
Diseño de una Bodega de Datos para la automatización de la asignación de gastos a regionales mediante BI en COOPIDROGAS

Autores

David Francisco Silva Venail
Erika Natalia Osorio Reyes

Director

Andres Julian Aristizabal Cardona

Co-Director

Mauricio Garcés Restrepo



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

	Página
Resumen	IV
Abstract	V
1. Introducción	1
2. Descripción del Problema	2
3. Estado del Arte	5
3.1. Fundamentos y enfoques de Business Intelligence (BI)	5
3.2. Aplicaciones de BI en la gestión financiera y contable, Automatización y precisión en decisiones financieras	6
3.3. Infraestructura y arquitectura para soluciones BI modernas	7
3.4. Diseño técnico de almacenes de datos y procesos ETL	8
4. Objetivos	9
4.1. Objetivo General	9
4.2. Objetivos Específicos	9
5. Marco teorico	10
5.1. La inteligencia de negocios o Business Inteligence	10
5.2. Las bodegas de datos y los Datamarts; El eje crucial del proyecto	12
6. Requerimientos	13
7. Solucion propuesta	14
7.1. Descripción general de la solución	14
7.2. Modelo conceptual	14
7.3. Estándares de la solución	19
7.4. Condiciones de diseño, propuesta de implementación y evaluación	21
8. Planeación del trabajo	23
8.1. Fase 1. Diagnóstico del proceso actual y definición de requerimientos	23
8.2. Fase II. Diseño del modelo dimensional y físico	24
8.3. Fase III. Desarrollo del proceso ETL (Extracción, transformación y carga)	24
8.4. Fase IV: Selección y justificación del sistema de gestión de base de datos (DBMS)	24
8.5. Fase V: Validación del modelo mediante herramientas de visualización	25
8.6. Descomposición de actividades WBS	25
8.7. Cronograma	26
9. Presupuesto	29
10. Conclusiones	30

Referencias	34
A. Anexos	35
A.1. Anexo A: Diferentes figuras que soportan el presente proyecto	35

Índice de figuras

1.	System Architecture of a Cost-Intelligent Cloud Data Warehouse	8
2.	A Business Intelligence System	11
3.	Ambiente de Producción	14
4.	Modelo logico	15
5.	Modelo relacional de la fuente	15
6.	Descipción medidas	16
7.	Descripción dimensiones	17
8.	Modelo dimensional	18
9.	Paleta de colores 1	20
10.	Paleta de colores 2	20
11.	Matriz de colores de riesgo	20
12.	Codificación paleta de colores	21
13.	Ejemplo Mockup	22
14.	Esquema Work Breakdown Structure oWBS	26
15.	Actividades fragmento 1	27
16.	Actividades fragmento 2	27
17.	Diagrama de Gantt fragmento 1	28
18.	Diagrama de Gantt fragmento 2	28
19.	Presupuesto nomina	29
20.	Presupuesto activos	29

Resumen

La cooperativa COOPIDROGAS gestiona la asignación de sus costos operativos mediante el costeo basado en actividades (ABC). Sin embargo, este proceso se realiza manualmente, extrayendo datos desde SAP, procesándolos en Excel y luego cargándolos nuevamente en SAP, lo que genera demoras, errores y afecta la confiabilidad de los informes gerenciales.

Este proyecto propone diseñar un modelo dimensional para una bodega de datos que automatice la asignación de gastos, utilizando herramientas de inteligencia de negocios (BI), procesos ETL y almacenamiento en un Data Warehouse. La implementación de esta solución optimizará los tiempos de análisis, reducirá errores y fortalecerá la calidad de la información financiera.

Se establecen los requerimientos de calidad, negocio, sistema y usuario, priorizando la integración con SAP y el uso de POWER BI para la gestión de datos. Este modelo permitirá la consolidación de información operativa y financiera, mejorando la toma de decisiones estratégicas en la cooperativa.

Abstract

The COOPIDROGAS cooperative manages the allocation of its operating costs using activity-based costing (ABC). However, this process is done manually, extracting data from SAP, processing it in Excel and then uploading it back to SAP, which generates delays, errors and affects the reliability of management reports.

This project proposes to design a dimensional model for a data warehouse that automates the allocation of expenses, using business intelligence (BI) tools, ETL processes and storage in a Data Warehouse. The implementation of this solution will optimize analysis times, reduce errors and strengthen the quality of financial information.

Quality, business, system and user requirements are established, prioritizing integration with SAP and the use of QLIK for data management. This model will allow the consolidation of operational and financial information, improving strategic decision making in the cooperative.

1. Introducción

En la actualidad, la eficiencia en la gestión de costos es fundamental para la sostenibilidad financiera de las organizaciones. La cooperativa COOPIDROGAS cuenta con cinco centros de distribución y una sede principal en Cota, desde donde se gestionan los gastos operativos. Actualmente, el área de costos y presupuestos asigna estos gastos utilizando la metodología de costeo basado en actividades (ABC), lo que permite una asignación más precisa de los costos indirectos a productos o servicios según su consumo de recursos [1] o, en el caso de COOPIDROGAS, líneas de negocio o centros de distribución. Sin embargo, el proceso es manual, basado en datos extraídos de SAP, procesados en Excel y posteriormente subidos nuevamente a SAP. Este flujo de trabajo genera demoras, errores y pérdida de credibilidad en los informes gerenciales, afectando la toma de decisiones.

Para mejorar esta situación, el presente proyecto propone el diseño de un sistema automatizado para la asignación de gastos, basado en herramientas de inteligencia de negocios (BI), bodegas de datos y automatización de procesos. La implementación de sistemas de BI permite a las empresas transformar datos en información útil para la toma de decisiones estratégicas [2]. Además, el uso de bodegas de datos facilita la consolidación y procesamiento de información histórica, mejorando la calidad y disponibilidad de los datos [3]. La automatización de este proceso contribuirá a reducir errores, optimizar tiempos de análisis y fortalecer la confiabilidad de los reportes financieros y gerenciales.

2. Descripción del Problema

La cooperativa COOPIDROGAS opera a través de cinco centros de distribución ubicados en Barranquilla, Bucaramanga, Cali, Medellín y Pereira. Su sede principal se encuentra en Cota, desde donde se coordina la operación central y se gestiona un centro de distribución adicional. La cooperativa distribuye los gastos operativos de la sede principal para dar como resultado la rentabilidad total de cada regional.

Para ello, se emplea la metodología de costeo ABC, que permite asignar los gastos a través de centros de costo. Sin embargo, dado que diversas áreas de la sede principal realizan actividades en beneficio de los centros de distribución, estos gastos se deben distribuir considerando varios inductores, como ventas, unidades vendidas, compras, asociados, personal, entre otros. Esto permite tener una visión clara de la viabilidad de tener operación en esas regionales y así poder tomar decisiones asertivas.

Al revisar el sistema de costos ABC, vemos que tiene una falencia para las empresas que se encargan de la producción y distribución en sus productos, ya que no se logra encontrar la unión entre los costos inductores y los costos de las actividades. Esto lleva, según explica E. V. Ramos Farroñan, L. A. Huacchillo Pardo, y P. Portocarrero Medina en el artículo, “El sistema de costos ABC como estrategia para la toma de decisiones empresarial,” a entender que; “Es por ello que la empresa dedicada al procesamiento y comercialización de productos hidrobiológicos ha evidenciado problemas relacionados con los costos de los inductores y de las actividades, impidiendo que el sistema sea una total solución para conocer los costos incurridos en el proceso [...] presentando registros de costos con datos irreales, sin claridad, deficientes, para luego los resultados que se obtienen no permitan tomar las mejores decisiones [4]. Viendo esto, el proyecto se enfrenta, dentro de COOPIDROGAS a tener un costeo con datos ineficientes, que afectan la toma de decisiones por parte de la gerencia, ya que no se sabe que tan exactos son.

Esta problemática, se registra en un momento donde la cooperativa busca, desde hace unos años, consolidar su fuerza laboral y crear unos equipos de trabajo altamente productivos. Esto, gracias a un liderazgo presente y una línea directiva que esté integrada con las áreas, productos y sistemas dentro de la organización, con el fin de mejorar la toma de decisiones.

En el artículo de grado de Luis Guillermo Torres Enríquez, para la Cooperativa Universitaria Minuto de Dios, de mayo del 2024, se abre la puerta al análisis de un aumento en el número de la competencia en el sector de los distribuidores para los droguistas detallistas, viendo desde el año 2019, la llegada de cadenas de droguerías detallistas como Cruz Verde o Audifarma. Este análisis refleja que Coopidrogas, al verse en un modelo de liderazgo ya anticuado y con posibilidades de ser sobrepasado en números, por sus competidores, emprendió una nueva estrategia interna, al cual se explica de la siguiente forma; En los últimos años la venta de medicamentos y productos populares ha tenido la entrada de diferentes cadenas de farmacia que cada vez cobran más relevancia con los clientes [...] Esto ha generado que Coopidrogas desarrolle diferentes estrategias como lo son: nuevas sedes de distribución, implementación de la tecnología EWM, reestructuraciones internas, entre otras, todo esto en búsqueda de que sus asociados mantengan su competitividad en el mercado. Estos cambios ejecutados en la Cooperativa, ha generado el incremento de áreas y colaboradores, por lo que se hace necesario generar estrategias para el desarrollo interno de las mismas, y por supuesto de todas las partes interesadas (asociados, proveedores, clientes), esto permitirá el alcance de forma eficiente de las metas establecidas en cada grupo de trabajo y por supuesto, del posicionamiento de la

Cooperativa en el mercado [5]. Por ende, el proyecto actual busca el objetivo de inscribirse en esta nueva problemática hallada dentro de la cooperativa, proponiendo una nueva solución de apoyo al sistema de costeo ABC ya implementado.

Al revisar también las necesidades del sector de droguistas, quienes son los asociados a la cooperativa, se evidencia que el proyecto busca encontrar una solución a la problemática del sistema de costos. Pero, indirectamente da solución a otras problemáticas de los droguistas se puede contemplar, al mejorar el proceso de tomas de decisiones, esta solución llevara la revisión de otras problemáticas presentes dentro del sector detallista, como estrategias de ventas, costos y marketing, entre otras. En el artículo de Gutierrez Gomez y Vargas Montes del 2024 [6] , se puede apreciar que uno de los principales ejes de atender problemáticas enfocadas hacia la mejora de la toma de decisiones, según nos muestra el artículo de la siguiente manera ;” [...] el grado de importancia que tiene crear una propuesta comercial que aumente de manera significativa las ventas en toda la cadena de droguerías a nivel nacional, ya que se basa el creación de estrategias de marketing que permitan el sostenimiento y/o fidelizar al cliente, aumente la demanda y el consumo de productos farmacéuticos, para aumentar y obtener clientes nuevos y mayor posicionamiento en el mercado droguista”[6].

Ahora, continuando con el área de costos, se encuentra que el sistema de costos ABC se complementa actualmente en el área, con la información de planos de costos que se obtiene desde la aplicación SAP, cada mes.

Esto provoca que el valor de los inductores cambie. Hurtado Obando explica en su estudio de las droguerías PHAMASALUDERO del 2024 [7] que los archivos de costos e históricos de gastos, realizados en Excel de manera manual, producen cambios en los inductores y en los vínculos de los datos, que se subirán al sistema SAP, lo cual generará reportes poco precisos e imposibles de actualizar en tiempo real (ya que debe hacerse un proceso manual en Excel, antes de subirlo a SAP), ya que una actualización de datos, puede demorar hasta cinco horas de trabajo de parte de los analistas de costo, dado que se deben de distribuir los gastos de 98 centros droguistas en 23 tipos y subtipos de costos y gastos. Generando mayor complejidad a la hora de entregar a tiempo los análisis gerenciales, teniendo en cuenta que, en estos mismo reportes hay incongruencias entre los datos, dificultando la toma de decisiones.

Ahora, complementando el caso de Hurtado Obando, en el caso expuesto por [8] Toro Restrepo en “Plan de practica de Supertiendas y Droguerías Olimpica” (2024), explica como la competencia directa de la Cooperativa, encontró un equilibrio entre la fuente de datos manuales de Excel y los sistemas de reportes y visualizaciones, al haber adquiridos sistemas robustos, que permiten tomar los archivos en Excel, actualizados a diario por los centros regionales, subirlos al motor de base datos y realizar las visualizaciones con un software de Data analytics , todo en el menor tiempo posible. Permitiendo así una mejora Enel flujo de información entre áreas y siendo parte de la estrategia de innovación interna de la empresa Supertiendas y Droguerías Olimpica.

Justificando esta última idea y revisando el artículo de La Universidad de los Andes “Construcción de una bodega de datos para analítica descriptiva y predictiva en el observatorio de estudiantes de la Universidad de los Andes”, se puede ver como las bodegas de datos se enfrentan a problemáticas ligadas a la toma de decisiones, dado que el autor Manuel Felipe Porras Tascón explica en la siguiente cita las problemáticas a las que se enfrenta el diseño de una bodega de datos dentro de las organizaciones con parámetros ya establecidos: “Para abordar esta problemática, es indispensable establecer un modelo subyacente que organice y estructure los datos de manera eficiente. Este modelo no solo permitirá la creación de reportes replica-

bles, sino que también sentará las bases para implementar herramientas avanzadas de analítica descriptiva y predictiva, facilitando la caracterización de las poblaciones universitarias y el desarrollo de estrategias que respondan a sus necesidades específicas. En este contexto, surge la importancia de la Inteligencia de Negocios, que se refiere al conjunto de técnicas y herramientas diseñadas para ayudar a las organizaciones a tomar decisiones informadas a partir de sus propios datos. Estas técnicas incluyen minería de datos, visualización de información, análisis de negocios e infraestructura de datos, entre otras” [9].

Porras Tascón [9], proporciona el último tramo de la problemática que resuelve este proyecto; El ligar aplicativos de Data Analytics con la Bodega de datos.

SAP junto a su producto SAP ERP (Enterprise Resources Planning) es un sistema de gestión empresarial utilizado por organizaciones del mundo, con el fin de administrar y automatizar diversos procesos empresariales. En la empresa Coopidrogas, se realizan análisis de datos y reportes, los cuales son transmitidos a la gerencia, ayudando así, a su toma de decisiones. Es de vital importancia entender que la principal problemática entre los sistemas de costos y el programa SAP, como nos lo explica [10] Reyes Osore y Valle en su plan de migración de datos a SAP S4 HANA, la clave de cualquier integración de dos sistemas de procesamiento de datos, es la veracidad y la velocidad de entrega de estos desde la fuente, hacia el área donde se realiza el proceso de Visualización de informes finales y Reporting.

3. Estado del Arte

Para realizar la automatización de distribución de gastos a las regionales de una empresa con presencia en todo el territorio nacional, como lo es La Cooperativa Nacional de Droguistas Detallistas o COOPIDROGAS, Esta revisión incluyó estudios, investigaciones y propuestas aplicadas en distintos sectores, con el fin de identificar buenas prácticas y enfoques metodológicos en la implementación de sistemas de Business Intelligence (BI). En los documentos expuestos a continuación se resalta la importancia de integrar tecnologías de Big Data y procesos ETL avanzados en sistemas de BI, con el objetivo de optimizar la gestión financiera y mejorar la calidad en la toma de decisiones estratégicas. Estos aportes resultan fundamentales para sustentar el desarrollo del presente proyecto, ya que evidencian el potencial transformador de la analítica de datos en entornos organizacionales complejos y de alta exigencia operativa como el de COOPIDROGAS.

3.1. Fundamentos y enfoques de Business Intelligence (BI)

- Evolución de la analítica de datos y BI

Con el paso de los años, la analítica de datos ha adquirido una creciente relevancia en el ámbito empresarial, ya que esto permite identificar tendencias, oportunidades de negocio y ayuda a la toma de decisiones informadas que impulsan el crecimiento de las organizaciones. Los volúmenes de datos cada vez son más robustos y las organizaciones dependen de sistemas de información para llevar a cabo estos análisis. En este contexto, surge la necesidad de implementación de herramientas de Business Intelligence (BI). Sin embargo, la implementación se enfrenta a una serie de desafíos. En un estudio realizado por Maghsoudi y Nezafati, se analizan dichos obstáculos, parten de dos metodologías, la implementación tradicional y la de autoservicio. Los modelos tradicionales son los realizados por los departamentos de tecnología de la información (TI) que toman el control del proceso de desarrollo del sistema de **Business Intelligence (BI)** el cual está centralizado allí, este desarrollo está dividido en varias etapas como la ETL (Extracción, transformación y carga), el modelado de datos, la generación de informes, el análisis y distribución de la información. Estos procesos a menudo requieren conocimientos especializados como un modelado dimensional, extracción de varias fuentes, etc. Mientras que los modelos de autoservicios están dentro del área o usuario final que lo requiere, lo que les permite tener control y acceso a los datos cuando lo necesitan, enfocándose también en sus necesidades específicas. El resultado de este estudio indica que la estrategia de autoservicio tiene mayores niveles de aceptación organizacional y demostró ser un 30 por ciento más efectiva que la tradicional. Sin embargo, la implementación tradicional puede mejorar si se abordan factores como el conocimiento del dominio, la cooperación con los solicitantes y el desarrollo de habilidades blandas[11]. Inteligencia negocios como ventaja competitiva, es el planteamiento de Pałys y Pałys, ya que las empresas con el pasar del tiempo deben procesar cada vez más datos, y el enfoque de BI tradicional puede no ser suficiente, los autores sugieren el uso del enfoque de Self-Service Business Intelligence (SSBI). El objeto de este estudio es [...] sintetizar el conocimiento sobre los beneficios y desafíos de la implementación de SSBI [...] [12]. Los autores encontraron varios beneficios de este enfoque como acceso fácil y rápido a la información, asignación de capacidades y la promoción de una cultura organizacional enfocada en los datos. Por otro lado, encontraron retos relacionados con el acceso y el uso de datos, la autosuficiencia de los usuarios y la preparación organizacional. Adicional a otras

nueve subcategorías que describen cada uno de los retos. Sin embargo, es una revisión literaria, por tanto, se recomienda tener investigaciones que incluyan estudios de caso para validar los hallazgos [12].

- Tipos de uso y madurez del BI

En la actualidad se evidencia como áreas de gestión de contabilidad de costos y gastos se ven obligadas a utilizar herramientas que permitan el análisis de información sea de una manera rutinaria o avanzada, En este contexto, Maldonado y Guzmán, analizan el impacto que genera el uso de herramientas de BI que utilizan funciones avanzadas como análisis predictivos, visualizaciones interactivas, funciones de desglose, informes de paneles, optimización, herramientas de construcción de consultas y modelado predictivo impulsan la mejora de la productividad y la calidad de la toma de decisiones. El estudio también profundiza en los factores que determinan si un profesional opta por un uso rutinario o avanzado de estas herramientas, señalando que dicha decisión depende de elementos como la complejidad de las tareas, las características del sistema y, especialmente, la autoeficacia del usuario. De esta manera, se concluye que aquellos usuarios que desarrollan una mayor autoeficacia para el uso avanzado de BI obtienen mejores resultados siempre y cuando las tareas a realizar sean lo suficientemente complejas como para justificar su aplicación. Estos hallazgos son relevantes en entornos organizacionales que manejan una alta carga operativa, como COOPIDROGAS [13].

3.2. Aplicaciones de BI en la gestión financiera y contable, Automatización y precisión en decisiones financieras

Du propone un sistema de gestión financiera inteligente que integra algoritmos de árbol de decisión y técnicas de minería de datos para apoyar la toma de decisiones en entornos financieros. Para su validación, el estudio utiliza datos de una ejecución presupuestal del período 2017-2021 para mejorar los procesos de gestión del presupuesto, implementando tecnologías que automatizan tareas como la detección de anomalías y la predicción de datos financieros. Los resultados evidencian una mejora en la eficiencia operativa y en la precisión del análisis financiero, permitiendo decisiones más ágiles y fundamentadas. Aunque los autores reconocen que el conjunto de datos de prueba es limitado y puede introducir cierto margen de error, el estudio aporta al presente proyecto al demostrar el potencial transformador de las tecnologías inteligentes en la gestión contable y financiera, respondiendo a la creciente necesidad de herramientas avanzadas para el análisis automatizado y predictivo [14]. Azevedo, Duarte y Santos presentan un estudio en el que desarrollan una solución de contabilidad de costos basada en herramientas de Business Intelligence (BI) para un sistema hospitalario. El trabajo abarca todas las etapas necesarias para construir un modelo relacional alineado con los requerimientos del negocio, incluyendo el diseño de múltiples mockups para el front-end que permiten visualizar métricas clave enfocadas en la toma de decisiones. Los autores destacan que se trata de un proyecto en curso con potencial de ser replicado en distintas áreas funcionales de la organización. Este aporte es importante para el presente proyecto ya que se demuestra como el diseño estructurado de soluciones BI puede adaptarse a diferentes contextos específicos para la toma de decisiones como en sectores de la salud que son de alta complejidad [15]. Por otro lado, el trabajo titulado .Automatización de un proceso ETL para la industria farmacéutica mediante un modelo predictivo basado en Power BI.^aborda la necesidad de mejorar la gestión

y análisis de datos en el sector farmacéutico. Los autores desarrollaron una solución que automatiza el proceso ETL (Extracción, Transformación y Carga) utilizando Power BI, integrando además un modelo predictivo. Implementaron un proceso ETL automatizado que facilita la integración y transformación de datos provenientes de diversas fuentes dentro de la industria farmacéutica. Posteriormente, desarrollaron un modelo predictivo en Power BI para analizar y visualizar la información procesada, permitiendo la toma de decisiones, según los autores, dicha automatización redujo significativamente el tiempo y los errores asociados a la gestión manual de datos. El modelo predictivo implementado en Power BI proporcionó herramientas de visualización y análisis avanzadas, mejorando la capacidad de la industria para anticipar tendencias y optimizar procesos internos [16]. El proyecto titulado "Propuesta de implementación de un sistema de BI predictivo basado en Big Data y ETL avanzado para la mejora de la toma de decisiones empresariales." aborda la necesidad de implementar un sistema de inteligencia de negocios en Practisistemas S.A.S. debido a la falta de integración y precisión en el análisis de datos en áreas críticas como la gestión de ventas, seguimiento de clientes y control de cartera. Los autores proponen una solución que integra grandes volúmenes de datos en un almacén centralizado, facilitando su análisis y proyección futura. La metodología incluye la transformación y depuración de datos mediante procesos ETL avanzados, así como la implementación de un sistema de análisis predictivo. Se espera que la implementación de este sistema mejore la capacidad de respuesta de la empresa en un entorno competitivo, permitiendo una toma de decisiones más informada y efectiva, reducción de costos y una ventaja competitiva sostenible [17].

3.3. Infraestructura y arquitectura para soluciones BI modernas

- Optimización de costos en la nube

Zhang, Lio y Yan introducen el concepto de inteligencia de costos en el análisis de datos en la nube, parten del hecho de que no solo se debe tener en cuenta el rendimiento en un almacén de datos en la nube sino también el costo que eso genera para las organizaciones. Los principales aportes del estudio abarcan un enfoque dual de optimización del costo y el rendimiento, proporcionan una arquitectura de almacén de datos en la nube enfocada y diseñada para la inteligencia de costos en donde enfrentan dos desafíos, el despliegue automático de recursos y la autoajuste orientado al costo. También distingue entre el costo del usuario y el proveedor con el fin de entender como esto impacta en la rentabilidad de ambos [18].

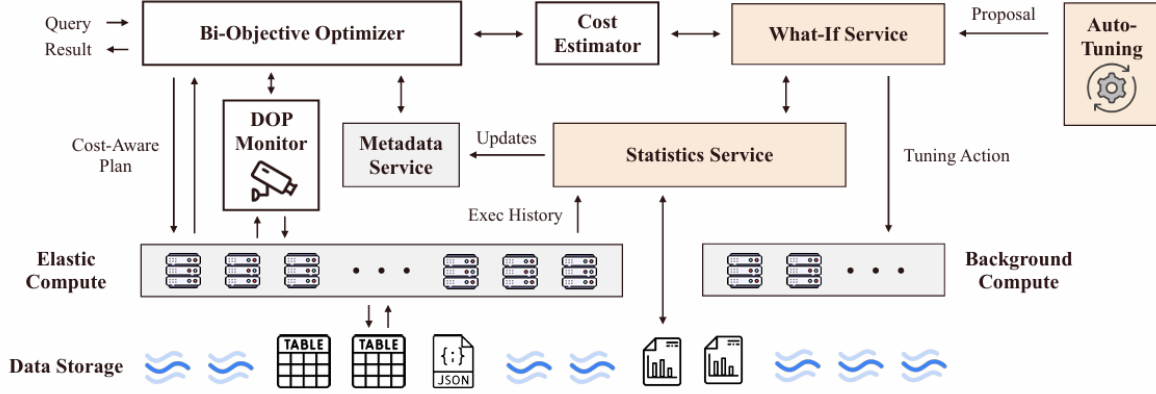


Figura 1: System Architecture of a Cost-Intelligent Cloud Data Warehouse

[18]

3.4. Diseño técnico de almacenes de datos y procesos ETL

Las herramientas de BI se aplican a cualquier área de negocio, por ello en el presente documento se han expuesto diferentes artículos en los que se aborda el tema, desde sectores como la salud hasta el sector marítimo. Gonçalves y Teixeira se enfrentan al desafío de [...] apoyar el proceso de transición digital y ecológica en el transporte, la logística y las operaciones portuarias a través de un sistema de apoyo a la decisión desarrollado utilizando Power BI para estructurar y acelerar el proceso de análisis de datos. Gracias al crecimiento de la actividad portuaria los autores ven la necesidad de acceder a estas herramientas ya que han identificado problemas como dificultad para planificar o estructurar los servicios logísticos por el desconocimiento de la demanda y datos dispersos sobre todo lo inherente al negocio. Para la solución de estos problemas los autores usan una metodología de ciencia de datos llamada CRISP-DM, en la que desarrollan todos los pasos de la metodología para el desarrollo de una bodega de datos. La herramienta desarrollada permitió que los interesados visualizaran diferentes indicadores de rendimiento del puerto, lo cual les otorga la capacidad de tomar decisiones asertivas y reducir el esfuerzo de un análisis manual [19].

4. Objetivos

4.1. Objetivo General

Diseñar un modelo dimensional para una bodega de datos que permita estructurar, organizar e integrar la información operativa y financiera de la cooperativa, facilitando la automatización del proceso de costeo y la generación de análisis gerenciales confiables para la toma de decisiones.

4.2. Objetivos Específicos

- Realizar un análisis de las fuentes de datos que documente los elementos clave del modelo de la bodega de datos, incluyendo dimensiones, hechos, atributos e inductores, para garantizar un entendimiento claro de la estructura de los datos y su integración en el proceso de almacenamiento y análisis.
- Proponer el modelo dimensional de la bodega de datos con base en los requerimientos del usuario, estableciendo un esquema estrella o copo de nieve que estructure los hechos y dimensiones, asegurando la cobertura de al menos once procesos clave de la cooperativa.
- Definir y desarrollar un proceso ETL (Extracción, Transformación y Carga) que automatice la integración de datos desde diversas fuentes, asegurando su calidad y consistencia permitiendo una mejora en la tasa de cumplimiento de plazos.
- Seleccionar y justificar el sistema de gestión de base de datos (DBMS) más adecuado para la implementación del modelo, evaluando su rendimiento, escalabilidad e integración con software de análisis de datos.
- Validar el diseño de la bodega de datos mediante la utilización de herramientas de visualización como Tableau o Power BI, para asegurar que los datos almacenados se presenten de manera clara, precisa

5. Marco teorico

Para realizar la automatización de distribución de gastos a las regionales de una empresa con presencia en todo el territorio nacional, como lo es La Cooperativa Nacional de Droguistas Detallistas o COOPIDROGAS, se revisaron estudios, investigaciones y textos, unidos con tres temas puntuales: La inteligencia de negocios (BI) junto con la automatización de los procesos a través del BI, las bodegas de datos, metadatos y Datamarts.

5.1. La inteligencia de negocios o Business Intelligence

La inteligencia de negocios o BI, en términos generales suena a una palabra moderna, de reciente uso y aplicación en los negocios del siglo XXI. Lo sorprendente de este término, es que fue acuñado en 1958, gracias al investigador de IBM, Hans Peter Luhn, al describirlo en su artículo A Business Intelligence System, definiéndolo como “[...] un sistema automático para difundir información a las distintas secciones de cualquier organización industrial, científica o gubernamental. Este sistema de inteligencia utilizará máquinas de procesamiento de datos para la auto extracción y codificación de documentos, así como para la creación de perfiles de interés para cada uno de los puntos de acción de una organización.” [20]. Es así como se crea el primer diagrama de bloques para BI, soportado en la creación de los Sistemas de Soporte de Decisiones (DSS), enfocados hacia las áreas gerenciales y basados en los análisis de datos. Estos sistemas DSS son la base para el modelo que se desarrollara en el proyecto de COOPIDROGAS.

Ahora, en base a ellos vemos que, IBM, la empresa a la que pertenecía Luhn, desarrolló esta idea de 1958 y permite entender, hoy en día, que la inteligencia de negocios moderna se basa en un conjunto de herramientas tipo software, que permite a los usuarios empresariales, desglosar los datos del negocio y presentarlos de manera sencilla, usando vistas, gráficos, tablas o históricos. Todo esto, con el fin de optimizar la toma de decisiones y simplificar los procesos del día a día en las compañías, volviendo a dar uso a los sistemas DSS creados por IBM y Luhn.

omitir las ideas y teorías de Howard Dresner, quien, en 1989 [21], fue quien siguió la teoría de “A Business Intelligence system” y dio paso a la creación de los sistemas de Data Warehouse y los procesos ETL ya mencionados.

5.2. Las bodegas de datos y los Datamarts; El eje crucial del proyecto

Si bien Hans Luhn define la inteligencia de negocios, fue Howard Dresner quien la aplicó a mayor escala en las compañías. Al revisar sus escritos, el autor decide reorientar las bases del BI, indicando que “La inteligencia de negocios o BI es un paraguas bajo el que se incluye un conjunto de conceptos y metodologías cuya misión consiste en mejorar el proceso de toma de decisiones en los negocios basándose en hechos y sistemas que trabajan con hechos” según el Gartner Group en 1989.

Fue a raíz de los análisis de Dresner, que se inicia la creación y uso de los almacenes de datos o Data Warehouses, junto al científico Bill Inmon. Este último define a los almacenes de datos como “[...] una recopilación de datos no volátil, variable en el tiempo, integrada y orientada a temas, que respalda las decisiones de gestión” según Bill Inmon en “Building the Data Warehouse”, 1992 [22]. O, en otras palabras, los almacenes de datos o Data Warehouses son un sistema de almacenamiento, que se diseña con el fin de consolidar, integrar y analizar grandes volúmenes de información provenientes de varias fuentes. Todo, con el fin de facilitar la toma de decisiones en la organización, haciendo uso del análisis de los datos pasados y conmutativos. En el siguiente gráfico se puede apreciar un ejemplo de la estructura base de una bodega de datos en base al modelo de Bill Inmon de 1990;

Para el caso del proyecto de las regionales de costos de COOPIDROGAS, el establecer un almacén de datos dedicado a recopilar en un solo sistema los datos primarios de cada ciudad o regional, facilita al área de análisis su extracción, procesamiento y carga (ETL), dando así paso a los metadatos que se incorporaran en los Datamarts.

Siguiendo con este último tema de metadatos y Datamarts, se puede unir a los sistemas de almacenes de datos, ya que son subsistemas de este, que se crean una vez el software tiene todo el modelo base de almacenamiento creado y los procesos ETL en ejecución.

Pero, ante la incógnita del manejo de los metadatos creados a partir de los datos primarios, se hace uso del artículo de Senso y De la Rosa [23] definen los metadatos como “Según Howe (1993), el término fue acuñado por Jack Myers en la década de los 60 para describir conjuntos de datos. La primera acepción que se le dio (y actualmente la más extendida) fue la de dato sobre el dato, ya que proporcionaban la información mínima necesaria para identificar un recurso”.

En base a estos autores, el modelo que se realizará en su estado del arte a las ideas expuestas por ellos, ya que, al momento de agrupar un gran volumen de datos, esto permitirá hacer una catalogación de los primarios y por consiguiente creará los metadatos necesarios para el análisis y posteriormente la toma de decisiones para las gerencias.

Por último y no menos importante, al haber revisado los metadatos como un producto que nace de la catalogación de los datos primarios de la sede principal en COOPIDROGAS, se debe revisar en la definición y creación de Datamarts. En este caso, Bill Inmon en “Building the Data Warehouse” de 1992 vs Ralph Kimball en “The Data Warehouse Toolkit” de 1996, ambos definen los Datamarts como un subconjunto del almacén de datos general, los cuales están enfocados en dar un fácil acceso a los datos relevantes para un departamento o área específica.

6. Requerimientos

- Requerimientos de calidad.

La bodega de datos o Data Warehouse que se creará con el modelo para la empresa COOPIDROGAS, debe tener sus procesos y sistemas informáticos, regulados bajo el sistema de gestión de calidad ISO 90001. El almacenamiento de datos debe adaptarse a los estándares exigidos por la norma (la cual vigila los estándares de calidad de la empresa), promoviendo la eficiencia y calidad en los resultados que brinda. Es por ello que esta normativa debe aplicarse con rigor principalmente en los procesos ETL (Extracción, Procesamiento y Carga), los cuales darán origen a los metadatos que se usan en las consultas.

- Requerimientos del negocio/ sistema.

Para la creación de la bodega de datos, COOPIDROGAS tiene ciertos parámetros que debe de incluir el proyecto. Relacionado con el sistema, es requisito el tener una ubicación clara del administrador de la bodega, donde haya una red estable, Sistema operativo Windows, MAC o Linux (ya que la empresa usa las tres), un servidor dedicado con dirección IP estática, antivirus Kaspersky Security Network (KSN), puertos disponibles para la CPU dedicada a la bodega de datos, DNS, copias de respaldo en nube y en discos duros o memorias. Además, es prioridad por parte de la Cooperativa, el contratar los servicios de la compañía sueca de Auto Data Warehouse QLIK, quienes proporcionarán una plataforma de integración de datos junto con un software para los análisis, respaldados por inteligencia artificial. Esta contratación será anual, realizada por el departamento de compras, bajo órdenes de la gerencia e integrada al presupuesto anual de costos de la compañía.

- Requerimientos de los usuarios.

Al tener una bodega de datos esta debe tener un administrador del sistema, el cual cataloga y crea los usuarios junto con los permisos que necesita cada perfil. Cada usuario tendrá su “username”, contraseña y se contará con la capacidad de permitir el acceso seguro a cada uno de ellos. Igualmente, como requerimiento fundamental al usuario es que el proyecto tenga la capacidad de consolidar todos los datos en un solo Data Warehouse, además de ser fácil de gestionar y001 administrar.

- Requerimientos funcionales:

Extracción de datos desde sistemas actuales: Definir estrategias para la migración desde sistemas legados sin pérdida de información.

Integración de datos, La propuesta de bodega de datos debe permitir la integración entre los diferentes inductores o hechos que juegan como base para la distribución de un gasto.

Procesos ETL automatizados, el sistema debe extraer, transformar y cargar datos de manera eficiente desde SAP.

Estructura de datos optimizada, Debe soportar modelos multidimensionales (OLAP) con tablas de hechos y dimensiones y soporte para procesamiento eficiente de grandes volúmenes de datos (indexación, particionamiento).

7. Solucion propuesta

7.1. Descripción general de la solución

El diseño de la Bodega de datos busca automatizar el proceso de revisión y reportes de los gastos hacia las regionales de la empresa COOPIDROGAS. La automatización se realizará a través de la compra e implementación de un servidor en la nube de base de datos (con escalabilidad para la formación de un Datawarehouse), la creación de procesos de ETL eficientes, que garanticen la actualización diaria de los datos, desde las regionales de gastos. Permitiendo así mejorar la comunicación entre el área de costos y la gerencia general, demostrando la capacidad de crear un nuevo sistema de procesamiento de datos robusto, rápido y que se actualiza diariamente, permitiendo así al área tener nuevas funciones, contratar nuevos empleados capacitados, los cuales garantizarán el funcionamiento del sistema y la rapidez en las consultas de la información requerida por los usuarios de la empresa. Esto es, con el fin de garantizar un flujo de datos optimo, una visualización de los reportes de manera ágil y certera, evitando datos erróneos o entorpecimiento en la toma de decisiones futuras por parte de la gerencia.

7.2. Modelo conceptual

■ Diagrama de arquitectura

En la figura a continuacion, se encuentra el diagrama de arquitectura el cual representa de forma visual la estructura y los componentes del sistema con el fin de facilitar la comprensión y comunicación del proyecto.

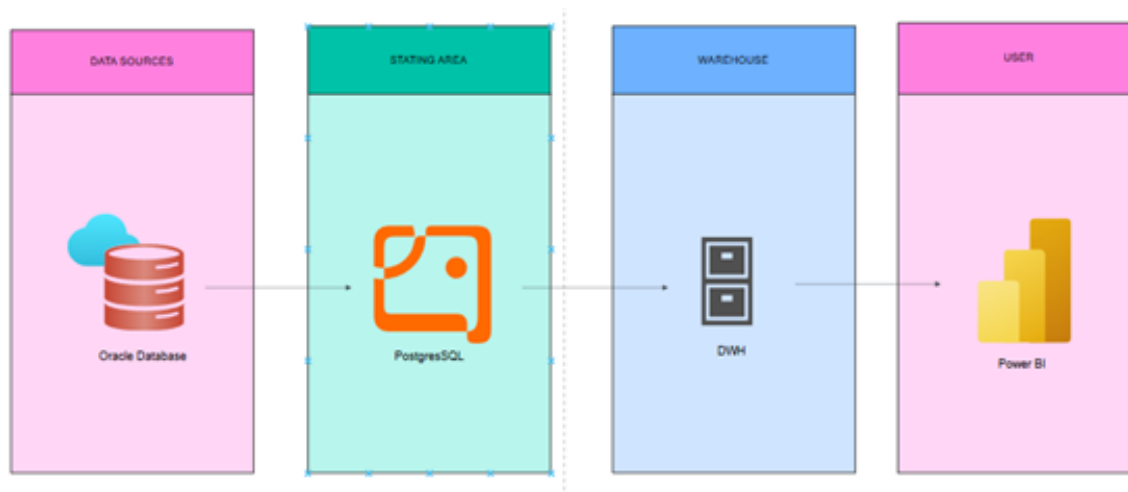


Figura 3: Ambiente de Producción

Fuente: Oracle XE + archivos CSV (datos adicionales).

Recursos: 16GB RAM, 200GB SSD.

Staging: PostgreSQL en Docker (para limpieza con Pentaho).

Target: DWH dimensional: 50GB en tablespace DWH (modelo estrella)

BI: Power BI Desktop conectado via ODBC

Granularidad: Por cada combinación de centro de costo, colaborador, cuenta contable, asociado (droguería), producto, ingreso, fecha y regional, se almacena el valor distribuido de los gastos, unidades vendidas, costos, ingresos y otros indicadores relacionados con la operación y asignación de recursos.

■ Medidas

FACT_GASTOS	Tipo	Descripción
Gastos	Medida	Representa el total de gastos registrados en un periodo determinado. Puede incluir gastos operativos, personal, administrativos, logísticos, etc.
Estructura gastos	Medida	Indica cómo se distribuyen o clasifican los gastos dentro de diferentes categorías o centros de costos. Teniendo en cuenta alguno de los indicadores que como las ventas, los asociados, etc.
Valor gasto distribuido	Medida	Monto de gasto asignado proporcionalmente entre distintas áreas, productos o unidades según criterios definidos.
Unidades vendidas	Medida	Total de unidades de productos o servicios vendidos durante el periodo de análisis.
Costo de venta	Medida	Representa el costo directo asociado a la producción o adquisición de los bienes vendidos
Venta	Medida	Total de ingresos generados por ventas brutas sin considerar devoluciones o descuentos.
Devolución	Medida	Monto o número de productos devueltos por los clientes, lo cual afecta negativamente los ingresos.
Teleferia	Medida	Descuento al asociado, porcentaje para alcanzar el precio del mercado, entre otros
Cantidad asociados	Medida	Número de asociados por regional, territorio, etc.
Cantidad droguerías	Medida	Total de droguerías
Cantidad empleados	Medida	Número total de empleados registrados en la organización durante el periodo de análisis.
Ingresos	Medida	Otros ingresos

Figura 6: Descripción medidas

■ Dimensiones

Dimensión	Descripción
DIM_PERSONAL	Contiene información detallada del personal, nombres, cargos, áreas de trabajo y otros atributos relevantes para el análisis del recurso humano.
DIM_ASOCIADO	Almacena los datos de los asociados, preasociados o terceros, incluye identificadores, nombres, tipo de asociado, fecha de vinculación y características demográficas o comerciales.
DIM_CUENTA	Describe las cuentas contables o financieras utilizadas para clasificar ingresos, gastos y transacciones. Incluye códigos de cuenta, descripciones, categorías contables y naturaleza
DIM_TIEMPO	Proporciona el contexto temporal para el análisis, desglosando fechas en componentes como año, mes día
DIM_PRODUCTO	Contiene la información descriptiva de los productos ofrecidos incluye código de producto, nombre, categoría, entre otros.
DIM_REGIONAL	Representa la estructura geográfica o regional. Facilita el análisis territorial o por áreas de cobertura.
DIM_INGRESO	Describe las diferentes fuentes o tipos de ingreso registrados, permitiendo clasificar y analizar los ingresos por origen (por ejemplo: venta directa, convenios, aportes, etc.).
DIM_CENTRO_DE_COSTO	Agrupar los diferentes centros de costo de la organización, identificando unidades responsables del consumo de recursos. Incluye códigos, gerencia y posibles responsables.

Figura 7: Descripción dimensiones

■ Modelo logico

Este modelo dimensional ha sido diseñado con el propósito de apoyar el proceso de distribución de gastos operativos desde la sede principal hacia los diferentes centros de distribución de las regionales (CEDIS). Su estructura permite analizar de manera detallada cómo se asignan los gastos en función de múltiples variables críticas para la organización.

La tabla de hechos central, FACTGASTOS, recopila información sobre montos de gasto, ventas, costos, devoluciones, ingresos y otras métricas clave, y se encuentra relacionada con diversas dimensiones. Este modelo se convierte en una herramienta clave para el análisis financiero, el control de costos y la toma de decisiones gerenciales.

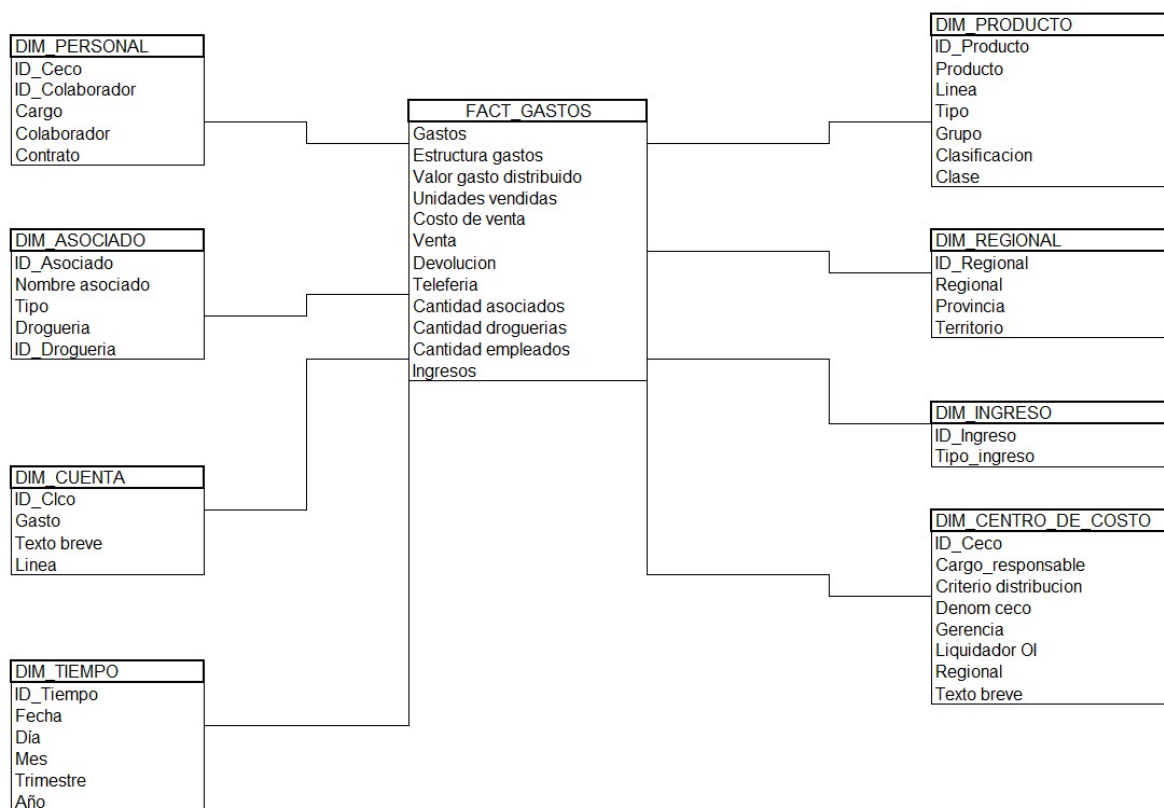


Figura 8: Modelo dimensional

■ Mapeo fuente-destino

En el Anexo, se encuentra el proceso de integración de datos desde la fuente al modelo dimensional, lo cual es necesario para transferir, consumir, procesar y gestionar los datos con el fin de cargar de manera correcta el modelo. En el anexo, en la hoja “MAPEO”, se encuentra cada tabla con el campo origen, destino y transformación y/o observaciones al respecto.

■ Creación del reporte

El propósito del reporte visual es el de describir de qué forma se presentarán la línea gráfica y el estilo visual de los dashboards y reportes generados a partir de las consultas realizadas usando como fuente la bodega de datos. Ahora, al revisar el alcance de este documento, se deduce que abarca todos los elementos visuales de la solución de BI (Business Intelligence), incluyendo las vistas, los indicadores, los reportes creados a través de las consultas y el tipo de presentaciones que se establezcan para los usuarios y/o la gerencia. A continuación, algunos indicadores a tener en cuenta para el reporte:

Gasto total distribuido: Suma de todos los gastos asignados a cada CEDIS.

Gasto por centro de costo (CECO): Gasto total distribuido por cada CECO.

Gasto por cuenta contable (CLCO): Agrupación y total de gastos según cuentas contables.

porcentaje participación del gasto por regional: Porcentaje del gasto total que representa cada región.

Gasto logístico por CEDIS: Gasto total relacionado con distribución por centro.

porcentaje del gasto de sede principal asignado por región : Qué proporción del gasto central fue absorbido por cada región.

7.3. Estándares de la solución

Para los estándares de la solución del proyecto de la Bodega de datos para COOPIDROGAS, se revisó la normativa para los nombres de las tablas, la documentación requerida por la empresa para solicitar la aprobación e inicio de la fase I del proyecto y los lineamientos visuales de los datos y reportes

■ Normativa para los nombres de tablas y columnas

En este caso, la bodega llevará las tablas con la nomenclatura “DIM” o “FACT” esto significa Dimensión, e irá inscrita antes del nombre de la tabla, la cual es traída desde la fuente de la regional de costos. El nombre de cada se verá de la siguiente forma “DIMNombreTabla” o “FACTGASTOS” Los campos que se muestran en las tablas dan forma a las columnas que llevarán los datos sobre los gastos regionales, los nombres de cada ciudad o región, gastos operativos, costos del proveedor y otros gastos agregados, entre otros. Estos datos irán en mayúscula, comprenderán letras y números, todo unido, para que se puedan leer y entender fácilmente, por ejemplo: “GASTOREGIONAL”.

■ Documentación administrativa requerida para el proyecto

COOPIDROGAS como empresa constituida y con un organigrama de tipo jerárquico, sigue unos lineamientos para la toma de decisiones y la aprobación de los proyectos. Para poder iniciar la fase 1 del proyecto de bodega de datos para el área de costos, se debe presentar una carta de solicitud, junto con la definición de la problemática, a la gerencia del área implicada. Estos documentos, se compartirán con la gerencia general, quien dará el aval para iniciar con el proyecto y las entrevistas a los usuarios. Entre la fase 2 y 3 del proyecto, se debe realizar una presentación de tipo dinámica, a la gerencia general y del área, sobre los resultados del estudio y el prototipo de la solución esperada, En caso de ser aprobado este prototipo, se debe mandar el Formato de Excepciones en el presupuesto general año 2025, para hacer la solicitud de los fondos del proyecto, Según la urgencia y las necesidades que se resolverán con el proyecto, se pide al área contable y financiera, el aval para el desembolso de los fondos, dando lugar a la ejecución del proyecto y evaluación de este.

■ Lineamientos visuales de los datos y reportes

El propósito de los lineamientos es el describir de qué forma se presentarán la línea gráfica y el estilo visual de los dashboards y reportes generados a partir de las consultas realizadas usando como fuente la bodega de datos. Ahora, si se revisa el tema del alcance de este documento, se puede deducir que abarca todos los elementos visuales de la solución de BI (Business Intelligence), incluyendo las vistas, los reportes creados a través de las consultas y el tipo de presentaciones que se establezcan para los usuarios y/o la gerencia.

■ Tipografía de la base de datos

Para la fuente principal de la bodega de datos, se estableció del tipo Arial. Por consiguiente, los tamaños de fuente se ejecutarán de la siguiente manera:

1. Los Títulos tendrán un tamaño de 18–24 pt
 2. Los Subtítulos tendrán un tamaño de: 14–16 pt
 3. El texto general se realizará en tamaño de: 12 pt
 4. Las notas o pies de página para las presentaciones y reportes serán de: 8 pt
- Como regla de estilos se permite en la bodega de datos:

5. La Negrita para indicar los títulos.
6. Se debe evitar el subrayado en todo los reportes o presentaciones.

■ Colores corporativos

Paleta de colores corporativos: Los colores corporativos de COOPIDROGAS, usados para esta bodega de datos son el azul, el rojo carmesí, el blanco, el amarillo y el naranja.



Figura 9: Paleta de colores 1

Ahora, la paleta de colores para las alertas que lleva la bodega de datos en sus dashboards o reportes, sirven como herramienta de monitoreo, que indican el estado de eventos como la claidad de los datos, las actualizaciones de ETL, entre otros

Se usa una categorización semaforica:



Figura 10: Paleta de colores 2

Matriz de colores de riesgo

Color	Significado	Ejemplo de mensaje de la alerta
Verde	Sin riesgo	Carga ETL completada, datos cargados correctamente
Amarillo	Riesgo moderado	Proceso de carga lento, pero completado/ Existen varios valores nulos en los datos cargados.
Rojo	Error - Riesgo crítico	Fallo del proceso de ETL/ datos incoherentes/ Sin internet o perdida de conexión a los datos.

Figura 11: Matriz de colores de riesgo

Codificación paleta de colores para la bodega de datos:

Uso	Color	Código Hex
Fondo	Blanco	#FEFEFE
Texto	Azul	#003459
Logo	Rojo carmesí	#ED624E
Logo	Naranja	#FF7043
Advertencia y logo	Amarillo	#FFC22E
Error	Rojo	#FF6016
Éxito	Verde	#316C40

Figura 12: Codificación paleta de colores

7.4. Condiciones de diseño, propuesta de implementación y evaluación

▪ Selección e implementación

Para la selección e implementación del motor de bases de datos, el software de integración y el de visualización Front End, que se usarán en el proyecto. Se decidió realizar tres fichas técnicas comparativas [24] [25] [26], cada una entre dos ofertas, mostrando el programa o software, las ventajas y características de cada uno y entregando una conclusión de la opción escogida. (Revisar Anexo, hoja MOTORESBD)

Al revisar ambos motores de bases de datos, se puede apreciar que el proyecto necesita un programa robusto, con sistema OLAP integrado, que tenga una fácil integración y manejo para las consultas a través de SQL. Este motor debe tener escalabilidad y una rápida respuesta frente a millones de datos por analizar. Por tema de facilidad de lectura priorizamos el uso de vistas materializadas y la posibilidad de construir con ROLAP dimensionados. No obstante, ya que motor debe tener compatibilidad con Windows, acorde con los activos o equipos se decide que la mejor opción de motor en cuanto a costos, integración, funciones que brinda desde la versión estándar y demás es: **ORACLE AUTONOMOUS DATABASE con Oracle OLAP**.

▪ Ficha técnica comparativa entre las ofertas de software de integración de datos.

Al revisar ambos softwares de ETL (Revisar Anexo, hoja SWETL) [27], el proyecto necesita una herramienta robusta, de fácil manejo y acceso. Esta debe ir a la par con el motor de base de datos Oracle DB con OLAP. Por lo cual debe ser segura en su uso y debe tener un factor de escalabilidad con el fin de que no se vuelva lento o se perjudique su funcionamiento al momento de procesar y transformar un gran volumen de datos. Al tener dos opciones similares, que son compatibles con el motor de Oracle, el proyecto opta por tomar la opción que tiene una mejor

Diseño de una Bodega de Datos para la automatización de la asignación de gastos a regionales mediante BI en COOPIDROGAS

interfaz, con funciones de transformations (transformación de flujo de datos) y Jobs (ejecutar tareas), la cual es: **Pentaho Data Integration o KETTLE**.

■ Ficha técnica comparativa entre dos herramientas de visualización Front End

Al revisar ambas herramientas de visualización Front End (Revisar Anexo, hoja VISUALIZACION)[28], se puede deducir que los requisitos se basan en ser intuitivos y que pueda manejar un gran volumen de datos, que logre varios tipos de visualización, facilidad de análisis y programación para las consultas más complejas. Es por ello que se opta por tomar la opción de: **Power BI PRO**.

Gastos		Actual	Budget	Variación	% Ejecución	Año anterior	Var Año anterior	% Participación Ventas
		Ventas	1.444.761.361	1.458.172.222	-13.410.861	99,1 %	1.290.083.399	12,0 %
		Gastos	124.964.999	137.796.822	12.831.824	90,7 %	117.657.401	6,2 %
Mes								
Ene	Feb	Mar	Abr	May	Jun	Jul	Ago	Sep
Oct	Nov	Dic						
Ejecución		Gerencia				Gasto		
		Administrativa	Desarrollo Empresarial	General	Jurídica	Tecnología y Proyectos	Actual	Budget
Mayor a ventas							49.582.403	52.528.453
Menor a ventas							767.070	959.193
							944.151	1.200.936
							19.892.819	19.243.303
							3.657.826	3.646.423
							4.062.141	4.465.429
							606.344	624.801
							12.464.933	12.886.074
							5.096.745	4.925.124
							175.073	168.591
							451.949	654.769
							422.389	334.998
							47.749	56.180
							904.819	987.063
							293.374	286.317
							5.072.834	5.473.032
							4.636.407	7.157.339
							36.933	20.158
							463.616	527.712
							2.629.010	2.803.835
							116.520.417	128.325.415
							8.444.581	9.471.407
							124.964.999	137.796.822
							12.831.824	90,7 %
							117.657.401	6,2 %

Figura 13: Ejemplo Mockup

8. Planeación del trabajo

Para realizar el presente proyecto, se hace uso de la metodología de investigación de tipo aplicada ya que se busca resolver un problema concreto dentro de la organización COOPIDROGAS, mediante la implementación de una herramienta de Datawarehouse, diseñado una bodega de datos bajo unos requerimientos puntuales exigidos por el negocio y los estándares de calidad y privacidad de los datos. Este tipo de investigación posee un enfoque mixto en materia de métodos de análisis al hacer uso de métodos cualitativos para el análisis de los requerimientos de la empresa y los usuarios combinados con métodos cuantitativos para su validación y evaluación.

Este enfoque está estructurado en fases, cada una corresponde a un objetivo específico del proyecto, de esta manera se organiza de una manera lógica y secuencial el desarrollo de la bodega de datos, lo que nos facilita su diseño, implementación y validación.

8.1. Fase 1. Diagnóstico del proceso actual y definición de requerimientos

Esta fase preliminar tiene como objetivo comprender el proceso actual de costeo en la cooperativa, identificar las necesidades del cliente interno y establecer los lineamientos iniciales del proyecto. Se hace un diagnóstico del proceso actual de costos: Revisión de la metodología ABC utilizada actualmente y evaluación de la productividad. Para esto es necesario realizar un análisis documental y análisis de datos históricos cargados en SAP.

Posteriormente, se define el alcance del proyecto definiendo y especificando las preguntas de negocio que dan origen al proyecto. A partir de ello, se plantea una solución basada en las necesidades y se presenta la idea ante la dirección de costos y gerencia financiera. Para poder entender el proyecto, se debe establecer los factores clave del negocio que se quieren plasmar en la bodega de datos, para ello se debe hacer uso de herramientas como:

Matriz de requerimientos en donde se presentan los diferentes elementos de información necesarios para responder a los requerimientos, estos elementos están organizados en tres categorías: atributos, indicadores y medidas. Los atributos son datos descriptivos utilizados para segmentar o clasificar la información, los indicadores representan valores acumulados relevantes para el análisis y solución de los requerimientos. Por último, las medidas son métricas calculadas que permiten realizar análisis comparativos o proporcionales. Esta matriz facilita la identificación de los elementos clave requeridos por cada requerimiento y sirve como guía para el diseño del modelo de datos e implementación de reportes orientados a la toma de decisiones. Se pueden llegar a realizar una o varias matrices hasta lograr la que mejor se acople a lo requerido.

Modelo de alto nivel en este modelo se representa de forma abstracta y sin especificaciones técnicas o físicas el diseño de la bodega de datos, en donde se organizan y relacionan los datos a gran escala.

Finalmente se debe realizar un **análisis de fuente** con el fin de revisar la documentación existente, tiene como propósito identificar y documentar los elementos clave de los datos operativos y financieros que alimentarán la bodega de datos. Identificando fuentes de datos internas en SAP y/o Excel.

8.2. Fase II. Diseño del modelo dimensional y físico

En esta fase se define la estructura lógica de la bodega de datos, estableciendo los hechos y dimensiones de acuerdo con el enfoque más adecuado (modelo estrella o copo de nieve), el tema de análisis y el nivel de granularidad, es decir, el detalle con el que se registrarán los hechos. El modelo dimensional se diseña con el objetivo de representar el negocio de forma intuitiva, facilitando el análisis por parte de los usuarios finales. Este modelo debe ser validado con los usuarios del negocio antes de avanzar, para garantizar que responde a sus necesidades analíticas.

Una vez validado, se diseña el modelo físico, donde se especifican los detalles necesarios para la implementación en un sistema gestor de bases de datos, incluyendo tipos de datos específicos, índices, particiones, estrategias de almacenamiento y políticas de mantenimiento, todo orientado a garantizar un alto rendimiento y facilidad de administración.

Durante esta fase, se recomienda el uso de herramientas como Lucidchart, dbdiagram, Oracle SQL Data Modeler, entre otras, para representar gráficamente los modelos. La correcta transición desde el modelo dimensional hasta el físico es esencial para asegurar una bodega de datos eficiente, escalable y alineada con los objetivos del negocio.

8.3. Fase III. Desarrollo del proceso ETL (Extracción, transformación y carga)

Se construye el flujo que permite integrar datos desde diversas fuentes hacia el modelo dimensional, garantizando su calidad, extracción de datos desde fuentes identificadas, aplicación de reglas de transformación, carga de datos al esquema definido, pruebas de calidad e integridad. el diseño del proceso de carga debe facilitar su posterior implementación y garantizar al cliente que sea consistente con las reglas de negocio de la organización, esto nos lleva a indicar que conceptos se tendrán en cuenta para el desarrollo de la ETL:

Fuentes: Qué tablas se utilizan para cargar cada objeto

Relaciones: Cuáles son las relaciones de dichas tablas

Mapeos: Cómo deben poblarse atributos y dimensiones a partir de los campos de las tablas fuente

Flujos de datos: Cómo debe poblarse la totalidad del modelo a partir de las tablas fuente

Flujos de proceso: En qué orden deberían cargarse las tablas del modelo

Estrategia de carga: Con qué técnica se debe introducir nueva información en las tablas

Periodicidad: Con qué regularidad debería actualizarse cada tabla del modelo

Manejo de nulos y errores: Qué comportamiento tiene el modelo ante valores inesperados

Se puede hacer uso de diferentes herramientas como Pentaho, Talend, Python, SQL Server Integration Services (SSIS).

8.4. Fase IV: Selección y justificación del sistema de gestión de base de datos (DBMS)

Se comparan diferentes tecnologías para seleccionar la más adecuada para la implementación del modelo garantizando la compatibilidad con el entorno tecnológico existente en la cooperativa.

Se deben **definir los criterios de evaluación** en donde se establecen factores clave para la toma de decisión, tales como rendimiento en la ejecución de consultas complejas, escalabilidad

para el crecimiento futuro, compatibilidad con plataformas, disponibilidad de soporte técnico, y costos asociados de licenciamiento e infraestructura.

Posteriormente se debe realizar una **evaluación comparativa de tecnologías** lo ideal es evaluar por lo menos tres alternativas tecnológicas viables y con ello hacer una matriz de decisión que califique cada opción en función de los criterios definidos. Esta evaluación se complementa con revisión bibliográfica, estudios técnicos y estudios recientes disponibles. Con base en los resultados de la evaluación, se selecciona la alternativa que mejor se ajuste a los objetivos.

8.5. Fase V: Validación del modelo mediante herramientas de visualización

Esta fase tiene como objetivo evaluar la efectividad del modelo dimensional implementado en la bodega de datos a través de su integración con herramientas de inteligencia de negocios (BI), asegurando que la información sea útil y fácilmente interpretable para la toma de decisiones gerenciales.

Se establecen conexiones entre la bodega de datos y plataformas de visualización como Power BI o Tableau, seleccionadas por su capacidad de integración y facilidad de uso para los analistas de la cooperativa. Allí se diseñan **dashboards** y se construyen visualizaciones que representen los procesos clave definidos durante el diseño del modelo, después se realiza una presentación para que los usuarios finales puedan revisar y aprobar, para obtener una **retroalimentación y ajustes**.

Esta fase permite comprobar la funcionalidad del modelo en un entorno operativo real, asegurando que los datos integrados y transformados sean aprovechables desde una perspectiva gerencial. Además, fortalece la confianza de los usuarios en la solución, incrementando su adopción y efectividad dentro de la cooperativa.

8.6. Descomposición de actividades WBS

Para poder descomponer el proyecto del diseño de una Bodega de Datos para la automatización de la asignación de gastos a regionales mediante BI en COOPIDROGAS, la directiva, autorizó al líder del proyecto a realizar un esquema WBS o Work Breakdown Structure, siendo este esquema mucho más fácil de entender y mucho más sencillo de explicar a los encargados de la toma de decisiones dentro de la cooperativa.

Al iniciar cada fase o rama de “Tasks”, se encontrará el recuadro del “Work Package”. Este funciona como cuadro resumen, donde se ve el resultado esperado para cada fase, el tiempo que se destinó para su elaboración, costos o monto utilizado desde el presupuesto y la o las personas encargadas de realizar la fase.

Siguiendo con este planteamiento, el WBS ejecutado muestra un modelo de 5 fases, el cual inicia con el planteamiento de los requerimientos, y, culmina con la implementación y pruebas de funcionamiento del Datawarehouse. Cada fase contiene unas subdivisiones específicas, llamadas actividades o “Tasks”. Estas se ven representada por dos condiciones; Con el color azul, cuando son lineales y dependen de que la anterior actividad se culmine para poder ejecutar la siguiente. O, por el contrario, si son de color naranja, significa que la actividad 1 y la actividad 2 (por ejemplo), pueden ejecutarse de manera paralela.

Para terminar con la explicación del esquema WBS, se debe aclarar que COOPIDROGAS, exige que sus proyectos preexistentes se puedan medir en base a los entregables que vayan

realizándose, según los tiempos establecidos en el diseño del proyecto. No obstante, ya que este proyecto es de nueva generación (proyecto innovador, que trae una mejora nunca vista dentro del área en la que se ejecuta), recibió el aval de la dirección del área de costos, para que se maneje por fases, al tener que ser diseñado, probado y ejecutado, en el menor tiempo posible.

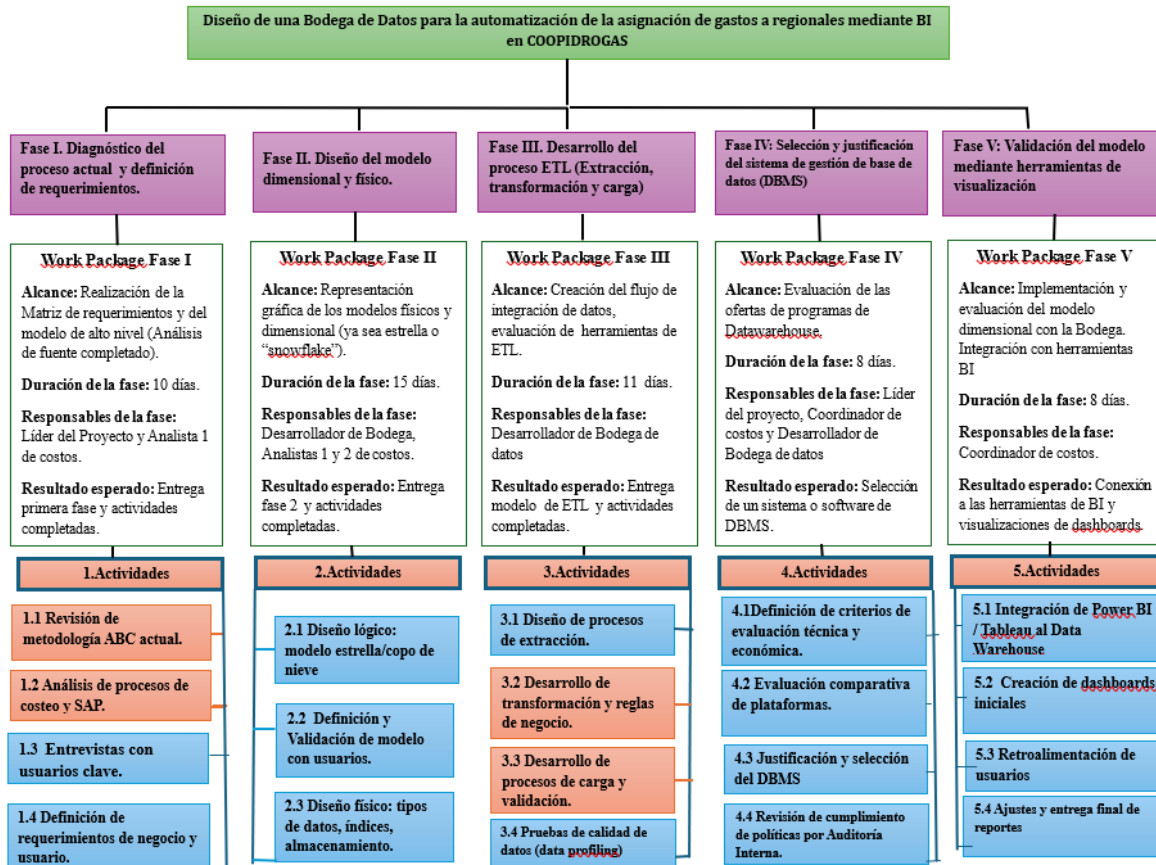


Figura 14: Esquema Work Breakdown Structure oWBS

8.7. Cronograma

A continuación, se presentan las actividades generales a realizar por cada fase del proyecto y su duración estimada:

Diseño de una Bodega de Datos para la automatización de la asignación de gastos a regionales mediante BI en COOPIDROGAS

Fase/ Entregable	Código	Actividad/ Subactividad	Responsable(s)	h	Total	Entregable	Precedencia
Fase I. Diagnóstico y requerimientos	1.1	Revisión metodología actual de costeo ABC y procesos en SAP	Todo el equipo	16	\$ 2.160.000	Informe técnico diagnóstico	Inicio (no depende de otras)
	1.2	Análisis documental histórico (planos de costos, Excel, SAP)	Analista del área de costos	20	\$ 300.000	Matriz de análisis de entradas	Depende de 1.1
	1.3	Entrevistas con usuarios clave (áreas de costos, presupuesto, TI)	Administrador de la Bodega de Datos - Coordinador del área de costos	24	\$ 1.560.000	Actas y requerimientos de usuario	Puede ejecutarse en paralelo con 1.2
	1.4	Definición de requerimientos del negocio y usuario final	Coordinador del área de costos - Analista del área de costos	24	\$ 1.080.000	Documento de requerimientos funcionales	Depende de 1.2 y 1.3
	1.5	Elaboración de matriz de requerimientos (atributos, indicadores, medidas)	Desarrollador de Base de Datos - Analista del área de costos	16	\$ 560.000	Matriz de requerimientos	Depende de 1.4
	1.6	Diseño del modelo de alto nivel	Desarrollador de Base de Datos - Administrador de la Bodega de Datos	24	\$ 1.320.000	Diagrama conceptual	Depende de 1.5
	1.7	Aprobación del diagnóstico por la Gerencia de Costos	Coordinador del área de costos	8	\$ 240.000		Depende de 1.6
Fase II. Diseño del modelo de datos	2.1	Diseño del modelo dimensional (estrella o copo de nieve)	Desarrollador de Base de Datos	32	\$ 640.000	Diagrama ER dimensional	Depende de 1.7
	2.2	Validación del modelo con usuarios clave	Administrador de la Bodega de Datos	16	\$ 560.000	Acta de validación	Depende de 2.1
	2.3	Diseño físico (tipos de datos, índices, normalización)	Administrador de la Bodega de Datos - Desarrollador de Base de Datos	24	\$ 1.320.000	Documento técnico de diseño físico	Depende de 2.2
	2.4	Selección de herramientas de modelado (Lucidchart, SQL Modeler, etc.)	Desarrollador de Base de Datos	8	\$ 160.000	Comparativa de herramientas	Puede ejecutarse en paralelo con 2.3
	2.5	Documentación técnica del modelo	Administrador de la Bodega de Datos	16	\$ 560.000		Depende de 2.3 y 2.4
Fase III. Desarrollo de procesos ETL	3.1	Identificación de fuentes de datos y extracción desde SAP y Excel	Desarrollador de Base de Datos - Analista del área de costos	24	\$ 840.000	Documento de mapeo de fuentes	Depende de 2.5
	3.2	Diseño de reglas de transformación y flujos	Desarrollador de Base de Datos	24	\$ 480.000	Diagrama de flujos de transformación	Depende de 3.1
	3.3	Desarrollo de scripts de carga al Data Warehouse	Desarrollador de Base de Datos	40	\$ 800.000	Scripts SQL / Talend / SSIS	Depende de 3.2
	3.4	Puebas de calidad de datos (data profiling, manejo de nulos y errores)	Desarrollador de Base de Datos - Auditor de datos	24	\$ 960.000	Reporte de calidad de datos	Depende de 3.3
	3.5	Documentación del flujo ETL completo	Desarrollador de Base de Datos	16	\$ 320.000	Manual técnico ETL	Depende de 3.4
	3.6	Manual técnico y operativo del proceso ETL	Desarrollador de Base de Datos - Coordinador del área de costos	12	\$ 600.000		Depende de 3.5

Figura 15: Actividades fragmento 1

Fase/ Entregable	Código	Actividad/ Subactividad	Responsable(s)	h	Total	Entregable	Precedencia
Fase IV. Selección y justificación DBMS	4.1	Definición de criterios de evaluación técnica y económica	Administrador de la Bodega de Datos - Coordinador del área de costos	8	\$ 520.000	Lista de criterios	Puede iniciar en paralelo con fase II o III
	4.2	Evaluación comparativa de plataformas (SAP BW, SQL Server, PostgreSQL, etc.)	Administrador de la Bodega de Datos - Desarrollador de Base de Datos	24	\$ 1.320.000	Matriz comparativa DBMS	Depende de 4.1
	4.3	Justificación y selección del DBMS	Coordinador del área de costos - Analista del área de costos	12	\$ 540.000	Documento de selección	Depende de 4.2
	4.4	Revisión de cumplimiento de políticas por Auditoría Interna	Auditor de datos	8	\$ 160.000		Depende de 4.3
Fase V. Validación con BI	5.1	Integración de Power BI / Tableau al Data Warehouse	Administrador de la Bodega de Datos - Analista BI	24	\$ 1.200.000	Conexiones BI configuradas	Depende de 3.6 y 4.4
	5.2	Creación de dashboards iniciales (KPIs, inductores, gastos por regional)	Analista BI	32	\$ 480.000	Dashboard BI prototipo	Depende de 5.1
	5.3	Revisión y retroalimentación de usuarios	Todo el equipo - Auditor de datos	16	\$ 2.480.000	Actas de retroalimentación	Depende de 5.2
	5.4	Ajustes y entrega final de reportes	Desarrollador de Base de Datos - Administrador de la Bodega de Datos	16	\$ 880.000	Dashboards definitivos	Depende de 5.3
	5.5	Manual de uso de dashboards y visualizaciones	Analista del área de costos - Administrador de la Bodega de Datos	12	\$ 600.000		Depende de 5.4
Fase VI. Cierre y entrega final	6.1	Presentación del proyecto y capacitación del personal	Todo el equipo - Auditor de datos	16	\$ 2.480.000	Informe final + Actas + Manuales	Depende de 5.5
					\$ 25.120.000		

Figura 16: Actividades fragmento 2

Con las actividades ya planteadas, se realiza el diagrama de gantt, el cual tambien esta dividido en dos fragmentos:

Diseño de una Bodega de Datos para la automatización de la asignación de gastos a regionales mediante BI en COOPIDROGAS

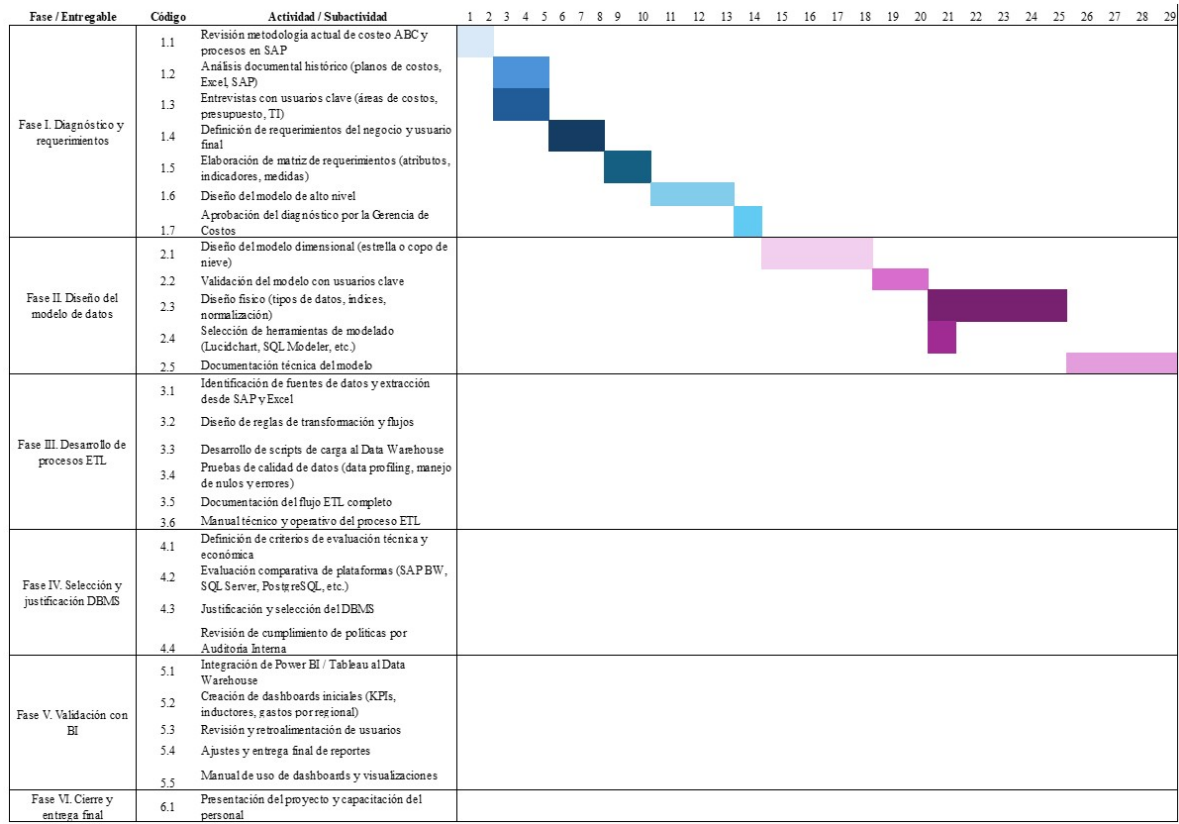


Figura 17: Diagrama de Gantt fragmento 1

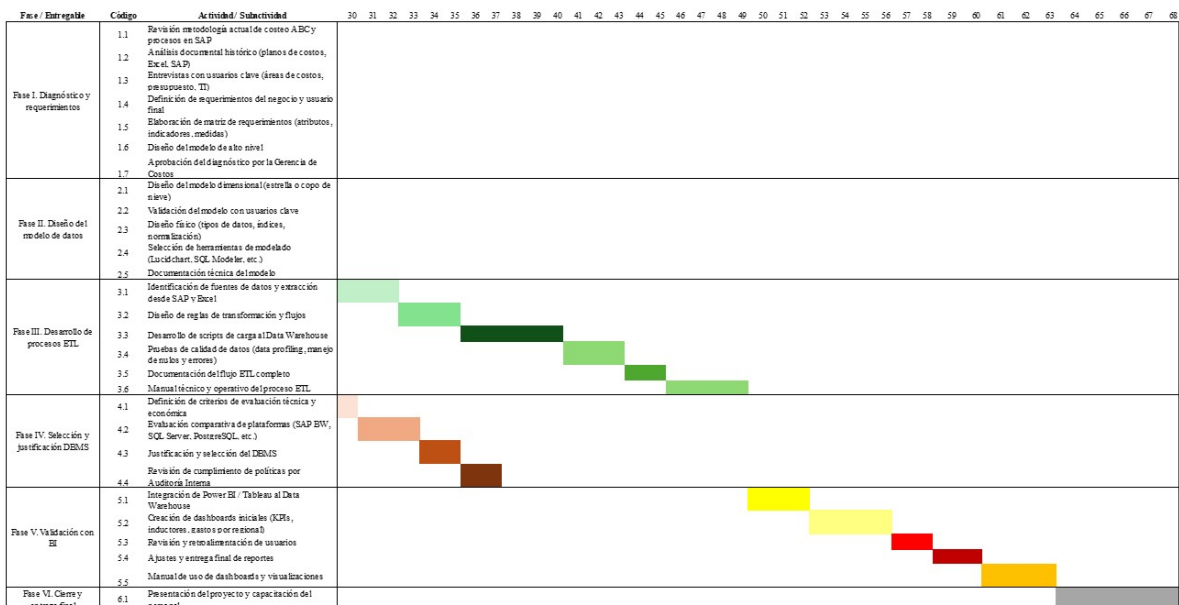


Figura 18: Diagrama de Gantt fragmento 2

9. Presupuesto

Presupuesto de nómina por semana del Proyecto Bodega de datos para el área de costos para COOPIDROGAS.				
Cargo / ROL	Q h/sem	Precio Unitario (COP/h)	Costo de contratación (carga en el primer mes del contrato)	Precio Total (COP/sem)
Cargos relacionados al funcionamiento de la bodega de datos.				
Administrador de la Bodega de datos como Líder del proyecto.	20	\$ 35.000,00	\$ 155.696,00	\$ 700.000,00
Administrador de la Bodega de datos como DBA de COOPIDROGAS	22	\$ 35.000,00	\$ 155.696,00	\$ 770.000,00
Desarrollador de base de datos, como desarrollador del Datawarehouse.	25	\$ 20.000,00	\$ 102.589,00	\$ 500.000,00
Desarrollador de base de datos como encargado del mantenimiento semanal del Datawarehouse.	17	\$ 20.000,00	\$ 102.589,00	\$ 340.000,00
		SUBTOTAL	\$ 516.570,00	\$ 2.310.000,00
Área de costos.				
Analista como colaborador del área de costos	21	\$ 15.000,00	Contratado en Diciembre/2023	\$ 315.000,00
Analista como creador de estrategias en Power BI.	21	\$ 15.000,00	Contratado en Diciembre/2023	\$ 315.000,00
Analista para la revisión de los reportes diarios de las regionales de costos (A Tiempo Completo)	42	\$ 17.500,00	Contratado en Diciembre/2023	\$ 735.000,00
Coordinador área de costos como coordinador en función	21	\$ 30.000,00	Contratado en Julio/2019	\$ 630.000,00
Coordinador área de costos como encargado de las visualizaciones y reportes a la gerencia en PowerBI.	21	\$ 30.000,00	Contratado en Julio/2019	\$ 630.000,00
		SUBTOTAL	\$ -	\$ 2.625.000,00
Área de Auditoría Interna				
Controller de datos o Auditor para las normas de tratamientos de datos y revisión del proceso	42	\$ 20.000,00	\$ 102.589,00	\$ 840.000,00
		SUBTOTAL	\$ 102.589,00	\$ 840.000,00
TOTAL PRESUPUESTO DE NÓMINA SEMANAL PARA EL PROYECTO.				\$ 6.394.159

Figura 19: Presupuesto nomina

Presupuesto de activos y equipos para el Proyecto COOPIDROGRAS			
Mensualidades sobre las licencias necesarias para el funcionamiento del Datawarehouse en la nube			
ITEM	Unidades necesarias	Precio Unitario (COP)	Precio Total (COP)
Servidor Datawarehouse en la nube con Amazon Redshift RA3 con 4 nodos. Debe incluirse el Concurrency Scaling, el Redshift Spectrum para datos históricos en S3 y la integración con AWS Glue para los procesos de ETL.	RA3.4xlarge por 4 nodos	\$ 4.474.000,00	\$ 17.896.000,00
Almacenamiento de 5 TB en la nube de Amazon	5 TB	\$ 99.698,00	\$ 498.490,00
Redshift Spectrum más el Concurrency Scaling	1 licencia conjunta de ambas funcionalidades	\$ 1.645.200,00	\$ 1.645.200,00
Licencia para el ETL AWS Glue ETL Jobs (basados en Spark)	Se debe de ejecutar 1 job diario por 30 minutos, teniendo en cuenta los 5 DPU's	\$ 135.729,00	\$ 135.729,00
		SUBTOTAL	\$ 20.175.419,00
Equipos necesarios para el funcionamiento del Datawarehouse en la nube. (Estos equipos se compran solo una vez, después de recibir el aprobado de las directivas. Se deprecian en valor constante anual, según los procedimientos del área de contabilidad)			
Equipo Lenovo ThinkCentre M920Q Tiny Desktop PC	2	\$ 1.377.462,00	\$ 2.754.924,00
Equipo HP EliteDesk 800 G4 SFF Business PC	4	\$ 1.078.975,00	\$ 4.315.900,00
Licencia de uso para Qlik Sense Enterprise (SaaS o On-Premise)	2	\$ 4.037.669,00	\$ 8.075.338,00
		SUBTOTAL	\$ 15.146.162,00
TOTAL PRESUPUESTO DE LICENCIAS Y EQUIPOS			\$ 35.321.581,00

Figura 20: Presupuesto activos

10. Conclusiones

El proyecto para diseñar e implementar una bodega de datos para la empresa Coopidrogas, se constituye como una propuesta clave para cumplir con los objetivos de transformación digital e innovación dentro de esta compañía colombiana. La ejecución del proyecto en dos meses y medio, logro consolidar diversas fuentes de datos desde las regionales de gastos, en un modelo estructurado, el cual identifica las dimensiones, hechos, atributos clave e inductores, fundamentales para diseñar un modelo dimensional sólido. Este modelo trajo como consecuencia directa de su implementación, la creación de un área encargada para el manejo y mantenimiento de la bodega de datos, ligados con una nueva área de reportes, auditoria de datos y visualizaciones en Power BI PRO. Al adquirir licencias para Oracle Autonomous Database y con la compra del software de integración de datos para los proceso de ETL, Pentaho DATA, se dio inicio a una nueva rutina, donde los analistas de costos, ahora realizan proceso de BI, se adquirió un personal más calificado, acorde a las necesidades del sistema de Datawarehouse, junto con la coordinación del área, quien ahora tiene un contacto continuo con los tomadores de decisiones, siendo esta la encargada de visualizar y mostrar que la automatización del sistema de procesamiento de datos (los cuales antes eran procesos manuales en Excel), les permite acceder a informes consolidado de Power BI PRO, con paneles interactivos, los cuales facilitan el análisis de la situación actual, permitiendo crear y ejecutar con mayor facilidad y fluidez, las estrategias de costos que repercutirán en una empresa con más de 1000 droguerías asociadas a nivel nacional. La implementación de la bodega de datos en COOPIDROGAS no solo puede automatizar la distribución de gastos, sino que sentó las bases para una cultura data-driven. Con Oracle Autonomous DB y Pentaho, se garantizó escalabilidad para procesar millones de registros, mientras que Power BI empoderó a los analistas con paneles interactivos. Este proyecto, alineado con la estrategia digital de la cooperativa, abre la puerta a futuras innovaciones como modelos predictivos de costos. La inversión en tecnología y talento especializado refleja el compromiso de COOPIDROGAS con la excelencia operativa en un sector farmacéutico cada vez más competitivo

Referencias

- [1] S. Vásquez, C. Quispe, D. Gonzales y Z. Hilario, «Hechos contables en la administración financiera: un enfoque conceptual,» 2021, [Accedido: 22-mar-2025]. dirección: <https://educas.com.pe/index.php/hechoscontables/article/view/74/302>.
- [2] T. Mucci y S. Cole, *¿Qué es la automatización de procesos empresariales?* [Accedido: 22-mar-2025], 2024. dirección: <https://www.ibm.com/es-es/topics/business-process-automation>.
- [3] *Automatización de procesos: la clave para la eficiencia*, [Accedido: 22-mar-2025], 2024. dirección: <https://www.sap.com/latinamerica/products/technology-platform/process-automation/what-is-process-automation.html>.
- [4] E. V. R. Farroñan, L. A. H. Pardo y P. P. Medina, «El sistema de costos ABC como estrategia para la toma de decisiones empresarial,» *Universidad y Sociedad*, vol. 12, n.º 2, págs. 178-183, 2020. dirección: <http://scielo.sld.cu/pdf/rus/v12n2/2218-3620-rus-12-02-178.pdf>.
- [5] L. G. T. Enríquez, *Influencia de los tipos de liderazgo para la consolidación de equipos altamente productivos en Coopidrogas*. Bogotá, Colombia: Corporación Universitaria Minuto de Dios, 2021. dirección: <https://repository.uniminuto.edu/server/api/core/bitstreams/1f1211de-bd5c-4a68-b71f-b9ba47a84a50/content>.
- [6] A. F. G. Gómez y A. K. V. Montes, *Propuesta de mejora en el proceso de expedición de pedidos en el área de despacho en la empresa Coopidrogas*. Bogotá, Colombia: Universidad El Bosque, 2024. dirección: <https://repositorio.unbosque.edu.co/items/cb0a3288-5df7-4ce7-b473-6c337132c817>.
- [7] E. H. Obando, «Estrategias comerciales para el mejoramiento de la satisfacción del cliente y los indicadores de ventas de las droguerías en Pereira Phamasaludero,» Universidad Católica de Pereira, 2024. dirección: <https://repositorio.ucp.edu.co/server/api/core/bitstreams/8ccd2371-4243-4152-a13e-a8b264df3188/content>.
- [8] V. T. Restrepo, «Plan de práctica de Supertiendas y Droguerías Olímpica,» Tecnológico de Antioquia, inf. téc., 2024. dirección: <https://repositorio.tdea.edu.co/bitstream/handle/tdea/5607/Plan%20de%20pr%C3%A1ctica%20Veronica%20Toro%20Restrepo.pdf?sequence=1&isAllowed=yDspace%20TDEA>.
- [9] M. F. P. Tascón, «Construcción de una bodega de datos para analítica descriptiva y predictiva en el observatorio de estudiantes de la Universidad de los Andes,» Tesis de mtría., Universidad de los Andes, Bogotá, Colombia, 2024. dirección: <https://repositorio.uniandes.edu.co/server/api/core/bitstreams/9121cc87-f2b7-41ce-82ce-7bcadf4806f8/content>.
- [10] M. D. R. Osoreo y M. E. T. Valle, «Plan de migración de datos desde los sistemas OFISIS e INFOREST al sistema SAP S/4 HANA en una empresa del sector gastronómico,» inf. téc., 2024. dirección: https://repositorioacademico.upc.edu.pe/bitstream/handle/10757/674833/Torre_VM.pdf?sequence=1.

- [11] M. Al-Omari, M. Al-Azzam y M. Al-Azzam, «Navigating the acceptance of implementing business intelligence in organizations: A system dynamics approach,» *Heliyon*, vol. 9, n.º 3, e15514, 2023, [Accedido: 11-may-2025]. dirección: <https://www.sciencedirect.com/science/article/pii/S2772503023000300>.
- [12] A. M. Alharthi, A. Alzahrani y M. A. Alharthi, «Benefits and Challenges of Self-Service Business Intelligence Implementation: A Systematic Literature Review,» *Procedia Computer Science*, vol. 215, págs. 1234-1243, 2023, [Accedido: 11-may-2025]. dirección: <https://www.sciencedirect.com/science/article/pii/S1877050923012243>.
- [13] M. M. Al-Omari, M. A. Al-Azzam y M. A. Al-Azzam, «Determinants and consequences of routine and advanced use of business intelligence systems: An empirical study,» *Information & Management*, vol. 61, n.º 1, 2024, [Accedido: 11-may-2025]. dirección: <https://www.sciencedirect.com/science/article/pii/S0378720623001362>.
- [14] N. Xu, «Intelligent financial management system based on data mining,» en *Proc. SPIE 13090, International Conference on Computer Application and Information Security (ICCAIS 2023)*, [Accedido: 11-may-2025], Wuhan, China, 2023. dirección: <https://doi.org/10.1117/12.3025671>.
- [15] J. Azevedo, J. Duarte y M. Santos, «Implementing a business intelligence cost accounting solution in a healthcare setting,» en *Procedia Computer Science*, [Accedido: 11-may-2025], vol. 198, 2022. dirección: <https://www.sciencedirect.com/science/article/pii/S1877050921024881>.
- [16] P. K. C. Palma y J. T. H. Guamán, «Automatización de un proceso ETL para la industria farmacéutica mediante un modelo predictivo basado en Power BI,» [Accedido: 11-may-2025], Tesis de mtría., Universidad Politécnica Salesiana, Guayaquil, Ecuador, 2024. dirección: <https://dspace.ups.edu.ec/bitstream/123456789/27873/1/UPS-GT005361.pdf>.
- [17] A. Godoy, C. González, J. Maldonado y D. Castaño, «Propuesta para la implementación de una bodega de datos como fuente de información en el Banco Verde,» inf. téc., 2018, [Accedido: 22-mar-2025]. dirección: <https://alejandria.poligran.edu.co/bitstream/handle/10823/1682/Propuesta%20para%20la%20implementaci%C3%B3n%20de%20una%20bodega%20de%20datos%20como%20fuente%20de%20informaci%C3%B3n%20en%20el%20Banco%20Verde.pdf?sequence=1&isAllowed=y>.
- [18] H. Zhang, Y. Liu y J. Yan, «Cost-Intelligent Data Analytics in the Cloud,» *arXiv preprint arXiv:2308.09569*, 2023, [Accedido: 11-may-2025]. dirección: <https://arxiv.org/abs/2308.09569>.
- [19] M. Gonçalves y L. Teixeira, «Decision support system for port terminals: Design and development of a business intelligence tool,» en *Procedia Computer Science*, [Accedido: 11-may-2025], vol. 253, 2025. dirección: <https://www.sciencedirect.com/science/article/pii/S1877050925003126>.
- [20] L. H. P., «A business intelligence system,» inf. téc., 1958, [Accedido: 22-mar-2025], págs. 314-319. dirección: <http://altaplana.com/ibm-luhn58-BusinessIntelligence.pdf>.

- [21] *¿Cómo y cuándo surge el business intelligence?* [Accedido: 22-mar-2025], 2015. dirección: <https://directortic.es/tecnologia-2/como-y-cuando-surge-el-business-intelligence-2015031813316.htm>.
- [22] *Bill Inmon: Biografía*, [Accedido: 22-mar-2025], 2024. dirección: <https://datascientest.com/es/bill-inmon-biografia>.
- [23] J. Senso y A. de la Rosa Piñero, «El concepto de metadato. Algo más que descripción de recursos electrónicos,» 2003, [Accedido: 22-mar-2025]. dirección: <https://www.scielo.br/j/ci/a/ZHtZZfYnJfKqVn4tGNSw4yv/?format=pdf&lang=es>.
- [24] *Getting Started with Oracle OLAP*, [Accessed: Apr. 11, 2025], Oracle, 2024. dirección: https://docs.oracle.com/cd/E11882_01/olap.112/e17123/start.htm#OLAUG9124.
- [25] *Oracle: Visión general*, [Accessed: Apr. 11, 2025], 2024. dirección: <https://www.ibm.com/mx-es/topics/oracle>.
- [26] *ClickHouse — Open Source Columnar Database*, [Accessed: Apr. 11, 2025], 2024. dirección: <https://clickhouse.com/clickhouse>.
- [27] *Talend Open Studio*, [Accessed: Apr. 11, 2025], Talend Data Integration Products, 2024. dirección: <https://www.talend.com/products/talend-open-studio/>.
- [28] *Especificaciones técnicas del producto*, [Accessed: Apr. 11, 2025], Tableau Software, 2024. dirección: <http://tableau.com/es-es/products/techspecs>.
- [29] L. G. Z. Chamorro y A. G. O. Audante, «Dashboards en SAP Business Intelligence para la Toma de Decisiones en la Industria Automotriz: Una Revisión Sistemática,» *Epistemia Revista Científica*, vol. 9, e2863-e2863, 2025. dirección: <https://revistas.uss.edu.pe/index.php/EPT/article/download/2863/3211>.
- [30] M. F. S. Bello, «Análisis de la implementación de SAP QM en sistema de gestión de la calidad para una empresa certificada en ISO 9001: 2015,» *inf. téc.*, 2024. dirección: <https://repositorioinstitucional.buap.mx/server/api/core/bitstreams/af84d92a-1e51-4656-944e-5429aa953050/content>.
- [31] M. J. P. Guerrero y J. L. D. Ugarte, «Implementación de SAP Cloud for Customer para optimizar el proceso de gestión de ventas y servicios en una empresa industrial,» *inf. téc.*, 2024. dirección: https://repositorioacademico.upc.edu.pe/bitstream/handle/10757/674682/Paiva_GM.pdf?sequence=1.
- [32] Pérez, «Estudio sobre Inteligencia de Negocios,» *Ciencias de la Información*, vol. 10, n.º 2, págs. 50-60, 2003, [Accedido: 22-mar-2025]. dirección: <https://www.scielo.br/j/ci/a/ZHtZZfYnJfKqVn4tGNSw4yv/?format=pdf&lang=es>.
- [33] *What is a Data Warehouse?* [Accedido: 22-mar-2025], 2024. dirección: <https://www.oracle.com/co/database/what-is-a-data-warehouse/>.
- [34] «Inteligencia de Negocios,» CEPAL, *inf. téc.*, 2024, [Accedido: 22-mar-2025]. dirección: https://www.cepal.org/sites/default/files/events/files/inteligencia_de_negocios.pdf.
- [35] «IBM Luhn58 - Business Intelligence,» AltaPlana, *inf. téc.*, 2024, [Accedido: 22-mar-2025]. dirección: <http://altaplana.com/ibm-luhn58-BusinessIntelligence.pdf>.

- [36] *Historia de la Inteligencia de Negocios*, [Accedido: 22-mar-2025], 2024. dirección: <https://reddesignsystems.com/2024/05/30/historia-inteligencia-negocios/>.
- [37] *Business Intelligence*, [Accedido: 22-mar-2025], 2024. dirección: <https://www.ibm.com/mx-es/topics/business-intelligence>.
- [38] *Business Intelligence*, [Accedido: 22-mar-2025], 2024. dirección: <https://www.tableau.com/es-es/learn/articles/business-intelligence>.
- [39] R. Kimball, *Dimensional Modeling Techniques*, [Accedido: 22-mar-2025], 2024. dirección: <https://www.kimballgroup.com/data-warehouse-business-intelligence-resources/kimball-techniques/dimensional-modeling-techniques/>.
- [40] A. Abelló, J. Curto, A. Rius, M. Serra, J. Samos y J. Vidal, «Introducción al Data Warehouse,» inf. téc., 2024, [Accedido: 22-mar-2025]. dirección: https://openaccess.uoc.edu/bitstream/10609/136246/5/Disen%C2%BFo%20y%20construccio%C2%BFn%20de%20un%20almace%C2%BFn%20de%20datos_Mo%C2%BFdulo1_Introduccio%C2%BFn%20al%20Data%20Warehouse.pdf.
- [41] S. Barrero, «Propuesta de implementación de un sistema de BI predictivo basado en big data y ETL avanzado para la mejora de la toma de decisiones empresariales,» inf. téc., 2024, [Accedido: 22-mar-2025]. dirección: <https://expeditiorepositorio.utadeo.edu.co/bitstream/handle/20.500.12010/36148/Propuesta%20de%20implementaci%C3%B3n%20de%20un%20sistema%20de%20BI%20predictivo%20basado%20en%20big%20Data%20y%20ETL%20avanzado%20para%20la%20mejora%20de%20la%20toma%20de%20decisiones%20empresariales.pdf?sequence=2&isAllowed=y>.
- [42] L. Bustos, R. Moreno y D. Duque, «Modelo de una bodega de datos para el soporte a la investigación bioinformática,» 2011, [Accedido: 22-mar-2025]. dirección: <https://www.redalyc.org/pdf/849/84922625025.pdf>.
- [43] J. Rodríguez, M. Pérez y L. Gómez, «A data warehouse model for the analysis of water quality in the Bogotá River,» *Data Science Journal*, vol. 15, n.º 3, págs. 45-58, 2023, [Accedido: 22-mar-2025]. dirección: <https://doi.org/10.1016/j.ds.j.2023.03.005>.
- [44] J. M. P. Rodríguez, J. M. G. García y J. A. M. Martínez, «Sistema de soporte a la decisión para terminales portuarios: diseño y desarrollo,» en *Procedia Computer Science*, [Accedido: 11-may-2025], vol. 187, 2025, págs. 123-130. dirección: <https://www.sciencedirect.com/science/article/pii/S1877050925003126>.

A. Anexos

A.1. Anexo A: Diferentes figuras que soportan el presente proyecto