Time Series Forecasting

Walmart Store Sales

OPIM 5671
Data Mining and
Business Intelligence

Group 6

- Kamal Kannan Krishnan
- Deexith Reddy
- Mukund Ramachandran
- Namita Singh
- Xiaoyue Wang

Walmart

Agenda

- Introduction
- Problem Statement
- Data Description
- Data Exploration (EDA)
- Data Modification
- Modeling
    - Forecasting Total sales of all departments in Store 34
    - Forecasting 5 individual department sales in Store 34
- Business Findings & Recommendations

Introduction

- Sales Forecasting for Walmart using Time series data
- Data available from Feb 2010 to Nov 2012
- For 45 stores and 99 departments
- **Objective**
  - To forecast weekly sales for Store 34 which has 78 departments.
  - To forecast sales for 5 different departments under Store 34
- **Why STORE 34 ?**
  - Gasoline - Sold exclusively ++
  - Sales of Store 34 is near to Avg Sales of all 45 stores.
- **Methodology**: **SEMMA** Approach

Problem Statement

Retail Industry -Highly competitive, tighter margins.

Multiple Interface- Brick & Mortar and Online Ecommerce- Problem of Plenty.

Importance of Forecasting Sales

- Operational Efficiency -Supply chain & Inventory management
- Demand Forecasting- Just in Time
- Predicting potential Strategic shifts in Marketing and customer Promotions
- Financial planning & Internal controls

Data Description

16 Variables

- ***Time Variable:***
  - **Date -** the end date of weekly cycle from Feb 2010 to Nov 2012.
- ***Dependent Variable:***
  - **Weekly Sales -** Sales for a given department in the given store in dollars($).
- ***Independent Variables (Used in Modeling):***
  - **Is Holiday -** Whether the week is a holiday week where **1 - Holiday** and **0 - Not a Holiday**. Holidays are Labor day, Valentines day, Super Bowl, Thanksgiving and Christmas.
  - **Temperature(Fahrenheit) -** Temperature in the region where the store is located.
  - **Fuel Price -** Cost of fuel in dollars($) in the region where the store is located.
  - **Mark Down1 -** 1st round of Promotional Markdown in dollars($).
  - **MarkDown2 -** 2nd round of Promotional Markdown in dollars($).
  - **Mark Down3 -** 3rd round of Promotional Markdown in dollars($).
  - **Mark Down4 -** 4th round of Promotional Markdown in dollars($).
  - **Mark Down5 -** 5th round of Promotional Markdown in dollars($) .
  - **CPI -** Consumer Price Index.
  - **Unemployment -** Unemployment rate in percentage.
- ***Independent Variables (Not used in Modeling)***
  - **Dept -** The department number numbered from **1-99** in Walmart.
  - **Store -** Store numbered from **1-45** in Walmart.
  - **Size -** Sizes in Square Feet for each Store.
  - **Type -** Types of Stores labeled with A,B and C.

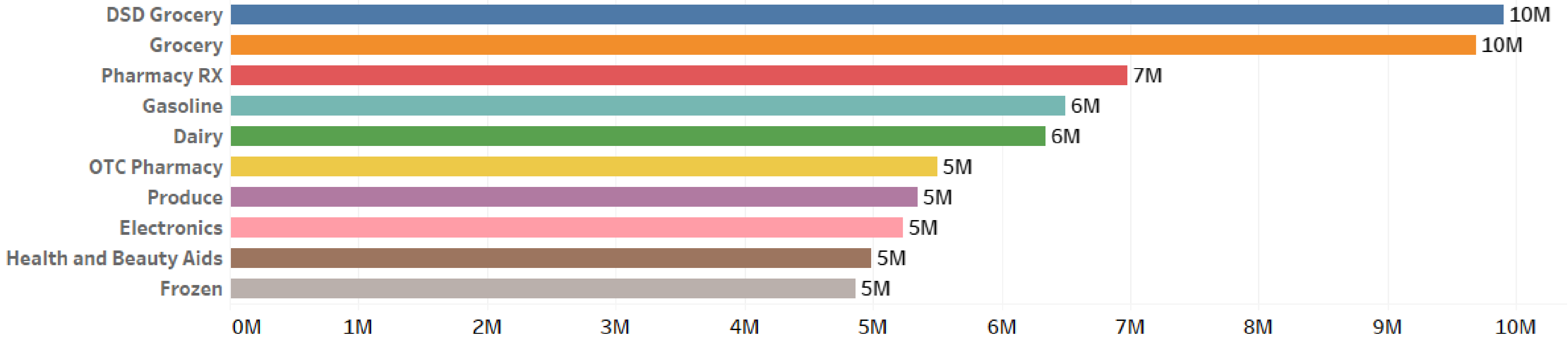
Data Exploration of Store 34

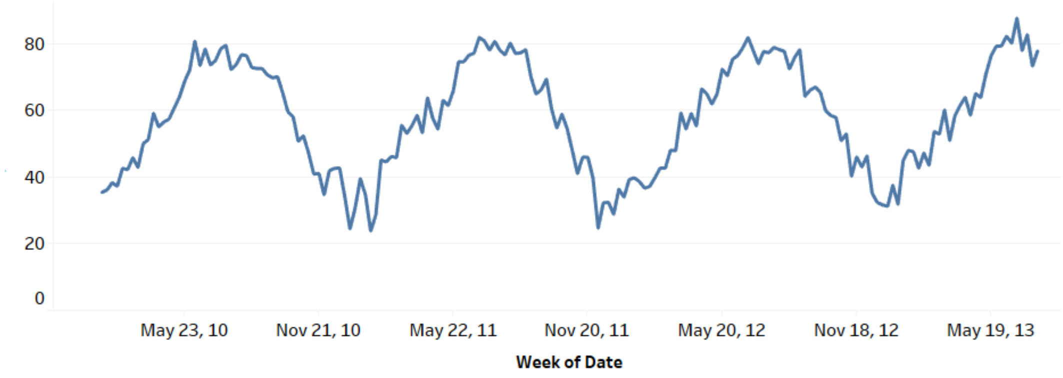
Store Size

158,114

Square Feet

Top 10 departments by Sales



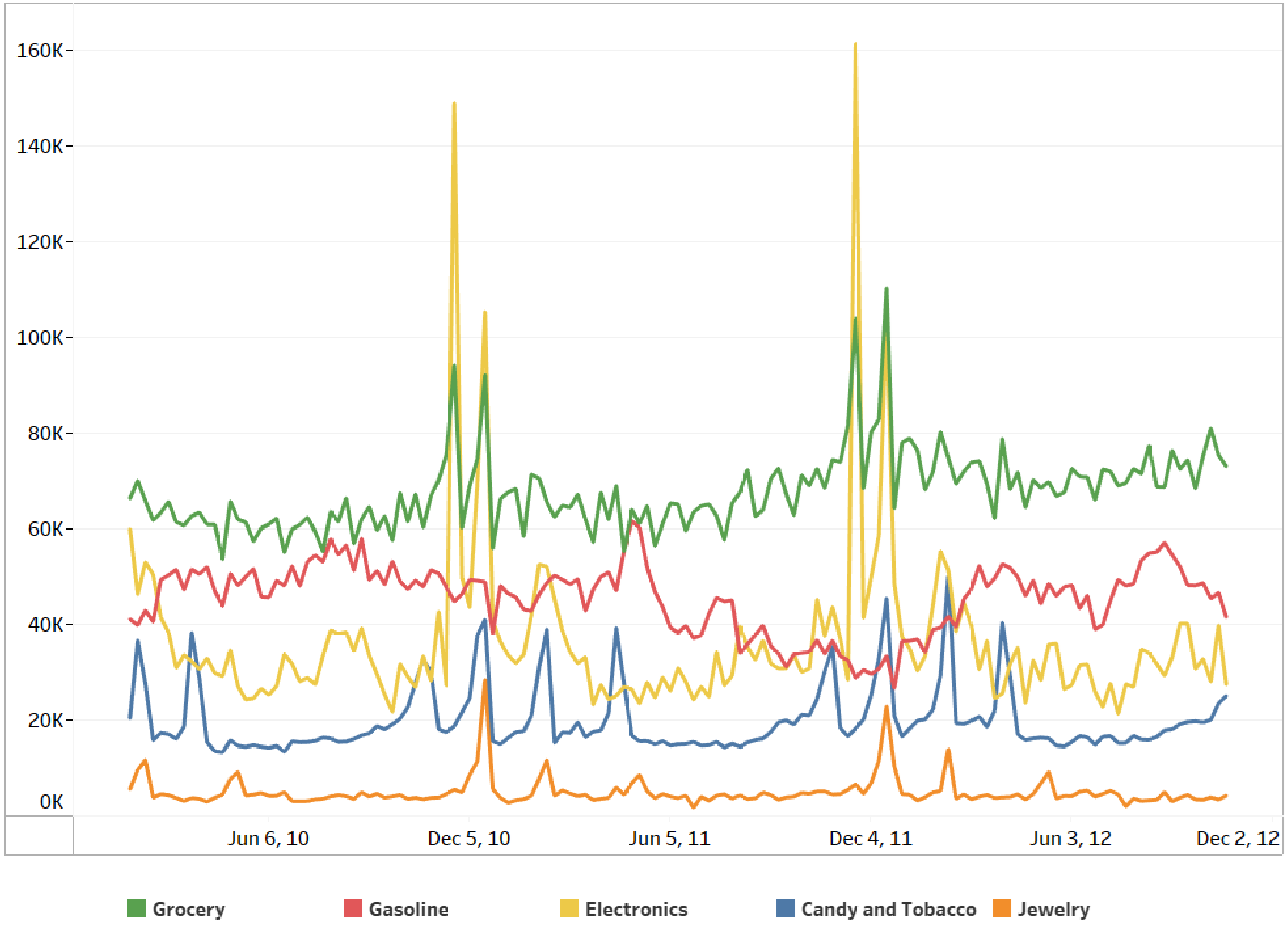


From Temperature Graph, Store Location seems to be in Northern states of the USA (Chicago / Boston / Seattle Area).

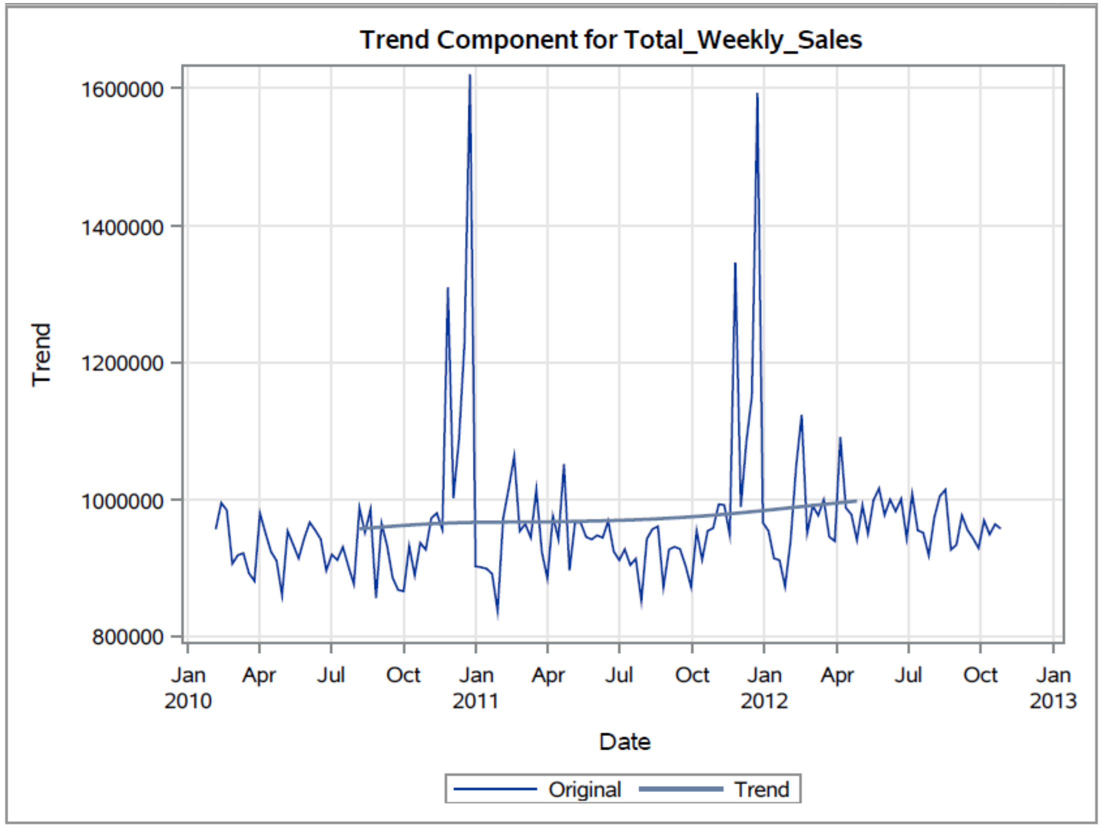
Exact location information is Not Available.

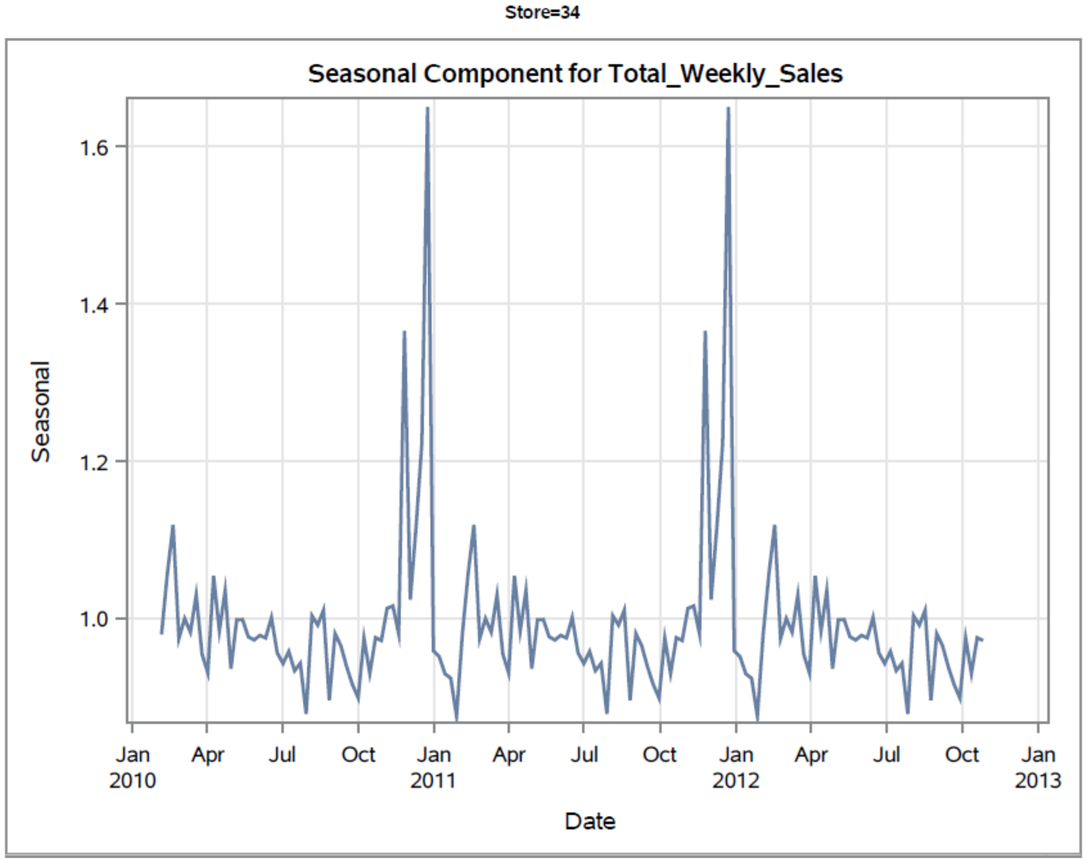
Data Exploration - 5 Departments in Store 34

- ***Grocery*** - Highest Sales. Slightly positive trend with seasonality
- ***Gasoline*** - Opposite trend during Christmas. No seasonality.
- ***Electronics*** - No trend but Seasonality. Huge Spike during Christmas season.
- ***Candy and Tobacco*** - Only series with 4 clear spikes with seasonality. Seasonality but no trend.
- ***Jewelry*** - Lowest Sales. Seasonality but no trend.



Data Seasonality





Data Modification

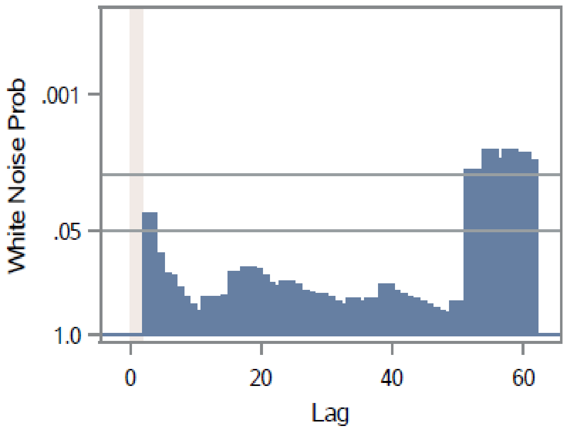
- Filled missing values in “Markdown” replaced with Zero
- Accumulated by summing the Sales of all individual departments to calculate Total weekly sales of Store 34 per week;
- Number of rows : 143
    - Fit Sample: 123 weeks
    - Holdout Sample: 20 weeks (from 06-2012 to 10-2012)
- Used only 20 rows for Test sample; since we cannot use more than 25% percent as test sample.
- Wanted more Rows for Fit sample in order to model the seasonality accurately
- Cannot use stratified sample as time series requires continuous data to test

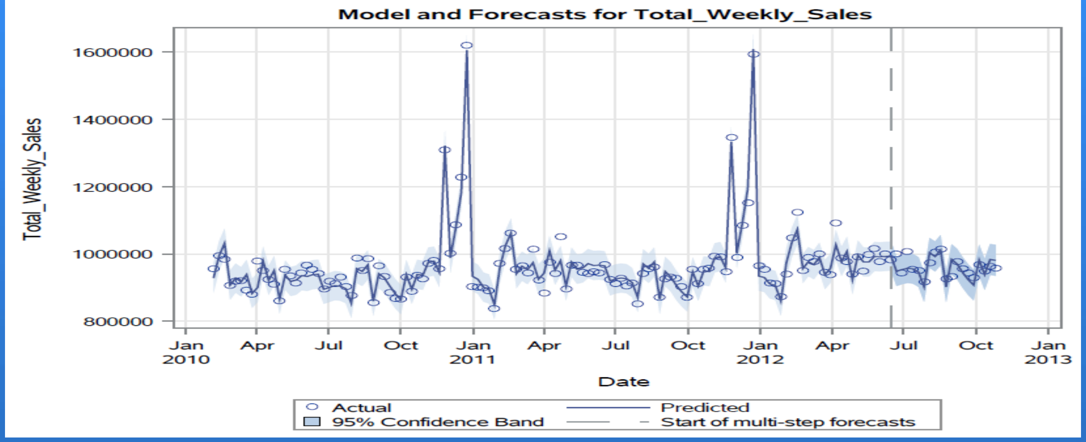
ESM Models - Total Sales of All Departments

Residuals

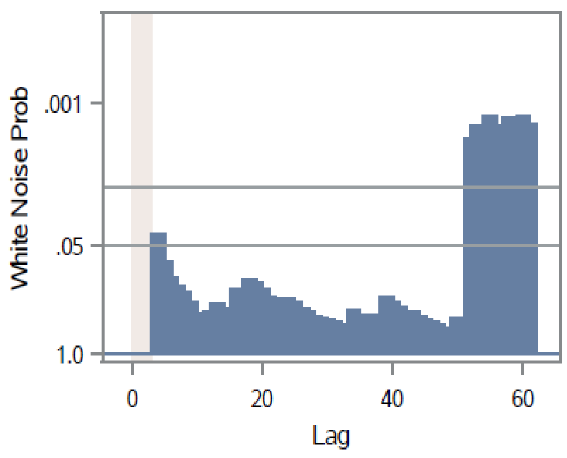
Model

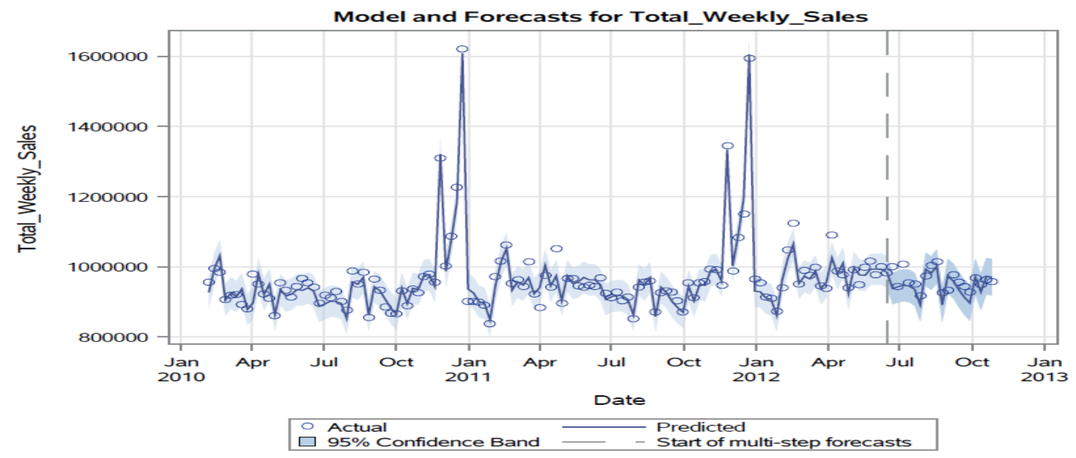
- *Additive Seasonal*
  - *MAPE = 2.3%*
  - *WMAE = 2152*





- *Additive Winters*
  - *MAPE = 2.017%*
  - *WMAE = 17289*





Arimax Models - Total Sales of all Departments

Methodology

- Determining order of differencing
  - dif=0
  - sdif=1
- Determining p,q values
  - possible p:1,4,5
  - possible q:1,2,4
- Dealing with autocorrelation in explanatory variables
- Check cross correlation between target variable and explanatory variables
  - 8 variables: Temperature, Fuel_Price, CPI, Unemployment IsHoliday, Mark Down 2, Mark Down 3, Mark Down 5
- Stepwise variable selection
- Final model

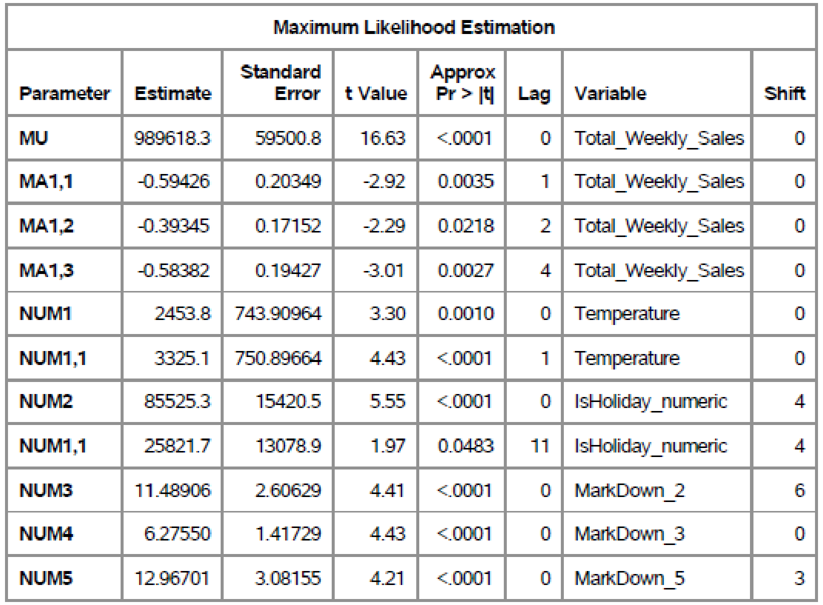
ARIMAX Model 1 - Total Sales of all Departments

Model Parameters

- dif=0, sdif=0
- p=0
- q=1,2,4
- inputs
  - (1)Temperature
  - 4 $ (11)IsHoliday
  - 6 $ MarkDown2
  - MarkDown3
  - 3 $ MarkDown5

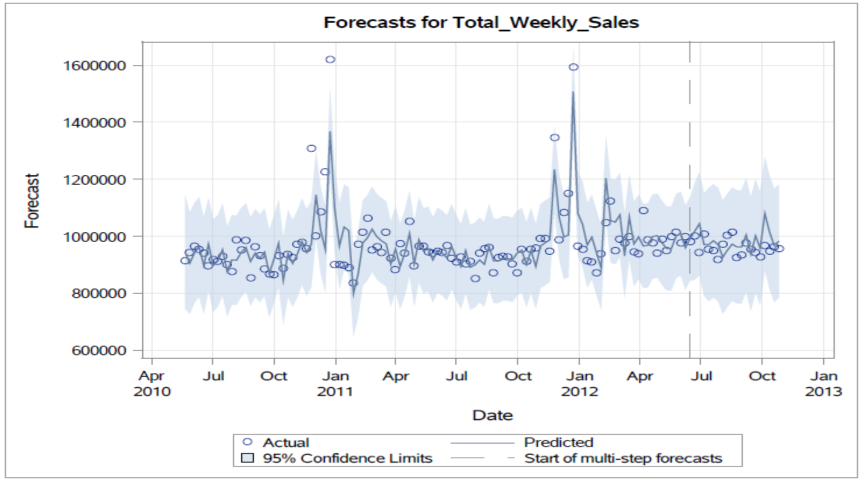
Accuracy Statistics

- MAPE=6.13%
- WMAE=60290

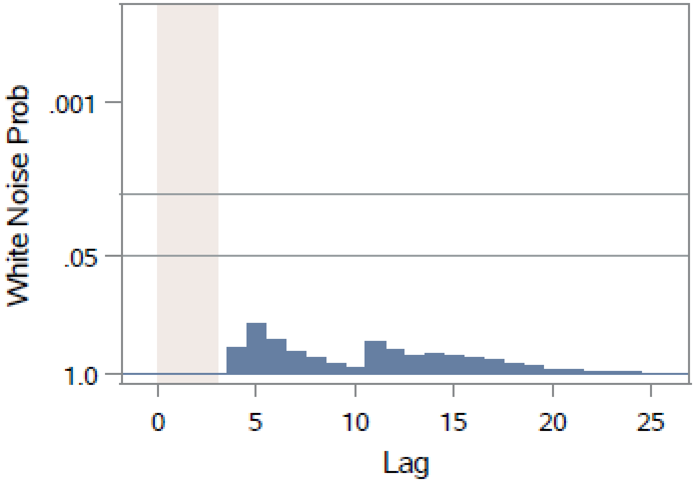


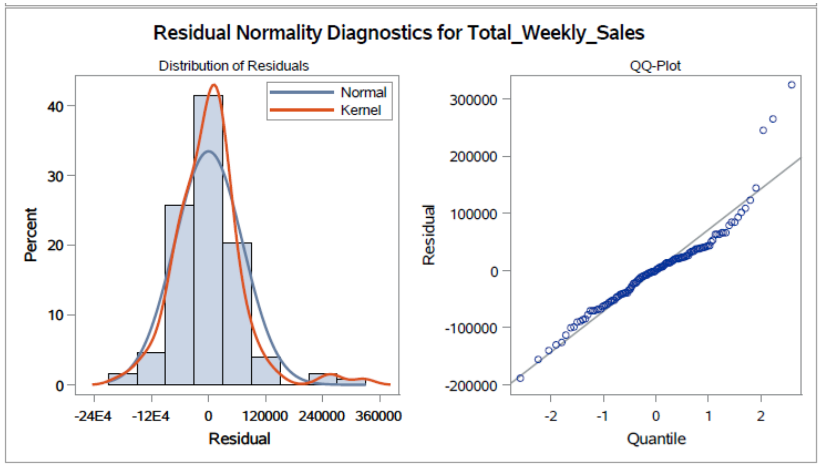
Maximum Likelihood Estimation

Parameter	Estimate	Standard Error	t Value	Approx Pr > |t|	Lag	Variable	Shift
MU	989618.3	59500.8	16.63	<.0001	0	Total_Weekly_Sales	0
MA1,1	-0.59426	0.20349	-2.92	0.0035	1	Total_Weekly_Sales	0
MA1,2	-0.39345	0.17152	-2.29	0.0218	2	Total_Weekly_Sales	0
MA1,3	-0.58382	0.19427	-3.01	0.0027	4	Total_Weekly_Sales	0
NUM1	2453.8	743.90964	3.30	0.0010	0	Temperature	0
NUM1,1	3325.1	750.89664	4.43	<.0001	1	Temperature	0
NUM2	85525.3	15420.5	5.55	<.0001	0	IsHoliday_numeric	4
NUM1,1	25821.7	13078.9	1.97	0.0483	11	IsHoliday_numeric	4
NUM3	11.48906	2.60629	4.41	<.0001	0	MarkDown_2	6
NUM4	6.27550	1.41729	4.43	<.0001	0	MarkDown_3	0
NUM5	12.96701	3.08155	4.21	<.0001	0	MarkDown_5	3



Residuals





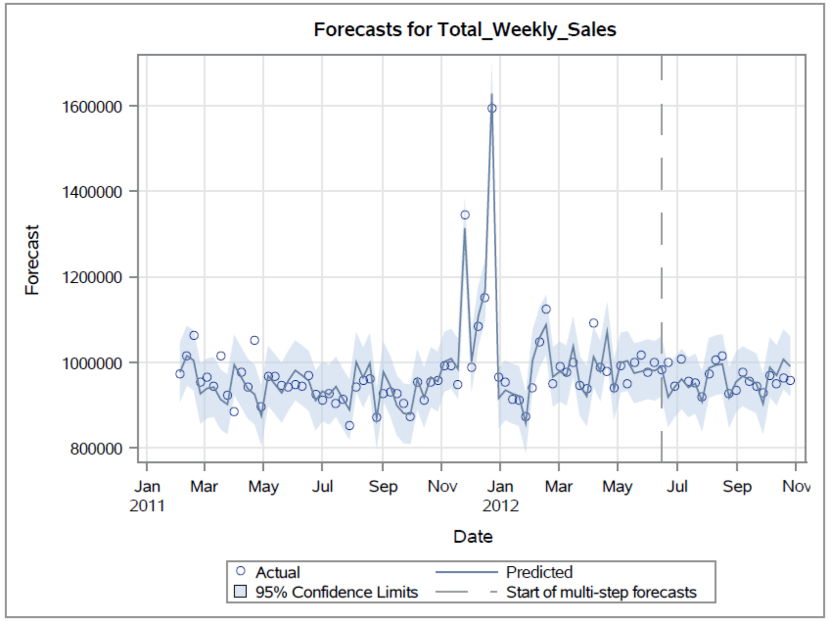
ARIMAX Model 2 - Total Sales of all Departments

Model Parameters

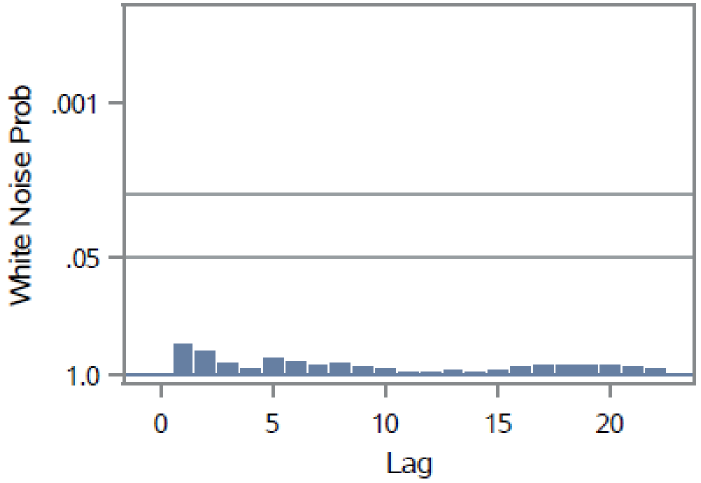
- dif=0, sdif=1
- p=0
- q=0
- Inputs
  - Unemployment
  - 3 $ MarkDown3

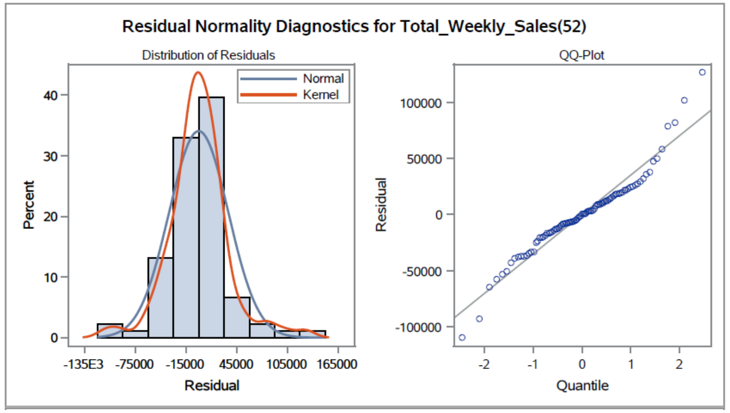
Accuracy Statistics

- MAPE=2.05%
- WMAE=18133

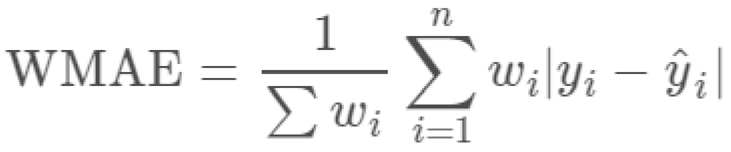


Residuals

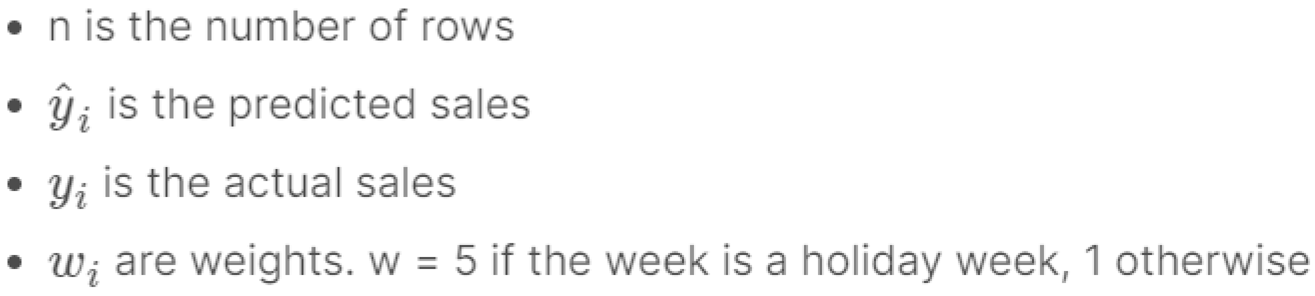




Model Comparison - Total Sales of All Departments



$$\text{WMAE} = \frac{1}{\sum w_i} \sum_{i=1}^{n} w_i |y_i - \hat{y}_i|$$



- n is the number of rows
- $\hat{y}_i$ is the predicted sales
- y_i is the actual sales
- w_i are weights. w = 5 if the week is a holiday week, 1 otherwise

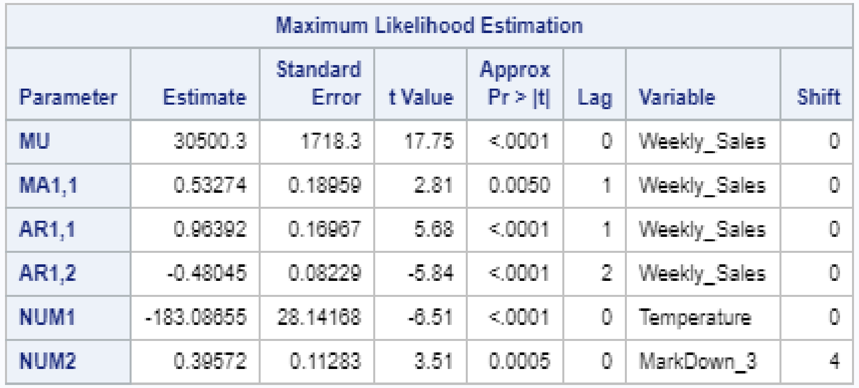
Model	MAPE	WMAE
Additive Seasonal	2.3%	21527
Additive Winters	2.017%	17289
Arimax Model 1	6.13%	60290
Arimax Model 2	2.05%	18133

Candy & Tobacco Department Sales

STORE 34

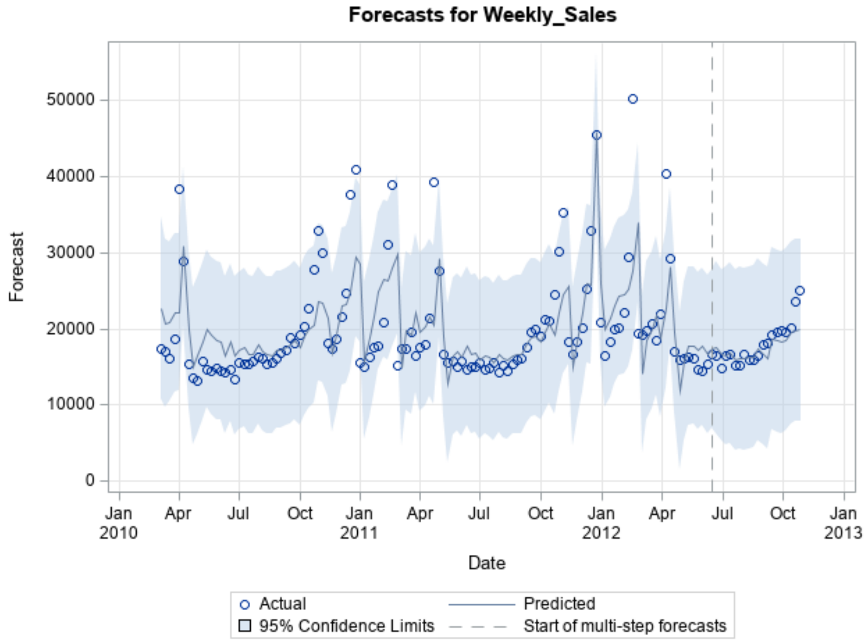
- **Seasonal Data; High sales during each Holidays in the year**
- **4 peak seen in “trend graph”** for each 12 month period.

- ***Independent Variables:***
  - Correlated with AR ; lag 1 and 2
  - **Temperature is negatively correlated**
  - **Markdown 3 is positively correlated after 4 weeks**



Maximum Likelihood Estimation

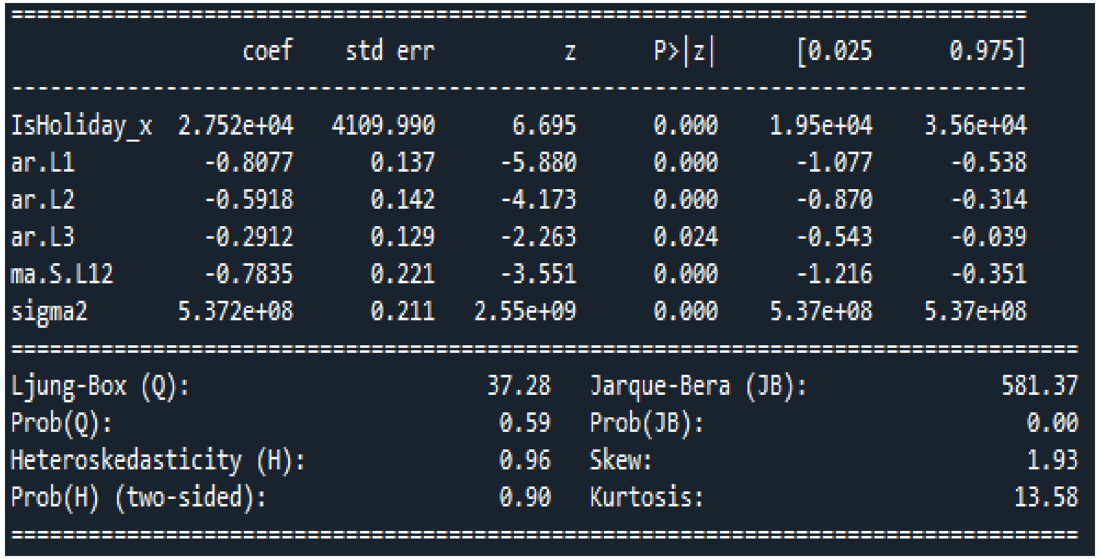
Parameter	Estimate	Standard Error	t Value	Approx Pr > |t|	Lag	Variable	Shift
MU	30500.3	1718.3	17.75	<.0001	0	Weekly_Sales	0
MA1,1	0.53274	0.18959	2.81	0.0050	1	Weekly_Sales	0
AR1,1	0.96392	0.16967	5.68	<.0001	1	Weekly_Sales	0
AR1,2	-0.48045	0.08229	-5.84	<.0001	2	Weekly_Sales	0
NUM1	-183.08855	28.14168	-6.51	<.0001	0	Temperature	0
NUM2	0.39572	0.11283	3.51	0.0005	0	MarkDown_3	4



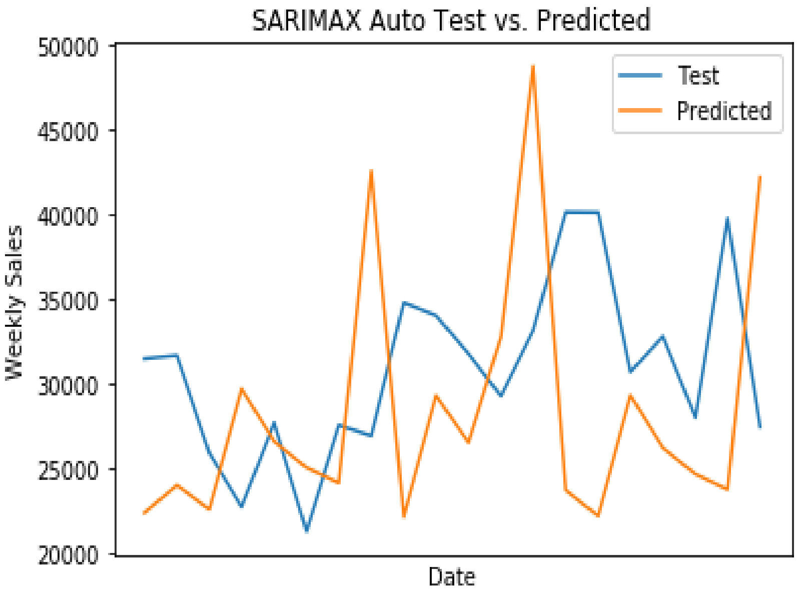
Electronics Department Sales

STORE 34

- **Seasonal Data**
- **High sales during end of year holidays (Black Friday sales and Christmas sales)**
- From estimates we see confirm “IsHoliday” is significant
- With every holiday, the sales increases by approx. 27,500



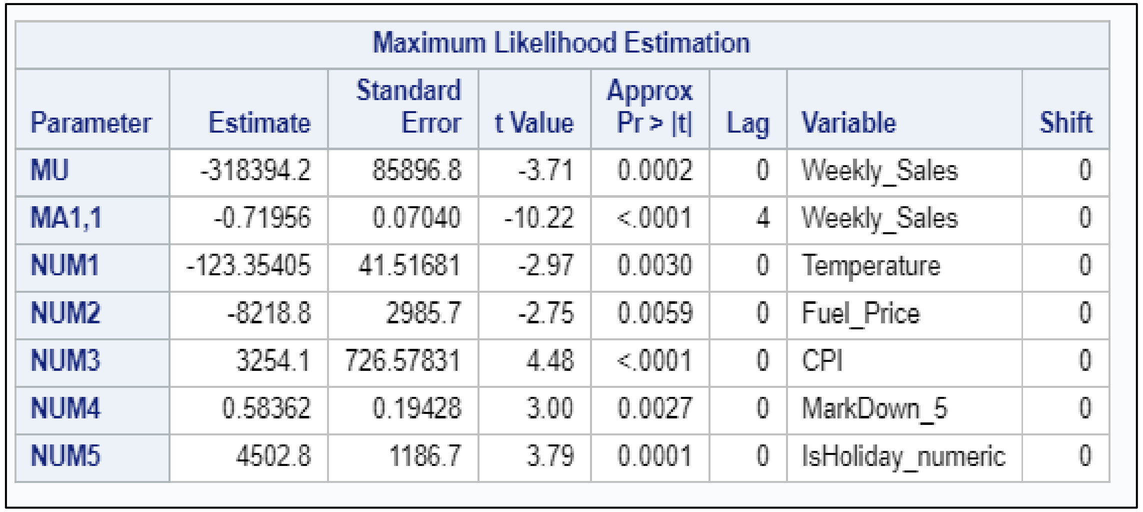
```
==============================================================================
                 coef    std err          z      P>|z|      [0.025      0.975]
------------------------------------------------------------------------------
IsHoliday_x  2.752e+04   4109.990      6.695      0.000    1.95e+04    3.56e+04
ar.L1          -0.8077      0.137     -5.880      0.000      -1.077      -0.538
ar.L2          -0.5918      0.142     -4.173      0.000      -0.870      -0.314
ar.L3          -0.2912      0.129     -2.263      0.024      -0.543      -0.039
ma.S.L12       -0.7835      0.221     -3.551      0.000      -1.216      -0.351
sigma2       5.372e+08      0.211   2.55e+09      0.000    5.37e+08    5.37e+08
===================================================================================
Ljung-Box (Q):                       37.28   Jarque-Bera (JB):               581.37
Prob(Q):                              0.59   Prob(JB):                         0.00
Heteroskedasticity (H):               0.96   Skew:                             1.93
Prob(H) (two-sided):                  0.90   Kurtosis:                        13.58
===================================================================================
```



Groceries Department Sales

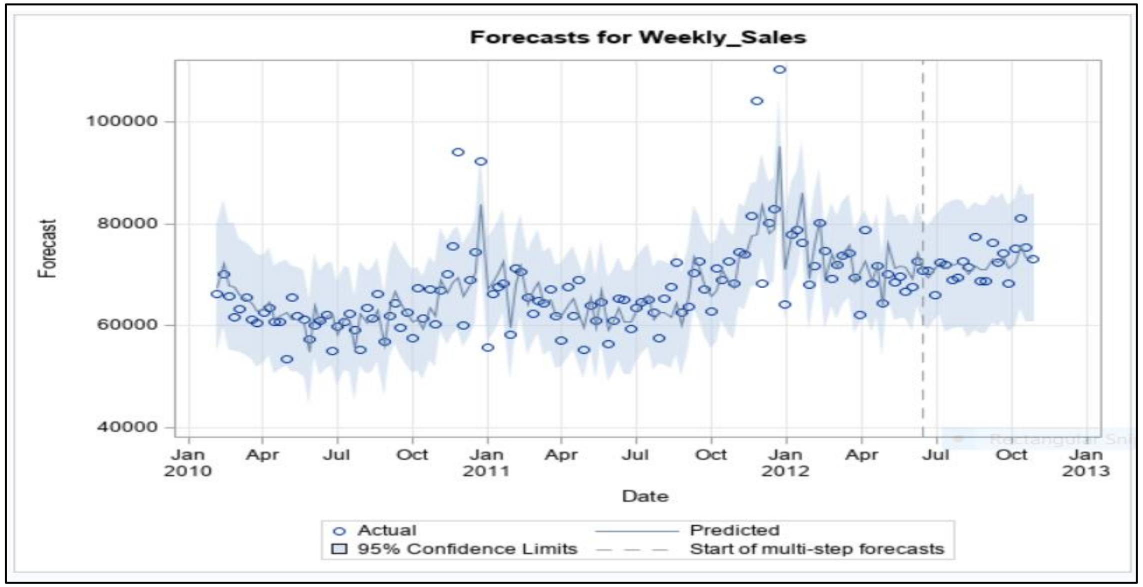
STORE 34

- **Seasonal Data with slightly positive Trend.**
- **Fuel price and Temperature are negatively correlated (lag 0) to weekly sales.**
- The effect of change in Temperature and fuel price takes place with immediate effect in Concurrent lag.



Maximum Likelihood Estimation

Parameter	Estimate	Standard Error	t Value	Approx Pr > |t|	Lag	Variable	Shift
MU	-318394.2	85896.8	-3.71	0.0002	0	Weekly_Sales	0
MA1,1	-0.71956	0.07040	-10.22	<.0001	4	Weekly_Sales	0
NUM1	-123.35405	41.51681	-2.97	0.0030	0	Temperature	0
NUM2	-8218.8	2985.7	-2.75	0.0059	0	Fuel_Price	0
NUM3	3254.1	726.57831	4.48	<.0001	0	CPI	0
NUM4	0.58362	0.19428	3.00	0.0027	0	MarkDown_5	0
NUM5	4502.8	1186.7	3.79	0.0001	0	IsHoliday_numeric	0

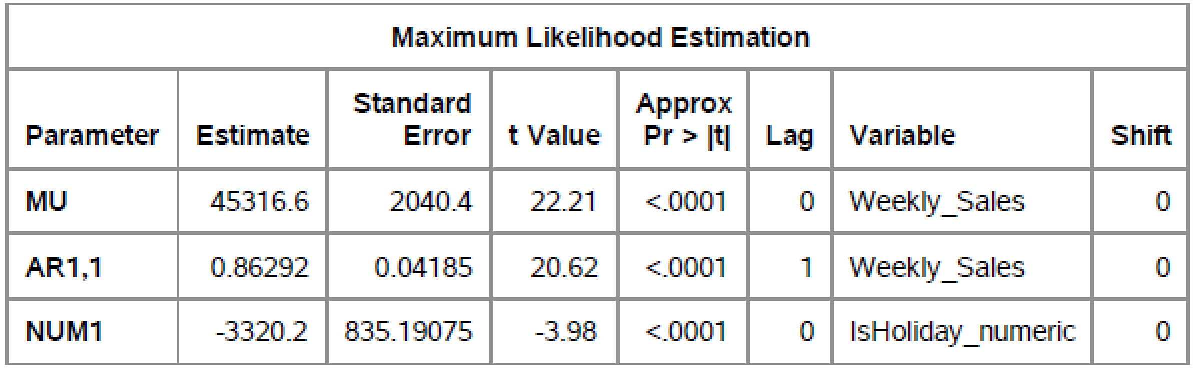


Gasoline Department Sales

STORE 34

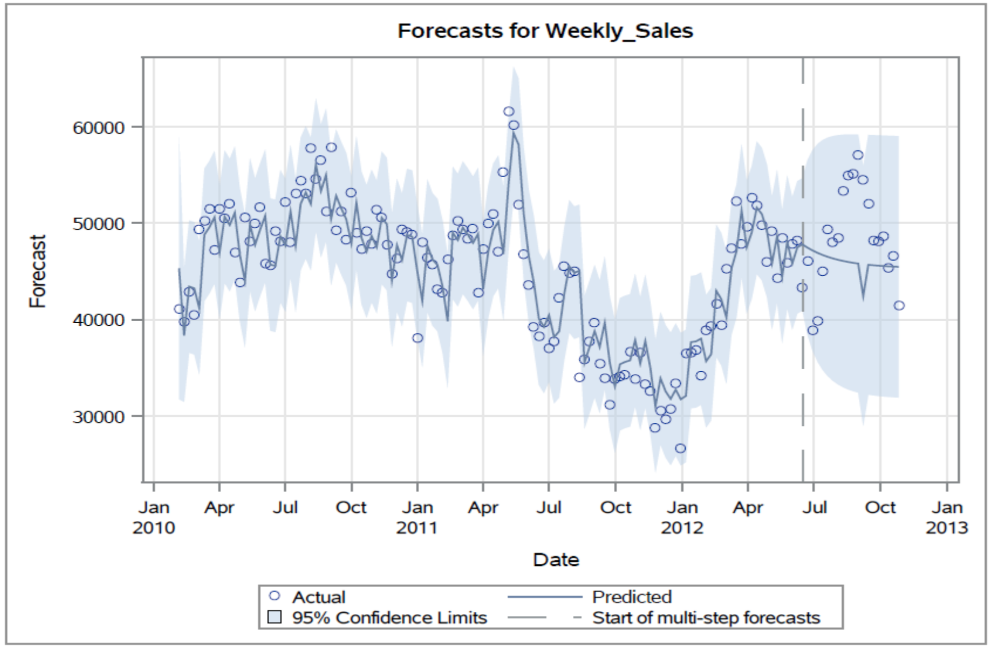
- **Consumers tend to consume less gasoline during holidays** than in ordinary workdays.

- In general, **gasoline sales of the week is positively correlated with sales of the previous week.**



Maximum Likelihood Estimation

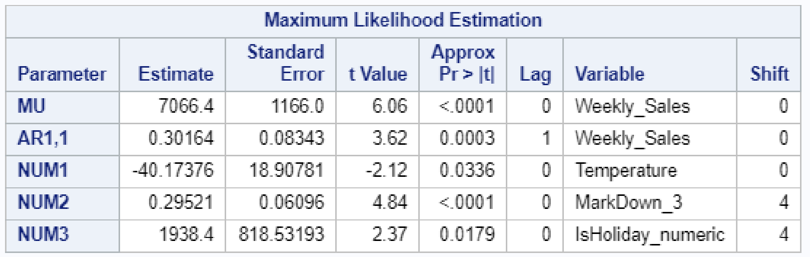
Parameter	Estimate	Standard Error	t Value	Approx Pr > |t|	Lag	Variable	Shift
MU	45316.6	2040.4	22.21	<.0001	0	Weekly_Sales	0
AR1,1	0.86292	0.04185	20.62	<.0001	1	Weekly_Sales	0
NUM1	-3320.2	835.19075	-3.98	<.0001	0	IsHoliday_numeric	0



Jewelry Department Sales

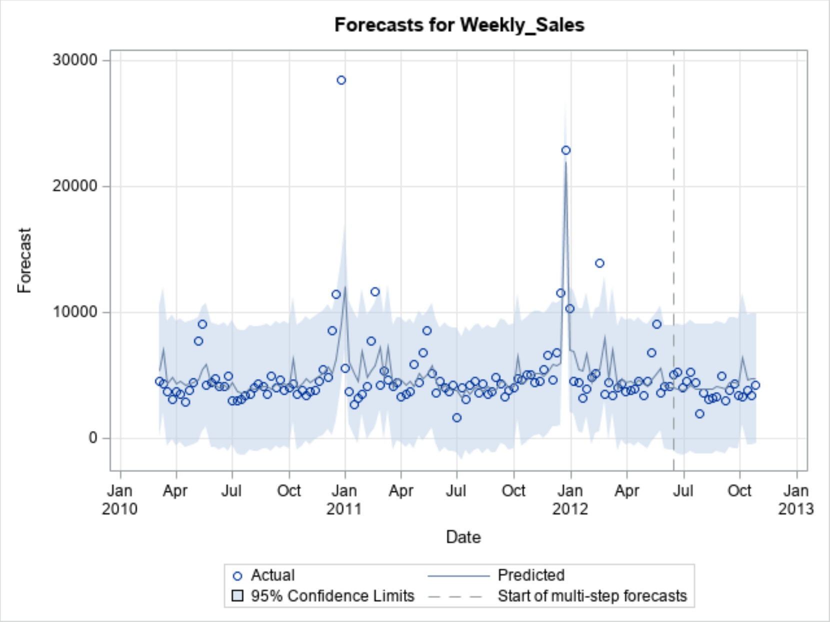
STORE 34

- People tend to **spend more on Jewellery during holidays and positive impact** is shown **after 4 weeks.**
- **Temperature is negatively correlated** with weekly sales. When temperature increases, weekly sales drop.
- **Markdown 3 has positive correlation** with weekly sales. Positive impact shows **after 4 weeks.**



Maximum Likelihood Estimation

Parameter	Estimate	Standard Error	t Value	Approx Pr > |t|	Lag	Variable	Shift
MU	7066.4	1166.0	6.06	<.0001	0	Weekly_Sales	0
AR1,1	0.30164	0.08343	3.62	0.0003	1	Weekly_Sales	0
NUM1	-40.17376	18.90781	-2.12	0.0336	0	Temperature	0
NUM2	0.29521	0.06096	4.84	<.0001	0	MarkDown_3	4
NUM3	1938.4	818.53193	2.37	0.0179	0	IsHoliday_numeric	4



Business Findings & Recommendations

- ***"Markdown" Variable:***
  - Markdown 3 is negatively impacting the Total Weekly Sales of Store 34.
  - Different Markdowns have positive impact on individual departments; May be due to selective items being marked down.
- ***"IsHoliday" Variable:***
  - Positively correlated to the Weekly Sales of Grocery, Jewellery, Electronics, Candy.
  - Negatively correlated to Gasoline; May be many people prefer staying home.
- ***"Temperature" Variable:*** Show negative correlation; may be due to seasonal trend of sales on in Store 34 is reflected through "temperature" variable
- ***"Unemployment" Variable***: Has Negative effect on the Total Sales of Store 34
- ***"CPI, Fuel Price" Variable:*** No effect on the Total Sales of Store 34

Business Findings & Recommendations

- Recommendation is to not have Markdown 3 for Store 34, as it negatively affects Total Sales after 3 weeks.
- More intelligent analysis can be done when more information made available like:
  - Actual “Store Location Information”
  - Number of Persons visiting stores each week
  - Median income of neighbourhood
  - Demographic of customers in neighbourhood

Thank You

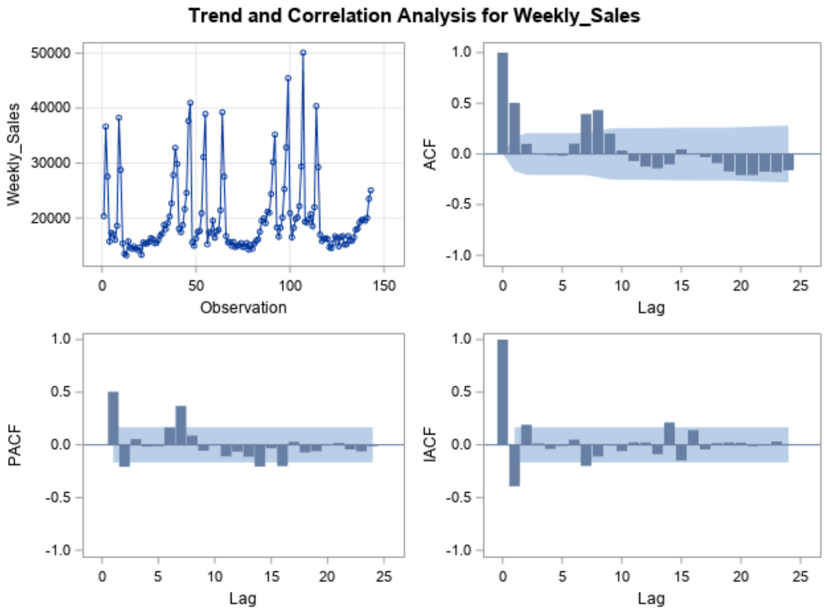
Questions ?

Appendix

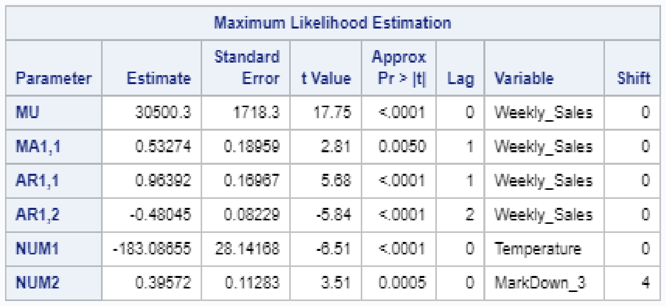
Walmart

Appendix - Candy Department Analysis

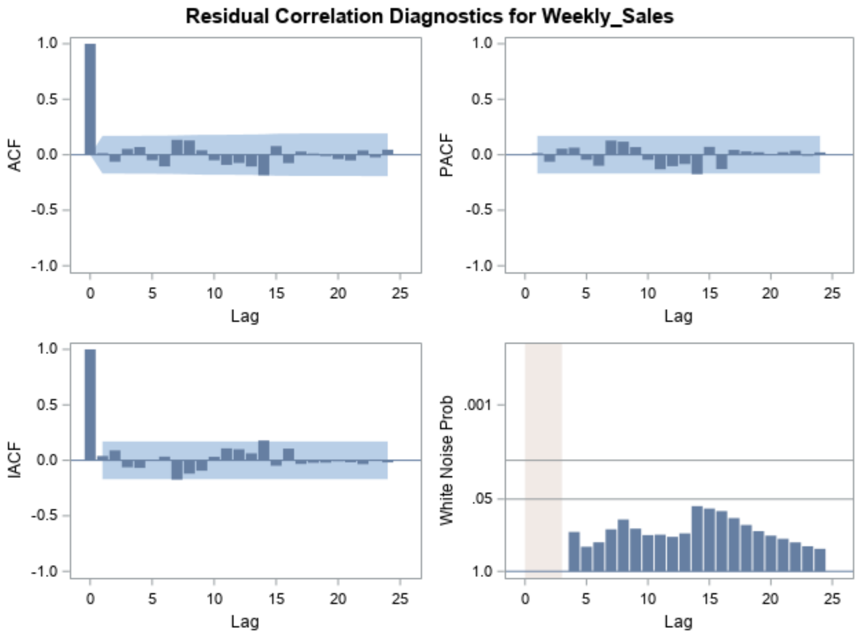
STORE 34



Trend and Correlation Analysis for Weekly_Sales

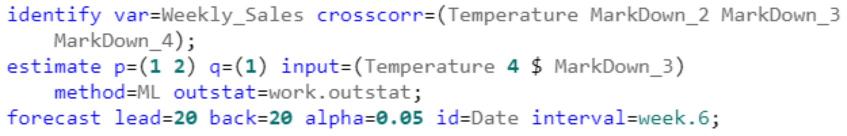


Maximum Likelihood Estimation							
Parameter	Estimate	Standard Error	t Value	Approx Pr > |t|	Lag	Variable	Shift
MU	30500.3	1718.3	17.75	<.0001	0	Weekly_Sales	0
MA1,1	0.53274	0.18959	2.81	0.0050	1	Weekly_Sales	0
AR1,1	0.96392	0.16967	5.68	<.0001	1	Weekly_Sales	0
AR1,2	-0.48045	0.08229	-5.84	<.0001	2	Weekly_Sales	0
NUM1	-183.08655	28.14168	-6.51	<.0001	0	Temperature	0
NUM2	0.39572	0.11283	3.51	0.0005	0	MarkDown_3	4



Residual Correlation Diagnostics for Weekly_Sales

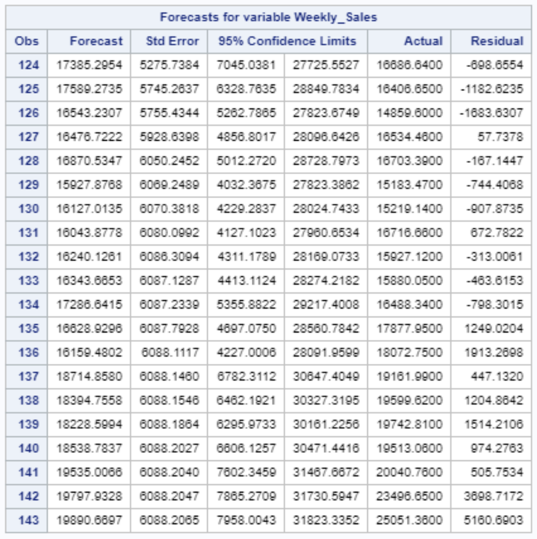
- **Seasonal Data; High sales during Holidays**
- **Best fit is ARMAX (p=2,q=1) model**
- **Dependent Var: Temperature, Markdown 3**



```
identify var=Weekly_Sales crosscorr=(Temperature MarkDown_2 MarkDown_3
    MarkDown_4);
estimate p=(1 2) q=(1) input=(Temperature 4 $ MarkDown_3)
    method=ML outstat=work.outstat;
forecast lead=20 back=20 alpha=0.05 id=Date interval=week.6;
```

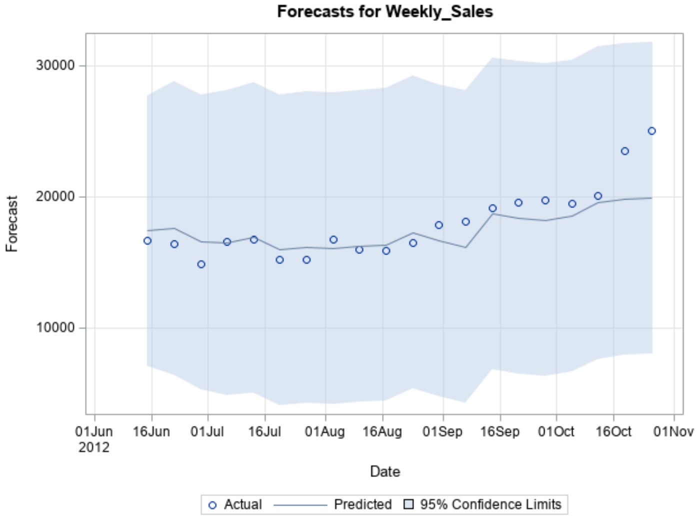
Appendix - Candy Department Model Output

STORE 34

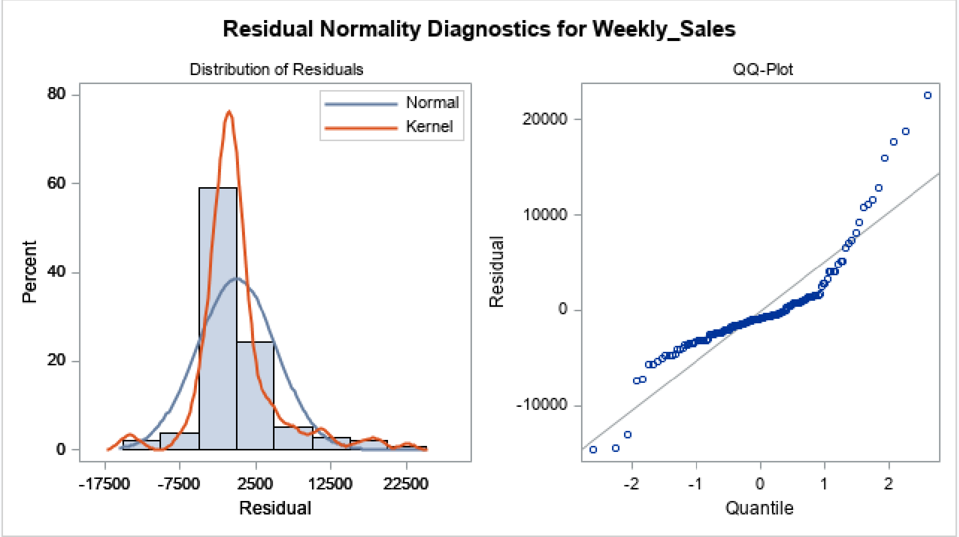


Forecasts for variable Weekly_Sales

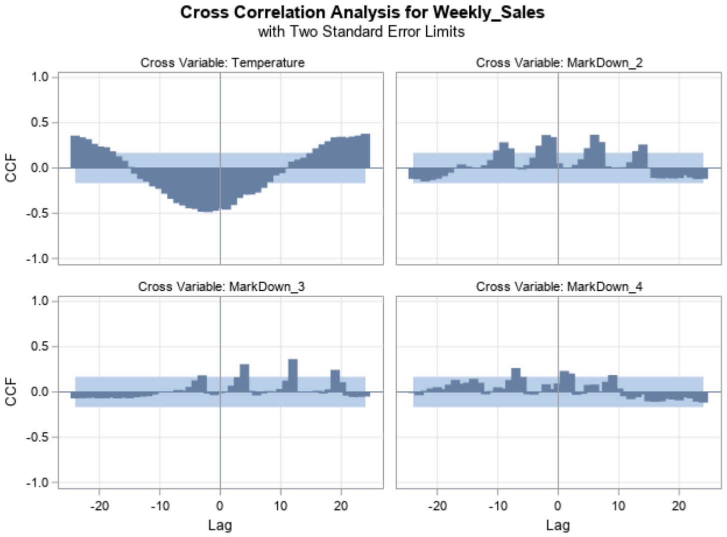
Obs	Forecast	Std Error	95% Confidence Limits		Actual	Residual
124	17385.2954	5275.7384	7045.0381	27725.5527	16686.6400	-698.6554
125	17589.2735	5745.2637	6328.7635	28849.7834	16406.6500	-1182.6235
126	16543.2307	5755.4344	5262.7865	27823.6749	14859.6000	-1683.6307
127	16476.7222	5928.6398	4856.8017	28096.6426	16534.4600	57.7378
128	16870.5347	6050.2452	5012.2720	28728.7973	16703.3900	-167.1447
129	15927.8768	6069.2489	4032.3675	27823.3862	15183.4700	-744.4068
130	16127.0135	6070.3818	4229.2837	28024.7433	15219.1400	-907.8735
131	16043.8778	6080.0992	4127.1023	27960.6534	16716.6600	672.7822
132	16240.1261	6086.3094	4311.1789	28169.0733	15927.1200	-313.0061
133	16343.6653	6087.1287	4413.1124	28274.2182	15880.0500	-463.6153
134	17286.6415	6087.2339	5355.8822	29217.4008	16488.3400	-798.3015
135	16628.9296	6087.7928	4697.0750	28560.7842	17877.9500	1249.0204
136	16159.4802	6088.1117	4227.0006	28091.9599	18072.7500	1913.2698
137	18714.8580	6088.1460	6782.3112	30647.4049	19161.9900	447.1320
138	18394.7558	6088.1546	6462.1921	30327.3195	19599.6200	1204.8642
139	18228.5994	6088.1864	6295.9733	30161.2256	19742.8100	1514.2106
140	18538.7837	6088.2027	6606.1257	30471.4416	19513.0600	974.2763
141	19535.0066	6088.2040	7602.3459	31467.6672	20040.7600	505.7534
142	19797.9328	6088.2047	7865.2709	31730.5947	23496.6500	3698.7172
143	19890.6697	6088.2065	7958.0043	31823.3352	25051.3600	5160.6903



Forecasts for Weekly_Sales



Residual Normality Diagnostics for Weekly_Sales

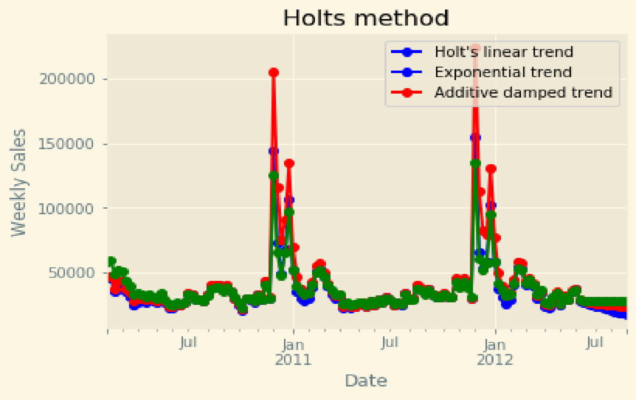


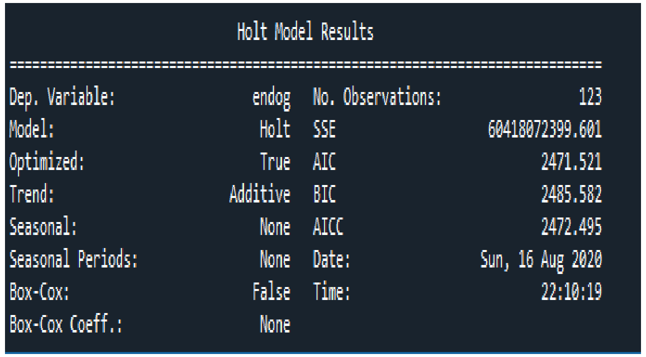
Cross Correlation Analysis for Weekly_Sales

with Two Standard Error Limits

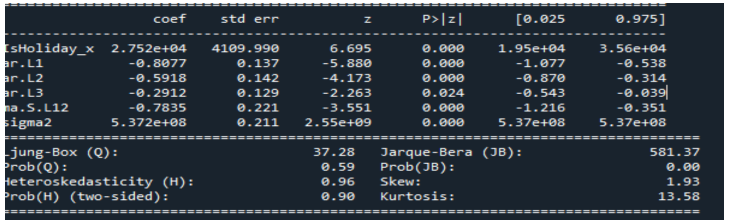
Appendix - Electronics Department Sales

STORE 34

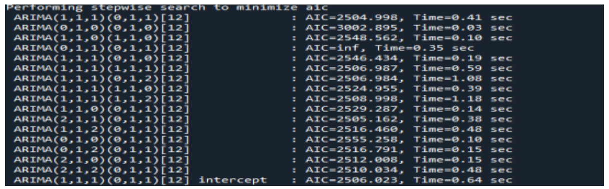




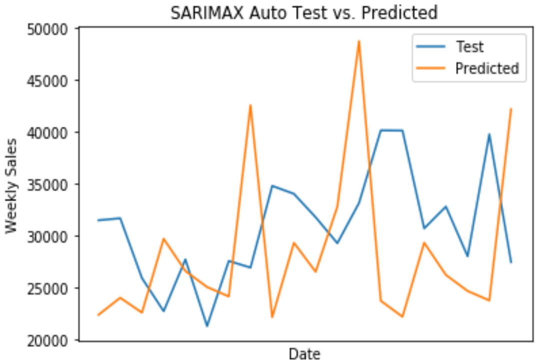
```
                        Holt Model Results
==============================================================================
Dep. Variable:                endog   No. Observations:                    123
Model:                         Holt   SSE                   60418072399.601
Optimized:                     True   AIC                          2471.521
Trend:                     Additive   BIC                          2485.582
Seasonal:                      None   AICC                         2472.495
Seasonal Periods:              None   Date:                 Sun, 16 Aug 2020
Box-Cox:                      False   Time:                         22:10:19
Box-Cox Coeff.:                None
```

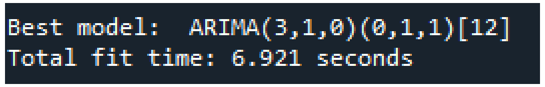


```
==============================================================================
                 coef    std err          z      P>|z|      [0.025      0.975]
------------------------------------------------------------------------------
IsHoliday_x  2.752e+04  4109.990      6.695      0.000    1.95e+04    3.56e+04
ar.L1          -0.8077     0.137     -5.880      0.000      -1.077      -0.538
ar.L2          -0.5918     0.142     -4.173      0.000      -0.870      -0.314
ar.L3          -0.2912     0.129     -2.263      0.024      -0.543      -0.039
ma.S.L12       -0.7835     0.221     -3.551      0.000      -1.216      -0.351
sigma2       5.372e+08     0.211   2.55e+09      0.000    5.37e+08    5.37e+08
===================================================================================
Ljung-Box (Q):                       37.28   Jarque-Bera (JB):               581.37
Prob(Q):                              0.59   Prob(JB):                         0.00
Heteroskedasticity (H):               0.96   Skew:                             1.93
Prob(H) (two-sided):                  0.90   Kurtosis:                        13.58
===================================================================================
```

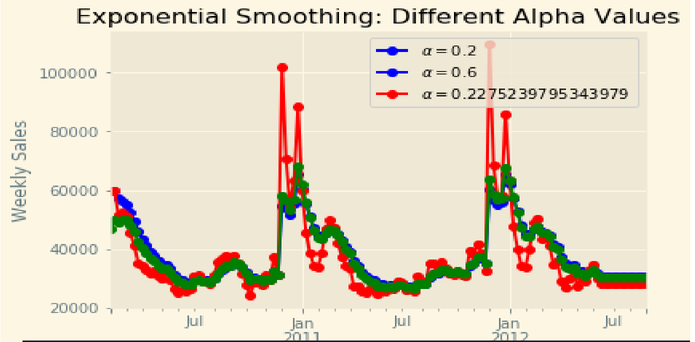


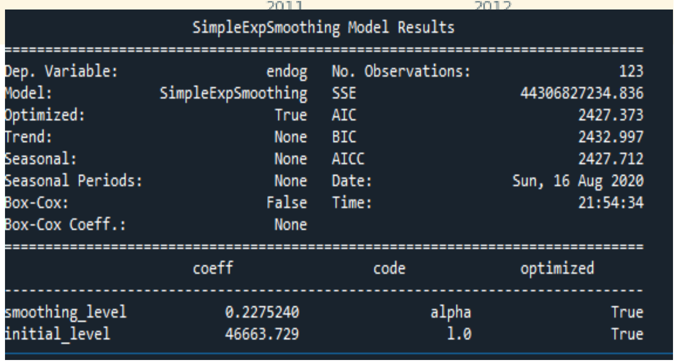
```
Performing stepwise search to minimize aic
 ARIMA(1,1,1)(0,1,1)[12]             : AIC=2504.998, Time=0.41 sec
 ARIMA(0,1,0)(0,1,0)[12]             : AIC=3002.895, Time=0.03 sec
 ARIMA(1,1,0)(1,1,0)[12]             : AIC=2548.562, Time=0.10 sec
 ARIMA(0,1,1)(0,1,1)[12]             : AIC=inf, Time=0.35 sec
 ARIMA(1,1,1)(0,1,0)[12]             : AIC=2546.434, Time=0.19 sec
 ARIMA(1,1,1)(1,1,1)[12]             : AIC=2506.987, Time=0.59 sec
 ARIMA(1,1,1)(0,1,2)[12]             : AIC=2506.984, Time=1.08 sec
 ARIMA(1,1,1)(1,1,0)[12]             : AIC=2524.955, Time=0.39 sec
 ARIMA(1,1,1)(1,1,2)[12]             : AIC=2508.998, Time=1.18 sec
 ARIMA(1,1,0)(0,1,1)[12]             : AIC=2529.287, Time=0.14 sec
 ARIMA(2,1,1)(0,1,1)[12]             : AIC=2505.162, Time=0.38 sec
 ARIMA(1,1,2)(0,1,1)[12]             : AIC=2516.460, Time=0.48 sec
 ARIMA(0,1,0)(0,1,1)[12]             : AIC=2555.258, Time=0.10 sec
 ARIMA(0,1,2)(0,1,1)[12]             : AIC=2516.791, Time=0.15 sec
 ARIMA(2,1,0)(0,1,1)[12]             : AIC=2512.008, Time=0.15 sec
 ARIMA(2,1,2)(0,1,1)[12]             : AIC=2510.034, Time=0.48 sec
 ARIMA(1,1,1)(0,1,1)[12] intercept   : AIC=2506.023, Time=0.64 sec
```





```
Best model:  ARIMA(3,1,0)(0,1,1)[12]
Total fit time: 6.921 seconds
```



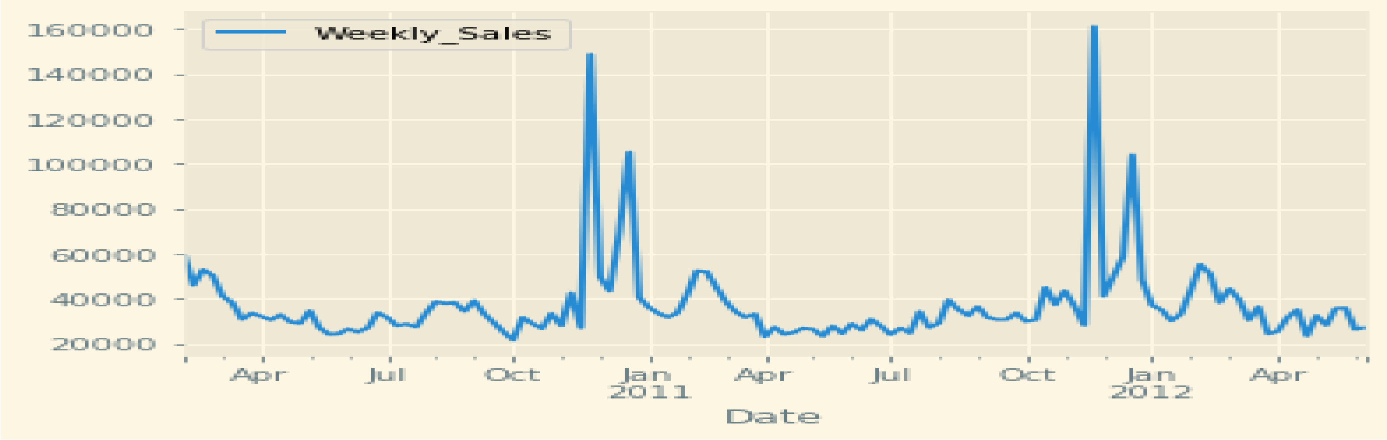


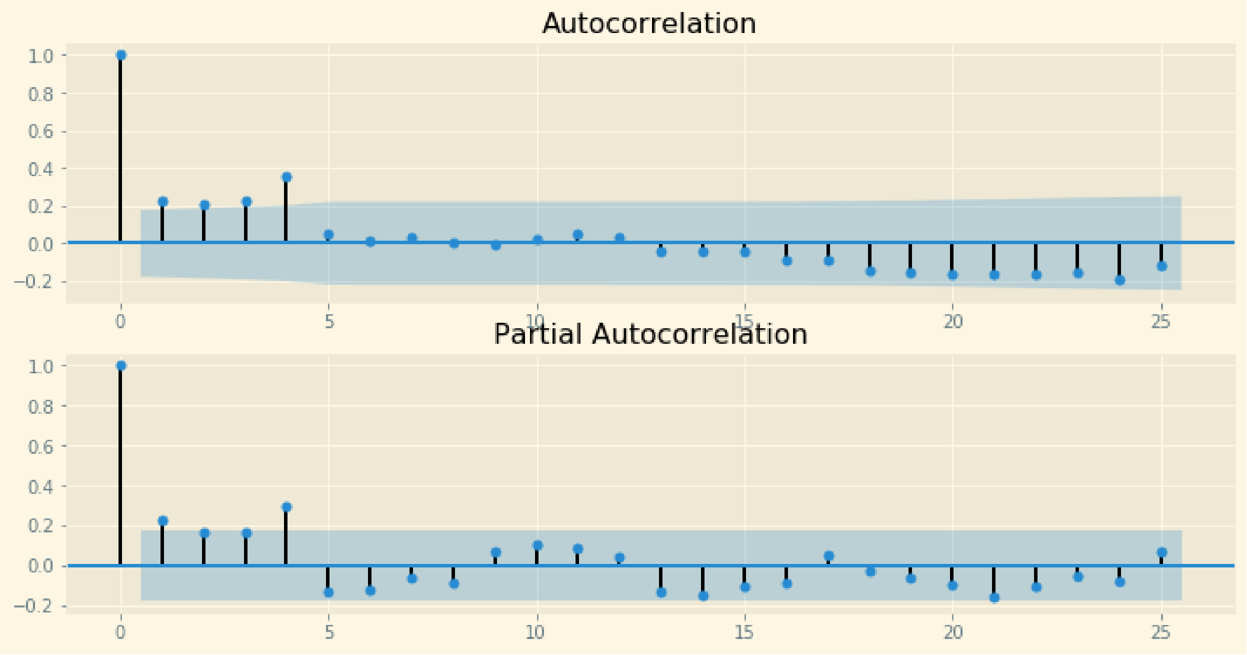
```
                       SimpleExpSmoothing Model Results
==============================================================================
Dep. Variable:                endog   No. Observations:                    123
Model:           SimpleExpSmoothing   SSE                   44306827234.836
Optimized:                     True   AIC                          2427.373
Trend:                         None   BIC                          2432.997
Seasonal:                      None   AICC                         2427.712
Seasonal Periods:              None   Date:                 Sun, 16 Aug 2020
Box-Cox:                      False   Time:                         21:54:34
Box-Cox Coeff.:                None
==============================================================================
                       coeff                 code              optimized
------------------------------------------------------------------------------
smoothing_level            0.2275240                alpha                 True
initial_level              46663.729                  1.0                 True
------------------------------------------------------------------------------
```

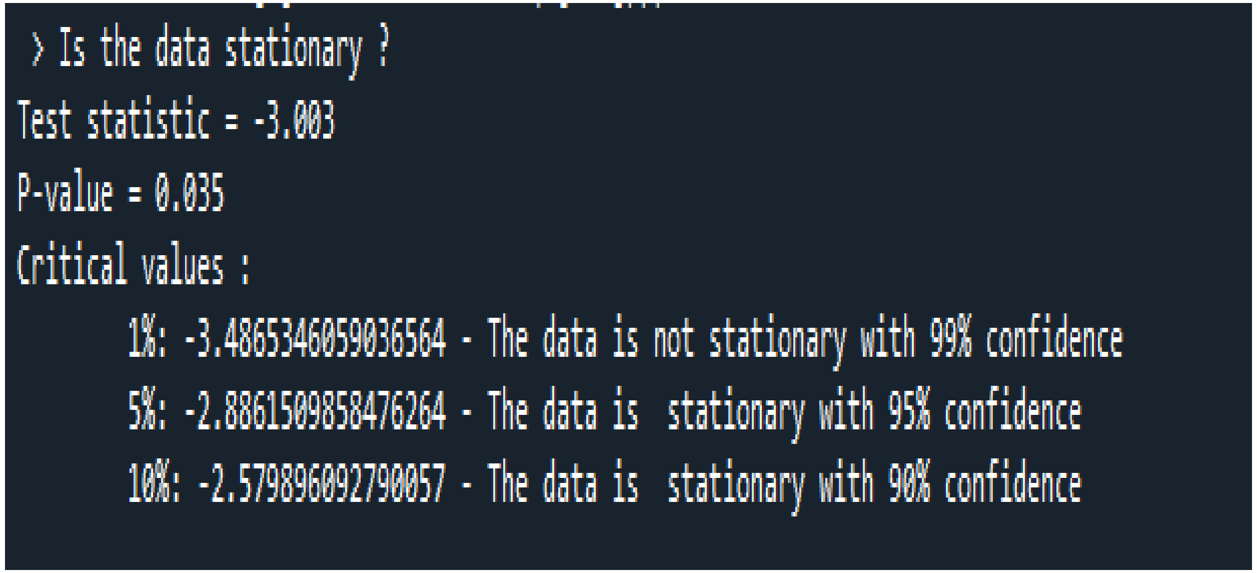
- **Seasonal Data; High sales during Holidays**
- **High sales end of year during December**
- **Best fit is Simple Exponential Smoothing**
- **Independent Var: IsHoliday**
- **AIC: 2427**

Appendix - Electronics Department Sales

STORE 34



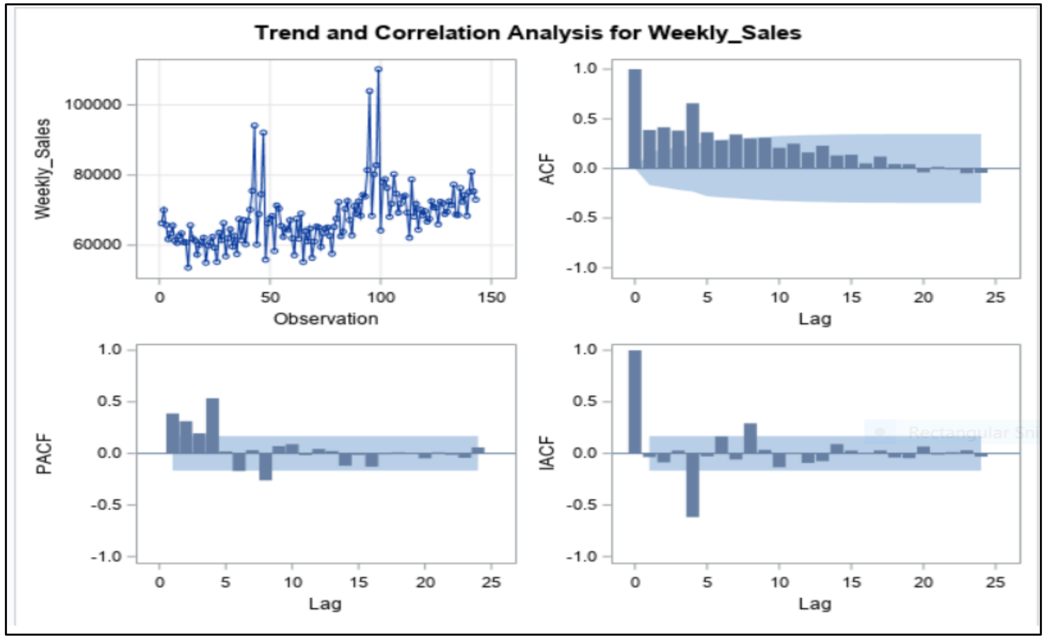


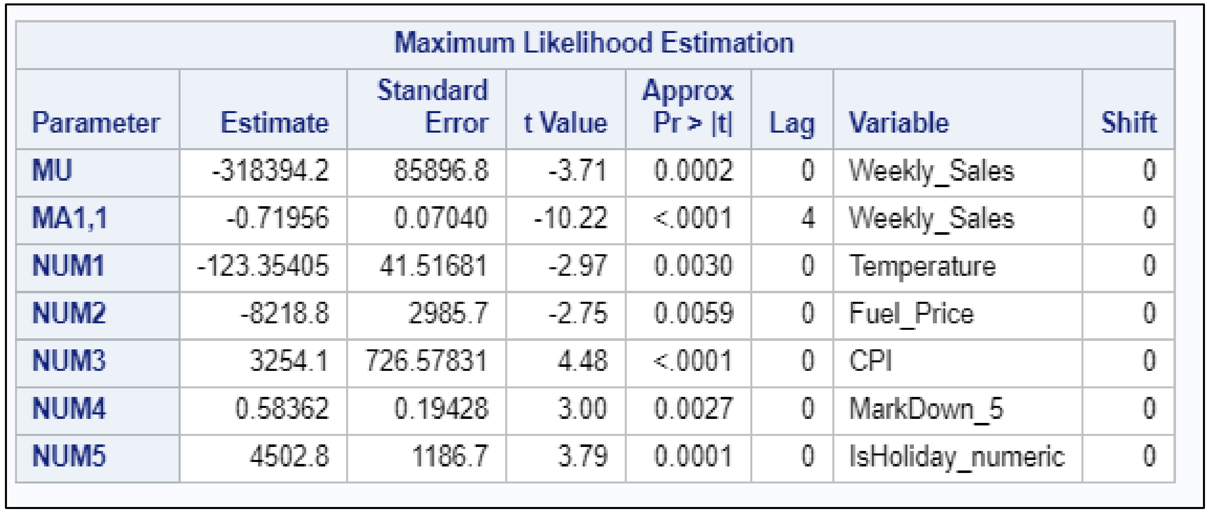


```
 > Is the data stationary ?
Test statistic = -3.003
P-value = 0.035
Critical values :
        1%: -3.4865346059036564 - The data is not stationary with 99% confidence
        5%: -2.8861509858476264 - The data is  stationary with 95% confidence
        10%: -2.579896092790057 - The data is  stationary with 90% confidence
```

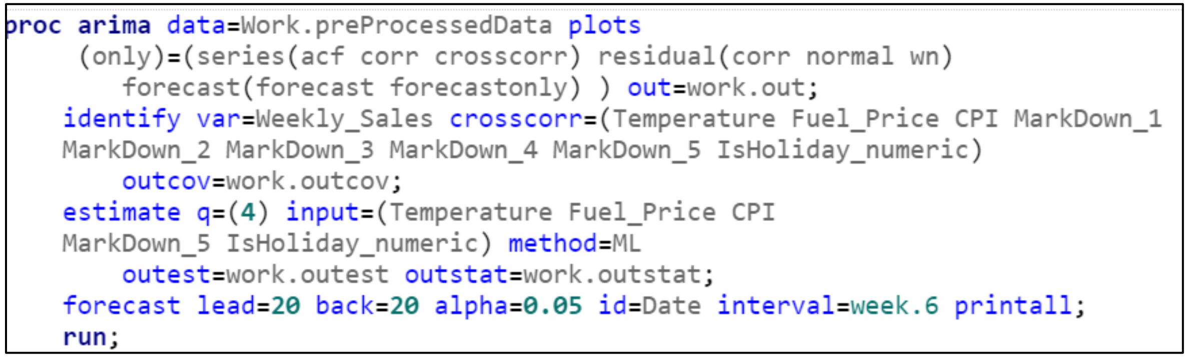
Appendix - Groceries Department Sales

STORE 34


Trend and Correlation Analysis for Weekly_Sales



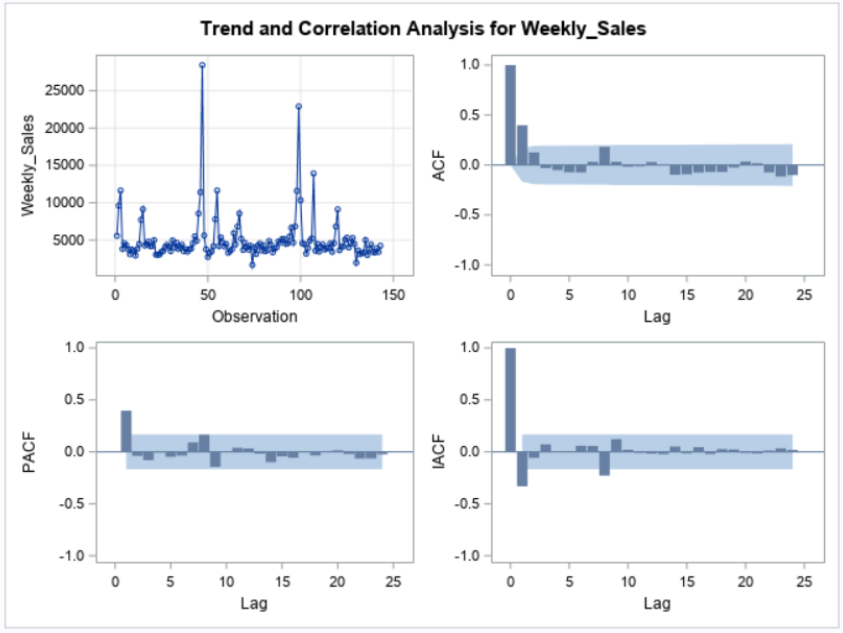
Maximum Likelihood Estimation							
Parameter	Estimate	Standard Error	t Value	Approx Pr > |t|	Lag	Variable	Shift
MU	-318394.2	85896.8	-3.71	0.0002	0	Weekly_Sales	0
MA1,1	-0.71956	0.07040	-10.22	<.0001	4	Weekly_Sales	0
NUM1	-123.35405	41.51681	-2.97	0.0030	0	Temperature	0
NUM2	-8218.8	2985.7	-2.75	0.0059	0	Fuel_Price	0
NUM3	3254.1	726.57831	4.48	<.0001	0	CPI	0
NUM4	0.58362	0.19428	3.00	0.0027	0	MarkDown_5	0
NUM5	4502.8	1186.7	3.79	0.0001	0	IsHoliday_numeric	0

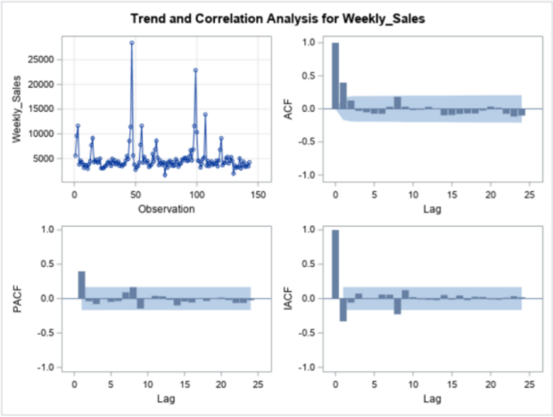


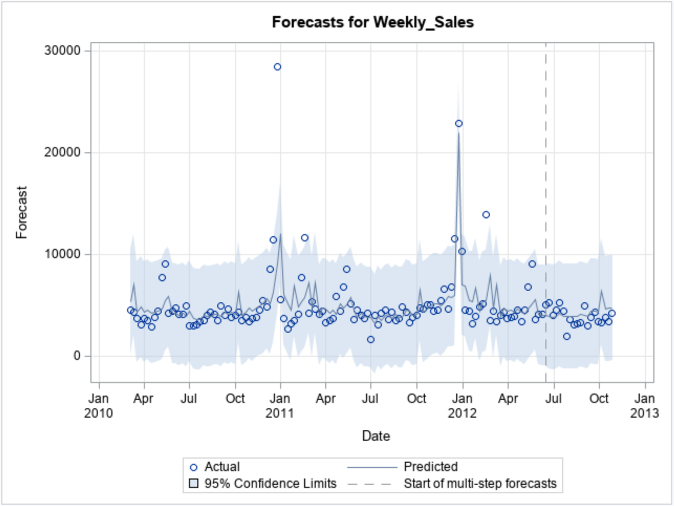
```
proc arima data=Work.preProcessedData plots
     (only)=(series(acf corr crosscorr) residual(corr normal wn)
        forecast(forecast forecastonly) ) out=work.out;
    identify var=Weekly_Sales crosscorr=(Temperature Fuel_Price CPI MarkDown_1
    MarkDown_2 MarkDown_3 MarkDown_4 MarkDown_5 IsHoliday_numeric)
        outcov=work.outcov;
    estimate q=(4) input=(Temperature Fuel_Price CPI
    MarkDown_5 IsHoliday_numeric) method=ML
        outest=work.outest outstat=work.outstat;
    forecast lead=20 back=20 alpha=0.05 id=Date interval=week.6 printall;
    run;
```

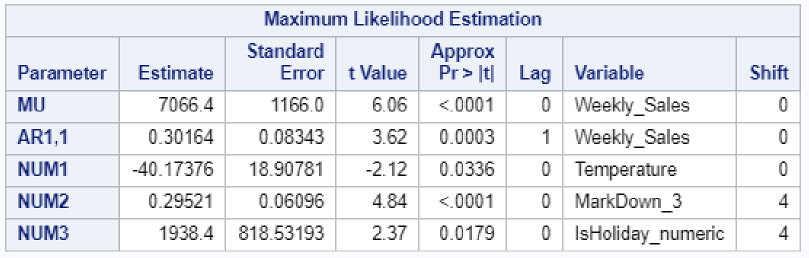
Appendix: Jewelry Department Sales

STORE 34



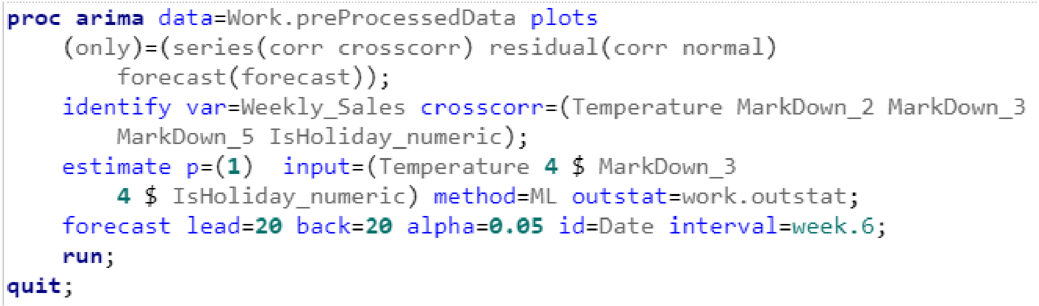






Maximum Likelihood Estimation

Parameter	Estimate	Standard Error	t Value	Approx Pr > |t|	Lag	Variable	Shift
MU	7066.4	1166.0	6.06	<.0001	0	Weekly_Sales	0
AR1,1	0.30164	0.08343	3.62	0.0003	1	Weekly_Sales	0
NUM1	-40.17376	18.90781	-2.12	0.0336	0	Temperature	0
NUM2	0.29521	0.06096	4.84	<.0001	0	MarkDown_3	4
NUM3	1938.4	818.53193	2.37	0.0179	0	IsHoliday_numeric	4



```
proc arima data=Work.preProcessedData plots
    (only)=(series(corr crosscorr) residual(corr normal)
        forecast(forecast));
    identify var=Weekly_Sales crosscorr=(Temperature MarkDown_2 MarkDown_3
        MarkDown_5 IsHoliday_numeric);
    estimate p=(1)  input=(Temperature 4 $ MarkDown_3
        4 $ IsHoliday_numeric) method=ML outstat=work.outstat;
    forecast lead=20 back=20 alpha=0.05 id=Date interval=week.6;
    run;
quit;
```