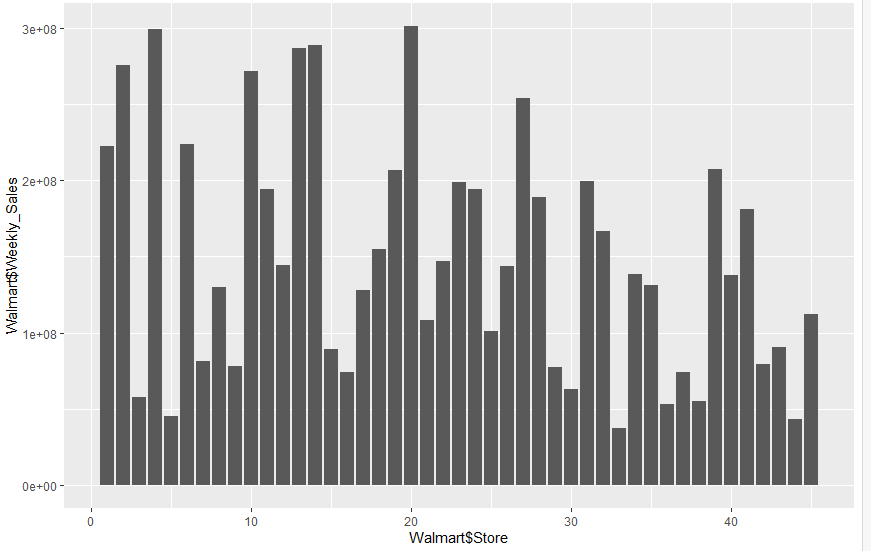
## Canisha Barron

## 3/6/2021

## Retail Analysis with Walmart

# Screenshot of Visualizations

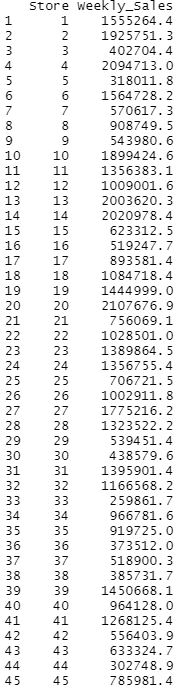
The following images are screenshots of the plots generated from the retail analysis with Walmart. Note that there is missing information from the dataset such as sales for a particular month. The analysis still provides a detailed synopsis of weekly sales for the 45 Walmart stores over the course of 2010-2012.



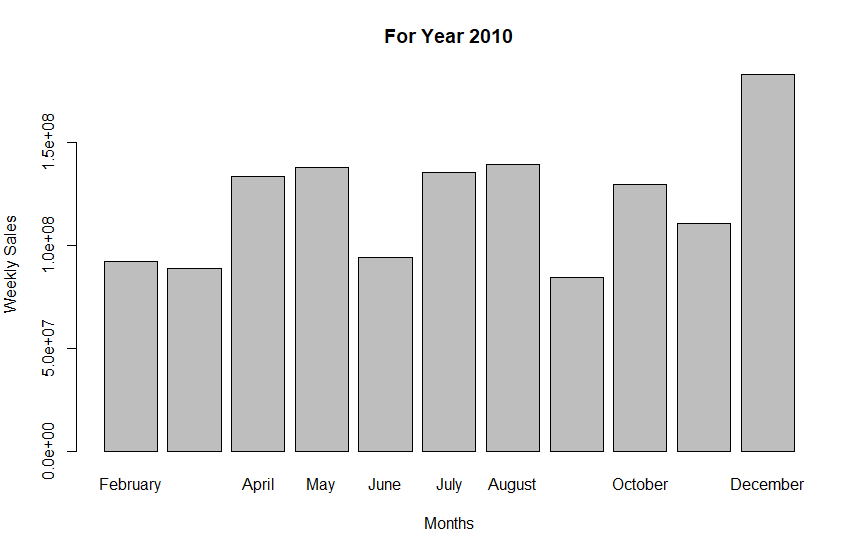
**Image 1** shows a bar chart for the weekly sales of all 45 Walmart stores based on the dataset provided. It shows that store 20 has the maximum weekly sales.

ggplot(walmart, aes(x=Store , y=Weekly\_Sales)) +

geom\_bar(stat="identity")

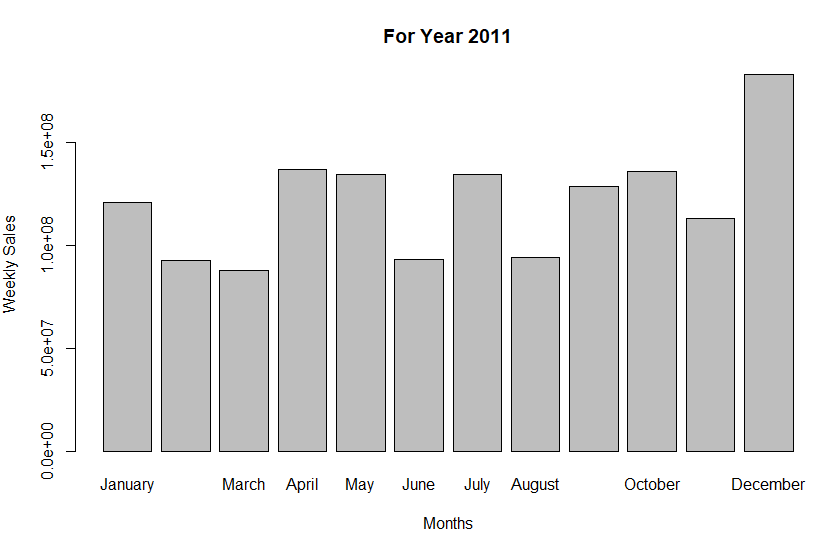


**Image 2** shows a list of all stores, providing a clearer view of actual value for weekly sales of all 45 stores. It also shows the range of the weekly sales and we can see that store 4 also has high weekly sales.



**Image 3** reflects the weekly sales per month for the year 2010. December is leading significantly in weekly sales. January data is not included based on the dataset provided.

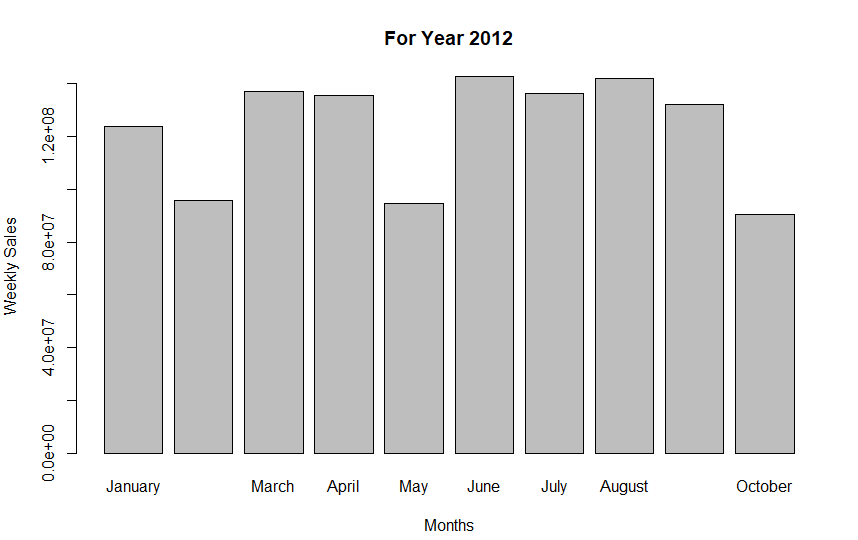
barplot(ws,names.arg=cat,xlab="Months", ylab="Weekly Sales", main="For Year 2010")



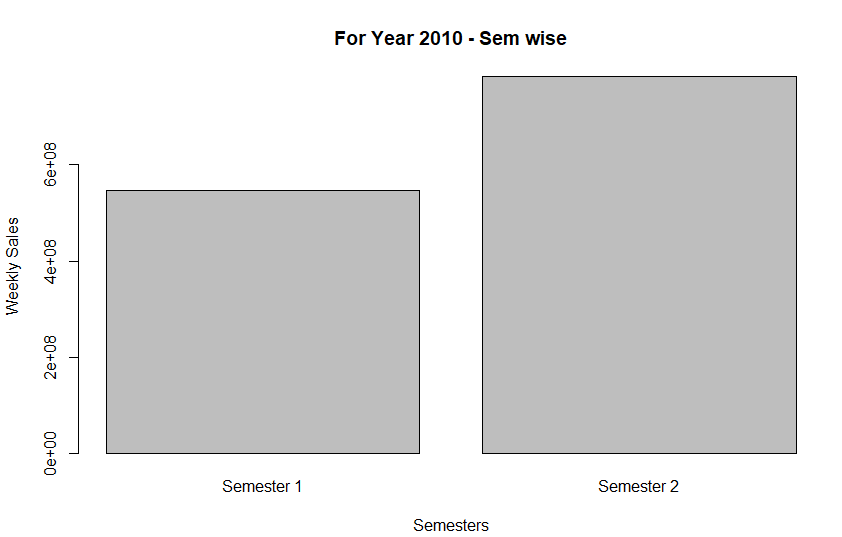
**Image 4** reflects the weekly sales per month for the year 2011. December is leading significantly in weekly sales.

barplot(ws,names.arg=cat,xlab="Months", ylab="Weekly Sales", main="For Year 2011")

**Image 5** reflects the weekly sales per month for the year 2012. To bring some relevance to this image. There is no data for November and December 2012, so it was excluded from this graph. However, the weekly sales were still analyzed and reflects that June and August produced highest with significantly high weekly sales in March and April.

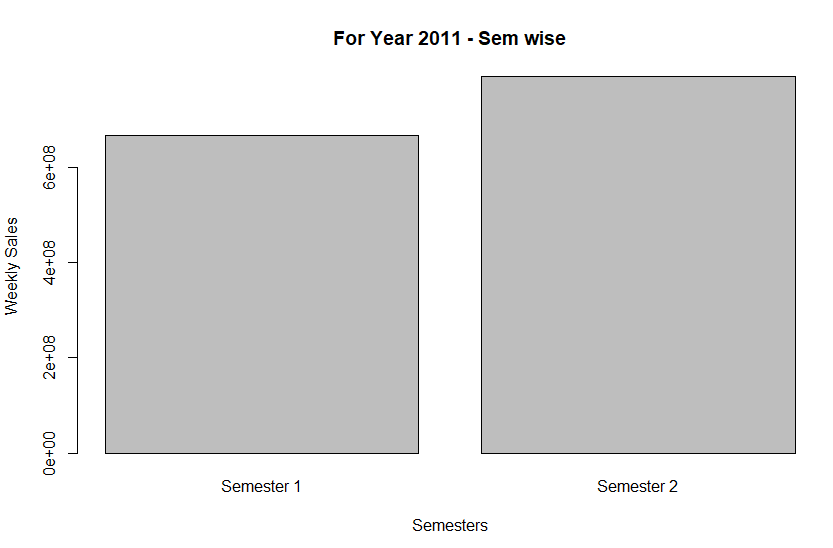


## barplot(ws,names.arg=cat,xlab="Months", ylab="Weekly Sales", main="For Year 2012")



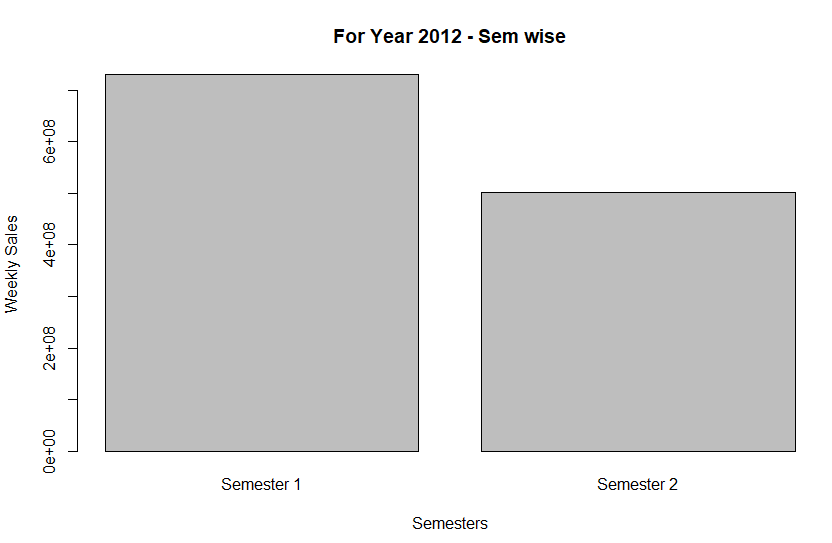
**Image 6** reflects semester weekly sales. In 2010, semester two has higher weekly sales which justifies that December falls into this category and did have the highest monthly sales for 2010.

barplot(ws,names.arg=cat,xlab="Semesters", ylab="Weekly Sales", main="For Year 2010 - Sem wise")



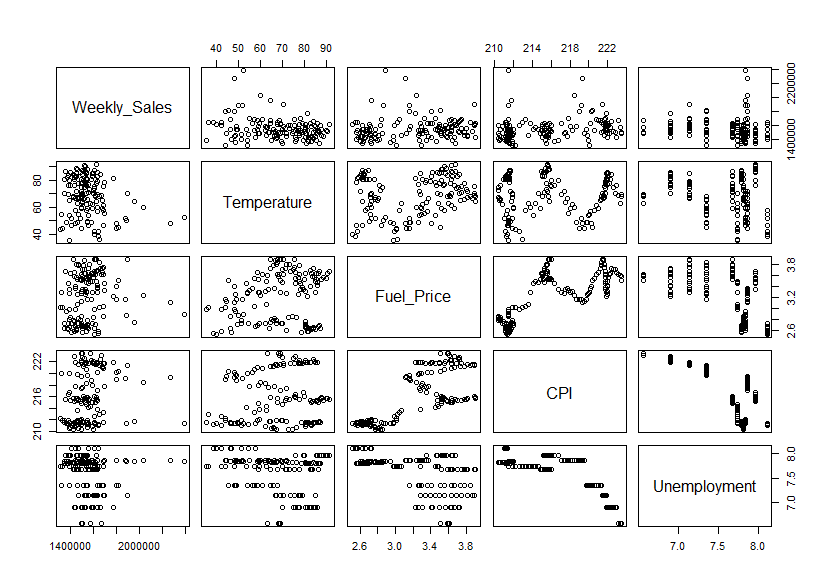
**Image 7** reflects semester weekly sales. In 2011, semester two has higher weekly sales which justifies that December falls into this category and did have the highest monthly sales for 2011.

barplot(ws,names.arg=cat,xlab="Semesters", ylab="Weekly Sales", main="For Year 2011 - Sem wise")



**Image 8** reflects semester weekly sales. In 2012, semester one has higher weekly sales which is expected due to absence of November and December months being omitted and with significantly high weekly sales in the months of March and April.

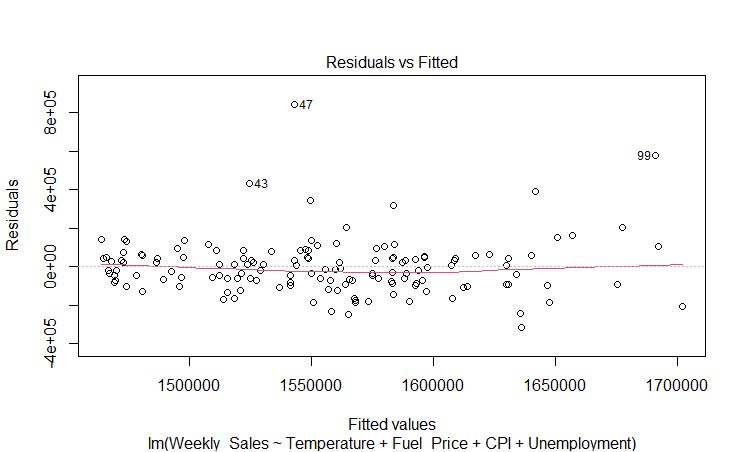
barplot(ws,names.arg=cat,xlab="Semesters", ylab="Weekly Sales", main="For Year 2012 - Sem wise")



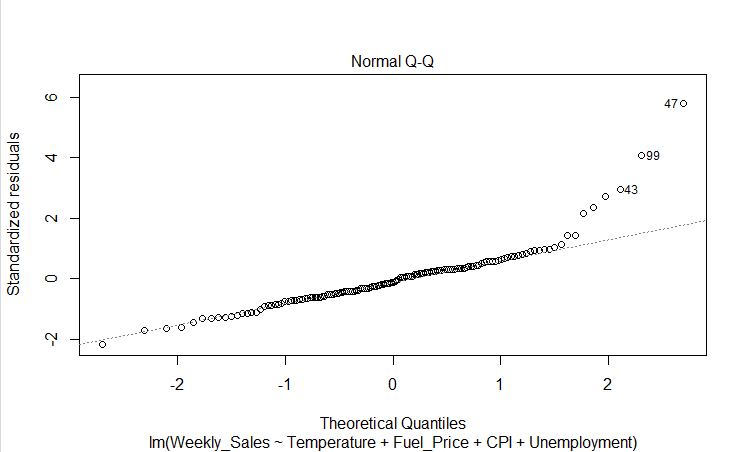
**Image 9** shows a linear model of how independent variables affect weekly sales. Weekly sales show correlation with temperature, fuel price, CPI and unemployment.

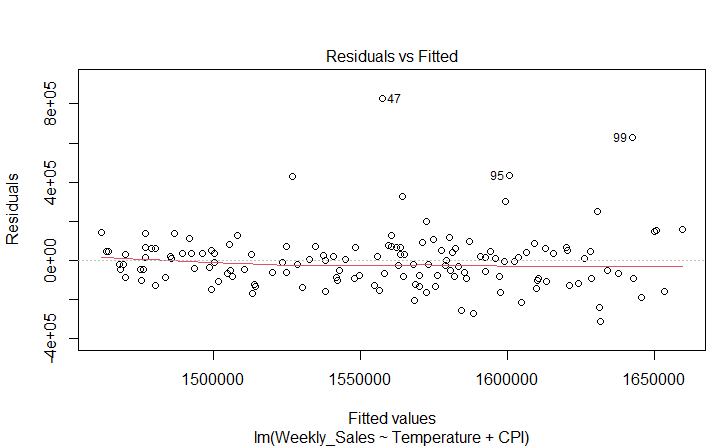
model<-lm(Weekly\_Sales~Temperature+Fuel\_Price+CPI+Unemployment, data = input)

print(model)

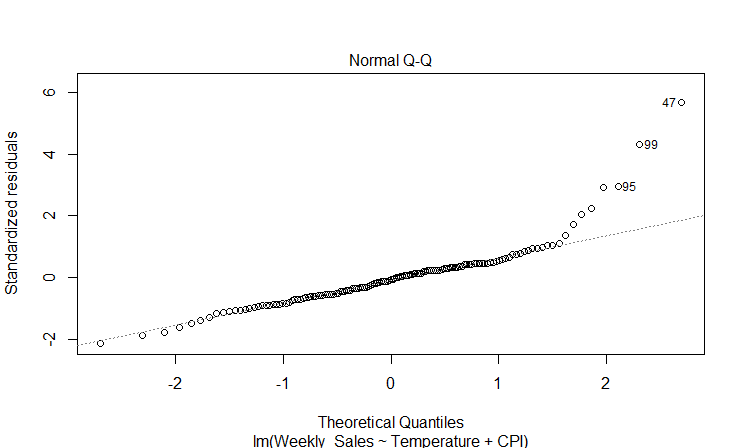


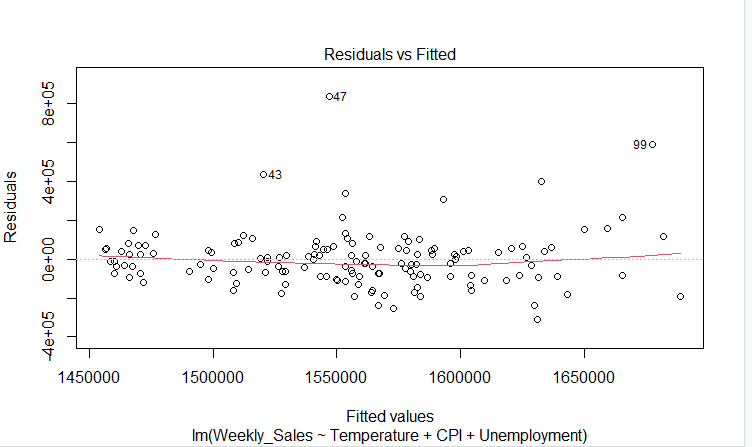
**Image 10 and 11** shows a plot of residuals vs fitted values which indicates homoskedasticity among the errors and the NormalQ-Q plot shows that the points are mainly on the line thus normally distributed.





**Image 12 and 13** shows a plot of residuals vs fitted values which indicates homoskedasticity among the errors and the NormalQ-Q plot shows that the points are mainly on the line thus normally distributed.





**Image 14 and 15** shows a plot of residuals vs fitted values which indicates homoskedasticity among the errors and the NormalQ-Q plot shows that the points are mainly on the line thus normally distributed.

