comcast-telecom-project

November 13, 2021

0.1 DESCRIPTION

Comcast is an American global telecommunication company. The firm has been providing terrible customer service. They continue to fall short despite repeated promises to improve. Only last month (October 2016) the authority fined them a $2.3 million, after receiving over 1000 consumer complaints. The existing database will serve as a repository of public customer complaints filed against Comcast. It will help to pin down **what is wrong with Comcast's customer service**.

[1]:
```
import pandas as pd
import numpy as np
```

[2]:
```
df = pd.read_csv('Comcast_telecom_complaints_data.csv')
df.head()
```

[2]:

	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

[3]:
```
df.dtypes
```

[3]:
```
Ticket #                         object
Customer Complaint               object
Date                             object
Date_month_year                  object
Time                             object
Received Via                     object
City                             object
State                            object
Zip code                          int64
Status                           object
Filing on Behalf of Someone      object
dtype: object
```

[4]:
```
# change data types to datetime
```

[5]:
```
df['Date'] = pd.to_datetime(df['Date'])
df['Date_month_year'] = pd.to_datetime(df['Date_month_year'])
df['Time'] = pd.to_datetime(df['Time'])
df.dtypes
```

[5]:
```
Ticket #                                object
Customer Complaint                      object
Date                           datetime64[ns]
Date_month_year                datetime64[ns]
Time                           datetime64[ns]
Received Via                            object
City                                    object
State                                   object
Zip code                                 int64
Status                                  object
Filing on Behalf of Someone             object
dtype: object
```

[6]:
```
# change 'Zip code' to object type from int64 since it's a categorical variable␣
 ↪and not numerical
df['Zip code'] = df['Zip code'].astype('object')
df.dtypes
```

[6]:
```
Ticket #                                object
Customer Complaint                      object
Date                           datetime64[ns]
Date_month_year                datetime64[ns]
Time                           datetime64[ns]
Received Via                            object
City                                    object
State                                   object
Zip code                                object
```

```
Status                                object
Filing on Behalf of Someone           object
dtype: object
```

[7]:
```
# check for null values
df.isna().sum()
```

[7]:
```
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
```

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

[8]:

	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	2015-04-22	2015-04-22	2021-11-13 15:53:50	Customer Care Call	Abingdon	Maryland	21009	Closed	No
1	223441	Payment disappear - service got disconnected	2015-04-08	2015-08-04	2021-11-13 10:22:56	Internet	Acworth	Georgia	30102	Closed	No
2	242732	Speed and Service	2015-04-18	2015-04-18	2021-11-13 09:55:47	Internet	Acworth	Georgia	30101	Closed	Yes
3	277946	Comcast Imposed a New Usage Cap of 300GB that …	2015-05-07	2015-07-05	2021-11-13 11:59:35	Internet	Acworth	Georgia	30101	Open	Yes
4	307175	Comcast not working and no service to boot	2015-05-26	2015-05-26	2021-11-13 13:25:26	Internet	Acworth	Georgia	30101	Solved	No

[9]:
```
# remove date prepended to 'Time' variable
df['Time'] = df['Time'].dt.time
df.head()
```

[9]:
```
  Ticket #                                  Customer Complaint       Date  \
0   250635                       Comcast Cable Internet Speeds 2015-04-22
1   223441       Payment disappear - service got disconnected 2015-04-08
2   242732                                    Speed and Service 2015-04-18
3   277946  Comcast Imposed a New Usage Cap of 300GB that … 2015-05-07
4   307175         Comcast not working and no service to boot 2015-05-26

  Date_month_year      Time        Received Via      City     State Zip code  \
0      2015-04-22  15:53:50  Customer Care Call  Abingdon  Maryland    21009
1      2015-08-04  10:22:56            Internet   Acworth   Georgia    30102
2      2015-04-18  09:55:47            Internet   Acworth   Georgia    30101
3      2015-07-05  11:59:35            Internet   Acworth   Georgia    30101
4      2015-05-26  13:25:26            Internet   Acworth   Georgia    30101

   Status Filing on Behalf of Someone
0  Closed                         No
1  Closed                         No
2  Closed                        Yes
3    Open                        Yes
4  Solved                         No
```

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

0.1.2 Monthly granularity trend chart

[10]:
```
import altair as alt
```

[11]:
```
# df to hold monthly complaints
monthly_complaints = pd.DataFrame(df['Customer Complaint'].
 ↪groupby(df['Date_month_year'].dt.month).count())
print(monthly_complaints)
```

```
                Customer Complaint
Date_month_year
1                               55
2                               59
3                               45
4                              375
5                              317
6                             1046
7                               49
8                               67
9                               55
10                              53
11                              38
12                              65
```

[12]:
```
monthly_complaints.reset_index(inplace=True)
print(monthly_complaints.columns)
```

```
Index(['Date_month_year', 'Customer Complaint'], dtype='object')
```

[13]:
```
monthly_complaints.rename(columns={'Date_month_year': 'month', 'Customer␣
↪Complaint': 'complaints'}, inplace=True)
monthly_complaints
```

[13]:

	month	complaints
0	1	55
1	2	59
2	3	45
3	4	375
4	5	317
5	6	1046
6	7	49
7	8	67
8	9	55
9	10	53
10	11	38
11	12	65

[14]:
```
alt.Chart(monthly_complaints).mark_line().encode(x='month', y='complaints',␣
↪tooltip=['month', 'complaints']).interactive()
```

[14]:
```
alt.Chart(…)
```

0.1.3 Daily granularity trend chart

[15]:
```
# df to hold daily complaints
daily_complaints = pd.DataFrame(df['Customer Complaint'].groupby(df['Date'].dt.
↪day).count())
print(daily_complaints)
```

```
      Customer Complaint
Date
1                     55
2                     59
3                     45
4                     36
5                     49
6                     38
7                     49
8                     67
9                     55
10                    53
11                    38
```

```
12                    65
13                    68
14                    54
15                    58
16                    65
17                    60
18                    69
19                    50
20                    51
21                    41
22                    66
23                   225
24                   249
25                   126
26                    90
27                    81
28                    79
29                    87
30                    86
31                    10
```

[16]:
```
daily_complaints.reset_index(inplace=True)
print(daily_complaints.columns)
```

```
Index(['Date', 'Customer Complaint'], dtype='object')
```

[17]:
```
daily_complaints.rename(columns={'Date': 'day', 'Customer Complaint':␣
↪'complaints'}, inplace=True)
daily_complaints
```

[17]:

	day	complaints
0	1	55
1	2	59
2	3	45
3	4	36
4	5	49
5	6	38
6	7	49
7	8	67
8	9	55
9	10	53
10	11	38
11	12	65
12	13	68
13	14	54
14	15	58
15	16	65
16	17	60

```
17   18          69
18   19          50
19   20          51
20   21          41
21   22          66
22   23         225
23   24         249
24   25         126
25   26          90
26   27          81
27   28          79
28   29          87
29   30          86
30   31          10
```

[18]:
```
alt.Chart(daily_complaints).mark_line().encode(x='day:O', y='complaints',␣
↪tooltip=['day', 'complaints']).interactive()
```

[18]:
```
alt.Chart(…)
```

We can see that the 24th day of each month has the most complaints for all the months. This is something that can be drilled down to find out what Comcast carries out around end of month with regards to their services.

0.1.4 Frequency of complaint types

[19]:
```
df.columns
```

[19]:
```
Index(['Ticket #', 'Customer Complaint', 'Date', 'Date_month_year', 'Time',
       'Received Via', 'City', 'State', 'Zip code', 'Status',
       'Filing on Behalf of Someone'],
      dtype='object')
```

[20]:
```
complaint_frequency = pd.DataFrame(df['Customer Complaint'].value_counts())
complaint_frequency
```

[20]:

	Customer Complaint
Comcast	83
Comcast Internet	18
Comcast Data Cap	17
comcast	13
Data Caps	11
…	…
COMCAST SERVICE	1
Comcast Internet and cable deals	1
Comcast trial of putting cap (300 GB / month) o…	1
Comcast Internet problems	1
credit for service outage 6/23/15	1

```
[1841 rows x 1 columns]
```

'Internet' and 'data cap' are the domains with the most complaints aisde from 'Comcast' itself according to the output above.

0.1.5 Create a new categorical variable with value as Open and Closed

[21]:
```
df.head()
```

[21]:

	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	2015-04-22	2015-04-22	15:53:50	Customer Care Call	Abingdon	Maryland	21009	Closed	No
1	223441	Payment disappear - service got disconnected	2015-04-08	2015-08-04	10:22:56	Internet	Acworth	Georgia	30102	Closed	No
2	242732	Speed and Service	2015-04-18	2015-04-18	09:55:47	Internet	Acworth	Georgia	30101	Closed	Yes
3	277946	Comcast Imposed a New Usage Cap of 300GB that …	2015-05-07	2015-07-05	11:59:35	Internet	Acworth	Georgia	30101	Open	Yes
4	307175	Comcast not working and no service to boot	2015-05-26	2015-05-26	13:25:26	Internet	Acworth	Georgia	30101	Solved	No

[22]:
```
# address duplicate
df.loc[df['State'] == 'District of Columbia', 'State'] = 'District Of Columbia'
```

[23]:
```
# create new categorical variable
df['Status_trans'] = df['Status']
df.head()
```

[23]:

	Ticket #	Customer Complaint	Date	Date_month_year	Time	Received Via	City	State	Zip code	Status	Filing on Behalf of Someone	Status_trans
0	250635	Comcast Cable Internet Speeds	2015-04-22	2015-04-22	15:53:50	Customer Care Call	Abingdon	Maryland	21009	Closed	No	Closed
1	223441	Payment disappear - service got disconnected	2015-04-08	2015-08-04	10:22:56	Internet	Acworth	Georgia	30102	Closed	No	Closed
2	242732	Speed and Service	2015-04-18	2015-04-18	09:55:47	Internet	Acworth	Georgia	30101	Closed	Yes	Closed
3	277946	Comcast Imposed a New Usage Cap of 300GB that …	2015-05-07	2015-07-05	11:59:35	Internet	Acworth	Georgia	30101	Open	Yes	Open
4	307175	Comcast not working and no service to boot	2015-05-26	2015-05-26	13:25:26	Internet	Acworth	Georgia	30101	Solved	No	Solved

```
2      2015-04-18  09:55:47           Internet   Acworth   Georgia    30101
3      2015-07-05  11:59:35           Internet   Acworth   Georgia    30101
4      2015-05-26  13:25:26           Internet   Acworth   Georgia    30101

   Status Filing on Behalf of Someone Status_trans
0  Closed                         No       Closed
1  Closed                         No       Closed
2  Closed                        Yes       Closed
3    Open                        Yes         Open
4  Solved                         No       Solved
```

[24]:
```
df.info()
```

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 2224 entries, 0 to 2223
Data columns (total 12 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   datetime64[ns]
 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   object
 9   Status                        2224 non-null   object
 10  Filing on Behalf of Someone   2224 non-null   object
 11  Status_trans                  2224 non-null   object
dtypes: datetime64[ns](2), object(10)
memory usage: 208.6+ KB
```

[25]:
```
# change Pending to Open
df.loc[df['Status'] == 'Pending', 'Status_trans'] = 'Open'
```

[26]:
```
df.head(10)
```

[26]:
```
  Ticket #                                Customer Complaint       Date  \
0   250635                    Comcast Cable Internet Speeds 2015-04-22
1   223441       Payment disappear - service got disconnected 2015-04-08
2   242732                                 Speed and Service 2015-04-18
3   277946  Comcast Imposed a New Usage Cap of 300GB that … 2015-05-07
4   307175         Comcast not working and no service to boot 2015-05-26
5   338519  ISP Charging for arbitrary data limits with ov… 2015-06-12
6   361148      Throttling service and unreasonable data caps 2015-06-24
7   359792  Comcast refuses to help troubleshoot and corre… 2015-06-23
8   318072                          Comcast extended outages 2015-06-01
```

```
9   371214  Comcast Raising Prices and Not Being Available… 2015-06-28

  Date_month_year      Time        Received Via      City       State  \
0      2015-04-22  15:53:50  Customer Care Call  Abingdon    Maryland
1      2015-08-04  10:22:56            Internet   Acworth     Georgia
2      2015-04-18  09:55:47            Internet   Acworth     Georgia
3      2015-07-05  11:59:35            Internet   Acworth     Georgia
4      2015-05-26  13:25:26            Internet   Acworth     Georgia
5      2015-12-06  21:59:40            Internet   Acworth     Georgia
6      2015-06-24  10:13:55  Customer Care Call   Acworth     Georgia
7      2015-06-23  18:56:14            Internet    Adrian    Michigan
8      2015-01-06  23:46:30  Customer Care Call   Alameda  California
9      2015-06-28  18:46:31  Customer Care Call   Alameda  California

  Zip code   Status Filing on Behalf of Someone Status_trans
0    21009   Closed                          No       Closed
1    30102   Closed                          No       Closed
2    30101   Closed                         Yes       Closed
3    30101     Open                         Yes         Open
4    30101   Solved                          No       Solved
5    30101   Solved                          No       Solved
6    30101  Pending                          No         Open
7    49221   Solved                          No       Solved
8    94502   Closed                          No       Closed
9    94501     Open                         Yes         Open
```

[27]:
```
# change Solved to Closed
df.loc[df['Status'] == 'Solved', 'Status_trans'] = 'Closed'
```

[28]:
```
df.head(10)
```

[28]:
```
  Ticket #                                Customer Complaint       Date  \
0   250635                    Comcast Cable Internet Speeds 2015-04-22
1   223441      Payment disappear - service got disconnected 2015-04-08
2   242732                                 Speed and Service 2015-04-18
3   277946  Comcast Imposed a New Usage Cap of 300GB that … 2015-05-07
4   307175        Comcast not working and no service to boot 2015-05-26
5   338519  ISP Charging for arbitrary data limits with ov… 2015-06-12
6   361148     Throttling service and unreasonable data caps 2015-06-24
7   359792  Comcast refuses to help troubleshoot and corre… 2015-06-23
8   318072                           Comcast extended outages 2015-06-01
9   371214  Comcast Raising Prices and Not Being Available… 2015-06-28

  Date_month_year      Time        Received Via      City       State  \
0      2015-04-22  15:53:50  Customer Care Call  Abingdon    Maryland
1      2015-08-04  10:22:56            Internet   Acworth     Georgia
2      2015-04-18  09:55:47            Internet   Acworth     Georgia
```

```
3      2015-07-05  11:59:35            Internet   Acworth     Georgia
4      2015-05-26  13:25:26            Internet   Acworth     Georgia
5      2015-12-06  21:59:40            Internet   Acworth     Georgia
6      2015-06-24  10:13:55  Customer Care Call   Acworth     Georgia
7      2015-06-23  18:56:14            Internet    Adrian    Michigan
8      2015-01-06  23:46:30  Customer Care Call   Alameda  California
9      2015-06-28  18:46:31  Customer Care Call   Alameda  California
```

	Zip code	Status	Filing on Behalf of Someone	Status_trans
0	21009	Closed	No	Closed
1	30102	Closed	No	Closed
2	30101	Closed	Yes	Closed
3	30101	Open	Yes	Open
4	30101	Solved	No	Closed
5	30101	Solved	No	Closed
6	30101	Pending	No	Open
7	49221	Solved	No	Closed
8	94502	Closed	No	Closed
9	94501	Open	Yes	Open

0.1.6 Provide state wise status of complaints in a stacked bar chart. Use the categorized variable from pervious step. Provide insights on

- Which state has the maximum complaints
- Which state has the highest percentage of unresolved complaints

[29]:
```
# create df with state and Status_trans
state_df = pd.DataFrame(df[['State', 'Status_trans']].value_counts())
```

[30]:
```
state_df.head(10)
```

[30]:

		0
State	Status_trans	
Georgia	Closed	208
Florida	Closed	201
California	Closed	159
Illinois	Closed	135
Pennsylvania	Closed	110
Tennessee	Closed	96
Michigan	Closed	92
Georgia	Open	80
Washington	Closed	75
Maryland	Closed	63

[31]:
```
state_df.reset_index(inplace=True)
```

[32]:
```
state_df.rename(columns={0:'total complaints'}, inplace=True)
```

[33]:
```
state_df.head(10)
```

[33]:

	State	Status_trans	total complaints
0	Georgia	Closed	208
1	Florida	Closed	201
2	California	Closed	159
3	Illinois	Closed	135
4	Pennsylvania	Closed	110
5	Tennessee	Closed	96
6	Michigan	Closed	92
7	Georgia	Open	80
8	Washington	Closed	75
9	Maryland	Closed	63

[34]:
```
statewise_complaints = pd.DataFrame(state_df.groupby(['State','Status_trans']).
↪aggregate(sum))
```

[35]:
```
statewise_complaints.reset_index(inplace=True)
statewise_complaints.head(10)
```

[35]:

	State	Status_trans	total complaints
0	Alabama	Closed	17
1	Alabama	Open	9
2	Arizona	Closed	14
3	Arizona	Open	6
4	Arkansas	Closed	6
5	California	Closed	159
6	California	Open	61
7	Colorado	Closed	58
8	Colorado	Open	22
9	Connecticut	Closed	9

[36]:
```
alt.Chart(statewise_complaints).mark_bar().encode(
    x='State',
    y='total complaints',
    color=alt.Color('Status_trans'),
    tooltip=['State', 'total complaints', 'Status_trans']).interactive()
```

[36]:
```
alt.Chart(…)
```

0.1.7 Which state has the maximum complaints?

- Georgia has the maximum complaints 288 (208 + 80: Closed,Open)

[37]:
```
state_df.groupby('State').aggregate('total complaints').sum()
```

[37]:
```
State
Alabama                 26
```

```
Arizona                  20
Arkansas                  6
California              220
Colorado                 80
Connecticut              12
Delaware                 12
District Of Columbia     17
Florida                 240
Georgia                 288
Illinois                164
Indiana                  59
Iowa                      1
Kansas                    2
Kentucky                  7
Louisiana                13
Maine                     5
Maryland                 78
Massachusetts            61
Michigan                115
Minnesota                33
Mississippi              39
Missouri                  4
Montana                   1
Nevada                    1
New Hampshire            12
New Jersey               75
New Mexico               15
New York                  6
North Carolina            3
Ohio                      3
Oregon                   49
Pennsylvania            130
Rhode Island              1
South Carolina           18
Tennessee               143
Texas                    71
Utah                     22
Vermont                   3
Virginia                 60
Washington               98
West Virginia            11
Name: total complaints, dtype: int64
```

0.1.8 Which state has the highest percentage of unresolved complaints?

[38]:
```
df_new = df
df_new.head(10)
```

[38]:

	Ticket #	Customer Complaint	Date	Date_month_year	Time	Received Via	City	State	Zip code	Status	Filing on Behalf of Someone	Status_trans
0	250635	Comcast Cable Internet Speeds	2015-04-22	2015-04-22	15:53:50	Customer Care Call	Abingdon	Maryland	21009	Closed	No	Closed
1	223441	Payment disappear - service got disconnected	2015-04-08	2015-08-04	10:22:56	Internet	Acworth	Georgia	30102	Closed	No	Closed
2	242732	Speed and Service	2015-04-18	2015-04-18	09:55:47	Internet	Acworth	Georgia	30101	Closed	Yes	Closed
3	277946	Comcast Imposed a New Usage Cap of 300GB that …	2015-05-07	2015-07-05	11:59:35	Internet	Acworth	Georgia	30101	Open	Yes	Open
4	307175	Comcast not working and no service to boot	2015-05-26	2015-05-26	13:25:26	Internet	Acworth	Georgia	30101	Solved	No	Closed
5	338519	ISP Charging for arbitrary data limits with ov…	2015-06-12	2015-12-06	21:59:40	Internet	Acworth	Georgia	30101	Solved	No	Closed
6	361148	Throttling service and unreasonable data caps	2015-06-24	2015-06-24	10:13:55	Customer Care Call	Acworth	Georgia	30101	Pending	No	Open
7	359792	Comcast refuses to help troubleshoot and corre…	2015-06-23	2015-06-23	18:56:14	Internet	Adrian	Michigan	49221	Solved	No	Closed
8	318072	Comcast extended outages	2015-06-01	2015-01-06	23:46:30	Customer Care Call	Alameda	California	94502	Closed	No	Closed
9	371214	Comcast Raising Prices and Not Being Available…	2015-06-28	2015-06-28	18:46:31	Customer Care Call	Alameda	California	94501	Open	Yes	Open

[39]:
```
# create dummy variables for Status_trans in order to convert to numerical.
df_new[['Closed', 'Open']] = pd.get_dummies(df_new['Status_trans'])
df_new.head()
```

[39]:

	Ticket #	Customer Complaint	Date	Date_month_year	Time	Received Via	City	State	Zip code	Status	Filing on Behalf of Someone	Status_trans	Closed	Open
0	250635	Comcast Cable Internet Speeds	2015-04-22	2015-04-22	15:53:50	Customer Care Call	Abingdon	Maryland	21009	Closed	No	Closed	1	0
1	223441	Payment disappear - service got disconnected	2015-04-08	2015-08-04	10:22:56	Internet	Acworth	Georgia	30102	Closed	No	Closed	1	0

```
2   242732                                  Speed and Service 2015-04-18
3   277946  Comcast Imposed a New Usage Cap of 300GB that … 2015-05-07
4   307175         Comcast not working and no service to boot 2015-05-26

  Date_month_year      Time        Received Via      City     State Zip code  \
0      2015-04-22  15:53:50  Customer Care Call  Abingdon  Maryland    21009
1      2015-08-04  10:22:56            Internet   Acworth   Georgia    30102
2      2015-04-18  09:55:47            Internet   Acworth   Georgia    30101
3      2015-07-05  11:59:35            Internet   Acworth   Georgia    30101
4      2015-05-26  13:25:26            Internet   Acworth   Georgia    30101

   Status Filing on Behalf of Someone Status_trans  Closed  Open
0  Closed                          No       Closed       1     0
1  Closed                          No       Closed       1     0
2  Closed                         Yes       Closed       1     0
3    Open                         Yes         Open       0     1
4  Solved                          No       Closed       1     0
```

[40]:
```
# separate the complaints
closed_complaints = df_new.groupby('State').aggregate('Closed').sum()
open_complaints = df_new.groupby('State').aggregate('Open').sum()
print(closed_complaints.shape)
print(open_complaints.shape)
```

```
(42,)
(42,)
```

[41]:
```
open_complaints.reset_index()
closed_complaints.reset_index()
```

[41]:

	State	Closed
0	Alabama	17
1	Arizona	14
2	Arkansas	6
3	California	159
4	Colorado	58
5	Connecticut	9
6	Delaware	8
7	District Of Columbia	15
8	Florida	201
9	Georgia	208
10	Illinois	135
11	Indiana	50
12	Iowa	1
13	Kansas	1
14	Kentucky	4
15	Louisiana	12
16	Maine	3

```
17              Maryland      63
18         Massachusetts      50
19              Michigan      92
20             Minnesota      29
21           Mississippi      23
22              Missouri       3
23               Montana       1
24                Nevada       1
25         New Hampshire       8
26            New Jersey      56
27            New Mexico      11
28              New York       6
29        North Carolina       3
30                  Ohio       3
31                Oregon      36
32          Pennsylvania     110
33          Rhode Island       1
34        South Carolina      15
35             Tennessee      96
36                 Texas      49
37                  Utah      16
38               Vermont       2
39              Virginia      49
40            Washington      75
41         West Virginia       8
```

[42]:
```
# calculate the ratio of open complaints to closed complaints
for state in open_complaints.index:
    print('{}: {:.4}'.format((state), (open_complaints[state] /␣
 ↪closed_complaints[state]) * 100))
```

```
Alabama: 52.94
Arizona: 42.86
Arkansas: 0.0
California: 38.36
Colorado: 37.93
Connecticut: 33.33
Delaware: 50.0
District Of Columbia: 13.33
Florida: 19.4
Georgia: 38.46
Illinois: 21.48
Indiana: 18.0
Iowa: 0.0
Kansas: 100.0
Kentucky: 75.0
Louisiana: 8.333
Maine: 66.67
```

```
Maryland: 23.81
Massachusetts: 22.0
Michigan: 25.0
Minnesota: 13.79
Mississippi: 69.57
Missouri: 33.33
Montana: 0.0
Nevada: 0.0
New Hampshire: 50.0
New Jersey: 33.93
New Mexico: 36.36
New York: 0.0
North Carolina: 0.0
Ohio: 0.0
Oregon: 36.11
Pennsylvania: 18.18
Rhode Island: 0.0
South Carolina: 20.0
Tennessee: 48.96
Texas: 44.9
Utah: 37.5
Vermont: 50.0
Virginia: 22.45
Washington: 30.67
West Virginia: 37.5
```

As we can see from the output above Kansas has the highest percentage of unresolved complaints **100**. This might be due to the low volume of complaints received and is not accurate representation of their customer service.

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

[43]:
```
df_new['Received Via'].value_counts()
```

[43]:
```
Customer Care Call    1119
Internet              1105
Name: Received Via, dtype: int64
```

[44]:
```
# define variables to hold total counts of complaints received for each␣
↪ingestion medium.
cust_care_call_total = 1119
internet_total = 1105
```

[45]:
```
print('Customer Care Call: \n{}'.format(df_new.loc[df['Received Via'] ==␣
↪'Customer Care Call'].aggregate('Status_trans').value_counts()))
print('\nInternet \n{}'.format(df_new.loc[df['Received Via'] == 'Internet'].
↪aggregate('Status_trans').value_counts()))
```

```
Customer Care Call:
Closed    864
Open      255
Name: Status_trans, dtype: int64

Internet
Closed    843
Open      262
Name: Status_trans, dtype: int64
```

[46]:
```
# create variable to hold all resolved complaints received through Internet
cust_care_closed = 864
Internet_closed = 843
```

[47]:
```
print('Customer Care Call percentage of complaints resolved: {:.2}'.
 ↪format(cust_care_closed / cust_care_call_total))
print('Internet percentage of complaints resolved: {:.2}'.
 ↪format(Internet_closed / internet_total))
```

```
Customer Care Call percentage of complaints resolved: 0.77
Internet percentage of complaints resolved: 0.76
```

[ ]: