**Proyecto Final: CURSO DATA ANALYTICS**

**Alumno:** Galo Badaró

**Comisión:** #73275

Índice

[INTRODUCCIÓN 3](#_Toc206493247)

[Objetivo General 4](#_Toc206493248)

[Hipótesis Propuesta 5](#_Toc206493249)

[Descripción de Tablas y Relaciones 5](#_Toc206493250)

[1.Tabla: stores 6](#_Toc206493251)

[2. Tabla: sales 7](#_Toc206493252)

[3. Tabla: features 7](#_Toc206493253)

[Transformaciones Realizadas 9](#_Toc206493254)

[Medidas Calculadas Generadas y sus Fórmulas: 10](#_Toc206493255)

[Descripción de las Solapas del Análisis: 12](#_Toc206493256)

[Análisis Global 12](#_Toc206493257)

[Análisis de Ventas 13](#_Toc206493258)

[Análisis Detallado 14](#_Toc206493259)

[Conclusión Final y Posibles Iniciativas a Futuro: 15](#_Toc206493260)

# 

# INTRODUCCIÓN

El presente análisis se basa en una muestra de datos históricos de ventas semanales correspondientes a 45 tiendas minoristas distribuidas en distintas regiones.

El estudio se enfoca en comprender el comportamiento de las ventas por departamento, considerando variables clave como promociones estacionales (rebajas o *markdowns*), fechas festivas y factores económicos regionales como el precio del combustible, el índice de precios al consumidor (CPI) y la tasa de desempleo.

Este conjunto de datos permite abordar con profundidad las variaciones en el consumo según la temporalidad, los efectos de las promociones antes de fechas clave (como el Super Bowl, Labor Day, Thanksgiving y Navidad), y las diferencias de comportamiento entre tiendas y tipos de departamentos. Al mismo tiempo, se exploran las implicancias que tienen estas variables en la toma de decisiones estratégicas a nivel comercial y operativo.

El dataset utilizado fue publicado en la plataforma [Kaggle](https://www.kaggle.com/datasets/manjeetsingh/retaildataset), y está compuesto por tres tablas principales: **Stores**, **Features** y **Sales**, que contienen información detallada sobre cada tienda, las condiciones externas y el historial de ventas respectivamente. Este recurso abierto ofrece una base sólida para desarrollar modelos predictivos y proponer recomendaciones orientadas a la mejora del rendimiento comercial.

## Objetivo General

Analizar el comportamiento de las ventas semanales en tiendas minoristas distribuidas en diferentes regiones, considerando el impacto de eventos especiales y rebajas promocionales.

A partir de este análisis, se busca generar conocimiento útil que permita predecir patrones de venta y proponer acciones estratégicas que optimicen el rendimiento comercial.

**Objetivos Específicos:**

* Evaluar la influencia de las semanas festivas (Super Bowl, Labor Day, Thanksgiving y Navidad) en el volumen de ventas semanales por departamento y tienda.
* Analizar la relación entre las promociones (markdowns) y el comportamiento de ventas para identificar en qué departamentos y en qué medida se ven afectados.
* Desarrollar un modelo predictivo que permita estimar las ventas por departamento en cada tienda para el año siguiente.
* Proponer recomendaciones estratégicas basadas en los patrones identificados, priorizando aquellas con mayor impacto en el negocio.

## 

## Hipótesis Propuesta

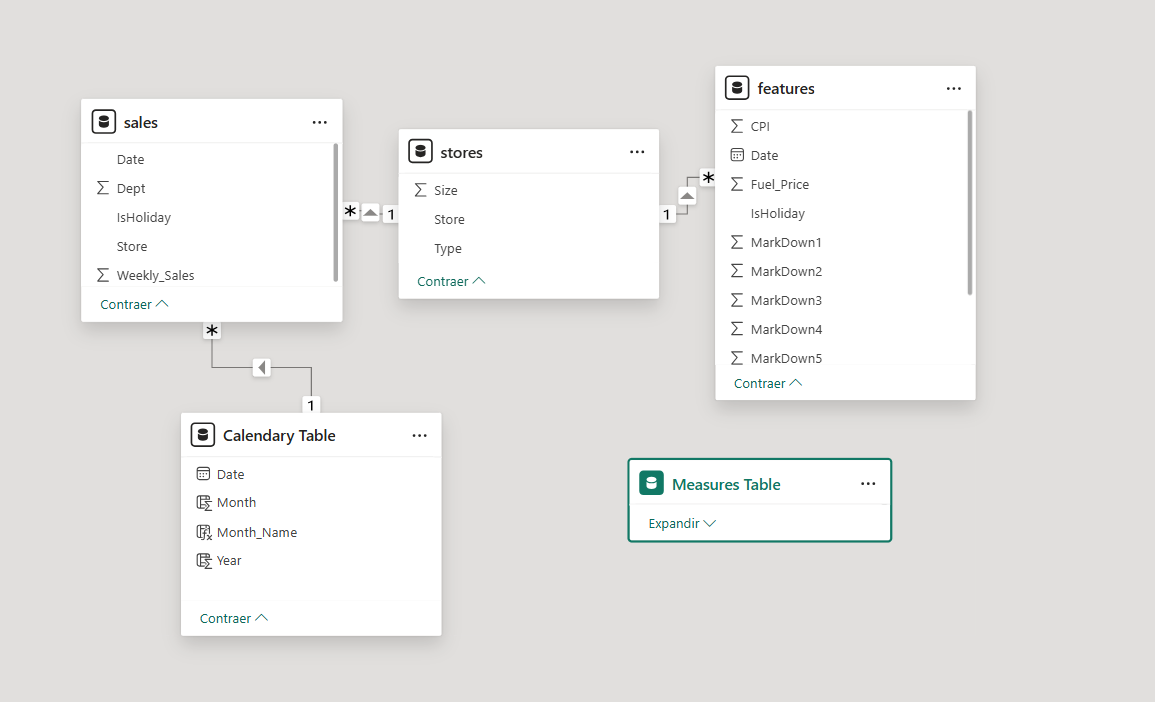
Las rebajas promocionales tienen un impacto significativo en las ventas semanales de los departamentos durante las semanas festivas, especialmente en fechas clave como el Super Bowl, Labor Day, Thanksgiving y Navidad.

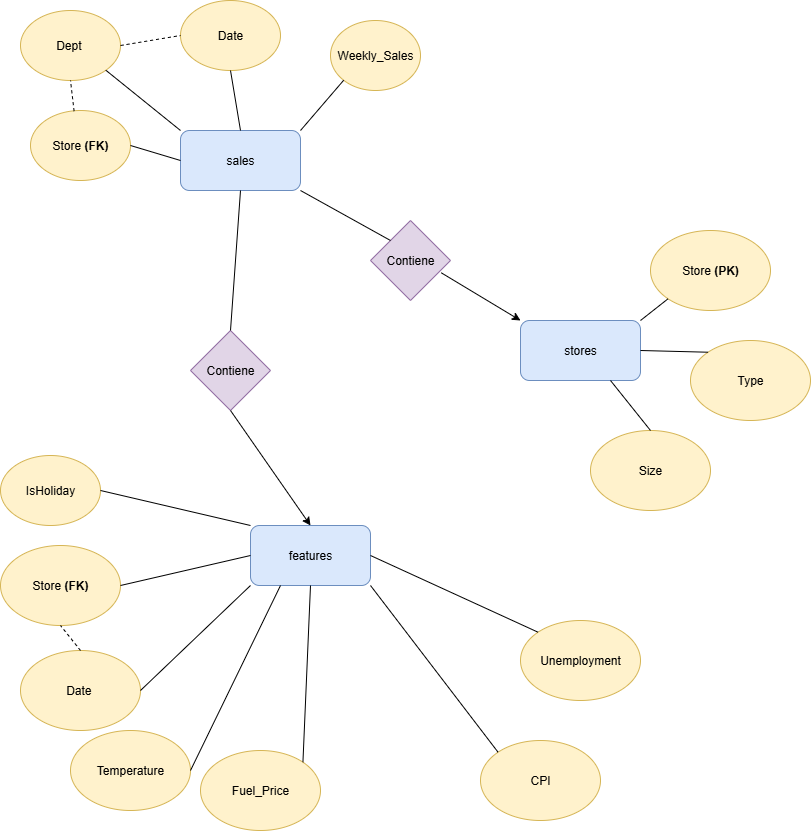
Se espera que los departamentos que participan en promociones experimenten un aumento notable en sus ventas, mientras que aquellos que no se ven beneficiados por descuentos mostrarán una tendencia estable o decreciente.

Asimismo, se plantea que factores como la temperatura, el precio del combustible, el índice de precios al consumidor (CPI) y la tasa de desempleo también influyen en el comportamiento de compra, especialmente en regiones y tiendas de mayor tamaño.

Este patrón permitirá generar recomendaciones personalizadas para maximizar las ventas según la temporada y características regionales.

# Descripción de Tablas y Relaciones





## Tabla: stores

**Descripción:**

Contiene información estática sobre cada tienda, como el tipo de tienda y su tamaño.

**Columnas principales:**

* **Store** (INT): Identificador único de la tienda.
* **Type** (CHAR): Tipo de tienda (por ejemplo: A, B o C).
* **Size** (INT): Tamaño de la tienda en pies cuadrados.

**Llave primaria:**

* Store

**Relaciones:**

* Se relaciona con sales mediante la columna Store.
* Se relaciona con features mediante la columna Store.

## 2. Tabla: sales

**Descripción:**

Registra datos de ventas semanales por tienda y departamento, incluyendo si la semana fue festiva.

**Columnas principales:**

* **Store** (INT): ID de la tienda.
* **Dept** (INT): ID del departamento dentro de la tienda.
* **Date** (DATE): Fecha de la semana.
* **Weekly\_Sales** (DECIMAL): Monto total de ventas para ese departamento y semana.

**Llave primaria:**

* Compuesta por: La combinación de Store, Dept, Date

**Llaves foráneas:**

* Store → Referencia a stores(Store)

**Relaciones:**

* Se relaciona con stores por Store.
* Se relaciona con features por Store y Date.

## 3. Tabla: features

**Descripción:**

Contiene variables externas semanales para cada tienda, como clima, precio del combustible, desempleo, etc.

**Columnas principales:**

* **Store** (INT): ID de la tienda.
* **Date** (DATE): Fecha de la semana.
* **Temperature** (FLOAT): Temperatura promedio semanal.
* **Fuel\_Price** (FLOAT): Precio promedio del combustible.
* **CPI** (FLOAT): Índice de precios al consumidor.
* **Unemployment** (FLOAT): Porcentaje de desempleo.
* **IsHoliday** (BOOLEAN): Indica si fue una semana festiva.

**Llave primaria:**

* Compuesta por: La combinación de Store, Date

**Llave foránea:**

* Store → Referencia a stores(Store)

**Relaciones:**

* Se relaciona con stores por Store.
* Se relaciona con sales por Store y Date.

# Transformaciones Realizadas

**En cuanto a la Entrega N°2:**

* Se corrigió el Diagrama Entidad Relación.
* Se unificaron todos los archivos xslx en uno solo, conteniendo un único archivo las tres tablas del dataset.
* Se modificaron los nombres de las tablas.
* Incorporación de Tabla Calendario
* Incorporación de Tabla de Medidas

**En cuanto a la Entrega N°3:**

* Se crearon todas las solapas requeridas en la consigna del Proyecto Final, incluidos filtros, botones de navegación, gráficos para cada tipo de análisis:
  + Se implementó una solapa de Storytelling y una solapa de Glosario
  + Las solapas de análisis se dividen en Análisis Global, Análisis de Ventas y Análisis Detallado; es decir, de lo global al detalle.
* Se crearon más medidas calculadas, las cuales ya han sido agregadas a la documentación de respaldo.
* Se agrega a la documentación de respaldo una breve explicación de las solapas de análisis mencionadas.

**Modificaciones para la Entrega Final:**

**1- Modelo Relacional: -IsHoliday aparece en más de una tabla, eliminar de la tabla Sales.(HECHO)**

**-Stores es FK en la tabla Features, pero cual es su PK?**

**-Eliminar las 5 Columnas de Markdowns (HECHO)**

**2-Archivo PowerBI:**

**-Análisis Detallado:**

-El gráfico de Embudo no tiene información**(CORREGIDO)**

-Hacer un top 5 o 10 departamentos **(CORREGIDO)**

-Departamentos dentro de cada tienda NO es claro**(CORREGIDO)**

-Los Segmentadores no van sobre el margen derecho (Van en el margen izquierdo o abajo del titulo) **(CORREGIDO)**

**-Análisis de Ventas:**

-Ambos gráficos de evolución son difíciles de entender **(CORREGIDO)**

-Para el filtro Fecha, departamento, tienda: destildar date y agregar filtro de mes y año, hacer otro de tienda y departamento. **(HECHO)**

-Todos los gráficos deben contar con etiqueta de datos **(CORREGIDO)**

## Medidas Calculadas Generadas y sus Fórmulas:

* **CPI Promedio:**

AvgCPI = average(features[CPI])

* **Número de Tiendas:**

Num\_Stores = count(stores[Store])

* **Ventas en Vacaciones:**

Sales\_On\_Holidays = calculate([Sum\_WeeklySales],sales[IsHoliday] = TRUE())

* **Suma de Ventas Semanales:**

Sum\_WeeklySales = sum(sales[Weekly\_Sales])

* **(VAR) Ventas Promedio POR TIENDA:**

Var\_SalesSummary =

VAR totalsales = SUM(sales[Weekly\_Sales])

VAR storecount = DISTINCTCOUNT(sales[Store])

VAR avgsalesperstore = DIVIDE(totalsales,storecount,0)

RETURN

Avgsalesperstore

* **(AVG) Ventas Promedio:**

AVG\_Sales = AVERAGE(sales[Weekly\_Sales])

* **(AVG) CPI Promedio:**

AvgCPI = average(features[CPI])

* **(AVGEX) Comparación entre CPI y ventas:**

AVGCPI\_vs\_sales = AVERAGEX(VALUES(features[CPI]),calculate([Sum\_WeeklySales]))

* **Porcentaje de participación de cada Departamento:**

dept\_participation = divide([Sum\_WeeklySales],calculate([Sum\_WeeklySales],all(sales[Dept])))

* **Ranking de Departamentos:**

dept\_ranking = rankx(all(sales[Dept]),[Salesperdept],,DESC,Dense)

* **Valor Estimado del Ticket:**

estimated\_ticket = divide([Sum\_WeeklySales],COUNTROWS(sales))

* **Ventas del Año Anterior:**

Last\_year\_sales = calculate([Sum\_WeeklySales],SAMEPERIODLASTYEAR('Calendary Table'[Date]))

* **Crecimiento Porcentual:**

percentualgrownth = divide([Sum\_WeeklySales] - [Last\_year\_sales],[Last\_year\_sales])

* **Impacto de la Promoción:**

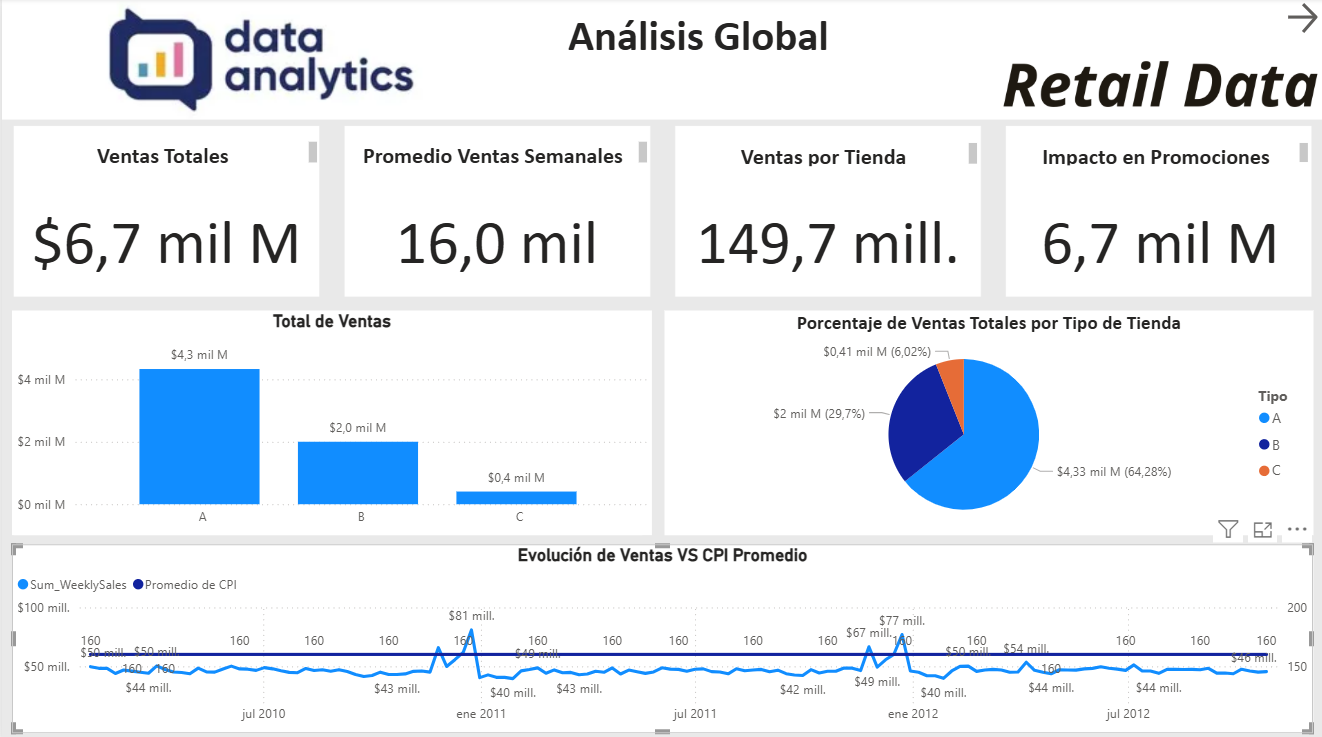
Promotion\_impact = CALCULATE([Sum\_WeeklySales],features[IsHoliday] = TRUE())

* **Ventas por Departamento:**

Salesperdept = calculate([Sum\_WeeklySales],ALLEXCEPT(sales,sales[Store],sales[Dept]))

# Descripción de las Solapas del Análisis:

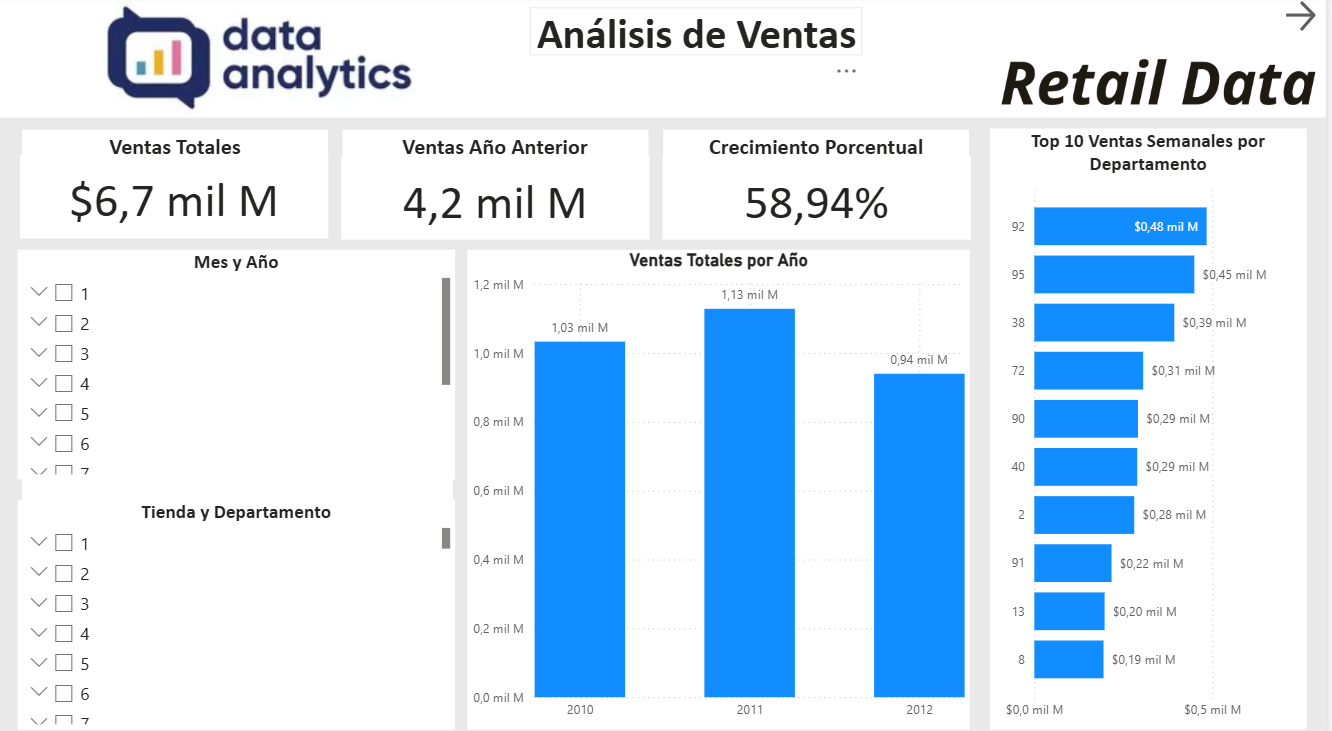
## Análisis Global

****

Enesta primera sección se presenta una visión general del negocio. El objetivo es mostrar las ventas en su contexto más amplio, considerando la evolución en el tiempo y la relación con variables externas como precios de combustible, índice de desempleo o el CPI.

Además, se analizan diferencias de desempeño según tipo y tamaño de tienda. Esta solapa responde a preguntas como: ¿Cómo se comporta la cadena en su conjunto? y ¿Qué factores macroeconómicos podrían estar influyendo en los resultados?

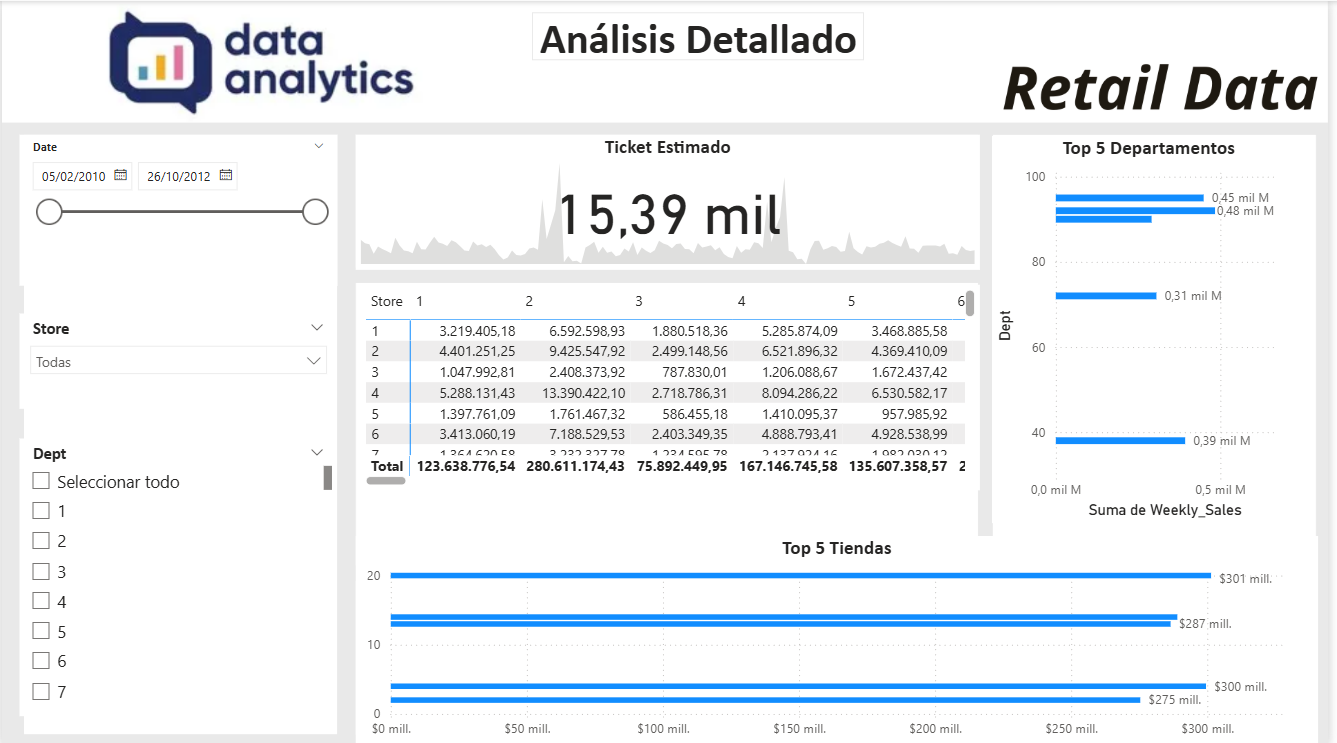
## Análisis de Ventas

****

La segunda solapa se enfoca en el desempeño comercial de manera más específica, observando el comportamiento de las ventas en distintos periodos y departamentos. Se comparan resultados interanuales y se identifican los departamentos con mayor volumen de ventas, así como su evolución en el tiempo.

Esta solapa responde a preguntas como: ¿Qué tan bien estamos vendiendo este año respecto al anterior? y ¿Cuáles son los departamentos más importantes para el negocio?

## Análisis Detallado

****

En la última solapa se profundiza a nivel de tiendas y departamentos individuales. Aquí se analiza la participación porcentual de cada tienda o departamento en las ventas totales, además de métricas más específicas como el ticket promedio. Se incluyen filtros interactivos para ajustar la vista por fechas, tienda o departamento, permitiendo un análisis granular.

Esta solapa responde a preguntas como: ¿Qué tiendas o departamentos están liderando las ventas?, ¿Dónde existen oportunidades de mejora? y ¿Cómo varía el ticket promedio en diferentes contextos?

# Conclusión Final y Posibles Iniciativas a Futuro:

Este proyecto constituye un primer paso en el análisis de ventas y desempeño de la cadena de tiendas. Sin embargo, existen diversas iniciativas que podrían complementar y enriquecer el trabajo realizado:

1. **Análisis de predicción de ventas**: aplicar modelos de Machine Learning para anticipar la demanda futura considerando estacionalidad, promociones y variables macroeconómicas.
2. **Optimización de promociones**: evaluar el impacto real de las campañas promocionales en las ventas y rentabilidad, con el fin de diseñar estrategias más eficientes.
3. **Segmentación de tiendas**: clasificar las tiendas según características como ubicación, tamaño y desempeño, para identificar patrones y adaptar estrategias comerciales a cada grupo.
4. **Análisis de rentabilidad**: incorporar costos y márgenes para complementar el análisis de ventas con una perspectiva de beneficios netos.
5. **Integración con datos externos**: sumar información de la competencia, clima o tendencias de consumo que puedan influir en el comportamiento de los clientes.

De esta forma, el proyecto no solo sirve para describir la situación actual, sino que también abre la puerta a nuevas herramientas de apoyo para la toma de decisiones estratégicas.

**FIN DEL PROYECTO**