lecom-consumer-complaints-project

June 18, 2024

[3]:
```
import math
import numpy as np
import matplotlib.pyplot as plt
import pandas as pd
import seaborn as sns
%matplotlib inline
```

[4]:
```
df=pd.read_csv('Comcast_telecom_complaints_data.csv')
```

[5]:
```
df
```

[5]:

	Ticket #	Customer Complaint	Date	Date_month_year	Time	Received Via	City	State
0	250635	Comcast Cable Internet Speeds	22-04-15	22-Apr-15	3:53:50 PM	Customer Care Call	Abingdon	Maryland
1	223441	Payment disappear - service got disconnected	04-08-15	04-Aug-15	10:22:56 AM	Internet	Acworth	Georgia
2	242732	Speed and Service	18-04-15	18-Apr-15	9:55:47 AM	Internet	Acworth	Georgia
3	277946	Comcast Imposed a New Usage Cap of 300GB that …	05-07-15	05-Jul-15	11:59:35 AM	Internet	Acworth	Georgia
4	307175	Comcast not working and no service to boot	26-05-15	26-May-15	1:25:26 PM	Internet	Acworth	Georgia
…	…	…	…	…	…	…	…	…
2219	213550	Service Availability	04-02-15	04-Feb-15	9:13:18 AM	Customer Care Call	Youngstown	Florida
2220	318775	Comcast Monthly Billing for Returned Modem	06-02-15	06-Feb-15	1:24:39 PM	Customer Care Call	Ypsilanti	Michigan
2221	331188	complaint about comcast	06-09-15	06-Sep-15	5:28:41 PM	Internet	Ypsilanti	Michigan
2222	360489	Extremely unsatisfied Comcast customer	23-06-15	23-Jun-15	11:13:30 PM	Customer Care Call	Ypsilanti	Michigan
2223	363614	Comcast, Ypsilanti MI Internet Speed	24-06-15	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]
```

[6]:
```
df.info
```

[6]:
```
<bound method DataFrame.info of      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]>
```

[7]:
```
df.head()
```

[7]:
```
  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
```

[8]:
```
df.tail()
```

[8]:
```
     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
```

Dropping any Duplicates

[9]:
```
df.drop_duplicates(inplace=True)
```

Renaming column with lowercase and replacing empty with _

[10]:
```
df.columns = list(map(lambda x: x.replace(' ','_').lower() , df.columns))
```

df

Renaming ticket_# with proper column name as 'ticketNo'

[11]:
```
df.rename(columns = {'ticket_#' : 'ticketNo'},inplace=True)
```

Replacing invalid entry in ticketNo column

[12]:
```
df.drop(index = df[df.ticketNo == 'comcas'].index,inplace=True)
```

Converting column data types

[13]:
```
df.ticketNo= df.ticketNo.astype(int)
df.zip_code= df.zip_code.astype(int)
df.customer_complaint=df.customer_complaint.astype(str)
```

Converting date column to dateTime format

[14]:
```
df["full_date"] = df["date_month_year"] + " " + df["time"]
df["date_month_year"] = pd.to_datetime(df["date_month_year"])
df["full_date"] = pd.to_datetime(df["full_date"])
```

Setting index to the newly created date

[15]:
```
updated_df = df.set_index(df["full_date"])
updated_df
```

[15]:
```
                     ticketNo  \
full_date
2015-04-22 15:53:50    250635
2015-08-04 10:22:56    223441
```

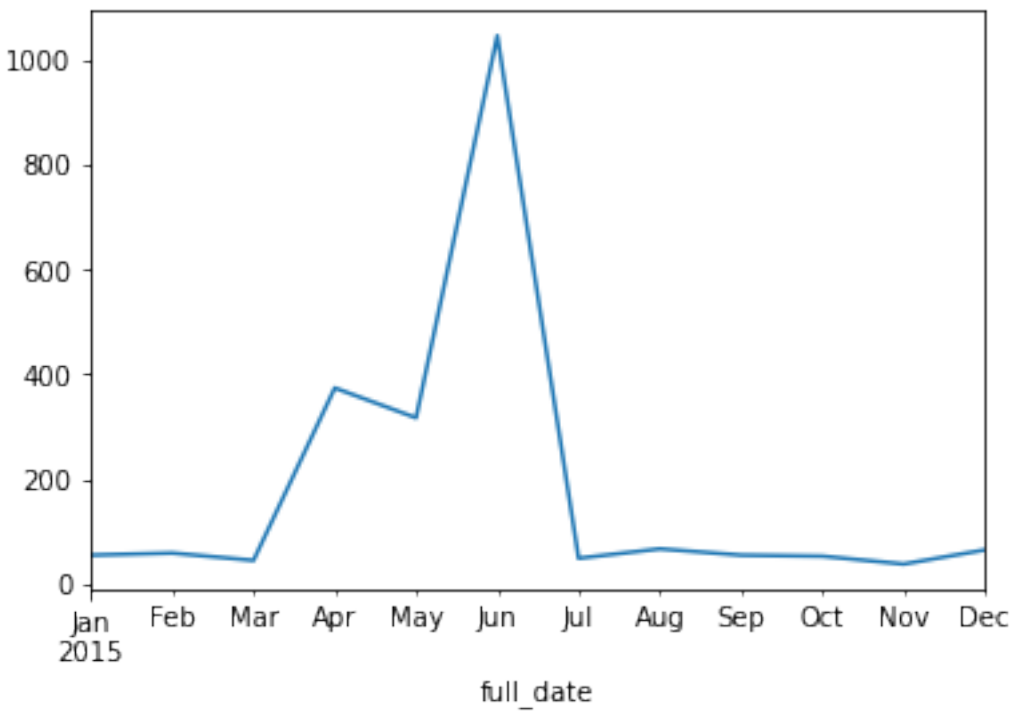
```
2015-04-18 09:55:47    242732
2015-07-05 11:59:35    277946
2015-05-26 13:25:26    307175
…                       …
2015-02-04 09:13:18    213550
2015-02-06 13:24:39    318775
2015-09-06 17:28:41    331188
2015-06-23 23:13:30    360489
2015-06-24 22:28:33    363614

                                                    customer_complaint  \
full_date
2015-04-22 15:53:50                        Comcast Cable Internet Speeds
2015-08-04 10:22:56       Payment disappear - service got disconnected
2015-04-18 09:55:47                                    Speed and Service
2015-07-05 11:59:35  Comcast Imposed a New Usage Cap of 300GB that …
2015-05-26 13:25:26         Comcast not working and no service to boot
…                                                                    …
2015-02-04 09:13:18                                  Service Availability
2015-02-06 13:24:39         Comcast Monthly Billing for Returned Modem
2015-09-06 17:28:41                              complaint about comcast
2015-06-23 23:13:30             Extremely unsatisfied Comcast customer
2015-06-24 22:28:33               Comcast, Ypsilanti MI Internet Speed

                        date date_month_year         time  \
full_date
2015-04-22 15:53:50  22-04-15      2015-04-22   3:53:50 PM
2015-08-04 10:22:56  04-08-15      2015-08-04  10:22:56 AM
2015-04-18 09:55:47  18-04-15      2015-04-18   9:55:47 AM
2015-07-05 11:59:35  05-07-15      2015-07-05  11:59:35 AM
2015-05-26 13:25:26  26-05-15      2015-05-26   1:25:26 PM
…                       …             …           …
2015-02-04 09:13:18  04-02-15      2015-02-04   9:13:18 AM
2015-02-06 13:24:39  06-02-15      2015-02-06   1:24:39 PM
2015-09-06 17:28:41  06-09-15      2015-09-06   5:28:41 PM
2015-06-23 23:13:30  23-06-15      2015-06-23  11:13:30 PM
2015-06-24 22:28:33  24-06-15      2015-06-24  10:28:33 PM

                           received_via        city     state  zip_code  \
full_date
2015-04-22 15:53:50  Customer Care Call    Abingdon  Maryland     21009
2015-08-04 10:22:56            Internet     Acworth   Georgia     30102
2015-04-18 09:55:47            Internet     Acworth   Georgia     30101
2015-07-05 11:59:35            Internet     Acworth   Georgia     30101
2015-05-26 13:25:26            Internet     Acworth   Georgia     30101
…                                   …         …       …        …
2015-02-04 09:13:18  Customer Care Call  Youngstown   Florida     32466
```

```
2015-02-06 13:24:39  Customer Care Call   Ypsilanti  Michigan     48197
2015-09-06 17:28:41           Internet   Ypsilanti  Michigan     48197
2015-06-23 23:13:30  Customer Care Call   Ypsilanti  Michigan     48197
2015-06-24 22:28:33  Customer Care Call   Ypsilanti  Michigan     48198

                     status filing_on_behalf_of_someone           full_date
full_date
2015-04-22 15:53:50  Closed                          No 2015-04-22 15:53:50
2015-08-04 10:22:56  Closed                          No 2015-08-04 10:22:56
2015-04-18 09:55:47  Closed                         Yes 2015-04-18 09:55:47
2015-07-05 11:59:35    Open                         Yes 2015-07-05 11:59:35
2015-05-26 13:25:26  Solved                          No 2015-05-26 13:25:26
…                     …                         …                 …
2015-02-04 09:13:18  Closed                          No 2015-02-04 09:13:18
2015-02-06 13:24:39  Solved                          No 2015-02-06 13:24:39
2015-09-06 17:28:41  Solved                          No 2015-09-06 17:28:41
2015-06-23 23:13:30  Solved                          No 2015-06-23 23:13:30
2015-06-24 22:28:33    Open                         Yes 2015-06-24 22:28:33

[2223 rows x 12 columns]
```

Creating Monthwise trend chart

[16]:
```
chart_monthly=updated_df.groupby(pd.Grouper(freq="M"))
chart_monthly.size().plot()
```

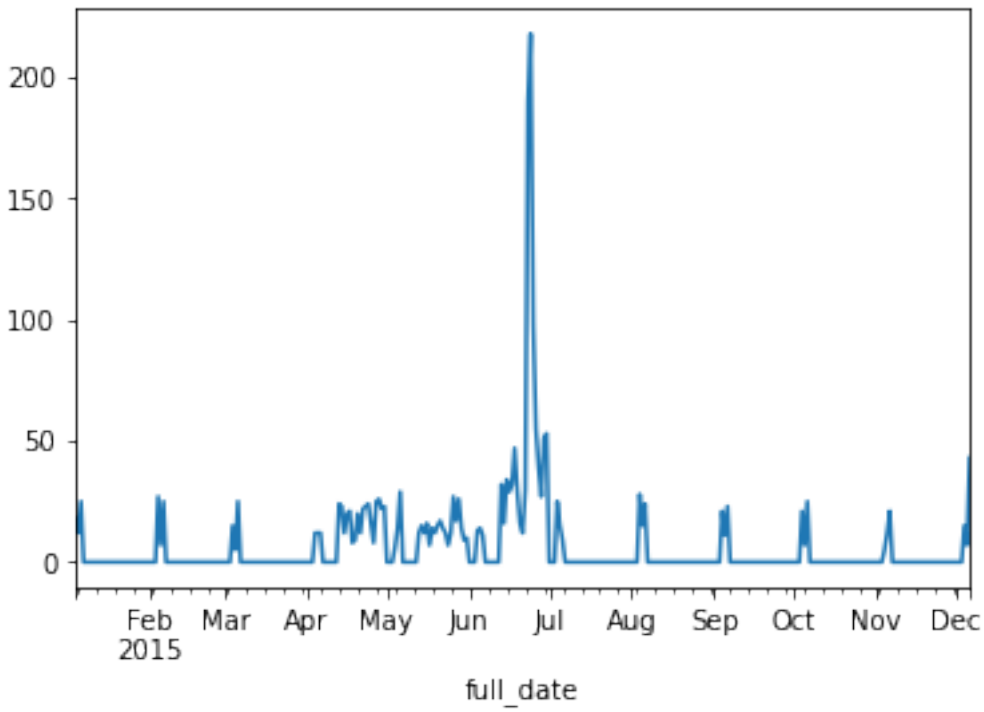
[16]:
```
<AxesSubplot:xlabel='full_date'>
```



Creating Daywise trend chart

[19]:
```
chart_daily=updated_df.groupby(pd.Grouper(freq="D"))
chart_daily.size().plot()
```

[19]: <AxesSubplot:xlabel='full_date'>



Creating a new categorical variable (new_status) out of existing status

[20]:
```
updated_df["new_status"] = ["Open" if i=="Open" or i=="Pending" else "Closed"␣
 ↪for i in updated_df['status']]

updated_df
```

[20]:
```
                     ticketNo  \
full_date
2015-04-22 15:53:50    250635
2015-08-04 10:22:56    223441
2015-04-18 09:55:47    242732
2015-07-05 11:59:35    277946
2015-05-26 13:25:26    307175
…                        …
2015-02-04 09:13:18    213550
2015-02-06 13:24:39    318775
2015-09-06 17:28:41    331188
2015-06-23 23:13:30    360489
2015-06-24 22:28:33    363614

                                                customer_complaint  \
full_date
2015-04-22 15:53:50                    Comcast Cable Internet Speeds
```

```
2015-08-04 10:22:56       Payment disappear - service got disconnected
2015-04-18 09:55:47                                    Speed and Service
2015-07-05 11:59:35  Comcast Imposed a New Usage Cap of 300GB that …
2015-05-26 13:25:26         Comcast not working and no service to boot
…                                                                    …
2015-02-04 09:13:18                                  Service Availability
2015-02-06 13:24:39         Comcast Monthly Billing for Returned Modem
2015-09-06 17:28:41                               complaint about comcast
2015-06-23 23:13:30             Extremely unsatisfied Comcast customer
2015-06-24 22:28:33               Comcast, Ypsilanti MI Internet Speed

                        date date_month_year         time  \
full_date
2015-04-22 15:53:50  22-04-15      2015-04-22   3:53:50 PM
2015-08-04 10:22:56  04-08-15      2015-08-04  10:22:56 AM
2015-04-18 09:55:47  18-04-15      2015-04-18   9:55:47 AM
2015-07-05 11:59:35  05-07-15      2015-07-05  11:59:35 AM
2015-05-26 13:25:26  26-05-15      2015-05-26   1:25:26 PM
…                       …             …          …
2015-02-04 09:13:18  04-02-15      2015-02-04   9:13:18 AM
2015-02-06 13:24:39  06-02-15      2015-02-06   1:24:39 PM
2015-09-06 17:28:41  06-09-15      2015-09-06   5:28:41 PM
2015-06-23 23:13:30  23-06-15      2015-06-23  11:13:30 PM
2015-06-24 22:28:33  24-06-15      2015-06-24  10:28:33 PM

                           received_via        city     state  zip_code  \
full_date
2015-04-22 15:53:50  Customer Care Call    Abingdon  Maryland     21009
2015-08-04 10:22:56            Internet     Acworth   Georgia     30102
2015-04-18 09:55:47            Internet     Acworth   Georgia     30101
2015-07-05 11:59:35            Internet     Acworth   Georgia     30101
2015-05-26 13:25:26            Internet     Acworth   Georgia     30101
…                                 …         …       …       …
2015-02-04 09:13:18  Customer Care Call  Youngstown   Florida     32466
2015-02-06 13:24:39  Customer Care Call   Ypsilanti  Michigan     48197
2015-09-06 17:28:41            Internet   Ypsilanti  Michigan     48197
2015-06-23 23:13:30  Customer Care Call   Ypsilanti  Michigan     48197
2015-06-24 22:28:33  Customer Care Call   Ypsilanti  Michigan     48198

                     status filing_on_behalf_of_someone           full_date  \
full_date
2015-04-22 15:53:50  Closed                          No 2015-04-22 15:53:50
2015-08-04 10:22:56  Closed                          No 2015-08-04 10:22:56
2015-04-18 09:55:47  Closed                         Yes 2015-04-18 09:55:47
2015-07-05 11:59:35    Open                         Yes 2015-07-05 11:59:35
2015-05-26 13:25:26  Solved                          No 2015-05-26 13:25:26
…                     …                         …                 …
```

```
2015-02-04 09:13:18  Closed                          No 2015-02-04 09:13:18
2015-02-06 13:24:39  Solved                          No 2015-02-06 13:24:39
2015-09-06 17:28:41  Solved                          No 2015-09-06 17:28:41
2015-06-23 23:13:30  Solved                          No 2015-06-23 23:13:30
2015-06-24 22:28:33    Open                         Yes 2015-06-24 22:28:33

                     new_status
full_date
2015-04-22 15:53:50     Closed
2015-08-04 10:22:56     Closed
2015-04-18 09:55:47     Closed
2015-07-05 11:59:35       Open
2015-05-26 13:25:26     Closed
…                          …
2015-02-04 09:13:18     Closed
2015-02-06 13:24:39     Closed
2015-09-06 17:28:41     Closed
2015-06-23 23:13:30     Closed
2015-06-24 22:28:33       Open

[2223 rows x 13 columns]
```

Distribution of statewise complaint count

[21]:
```
state_sorted=updated_df.groupby(["state"]).size().sort_values(ascending=False)
state_sorted=state_sorted.reset_index()
state_sorted=state_sorted.rename({0: "count"}, axis=1)
state_sorted
```

[21]:

	state	count
0	Georgia	288
1	Florida	240
2	California	220
3	Illinois	164
4	Tennessee	142
5	Pennsylvania	130
6	Michigan	115
7	Washington	98
8	Colorado	80
9	Maryland	78
10	New Jersey	75
11	Texas	71
12	Massachusetts	61
13	Virginia	60
14	Indiana	59
15	Oregon	49
16	Mississippi	39

```
17              Minnesota     33
18                Alabama     26
19                   Utah     22
20                Arizona     20
21         South Carolina     18
22  District Of Columbia     16
23             New Mexico     15
24              Louisiana     13
25          New Hampshire     12
26            Connecticut     12
27               Delaware     12
28          West Virginia     11
29               Kentucky      7
30               New York      6
31               Arkansas      6
32                  Maine      5
33               Missouri      4
34         North Carolina      3
35                Vermont      3
36                   Ohio      3
37                 Kansas      2
38  District of Columbia      1
39           Rhode Island      1
40                Montana      1
41                   Iowa      1
42                 Nevada      1
```

State with maximum complaints

[22]:
```
print(state_sorted.max()['state'],'has maximum open complaints with count :
 ↪',state_sorted.max()['count'])
```

```
West Virginia has maximum open complaints with count : 288
```

Distribution of statewise based on new status

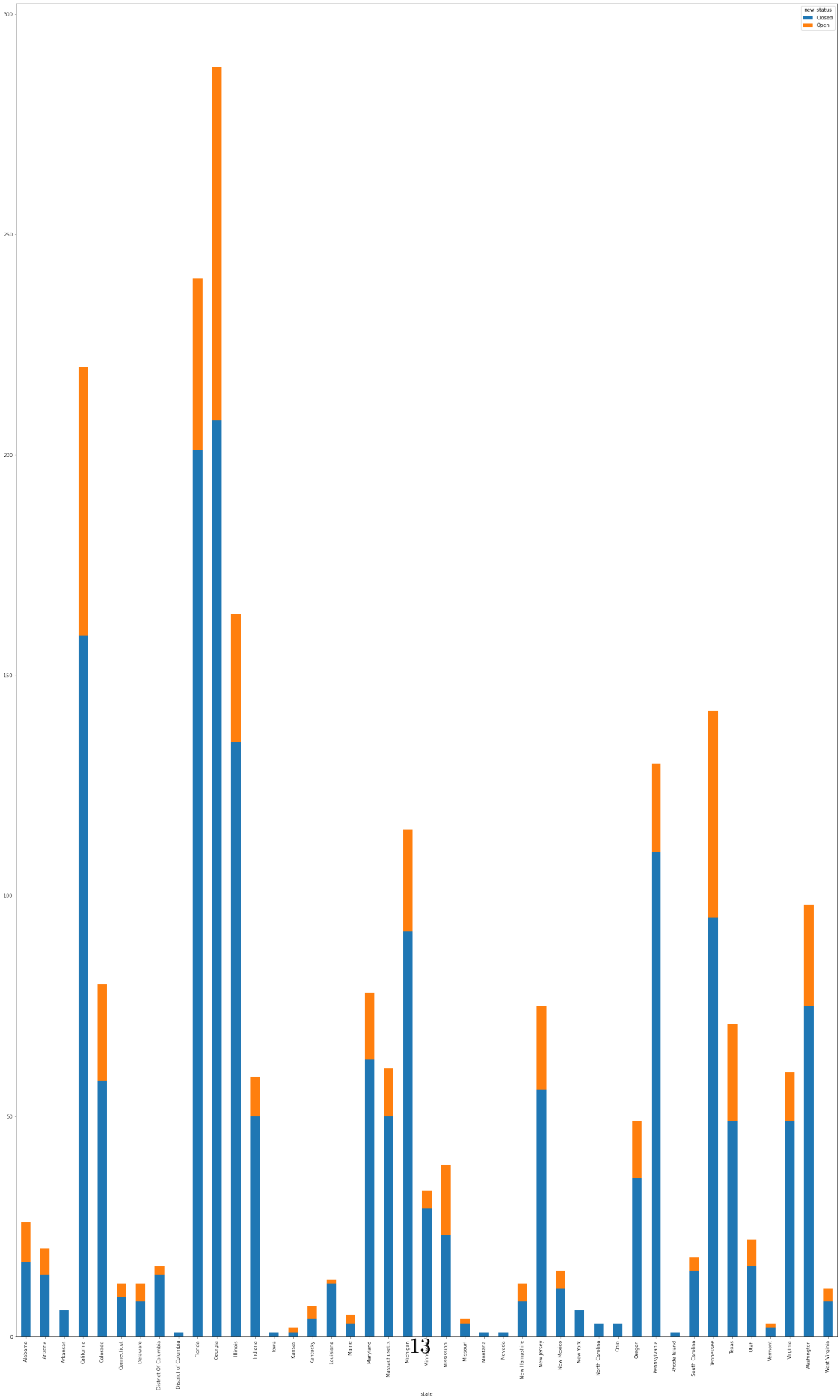
[23]:
```
complaint_status = updated_df.groupby(["state","new_status"]).size().unstack().
 ↪fillna(0)
complaint_status
```

[23]:
```
new_status            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
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               95.0  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
```

[24]:
```
complaint_status.plot(kind="bar", figsize=(30,50), stacked=True)
```

[24]:
```
<AxesSubplot:xlabel='state'>
```


new_status
Closed
Open
300
250
200
150
100
50
0
Alabama
Arizona
Arkansas
California
Colorado
Connecticut
Delaware
District Of Columbia
District of Columbia
Florida
Georgia
Illinois
Indiana
Iowa
Kansas
Kentucky
Louisiana
Maine
Maryland
Massachusetts
Michigan
Minnesota
Mississippi
Missouri
Montana
Nevada
New Hampshire
New Jersey
New Mexico
New York
North Carolina
Ohio
Oregon
Pennsylvania
Rhode Island
South Carolina
Tennessee
Texas
Utah
Vermont
Virginia
Washington
West Virginia
state

state wise all complaint percentages

[25]:
```
state_sorted['cumulative_sum'] = state_sorted['count'].cumsum()
state_sorted['percent_contribution']=100*state_sorted['count'].round(1)/
 ↪state_sorted['cumulative_sum'].max()
state_sorted
```

[25]:

	state	count	cumulative_sum	percent_contribution
0	Georgia	288	288	12.955466
1	Florida	240	528	10.796221
2	California	220	748	9.896536
3	Illinois	164	912	7.377418
4	Tennessee	142	1054	6.387764
5	Pennsylvania	130	1184	5.847953
6	Michigan	115	1299	5.173189
7	Washington	98	1397	4.408457
8	Colorado	80	1477	3.598740
9	Maryland	78	1555	3.508772
10	New Jersey	75	1630	3.373819
11	Texas	71	1701	3.193882
12	Massachusetts	61	1762	2.744040
13	Virginia	60	1822	2.699055
14	Indiana	59	1881	2.654071
15	Oregon	49	1930	2.204229
16	Mississippi	39	1969	1.754386
17	Minnesota	33	2002	1.484480
18	Alabama	26	2028	1.169591
19	Utah	22	2050	0.989654
20	Arizona	20	2070	0.899685
21	South Carolina	18	2088	0.809717
22	District Of Columbia	16	2104	0.719748
23	New Mexico	15	2119	0.674764
24	Louisiana	13	2132	0.584795
25	New Hampshire	12	2144	0.539811
26	Connecticut	12	2156	0.539811
27	Delaware	12	2168	0.539811
28	West Virginia	11	2179	0.494827
29	Kentucky	7	2186	0.314890
30	New York	6	2192	0.269906
31	Arkansas	6	2198	0.269906
32	Maine	5	2203	0.224921
33	Missouri	4	2207	0.179937
34	North Carolina	3	2210	0.134953
35	Vermont	3	2213	0.134953
36	Ohio	3	2216	0.134953

37	Kansas	2	2218	0.089969
38	District of Columbia	1	2219	0.044984
39	Rhode Island	1	2220	0.044984
40	Montana	1	2221	0.044984
41	Iowa	1	2222	0.044984
42	Nevada	1	2223	0.044984

state wise open complaint percentages

[26]:
```
status_summary=updated_df.groupby(["state","new_status"]).size().unstack().
 ↪fillna(0)
total_open_complaints=status_summary['Open'].sum()
```

Analysis comment

[27]:
```
print('Total open complaints',total_open_complaints)
```

```
Total open complaints 517.0
```

[28]:
```
complaint_status['percent_open_complaints']=(100*complaint_status['Open']/
 ↪total_open_complaints).round(1)
complaint_status
```

[28]:

new_status	Closed	Open	percent_open_complaints
state			
Alabama	17.0	9.0	1.7
Arizona	14.0	6.0	1.2
Arkansas	6.0	0.0	0.0
California	159.0	61.0	11.8
Colorado	58.0	22.0	4.3
Connecticut	9.0	3.0	0.6
Delaware	8.0	4.0	0.8
District Of Columbia	14.0	2.0	0.4
District of Columbia	1.0	0.0	0.0
Florida	201.0	39.0	7.5
Georgia	208.0	80.0	15.5
Illinois	135.0	29.0	5.6
Indiana	50.0	9.0	1.7
Iowa	1.0	0.0	0.0
Kansas	1.0	1.0	0.2
Kentucky	4.0	3.0	0.6
Louisiana	12.0	1.0	0.2
Maine	3.0	2.0	0.4
Maryland	63.0	15.0	2.9
Massachusetts	50.0	11.0	2.1
Michigan	92.0	23.0	4.4
Minnesota	29.0	4.0	0.8
Mississippi	23.0	16.0	3.1

```
Missouri                3.0   1.0                     0.2
Montana                 1.0   0.0                     0.0
Nevada                  1.0   0.0                     0.0
New Hampshire           8.0   4.0                     0.8
New Jersey             56.0  19.0                     3.7
New Mexico             11.0   4.0                     0.8
New York                6.0   0.0                     0.0
North Carolina          3.0   0.0                     0.0
Ohio                    3.0   0.0                     0.0
Oregon                 36.0  13.0                     2.5
Pennsylvania          110.0  20.0                     3.9
Rhode Island            1.0   0.0                     0.0
South Carolina         15.0   3.0                     0.6
Tennessee              95.0  47.0                     9.1
Texas                  49.0  22.0                     4.3
Utah                   16.0   6.0                     1.2
Vermont                 2.0   1.0                     0.2
Virginia               49.0  11.0                     2.1
Washington             75.0  23.0                     4.4
West Virginia           8.0   3.0                     0.6
```

Top 3 states with open complaint %

[29]:
```
complaint_status.
 ↪sort_values(by='percent_open_complaints',ascending=False,inplace=True)
complaint_status.head(3)
```

[29]:
```
new_status  Closed  Open  percent_open_complaints
state
Georgia      208.0  80.0                    15.5
California   159.0  61.0                    11.8
Tennessee     95.0  47.0                     9.1
```

Analysis comment

[30]:
```
print('State ',complaint_status.iloc[0].name, ' has maximum open cases of␣
 ↪',complaint_status.iloc[0]['Open'])
```

```
State  Georgia  has maximum open cases of  80.0
```

Percentage of complaints resolved till date, which were received through the Internet and customer care calls.

[31]:
```
recieved_via_call_or_internet = updated_df.
 ↪groupby(['received_via','new_status']).size().unstack().fillna(0)
total_closed=recieved_via_call_or_internet['Closed'].sum()
recieved_via_call_or_internet['percentage_closed']=(100*recieved_via_call_or_internet['Closed'
 ↪total_closed).round(1)
```

```
recieved_via_call_or_internet
```

[31]:
```
new_status          Closed  Open  percentage_closed
received_via
Customer Care Call     863   255               50.6
Internet               843   262               49.4
```

Analysis comment

[33]:
```
print('Complaints recieved by ',recieved_via_call_or_internet.iloc[0].name, '␣
↪has maximum closed percentage of ',recieved_via_call_or_internet.iloc[0].
↪name)
```

```
Complaints recieved by  Customer Care Call  has maximum closed percentage of
Customer Care Call
```

New categorical variable: Complaint type

[34]:
```
updated_df.customer_complaint=updated_df.customer_complaint.str.lower()
wordcounts=updated_df.customer_complaint.str.split(expand=True).stack().
↪value_counts()
wordcounts
```

[34]:
```
comcast             1160
internet             508
service              411
and                  277
billing              273
                    …
same-day               1
practices-unfair       1
services/billing       1
charge.                1
it.                    1
Length: 1806, dtype: int64
```

[ ]:
```

```