

## Importing Necessary Libraries

```
In [53]: import numpy as np
import pandas as pd
import matplotlib.pyplot as plt
import seaborn as sns
```

## Task 1: Import data into Python environment.

### Import data set

```
In [54]: dataset = pd.read_csv("Comcast_telecom_complaints_data.csv")
```

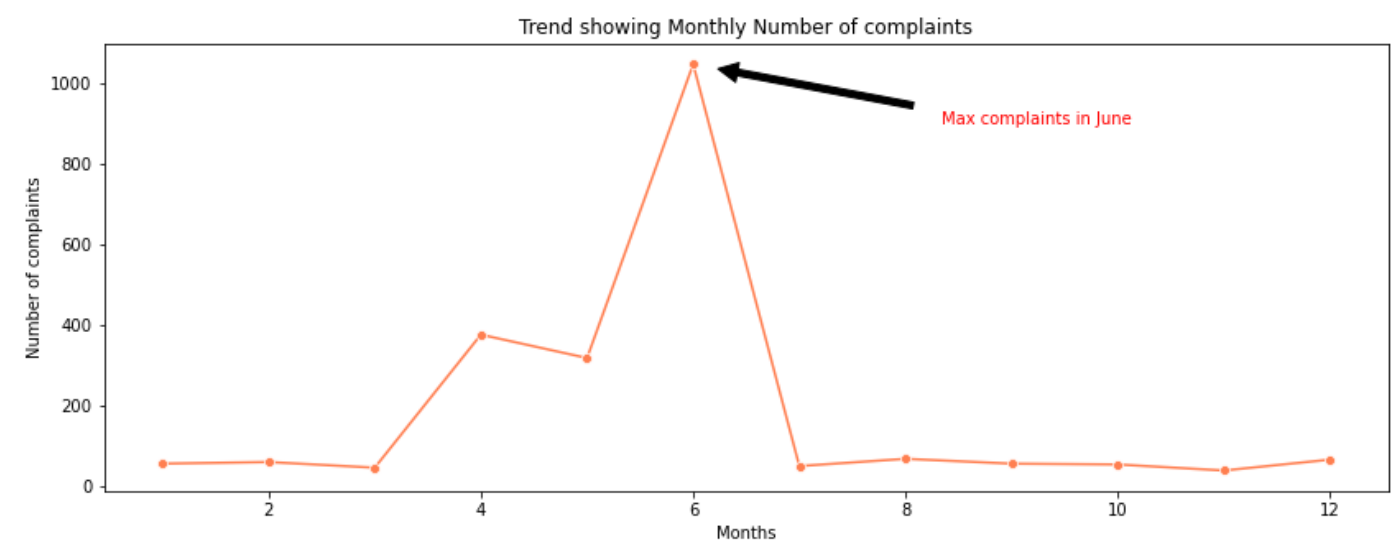
### Check whether data imported or not

```
In [55]: dataset.head()
```

Out[55]:

	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

```
In [66]: plt.figure(figsize=(14,5))
sns.lineplot(data= dataset_month_wise, color = "coral", marker = 'o')
plt.title('Trend showing Monthly Number of complaints')
plt.ylabel('Number of complaints')
plt.xlabel('Months')
plt.annotate('Max complaints in June', color='red',
            xy=(6, 1046), xycoords='data',
            xytext=(0.8, 0.85), textcoords='axes fraction',
            arrowprops=dict(facecolor='black', shrink=0.1),
            horizontalalignment='right', verticalalignment='top')
plt.show()
```



```
In [70]: complain_type = dataset['Customer Complaint'].str.lower().value_counts().sort_values(ascending=False)
#Lower case the srting as Python is case sensitive and it will read COMCAST,Comcast,comcast as 3 different complaints
complain_type.head()
```

```
Out[70]: Comcast          102
Comcast data cap        30
Comcast internet        29
Comcast data caps        21
Comcast billing          18
Name: Customer Complaint, dtype: int64
```

Top 10 Complaints are

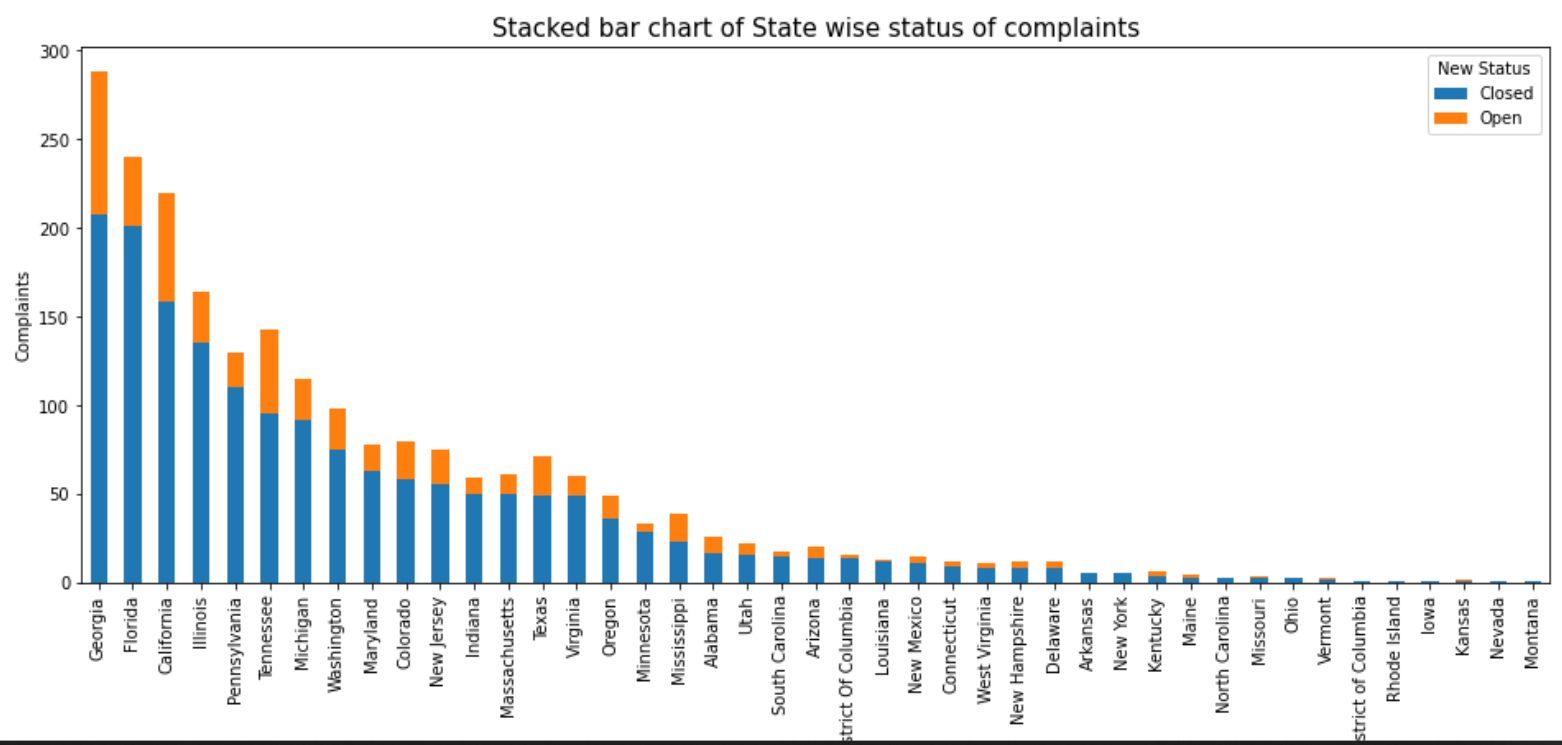
```
In [71]: complain_type.head(10)
```

```
Out[71]: Comcast          102
Comcast data cap        30
Comcast internet        29
Comcast data caps        21
Comcast billing          18
Comcast service          15
internet speed           15
data caps                13
unfair billing practices  13
data cap                 12
Name: Customer Complaint, dtype: int64
```

```
In [72]: complain_type_head = complain_type.head(10)
plt.figure(figsize=(10,10))
complain_type_head.plot.bar()
```

```
In [78]: complaint_by_state.sort_values('Closed',axis = 0,ascending=False).plot(kind="bar", figsize=(16,6), stacked=True)
plt.ylabel("Complaints")
plt.title('Stacked bar chart of State wise status of complaints', fontsize = 15)
```

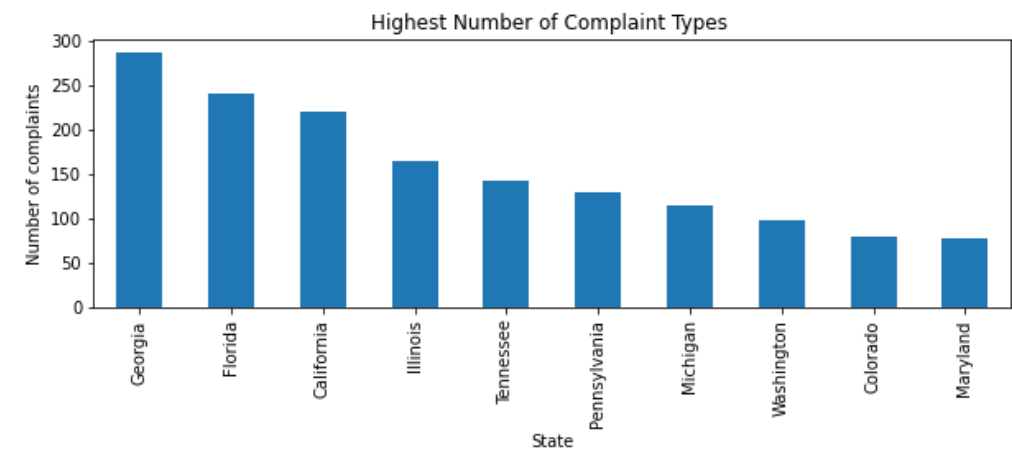
Out[78]: Text(0.5, 1.0, 'Stacked bar chart of State wise status of complaints')



Run Markdown

```
Pennsylvania 130
Michigan 115
Washington 98
Colorado 80
Maryland 78
Name: State, dtype: int64
```

```
In [80]: plt.figure(figsize=(10,3))
dataset['State'].value_counts().head(10).plot.bar()
plt.title('Highest Number of Complaint Types ')
plt.xlabel('State')
plt.ylabel('Number of complaints')
plt.show()
```



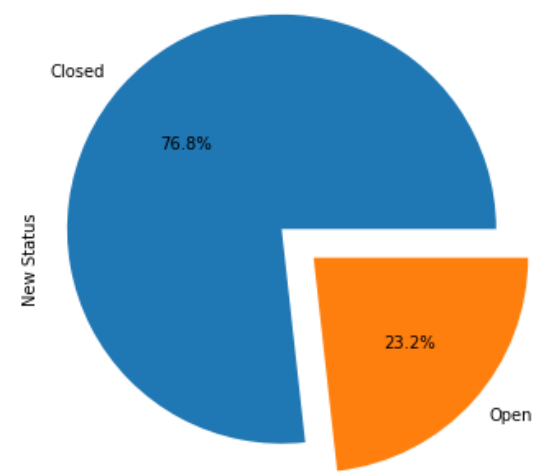
Georgia has the maximum complaints (288).

Run Markdown

```
plt.title('Complaints Status through the Internet & Customer Care Calls\n')
dataset['New Status'].value_counts().plot(kind='pie',explode = myexplode,autopct='%1.1f%%',
figsize = (20,6))
```

Out[95]: <AxesSubplot:title={'center': 'Complaints Status through the Internet & Customer Care Calls\n'}, ylabel='New Status'>

Complaints Status through the Internet & Customer Care Calls



Percentage of complaints resolved till date is 76.8%