Vivek A

Click here to view the github

Walmart

Sales Analysis

Data Analysis project

Problem Statement

A retail store that has multiple outlets across the country are facing issues in managing the inventory - to match the demand with respect to supply.

a. If the weekly sales are affected by the unemployment rate, if yes - which stores are suffering the most?

b. If the weekly sales show a seasonal trend, when and what could be the reason?

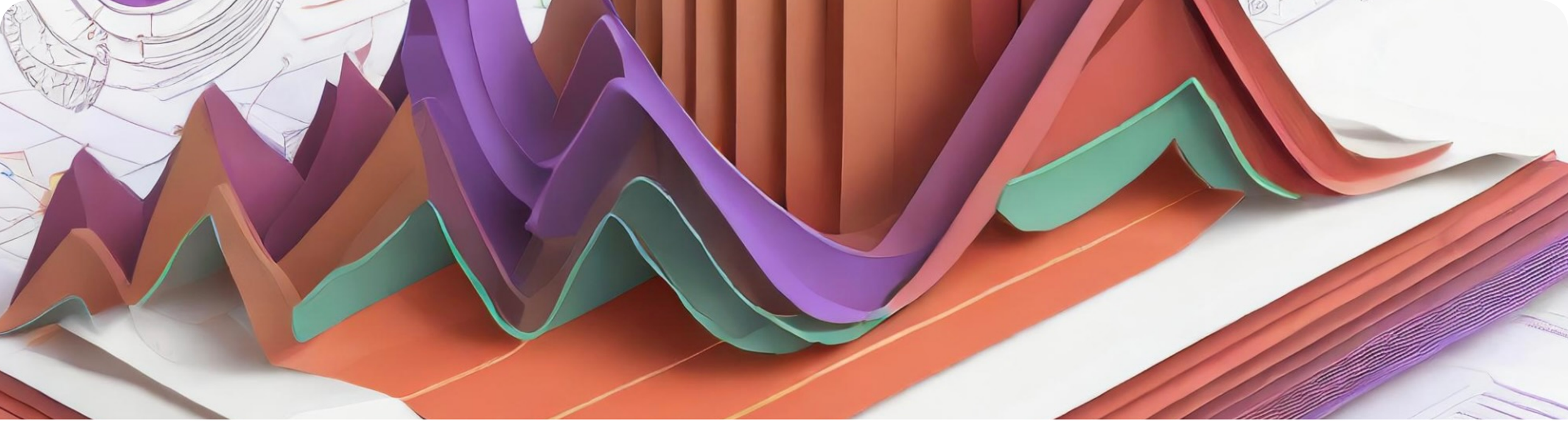
c. Does temperature affect the weekly sales in any manner?

Problem Statement

d. How is the Consumer Price index affecting the weekly sales of various stores? e. Top performing stores according to the historical data.

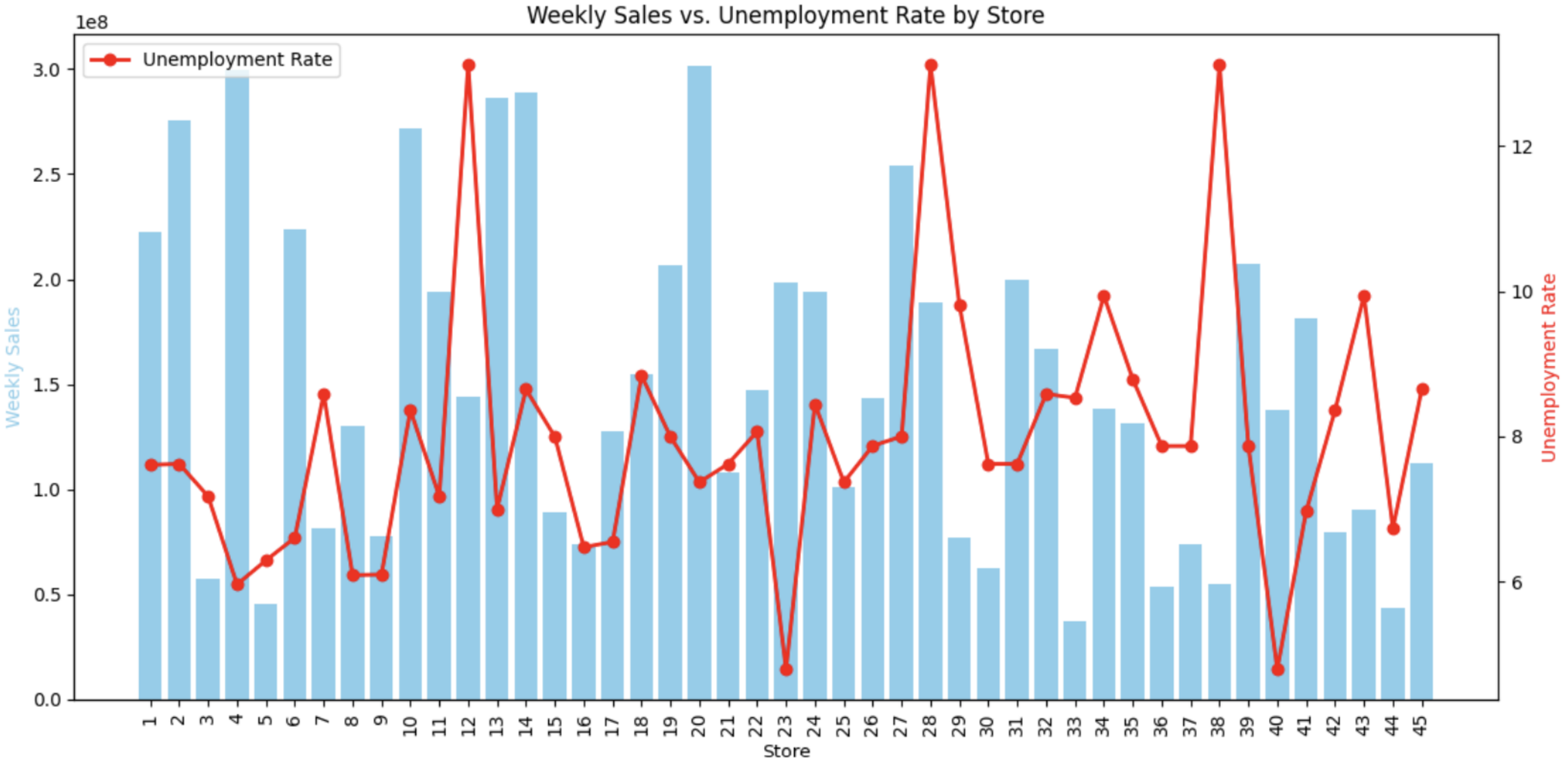
f. The worst performing store, and how significant is the difference between the highest and lowest performing stores.

Use predictive modeling techniques to forecast the sales for each store for the next 12 weeks.



Data Analysis & Visualizations

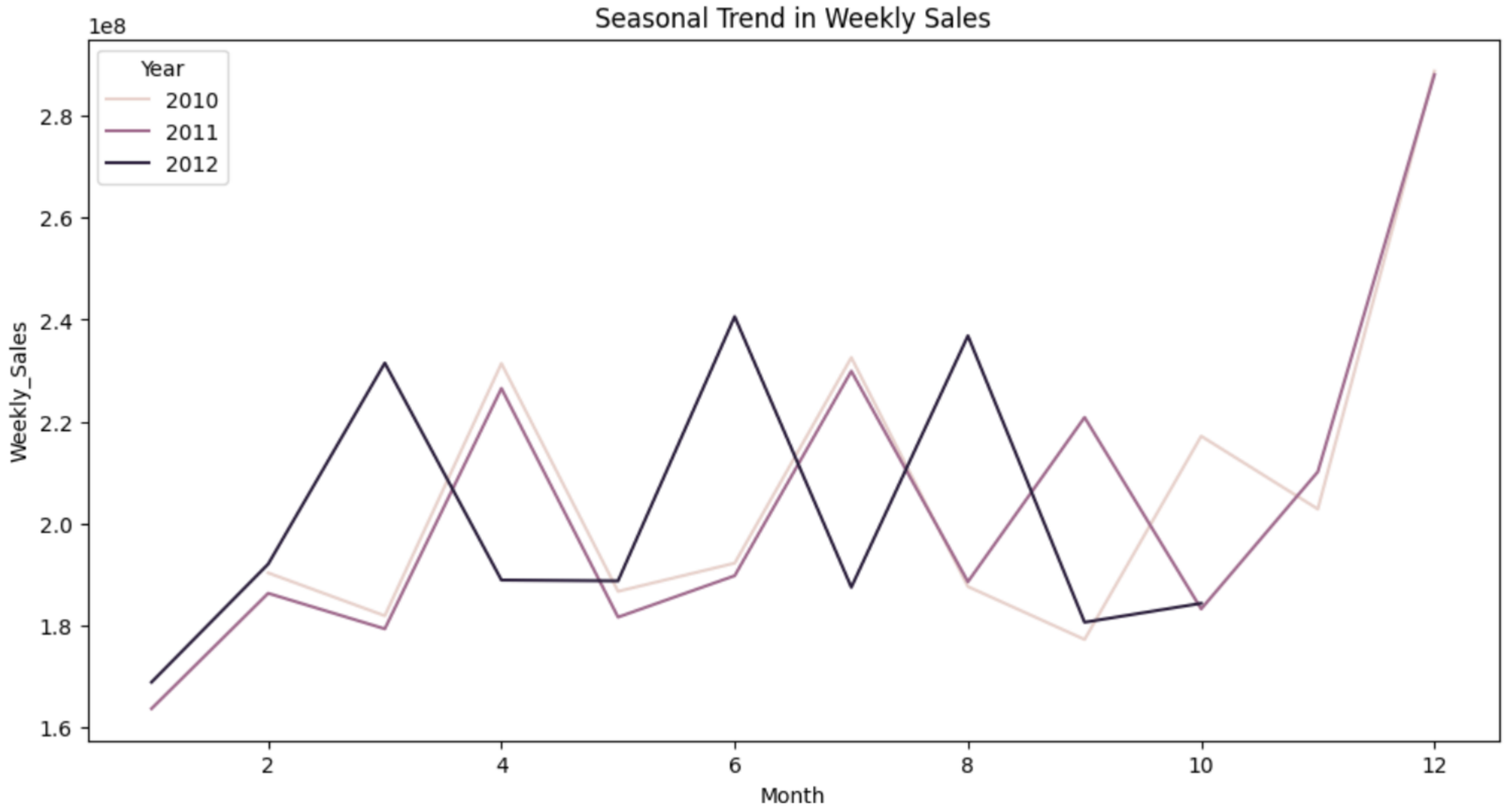
Weekly Sales vs. Unemployment Rate



Weekly Sales vs. Unemployment Rate

- Most of the time, the unemployment rate is stable accross the store
- store 38 suffered high unemployment maybe because of the low sales in weeks
- store 23 and 40 have very low unemployment rate, but with mediocore sales
- surprisingly, store 12 and 28 also have high unemployment even with good weekly sales

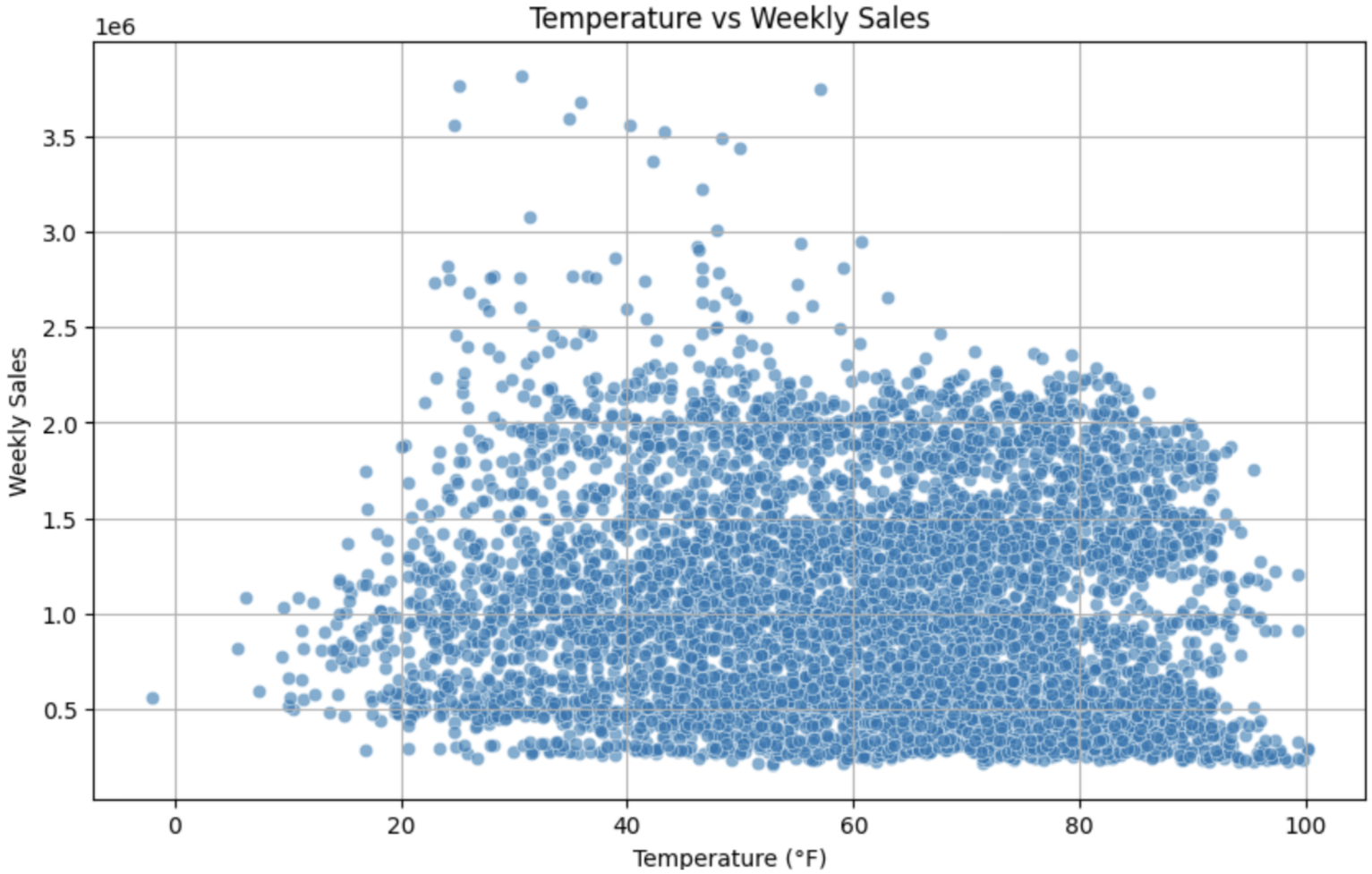
Seasonal Trend in Weekly Sales



Seasonal Trend in Weekly Sales

- Weekly sales have peaked in december
- The reason for that maybe more people comimg to shop because of the christmas and new year. Year ending sales may also be the reason

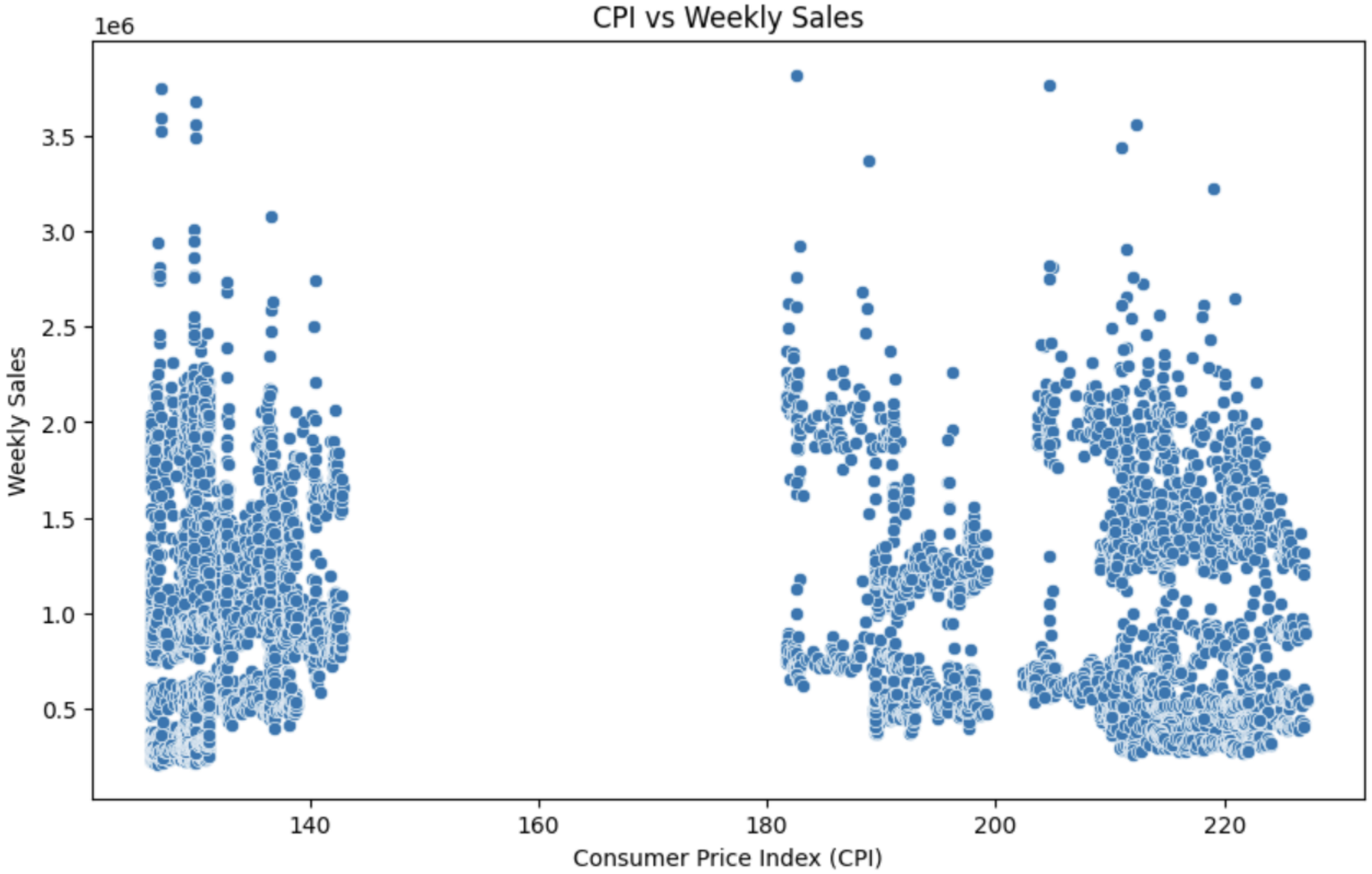
Effect of Temperature on Weekly Sales



Effect of Temperature on Weekly Sales

- A correlation of -0.0638 indicates a very weak negative relationship, suggesting that temperature has very little impact on weekly sales. This could be due to other factors influencing sales more strongly than temperature.

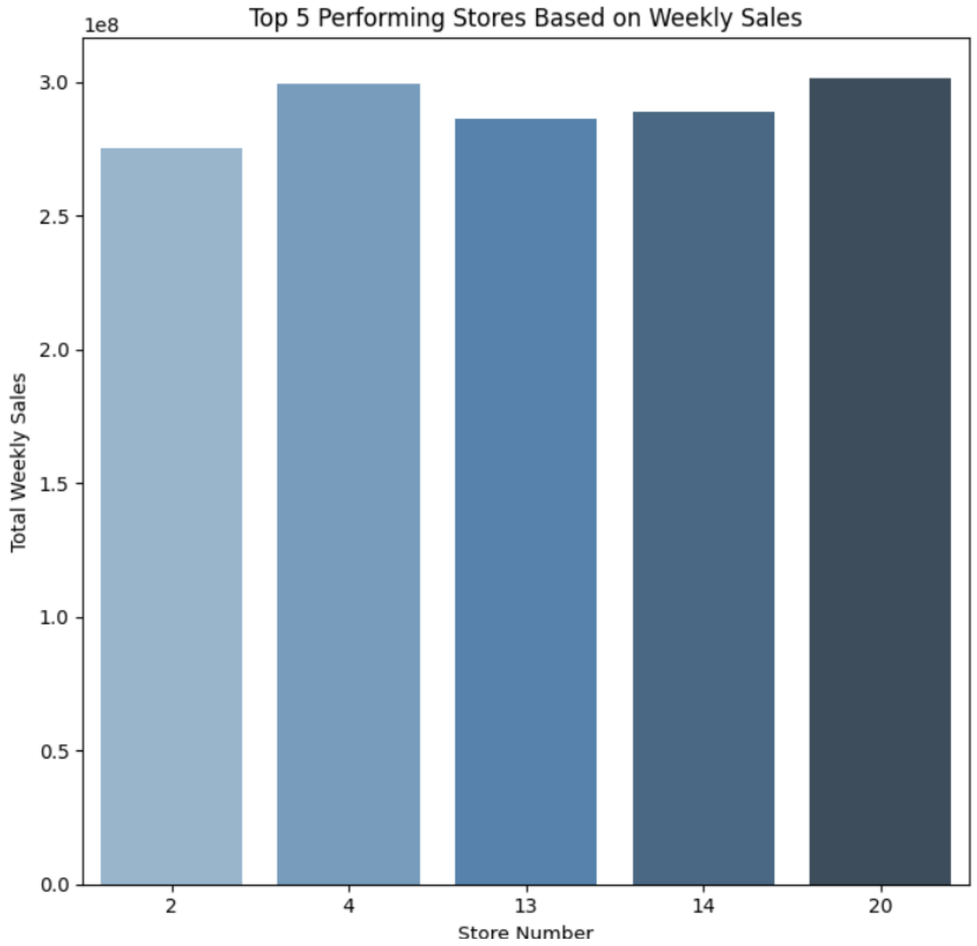
CPI vs. Weekly Sales

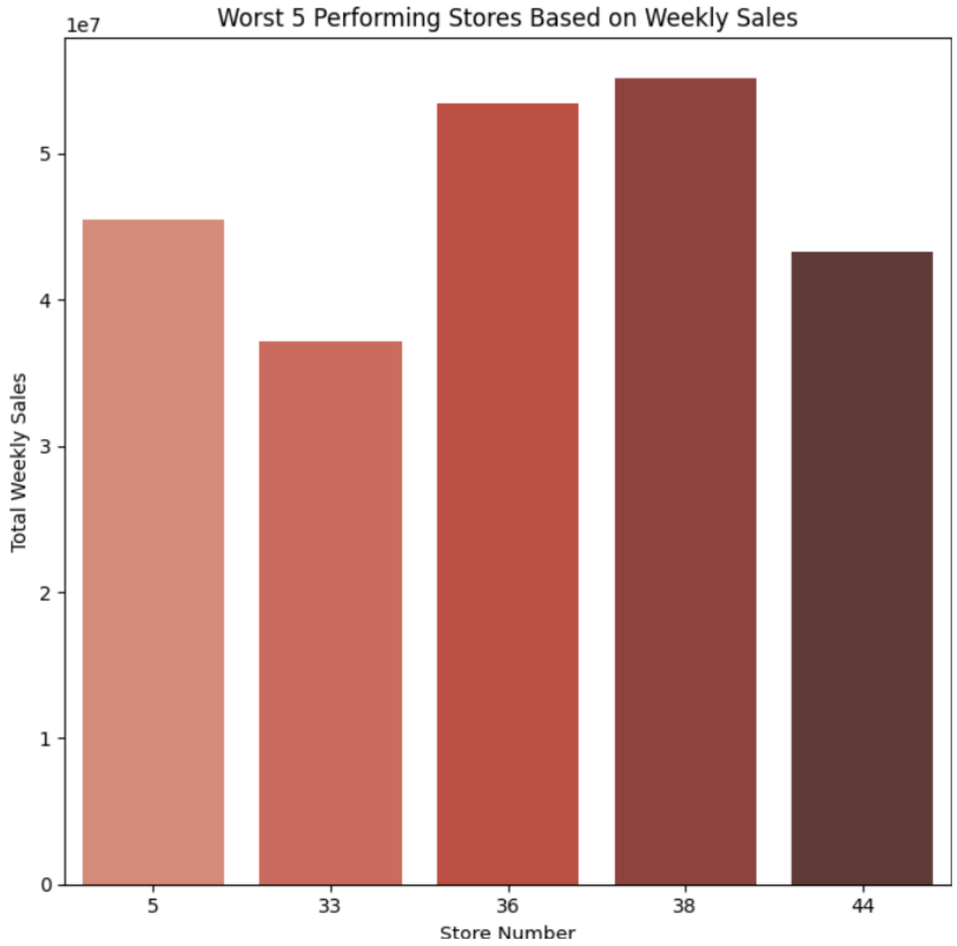


CPI vs. Weekly Sales

- CPI have very little direct impact on the weekly sales in this dataset
- There is no data with the cpi from 140 to 180

Top and worse Performing Stores





Top and worse Performing Stores

```
highest_performing_store = store_sales_sorted.iloc[0]

lowest_performing_store = store_sales_sorted.iloc[-1]

difference_in_sales = highest_performing_store['Weekly_Sales'] - lowest_performing_store['Weekly_Sales']

percentage_difference = (difference_in_sales / highest_performing_store['Weekly_Sales']) * 100

# results
print(f"Highest Performing Store (Store {highest_performing_store['Store']}): {highest_performing_store['Weekly_Sales']}")
print(f"Lowest Performing Store (Store {lowest_performing_store['Store']}): {lowest_performing_store['Weekly_Sales']}")
print(f"Absolute Difference in Sales: {difference_in_sales:.2f}")
print(f"Percentage Difference: {percentage_difference:.2f}%")
```

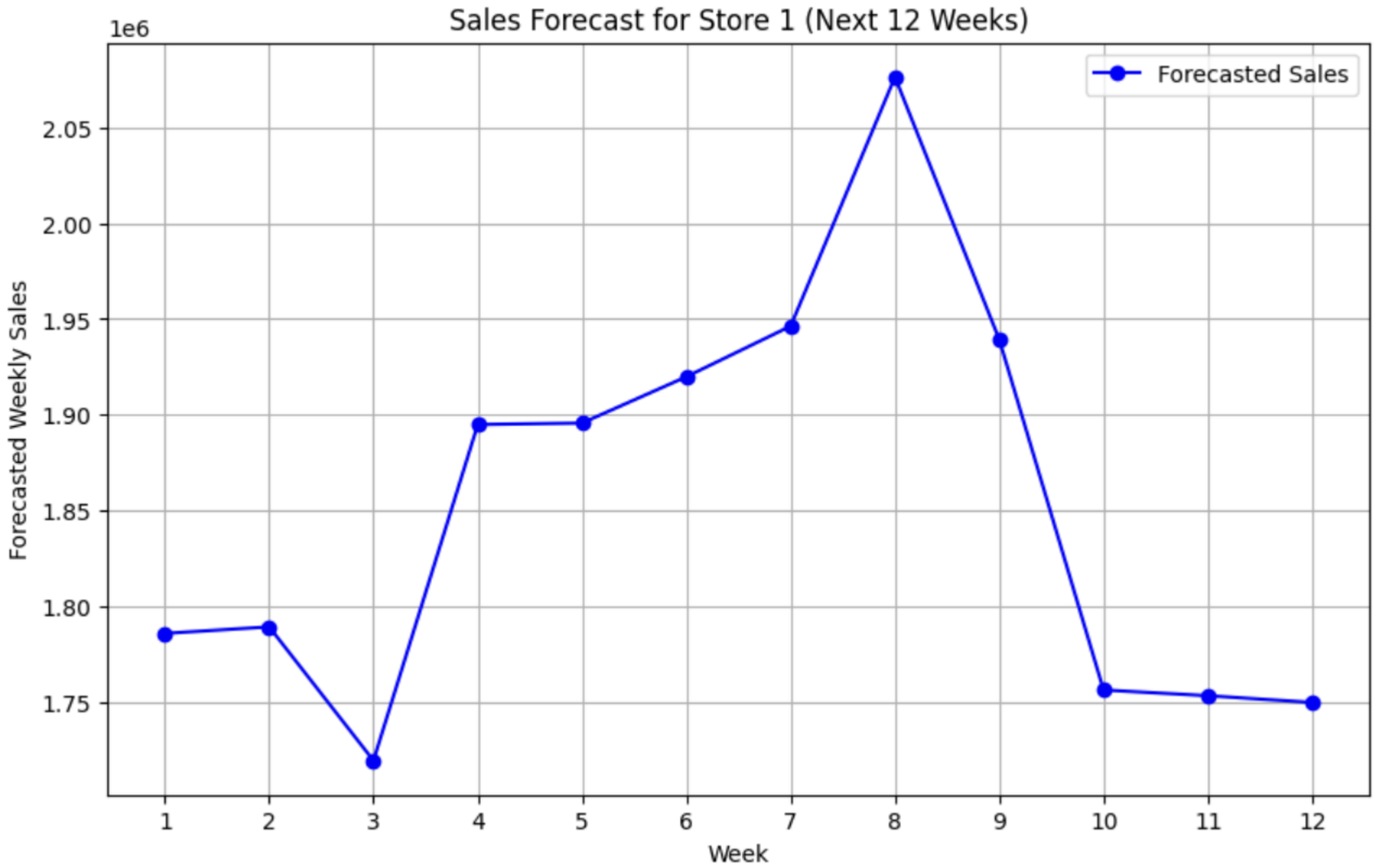
```
Highest Performing Store (Store 20.0): 301397792.46
Lowest Performing Store (Store 33.0): 37160221.96
Absolute Difference in Sales: 264237570.50
Percentage Difference: 87.67%
```

- stores 2,4,13,14,20 are the top perdorming stores
- Stores 5,33,36,38 and 44 are the worsr performing stores
- There is a 87.67% difference between top and worse performing stores

Top and worse Performing Stores

- stores 2,4,13,14,20 are the top perdorming stores
- Stores 5,33,36,38 and 44 are the worsr performing stores
- There is a 87.67% difference between top and worse performing stores

Forecasting the sales for another 12 weeks



Forecasting the sales for another 12 weeks

- Seems accurate becasue of the jump in december month. The rest of the store forecast recoded as subplots in next python file

Thanks

Thank you everyone who been with me with this journey.
Check the github repository for more info about the project