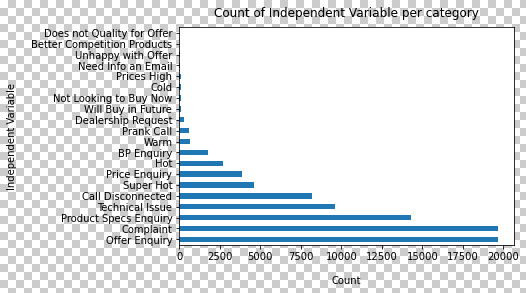
**Assignment #1:**

**Analyze the contact center data (attached herewith) to understand the causes of customer calls and come up with recommendations on reducing call volumes and handling calls better to increase customer sales. You should:**

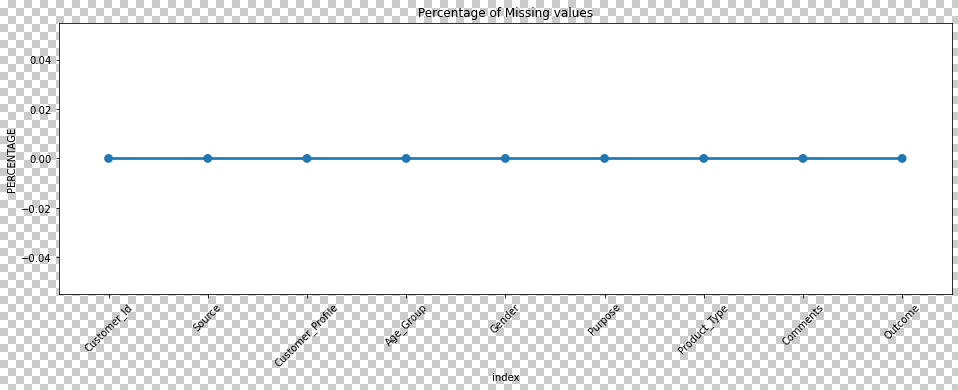
1. **Conduct EDAs to analyze the data (including mining/analysis of text fields and inputs) and better understand call reasons and outcomes**
2. **Summarize your findings using appropriate visualizations (feel free to use any tools for this) For this assignment, please provide crisp, data-backed and to-the-point recommendations for improvement, based on your analysis.**

Analysis

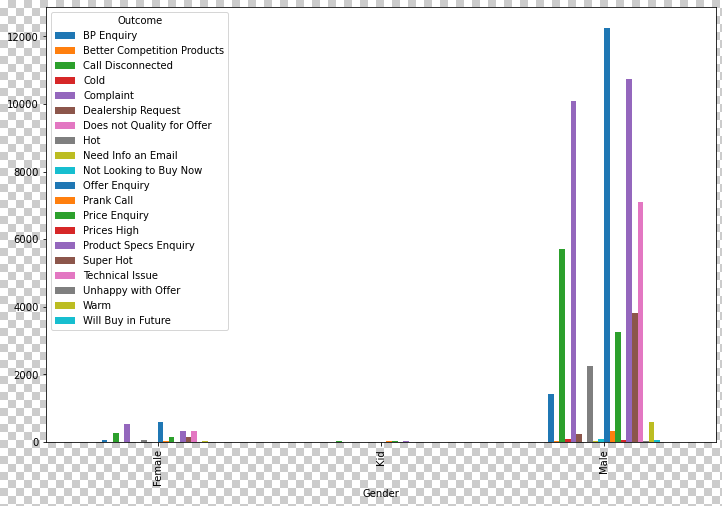
1. Convert the xlsb file into xlsx file
2. With the help of bar plot we can to know the distribution of Outcome variable.



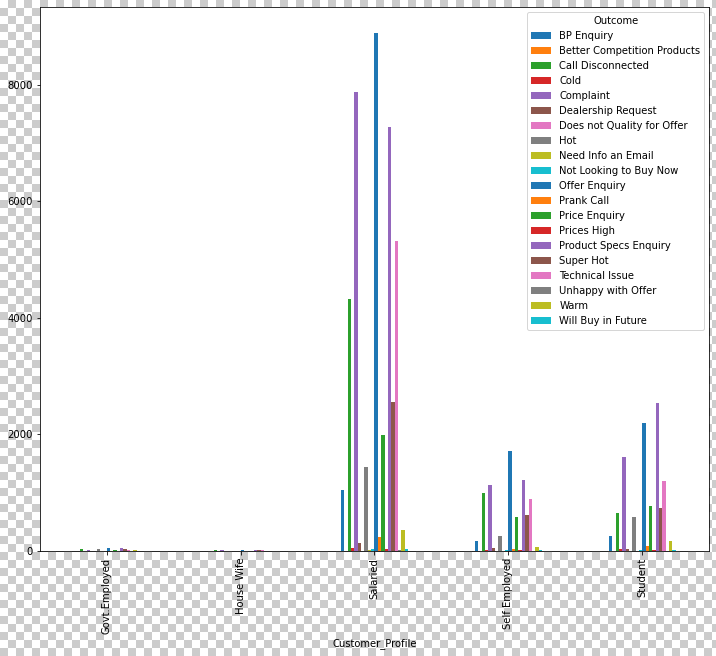
1. Checking for the missing value if it is present



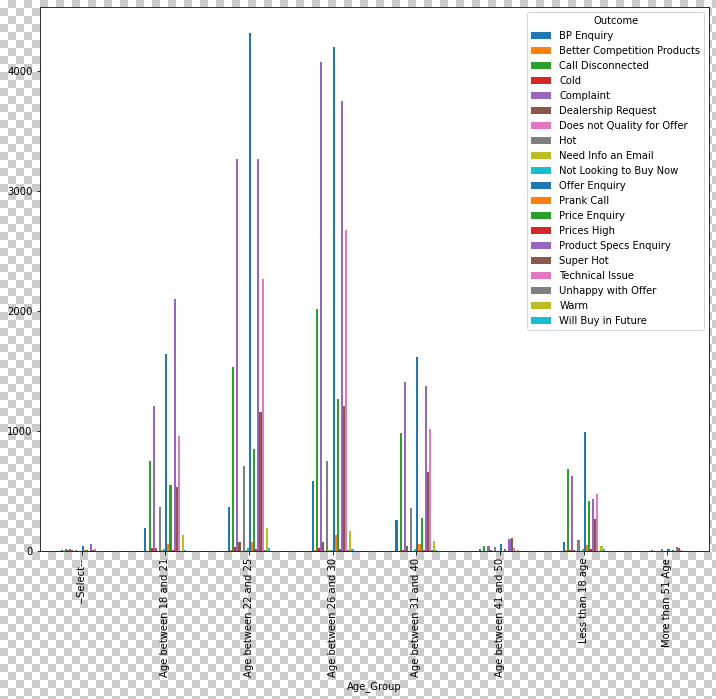
1. Checking for the missing value if it is present
2. Create a copy of the dataframe so that we can perform some EDA on top of it.
3. We can see how the customers are divided based on Age\_Group distribution and similarly with Customer\_profile variable.
4. Since we can clearly see that there are 8028 column values are blank in Age\_Group column and similarly more than 10k column values are blank in Customer\_profile column.
5. A huge number of records are blank so instead of dropping we can some insights from it as dropping these columns would lead to loss of data, on the other hand we can drop these column or replace these column with Mode values and these decisions should be taken wisely with respect to the business ask.
6. Extracted the blank records from both the columns and got some insights such as
   * 1. As we can most of the Female customers do make a complaint regarding Net Surfing issue in the notebook. Probably we can do something in order to improve in this area. On the other hand Male customers also make complaint regarding Net Surfing issue in the notebook.
   * 2. Compare to female, Male customers make more calls and make complaints.
7. there are records where Customer\_Profile is not given but the age group are given
   * 1. As we can most of the Female customers do make a complaint regarding Net Surfing issue in the notebook. Probably we can do something in order to improve in this area. On the other hand Male customers also make complaint regarding Net Surfing issue in the notebook.
   * 2. Compare to female, Male customers make more calls and make complaints.
8. Discarding all the blank records which are present in the dataframe and doing separate analysis. Again there is no hard and fast rule to do separate analysis but in order to show some insights we did it in this way.
   * + We got total number of Outcomes based on Gender column
     + total number of outcomes based on Gender and Customer\_Profile
     + total number of outcomes based on Gender, Purpose and Product\_Type
     + As we can see from the above observations that most of the calls are received from the Male customers then female and kids.
     + From the above observations most of the calls done by Male customers are for Offer Enquiry, Complaint, Product Specs Enquiry Technical Issue and Call Disconnected.
     + Also we can see that most of the male customers are salaried employees.
     + There are some male customers whose customer profile is not listed which we do not know what category they belong.
     + Male customers who are salaried employees make a lot of call regarding complaint, call disconnect issues, Technical issues on the other hand Male customers who are salaried employees make a call regarding offer enquiry, price enquiry, Product Specs Enquiry.
     + Based on the above observation we got to know that Most of the male customers who make a call regarding Complaint are mostly within 22-30 age groups **so basically in order to reduce the call volumes we should target these people and come up with some solutions**.
     + As we can see Male customers needed laptop for gaming, net surfing and for many more purposes, if we provide proper information regarding GPU, ram size, rom size etc on the website in a proper way probably we can a sudden drop of calls with respect to such enquiry.
     + As we can see kids below 18 age do a prank call which unnecessary increase the amount of calls we can do something to avoid this
     + On the other Female employees who are salaried and age between 26 and 30 do a lot of complaints and so do male customers.
9. As we can see Male employees are often making a call in order to know if any offer is going on



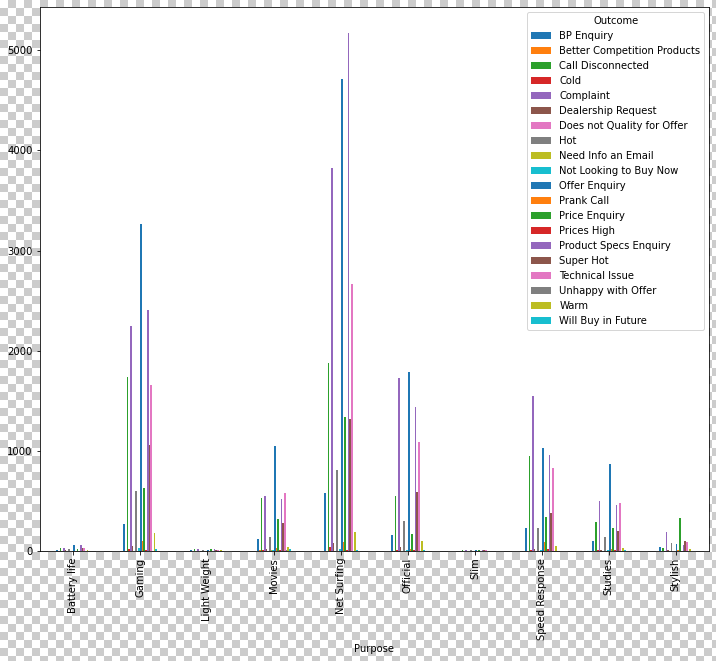
1. Salaried Employees are more often to make offer enquiry



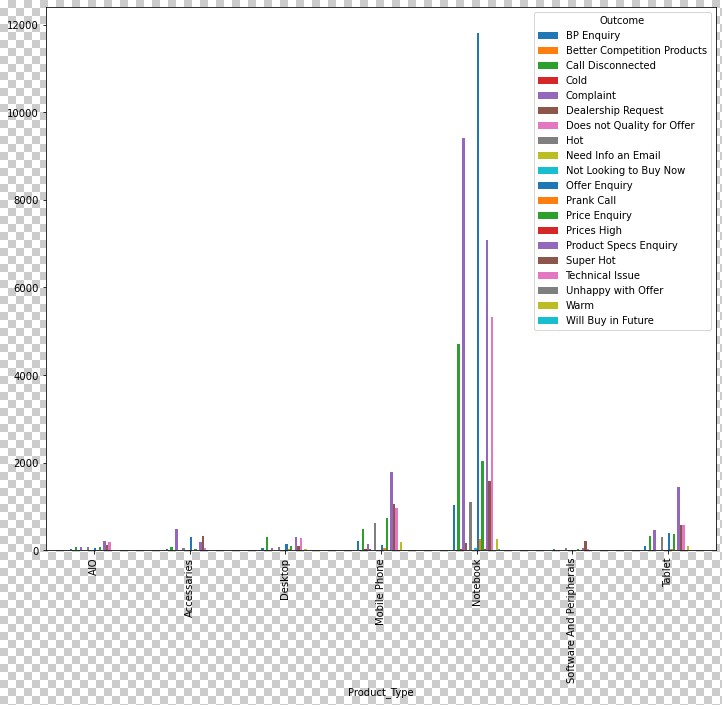
1. Age between 22 to 25 and 26 to 30 are more likely to call.



1. Customers are more likely to call for Net surfing and Gaming purposes.



1. Customers are more likely to call regarding Notebook, probably based on the situation we can put all necessary amount of information on the website in order to receive less number of calls.



Note: - We can come up with another set of ideas for the problem statement.

**Assignment 2:**

**Create a forecast for what the visits would look like for the next year based on the historic data points from the ‘Data for Forecast’ file (attached herewith). You should:**

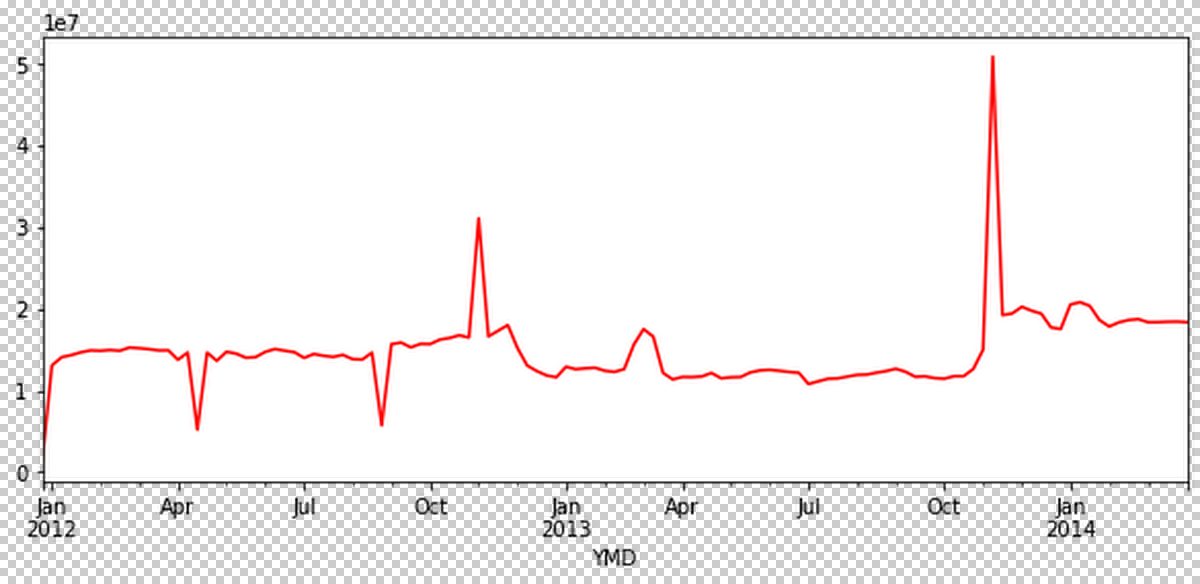
**1. Consider seasonality and normalize the anomalies, if any**

**2. Choose any apt methodology/tool/technique to create and present the analysis (based on your current skillsets and abilities)**

**For this assignment, we will be looking for creativity around "visual representation" of forecast, as well as details on the methodology used.**

1. Reading the input file
2. Splitting the Date column
3. Creating a dictionary for mapping the months
4. converting the date in the required format
5. Dropping the unnecessary columns for modeling purpose
6. Converting the YMD column into DatetimeIndex format. In order to make the data ready for modeling purpose.

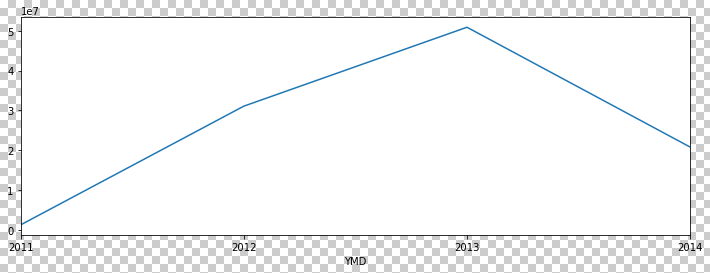
**Checking the trend of the data:-**

****

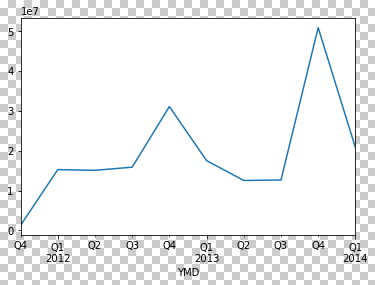
**By looking at the above figure we can observe that in 2013 between October to December month we can see some sudden growth and can be identified as an outlier/Anomalies present.**

**# Time Resampling**

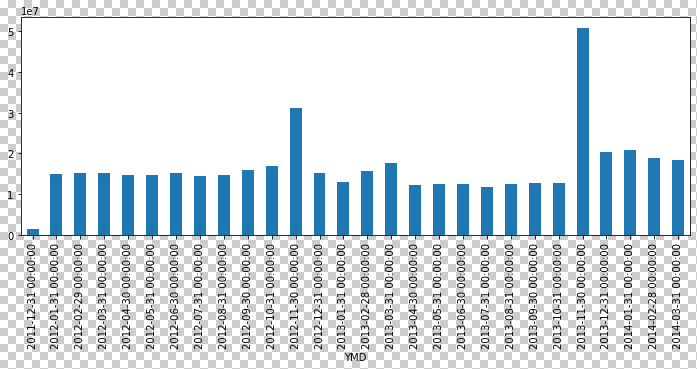
**Resampling is used in time series data. This is a convenience method for frequency conversion and resampling of time series data.**

****

**(Rule-A) As we can the trend of the data how it is increasing from 2011 to 2013 and then there is sudden decrease after 2013.**

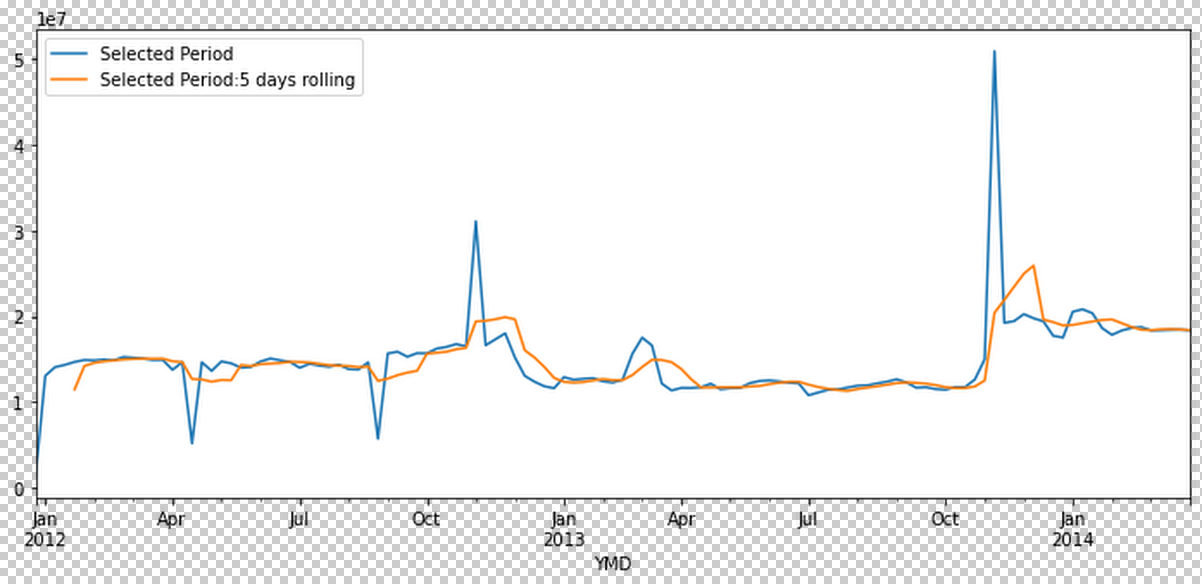
****

**(Rule-QS) With the help of rule QS we can see which particular quarter has seen sudden increase like in the above plot we can clearly say that in Q3 there is a sudden increase.**

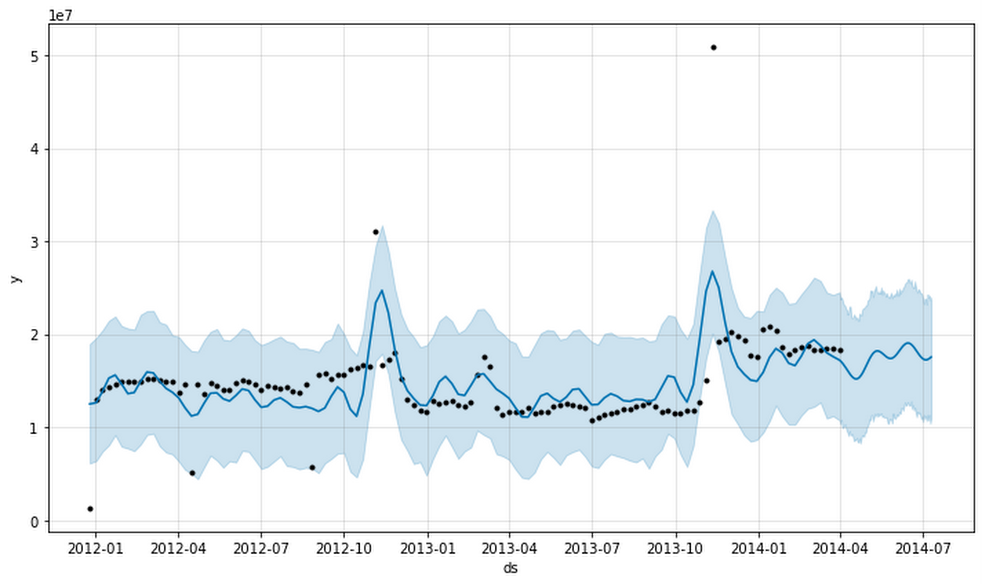
****

**(Rule-M) As in the above Bar plot we can clearly see that in November 2013 there is sudden increase in Data points which also leads to outliers or Anomalies.**

**We also used the concept of Rolling window in order to smoothen the curve below is the image attached**

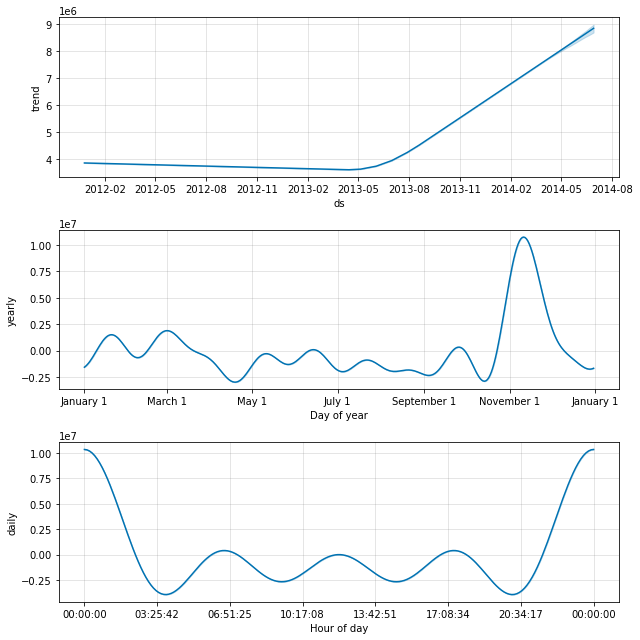
****

**Used FbProphet for training purpose could use other models also since the data was small so I used Fbprophet.**

****

**# BLUE LINE PREDICTS THE PREDICTION OR THE YHAT VALUE**

**# THE BLUE SHADED REGION INDICATES UPPER AND LOWER ESTIMATES OF YHAT VALUE**

****