Retail Analysis with Walmart Data

The aim of this project is the leading retail stores in the US, Walmart, would like to predict the sales and demand accurately.

Business problem

The leading retail stores in the US, Walmart, would like to predict the sales and demand accurately. Build a baseline model to predict the forecast demand by Linear Regression.

Analytics Problem

In this problem, we have analysed a report about the data explaining the distribution of the sales and store. And the related factors using the steps listed below:

Basic Statistics tasks solution

- Which store has maximum sales

  - Store - 20 has a sales maximum = $301,397,792

- Which store has maximum standard deviation i.e., the sales vary a lot. Also, find out the coefficient of mean to standard deviation

  - Store - 14 has a maximum standard deviation = $317,569.949
  - Mean to Standard Deviation = 15.71%

- Which store/s has good quarterly growth rate in Q3’2012

  - No store has shown quarterly growth in quarter'3 2012.

- Some holidays have a negative impact on sales. Find out holidays which have higher sales than the mean sales in non-holiday season for all stores together

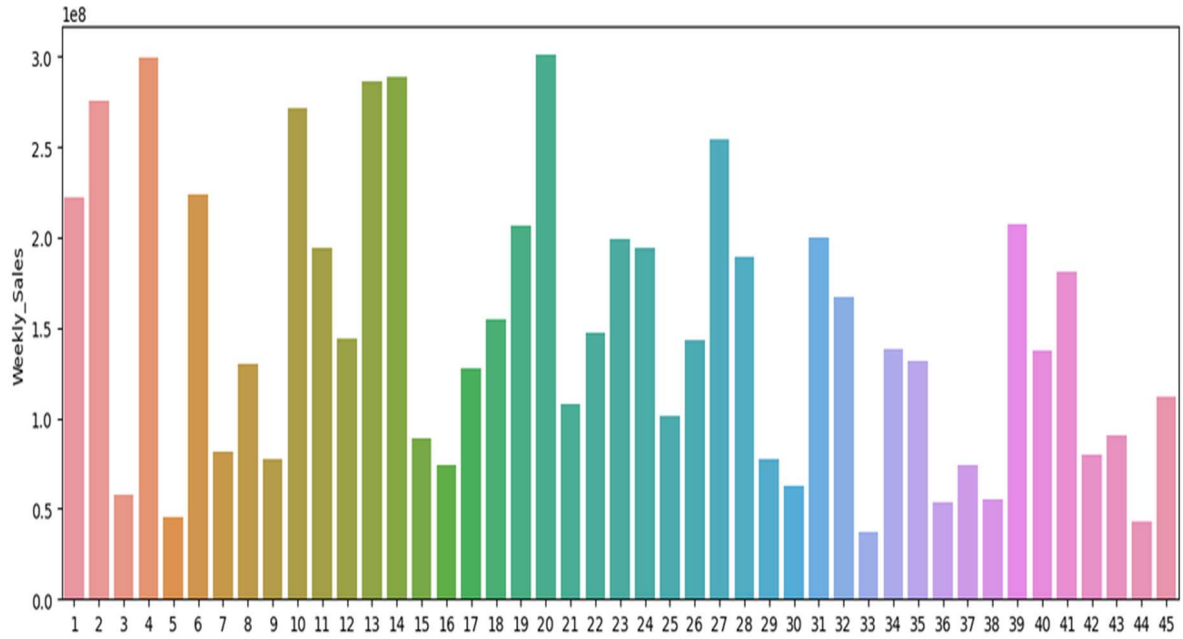
  - Thanksgiving, Super Bowl, Labour Day Sales has the sales > non-holiday sales.

- Provide a monthly and semester view of sales in units and give insights

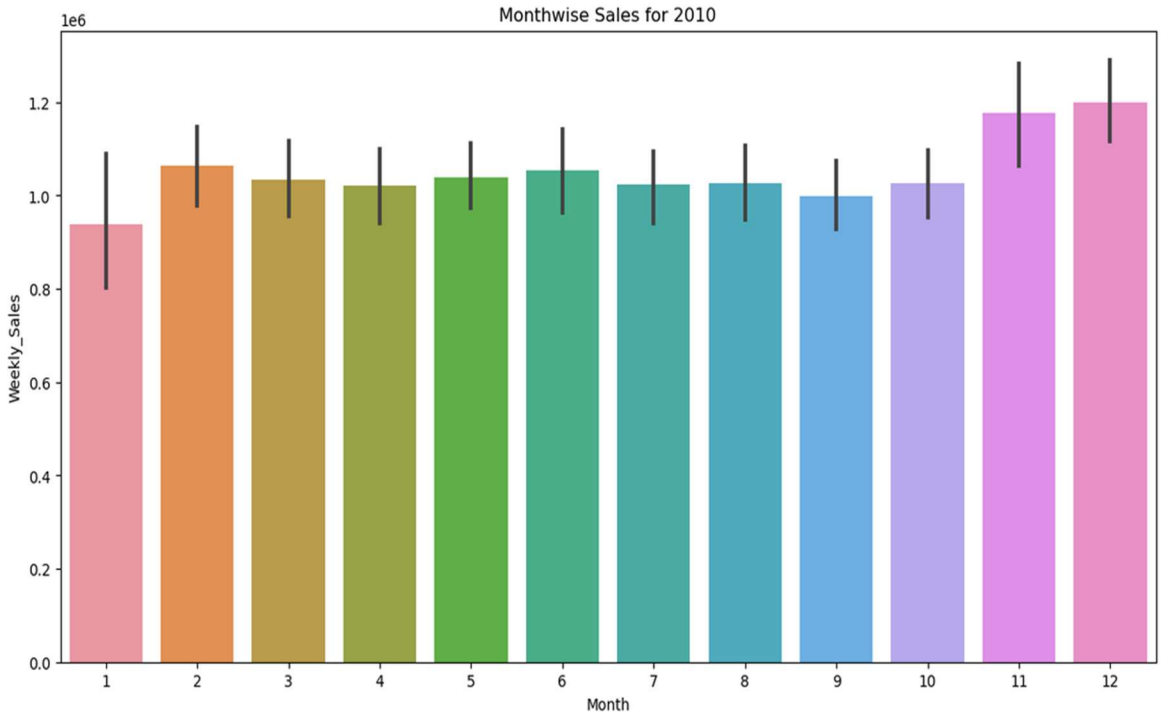
Observation

Study the store weakly sales

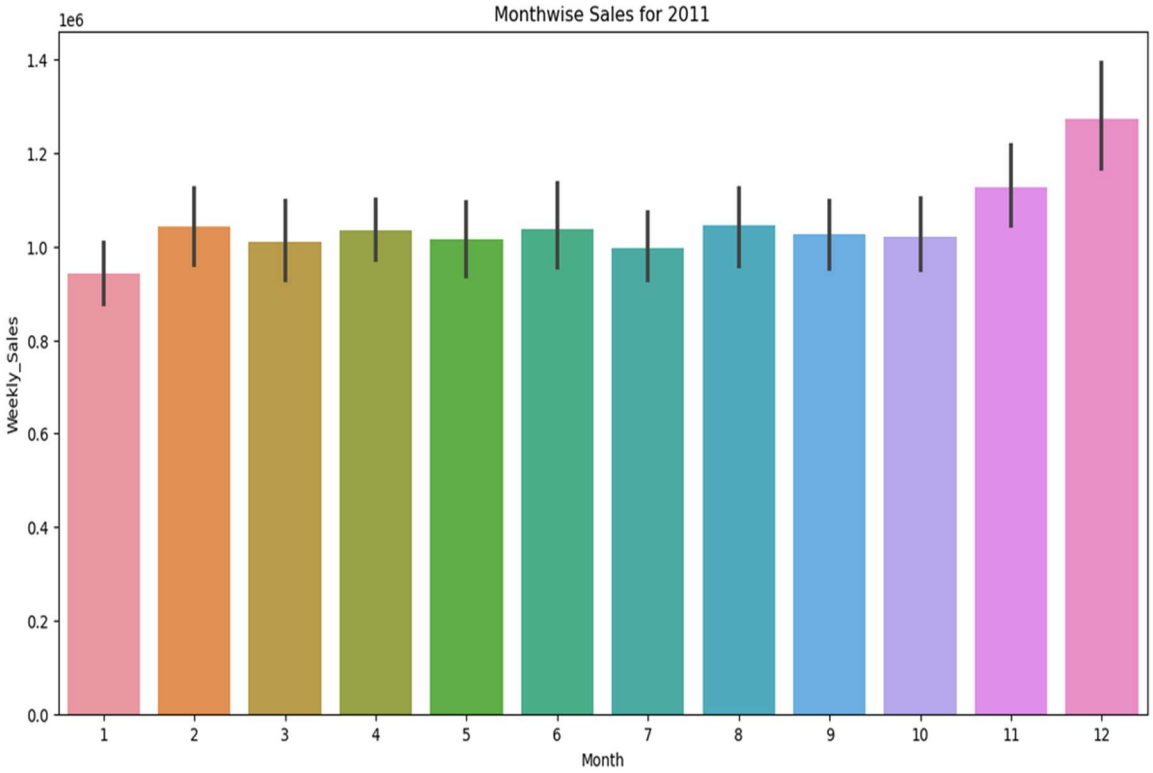
```
plt.figure(figsize=(15,5))
sns.barplot(x=df1,y=t1['Weekly_Sales'])
plt.savefig('p1.png')
```



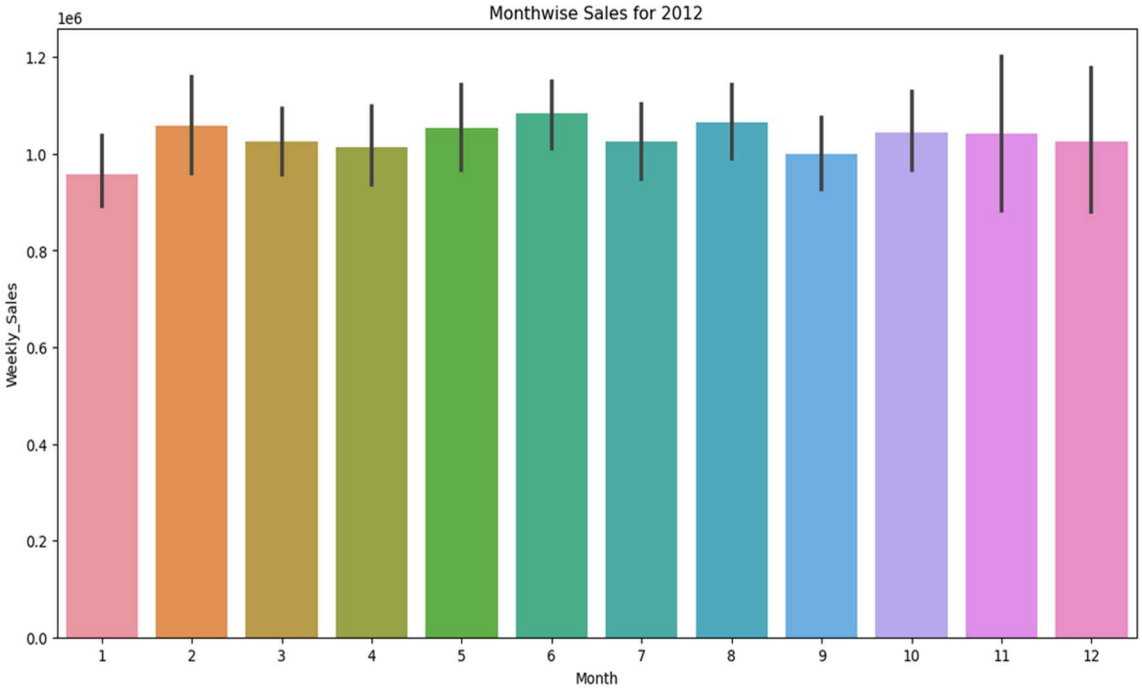
Monthwise sale for 2010



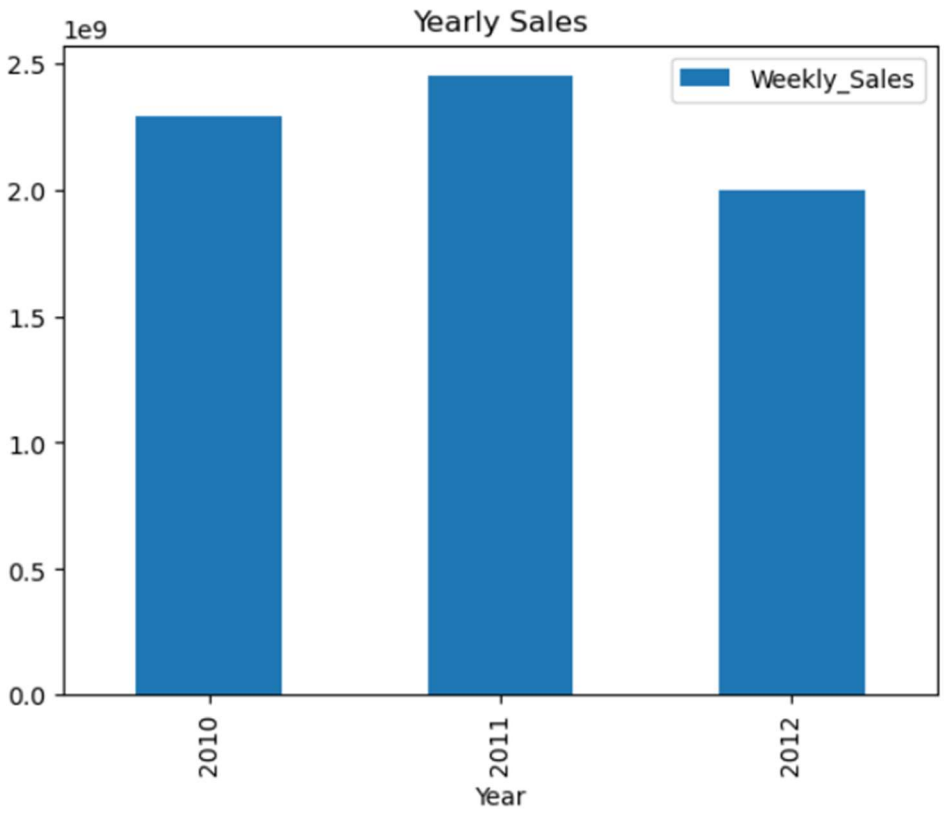
Monthwise sale for 2011



Monthwise sale for 2012



Yearly sales



Year 2011 has the highest weekly sales.

Conclusion

Linear Regression to predict the weekly sales and got below result

Mean Absolute Error: 435460.651

Mean Squared Error: 275296379013.932

Root Mean Squared Error: 524686.934