Project 03: NYC 311 Service Request Analysis

You don't need to limit yourself to the number of rows/cells provided. You can add additional rows in each section to add more lines of code.

Happy coding!

In [1]:
```
%matplotlib inline
import pandas as pd
import numpy as np
import matplotlib.pyplot as plt
import seaborn as sns
from pandas import Series, DataFrame
```

In [5]:
```
df=pd.read_csv("311_Service_Requests_from_2010_to_Present.csv")
```
```
/anaconda3/lib/python3.7/site-packages/IPython/core/interactiveshell.py:2785: DtypeWarning: Columns (48,49) have mixed types. Specify dtype option on import or set low_memory=False.
  interactivity=interactivity, compiler=compiler, result=result)
```

In [6]:
```
df.head()
```
Out[6]:

	Unique Key	Created Date	Closed Date	Agency	Agency Name	Complaint Type	Descriptor	Location Type	Incident Zip	Incident Address
0	32310363	12/31/2015 11:59:45 PM	01/01/2016 12:55:15 AM	NYPD	New York City Police Department	Noise - Street/Sidewalk	Loud Music/Party	Street/Sidewalk	10034.0	71 VERMILYEA AVENUE
1	32309934	12/31/2015 11:59:44 PM	01/01/2016 01:26:57 AM	NYPD	New York City Police Department	Blocked Driveway	No Access	Street/Sidewalk	11105.0	27-07 23 AVENUE
2	32309159	12/31/2015 11:59:29 PM	01/01/2016 04:51:03 AM	NYPD	New York City Police Department	Blocked Driveway	No Access	Street/Sidewalk	10458.0	2897 VALENTINE AVENUE
3	32305098	12/31/2015 11:57:46 PM	01/01/2016 07:43:13 AM	NYPD	New York City Police Department	Illegal Parking	Commercial Overnight Parking	Street/Sidewalk	10461.0	2940 BAISLEY AVENUE
4	32306529	12/31/2015 11:56:58 PM	01/01/2016 03:24:42 AM	NYPD	New York City Police Department	Illegal Parking	Blocked Sidewalk	Street/Sidewalk	11373.0	87-14 57 ROAD

5 rows × 53 columns

In [7]:
```
df.describe()
```
Out[7]:

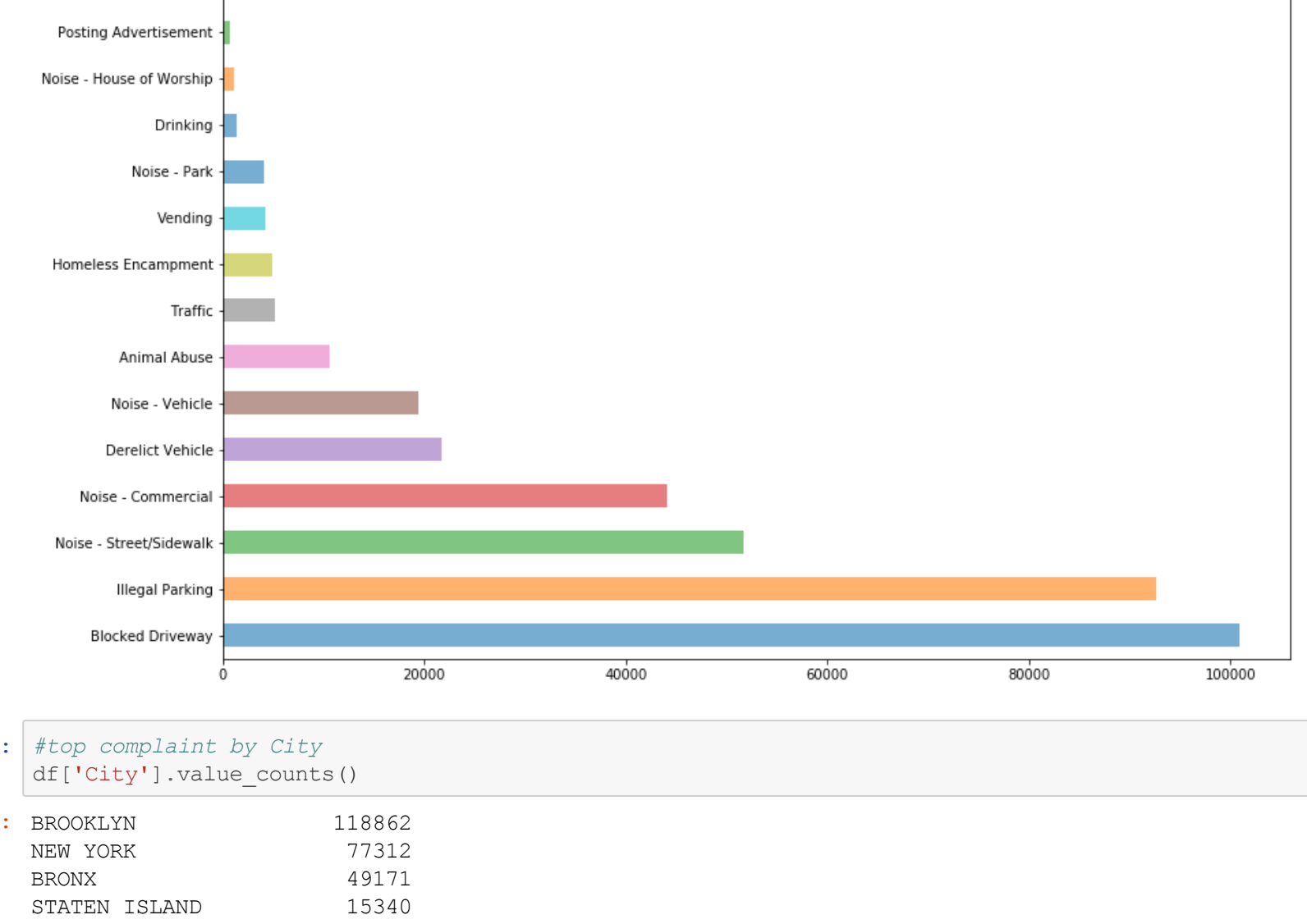
	Unique Key	Incident Zip	X Coordinate (State Plane)	Y Coordinate (State Plane)	School or Citywide Complaint	Vehicle Type	Taxi Company Borough	Taxi Pick Up Location	Garage Lot Name	
count	3.645580e+05	361560.000000	3.605280e+05	360528.000000	0.0	0.0	0.0	0.0	0.0	360528
mean	3.106595e+07	10858.496659	1.005043e+06	203425.305782	NaN	NaN	NaN	NaN	NaN	40
std	7.331531e+05	578.263114	2.196362e+04	29842.192857	NaN	NaN	NaN	NaN	NaN	0
min	2.960737e+07	83.000000	9.133570e+05	121185.000000	NaN	NaN	NaN	NaN	NaN	40
25%	3.049938e+07	10314.000000	9.919460e+05	182945.000000	NaN	NaN	NaN	NaN	NaN	40
50%	3.108795e+07	11209.000000	1.003470e+06	201023.000000	NaN	NaN	NaN	NaN	NaN	40
75%	3.167433e+07	11238.000000	1.019134e+06	222790.000000	NaN	NaN	NaN	NaN	NaN	40
max	3.231065e+07	11697.000000	1.067186e+06	271876.000000	NaN	NaN	NaN	NaN	NaN	40

In [8]:
```
df.shape
```
Out[8]:
```
(364558, 53)
```

In [9]:
```
df.columns
```
Out[9]:
```
Index(['Unique Key', 'Created Date', 'Closed Date', 'Agency', 'Agency Name',
       'Complaint Type', 'Descriptor', 'Location Type', 'Incident Zip',
       'Incident Address', 'Street Name', 'Cross Street 1', 'Cross Street 2',
       'Intersection Street 1', 'Intersection Street 2', 'Address Type',
       'City', 'Landmark', 'Facility Type', 'Status', 'Due Date',
       'Resolution Description', 'Resolution Action Updated Date',
       'Community Board', 'Borough', 'X Coordinate (State Plane)',
       'Y Coordinate (State Plane)', 'Park Facility Name', 'Park Borough',
       'School Name', 'School Number', 'School Region', 'School Code',
       'School Phone Number', 'School Address', 'School City', 'School State',
       'School Zip', 'School Not Found', 'School or Citywide Complaint',
       'Vehicle Type', 'Taxi Company Borough', 'Taxi Pick Up Location',
       'Bridge Highway Name', 'Bridge Highway Direction', 'Road Ramp',
       'Bridge Highway Segment', 'Garage Lot Name', 'Ferry Direction',
       'Ferry Terminal Name', 'Latitude', 'Longitude', 'Location'],
      dtype='object')
```

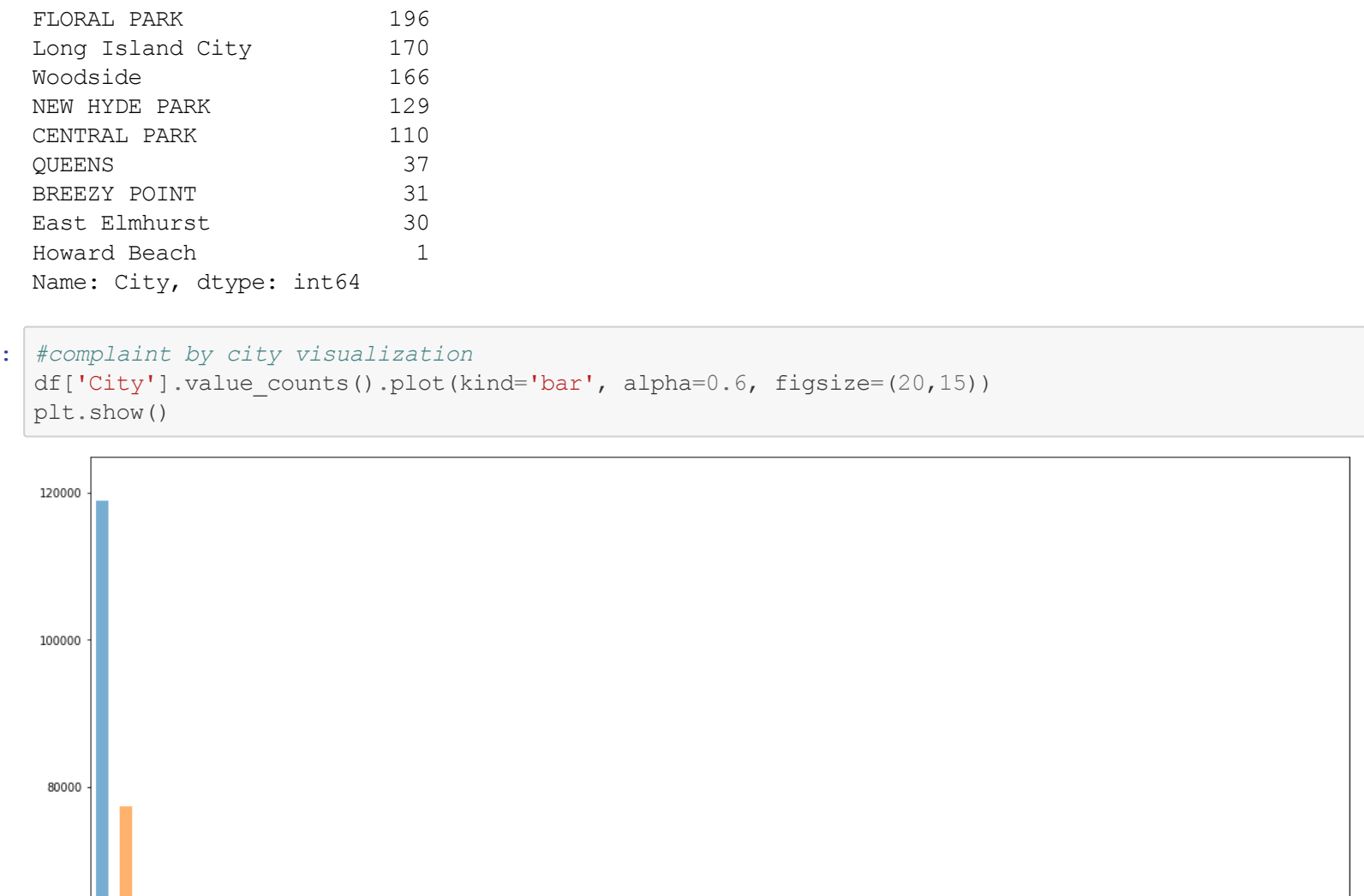
In [10]:
```
df.info()
```
```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 364558 entries, 0 to 364557
Data columns (total 53 columns):
Unique Key                        364558 non-null int64
Created Date                      364558 non-null object
Closed Date                       362177 non-null object
Agency                            364558 non-null object
Agency Name                       364558 non-null object
Complaint Type                    364558 non-null object
Descriptor                        358057 non-null object
Location Type                     364425 non-null object
Incident Zip                      361560 non-null float64
Incident Address                  312859 non-null object
Street Name                       312859 non-null object
Cross Street 1                    307370 non-null object
Cross Street 2                    306753 non-null object
Intersection Street 1             51120 non-null object
Intersection Street 2             50512 non-null object
Address Type                      361306 non-null object
City                              361561 non-null object
Landmark                          375 non-null object
Facility Type                     362169 non-null object
Status                            364558 non-null object
Due Date                          364555 non-null object
Resolution Description            364558 non-null object
Resolution Action Updated Date    362156 non-null object
Community Board                   364558 non-null object
Borough                           364558 non-null object
X Coordinate (State Plane)        360528 non-null float64
Y Coordinate (State Plane)        360528 non-null float64
Park Facility Name                364558 non-null object
Park Borough                      364558 non-null object
School Name                       364558 non-null object
School Number                     364558 non-null object
School Region                     364557 non-null object
School Code                       364557 non-null object
School Phone Number               364558 non-null object
School Address                    364558 non-null object
School City                       364558 non-null object
School State                      364558 non-null object
School Zip                        364557 non-null object
School Not Found                  364558 non-null object
School or Citywide Complaint      0 non-null float64
Vehicle Type                      0 non-null float64
Taxi Company Borough              0 non-null float64
Taxi Pick Up Location             0 non-null float64
Bridge Highway Name               297 non-null object
Bridge Highway Direction          297 non-null object
Road Ramp                         262 non-null object
Bridge Highway Segment            262 non-null object
Garage Lot Name                   0 non-null float64
Ferry Direction                   1 non-null object
Ferry Terminal Name               2 non-null object
Latitude                          360528 non-null float64
Longitude                         360528 non-null float64
Location                          360528 non-null object
dtypes: float64(10), int64(1), object(42)
memory usage: 147.4+ MB
```

In [12]:
```
#complaint type breakdown with bars plot to figure out the top 10 complaints
df['Complaint Type'].value_counts().plot(kind='barh',alpha=0.6, figsize=(14,15))
plt.show()
```



In [13]:
```
#top complaint by City
df['City'].value_counts()
```
Out[13]:
```
BROOKLYN               118862
NEW YORK                77312
BRONX                   49171
STATEN ISLAND           15340
JAMAICA                  8932
ASTORIA                  7991
FLUSHING                 7487
RIDGEWOOD                6392
CORONA                   5383
WOODSIDE                 4357
EAST ELMHURST            3558
OZONE PARK               3446
ELMHURST                 3438
SOUTH RICHMOND HILL      3431
MASPETH                  3118
WOODHAVEN                3103
LONG ISLAND CITY         3028
SOUTH OZONE PARK         2668
FRESH MEADOWS            2453
RICHMOND HILL            2335
MIDDLE VILLAGE           2291
QUEENS VILLAGE           2251
FOREST HILLS             2122
JACKSON HEIGHTS          2106
REGO PARK                1807
BAYSIDE                  1550
COLLEGE POINT            1544
FAR ROCKAWAY             1397
WHITESTONE               1369
HOLLIS                   1231
HOWARD BEACH             1144
SPRINGFIELD GARDENS      1094
ROSEDALE                 1091
SAINT ALBANS             1047
KEW GARDENS              1008
SUNNYSIDE                 944
Astoria                   906
ROCKAWAY PARK             831
OAKLAND GARDENS           717
LITTLE NECK               712
CAMBRIA HEIGHTS           617
BELLEROSE                 487
GLEN OAKS                 361
ARVERNE                   259
FLORAL PARK               196
Long Island City          170
Woodside                  166
NEW HYDE PARK             129
CENTRAL PARK              110
QUEENS                     37
BREEZY POINT               31
East Elmhurst              30
Howard Beach                1
Name: City, dtype: int64
```

In [15]:
```
#complaint by city visualization
df['City'].value_counts().plot(kind='bar', alpha=0.6, figsize=(20,15))
plt.show()
```



In [16]:
```
#complaint type with city
df[['Complaint Type','City']].head(50)
```
Out[16]:

	Complaint Type	City
0	Noise - Street/Sidewalk	NEW YORK
1	Blocked Driveway	ASTORIA
2	Blocked Driveway	BRONX
3	Illegal Parking	BRONX
4	Illegal Parking	ELMHURST
5	Illegal Parking	BROOKLYN
6	Illegal Parking	NEW YORK
7	Blocked Driveway	BRONX
8	Illegal Parking	KEW GARDENS
9	Blocked Driveway	BROOKLYN
10	Blocked Driveway	JACKSON HEIGHTS
11	Blocked Driveway	BRONX
12	Noise - Street/Sidewalk	BRONX
13	Illegal Parking	BROOKLYN
14	Derelict Vehicle	MIDDLE VILLAGE
15	Blocked Driveway	REGO PARK
16	Blocked Driveway	SAINT ALBANS
17	Noise - Commercial	BROOKLYN
18	Noise - Commercial	BROOKLYN
19	Noise - Street/Sidewalk	NEW YORK
20	Illegal Parking	BRONX
21	Illegal Parking	MIDDLE VILLAGE
22	Noise - Commercial	BROOKLYN
23	Illegal Parking	NEW YORK
24	Blocked Driveway	BRONX
25	Blocked Driveway	JAMAICA
26	Noise - House of Worship	NEW YORK
27	Blocked Driveway	SOUTH RICHMOND HILL
28	Illegal Parking	NEW YORK
29	Noise - Commercial	BROOKLYN
30	Noise - Commercial	BROOKLYN
31	Noise - Commercial	BROOKLYN
32	Blocked Driveway	BRONX
33	Illegal Parking	NaN
34	Blocked Driveway	RIDGEWOOD
35	Illegal Parking	HOWARD BEACH
36	Blocked Driveway	BROOKLYN
37	Illegal Parking	FOREST HILLS
38	Noise - Street/Sidewalk	BROOKLYN
39	Posting Advertisement	STATEN ISLAND
40	Noise - Commercial	NEW YORK
41	Noise - Commercial	BROOKLYN
42	Posting Advertisement	STATEN ISLAND
43	Noise - Commercial	BROOKLYN
44	Blocked Driveway	OZONE PARK
45	Noise - Commercial	NEW YORK
46	Posting Advertisement	STATEN ISLAND
47	Illegal Parking	NEW YORK
48	Noise - Commercial	NEW YORK
49	Posting Advertisement	STATEN ISLAND

In [17]:
```
#checking for null values
df.isnull().sum()
```
Out[17]:
```
Unique Key                             0
Created Date                           0
Closed Date                         2381
Agency                                 0
Agency Name                            0
Complaint Type                         0
Descriptor                          6501
Location Type                        133
Incident Zip                        2998
Incident Address                   51699
Street Name                        51699
Cross Street 1                     57188
Cross Street 2                     57805
Intersection Street 1             313438
Intersection Street 2             314046
Address Type                        3252
City                                2997
Landmark                          364183
Facility Type                       2389
Status                                 0
Due Date                               3
Resolution Description                 0
Resolution Action Updated Date      2402
Community Board                        0
Borough                                0
X Coordinate (State Plane)          4030
Y Coordinate (State Plane)          4030
Park Facility Name                     0
Park Borough                           0
School Name                            0
School Number                          0
School Region                          1
School Code                            1
School Phone Number                    0
School Address                         0
School City                            0
School State                           0
School Zip                             1
School Not Found                       0
School or Citywide Complaint      364558
Vehicle Type                      364558
Taxi Company Borough              364558
Taxi Pick Up Location             364558
Bridge Highway Name               364261
Bridge Highway Direction          364261
Road Ramp                         364296
Bridge Highway Segment            364296
Garage Lot Name                   364558
Ferry Direction                   364557
Ferry Terminal Name               364556
Latitude                            4030
Longitude                           4030
Location                            4030
dtype: int64
```

In [19]:
```
#fix blank values in City column and shape
df['City'].dropna(inplace=True)
```

In [21]:
```
df['City'].shape
```
Out[21]:
```
(361561,)
```

In [24]:
```
#group by complaint
groupByComplainType=df.groupby('Complaint Type')
```

In [26]:
```
grp_data=groupByComplainType.get_group('Illegal Parking')
grp_data.shape
```
Out[26]:
```
(92679, 53)
```

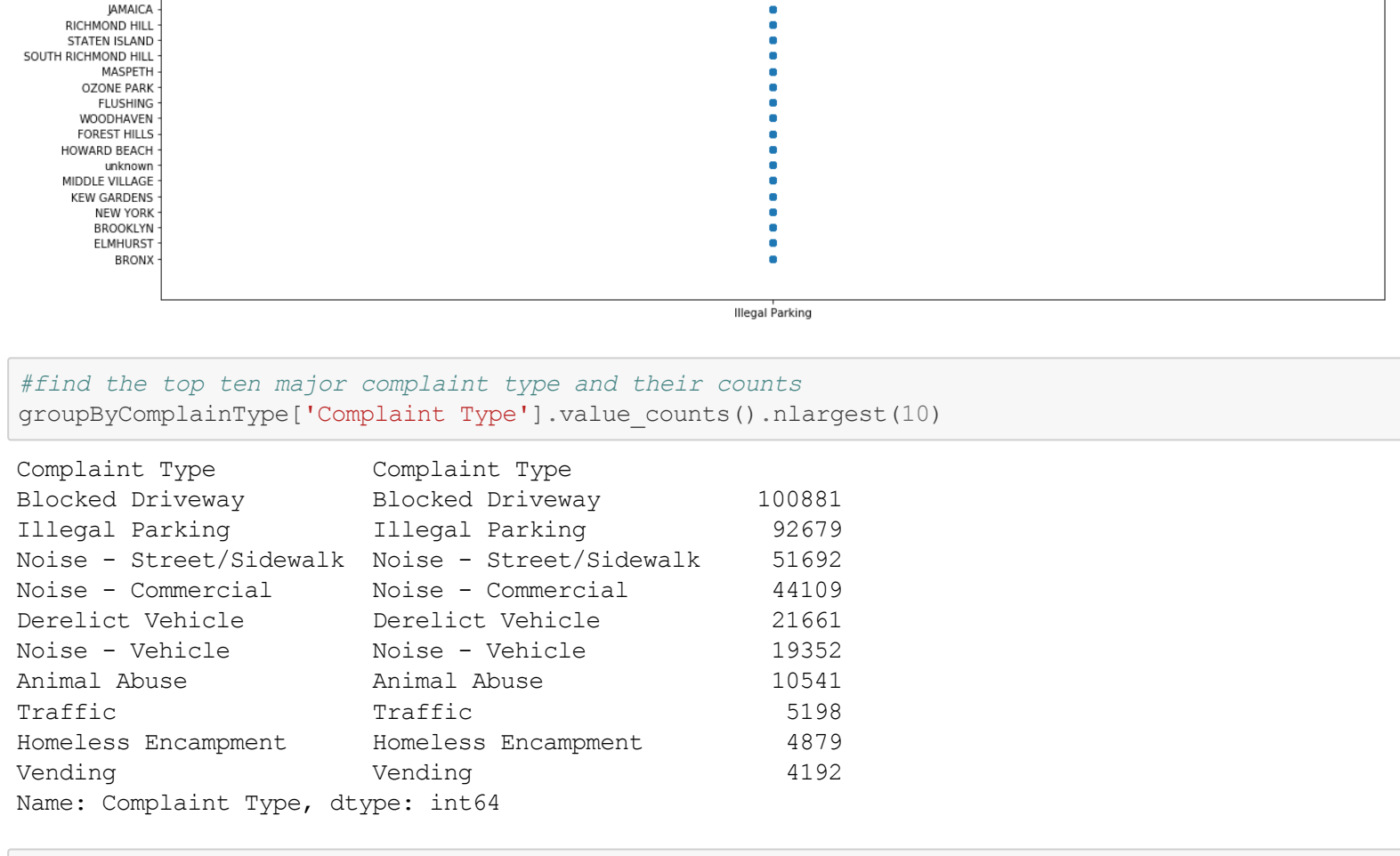
In [27]:
```
df.isnull().sum()
```
Out[27]:
```
Unique Key                             0
Created Date                           0
Closed Date                         2381
Agency                                 0
Agency Name                            0
Complaint Type                         0
Descriptor                          6501
Location Type                        133
Incident Zip                        2998
Incident Address                   51699
Street Name                        51699
Cross Street 1                     57188
Cross Street 2                     57805
Intersection Street 1             313438
Intersection Street 2             314046
Address Type                        3252
City                                2997
Landmark                          364183
Facility Type                       2389
Status                                 0
Due Date                               3
Resolution Description                 0
Resolution Action Updated Date      2402
Community Board                        0
Borough                                0
X Coordinate (State Plane)          4030
Y Coordinate (State Plane)          4030
Park Facility Name                     0
Park Borough                           0
School Name                            0
School Number                          0
School Region                          1
School Code                            1
School Phone Number                    0
School Address                         0
School City                            0
School State                           0
School Zip                             1
School Not Found                       0
School or Citywide Complaint      364558
Vehicle Type                      364558
Taxi Company Borough              364558
Taxi Pick Up Location             364558
Bridge Highway Name               364261
Bridge Highway Direction          364261
Road Ramp                         364296
Bridge Highway Segment            364296
Garage Lot Name                   364558
Ferry Direction                   364557
Ferry Terminal Name               364556
Latitude                            4030
Longitude                           4030
Location                            4030
dtype: int64
```

In [29]:
```
#count of nulls in group city data
grp_data['City'].isnull().sum()
```
Out[29]:
```
1267
```

In [60]:
```
#filling nan with unknown
grp_data['City'].fillna('unknown city',inplace=True)
```
```
/anaconda3/lib/python3.7/site-packages/pandas/core/generic.py:5434: SettingWithCopyWarning:
A value is trying to be set on a copy of a slice from a DataFrame

See the caveats in the documentation: http://pandas.pydata.org/pandas-docs/stable/indexing.html#indexing-view-versus-copy
  self._update_inplace(new_data)
```

In [39]:
```
 #plot to showing the city raising 'illegal parking 'complaint
plt.figure(figsize=(20,15))
plt.scatter(grp_data['Complaint Type'],grp_data['City'])
plt.title('Illegal Parking complaint raise by different city')
plt.show()
```



In [40]:
```
#find the top ten major complaint type and their counts
groupByComplainType['Complaint Type'].value_counts().nlargest(10)
```
Out[40]:
```
Complaint Type           Complaint Type
Blocked Driveway         Blocked Driveway           100881
Illegal Parking          Illegal Parking             92679
Noise - Street/Sidewalk  Noise - Street/Sidewalk     51692
Noise - Commercial       Noise - Commercial          44109
Derelict Vehicle         Derelict Vehicle            21661
Noise - Vehicle          Noise - Vehicle             19352
Animal Abuse             Animal Abuse                10541
Traffic                  Traffic                      5198
Homeless Encampment      Homeless Encampment          4879
Vending                  Vending                      4192
Name: Complaint Type, dtype: int64
```

In [63]:
```
#fix location type NaN with unknown loc
grp_data['Location Type'].fillna('Unknown loc',inplace=True)
```
```
/anaconda3/lib/python3.7/site-packages/pandas/core/generic.py:5434: SettingWithCopyWarning:
A value is trying to be set on a copy of a slice from a DataFrame

See the caveats in the documentation: http://pandas.pydata.org/pandas-docs/stable/indexing.html#indexing-view-versus-copy
  self._update_inplace(new_data)
```

In [64]:
```
grp_data['Location Type'].values
```
Out[64]:
```
array(['Street/Sidewalk', 'Street/Sidewalk', 'Street/Sidewalk', ...,
       'Street/Sidewalk', 'Street/Sidewalk', 'Street/Sidewalk'],
      dtype=object)
```

In [65]:
```
#count all null values in grouped location type column data
grp_data['Location Type'].isnull().sum()
```
Out[65]:
```
0
```

In [66]:
```
#plot major complaint type 'Illegal parking' against location type for any pattern
plt.figure(figsize=(3,3))
plt.scatter(grp_data['Complaint Type'],grp_data['Location Type'])
plt.title='Plot complaint type Heating against location type'
plt.xlabel='Complaint Type'
plt.ylabel='Location Type'
plt.show()
```

In [ ]:
```
#Plot above gives us a clear picture of the fact that all the complaints rasied of type "Illeagal parking" .
#occured only in Street/Sidewalk! This shows that majority of complaints recorded was from unknown location!
```