Comcast Telecom Customer Complaints

June 15, 2022

1 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
- Zipcode: 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.

Import Libraries

[1]:
```
import numpy as np
import pandas as pd
```

Task1: Import data into Python environment

[2]:
```
com_data=pd.read_csv("Comcast_telecom_complaints_data.csv")
```

[3]:
```
# show data
com_data
```

[3]:

	Ticket #	Customer Complaint	Date	Date_month_year	Time	Received Via	City	State	Zip code	Status	Filing on Behalf of Someone
0	250635	Comcast Cable Internet Speeds	22-04-15	22-Apr-15	3:53:50 PM	Customer Care Call	Abingdon	Maryland	21009	Closed	No
1	223441	Payment disappear - service got disconnected	04-08-15	04-Aug-15	10:22:56 AM	Internet	Acworth	Georgia	30102	Closed	No
2	242732	Speed and Service	18-04-15	18-Apr-15	9:55:47 AM	Internet	Acworth	Georgia	30101	Closed	Yes
3	277946	Comcast Imposed a New Usage Cap of 300GB that …	05-07-15	05-Jul-15	11:59:35 AM	Internet	Acworth	Georgia	30101	Open	Yes
4	307175	Comcast not working and no service to boot	26-05-15	26-May-15	1:25:26 PM	Internet	Acworth	Georgia	30101	Solved	No
…	…	…	…	…	…	…	…	…	…	…	…
2219	213550	Service Availability	04-02-15	04-Feb-15	9:13:18 AM	Customer Care Call	Youngstown	Florida	32466	Closed	No
2220	318775	Comcast Monthly Billing for Returned Modem	06-02-15	06-Feb-15	1:24:39 PM	Customer Care Call	Ypsilanti	Michigan	48197	Solved	No
2221	331188	complaint about comcast	06-09-15	06-Sep-15	5:28:41 PM	Internet	Ypsilanti	Michigan	48197	Solved	No
2222	360489	Extremely unsatisfied Comcast customer	23-06-15	23-Jun-15	11:13:30 PM	Customer Care Call	Ypsilanti	Michigan			
2223	363614	Comcast, Ypsilanti MI Internet Speed	24-06-15	24-Jun-15	10:28:33 PM	Customer Care Call	Ypsilanti	Michigan			

```
2222     48197  Solved                            No
2223     48198    Open                           Yes

[2224 rows x 11 columns]
```

[4]:
```
# 5 five entries of the data
com_data.head()
```

[4]:
```
  Ticket #                                  Customer Complaint      Date  \
0   250635                       Comcast Cable Internet Speeds  22-04-15
1   223441       Payment disappear - service got disconnected  04-08-15
2   242732                                    Speed and Service  18-04-15
3   277946  Comcast Imposed a New Usage Cap of 300GB that …  05-07-15
4   307175         Comcast not working and no service to boot  26-05-15

  Date_month_year         Time        Received Via      City     State  \
0       22-Apr-15   3:53:50 PM  Customer Care Call  Abingdon  Maryland
1       04-Aug-15  10:22:56 AM            Internet   Acworth   Georgia
2       18-Apr-15   9:55:47 AM            Internet   Acworth   Georgia
3       05-Jul-15  11:59:35 AM            Internet   Acworth   Georgia
4       26-May-15   1:25:26 PM            Internet   Acworth   Georgia

   Zip code  Status Filing on Behalf of Someone
0     21009  Closed                           No
1     30102  Closed                           No
2     30101  Closed                          Yes
3     30101    Open                          Yes
4     30101  Solved                           No
```

[5]:
```
# data shape
com_data.shape
```

[5]:
```
(2224, 11)
```

[6]:
```
# data info
com_data.info()
```

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 2224 entries, 0 to 2223
Data columns (total 11 columns):
 #   Column                      Non-Null Count  Dtype
---  ------                      --------------  -----
 0   Ticket #                    2224 non-null   object
 1   Customer Complaint          2224 non-null   object
 2   Date                        2224 non-null   object
 3   Date_month_year             2224 non-null   object
 4   Time                        2224 non-null   object
```

```
 5   Received Via                  2224 non-null   object
 6   City                          2224 non-null   object
 7   State                         2224 non-null   object
 8   Zip code                      2224 non-null   int64
 9   Status                        2224 non-null   object
 10  Filing on Behalf of Someone   2224 non-null   object
dtypes: int64(1), object(10)
memory usage: 191.2+ KB
```

[7]:
```
# column names
com_data.columns
```

[7]:
```
Index(['Ticket #', 'Customer Complaint', 'Date', 'Date_month_year', 'Time',
       'Received Via', 'City', 'State', 'Zip code', 'Status',
       'Filing on Behalf of Someone'],
      dtype='object')
```

[8]:
```
# check for null values
com_data.isnull().sum()
```

[8]:
```
Ticket #                      0
Customer Complaint            0
Date                          0
Date_month_year               0
Time                          0
Received Via                  0
City                          0
State                         0
Zip code                      0
Status                        0
Filing on Behalf of Someone   0
dtype: int64
```

Task 2: Provide the trend chart for the number of complaints at monthly and daily granularity levels

Import libraries for graphs

[9]:
```
import matplotlib.pyplot as plt
import seaborn as sns
```

[10]:
```
# convert 'Date_month_year' to datetime format
com_data['Date_month_year'] = pd.to_datetime(com_data['Date_month_year'])
com_data['Month'] =  com_data['Date_month_year'].apply(lambda x: x.month)
com_data['Day'] = com_data['Date_month_year'].apply(lambda x: x.day)
com_data['Day of Week'] = com_data['Date_month_year'].apply(lambda x: x.
 ↪dayofweek)
```

```
com_data['Day of Week']=com_data['Day of Week'].map({0:'Mon',1:'Tue',2:'Wed',3:
↪'Thur',4:'Fri',5:'Sat',6:'Sun'})
```

[11]:
```
com_data
```

[11]:

	Ticket #	Customer Complaint	Date	Date_month_year	Time	Received Via	City	State	Zip code	Status	Filing on Behalf of Someone	Month	Day	Day of Week
0	250635	Comcast Cable Internet Speeds	22-04-15	2015-04-22	3:53:50 PM	Customer Care Call	Abingdon	Maryland	21009	Closed	No	4	22	Wed
1	223441	Payment disappear - service got disconnected	04-08-15	2015-08-04	10:22:56 AM	Internet	Acworth	Georgia	30102	Closed	No	8	4	Tue
2	242732	Speed and Service	18-04-15	2015-04-18	9:55:47 AM	Internet	Acworth	Georgia	30101	Closed	Yes	4	18	Sat
3	277946	Comcast Imposed a New Usage Cap of 300GB that …	05-07-15	2015-07-05	11:59:35 AM	Internet	Acworth	Georgia	30101	Open	Yes	7	5	Sun
4	307175	Comcast not working and no service to boot	26-05-15	2015-05-26	1:25:26 PM	Internet	Acworth	Georgia	30101	Solved	No	5	26	Tue
…	…	…	…	…	…	…	…	…	…	…	…	…	…	…
2219	213550	Service Availability	04-02-15	2015-02-04	9:13:18 AM	Customer Care Call	Youngstown	Florida	32466	Closed	No	2	4	Wed
2220	318775	Comcast Monthly Billing for Returned Modem	06-02-15	2015-02-06	1:24:39 PM	Customer Care Call	Ypsilanti	Michigan	48197	Solved	No	2	6	Fri
2221	331188	complaint about comcast	06-09-15	2015-09-06	5:28:41 PM	Internet	Ypsilanti	Michigan	48197	Solved	No	9	6	Sun
2222	360489	Extremely unsatisfied Comcast customer	23-06-15	2015-06-23	11:13:30 PM	Customer Care Call	Ypsilanti	Michigan	48197	Solved	No	6	23	Tue
2223	363614	Comcast, Ypsilanti MI Internet Speed	24-06-15	2015-06-24	10:28:33 PM	Customer Care Call	Ypsilanti	Michigan	48198	Open	Yes	6	24	Wed

[2224 rows x 14 columns]

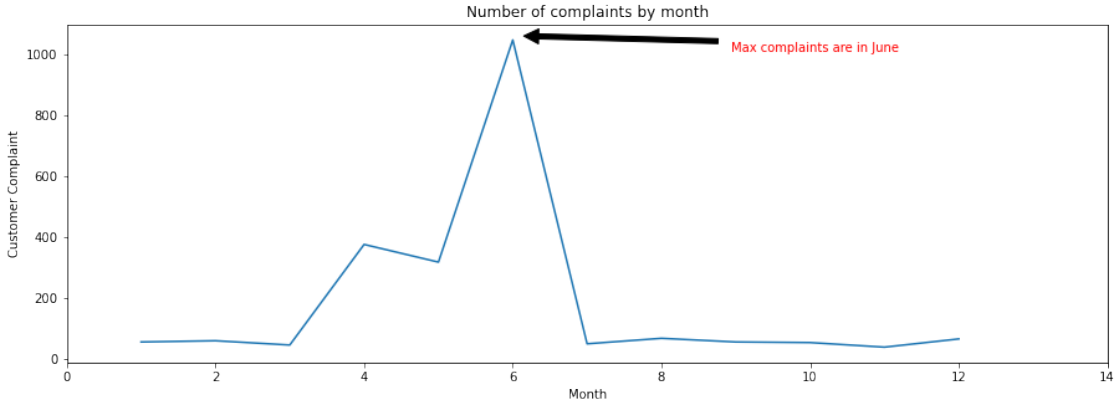
Provide the trend chart for the number of complaints at monthly granularity level

[12]:
```
# find the month with the most complaints
month = com_data.groupby('Month')["Ticket #"].count()
print("The month with the most complaints was : ", month.idxmax()," with␣
 ↪",month.max())
# months: 1:'January',2:'February',3:'March',4:'April',5:'May',6:'June',7:
 ↪'July',8:'August',9:'September',10:'October',11:'November',12:'December'
```

```
The month with the most complaints was :  6  with  1046
```

[13]:
```
# trend chart for monthly complaints
plt.figure(figsize=(15,5))
month = com_data.groupby('Month').count().reset_index()
lp = sns.lineplot(x='Month', y= 'Customer Complaint', data = month)
plt.title('Number of complaints by month')
ax = lp.axes
ax.set_xlim(0,14)
ax.annotate('Max complaints are in June', color='red',
            xy=(6, 1060), xycoords='data',
            xytext=(0.8, 0.95), textcoords='axes fraction',
            arrowprops=dict(facecolor='black', shrink=0.05),
            horizontalalignment='right', verticalalignment='top')
```

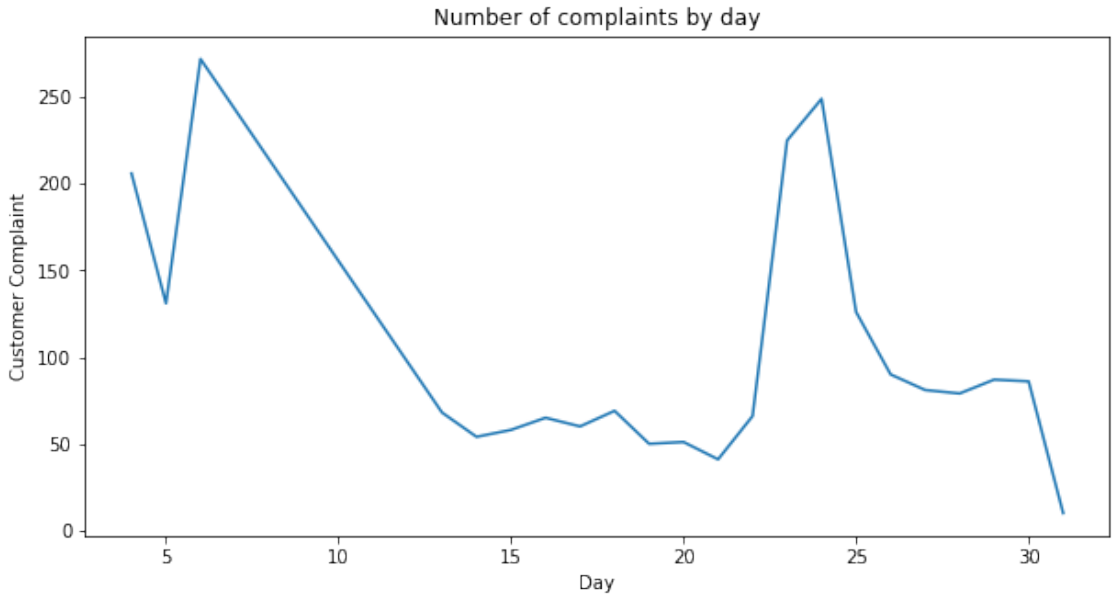
[13]:
```
Text(0.8, 0.95, 'Max complaints are in June')
```



Provide the trend chart for the number of complaints at daily granularity level

[14]:
```
plt.figure(figsize=(10,5))
day = com_data.groupby('Day').count().reset_index()
lp = sns.lineplot(x='Day', y= 'Customer Complaint', data = day)
plt.title('Number of complaints by day')
ax.set_xlim(0,35)
```

[14]:
```
(0.0, 35.0)
```



Task 3: Provide a table with the frequency of complaint types

[15]:
```
ct_freq = com_data['Customer Complaint'].value_counts().to_frame().reset_index()
ct_freq
```

[15]:

	index	Customer Complaint
0	Comcast	83
1	Comcast Internet	18
2	Comcast Data Cap	17
3	comcast	13
4	Comcast Data Caps	11
…	…	…
1836	double billing after change of service	1
1837	Un able to access my email address after reloc…	1
1838	Comcast Corporation	1
1839	Complaint against Comcast Corporation	1
1840	Comcast high prices and throttling speeds	1

[1841 rows x 2 columns]

Task 4: Find which complaint types are maximum i.e., around internet, network issues, or across any other domains

[16]:
```
com_data['Customer Complaint'].value_counts().head(5)
```

[16]:
```
Comcast              83
Comcast Internet     18
Comcast Data Cap     17
```

```
comcast               13
Comcast Data Caps     11
Name: Customer Complaint, dtype: int64
```

Task 5: 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.

[17]:
```
com_data['Final_Status'] = ["Open" if Status=="Open" or Status=="Pending" else␣
↪"Closed" for Status in com_data["Status"]]
```

[18]:
```
com_data.head()
```

[18]:

	Ticket #	Customer Complaint	Date	Date_month_year	Time	Received Via	City	State	Zip code	Status	Filing on Behalf of Someone	Month	Day	Day of Week	Final_Status
0	250635	Comcast Cable Internet Speeds	22-04-15	2015-04-22	3:53:50 PM	Customer Care Call	Abingdon	Maryland	21009	Closed	No	4	22	Wed	Closed
1	223441	Payment disappear - service got disconnected	04-08-15	2015-08-04	10:22:56 AM	Internet	Acworth	Georgia	30102	Closed	No	8	4	Tue	Closed
2	242732	Speed and Service	18-04-15	2015-04-18	9:55:47 AM	Internet	Acworth	Georgia	30101	Closed	Yes	4	18	Sat	Closed
3	277946	Comcast Imposed a New Usage Cap of 300GB that …	05-07-15	2015-07-05	11:59:35 AM	Internet	Acworth	Georgia	30101	Open	Yes	7	5	Sun	Open
4	307175	Comcast not working and no service to boot	26-05-15	2015-05-26	1:25:26 PM	Internet	Acworth	Georgia	30101	Solved	No	5	26	Tue	Closed

Task 6: 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

Check how many closed and open tickets there are by state

[19]:
```
# final ticket status by state
com_data_state_status = pd.crosstab(com_data["State"],com_data["Final_Status"])
```

```
com_data_state_status
```

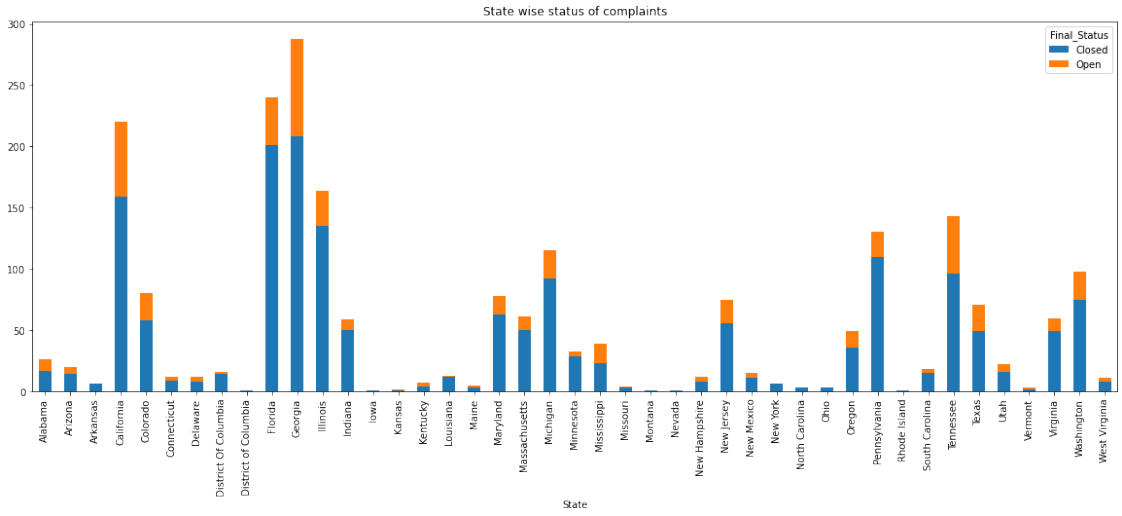
[19]:

State	Closed	Open
Alabama	17	9
Arizona	14	6
Arkansas	6	0
California	159	61
Colorado	58	22
Connecticut	9	3
Delaware	8	4
District Of Columbia	14	2
District of Columbia	1	0
Florida	201	39
Georgia	208	80
Illinois	135	29
Indiana	50	9
Iowa	1	0
Kansas	1	1
Kentucky	4	3
Louisiana	12	1
Maine	3	2
Maryland	63	15
Massachusetts	50	11
Michigan	92	23
Minnesota	29	4
Mississippi	23	16
Missouri	3	1
Montana	1	0
Nevada	1	0
New Hampshire	8	4
New Jersey	56	19
New Mexico	11	4
New York	6	0
North Carolina	3	0
Ohio	3	0
Oregon	36	13
Pennsylvania	110	20
Rhode Island	1	0
South Carolina	15	3
Tennessee	96	47
Texas	49	22
Utah	16	6
Vermont	2	1
Virginia	49	11
Washington	75	23
West Virginia	8	3

Find the state with the most complaints

[20]:
```
# state wise status of complaints in a stacked bar chart
pd.crosstab(com_data["State"],com_data["Final_Status"]).
 ↪plot(kind='bar',figsize=(20,7),stacked=True)
plt.title('State wise status of complaints')
```

[20]:
```
Text(0.5, 1.0, 'State wise status of complaints')
```



[21]:
```
com_data.groupby(["State"]).size().sort_values(ascending=False).to_frame().
 ↪rename({0: "Complaints"}, axis=1).head(1)
```

[21]:
```
         Complaints
State
Georgia         288
```

As we see both from the stacked plot bar and the code above, Georgia is the state with the most complaints.

Find the state with the highest percentage of unresolved complaints

[22]:
```
State_Unsolved=com_data.loc[com_data['Final_Status']=='Open',['State']].
 ↪value_counts()
State_Unsolved.head(1)/State_Unsolved.sum()*100
```

[22]:
```
State
Georgia    15.473888
dtype: float64
```

Task 7: Provide the percentage of complaints resolved till date, which were received through the Internet and customer care calls.

[23]:
```
com_data[com_data['Final_Status']=='Closed'].groupby('Final_Status')['Received␣
↪Via'].value_counts(normalize=True)*100
```

[23]:
```
Final_Status  Received Via
Closed        Customer Care Call    50.615114
              Internet              49.384886
Name: Received Via, dtype: float64
```