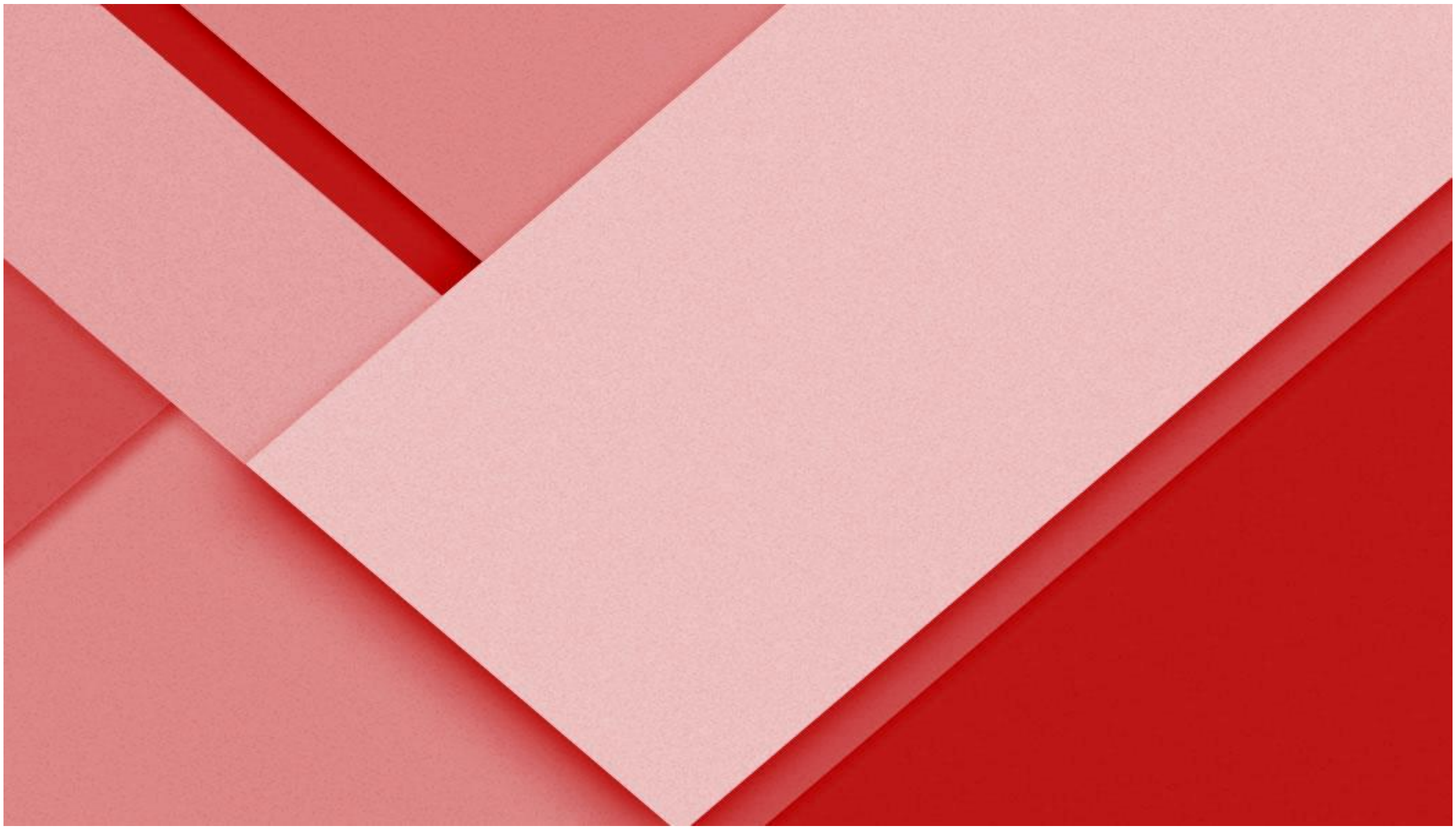


Walmart Sales Prediction

02/08/2018

ACKNOWLEDGEMENT

It is a great pleasure to have the opportunity to extend our heartiest felt gratitude to everybody who helped us throughout the course of this project.

It is distinct pleasure to express our deep sense of gratitude and indebtedness to our learned professor, **Dr. Anol Bhattacharjee** for his invaluable guidance and encouragement. With their continuous inspiration only, it becomes possible to complete this project.

Members:

Atin Maiti

Deepesh Puthran

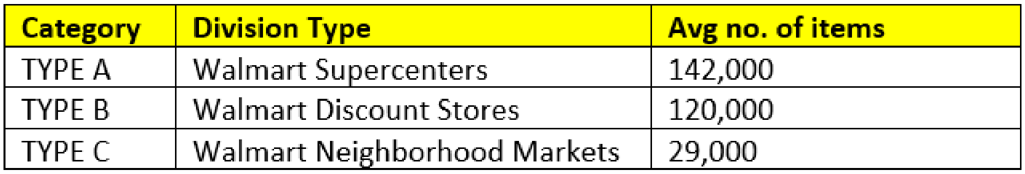
Pankaj Kumar

Siva Prasad Sahoo

Table of Contents

1.0 Overview:

Walmart Inc. is an American multinational retail corporation. It is one of the largest retail corporation in the world, which was founded by Sam Walton in year 1962. And, it has revenue over $485.87 billion dollars recorded in 2016 [1]. Walmart runs three types of stores based on number items namely hypermarkets, discount department stores, and grocery stores.



Category	Division Type	Avg no. of items
TYPE A	Walmart Supercenters	142,000
TYPE B	Walmart Discount Stores	120,000
TYPE C	Walmart Neighborhood Markets	29,000

FIGURE 1: DIVISION TYPE

Since it is a huge competitor in the retail sector it is intriguing for us to find out what are the factors that drives the sales. Recently, we came across Walmart sales data on Kaggle. This data was for a competition posted by Walmart for recruiting. The challenge was to predict weekly sales of store located in 45 difference regions. Based on this data containing details of weekly sales, store size, department code, consumer price index, unemployment of that region and promotional markdown we would predict the sales of stores by department. Here we have 99 different departments such as Dry Grocery, Sporting Goods, Frozen Goods etc.

2.0 Problem Significance:

News is ripe with the success of Walmart being the major retail in the US. Our motivation is to find out what are the factors that affect the weekly sales of the Walmart.

3.0 Data Dictionary:

We wanted to segregate data into dependent and Independent variable and further classify the data into cross-sectional and Time series data for better understanding of the data.

Dependent Variables:

- Weekly_Sales (Numeric - Continuous)– It is the totals sales in units (proprietary data, the units are unknown) of different departments in a store recorded on every Friday for that week from Feb 2010 to November 2012.

Independent Variables:

- Temperature (Numeric - Continuous) – Average temperature recorded for a week of that region in Fahrenheit.
- Size (Numeric - Discrete) – Number of items in that store.
- CPI (Numeric - Continuous) – Weekly consumer price index of that region.
- Unemployment (Numeric - Continuous)- the unemployment rate of that region.
- IsHoliday**(Logical - Discrete) - Whether the week has a special holiday.
- Fuel_Price (Numeric - Continuous) - Cost of fuel in the region per gallon.

Multilevel Data:

- Date(Panel) – Date recorded on Fridays of every week.
- Type (Cross – Sectional) – Type of division as discussed previously.
- Store (Cross – Sectional) – Store number.
- Dept (Cross – Sectional) – Departments in Walmart. There are 99 department. Names of all the department are shared in the appendix.

**

- Super Bowl: 12-Feb-10, 11-Feb-11, 10-Feb-12, 8-Feb-13
- Labor Day: 10-Sep-10, 9-Sep-11, 7-Sep-12, 6-Sep-13
- Thanksgiving: 26-Nov-10, 25-Nov-11, 23-Nov-12, 29-Nov-13
- Christmas: 31-Dec-10, 30-Dec-11, 28-Dec-12, 27-Dec-13

4.0 Assumptions:

The sales of Walmart depend on other factors apart from the predictors that we have considered such as the population in the area, the site of the store, employee services, stock of items etc. But the dataset doesn't cover such predictors, so we'll restrict our analysis only to the above-mentioned predictors.

5.0 Data Preprocessing:

As part of data preprocessing we loaded the datasets in SQL developer and merged the table for stores, department and weekly sales.

6.0 Hypothesis:

1. Department being categorical, markdown with respect to department will affect the weekly sales value.

   H_{b1} = With increase in markdown i.e. promotions for the departments which have average sales, the weekly sales figure increases.

2. Holiday along with markdown will affect the weekly sales value.

   H_{b2} = Interaction of holiday and markdown i.e. with promotional offers during holiday, the weekly sales value increases.

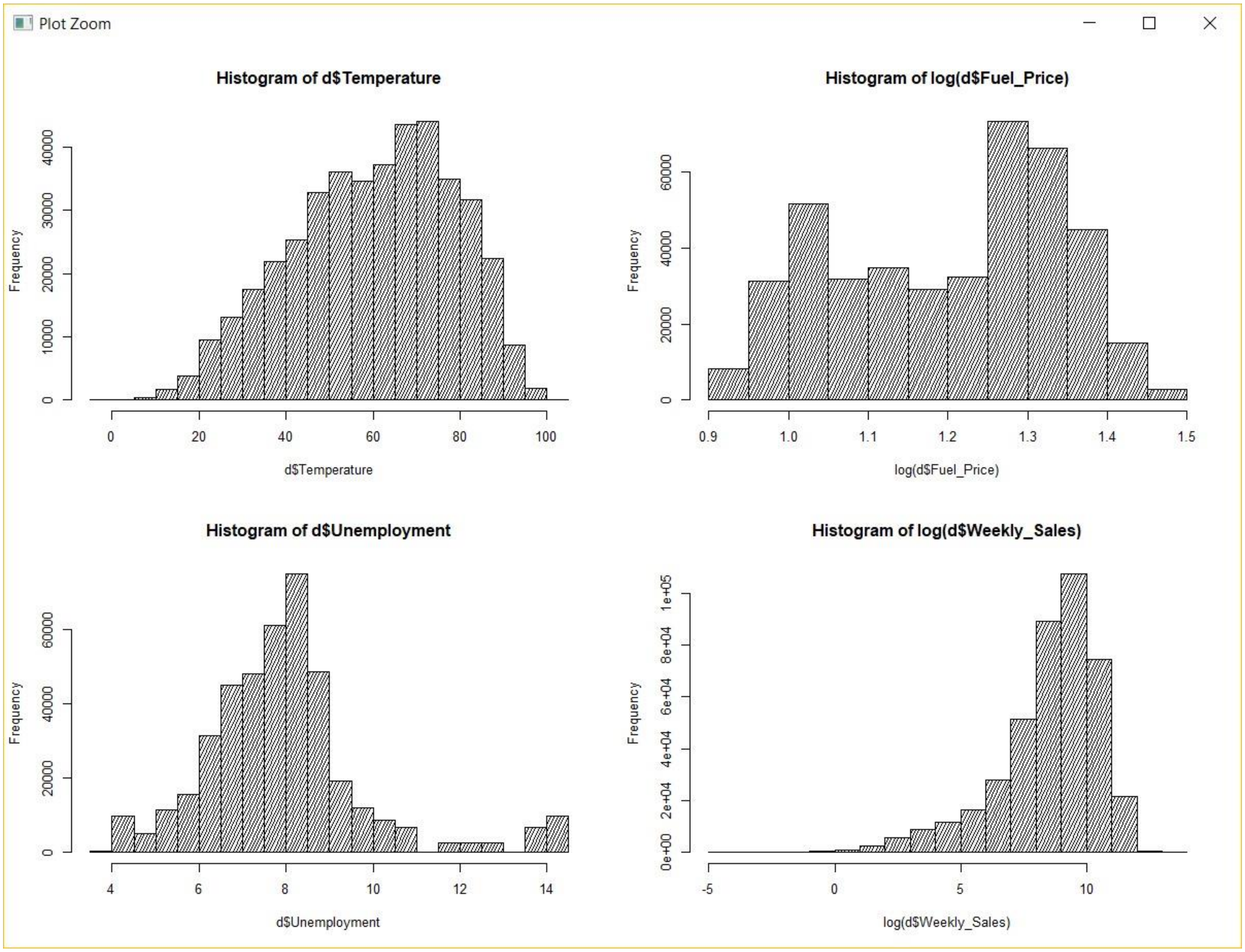
3. Unemployment will have negative effect on the weekly sales values.

   H_{b3} = Higher the unemployment of the region lower will be the sales.

7.0 Exploratory Data Analysis:

1. The histogram plot of temperature, log(fuel_price), unemployment and log(Weekly_Sales) is normal as shown below: [Plot-1] shows the histogram of Weekly Sales.

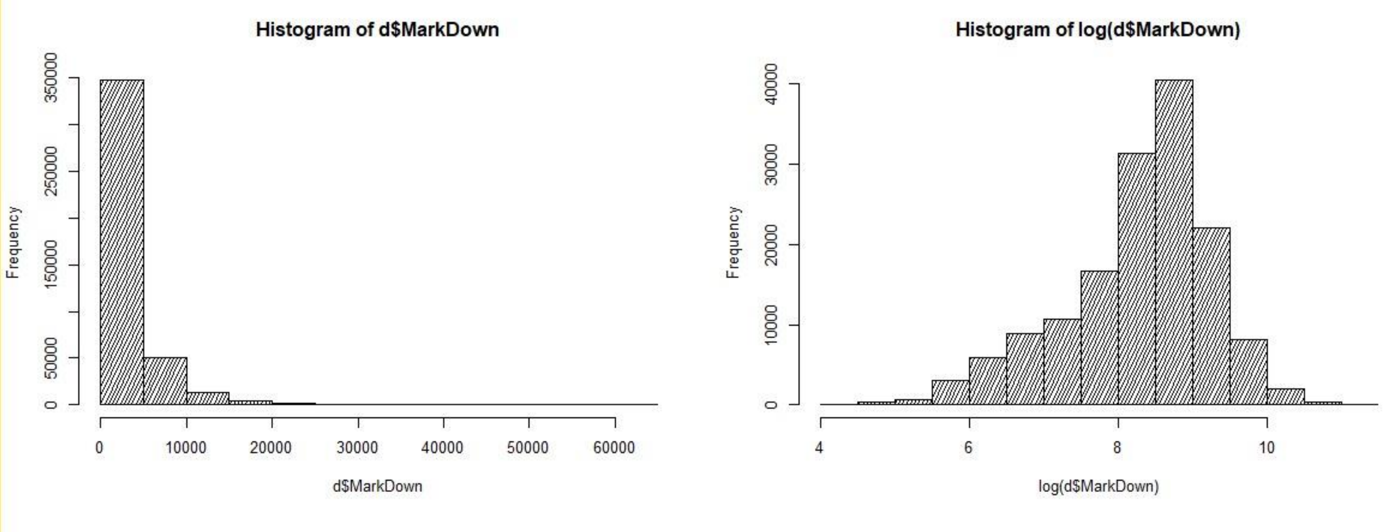
[Plot-1: Histogram of Weekly Sales]



The variables of interest are normally distributed as seen above. Among the above for the one's which were not normally distributed are transformed using log transformation.

2. Histogram plot of the aggregated markdown. **[Plot-4]** show the aggregated markdown.

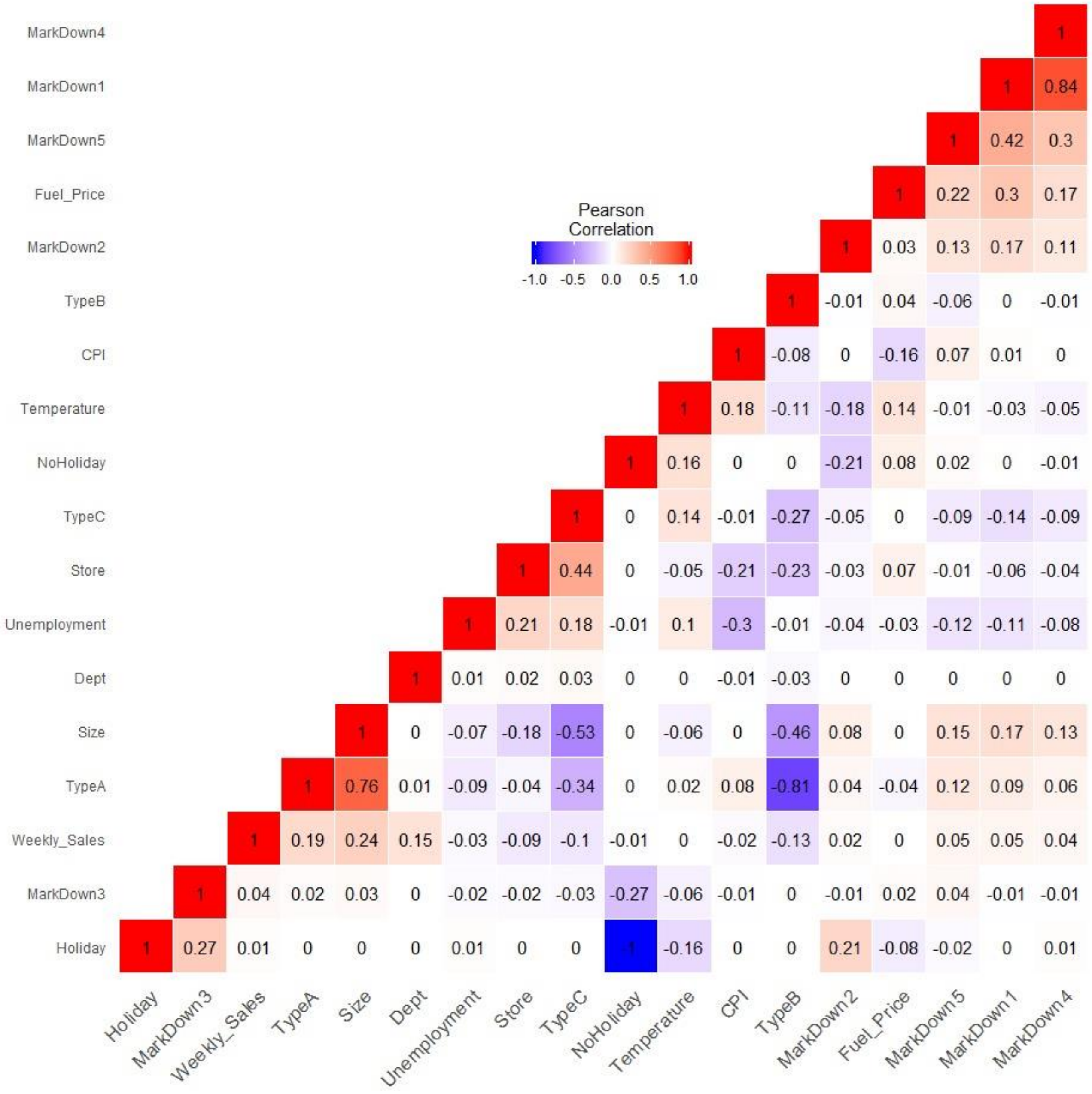
[Plot-4: Aggregated markdown]



The above histograms with log transformed value of aggregate MarkDown shows a normal distribution.

3. Computing correlation matrix. **[Plot-4]** shows the correlation matrix.

[Plot-4: Correlation matrix]

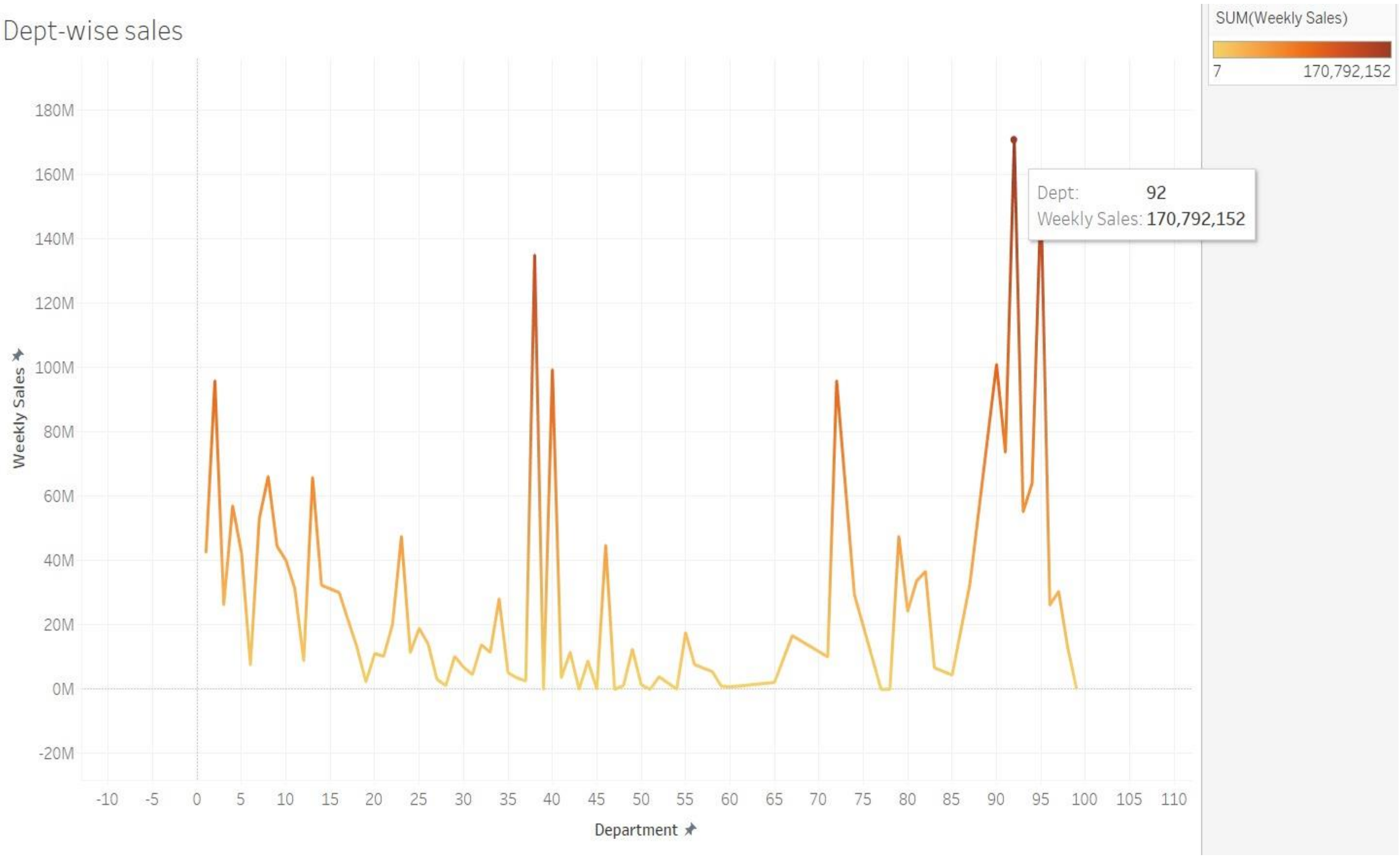


MarkDown4																		1
MarkDown1																	1	0.84
MarkDown5																1	0.42	0.3
Fuel_Price															1	0.22	0.3	0.17
MarkDown2														1	0.03	0.13	0.17	0.11
TypeB													1	-0.01	0.04	-0.06	0	-0.01
CPI												1	-0.08	0	-0.16	0.07	0.01	0
Temperature											1	0.18	-0.11	-0.18	0.14	-0.01	-0.03	-0.05
NoHoliday										1	0.16	0	0	-0.21	0.08	0.02	0	-0.01
TypeC									1	0	0.14	-0.01	-0.27	-0.05	0	-0.09	-0.14	-0.09
Store								1	0.44	0	-0.05	-0.21	-0.23	-0.03	0.07	-0.01	-0.06	-0.04
Unemployment							1	0.21	0.18	-0.01	0.1	-0.3	-0.01	-0.04	-0.03	-0.12	-0.11	-0.08
Dept						1	0.01	0.02	0.03	0	0	-0.01	-0.03	0	0	0	0	0
Size					1	0	-0.07	-0.18	-0.53	0	-0.06	0	-0.46	0.08	0	0.15	0.17	0.13
TypeA				1	0.76	0.01	-0.09	-0.04	-0.34	0	0.02	0.08	-0.81	0.04	-0.04	0.12	0.09	0.06
Weekly_Sales			1	0.19	0.24	0.15	-0.03	-0.09	-0.1	-0.01	0	-0.02	-0.13	0.02	0	0.05	0.05	0.04
MarkDown3		1	0.04	0.02	0.03	0	-0.02	-0.02	-0.03	-0.27	-0.06	-0.01	0	-0.01	0.02	0.04	-0.01	-0.01
Holiday	1	0.27	0.01	0	0	0	0.01	0	0	-1	-0.16	0	0	0.21	-0.08	-0.02	0	0.01
	Holiday	MarkDown3	Weekly_Sales	TypeA	Size	Dept	Unemployment	Store	TypeC	NoHoliday	Temperature	CPI	TypeB	MarkDown2	Fuel_Price	MarkDown5	MarkDown1	MarkDown4

The above correlation matrix shows that markdown1 and markdown5 has high correlation. So, for our analysis we have taken average of markdown1 and markdown5.

4. We would plot department-wise sales trend. **[Plot-3]** shows the department-wise sales trend.

[Plot-3: Department-wise sales trend]

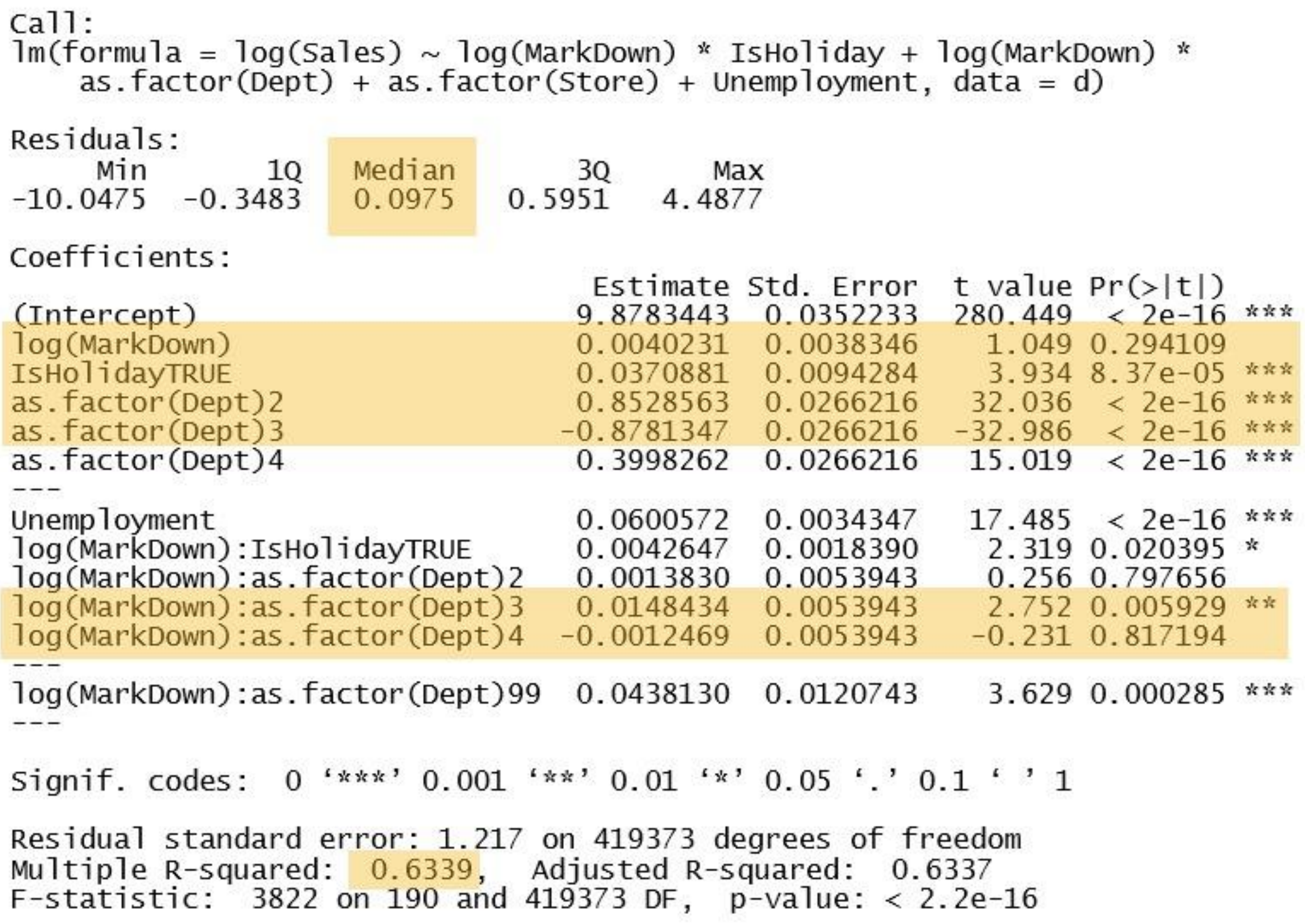


The above plot shows that **department 92** i.e. Dried Grocery has the higher weekly sales among all the departments.

8.0 Model

8.1 Preliminary Fixed Effects Model

1. For testing our first two hypothesis we have built a linear model taking IsHoliday and log(Aggregate Markdown) as well as interaction between them since on holiday if we give promotional offer the sales value will increase.



```
Call:
lm(formula = log(Sales) ~ log(MarkDown) * IsHoliday + log(MarkDown) *
    as.factor(Dept) + as.factor(Store) + Unemployment, data = d)

Residuals:
     Min       1Q   Median       3Q      Max
-10.0475  -0.3483   0.0975   0.5951   4.4877

Coefficients:
                              Estimate Std. Error  t value Pr(>|t|)
(Intercept)                  9.8783443  0.0352233  280.449  < 2e-16 ***
log(MarkDown)                0.0040231  0.0038346    1.049 0.294109
IsHolidayTRUE                0.0370881  0.0094284    3.934 8.37e-05 ***
as.factor(Dept)2             0.8528563  0.0266216   32.036  < 2e-16 ***
as.factor(Dept)3            -0.8781347  0.0266216  -32.986  < 2e-16 ***
as.factor(Dept)4             0.3998262  0.0266216   15.019  < 2e-16 ***
---
Unemployment                 0.0600572  0.0034347   17.485  < 2e-16 ***
log(MarkDown):IsHolidayTRUE  0.0042647  0.0018390    2.319 0.020395 *
log(MarkDown):as.factor(Dept)2  0.0013830  0.0053943    0.256 0.797656
log(MarkDown):as.factor(Dept)3  0.0148434  0.0053943    2.752 0.005929 **
log(MarkDown):as.factor(Dept)4 -0.0012469  0.0053943   -0.231 0.817194
---
log(MarkDown):as.factor(Dept)99  0.0438130  0.0120743    3.629 0.000285 ***
---

Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 1.217 on 419373 degrees of freedom
Multiple R-squared:  0.6339,    Adjusted R-squared:  0.6337
F-statistic:  3822 on 190 and 419373 DF,  p-value: < 2.2e-16
```

Interpretation:

As markdown increases by 100% sales increases by 0.4%. For the week comprising holidays sales is 3.7% higher compared to normal days. For department 2 (Health and beauty) the slope explains a sales increase of 85%. The interaction between department 3 (Stationary) and markdown, the slope explains a sales increase of 1.4%.

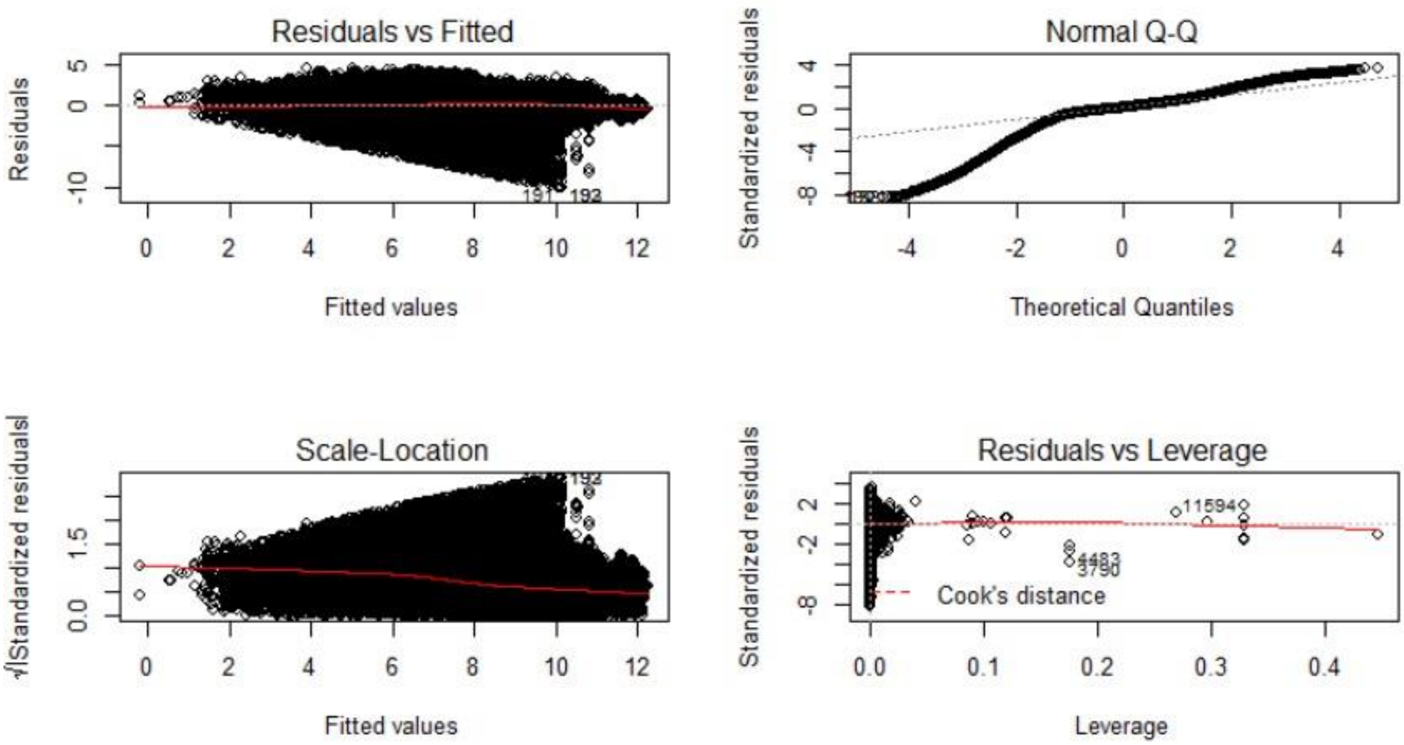
Hypothesis test results:

1. Increase in Markdown i.e. promotion have significant effect for department 3 i.e. Stationary but not on department 4 i.e. Paper Goods. So here our hypothesis stands true that markdown do not have significant effect on all the departments.

2. Our 2nd hypothesis that IsHoliday have significant effect on Sales stands true.

3. Our 3rd hypothesis that Unemployment have significant effect on Sales also stands true.

Residual Diagnostic:



1. As we could see from the residual vs Fitted values plot that the variance is not constant, and the relationship is heteroskedastic.

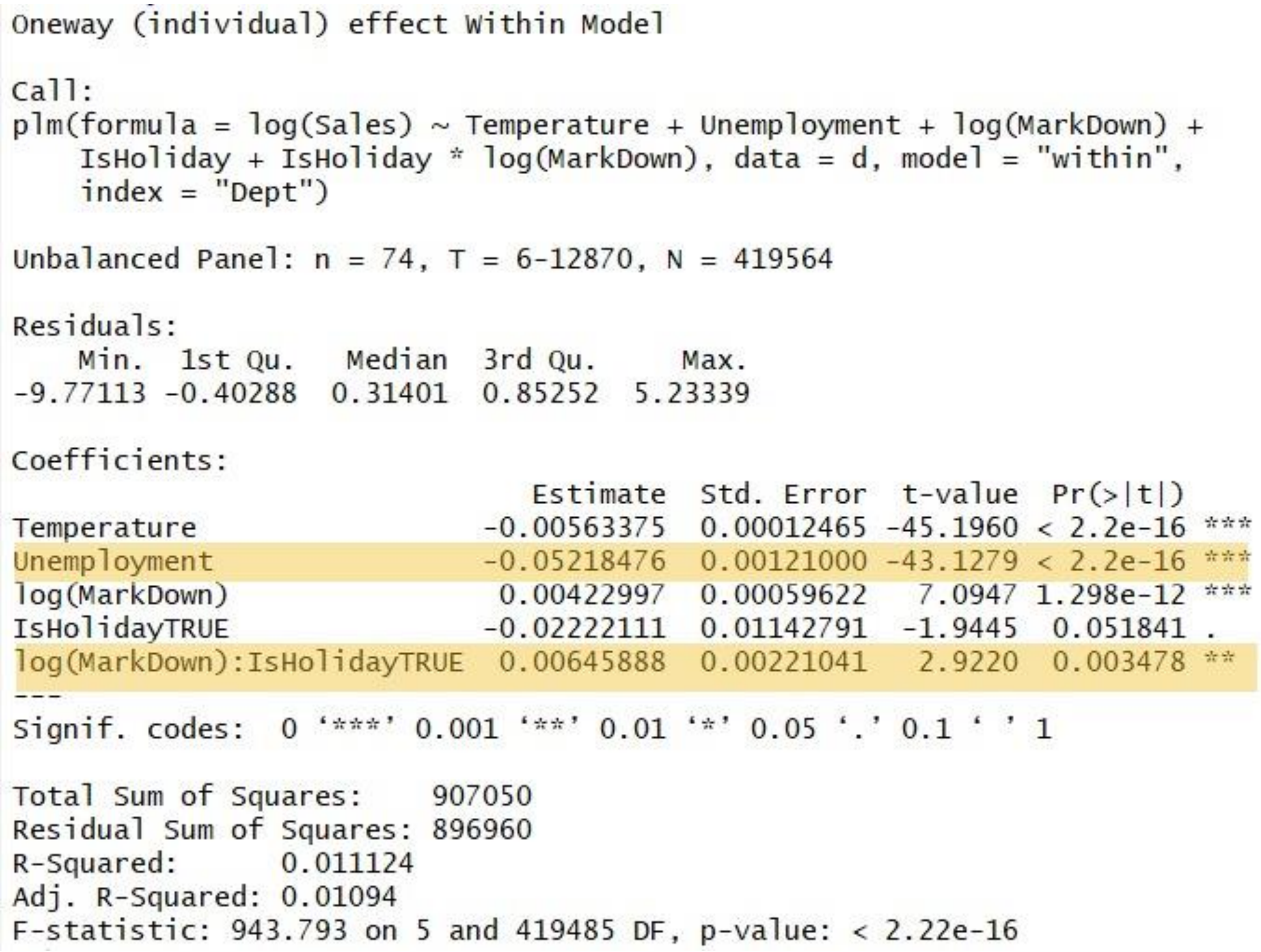
2. The normal q-q plot shows that the residual does not follow the straight line. This indicates that the error is not normally distributed

3. The Residual vs Leverage plot shows us that there are no leverage points as there are no residual beyond the Cook's distance.

8.2 Fixed Effect Model:

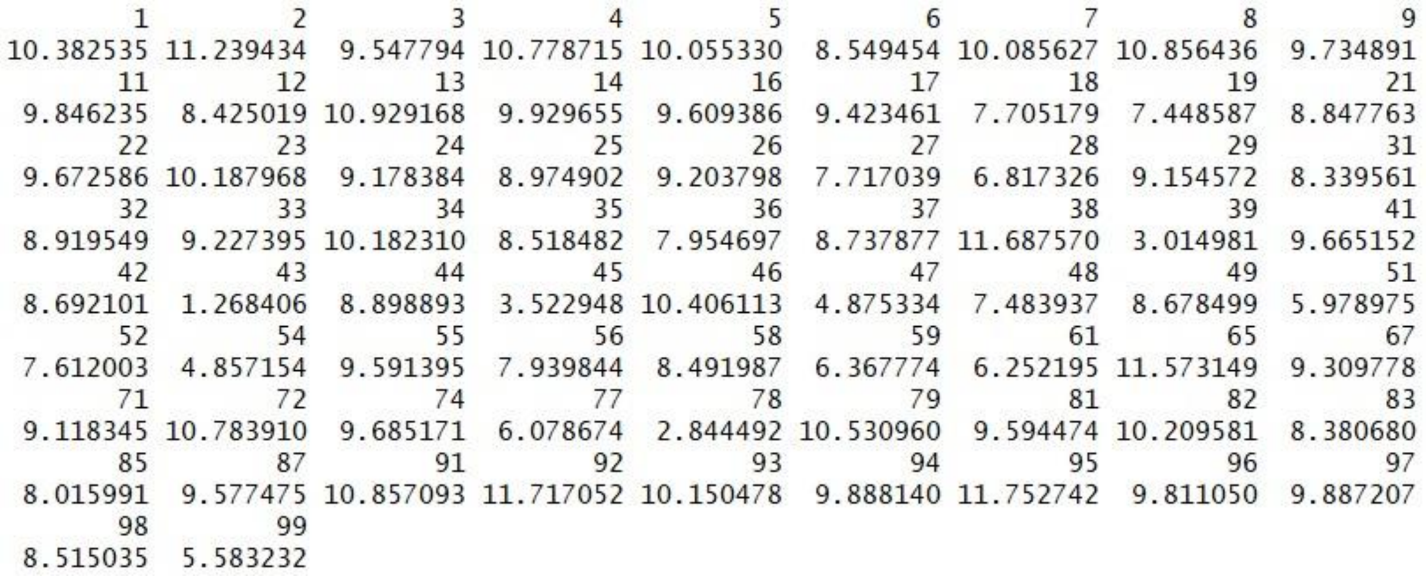
Since the dataset is multilevel, we have resorted to fixed effect analysis. As part of fixed effect modelling we have done separate modelling based on department and store.

Model1(Department):



```
Oneway (individual) effect Within Model

Call:
plm(formula = log(Sales) ~ Temperature + Unemployment + log(MarkDown) +
    IsHoliday + IsHoliday * log(MarkDown), data = d, model = "within",
    index = "Dept")

Unbalanced Panel: n = 74, T = 6-12870, N = 419564

Residuals:
    Min.  1st Qu.   Median  3rd Qu.     Max.
-9.77113 -0.40288  0.31401  0.85252  5.23339

Coefficients:
                              Estimate  Std. Error  t-value  Pr(>|t|)
Temperature                -0.00563375  0.00012465 -45.1960 < 2.2e-16 ***
Unemployment               -0.05218476  0.00121000 -43.1279 < 2.2e-16 ***
log(MarkDown)               0.00422997  0.00059622   7.0947 1.298e-12 ***
IsHolidayTRUE              -0.02222111  0.01142791  -1.9445  0.051841 .
log(MarkDown):IsHolidayTRUE  0.00645888  0.00221041   2.9220  0.003478 **
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Total Sum of Squares:    907050
Residual Sum of Squares: 896960
R-Squared:      0.011124
Adj. R-Squared: 0.01094
F-statistic: 943.793 on 5 and 419485 DF, p-value: < 2.22e-16
```

Fixef:



```
        1         2         3         4         5         6         7         8         9 
10.382535 11.239434  9.547794 10.778715 10.055330  8.549454 10.085627 10.856436  9.734891 
       11        12        13        14        16        17        18        19        21 
 9.846235  8.425019 10.929168  9.929655  9.609386  9.423461  7.705179  7.448587  8.847763 
       22        23        24        25        26        27        28        29        31 
 9.672586 10.187968  9.178384  8.974902  9.203798  7.717039  6.817326  9.154572  8.339561 
       32        33        34        35        36        37        38        39        41 
 8.919549  9.227395 10.182310  8.518482  7.954697  8.737877 11.687570  3.014981  9.665152 
       42        43        44        45        46        47        48        49        51 
 8.692101  1.268406  8.898893  3.522948 10.406113  4.875334  7.483937  8.678499  5.978975 
       52        54        55        56        58        59        61        65        67 
 7.612003  4.857154  9.591395  7.939844  8.491987  6.367774  6.252195 11.573149  9.309778 
       71        72        74        77        78        79        81        82        83 
 9.118345 10.783910  9.685171  6.078674  2.844492 10.530960  9.594474 10.209581  8.380680 
       85        87        91        92        93        94        95        96        97 
 8.015991  9.577475 10.857093 11.717052 10.150478  9.888140 11.752742  9.811050  9.887207 
       98        99 
 8.515035  5.583232 
```

Department Intercept Interpretation:

These are the effects of respective department on sales. Using these values, we can realize regression equation for each department which can be used for future prediction.

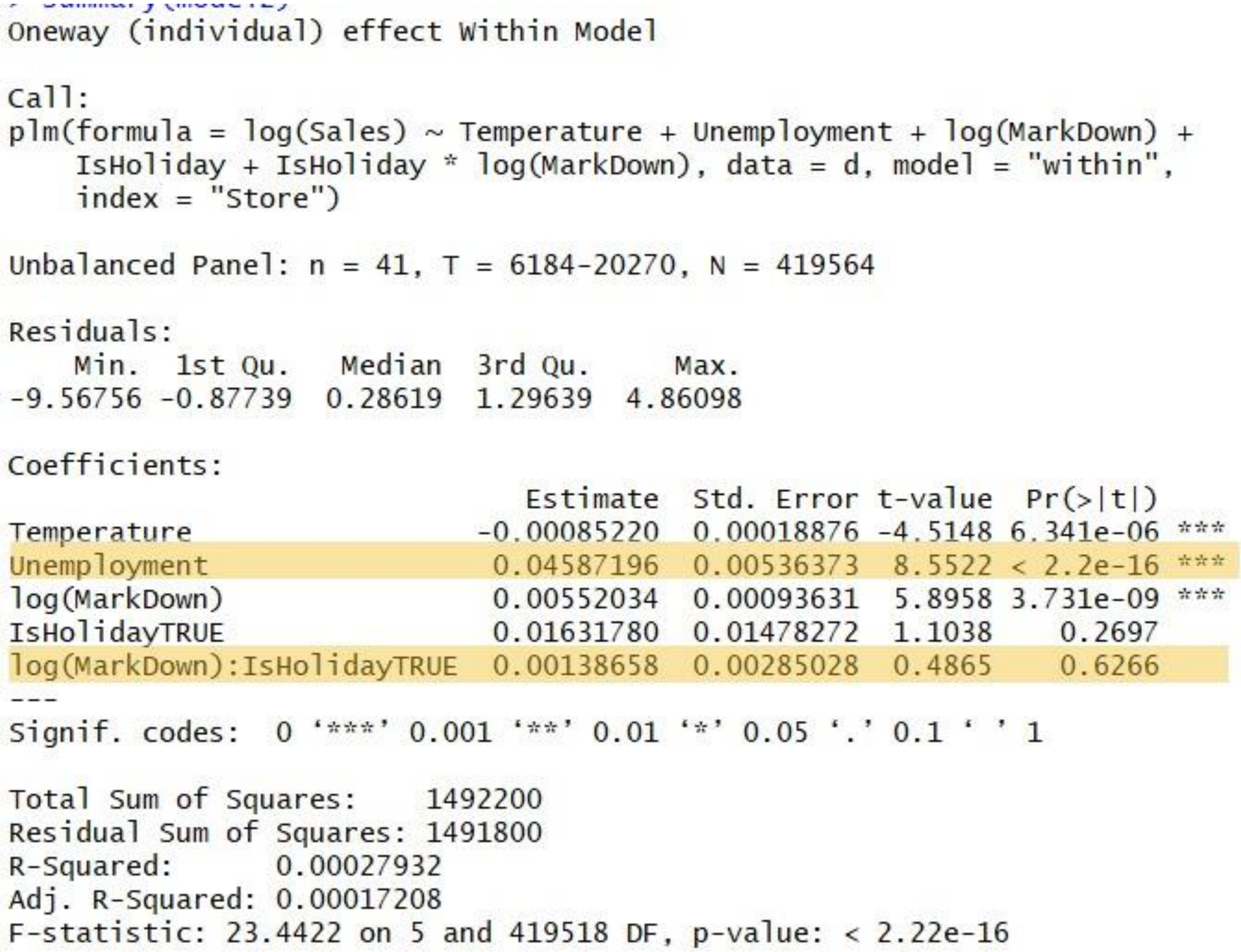
Model Interpretation:

1. As markdown increases by 100% sales increases by 0.4%.
2. The interaction between markdown and holiday has a significant effect i.e. sales increases by 0.6%.

Hypothesis test results:

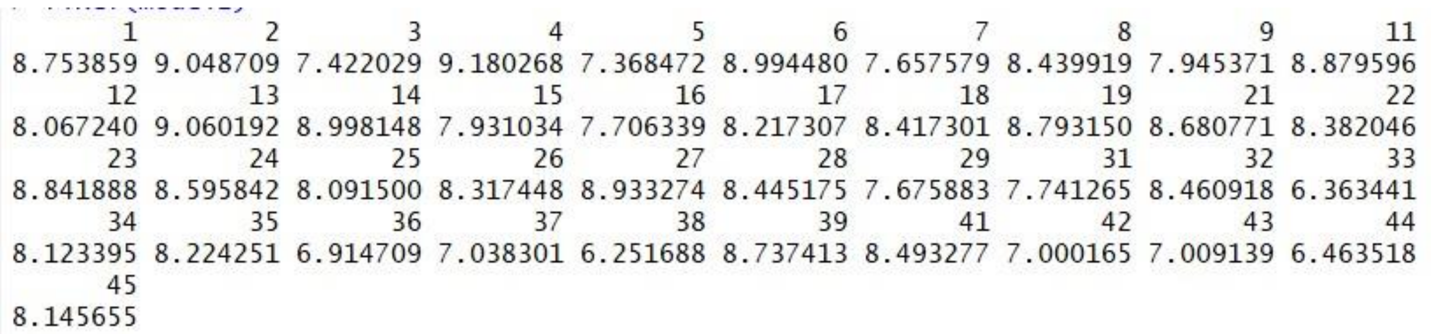
As we could see markdown and holiday have positive effect on sales. So, hypothesis stands true.

Model2(Store):



```
> summary(model2)
Oneway (individual) effect Within Model

Call:
plm(formula = log(Sales) ~ Temperature + Unemployment + log(MarkDown) +
    IsHoliday + IsHoliday * log(MarkDown), data = d, model = "within",
    index = "Store")

Unbalanced Panel: n = 41, T = 6184-20270, N = 419564

Residuals:
    Min.  1st Qu.   Median  3rd Qu.     Max.
-9.56756 -0.87739  0.28619  1.29639  4.86098

Coefficients:
                              Estimate  Std. Error t-value  Pr(>|t|)
Temperature                -0.00085220  0.00018876 -4.5148 6.341e-06 ***
Unemployment                0.04587196  0.00536373  8.5522 < 2.2e-16 ***
log(MarkDown)               0.00552034  0.00093631  5.8958 3.731e-09 ***
IsHolidayTRUE               0.01631780  0.01478272  1.1038    0.2697
log(MarkDown):IsHolidayTRUE 0.00138658  0.00285028  0.4865    0.6266
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Total Sum of Squares:    1492200
Residual Sum of Squares: 1491800
R-Squared:      0.00027932
Adj. R-Squared: 0.00017208
F-statistic: 23.4422 on 5 and 419518 DF, p-value: < 2.22e-16
```

Fixef:



```
> fixef(model2)
       1        2        3        4        5        6        7        8        9       11
8.753859 9.048709 7.422029 9.180268 7.368472 8.994480 7.657579 8.439919 7.945371 8.879596
      12       13       14       15       16       17       18       19       21       22
8.067240 9.060192 8.998148 7.931034 7.706339 8.217307 8.417301 8.793150 8.680771 8.382046
      23       24       25       26       27       28       29       31       32       33
8.841888 8.595842 8.091500 8.317448 8.933274 8.445175 7.675883 7.741265 8.460918 6.363441
      34       35       36       37       38       39       41       42       43       44
8.123395 8.224251 6.914709 7.038301 6.251688 8.737413 8.493277 7.000165 7.009139 6.463518
      45
8.145655
```

Store Intercept Interpretation:

These are the effects of respective stores on sales. Using these values, we can realize regression equation for each store which can be used for future prediction.

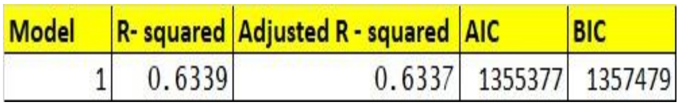
Interpretation:

1. As unemployment increases by 1-unit sales increases by 4.5%.
2. The interaction between markdown and holiday have a significant effect i.e. sales increases by 0.13%.
3. As markdown increases by 100% sales increases by 0.55%.

Hypothesis test results:

As we could see markdown, employment, holiday have positive effect on sales. So, hypothesis stands true.

8.3 Model Comparison:



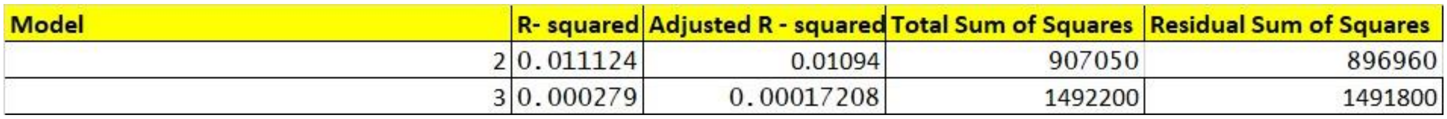
Model	R- squared	Adjusted R - squared	AIC	BIC
1	0.6339	0.6337	1355377	1357479

Interpretation:

1. Model 1 is not a parsimonious model but has a high R-squared value.
2. Unwieldy and eat 190 degrees of freedom.
3. Dummies may absorb most of the explanatory power, leaving little variance for predictors of interest.

Conclusion:

For the above-mentioned reason.



Model	R- squared	Adjusted R - squared	Total Sum of Squares	Residual Sum of Squares
2	0.011124	0.01094	907050	896960
3	0.000279	0.00017208	1492200	1491800

Interpretation:

1. Adjusted R - squared of model 2 is higher than model 3.
2. Total sum of residual of model 2 is less than model 3.
3. These models have more explanatory power than model 1.

Conclusion:

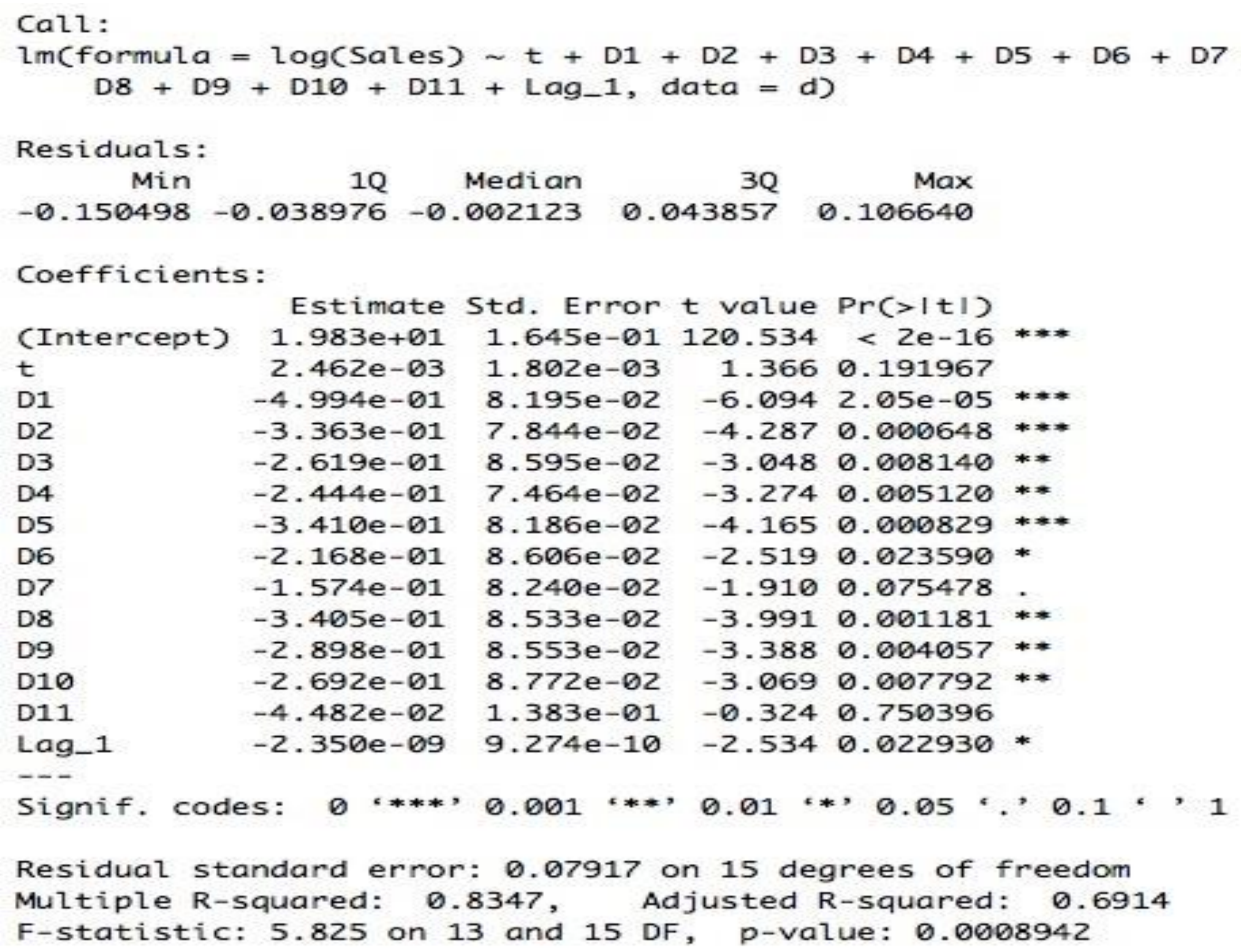
For the above-mentioned reasons, we would choose model 2 over model 3.

8.4 Time Series Analysis:

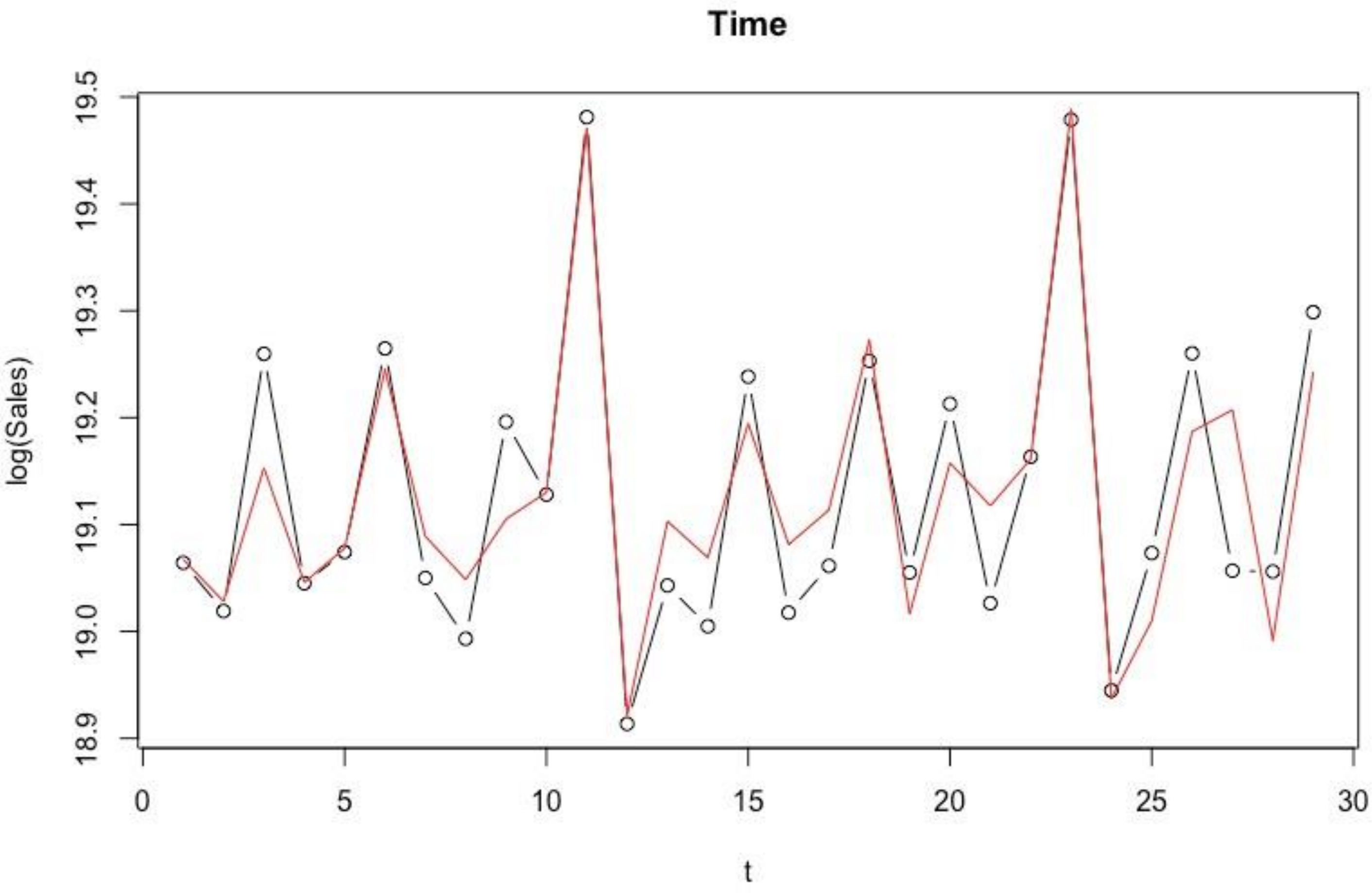
As part of time series analysis, we have designed model to forecast the monthly sales of Walmart across all the stores and department. From our analysis we found the data have seasonal trend.

Firstly, we executed linear regression for monthly sales with respect to time. The resultant R-squared value was 0.1377. Then on addition of monthly seasonality, the resultant adjusted R – squared value improved to 0.5495. Post that we added lag-1 to the data for which multiple R – squared was 0.7586 and adjusted R- Squared was 0.5494. As we could see below, lag-1 variable is marginally significant with Pr (>|t|) = 0.042. Similarly, we performed our analysis by adding lag-2, lag-3, lag-4 but this resulted in lag-1 variable losing its significance along with all the other lag variables that we tested. Due to this reason we finalized to move ahead with lag-1. Below is our model along with its plot.

[Table – 8: Time Series Analysis model]

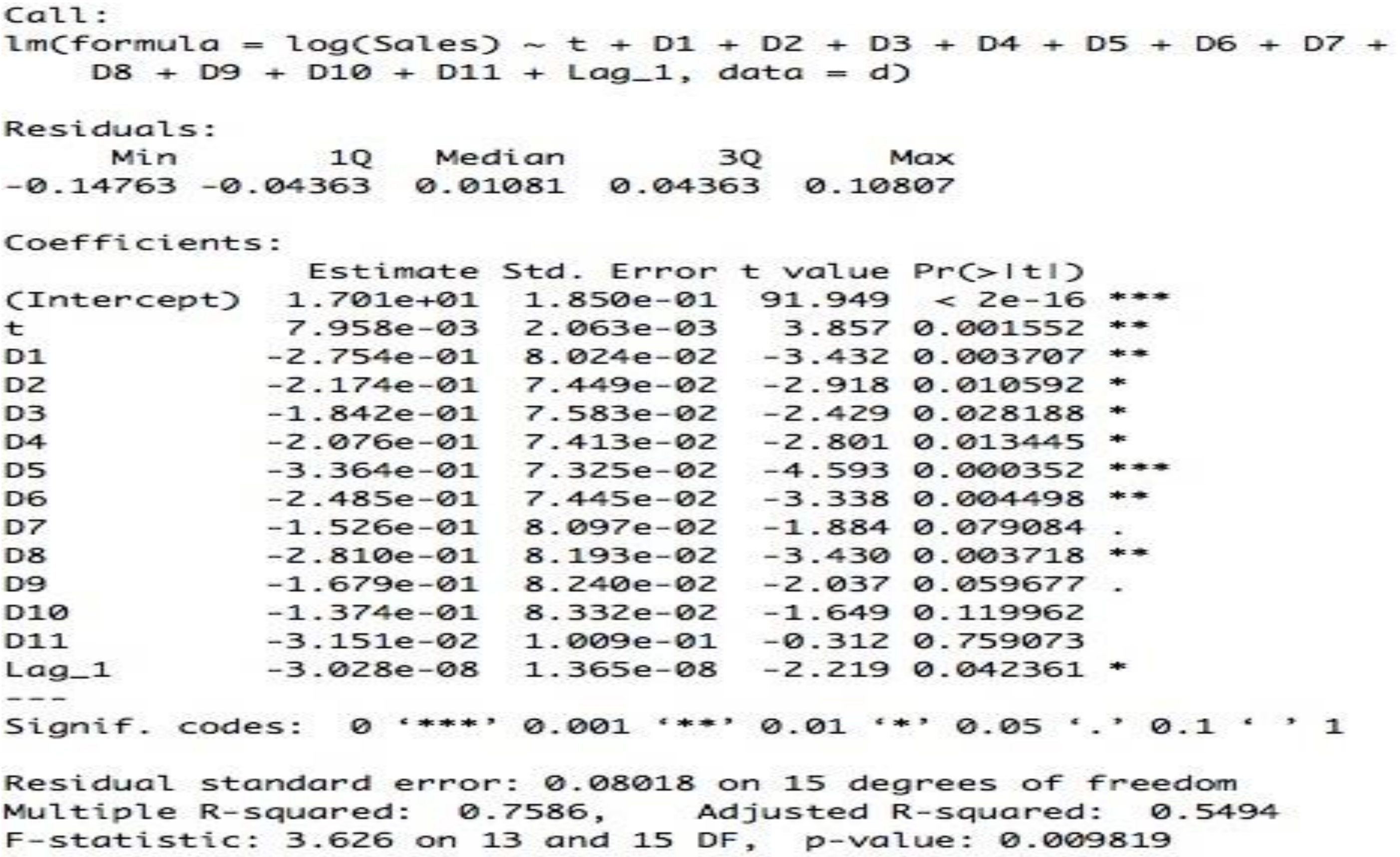


```
Call:
lm(formula = log(Sales) ~ t + D1 + D2 + D3 + D4 + D5 + D6 + D7 +
    D8 + D9 + D10 + D11 + Lag_1, data = d)

Residuals:
      Min        1Q    Median        3Q       Max
-0.150498 -0.038976 -0.002123  0.043857  0.106640

Coefficients:
              Estimate Std. Error t value Pr(>|t|)
(Intercept)  1.983e+01  1.645e-01 120.534  < 2e-16 ***
t            2.462e-03  1.802e-03   1.366 0.191967
D1          -4.994e-01  8.195e-02  -6.094 2.05e-05 ***
D2          -3.363e-01  7.844e-02  -4.287 0.000648 ***
D3          -2.619e-01  8.595e-02  -3.048 0.008140 **
D4          -2.444e-01  7.464e-02  -3.274 0.005120 **
D5          -3.410e-01  8.186e-02  -4.165 0.000829 ***
D6          -2.168e-01  8.606e-02  -2.519 0.023590 *
D7          -1.574e-01  8.240e-02  -1.910 0.075478 .
D8          -3.405e-01  8.533e-02  -3.991 0.001181 **
D9          -2.898e-01  8.553e-02  -3.388 0.004057 **
D10         -2.692e-01  8.772e-02  -3.069 0.007792 **
D11         -4.482e-02  1.383e-01  -0.324 0.750396
Lag_1       -2.350e-09  9.274e-10  -2.534 0.022930 *
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 0.07917 on 15 degrees of freedom
Multiple R-squared:  0.8347,    Adjusted R-squared:  0.6914
F-statistic: 5.825 on 13 and 15 DF,  p-value: 0.0008942
```

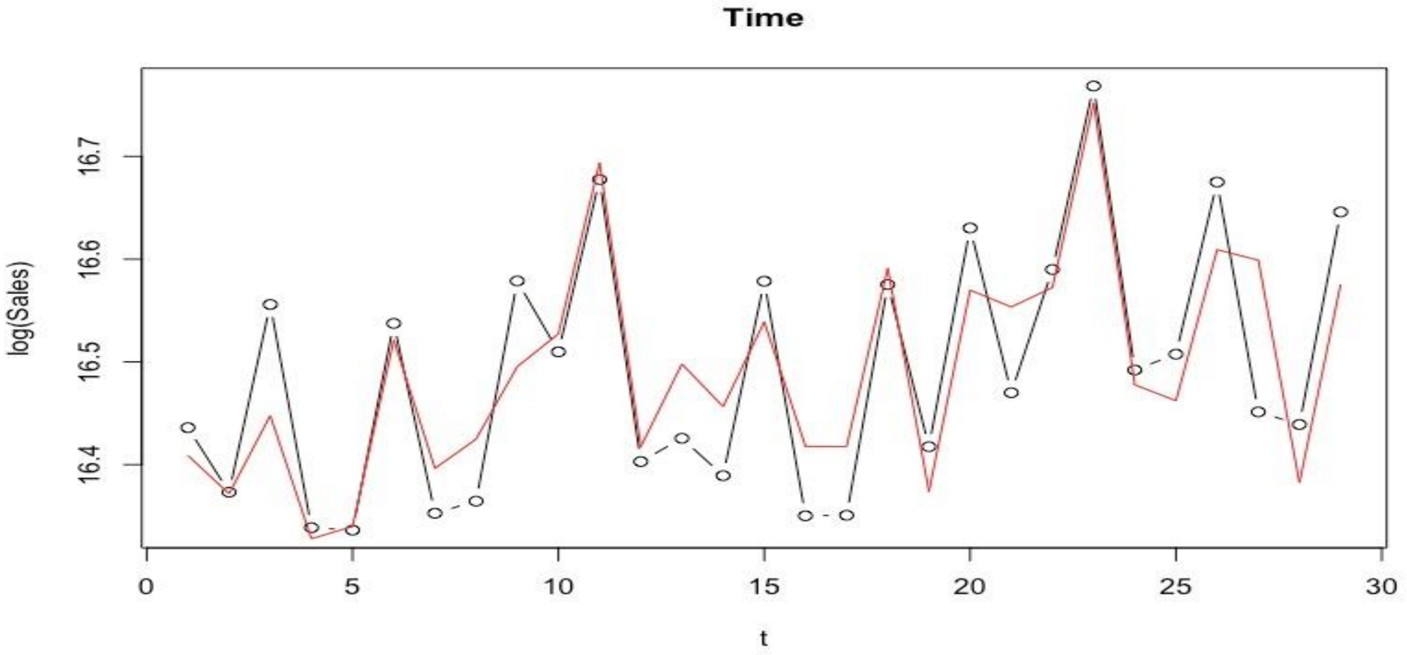


From our prior analysis shown in **[Plot-3]** and through result of Fixed affect model analysis department no. 92 i.e Dry Grocery have major share in the overall Sales across majority of the store. This led to perform time series analysis of sales for this particular department. We performed similar approach as above. Below is our model along with its plot.

[Table – 9: Time Series Analysis model]



```
Call:
lm(formula = log(Sales) ~ t + D1 + D2 + D3 + D4 + D5 + D6 + D7 +
    D8 + D9 + D10 + D11 + Lag_1, data = d)

Residuals:
     Min       1Q   Median       3Q      Max
-0.14763 -0.04363  0.01081  0.04363  0.10807

Coefficients:
              Estimate Std. Error t value Pr(>|t|)
(Intercept)  1.701e+01  1.850e-01  91.949  < 2e-16 ***
t            7.958e-03  2.063e-03   3.857 0.001552 **
D1          -2.754e-01  8.024e-02  -3.432 0.003707 **
D2          -2.174e-01  7.449e-02  -2.918 0.010592 *
D3          -1.842e-01  7.583e-02  -2.429 0.028188 *
D4          -2.076e-01  7.413e-02  -2.801 0.013445 *
D5          -3.364e-01  7.325e-02  -4.593 0.000352 ***
D6          -2.485e-01  7.445e-02  -3.338 0.004498 **
D7          -1.526e-01  8.097e-02  -1.884 0.079084 .
D8          -2.810e-01  8.193e-02  -3.430 0.003718 **
D9          -1.679e-01  8.240e-02  -2.037 0.059677 .
D10         -1.374e-01  8.332e-02  -1.649 0.119962
D11         -3.151e-02  1.009e-01  -0.312 0.759073
Lag_1       -3.028e-08  1.365e-08  -2.219 0.042361 *
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 0.08018 on 15 degrees of freedom
Multiple R-squared:  0.7586,    Adjusted R-squared:  0.5494
F-statistic: 3.626 on 13 and 15 DF,  p-value: 0.009819
```



9.0 Conclusion:

Markdown i.e. promotion can only be effective for departments where the sales are average.

Appendix:

Dept	Department Name
1	Candy and Tobacco
2	Health and Beauty Aids
3	Stationery
4	Paper Goods
5	Media and Gaming
6	Cameras and Supplies
7	Toys
8	Pet Supplies
9	Sporting Goods
10	Automotive
11	Hardware
12	Paint and Accessories
13	Household Chemicals
14	Kitchen and Dining
15	Small Appliances (defunct)
16	Outdoor Living
17	Home Décor
18	Seasonal
19	Crafts and Fabric
20	Bath and Shower
21	Books and Magazines
22	Bedding
23	Menswear
24	Boyswear
25	Shoes
26	Infant Apparel
27	Ladies' Socks
28	Hosiery
29	Sleepwear and Scrubs

30	Bras and Shapewear
31	Ladies Accessories
32	Jewelry
33	Girlswear
34	Ladieswear
35	Plus Size and Maternity
36	Ladies' Outerwear/Swimwear/Seasonal Apparel
37	Service Income (TLE)
38	Prescription Pharmacy
39	Radio Grill (defunct)
40	Pharmacy OTC
41	Team Sports Apparel
42	TLE Oil
43	Models
44	Crafts (expanded)
45	Sporting Goods (expanded)
46	Cosmetics
48	Firearms
49	Optical
51	Fishing
52	Floral (artificial)
55	Media and Gaming (expanded)
56	Live Plants and Flowers (Lawn and Garden)
57	Seasonal Pool and Spa Chemicals
58	Contract Wireless
59	Service Plans
60	Concept Stores and Stamps
67	Celebrations
69	Dot-Com (S2S)
71	Furniture
72	Electronics
74	Home Management and Luggage
77	Large Appliances (defunct)
79	Infant Consumables and Hardlines
80	Service Deli
81	Commercial Bread
82	Impulse and Checkout
83	Fresh/Frozen Seafood
84	Balloons and Fresh Flowers (defunct)

85	Photo Center
86	Financial Services
87	Wireless
88	In-Store Signage
90	Dairy
91	Frozen Foods
92	Dry Grocery
93	Fresh/Frozen Meat
94	Fresh Produce
95	Snacks and Beverages
96	Liquor
97	Packaged Deli
98	Fresh Bakery
99	Store Supplies