project : Comcast Telecom Consumer Complaints

```
In [1]: #importing usefull library

        import pandas as pd
        import numpy as np
        import seaborn as sns
        from matplotlib import pyplot as plt
```

```
In [2]: df= pd.read_csv('telecom.csv')    #importing dataset using pandas library !!!!!
```

```
In [3]: df.head()
```

Out[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

```
In [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
```

```
In [5]: df['Date']= pd.to_datetime(df['Date'])   #converting date , str to date time formate!!!!
```

```
In [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
```

```
In [7]: df['month']= df['Date'].dt.month_name()   #assign a new col. (month) using date col.!!
```

```python
In [8]:   df.head()
```

Out[8]:

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

```python
In [9]:   daily= df.groupby(['Date']).size()
```

```python
In [10]:  daily
```

```
Out[10]: Date
         2015-04-01    18
         2015-04-02    27
         2015-04-03    15
         2015-04-04    12
         2015-04-05     6
                       ..
         2015-06-26    55
         2015-06-27    39
         2015-06-28    27
         2015-06-29    51
         2015-06-30    53
         Length: 91, dtype: int64
```

```python
In [11]:  daily=pd.DataFrame(daily).reset_index()
```

```python
In [12]:  daily
```

Out[12]:

	Date	0
0	2015-04-01	18
1	2015-04-02	27
2	2015-04-03	15
3	2015-04-04	12
4	2015-04-05	6
...	...	...
86	2015-06-26	55
87	2015-06-27	39
88	2015-06-28	27
89	2015-06-29	51
90	2015-06-30	53

91 rows × 2 columns

```python
In [13]:  daily= daily.rename(columns={0:'count'})
```

```
In [14]:  daily
```

Out[14]:

	Date	count
0	2015-04-01	18
1	2015-04-02	27
2	2015-04-03	15
3	2015-04-04	12
4	2015-04-05	6
...	...	...
86	2015-06-26	55
87	2015-06-27	39
88	2015-06-28	27
89	2015-06-29	51
90	2015-06-30	53

91 rows × 2 columns

```
In [15]:  daily.plot(x='Date',y='count',kind='line')  #creating lineplot to analysis of no. of complains received according to month!!
          plt.show()
```
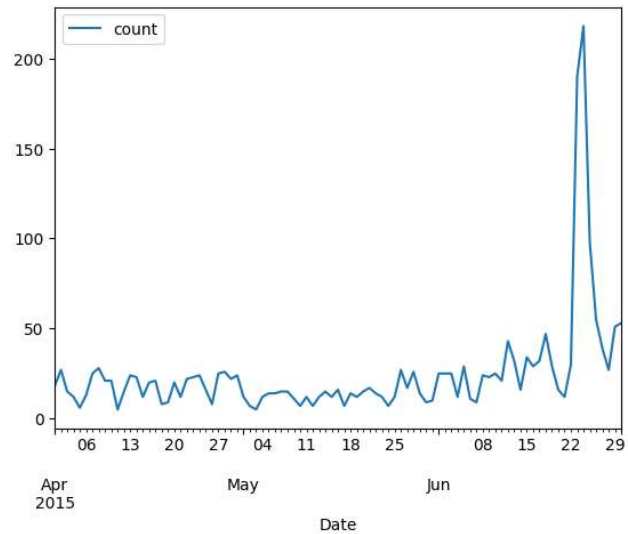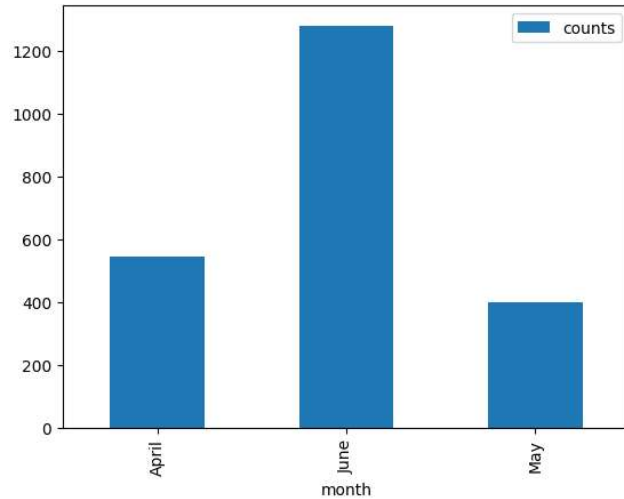


```
In [16]:  monthly= df.groupby(['month']).size()
```

```
In [17]:  monthly        # max complains received in june month !!!
```

Out[17]:
```
month
April     545
June     1280
May       399
dtype: int64
```

```
In [18]:  monthly=pd.DataFrame(monthly).reset_index()
          monthly=monthly.rename(columns={0:'counts'})
```

```
In [19]:  monthly.plot(x='month',y='counts',kind='bar')  #creating the bar plot to visualization the monthly data
          plt.show()
```



```
In [20]:  df.head(10)
```

Out[20]:

	Ticket #	Customer Complaint	Date	Date_month_year	Time	Received Via	City	State	Zip code	Status	Filing on Behalf of Someone	month
0	250635	Comcast Cable Internet Speeds	2015-04-22	22-Apr-15	3:53:50 PM	Customer Care Call	Abingdon	Maryland	21009	Closed	No	April
1	223441	Payment disappear - service got disconnected	2015-04-08	04-Aug-15	10:22:56 AM	Internet	Acworth	Georgia	30102	Closed	No	April
2	242732	Speed and Service	2015-04-18	18-Apr-15	9:55:47 AM	Internet	Acworth	Georgia	30101	Closed	Yes	April
3	277946	Comcast Imposed a New Usage Cap of 300GB that ...	2015-05-07	05-Jul-15	11:59:35 AM	Internet	Acworth	Georgia	30101	Open	Yes	May
4	307175	Comcast not working and no service to boot	2015-05-26	26-May-15	1:25:26 PM	Internet	Acworth	Georgia	30101	Solved	No	May
5	338519	ISP Charging for arbitrary data limits with ov...	2015-06-12	06-Dec-15	9:59:40 PM	Internet	Acworth	Georgia	30101	Solved	No	June
6	361148	Throttling service and unreasonable data caps	2015-06-24	24-Jun-15	10:13:55 AM	Customer Care Call	Acworth	Georgia	30101	Pending	No	June
7	359792	Comcast refuses to help troubleshoot and corre...	2015-06-23	23-Jun-15	6:56:14 PM	Internet	Adrian	Michigan	49221	Solved	No	June
8	318072	Comcast extended outages	2015-06-01	06-Jan-15	11:46:30 PM	Customer Care Call	Alameda	California	94502	Closed	No	June
9	371214	Comcast Raising Prices and Not Being Available...	2015-06-28	28-Jun-15	6:46:31 PM	Customer Care Call	Alameda	California	94501	Open	Yes	June

```
In [21]:  df['Customer Complaint'].value_counts()    #analysing the types and frequency of complaints by the customers .
```

```
Out[21]:  Comcast                                          83
          Comcast Internet                                 18
          Comcast Data Cap                                 17
          comcast                                          13
          Comcast Billing                                  11
                                                           ..
          Improper Billing and non resolution of issues    1
          Deceptive trade                                  1
          intermittent internet                            1
          Internet Speed on Wireless Connection            1
          Comcast, Ypsilanti MI Internet Speed             1
          Name: Customer Complaint, Length: 1841, dtype: int64
```

```
In [22]:  #cleaning the data , by making first latter in upper case by using title function !!

          df['Customer Complaint']=df['Customer Complaint'].apply( lambda x :  x.title() )
```

```
In [23]:  df['Customer Complaint'].value_counts()
```

```
Out[23]: Comcast                                                  102
         Comcast Data Cap                                          30
         Comcast Internet                                          29
         Comcast Data Caps                                         21
         Comcast Billing                                           18
                                                                  ...
         Monthly Data Caps                                          1
         Comcast/Xfinity Poor Service, Fraudulent Billing And Collection   1
         Lost Emails/Billing                                        1
         Improper Billing And Non Resolution Of Issues              1
         Comcast, Ypsilanti Mi Internet Speed                       1
         Name: Customer Complaint, Length: 1740, dtype: int64
```

```
In [24]:  #creating the bar plot for visualization of frequency for complaints by the customer !!!
          df['Customer Complaint'].value_counts()[:20].plot.bar()
          plt.show()
```
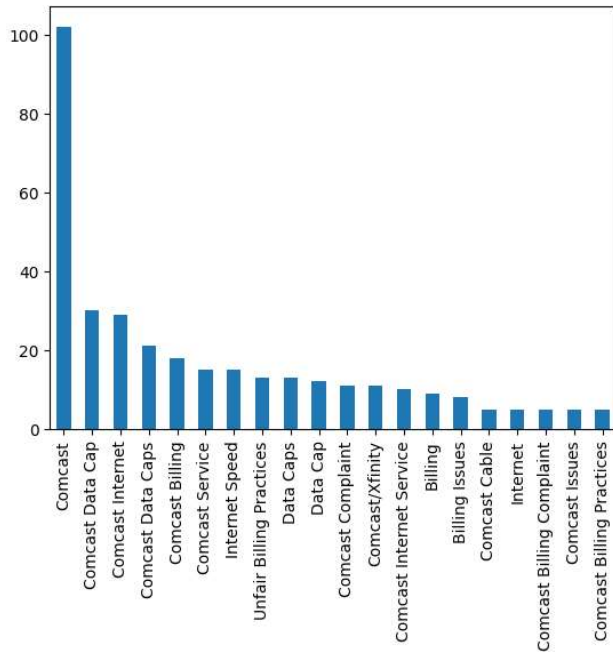


```
In [25]:  # analysing the different type of issues,
          #no of internet related issues !
          internet_issue_1= df[df['Customer Complaint'].str.contains('Internet')].count()
```

```
In [26]:  internet_issue_1
```

```
Out[26]: Ticket #                    531
         Customer Complaint          531
         Date                        531
         Date_month_year             531
         Time                        531
         Received Via                531
         City                        531
         State                       531
         Zip code                    531
         Status                      531
         Filing on Behalf of Someone 531
         month                       531
         dtype: int64
```

```
In [27]:  internet_issue_2=df[df['Customer Complaint'].str.contains('Speed')].count()
```

```
In [28]:  internet_issue_2
```

```
Out[28]: Ticket #                    192
         Customer Complaint          192
         Date                        192
         Date_month_year             192
         Time                        192
         Received Via                192
         City                        192
         State                       192
         Zip code                    192
         Status                      192
         Filing on Behalf of Someone 192
         month                       192
         dtype: int64
```

```
In [29]:  internet_issue_3=df[df['Customer Complaint'].str.contains('Network')].count()
```

```
In [30]:  internet_issue_3
```

```
Out[30]: Ticket #                    2
         Customer Complaint          2
         Date                        2
         Date_month_year             2
         Time                        2
         Received Via                2
         City                        2
         State                       2
         Zip code                    2
         Status                      2
         Filing on Behalf of Someone 2
         month                       2
         dtype: int64
```

```
In [31]:  internet_issue= internet_issue_1+internet_issue_2+internet_issue_3
```

```
In [32]:  internet_issue # total no of internet related issues !
```

```
Out[32]: Ticket #                    725
         Customer Complaint          725
         Date                        725
         Date_month_year             725
         Time                        725
         Received Via                725
         City                        725
         State                       725
         Zip code                    725
         Status                      725
         Filing on Behalf of Someone 725
         month                       725
         dtype: int64
```

```
In [33]:  #no of service related issues !
          service_issue_1= df[df['Customer Complaint'].str.contains('Service')].count()
```

```
In [34]:    service_issue_1
```

```
Out[34]: Ticket #                   489
         Customer Complaint         489
         Date                       489
         Date_month_year            489
         Time                       489
         Received Via               489
         City                       489
         State                      489
         Zip code                   489
         Status                     489
         Filing on Behalf of Someone 489
         month                      489
         dtype: int64
```

```
In [35]:    service_issue_2= df[df['Customer Complaint'].str.contains('Customer')].count()
```

```
In [36]:    service_issue_2
```

```
Out[36]: Ticket #                   95
         Customer Complaint         95
         Date                       95
         Date_month_year            95
         Time                       95
         Received Via               95
         City                       95
         State                      95
         Zip code                   95
         Status                     95
         Filing on Behalf of Someone 95
         month                      95
         dtype: int64
```

```
In [37]:    service_issue= service_issue_1+service_issue_2
```

```
In [38]:    service_issue # total no of internet related issues !
```

```
Out[38]: Ticket #                   584
         Customer Complaint         584
         Date                       584
         Date_month_year            584
         Time                       584
         Received Via               584
         City                       584
         State                      584
         Zip code                   584
         Status                     584
         Filing on Behalf of Someone 584
         month                      584
         dtype: int64
```

```
In [39]:  #no of billing related issues !
          biling_issue_1= df[df['Customer Complaint'].str.contains('Billing')].count()
          biling_issue_2=df[df['Customer Complaint'].str.contains('Charges')].count()
```

```
In [40]:    biling_issue= biling_issue_1+biling_issue_2
```

```
In [41]:   biling_issue  # total no of internet related issues !
```

```
Out[41]:  Ticket #                     365
          Customer Complaint           365
          Date                         365
          Date_month_year              365
          Time                         365
          Received Via                 365
          City                         365
          State                        365
          Zip code                     365
          Status                       365
          Filing on Behalf of Someone  365
          month                        365
          dtype: int64
```

```
In [42]:   df.shape
```

```
Out[42]:  (2224, 12)
```

```
In [43]:   other_issue= 2224-(biling_issue+internet_issue+service_issue)
```

```
In [44]:   other_issue   #other no. of issues which are facing by the customers/
```

```
Out[44]:  Ticket #                     550
          Customer Complaint           550
          Date                         550
          Date_month_year              550
          Time                         550
          Received Via                 550
          City                         550
          State                        550
          Zip code                     550
          Status                       550
          Filing on Behalf of Someone  550
          month                        550
          dtype: int64
```

```
In [45]:   df.head()
```

Out[45]:

	Ticket #	Customer Complaint	Date	Date_month_year	Time	Received Via	City	State	Zip code	Status	Filing on Behalf of Someone	month
0	250635	Comcast Cable Internet Speeds	2015-04-22	22-Apr-15	3:53:50 PM	Customer Care Call	Abingdon	Maryland	21009	Closed	No	April
1	223441	Payment Disappear - Service Got Disconnected	2015-04-08	04-Aug-15	10:22:56 AM	Internet	Acworth	Georgia	30102	Closed	No	April
2	242732	Speed And Service	2015-04-18	18-Apr-15	9:55:47 AM	Internet	Acworth	Georgia	30101	Closed	Yes	April
3	277946	Comcast Imposed A New Usage Cap Of 300Gb That ...	2015-05-07	05-Jul-15	11:59:35 AM	Internet	Acworth	Georgia	30101	Open	Yes	May
4	307175	Comcast Not Working And No Service To Boot	2015-05-26	26-May-15	1:25:26 PM	Internet	Acworth	Georgia	30101	Solved	No	May

```
In [46]: ▼ #now chacking the complains status
           df['Status'].unique()
```

```
Out[46]:  array(['Closed', 'Open', 'Solved', 'Pending'], dtype=object)
```

```
In [47]: ▼ #now coverting status :( open & pending to "Open" )  and   ( closed & solved to "Closed")

           df['New_status']= [  'Open'  if i=='Open' or i=='Pending' else 'Closed'  for i in df['Status']                ]
```

```
In [48]:   df.head()
```

Out[48]:

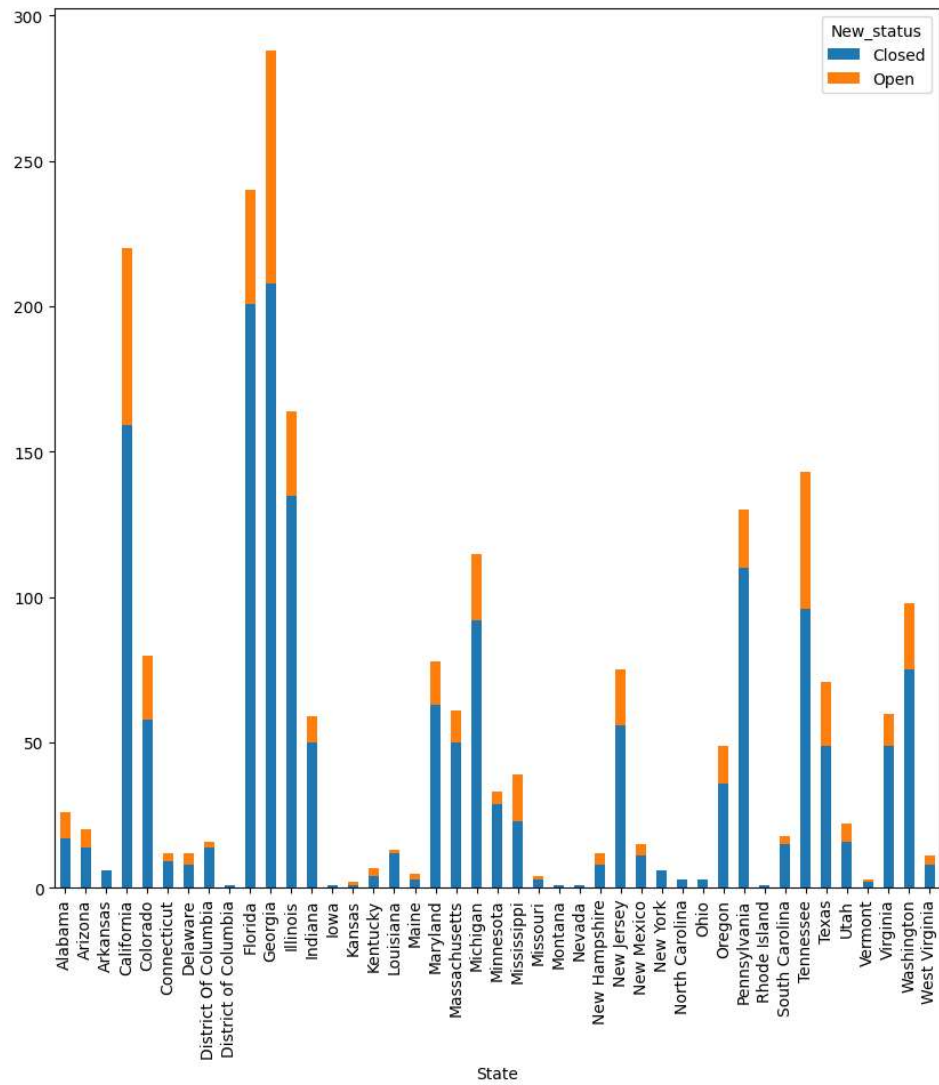|	Ticket #	Customer Complaint	Date	Date_month_year	Time	Received Via	City	State	Zip code	Status	Filing on Behalf of Someone	month	New_status
0	250635	Comcast Cable Internet Speeds	2015-04-22	22-Apr-15	3:53:50 PM	Customer Care Call	Abingdon	Maryland	21009	Closed	No	April	Closed
1	223441	Payment Disappear - Service Got Disconnected	2015-04-08	04-Aug-15	10:22:56 AM	Internet	Acworth	Georgia	30102	Closed	No	April	Closed
2	242732	Speed And Service	2015-04-18	18-Apr-15	9:55:47 AM	Internet	Acworth	Georgia	30101	Closed	Yes	April	Closed
3	277946	Comcast Imposed A New Usage Cap Of 300Gb That ...	2015-05-07	05-Jul-15	11:59:35 AM	Internet	Acworth	Georgia	30101	Open	Yes	May	Open
4	307175	Comcast Not Working And No Service To Boot	2015-05-26	26-May-15	1:25:26 PM	Internet	Acworth	Georgia	30101	Solved	No	May	Closed

```
In [49]:   State_complains=df.groupby(['State','New_status']).size().unstack().fillna(0)
```

```
In [50]:   #analysing the state in which the most no of complains are open !!
           State_complains['Open' ].sort_values()
```

Out[50]: State
          New York   0.0
          Nevada   0.0
          Arkansas   0.0
          Rhode Island   0.0
          Ohio   0.0
          Iowa   0.0
          District of Columbia   0.0
          North Carolina   0.0
          Montana   0.0
          Louisiana   1.0
          Kansas   1.0
          Missouri   1.0
          Vermont   1.0
          District Of Columbia   2.0
          Maine   2.0
          Connecticut   3.0
          Kentucky   3.0
          South Carolina   3.0
          West Virginia   3.0
          New Hampshire   4.0
          New Mexico   4.0
          Minnesota   4.0
          Delaware   4.0
          Arizona   6.0
          Utah   6.0
          Alabama   9.0
          Indiana   9.0
          Virginia   11.0
          Massachusetts   11.0
          Oregon   13.0
          Maryland   15.0
          Mississippi   16.0
          New Jersey   19.0
          Pennsylvania   20.0
          Colorado   22.0
          Texas   22.0
          Washington   23.0
          Michigan   23.0
          Illinois   29.0
          Florida   39.0
          Tennessee   47.0
          California   61.0
          Georgia   80.0
          Name: Open, dtype: float64
```

`#creating the bar plot for visulaization the open and close complains according to state.`
`State_complains.plot.bar(stacked=True,figsize=(10,10))`

Out[51]: <AxesSubplot:xlabel='State'>

```
In [52]:    State_complains
```

Out[52]:

State	New_status Closed	Open
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
Michigan	92.0	23.0
Minnesota	29.0	4.0
Mississippi	23.0	16.0
Missouri	3.0	1.0
Montana	1.0	0.0
Nevada	1.0	0.0
New Hampshire	8.0	4.0
New Jersey	56.0	19.0
New Mexico	11.0	4.0
New York	6.0	0.0
North Carolina	3.0	0.0
Ohio	3.0	0.0
Oregon	36.0	13.0
Pennsylvania	110.0	20.0
Rhode Island	1.0	0.0
South Carolina	15.0	3.0
Tennessee	96.0	47.0
Texas	49.0	22.0
Utah	16.0	6.0
Vermont	2.0	1.0
Virginia	49.0	11.0
Washington	75.0	23.0
West Virginia	8.0	3.0

```
In [53]:  df['State'].value_counts()[:5] #top 5 state according to complains
```

```
Out[53]:  Georgia      288
          Florida      240
          California   220
          Illinois     164
          Tennessee    143
          Name: State, dtype: int64
```

```
In [54]:  unresolved_complains= df.groupby(['State','New_status']).size().unstack().fillna(0).sort_values(by='Open', ascending=False)
```

```
In [55]:  unresolved_complains
```

Out[55]:

State	New_status Closed	Open
Georgia	208.0	80.0
California	159.0	61.0
Tennessee	96.0	47.0
Florida	201.0	39.0
Illinois	135.0	29.0
Washington	75.0	23.0
Michigan	92.0	23.0
Colorado	58.0	22.0
Texas	49.0	22.0
Pennsylvania	110.0	20.0
New Jersey	56.0	19.0
Mississippi	23.0	16.0
Maryland	63.0	15.0
Oregon	36.0	13.0
Virginia	49.0	11.0
Massachusetts	50.0	11.0
Alabama	17.0	9.0
Indiana	50.0	9.0
Utah	16.0	6.0
Arizona	14.0	6.0
New Hampshire	8.0	4.0
New Mexico	11.0	4.0
Minnesota	29.0	4.0
Delaware	8.0	4.0
West Virginia	8.0	3.0
Connecticut	9.0	3.0
Kentucky	4.0	3.0
South Carolina	15.0	3.0
Maine	3.0	2.0
District Of Columbia	14.0	2.0
Kansas	1.0	1.0
Vermont	2.0	1.0
Missouri	3.0	1.0
Louisiana	12.0	1.0
Montana	1.0	0.0
Rhode Island	1.0	0.0
Ohio	3.0	0.0
District of Columbia	1.0	0.0
North Carolina	3.0	0.0
New York	6.0	0.0
Nevada	1.0	0.0
Arkansas	6.0	0.0
Iowa	1.0	0.0

```
In [56]:   #making another col. for unresolved complains percentage for further analysis
           unresolved_complains['unresolved_complains_percentage']=unresolved_complains["Open"]/unresolved_complains["Open"].sum()*100
```

```
In [57]:   unresolved_complains
```

Out[57]:

State	New_status Closed	Open	unresolved_complains_percentage
Georgia	208.0	80.0	15.473888
California	159.0	61.0	11.798839
Tennessee	96.0	47.0	9.090909
Florida	201.0	39.0	7.543520
Illinois	135.0	29.0	5.609284
Washington	75.0	23.0	4.448743
Michigan	92.0	23.0	4.448743
Colorado	58.0	22.0	4.255319
Texas	49.0	22.0	4.255319
Pennsylvania	110.0	20.0	3.868472
New Jersey	56.0	19.0	3.675048
Mississippi	23.0	16.0	3.094778
Maryland	63.0	15.0	2.901354
Oregon	36.0	13.0	2.514507
Virginia	49.0	11.0	2.127660
Massachusetts	50.0	11.0	2.127660
Alabama	17.0	9.0	1.740812
Indiana	50.0	9.0	1.740812
Utah	16.0	6.0	1.160542
Arizona	14.0	6.0	1.160542
New Hampshire	8.0	4.0	0.773694
New Mexico	11.0	4.0	0.773694
Minnesota	29.0	4.0	0.773694
Delaware	8.0	4.0	0.773694
West Virginia	8.0	3.0	0.580271
Connecticut	9.0	3.0	0.580271
Kentucky	4.0	3.0	0.580271
South Carolina	15.0	3.0	0.580271
Maine	3.0	2.0	0.386847
District Of Columbia	14.0	2.0	0.386847
Kansas	1.0	1.0	0.193424
Vermont	2.0	1.0	0.193424
Missouri	3.0	1.0	0.193424
Louisiana	12.0	1.0	0.193424
Montana	1.0	0.0	0.000000
Rhode Island	1.0	0.0	0.000000
Ohio	3.0	0.0	0.000000
District of Columbia	1.0	0.0	0.000000
North Carolina	3.0	0.0	0.000000
New York	6.0	0.0	0.000000
Nevada	1.0	0.0	0.000000
Arkansas	6.0	0.0	0.000000
Iowa	1.0	0.0	0.000000

```
In [58]:  df.head()
```

Out[58]:

	Ticket #	Customer Complaint	Date	Date_month_year	Time	Received Via	City	State	Zip code	Status	Filing on Behalf of Someone	month	New_status
0	250635	Comcast Cable Internet Speeds	2015-04-22	22-Apr-15	3:53:50 PM	Customer Care Call	Abingdon	Maryland	21009	Closed	No	April	Closed
1	223441	Payment Disappear - Service Got Disconnected	2015-04-08	04-Aug-15	10:22:56 AM	Internet	Acworth	Georgia	30102	Closed	No	April	Closed
2	242732	Speed And Service	2015-04-18	18-Apr-15	9:55:47 AM	Internet	Acworth	Georgia	30101	Closed	Yes	April	Closed
3	277946	Comcast Imposed A New Usage Cap Of 300Gb That ...	2015-05-07	05-Jul-15	11:59:35 AM	Internet	Acworth	Georgia	30101	Open	Yes	May	Open
4	307175	Comcast Not Working And No Service To Boot	2015-05-26	26-May-15	1:25:26 PM	Internet	Acworth	Georgia	30101	Solved	No	May	Closed

```
In [59]:  #getting information about complains  were received through the Internet and customer care calls.
          resolved_data_via= df.groupby(['Received Via','New_status']).size().unstack()
```

```
In [60]:  resolved_data_via
```

Out[60]:

New_status	Closed	Open
Received Via		
Customer Care Call	864	255
Internet	843	262

```
In [61]:  # percentage complains  were received through the Internet and customer care calls .
          resolved_data_via['percent_solved_complains']= resolved_data_via['Closed']/resolved_data_via['Closed'].sum()*100
```

```
In [62]:  resolved_data_via
```

Out[62]:

New_status	Closed	Open	percent_solved_complains
Received Via			
Customer Care Call	864	255	50.615114
Internet	843	262	49.384886

```
In [ ]:
```