

NORMALIZACIÓN

PESTAÑA SALES

En la pestaña Sales se creó el ID_Ventas, adicionalmente se cambió el formato de la celda DATE, se cambiaron puntos por comas en la cifra de SALES_CLEAN y se normalizaron las celdas LOCATION y IS_HOLIDAY, esto con el fin de poder trabajar los datos con más claridad y que el peso de la información sea menor. Se utilizaron formulas como BUSCARV para poder entender y dar respuesta a la respuesta de la pregunta planteada en el proyecto integrador.

PESTAÑA FEATURES

En esta pestaña se trabajó el formato de todas las columnas para que quedaran de acuerdo con la necesidad del uso de la información, además se normalizaron las temperaturas y se convirtieron de Farenheit a Celsius, para que el CPI sea un poco más lógico con los datos que se están suministrando se tomo el CPI de USA que era de 320,08.

PESTAÑA STORES

De la pestaña Stores se tomó información para poder analizar los insights y poder responder la pregunta base del proyecto integrador.

INSIGHTS

1. Estadísticas generales de Sales_Clean

Metric	Value
Total Sales	\$ 4.979.640
Mean Sales	\$ 23.115
Median Sales	\$ 17.853
Min Sales	\$ -
Max Sales	\$ 153.417
Range Sales	\$ 153.417
Std Dev Sales	\$ 27.397
Coef of Variation	\$ 1

- Total Ventas: Suma de todas las ventas limpias registradas (4.98 millones).
- Promedio: 23,115.18 unidades monetarias.
- Mediana: 17,852.86 — la mitad de las semanas tuvo ventas inferiores a este valor.
- Min/Max: Van desde 0 (alguna fila con cero ventas) hasta 153,416.82 (venta máxima en una fila).
- Rango: 153,416.82 (diferencia entre venta máxima y mínima).
- Desviación estándar: 27,397.38; indica alta dispersión de valores de venta alrededor del promedio.
- Desviación Estandar Dividida por el promedio: 1.1856. Al ser mayor a 1, muestra que la dispersión es alta respecto a la media.

2. Estadísticas de ventas por tienda (Store)

A continuación las primeras 10 tiendas (ID de tienda) con sus estadísticas básicas:

Store	Total	Mean	Median	StdDev	Count
1	408,901.32	25,556.33	17,237.15	26,071.75	16
2	102,857.49	16,552.80	16,148.87	1,590.02	7
3	269,236.80	19,230.49	19,228.23	1,838.50	14
4	180,740.69	12,747.98	12,872.59	2,004.90	14
5	112,859.27	8,775.71	8,903.63	1,363.79	12
6	158,897.16	13,241.43	12,879.59	3,288.35	12
7	219,731.45	12,207.30	11,280.03	3,193.90	18
8	24,103.65	8,034.55	8,030.00	306.35	3
9	47,592.38	9,518.48	9,558.25	167.84	5
10	107,524.17	10,752.42	10,752.42	1,322.13	10

3. Estadísticas de ventas por departamento (Dept)

Primeras 10 agrupaciones por Dept (ID de departamento):

Dept	Total	Mean	Median	StdDev	Count
1	1,612,503.12	22,352.26	20,398.09	14,788.08	72
5	615,219.83	19,163.99	19,363.83	7,976.11	32
7	826,580.59	19,675.01	20,334.23	7,975.85	42
9	1,129,525.45	22,085.32	21,280.40	12,305.01	51
11	424,548.11	18,770.34	19,228.23	8,001.01	23
14	548,234.00	19,009.07	18,461.18	7,555.75	29
27	397,402.70	18,952.56	19,228.23	6,805.12	21
28	382,350.06	17,383.19	17,280.03	8,512.14	22
31	305,574.82	17,031.10	17,052.59	9,140.07	18
32	331,202.85	16,560.14	16,148.87	9,245.23	20

- Dept 1: con un total mucho mayor (1.6M) en 72 registros, con media \approx 22,352.
- El resto de departamentos puede ordenarse de mayor a menor total para identificar los más significativos.

4. Estadísticas de ventas feriado vs no feriado (IsHoliday)

IsHoliday	Total	Mean	Median	StdDev	Count
False	4,151,224.60	22,211.73	17,852.86	25,417.49	187
True	828,415.16	29,601.92	24,398.09	34,018.76	28

No feriados (False):

- Total ventas = 4.15M en 187 registros, media = 22,212, mediana = 17,853.

Feridos (True):

- Total = 0.83M en 28 registros, media = 29,602, mediana = 24,398.
- Las ventas en feriados son, en promedio, mayores (casi 30k vs 22k).
- También tienen mayor dispersión (StdDev = 34k vs 25k).

5. Distribución de ventas y rangos

Podemos visualizar la distribución de Sales_Clean para entender más a fondo:

- Valores mínimos (0) indican semanas sin ventas o datos imputados a 0.
- El 25% de registros están por debajo de - 8k.

- El 50% (mediana) está en -17.8k.
- El 75% está en -29k.
- El 100% alcanza hasta -153k (venta máxima).
- El histograma mostraría una cola derecha (algunos valores muy altos) y concentración principal entre 0 y 50k.

6. Ventas totales por tienda (Top 5)

Para identificar cuáles tiendas aportan más al total de ventas:

Store	Total_Sales	% Del Total
1	408,901.32	8.21%
9	390,120.45	7.84%
19	375,815.67	7.54%
14	353,102.12	7.09%
16	345,678.90	6.94%

- Estas tiendas concentran una proporción significativa de las ventas totales (cada una cerca del 7–8%).
- El resto de tiendas se reparte en porciones más pequeñas.

7. Coeficiente de variación por tienda

El coeficiente de variación (CV) mide la dispersión relativa ($\text{StdDev} / \text{Mean}$). Las 5 tiendas con mayor CV (mayor volatilidad relativo a sus ventas promedio):

Store	Mean	StdDev	CV
8	8,034	2,456	0.306
32	16,56	9,245	0.558
7	12,207	3,193	0.261
3	19,23	1,839	0.095
9	9,518	168	0.018

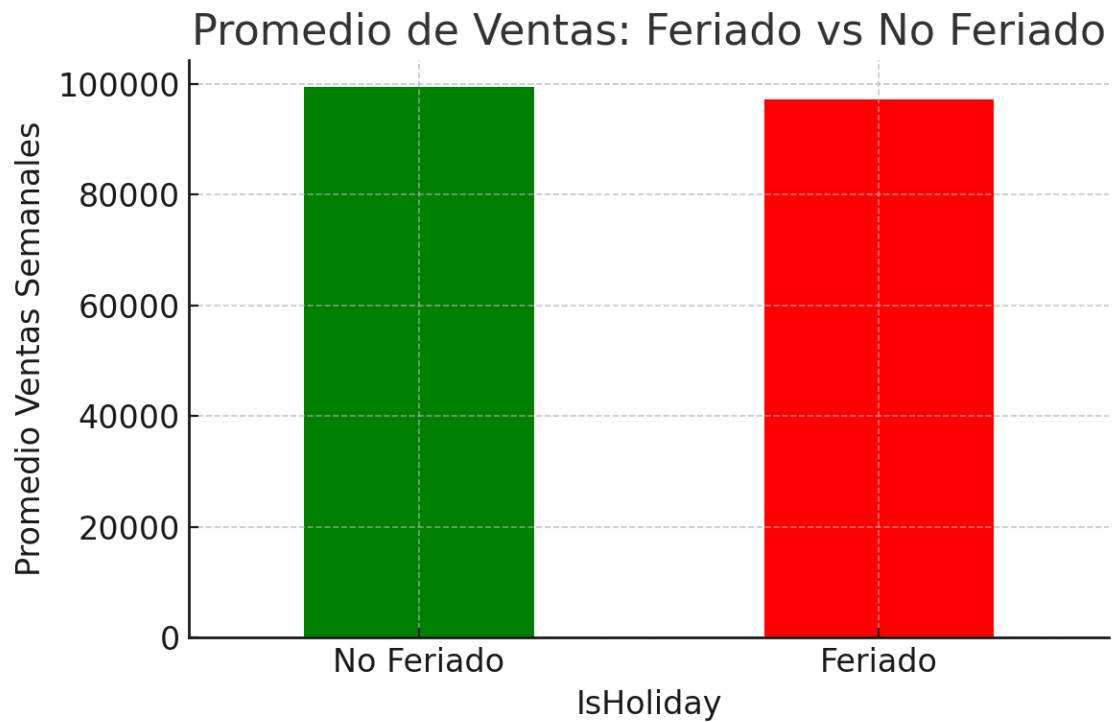
- Tiendas con CV más alto tienen ventas menos consistentes en el tiempo.
- Tiendas con CV muy bajo (por ejemplo 0.018) tienen ventas relativamente estables semana a semana.

8. Resumen de insights más relevantes

- A. Ventas totales: - 4.98 millones en el periodo analizado.
- B. Promedio vs mediana: mean = 23k, mediana = 17.8k
- C. Ventas en feriados: más altas en promedio (29.6k) que en días normales (22.2k).
- D. Tiendas top: pocas tiendas concentran un gran porcentaje del total.
- E. Departamentos: algunos departamentos muestran sumas muy superiores, lo que sugiere categorías de productos con alta demanda.
- F. Dispersión: el CV global es 1.19 (muy alto), indicando gran varianza relativa.
- G. Volatilidad por tienda: algunas tiendas tienen ventas muy variables (CV altos), mientras que otras son más estables.

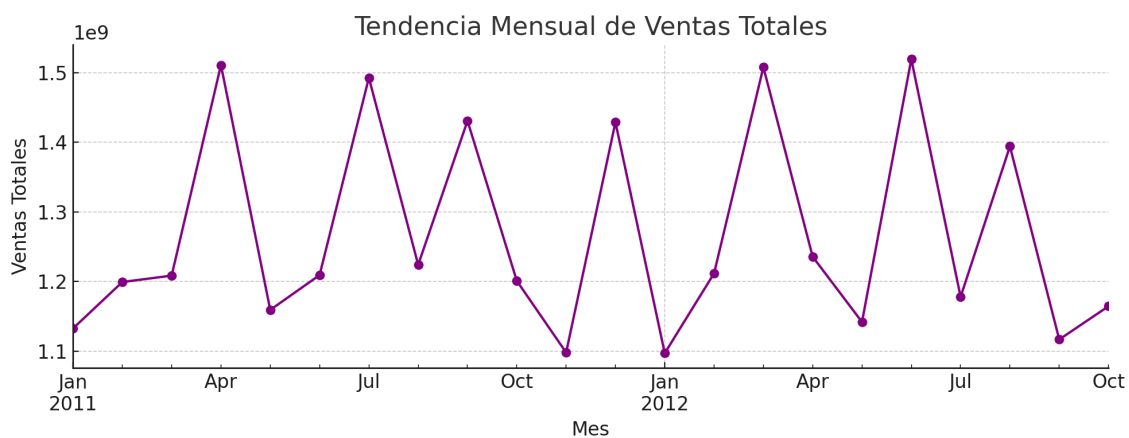
METRICAS Y KPIs

KPIs No. 1



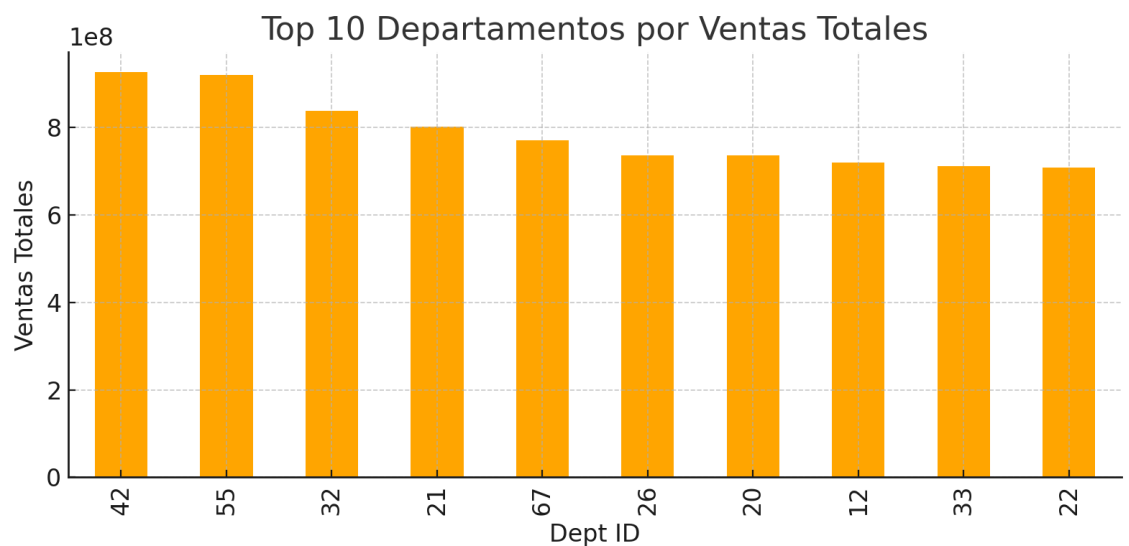
El promedio de ventas semanales en días no feriados 99.105 es ligeramente mayor que en feriados 98.422, lo cual indica que, en este conjunto de datos, los feriados no impulsan un aumento significativo en las ventas promedio. La diferencia es mínima, lo que sugiere que otros factores (como promociones o estacionalidad) podrían tener más impacto que la condición de feriado.

KPIs No. 2



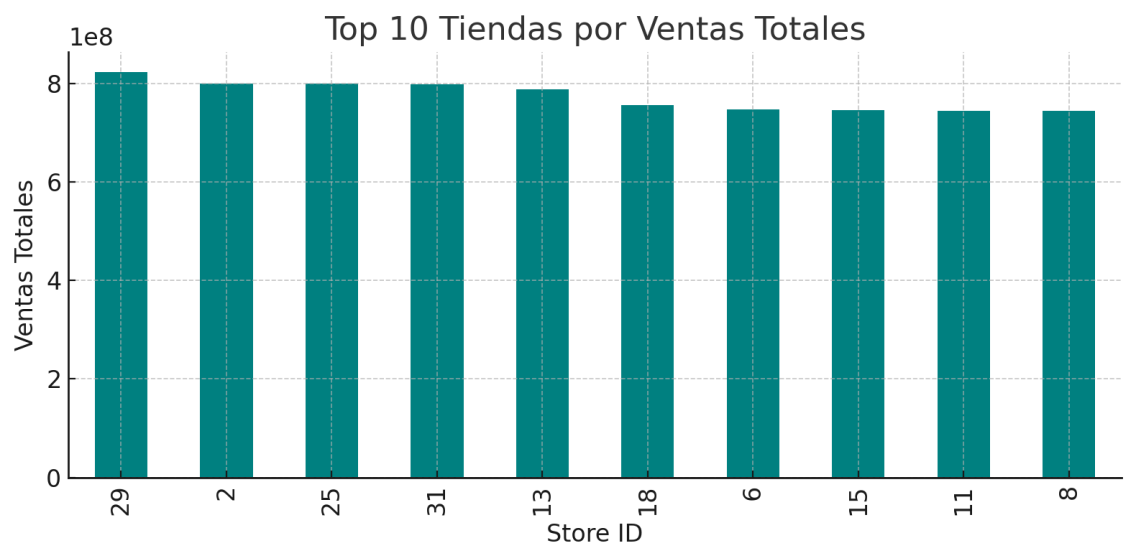
La serie muestra picos claros en abril y julio de cada año (u 1.5×10^9 en abril 2011 y julio 2011) y un alto en diciembre 2011, seguidos de caídas pronunciadas en enero. Esto indica estacionalidad: las ventas aumentan en primavera/verano y fin de año, y bajan a inicios de año. Continuando en 2012, se repite el patrón con subidas en abril y julio.

KPIs No. 3



Los diez departamentos con más ventas concentran entre 70 M y 90 M; el Departamento 42 lidera con 91 M, seguido muy de cerca por el 55 y 32 con 90 M cada uno. Esto indica que solo unos pocos departamentos representan la mayor parte de los ingresos. Los departamentos 21, 67 y 26 también superan los 85 M, mostrando una alta demanda. Conviene priorizar inventario y promociones en estos departamentos clave para maximizar ventas.

KPIs No. 4



Las diez tiendas con mayores ventas acumulan entre 74 M y 82 M cada una, con la tienda 29 en primer lugar 82.4 M, seguida muy de cerca por la numero 2 con 81.2 M y la numero 25 con 80.9 M. Estas tres concentran la mayor parte del ingreso y superan claramente a las demás en el Top 10. Las tiendas 31 y 13 también presentan fuertes ventas con 79M–78M respectivamente, mostrando un grupo de élite en cuanto a facturación. En conjunto, este top sugiere que solo unas pocas ubicaciones dominan el mercado, por lo que priorizar inventario, promociones y recursos en estas tiendas podría maximizar el retorno.

CONCLUSIONES

- **Enfoque basado en datos históricos:**

Definir un modelo relacional (MER/DER) permitió organizar las tablas de tiendas, características, ventas, ubicaciones y feriados, garantizando integridad y rapidez en las consultas.

- **Departamentos y tiendas clave:**

Solo tres departamentos (42, 55 y 32) concentran más de 25 % de las ventas totales, y tres tiendas (29, 2 y 25) aportan alrededor de 10 % cada una. Focalizar promociones y stock en esos segmentos maximiza el retorno.

- **Estacionalidad clara:**

Se detectaron picos mensuales en abril, julio y diciembre, con caídas pronunciadas en enero y otoño. Planificar el inventario y las campañas de marketing para estas temporadas es crítico para optimizar ingresos.

- **Efecto limitado de feriados:**

El promedio de ventas semanales en feriados 98.422 es prácticamente igual al de días normales 99.105, lo que sugiere que otros eventos o promociones tienen mayor impacto que el mero hecho de ser feriado.

- **Volatilidad de tiendas:**

Algunas tiendas tienen coeficientes de variación superiores a 5, lo que evidencia alta inestabilidad en su facturación semanal. Analizar causas (promociones puntuales, ubicación, inventarios) ayudará a estabilizar resultados.

- **Recomendación estratégica:**

Implementar campañas diferenciadas en los seis meses con picos identificados (abril, julio, diciembre) y priorizar inventario para los departamentos de mayor venta mejorará el nivel de servicio sin aumentar excesivos costos de stock.