**Product market prediction**

Sree Kurapati Girija Kondapally

University of North Texas University of North Texas

[SreeKurapati@my.unt.](mailto:SreeKurapati@my.unt.)edu kondapallygirija@gmail.com

Sai Sidharth Chunduri Ranganadha Pawan Parankusam

University of North Texas University of North Texas

SaiSiddharthaChunduri@my.unt.edu PawanParankusam@my.unt.edu

Abstract

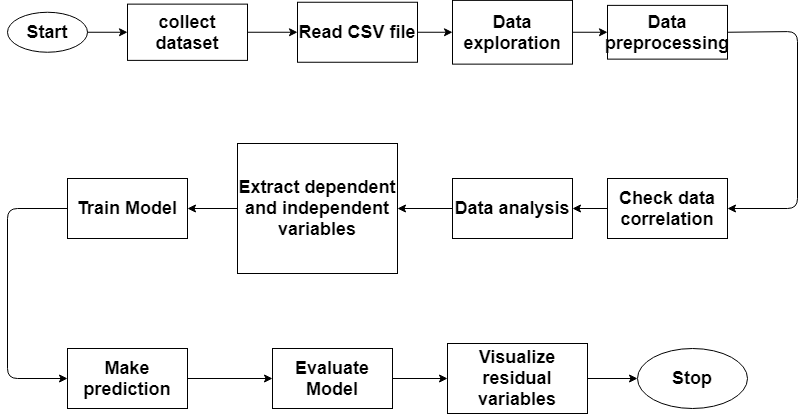
The stock market is essential to the corporate world. Physical stock exchanges allow for order placement from any location, or they can be electronic (trading over computer) done. We can estimate a company's future value with the aid of stock market predictions. With the help of a supervised learning process called linear regression, we can predict stock values using machine learning algorithms. The open, closed, and highest stock prices for each day are included in the data set we chose for this project. This data set is divided into training and testing datasets, and the data normalization process is carried out. After preprocessing the data, a liner regression model is trained and predictions are made, the model performance is evaluated to check on model performance.

Introduction

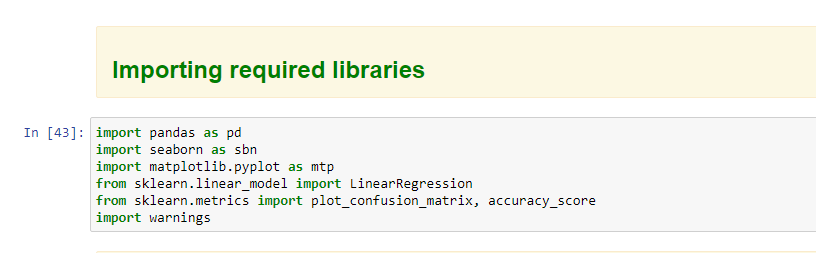
Stock market prediction Is a process that employs mathematical modelling to predict future market value of the product. The project aims in predicting the product volume using supervised machine learning linear regression model. The model is intended to work by fitting clean data to the regression model, the model will then extrapolate data patterns, understand variable relations and predict target value. The project evaluate model performance to check its accuracy.

Methodology and Result

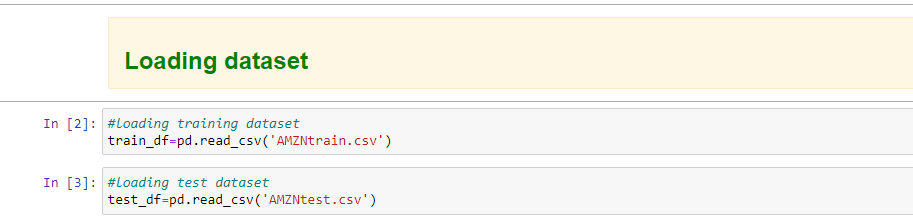
Below is our data analysis pipeline.



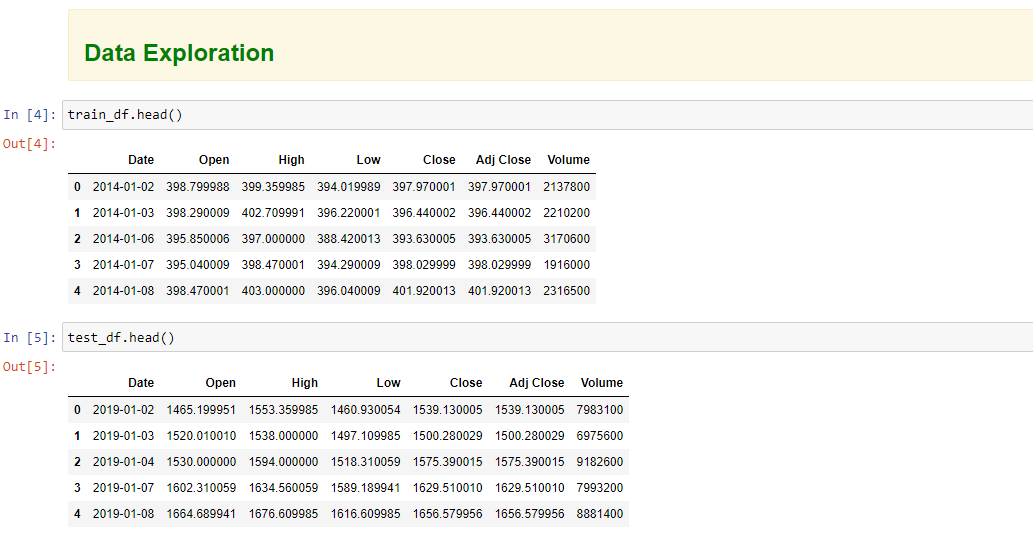
To analyze the data, we imported all the required libraries into the program.



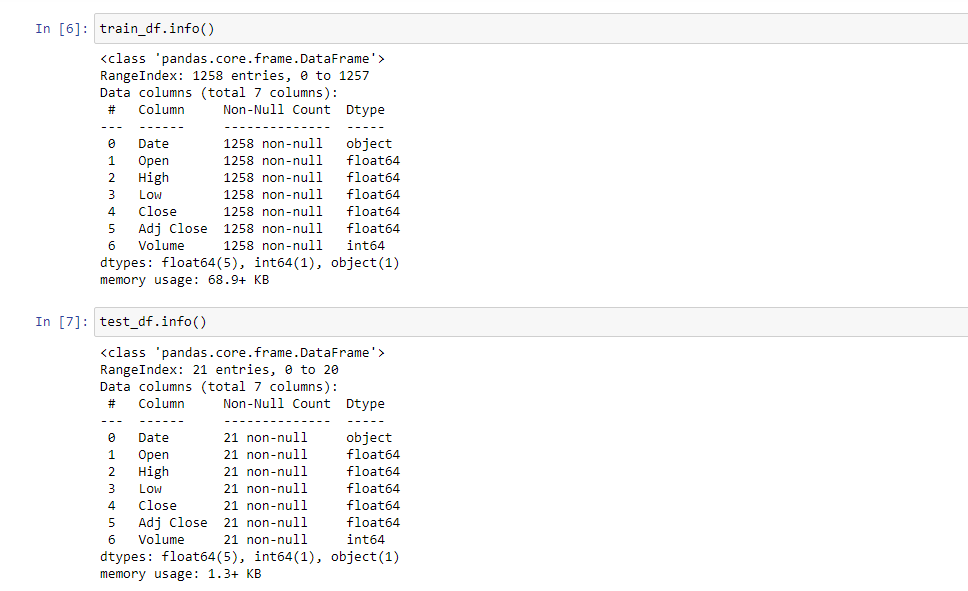
We loaded the datasets into the program and assigned it to the training and test data frames.



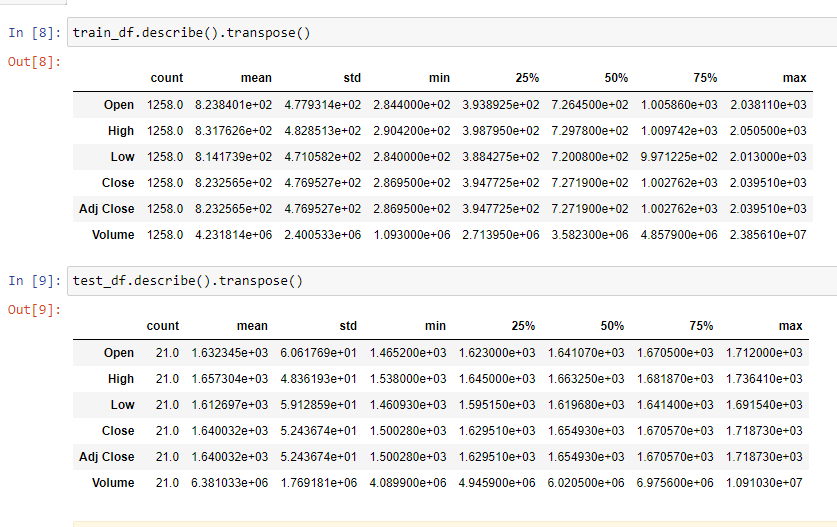
We explored the datasets by running the head function to explore the column headers.



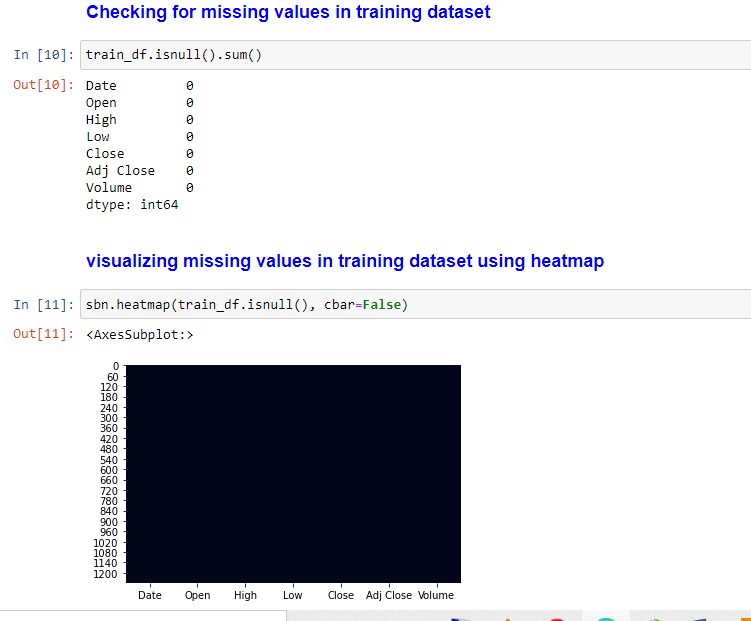
To check the number of rows and columns we ran info function



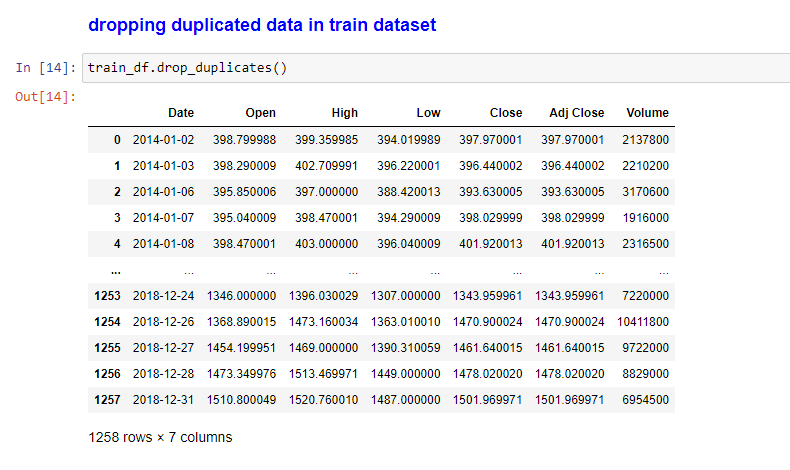
. To find statistical distribution of the data, we ran a describe function.



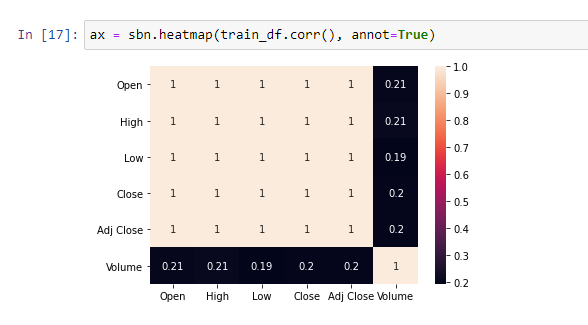
Data cleaning involves filling missing values, dropping duplicated data and converting data format into a correct one. From heat map, it was clear that the dataset had no missing values.



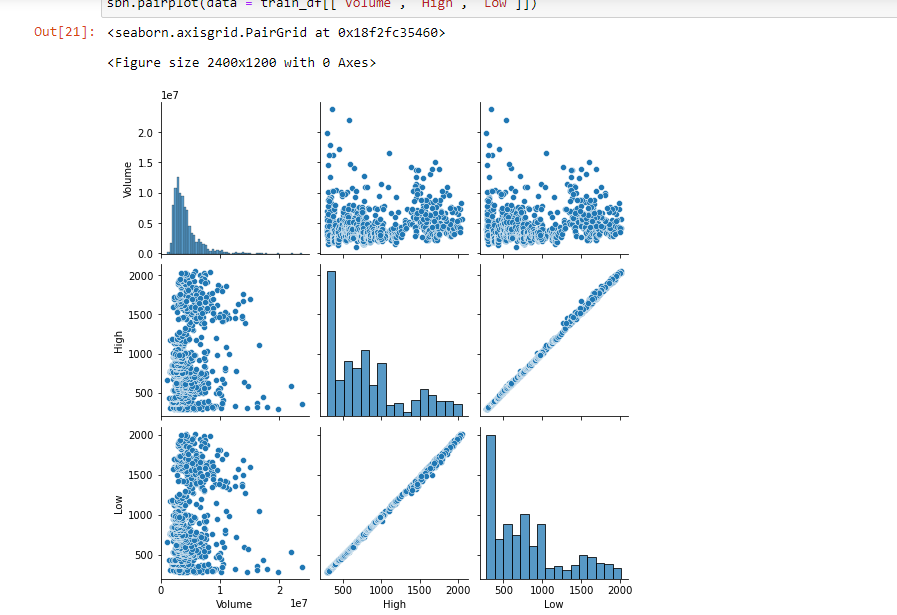
We dropped all rows with duplicated values to remove redundancy.



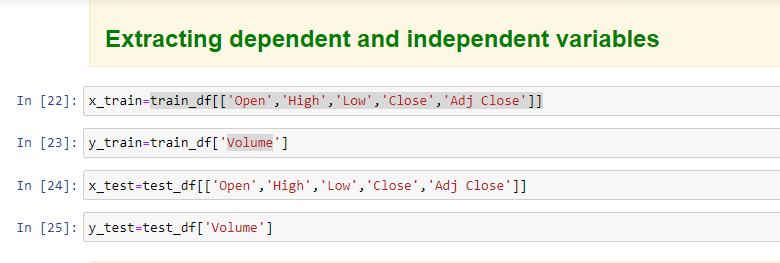
We checked the correlation between dependent and independent variables. Our dependent variable was Volume and it had the following correlation value with independent variables:



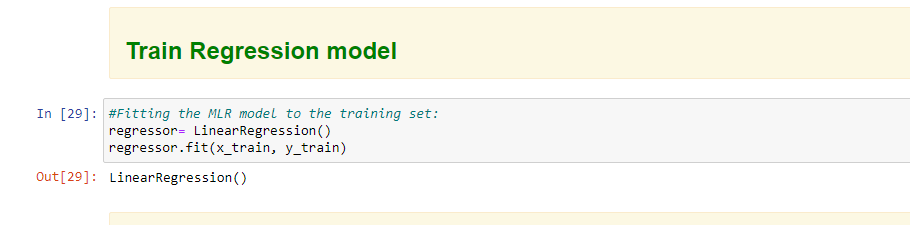
We ran a pair plot to analyze variable distribution and we found out that, the relationship between volumes and high is skewed, and the relationship between volume and low is also skewed.



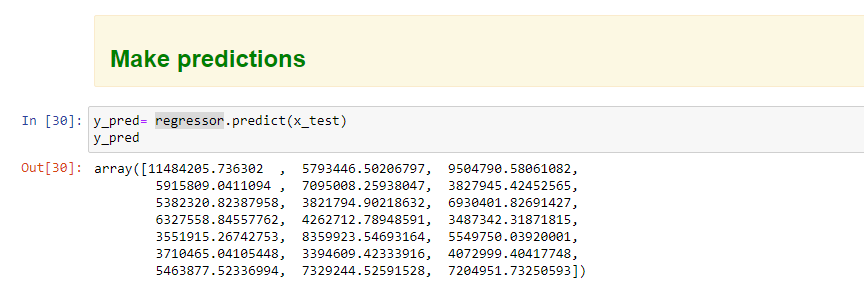
To train a predictive model, we had to extract dependent and independent variables.



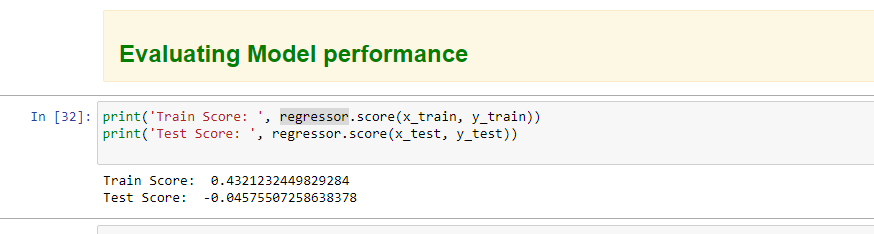
We trained the predictive model by fitting it with training dataset.



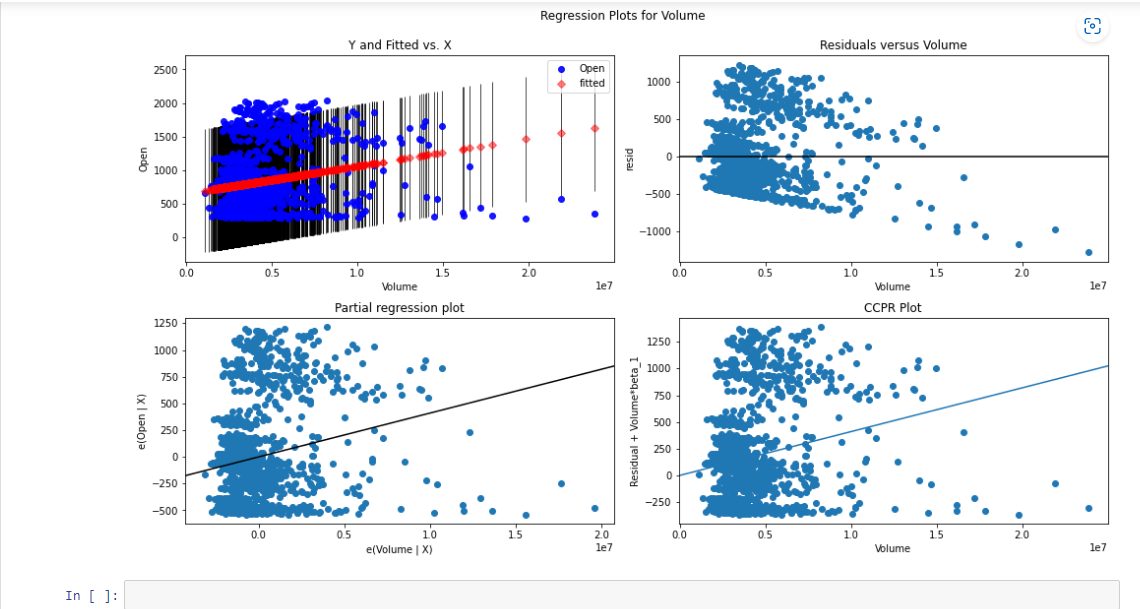
We made predictions by running predict function on test set.



We evaluated the model performance and found out that it had a Train Score: 0.4321232449829284.



We plotted the outliers using residual plot.



Conclusion

Stock market prediction is a key tool that if embraced by marketing company can foster its growth by foretelling future market trends.