

Walmart Store Sales

Business Forecasting
Group 9

Walmart
Save money. Live better.



Vaibhavi Mukadam, Aditya Pai, Cameron Clark

Data Description

Walmart is a renowned retail corporation that operates a chain of hypermarkets.

Walmart has provided a data combining of 45 stores including store information and weekly sales of year 2010, 2011 and 2012.

```
> summary(df)
```

Store	Dept	Date	Weekly_Sales	IsHoliday
Min. : 1.0	Min. : 1.00	Min. : 2010-02-05	Min. : -4989	Mode :logical
1st Qu.: 11.0	1st Qu.: 18.00	1st Qu.: 2010-10-08	1st Qu.: 2080	FALSE:391909
Median : 22.0	Median : 37.00	Median : 2011-06-17	Median : 7612	TRUE :29661
Mean : 22.2	Mean : 44.26	Mean : 2011-06-18	Mean : 15981	
3rd Qu.: 33.0	3rd Qu.: 74.00	3rd Qu.: 2012-02-24	3rd Qu.: 20206	
Max. : 45.0	Max. : 99.00	Max. : 2012-10-26	Max. : 693099	

We aggregated the data to get consistent interval of time within the dataset.

```
> summary(df_date)
```

Date	Weekly_Sales	Weekly_Sales_norm
Min. : 2010-02-05	Min. : 39599853	Min. : 0.0000
1st Qu.: 2010-10-11	1st Qu.: 44880588	1st Qu.: 0.1278
Median : 2011-06-17	Median : 46243900	Median : 0.1607
Mean : 2011-06-17	Mean : 47113419	Mean : 0.1818
3rd Qu.: 2012-02-20	3rd Qu.: 47792025	3rd Qu.: 0.1982
Max. : 2012-10-26	Max. : 80931416	Max. : 1.0000

Holidays

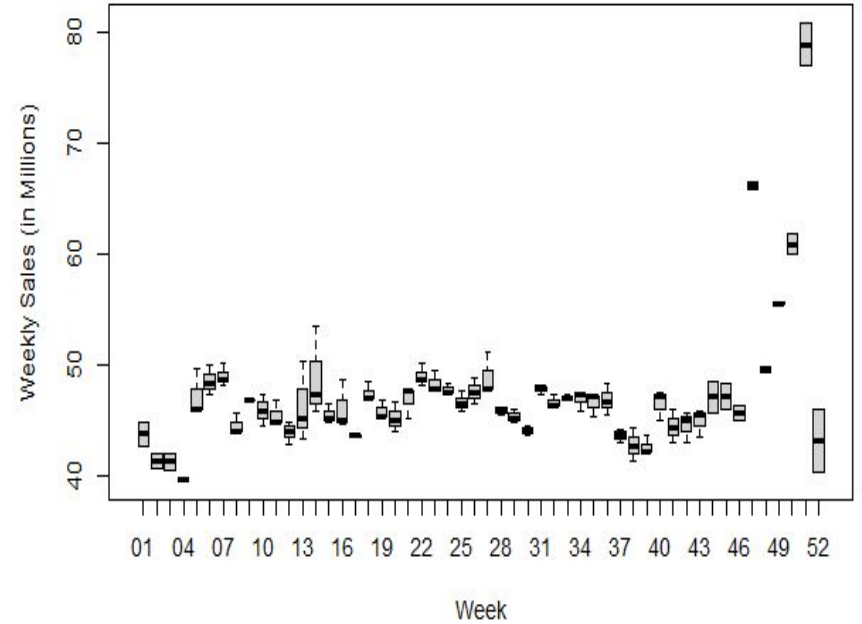
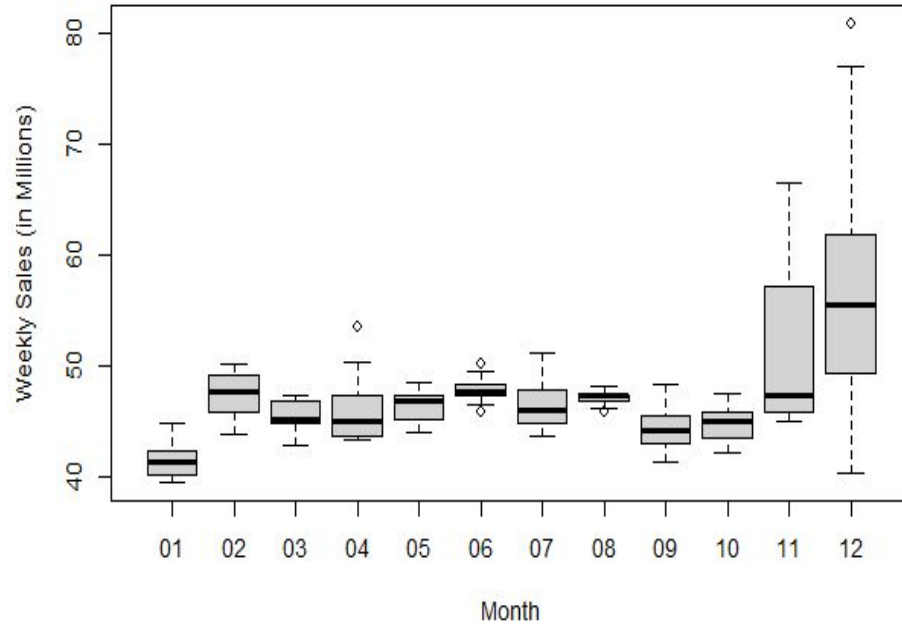
Super Bowl: ~ 2nd Week of February

Labor Day: ~ 2nd Week of September

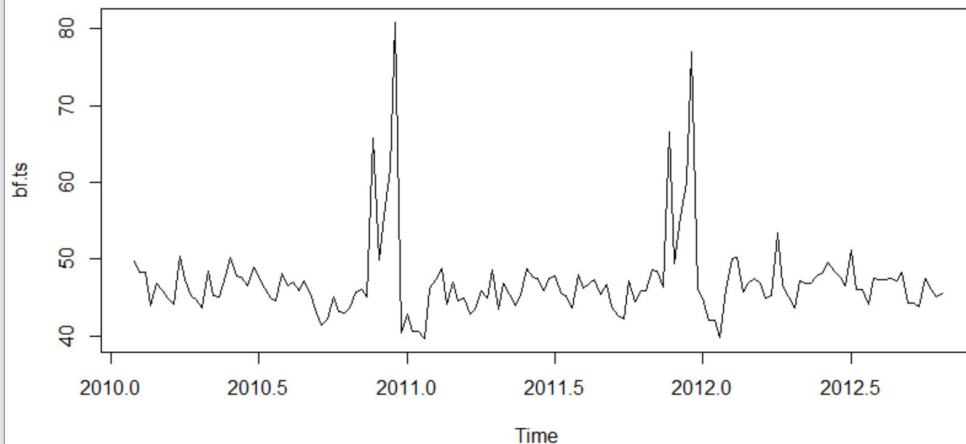
Thanksgiving/Black Friday: ~4th week of November

Christmas: ~ 4th Week of December

Data in Millions Aggregated



Time Series -

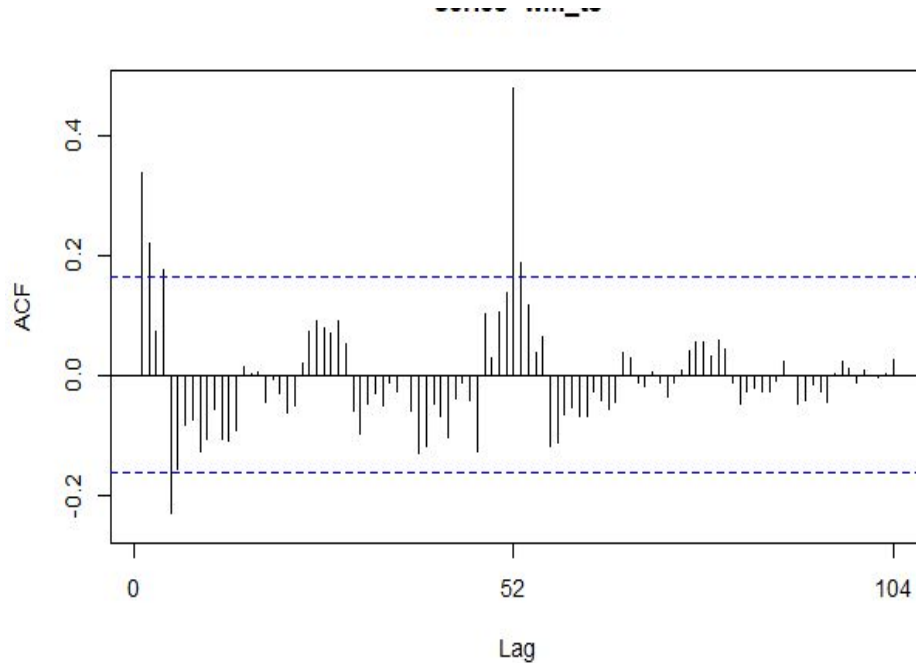


- Weekly sales are in Millions
- There is a spike late in the year
- The data shows seasonality at the end of each year
- There is dip in the initial months of the year

```
> summary(bf.ts)
```

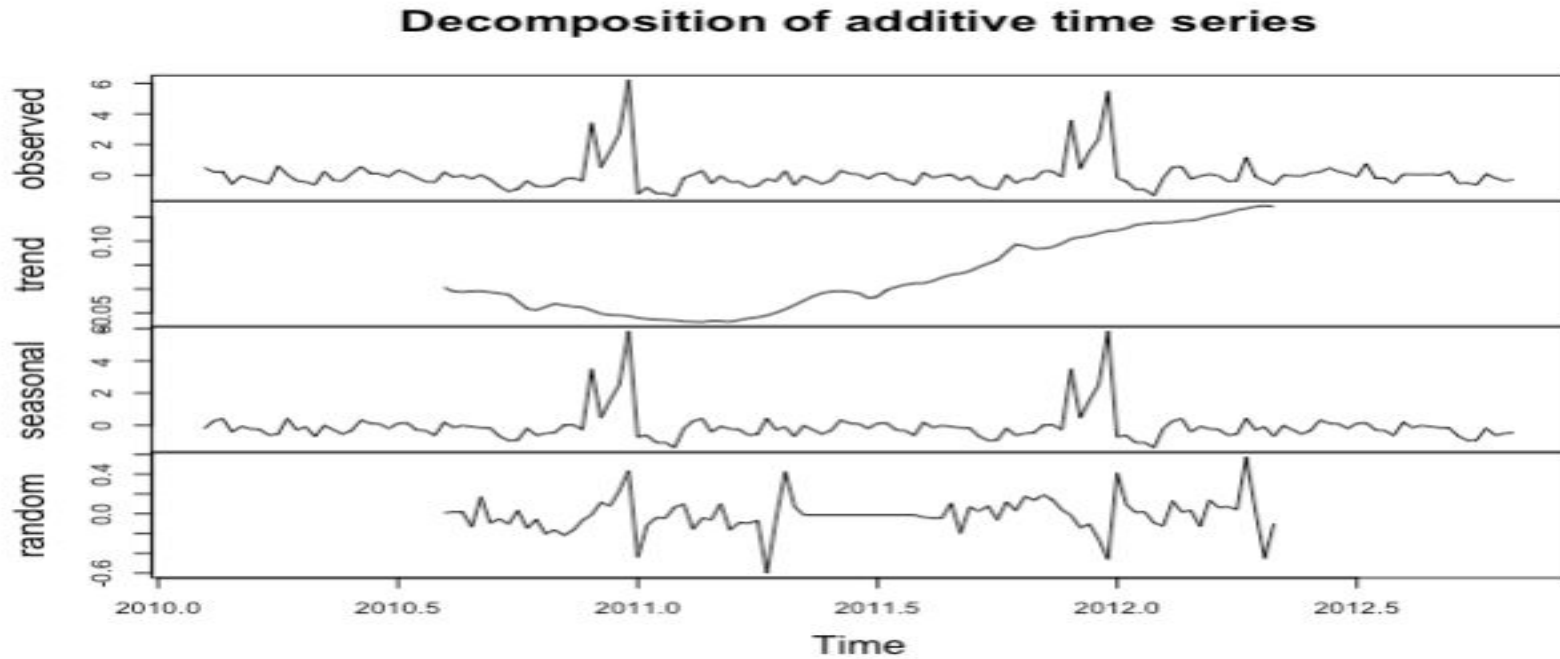
Min.	1st Qu.	Median	Mean	3rd Qu.	Max.
39.60	44.88	46.24	47.11	47.79	80.93

Autocorrelation



Autocorrelation represents the degree of similarity between a given time series and a lagged version of itself over successive time intervals.

Decomposition



Accuracy Measure: MAPE

$$\text{MAPE} = \frac{100\%}{n} \sum_{t=1}^n \left| \frac{A_t - F_t}{A_t} \right|$$

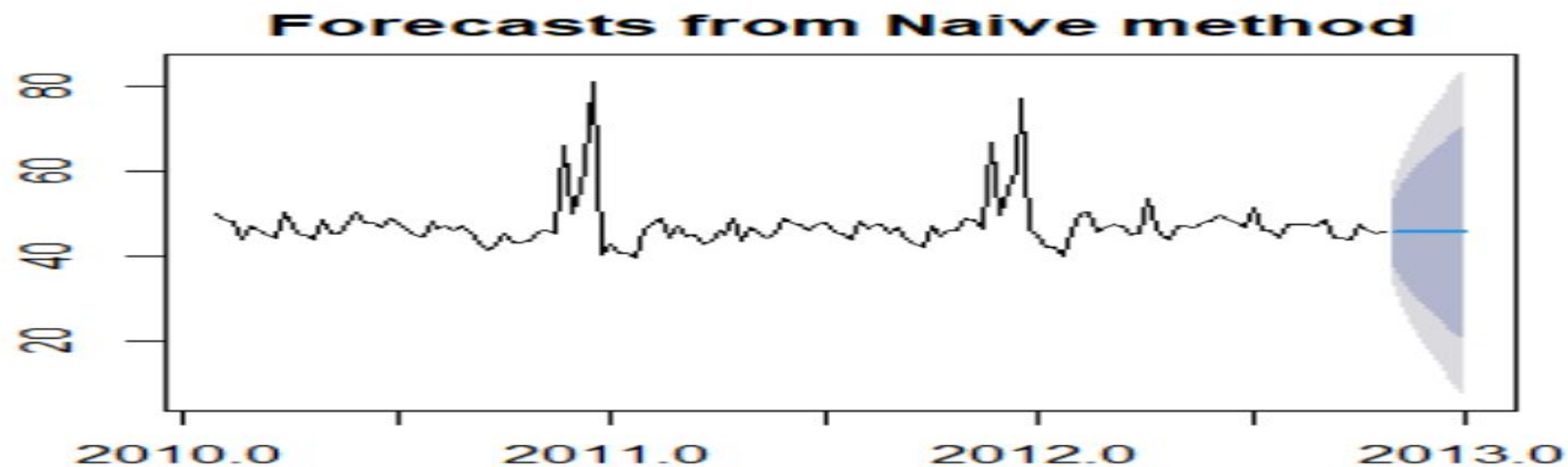
Mean absolute percentage error :

Actual data is non zero, that is, the aggregated data which we are analysing does not have zero as a value in Weekly Sales.

MAPE puts a heavier penalty on negative errors than on positive errors.

We cannot use MAE and RMSE as it depends upon the scale and also it is difficult to make comparison for a different time interval.

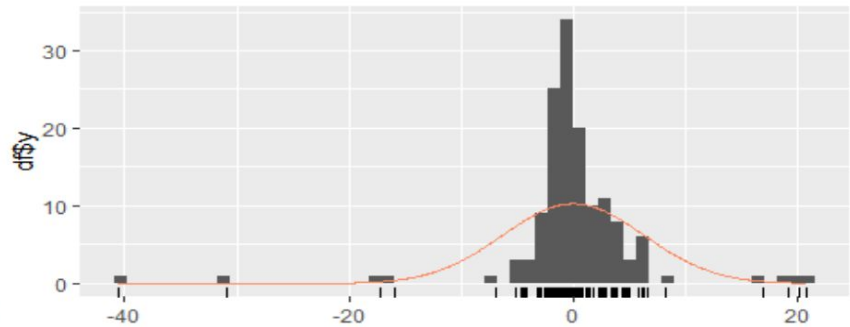
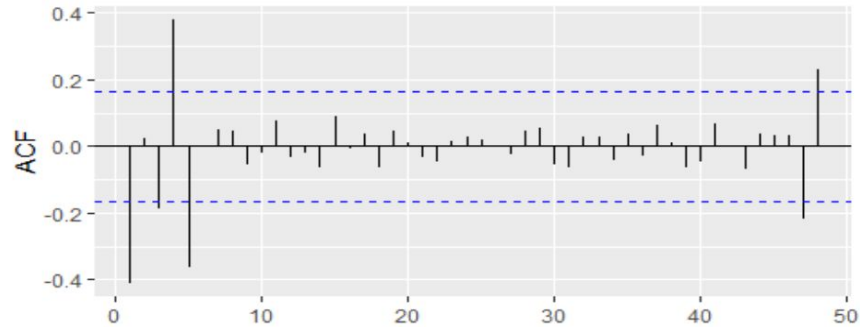
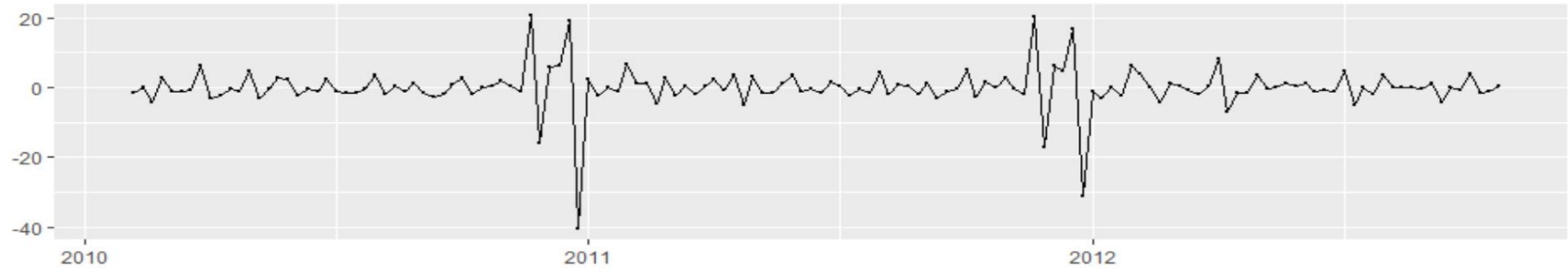
Naive Forecast



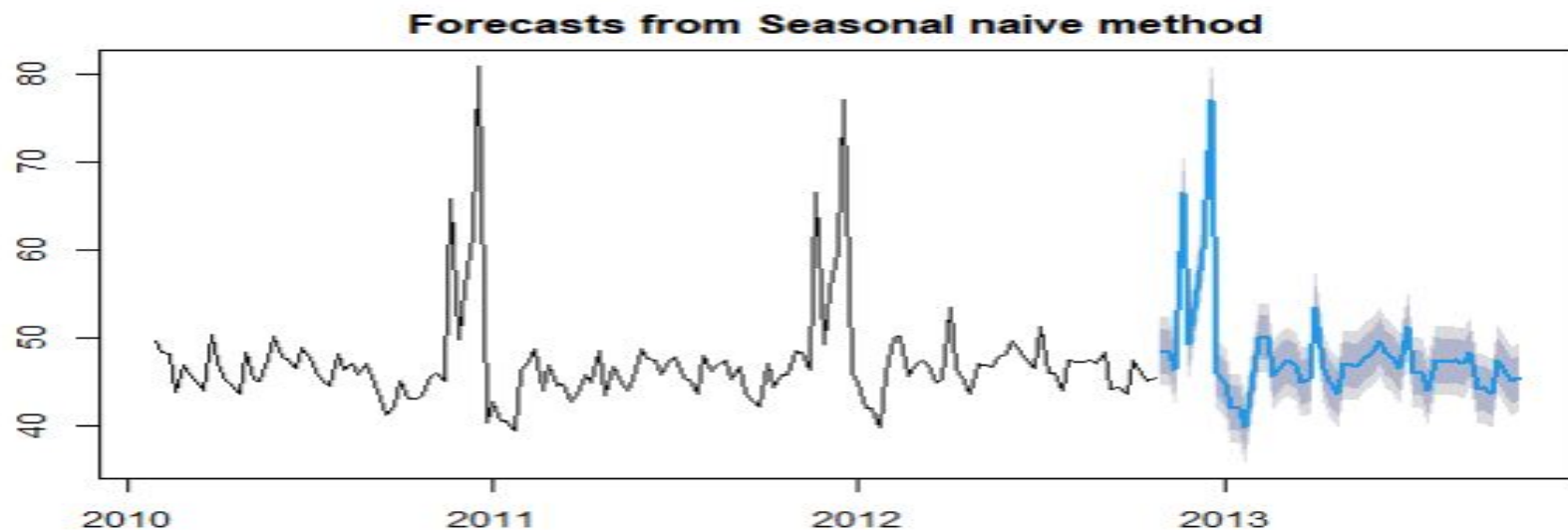
ME	RMSE	MAE
-0.02962411	6.26193	3.144517
MPE	MAPE	MASE
-0.7290743	6.365846	2.194441
ACF1		
-0.4120562		

Naive : Residual analysis

Residuals from Naive method



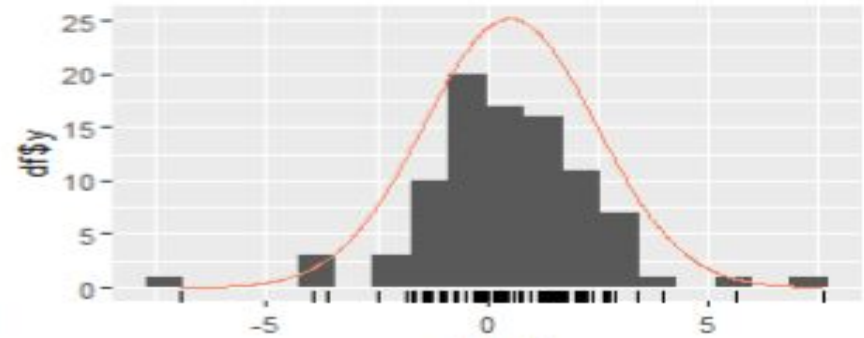
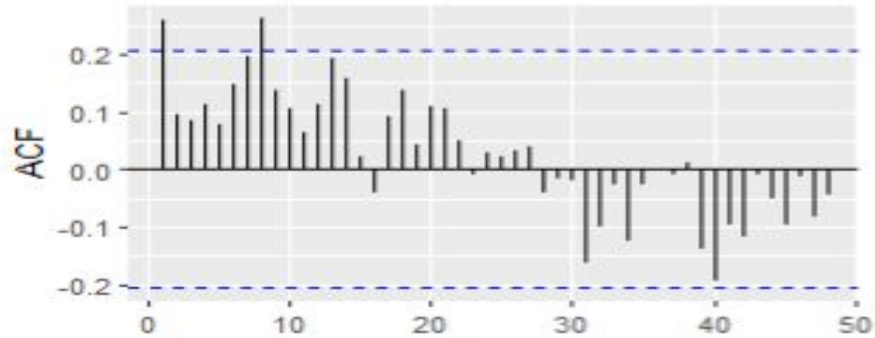
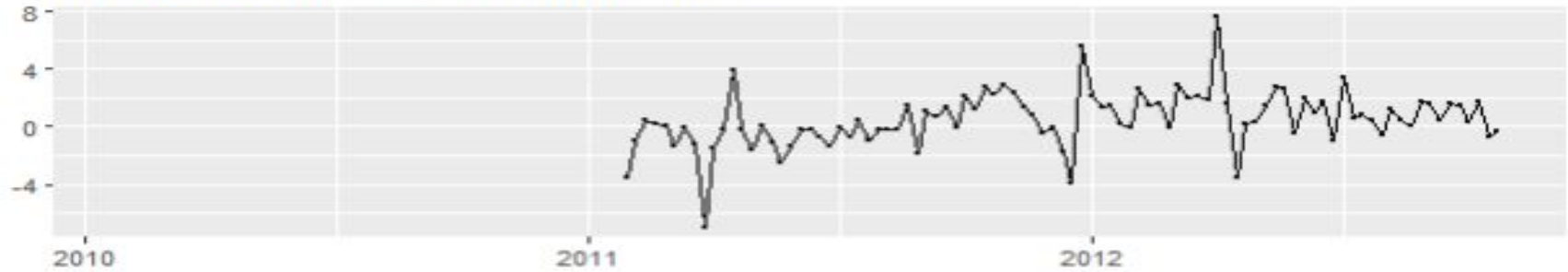
Seasonal Naive Forecast



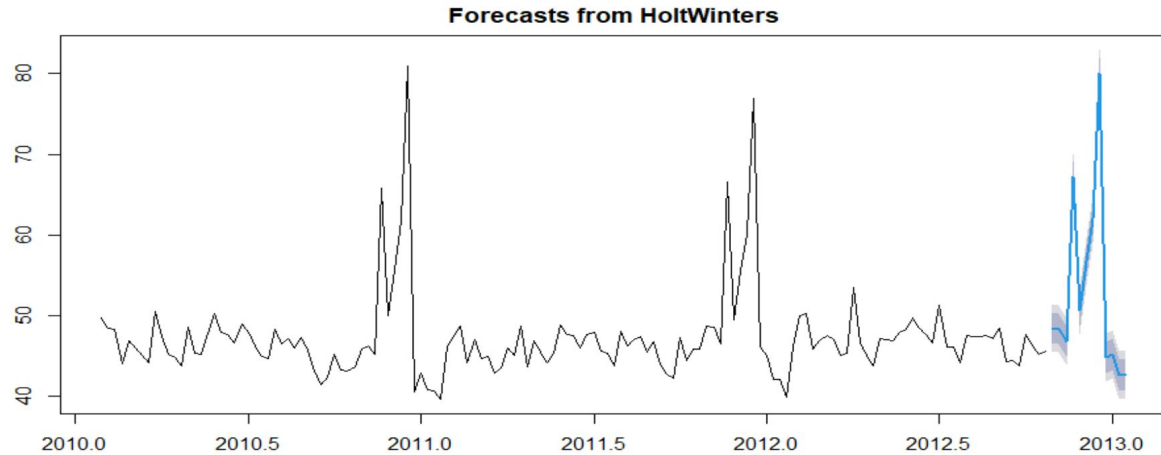
	ME	RMSE	MAE	MPE	MAPE	MASE	ACF1
Training set	0.5031124	1.998564	1.432947	1.067447	3.017517	1	0.2586785

Snaive: Residual Analysis

Residuals from Seasonal naive method



Holts Winter Forecasting Model

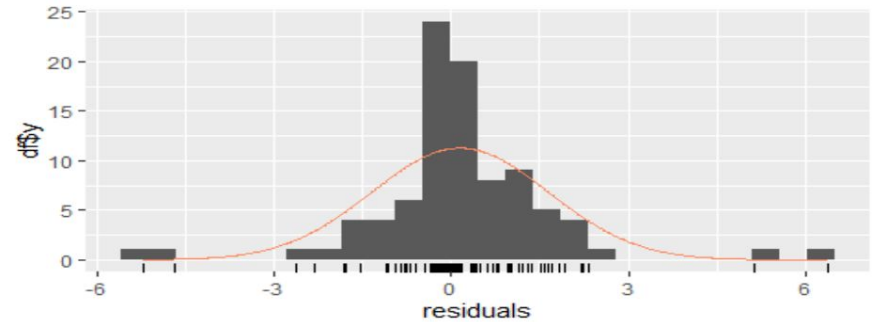
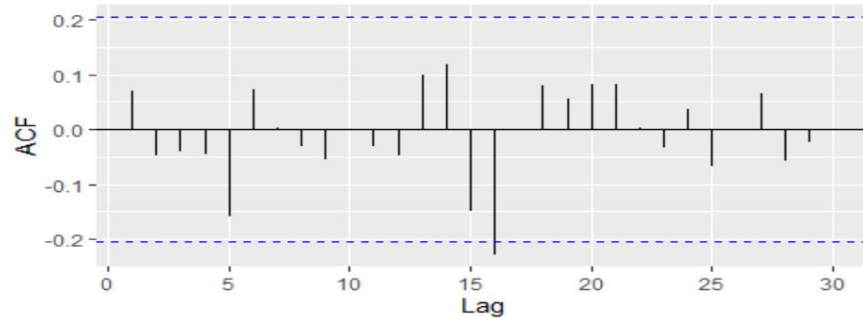
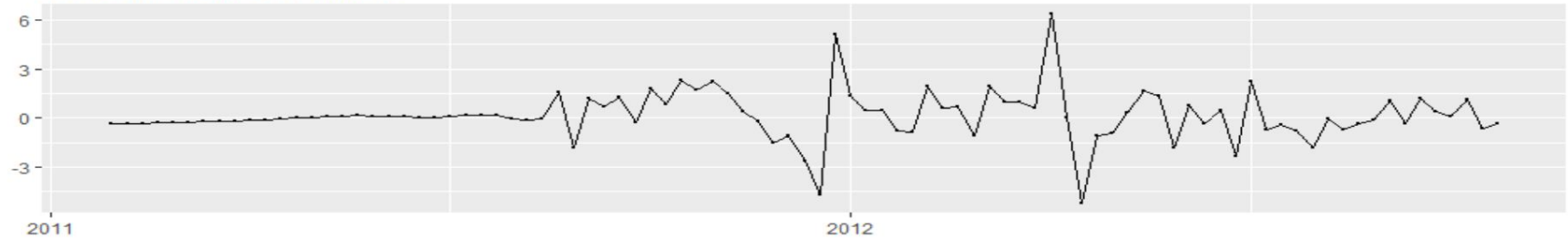


Smoothing parameters:
alpha: 0.06058175
beta : 0.008479722
gamma: 0.557241

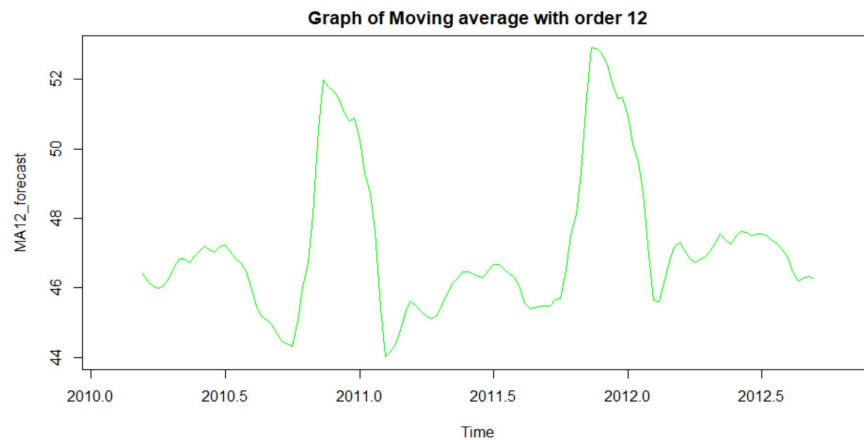
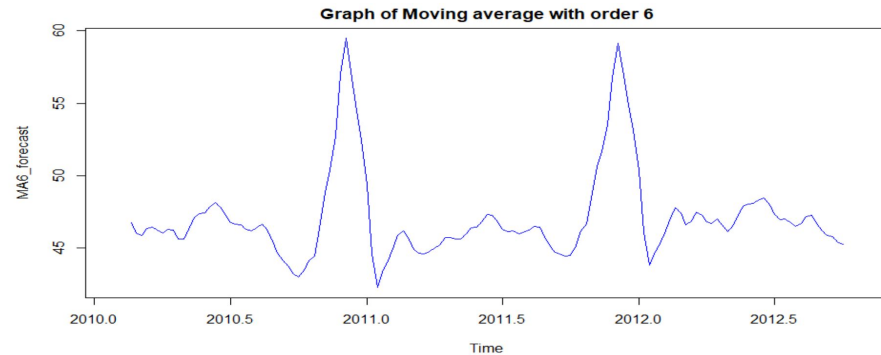
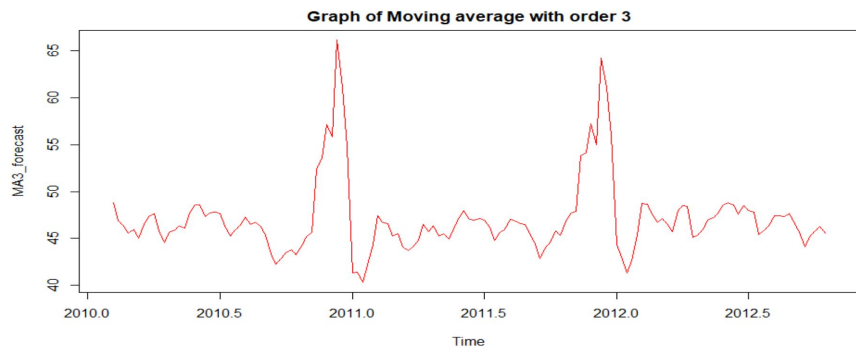
Training set	ME	RMSE	MAE
	0.1509947	1.494934	0.9357783
Training set	MPE	MAPE	MASE
	0.3562118	1.932219	0.6530448
Training set	ACF1		
	0.06952215		

Holts Winter: Residual Analysis

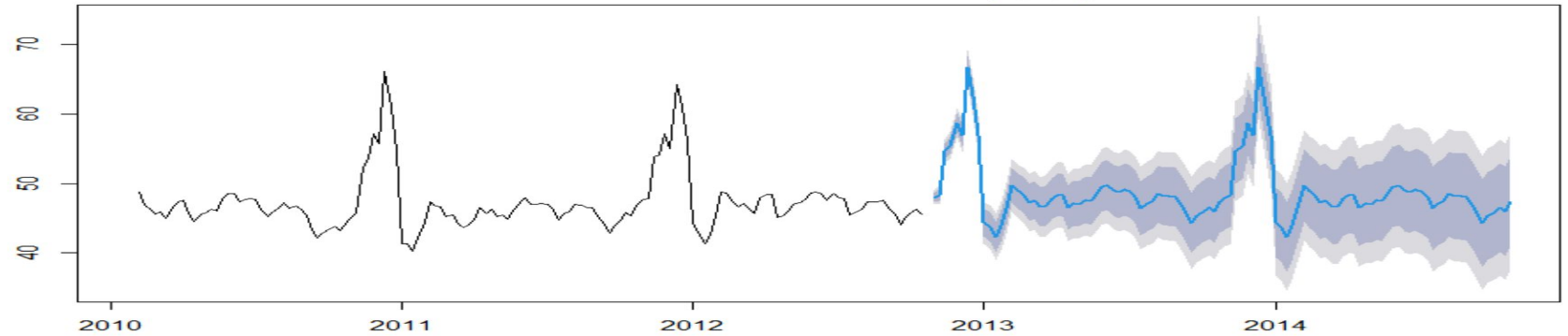
Residuals from HoltWinters



Moving Averages Forecast

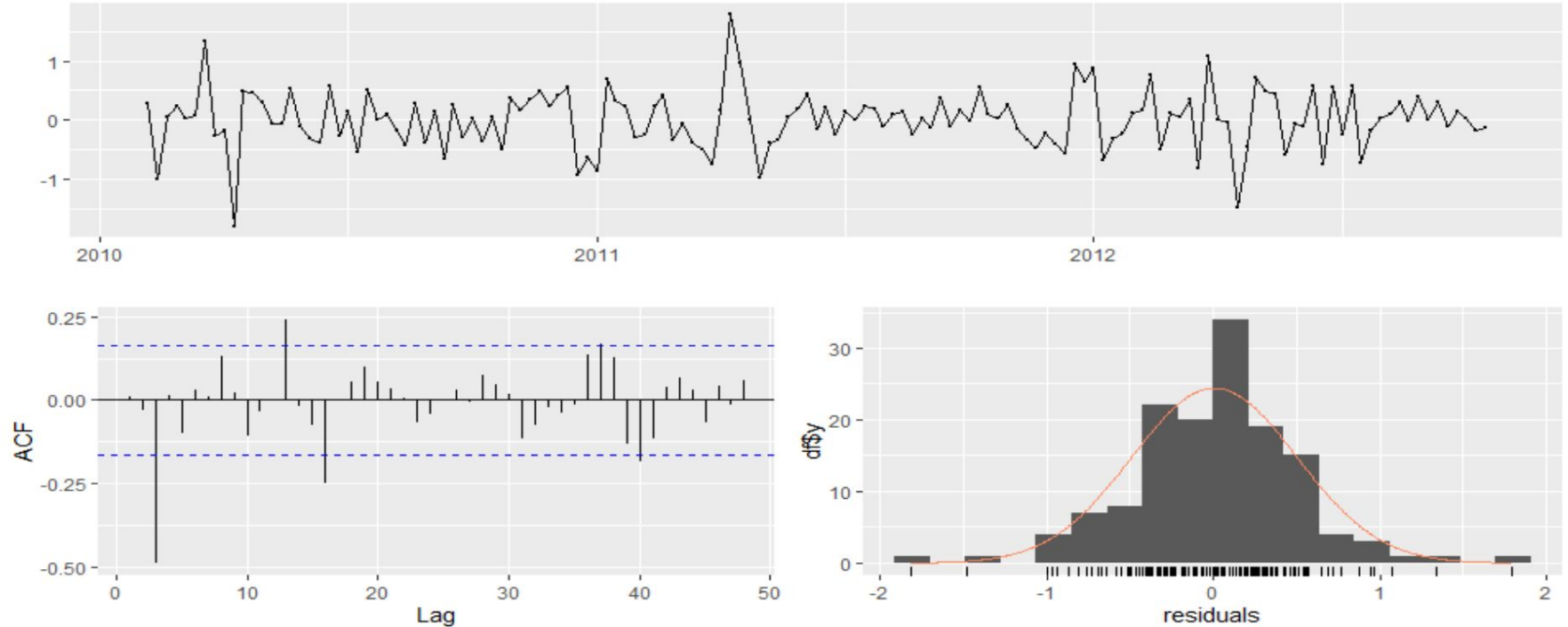


Moving average order = 3

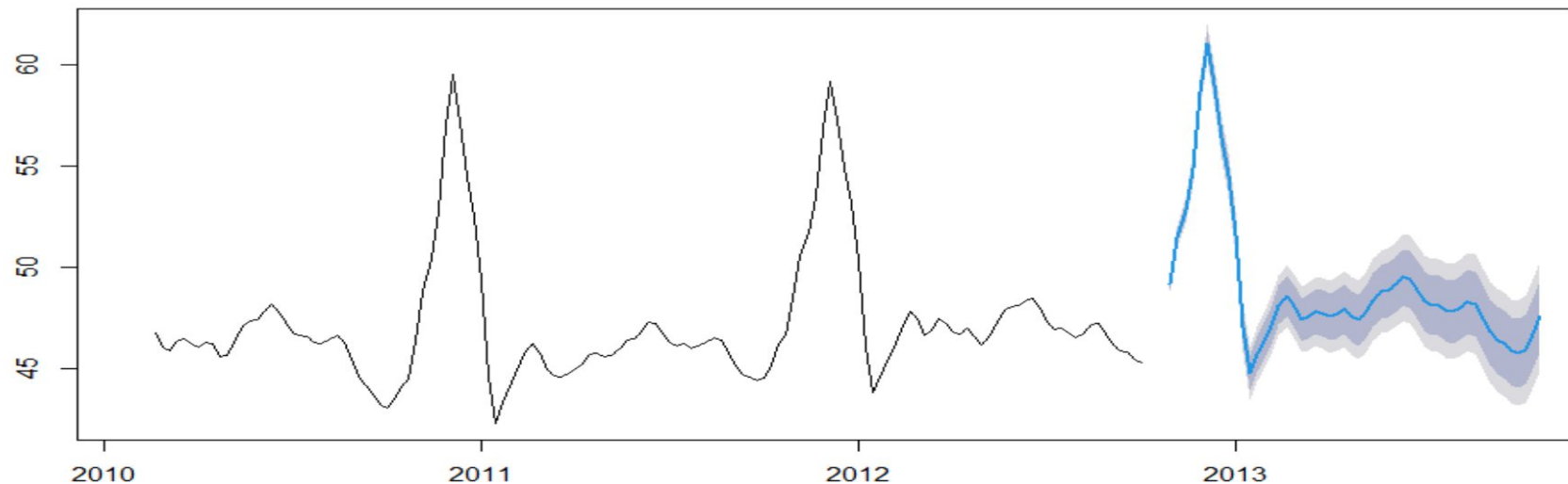


	ME	RMSE	MAE
Training set	0.006005347	0.4930776	0.3657756
	MPE	MAPE	MASE
Training set	0.00676566	0.775721	0.3180166
	ACF1		
Training set	0.008710816		

Moving Average: Residual Analysis

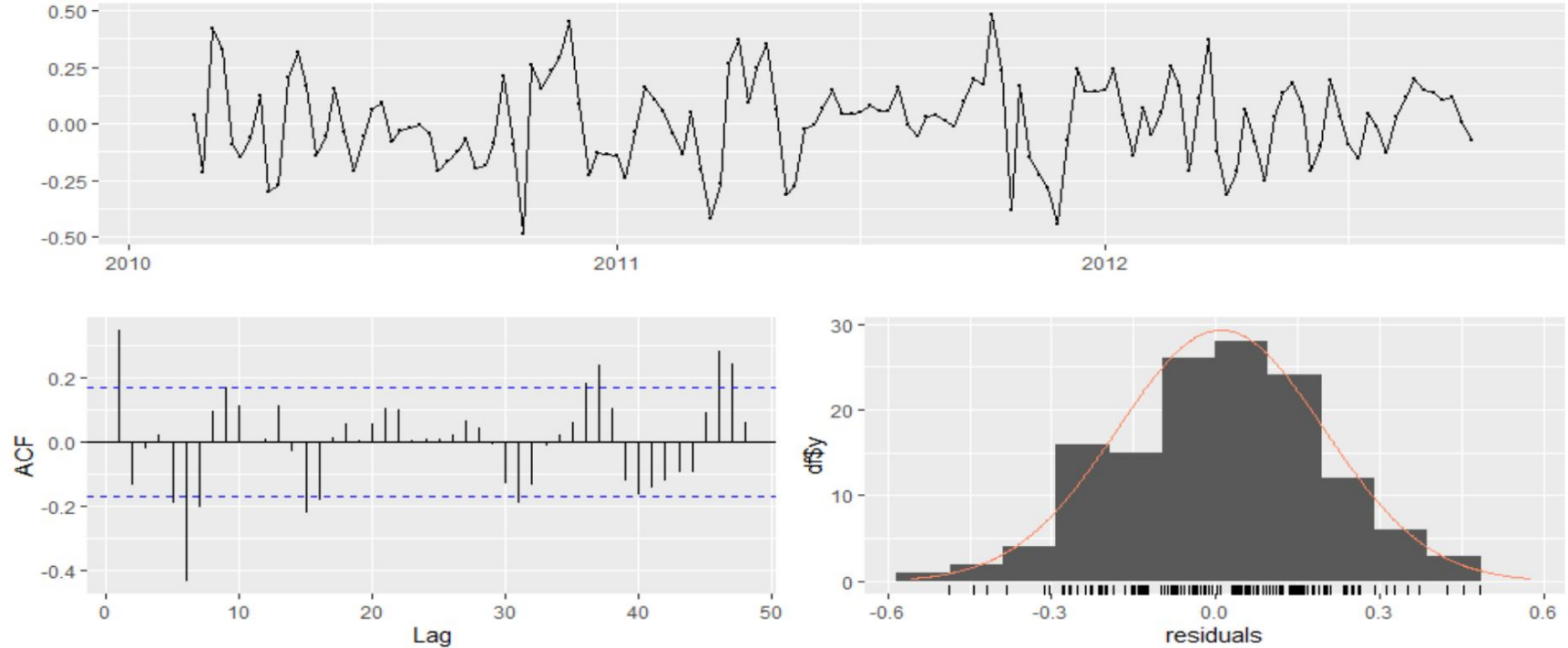


Moving average order = 6



	ME	RMSE	MAE
Training set	0.0096508	0.1885963	0.1522132
	MPE	MAPE	MASE
Training set	0.01988877	0.3223557	0.150448
	ACF1		
Training set	0.3467995		

Moving Average: Residual Analysis



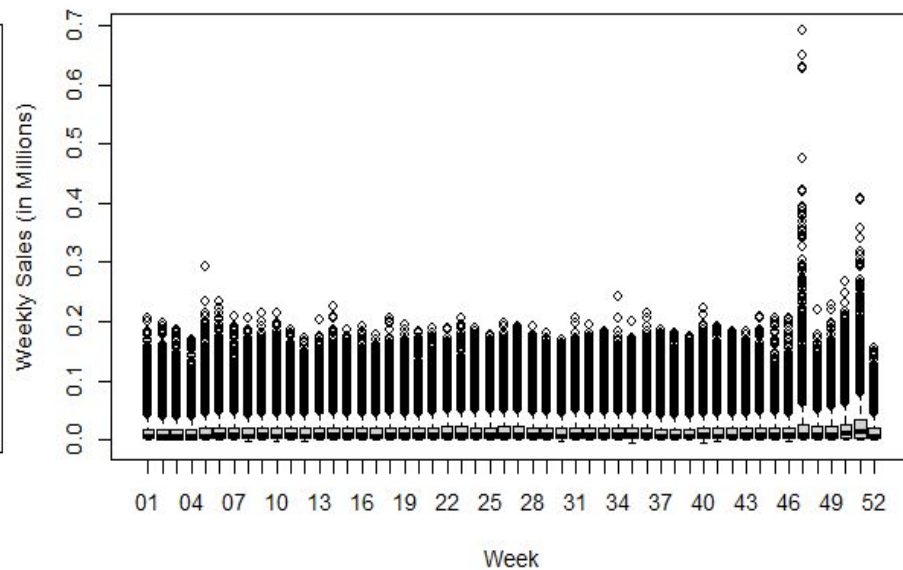
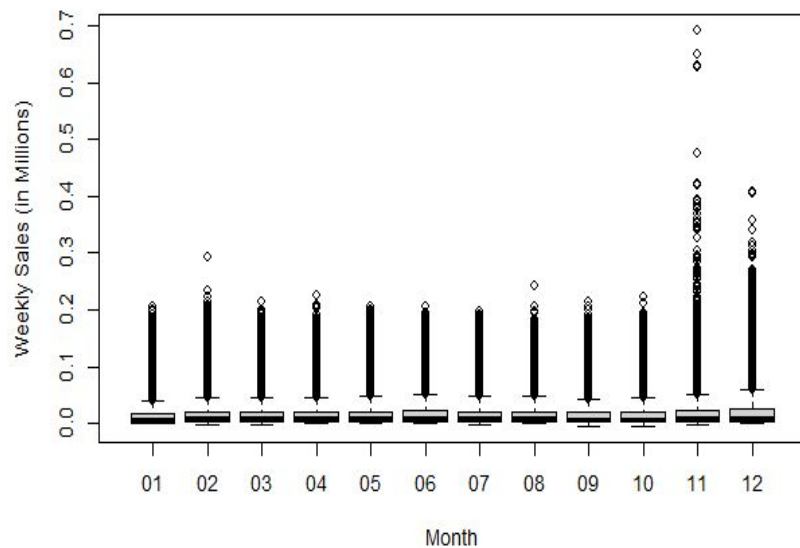
Accuracy Measures

Forecasting Models	Accuracy (MAPE)	ACF
Naive Model	6.365	- 0.412
Seasonal Naive Model	3.0175	0.258
Moving Average (O=3)	0.775721	0.008710
Moving Average (O=6)	0.3223557	0.3467995
Holts Winter Forecasting	1.9322	0.069522
STL + ETS	1.6242	0.0067
Arima Model	1.72022	0.0067

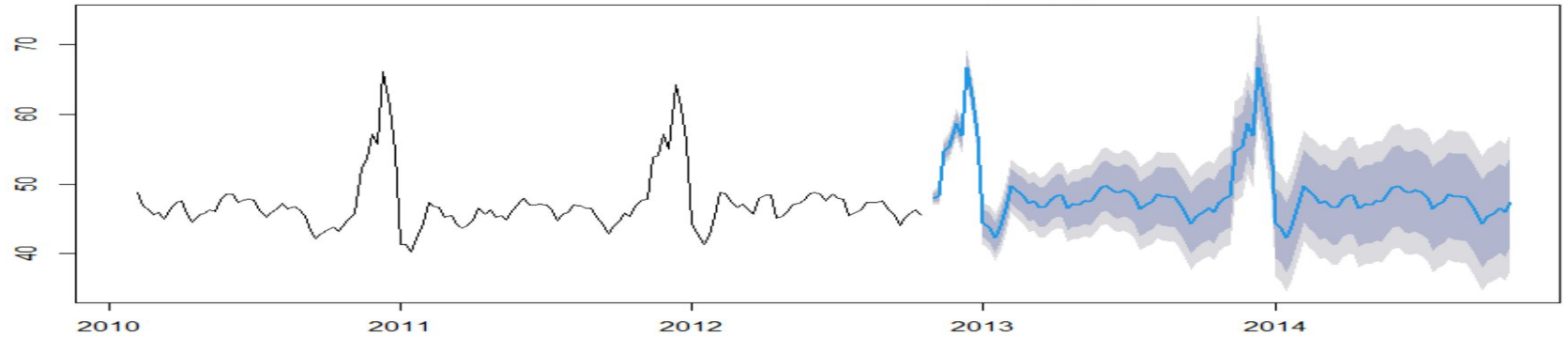
Insights

- Walmart may need to hire more people to work during high periods of sales
- Walmart may want to capitalise on the holidays other than end of the year.
Have more effective sales during February and September other than Thanksgiving and Christmas.
- Using the model consistently, it can hire part time worker if they see growth in the business 1 month prior.

Data in Millions



Moving average order = 3



	ME	RMSE	MAE
Training set	0.006005347	0.4930776	0.3657756
	MPE	MAPE	MASE
Training set	0.00676566	0.775721	0.3180166
	ACF1		
Training set	0.008710816		

Patterns in the data

1. Walmart week pattern for the year
2. First week of every month except november and december have more sales
3. Customers show pattern in buying, there is more sales on weekends than weekdays, first week of every month has more sales, there are few months which can be capitalised majorly using these patterns.

February and August

- In the month of february we can see a spike in the graph, notable events are 14th february Valentine's day and Super bowl in the same week, here we have two weeks to target, the first week, and mainly the second week. Departments such as beauty, clothing, jewellery, shoes, bag and accessories can be very lucrative and should be targeted.
- August has an unusual spike, as there are no holidays or notable events around that month. After some research we found that around the first-second week of august most of schools and colleges reopen. Departments such as stationery, clothing, electronics can be targeted.

October end - December end

These months are very unpredictable, it cannot be said surely on how high the sales for these months might go. These are the most important months of the year, and so the main things have to be seriously taken care of. Firstly, the work force has to be increased according to the needs. The inventory has to be closely monitored as products go out of stock quickly in these months. Starting from halloween in october end, thanksgiving in third week of november, black friday in the same week. Christmas and New year in the end of december to complete the holiday season. All departments should be equally monitored.

Targeted Departments

1. The products in these departments should be displayed more, as people have spending patterns, the more the products on display the more will be the sales.
2. The staff should be trained for specific work in these departments, as these departments will increase the sales.
3. Rollback offers that Walmart offers and claims to be lowest in the market can be put on products in the untargeted departments, so that even those departments make more sales.
4. Stores located in colder areas have better chance of selling winter goods, so they should be targeted in the months of October, November and December.

Thank you

Questions?