

Project on Python

Comcast Telecom Consumer Complaints

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.

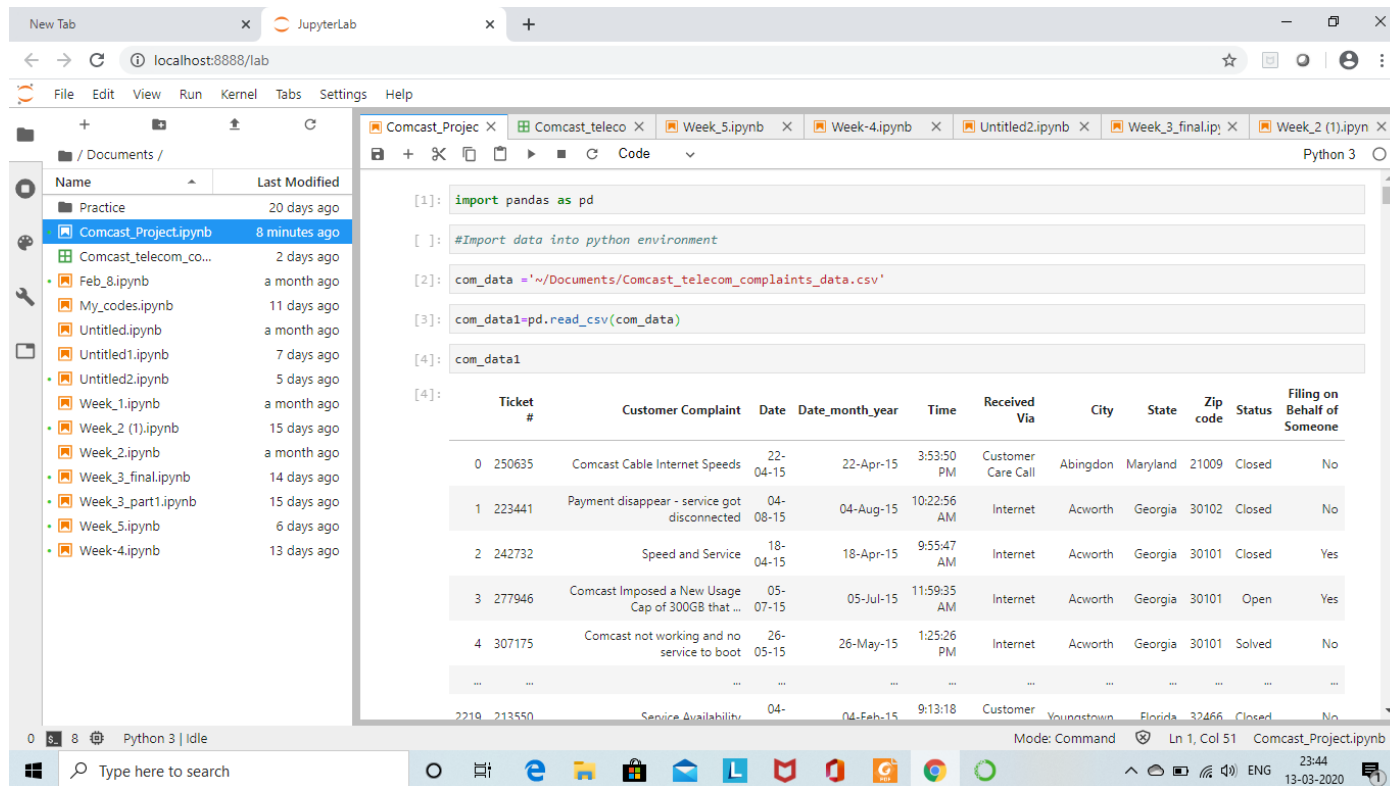
The existing database will serve as a repository of public customer complaints filed against Comcast.

It will help to pin down what is wrong with Comcast's customer service.

In this project we will analyse the dataset provided by the client to derive meaningful output as per the client's demand.

Analysis Task 1

- Import data into Python environment.



The screenshot shows a JupyterLab environment with a file explorer on the left and a code editor on the right. The file explorer lists files in the /Documents/ directory, including Comcast_Project.ipynb, Comcast_telecom_co..., Feb_8.ipynb, My_codes.ipynb, Untitled.ipynb, Untitled1.ipynb, Untitled2.ipynb, Week_1.ipynb, Week_2 (1).ipynb, Week_2.ipynb, Week_3_final.ipynb, Week_3_part1.ipynb, Week_5.ipynb, and Week_4.ipynb. The code editor shows the following code:

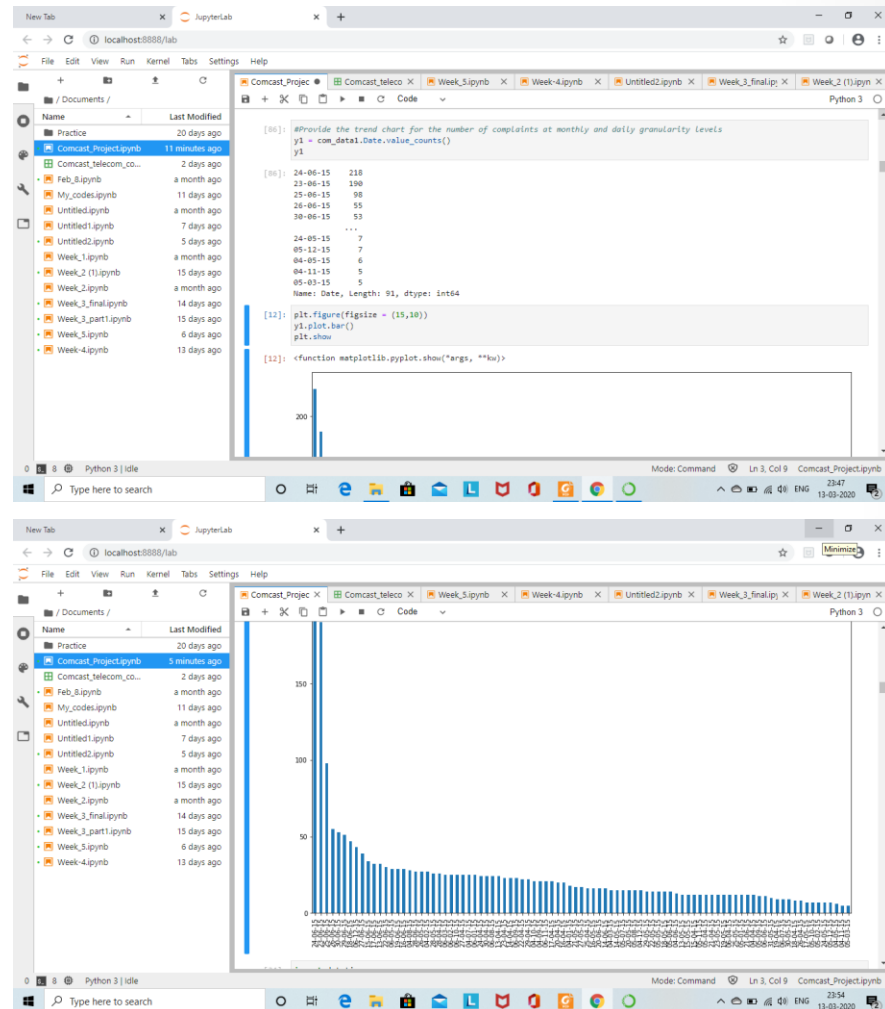
```
[1]: import pandas as pd
[ ]: #Import data into python environment
[2]: com_data = '~/Documents/Comcast_telecom_complaints_data.csv'
[3]: com_data1=pd.read_csv(com_data)
[4]: com_data1
```

The output of the code is a table with 11 columns: Ticket #, Customer Complaint, Date, Date_month_year, Time, Received Via, City, State, Zip code, Status, and Filing on Behalf of Someone. The table contains 5 rows of data, with the last row being a summary row for Service Availability.

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
...
2219_213550	Service Availability	04-04-15	04-Feb-15	9:13:18	Customer	Younstown	Florida	32466	Closed	No

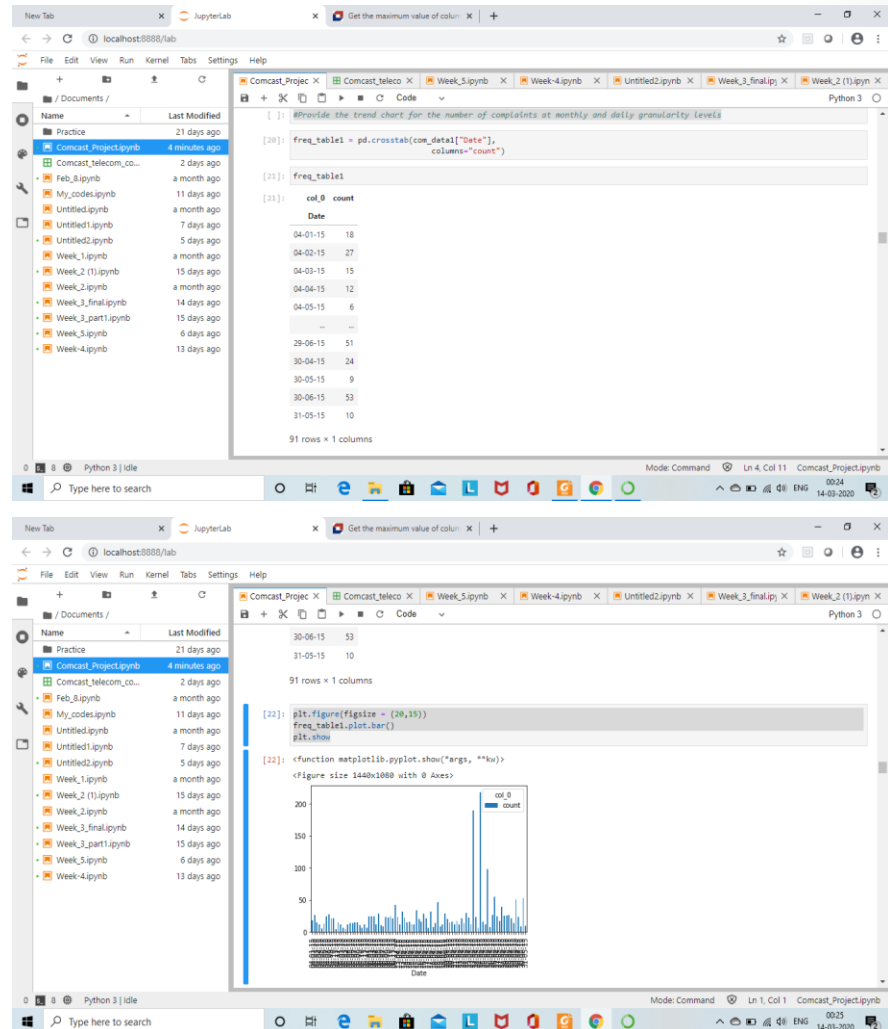
Analysis Task 2

- Provide the trend chart for the number of complaints at daily granularity levels



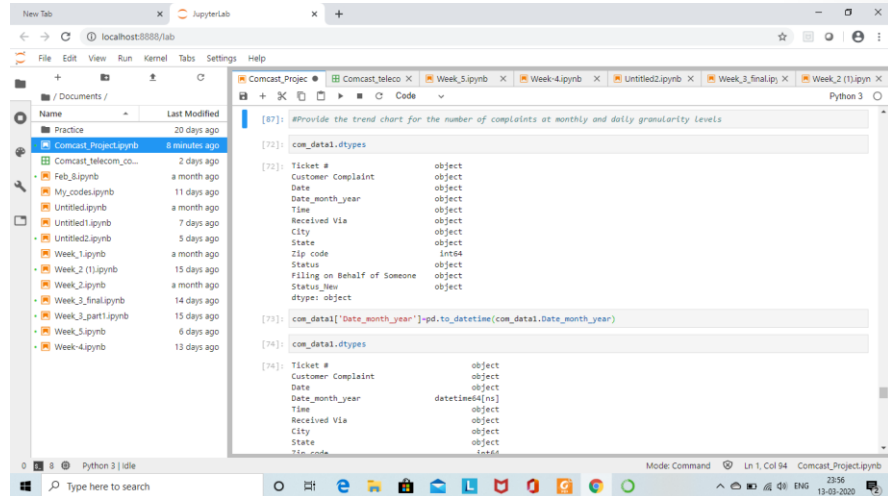
Analysis Task 2

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



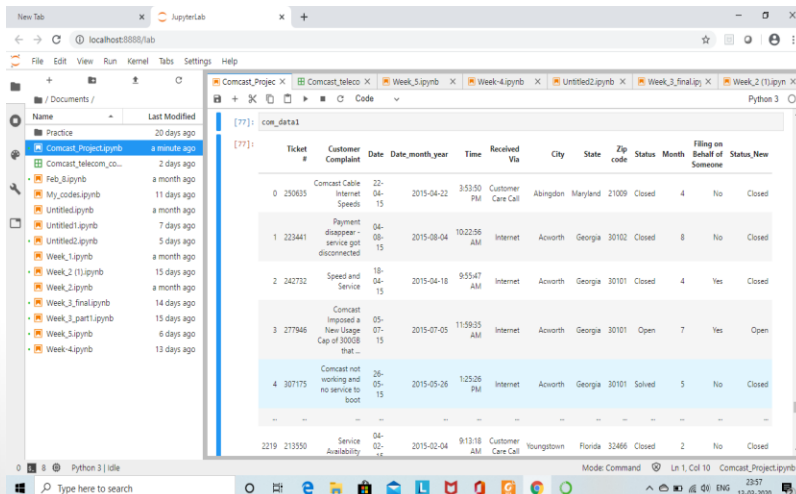
Analysis Task 2

- Provide the trend chart for the number of complaints at monthly granularity levels



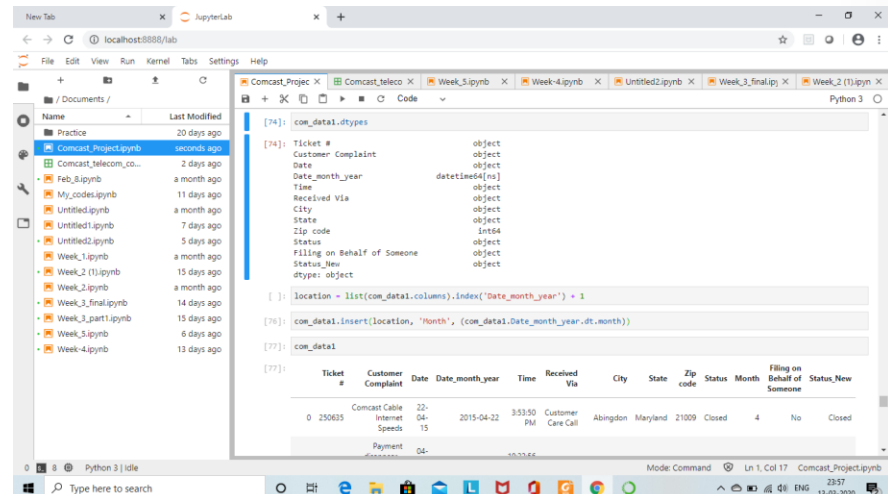
The screenshot shows a JupyterLab interface with a file explorer on the left and a code editor on the right. The code editor contains the following Python code:

```
[67]: #Provide the trend chart for the number of complaints at monthly and daily granularity levels
[72]: com_data.dtypes
[73]: 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
      Status_New        object
      dtype: object
[74]: com_data['Date_month_year'].pd.to_datetime(com_data.Date_month_year)
[75]: com_data.dtypes
[76]: Ticket #      object
      Customer Complaint  object
      Date              object
      Date_month_year    datetime64[ns]
      Time              object
      Received Via      object
      City              object
      State             object
      Zip code          int64
      dtype: object
```



The screenshot shows a JupyterLab interface with a file explorer on the left and a table view of complaint data on the right. The table has the following columns: Ticket #, Customer Complaint, Date, Date_month_year, Time, Received Via, City, State, Zip code, Status, Months, Filing on Behalf of Someone, Status_New, and Status. The data is as follows:

Ticket #	Customer Complaint	Date	Date_month_year	Time	Received Via	City	State	Zip code	Status	Months	Filing on Behalf of Someone	Status_New	Status
0 250635	Comcast Cable Internet Speeds	22-04-15	2015-04-22	3:53:50 PM	Customer Care Call	Abingdon	Maryland	21009	Closed	4	No	Closed	Closed
1 223441	Payment disapp... service got disconnected	04-08-15	2015-08-04	10:22:56 AM	Internet	Aciworth	Georgia	30102	Closed	8	No	Closed	Closed
2 242732	Speed and Service	18-04-15	2015-04-18	9:55:47 AM	Internet	Aciworth	Georgia	30101	Closed	4	Yes	Closed	Closed
3 277948	Comcast imposed a New Usage Cap of 300GB that...	05-07-15	2015-07-05	11:59:35 AM	Internet	Aciworth	Georgia	30101	Open	7	Yes	Open	Open
4 307175	Comcast not working and no service to boot	26-05-15	2015-05-26	1:25:28 PM	Internet	Aciworth	Georgia	30101	Solved	5	No	Closed	Closed
2219 213550	Service Availability	04-02-16	2015-02-04	9:13:18 AM	Customer Care Call	Youngstown	Florida	32466	Closed	2	No	Closed	Closed



The screenshot shows a JupyterLab interface with a file explorer on the left and a code editor on the right. The code editor contains the following Python code:

```
[74]: com_data.dtypes
[75]: Ticket #      object
      Customer Complaint  object
      Date              object
      Date_month_year    datetime64[ns]
      Time              object
      Received Via      object
      City              object
      State             object
      Zip code          int64
      Status            object
      Filing on Behalf of Someone  object
      Status_New        object
      dtype: object
[76]: location = list(com_data.columns).index('Date_month_year') + 1
[77]: com_data.insert(location, 'Month', (com_data.Date_month_year.dt.month))
[78]: com_data
[79]: Ticket #      Customer Complaint  Date  Date_month_year  Time  Received Via  City  State  Zip code  Status  Months  Filing on Behalf of Someone  Status_New  Status
0 250635  Comcast Cable Internet Speeds  22-04-15  2015-04-22  3:53:50 PM  Customer Care Call  Abingdon  Maryland  21009  Closed  4  No  Closed
```

JupyterLab interface showing a file browser on the left and a code editor on the right. The file browser displays a list of files in the /Documents/ directory, including Comcast_Project.ipynb, Comcast_telecom_co..., Feb_8.ipynb, My_codes.ipynb, and several Week_*.ipynb files. The code editor shows the following code:

```
[78]: freq_table3 = pd.crosstab(com_data1["Month"],
                                columns="count")

[79]: freq_table3

[79]:
```

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

```
[80]: plt.figure(figsize = (20,15))
      freq_table3.plot.bar()
```

The bottom status bar indicates the mode is Command, the cursor is at Ln 1, Col 12, and the file is Comcast_Project.ipynb. The system clock shows 23:58 on 13-03-2020.

JupyterLab interface showing the same file browser and code editor. The code editor now displays the following code:

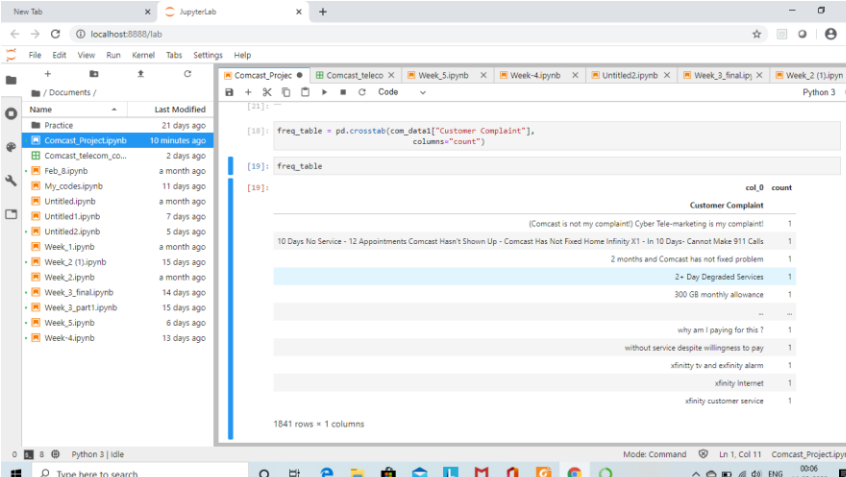
```
[80]: plt.figure(figsize = (20,15))
      freq_table3.plot.bar()
      plt.show

[80]: <function matplotlib.pyplot.show(*args, **kw)>
      <Figure size 1440x1080 with 0 Axes>
```

The bottom status bar indicates the mode is Command, the cursor is at Ln 1, Col 1, and the file is Comcast_Project.ipynb. The system clock shows 23:58 on 13-03-2020.

Analysis Task 3

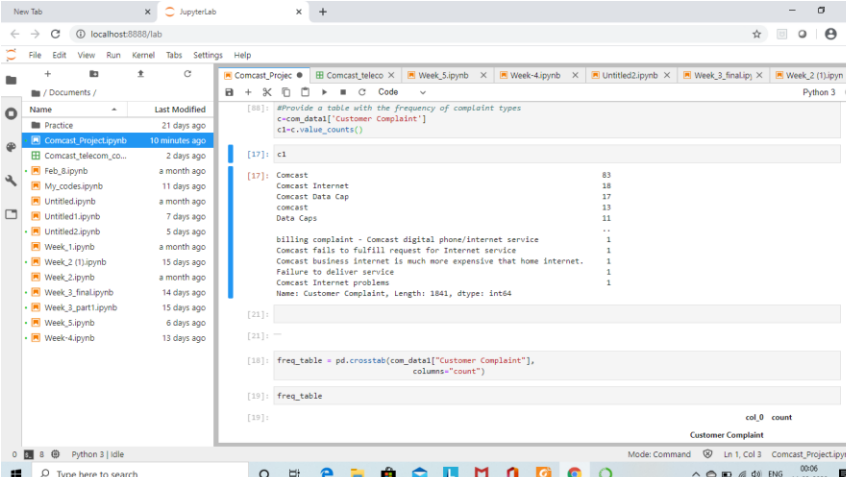
- Provide a table with the frequency of complaint types



```
[18]: freq_table = pd.crosstab(com_data["Customer Complaint"],
                               columns="count")
[19]: freq_table
```

	col_0	count
Customer Complaint		
(Comcast is not my complaint) Cyber Tele-marketing is my complaint!		1
10 Days No Service - 12 Appointments Comcast Hasn't Shown Up - Comcast Has Not Fixed Home Infinity X1 - In 10 Days- Cannot Make 911 Calls		1
2 months and Comcast has not fixed problem		1
2+ Day Degraded Services		1
300 GB monthly allowance		1
...		...
why am I paying for this?		1
without service despite willingness to pay		1
infinity tv and infinity alarm		1
infinity Internet		1
infinity customer service		1

1841 rows x 1 columns



```
[16]: #Provide a table with the frequency of complaint types
c=com_data["Customer Complaint"]
ci=c.value_counts()
[17]: ci
```

	count
Comcast	83
Comcast Internet	18
Comcast Data Cap	17
comcast	13
Data Caps	11
...	...
killing complaint - Comcast digital phone/internet service	1
Comcast fails to fulfill request for Internet service	1
Comcast business Internet is much more expensive than home Internet.	1
Failure to deliver service	1
Comcast Internet problems	1

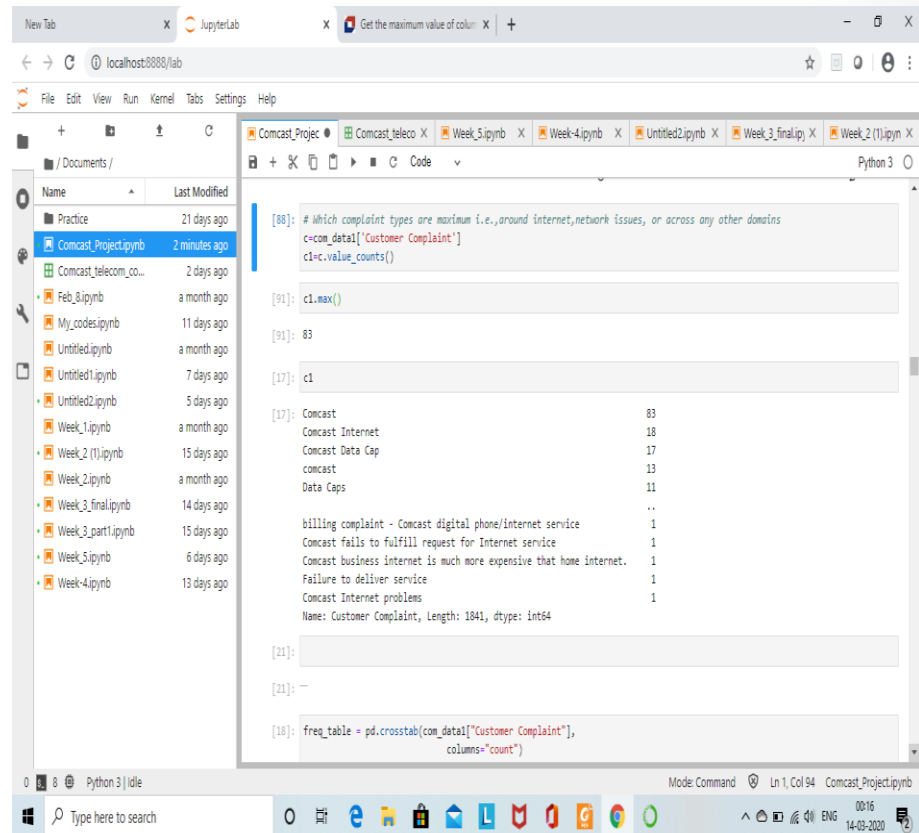
Name: Customer Complaint, Length: 1841, dtype: int64

```
[21]:
[22]:
[23]: freq_table = pd.crosstab(com_data["Customer Complaint"],
                               columns="count")
[24]: freq_table
```

	col_0	count
Customer Complaint		

Analysis Task 4

- Which complaint types are maximum i.e., around internet, network issues, or across any other domains



```
[08]: # Which complaint types are maximum i.e., around internet, network issues, or across any other domains
c=com_data1['Customer Complaint']
c1=c.value_counts()

[91]: c1.max()

[91]: 83

[17]: c1

[17]: Comcast                                     83
Comcast Internet                               18
Comcast Data Cap                               17
comcast                                         13
Data Caps                                      11
..
billing complaint - Comcast digital phone/Internet service 1
Comcast fails to fulfill request for Internet service       1
Comcast business internet is much more expensive than home internet. 1
Failure to deliver service                                  1
Comcast Internet problems                                  1
Name: Customer Complaint, Length: 1841, dtype: int64

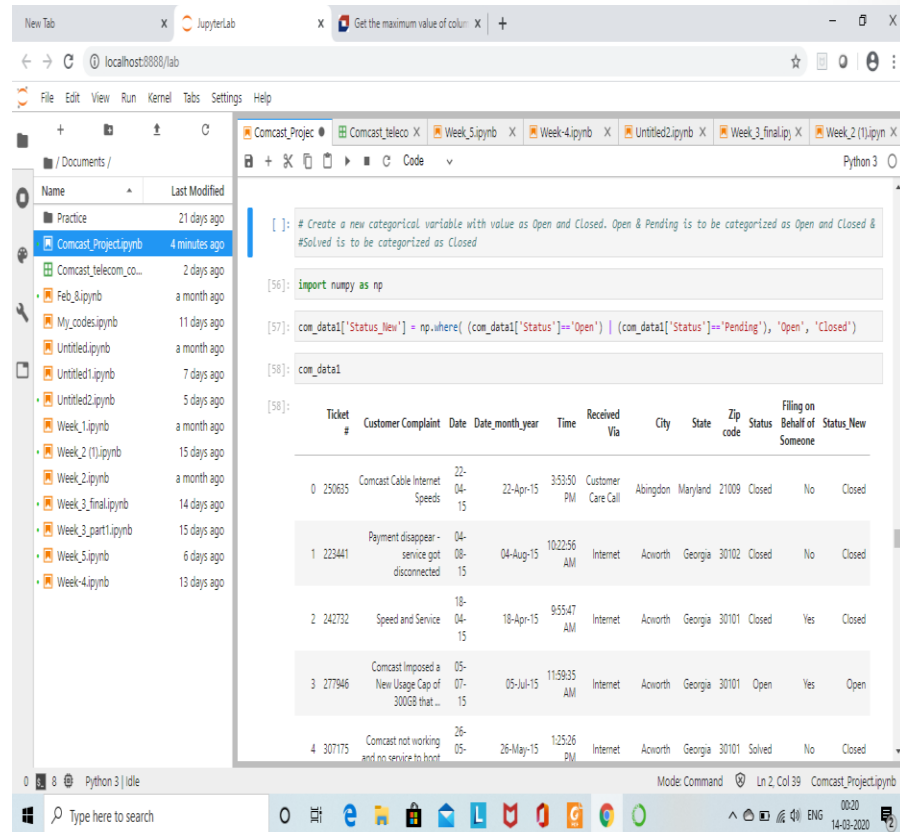
[21]:

[21]: --

[18]: freq_table = pd.crosstab(com_data1['Customer Complaint'],
                               columns='count')
```

Analysis Task 5

- Create a new categorical variable with value as Open and Closed. Open & Pending is to be categorized as Open and Closed & Solved is to be categorized as Closed



The screenshot shows a JupyterLab environment with the following components:

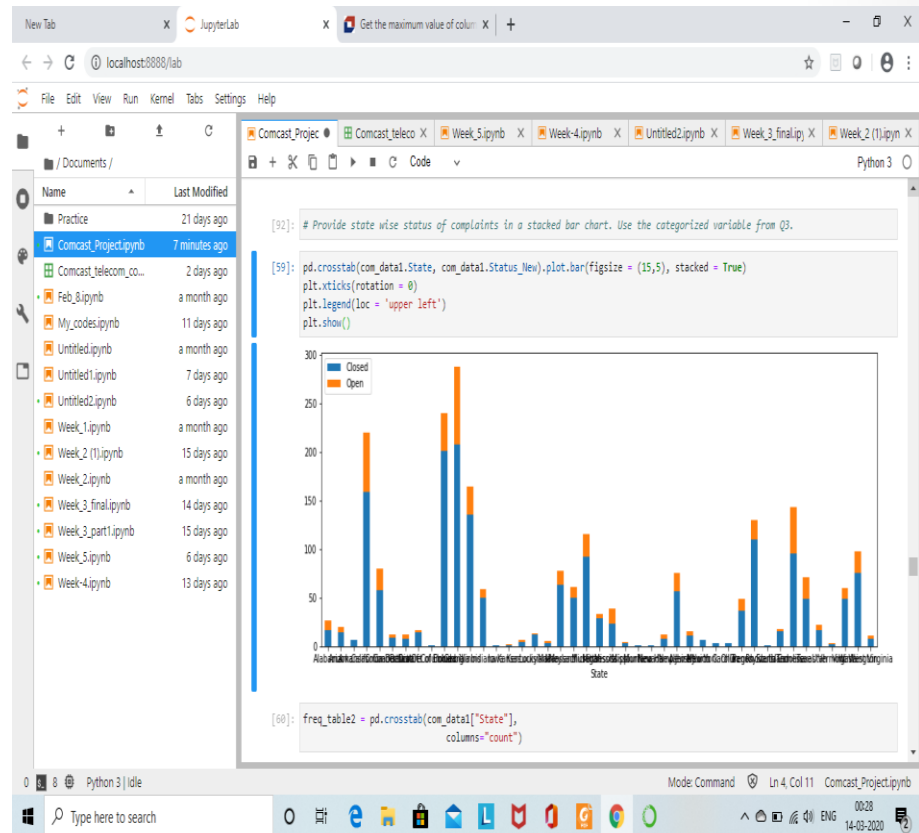
- File Explorer (Left):** Displays a list of files in the 'Documents' directory, including 'Practice', 'Comcast_Project.ipynb', 'Comcast_telecom_co...', 'Feb_8.ipynb', 'My_codes.ipynb', 'Untitled1.ipynb', 'Untitled2.ipynb', 'Week_1.ipynb', 'Week_2 (1).ipynb', 'Week_2.ipynb', 'Week_3_final.ipynb', 'Week_3_part1.ipynb', 'Week_5.ipynb', and 'Week_4.ipynb'.
- Code Editor (Center):** Contains the following Python code:

```
[ ]: # Create a new categorical variable with value as Open and Closed. Open & Pending is to be categorized as Open and Closed & #Solved is to be categorized as Closed
[56]: import numpy as np
[57]: com_data1['Status_New'] = np.where( (com_data1['Status']=='Open') | (com_data1['Status']=='Pending'), 'Open', 'Closed')
[58]: com_data1
```
- Table View (Right):** Displays a table with the following data:

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

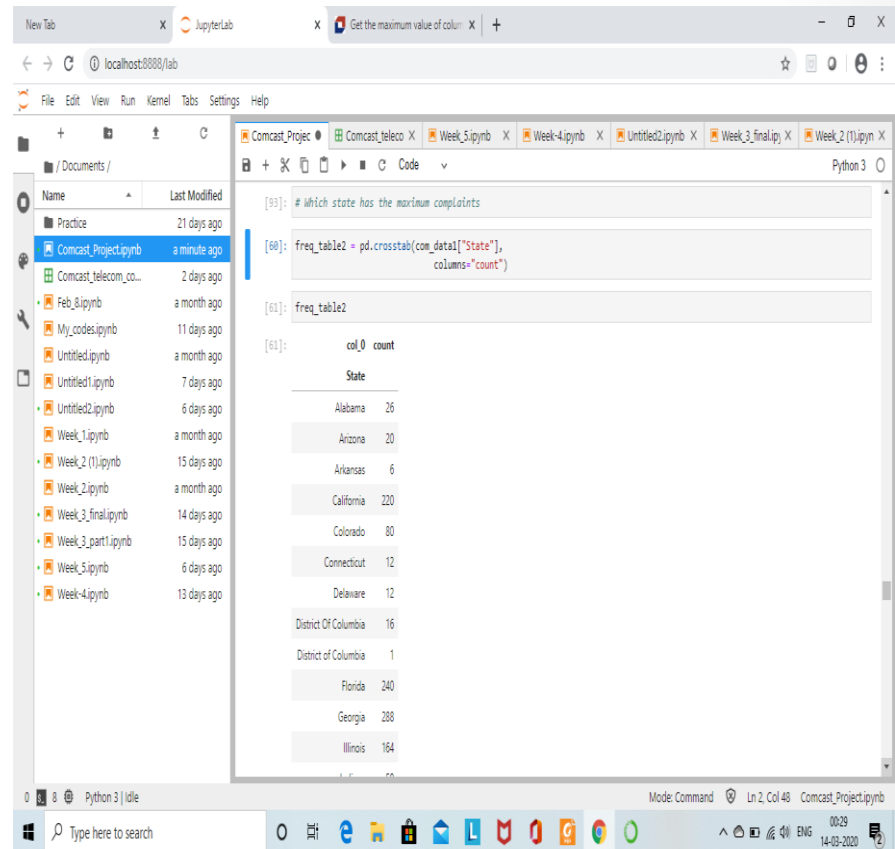
Analysis Task 6

- Provide state wise status of complaints in a stacked bar chart. Use the categorized variable from Q3.



Analysis Task 7

- Which state has the maximum complaints
- As per the frequency table, Georgia has maximum number of complaints that is 288.



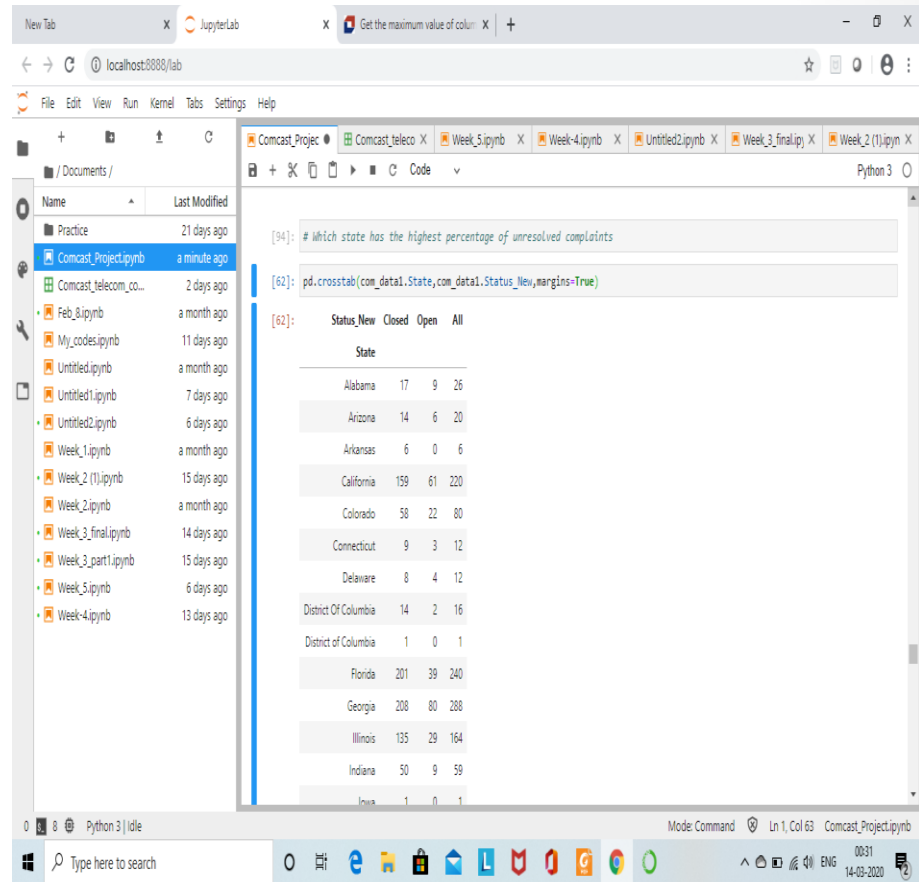
The screenshot shows a JupyterLab environment with a file explorer on the left and a code editor on the right. The file explorer lists several files, including 'Comcast_Project.ipynb'. The code editor displays a Python script that uses pandas to create a frequency table of complaints by state. The output of the script is a table showing the count of complaints for each state.

```
[93]: # Which state has the maximum complaints
[60]: freq_table2 = pd.crosstab(com_data["State"],
[        columns="count")
[61]: freq_table2
[61]:
```

col_0	count
State	
Alabama	26
Arizona	20
Arkansas	6
California	220
Colorado	80
Connecticut	12
Delaware	12
District Of Columbia	16
District of Columbia	1
Florida	240
Georgia	288
Illinois	164

Analysis Task 8

- Which state has the highest percentage of unresolved complaints
- As per frequency table, Georgia has highest percentage of unresolved complaints:
- Total complaints=2224
- Unresolved complaints(Georgia)=80
- Percentage = $(80/2224)*100=3.59\%$



The screenshot shows a JupyterLab environment with a file explorer on the left and a code editor on the right. The code editor contains the following Python code:

```
[94]: # Which state has the highest percentage of unresolved complaints
[62]: pd.crosstab([com_data1.State, com_data1.Status_New], margins=True)
```

The output of the code is a crosstab table showing the count of complaints for each state, categorized by 'Status_New' (Closed, Open, All). The table is as follows:

State	Status_New	Closed	Open	All
Alabama		17	9	26
Arizona		14	6	20
Arkansas		6	0	6
California		159	61	220
Colorado		58	22	80
Connecticut		9	3	12
Delaware		8	4	12
District Of Columbia		14	2	16
District of Columbia		1	0	1
Florida		201	39	240
Georgia		208	80	288
Illinois		135	29	164
Indiana		50	9	59
Iowa		1	0	1

Analysis Task 9

- Provide the percentage of complaints resolved till date, which were received through the Internet and customer care calls.
- On the basis of above data, the percentage of complaints resolved till date is:
- Total Closed Complaints= 1707
- Total Complaints= 2224
- Percentage = $1707/2224 = 76.75\%$

The top screenshot shows the JupyterLab interface with a file explorer on the left and a code editor on the right. The code editor displays a pandas DataFrame with columns: Status, New, Closed, Open, and All. The data is sorted by the 'All' column in descending order.

State	New	Closed	Open	All
Alabama	17	9	26	
Arizona	14	6	20	
Arkansas	6	0	6	
California	159	61	220	
Colorado	58	22	80	
Connecticut	9	3	12	
Delaware	8	4	12	
District Of Columbia	14	2	16	
District of Columbia	1	0	1	
Florida	201	39	240	
Georgia	208	80	288	
Illinois	135	29	164	
Indiana	50	9	59	

The bottom screenshot shows the same JupyterLab interface, but the code editor displays a different DataFrame with columns: State, New, Closed, Open, and All. The data is sorted by the 'All' column in descending order.

State	New	Closed	Open	All
New Jersey	56	19	75	
New Mexico	11	4	15	
New York	6	0	6	
North Carolina	3	0	3	
Ohio	3	0	3	
Oregon	36	13	49	
Pennsylvania	110	20	130	
Rhode Island	1	0	1	
South Carolina	15	3	18	
Tennessee	96	47	143	
Texas	49	22	71	
Utah	16	6	22	
Vermont	2	1	3	
Virginia	49	11	60	
Washington	75	23	98	
West Virginia	8	3	11	
All	1707	517	2224	