# PROJECT: COMCAST TELECOM CONSUMER COMPLAINTS

**DESCRIPTION:** Comcast is an American global telecommunication company. The firm has been providing terrible customer service. They continue to fall short despite repeated promises to improve. Only last month (October 2016) the authority fined them a $2.3 million, after receiving over 1000 consumer complaints.  
The existing database will serve as a repository of public customer complaints filed against Comcast.  
It will help to pin down what is wrong with Comcast's customer service.

**Data Dictionary :**

* Ticket #: Ticket number assigned to each complaint
* Customer Complaint: Description of complaint
* Date: Date of complaint
* Time: Time of complaint
* Received Via: Mode of communication of the complaint
* City: Customer city
* State: Customer state
* Zip code: Customer zip
* Status: Status of complaint
* Filing on behalf of someone

**Analysis Task :** To perform these tasks, you can use any of the different Python libraries such as NumPy, SciPy, Pandas, scikit-learn, matplotlib, and BeautifulSoup.

- Import data into Python environment.  
- Provide the trend chart for the number of complaints at monthly and daily granularity levels.  
- Provide a table with the frequency of complaint types.

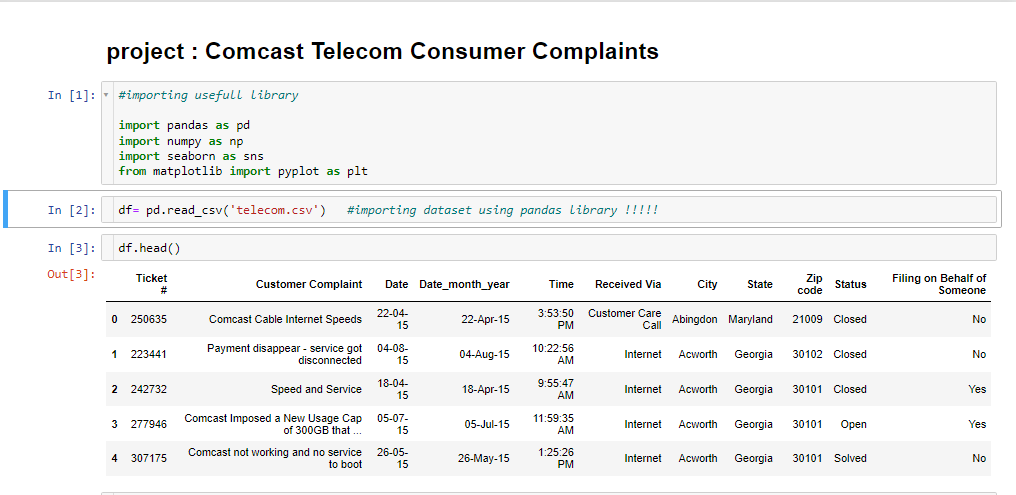
* Which complaint types are maximum i.e., around internet, network issues, or across any other domains.

- Create a new categorical variable with value as **Open**and **Closed**. Open & Pending is to be categorized as Open and Closed & Solved is to be categorized as Closed.  
- Provide state wise status of complaints in a stacked bar chart. Use the categorized variable from Q3. Provide insights on:

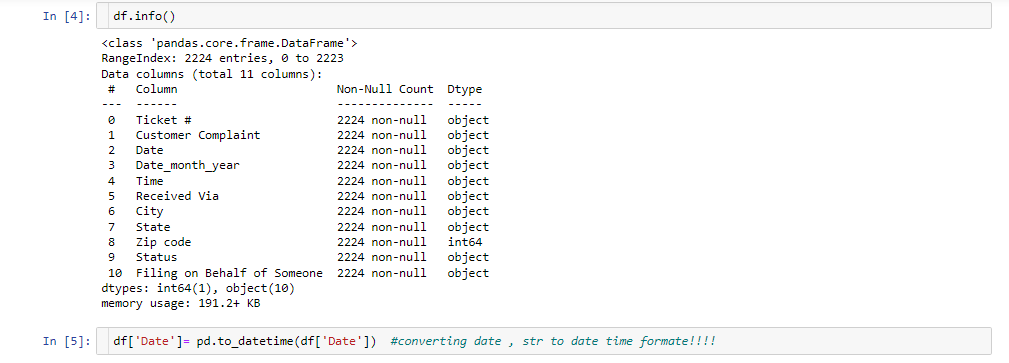
* Which state has the maximum complaints
* Which state has the highest percentage of unresolved complaints

- Provide the percentage of complaints resolved till date, which were received through the Internet and customer care calls.

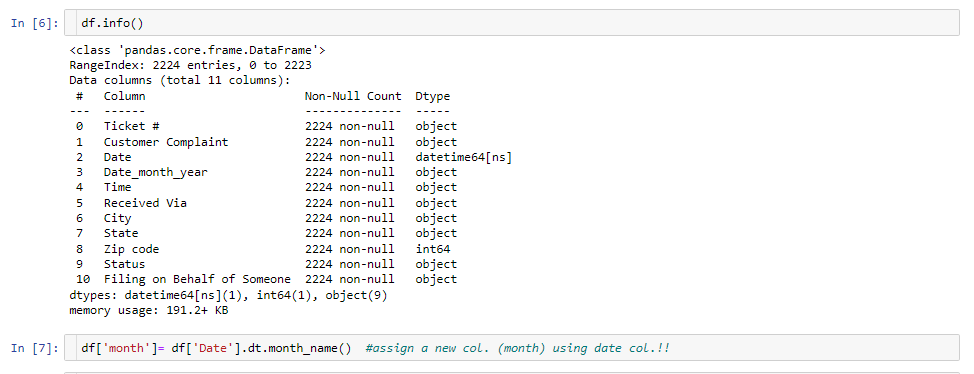
Step 1: Import the relevant data set and read the data and find out the column of the dataset to analyse.



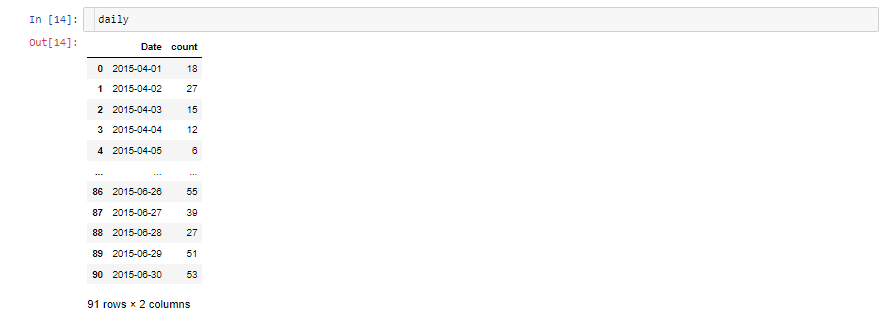
Step 2 : Checking null values in each column , **as we found that there is no null values in cumulus.**

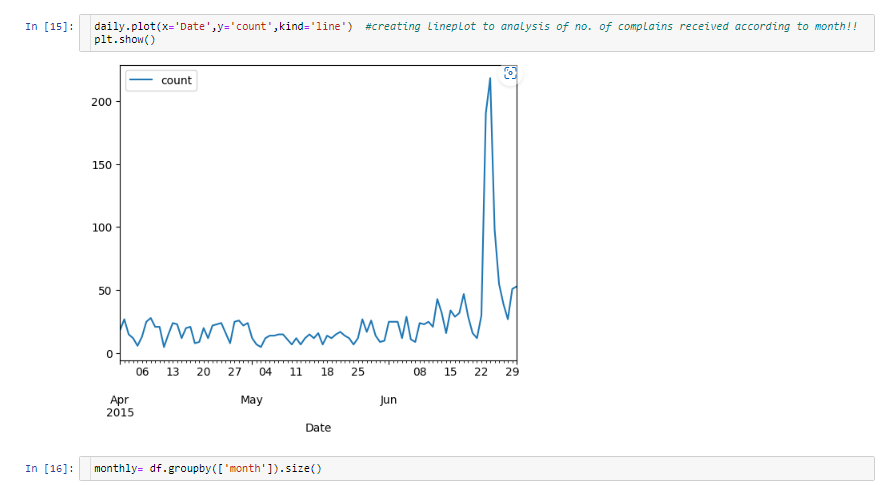
And converts ‘date' (column which is in string format to date column by predefine functions. 

Step 3: Assigning a new column using Date column, for further analysis



Step 4: Assigning a new variable “daily” , in which we found that no of complains daily basis.Graphical user interface, application, Word

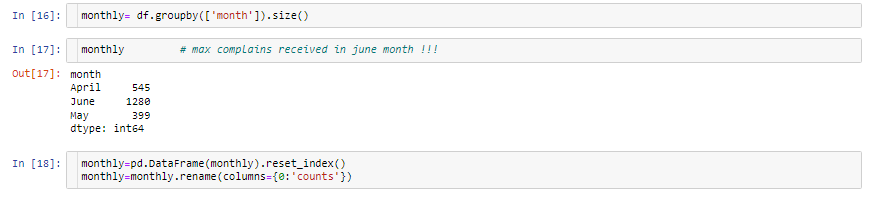
Description automatically generated

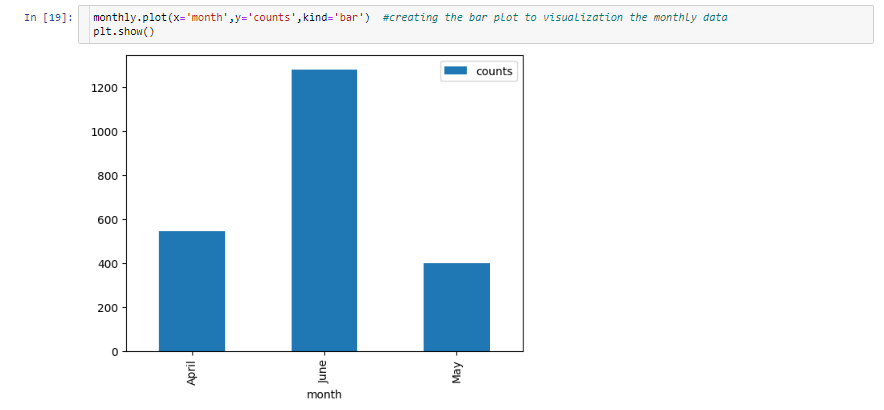


Step 5: in this step by creating a graph in which **we found maximum number of complaints in between 22 to 29 of june , according to daily basis.**

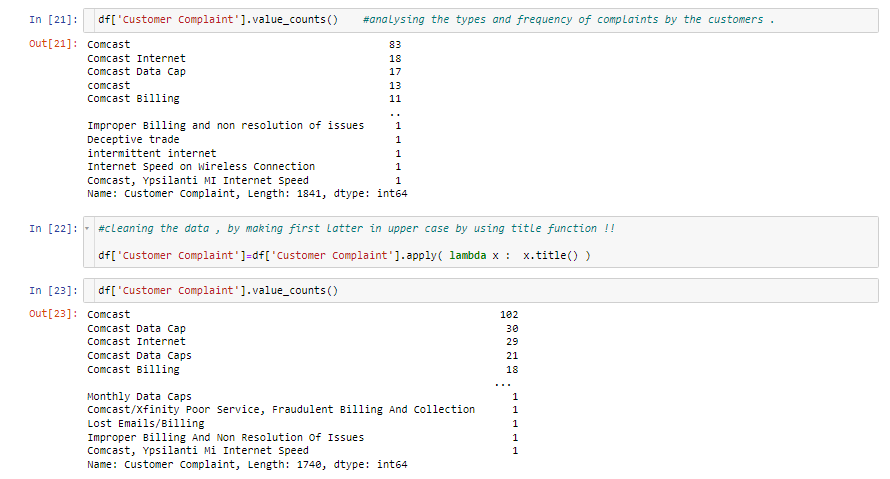
Step 5 : Grouping the data according to month to analysis the maximum no of complains getting ,

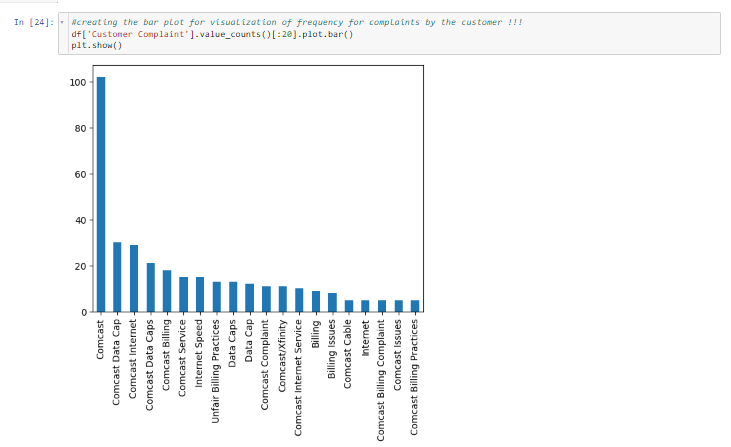
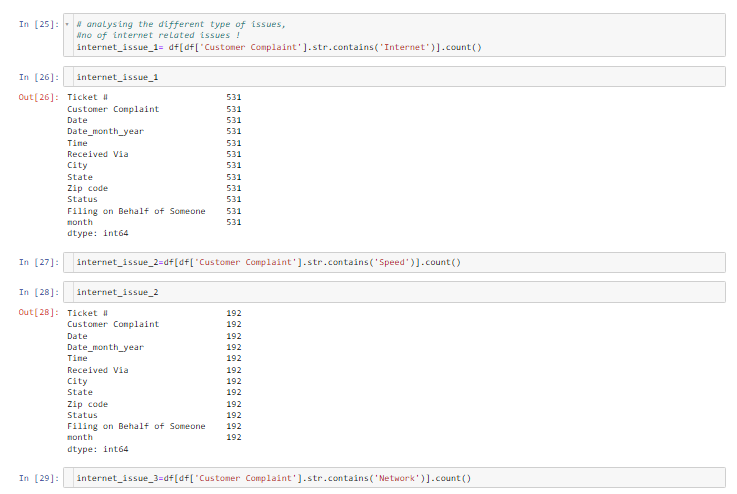
And **found that maximum no of complains in June month, which is 1280**

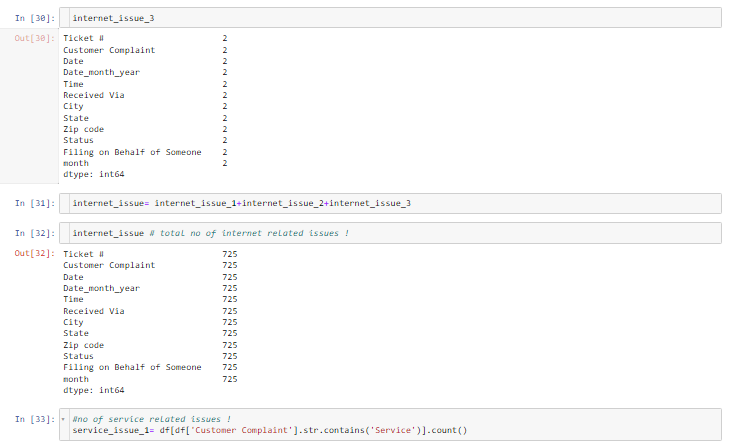




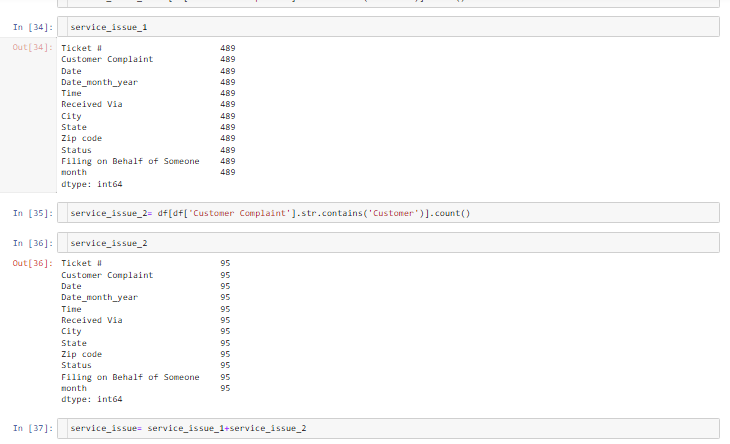
Step 5 : creating a bar chart using one smart according to their complaints,

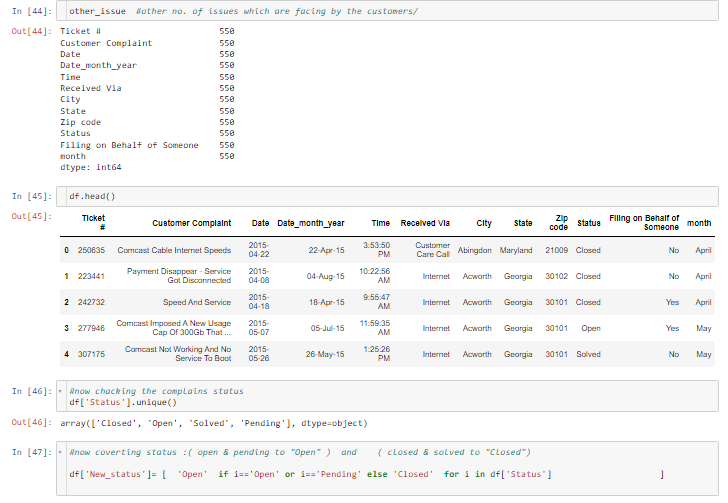
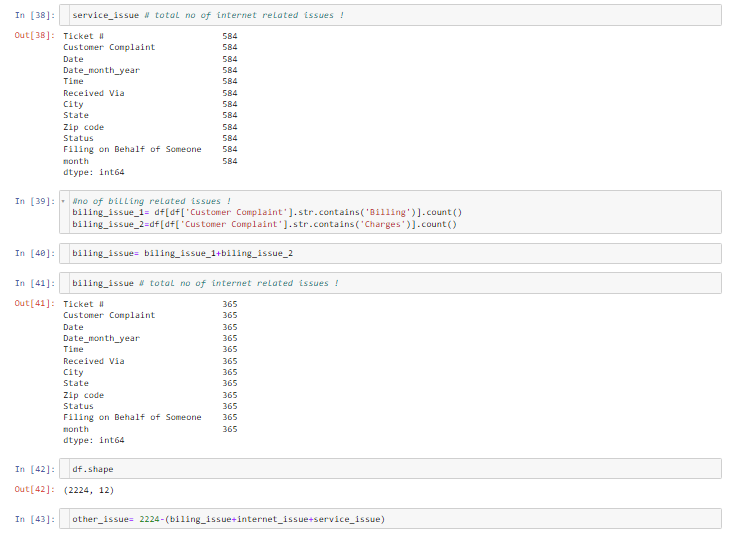
Step 6: creating a different type of variable according to complaints of the customers. Identify the complaint category by using value\_counts function. Using Lambda function to changing the string items upper case in each words.

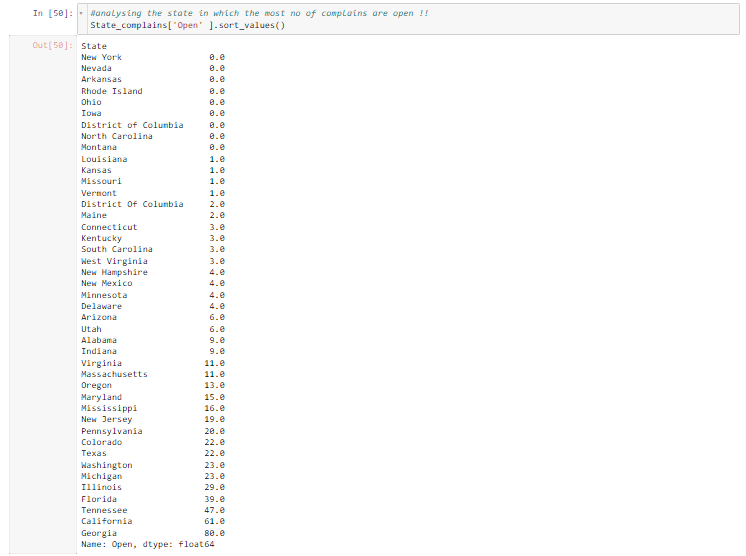
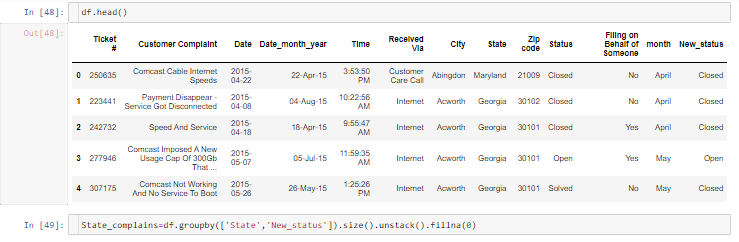
Step 7: For better understanding we would go for bar graph to verify the most category of customer complaint is “comcast” as per the graph.Step 8: Step 13: Identify the complaints by using string contains by keywords.

Step 9: As we find total number of complaints related to Internet issues are 725.

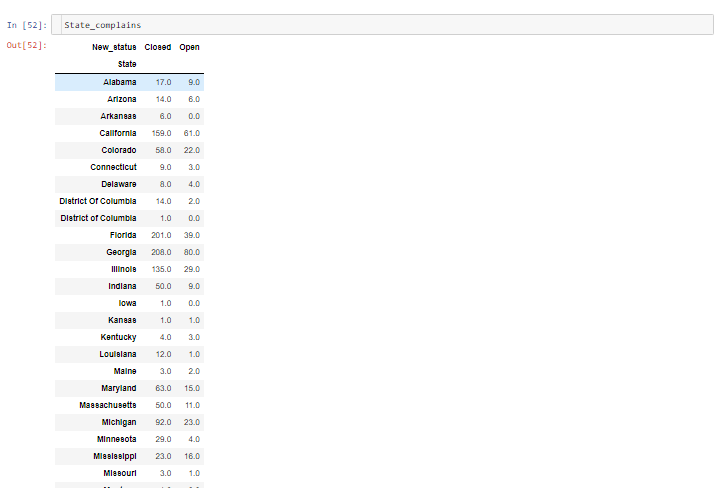
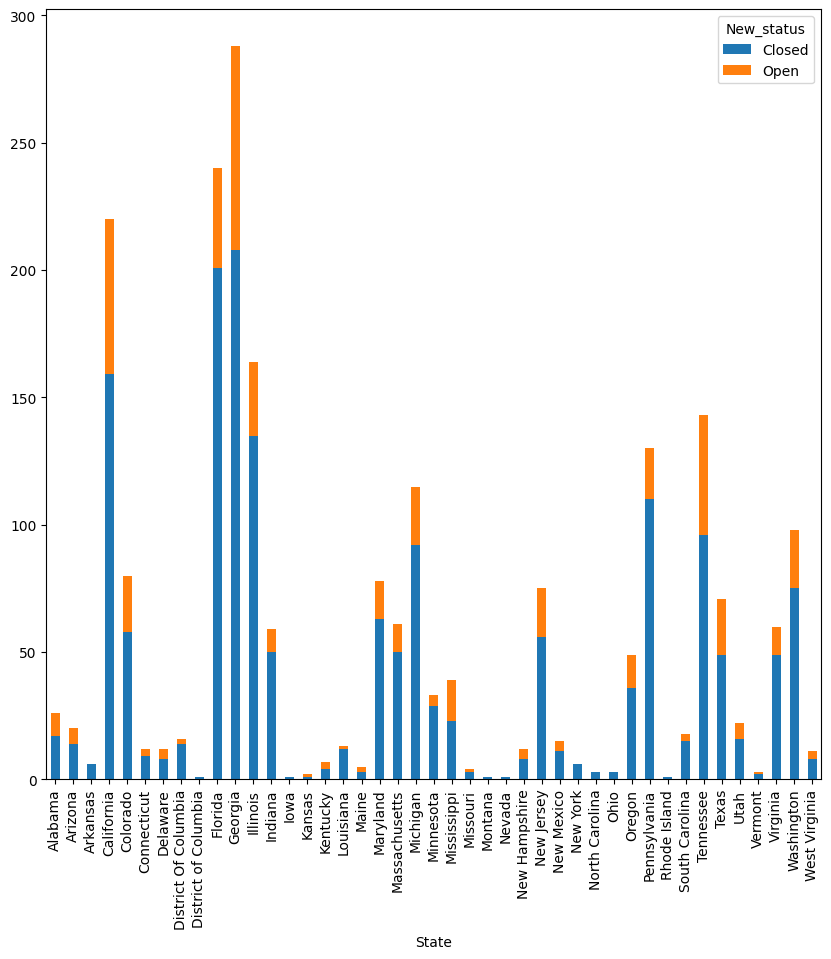
Step 10: now analysing the issues related to customer services. And found to total 584 customer service-related issues.



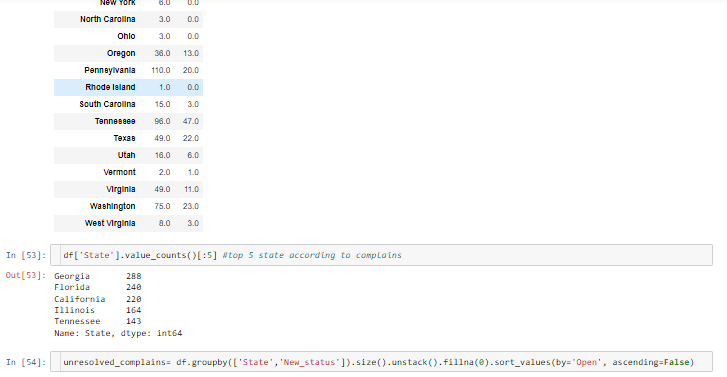
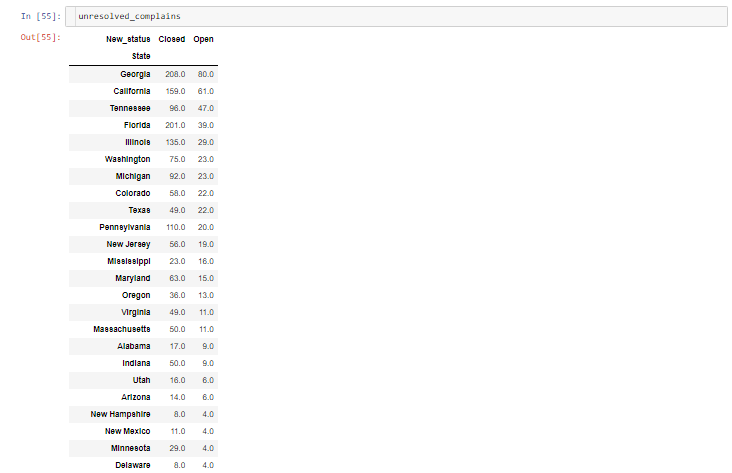
Step 11: Analysing the complaints related to billing issues and found total number of billing issues are 365. Remaining issues are put in other issues category, and total no. of other issues are 550.

Step12: Identify the “status” column unique values. Modifying the status column by using if else condition. Finding the status of complaint by state wise and fill the Nan by using fillna().

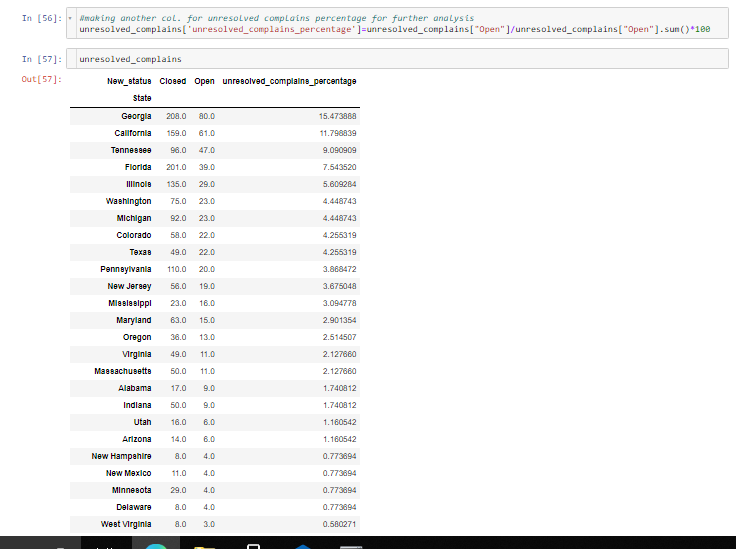
Step 13: Using Bar plot to find the most of the complaint are closed state is Georgia as per the bar graph.



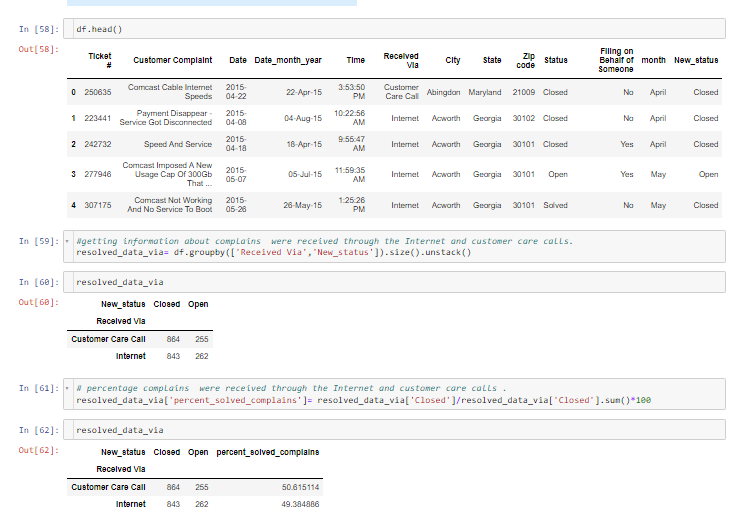
Step 14: finding top 5 states according to complaint status, age we found the Georgia is the most number of values in terms of status.

Step 15: finding number of unresolved complaints according to states and we analyse dad Georgia have most number of open complaints.



Step 16: : By using percentage Georgia has a highest number of unresolved complaints



Step 17 : Step 24: Identifying the resolved status, Grouping the data by using “Received Via” and “New Status” Column to find the percentage and Received Via input.

**Conclusion:**

Therefore, the Total percentage of Complaint have been resolved via “customer care call” and the percentage of resolved complains are **50.615114%.**