Project

Brian Park, Dylan Maray, Hangting Lu, He Li

2024-10-31

```r
file.exists("/Users/dylanmaray/Desktop/AMS 317/Walmart.csv")
```

```
## [1] TRUE
```

Uploading the Walmart data set and checking if there are any missing data.

```r
walmart = read.csv("/Users/dylanmaray/Desktop/AMS 317/Walmart.csv")
```

Check for missing data.

```r
sum(is.na(walmart))
```

```
## [1] 0
```

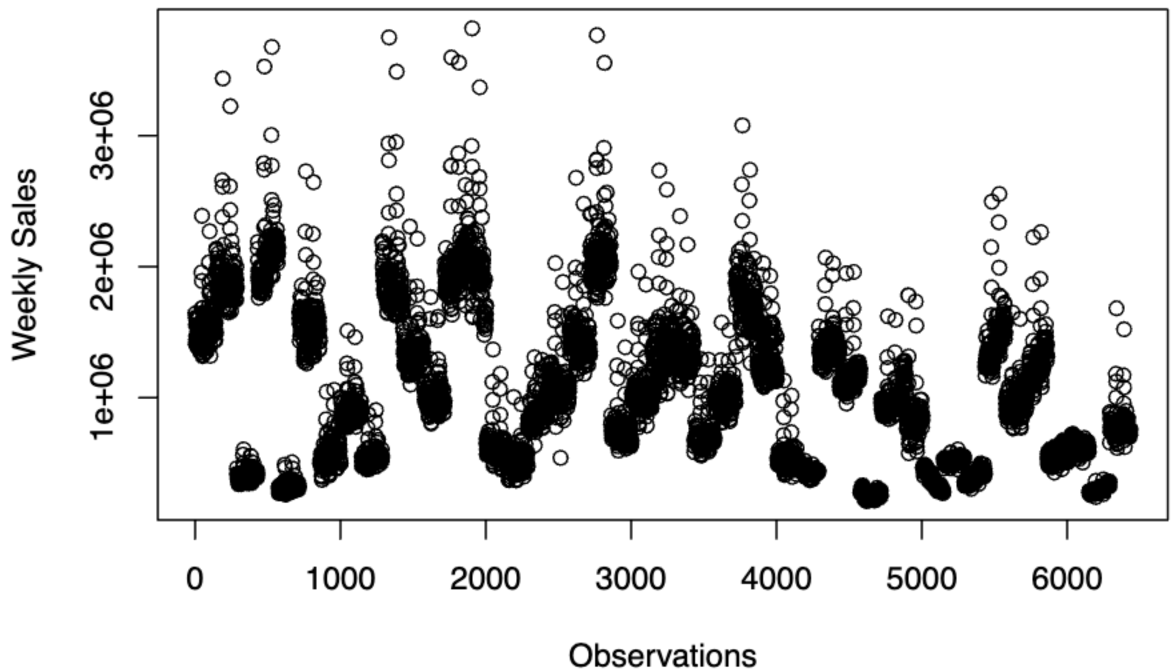
Checking the summary of this data set.

```r
summary(walmart)
```

```
##      Store         Date            Weekly_Sales      Holiday_Flag
##  Min.   : 1   Length:6435        Min.   : 209986   Min.   :0.00000
##  1st Qu.:12   Class :character   1st Qu.: 553350   1st Qu.:0.00000
##  Median :23   Mode  :character   Median : 960746   Median :0.00000
##  Mean   :23                      Mean   :1046965   Mean   :0.06993
##  3rd Qu.:34                      3rd Qu.:1420159   3rd Qu.:0.00000
##  Max.   :45                      Max.   :3818686   Max.   :1.00000
##   Temperature       Fuel_Price         CPI         Unemployment
##  Min.   : -2.06   Min.   :2.472   Min.   :126.1   Min.   : 3.879
##  1st Qu.: 47.46   1st Qu.:2.933   1st Qu.:131.7   1st Qu.: 6.891
##  Median : 62.67   Median :3.445   Median :182.6   Median : 7.874
##  Mean   : 60.66   Mean   :3.359   Mean   :171.6   Mean   : 7.999
##  3rd Qu.: 74.94   3rd Qu.:3.735   3rd Qu.:212.7   3rd Qu.: 8.622
##  Max.   :100.14   Max.   :4.468   Max.   :227.2   Max.   :14.313
```

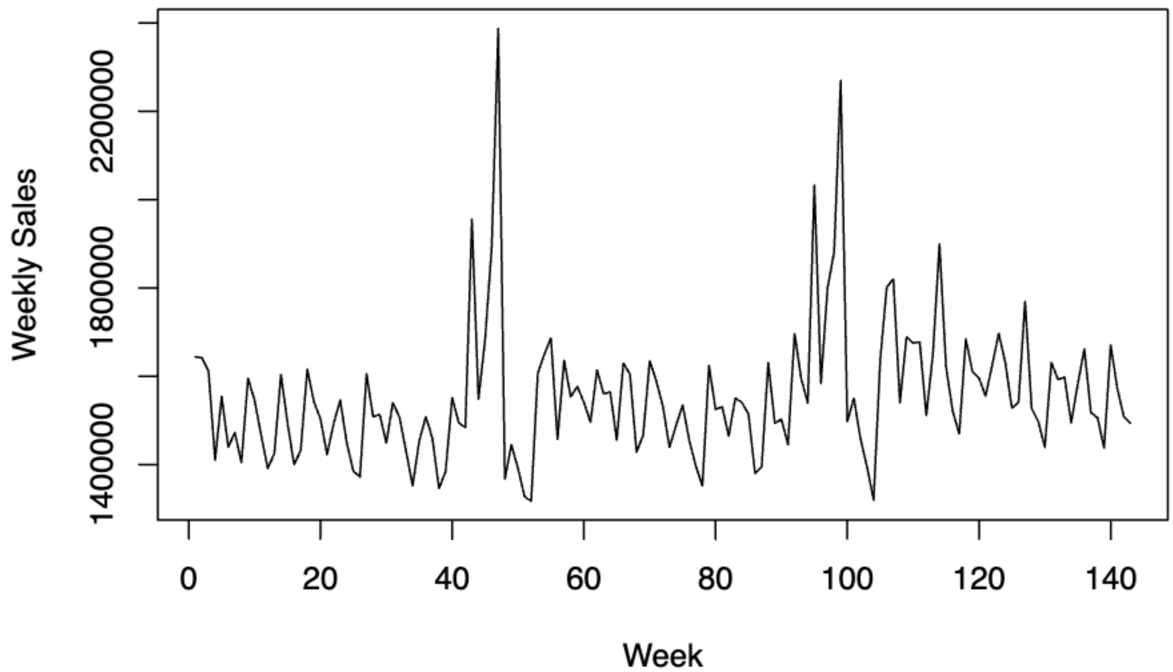
Plotting the Weekly Sales from the walmart data set.

```r
plot(walmart$Weekly_Sales, xlab = 'Observations', ylab = 'Weekly Sales')
```



Plotting the Weekly Sales from Store 1.

```
plot(walmart[walmart$Store == 1,]$Weekly_Sales, type = 'l', xlab = 'Week', ylab = 'Weekly Sales')
```



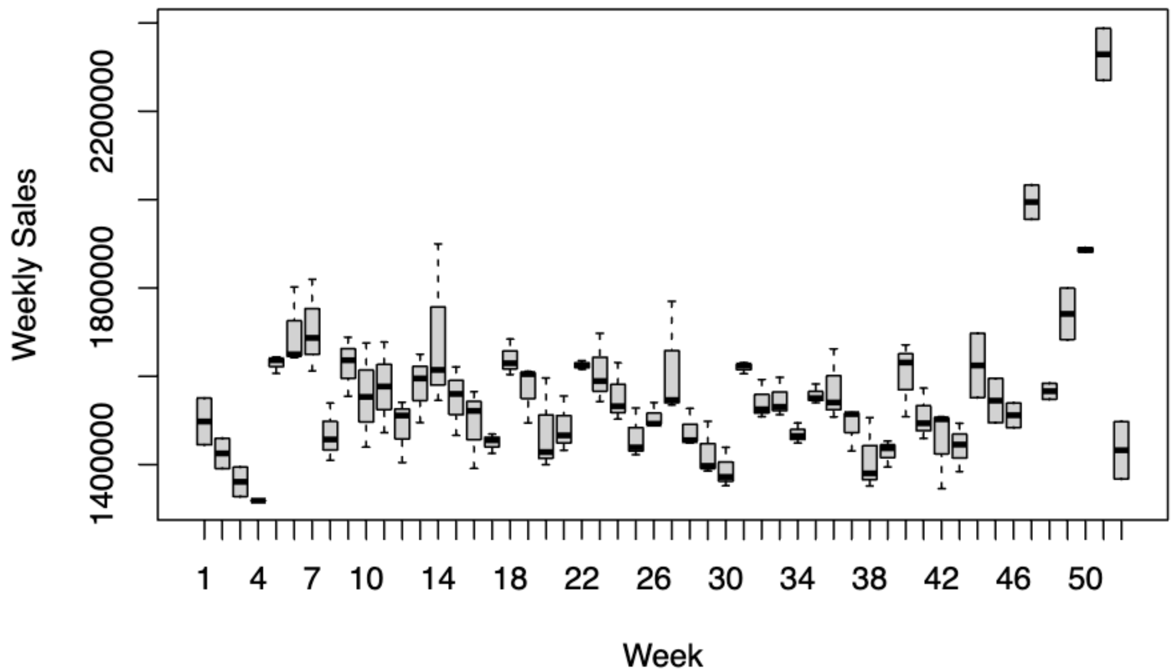

```r
walmart$Year = factor(substr(walmart$Date, 7, 10))

walmart_s1 = walmart[walmart$Store == 1, ]
# Focusing on Store 1

walmart_s1$Week =
factor(c((52-dim(walmart_s1[walmart_s1$Year=='2010',])[1]+1):52, 1:dim(walmart_s1[walmart_s1$Year=='201
# Order of Weeks within each one year
```

Here, each store has different group mean. Weekly Sales changes periodically, then for each different labeled week might have different group mean. Then, we focus on store 1 weekly sales to avoid multiple categorical variables.

```r
plot(walmart_s1$Week, walmart_s1$Weekly_Sales, xlab = 'Week', ylab = 'Weekly Sales')
```



We see there is different Weekly Sales range in each labeled week. Then we use Week as categorical variable, where week 1, 2, . . . , 52 as levels.

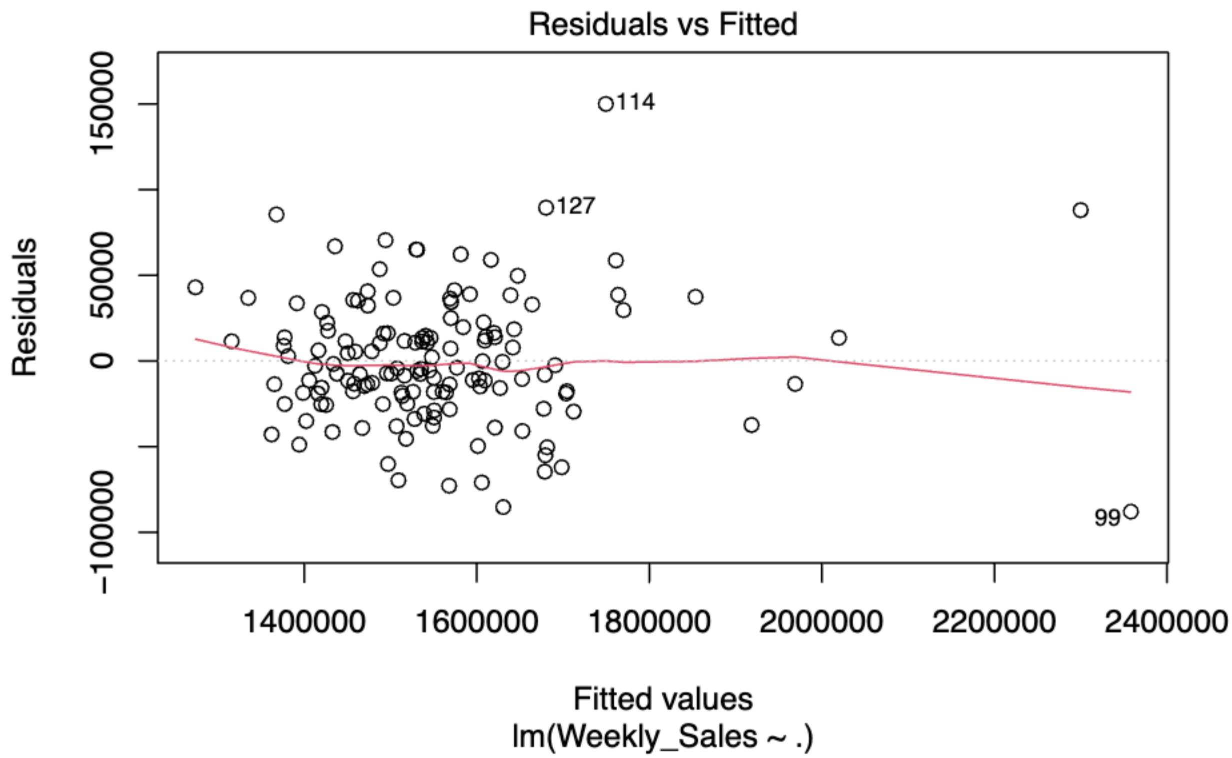
Now, we focus on Walmart store sales data specifically for Store 1 and organize it by week within each year.

We store those variables we are interested in and create a data frame.

```
Weekly_Sales = walmart_s1$Weekly_Sales
Week = walmart_s1$Week
walmart_s1t = data.frame(Weekly_Sales, walmart_s1[, 5:8], Week)
# Data for Store 1, weekly sales, categorical & numeric variables
```

Let's fit the linear model and plot the Residuals vs Fitted.

```
fit = lm(Weekly_Sales ~ ., data = walmart_s1t)
plot(fit, which = 1)
```



Let's check if the model passes the normal test.

```
shapiro.test(fit$residuals)
```

```
##
##  Shapiro-Wilk normality test
##
## data:  fit$residuals
## W = 0.97717, p-value = 0.01714
```

The W statistic is close to 1. However, the p-value is less than 0.05 level of significance. Then, we can conclude that the residuals is not normally distributed. This conveys that a transformation is needed.

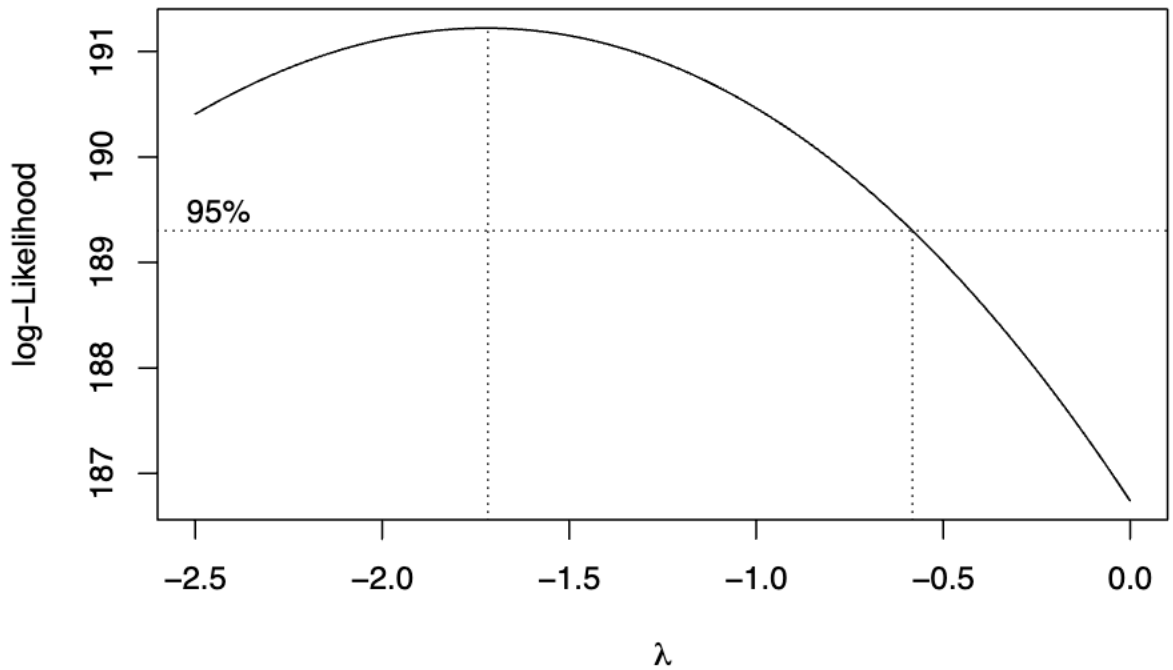
Transformation

Let's fit a linear model to predict weekly sales and then apply a Box-Cox transformation to find an optimal transformation for the Weekly Sales.

```
wmfit_trans = lm(Weekly_Sales ~ ., data = walmart_s1t)
require(MASS)
```

```
## Loading required package: MASS
```

```
b = boxcox(wmfit_trans, lambda = seq(-2.5, 0, by = 0.1))
```



Extracting the optimal lambda value from the Box-Cox transformation results.

```
b$x[which.max(b$y)]
```

```
## [1] -1.717172
```

Since -2 is within the 95% confidence interval, we use -2 as the power of transformation.

Model diagnostics

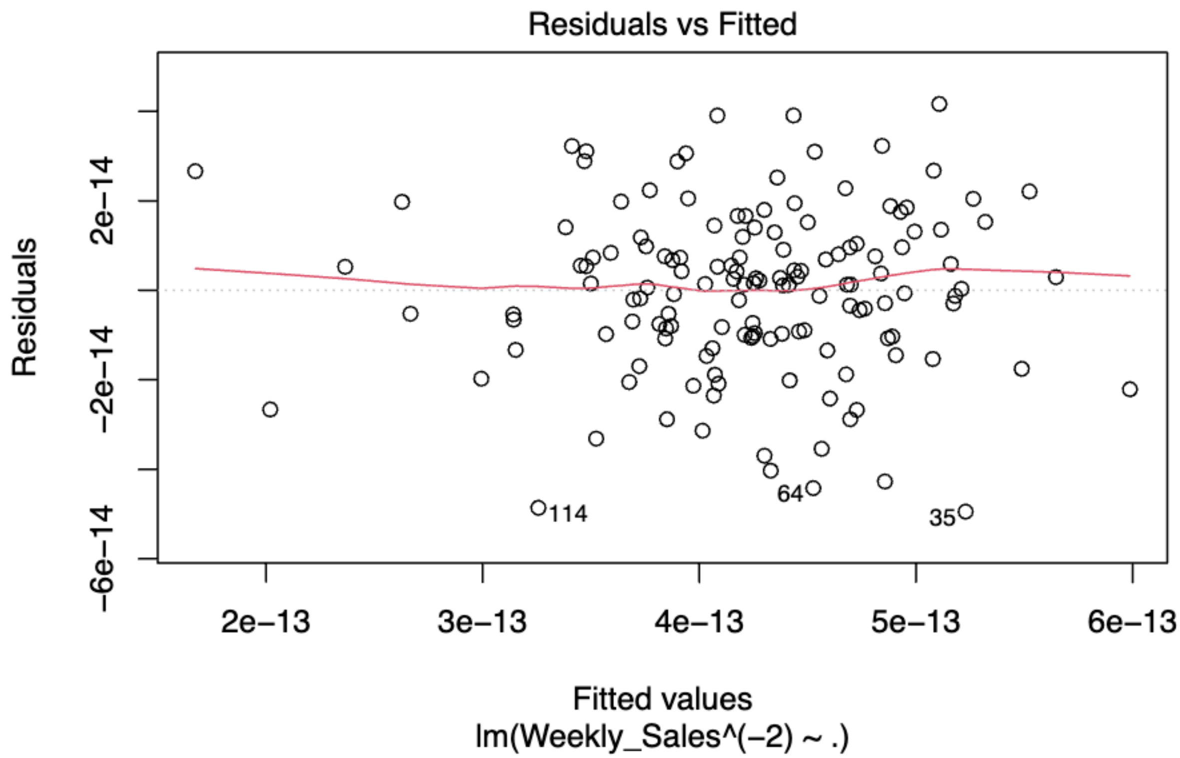
Fit a new linear model using a transformed version of the Weekly Sales variable, raising it to the power of -2.

```
wmfit = lm(Weekly_Sales^(-2) ~ ., data = walmart_s1t)
```

Homoscedasticity & Linearity

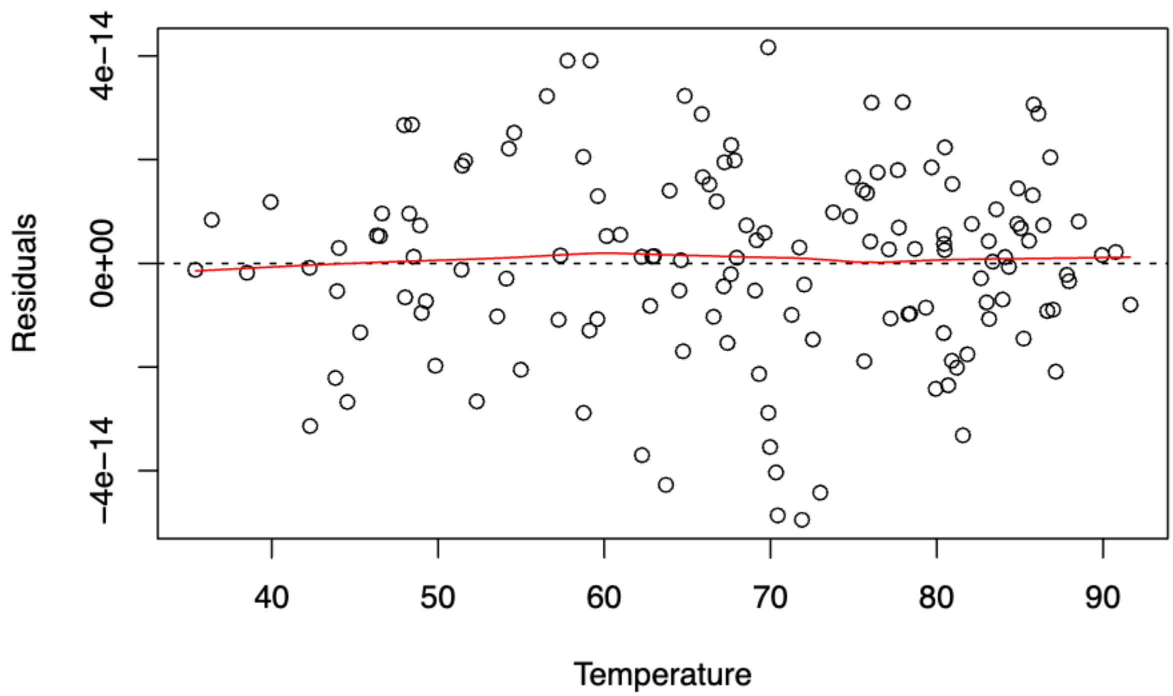
Plot the Residuals vs fitted value.

```
plot(wmfit, which = 1)
```



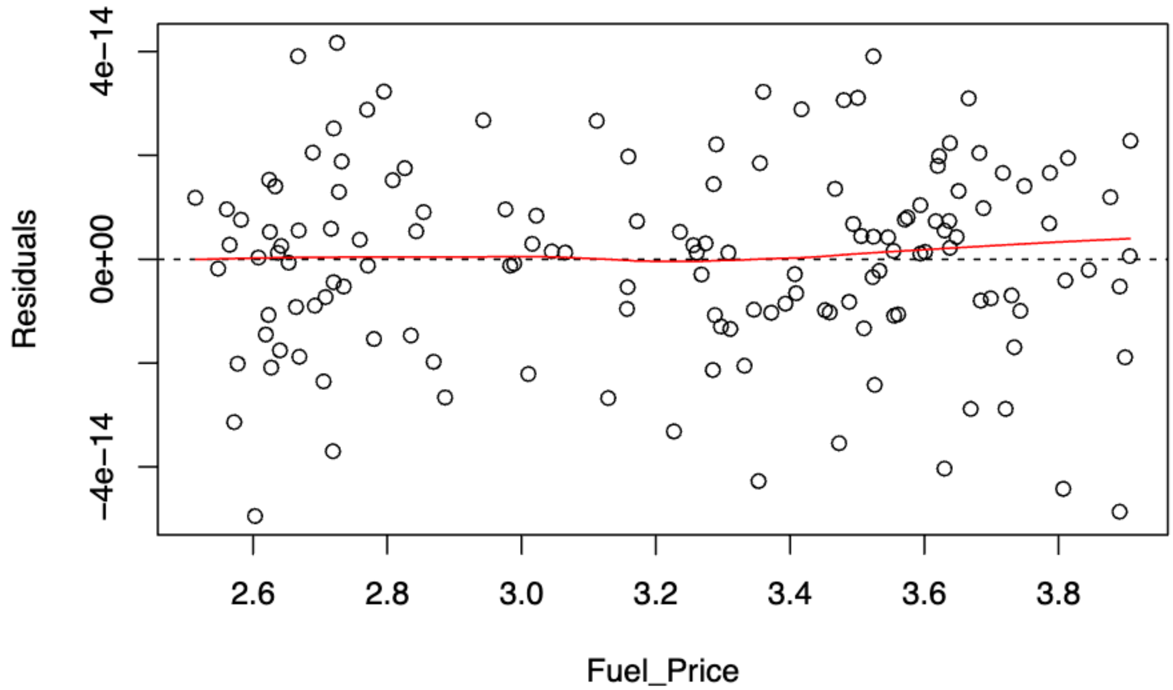
Plot the Residuals vs Temperature.

```
plot(walmart_s1t$Temperature, wmfit$residuals, xlab = 'Temperature', ylab = 'Residuals')
abline(h = 0, lty = 2)
lines(lowess(walmart_s1t$Temperature, residuals(wmfit)), col = 'red')
```



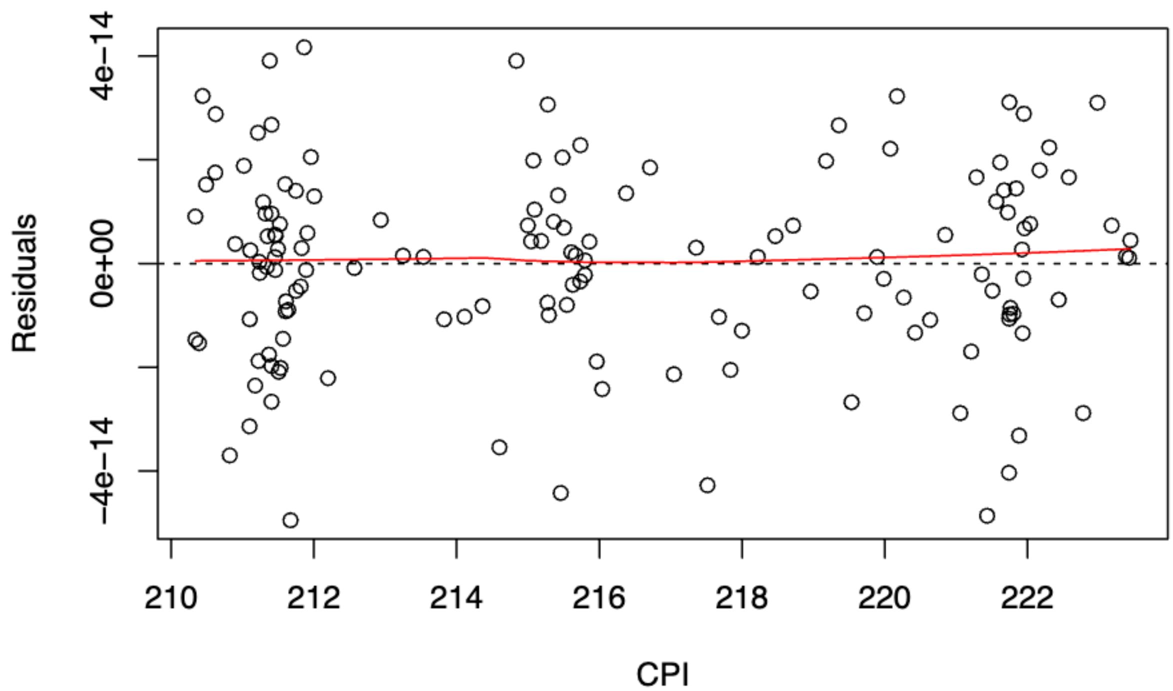
Plot the Residuals vs Fuel Price.

```
plot(walmart_s1t$Fuel_Price, wmfit$residuals, xlab = 'Fuel_Price', ylab = 'Residuals')
abline(h = 0, lty = 2)
lines(lowess(walmart_s1t$Fuel_Price, residuals(wmfit)), col = 'red')
```



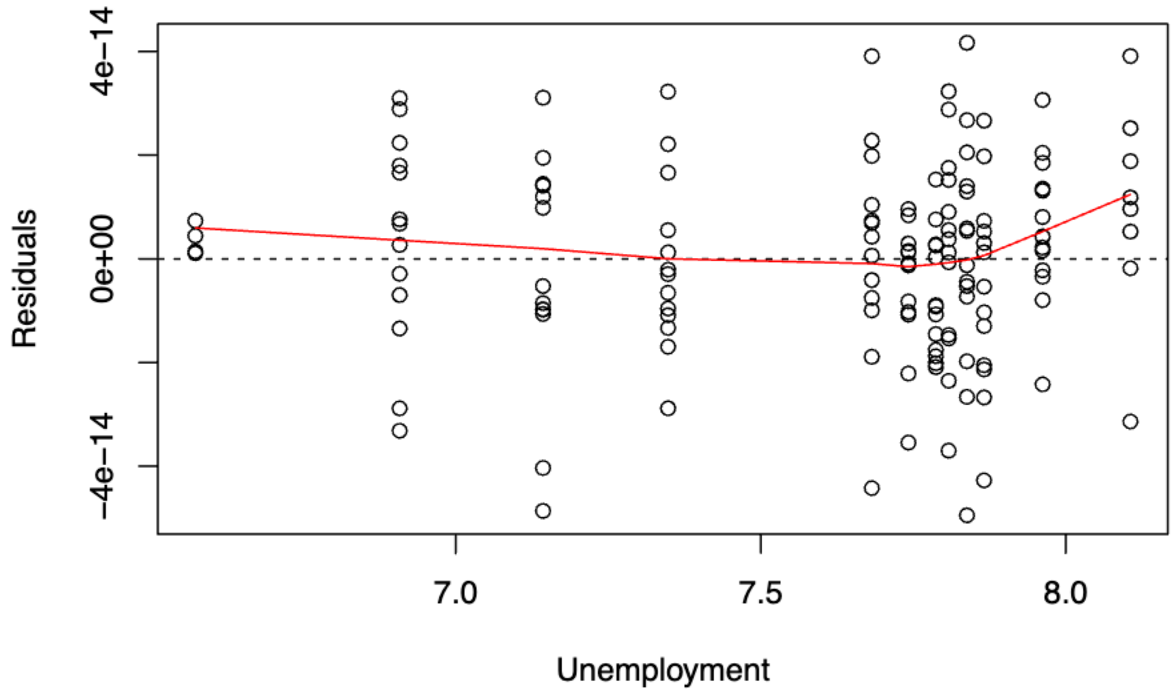
Plot the Residuals vs CPI.

```
plot(walmart_s1t$CPI, wmfit$residuals, xlab = 'CPI', ylab = 'Residuals')
abline(h = 0, lty = 2)
lines(lowess(walmart_s1t$CPI, residuals(wmfit)), col = 'red')
```



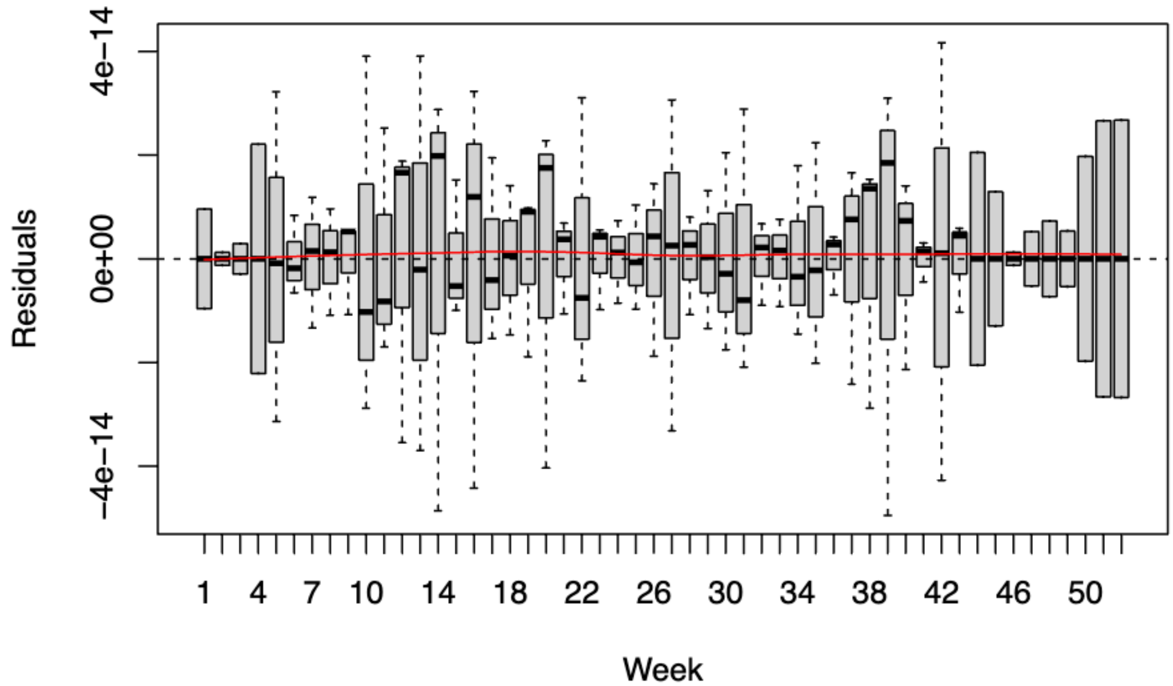
Plot the Residuals vs Unemployment.

```
plot(walmart_s1t$Unemployment, wmfit$residuals, xlab = 'Unemployment', ylab = 'Residuals')
abline(h = 0, lty = 2)
lines(lowess(walmart_s1t$Unemployment, residuals(wmfit)), col = 'red')
```



Plot the Residuals vs Week (Categorical Variable)

```
plot(walmart_s1t$Week, wmfit$residuals, xlab = 'Week', ylab = 'Residuals')
abline(h = 0, lty = 2)
lines(lowess(walmart_s1t$Week, residuals(wmfit)), col = 'red')
```



The data is linear, with constant variance.

Normality

Let's check if the model passes the normal test.

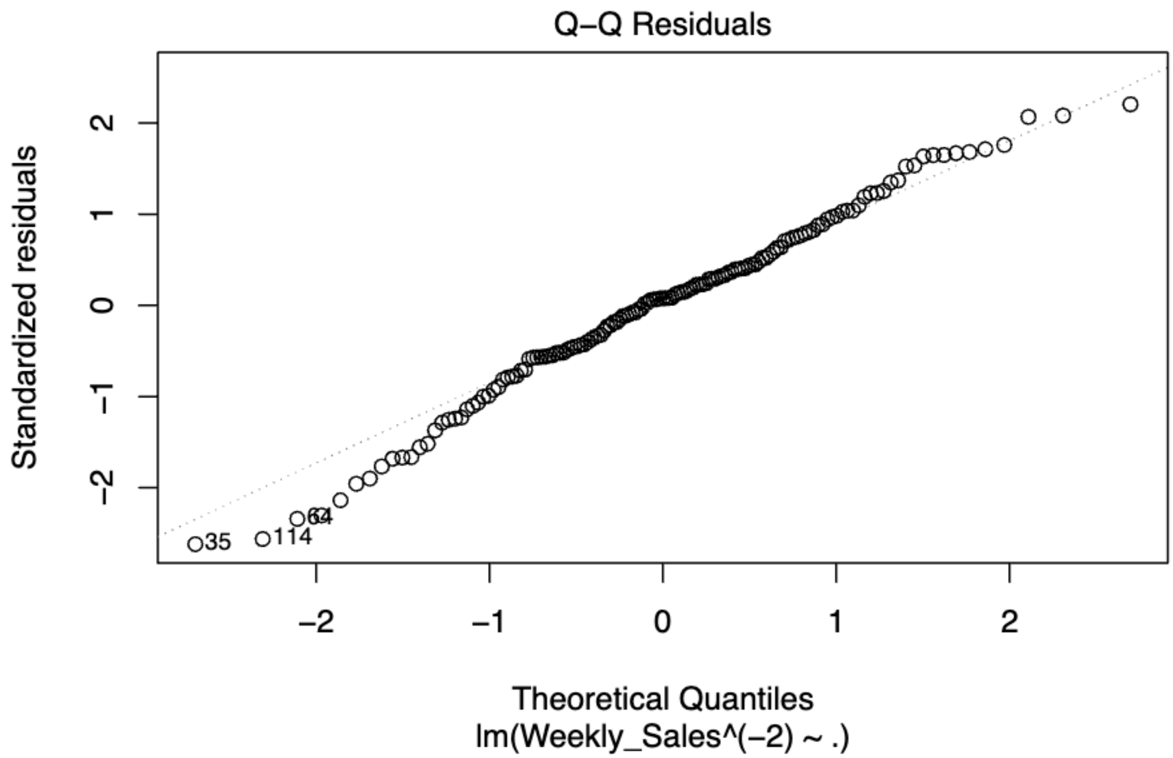
```
shapiro.test(wmfit$residuals)
```

```
##
##  Shapiro-Wilk normality test
##
## data:  wmfit$residuals
## W = 0.98846, p-value = 0.2829
```

The W statistic is close to 1. The p-value is greater than 0.05 level of significance. Then, we can conclude that the residuals is normally distributed.

Generating a plot for the linear model, focusing on the Normal Q-Q plot.

```
plot(wmfit, which = 2)
```



Independence

Performing the Durbin-Watson test on the residuals of the linear model to check for autocorrelation.

```
lmtest::dwtest(wmfit, alternative = 'two.sided')
```

```
##
##  Durbin-Watson test
##
## data:  wmfit
## DW = 1.9379, p-value = 0.5535
## alternative hypothesis: true autocorrelation is not 0
```

The DW statistic is around 2, suggesting no autocorrelation. The p-value is much higher than 0.05. Then, we conclude that the residuals is independently distributed.

Hence, we can apply linear regression to this data set.

Outliers

Identifying the most extreme standardized residual in the linear model.

```
stud = rstudent(wmfit)
abs(stud[which.max(abs(stud))])
```

```
##     35
## 2.7124
```

Calculating the threshold for identifying outliers.

```r
n = dim(walmart_s1t)[1]
p = length(wmfit$coefficients)
abs(qt((0.05/n)/2, n-p-1))
```

```
## [1] 3.723998
```

```r
# Divide 2 for two-sided test, alpha = 0.05
```

Since $2.7124 < 3.723998$, then observation 35 is not an outlier.

Leverage Points

Calculating the leverage values for each observation in the linear model and then summing them.

```r
h_diag = hatvalues(wmfit)
sum(h_diag)
```

```
## [1] 56
```

Number of parameters in the linear model.

```r
p
```

```
## [1] 56
```

Hence, sum(h_diag) = p.

Checking for leverage points.

```r
h_diag[which(h_diag > (2*p/n))]
```

```
## named numeric(0)
```

No leverage points.

Influential Points

Checking for influential points.

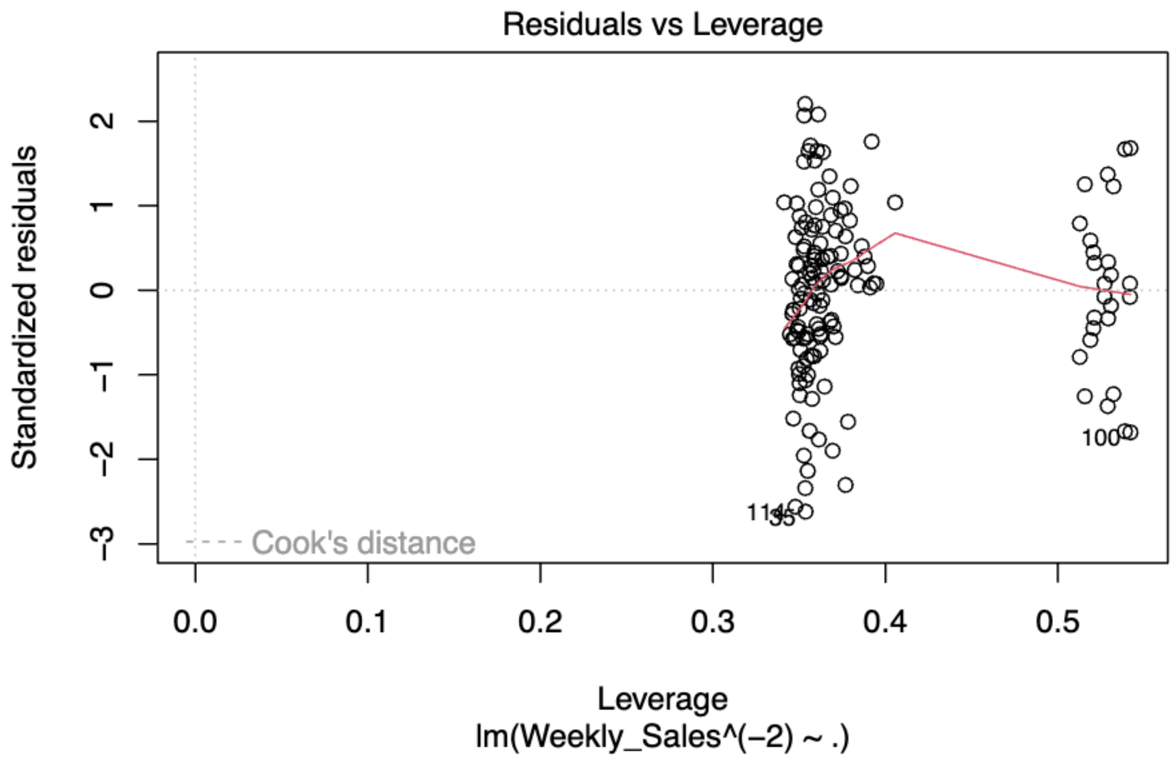
```r
cook = cooks.distance(wmfit)
cook[which(cook > 1)]
```

```
## named numeric(0)
```

No influential points.

Plotting Residuals vs Leverage.

```r
plot(wmfit, which = 5)
```



Model Selection

Due to the possibility of collinearity and model complexity, we don't want to choose all variables. Then we should take a look at vcv matrix first.

Variance Covariance Matrix of Numeric Variables

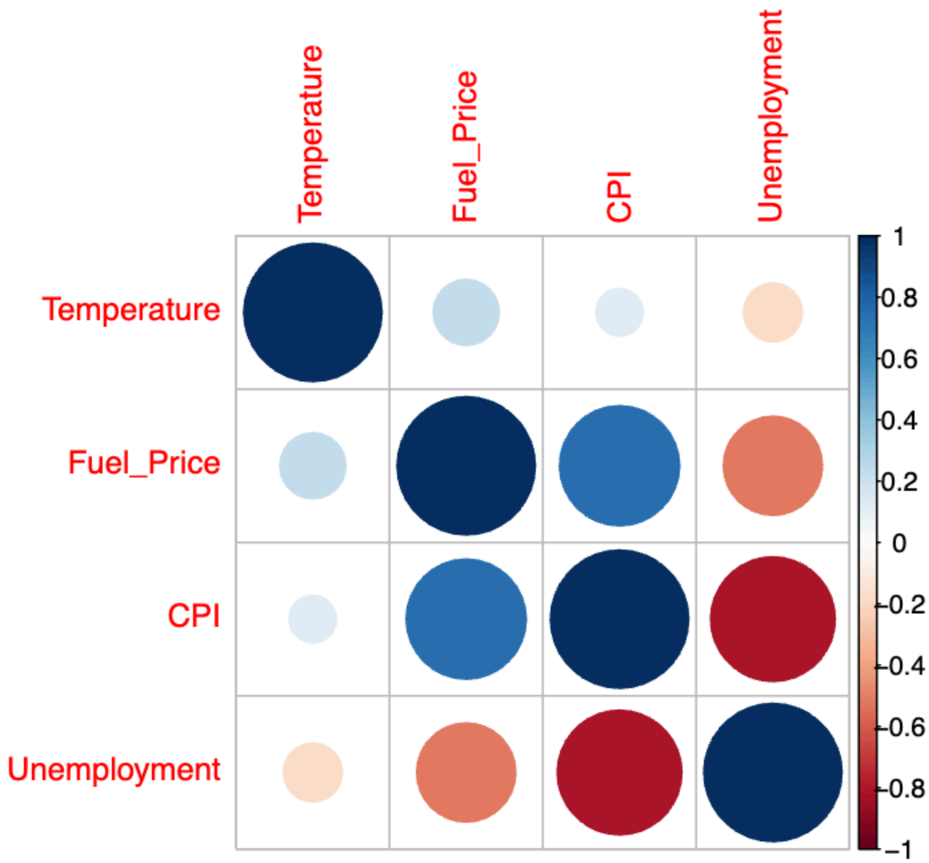
Calculating the correlation matrix.

```
round(cor(walmart_s1t[, 2:5]), 2)
```

```
##              Temperature Fuel_Price   CPI Unemployment
## Temperature         1.00       0.23  0.12        -0.18
## Fuel_Price          0.23       1.00  0.76        -0.51
## CPI                 0.12       0.76  1.00        -0.81
## Unemployment       -0.18      -0.51 -0.81         1.00
```

Generating the correlation plot.

```
corrplot::corrplot(cor(walmart_s1t[, 2:5]))
```



There is correlation among numeric variables, especially (Fuel_Price, CPI), (Fuel_Price, Unemployment), (CPI, Unemployment). Notice CPI might represent Fuel_Price and Unemployment well.

Then we use BIC to select our model:

```
wmfit_1 = lm(Weekly_Sales^(-2) ~ Week + Temperature, data = walmart_s1t)
wmfit_2 = lm(Weekly_Sales^(-2) ~ Week + CPI, data = walmart_s1t)
wmfit_3 = lm(Weekly_Sales^(-2) ~ Week + Temperature + CPI, data = walmart_s1t)
BIC(wmfit_1); BIC(wmfit_2); BIC(wmfit_3)
```

```
## [1] -8224.578
```

```
## [1] -8368.978
```

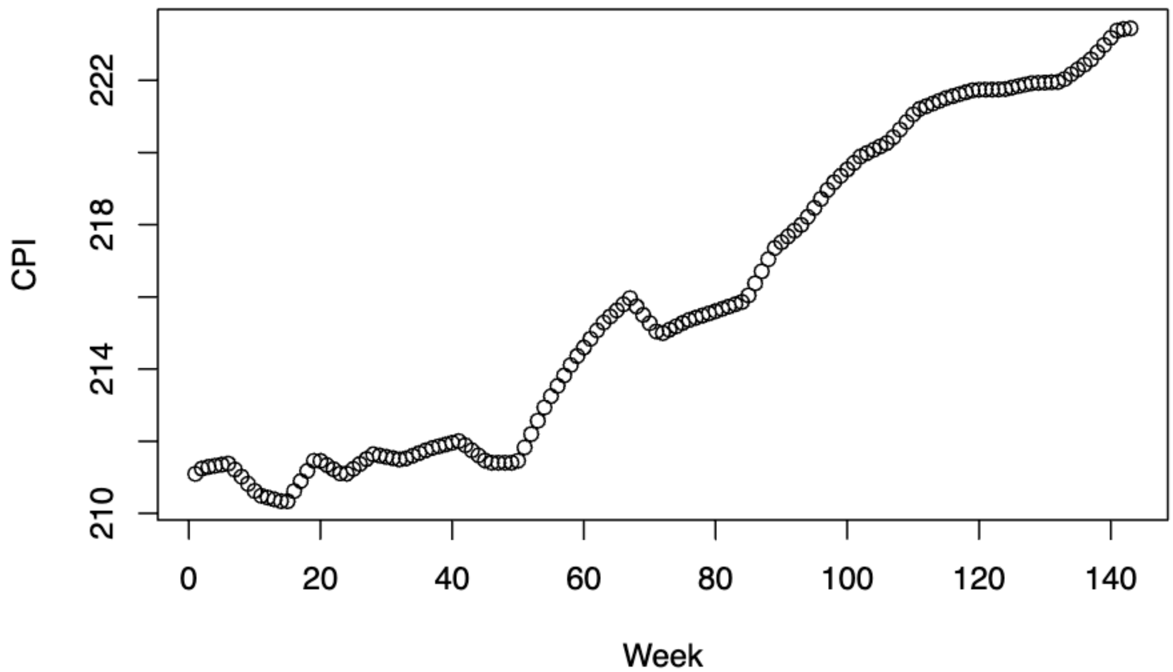
```
## [1] -8364.743
```

We choose the second model (with the lowest BIC value).

Hypothesis 1

Should we consider CPI for analyzing weekly sales?

```
plot(walmart_s1t$CPI, xlab = 'Week', ylab = 'CPI')
```



The CPI shows a increasing trend as time. But the Weekly Sales does not present a certain upward trend as time. Then we want to know if it is reasonable to include CPI as a variable in our model.

Null hypothesis: There is no effect from CPI.

Alternative hypothesis: There is effect from CPI.

Applying linear model to our categorical variable as our small model (wmfit_week). Summarizing our full model (wmfit_cpi).

```r
wmfit_week = lm(Weekly_Sales^(-2) ~ Week, data = walmart_s1t)
wmfit_cpi = wmfit_2
summary(wmfit_cpi)
```

```
##
## Call:
## lm(formula = Weekly_Sales^(-2) ~ Week + CPI, data = walmart_s1t)
##
## Residuals:
##        Min         1Q     Median         3Q        Max
## -4.958e-14 -1.121e-14  7.260e-16  1.246e-14  4.082e-14
##
## Coefficients:
##               Estimate Std. Error t value Pr(>|t|)
## (Intercept)  1.697e-12  9.996e-14  16.974  < 2e-16 ***
## Week2        4.620e-14  2.339e-14   1.976 0.051251 .
## Week3        9.537e-14  2.339e-14   4.078 9.80e-05 ***
## Week4        1.314e-13  2.339e-14   5.616 2.15e-07 ***
```

```
## Week5        -7.606e-14  2.135e-14  -3.562 0.000591 ***
## Week6        -1.032e-13  2.135e-14  -4.831 5.54e-06 ***
## Week7        -1.048e-13  2.135e-14  -4.909 4.06e-06 ***
## Week8         1.560e-14  2.135e-14   0.731 0.466881
## Week9        -6.952e-14  2.135e-14  -3.256 0.001592 **
## Week10       -3.004e-14  2.135e-14  -1.407 0.162878
## Week11       -4.110e-14  2.135e-14  -1.925 0.057384 .
## Week12        8.203e-15  2.135e-14   0.384 0.701697
## Week13       -4.426e-14  2.135e-14  -2.073 0.041034 *
## Week14       -8.684e-14  2.135e-14  -4.067 0.000102 ***
## Week15       -2.745e-14  2.135e-14  -1.286 0.201748
## Week16        6.266e-15  2.135e-14   0.294 0.769807
## Week17        3.033e-14  2.135e-14   1.421 0.158906
## Week18       -7.280e-14  2.135e-14  -3.410 0.000975 ***
## Week19       -3.782e-14  2.135e-14  -1.771 0.079903 .
## Week20        1.961e-14  2.135e-14   0.918 0.360831
## Week21        1.067e-14  2.135e-14   0.500 0.618532
## Week22       -6.592e-14  2.135e-14  -3.088 0.002683 **
## Week23       -5.670e-14  2.135e-14  -2.656 0.009355 **
## Week24       -2.973e-14  2.135e-14  -1.392 0.167207
## Week25        2.393e-14  2.135e-14   1.121 0.265401
## Week26       -3.958e-15  2.135e-14  -0.185 0.853336
## Week27       -5.713e-14  2.135e-14  -2.676 0.008854 **
## Week28        1.489e-14  2.135e-14   0.697 0.487486
## Week29        4.925e-14  2.135e-14   2.307 0.023363 *
## Week30        7.685e-14  2.135e-14   3.599 0.000521 ***
## Week31       -6.215e-14  2.135e-14  -2.911 0.004543 **
## Week32       -2.140e-14  2.135e-14  -1.002 0.318933
## Week33       -2.407e-14  2.135e-14  -1.127 0.262548
## Week34        2.125e-14  2.135e-14   0.995 0.322280
## Week35       -2.946e-14  2.135e-14  -1.380 0.171143
## Week36       -3.369e-14  2.135e-14  -1.578 0.118188
## Week37        1.219e-14  2.136e-14   0.571 0.569648
## Week38        6.470e-14  2.136e-14   3.029 0.003201 **
## Week39        5.211e-14  2.136e-14   2.439 0.016674 *
## Week40       -4.609e-14  2.136e-14  -2.157 0.033647 *
## Week41        4.419e-15  2.137e-14   0.207 0.836636
## Week42        4.261e-14  2.137e-14   1.994 0.049175 *
## Week43        4.767e-14  2.137e-14   2.230 0.028205 *
## Week44       -7.016e-14  2.339e-14  -3.000 0.003495 **
## Week45       -3.038e-14  2.339e-14  -1.299 0.197242
## Week46       -1.241e-14  2.339e-14  -0.531 0.596862
## Week47       -1.985e-13  2.339e-14  -8.488 3.99e-13 ***
## Week48       -4.197e-14  2.339e-14  -1.795 0.076053 .
## Week49       -1.186e-13  2.339e-14  -5.071 2.11e-06 ***
## Week50       -1.680e-13  2.339e-14  -7.184 1.89e-10 ***
## Week51       -2.639e-13  2.339e-14 -11.284  < 2e-16 ***
## Week52        4.234e-14  2.339e-14   1.810 0.073595 .
## CPI          -5.795e-15  4.573e-16 -12.671  < 2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 2.339e-14 on 90 degrees of freedom
## Multiple R-squared:  0.9298, Adjusted R-squared:  0.8893
```

```
## F-statistic: 22.94 on 52 and 90 DF,  p-value: < 2.2e-16
```

Comparing both models using the Anova Table.

```
anova(wmfit_week, wmfit_cpi)
```

```
## Analysis of Variance Table
##
## Model 1: Weekly_Sales^(-2) ~ Week
## Model 2: Weekly_Sales^(-2) ~ Week + CPI
##   Res.Df        RSS Df  Sum of Sq      F    Pr(>F)
## 1     91 1.3704e-25
## 2     90 4.9223e-26  1 8.7816e-26 160.56 < 2.2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

According to the ANOVA table, the F is greater than $F_{0.05,1,90}$, the p-value is less than 0.05 level of significance. Then, we reject the null hypothesis, concluding that the effect of CPI is significant.

Hypothesis 2

Is CPI effected by $week_i$? (i = 1, 2, . . . , 52)

Null hypothesis: There is no interaction between $week_i$ and CPI. (coefficient of $week_i$:CPI is not significant)

Alternative hypothesis: There is interaction between $week_i$ and CPI.

Applying the linear model with interaction and obtaining the summary.

```
wmfit_int = lm(Weekly_Sales^(-2) ~ Week * CPI, data = walmart_s1t)
summary(wmfit_int)
```

```
##
## Call:
## lm(formula = Weekly_Sales^(-2) ~ Week * CPI, data = walmart_s1t)
##
## Residuals:
##        Min         1Q     Median         3Q        Max
## -4.653e-14 -4.121e-15  0.000e+00  5.031e-15  4.191e-14
##
## Coefficients:
##               Estimate Std. Error t value Pr(>|t|)
## (Intercept)  2.084e-12  8.222e-13   2.534   0.0154 *
## Week2       -3.777e-13  1.155e-12  -0.327   0.7453
## Week3       -1.342e-13  1.175e-12  -0.114   0.9096
## Week4       -1.450e-12  1.196e-12  -1.212   0.2329
## Week5       -1.631e-12  1.079e-12  -1.512   0.1387
## Week6       -1.582e-13  1.086e-12  -0.146   0.8850
## Week7        8.703e-14  1.085e-12   0.080   0.9365
## Week8        1.632e-13  1.080e-12   0.151   0.8807
## Week9       -4.386e-13  1.075e-12  -0.408   0.6855
## Week10       9.499e-13  1.070e-12   0.888   0.3801
## Week11       4.741e-13  1.060e-12   0.447   0.6571
## Week12      -4.326e-13  1.052e-12  -0.411   0.6831
## Week13      -9.530e-13  1.044e-12  -0.913   0.3667
## Week14       1.156e-12  1.035e-12   1.116   0.2712
## Week15      -3.529e-14  1.029e-12  -0.034   0.9728
## Week16      -7.580e-14  1.026e-12  -0.074   0.9415
## Week17      -1.057e-12  1.023e-12  -1.033   0.3081
```

```
## Week18      -1.017e-12  1.020e-12  -0.997   0.3250
## Week19      -4.932e-13  1.019e-12  -0.484   0.6310
## Week20       7.038e-13  1.027e-12   0.685   0.4971
## Week21      -1.304e-13  1.035e-12  -0.126   0.9004
## Week22      -1.641e-12  1.044e-12  -1.573   0.1238
## Week23      -1.557e-13  1.052e-12  -0.148   0.8831
## Week24      -2.500e-13  1.051e-12  -0.238   0.8131
## Week25      -2.118e-13  1.046e-12  -0.203   0.8405
## Week26      -1.040e-12  1.041e-12  -0.999   0.3239
## Week27       4.235e-13  1.036e-12   0.409   0.6849
## Week28      -6.358e-13  1.035e-12  -0.614   0.5425
## Week29      -2.542e-15  1.039e-12  -0.002   0.9981
## Week30      -4.702e-13  1.044e-12  -0.450   0.6549
## Week31      -1.465e-12  1.049e-12  -1.397   0.1703
## Week32      -6.890e-13  1.054e-12  -0.654   0.5170
## Week33      -7.065e-13  1.049e-12  -0.673   0.5048
## Week34      -1.041e-12  1.043e-12  -0.997   0.3247
## Week35      -1.219e-12  1.037e-12  -1.175   0.2471
## Week36      -1.006e-13  1.032e-12  -0.098   0.9228
## Week37      -6.350e-13  1.028e-12  -0.617   0.5406
## Week38       5.281e-13  1.025e-12   0.515   0.6094
## Week39      -1.750e-12  1.022e-12  -1.712   0.0949 .
## Week40      -1.698e-13  1.019e-12  -0.167   0.8685
## Week41      -4.078e-13  1.015e-12  -0.402   0.6902
## Week42       5.006e-13  1.016e-12   0.493   0.6249
## Week43      -2.101e-13  1.017e-12  -0.207   0.8374
## Week44       7.895e-13  1.420e-12   0.556   0.5815
## Week45       2.962e-13  1.403e-12   0.211   0.8339
## Week46      -5.492e-13  1.355e-12  -0.405   0.6874
## Week47      -1.204e-12  1.306e-12  -0.922   0.3621
## Week48      -1.107e-12  1.264e-12  -0.876   0.3863
## Week49      -4.762e-13  1.226e-12  -0.388   0.7000
## Week50      -1.884e-12  1.203e-12  -1.566   0.1254
## Week51      -2.404e-12  1.189e-12  -2.023   0.0500 .
## Week52       7.636e-13  1.175e-12   0.650   0.5198
## CPI         -7.589e-15  3.814e-15  -1.990   0.0536 .
## Week2:CPI    1.967e-15  5.353e-15   0.367   0.7154
## Week3:CPI    1.066e-15  5.444e-15   0.196   0.8458
## Week4:CPI    7.321e-15  5.542e-15   1.321   0.1942
## Week5:CPI    7.236e-15  5.013e-15   1.443   0.1569
## Week6:CPI    2.500e-16  5.046e-15   0.050   0.9607
## Week7:CPI   -8.971e-16  5.039e-15  -0.178   0.8596
## Week8:CPI   -6.893e-16  5.014e-15  -0.137   0.8914
## Week9:CPI    1.712e-15  4.989e-15   0.343   0.7333
## Week10:CPI  -4.547e-15  4.963e-15  -0.916   0.3652
## Week11:CPI  -2.389e-15  4.915e-15  -0.486   0.6296
## Week12:CPI   2.045e-15  4.877e-15   0.419   0.6773
## Week13:CPI   4.214e-15  4.839e-15   0.871   0.3892
## Week14:CPI  -5.758e-15  4.801e-15  -1.199   0.2376
## Week15:CPI   3.800e-17  4.772e-15   0.008   0.9937
## Week16:CPI   3.824e-16  4.758e-15   0.080   0.9363
## Week17:CPI   5.038e-15  4.743e-15   1.062   0.2947
## Week18:CPI   4.375e-15  4.729e-15   0.925   0.3606
## Week19:CPI   2.112e-15  4.721e-15   0.447   0.6571
```

```
## Week20:CPI  -3.163e-15  4.758e-15  -0.665   0.5101
## Week21:CPI   6.572e-16  4.797e-15   0.137   0.8917
## Week22:CPI   7.295e-15  4.836e-15   1.509   0.1395
## Week23:CPI   4.624e-16  4.873e-15   0.095   0.9249
## Week24:CPI   1.024e-15  4.868e-15   0.210   0.8345
## Week25:CPI   1.095e-15  4.845e-15   0.226   0.8223
## Week26:CPI   4.799e-15  4.823e-15   0.995   0.3259
## Week27:CPI  -2.220e-15  4.801e-15  -0.462   0.6464
## Week28:CPI   3.015e-15  4.795e-15   0.629   0.5331
## Week29:CPI   2.449e-16  4.815e-15   0.051   0.9597
## Week30:CPI   2.536e-15  4.836e-15   0.524   0.6030
## Week31:CPI   6.492e-15  4.857e-15   1.336   0.1892
## Week32:CPI   3.092e-15  4.880e-15   0.634   0.5300
## Week33:CPI   3.160e-15  4.860e-15   0.650   0.5193
## Week34:CPI   4.913e-15  4.831e-15   1.017   0.3155
## Week35:CPI   5.502e-15  4.803e-15   1.145   0.2590
## Week36:CPI   3.175e-16  4.777e-15   0.066   0.9473
## Week37:CPI   2.996e-15  4.761e-15   0.629   0.5329
## Week38:CPI  -2.125e-15  4.745e-15  -0.448   0.6568
## Week39:CPI   8.313e-15  4.729e-15   1.758   0.0866 .
## Week40:CPI   5.837e-16  4.713e-15   0.124   0.9021
## Week41:CPI   1.911e-15  4.695e-15   0.407   0.6862
## Week42:CPI  -2.088e-15  4.696e-15  -0.445   0.6591
## Week43:CPI   1.202e-15  4.700e-15   0.256   0.7996
## Week44:CPI  -4.006e-15  6.602e-15  -0.607   0.5475
## Week45:CPI  -1.524e-15  6.519e-15  -0.234   0.8164
## Week46:CPI   2.492e-15  6.293e-15   0.396   0.6943
## Week47:CPI   4.672e-15  6.065e-15   0.770   0.4458
## Week48:CPI   4.948e-15  5.867e-15   0.843   0.4042
## Week49:CPI   1.659e-15  5.694e-15   0.291   0.7724
## Week50:CPI   7.967e-15  5.582e-15   1.427   0.1615
## Week51:CPI   9.937e-15  5.516e-15   1.801   0.0794 .
## Week52:CPI  -3.348e-15  5.453e-15  -0.614   0.5428
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 2.241e-14 on 39 degrees of freedom
## Multiple R-squared:  0.9721, Adjusted R-squared:  0.8984
## F-statistic: 13.18 on 103 and 39 DF,  p-value: 1.302e-14
```

Since the p-value is greater than 0.05, we are not able to reject the null hypothesis. Hence, the interaction between $week_i$ and CPI is not significant.

Comparing both models using the Anova Table.

```r
anova(wmfit_cpi, wmfit_int)
```

```
## Analysis of Variance Table
##
## Model 1: Weekly_Sales^(-2) ~ Week + CPI
## Model 2: Weekly_Sales^(-2) ~ Week * CPI
##   Res.Df        RSS Df  Sum of Sq      F Pr(>F)
## 1     90 4.9223e-26
## 2     39 1.9584e-26 51 2.9639e-26 1.1574 0.3199
```

The Final Model:

```r
print(paste("Adjusted R-squared::",summary(wmfit_cpi)$adj.r.squared))
```

```
## [1] "Adjusted R-squared:: 0.88929210084957"
```

Calculate Root Mean Squared Error (RMSE)

```r
rmse <- sqrt(mean(wmfit_int$residuals^2))
print(paste("RMSE:", rmse))
```

```
## [1] "RMSE: 1.17025597389214e-14"
```

Calculate Mean Absolute Error (MAE)

```r
mae <- mean(abs(wmfit_int$residuals))
print(paste("MAE:", mae))
```

```
## [1] "MAE: 7.59780565175573e-15"
```