

Project: Comcast Telecom Consumer Complaints

Report and Source Code:

Import the dataset 'Comcast_telecom_complaints_data.csv' from the location it saved in your computer

```
import pandas as pd
import matplotlib.pyplot as plt
df = pd.read_csv("Comcast_telecom_complaints_data.csv")
df
dfCopy = df.copy(deep=True)
```

Monthly and Daily Trends

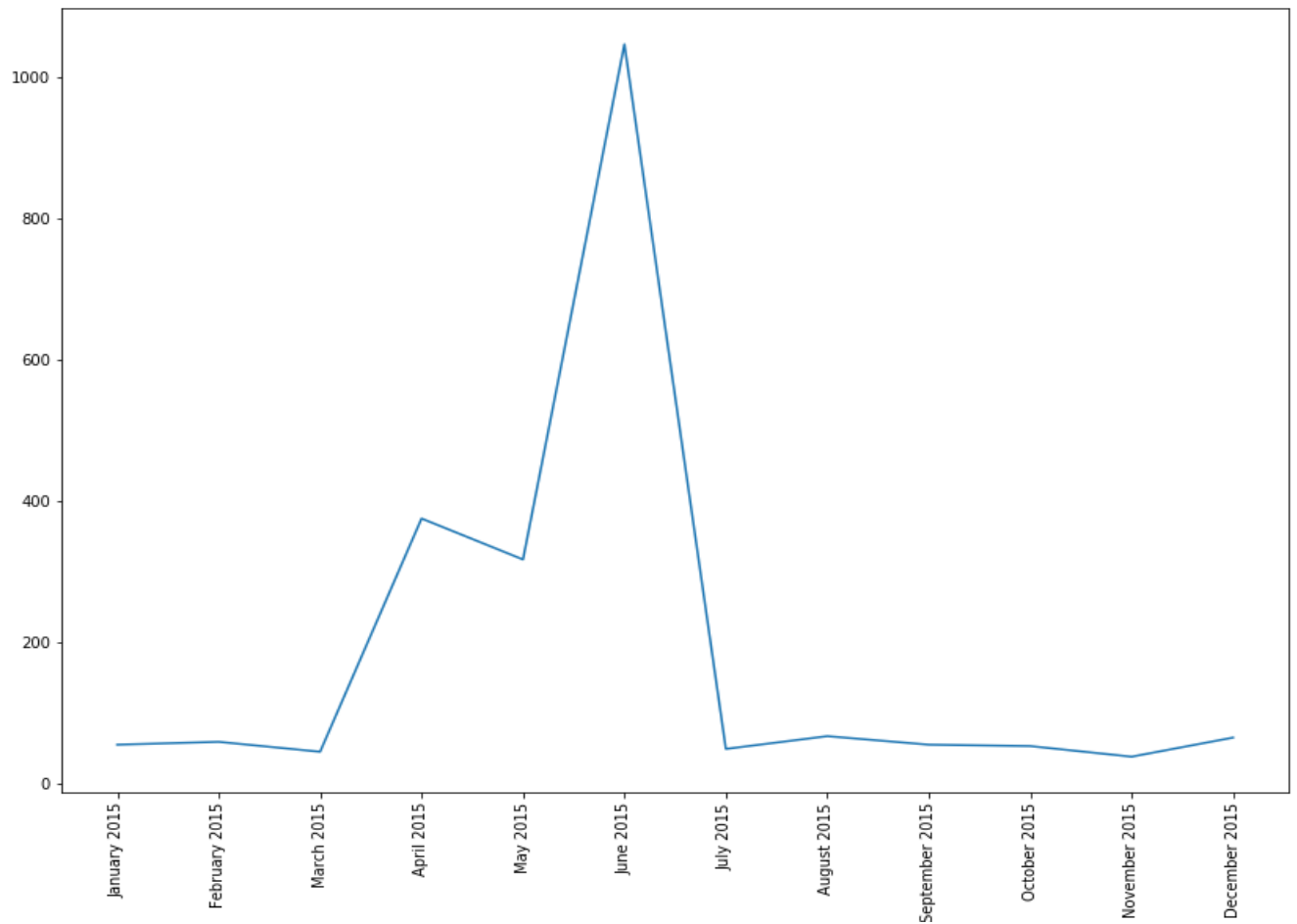
Trend chart for the number of complaints at monthly and daily granularity levels.

```
dfCopy["Date_month_year"] = pd.to_datetime(dfCopy["Date_month_year"])
dfCopy.set_index(dfCopy["Date_month_year"], inplace=True)
dfCopy.index
monthlyComplaints = dfCopy["Ticket #"].resample('M', convention='start').count()
dailyComplaints = dfCopy["Ticket #"].groupby("Date_month_year").count()
monthlyComplaints.count()
```

Maximum number of complaints noted in the month of June on 24th.

Minimum number of complaints noted in the month of November

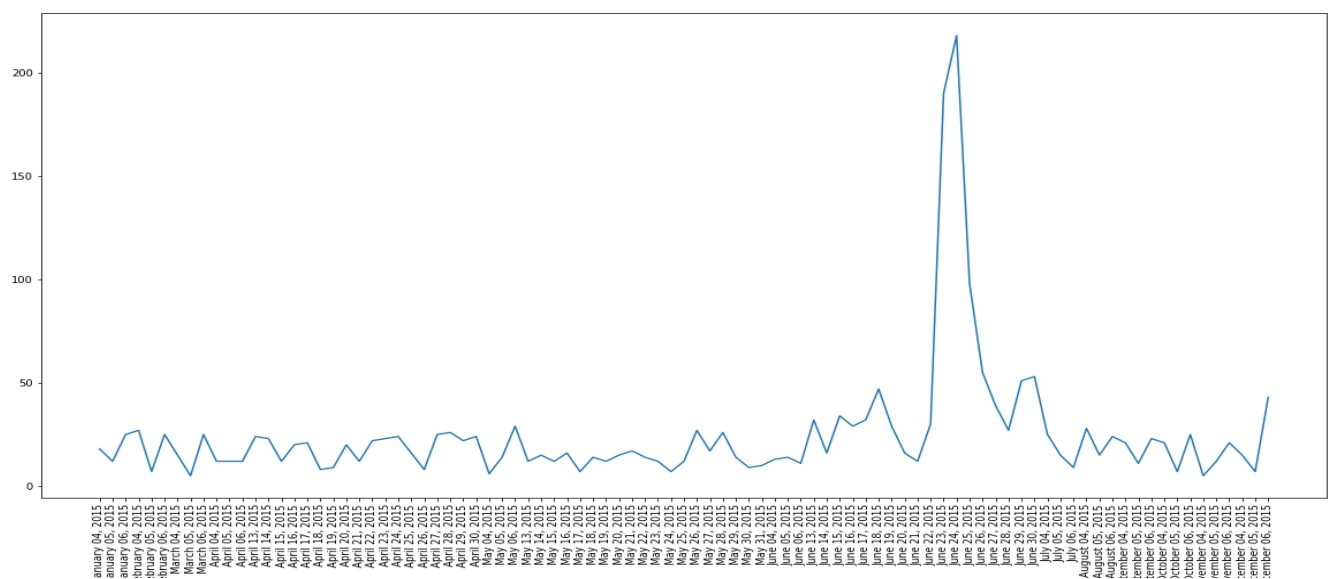
```
plt.figure(figsize=(14,10))
plt.xticks(rotation=90)
plt.plot(monthlyComplaints.index.strftime("%B %Y"), monthlyComplaints.values)
```



```
plt.figure(figsize=(20,10))
```

```
plt.xticks(rotation=90)
```

```
plt.plot(dailyComplaints.index.strftime("%B %d, %Y"),dailyComplaints.values)
```



Frequency of Complaint Types

Table with the frequency of complaint types and which complaint types are maximum i.e., around internet, network issues, or across any other domains.

```
df["Customer Complaint"].str.title().value_counts()
df["Customer Complaint"].str.title().value_counts().idxmax()
```

The complaint type 'Comcast' is maximum with frequency of 102

Categorical Variables

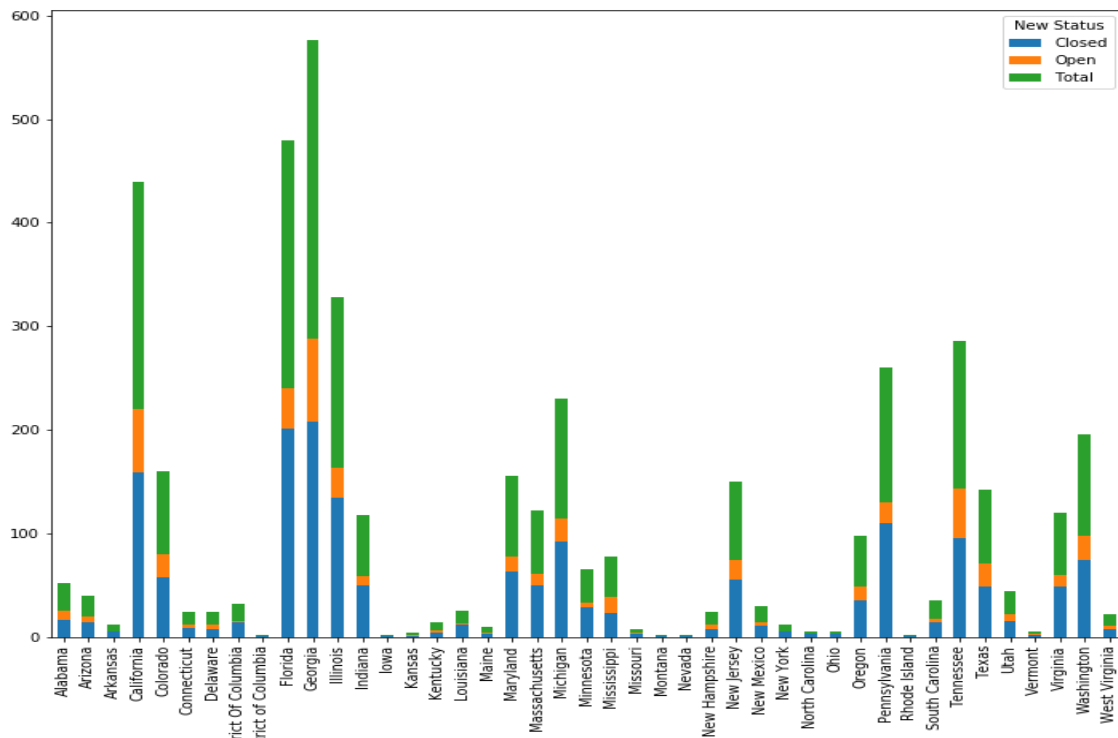
Created a new categorical variable with value as **Open** and **Closed**. Open & Pending is to be categorized as Open and Closed & Solved is to be categorized as Closed.

```
dfCategorical = df.copy(deep=True)
dfCategorical["New Status"] = dfCategorical["Status"]
dfCategorical
dfCategorical.loc[(dfCategorical["New Status"] == "Pending"), "New Status"] = "Open"
dfCategorical.loc[(dfCategorical["New Status"] == "Solved"), "New Status"] = "Closed"
```

Bar Charts

State wise status of complaints in a stacked bar chart using the new categorical variable

```
pivotTable = dfCategorical.pivot_table(index="State", columns="New Status", values="Status",
aggfunc='count').fillna(0)
pivotTable.plot.bar(stacked=True, figsize=(12,10))
pivotTable["Total"] = pivotTable["Closed"]+pivotTable["Open"]
```



State with Maximum Complaints

`pivotTable["Total"].idxmax()`

'Georgia' is the state with maximum complaints

State with Maximum Percentage of Unresolved Complaints

`(pivotTable["Open"]/pivotTable["Total"]).idxmax()`

'Kansas' is the state with maximum percentage of unresolved complaints

Percentage of Complaints Resolved till date through Internet and Customer Care

```
dfpivotTable2 = dfCategorical.pivot_table(index="Received Via", columns="New Status",
values="Status", aggfunc='count')
```

```
dfpivotTable2["Percent Resolved"] =
(dfpivotTable2["Closed"]/(dfpivotTable2["Closed"]+dfpivotTable2["Open"]))*100
```

dfpivotTable2

New Status	Closed	Open	Percent Resolved
Received Via			
Customer Care Call	864	255	77.211796
Internet	843	262	76.289593