

```
#-- Importação de bibliotecas
import numpy as np
from datetime import date,datetime
import pandas as pd
import matplotlib.pyplot as plt
import seaborn as sns
```

```
df = pd.read_csv('/tmp/Walmart.csv')
####
```

```
Incluindo novas colunas ano e semana
df['Date'] = pd.to_datetime(df['Date'], format = '%d-%m-%Y')
df['Week'] = df['Date'].dt.isocalendar().week
df['year'] = df['Date'].dt.isocalendar().year
```

```
print(df)
####
```

```
Qual o preço médio de venda semanal por loja?
df.groupby(['Store','Week'])['Weekly_Sales'].mean()
```

```
Qual o atributo que tem maior correlação (positiva ou negativa) com o
número de
vendas? (E que faça sentido)
Year e Fuel_Price 0,07
Holiday e Weekly_Sales -0.04
###
```

```
Da loja com maior venda acumulada (soma de vendas de todo o período),
quantas semanas do ano ele ultrapassou a média do período?
```

```
Super Bowl 12-Feb-10, 11-Feb-11, 10-Feb-12 Labour Day 10-Sep-10, 09-Sep-
11, 07-Sep-12
Thanksgiving 26-Nov-10,
25-Nov-11, 23-Nov-12
Christmas 31-Dec-10, 30-Dec-11, 28-Dec-12
```

```
Faça um gráfico com as vendas máximas, mínimas e médias de todas as lojas
ao longo do tempo.
Ressalte de alguma forma, o período de feriados mencionados acima.
```

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

```
# Carrega os dados das vendas em um DataFrame
df =df
sns.set_style('whitegrid')
plt.subplots(figsize=(18,6))
# Converte a coluna de data para o tipo datetime
df['Date'] = pd.to_datetime(df['Date'])
```

```
# Agrupa os dados por data e calcula as vendas máximas, mínimas e médias
para cada dia
df_grouped = df.groupby('Date')['Weekly_Sales'].agg(['max', 'min',
'mean'])
```

```
# Cria um gráfico de linha com as vendas máximas, mínimas e médias
destacadas durante os períodos de feriado
plt.plot(df_grouped.index, df_grouped['mean'], label='Média')
```

```

plt.plot(df_grouped.index, df_grouped['max'], label='Máximo')
plt.plot(df_grouped.index, df_grouped['min'], label='Mínimo')

# Destaca os períodos de feriado no gráfico
feriados = ['2010-02-12', '2011-02-11', '2012-02-11',
            '2010-09-10', '2011-09-09', '2012-09-07',
            '2010-11-26', '2011-11-25', '2012-11-23',
            '2010-12-31', '2011-12-30', '2012-12-2'
            ]

for feriado in feriados:
    plt.axvline(x=pd.to_datetime(feriado), color='red', linestyle='--')

# Configurações do gráfico
plt.xlabel('Data')
plt.ylabel('Vendas')
plt.title('Vendas máximas, mínimas e médias de todas as lojas')
plt.legend()
plt.show()

#Pensando na expansão de uma das lojas, que loja você escolheria e o por
que?
df3 =
df.groupby(['Store', 'year'])['Weekly_Sales'].agg(['max', 'min', 'mean', 'sum
'])
print(df3['max'].max())

df3[df3["max"] == 3818686.45]
print(df[df["Store"]==14])
Escolhe loja 14 por que foi a loja que vendeu mais em 2010.

Análise final + Sugestão de imóveis
Minha sugestao e a loja 14 em 2010 deve a maior soma de vendas em relacao
as outras lojas.

```