

# AIR QUALITY ANALYSIS

## RESULTS ANALYSIS

- Box plot is one of common graphical systems utilized in EDA. A box plot or boxplot is a helpful method for graphically portraying gatherings of numerical information through their quartiles. Box plots may likewise have lines
- broadening vertically from the containers (whiskers) demonstrating inconsistency outside the upper and lower quartiles, henceforth the terms box-and-whisker plot and box-and-whisker graph. Exceptions might be plotted as individual focuses.
- Box Plot gives fundamental data about a dispersion. It graphically delineates a gathering of numerical information as indicated.

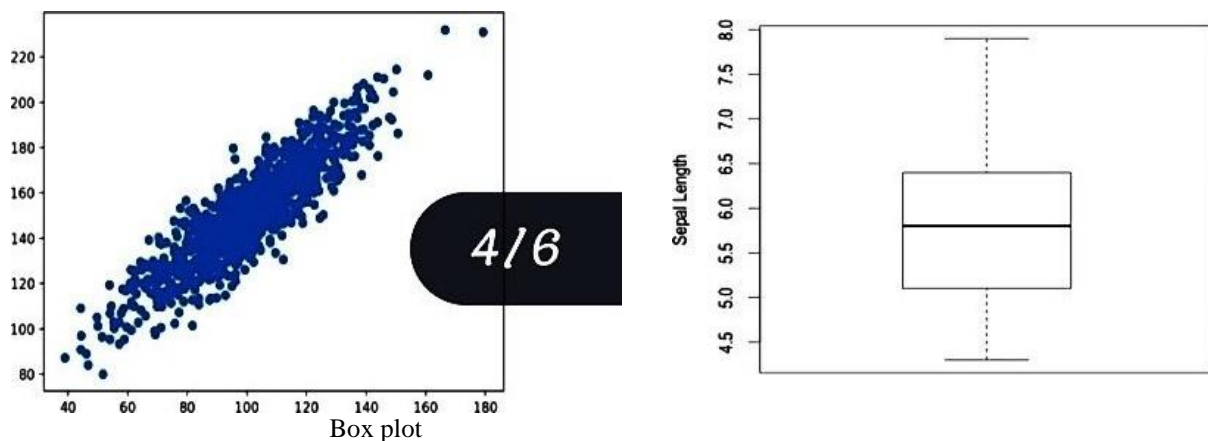
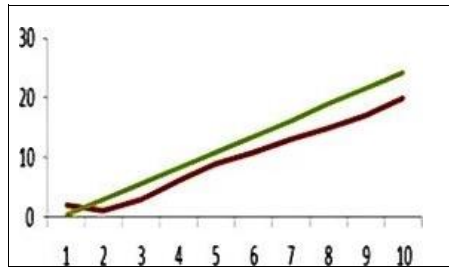


Figure Box-Plot analysis

- By this data analysis we came to know that there are seasonal variations and trend, in order to reduce these metrics, we resample the data month wise to predict it month wise.
- By resampling the data, we can reduce the outlier more efficiently than raw data. After removing the outlier's linear regression is applied to the filtered data and to fit the trend line on the data points gradient descent hyper parameters are used to optimize the model.

## LINEAR REGRESSION

- While doing straight relapse our goal is to fit a line through the dissemination which is closest to the majority of the focuses.
- Subsequently lessening the separation (mistake term) of information focuses from the fitted line.



## GRADIENT BOOST ALGORITHM

- The principle issue influenced by individuals is air contamination since air contains numerous substances which might be made by manmade or regular procedure.
- The Air substances present most organic atoms, points of interest and perilous material into the air.
- Boosting Algorithm is a victor among the most prevalent learning insights showed over the most recent twenty years.

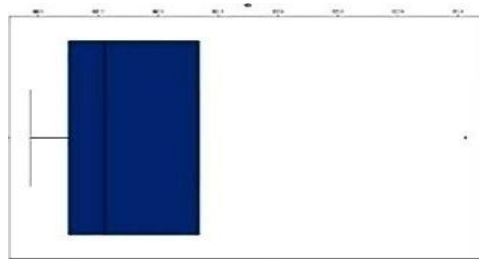


Figure Outlier removal using BPA

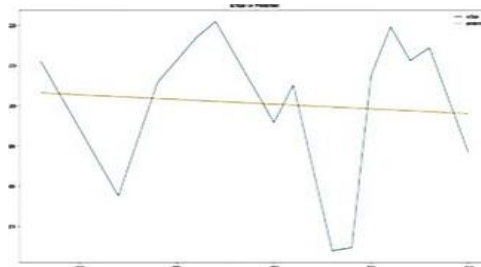


Figure Actual and predicted values

## CONCLUSION AND FUTURE ENHANCEMENTS

- Since our model is capable of predicting the current data with 95% accuracy it will successfully predict the upcoming air quality index of any particular data within a given region.
- With this model we can forecast the AQI and alert the respected region of the country also it a progressive learning model it is capable of tracing back to the particular location needed attention provided the time series data of every possible region needed attention.
- The air quality information utilized in this paper originates from the china air quality checking and investigation stage, and incorporates the normal every day fine particulate issue (PM2.5),