

# Análisis de Ventas Semanales en una Cadena de Tiendas

## 1. Introducción

Este análisis tiene como objetivo explorar y analizar los datos proporcionados por una cadena de tiendas para efectos de entender que factores influyen en su comportamiento de ventas y poder dar conclusiones al respecto. Los datos utilizados provienen de la plataforma Kaggle, específicamente del conjunto de datos Retail Data Analytics.

## 2. Procesamiento Inicial de Datos

### 2.1. Fuente de Datos

Se descargaron tres archivos en formato CSV que contenían la información:

Stores data-set: Información sobre las tiendas.

Sales data-set: Ventas semanales por tienda.

Features data-set: Información adicional como precios de combustible, descuentos, índices económicos y tasas de desempleo.

### 2.2. Preparación de los Datos

Los archivos de ventas y características se convirtieron a formato Excel para estandarizar los formatos de fecha. Posteriormente, se importaron a Python para realizar el procesamiento y análisis.

## 3. Exploración y Limpieza de Datos

### 3.1. Estructura de las Tablas

Se tuvieron tres tablas iniciales:

#### Tabla Stores

**Store:** Identificador único de cada tienda.

**Store\_Type:** Categoría o tipo de tienda. (A, B, C)

**Size:** Tamaño físico de la tienda.

#### Tabla Sales

**Store:** Identificador único de la tienda.

**Dept:** Departamento dentro de la tienda.

**Sales\_Week\_Date:** Fecha correspondiente a una semana específica.

AUTOR: RHEYJACH ALI ARRIETA DELGADO

**Weekly\_Sales:** Ventas totales generadas en la semana.

**IsHoliday:** Indica si la semana incluye un día festivo. (True/False)

### Tabla Features

**Store:** Identificador único de la tienda.

**Sales\_Week\_Date:** Fecha correspondiente a una semana específica.

**Temperature:** Temperatura promedio semanal en la región de la tienda.

**Fuel\_Price:** Precio promedio del combustible durante la semana.

**MarkDown1 - MarkDown5:** Descuentos promocionales en productos.

**CPI:** Índice de Precios al Consumidor.

**Unemployment:** Tasa de desempleo en la región.

**IsHoliday:** Indica si la semana incluye un día festivo. (True/False)

## 3.2. Limpieza de Datos

### Tabla Stores

Codigo:

```
null_stores=stores.isnull().sum()
duplicated_stores=stores.duplicated(subset=["Store"],keep=False).sum()
```

Sin datos nulos ni duplicados.

### Tabla Sales

Codigo:

```
sales=sales.drop(columns=["Dept","IsHoliday"],axis=1)
sales['Sales_Week_Date'] = pd.to_datetime(sales['Sales_Week_Date'])
null_sales=sales.isnull().sum()
sales=sales.groupby(["Store","Sales_Week_Date"])["Weekly_Sales"].sum().reset_index()
duplicated_sales=sales.duplicated(subset=["Store","Sales_Week_Date"],keep=False).sum()
```

Se eliminaron las columnas:

- Dept: No identificaba al departamento.
- IsHoliday: Redundante con la tabla Features.

Se convirtió la columna Sales\_Week\_Date a formato datetime.

Se agruparon las ventas por tienda y semana.

Sin datos nulos ni duplicados

AUTOR: RHEYJACH ALI ARRIETA DELGADO

## Tabla Features

Codigo:

[illegible]

Los datos nulos en las columnas Markdown1-5 se reemplazaron por 0.

Se convirtió Sales\_Week\_Date a formato datetime.

No quedaron datos nulos ni duplicados

## 4. Consolidación de Datos

Codigo:

```
sales_features = pd.merge(sales, features, on=['Store', 'Sales_Week_Date'], how='inner')
final_data = pd.merge(sales_features, stores, on='Store', how='inner')
duplicated_final_data = final_data.duplicated(subset=['Store', 'Sales_Week_Date'], keep=False).sum()
```

Se unieron las tablas utilizando operaciones de merge en Python:

Sales y Features: Se unieron por Store y Sales\_Week\_Date.

Tabla resultante con Stores: Se utilizó un método inner para mantener solo las filas coincidentes.

El resultado fue una tabla final llamada `final_data`, libre de duplicados y lista para análisis.

## 5. Análisis

### 5.1 Análisis Exploratorio

```
exploration_final_data=final_data.describe(include=[int,float])
```

```
final_data['MarkDown2']=final_data['MarkDown2'].apply(lambda x:0 if x <0 else x)
final_data['MarkDown3']=final_data['MarkDown3'].apply(lambda x:0 if x <0 else x)
```

	Weekly_Sales	Temperature	Fuel_Price	CPI	Unemployment	Size
count	6435	6435	6435	6435	6435	6435
mean	1046964,878	60,66378244	3,358606838	171,5783938	7,999151049	130287,6
std	564366,6221	18,44493288	0,459019707	39,3567123	1,875884782	63117,02247
min	209986,25	-2,06	2,472	126,064	3,879	34875
max	3818686,45	100,14	4,468	227,2328068	14,313	219622

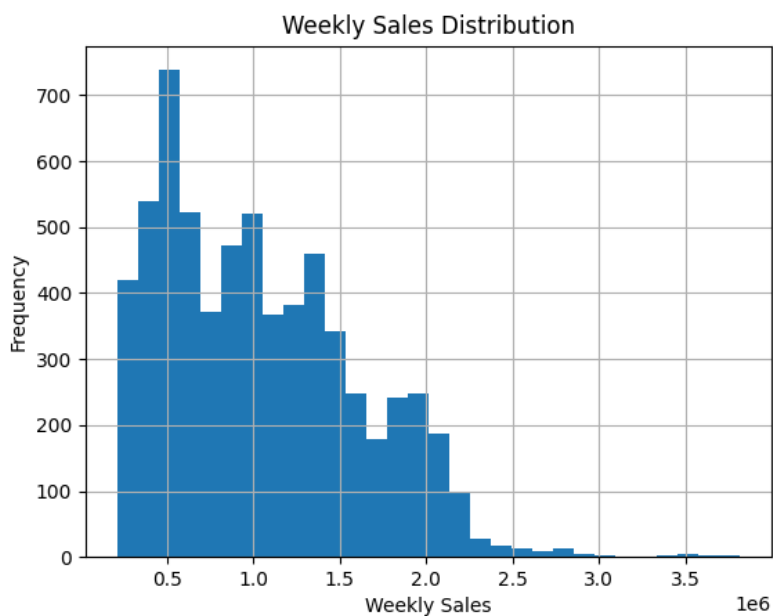
	MarkDown1	MarkDown2	MarkDown3	MarkDown4	MarkDown5
count	6435	6435	6435	6435	6435
mean	2429,019322	818,9616208	429,1888811	1008,870435	1581,806813
std	5871,477726	4879,24085	5273,522766	3753,943314	4098,19854
min	0	0	0	0	0
max	88646,76	104519,54	141630,61	67474,85	108519,28

Se realizó una exploración inicial de los datos utilizando el método describe en Python, lo cual permitió identificar dos aspectos relevantes:

1. La desviación estándar de las variables era considerablemente alta en comparación con los valores máximos y mínimos de los datos, lo que llevó a concluir que estos presentan una alta dispersión.
2. Gracias a que el método describe muestra los valores mínimos de cada columna, se detectaron valores negativos en las columnas MarkDown2 y MarkDown3. Estos valores fueron reemplazados por 0, ya que no correspondían con la naturaleza de los datos, los cuales estaban destinados a representar posibles descuentos.

## Análisis de distribución de ventas

```
final_data['Weekly_Sales'].hist(bins=30)
plt.title('Weekly Sales Distribution')
plt.xlabel('Weekly Sales')
plt.ylabel('Frequency')
plt.show()
```

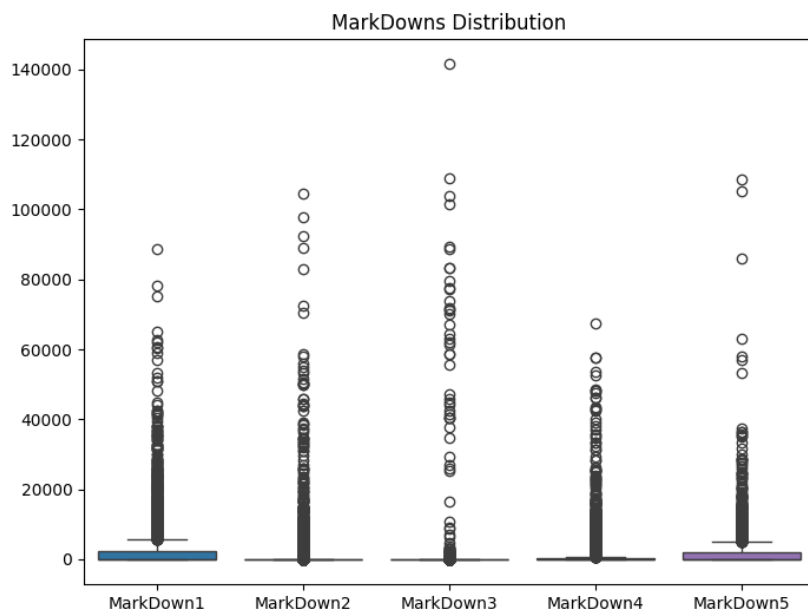


Se generó un histograma para analizar la distribución de las ventas, llegando a las siguientes conclusiones:

El pico más alto de ventas es el menos frecuente, mientras que las ventas más bajas son las más comunes entre todas las tiendas. Esto será clave para determinar las condiciones que permitieron a ciertas tiendas alcanzar niveles tan altos de ventas y cómo estas podrían replicarse en otras tiendas a lo largo del tiempo.

## Análisis de distribución de los datos en los Markdowns

```
plt.figure(figsize=(8, 6))
sns.boxplot(data=final_data[['MarkDown1', 'MarkDown2', 'MarkDown3', 'MarkDown4', 'MarkDown5']])
plt.title('MarkDowns Distribution')
plt.show()
```



Se creó un diagrama de caja para visualizar la distribución de los datos en las variables MarkDown, dado que manejan rangos de valores similares, lo cual facilita su evaluación conjunta. Este análisis fue motivado por el hallazgo y la posterior corrección de valores negativos. Las conclusiones obtenidas fueron:

1. En todos los MarkDown, la mayoría de los datos se encuentran próximos entre sí, abarcando un rango aproximado de 0 a 45,000.
2. Con excepción de MarkDown3, la dispersión de los datos en los MarkDown comienza a partir de los 45,000. Aunque los valores superiores a este rango son menos frecuentes, están significativamente más dispersos.
3. El MarkDown3 muestra un comportamiento distinto en comparación con los otros. En primer lugar, presenta el valor máximo, superior a 140,000. Además, la dispersión de datos inicia antes de los 45,000, aproximadamente en 10,000, y los datos están altamente dispersos hasta alcanzar el máximo. Sin embargo, hay una notable concentración entre 0 y 10,000, lo que explica por qué su desviación estándar no es la mayor, siendo la segunda más alta después del MarkDown1.

## Matriz de correlación

```
final_data["IsHoliday"] = final_data["IsHoliday"].astype(int)
selectionData = final_data.select_dtypes(include=["int", "float"])
Correlation_matrix = selectionData.corr()
CorrelationWeekly_Sales = Correlation_matrix["Weekly_Sales"].reset_index(name="Weekly_Sales")
positive_correlation = CorrelationWeekly_Sales[CorrelationWeekly_Sales["Weekly_Sales"] > 0].sort_values(by="Weekly_Sales", ascending=False)
negative_correlation = CorrelationWeekly_Sales[CorrelationWeekly_Sales["Weekly_Sales"] <= 0].sort_values(by="Weekly_Sales").reset_index(drop=True)
```

Positively correlated variables	Weekly_Sales_Correlation
Size	0,810468495
MarkDown1	0,179107167
MarkDown5	0,173272609
MarkDown4	0,139194928
MarkDown3	0,120289336
MarkDown2	0,080164014
IsHoliday	0,036890968
Fuel_Price	0,009463786

Negatively correlated variables	Weekly_Sales_Correlation
Store	-0,335332015
Unemployment	-0,10617609
CPI	-0,072634162
Temperature	-0,063810013

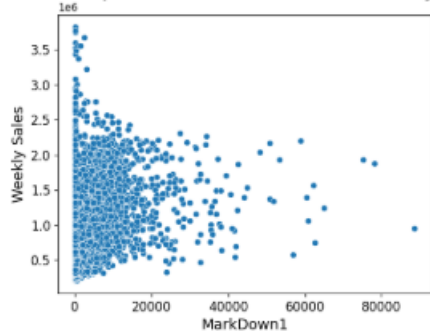
Se generó una matriz de correlación focalizada en la variable objetivo, Weekly\_Sales. Los resultados fueron los siguientes:

1. La correlación puede ayudar a identificar relaciones entre variables, pero depende del contexto y las consideraciones específicas. La matriz mostró que 7 variables tienen una correlación positiva con la variable objetivo, pero en un rango de 0.2 a 0, indicando una relación positiva muy débil. La variable Size presentó una correlación de aproximadamente 0.81, indicando una influencia considerable. Este resultado es lógico, ya que el tamaño de la tienda afecta la capacidad para atender clientes y, por ende, las ventas totales.
2. Cuatro variables mostraron correlación negativa, en un rango de -0.33 a 0, indicando relaciones débiles. Estas variables fueron Unemployment, CPI, Temperature y Store. Aunque las tres primeras influyen negativamente en las ventas, su impacto no es determinante. La variable Store tuvo una correlación de -0.33, ligeramente más alta, pero esto es natural dado que representa el identificador de la tienda, y las ventas varían según la tienda.
3. Es necesario explorar otros métodos para identificar relaciones o influencias más precisas en la variable Weekly\_Sales.

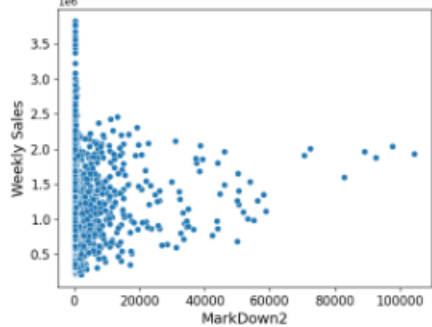
## 5.2 Análisis Relacional

### Gráficos de dispersión

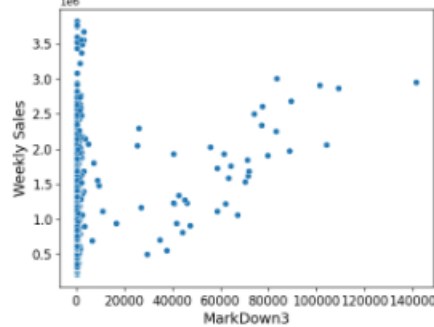
Relationship between MarkDown1 and Weekly Sale



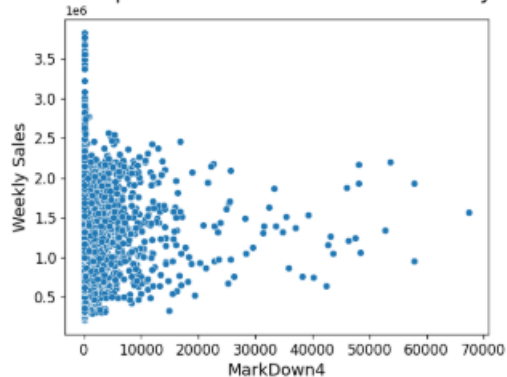
Relationship between MarkDown2 and Weekly Sale



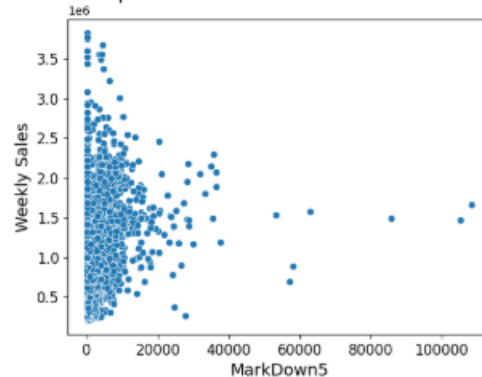
Relationship between MarkDown3 and Weekly Sale



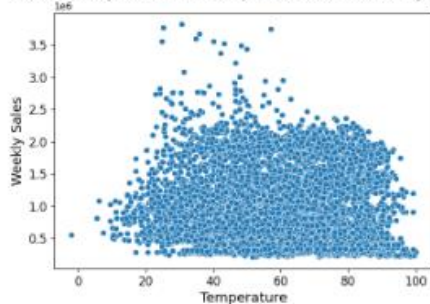
Relationship between MarkDown4 and Weekly Sale



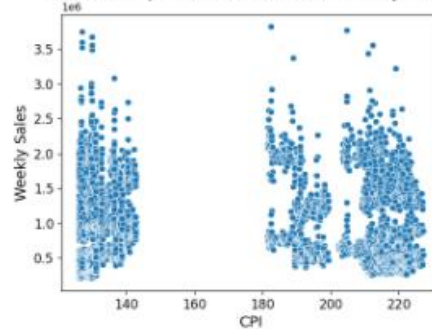
Relationship between MarkDown5 and Weekly Sale



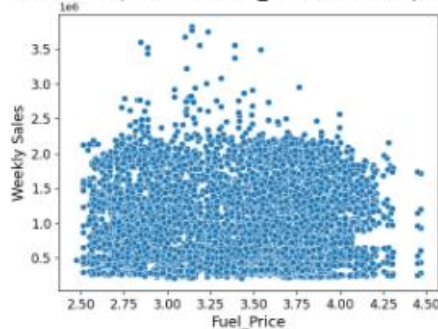
Relationship between Temperature and Weekly Sales



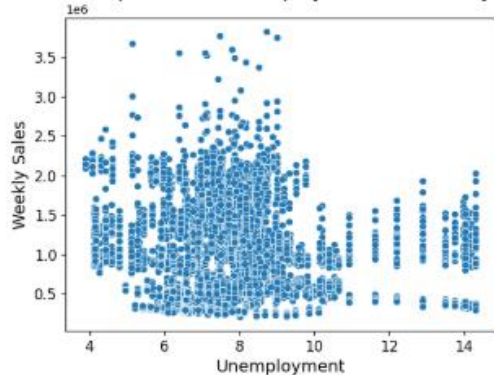
Relationship between CPI and Weekly Sales



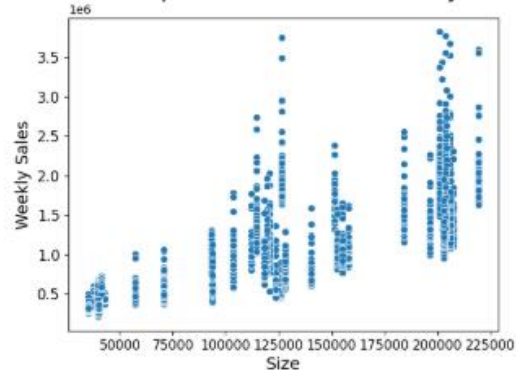
Relationship between Fuel\_Price and Weekly Sales



Relationship between Unemployment and Weekly Sales



Relationship between Size and Weekly Sales





Codigo:

```
for column in ['Temperature', 'Fuel_Price', 'CPI', 'Unemployment', 'Markdown1', 'Markdown2', 'Markdown3', 'Markdown4', 'Markdown5', 'Size']:
    plt.figure()
    sns.scatterplot(data=final_data, x=column, y='Weekly_Sales')
    plt.title(f'Relationship between {column} and Weekly Sales', fontsize=18)
    plt.xlabel(column, fontsize=14)
    plt.ylabel('Weekly Sales', fontsize=14)
    plt.xticks(fontsize=11)
    plt.yticks(fontsize=12)
    plt.show()
```

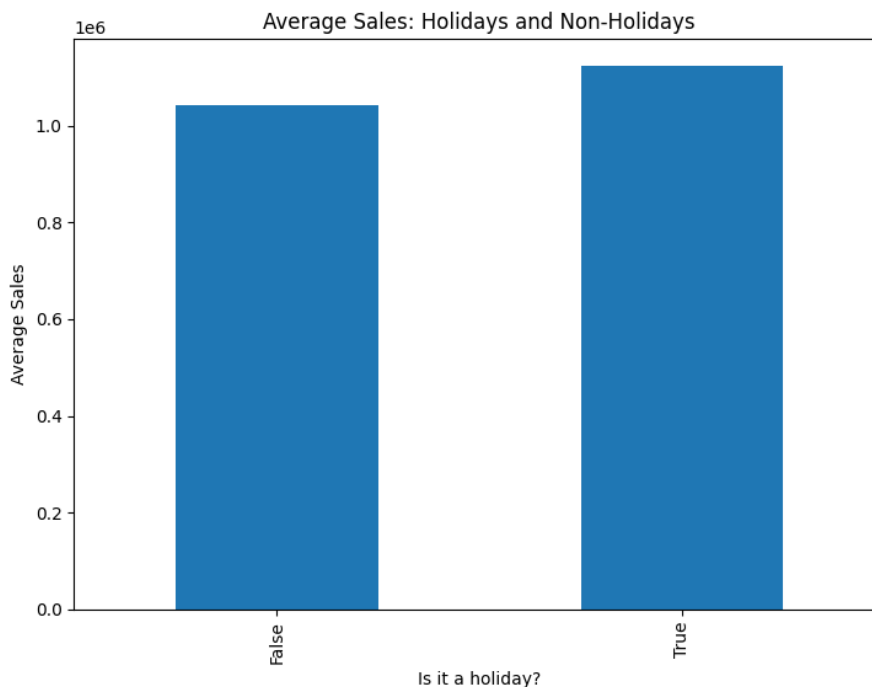
Se procedió a realizar un análisis relacional con gráficos de dispersion para identificar tendencias y relaciones entre cada variable y la variable objetivo Weekly\_Sales, lo que dio a lugar las siguientes consideraciones.

1. Se realizó una gráfica de dispersión entre los Markdowns y Weekly\_Sales. Para cada uno, la tendencia mostró que la mayoría de los datos de ventas se dieron con un Markdown cercano a cero, es decir, no hubo muchos descuentos con valores altos. El punto máximo de ventas se dio con un Markdown en 0 para cada uno, pero aun así, con ese mismo valor, se registraron diversas cantidades de ventas, incluyendo valores mínimos. En ciertos Markdowns, como el 2, 3 y 5, se observó que, para valores altos, la cantidad mínima de ventas era considerablemente superior a la mayoría de los datos mínimos. Sin embargo, en estos casos, la cantidad de datos fue aproximada a 15 para cada Markdown, lo que, en comparación con los más de 6000 registros totales, no representa una tendencia clara y comprobada.
2. Se realizó una gráfica de dispersión entre Temperature y Weekly\_Sales para determinar el comportamiento y la relación entre ambas. Siguiendo lo concluido con la matriz de correlación, en este caso no se encontró una gran influencia de la variable Temperature respecto a las ventas. Los datos mínimos de ventas se presentaron tanto en registros mínimos como máximos de temperatura. En el caso de los datos máximos de ventas, a una misma temperatura se encontraron registros con diferentes cantidades de ventas, incluidas algunas con valores mínimos, lo que indica que no existe una gran influencia ni una tendencia que relacione la temperatura con la cantidad de ventas.
3. Se realizó una gráfica de dispersión entre CPI y Weekly\_Sales, donde tampoco se evidenció una influencia significativa entre el CPI y la cantidad de ventas. Hubo valores máximos y mínimos para la variable CPI, con alta variabilidad en las ventas. Incluso los valores máximos de ventas se dieron con un determinado CPI, pero ese mismo valor también mostró ventas variadas, llegando a mínimos. Esto concuerda a lo indicado por la matriz de correlación, esta variable no tiene una gran influencia en la cantidad de ventas.
4. Se realizó una gráfica de dispersión entre Fuel\_Price y Weekly\_Sales, donde no se mostró una influencia significativa entre el precio del combustible y la cantidad de ventas. Se observaron valores máximos y mínimos para la variable Fuel\_Price, con alta variabilidad en las ventas. Incluso los valores máximos de ventas se dieron a un

determinado precio, pero ese mismo precio también presentó valores mínimos. Esto concuerda con lo señalado por la matriz de correlación: esta variable no influye considerablemente en la cantidad de ventas.

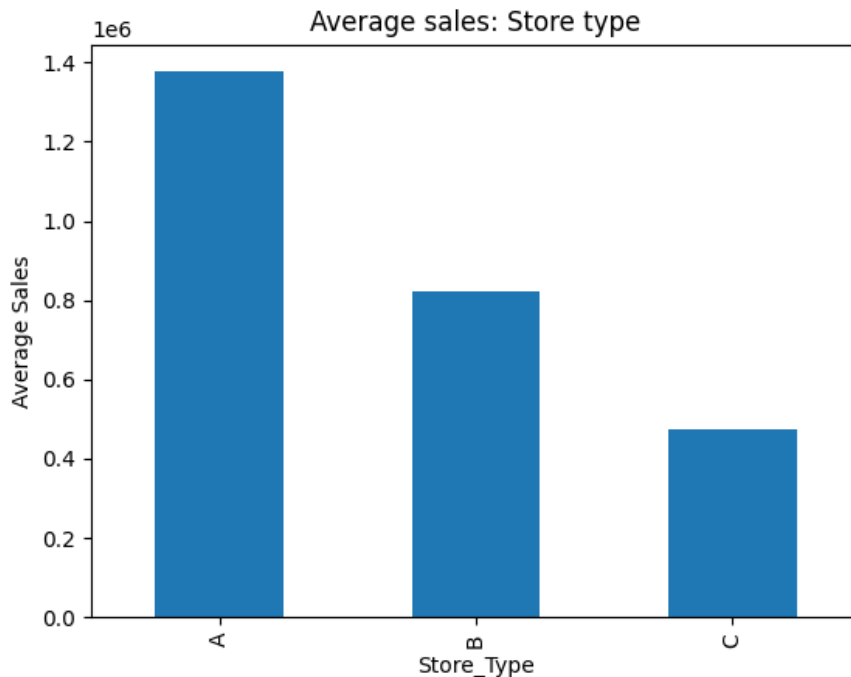
5. Se realizó una gráfica de dispersión entre Unemployment y Weekly\_Sales. En este caso, a pesar de que hubo gran variabilidad entre los valores máximos y mínimos de ventas respecto a la tasa de desempleo, se detectó que aproximadamente a partir de una tasa de desempleo de 10 en adelante, las ventas alcanzaban un rango máximo que no sobrepasaban. En contraste, con tasas más bajas, aunque también había valores mínimos, existía mayor variación en las ventas, llegando a valores mucho más altos. Por tanto, se concluyó que, para evitar encapsular el rango de ventas en aproximadamente la mitad del rango total, es mejor no superar una tasa de desempleo de 10. Sin embargo, esto no garantiza que un valor menor de esta variable asegure un desempeño alto en ventas.
6. Se realizó una gráfica de dispersión entre Size y Weekly\_Sales. Se detectó que las ventas mínimas más considerables ocurrieron en tiendas con un Size aproximadamente superior a 220,000, con un rango de valores que incluso bordeaba los máximos. También se observó que, cuanto menor era el Size, existía una tendencia hacia rangos bajos de ventas. A medida que aumentaba el Size, el rango de ventas seguía siendo variable, pero con una tendencia hacia un valor máximo y un valor mínimo mayor que los anteriores. Esto refuerza lo indicado por la matriz de correlación, donde Size mostró una correlación de 0.8, lo que comprueba que, a mayor tamaño de la tienda, incrementa la posibilidad de obtener mayores ventas, y los mínimos de ventas son más aceptables.

### Gráficos de barras



```
holiday_sales = final_data.groupby('IsHoliday')['Weekly_Sales'].mean()
holiday_sales.plot(kind='bar')
plt.title('Average Sales: Holidays and Non-Holidays')
plt.xlabel('Is it a holiday?')
plt.ylabel('Average Sales')
plt.show()
```

Se realizó una gráfica de barras entre IsHoliday y Weekly\_Sales. Dado el tipo de variable que es IsHoliday, se consideró que lo más adecuado era realizar un gráfico de barras que mostrara el promedio de ventas en semanas con festivos y sin festivos. Los resultados no mostraron una gran diferencia: se observó un mayor promedio de número de ventas cuando había festivo, pero no muy distante de cuando no lo había. En el futuro, si se espera incrementar las ventas, lo ideal sería que coincidieran con semanas festivas, aunque esta diferencia no es considerable.

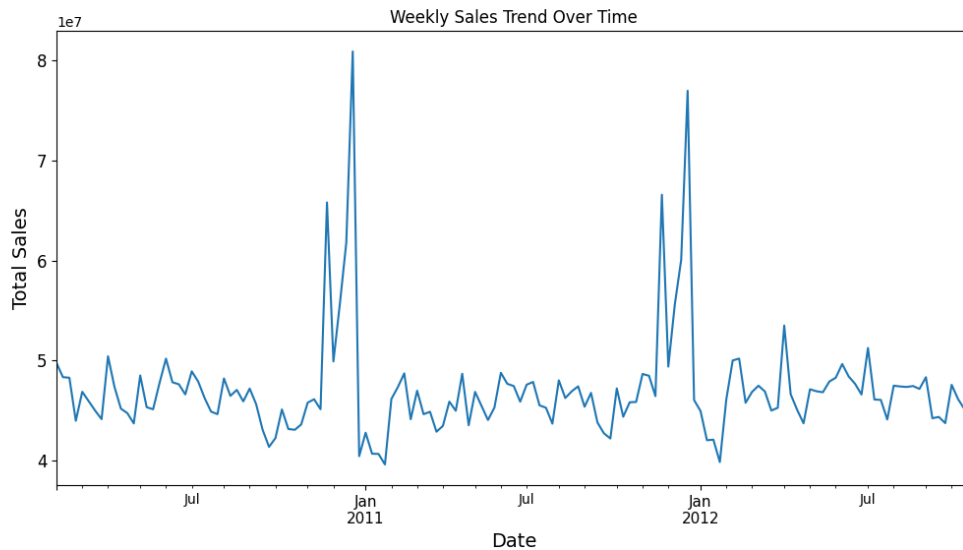


```
storeType_Sales=final_data.groupby("Store_Type")["Weekly_Sales"].mean()
storeType_Sales.plot(kind='bar')
plt.title('Average sales: Store type')
plt.xlabel('Store_Type')
plt.ylabel('Average Sales')
plt.show()
```

Se realizó una gráfica de barras entre Store\_Type y Weekly\_Sales. Debido a que Store\_Type es una variable de tipo cadena, no se incluyó en la matriz de correlación. Sin embargo, considerando que la variable Size ha mostrado gran influencia en las ventas promedio, y dado que ambas provienen de la misma tabla, era necesario buscar una relación entre Store\_Type

y Weekly\_Sales. Para ello, se agruparon los datos por tipo de tienda y se calculó el promedio de ventas correspondiente. Los resultados mostraron que las tiendas tipo A tienen un promedio de ventas superior, incluso mayor que la suma de los promedios de las tiendas tipo B y C. Por lo tanto, es evidente que las tiendas tipo A son preferibles para mantener un promedio elevado de ventas, independientemente de lo que represente la categoría o tipo especial de tienda.

### Gráfico de líneas

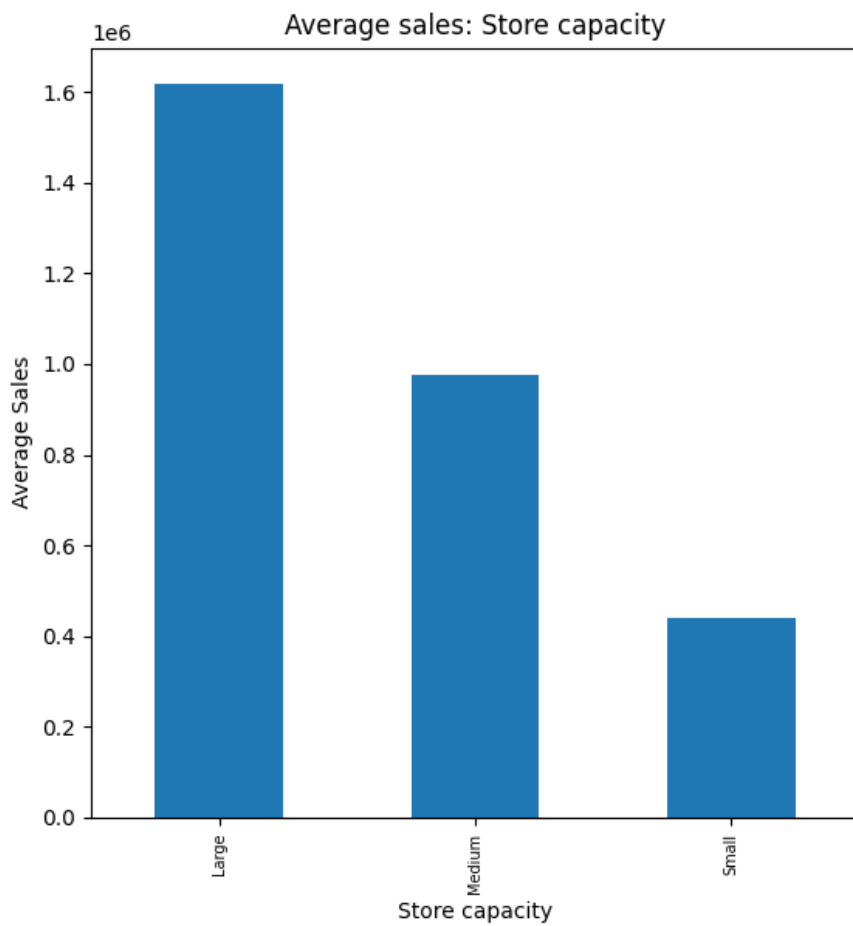


```
sales_by_date = final_data.groupby('Sales_Week_Date')['Weekly_Sales'].sum()
sales_by_date.plot(figsize=(12, 6))
plt.title('Weekly Sales Trend Over Time')
plt.xlabel('Date', fontsize=14)
plt.ylabel('Total Sales', fontsize=14)
plt.xticks(fontsize=11)
plt.yticks(fontsize=12)
plt.show()
```

Se realizó un gráfico de líneas entre Sales\_Week\_Date y Weekly\_Sales. En esta gráfica, primero se agrupó por fecha y se sumaron las ventas totales en todas las tiendas para cada fecha, con el objetivo de observar la tendencia de ventas a lo largo del tiempo, independientemente del tipo de tienda o venta. En términos generales, se mostró una tendencia constante en valores bajos, pero que aumentaba hasta alcanzar sus máximos en noviembre y diciembre de 2010 y 2011. Aunque el año 2012 no llegó a esos meses en los datos, si se sigue la tendencia observada, es razonable suponer que también alcanzaría valores máximos en esas fechas. Esto tiene sentido, ya que para esos meses los negocios tienden a registrar mayor flujo de clientes y ventas.

### 5.3 Análisis Especifico

#### Categorización de variable Size

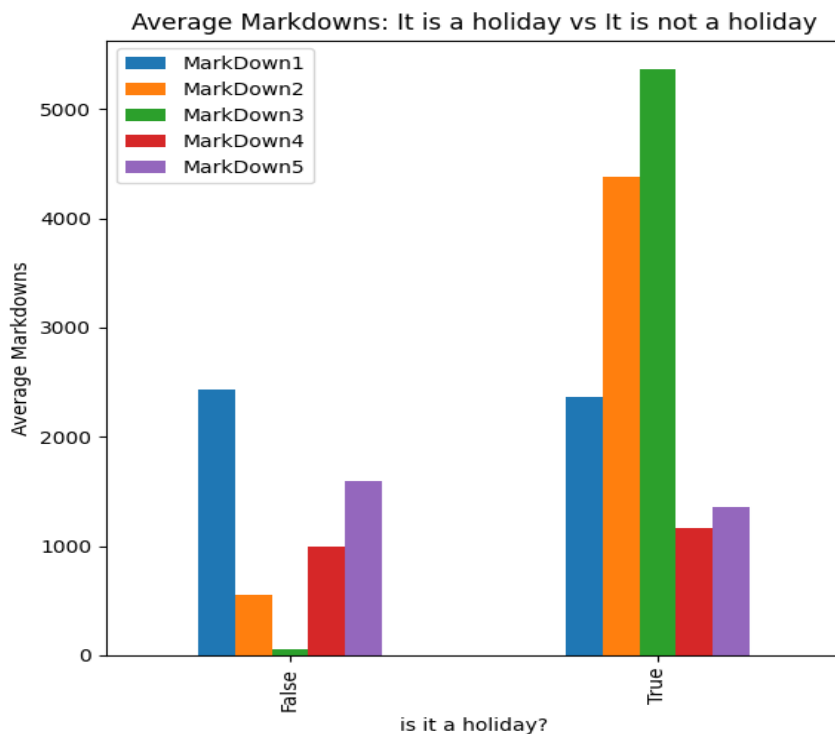


```
def Capacity_Type(x):  
    if x < 80000:  
        return 'Small'  
    elif x > 160000:  
        return 'Large'  
    else:  
        return 'Medium'  
  
final_data['Capacity_Type']=final_data["Size"].apply(Capacity_Type)  
sales_by_Capacity=final_data.groupby("Capacity_Type")["Weekly_Sales"].mean()  
sales_by_Capacity.plot(kind='bar')  
plt.title('Average sales: Store capacity')  
plt.xlabel('Store capacity',fontsize=10)  
plt.ylabel('Average Sales',fontsize=10)  
plt.xticks(fontsize=7)  
plt.yticks(fontsize=10)  
plt.show()
```

Debido a la gran influencia detectada de la variable Size en la variable Weekly\_Sales, se decidió categorizar las tiendas según su tamaño para tener una mejor idea de cuánto pueden influir en las ventas. Para ello, se realizó lo siguiente:

- Se creó una nueva columna llamada Capacity\_Type, a la que se le aplicó una función que devolvía la categoría small para capacidades inferiores a 80,000, medium para capacidades entre 80,000 y 160,000, y, por último, large para capacidades superiores a 160,000. Estos límites fueron determinados gracias a que, anteriormente, con el método describe, se pudo identificar el valor máximo de Size y, con base en ello, se definieron los límites apropiados.
- Posteriormente, se agrupó por la variable Capacity\_Type para calcular el promedio de ventas y se realizó un gráfico de barras.
- En la gráfica se puede observar que el promedio de ventas es significativamente más alto en la categoría large que en las demás, similar a lo que ocurría con la variable Store\_Type. Prácticamente, la suma de los promedios de las otras dos categorías no supera el promedio de la primera.

### IsHoliday vs Markdowns



```
isholiday_markdowns=final_data.groupby("IsHoliday")[['MarkDown1', 'MarkDown2', 'MarkDown3', 'MarkDown4', 'MarkDown5']].mean()
isholiday_markdowns.plot(kind="bar")
plt.title('Average Markdowns: It is a holiday vs It is not a holiday')
plt.xlabel('is it a holiday?')
plt.ylabel('Average Markdowns')
plt.show()
```

Debido a que la variable IsHoliday mostró una ligera tendencia a que, cuando era festivo, el promedio de ventas aumentaba ligeramente, y considerando que los Markdowns demostraron poca influencia en las ventas, aunque con algunos datos que indicaban cierta señal de influencia mínima, se procedió a analizarlas en conjunto. Para ello, se realizó lo siguiente:

- Se agrupó por la variable IsHoliday y se calculó el promedio de los Markdowns según si era festivo o no para cada uno. Posteriormente, se realizó un gráfico de barras.
- En la gráfica se puede observar que los Markdowns 2, 3 y 4 incrementan sus promedios cuando hay festivo en comparación con cuando no lo hay, siendo los aumentos del 2 y el 3 considerablemente destacados. Por otro lado, los Markdowns 1 y 5 disminuyen sus promedios cuando hay festivo en comparación con cuando no lo hay, pero estas disminuciones son bajas. Esto sugiere que estas variables están orientadas a representar los descuentos según los productos.

## 6. Conclusiones

- ✓ La variable Size, que indica el tamaño de la tienda, demostró ser la que más influye en la cantidad de ventas. Por lo tanto, se debe adoptar un enfoque que apunte a contar con un espacio mínimo de 160,000, como se demostró con anterioridad al categorizar esta variable. A partir de este valor se tiene el mayor promedio de ventas dentro de la variable.
- ✓ Store\_Type mostró que las tiendas tipo A son las que cuentan con el mayor promedio de ventas, significativamente por encima de los otros dos tipos. En el futuro, nuevos inversores en el mundo de las tiendas deberían considerar las tiendas tipo A como su opción para montar un negocio. En el caso de las tiendas que ya están instaladas, se debería realizar una transición progresiva hacia este tipo de tiendas.
- ✓ La variable IsHoliday mostró que existe un aumento en el promedio de ventas cuando es festivo en comparación con cuando no lo es. Aunque la diferencia no es muy alta, en el mundo de la inversión y la rentabilidad, obtener la mayor cantidad de ingresos siempre es un factor para tener en cuenta. Por tal motivo, los encargados de las tiendas deben considerar que se espera un mayor volumen de ventas en las semanas con festivos y, con base en esto, planificar sus recursos para estar preparados para recibir un mayor número de ventas.
- ✓ Los Markdowns mostraron poca influencia en cuanto a la cantidad de ventas y una tendencia a aumentar en su mayoría durante los días festivos. Teniendo en cuenta todo esto, una alternativa para reducir el impacto de los descuentos sería no aumentarlos excesivamente, dado que no influyen considerablemente en las ventas. Incluso, sería más conveniente que en las semanas con festivos no se ofrezcan descuentos y, en cambio, se reserven para semanas sin festivos como una estrategia para captar la atención de los consumidores.
- ✓ Las ventas en su mayoría aumentan entre los meses de noviembre y diciembre, como es lógico por el fin de año. La diferencia es considerable respecto a los demás meses. Teniendo esto en cuenta, las tiendas deberían enfocar sus esfuerzos en aprovechar estos meses para obtener mayores rendimientos financieros, ya que la afluencia de clientes es mayor. En el resto de los meses, las tiendas deberían implementar estrategias de ventas alternativas para atraer a los clientes y aumentar sus ventas.

Estas estrategias podrían incluir campañas de marketing, descuentos o incluso una reubicación en la zona donde se encuentran.

- ✓ Las variables Temperature, Fuel\_Price, CPI y Unemployment no demostraron gran influencia en las ventas. Sin embargo, esto no debe considerarse una mala noticia, ya que son variables que no se pueden controlar, pues dependen de factores como la naturaleza o las políticas gubernamentales. En el futuro, deberían tomarse en cuenta como factores relevantes, aunque con menor prioridad en comparación con otras variables más influyentes.
- ✓ Si es posible, se deberían realizar remodelaciones en la infraestructura de la tienda para aumentar su capacidad y, por ende, las ventas. Para generar un impacto, esta reestructuración debería estar lista para los meses de noviembre y diciembre, que son los meses con mayor flujo de ventas.

## 7. Anexos

**Librerías utilizadas en Python para el desarrollo del proyecto:**

```
import pandas as pd
import seaborn as sns
import matplotlib.pyplot as plt
```

**Versión de Python:** Python 3.13.0

**Enlace de los datos originales:**

<https://www.kaggle.com/datasets/manjeetsingh/retaildataset/discussion?sort=hotness>