

```
import numpy as np
import pandas as pd
```

```
data=pd.read_csv('Comcast_telecom_complaints_data.csv')
```

```
df_data=pd.DataFrame(data)
```

```
df_data.head()
```

	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

```
#print the number of rows and columns of the Data Frame
```

```
df_data.shape
```

```
(2224, 11)
```

```
#print the names of all columns.
```

```
df_data.columns
```

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

```
#print the last n rows of the Data Frame
df_data.tail()
```

	Ticket #	Customer Complaint	Date \
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 \				
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
2219	32466	Closed	No
2220	48197	Solved	No
2221	48197	Solved	No
2222	48197	Solved	No
2223	48198	Open	Yes

```
#give Index, Datatype and Memory information
df_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

```

#Dropping the column as it has too many null values

```
df_data.isnull().sum()
```

#There is no missing value in our dataset

```

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

```

#Replacing missing values with median values.

```
df_data.fillna(df_data.median())
```

	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]

#Replacing missing values with median values.

df_data.fillna(df_data.median())

	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]

#Replacing 0 values with median values.

```
df_data=df_data.mask(df_data==0).fillna(df_data.mean())
df_data
```

	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]

#Post Pandas Profiling

```

from pandas_profiling import ProfileReport
profile = ProfileReport(df_data)
profile

```

```

{"version_major":2,"version_minor":0,"model_id":"ef5568f3ec49476a92405
fee476438d5"}

```

```

{"version_major":2,"version_minor":0,"model_id":"c20e1cae8fda4454b7aae
3f699f89f90"}

```

```

{"version_major":2,"version_minor":0,"model_id":"8f2650f5bc1448c79f067
cb01f2cfcad"}

```

<IPython.core.display.HTML object>

#Installing packages

#Commonly used packages are: seaborn, matplotlib, bokeh, plotly

```

!pip install seaborn matplotlib bokeh plotly
import seaborn as sns
import matplotlib

```

```
import matplotlib.pyplot as plt
%matplotlib inline
```

Defaulting to user installation because normal site-packages is not writeable

```
Requirement already satisfied: seaborn in
/usr/local/lib/python3.7/site-packages (0.11.2)
Requirement already satisfied: matplotlib in
/usr/local/lib/python3.7/site-packages (3.5.1)
Requirement already satisfied: bokeh in /usr/local/lib/python3.7/site-
packages (2.4.2)
Requirement already satisfied: plotly in
/usr/local/lib/python3.7/site-packages (5.5.0)
Requirement already satisfied: scipy>=1.0 in
/usr/local/lib/python3.7/site-packages (from seaborn) (1.4.1)
Requirement already satisfied: pandas>=0.23 in
/usr/local/lib/python3.7/site-packages (from seaborn) (1.1.5)
Requirement already satisfied: numpy>=1.15 in
/usr/local/lib/python3.7/site-packages (from seaborn) (1.21.5)
Requirement already satisfied: python-dateutil>=2.7 in
/usr/local/lib/python3.7/site-packages (from matplotlib) (2.8.1)
Requirement already satisfied: packaging>=20.0 in
/usr/local/lib/python3.7/site-packages (from matplotlib) (21.0)
Requirement already satisfied: fonttools>=4.22.0 in
/usr/local/lib/python3.7/site-packages (from matplotlib) (4.28.5)
Requirement already satisfied: cycler>=0.10 in
/usr/local/lib/python3.7/site-packages (from matplotlib) (0.10.0)
Requirement already satisfied: kiwisolver>=1.0.1 in
/usr/local/lib/python3.7/site-packages (from matplotlib) (1.2.0)
Requirement already satisfied: pillow>=6.2.0 in
/usr/local/lib/python3.7/site-packages (from matplotlib) (7.1.1)
Requirement already satisfied: pyparsing>=2.2.1 in
/usr/local/lib/python3.7/site-packages (from matplotlib) (2.4.6)
Requirement already satisfied: typing-extensions>=3.10.0 in
/usr/local/lib/python3.7/site-packages (from bokeh) (4.0.1)
Requirement already satisfied: PyYAML>=3.10 in
/usr/local/lib/python3.7/site-packages (from bokeh) (5.3.1)
Requirement already satisfied: Jinja2>=2.9 in
/usr/local/lib/python3.7/site-packages (from bokeh) (2.11.1)
Requirement already satisfied: tornado>=5.1 in
/usr/local/lib/python3.7/site-packages (from bokeh) (6.1)
Requirement already satisfied: tenacity>=6.2.0 in
/usr/local/lib/python3.7/site-packages (from plotly) (8.0.1)
Requirement already satisfied: six in /usr/local/lib/python3.7/site-
packages (from plotly) (1.14.0)
Requirement already satisfied: MarkupSafe>=0.23 in
/usr/local/lib/python3.7/site-packages (from Jinja2>=2.9->bokeh)
(1.1.1)
Requirement already satisfied: pytz>=2017.2 in
/usr/local/lib/python3.7/site-packages (from pandas>=0.23->seaborn)
```


(2019.3)

WARNING: You are using pip version 22.0.3; however, version 22.0.4 is available.

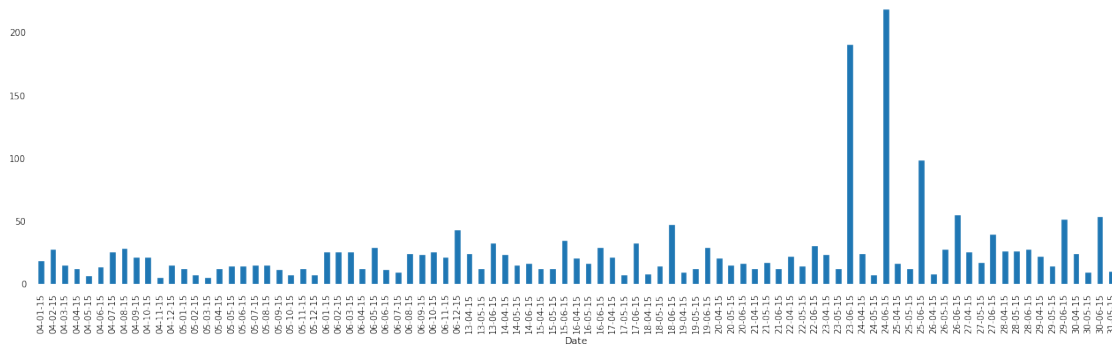
You should consider upgrading via the '/usr/local/bin/python3 -m pip install --upgrade pip' command.

#Provide the trend chart for the number of complaints at monthly and daily granularity levels

#Lookout complaints over the date

```
df_data['Customer  
Complaint'].groupby(df_data.Date).count().plot(kind='bar',figsize=(22,  
6))
```

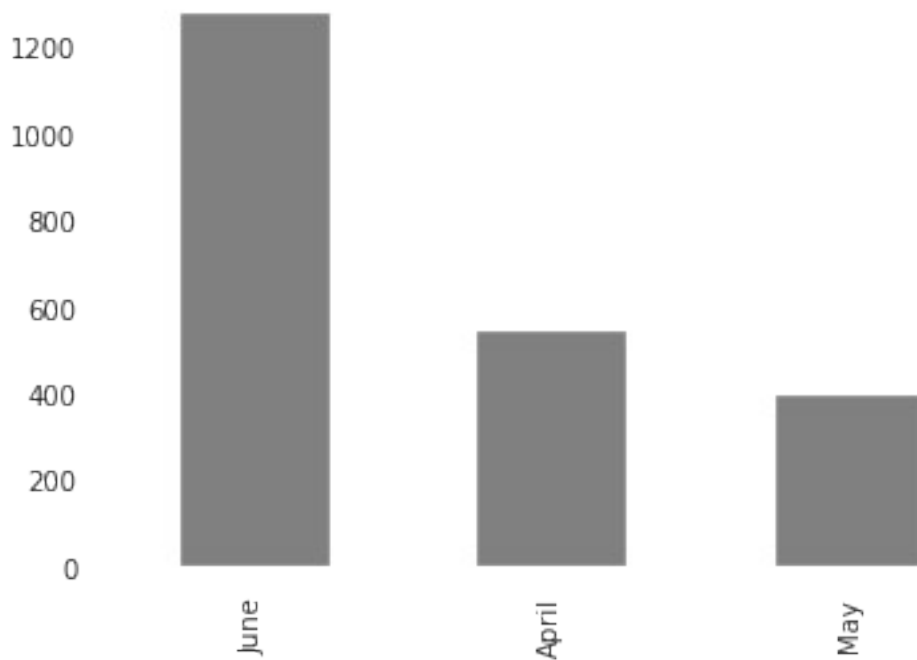
<AxesSubplot:xlabel='Date'>



#Look out complaints over the month

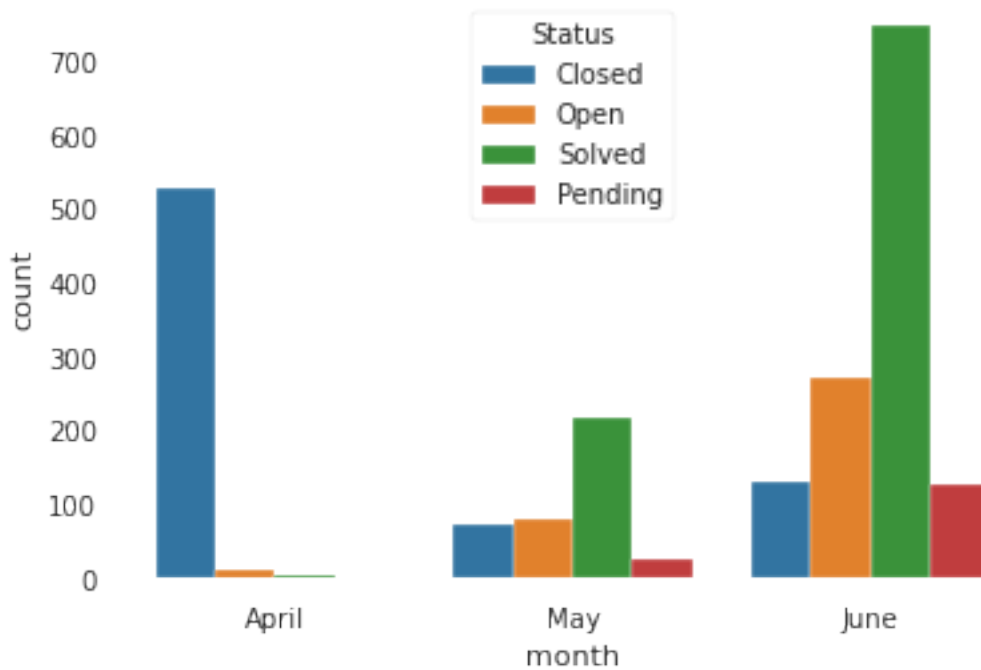
```
from datetime import datetime as dt  
df_data['month'] = pd.DatetimeIndex(df_data['Date']).month  
df_data['month']=df_data['month'].replace([1,2,3,4,5,6,7,8,9,10,11,12]  
,  
['Jan','Feb','March','April','May','June','July','Aug','Sep','Oct','No  
v','Dec'])  
df_data['month'].value_counts().plot(kind='bar',color='grey')
```

<AxesSubplot:>



```
# Look out the status of complaints over the months
sns.countplot(x='month', hue='Status', data=df_data)

<AxesSubplot:xlabel='month', ylabel='count'>
```



```
#No. of complaints statewise
df_state=df_data[['Ticket #',
'State']].groupby('State').count().sort_values(by='Ticket
```

```
#',ascending=False)
df_state.describe
```

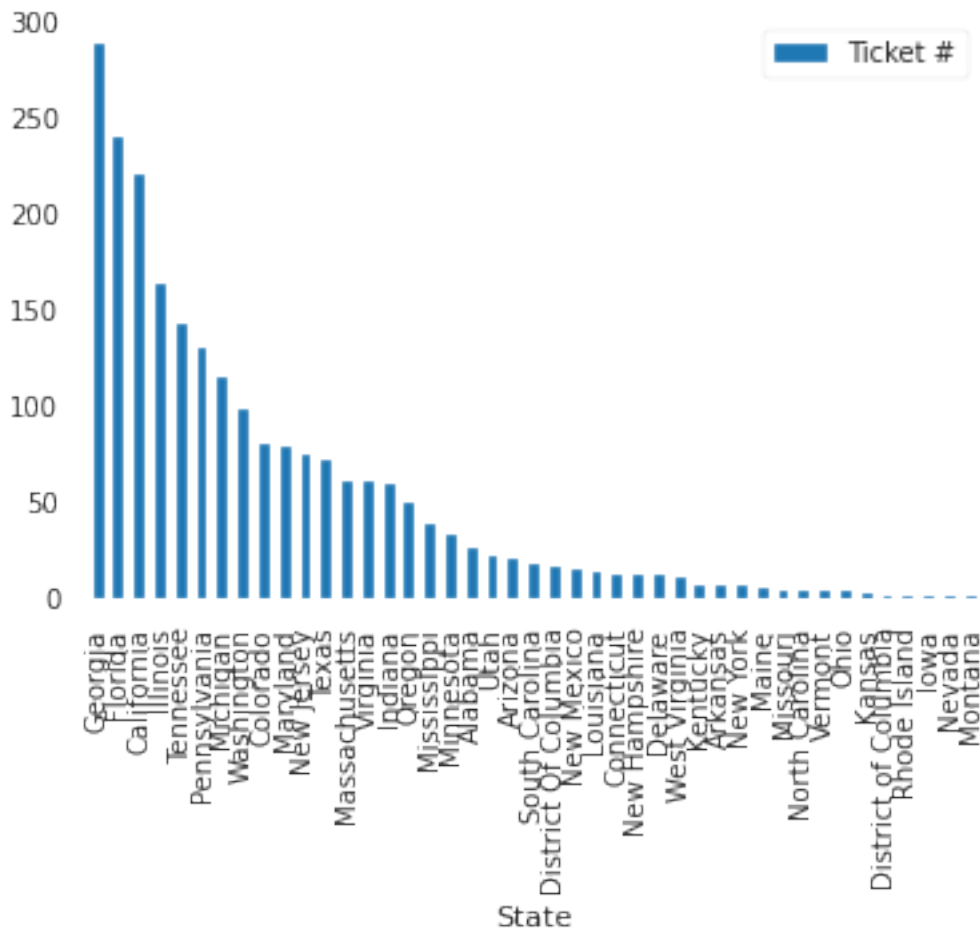
```
<bound method NDFrame.describe of
State
```

Ticket #

Georgia	288
Florida	240
California	220
Illinois	164
Tennessee	143
Pennsylvania	130
Michigan	115
Washington	98
Colorado	80
Maryland	78
New Jersey	75
Texas	71
Massachusetts	61
Virginia	60
Indiana	59
Oregon	49
Mississippi	39
Minnesota	33
Alabama	26
Utah	22
Arizona	20
South Carolina	18
District Of Columbia	16
New Mexico	15
Louisiana	13
Connecticut	12
New Hampshire	12
Delaware	12
West Virginia	11
Kentucky	7
Arkansas	6
New York	6
Maine	5
Missouri	4
North Carolina	3
Vermont	3
Ohio	3
Kansas	2
District of Columbia	1
Rhode Island	1
Iowa	1
Nevada	1
Montana	1>

```
df_state.plot(kind='bar')
```

<AxesSubplot:xlabel='State'>



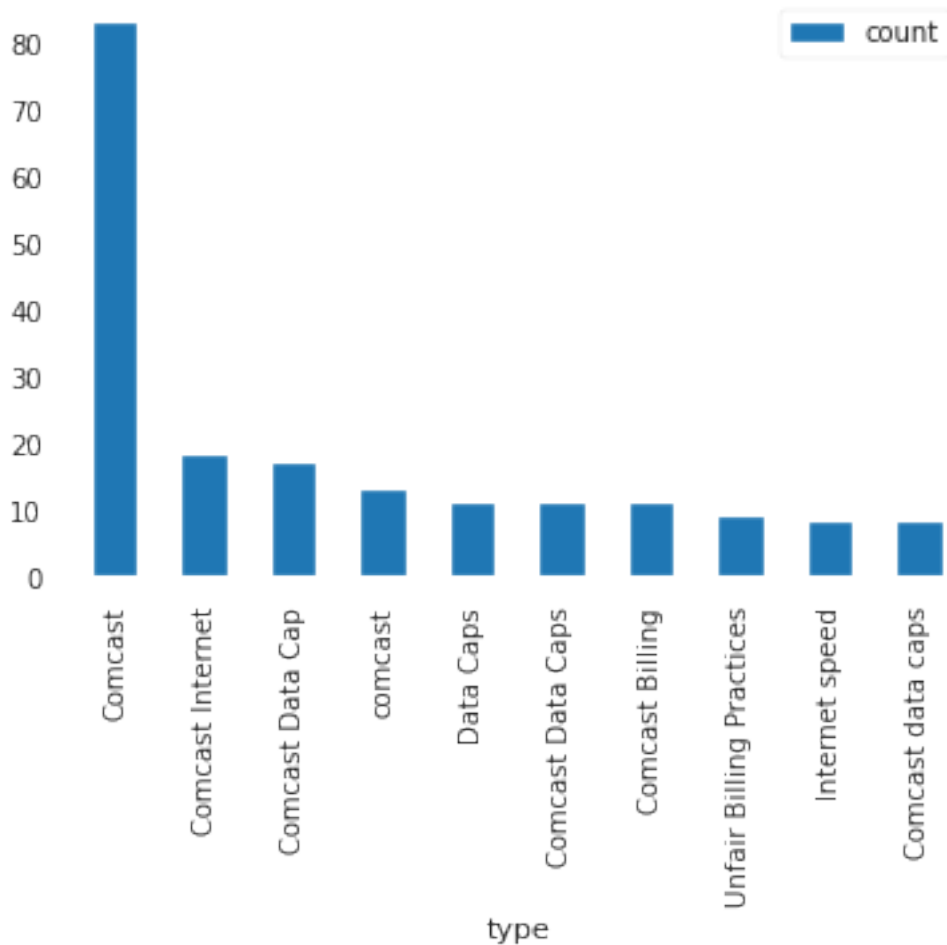
```
#Find the top 10 complaint types
from collections import Counter
c_count = Counter(df_data['Customer Complaint'])
x=c_count.most_common(10)
x=pd.DataFrame(x,columns=['type','count'])
x
```

	type	count
0	Comcast	83
1	Comcast Internet	18
2	Comcast Data Cap	17
3	comcast	13
4	Data Caps	11
5	Comcast Data Caps	11
6	Comcast Billing	11
7	Unfair Billing Practices	9
8	Internet speed	8
9	Comcast data caps	8

```
#Complaint TypePlots
```

```
x.plot(x='type',kind='bar')
```

```
<AxesSubplot:xlabel='type'>
```



```
#Calculating statewide complaint resolving rate
```

```
df_data.Status.unique()
```

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

```
df_data['open/closed']=df_data['Status'].apply(lambda x: 'Closed' if
(x=='Solved'or x== 'Closed') else 'Open')
```

```
State_solv_rate=df_data.groupby(['State','open/closed'])['open/
closed'].count().unstack().fillna(0)
```

```
State_solv_rate=pd.DataFrame(Statewise_solving_rate)
```

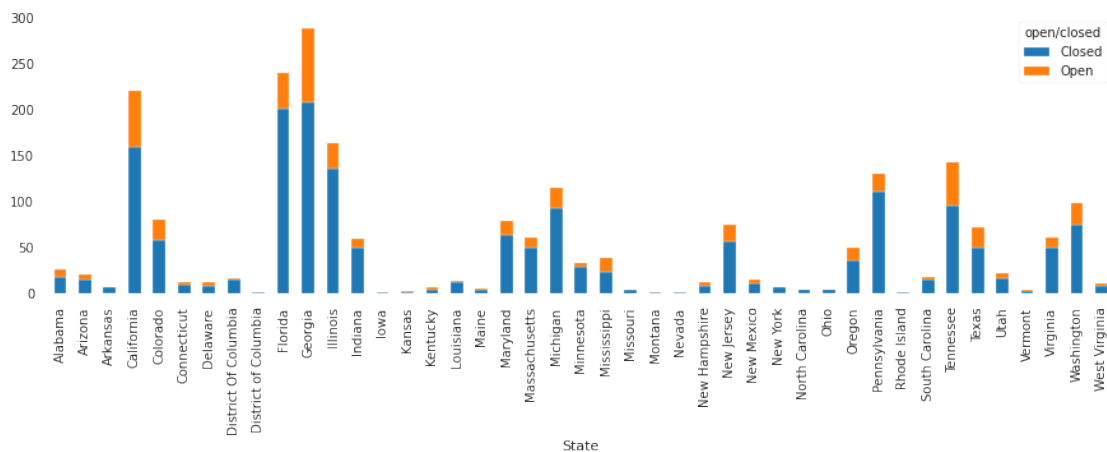
```
State_solv_rate.head(20)
```

open/closed	Closed	Open
State		
Alabama	17.0	9.0
Arizona	14.0	6.0

Arkansas	6.0	0.0
California	159.0	61.0
Colorado	58.0	22.0
Connecticut	9.0	3.0
Delaware	8.0	4.0
District Of Columbia	14.0	2.0
District of Columbia	1.0	0.0
Florida	201.0	39.0
Georgia	208.0	80.0
Illinois	135.0	29.0
Indiana	50.0	9.0
Iowa	1.0	0.0
Kansas	1.0	1.0
Kentucky	4.0	3.0
Louisiana	12.0	1.0
Maine	3.0	2.0
Maryland	63.0	15.0
Massachusetts	50.0	11.0

```
State_solv_rate.plot(kind='bar',stacked=True,figsize=(15,4))
```

```
<AxesSubplot:xlabel='State'>
```



```
#% of complaints resolved till date
```

```
df_compl=df_data.groupby(['Received Via','open/closed'])['Received  
Via'].count().unstack()  
df_compl
```

open/closed	Closed	Open
Received Via		
Customer Care Call	864	255
Internet	843	262

```
compl=df_data['open/closed'].value_counts()  
compl
```

```
Closed    1707
Open       517
Name: open/closed, dtype: int64
```

```
df_totalcompl=len(df_data)
df_totalcompl
```

```
2224
```

```
for x in compl.values:
    l=(x/df_totalcompl*100)
    print(l)
```

```
76.75359712230215
23.246402877697843
```