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

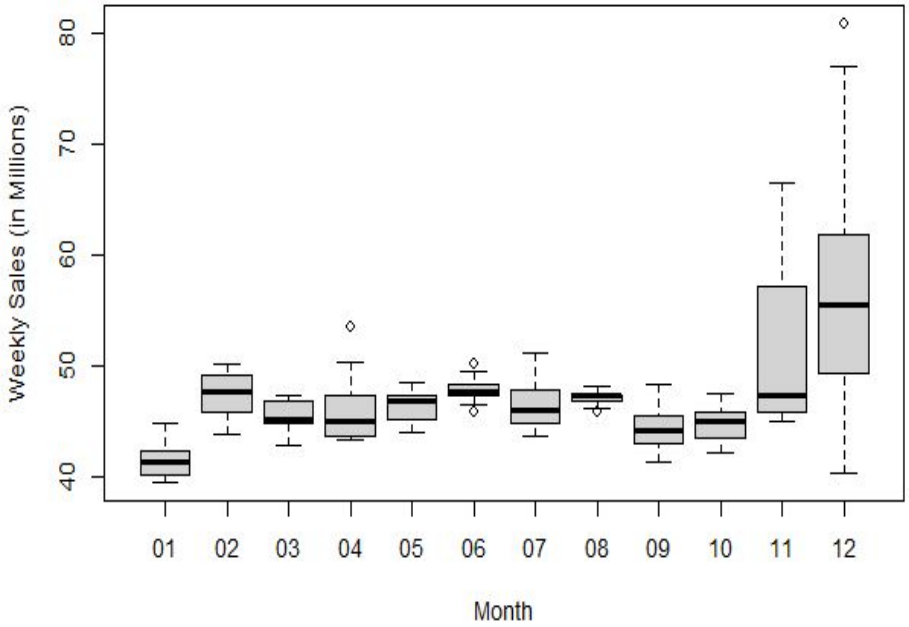
Super Bowl: ~ 2nd Week of February

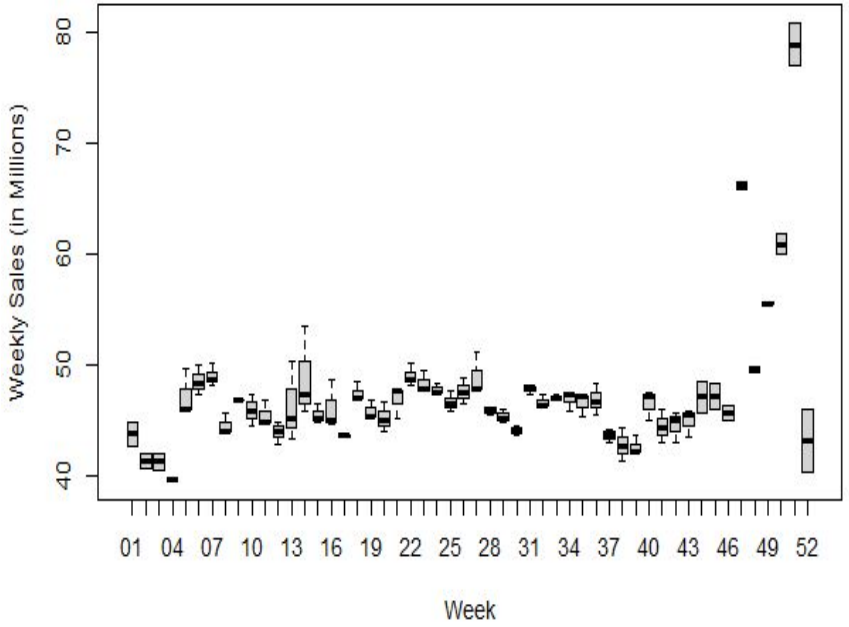
Labor Day: ~ 2nd Week of September

Thanksgiving/Black Friday: ~4th week of November

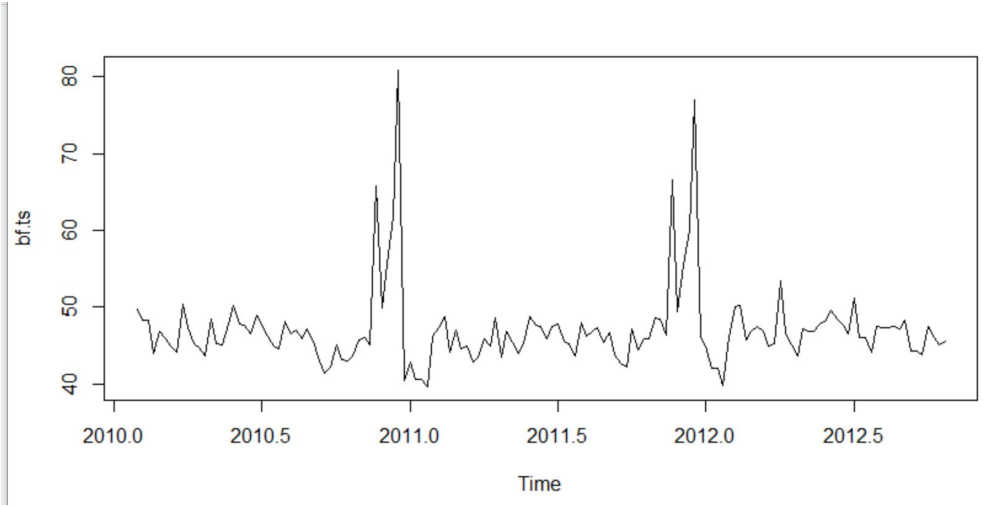
Christmas: ~ 4th Week of December

Data in Millions Aggregated





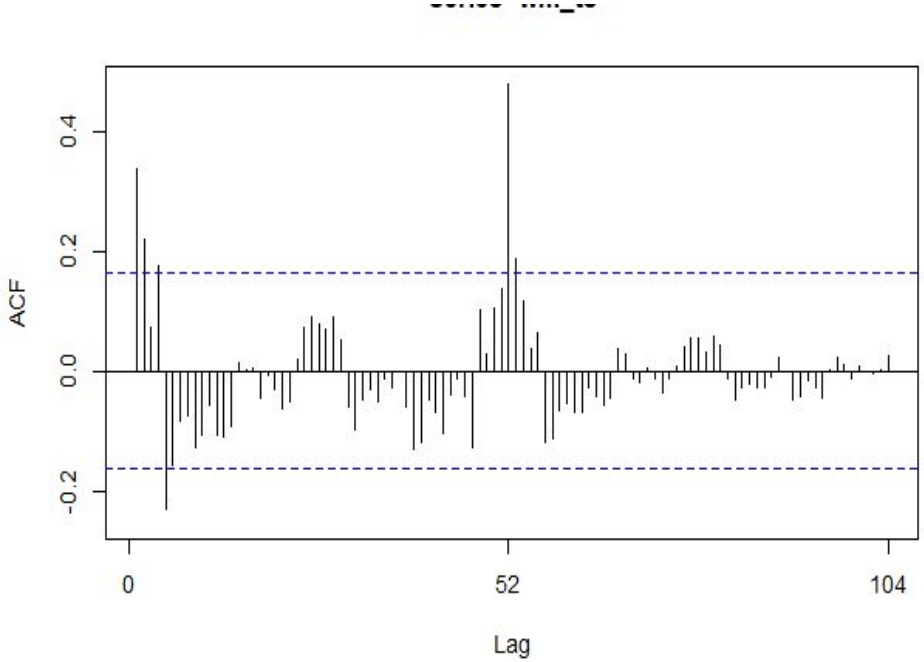
Time Series -




```
> summary(bf.ts)
   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
  39.60   44.88   46.24   47.11   47.79   80.93
>
```

- 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

Autocorrelation



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

Decomposition