Customer Service request Analysis

November 30, 2022

[1]:
```
import numpy as np
import pandas as pd
import matplotlib.pyplot as plt
%matplotlib inline
```

[2]:
```
data = pd.read_csv('311_Service_Requests_from_2010_to_Present.csv')
```

[3]:
```
data.head()
```

[3]:

	Unique Key	Created Date	Closed Date	Agency	Agency Name	Complaint Type	Descriptor	Location Type	Incident Zip	Incident Address	…	Bridge Highway Name	Bridge Highway Direction
0	32310363	12/31/2015 11:59:45 PM	01-01-16 0:55	NYPD	New York City Police Department	Noise - Street/Sidewalk	Loud Music/Party	Street/Sidewalk	10034.0	71 VERMILYEA AVENUE	…	NaN	NaN
1	32309934	12/31/2015 11:59:44 PM	01-01-16 1:26	NYPD	New York City Police Department	Blocked Driveway	No Access	Street/Sidewalk	11105.0	27-07 23 AVENUE	…	NaN	NaN
2	32309159	12/31/2015 11:59:29 PM	01-01-16 4:51	NYPD	New York City Police Department	Blocked Driveway	No Access	Street/Sidewalk	10458.0	2897 VALENTINE AVENUE	…	NaN	NaN
3	32305098	12/31/2015 11:57:46 PM	01-01-16 7:43	NYPD	New York City Police Department	Illegal Parking	Commercial Overnight Parking	Street/Sidewalk	10461.0	2940 BAISLEY AVENUE	…	NaN	NaN
4	32306529	12/31/2015 11:56:58 PM	01-01-16 3:24	NYPD	New York City Police Department	Illegal Parking	Blocked Sidewalk	Street/Sidewalk	11373.0	87-14 57 ROAD	…	NaN	NaN

```
  Road Ramp Bridge Highway Segment Garage Lot Name Ferry Direction  \
0                       NaN                      NaN              NaN
1                       NaN                      NaN              NaN
2                       NaN                      NaN              NaN
3                       NaN                      NaN              NaN
4                       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 53 columns]
```

1 SHAPE OF DATA SET

[5]:
```
data.info()
```

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 77284 entries, 0 to 77283
Data columns (total 53 columns):
 #   Column                          Non-Null Count  Dtype
---  ------                          --------------  -----
 0   Unique Key                      77284 non-null  int64
 1   Created Date                    77284 non-null  object
 2   Closed Date                     76834 non-null  object
 3   Agency                          77284 non-null  object
 4   Agency Name                     77284 non-null  object
 5   Complaint Type                  77284 non-null  object
 6   Descriptor                      76089 non-null  object
 7   Location Type                   77284 non-null  object
 8   Incident Zip                    76780 non-null  float64
 9   Incident Address                67719 non-null  object
 10  Street Name                     67719 non-null  object
 11  Cross Street 1                  66654 non-null  object
 12  Cross Street 2                  66569 non-null  object
```

```
 13  Intersection Street 1           9457 non-null   object
 14  Intersection Street 2           9365 non-null   object
 15  Address Type                    76707 non-null  object
 16  City                            76779 non-null  object
 17  Landmark                        47 non-null     object
 18  Facility Type                   76839 non-null  object
 19  Status                          77284 non-null  object
 20  Due Date                        77284 non-null  object
 21  Resolution Description          77283 non-null  object
 22  Resolution Action Updated Date  76846 non-null  object
 23  Community Board                 77283 non-null  object
 24  Borough                         77283 non-null  object
 25  X Coordinate (State Plane)      76603 non-null  float64
 26  Y Coordinate (State Plane)      76603 non-null  float64
 27  Park Facility Name              77283 non-null  object
 28  Park Borough                    77283 non-null  object
 29  School Name                     77283 non-null  object
 30  School Number                   77283 non-null  object
 31  School Region                   77283 non-null  object
 32  School Code                     77283 non-null  object
 33  School Phone Number             77283 non-null  object
 34  School Address                  77283 non-null  object
 35  School City                     77283 non-null  object
 36  School State                    77283 non-null  object
 37  School Zip                      77283 non-null  object
 38  School Not Found                77283 non-null  object
 39  School or Citywide Complaint    0 non-null      float64
 40  Vehicle Type                    0 non-null      float64
 41  Taxi Company Borough            0 non-null      float64
 42  Taxi Pick Up Location           0 non-null      float64
 43  Bridge Highway Name             77 non-null     object
 44  Bridge Highway Direction        77 non-null     object
 45  Road Ramp                       66 non-null     object
 46  Bridge Highway Segment          66 non-null     object
 47  Garage Lot Name                 0 non-null      float64
 48  Ferry Direction                 0 non-null      float64
 49  Ferry Terminal Name             0 non-null      float64
 50  Latitude                        76603 non-null  float64
 51  Longitude                       76603 non-null  float64
 52  Location                        76603 non-null  object
dtypes: float64(12), int64(1), object(40)
memory usage: 31.3+ MB
```

2 Finding Null values

[6]:
```
data.isna().any()
```

[6]:
```
Unique Key                      False
Created Date                    False
Closed Date                      True
Agency                          False
Agency Name                     False
Complaint Type                  False
Descriptor                       True
Location Type                   False
Incident Zip                     True
Incident Address                 True
Street Name                      True
Cross Street 1                   True
Cross Street 2                   True
Intersection Street 1            True
Intersection Street 2            True
Address Type                     True
City                             True
Landmark                         True
Facility Type                    True
Status                          False
Due Date                        False
Resolution Description           True
Resolution Action Updated Date   True
Community Board                  True
Borough                          True
X Coordinate (State Plane)       True
Y Coordinate (State Plane)       True
Park Facility Name               True
Park Borough                     True
School Name                      True
School Number                    True
School Region                    True
School Code                      True
School Phone Number              True
School Address                   True
School City                      True
School State                     True
School Zip                       True
School Not Found                 True
School or Citywide Complaint     True
Vehicle Type                     True
Taxi Company Borough             True
Taxi Pick Up Location            True
```

```
Bridge Highway Name                True
Bridge Highway Direction           True
Road Ramp                          True
Bridge Highway Segment             True
Garage Lot Name                    True
Ferry Direction                    True
Ferry Terminal Name                True
Latitude                           True
Longitude                          True
Location                           True
dtype: bool
```

3 Check how many null values are there

[7]:
```
data.isnull().sum()
```

[7]:
```
Unique Key                         0
Created Date                       0
Closed Date                      450
Agency                             0
Agency Name                        0
Complaint Type                     0
Descriptor                      1195
Location Type                      0
Incident Zip                     504
Incident Address                9565
Street Name                     9565
Cross Street 1                 10630
Cross Street 2                 10715
Intersection Street 1          67827
Intersection Street 2          67919
Address Type                     577
City                             505
Landmark                       77237
Facility Type                    445
Status                             0
Due Date                           0
Resolution Description             1
Resolution Action Updated Date   438
Community Board                    1
Borough                            1
X Coordinate (State Plane)       681
Y Coordinate (State Plane)       681
Park Facility Name                 1
Park Borough                       1
```

```
School Name                          1
School Number                        1
School Region                        1
School Code                          1
School Phone Number                  1
School Address                       1
School City                          1
School State                         1
School Zip                           1
School Not Found                     1
School or Citywide Complaint     77284
Vehicle Type                     77284
Taxi Company Borough             77284
Taxi Pick Up Location            77284
Bridge Highway Name              77207
Bridge Highway Direction         77207
Road Ramp                        77218
Bridge Highway Segment           77218
Garage Lot Name                  77284
Ferry Direction                  77284
Ferry Terminal Name              77284
Latitude                           681
Longitude                          681
Location                           681
dtype: int64
```

4 Drop the empty columns

[8]:
```
data = data.drop(['School or Citywide Complaint', 'Vehicle Type', 'Taxi Company␣
↪Borough', 'Taxi Pick Up Location', 'Garage Lot Name'], axis=1)
data.head()
```

[8]:
```
   Unique Key            Created Date    Closed Date Agency  \
0    32310363  12/31/2015 11:59:45 PM  01-01-16 0:55   NYPD
1    32309934  12/31/2015 11:59:44 PM  01-01-16 1:26   NYPD
2    32309159  12/31/2015 11:59:29 PM  01-01-16 4:51   NYPD
3    32305098  12/31/2015 11:57:46 PM  01-01-16 7:43   NYPD
4    32306529  12/31/2015 11:56:58 PM  01-01-16 3:24   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]
```

5 Show Statistics for numerical Columns

[9]:
```
data.describe()
```

[9]:
```
         Unique Key  Incident Zip  X Coordinate (State Plane)  \
count  7.728400e+04  76780.000000                7.660300e+04
mean   3.204443e+07  10871.925046                1.004348e+06
```

```
std    1.546407e+05    576.734611                2.228461e+04
min    3.176300e+07     83.000000                9.133570e+05
25%    3.190798e+07  10314.000000                9.907355e+05
50%    3.204367e+07  11210.000000                1.002480e+06
75%    3.218296e+07  11249.000000                1.019316e+06
max    3.231065e+07  11697.000000                1.067154e+06

       Y Coordinate (State Plane)  Ferry Direction  Ferry Terminal Name  \
count                76603.000000              0.0                  0.0
mean                201801.470465              NaN                  NaN
std                  29281.850135              NaN                  NaN
min                 121411.000000              NaN                  NaN
25%                 181093.000000              NaN                  NaN
50%                 200020.000000              NaN                  NaN
75%                 219050.000000              NaN                  NaN
max                 271391.000000              NaN                  NaN

           Latitude     Longitude
count  76603.000000  76603.000000
mean      40.720524    -73.927455
std        0.080368      0.080371
min       40.499673    -74.254937
25%       40.663709    -73.976611
50%       40.715622    -73.934250
75%       40.767869    -73.873489
max       40.911533    -73.700837
```

6 Display City Names

[10]:
```
data['City'].unique()
```

[10]:
```
array(['NEW YORK', 'ASTORIA', 'BRONX', 'ELMHURST', 'BROOKLYN',
       'KEW GARDENS', 'JACKSON HEIGHTS', 'MIDDLE VILLAGE', 'REGO PARK',
       'SAINT ALBANS', 'JAMAICA', 'SOUTH RICHMOND HILL', nan, 'RIDGEWOOD',
       'HOWARD BEACH', 'FOREST HILLS', 'STATEN ISLAND', 'OZONE PARK',
       'RICHMOND HILL', 'WOODHAVEN', 'FLUSHING', 'CORONA',
       'QUEENS VILLAGE', 'OAKLAND GARDENS', 'HOLLIS', 'MASPETH',
       'EAST ELMHURST', 'SOUTH OZONE PARK', 'WOODSIDE', 'FRESH MEADOWS',
       'LONG ISLAND CITY', 'ROCKAWAY PARK', 'SPRINGFIELD GARDENS',
       'COLLEGE POINT', 'BAYSIDE', 'GLEN OAKS', 'FAR ROCKAWAY',
       'BELLEROSE', 'LITTLE NECK', 'CAMBRIA HEIGHTS', 'ROSEDALE',
       'SUNNYSIDE', 'WHITESTONE', 'ARVERNE', 'FLORAL PARK',
       'NEW HYDE PARK', 'CENTRAL PARK', 'BREEZY POINT', 'QUEENS',
       'Astoria', 'Long Island City', 'Woodside'], dtype=object)
```

7 Change all city names to title Format

[11]:
```
def to_title(city):
    try:
        city = city.title()
        return city
    except:
        return np.nan

data['City'] = data['City'].apply(to_title)
data['City'].value_counts()
```

[11]:
```
Brooklyn               26169
New York               16073
Bronx                   9311
Staten Island           3489
Astoria                 1891
Flushing                1829
Jamaica                 1813
Ridgewood               1511
Corona                  1190
Woodside                 932
East Elmhurst            801
Elmhurst                 761
South Richmond Hill      759
Fresh Meadows            732
Maspeth                  694
Ozone Park               683
Woodhaven                681
Long Island City         656
South Ozone Park         510
Queens Village           496
Middle Village           494
Forest Hills             493
Richmond Hill            490
Rego Park                480
Jackson Heights          421
College Point            360
Bayside                  325
Whitestone               256
Far Rockaway             238
Saint Albans             231
Howard Beach             219
Hollis                   212
Kew Gardens              206
Springfield Gardens      205
Sunnyside                186
```

```
Rosedale                174
Little Neck             149
Oakland Gardens         145
Rockaway Park           112
Cambria Heights         108
Glen Oaks                88
Bellerose                68
Arverne                  48
Floral Park              34
New Hyde Park            25
Central Park             24
Queens                    5
Breezy Point              2
Name: City, dtype: int64
```

8 Display Complaint types

[12]:
```
data['Complaint Type'].value_counts()
```

[12]:
```
Blocked Driveway            23906
Illegal Parking             21668
Noise - Commercial          10307
Noise - Street/Sidewalk      7403
Derelict Vehicle             4586
Noise - Vehicle              3341
Animal Abuse                 1855
Traffic                      1117
Homeless Encampment           939
Vending                       748
Noise - Park                  370
Posting Advertisement         348
Drinking                      235
Noise - House of Worship      156
Urinating in Public            97
Bike/Roller/Skate Chronic      80
Panhandling                    72
Disorderly Youth               26
Graffiti                       23
Illegal Fireworks               7
Name: Complaint Type, dtype: int64
```

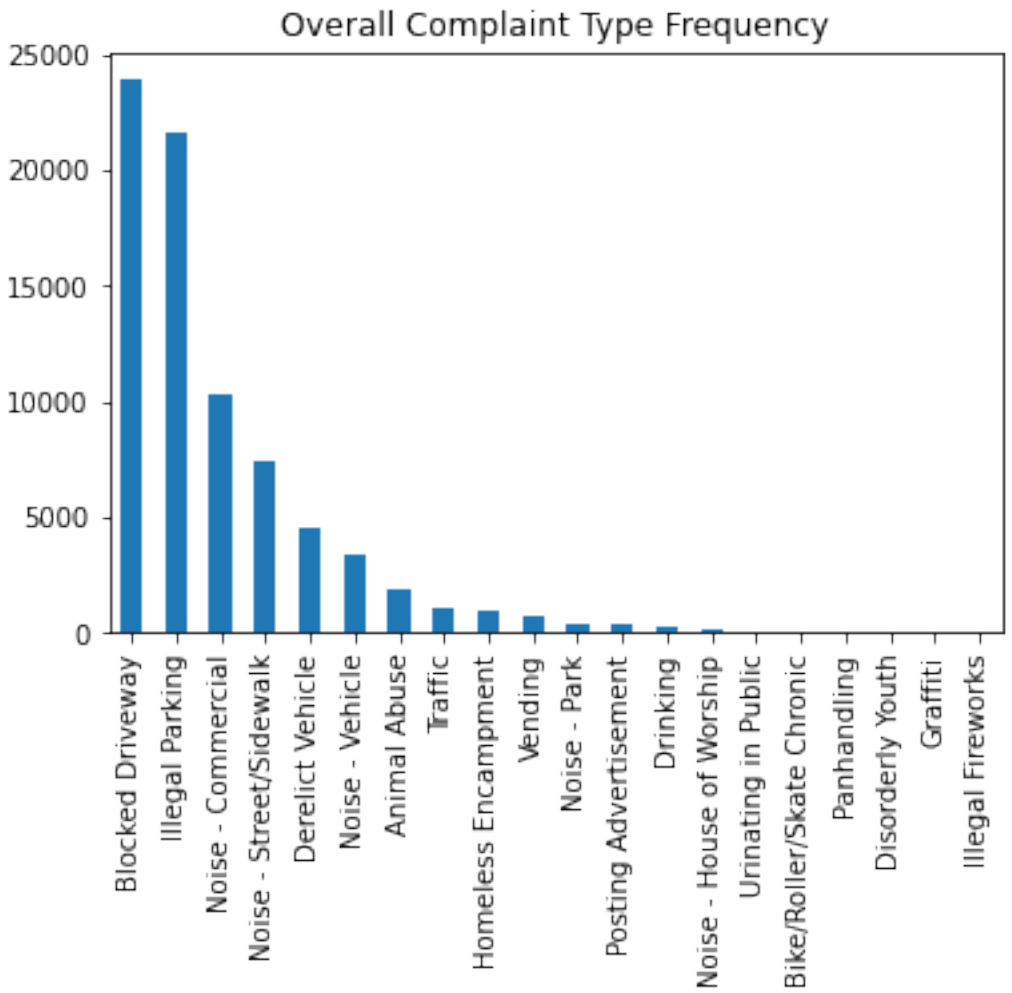
9 Count Complain types by city

[13]:
```
data.groupby(['City','Complaint Type']).size()
```

[13]:
```
City      Complaint Type
Arverne   Animal Abuse                6
          Blocked Driveway            9
          Derelict Vehicle            8
          Homeless Encampment         1
          Illegal Parking            17
                                     ..
Woodside  Noise - Street/Sidewalk    22
          Noise - Vehicle            27
          Traffic                    15
          Urinating in Public         1
          Vending                     6
Length: 533, dtype: int64
```

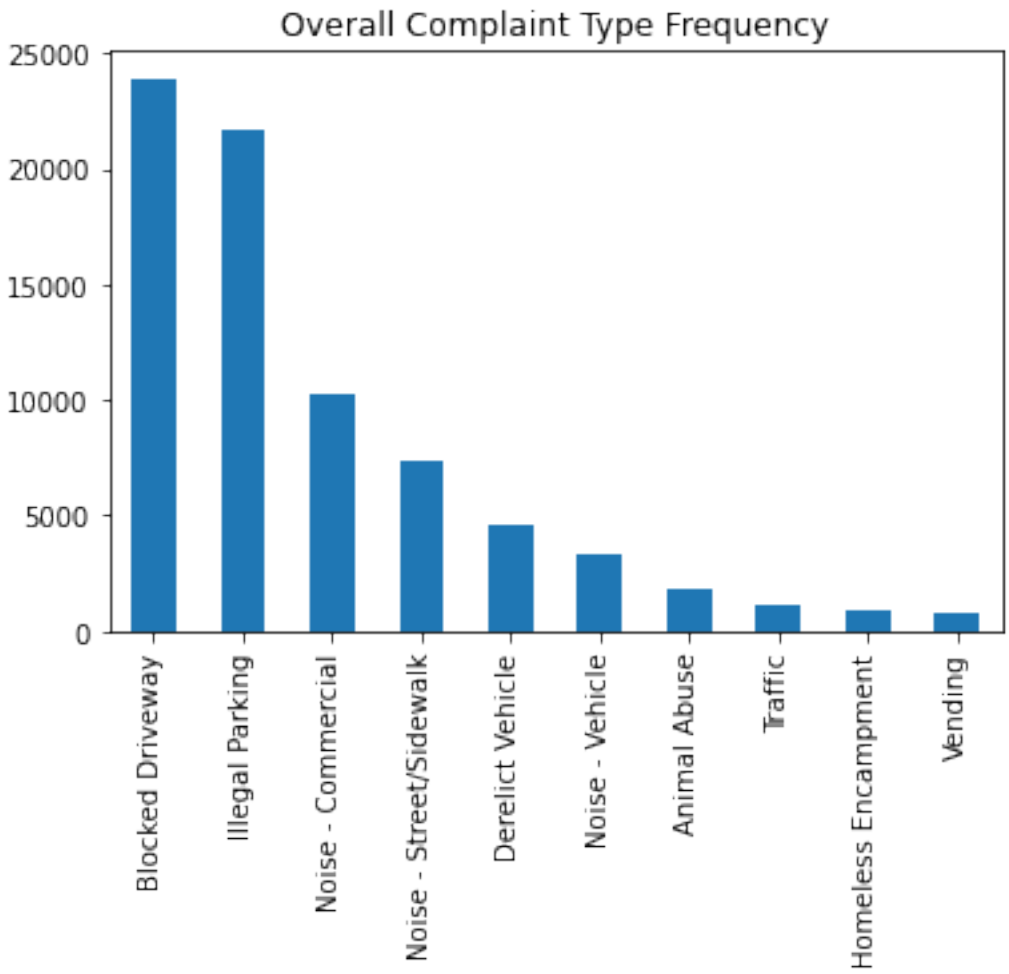
10 Display frequency of Overall Complaint types

[14]:
```
data['Complaint Type'].value_counts().plot(kind = 'bar', title = 'Overall␣
↪Complaint Type Frequency');
```



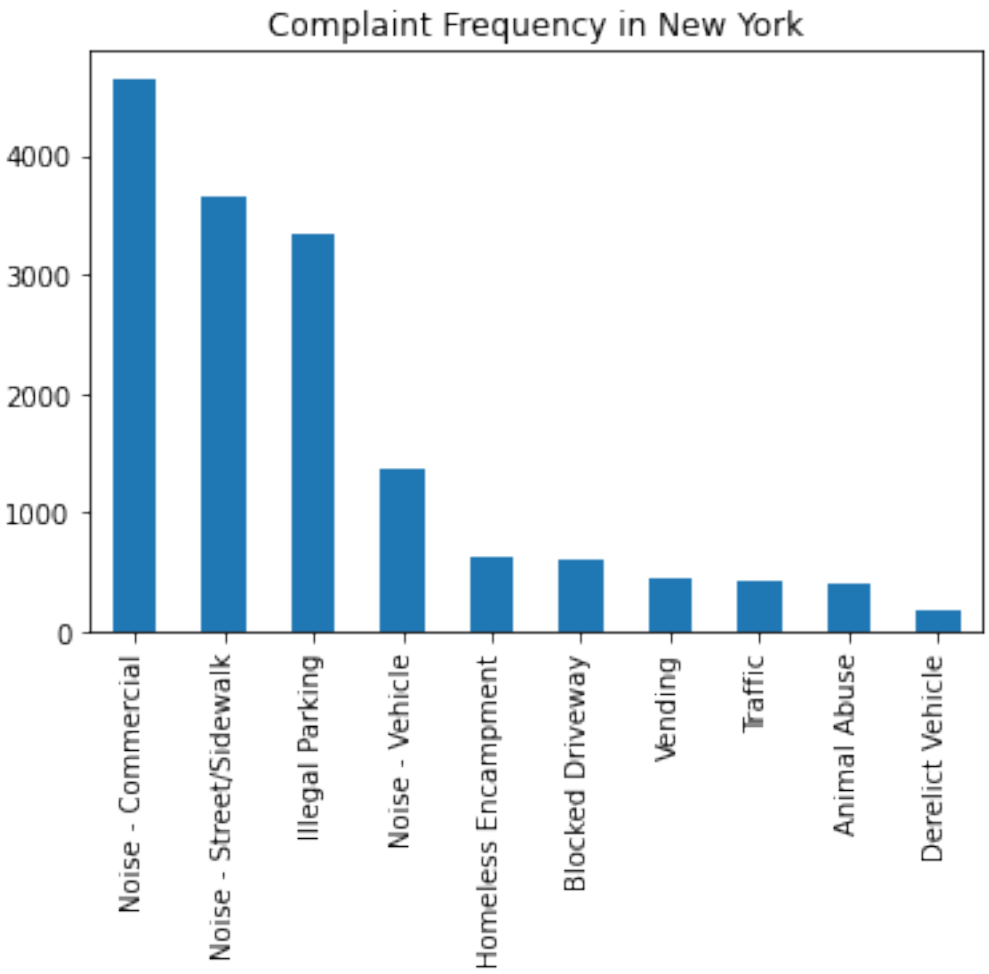
11 Display frequency of top 10 overall complaint types

[16]:
```
data['Complaint Type'].value_counts().head(10).plot(kind = 'bar', title =␣
↪'Overall Complaint Type Frequency');
```



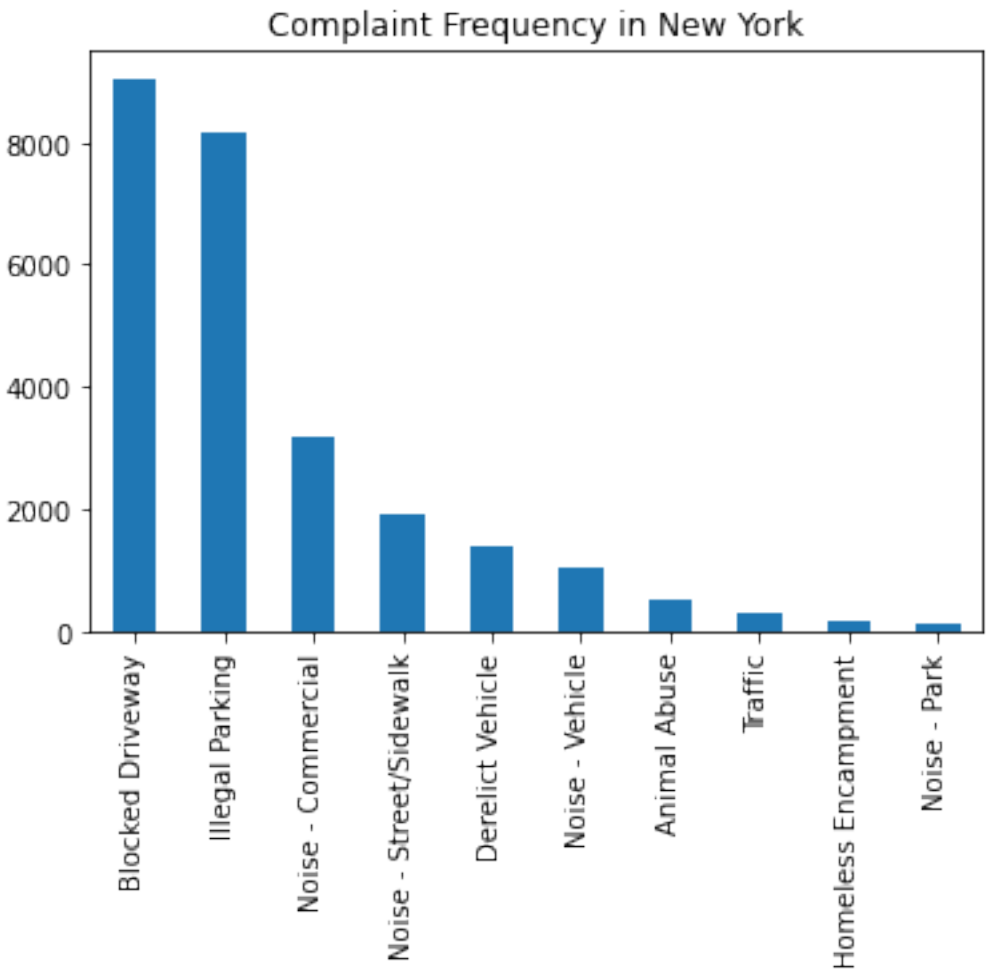
12 Frequency chart for the ten most common complaints in New York

[15]:
```
data_ny = data.loc[data['City'] == 'New York']
data_ny['Complaint Type'].value_counts().head(10).plot(kind='bar', title =␣
↪'Complaint Frequency in New York');
```



13 Frequency chart for the ten most common complaints in Brooklyn

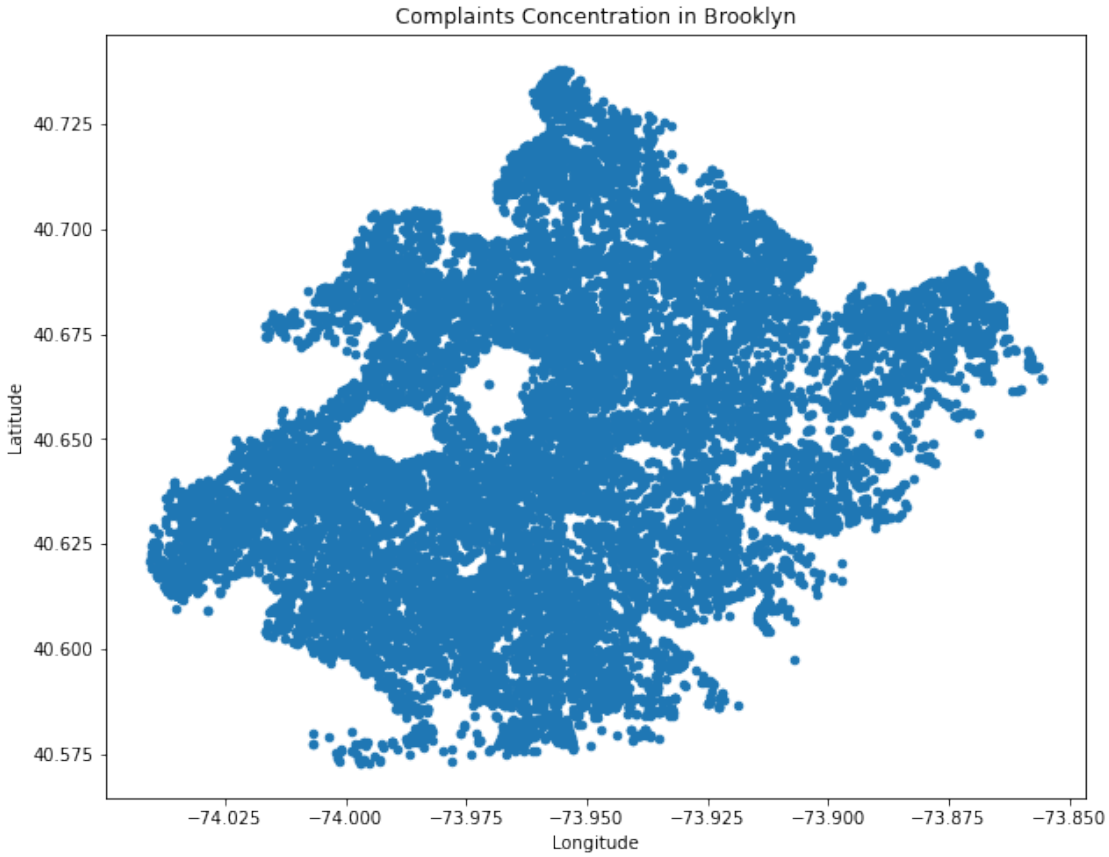
[17]:
```
data_bkn = data.loc[data['City'] == 'Brooklyn']
data_bkn['Complaint Type'].value_counts().head(10).plot(kind='bar', title =␣
↪'Complaint Frequency in New York');
```



14 From the above diagrams we can see that noise complaints are more common in New York, whereas issues related to parking are more prevalanet in Brooklyn.

15 Draw scatter plot for complaint concentration across Brooklyn

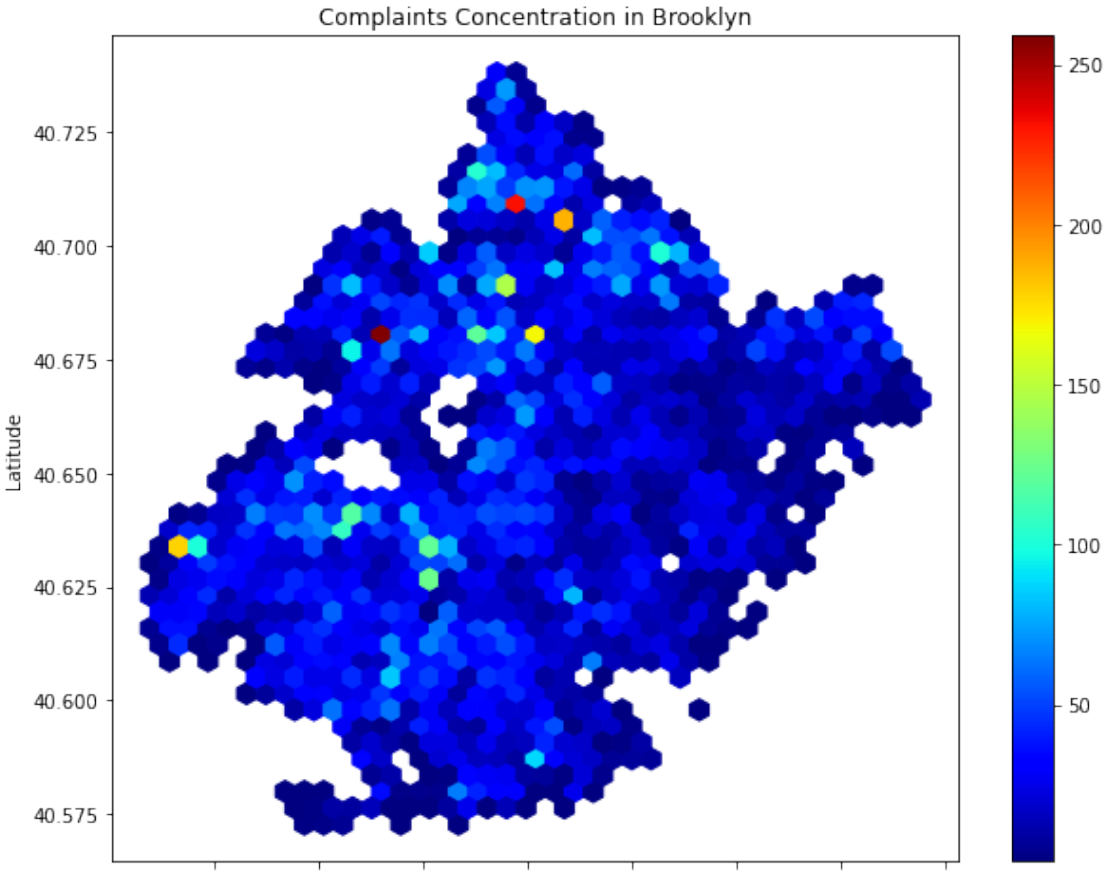
[18]:
```
data_bkn[['Longitude', 'Latitude']].plot(kind = 'scatter', x='Longitude',␣
↪y='Latitude', title = 'Complaints Concentration in Brooklyn', figsize = (10,␣
↪8));
```



16 Scatter plot is hard to read, let's draw a hexbin plot.

17 Draw hexbin plot for complaint concentration across Brooklyn

[20]:
```
data_bkn[['Longitude', 'Latitude']].plot(kind = 'hexbin', x='Longitude',␣
↪y='Latitude', gridsize=40,
   colormap = 'jet', mincnt=1, title = 'Complaints Concentration in Brooklyn',␣
↪figsize = (10, 8));
```



18 Analysing the response time.

19 Change date columns to YYYY-MM-DD format

[21]:
```
data['Created Date'] = pd.to_datetime(data['Created Date'])
data['Closed Date'] = pd.to_datetime(data['Closed Date'])
```

20 Check whether the date are in the correct order

[22]:
```
data.loc[data['Created Date']>=data['Closed Date']].shape
```

[22]: (0, 48)

[26]:
```
# Drop rows where 'Closed Date' is empty
data = data[data['Closed Date'].notna()]
```

21 Calculate resolution time in terms of days

[28]:
```
data['Resolution Time'] = (data['Closed Date'] - data['Created Date']).dt.days
data.head()
```

[28]:
```
   Unique Key        Created Date         Closed Date Agency  \
0    32310363 2015-12-31 23:59:45 2016-01-01 00:55:00   NYPD
1    32309934 2015-12-31 23:59:44 2016-01-01 01:26:00   NYPD
2    32309159 2015-12-31 23:59:29 2016-01-01 04:51:00   NYPD
3    32305098 2015-12-31 23:57:46 2016-01-01 07:43:00   NYPD
4    32306529 2015-12-31 23:56:58 2016-01-01 03:24:00   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  … Bridge Highway Name Bridge Highway Direction  \
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
4          87-14 57 ROAD  …                 NaN                      NaN

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

    Latitude  Longitude                                  Location  \
0  40.865682 -73.923501   (40.86568153633767, -73.92350095571744)
1  40.775945 -73.915094  (40.775945312321085, -73.91509393898605)
2  40.870325 -73.888525  (40.870324522111424, -73.88852464418646)
3  40.835994 -73.828379   (40.83599404683083, -73.82837939584206)
4  40.733060 -73.874170  (40.733059618956815, -73.87416975810375)
```

```
  Resolution Time
0               0
1               0
2               0
3               0
4               0

[5 rows x 49 columns]
```

22 Resolution time according to complaint type

[29]:
```
data.groupby('Complaint Type')['Resolution Time'].mean().sort_values()
```

[29]:
```
Complaint Type
Disorderly Youth              0.000000
Graffiti                      0.000000
Illegal Fireworks             0.000000
Posting Advertisement         0.002882
Noise - House of Worship      0.006410
Noise - Park                  0.008130
Homeless Encampment           0.009585
Noise - Commercial            0.009957
Urinating in Public           0.010309
Bike/Roller/Skate Chronic     0.012821
Noise - Vehicle               0.014701
Noise - Street/Sidewalk       0.015391
Traffic                       0.017025
Vending                       0.017380
Illegal Parking               0.024838
Blocked Driveway              0.028389
Drinking                      0.030172
Animal Abuse                  0.049110
Panhandling                   0.083333
Derelict Vehicle              0.095469
Name: Resolution Time, dtype: float64
```

[ ]:
```

```