**Asignatura:** Gestión De Proyectos

**Sección:** (TI2072/IEI(1)-171-N7/D Temuco IEI)

**Nombre del académico:** Roberto Conejeros Gómez

**Nombre de los integrantes del grupo:** Fabián Quezada, Rafael Abello

**Fecha de entrega:** 22/04/25

**Gestión de Proyectos Informáticos**

**“Presente Ausente”**

Contenido

[I. Introducción 3](#_Toc202988783)

[II. Ingeniería de Datos 4](#_Toc202988784)

[a. Procesamiento de Datos y Nuevas Variables 4](#_Toc202988785)

[b. Variables creadas 4](#_Toc202988786)

[c. Proceso ETL implementado 5](#_Toc202988787)

[d. Beneficios del Proceso 6](#_Toc202988788)

[e. Visualización de nuevas columnas 6](#_Toc202988789)

[III. Visualización de datos 7](#_Toc202988790)

[a. Herramientas de Business Intelligence Implementadas 7](#_Toc202988791)

[b. Microsoft Power BI 7](#_Toc202988792)

[c. Google Looker Studio 11](#_Toc202988793)

[d. Comparación resumida de herramientas: 13](#_Toc202988794)

[IV. Redes neuronales y Chat Inteligentes 14](#_Toc202988795)

[a. Implementación del Chatbot 14](#_Toc202988796)

[b. Arquitectura del sistema 14](#_Toc202988797)

[c. Funcionalidades implementadas 15](#_Toc202988798)

[d. Experiencia de Usuario 15](#_Toc202988799)

[e. Beneficios del Chatbot 16](#_Toc202988800)

[V. Conclusiones 16](#_Toc202988801)

[VI. Anexos 18](#_Toc202988802)

[VII. Bibliografía 19](#_Toc202988803)

1. Introducción

El presente documento detalla el desarrollo e implementación de un sistema integral de análisis de mermas para supermercados, abarcando las actividades realizadas durante las semanas 9, 10 , 11 y 12 del proyecto. Este desarrollo combina técnicas avanzadas de ingeniería de datos, visualización empresarial e inteligencia artificial para proporcionar información sumamente valiosa permite comprender, analizar y gestionar las pérdidas de productos en el retail.

El desarrollo se enfoca en la creación de nuevas variables derivadas mediante procesos ETL, el desarrollo de dashboards interactivos utilizando herramientas de Business Intelligence, y la implementación de un chatbot inteligente capaz de responder preguntas de negocio sobre los datos de mermas. Esta aproximación multidisciplinaria busca transformar datos operacionales en insights accionables para la toma de decisiones estratégicas.

1. Ingeniería de Datos
   1. Procesamiento de Datos y Nuevas Variables

Como parte del proceso de ingeniería de datos, se implementó un sistema automatizado para enriquecer la base de datos existente con nuevas variables derivadas que mejoran significativamente la capacidad de análisis y predicción de mermas.

* 1. Variables creadas

Se desarrollaron nueve nuevas columnas que aportan valor analítico al dataset:

Tabla 1: Resumen de variables creadas

|  |  |
| --- | --- |
| **Columna Nueva** | **Descripción** |
| tamano\_producto | Clasifica como pequeño, mediano o grande |
| es\_a\_granel | Indica si el producto es a granel |
| nivel\_merma | Nivel general de merma (bajo/medio/alto) |
| dia\_semana | Día numérico de la semana |
| estacion | Estación del año (invierno, verano, etc) |
| tipo\_producto\_estacional | Si es típico de una temporada |
| riesgo\_perecibilidad | Riesgo según categoría del producto |
| valor\_perdida | Valor relativo de la pérdida |
| riesgo\_categoria | Clasificación de riesgo general de categoría |

* 1. Proceso ETL implementado

El proceso ETL (Extract, Transform, Load) desarrollado presenta las siguientes características:

**Extracción:** Conexión automatizada a la base de datos MariaDB para recuperar registros sin análisis previo.

**Transformación:** Utilización de la API de ChatGPT (modelo GPT-4.1-nano) para generar análisis automático de productos y crear las nuevas variables mediante procesamiento de lenguaje natural.

**Carga:** Actualización batch de la base de datos con los resultados del análisis, procesando 1000 registros simultáneamente.

Imagen 1, prompt utilizado

Texto

El contenido generado por IA puede ser incorrecto.

* 1. Beneficios del Proceso

**Automatización completa:** Análisis de miles de registros sin intervención manual

**Escalabilidad:** Capacidad de procesar grandes volúmenes de datos históricos

**Precisión:** Clasificación consistente basada en criterios predefinidos

**Eficiencia:** Procesamiento de toda la base de datos en 2 horas con un costo de $1,000 CLP

* 1. Visualización de nuevas columnas

Imagen 2, nuevas columnas agregadas

Captura de pantalla de un celular

El contenido generado por IA puede ser incorrecto.

1. Visualización de datos
   1. Herramientas de Business Intelligence Implementadas

Para la visualización de datos se implementaron dos herramientas de BI diferentes con el objetivo de comparar capacidades, flexibilidad y facilidad de uso:

* 1. Microsoft Power BI

Para la visualización de datos se implementaron dos herramientas de BI diferentes con el objetivo de comparar capacidades, flexibilidad y facilidad de uso:

**Pantalla 1: Resumen General de Mermas** esta primera pantalla está diseñada para ofrecer una visión panorámica del estado general de las pérdidas en toda la cadena. En la parte superior se presentan dos tarjetas que muestran los totales de pérdida en dinero y en unidades, expresadas en miles, permitiendo al usuario tener una idea rápida del impacto financiero.

Debajo, se incluyen dos visualizaciones complementarias: a la izquierda, un gráfico de barras que muestra los productos con mayores pérdidas económicas, permitiendo identificar rápidamente qué ítems generan más impacto; y a la derecha, un gráfico de torta que representa la proporción de motivos asociados a las mermas (interno, clientes, vencimiento y proveedor). Todo el diseño se encuentra reforzado visualmente con bordes rojos y el logotipo de Santa Isabel, para mantener coherencia estética con la empresa de ejemplo elegida.

Imagen 3, Pantalla 1; resumen general de mermas en power BI

Interfaz de usuario gráfica

El contenido generado por IA puede ser incorrecto.

El **gráfico de barras horizontales** muestra los productos con mayores pérdidas económicas. Esta visualización no solo ordena los productos por impacto financiero, sino que además es **interactiva**: al hacer clic sobre un producto específico, por ejemplo, ‘Azúcar’, todas las demás visualizaciones de la página (incluyendo las tarjetas y el gráfico de motivos) se actualizan automáticamente, mostrando la información **exclusivamente relacionada con ese producto**. Esto permite a los usuarios realizar un análisis puntual de un ítem específico y entender, por ejemplo, cuánto de la pérdida total se debe a él y cuáles fueron las causas principales asociadas.

**Pantalla 2: Detalles de Mermas por Tienda**, esta pantalla permite al usuario realizar un análisis más específico. En la parte superior se dispone un segmentador que permite seleccionar una tienda (por ejemplo: Angol, Temuco II, etc.). Al seleccionar una, todas las visualizaciones del panel se actualizan automáticamente. Inmediatamente debajo, dos tarjetas muestran el total de dinero y unidades perdidas en esa tienda.

Imagen 4, Pantalla 2; Detalles de mermas por tienda en Power BI

Interfaz de usuario gráfica

El contenido generado por IA puede ser incorrecto.

Debajo del segmentador, se presenta en primer lugar un resumen del dinero y unidades perdidas y en segundo lugar un gráfico de columnas que detalla los distintos motivos de pérdida registrados en la tienda seleccionada, ayudando a entender qué causas son más frecuentes en cada local.

Finalmente, en el lado derecho se dispone una tabla con el detalle de productos perdidos, incluyendo su categoría y nivel de riesgo, entregando un complemento informativo para el análisis puntual de los productos más afectados.

Esta estructura permite abordar tanto una lectura general como una exploración focalizada, adaptando la información al tipo de análisis requerido. Además, la navegación entre páginas se facilita mediante botones superiores, generando una experiencia de usuario fluida.

En conjunto, Power BI permitió construir un panel visualmente atractivo, interactivo y útil para explorar información compleja de manera simplificada. Esta es la primera parte del análisis, que será complementada con el uso de otra herramienta de BI con el fin de comparar capacidades, flexibilidad y facilidad de uso.

Ejemplo de que los datos son actualizados dependiendo de la tienda seleccionada :

Interfaz de usuario gráfica

El contenido generado por IA puede ser incorrecto.

**Conclusion de PowerBI**

El uso de Power BI permitió una comprensión ágil y visualmente clara de los datos de mermas. Su facilidad para integrar diferentes tipos de visualizaciones en páginas diferenciadas (visión general y análisis específico) permitió identificar tanto el panorama global como los detalles relevantes por tienda.

Power BI demostró ser una herramienta adecuada para la exploración inicial y presentación de hallazgos sobre este tipo de información operacional.

* 1. Google Looker Studio

Como segunda herramienta se ocupó looker studio, esta herramienta de Google permite crear rápidamente dashboards a partir de datos, su estructura es la siguiente:

Imagen 5. Panel General (Información de Mermas)

Interfaz de usuario gráfica, Aplicación

El contenido generado por IA puede ser incorrecto.

Este panel está diseñado para proporcionar una visión de alto nivel y un resumen conciso de las mermas registradas en los supermercados, permitiendo identificar rápidamente el impacto y las tendencias principales

Luego seubican en la parte superior KPIs que proporcionan una instantánea concisa del impacto total de las pérdidas:

* **Cantidad de Mermas (Unidades):** Muestra el volumen total de productos perdidos. (Nota: Asegúrate de que este valor sea positivo, como 15.900,9).
* **Dinero Perdido ($):** Representa el costo monetario total de las mermas. (Nota: Este valor también debe ser positivo y con formato de moneda, por ejemplo, $ 13.330.034).

Estos indicadores son esenciales para dimensionar la escala y el impacto económico general de las mermas.

**2. Tendencia de Dinero Perdido por Fecha**

Este gráfico de líneas visualiza cómo ha evolucionado el monto total de dinero perdido por mermas a lo largo del tiempo. Su propósito es identificar patrones temporales, picos o reducciones, ayudando a entender la dinámica de las pérdidas y su relación con eventos específicos o periodos.

**3. Dinero Perdido por Motivo de Merma**

Este gráfico de barras desglosa el costo monetario de las mermas según su motivo (Interno, Clientes, Vencimiento, Proveedor). Permite identificar rápidamente qué causas generan las mayores pérdidas económicas, dirigiendo las acciones correctivas hacia las áreas de mayor impacto.

**4. Gráfico de Cantidad de Valores de Pérdidas**

Este gráfico de barras muestra la frecuencia de los incidentes de merma según su nivel\_merma (alto, bajo, medio). Es útil para comprender si las mermas son más frecuentes en categorías de alto o bajo impacto, complementando el análisis monetario con una perspectiva de ocurrencia.

**Interactividad Clave:** Un **Control de Período** (fecha) permitirá filtrar todos los gráficos por un rango de tiempo específico. Además, la función de **"Filtro cruzado"** en los gráficos facilitará la exploración detallada al seleccionar un elemento (ej. un motivo de merma), actualizando el resto del panel con datos relacionados.

* 1. Comparación resumida de herramientas:

**Microsoft Power BI:**

* Superior conectividad con bases de datos
* Mayor flexibilidad en interactividad avanzada
* Funcionalidades más robustas para análisis complejo
* Mejor rendimiento con grandes volúmenes de datos

**Google Looker Studio:**

* Herramienta gratuita y de fácil acceso
* Interfaz más simple para usuarios básicos
* Limitaciones en integración con bases de datos locales
* Adecuado para dashboards sencillos y análisis rápidos

La comparación de ambas herramientas en el contexto del análisis de arroja que, para las necesidades específicas de este proyecto, Microsoft Power BI se posiciona como la herramienta más robusta y versátil. Su superioridad en la conectividad con bases de datos, la amplitud de sus funcionalidades y la riqueza de su interactividad, le permitieron manejar la complejidad del análisis de mermas de una manera más eficiente y completa.

Google Looker Studio, aunque una opción gratuita y sencilla, mostró limitaciones en su integración con bases de datos y una menor flexibilidad en la interactividad avanzada para este caso particular. Sin embargo, su simplicidad lo convierte en una opción viable para dashboards más sencillos, análisis rápidos con fuentes de datos ya preparadas, o cuando la conectividad a bases de datos locales no es un requisito prioritario.

1. Redes neuronales y Chat Inteligentes
   1. Implementación del Chatbot

Se desarrolló un chatbot inteligente utilizando Python y la API de OpenAI (GPT-3.5-turbo) para responder preguntas de negocio sobre los datos de mermas de manera natural y conversacional.

* 1. Arquitectura del sistema

La **Estructura de la Base de Datos:** La tabla mermasdb contiene 29 campos incluyendo:

* Información del producto (código, descripción, categoría, línea)
* Datos de ubicación (tienda, comuna, región, zonal)
* Información temporal (fecha, mes, año, semestre, día\_semana, estación)
* Métricas de merma (unidades, monto, porcentajes)
* Variables derivadas (tamaño\_producto, es\_a\_granel, nivel\_merma, etc.)

**Componentes del Chatbot:**

1. **Generador de Consultas SQL:** Convierte preguntas en lenguaje natural a consultas SQL optimizadas para MariaDB
2. **Ejecutor de Consultas:** Maneja la conexión a la base de datos y ejecuta las consultas generadas
3. **Generador de Respuestas:** Transforma los resultados SQL en respuestas comprensibles para usuarios de negocio
4. **Sistema de Historial:** Mantiene un registro de las últimas 5 preguntas y respuestas
   1. Funcionalidades implementadas

**Procesamiento de Lenguaje Natural:**

* Interpretación de preguntas complejas sobre mermas
* Generación automática de consultas SQL optimizadas
* Manejo de búsquedas insensibles a mayúsculas y minúsculas
* Soporte para agregaciones, agrupaciones y ordenamientos

**Ejemplos de Preguntas Soportadas:**

* "¿Cuáles son las mermas más altas por región?"
* "¿Qué productos tuvieron mayor merma en el mes de enero?"
* "¿Cuánto fue el monto total de mermas en 2023?"
* "¿En qué tienda se registró la mayor merma de unidades?"
* "¿Cuáles son las categorías con mayor riesgo de perecibilidad?"

**Características Técnicas:**

* Límite de 100 resultados por consulta para optimizar rendimiento
* Manejo de valores nulos mediante funciones SQL apropiadas
* Formato de respuestas en lenguaje empresarial sin terminología técnica
* Visualización tabular de resultados para mejor comprensión
  1. Experiencia de Usuario

**Interfaz de Conversación:**

* Entrada de preguntas en lenguaje natural
* Comandos especiales: 'ejemplos', 'historial', 'salir'
* Respuestas estructuradas y profesionales
* Manejo de errores con mensajes informativos

**Calidad de Respuestas:**

* Lenguaje claro y profesional adaptado para usuarios de negocio
* Formato monetario en pesos chilenos con separadores de miles
* Estructuración organizada de datos complejos
* Respuestas precisas sin análisis no solicitado
  1. Beneficios del Chatbot

**Accesibilidad:** Permite a usuarios no técnicos acceder a insights complejos

**Eficiencia:** Respuestas inmediatas sin necesidad de reportes programados

**Flexibilidad:** Capacidad de responder preguntas ad-hoc y específicas

**Escalabilidad:** Manejo de consultas simultáneas sin degradación del servicio

1. Conclusiones

El desarrollo del sistema integral de análisis de mermas ha demostrado la efectividad de combinar técnicas avanzadas de ingeniería de datos, visualización empresarial e inteligencia artificial para crear una solución completa de business intelligence.

**Ingeniería de Datos**

La implementación del proceso ETL automatizado ha permitido enriquecer significativamente el dataset original, creando nueve nuevas variables que mejoran la capacidad de análisis y predicción. El uso de inteligencia artificial para la generación automática de estas variables ha demostrado ser una aproximación eficiente y escalable, procesando miles de registros con consistencia y precisión.

**Visualización de Datos**

La comparación entre Microsoft Power BI y Google Looker Studio ha revelado que, para proyectos complejos que requieren conectividad robusta con bases de datos y funcionalidades avanzadas de interactividad, Power BI se posiciona como la herramienta más adecuada. Sin embargo, Looker Studio mantiene su valor como una alternativa gratuita y accesible para análisis más básicos.

**Chatbot Inteligente**

La implementación del chatbot ha democratizado el acceso a los datos de mermas, permitiendo que usuarios no técnicos puedan obtener insights valiosos mediante preguntas en lenguaje natural. La calidad de las respuestas y la capacidad de manejar consultas complejas posicionan a esta herramienta como un componente valioso para la toma de decisiones operacionales.

1. Anexos
2. Bibliografía

Microsoft Corporation. (2024). Power BI Documentation. Microsoft Learn.

Google LLC. (2024). Looker Studio Help Center. Google Cloud Documentation.

OpenAI. (2024). GPT API Documentation. OpenAI Platform.

Oracle Corporation. (2024). MySQL Connector/Python Documentation. MySQL Documentation.

Python Software Foundation. (2024). Python 3.x Documentation. Python.org.