Comcast_Telecom

November 27, 2021

1 Import data into Python environment.

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

[2]:
```
df=pd.read_csv('Comcast_telecom_complaints_data.csv')
```

[3]:
```
df
```

[3]:
```
     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
…       …                                                …       …
2219   213550                               Service Availability  04-02-15
2220   318775         Comcast Monthly Billing for Returned Modem  06-02-15
2221   331188                            complaint about comcast  06-09-15
2222   360489             Extremely unsatisfied Comcast customer  23-06-15
2223   363614               Comcast, Ypsilanti MI Internet Speed  24-06-15

     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
…              …          …                 …         …       …
2219       04-Feb-15   9:13:18 AM  Customer Care Call  Youngstown   Florida
2220       06-Feb-15   1:24:39 PM  Customer Care Call   Ypsilanti  Michigan
2221       06-Sep-15   5:28:41 PM            Internet   Ypsilanti  Michigan
2222       23-Jun-15  11:13:30 PM  Customer Care Call   Ypsilanti  Michigan
2223       24-Jun-15  10:28:33 PM  Customer Care Call   Ypsilanti  Michigan

      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
…        …      …                         …
2219     32466  Closed                           No
2220     48197  Solved                           No
2221     48197  Solved                           No
2222     48197  Solved                           No
2223     48198    Open                          Yes

[2224 rows x 11 columns]
```

[4]:

```
df.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
```

[5]:

```
df['Date']=pd.to_datetime(df['Date']) #convereting the date column into date time object as it was only object
```

[6]:

```
df.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   datetime64[ns]
 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: datetime64[ns](1), int64(1), object(9)
memory usage: 191.2+ KB
```

2 - Provide the trend chart for the number of complaints at monthly and daily granularity levels.

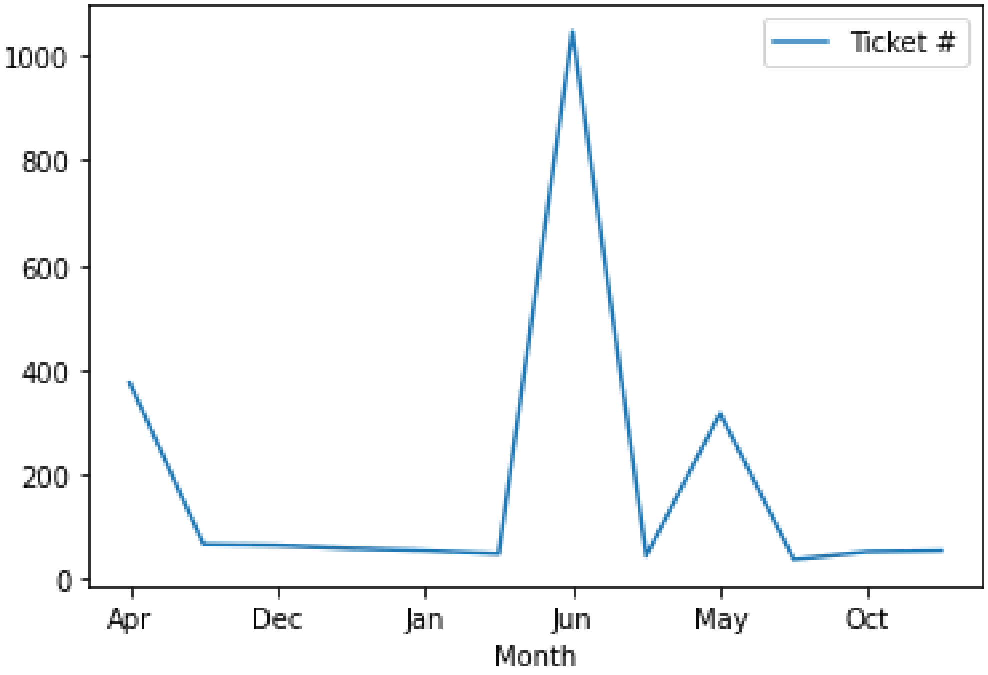
[7]:
```python
def month(x):
    month = x.split('-')[1]
    return month
```

[8]:
```python
df['Month'] = df['Date_month_year'].apply(month)
```

[9]:
```python
df.groupby('Month').count()[['Ticket #']].plot()
```

```
/usr/local/lib/python3.7/site-packages/pandas/plotting/_matplotlib/core.py:1192:
UserWarning: FixedFormatter should only be used together with FixedLocator
  ax.set_xticklabels(xticklabels)
```

[9]:
```
<AxesSubplot:xlabel='Month'>
```



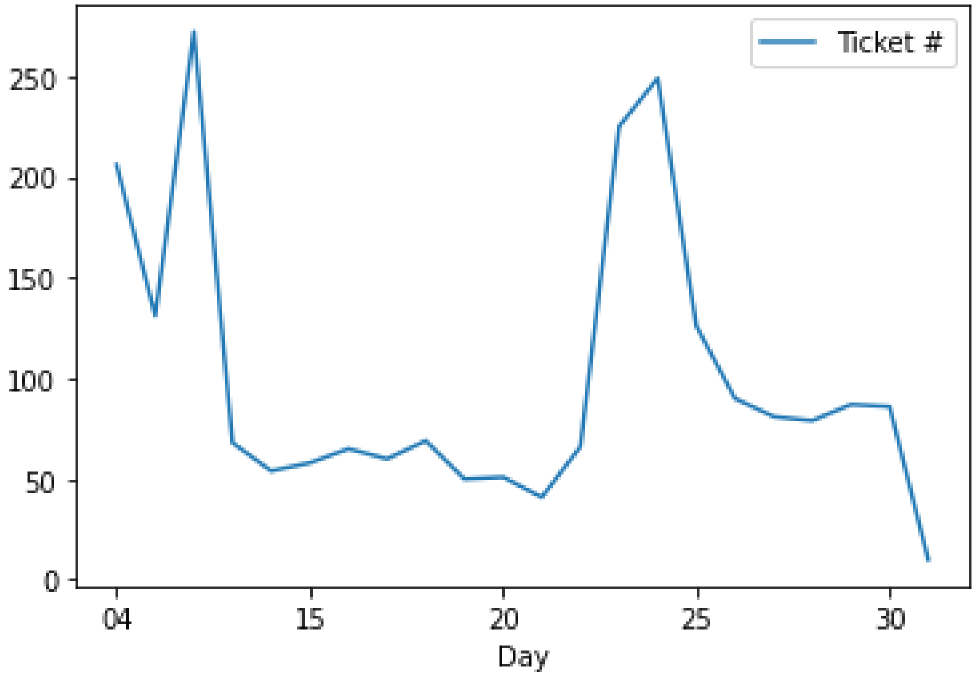
[10]:
```
def month(x):
    month = x.split('-')[0]
    return month
```

[11]:
```
df['Day'] = df['Date_month_year'].apply(month)
```

[12]:
```
df.groupby('Day').count()[['Ticket #']].plot()
```

```
/usr/local/lib/python3.7/site-packages/pandas/plotting/_matplotlib/core.py:1192:
UserWarning: FixedFormatter should only be used together with FixedLocator
  ax.set_xticklabels(xticklabels)
```

[12]:
```
<AxesSubplot:xlabel='Day'>
```



3 - Provide a table with the frequency of complaint types.

```
[13]: df['Customer Complaint'].value_counts()
```

```
[13]: Comcast                                                           83
      Comcast Internet                                                  18
      Comcast Data Cap                                                  17
      comcast                                                           13
      Data Caps                                                         11
                                                                        ..
      Comcast bundling of internet and tv services and deceptive sales practices     1
      Comcast refuses to recognize our same sex marriage                 1
      Comcast speed and billing                                          1
      Internet being slowed.                                             1
      Modem rental fees                                                  1
      Name: Customer Complaint, Length: 1841, dtype: int64
```

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

```
[14]: df['Received Via'].value_counts()
```

```
[14]: Customer Care Call    1119
      Internet              1105
      Name: Received Via, dtype: int64
```

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.

```
[15]: df['Status'].unique()
```

```
[15]: array(['Closed', 'Open', 'Solved', 'Pending'], dtype=object)
```

```
[16]: def status(x):
          if x in['Open' , 'Pending']:
              return 'Open'
          if x in['Closed' , 'Solved']:
              return 'Closed'
```

```
[17]: df['New_Status'] = df['Status'].apply(status)
```

```
[18]: df  #column for new status appeared with open and close values
```

```
[18]:      Ticket #                                 Customer Complaint       Date  \
      0      250635                      Comcast Cable Internet Speeds 2015-04-22
      1      223441       Payment disappear - service got disconnected 2015-04-08
      2      242732                                  Speed and Service 2015-04-18
      3      277946  Comcast Imposed a New Usage Cap of 300GB that … 2015-05-07
      4      307175         Comcast not working and no service to boot 2015-05-26
      …       …                                                …        …
      2219   213550                               Service Availability 2015-04-02
      2220   318775         Comcast Monthly Billing for Returned Modem 2015-06-02
      2221   331188                            complaint about comcast 2015-06-09
      2222   360489             Extremely unsatisfied Comcast customer 2015-06-23
      2223   363614               Comcast, Ypsilanti MI Internet Speed 2015-06-24

           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
…              …          …                     …         …         …
2219       04-Feb-15   9:13:18 AM  Customer Care Call  Youngstown    Florida
2220       06-Feb-15   1:24:39 PM  Customer Care Call   Ypsilanti   Michigan
2221       06-Sep-15   5:28:41 PM            Internet   Ypsilanti   Michigan
2222       23-Jun-15  11:13:30 PM  Customer Care Call   Ypsilanti   Michigan
2223       24-Jun-15  10:28:33 PM  Customer Care Call   Ypsilanti   Michigan

      Zip code  Status Filing on Behalf of Someone Month Day New_Status
0        21009  Closed                          No   Apr  22     Closed
1        30102  Closed                          No   Aug  04     Closed
2        30101  Closed                         Yes   Apr  18     Closed
3        30101    Open                         Yes   Jul  05       Open
4        30101  Solved                          No   May  26     Closed
…          …      …                      …   …   ..         …
2219     32466  Closed                          No   Feb  04     Closed
2220     48197  Solved                          No   Feb  06     Closed
2221     48197  Solved                          No   Sep  06     Closed
2222     48197  Solved                          No   Jun  23     Closed
2223     48198    Open                         Yes   Jun  24       Open

[2224 rows x 14 columns]
```

6 - Provide state wise status of complaints in a stacked bar chart. Use the categorized variable from Q3. Provide insights on:

[20]:

```
lm = df[['State' , 'Ticket #' , 'New_Status']].groupby(['State' ,␣
↪'New_Status']).count().sort_values('Ticket #')
```

[21]:

```
lm
```

[21]:

```
                                 Ticket #
State                New_Status
Nevada               Closed             1
District of Columbia Closed             1
Montana              Closed             1
Iowa                 Closed             1
Kansas               Closed             1
…                                    …
Pennsylvania         Closed           110
Illinois             Closed           135
California           Closed           159
Florida              Closed           201
Georgia              Closed           208
```

```
[77 rows x 1 columns]
```

```
[22]: lm.reset_index()
```

```
[22]:                  State New_Status  Ticket #
      0               Nevada     Closed         1
      1   District of Columbia     Closed         1
      2              Montana     Closed         1
      3                 Iowa     Closed         1
      4               Kansas     Closed         1
      ..                 …          …        …
      72        Pennsylvania     Closed       110
      73            Illinois     Closed       135
      74          California     Closed       159
      75             Florida     Closed       201
      76             Georgia     Closed       208

      [77 rows x 3 columns]
```

```
[23]: lm = lm.pivot_table(index = 'State' , columns = 'New_Status' , values = 'Ticket␣
       ↪#').fillna(0).reset_index()
```

```
[24]: lm
```

```
[24]: New_Status                  State  Closed  Open
      0                         Alabama    17.0   9.0
      1                         Arizona    14.0   6.0
      2                        Arkansas     6.0   0.0
      3                      California   159.0  61.0
      4                        Colorado    58.0  22.0
      5                     Connecticut     9.0   3.0
      6                        Delaware     8.0   4.0
      7            District Of Columbia    14.0   2.0
      8            District of Columbia     1.0   0.0
      9                         Florida   201.0  39.0
      10                        Georgia   208.0  80.0
      11                       Illinois   135.0  29.0
      12                        Indiana    50.0   9.0
      13                           Iowa     1.0   0.0
      14                         Kansas     1.0   1.0
      15                       Kentucky     4.0   3.0
      16                      Louisiana    12.0   1.0
      17                          Maine     3.0   2.0
      18                       Maryland    63.0  15.0
      19                  Massachusetts    50.0  11.0
      20                       Michigan    92.0  23.0
```

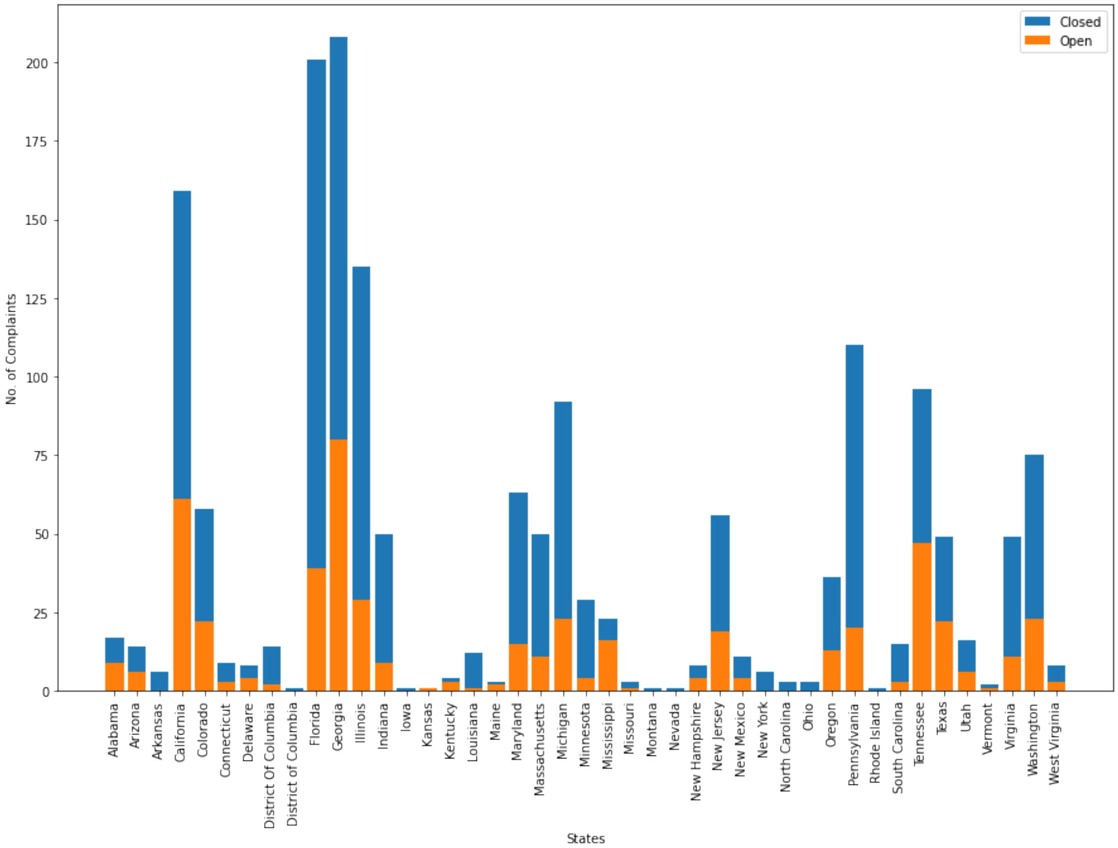
```
21                    Minnesota    29.0   4.0
22                  Mississippi    23.0  16.0
23                     Missouri     3.0   1.0
24                      Montana     1.0   0.0
25                       Nevada     1.0   0.0
26                New Hampshire     8.0   4.0
27                   New Jersey    56.0  19.0
28                   New Mexico    11.0   4.0
29                     New York     6.0   0.0
30               North Carolina     3.0   0.0
31                         Ohio     3.0   0.0
32                       Oregon    36.0  13.0
33                 Pennsylvania   110.0  20.0
34                 Rhode Island     1.0   0.0
35               South Carolina    15.0   3.0
36                    Tennessee    96.0  47.0
37                        Texas    49.0  22.0
38                         Utah    16.0   6.0
39                      Vermont     2.0   1.0
40                     Virginia    49.0  11.0
41                   Washington    75.0  23.0
42                West Virginia     8.0   3.0
```

[25]:

```
import matplotlib.pyplot as plt
```

[26]:

```
fig = plt.figure(figsize = (15 , 10))
plt.bar(lm['State'] , lm['Closed'] , label = 'Closed')
plt.bar(lm['State'] , lm['Open'] , label = 'Open')
plt.xticks(rotation = 90)
plt.legend()
plt.xlabel('States')
plt.ylabel('No. of Complaints')
plt.show()
```



7 Which state has the maximum complaints

[28]:
```
column = lm['Open']
max_value = column.max()
print(max_value)
```

```
80.0
```

8 Which state has the highest percentage of unresolved complaints

[29]:
```
df['Open/Closed'] = df['Status']
```

[31]:
```
sc = df.groupby(['State', 'Open/Closed'])['Open/Closed'].count().unstack().
 ↪fillna(0)
sc = pd.DataFrame(sc)
sc.head()
```

```
[31]: Open/Closed  Closed  Open  Pending  Solved
      State
      Alabama         8.0   4.0      5.0     9.0
      Arizona         6.0   2.0      4.0     8.0
      Arkansas        1.0   0.0      0.0     5.0
      California     72.0  47.0     14.0    87.0
      Colorado       23.0  12.0     10.0    35.0
```

```
[32]: sc['unresolved complaint'] = (sc['Open']/sc['Closed']*100)
```

```
[33]: sc.head().max()
```

```
[33]: Open/Closed
      Closed                  72.000000
      Open                    47.000000
      Pending                 14.000000
      Solved                  87.000000
      unresolved complaint    65.277778
      dtype: float64
```

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

```
[41]: resolved_complaint = df.groupby(['Received Via','Open/Closed'])['Received Via'].
       ↪count().unstack().fillna(0)
```

```
[42]: resolved_complaint
```

```
[42]: Open/Closed         Closed  Open  Pending  Solved
      Received Via
      Customer Care Call     387   186       69     477
      Internet               347   177       85     496
```

```
[43]: t_resolved_complaint = len(df)
      t_resolved_complaint
```

```
[43]: 2224
```

```
[44]: for x in resolved_complaint.values:
          l=(x/t_resolved_complaint*100)
          print(l)
```

```
[17.40107914  8.36330935  3.10251799 21.44784173]
[15.60251799  7.95863309  3.82194245 22.30215827]
```

9.1 Thank You

```
[ ]:
```