

SDA Project Report

Lushaank K – 20170010082

Vineesh S – 20170010132

Hemanth P G – 20170010048

Sumanth J – 20170010135

Dataset: Metro Interstate Traffic Volume Data Set Minneapolis-St Paul, MN traffic volume for westbound I-94 along with the weather patterns from 2012 to 2018.

It has the following columns:

1. holiday Categorical US National holidays plus regional holiday, Minnesota State Fair
2. temp Numeric Average temp in kelvin
3. rain_1h Numeric Amount in mm of rain that occurred in the hour
4. snow_1h Numeric Amount in mm of snow that occurred in the hour
5. clouds_all Numeric Percentage of cloud cover
6. weather_main Categorical Short textual description of the current weather
7. weather_description Categorical Longer textual description of the current weather
8. date_time DateTime Hour of the data collected in local CST time
9. traffic_volume Numeric Hourly I-94 ATR 301 reported westbound traffic volume

We are working on Jupyter Notebook which uses Python Language and here are the packages you need to install before running the code in the Shell

1. pip install numpy
2. pip install pandas
3. pip install matplotlib
4. pip install seaborn
5. pip install statsmodels

From the data, we could identify number of holidays present in the datasets. For our better understanding and lack of confusion the holiday is only at the time 00:00:00, as the data was collected hourly, we would get 24 holidays for the data in a day.

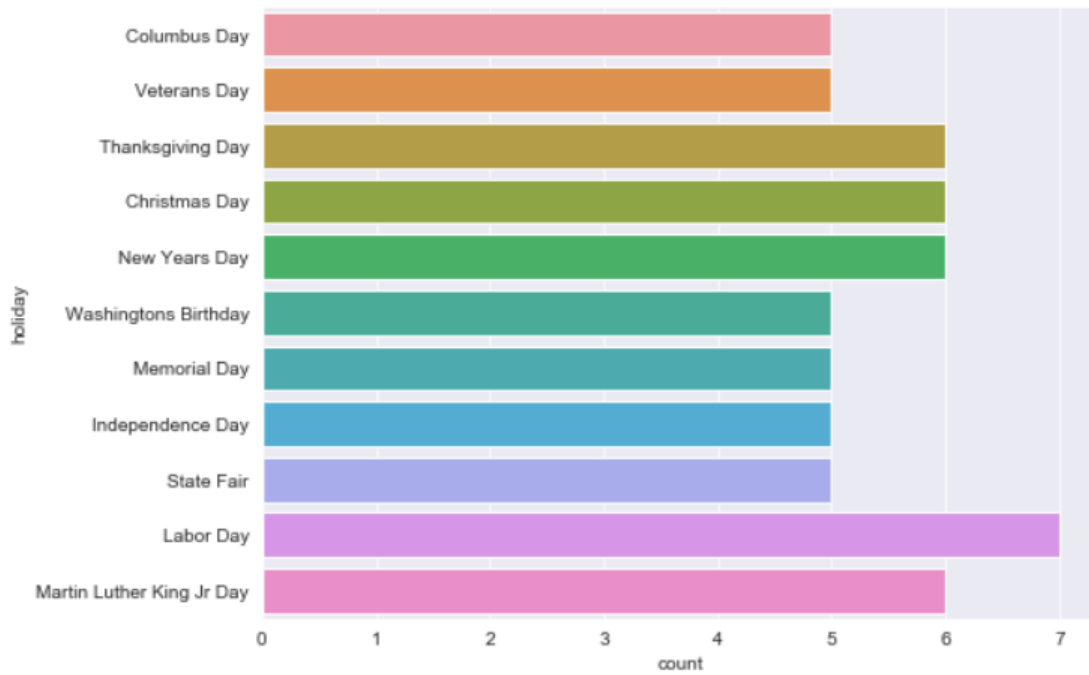


Fig-1: Holidays according to the data

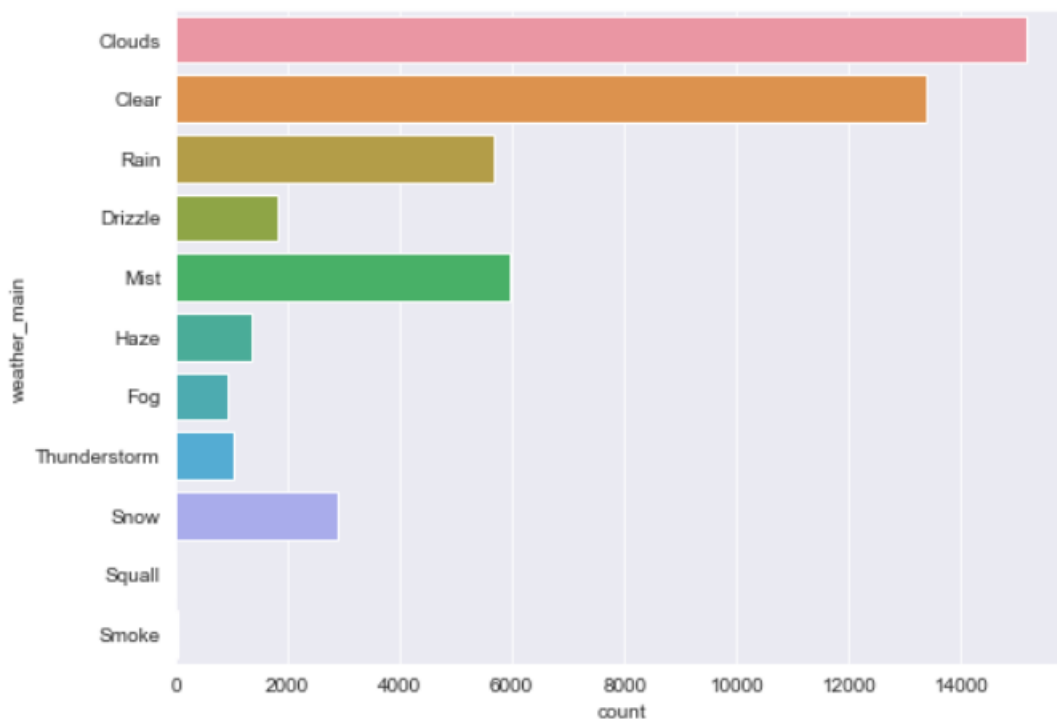


Fig-2: Weather according to the data

We could check the probabilistic weather for the future according to the previous dates.

We have used Five different types of Forecasting methods for our model

1. Auto Regression: The autoregression (AR) method models the next step in the sequence as a linear function of the observations at prior time steps.
2. Moving Average: The moving average (MA) method models the next step in the sequence as a linear function of the residual errors from a mean process at prior time steps. A moving average model is different from calculating the moving average of the time series.
3. Autoregressive Integrated Moving Average: The Autoregressive Integrated Moving Average (ARIMA) method models the next step in the sequence as a linear function of the differenced observations and residual errors at prior time steps. It combines both Autoregression (AR) and Moving Average (MA) models as well as a differencing pre-processing step of the sequence to make the sequence stationary, called integration (I).
4. Simple Exponential Smoothing: The Simple Exponential Smoothing (SES) method models the next time step as an exponentially weighted linear function of observations at prior time steps.
5. Holt Winter's Exponential Smoothing: The Holt Winter's Exponential Smoothing (HWES) also called the Triple Exponential Smoothing method models the next time step as an exponentially weighted linear function of observations at prior time steps, taking trends and seasonality into account. The method is suitable for univariate time series with trend and/or seasonal components.

```

AutoRegression: {'temp': -23.269830263988684, 'rain_1h': 0, 'snow_1h': 0, 'clouds_all': 12.64747109586271, 'traffic_volume': 1704.0354035605885}
Moving Average: {'temp': -18.616134327475407, 'rain_1h': 0, 'snow_1h': 0, 'clouds_all': 23.44955512148149, 'traffic_volume': 2162.9316063669744}
Autoregressive Integrated Moving Average: {'temp': -23.776878511432663, 'rain_1h': 0, 'snow_1h': 0, 'clouds_all': 3.6649648968429904, 'traffic_volume': 1385.133352451227}
Simple Exponential Smoothing: {'temp': -23.69999999925493, 'rain_1h': 0, 'snow_1h': 0, 'clouds_all': 1.174895234178439, 'traffic_volume': 1582.3160811661119}
Holt Winter's Exponential Smoothing {'temp': -23.69999999925493, 'rain_1h': 0, 'snow_1h': 0, 'clouds_all': 1.174895234178439, 'traffic_volume': 1582.3160811661119}

```

We then calculated Temperature vs Traffic graph, Weather vs Traffic and Converting the data into weekdays and plotting Traffic vs Weekday graph with Holiday vs Traffic respectively.

