende es ser independiente al modelo actual que se tiene, para ello se implementará un espacio de trabajo independiente (Custom WorkSpace) en donde se realizará la construcción de Data Mart y al cual solo los usuarios de la gerencia puedan acceder. La bodega de datos se encuentra implementada en Teradata, que es una base de datos robusta que permite manejar grandes volúmenes de información de manera rápida y eficiente; además cuenta con la herramienta como SAP Data Service en donde se desarrollarán de procesos de ETL (Extracción, Transformación y Carga) y permitirán el almacenamiento de la información desde las fuentes primarias.

La definición y el diseño del modelo se encuentra dirigida a la información que maneja la gerencia operativa de canales y oficinas en donde se debe entregar el modelo construido que permita a los usuarios acceder al mismo y tomar los datos para realizar sus labores como reportes o tableros de visualización para la toma de decisiones.

## Abstract

---

The following research project proposes the design and implementation of a Data Mart tailored to the operational area of channels and offices at Banco Agrario de Colombia. Although the bank currently has a Data Warehouse in Teradata, it does not fully meet the information needs of this area, resulting in recurring requests to the IT department. To address this, a dedicated, independent Data Mart will be developed to provide direct access to relevant data for reporting and decision-making. The solution integrates information from production databases (such as AS400, AgroBAC, and Cobis) and external flat files using SAP Data Services for ETL processes. The Data Mart will adopt a star schema with a historical facts table and several dimension tables (e.g., customers, transactions, channels, products, offices, time, municipalities, and departments). This model follows the bottom-up approach proposed by Kimball, ensuring efficient and scalable data modeling. The project focuses on ensuring data quality, controlling user access, and optimizing performance. Ultimately, the proposed Data Mart will enhance data availability, reduce IT dependency, and strengthen self-service analytics in the operational area, enabling faster and more informed decision-making within the bank.

## 1. Introducción

Las bodegas de datos (Data Warehouse) tienen un papel fundamental en las entidades financieras, como los bancos. Tal ha sido su impacto que las empresas financieras han invertido tanto su capital humano (para ser capacitado) como su capital financiero (para la adquisición de herramientas de software y hardware) con el fin de implementar su uso ya que permiten ser una herramienta de apoyo fundamental para la toma de decisiones ya que con la información que se almacena allí permiten a estas empresas sobresalir en el mercado financiero y aumentar su productividad e ingresos monetarios.

El Banco Agrario de Colombia no ha sido ajeno a este fenómeno ya que actualmente cuenta con una Bodega de Datos. Sin embargo, hay información que actualmente no se tiene almacenada allí; para lo cual la gerencia de Inteligencia de Negocios y Analítica del Banco, en su cultura organizacional de la auto atención de solicitudes de información, ha dado vía libre a las áreas internas del banco soliciten la centralización de la información en la bodega de datos, para este caso se toma como referencia la gerencia operativa de canales y oficinas.

## 2. Descripción del Problema

Actualmente las entidades financieras, tanto bancos como entidades de cobranza y recuperación de cartera, han enfocado sus esfuerzos en la innovación de estrategias para el crecimiento de sus ingresos mediante herramientas que logren objetivos tales como la atracción de clientela a razón de la adquisición de sus productos sino también ayudan a la retención y fidelización de sus clientes. Para lo cual los almacenes de datos (o data Warehouse) como se les conoce son una herramienta que se encuentra a la vanguardia debido a que de allí se puede tomar información histórica que permite realizar análisis históricos de información que permitan resolver las preguntas que una organización financiera requiera para cumplir con los objetivos mencionados.

Otro de los enfoques que podemos darles a los conjuntos de datos es investigativo, es así como Marta Fernández-Olmos, Jorge Fleta-Asin, Talía Gómez-Aguas, Fernando Muñoz y Carlos Sáenz-Royo, [1] investigadores de la Universidad de Zaragoza (España) y la Fundación SIP presentaron una base de datos mejorada sobre asociaciones público privadas (APP's) en donde se tomó una recopilación de 10,958 proyectos realizados entre 1900 y 2021 en 130 países, incorporando 49 variables nuevas entre económicas, institucionales y de conflicto, así como los ISO's por cada país. Esta actualización permitió análisis más precisos en contexto donde se desarrollaron los proyectos facilitando estudios sobre el impacto de la gobernanza, la violencia en el éxito de las APP's; los autores concluyeron que es una herramienta esencial para el estudio y la toma de decisiones en políticas e investigación aplicada en infraestructura.

Otros autores como Pelumi Oguntunde, Hilary Okagbue, Patience Adamu, Omoleye Oguntunde, Sola Oluwatunde y Abiodun Opanuga [2] dieron buen uso a los almacenes de datos enfocados en negocios financieros, en este caso particular el conjunto de datos que desarrollaron fue utilizado para el análisis estadístico de los depósitos bancarios en seis tipos de cuenta distintos en un banco líder en Nigeria (la identidad del banco se reserva); para este caso este conjunto de datos fue llevado a ciertas herramientas estadísticas como ANOVA en donde los autores si existen diferencias significativas en cuanto a monto medio depositado en cada tipo de cuenta; los resultados al final mostraron que las cuentas corrientes (tipo 2) y las corporativas (tipo 3) reciben significativamente mayores montos que otros tipos de cuenta mientras que las cuentas de ahorro y los otros tipos de cuenta (4,5 y 6) no presentan diferencias significativas entre sí y esto sugiere un comportamiento diferencial en el uso de los servicios bancarios, posiblemente ligado al perfil de los usuarios (empresarial vs. individual).

Los investigadores Damilola Felix Eluyela, Olamide Oluwabusola Akintimehin, Wisdom Okere, Emmanuel Ozordi, Godswill Osagie Osuma, Simon Osiregbemhe Ilogho y Olufemi Adebayo Oladipo [3], afiliados a las Landmark University, Covenant University y Bells University of Technology, todas en Nigeria, en su artículo "Datasets for board meeting frequency and financial performance of Nigerian deposit money banks" realizaron un estudio en donde utilizaron datos de 14 bancos comerciales nigerianos cotizados en la bolsa de valores de Nigeria recopilados entre 2010 y 2016 para analizar la relación entre la frecuencia del consejo de administración y el desempeño financiero de los bancos, mediante el indicador Tobin Q. En el concluyeron que Existe una relación positiva entre la frecuencia de las reuniones del consejo y el rendimiento financiero de los bancos, lo cual apoya la teoría de agencia: más reuniones permiten una mejor

supervisión, asesoramiento y control interno; Se destaca que los datos son útiles para futuras investigaciones sobre gobernanza corporativa y otras características del consejo directivo y el estudio proporciona una base empírica importante para responsables de políticas, analistas financieros y académicos interesados en la efectividad de la gobernanza bancaria en economías emergentes como la de Nigeria.

En ese mismo sentido, los investigadores Omáš Pokorný, Adam Ruschka, y Helena Chytilová [4] de la Universidad de Economía y negocios de Praga (República Checa) realizaron un estudio tomando en cuenta un conjunto de datos sobre la independencia legal de los bancos centrales de 21 países de la OCDE (Organización para la Cooperación y el Desarrollo Económico), excluyendo a los países de la Eurozona. El periodo cubierto es de 2010, 2015 y 2020; este estudio propone un índice que evalúa cuatro dimensiones clave: Independencia personal, funcional e institucional, financiera y presupuestaria; combinan legislación oficial de bancos centrales con base de datos del FMI (Fondo Monetario Internacional); se aplican dos enfoques de ponderación para construir el índice: Pesos iguales y pesos expertos y la independencia se mide con una escala normalizada entre 0 y 1. Al finalizar este estudio se concluyó lo siguiente:

1. Chile y Suecia presentan los niveles más altos de independencia legal en los tres años analizados.
2. Corea del Sur y Nueva Zelanda tienen niveles comparativamente bajos.
3. La independencia Financiera es generalmente la más débil entre las dimensiones analizadas, debido a provisiones legales que permiten préstamos de gobierno.
4. Turquía mostró una caída significativa en independencia entre 2015 y 2020, mientras que Costa Rica e Islandia aumentaron su autonomía.
5. Los autores recomiendan enfocar los esfuerzos legislativos en fortalecer la independencia funcional e institucional, por su impacto crítico en la autonomía efectiva de los bancos centrales.

Otros autores como Arfan Wiraguna, Rofikoh Rokhim, Buddi Wibowo, Roy Sembel [5] investigadores de la Universitas Indonesia, Universitas Prasteliya Mulya e IPMI International Business School de Indonesia, en su estudio muestran un nuevo conjunto de datos que analiza el efecto directo e indirecto de la securitización de préstamos para micro, pequeñas y medianas empresas (MSMEs) sobre la estabilidad de los bancos europeos. La muestra incluye 440 operaciones de securitización realizadas por 48 bancos originadores en 8 países europeos entre 2004 y 2022. En este contenido de datos se toman los siguientes aspectos: Medidas de estabilidad bancaria; Variables mediadoras como el costo de financiamiento, capital regulatorio, riesgo y liquidez; datos de dimensiones de securitización de préstamos MSMEs y Variables de control macroeconómicas e institucionales.

En conclusión, este estudio arrojó los siguientes datos:

1. La securitización de préstamos MSME puede influir en la estabilidad de los bancos, tanto de forma directa como a través de variables intermedias como el riesgo, la liquidez y el

costo de financiación.

2. El Z-score fue utilizado como medida principal de estabilidad bancaria. Un Z-score más bajo indica mayor probabilidad de insolvencia.

3. También se usaron ratios de capital, liquidez y préstamos morosos como indicadores alternativos de estabilidad.

4. Los resultados apoyan la importancia de monitorear la securitización como herramienta de financiamiento alternativa, que puede tener efectos tanto beneficiosos como riesgosos sobre la banca.

5. El dataset proporciona una base valiosa para futuros estudios empíricos sobre mercados financieros y regulación bancaria en Europa.

Para reforzar esta idea se toma una investigación realizada por Prashant Gala y Abishek Balakumar [6] quienes afirman que en un estudio realizado que el sector bancario para 2022, el capital mínimo de Nivel 1 para un banco en el Top 1000 ahora se sitúa en (signopesos)556 millones, un aumento desde \$547 millones el año pasado. Además, los activos agregados de Nivel 1 han alcanzado su nivel más alto de la historia, creciendo un 3.8% a \$154.21 billones.

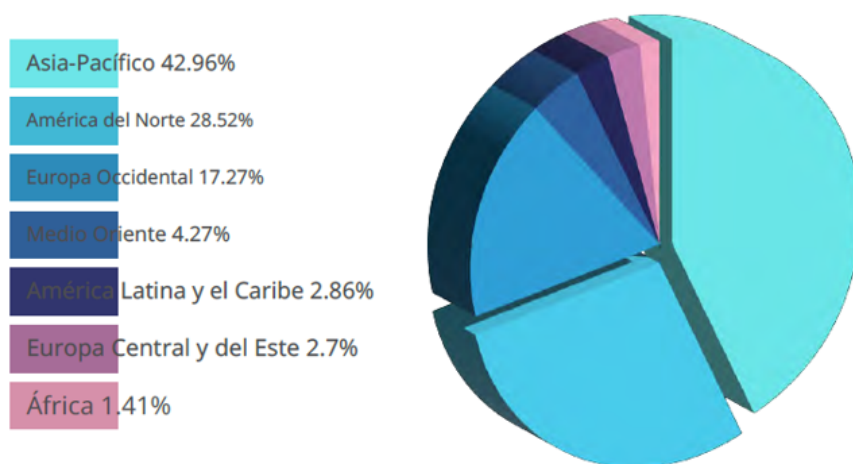


Figura 1: Datos de Ingresos Netos.

Los bancos chinos ahora constituyen el 65% del capital agregado de Nivel 1 en la región de Asia-Pacífico, un aumento desde el 62% el año pasado. Por otro lado, la participación de los bancos norteamericanos en los ingresos netos ha aumentado del 23.9% al 28.5%, mientras que la participación de Europa Occidental ha aumentado del 10.3% al 17.3%. Además, otros países han tomado la iniciativa de introducir características novedosas destinadas a atraer nueva clientela mientras retienen su base de clientes existente.

Otro de los enfoques que se puede dar en el uso, incluso de los Data Warehouse en las entidades financieras va enfocado hacia la seguridad, más específicamente si se aplica para la detección de irregularidades financieras. En este ámbito queremos citar a Qian Du y a Jieping Zai [7] quienes mediante el desarrollo de un almacén de datos pudieron aplicar un algoritmo de bosque de datos el cual permita ser más preciso en el reconocimiento de fraudes financieros. En los siguientes gráficos se evidencian los resultados de dicha operación:

**Resultados de la identificación de modelos de bosque aleatorio.**

Opciones	0	1	Precisión	Media ponderada
Precisión	0.94	0.79	-	0.88
Tasa de recuperación	0.89	0.88	-	0.89
Puntuación FL	0.92	0.82	0.89	0.87
Índice de aprobación	711	331	1033	1033

Figura 2: Resultados del modelo.

**Tabla de matriz de confusión de bosque aleatorio.**

Opciones	El modelo predice si se produce fraude financiero		Precisión de la predicción
	0	1	
Si se ha producido realmente un fraude financiero en la muestra	0 735	80	0.90
	1 121	352	0.74
La precisión real de la predicción es			0.8849

Figura 3: Matriz de confusión.

En la actualidad, el Banco Agrario de Colombia cuenta con una Bodega de Datos que almacena información de sus sistemas productivos con periodicidad diaria, semanal y mensual. Esta información es utilizada tanto por las áreas de ingeniería de datos y analítica como por diversos usuarios de áreas administrativas y gerenciales, quienes tienen acceso controlado para realizar extracciones de información según sus necesidades operativas.

Sin embargo, se ha identificado que no toda la información requerida por los usuarios se encuentra almacenada en la Bodega de Datos. Como parte de la estrategia de la Gerencia de Inteligencia de Negocios y Analítica, se ha promovido la cultura de auto atención en la consulta de datos, brindando acceso controlado a los usuarios para que puedan realizar sus propias consultas de información.

A pesar de esta iniciativa, algunas áreas del banco continúan escalando solicitudes de información al área de Tecnología (IT) debido a la ausencia de ciertos datos en la Bodega. Un ejemplo representativo es el caso de la Gerencia Operativa de Canales y Oficinas, que aún depende de IT para obtener información específica. Este flujo de solicitudes ha generado una alta acumulación de trabajo pendiente en el área de IT, afectando los tiempos de respuesta y la eficiencia operativa.

Dado este contexto, surge la necesidad de analizar y proponer estrategias que optimicen la gestión y disponibilidad de la información en la Bodega de Datos, con el objetivo de mejorar la auto atención de los usuarios y reducir la carga operativa sobre el área de IT.

### **3. Objetivos**

#### **3.1. Objetivo General**

Entregar la propuesta de diseño y desarrollo de un DATA MART orientado al área de operativa de canales y oficinas del Banco Agrario de Colombia en el cual los usuarios de dicha área puedan consultar la información de los canales transaccionales del Banco y poder realizar reportes o tableros de visualización analíticos y preventivos que permitan ser el apoyo para la toma de decisiones en cuanto al manejo de los canales transaccionales del Banco (Oficinas, cajeros y transacciones).

#### **3.2. Objetivos Específicos**

1. Realizar sesiones de entendimiento con el usuario con el propósito de conocer la información que dicha área maneja actualmente.
2. Conocer las fuentes primarias de donde los usuarios del área operativa de canales y oficinas obtienen la información para realizar sus tareas de reportes o tableros de visualización.
3. Definir la metodología que se utilizará para los procesos de extracción y carga de la información al data mart teniendo como base a las fuentes primarias.
4. Diseñar el modelo del data mart para el cargue de la información.
5. Construir el diseño del modelo en la Bodega de Datos cumpliendo los requerimientos del área operativa de canales y oficinas.

## **4. Requerimientos**

### **4.1. Requerimientos del negocio**

La gerencia de inteligencia de negocios y analítica en la búsqueda de implementar en todas las áreas del Banco Agrario de Colombia la cultura de la auto atención, en el caso del área operativa de canales y oficinas se propone implementar un modelo de datos bajo los siguientes propósitos:

1. Custodia de la información correspondiente al área.
2. Disponibilidad de la información.
3. Agilidad en los procesos del área usuaria.
4. Reducir las solicitudes de información a demanda al área de IT

### **4.2. Requerimientos Funcionales**

De acuerdo con las reuniones sostenidas con el área operativa y de canales y teniendo en cuenta tanto el levantamiento de información como las historias de usuario es generar un modelo que permita responder las siguientes preguntas:

1. Número de transacciones en los diferentes canales del Banco Agrario.
2. Disponibilidad de canales físicos (Oficinas, cajeros automáticos, corresponsales bancarios).
3. Número de transacciones con las diferentes franquicias.
4. Información histórica de clientes activos, inactivos e histórico de servicios contratados en el Banco Agrario (Cuentas corrientes, cuentas de ahorro, solicitudes de crédito, depósitos judiciales, etc.).

### **4.3. Requerimientos de calidad**

Para llevar a cabo y con éxito la construcción del modelo de Data Mart que se implementará para el área operativa de canales y oficinas se debe tener en cuenta los siguientes puntos con base a cumplir la calidad de los datos:

1. Consistencia y coherencia de los datos desde la fuente hasta su paso al modelo.
2. Integridad de la información.
3. Identificación y corrección de registros erróneos.

#### **4.4. Requerimientos de Implementación**

El modelo a construir para el área operativa de canales y oficinas, se implementará dentro de la bodega de datos que actualmente maneja el Banco Agrario de Colombia en Teradata; sin embargo, este modelo será totalmente independiente al modelo que tiene actualmente implementado. Los procesos de Extracción y Carga de información se desarrollarán desde SAP DATA SERVICES ya que es la herramienta integradora que actualmente maneja el Banco Agrario de Colombia, su mayor ventaja es que permite realizar el procesamiento de cargue de grandes volúmenes de información en un tiempo óptimo.

## 5. Estado del Arte

En el mundo actual, las Bodegas de Datos (Data Warehouse) han sido una herramienta fundamental para organizaciones y empresas financieras, ya que no solo permiten el almacenamiento de grandes volúmenes de información, sino que también facilitan su uso con fines analíticos, prescriptivos y estratégicos. Estas capacidades brindan una ventaja competitiva al permitir proyecciones futuras y respaldar la toma de decisiones basada en datos.

Diversos autores han definido el concepto de Bodega de Datos. En particular, Néstor Darío Duque Méndez [8] la describe como “un conjunto de datos orientados a una materia que varían con el tiempo y no son transitorios, los cuales soportan el proceso de toma de decisiones en una organización”.

Casos de Éxito en la Implementación de Data Warehouse La eficacia de los Data Warehouse ha sido demostrada en múltiples implementaciones exitosas dentro del sector financiero. Uno de estos casos es el estudio de Yolanda Cecilia Mancha Paredes [9], quien desarrolló un Data Warehouse para la generación de campañas en una institución financiera. Su investigación se centró en la necesidad de automatizar los procesos de generación de campañas con el objetivo de captar clientes y aumentar la cartera de la entidad. La solución implementada permitió mantener la información centralizada y actualizada, logrando que los analistas de campañas dispusieran de mayor tiempo para evaluar estrategias competitivas y mejorar la toma de decisiones.

Otro caso relevante es el de Nicole Marie Jhonson Orihuela [10], quien en 2021 presentó una solución de Data Warehouse enfocada en optimizar la productividad de una entidad financiera en Arequipa, Perú. Su estudio abordó el problema de falta de comunicación vertical y horizontal dentro del área comercial de la empresa, debido a la ausencia de una herramienta que brindara información rápida y confiable para el seguimiento de metas y productividad. A través de un proceso de investigación en campo, se implementó un Data Warehouse que mejoró significativamente la disponibilidad y accesibilidad de la información para los equipos comerciales.

En otras líneas de investigación en las que se utilizaron dataset o bodegas de datos tiene que ver con el estudio realizado por Petr Jakubik y Bogdan Gabriel Moinescu [11] pertenecientes a la Charles University en Praga (República Checa) y al National Bank of Romania y Bucharest University of Economic Studies (Rumanía); en donde se utilizó con un enfoque micriprudencial personalizado para la retención de beneficios en bancos, basado en un marco de tolerancia al riesgo. Utiliza modelos de predicción inversa de fallas bancarias sobre un conjunto de datos de aproximadamente 700 bancos comerciales europeos (2012–2022), incluyendo 30 instituciones fallidas. El estudio incorpora indicadores macrofinancieros, variables específicas de bancos y buffers de capital para diseñar recomendaciones sobre políticas de dividendos alineadas con el perfil de riesgo de cada banco. Se sugiere que los bancos con mayor probabilidad de insolvencia deben retener utilidades, mientras que aquellos con perfiles más sólidos pueden distribuir dividendos. Este enfoque dinámico busca mejorar la estabilidad financiera y la eficiencia regulatoria. A las conclusiones que se llegó en este estudio fueron las siguientes:

1. Los modelos propuestos permiten alinear las expectativas prudenciales con la tolerancia al riesgo del supervisor.
2. Se fomenta una planificación de capital más flexible, basada en condiciones macroeconómicas.
3. El estudio representa una innovación en Europa al usar listas oficiales de bancos fallidos (EBA) para desarrollar sistemas de alerta temprana más efectivos.

Otro ejemplo encontrado se encuentra en Suiza. En este caso el investigador Ugo Panizza [12] del Geneva Graduate Institute (Y) CEPR de Suiza, realizó un almacén de datos el cual contó con un conjunto de datos único que cubre la propiedad de más de 6.500 bancos en 181 países durante el período 1995–2020. Este conjunto de datos permitió el análisis revela una reducción en la propiedad estatal y un aumento en la propiedad extranjera hasta la crisis financiera global, que interrumpió estas tendencias. A nivel país, no se encuentra una relación robusta entre la propiedad bancaria y el crecimiento del PIB o la profundidad financiera. Sin embargo, a nivel bancario, se observa que los bancos estatales son menos rentables y tienen una mayor proporción de préstamos en mora que los bancos privados o extranjeros. Además, los bancos estatales tienden a estabilizar el crédito ante choques domésticos, mientras que los bancos extranjeros amplifican choques externos. Este estudio arrojó las siguientes conclusiones:

1. La propiedad estatal está asociada a menor rentabilidad y mayor morosidad.
2. Los bancos estatales estabilizan el crédito ante choques internos.
3. Los bancos extranjeros no estabilizan choques internos, pero amplifican choques externos.
4. No hay relación sistemática entre propiedad bancaria y crecimiento económico a nivel país.

Otro de los estudios en donde los data Warehouse fueron fundamentales para los negocios financieros fue el efectuado por los investigadores Douglas A. Adu, Mohammad Zoynul Abedin, Vida Y. Saa y Frank Boateng [13] en donde emplearon un panel de datos bancarios con 2,785 observaciones banco-año de 220 bancos en 16 países del África Subsahariana (SSA), recopilados durante el periodo 2007–2022; los datos que tomaron de para este estudio fueron informes anuales de los bancos, bases de datos como BankScope y datos del Banco Mundial y del Fondo Monetario Internacional. Al finalizar el estudio, obtuvieron las siguientes conclusiones:

1. BSRF y BSCOM aumentan significativamente las iniciativas climáticas bancarias (BC-CI).
2. El índice de gobernanza (CORPGOVDISCIN) modera positivamente estas relaciones: cuanto mejor la gobernanza, más fuerte el impacto.
3. BSCOM mejora directamente el desempeño financiero (FP), mientras que BSRF no muestra efecto significativo por sí solo.

4. 4.Divulgar las iniciativas climáticas mejora el rendimiento financiero (ROA, ROE, NIM); sin embargo, invertir dinero real en ellas (BICCI) no muestra impacto a corto plazo.
5. El estudio demuestra que estas relaciones varían con el tiempo, especialmente en contextos de reformas, crisis financieras y COVID-19.

Contexto en el Banco Agrario de Colombia.

El Banco Agrario de Colombia cuenta con una Bodega de Datos desde el año 2010, enfocada en su modelo de negocio financiero. No obstante, como se mencionó en la definición del problema, el área operativa de canales y oficinas no tiene acceso a toda la información que requiere dentro de la Bodega de Datos. Esta situación los obliga a escalar solicitudes al área de IT, generando un incremento en el volumen de solicitudes y provocando que el equipo de desarrollo de aplicaciones del banco asuma un rol operativo adicional.

Ante este panorama, se hace necesario el desarrollo de un Data Mart dentro de la Bodega de Datos del Banco Agrario de Colombia. Esta solución permitirá que los usuarios del área operativa de canales y oficinas accedan directamente a la información relevante para sus procesos. A través del Data Mart, podrán realizar consultas, generar reportes y desarrollar tableros de visualización (BI) para la toma de decisiones estratégicas. Esto contribuirá al mejoramiento del servicio del Banco Agrario de Colombia, optimizando la gestión de información y reduciendo la dependencia del área de IT.

## 6. Marco Teórico

### Bodegas de Datos (Data Warehouse)

De acuerdo con Alejandro Vaisman y Esteban Zimanyi [14] una bodega de datos, o Data Warehouse como se le conoce en inglés, es una base de datos centralizada que permite el almacenamiento de grandes cantidades de datos históricos y no volátiles de múltiples fuentes para facilitar su análisis. Comúnmente los Data Warehouse son diseñados y construidos en diferentes empresas con el objetivo de realizar consultas que permitan responder a ciertas preguntas que requiera el negocio. Esta información que se almacena allí puede provenir de una amplia gama de fuentes de datos, desde Bases de Datos productivas de sistemas de información o aplicaciones, también de insumos de texto.

### Características de un Data Warehouse

Según Immon y Kimball [15] estas son las características más sobresalientes de un Data Warehouse:

1. Es orientado: Las bodegas pueden ser orientadas de acuerdo a las preguntas que se requieran responder sobre el negocio como ventas, clientes, productos.
2. Es Integrado: Porque combina datos de diferentes fuentes de datos con coherencia.
3. No Volátil: Los datos no cambia una vez cargados, solo se agregan nuevos.
4. Variante en el tiempo: Permite un análisis histórico ya sea por fecha o por periodo.

### Arquitectura de un Data Warehouse

Según Joe Reis y Matt Housley [16] indican que hay dos tipos de arquitectura que se destacan para los almacenes de datos: La Organizativa y la Técnica. En la primera, la arquitectura organizativa del almacén de datos organiza los datos a determinadas estructuras y procesos de los equipos especiales. Para la arquitectura Técnica refleja la naturaleza técnica del mismo, como es MPP (Procesamiento Masivo Paralelo). Es decir, una empresa puede tener un almacén de datos sin un MPP o utilizar un sistema MPP que no esté organizado como un almacén de datos. Estas arquitecturas tienen dos características principales:

1. Separa el procesamiento analítico en línea (Online Analytical Processing – OLAP) de las bases de datos de producción (Procesamiento de transacción en línea): Esta separación es fundamental porque permite trasladar los datos a un sistema físico independiente, aleja la carga de los sistemas de producción y mejora el rendimiento de los procesos analíticos.
2. Centraliza y organiza los datos: Los almacenes de datos extraen los datos de las bases de datos productivas mediante una aplicación ETL (Extracción, Transformación y Carga). Esta aplicación permite en primer nivel la extracción de los datos desde las fuentes primarias; en el segundo nivel de transformación, se realiza el proceso de estandarización y limpieza de los datos, organizándolos y poniendo la lógica del negocio en un formato

altamente moderado y en la fase de Carga introduce los datos a la base de datos del almacén de datos destino. Los datos se cargan en múltiples mercados de datos que sirven a las necesidades analíticas de líneas de negocio y departamentos específicos.

Otra arquitectura que se puede destacar es la planteada por Jim Holdsworth, Matthew Kosinski [17] en donde lo dividen en tres niveles que son:

### Nivel Inferior

En este nivel donde los datos son extraídos desde la fuente de datos dispar hacia un servidor de base de datos donde se almacenan. Los datos se mueven mediante procesos de ETL o en ocasiones se mueven mediante procesos ELT (Extraer, Cargar y Transformar). Estos procesos se ejecutan de maneras distintas, pero ambos utilizan la automatización para trasladar los datos a un almacén y prepararlos para uso en análisis. En general, las empresas utilizan en gran medida los procesos ETL ya que si se estructura bien debe ser eficiente y continuo.

### Nivel Medio

En este nivel se construye en torno a un modelo de sistema analítico en línea (OLAP) el cual ofrece análisis rápidos y velocidades de consulta; existen actualmente 3 tipos de modelos OLAP: El ROLAP (Procesamiento analítico relacional en línea) que permite el análisis de datos multidimensional de bases de datos relacionales. El MOLAP (Procesamiento Analítico Multidimensional en línea) que utiliza motores de almacenamiento multidimensionales basados en matrices para crear vistas multidimensionales de los datos. Y el HOLAP (Procesamiento analítico Híbrido en línea) que combina que combina las capacidades de los dos primeros sistemas.

### Nivel Superior

Se encuentra respaldado por una interfaz de usuario front-end. En este nivel intervienen las diferentes herramientas de visualización gráficas en donde los usuarios pueden plasmas sus tableros de visualización o también las herramientas de reportería, que permiten la generación de informes basados en datos históricos.

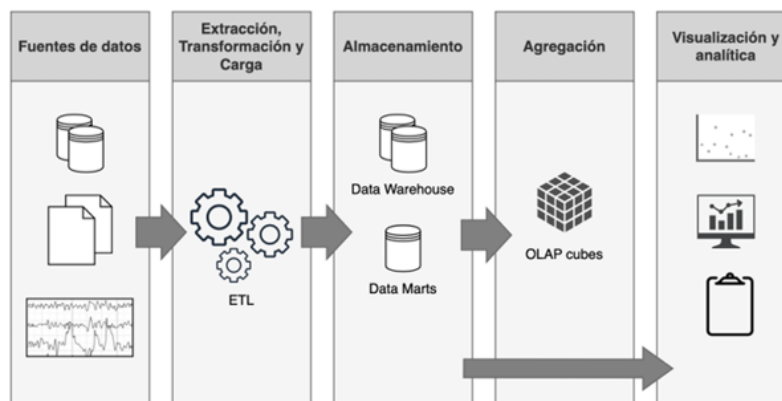


Figura 4: Arquitectura de un Data Warehouse.

Diseño de un Data Mart para el almacenamiento histórico de información de canales, oficinas y transacciones para el área operativa de canales y oficinas del Banco Agrario de Colombia

## Esquemas de Bodega de Datos

Cuando se habla de esquemas de Bodegas de datos, entendemos que es la forma en cómo se deben organizar los datos de acuerdo a la necesidad y al análisis previo antes de su construcción. En la actualidad hay dos tipos de estructuras que se utilizan en los almacenes de datos:

### Modelo Estrella

En este modelo podemos encontrar en la estructura que cuenta con una tabla principal a la cual la llamamos la tala de hechos, esta tabla contiene la información histórica medible según el modelo que se quiera construir (información de ventas, artículos que más se venden, tiendas que registran mayores, menores ventas, etc.); esta tabla se encuentra unida a otra serie de tablas que llamamos dimensiones o tablas maestras en donde se encuentra la información de referencia que ayuda como se pueden agrupar y organizar los hechos.

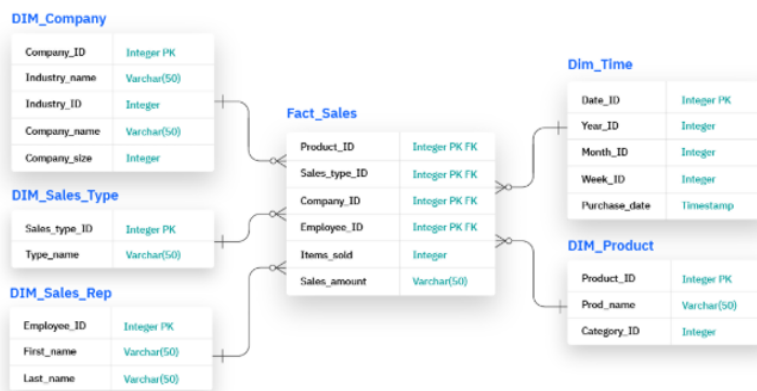


Figura 5: Ejemplo de Modelo Estrella.

### Modelo Copo de nieve (Snowflake)

Es un modelo algo similar al modelo estrella. Sin embargo, en este podemos evidenciar que la tabla de hechos (histórica) se encuentra conectada a varias tablas de dimensión y, de acuerdo al análisis realizado, las tablas de dimensión pueden estar unidas a otras tablas de dimensión ya normalizadas que contienen datos descriptivos. Debido a que este modelo es más normalizado, puede ocasionar un costo en el rendimiento de las consultas más lento.

Diseño de un Data Mart para el almacenamiento histórico de información de canales, oficinas y transacciones para el área operativa de canales y oficinas del Banco Agrario de Colombia

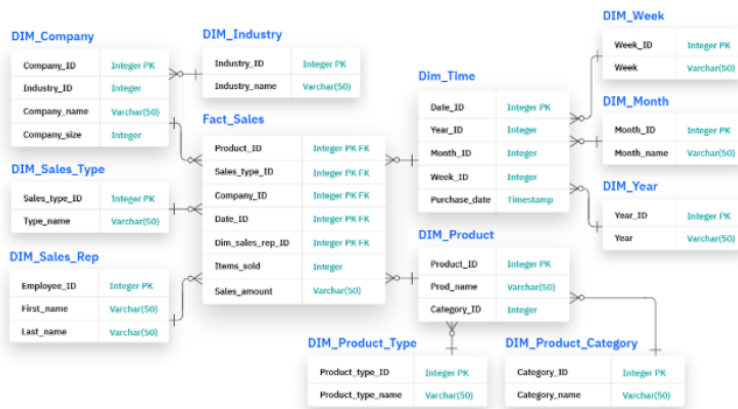


Figura 6: Ejemplo de Modelo Copo de Nieve (Snowflake).

### Modelos o enfoques de desarrollo de un sistema de Data Warehouse

Los modelos de sistemas de Data Warehouse, según Luis Joyanes Aguilar [18] son metodologías o enfoques para el desarrollo e implementación de una bodega de datos que dependiendo el desarrollador se elegirá uno u otro. En este caso, se tienen dos enfoques:

#### Metodología de desarrollo Descendente (Top-Down):

Conocido comúnmente como Modelo Immon o Modelo EDW. Esta la metodología atribuida a Bill Immon, consiste en adaptar herramientas tradicionales a las necesidades del desarrollo del Data Warehouse. Es comúnmente conocido como modelo EDW; Utiliza la metodología de desarrollo de bases de datos en las empresas y herramientas como el modelo entidad relación y una adaptación del modelo de desarrollo en espiral. En este enfoque no impide la creación de Data Mart y es ideal para enfoque corporativo.



Figura 7: Modelo Metodología Top-Down

## Metodología de desarrollo Ascendente (Botton-Up)

Es conocida como modelo Kimball. Este enfoque emplea un modelo dimensional conocido como enfoque de Data Mart. Esta es una versión descendente de un Data Warehouse, ya que se centra en las peticiones y necesidades de un departamento específico en una organización como ventas, finanzas o contabilidad. Este modelo aplica al modelo de datos dimensionados que comienza con tablas y una metodología que implica un enfoque descendente que implica la construcción de un Data Mart en cada momento que sea necesario.

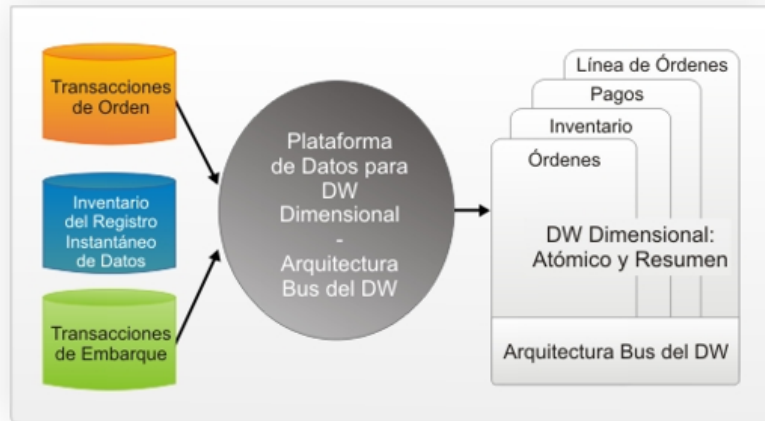


Figura 8: Modelo Metodología Botton-Up

## Data Mart

Un Data mart es, de acuerdo con [19] es un subconjunto de almacenes de datos cuyo objetivo principal es dar respuestas a una línea de negocio, departamento o área en particular. lo que les permite acceder rápidamente a información crítica sin perder tiempo buscando en todo un almacén de datos. Por ejemplo, muchas empresas pueden tener un Data Mart que se alinea con un departamento específico de la empresa, como finanzas, ventas o marketing. Normalmente, un Data Mart lo crea y gestiona el departamento empresarial específico que tiene la intención de utilizarlo. El proceso para diseñar un Data mart generalmente comprende los siguientes pasos:

1. Documentar los requisitos esenciales para comprender las necesidades empresariales y técnicas del Data mart.
2. Identificar las fuentes de datos en las que se basará su data mart para obtener información.
3. Determinar el subconjunto de datos, ya sea toda la información de un tema o campos específicos a un nivel más granular.
4. Crear el diseño lógico para el Data Mart eligiendo un esquema que se correlacione con el almacén de datos más grande.

Una vez hecho el trabajo preliminar, puede obtener el máximo valor de un Data Mart mediante el uso de herramientas de inteligencia empresarial especializadas, como Qlik o SiSense.

## **Beneficios de un Data Mart**

Dentro de los beneficios que se pueden tener de un Data Mart se encuentran los siguientes:

1. **Rentabilidad:** Hay muchos factores que considerar implementar un Data Mart, como el alcance, las integraciones y el proceso para extraer, transformar y cargar (ETL). Sin embargo, un Data Mart normalmente solo incurre en una fracción del costo de un almacén de datos.
2. **Acceso a datos simplificados:** Los Data Marts solo contienen un pequeño subconjunto de datos, por lo que los usuarios pueden extraer rápidamente los datos que necesitan con menos esfuerzo del que les tomaría al trabajar con un conjunto de datos más amplio de un almacén de datos.
3. **Acceso más rápido a insights:** La intuición obtenida de un almacén de datos respalda la toma de decisiones estratégicas a nivel empresarial, lo que afecta a toda la empresa. Un Data Mart impulsa la inteligencia empresarial y la analítica que influyen las decisiones a nivel de departamento. Los equipos pueden aprovechar los insights de datos enfocados con sus objetivos específicos en mente. A medida que los equipos identifican y extraen datos valiosos en menor tiempo, la empresa se beneficia de procesos de negocios acelerados y una mayor productividad.
4. **Mantenimiento de datos más sencillo:** Un almacén de datos contiene una gran cantidad de información empresarial, con capacidad para múltiples líneas de negocio. Los Data Marts se centran en una sola línea, con menos de 100 GB, lo que reduce el desorden y facilita el mantenimiento.
5. **Implementación más fácil y rápida:** Un almacén de datos implica un tiempo de implementación significativo, especialmente en una gran empresa, ya que recopila datos de una gran cantidad de fuentes internas y externas. Por otro lado, solo necesita un pequeño subconjunto de datos para configurar un Data Mart, por lo que la implementación tiende a ser más eficiente y más rápida.

## **Tipos de Data Marts**

Hay tres tipos de data Marts que se diferencian según su relación con el almacén de datos y las respectivas fuentes de datos de cada sistema.

### **Data Marts dependientes:**

Son segmentos particionados dentro de un almacén de datos empresarial. Este abordaje de arriba hacia abajo comienza con el almacenamiento de todos los datos empresariales en una ubicación central. Los Data Marts recién creados extraen un subconjunto definido de los datos primarios siempre que sea necesario para analizarlos.

**Data Marts independientes:**

Actúan como un sistema autónomo que no depende de un almacén de datos. Los analistas pueden extraer datos de un tema o proceso de negocios en particular de fuentes de datos internas o externas, procesarlos y luego almacenarlos en un repositorio de data mart hasta que el equipo los necesite.

**Data Marts híbridos:**

Combinan datos de almacenes de datos existentes y otras fuentes operativas. Este abordaje unificado aprovecha la velocidad y la interfaz intuitiva de un abordaje de arriba hacia abajo, y también ofrece la integración del método independiente a nivel empresarial.

## 7. Solución propuesta

### 7.1. Descripción general de la solución

a solución que se plantea es construir un modelo de Data Mart independiente que permita a los funcionarios del área operativa de canales y oficinas que permita, además de tener la custodia de su información, la disponibilidad de la misma para realizar sus actividades de reportería y elaboración de tableros de Visualización que permitan responder las preguntas concernientes al área. Para ello se tiene planteada la siguiente arquitectura: ??.

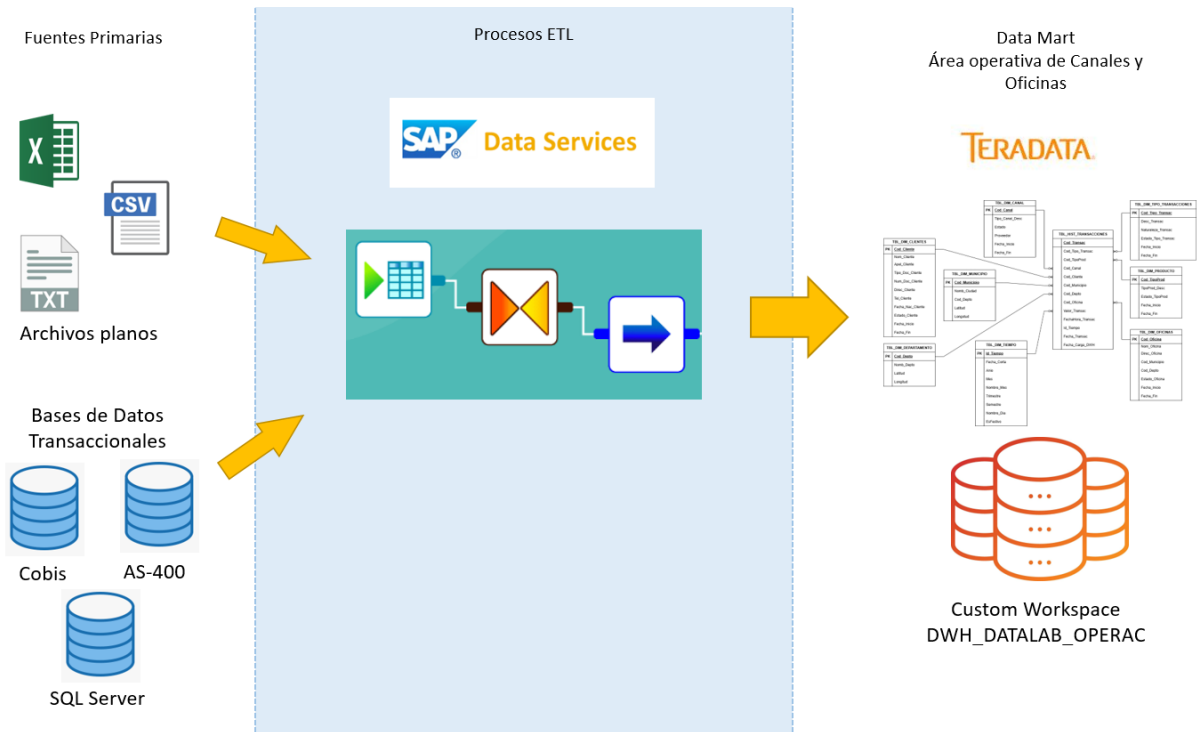


Figura 9: Arquitectura propuesta

El Banco Agrario de Colombia actualmente cuenta con sus Bases de Datos productivas (como AS400, Cobis, AgroBAC (SQL Server), entre otras.) en donde actualmente se encuentra la información inherente al negocio. Asimismo, también cuenta con su Bodega de Datos en donde actualmente se encuentra cierta parte de la información histórica de las fuentes primarias la cual permite dar respuesta a ciertas preguntas del negocio para las áreas internas del banco, como para desarrollo de algoritmos analíticos y envío de reportes de información a entidades gubernamentales. Adicionalmente, el área operativa de canales y oficinas maneja información que llega a partir de colaboradores tercerizados como proveedores de cajeros automáticos, informes de transaccionalidad por franquicias e información de transacciones virtuales. Esta información es recepcionada mediante archivos planos como archivos en Excel y archivos de texto (.txt y .csv).

De las fuentes primarias anteriormente descritas, se realizarán los procesos de extracción de la información hacia la Bodega de Datos Teradata a un nuevo Data Mart creado en donde

Diseño de un Data Mart para el almacenamiento histórico de información de canales, oficinas y transacciones para el área operativa de canales y oficinas del Banco Agrario de Colombia

se alojará la información correspondiente al área operativa de canales y oficinas. Los procesos de extracción y carga se desarrollarán en la herramienta SAP Data Services Designer. Una vez se desarrollen los Jobs de extracción y carga, se procederá con su automatización para que extraigan la información de acuerdo a los periodos que se indiquen (diario, semanal o mensual).

## 7.2. Modelo conceptual

El modelo que se desarrollará es un modelo estrella en donde se contará con una tabla de hechos (histórica) donde se almacenarán las transacciones y las tablas dimensión que se describirán a continuación. Cabe destacar que el Data Mart se construirá en un Custom Workspace (Espacio de trabajo) independiente llamado DWH\_DATALAB\_OPERAC:

### Tabla de Hechos (Histórica) de transacciones

La tabla de hechos en donde se encontrará la información histórica de las transacciones efectuadas por los clientes de acuerdo a los canales, tipos de transacción, montos y estados:

DWH_DATALAB_OPERAC.TBL_HIST_TRANSACCIONES		
Nombre del Campo	Descripción	Tipo de Dato
Cod_Transac (PK)	Código de Transacción	INTEGER
Cod_Tipo_Transac	Código del tipo de Transacción	INTEGER
Cod_TipoProd	Código del tipo de Producto	INTEGER
Cod_Canal	Código del canal de transacción	INTEGER
Cod_Municipio	Código del Municipio	INTEGER
Cod_Depto	Código del Departamento	INTEGER
Cod_Cliente	Código del Cliente	INTEGER
Cod_Oficina	Código de Oficina	INTEGER
Valor_Transaccion	Costo de la transacción	Decimal (20,2)
FechaHora_Transac	Fecha de la Transacción	DATETIME (YYYY-MM-DD HH:mm:ss)
Id_Tiempo	Código de fecha de transacción	INTEGER
Fecha_Transac	Fecha corta de la Transacción	DATE (YYYY-MM-DD)
Fecha_Carga_DWH	Fecha de cargue en Data Mart	DATE (YYYY-MM-DD)

Figura 10: Tabla Histórica de Transacciones

### Tablas Dimensión

Las siguientes son las tablas dimensión a construir en el Data Mart:

Diseño de un Data Mart para el almacenamiento histórico de información de canales, oficinas y transacciones para el área operativa de canales y oficinas del Banco Agrario de Colombia

### Tabla Dimensión Canal

DWH_DATA LAB_OPERAC.TBL_DIM_CANAL		
Nombre del Campo	Descripción	Tipo de Dato
Cod_Canal (PK)	Código del canal transaccional	INTEGER
Tipo_Canal_Desc	Descripción del canal transaccional	VARCHAR(30)
Estado	Estado del canal Transaccional	VARCHAR(30)
Proveedor	Descripción del Proveedor del canal	VARCHAR(30)
Fecha_Inicio	Fecha de inserción del registro	DATE (YYYY-MM-DD)
Fecha_Fin	Fecha de actualización del estado	DATE (YYYY-MM-DD)

Figura 11: Tabla dimensión canal

### Tabla Dimensión Clientes

DWH_DATA LAB_OPERAC.TBL_DIM_CLIENTES		
Nombre del Campo	Descripción	Tipo de Dato
Cod_Cliente (PK)	Código del cliente	INTEGER
Nom_Cliente	Nombre(s) del cliente	VARCHAR (30)
Apel_Cliente	Apellido(s) del cliente	VARCHAR (30)
Tipo_Doc_Cliente	Tipo de Documento del cliente	INTEGER
Num_Doc_Cliente	Numero de documento del cliente	VARCHAR(20)
Direc_Cliente	Dirección de residencia del cliente	VARCHAR(50)
Tel_Cliente	Teléfono de contacto del cliente	VARCHAR(20)
Fecha_Nac_Cliente	Fecha de nacimiento del cliente	DATE (YYYY-MM-DD)
Estado_Cliente	Estado del cliente	VARCHAR(20)
Fecha_Inicio	Fecha de vigencia del cliente	DATE (YYYY-MM-DD)
Fecha_Fin	Fecha de inactividad del cliente	DATE (YYYY-MM-DD)

Figura 12: Tabla dimensión clientes

Diseño de un Data Mart para el almacenamiento histórico de información de canales, oficinas y transacciones para el área operativa de canales y oficinas del Banco Agrario de Colombia

### Tabla Dimensión Oficinas

DWH_DATALAB_OPERAC.TBL_DIM_OFICINAS		
Nombre del Campo	Descripción	Tipo de Dato
Cod_Oficina	Código de la oficina	INTEGER
Nom_Oficina	Nombre de la Oficina	VARCHAR (30)
Direc_Oficina	Dirección de ubicación de la oficina	VARCHAR (50)
Cod_Ciudad	Código de la ciudad de ubicación (DANE)	INTEGER
Cod_Depto	Código del Departamento de ubicación	INTEGER
Estado_Oficina	Estado de la oficina	VARCHAR (20)
Fecha_Inicio	Fecha de creación de la oficina	DATE (YYYY-MM-DD)
Fecha_Fin	Fecha de inactividad de la oficina	DATE (YYYY-MM-DD)

Figura 13: Tabla dimensión oficinas

### Tabla Dimensión Tipo Transacción

DWH_DATALAB_OPERAC.TBL_DIM_TIPO_TRANSACCION		
Nombre del Campo	Descripción	Tipo de Dato
Cod_Tipo_Transac (PK)	Código del tipo de transacción	INTEGER
Desc_Tipo_Transac	Descripción del tipo de transacción	VARCHAR (30)
Naturaleza_Transac	Descripción de la naturaleza de la transacción	VARCHAR (30)
Estado_Tipo_Transac	Descripción del estado de la transacción	VARCHAR (30)
Fecha_Inicio	Fecha de activación del tipo de transacción	DATE (YYYY-MM-DD)
Fecha_Fin	Fecha de inactivación del tipo de transacción	DATE (YYYY-MM-DD)

Figura 14: Tabla dimensión tipo transacción

Diseño de un Data Mart para el almacenamiento histórico de información de canales, oficinas y transacciones para el área operativa de canales y oficinas del Banco Agrario de Colombia

### Tabla Dimensión tipo producto

DWH_DATA LAB_OPERAC.TBL_DIM_PRODUCTO		
Nombre del Campo	Descripción	Tipo de Dato
Cod_TipoProd (PK)	Código del producto	INTEGER
TipoProd_Desc	Descripción del tipo de producto	VARCHAR (100)
Estado_TipoProd	Estado del tipo de producto	VARCHAR (30)
Fecha_Inicio	Fecha de activación del producto	DATE (YYYY-MM-DD)
Fecha_Fin	Fecha de desactivación del Producto	DATE (YYYY-MM-DD)

Figura 15: Tabla dimensión producto

### Tabla Dimensión tiempo

DWH_DATA LAB_OPERAC.TBL_DIM_TIEMPO		
Nombre del Campo	Descripción	Tipo de Dato
Id_Tiempo	Código de la fecha	INTEGER
Fecha_Corta	Fecha corta en formato (YYYY-MM-DD)	DATE (YYYY-MM-DD)
Anio	Año	INTEGER
Mes	Mes	INTEGER
Nombre_Mes	Descripción del mes	VARCHAR (20)
Trimestre	Número del trimestre de acuerdo al mes	INTEGER
Semestre	Número del semestre de acuerdo a los meses	INTEGER
Nombre_Dia	Descripción del día de acuerdo a la fecha	VARCHAR (20)
EsFestivo	Indica si el día es festivo o no	BYEINT

Figura 16: Tabla dimensión tiempo

### Tabla Dimensión municipio

DWH_DATAALAB_OPERAC.TBL_DIM_MUNICIPIO		
Nombre del Campo	Descripción	Tipo de Dato
Id_Municipio	Código del Municipio (Código DANE)	INTEGER
Nomb_Municipio	Nombre del Municipio	DATE (YYYY-MM-DD)
Latitud	Latitud del Municipio	DECIMAL (20,2)
Longitud	Longitud del Municipio	DECIMAL (20,2)

Figura 17: Tabla dimensión municipio

### Tabla Dimensión departamento

DWH_DATAALAB_OPERAC.TBL_DIM_DEPARTAMENTO		
Nombre del Campo	Descripción	Tipo de Dato
Id_Depto	Código del Departamento (Código DANE)	INTEGER
Nomb_Depto	Nombre del Departamento	DATE (YYYY-MM-DD)
Latitud	Latitud del Departamento	DECIMAL (20,2)
Longitud	Longitud del Departamento	DECIMAL (20,2)

Figura 18: Tabla dimensión municipio

## 7.3. Estándares de la solución

### Estandarización de nombres de Tablas de hechos y Dimensión

Para la tabla de hechos, se identificará de la siguiente forma:

- DWH\_DATAALAB\_OPERAC.TBL\_HIST\_TRANSACCIONES

Las tablas dimensiones se identificarán de acuerdo a la siguiente nomenclatura:

- DWH\_DATAALAB\_OPERAC.TBL\_DIM\_CLIENTES
- DWH\_DATAALAB\_OPERAC.TBL\_DIM\_CANAL
- DWH\_DATAALAB\_OPERAC.TBL\_DIM\_OFICINAS
- DWH\_DATAALAB\_OPERAC.TBL\_DIM\_TIPO\_TRANSACCION
- DWH\_DATAALAB\_OPERAC.TBL\_DIM\_PRODUCTO
- DWH\_DATAALAB\_OPERAC.TBL\_DIM\_TIEMPO

Diseño de un Data Mart para el almacenamiento histórico de información de canales, oficinas y transacciones para el área operativa de canales y oficinas del Banco Agrario de Colombia

---

- **DWH\_DATA LAB\_OPERAC.TBL\_DIM\_MUNICIPIO**
- **DWH\_DATA LAB\_OPERAC.TBL\_DIM\_DEPARTAMENTO**

Los campos de cada una de las tablas llevarán la primera letra en mayúscula y las siguientes en minúsculas sin acentos; los espacios se reemplazarán mediante guion al piso (\_).

### **Especificaciones de seguridad y control de accesos**

Teniendo en cuenta que el Data Mart será enfocado para el uso exclusivo del área operativa de canales y oficinas, ellos serán a futuro los custodios de la información que allá se almacene. Los usuarios solo contarán con acceso a consultas sobre el modelo. En caso tal de encontrar inconsistencias en la información deberán reportarlas a la sub-área de ingeniería de datos con el objetivo de informarla para dar soporte y solución. En caso tal que un área independiente desee solicitar acceso al modelo, debe gestionarlo mediante la herramienta previa autorización de la gerencia operativa de canales y oficinas y se le brindarán únicamente permisos de lectura sobre los objetos desarrollados.

### **Calidad de Datos**

Con el objetivo de asegurar una la calidad de los datos se realizarán las siguientes actividades:

#### **Reglas de validación desde la fuente**

1. Estas tienen como fin asegurar que no lleguen registros nulos en campos clave.
2. Validación de fechas, estas deben venir ya sea en formato año-mes-día (AAAAMMDD) o en formato año-mes-día Hora-Minutos-Segundos (AAAAMMDD HH:mm:ss).
3. Control de registros duplicados.
4. Rechazo de registros inválidos.

Diseño de un Data Mart para el almacenamiento histórico de información de canales, oficinas y transacciones para el área operativa de canales y oficinas del Banco Agrario de Colombia

## Estructura del Modelo Dimensional

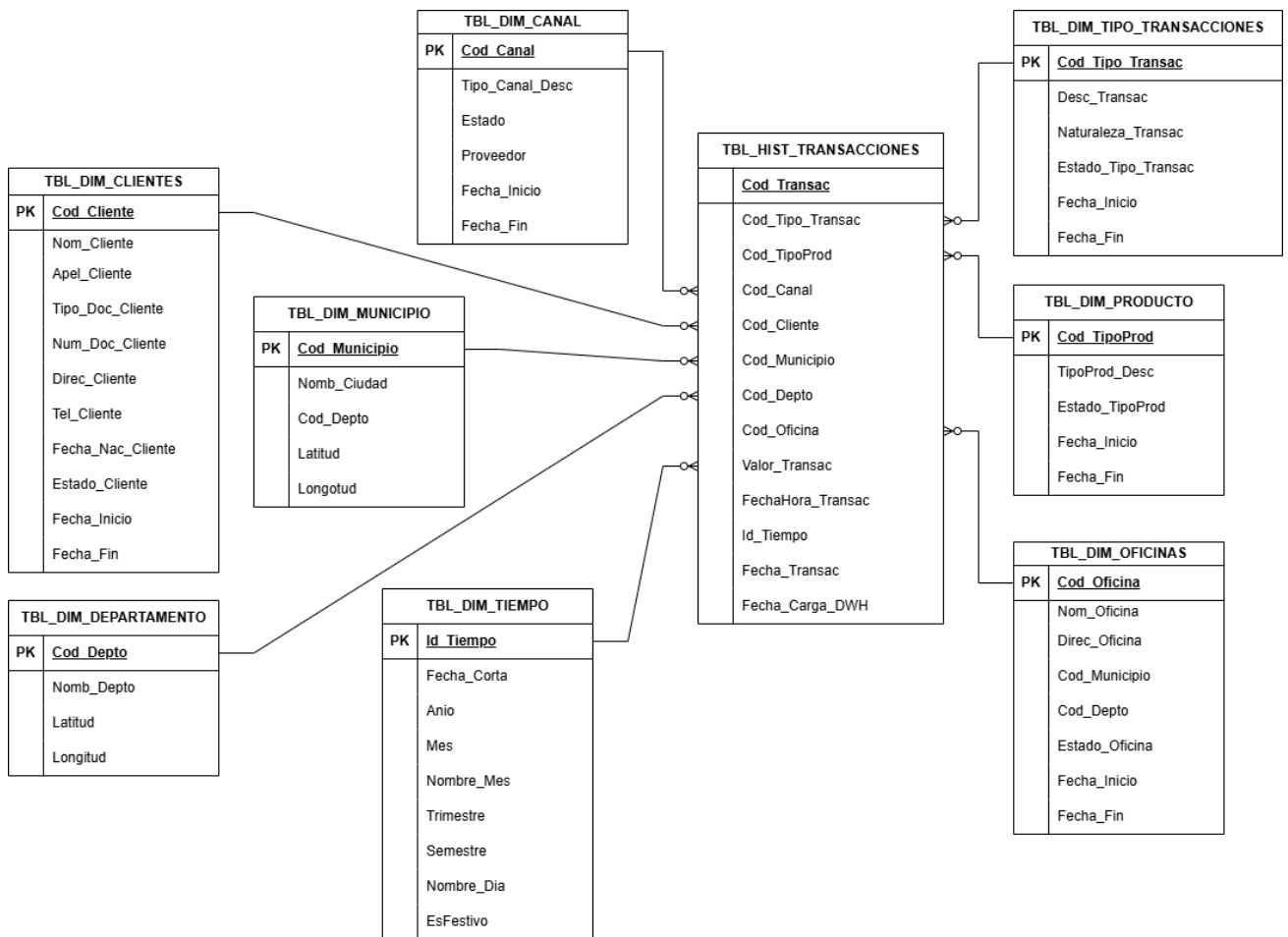


Figura 19: Modelo estrella de alto nivel

## 7.4. Condiciones de diseño, propuesta de implementación y evaluación

### Condiciones de Diseño

Las condiciones de diseño establecen los lineamientos técnicos y funcionales que guían la construcción del Data Mart:

1. Modelo de datos
  - Estructura en modelo estrella, para facilitar consultas analíticas.
  - Tabla de hechos: transacciones.
  - Tablas de dimensiones: canal, oficinas, tiempo, clientes, productos, municipios, departamentos, tiempo.
2. Requerimientos funcionales

- Capacidad de responder preguntas del negocio como:
    - Número de transacciones por canal, franquicia u oficina.
    - Disponibilidad de canales físicos y virtuales.
    - Histórico de clientes y productos contratados.
3. Fuentes de datos
- Bases de datos productivas (AS400, AgroBAC (SQL Server), Cobis, etc.).
  - Archivos planos recibidos por el área usuaria (.csv, .xlsx, .txt).
  - Otras tablas ya existentes en la Bodega de Datos Teradata.
4. Frecuencia de carga
- Procesos definidos como diarios, semanales o mensuales, según la naturaleza de los datos.
5. Herramientas
- Motor de datos: Teradata.
  - ETL: SAP Data Services Designer.
6. Seguridad y privacidad
- Acceso por roles definidos, permisos solo de lectura para usuarios analíticos.
  - Custodia de la información delegada al área operativa de canales y oficinas.

## **Propuesta de Implementación**

1. Diseño lógico y físico
- Construcción de modelo dimensional en Teradata.
  - Definición de claves primarias, índices y relaciones.
  - Establecimiento de metadata técnica y funcional.
2. Desarrollo de procesos ETL
- Extracción desde las fuentes identificadas.
  - Transformación: estandarización, validación y enriquecimiento de datos.
  - Carga al Data Mart según periodicidad definida.
3. Automatización
- Jobs ETL automatizados para garantizar ejecución periódica.
  - Programación según ventanas horarias de menor impacto operativo.
4. Gestión de calidad
- Verificación de integridad, consistencia y unicidad de datos.

- Bitácora de errores e informes de calidad.
5. Documentación
- Diccionario de datos.
  - Manual de consultas para usuarios.
  - Manual técnico para mantenimiento del modelo.
6. Capacitación
- Formación al equipo operativo sobre uso del Data Mart y consultas básicas.

## 8. Planeación del Trabajo

Para llevar a cabo el diseño y desarrollo de un Data Mart que tendrá la información del área operativa de canales y oficinas del Banco Agrario de Colombia, se tuvieron en cuenta 6 objetivos específicos. Cada objetivo se trabajará en etapas para llegar a cumplir con el objetivo principal, en donde los usuarios de dicha área puedan acceder a ella y realizar sus tareas de re-portería y desarrollo de tableros de visualización.

### 1. Realizar sesiones de entendimiento con el usuario con el propósito de conocer la información que dicha área maneja actualmente.

Las siguientes tareas se deberá desarrollar durante esta etapa:

#### 1. Sesión de entendimiento con el área de Ingeniería de Datos

En esta primera sesión se debe tener conocimiento por parte del área usuaria la necesidad de poder tener acceso a la bodega de datos del Banco Agrario de Colombia esto con el objetivo de obtener la información concerniente al área operativa de canales y oficinas.

De igual manera, se realiza un entendimiento sobre la información que actualmente maneja cada una de las sub-áreas en las que se divide la gerencia operativa de canales y oficinas.

#### 2. Definición del alcance del desarrollo.

En esta etapa se debe tener la aclaración por parte del usuario final el alcance que debe tener el proyecto, la cual es el desarrollo del Data Mart en donde los usuarios, tendrán el acceso a la información que se va a alojar allí.

Es importante tener en cuenta que los usuarios de dicha gerencia deban entender que ellos serán responsables de la información, en caso tal que cualquier gerencia ajena al área operativa de canales y oficinas requieran acceder a esta información, deban solicitar el visto bueno para solicitar su acceso ya que nuestra tarea al final será su custodia. En esta etapa se les solicitará diligenciar un formato de historias de usuario en donde definirán los requerimientos y un formato de lineamiento de información en donde deben dejar claro los campos con sus tipos de datos.

#### 3. Validación de usuarios: Para esta etapa, dentro de las sesiones se debe tener en cuenta la cantidad de usuarios que van a tener el acceso al Data Mart. Se definen los permisos que deben tener los usuarios para consultar el Data Mart.

El área usuaria se debe encargar de solicitar al área de ingeniería de datos del Banco Agrario la creación de los usuarios que deben tener acceso al Data Mart con permiso de lectura. Para ellos se les solicitará datos como el nombre del usuario, usuario de dominio (directorio activo), Vicepresidencia a la que pertenece y el tipo de permisos y esquema al que van a acceder.

#### 4. Definición de la periodicidad de la información:

Por parte del área usuaria se debe definir la periodicidad con la cual se deben realizar los procesos de extracción desde las fuentes primarias ya sea diario, semanal y mensual.

El usuario debe dar conocimiento sobre la estructura que desean ver al finalizar el desarrollo del Data Mart. Una vez certificada la información en la etapa de desarrollo se deberá procederá con el paso a producción y entrega del Data Mart desarrollado.

5. Elaboración de Historia de Usuario y FLI (Formato de Lineamiento de Información): Como parte del proceso interno del Banco Agrario de Colombia, se solicitará al usuario el diligenciamiento de dos documentos para dar inicio al desarrollo del Data Mart: La Historia de Usuario y el Formato de Lineamiento de Información.

En la Historia de Usuario se debe diligenciar la información que desde el área operativa de canales y oficinas desea llevar al Data Mart. En el formato de Lineamiento de información se deben diligenciar los campos de salida que se requieren por parte del área operativa de canales y oficinas, se deben tener en cuenta que campo deben ser calculados y las formulas a utilizar para los mismos.

## **2. Conocer las fuentes primarias de donde los usuarios del área operativa de canales y oficinas obtienen la información para realizar sus tareas de reportes o tableros de visualización.**

1. Conocimiento del flujo de proceso de solicitud de información al área de IT:

Durante las sesiones de entendimiento (primera etapa del primer objetivo específico) se solicita al usuario indicarnos el proceso por el cual ellos solicitan la información al área de Desarrollo del banco teniendo en cuenta la estructura que ellos definan y los campos a solicitar.

Este mismo proceso se realiza con las fuentes de información que no se encuentran en Bases de Datos, sino que provienen de archivos planos (Excel, .csv, .txt, etc.) se debe tener en cuenta la estructura con la que vienen dichos insumos y si los datos vienen estructurados.

2. Validación de fuentes de información primaria (Bases de Datos del BAC). Solicitar sesiones con las áreas de IT que suministran la información a los usuarios del área operativa de canales y oficinas con el objetivo de conocer los objetos (tablas) a consultar y los campos que se tienen en cuenta en los procesos de extracción de información que envían a dicha área.

En estos espacios es indispensable que el desarrollador indique de manera detallada la estructura de la información que envía a los usuarios del área usuaria y dentro de los campos a tener en cuenta que filtros se deben aplicar para que al momento de realizar la ejecución de los procesos de extracción de información sea consistente y coherente.

3. Validación de acceso a las fuentes primarias (Bases de datos productivas). Se debe solicitar al área usuaria acceso a las bases de datos productivas que harán parte del proceso de extracción de información hacia el nuevo Data Mart en las cuales desde ingeniería de datos no se tiene acceso, de igual forma se deben solicitar permisos a los objetos para realizar las respectivas consultas

Para ello se debe validar a que bases de datos se tiene acceso y que hacen parte del desarrollo, cuales bases de datos se puede acceder desde Bodega de Dato pero que no se tiene credenciales de acceso y solicitarlas y a que bases de datos no se tiene en definitivo

acceso hacia la Bodega de Datos y con ello se debe hacer una solicitud formal para generación de archivos planos.

4. Validación de fuentes de información (Archivos planos):

Se realizan sesiones con el área usuaria en donde nos muestran las fuentes primarias que son archivos planos (Excel, .csv, .txt, etc.) con el objetivo de entender su estructura y campos con los que cuenta cada insumo.

En esa validación se debe tener en claro el periodo en el cual el usuario tiene esta información a la mano para poder realizar sus procesos internos. Este proceso es importante, ya que desde ahí podemos tener el conocimiento del tiempo en el cual se debe tener disponibilidad la información para la ejecución del proceso de extracción de la información hacia el Data Mart.

5. Validación de pruebas de conexión con Fuentes de Datos y disponibilidad:

Teniendo en cuenta el punto anterior, se deben realizar pruebas de conectividad desde el programa de ETL en donde se van a desarrollar los procesos de extracción desde la fuente primaria de datos en este proceso se debe validar las credenciales entregada por el área de desarrollo para acceder a las bases de datos de pruebas y productivas.

En estas pruebas se debe tener en cuenta el tiempo de ejecución de cada uno de los procesos de extracción que se van a desarrollar y la hora en la que se deben ejecutar. Se realizarán pruebas de extracción por periodicidad diaria, semanal y mensual y validación de tiempos de extracción para así definir la hora en la que no interfiera en las operaciones diarias del Banco.

### **3. Definir la metodología que se utilizará para los procesos de extracción y carga de la información al DATA MART teniendo como base a las fuentes primarias.**

1. Evaluación de procesos del desarrollo de los procesos de extracción sobre fuentes primarias (Bases de datos productivas)

Se debe realizar una prueba de tiempos de ejecución de consultas sobre los objetos (tablas) que fueron suministradas y los campos y filtros solicitados por el área usuaria con el objetivo de validar los tiempos de ejecución.

En esta evaluación se ejecutarán las consultas en una hora específica en donde no se interfiera en la operación diaria del banco y se extraerá información por periodos diario, semanal y mensual para medir los tiempos de respuesta del servidor de base de datos.

2. Evaluación de la calidad de los datos en el proceso de extracción (Bases de datos productivas)

Se debe realizar la revisión de la calidad de los datos que vienen desde las bases de datos productivas, en ese sentido se debe validar si durante el proceso de extracción se encuentran datos que pueden presentar inconvenientes como en el caso de la traducción de caracteres alfabéticos (tildes apostrofes, etc.), en los caracteres numéricos (decimales, flotantes y enteros) y en los datos de tipo fecha.

Esta evaluación se debe realizar sobre cada uno de los objetos (tablas) campo a campo con el objetivo de que la información llegue de manera completa al Data Mart y que no se presenten fallas o pérdida de registros durante el proceso de extracción.

3. Evaluación de la calidad de los datos en el proceso de extracción (archivos planos)

Se debe realizar la revisión de la calidad de los datos que vienen desde los archivos planos, en ese sentido se debe validar si durante el proceso de extracción se encuentran datos que pueden presentar inconvenientes como en el caso de la traducción de caracteres alfabéticos (tildes apostrofes, etc.), en los caracteres numéricos (decimales, flotantes y enteros) y en los datos de tipo fecha, adicional en los separadores (si vienen con separador pipes "|" o separadores punto y coma ";").

Esta evaluación se debe realizar sobre cada uno de las estructuras de los archivos planos campo a campo con el objetivo de que la información llegue de manera completa al Data mart y que no se presenten fallas o pérdida de registros durante el proceso de extracción.

4. Definir el tipo de modelo de Data Mart a desarrollar

Este modelo se debe implementar teniendo en cuenta las fuentes primarias, desde las Bases de Datos productivas hasta archivos planos. Asimismo, se podría hablar de un modelo híbrido ya que algunas de las fuentes provienen del Data Mart que se encuentra en Bodega de Datos.

De acuerdo a esto, se debe definir si el Data Mart va a ser de un tipo dependiente, independiente o híbrido.

#### **4. Diseñar el modelo del DATA MART para el cargue de la información.**

1. Definición de las tablas dimensionales que se construirán a partir del análisis realizado.

Con base a los análisis realizados a las historias de usuario y teniendo en cuenta los campos solicitados por el usuario, se realizará la definición de las dimensiones y la tabla histórica que llevará el Data Mart a desarrollar.

Validar sobre cada una de las dimensiones cuales serán nuestros índices y llaves primarias esto con el objetivo de optimizar el modelo al momento en que los usuarios vayan a realizar consultas sobre el Data Mart y estas a su vez deben ser llevadas a la tabla histórica transaccional.

2. Realizar diseño y la documentación del mapeo Fuente destino.

Una vez realizado el análisis de las futuras dimensiones y la tabla histórica a construir, se procederá a realizar el mapeo fuente destino. En él se deben colocar las dimensiones ya definidas, los campos que van a pertenecer a cada una de ellas.

En este mapeo se debe colocar en cada uno de los campos de cada una de las dimensiones los tipos de datos que vienen definidos desde la fuente y como deben llegar en caso tal que se haga conversión.

3. Diseñar la estructura de la ETL.

Diseñar el proceso ETL teniendo en cuenta las herramientas del programa de desarrollo de ETLs. Cabe resaltar que en este proceso se debe tener en cuenta los tipos de datos que se definieron en el diseño del Data Mart.

## **5. Construir el diseño del modelo en la Bodega de Datos cumpliendo los requerimientos del área operativa de canales y oficinas.**

1. Realiza la construcción del Data Mart en la Bodega de Datos.

Se deben iniciar con la construcción de las tablas dimensión y la tabla histórica ya definidas en el diseño y mapeo fuente destino. En este proceso se debe tener en cuenta el mapeo fuente destino, los campos definidos para cada dimensión y la tabla histórica, los tipos de datos y los respectivos índices.

2. Realizar los procesos de relación de las tablas dimensión con la tabla Histórica.

Una vez construidas las tablas dimensión y la tabla histórica, se deben construir las relaciones entre las dimensiones y la tabla histórica teniendo como base el diseño construido en papel.

3. Desarrollo de los procesos ETL de las fuentes primarias (Bases de Datos productivas).

Se deben desarrollar los procesos ETLs en donde se debe tener en cuenta las tablas y campos a traer desde la fuente de datos primarias que dependen de Bases de Datos productivas. Se debe tener en cuenta el tipo de dato definido en el destino.

4. Desarrollo de los procesos ETLs de las fuentes primarias (Archivos planos).

Se deben desarrollar los procesos ETLs en donde se debe tener en cuenta la estructura, separadores y campos de los archivos planos (.Xlsx, .txt, .csv, etc.); Se debe tener en cuenta el tipo de dato definido en el destino.

5. Realizar pruebas de los Jobs de extracción.

Una vez finalizado el desarrollo de los Jobs de extracción y cargue, se deben realizar las pruebas de ejecución pertinentes, esto con el objetivo de certificar el funcionamiento de los job de extracción.

En estas pruebas se debe evaluar si los datos llegan de manera completa desde las fuentes primarias, asimismo validar campo a campo como llegan los datos de acuerdo a lo definido en diseño.

## **6. Validar y evaluar el desempeño del DATA MART desarrollado.**

1. Planificación de ejecución de procesos de extracción.

Una vez realizadas las pruebas de los procesos de extracción y cargue, se procederá con la planificación de las ejecuciones ya en ambiente productivo de las tablas dimensión y de la tabla histórica.

Estas ejecuciones se realizarán de acuerdo a la disponibilidad tanto de los las bases de datos productivas donde apuntan los procesos de extracción como de los archivos planos, asimismo validación de los tiempos de ejecución de las extracciones y los periodos que se van a tener en cuenta en las extracciones de acuerdo al volumen de información (diario, semanal y mensual).

2. Validación de cargue de las tablas dimensión.

Validación de los cargues de información en ambiente productivo sobre las tablas dimensión y la tabla histórica para asegurar el correcto funcionamiento del proceso ETL.

En este sentido se debe tener en cuenta la consistencia de los datos campo a campo en cada una de las tablas dimensiones y la tabla histórica.

3. Automatización y programación del proceso ETL.

Se debe realizar la programación de las ejecuciones de los Jobs de extracción y cargue de información desde las fuentes primarias hacia las tablas dimensión y tabla histórica.

Se deben programar de acuerdo a las horas indicadas de indisponibilidad definidas por parte del administrador de las bases de datos productivas y los tiempos de ejecución de las extracciones diarias, semanales y mensuales.

4. Documentación del modelo.

Se debe realizar la documentación del Data Mart ya construido, en donde se entregará al usuario un diccionario de datos en donde se plasmarán los objetos construidos (tablas dimensión y tabla histórica).

5. Capacitación a usuarios en el uso del Data Mart.

Se realizarán capacitaciones a los usuarios ya definidos que va a realizar la exploración del Data Mart en donde se les explicará mediante ejercicios de lenguaje modelado de consultas como deben hacer su uso. Asimismo, se les debe explicar la manera más óptima de ejecución de las consultas con el objetivo de asegurar el performance tanto del Data Mart como de la Bodega de Datos.

### 8.1. Descomposición de actividades WBS

En la siguiente figura, se muestra el diagrama WBS (Work Breakdown Structure)

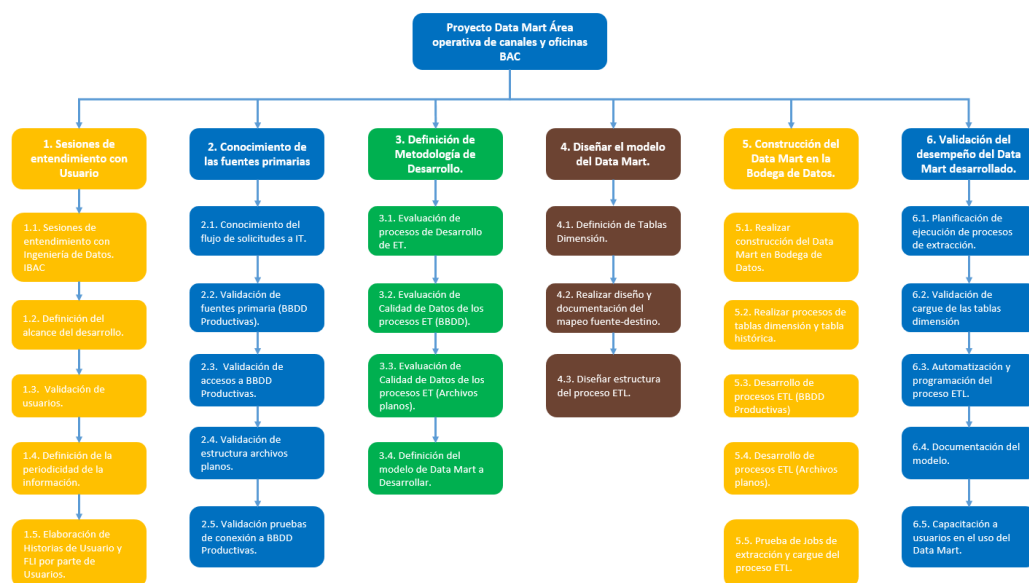


Figura 20: Figura n: Descomposición de Actividades.

## 8.2. Diagrama de Gantt

Las figuras 21 y 22 corresponden al diagrama de distribución de tiempo previsto para cada una de las actividades planteadas.

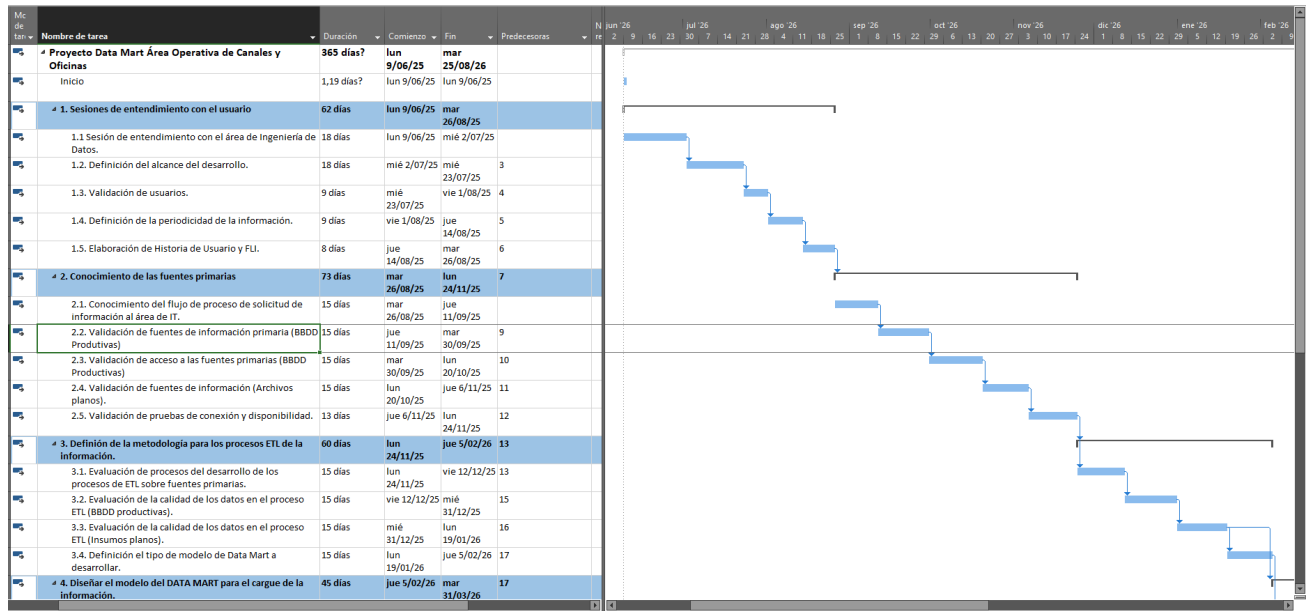


Figura 21: Figura 1: Diagrama - Parte 1

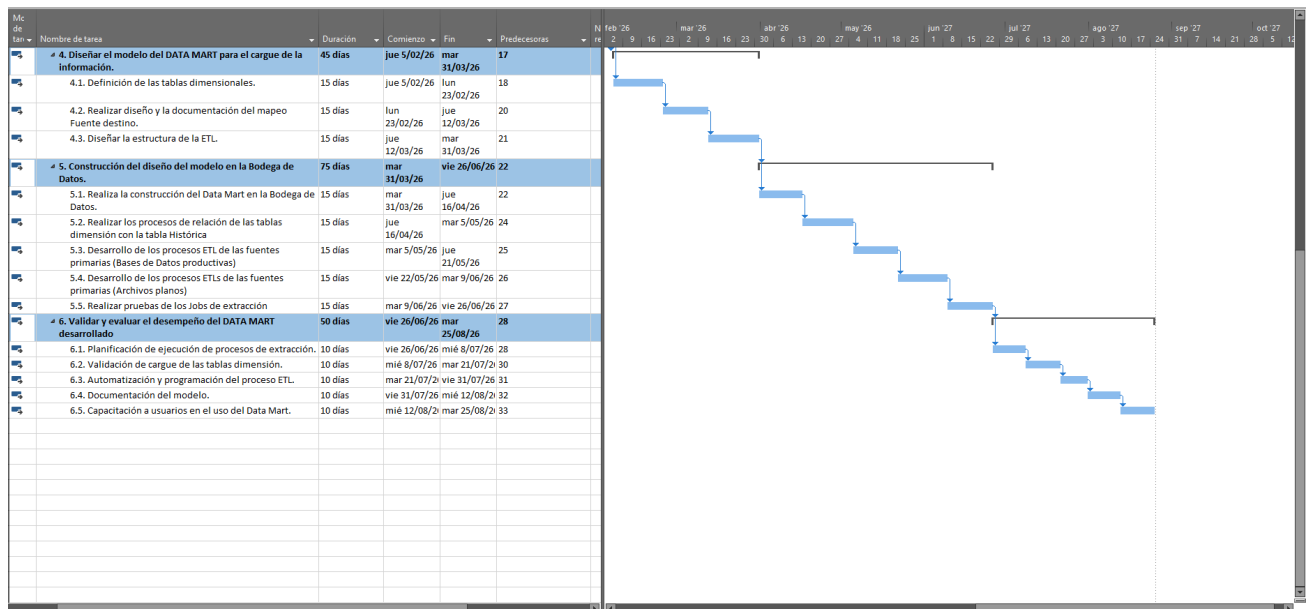


Figura 22: Figura 1: Diagrama - Parte 2

Diseño de un Data Mart para el almacenamiento histórico de información de canales, oficinas y transacciones para el área operativa de canales y oficinas del Banco Agrario de Colombia

---

## 9. Presupuesto

En la siguiente tabla 1 se indica que el presupuesto se toma con base a que el Banco Agrario ya cuenta con las herramientas (tanto de desarrollo de ETLs como la Base de Datos dedicada a Bodega de Datos licenciadas) y se toma únicamente tanto las horas de trabajo del estudiante, horas adicionales y horas de asesoría de la fábrica de desarrollo Teradata. Lo cual arroja un valor total de \$35,910,000(COP) para realización del proyecto.

Presupuesto			
Item	Cantidad	Precio Unitario(COP)	Total(COP)
Horas de Trabajo Estudiante (4 Horas diarias de jornada laboral)	1.460	\$20.000	\$29.200.000
Equipos de Computo	13	\$3.500.000	\$3.500.000
Horas de asesoría Fabrica de Desarrollo Teradata	10	\$300.000	\$3.000.000
Horas Adicionales	12	\$17.500	\$210.000
<b>TOTAL</b>			\$35.910.000

Tabla 1: Presupuesto

## 10. Conclusiones

De acuerdo con la propuesta de desarrollo del Data Mart dedicado al área operativa de canales y oficinas del Banco Agrario de Colombia, se espera que se cumplan los siguientes objetivos:

- Que los usuarios del área operativa de canales y oficinas puedan acceder de manera prioritaria a la información histórica para realizar sus actividades.
- Se pretende que se reduzca el escalamiento de solicitudes de información al área de IT. Ya que los usuarios podrán auto-atenderse para extrayendo la información desde su Data Mart.
- La integración de múltiples fuentes de información mediante procesos robustos de ETL, así como la implementación de un modelo estrella, aseguran una estructura de datos escalable, eficiente y alineada con las necesidades del negocio.
- Esta propuesta se basa en buenas prácticas de modelado dimensional y gobierno de datos, lo que sienta las bases para futuras iniciativas analíticas dentro de la organización.

## Referencias

- [1] M. Fernández-Olmos, J. Fleta-Asín, T. Gómez-Aguas, F. Muñoz y C. Sáenz-Royo, «Improved database of public-private partnerships from World Bank with imputed economic, institutional and conflict data,» *Data in Brief*, vol. 50, pág. 109742, 2025, Art. no. 109742. DOI: 10.1016/j.dib.2025.109742.
- [2] P. E. Oguntunde et al., «Statistical analysis of bank deposits dataset,» *Data in Brief*, vol. 18, págs. 864-872, 2018. DOI: 10.1016/j.dib.2018.03.096.
- [3] D. F. Eluyela et al., «Datasets for board meeting frequency and financial performance of Nigerian deposit money banks,» *Data in Brief*, vol. 19, págs. 1852-1855, 2018. DOI: 10.1016/j.dib.2018.06.044.
- [4] J. Pokorný, A. Ruschka y H. Chytilová, «Panel dataset on de jure central bank independence in 21 OECD countries (excluding the Eurozone),» *Data in Brief*, vol. 53, pág. 110094, 2024, Art. no. 110094. DOI: 10.1016/j.dib.2024.110094.
- [5] A. Wiraguna, R. Rokhim, B. Wibowo y R. Sembel, «Dataset on direct and indirect effects of MSME loan securitization issuances on stability of banks,» *Data in Brief*, vol. 54, pág. 110437, 2024, Art. no. 110437. DOI: 10.1016/j.dib.2024.110437.
- [6] P. Gala, Z. Abbas y A. Balakumar, *Data Modernization in BFSI Industry – From Legacy to Leading-edge*, Online, Available at: <https://www.indiumsoftware.com>, 2023. dirección: <https://www.indiumsoftware.com>.
- [7] Y. Wang, «Superposed Atomic Representation for Robust High-Dimensional Data Recovery of Multiple Low-Dimensional Structures,» en *Proceedings of the AAAI Conference on Artificial Intelligence*, vol. 38, AAAI Press, 2024, págs. 15735-15742. DOI: 10.1609/aaai.v38i14.29502.
- [8] N. D. D. Méndez y A. T. Alzate, *Data Warehouse (Bodega de Datos). Herramienta para la toma de decisiones. Parte I*, Online, Available: [archivo PDF proporcionado], Colombia, 2025.
- [9] Y. C. M. Paredes, *Generación de campañas en una institución financiera mediante el uso de Data Warehouse y Datalab*, Online, Trabajo de Suficiencia Profesional, Available at: <http://hdl.handle.net/10757/670134>, Lima, Perú, 2023. dirección: <http://hdl.handle.net/10757/670134>.
- [10] N. M. J. Orihuela, *Implementación de un Data Mart para la productividad del área comercial en una entidad financiera de Arequipa*, Online, Tesis de pregrado, Available at: <https://repositorio.ucsm.edu.pe/handle/20.500.12920/11216>, Arequipa, Perú, 2021. dirección: <https://repositorio.ucsm.edu.pe/handle/20.500.12920/11216>.
- [11] P. Jakubik y B. G. Moinescu, «Tailored microprudential recommendations for bank profit retention using a risk tolerance framework,» *International Review of Economics and Finance*, vol. 98, pág. 103951, feb. de 2025, Available at: <https://doi.org/10.1016/j.iref.2025.103951>. DOI: 10.1016/j.iref.2025.103951.
- [12] U. Panizza, «Bank ownership around the world,» *Journal of Banking and Finance*, vol. 166, pág. 107255, 2024, Art. no. 107255. Available at: <https://doi.org/10.1016/j.jbankfin.2024.107255>. DOI: 10.1016/j.jbankfin.2024.107255.

- [13] D. A. Adu, M. Z. Abedin, V. Y. Saa y F. Boateng, «Bank sustainability, climate change initiatives and financial performance: The role of corporate governance,» *International Review of Financial Analysis*, vol. 95, pág. 103 438, 2024, Art. no. 103438. Available at: <https://doi.org/10.1016/j.irfa.2024.103438>. DOI: 10.1016/j.irfa.2024.103438.
- [14] A. Vaisman y E. Zimányi, *Data Warehouse Systems: Design and Implementation*, 2.<sup>a</sup> ed. Berlín, Alemania: Springer, 2022, págs. i-xxvi, 1-696.
- [15] R. Kimball y M. Ross, *The Data Warehouse Toolkit: The Definitive Guide to Dimensional Modeling*, 3.<sup>a</sup> ed. Indianapolis, IN, EE. UU.: Wiley, 2013.
- [16] J. Reis y M. Housley, *Fundamentos de Ingeniería de Datos: Planifique y desarrolle sistemas robustos de datos*, 1.<sup>a</sup> ed. Bogotá: Marcombo/Alpha Editorial, 2024, págs. 108-112.
- [17] J. Holdsworth y M. Kosinski, *¿Qué es un almacén de datos?* Online, [Accedido: 12 de mayo de 2025], oct. de 2024. dirección: <https://www.ibm.com/es-es/topics/data-warehouse>.
- [18] L. J. Aguilar, *Inteligencia de negocios y analítica de datos*, 1.<sup>a</sup> ed. Ciudad de México: Alfaomega, 2019, págs. 165-167.
- [19] IBM, *¿Qué es un Data Mart?* Online, [Accedido: 24 de mayo de 2025], 2024. dirección: <https://www.ibm.com/mx-es/topics/data-mart>.