CustSerAnalysis-Part3

November 23, 2022

1   3.Find major types of complaints:

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

```
[2]: # read newDF.csv created in Customer service  part 2  ipnyb file
```

```
[ ]: newDF = pd.read_csv('/home/labsuser/Applied DS/newDF.csv')
```

```
[16]: newDF.shape
```

```
[16]: (364558, 49)
```

```
[17]: newDF.head()
```

```
[17]:    Unnamed: 0  Unique Key        Created Date         Closed Date Agency  \
       0           0    32310363  2015-12-31 11:59:45  2016-01-01 12:55:15   NYPD
       1           1    32309934  2015-12-31 11:59:44  2016-01-01 01:26:57   NYPD
       2           2    32309159  2015-12-31 11:59:29  2016-01-01 04:51:03   NYPD
       3           3    32305098  2015-12-31 11:57:46  2016-01-01 07:43:13   NYPD
       4           4    32306529  2015-12-31 11:56:58  2016-01-01 03:24:42   NYPD

                          Agency Name          Complaint Type  \
       0  New York City Police Department  Noise - Street/Sidewalk
       1  New York City Police Department         Blocked Driveway
       2  New York City Police Department         Blocked Driveway
       3  New York City Police Department          Illegal Parking
       4  New York City Police Department          Illegal Parking

                          Descriptor    Location Type  Incident Zip  … \
       0            Loud Music/Party  Street/Sidewalk       10034.0  …
       1                   No Access  Street/Sidewalk       11105.0  …
       2                   No Access  Street/Sidewalk       10458.0  …
       3  Commercial Overnight Parking  Street/Sidewalk     10461.0  …
       4            Blocked Sidewalk  Street/Sidewalk       11373.0  …

         School Not Found Bridge Highway Name Bridge Highway Direction Road Ramp  \
       0                N                  NaN                      NaN       NaN
```

```
1                N              NaN                     NaN        NaN
2                N              NaN                     NaN        NaN
3                N              NaN                     NaN        NaN
4                N              NaN                     NaN        NaN

  Bridge Highway Segment Ferry Direction Ferry Terminal Name   Latitude  \
0                   NaN            NaN                   NaN  40.865682
1                   NaN            NaN                   NaN  40.775945
2                   NaN            NaN                   NaN  40.870325
3                   NaN            NaN                   NaN  40.835994
4                   NaN            NaN                   NaN  40.733060

   Longitude                                 Location
0 -73.923501    (40.86568153633767, -73.92350095571744)
1 -73.915094   (40.775945312321085, -73.91509393898605)
2 -73.888525   (40.870324522111424, -73.88852464418646)
3 -73.828379    (40.83599404683083, -73.82837939584206)
4 -73.874170   (40.733059618956815, -73.87416975810375)

[5 rows x 49 columns]
```

[3]:
```python
# There is Unnamed: 0 columns which can be dropped
newDF.drop(['Unnamed: 0'], axis=1, inplace=True)
```

[19]:
```python
newDF.head()
```

[19]:
```
   Unique Key          Created Date          Closed Date Agency  \
0    32310363  2015-12-31 11:59:45  2016-01-01 12:55:15   NYPD
1    32309934  2015-12-31 11:59:44  2016-01-01 01:26:57   NYPD
2    32309159  2015-12-31 11:59:29  2016-01-01 04:51:03   NYPD
3    32305098  2015-12-31 11:57:46  2016-01-01 07:43:13   NYPD
4    32306529  2015-12-31 11:56:58  2016-01-01 03:24:42   NYPD

                     Agency Name          Complaint Type  \
0  New York City Police Department  Noise - Street/Sidewalk
1  New York City Police Department          Blocked Driveway
2  New York City Police Department          Blocked Driveway
3  New York City Police Department           Illegal Parking
4  New York City Police Department           Illegal Parking

                      Descriptor    Location Type  Incident Zip  \
0             Loud Music/Party  Street/Sidewalk        10034.0
1                    No Access  Street/Sidewalk        11105.0
2                    No Access  Street/Sidewalk        10458.0
3  Commercial Overnight Parking  Street/Sidewalk        10461.0
4             Blocked Sidewalk  Street/Sidewalk        11373.0
```

```
        Incident Address  … School Not Found Bridge Highway Name  \
0    71 VERMILYEA AVENUE  …               N                  NaN
1       27-07 23 AVENUE   …               N                  NaN
2  2897 VALENTINE AVENUE  …               N                  NaN
3   2940 BAISLEY AVENUE   …               N                  NaN
4          87-14 57 ROAD  …               N                  NaN


  Bridge Highway Direction Road Ramp Bridge Highway Segment Ferry Direction  \
0                      NaN       NaN                    NaN             NaN
1                      NaN       NaN                    NaN             NaN
2                      NaN       NaN                    NaN             NaN
3                      NaN       NaN                    NaN             NaN
4                      NaN       NaN                    NaN             NaN


  Ferry Terminal Name   Latitude  Longitude  \
0                 NaN  40.865682 -73.923501
1                 NaN  40.775945 -73.915094
2                 NaN  40.870325 -73.888525
3                 NaN  40.835994 -73.828379
4                 NaN  40.733060 -73.874170


                                Location
0   (40.86568153633767, -73.92350095571744)
1  (40.775945312321085, -73.91509393898605)
2  (40.870324522111424, -73.88852464418646)
3   (40.83599404683083, -73.82837939584206)
4  (40.733059618956815, -73.87416975810375)

[5 rows x 48 columns]
```
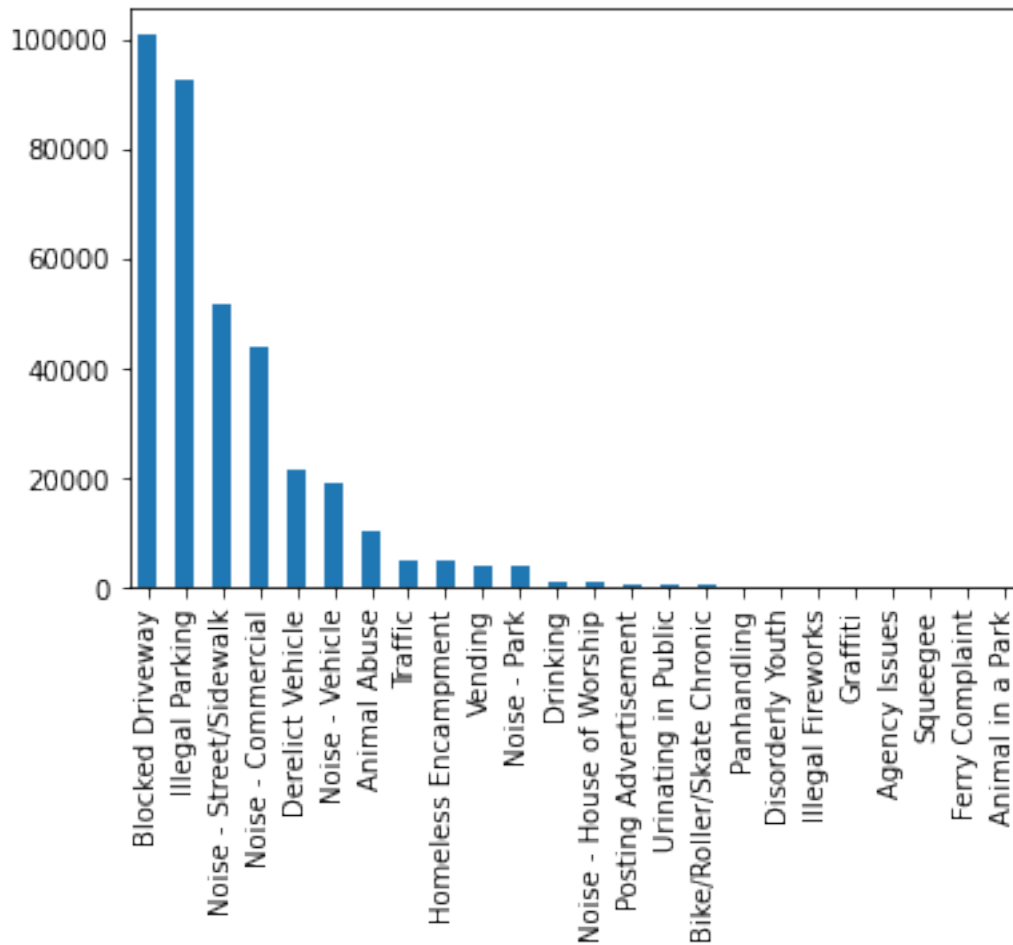
[27]: `# 1. Plot a bar graph of count vs. complaint types`

[4]: `newDF['Complaint Type'].value_counts().plot(kind='bar')`

[4]: `<AxesSubplot:>`

```
[ ]: #Find the top 10 types of complaints
```

```
[22]: top15TypeDF=newDF['Complaint Type'].value_counts()[:15]
```

```
[23]: top15TypeDF
```

```
[23]: Blocked Driveway         100881
      Illegal Parking           92679
      Noise - Street/Sidewalk   51692
      Noise - Commercial        44109
      Derelict Vehicle          21661
      Noise - Vehicle           19352
      Animal Abuse              10541
      Traffic                    5198
      Homeless Encampment        4879
      Vending                    4192
      Noise - Park               4109
```
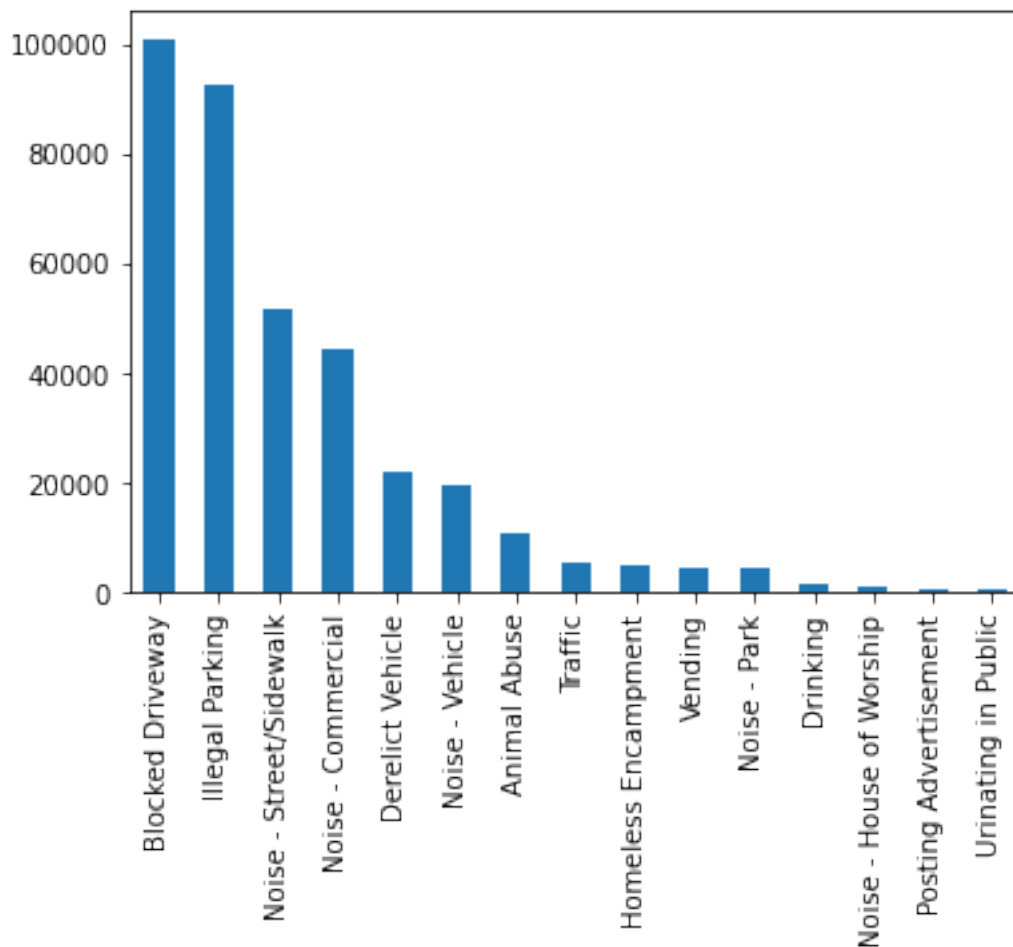
```
Drinking                       1409
Noise - House of Worship       1070
Posting Advertisement           681
Urinating in Public             641
Name: Complaint Type, dtype: int64
```

[35]: `top15TypeDF.plot(kind='bar')`

[35]: `<AxesSubplot:>`

[39]:
```python
# Display the types of complaints in each city in a separate dataset

df_type_city=newDF.groupby(['City','Complaint Type'])['Unique Key'].count().
 ↪unstack()
```

[40]: `df_type_city`

Complaint Type	Animal Abuse	Blocked Driveway	Derelict Vehicle \
City			
ARVERNE	46.0	50.0	32.0
ASTORIA	170.0	3436.0	426.0
Astoria	NaN	159.0	14.0
BAYSIDE	53.0	514.0	231.0
BELLEROSE	15.0	138.0	120.0
BREEZY POINT	2.0	3.0	3.0
BRONX	1971.0	17063.0	2403.0
BROOKLYN	3191.0	36447.0	6259.0
CAMBRIA HEIGHTS	15.0	177.0	148.0
CENTRAL PARK	NaN	NaN	NaN
COLLEGE POINT	35.0	597.0	223.0
CORONA	104.0	3597.0	72.0
EAST ELMHURST	85.0	1925.0	136.0
ELMHURST	59.0	1992.0	94.0
East Elmhurst	NaN	NaN	2.0
FAR ROCKAWAY	111.0	383.0	215.0
FLORAL PARK	7.0	33.0	74.0
FLUSHING	191.0	3641.0	532.0
FOREST HILLS	78.0	873.0	71.0
FRESH MEADOWS	66.0	682.0	347.0
GLEN OAKS	5.0	48.0	57.0
HOLLIS	39.0	442.0	162.0
HOWARD BEACH	51.0	215.0	172.0
Howard Beach	NaN	1.0	NaN
JACKSON HEIGHTS	50.0	703.0	41.0
JAMAICA	317.0	3621.0	1133.0
KEW GARDENS	26.0	429.0	16.0
LITTLE NECK	21.0	174.0	73.0
LONG ISLAND CITY	40.0	1052.0	220.0
Long Island City	NaN	55.0	4.0
MASPETH	56.0	1000.0	510.0
MIDDLE VILLAGE	36.0	663.0	366.0
NEW HYDE PARK	1.0	76.0	14.0
NEW YORK	1941.0	2707.0	695.0
OAKLAND GARDENS	29.0	177.0	117.0
OZONE PARK	72.0	1681.0	479.0
QUEENS	1.0	3.0	2.0
QUEENS VILLAGE	90.0	772.0	478.0
REGO PARK	33.0	780.0	94.0
RICHMOND HILL	55.0	1100.0	201.0
RIDGEWOOD	154.0	2162.0	507.0
ROCKAWAY PARK	33.0	80.0	19.0
ROSEDALE	44.0	270.0	247.0
SAINT ALBANS	43.0	318.0	248.0
SOUTH OZONE PARK	74.0	1202.0	425.0

```
SOUTH RICHMOND HILL           40.0        1946.0        356.0
SPRINGFIELD GARDENS           42.0         330.0        267.0
STATEN ISLAND                786.0        2845.0       2184.0
SUNNYSIDE                     40.0         278.0         17.0
WHITESTONE                    43.0         279.0        279.0
WOODHAVEN                     57.0        1364.0        369.0
WOODSIDE                     111.0        2038.0        298.0
Woodside                      NaN          27.0          8.0

Complaint Type      Disorderly Youth  Drinking  Graffiti  \
City
ARVERNE                          2.0       1.0       1.0
ASTORIA                          5.0      43.0       4.0
Astoria                          NaN       NaN       NaN
BAYSIDE                          2.0       1.0       3.0
BELLEROSE                        2.0       1.0       NaN
BREEZY POINT                     NaN       1.0       NaN
BRONX                           66.0     206.0      15.0
BROOKLYN                        79.0     291.0      60.0
CAMBRIA HEIGHTS                  NaN       NaN       NaN
CENTRAL PARK                     NaN       NaN       NaN
COLLEGE POINT                    1.0       1.0       2.0
CORONA                           6.0      34.0       4.0
EAST ELMHURST                    1.0       9.0       3.0
ELMHURST                         2.0      13.0       1.0
East Elmhurst                    NaN       NaN       NaN
FAR ROCKAWAY                     1.0       4.0       NaN
FLORAL PARK                      1.0       1.0       NaN
FLUSHING                         2.0      47.0       6.0
FOREST HILLS                     1.0       1.0       3.0
FRESH MEADOWS                    NaN       2.0       NaN
GLEN OAKS                        NaN       NaN       NaN
HOLLIS                           1.0       3.0       NaN
HOWARD BEACH                     1.0       4.0       NaN
Howard Beach                     NaN       NaN       NaN
JACKSON HEIGHTS                  NaN      10.0       1.0
JAMAICA                          9.0      40.0       3.0
KEW GARDENS                      NaN       1.0       NaN
LITTLE NECK                      2.0       1.0       NaN
LONG ISLAND CITY                 2.0       8.0       3.0
Long Island City                 NaN       NaN       NaN
MASPETH                          2.0       9.0       1.0
MIDDLE VILLAGE                   NaN       2.0       NaN
NEW HYDE PARK                    NaN       NaN       NaN
NEW YORK                        81.0     321.0      25.0
OAKLAND GARDENS                  1.0       2.0       NaN
OZONE PARK                       4.0      20.0       NaN
```

```
QUEENS                          NaN       NaN       NaN
QUEENS VILLAGE                  NaN       5.0       1.0
REGO PARK                       NaN       4.0       1.0
RICHMOND HILL                   NaN      10.0       1.0
RIDGEWOOD                       3.0      10.0       3.0
ROCKAWAY PARK                   4.0      23.0       NaN
ROSEDALE                        NaN       2.0       2.0
SAINT ALBANS                    1.0       3.0       NaN
SOUTH OZONE PARK                2.0      14.0       2.0
SOUTH RICHMOND HILL             2.0      25.0       NaN
SPRINGFIELD GARDENS             NaN       6.0       NaN
STATEN ISLAND                  25.0     188.0       6.0
SUNNYSIDE                       2.0      12.0       1.0
WHITESTONE                      1.0       3.0       1.0
WOODHAVEN                       NaN       4.0       NaN
WOODSIDE                        1.0      15.0       4.0
Woodside                        NaN       NaN       NaN
```

Complaint Type	Homeless Encampment	Illegal Parking	Noise - Commercial \
City			
ARVERNE	4.0	62.0	2.0
ASTORIA	32.0	1340.0	1653.0
Astoria	NaN	277.0	311.0
BAYSIDE	2.0	638.0	47.0
BELLEROSE	1.0	132.0	38.0
BREEZY POINT	NaN	16.0	4.0
BRONX	275.0	9889.0	2945.0
BROOKLYN	948.0	33533.0	13860.0
CAMBRIA HEIGHTS	6.0	113.0	19.0
CENTRAL PARK	NaN	5.0	NaN
COLLEGE POINT	3.0	449.0	38.0
CORONA	26.0	791.0	281.0
EAST ELMHURST	2.0	1093.0	41.0
ELMHURST	34.0	760.0	85.0
East Elmhurst	NaN	28.0	NaN
FAR ROCKAWAY	16.0	339.0	59.0
FLORAL PARK	NaN	72.0	3.0
FLUSHING	26.0	2250.0	222.0
FOREST HILLS	18.0	627.0	163.0
FRESH MEADOWS	6.0	1158.0	21.0
GLEN OAKS	NaN	95.0	84.0
HOLLIS	9.0	181.0	54.0
HOWARD BEACH	3.0	384.0	258.0
Howard Beach	NaN	NaN	NaN
JACKSON HEIGHTS	11.0	241.0	619.0
JAMAICA	93.0	1698.0	552.0
KEW GARDENS	5.0	276.0	203.0

Complaint Type			
City			
LITTLE NECK	NaN	322.0	77.0
LONG ISLAND CITY	10.0	988.0	269.0
Long Island City	NaN	64.0	19.0
MASPETH	11.0	1234.0	58.0
MIDDLE VILLAGE	5.0	1104.0	13.0
NEW HYDE PARK	NaN	32.0	4.0
NEW YORK	3060.0	14553.0	18692.0
OAKLAND GARDENS	1.0	337.0	2.0
OZONE PARK	8.0	774.0	125.0
QUEENS	2.0	10.0	6.0
QUEENS VILLAGE	19.0	669.0	49.0
REGO PARK	6.0	640.0	82.0
RICHMOND HILL	30.0	489.0	249.0
RIDGEWOOD	26.0	2235.0	491.0
ROCKAWAY PARK	4.0	337.0	72.0
ROSEDALE	4.0	326.0	28.0
SAINT ALBANS	11.0	237.0	36.0
SOUTH OZONE PARK	5.0	602.0	82.0
SOUTH RICHMOND HILL	12.0	596.0	223.0
SPRINGFIELD GARDENS	7.0	291.0	38.0
STATEN ISLAND	77.0	6224.0	784.0
SUNNYSIDE	12.0	167.0	238.0
WHITESTONE	NaN	631.0	21.0
WOODHAVEN	10.0	896.0	209.0
WOODSIDE	38.0	1083.0	256.0
Woodside	NaN	124.0	2.0

Complaint Type	Noise - House of Worship	…	Noise - Vehicle \
City		…	
ARVERNE	14.0	…	10.0
ASTORIA	21.0	…	236.0
Astoria	NaN	…	NaN
BAYSIDE	3.0	…	24.0
BELLEROSE	1.0	…	11.0
BREEZY POINT	NaN	…	1.0
BRONX	90.0	…	3556.0
BROOKLYN	389.0	…	5966.0
CAMBRIA HEIGHTS	2.0	…	100.0
CENTRAL PARK	NaN	…	NaN
COLLEGE POINT	2.0	…	140.0
CORONA	3.0	…	110.0
EAST ELMHURST	25.0	…	82.0
ELMHURST	6.0	…	69.0
East Elmhurst	NaN	…	NaN
FAR ROCKAWAY	1.0	…	83.0
FLORAL PARK	NaN	…	2.0
FLUSHING	5.0	…	147.0

City			
FOREST HILLS	1.0	…	70.0
FRESH MEADOWS	NaN	…	97.0
GLEN OAKS	NaN	…	4.0
HOLLIS	215.0	…	52.0
HOWARD BEACH	1.0	…	10.0
Howard Beach	NaN	…	NaN
JACKSON HEIGHTS	2.0	…	75.0
JAMAICA	15.0	…	337.0
KEW GARDENS	1.0	…	23.0
LITTLE NECK	NaN	…	8.0
LONG ISLAND CITY	NaN	…	124.0
Long Island City	NaN	…	NaN
MASPETH	2.0	…	26.0
MIDDLE VILLAGE	NaN	…	45.0
NEW HYDE PARK	NaN	…	2.0
NEW YORK	222.0	…	6295.0
OAKLAND GARDENS	NaN	…	7.0
OZONE PARK	4.0	…	81.0
QUEENS	1.0	…	2.0
QUEENS VILLAGE	2.0	…	54.0
REGO PARK	1.0	…	60.0
RICHMOND HILL	NaN	…	69.0
RIDGEWOOD	2.0	…	249.0
ROCKAWAY PARK	NaN	…	29.0
ROSEDALE	2.0	…	25.0
SAINT ALBANS	1.0	…	50.0
SOUTH OZONE PARK	5.0	…	97.0
SOUTH RICHMOND HILL	3.0	…	93.0
SPRINGFIELD GARDENS	1.0	…	48.0
STATEN ISLAND	18.0	…	424.0
SUNNYSIDE	NaN	…	53.0
WHITESTONE	NaN	…	31.0
WOODHAVEN	3.0	…	81.0
WOODSIDE	4.0	…	136.0
Woodside	NaN	…	NaN

Complaint Type	Panhandling	Traffic	Urinating in Public	Vending \
City				
ARVERNE	1.0	1.0	1.0	1.0
ASTORIA	2.0	60.0	10.0	57.0
Astoria	NaN	NaN	NaN	NaN
BAYSIDE	NaN	9.0	NaN	2.0
BELLEROSE	1.0	9.0	1.0	NaN
BREEZY POINT	NaN	NaN	NaN	NaN
BRONX	20.0	427.0	54.0	433.0
BROOKLYN	49.0	1258.0	155.0	575.0
CAMBRIA HEIGHTS	NaN	7.0	NaN	NaN

City				
CENTRAL PARK	NaN	NaN	NaN	NaN
COLLEGE POINT	NaN	16.0	NaN	1.0
CORONA	1.0	14.0	7.0	65.0
EAST ELMHURST	NaN	24.0	6.0	9.0
ELMHURST	3.0	18.0	10.0	25.0
East Elmhurst	NaN	NaN	NaN	NaN
FAR ROCKAWAY	NaN	11.0	1.0	10.0
FLORAL PARK	NaN	NaN	NaN	NaN
FLUSHING	2.0	59.0	12.0	37.0
FOREST HILLS	6.0	65.0	2.0	10.0
FRESH MEADOWS	1.0	15.0	1.0	1.0
GLEN OAKS	NaN	3.0	2.0	19.0
HOLLIS	NaN	11.0	2.0	NaN
HOWARD BEACH	2.0	9.0	NaN	5.0
Howard Beach	NaN	NaN	NaN	NaN
JACKSON HEIGHTS	1.0	13.0	3.0	86.0
JAMAICA	3.0	632.0	37.0	24.0
KEW GARDENS	NaN	10.0	3.0	1.0
LITTLE NECK	NaN	20.0	1.0	NaN
LONG ISLAND CITY	2.0	83.0	3.0	31.0
Long Island City	NaN	NaN	NaN	NaN
MASPETH	NaN	71.0	2.0	7.0
MIDDLE VILLAGE	NaN	14.0	NaN	NaN
NEW HYDE PARK	NaN	NaN	NaN	NaN
NEW YORK	206.0	1770.0	264.0	2639.0
OAKLAND GARDENS	NaN	6.0	NaN	2.0
OZONE PARK	7.0	21.0	4.0	1.0
QUEENS	NaN	2.0	1.0	NaN
QUEENS VILLAGE	1.0	27.0	5.0	2.0
REGO PARK	NaN	16.0	1.0	3.0
RICHMOND HILL	NaN	8.0	5.0	15.0
RIDGEWOOD	NaN	50.0	9.0	9.0
ROCKAWAY PARK	NaN	7.0	1.0	2.0
ROSEDALE	NaN	25.0	NaN	19.0
SAINT ALBANS	NaN	14.0	1.0	2.0
SOUTH OZONE PARK	NaN	36.0	2.0	5.0
SOUTH RICHMOND HILL	NaN	12.0	1.0	24.0
SPRINGFIELD GARDENS	2.0	12.0	3.0	1.0
STATEN ISLAND	13.0	229.0	19.0	25.0
SUNNYSIDE	NaN	17.0	2.0	15.0
WHITESTONE	NaN	32.0	NaN	1.0
WOODHAVEN	1.0	7.0	2.0	6.0
WOODSIDE	NaN	45.0	8.0	15.0
Woodside	NaN	NaN	NaN	NaN

Complaint Type   Bike/Roller/Skate Chronic Illegal Fireworks \
City

11

ARVERNE	NaN	NaN
ASTORIA	16.0	4.0
Astoria	NaN	NaN
BAYSIDE	NaN	NaN
BELLEROSE	1.0	1.0
BREEZY POINT	NaN	NaN
BRONX	22.0	24.0
BROOKLYN	124.0	61.0
CAMBRIA HEIGHTS	NaN	1.0
CENTRAL PARK	NaN	NaN
COLLEGE POINT	NaN	NaN
CORONA	NaN	NaN
EAST ELMHURST	1.0	NaN
ELMHURST	2.0	1.0
East Elmhurst	NaN	NaN
FAR ROCKAWAY	NaN	NaN
FLORAL PARK	NaN	NaN
FLUSHING	3.0	2.0
FOREST HILLS	6.0	1.0
FRESH MEADOWS	NaN	NaN
GLEN OAKS	NaN	NaN
HOLLIS	NaN	NaN
HOWARD BEACH	1.0	4.0
Howard Beach	NaN	NaN
JACKSON HEIGHTS	2.0	1.0
JAMAICA	3.0	4.0
KEW GARDENS	NaN	NaN
LITTLE NECK	NaN	NaN
LONG ISLAND CITY	3.0	NaN
Long Island City	NaN	NaN
MASPETH	1.0	1.0
MIDDLE VILLAGE	1.0	NaN
NEW HYDE PARK	NaN	NaN
NEW YORK	254.0	38.0
OAKLAND GARDENS	2.0	NaN
OZONE PARK	1.0	1.0
QUEENS	NaN	NaN
QUEENS VILLAGE	NaN	5.0
REGO PARK	NaN	NaN
RICHMOND HILL	NaN	4.0
RIDGEWOOD	3.0	2.0
ROCKAWAY PARK	NaN	NaN
ROSEDALE	2.0	NaN
SAINT ALBANS	NaN	NaN
SOUTH OZONE PARK	1.0	1.0
SOUTH RICHMOND HILL	1.0	2.0
SPRINGFIELD GARDENS	NaN	1.0

STATEN ISLAND		10.0	11.0
SUNNYSIDE		2.0	NaN
WHITESTONE		4.0	1.0
WOODHAVEN		2.0	NaN
WOODSIDE		5.0	1.0
Woodside		NaN	NaN

Complaint Type	Posting Advertisement	Squeegee	Animal in a Park
City			
ARVERNE	NaN	NaN	NaN
ASTORIA	3.0	NaN	NaN
Astoria	NaN	NaN	NaN
BAYSIDE	NaN	NaN	NaN
BELLEROSE	1.0	NaN	NaN
BREEZY POINT	NaN	NaN	NaN
BRONX	18.0	NaN	NaN
BROOKLYN	58.0	NaN	NaN
CAMBRIA HEIGHTS	NaN	NaN	NaN
CENTRAL PARK	NaN	NaN	NaN
COLLEGE POINT	NaN	NaN	NaN
CORONA	1.0	NaN	NaN
EAST ELMHURST	1.0	NaN	NaN
ELMHURST	1.0	NaN	NaN
East Elmhurst	NaN	NaN	NaN
FAR ROCKAWAY	NaN	NaN	NaN
FLORAL PARK	NaN	NaN	NaN
FLUSHING	1.0	NaN	NaN
FOREST HILLS	4.0	NaN	NaN
FRESH MEADOWS	NaN	NaN	NaN
GLEN OAKS	NaN	NaN	NaN
HOLLIS	NaN	NaN	NaN
HOWARD BEACH	NaN	NaN	NaN
Howard Beach	NaN	NaN	NaN
JACKSON HEIGHTS	1.0	NaN	NaN
JAMAICA	8.0	NaN	NaN
KEW GARDENS	1.0	NaN	NaN
LITTLE NECK	1.0	NaN	NaN
LONG ISLAND CITY	2.0	NaN	NaN
Long Island City	NaN	NaN	NaN
MASPETH	NaN	NaN	NaN
MIDDLE VILLAGE	NaN	NaN	NaN
NEW HYDE PARK	NaN	NaN	NaN
NEW YORK	49.0	4.0	NaN
OAKLAND GARDENS	NaN	NaN	NaN
OZONE PARK	3.0	NaN	NaN
QUEENS	NaN	NaN	1.0
QUEENS VILLAGE	1.0	NaN	NaN

```
REGO PARK                                NaN       NaN         NaN
RICHMOND HILL                            2.0       NaN         NaN
RIDGEWOOD                                1.0       NaN         NaN
ROCKAWAY PARK                            NaN       NaN         NaN
ROSEDALE                                 NaN       NaN         NaN
SAINT ALBANS                             NaN       NaN         NaN
SOUTH OZONE PARK                         1.0       NaN         NaN
SOUTH RICHMOND HILL                      NaN       NaN         NaN
SPRINGFIELD GARDENS                      2.0       NaN         NaN
STATEN ISLAND                          517.0       NaN         NaN
SUNNYSIDE                                3.0       NaN         NaN
WHITESTONE                               NaN       NaN         NaN
WOODHAVEN                                NaN       NaN         NaN
WOODSIDE                                 NaN       NaN         NaN
Woodside                                 NaN       NaN         NaN

[53 rows x 22 columns]
```

[25]:
```python
#Visualize the major types of complaints in each city
#get top 10 complaints in top 10 cities
top10_types = newDF['Complaint Type'].value_counts()[:10]
top10_city = newDF['City'].value_counts()[:10]
```

[26]:
```python
#data whose complaintType is in the top 10 complaint types
top10_typesDF = newDF[newDF['Complaint Type'].isin(top10_types.index)]
```

[27]:
```python
#data whose city and Types are in teh top 10 cities with top 10 complaints
top10cityTypeDF = top10_typesDF[top10_typesDF['City'].isin(top10_city.index)]
```
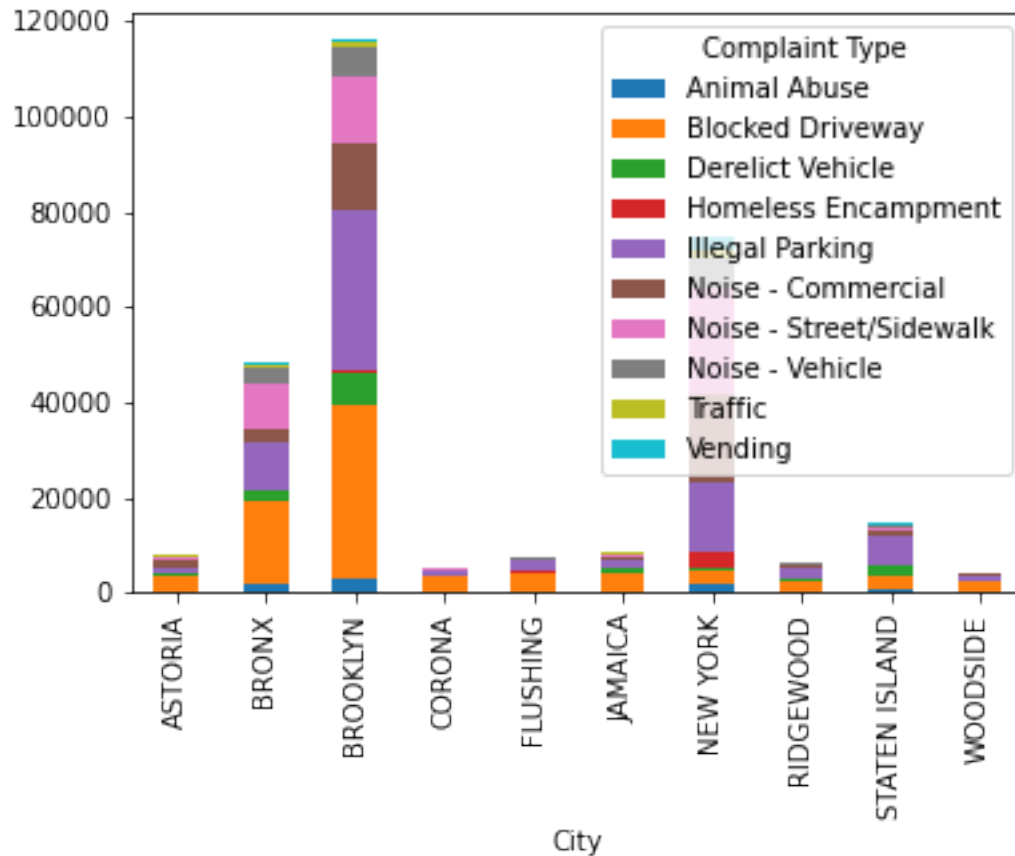
[28]:
```python
top10cityTypeDF.shape
```

[28]: (293051, 50)

[49]:
```python
#for categorical data lets do a cross tab
cityComplaints=pd.
 ↪crosstab(index=top10cityTypeDF['City'],columns=top10cityTypeDF['Complaint␣
 ↪Type'])
```

[51]:
```python
import matplotlib.pyplot as plt
plt.figure(figsize=(12,10))
cityComplaints.plot(kind='bar',stacked=True)
```

[51]: <AxesSubplot:xlabel='City'>

```
<Figure size 864x720 with 0 Axes>
```

2 Check if the average response time across various types of complaints

```
[8]: #measure  response time
     #first convert the Date columns to datetime object
     # lets convert them to dateTime objects
     newDF['Created Date'] = pd.to_datetime(newDF['Created Date']);
     newDF['Closed Date'] = pd.to_datetime(newDF['Closed Date']);
     newDF['Resolution Action Updated Date'] = pd.to_datetime(newDF['Resolution
     ↪Action Updated Date']);
     newDF['Due Date'] = pd.to_datetime(newDF['Due Date']);
```

```
[ ]: newDF.dtypes
```

```
[9]: import numpy as np
```

```
[10]: newDF['Response Time'] = newDF['Closed Date']-newDF['Created Date']
```

```
[12]: newDF['RespTime_Hrs']=newDF['Response Time']/np.timedelta64(1,'h')
```

```
[ ]: newDF.head()
```

3 Identify significant variables by performing a statistical analysis using p-values and chi-square values (Optional)

```
[13]: newDF['Complaint Type'].value_counts().head(5)
```

```
[13]: Blocked Driveway          100881
      Illegal Parking            92679
      Noise - Street/Sidewalk    51692
      Noise - Commercial         44109
      Derelict Vehicle           21661
      Name: Complaint Type, dtype: int64
```

```
[14]: # Check if the average response time across various types of complaints is same␣
      ↪or not
      newDF[['Complaint Type','RespTime_Hrs']]
      # h0: avg response time for all the complaints in population for type a=
      #     avg response time for type b =avg response time for type c=..
      # avg response time for type e.

      # ha: at least one statement is not true
      # In our scenario lets take top 5 complaint Types as samples and analysie␣
      ↪average response times
```

```
[14]:                  Complaint Type  RespTime_Hrs
      0       Noise - Street/Sidewalk     24.925000
      1              Blocked Driveway     13.453611
      2              Blocked Driveway     16.859444
      3               Illegal Parking     19.757500
      4               Illegal Parking     15.462222
      ...                         ...           ...
      364553          Illegal Parking     -1.703611
      364554           Noise - Vehicle     -9.657222
      364555  Noise - Street/Sidewalk      0.317500
      364556          Blocked Driveway     -9.318611
      364557          Blocked Driveway     -9.216667

      [364558 rows x 2 columns]
```

```
[15]: df1=newDF[newDF['Complaint Type']=='Blocked Driveway']['RespTime_Hrs'].dropna()␣
      ↪# lest remove NAN values from samples
```

```
[16]: df2=newDF[newDF['Complaint Type']=='Illegal Parking']['RespTime_Hrs'].dropna()␣
      ↪# lest remove NAN values from samples
```

```
[17]: df3=newDF[newDF['Complaint Type']=='Noise - Street/Sidewalk']['RespTime_Hrs'].
      ↪dropna() # lest remove NAN values from samples
```

```
[18]: df4=newDF[newDF['Complaint Type']=='Noise - Commercial']['RespTime_Hrs'].
      ↪dropna() # lest remove NAN values from samples
```

```
[19]: df5=newDF[newDF['Complaint Type']=='Derelict Vehicle']['RespTime_Hrs'].dropna()␣
      ↪# lest remove NAN values from samples
```

```
[84]: #NUll Hypothesis claim = average response time is same for all the complaints
```

```
[20]: from scipy import stats
```

```
[21]: # lets do #anova test test to perform null hypothesis
      stats.f_oneway(df1,df2,df3,df4,df5)
```

```
[21]: F_onewayResult(statistic=304.95779134697534, pvalue=2.649697471012518e-262)
```

```
[83]: #here p value is 2.64*** ^-262 ~ 0
      # alpha(statistsic)  = 304.9577
      #Since pvalue is less than alpha(risk) factor we will reject(do not accept null␣
      ↪hypothesis)
```

3.1 chi sq test

```
[ ]: # # any relationship between city and complaint type
```

```
[29]: top10cityTypeDF
```

```
[29]:        Unique Key        Created Date         Closed Date Agency  \
      0        32310363 2015-12-31 11:59:45 2016-01-01 12:55:15   NYPD
      1        32309934 2015-12-31 11:59:44 2016-01-01 01:26:57   NYPD
      2        32309159 2015-12-31 11:59:29 2016-01-01 04:51:03   NYPD
      3        32305098 2015-12-31 11:57:46 2016-01-01 07:43:13   NYPD
      5        32306554 2015-12-31 11:56:30 2016-01-01 01:50:11   NYPD
      ...           ...                 ...                 ...    ...
      364551   29607567 2015-01-01 12:06:02 2015-01-01 12:43:41   NYPD
      364552   29610051 2015-01-01 12:05:05 2015-01-01 01:22:10   NYPD
      364554   29608392 2015-01-01 12:04:28 2015-01-01 02:25:02   NYPD
      364555   29607589 2015-01-01 12:01:30 2015-01-01 12:20:33   NYPD
      364556   29610889 2015-01-01 12:01:29 2015-01-01 02:42:22   NYPD
```

```
                            Agency Name          Complaint Type  \
0        New York City Police Department   Noise - Street/Sidewalk
1        New York City Police Department          Blocked Driveway
2        New York City Police Department          Blocked Driveway
3        New York City Police Department           Illegal Parking
5        New York City Police Department           Illegal Parking
...                                  ...                       ...
364551   New York City Police Department   Noise - Street/Sidewalk
364552   New York City Police Department   Noise - Street/Sidewalk
364554   New York City Police Department           Noise - Vehicle
364555   New York City Police Department   Noise - Street/Sidewalk
364556   New York City Police Department          Blocked Driveway

                            Descriptor    Location Type   Incident Zip  \
0                     Loud Music/Party   Street/Sidewalk        10034.0
1                            No Access   Street/Sidewalk        11105.0
2                            No Access   Street/Sidewalk        10458.0
3           Commercial Overnight Parking Street/Sidewalk        10461.0
5           Posted Parking Sign Violation Street/Sidewalk       11215.0
...                                  ...             ...            ...
364551               Loud Music/Party   Street/Sidewalk        10453.0
364552               Loud Music/Party   Street/Sidewalk        10002.0
364554                 Car/Truck Horn   Street/Sidewalk        10468.0
364555               Loud Music/Party   Street/Sidewalk        10031.0
364556                      No Access   Street/Sidewalk        10466.0

            Incident Address  ... Bridge Highway Direction Road Ramp  \
0           71 VERMILYEA AVENUE  ...                      NaN       NaN
1               27-07 23 AVENUE  ...                      NaN       NaN
2         2897 VALENTINE AVENUE  ...                      NaN       NaN
3           2940 BAISLEY AVENUE  ...                      NaN       NaN
5                 260 21 STREET  ...                      NaN       NaN
...                        ... ...                      ...       ...
364551                     NaN  ...                      NaN       NaN
364552                     NaN  ...                      NaN       NaN
364554   2555 SEDGWICK AVENUE  ...                      NaN       NaN
364555     508 WEST 139 STREET  ...                      NaN       NaN
364556     931 EAST 226 STREET  ...                      NaN       NaN

        Bridge Highway Segment Ferry Direction Ferry Terminal Name   Latitude  \
0                          NaN             NaN                 NaN  40.865682
1                          NaN             NaN                 NaN  40.775945
2                          NaN             NaN                 NaN  40.870325
3                          NaN             NaN                 NaN  40.835994
5                          NaN             NaN                 NaN  40.660823
...                        ...             ...                 ...        ...
364551                     NaN             NaN                 NaN  40.848639
```

```
364552                NaN         NaN         NaN   40.721235
364554                NaN         NaN         NaN   40.867830
364555                NaN         NaN         NaN   40.821647
364556                NaN         NaN         NaN   40.886361

        Longitude                                Location     Response Time  \
0       -73.923501    (40.86568153633767, -73.92350095571744)    1 days 00:55:30
1       -73.915094   (40.775945312321085, -73.91509393898605)    0 days 13:27:13
2       -73.888525   (40.870324522111424, -73.88852464418646)    0 days 16:51:34
3       -73.828379    (40.83599404683083, -73.82837939584206)    0 days 19:45:27
5       -73.992568   (40.66082272389114, -73.99256786342693)    0 days 13:53:41
...            ...                                       ...                ...
364551  -73.911679   (40.84863947227845, -73.91167941229558)    0 days 00:37:39
364552  -73.987770   (40.72123468734571, -73.98777023226815)   -1 days +13:17:05
364554  -73.907178   (40.86782963689454, -73.90717786644662)   -1 days +14:20:34
364555  -73.950873  (40.821646626438095, -73.95087342885292)    0 days 00:19:03
364556  -73.853290   (40.88636077906953, -73.85329048666742)   -1 days +14:40:53

        RespTime_Hrs
0          24.925000
1          13.453611
2          16.859444
3          19.757500
5          13.894722
...              ...
364551      0.627500
364552    -10.715278
364554     -9.657222
364555      0.317500
364556     -9.318611

[293051 rows x 50 columns]
```

[32]:
```python
obj=pd.
 ↪crosstab(index=top10cityTypeDF['City'],columns=top10cityTypeDF['Complaint␣
 ↪Type'])
```

[33]:
```python
obj
```

[33]:
```
Complaint Type  Animal Abuse  Blocked Driveway  Derelict Vehicle  \
City
ASTORIA                  170              3436               426
BRONX                   1971             17063              2403
BROOKLYN                3191             36447              6259
CORONA                   104              3597                72
FLUSHING                 191              3641               532
JAMAICA                  317              3621              1133
```

```
NEW YORK                 1941            2707            695
RIDGEWOOD                 154            2162            507
STATEN ISLAND             786            2845           2184
WOODSIDE                  111            2038            298


Complaint Type  Homeless Encampment  Illegal Parking  Noise - Commercial  \
City
ASTORIA                          32             1340                1653
BRONX                           275             9889                2945
BROOKLYN                        948            33533               13860
CORONA                           26              791                 281
FLUSHING                         26             2250                 222
JAMAICA                          93             1698                 552
NEW YORK                       3060            14553               18692
RIDGEWOOD                        26             2235                 491
STATEN ISLAND                    77             6224                 784
WOODSIDE                         38             1083                 256


Complaint Type  Noise - Street/Sidewalk  Noise - Vehicle  Traffic  Vending
City
ASTORIA                            409              236       60       57
BRONX                             9146             3556      427      433
BROOKLYN                         13984             5966     1258      575
CORONA                             243              110       14       65
FLUSHING                           241              147       59       37
JAMAICA                            365              337      632       24
NEW YORK                         22252             6295     1770     2639
RIDGEWOOD                          448              249       50        9
STATEN ISLAND                      888              424      229       25
WOODSIDE                           261              136       45       15
```

[35]:
```
# h0: the city and the complaint types are independent of each other # Null␣
 ↪Hypothesis

# ha:the city and the complaint types are not independent of each other #␣
 ↪alternate Hypothesis
```

[36]:
```
chi_sq,pvalue,degrees,exp_value= stats.chi2_contingency(obj)
```

[47]:
```
doHypothesisTest(pvalue)
```

```
Dependent (reject H0)
```

[45]:
```python
def doHypothesisTest(pval):
    alpha = 0.05 # 50% risk factor
    if(pval <= alpha):
        print('Dependent (reject H0)')
```

```
    else:
        print('Independent (H0 holds true)')
```

[ ]: