A REPORT ON

TREND AND LAG ANALYSIS ON MIS AND MMS OF NTCL

BY

1. Ashish Sinha 2017AAPS0400G
2. Aarnav Dhanuka 2017B5A70945P
3. Ojit Mehta 2017A3PS0372G
4. Riddhesh Sawant 2017AAPS0261G

AT

NATIONAL TEXTILE CORPORATION LTD. WESTERN REGIONAL OFFICE

A PRACTICE SCHOOL - 1 STATION OF

BIRLA INSTITUTE OF TECHNOLOGY AND SCIENCE, PILANI

(JUNE, 2019)

A REPORT ON

TREND AND LAG ANALYSIS ON MIS AND MMS OF NTCL

BY

Ashish Sinha	2017AAPS0400G	B.E. ECE
Aarnav Dhanuka	2017B5A70945P	MSc. Physics ,B.E.CSE
Ojit Mehta	2017A3PS0372G	B.E. EEE
Riddhesh Sawant	2017AAPS0261G	B.E. ECE

AT

NATIONAL TEXTILE CORPORATION LTD. WESTERN REGIONAL OFFICE

A PRACTICE SCHOOL - 1 STATION OF

BIRLA INSTITUTE OF TECHNOLOGY AND SCIENCE, PILANI

(JUNE, 2019)

Acknowledgements

I would like to express my special thanks of gratitude to Mr. Manoj Kumar K. G. the head of the office(NTCL,WRO),Mr. Pawar the HR head of the office, Mr. D.V.G. Hariprasad Korupolu the joint IT manager of the company who is also our project mentor ,Deepika Maam from the HR department and Dr. Mukesh Kumar Rohil our PS instructor who helped us throughout the period. I would like to thank all of them for their guidance and for the knowledge that they have given us .I would also like to thank my project team for their constant support and determination towards completing the project.

BIRLA INSTITUTE OF TECHNOLOGY AND SCIENCE, PILANI(RAJASTHAN)
Practice School Division

Station: National Textile Corporation Ltd. Western Regional Office.

Duration:55 days **Date of Start:21st May,2019**

Date of Submission:12th July,2019

Title of Project:Trend and lag analysis on MMS and MIS of NTCL

Students:

Ashish Sinha	2017AAPS0400G	B.E. ECE
Aarnav Dhanuka	2017B5A70945P	MSc. Phsyics + B.E. CSE
Ojit Mehta	2017A3PS0372G	B.E. EEE
Riddhesh Sawant	2017AAPS0261G	B.E. ECE

PS Faculty: Dr.Mukesh Kumar Rohil

Project Areas: Data Analytics, Machine learning , Software Development

Abstract: The Project aims towards analysing the trend of sales of products(yarn and fabric) taking to consideration the various factors of inputs and observe the yearly trends of sales to create a model to predict future sales of the company and optimize the buying of raw products to avoid losses. And to from a interface to show different graphs and the predicted output.

Signature of Student **Signature of PS Faculty**

Table of Contents

Introduction

The Indian textile industry contributes about 14 percent to industrial production, 4 percent to the GDP and 13 percent to the country's export earnings. It provides direct employment to over 45 million people. The textile sector is the second largest provider of employment after agriculture. National Textile Corporation was established in 1968 to look after the functioning of 119 sick textile mills acquired through three Nationalisation Acts. The mills which were not in a position to revive have been closed subsequently with surplus manpower given MVRS and 23 mills revived to the extent possible. The current employee strength is around 7800.

The lower level of Modernization in most of the Mills is resulting in lower productivity & utilization. Mills are not able to produce high value yarn and also limiting value addition to the product for getting better margins. Higher Power cost due to non-availability of captive power,

Mills have to depend on State Electricity Boards for getting power at Apaying higher wages.

In such a scenario analysis of performance statistics of the corporation for maximising the overall turnover and satisfying the needs of the employees simultaneously to ensure overall growth of NTC is important. Analysing the sales data by taking into account the parameters affecting the sales will help in making sound decisions regarding the modernisation of mills, employment, actions on closed mills, leasing of land and other by the management and Directors of NTC.

Lag analysis of Sales data will help in predicting the future Sales by taking in the production, mills and other dependent parameters and will help in setting up the correct cost for maximising sales of the organisation.

The interface is formed so as to make the visualization of data easy and efficient. And so that any employee working on it can refer to it and make decisions.

Trend and Lag analysis

1)LAG ANALYSIS:

In statistics and econometrics, a distributed lag model is a model for time series data in which a regression equation is used to predict current values of a dependent variable based on both the current values of an explanatory variable and the lagged (past period) values of this explanatory variable.

Here assuming some dependency of sales on parameters mills , mill location , cost of raw materials , quantity produced, wages and salaries of employees in mills , power and transportation charges and other information available on MIS and MMS the sales of the company from every individual mill will be predicted.

2) TREND ANALYSIS:

Trend analysis is a probabilistic analysis in which the data is analysed to find correlation in two functions and give a possible future trend for both the parameters.

MIS and MMS

Management Information System (MIS) is a framework within the ERP of the company which logs the data at every step of the process of production of yarn and fabric starting at quantities and prices of raw materials for each mill followed by quantities of input and output at each step of production including waste generated at each step up until the final quantity of produce.

Marketing Management System (MMS) is the framework utilized by the sales and marketing department. It logs all the data about sale of produce, the price each vendor pays and the quantity they receive from individual mills

GUI

GUI stands for Graphical User Interface is a form of user interface that allows users to interact with electronic devices through graphical icons and visual indicators such as secondary notation, instead of text-based user interfaces, typed command labels or text navigation. It was made on demand of the company. We had planned to make it using a python library called Tkinter which had a few limitations and we then used another library called as PyQt. Where we have made a window which has widget for time spam as input where in we give it a start date and a end date and the mill name and a item name and a particular type of graph. After inputting all this we get the corresponding graph as the output to which we can add headings as well as zoom it for detailed view. There is another tab which gives us a prediction value of cost.

Methodology

The methodology to be followed for the project was divided into different steps. Those were:

1)Data collection

2)Data extraction

3)Data cleaning

4)Data visualization

5)Analysis

6)Building a prediction model

7)Predicting future sales

8)Forming a interface for displaying the results

1)Data collection: After discussion with the IT manager we came to this topic. He gave us access to the MIS and MMS servers of NTCL and we got access to the real time data form the past 2-3 years.

2)Data extraction: The data on the server was on Microsoft SQL which we had to obtain. We used the scrapping method where in we extracted the data from the SQL server and converted it into tabular form in excel.

3)Data cleaning: The data extracted has to be cleaned and pre processed to access it properly and to perform operations on it. This was done with the Pandas library of python. We converted the excel file to pandas dataframe and then performed operations on it.

4)Data visualization: To analyse the data it needs to be plotted so that we get a better idea of what and how are the trends in the market. We

used matplotlib and seaborn to libraries of python which are used for data plotting. We plotted different parameters against each other to get an idea of variation of sales and the inter-relation between them.

5)Analysis: analysing the trends that are being followed and to judge witch parameters are important in sales and work on the model according to that.

6) Building a prediction model: Build a predicting model using neural network by using keras, another python library. This prediction model will we test on a test data to check for its accuracy and will be optimized to increase the accuracy.

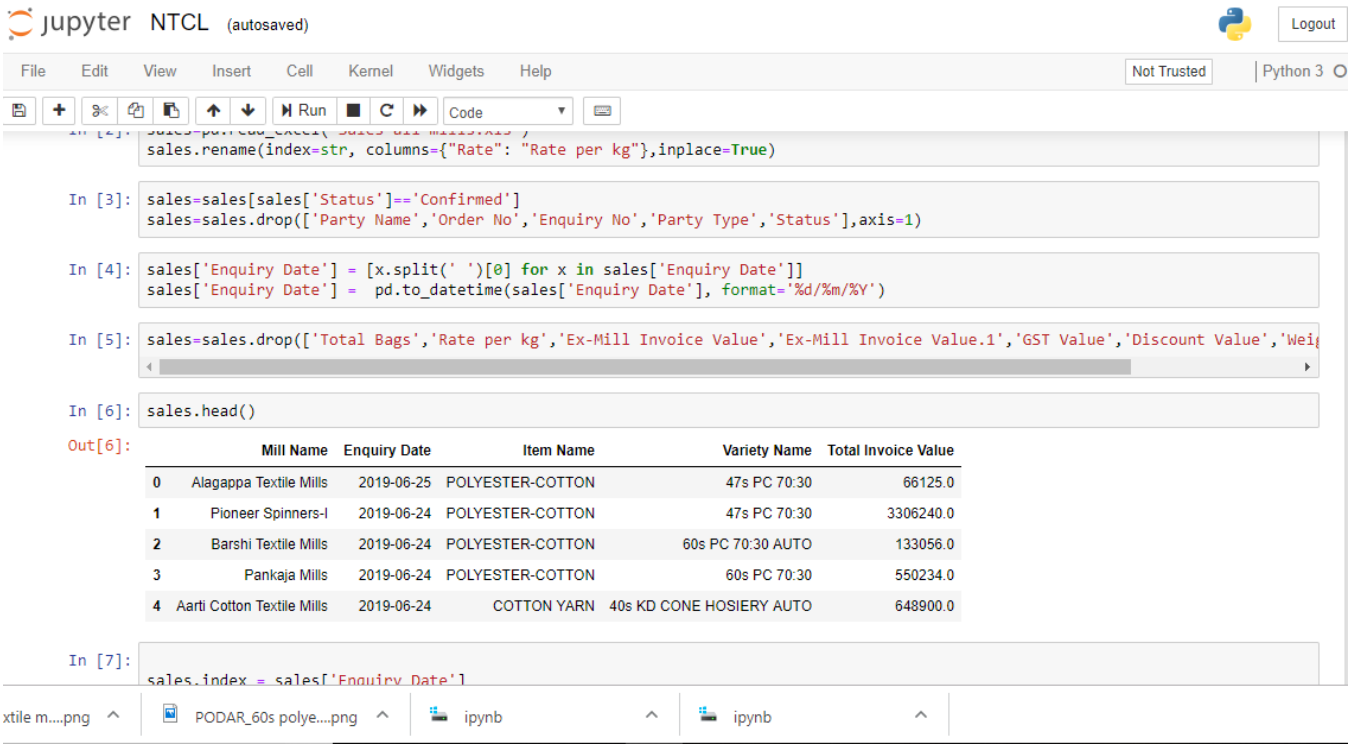
7)Predicting future sales: The model will be used to predict future sales where the input parameters will be given and it will output the optimum price for maximum sales.

8) Forming a interface for displaying the results: We formed a user interface to display the prediction results as well as the graphs for different time periods.

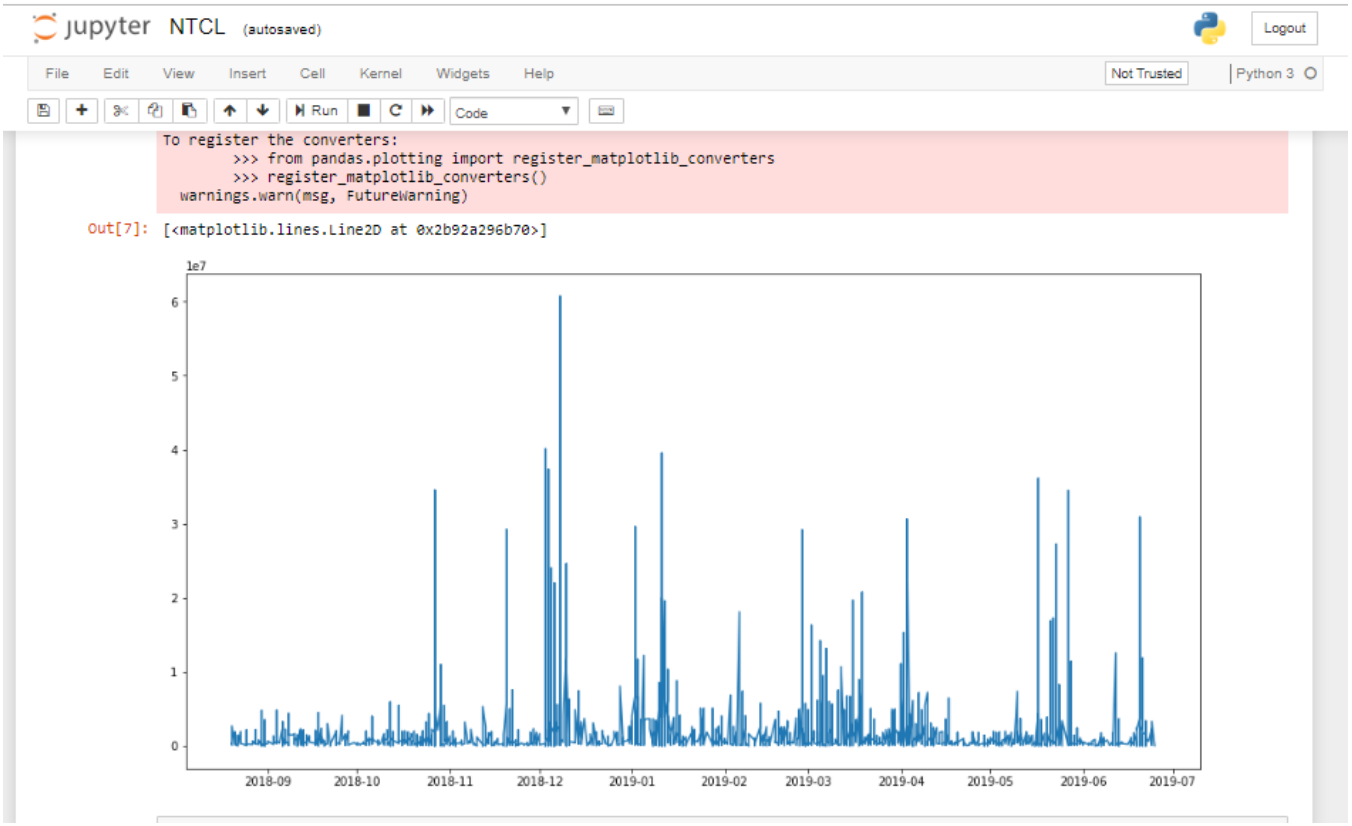
The Model

The prediction model is made using the machine learning algorithms where in we use different mathematical techniques to for a model that will predict the output in future. Different model had to be tried before getting a proper model which could predict the output properly i.e. a model with maximum accuracy. This model can be used to predict the future sales which will give the company an idea about how will the market behave.

The data we obtained from the IT department was raw and had to be cleaned as well as needed to be sliced and edited to make it algorithm ready. We used Pandas and Numpy libraries from Python to do these tasks. We had to drop different parameters which were not necessary and did not affect the final cost .Below we have show different parameters that we dropped as they were not contributing towards the output.

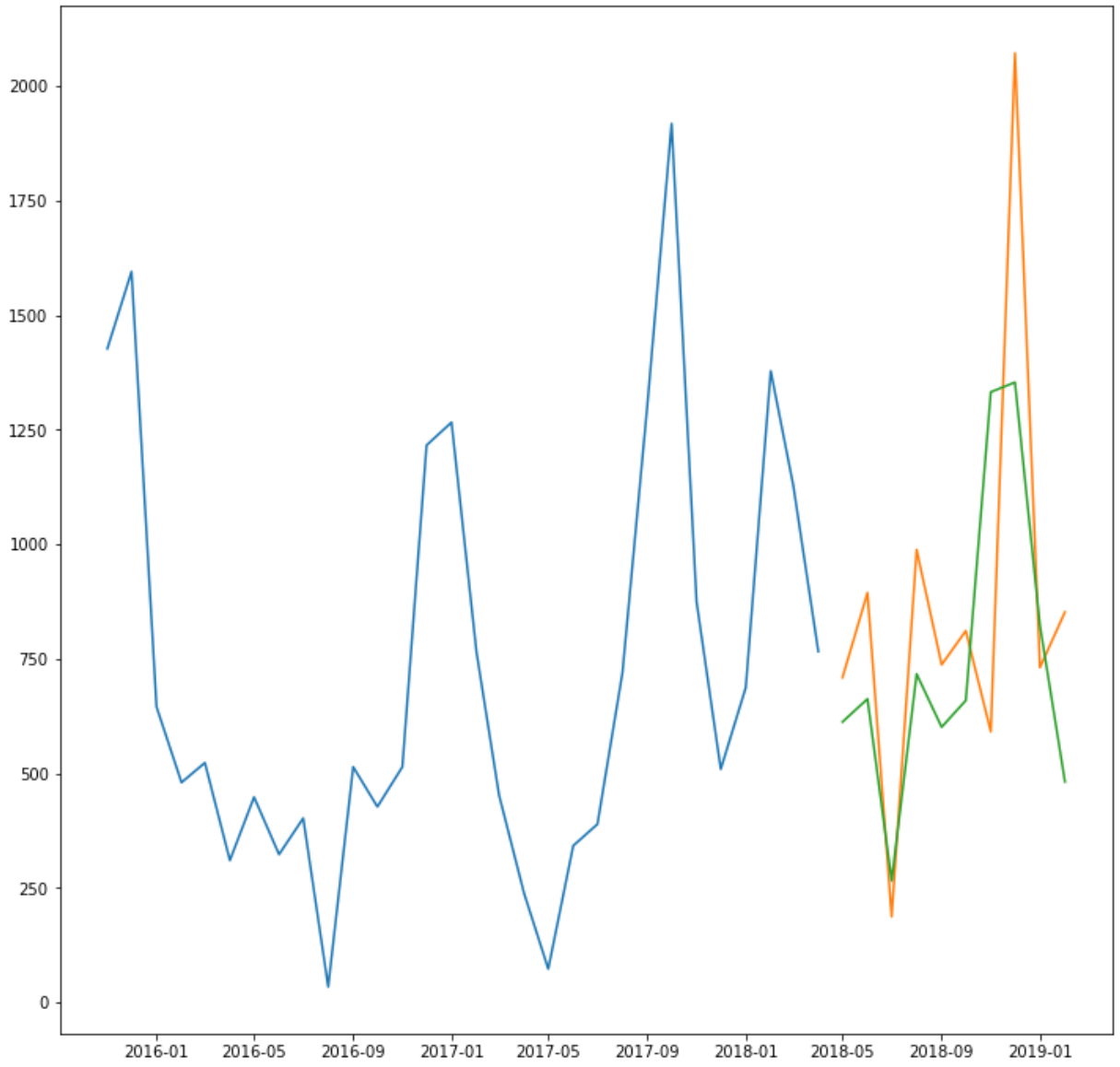


Further we plotted the data to get an insight of how is it varying with time and does it need to be standardized. This was done using another python library called Matplotlib which has different variations of graphs to be plotted. Below we have shown a graph of total invoice value vs time. As we could see there was no such trend so it was a bit difficult to decide which model to use.



We first tried the kNN regression model but the model was not at all accurate and was not the proper choice for the problem. Then we tried using neural network using keras but the dataset for that was small and it could not make accurate prediction on it.

As it was a time series data we used the ARIMAX model for prediction and we got good results from it.



Rajnagar textile mills 40s KD auto

The graph shown above is of a particular mill and a particular variety.

The accuracy is decent and shows peaks of sales and times when it is high.

6.1 ARIMAX

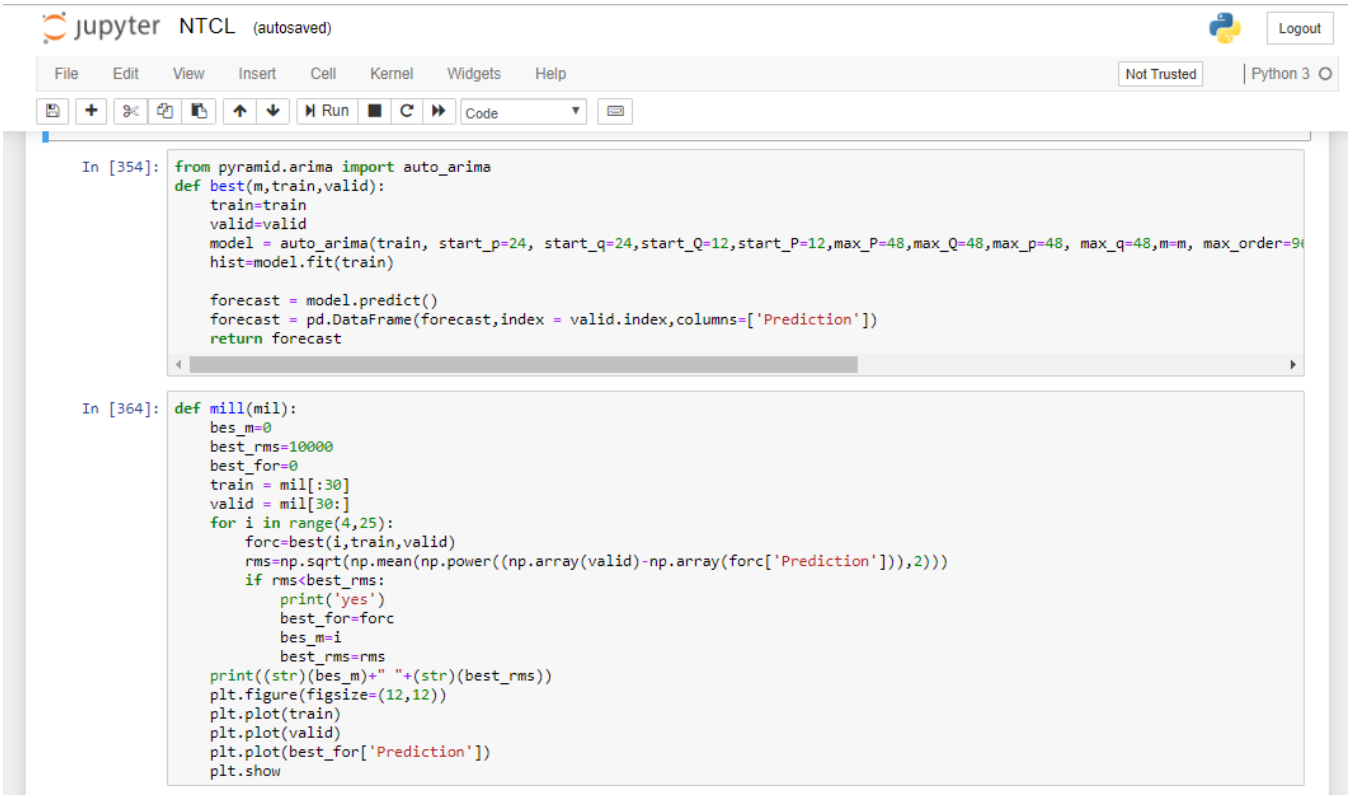
An Autoregressive Integrated Moving Average with Explanatory Variable (ARIMAX) model can be viewed as a multiple regression model with one or more autoregressive (AR) terms and/or one or more moving average (MA) terms. This method is suitable for forecasting when data is stationary/non stationary, and multivariate with any type of data pattern, i.e., level/trend /seasonality/cyclicity.

ARIMAX is related to the ARIMA technique but, while ARIMA is suitable for datasets that are univariate (see the article, entitled' What is ARIMA Forecasting and How Can it Be Used for Enterprise Analysis?'). ARIMAX is suitable for analysis where there are additional explanatory variables (multivariate) in categorical and/or numeric format.

Business Benefit: By analyzing the various combinations of predictor variables, the business can forecast product growth, trends, patterns and seasonality, if any. The enterprise can also identify any gap between

the targeted and estimated growth and develop an appropriate strategy to reduce this gap in order to achieve targets and results.

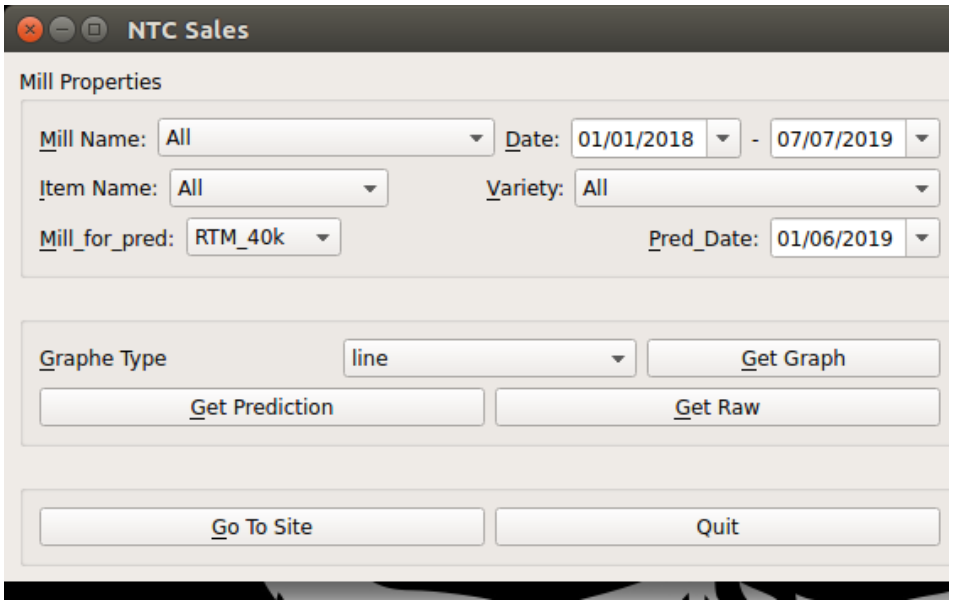
ARIMAX provides forecasted values of the target variables for user-specified time periods to clearly illustrate results for planning, production, sales and other factors.



This is the code for forming the model and fitting it to the given data.

The Interface

The Interface consists of a mainWindow on which a grid Layout is attached. The grid Layout consists of children widgets arranged relatively in the layout. Each Button, LineEditor, Date editBox and Pushbuttons is initialized as a widget for uniformity and attached in relative positions in the grid.



See Apendix for more.

The grid contains:

Widget Name	Widget Type	Widget Function
Mill Name, Item Name,Variety, Mill_for_pred, Graph_type	Combobox	Dropdown widget for selecting from list
Date, Pred_date	DateEdit	Date selection from Calendar widget
Get Prediction , Get Raw, Get Graph, Go To Site , Quit	QPushButton	Clickable Button Widget

Some Properties of Gui:

- ➔ Plotting Graphs based on the mill name, variety name, Item name and graph type and giving the trend of Rate and Bags produced. The options available for plotting are Line, Box, Bar, Histogram, Scatter, Area
- ➔ Getting the Raw data. The actual data based on the parameters set can be provided by the gui in Tabular form.

- ➔ Visiting the Company Site. The site of the company can be visited directly from the GUI and all the functionalities can be used.
- ➔ Getting the prediction of Total Invoice. The predicted value based on the parameters date and Mill name can be obtained in graphical form as well as raw numerical form.

Conclusion

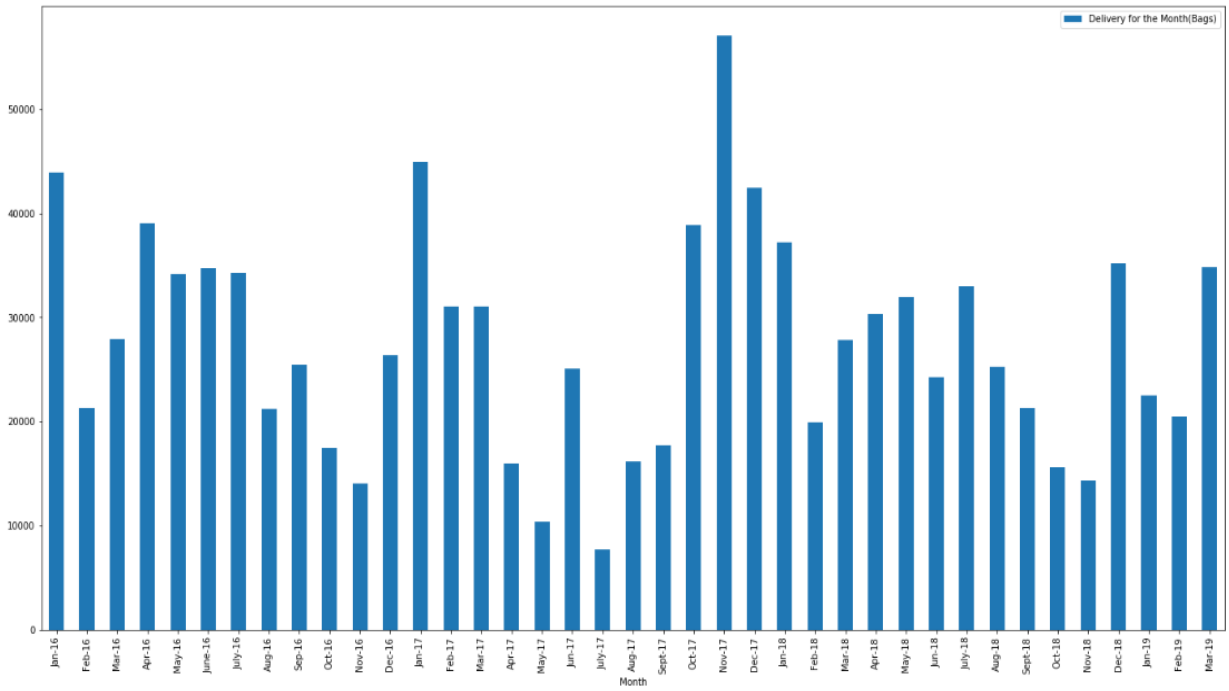
After the various observations we made through the various plots of different mills, we arrived at a few conclusions of how much each mill produces and sells, and what additional factors may be taken into consideration.

On an average the Winter period is generally the one with the highest sales. In addition to the time, the location of the mill highly influences the output, as mills located in not metropolitan cities tend to have a greater output and sales compared to the ones in metropolitan cities like Mumbai. This maybe down to the external factors pertaining to that area or region.

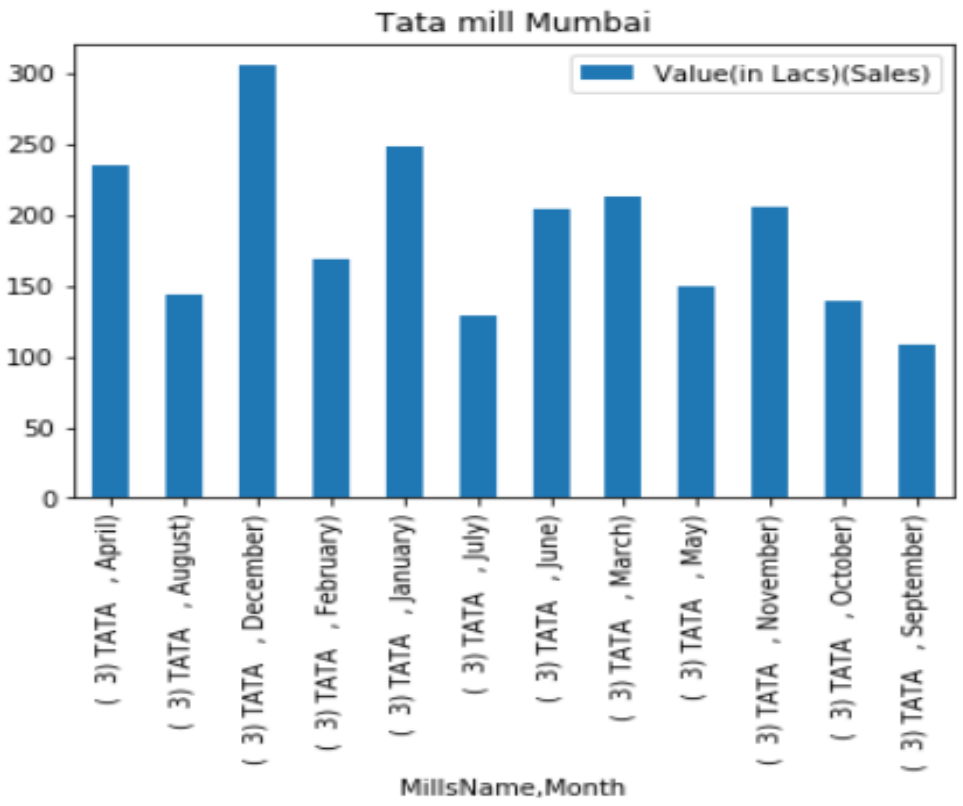
The model we build can help the organization estimate the amount of production and help them in optimizing the amount of raw product they buy, and hence save them from over or under buying raw products. And the interface contains different graphs

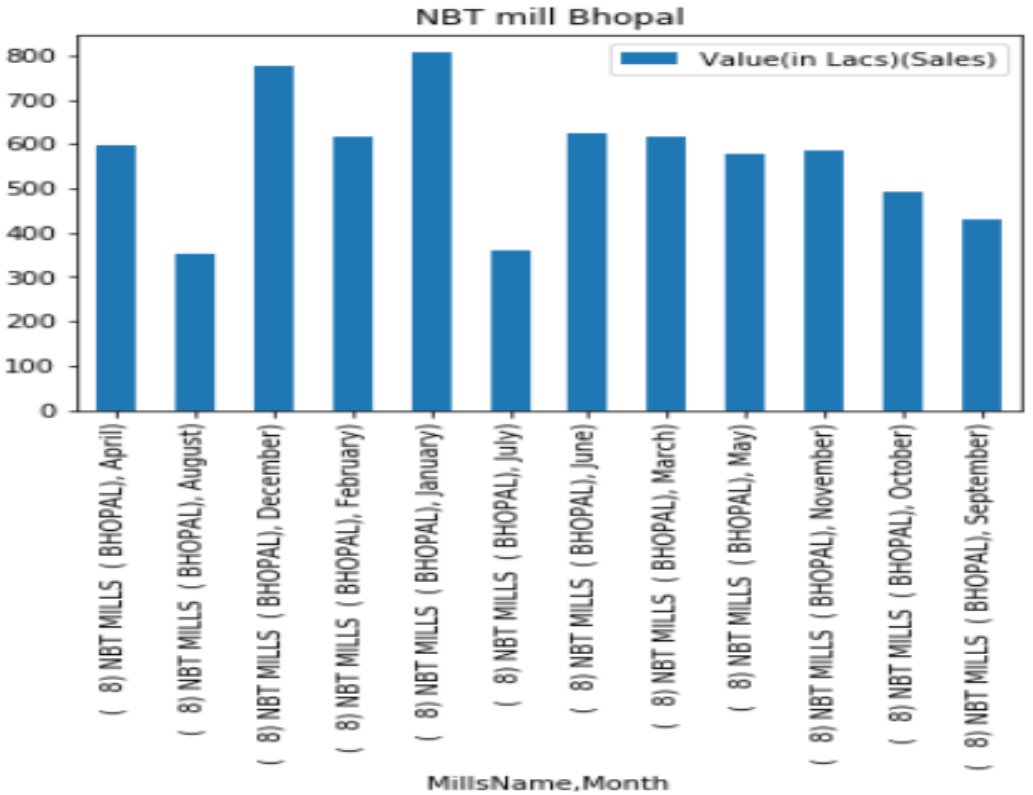
where in the user can select the given time period and the type of graph and will get the corresponding output, there is another tab that shows the output of the prediction model.

Appendix



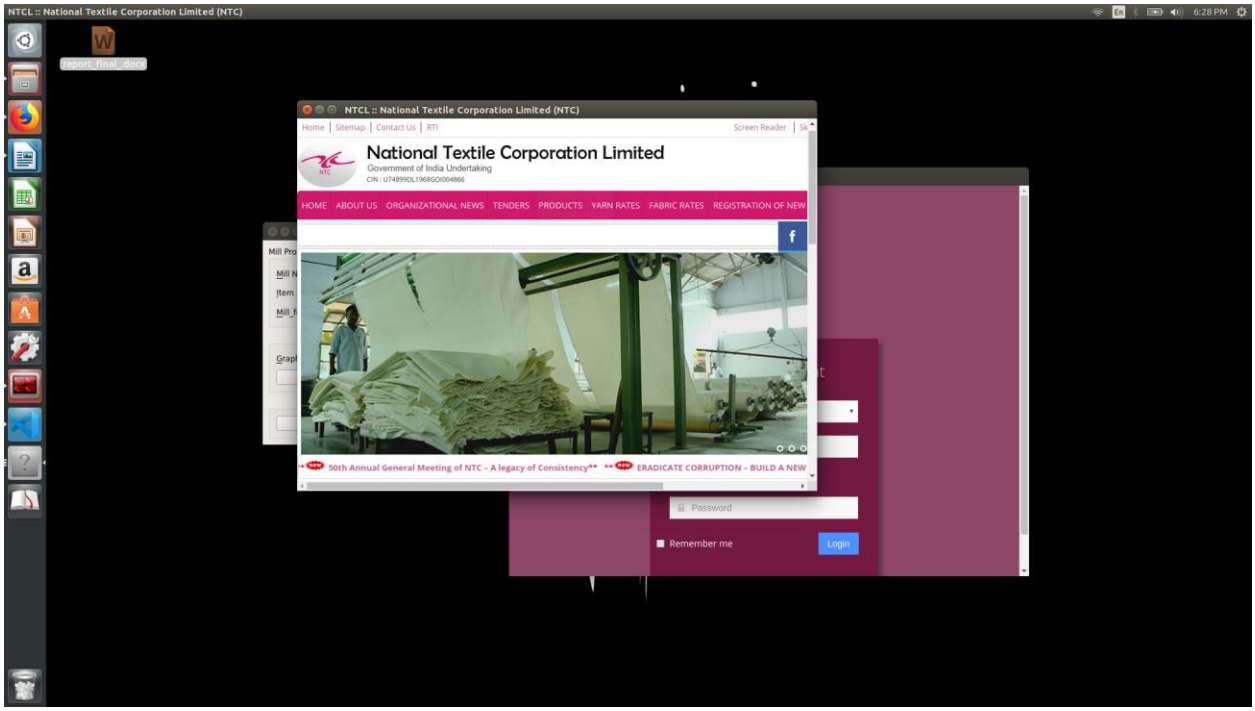
This is the total delivery data of NTCL of January,2016 to March,2019. As we can see every year the highest sales occur during the winter is the highest (i.e., the months of November, December, and January).


Tata mill Mumbai
Value(in Lacs)(Sales)
300
250
200
150
100
50
0
(3) TATA , April)
(3) TATA , August)
(3) TATA , December)
(3) TATA , February)
(3) TATA , January)
(3) TATA , July)
(3) TATA , June)
(3) TATA , March)
(3) TATA , May)
(3) TATA , November)
(3) TATA , October)
(3) TATA , September)
MillsName,Month


NBT mill Bhopal
Value(in Lacs)(Sales)
800
700
600
500
400
300
200
100
0
(8) NBT MILLS (BHOPAL), April)
(8) NBT MILLS (BHOPAL), August)
(8) NBT MILLS (BHOPAL), December)
(8) NBT MILLS (BHOPAL), February)
(8) NBT MILLS (BHOPAL), January)
(8) NBT MILLS (BHOPAL), July)
(8) NBT MILLS (BHOPAL), June)
(8) NBT MILLS (BHOPAL), March)
(8) NBT MILLS (BHOPAL), May)
(8) NBT MILLS (BHOPAL), November)
(8) NBT MILLS (BHOPAL), October)
(8) NBT MILLS (BHOPAL), September)
MillsName,Month

The above two graphs show the different ranges of production of two mills, one situated in Mumbai, and the other in Bhopal, which shows that the production in Bhopal mill is more than that one in Mumbai by a considerable amount. This shows us the importance of the mills location.

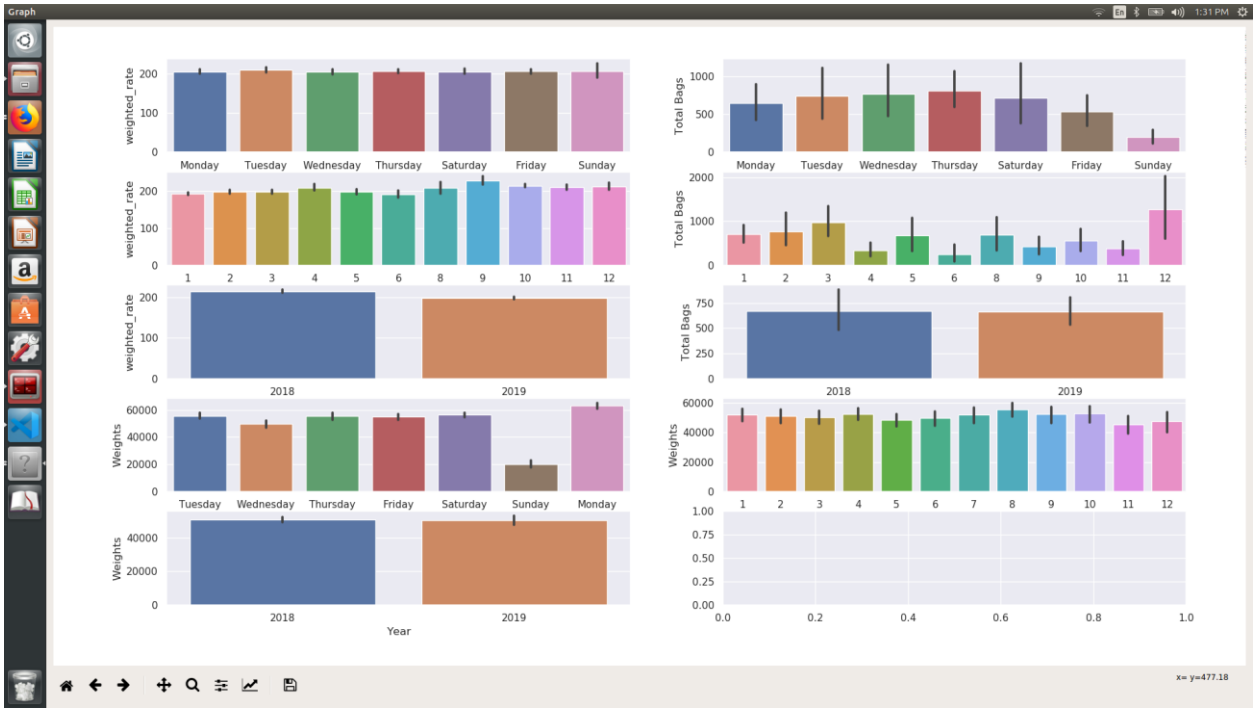
1)Site from Gui



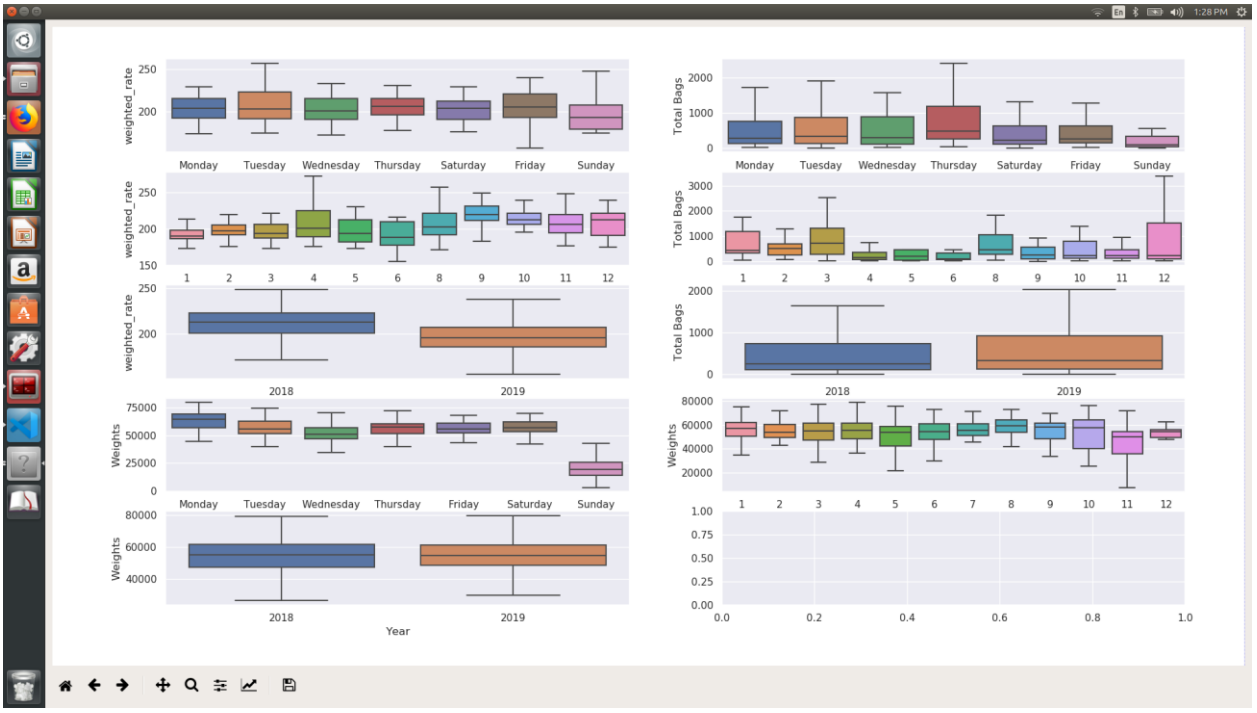
2)Line Plot Showing The trend

3)

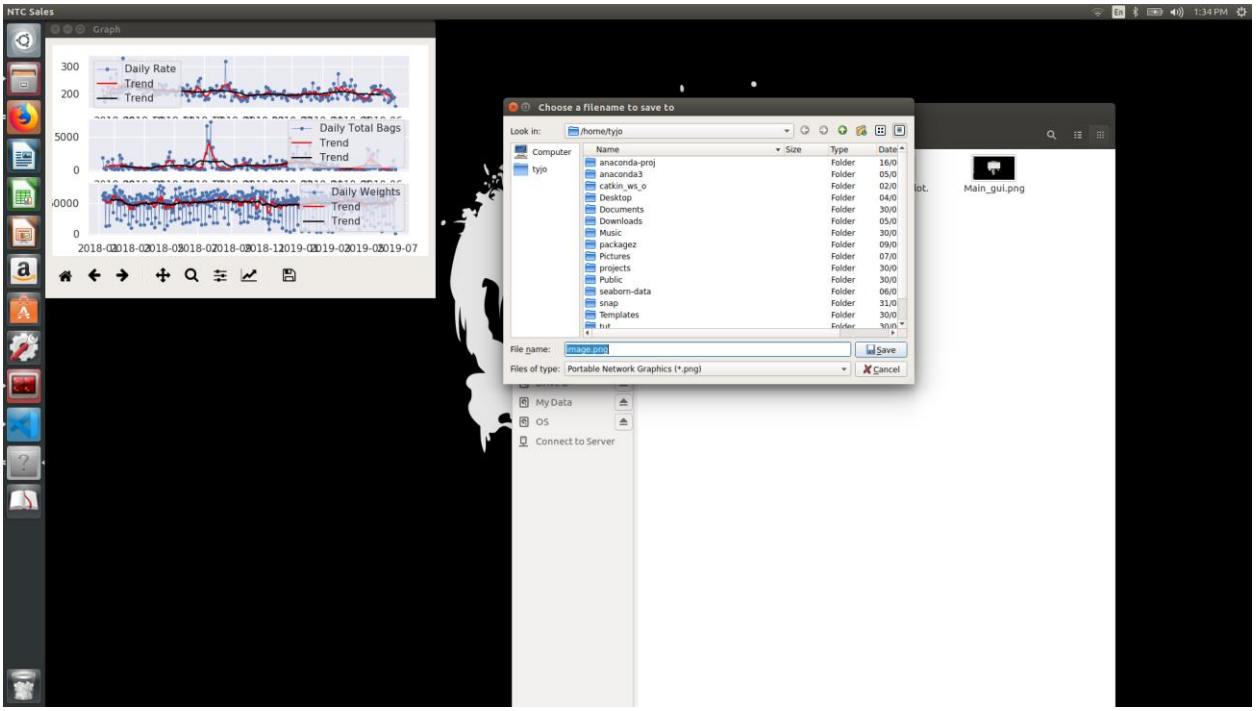
Bar Plot from gui



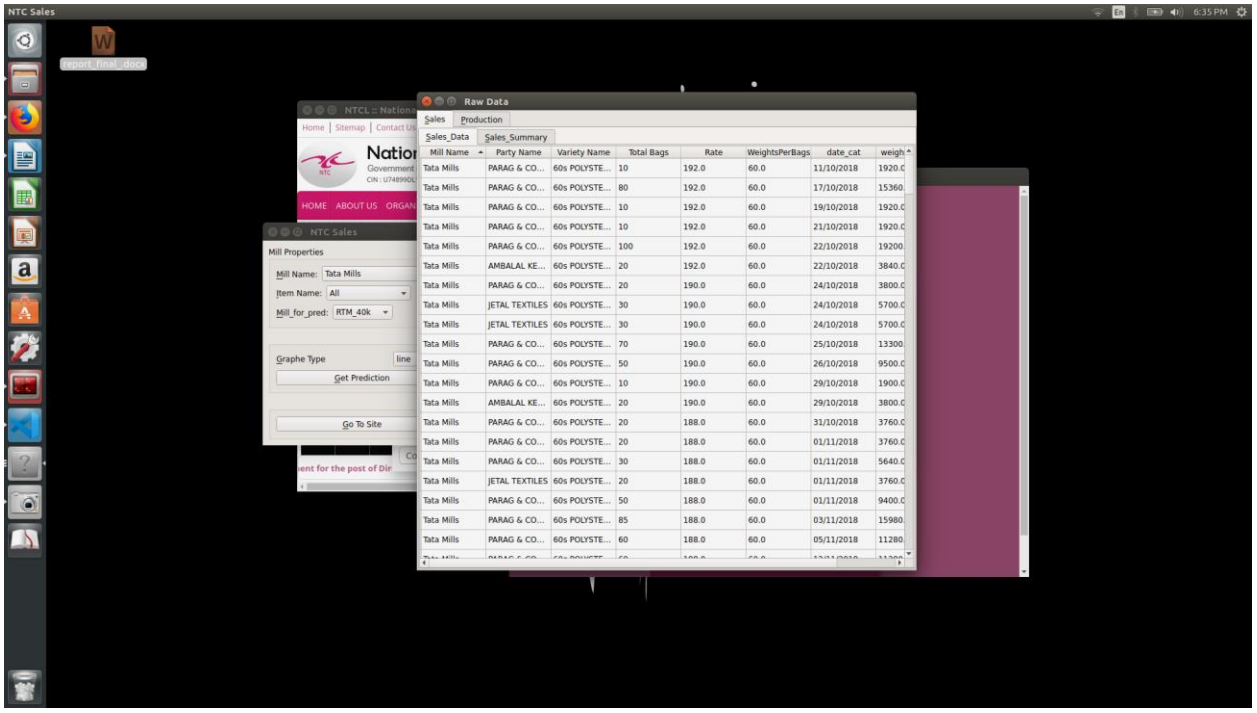
4)Box Plot using Gui



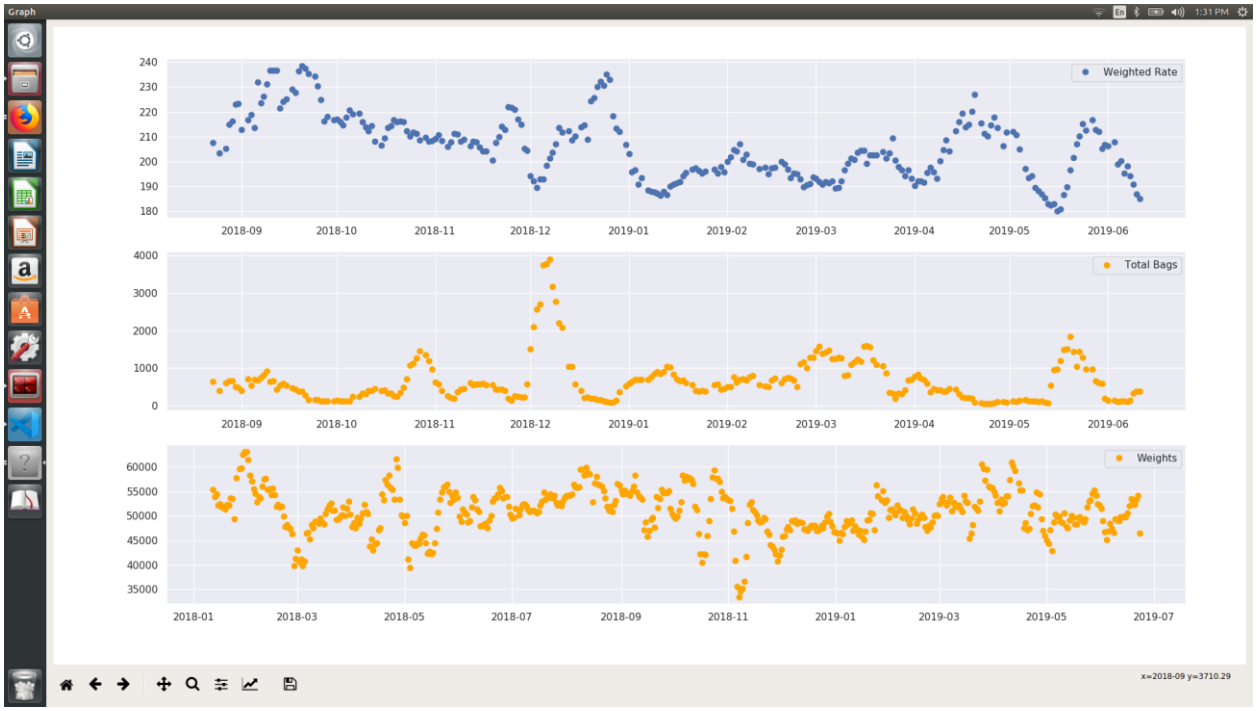
5)Saving Graph



6)Raw Data from Gui



7)Scatter Plot From Gui



References

1. Python - https://www.python.org/

2. Pandas library - https://pandas.pydata.org/

3. Numpy library - https://www.numpy.org/

4. Jupyter Notebook - https://jupyter.org/

5. Matplotlib library - https://matplotlib.org/

6. PyQt- https://pypi.org/project/PyQt5/

7. Pyramid-Arima- https://pypi.org/project/pyramid-arima/