CustomerRequestService

July 7, 2023

[1]:
```
import numpy as np
import pandas as pd
import matplotlib.pyplot as plt
import seaborn as sns
```

[2]:
```
## 1.1 Import the dataset
service_request_nyc= pd.read_csv('311_Service_Requests_from_2010_to_Present.
↪csv')
```

```
/usr/local/lib/python3.7/site-packages/IPython/core/interactiveshell.py:3063:
DtypeWarning: Columns (48,49) have mixed types.Specify dtype option on import or
set low_memory=False.
  interactivity=interactivity, compiler=compiler, result=result)
```

0.1 1.Understand the Data sets

[3]:
```
# 1.2 Visualize the dataset
service_request_nyc.info()
```

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 364558 entries, 0 to 364557
Data columns (total 53 columns):
 #   Column                          Non-Null Count   Dtype
---  ------                          --------------   -----
 0   Unique Key                      364558 non-null  int64
 1   Created Date                    364558 non-null  object
 2   Closed Date                     362177 non-null  object
 3   Agency                          364558 non-null  object
 4   Agency Name                     364558 non-null  object
 5   Complaint Type                  364558 non-null  object
 6   Descriptor                      358057 non-null  object
 7   Location Type                   364425 non-null  object
 8   Incident Zip                    361560 non-null  float64
 9   Incident Address                312859 non-null  object
 10  Street Name                     312859 non-null  object
 11  Cross Street 1                  307370 non-null  object
 12  Cross Street 2                  306753 non-null  object
```

```
 13  Intersection Street 1           51120 non-null   object
 14  Intersection Street 2           50512 non-null   object
 15  Address Type                    361306 non-null  object
 16  City                            361561 non-null  object
 17  Landmark                        375 non-null     object
 18  Facility Type                   362169 non-null  object
 19  Status                          364558 non-null  object
 20  Due Date                        364555 non-null  object
 21  Resolution Description          364558 non-null  object
 22  Resolution Action Updated Date  362156 non-null  object
 23  Community Board                 364558 non-null  object
 24  Borough                         364558 non-null  object
 25  X Coordinate (State Plane)      360528 non-null  float64
 26  Y Coordinate (State Plane)      360528 non-null  float64
 27  Park Facility Name              364558 non-null  object
 28  Park Borough                    364558 non-null  object
 29  School Name                     364558 non-null  object
 30  School Number                   364558 non-null  object
 31  School Region                   364557 non-null  object
 32  School Code                     364557 non-null  object
 33  School Phone Number             364558 non-null  object
 34  School Address                  364558 non-null  object
 35  School City                     364558 non-null  object
 36  School State                    364558 non-null  object
 37  School Zip                      364557 non-null  object
 38  School Not Found                364558 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             297 non-null     object
 44  Bridge Highway Direction        297 non-null     object
 45  Road Ramp                       262 non-null     object
 46  Bridge Highway Segment          262 non-null     object
 47  Garage Lot Name                 0 non-null       float64
 48  Ferry Direction                 1 non-null       object
 49  Ferry Terminal Name             2 non-null       object
 50  Latitude                        360528 non-null  float64
 51  Longitude                       360528 non-null  float64
 52  Location                        360528 non-null  object
dtypes: float64(10), int64(1), object(42)
memory usage: 147.4+ MB
```

[4]:
```
# 1.2 Visualize the dataset
service_request_nyc
```

[4]:

	Unique Key	Created Date	Closed Date	Agency	…
0	32310363	12/31/2015 11:59:45 PM	01/01/2016 12:55:15 AM	NYPD	
1	32309934	12/31/2015 11:59:44 PM	01/01/2016 01:26:57 AM	NYPD	
2	32309159	12/31/2015 11:59:29 PM	01/01/2016 04:51:03 AM	NYPD	
3	32305098	12/31/2015 11:57:46 PM	01/01/2016 07:43:13 AM	NYPD	
4	32306529	12/31/2015 11:56:58 PM	01/01/2016 03:24:42 AM	NYPD	
…	…	…	…	…	
364553	29609918	01/01/2015 12:04:44 AM	01/01/2015 10:22:31 AM	NYPD	
364554	29608392	01/01/2015 12:04:28 AM	01/01/2015 02:25:02 AM	NYPD	
364555	29607589	01/01/2015 12:01:30 AM	01/01/2015 12:20:33 AM	NYPD	
364556	29610889	01/01/2015 12:01:29 AM	01/01/2015 02:42:22 AM	NYPD	
364557	29611816	01/01/2015 12:00:50 AM	01/01/2015 02:47:50 AM	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	
…	…	…	
364553	New York City Police Department	Illegal Parking	
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	
364557	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	
4	Blocked Sidewalk	Street/Sidewalk	11373.0	
…	…	…	…	
364553	Blocked Hydrant	Street/Sidewalk	11421.0	
364554	Car/Truck Horn	Street/Sidewalk	10468.0	
364555	Loud Music/Party	Street/Sidewalk	10031.0	
364556	No Access	Street/Sidewalk	10466.0	
364557	No Access	Street/Sidewalk	11420.0	

	Incident Address	…	Bridge Highway Name	…
0	71 VERMILYEA AVENUE	…	NaN	
1	27-07 23 AVENUE	…	NaN	
2	2897 VALENTINE AVENUE	…	NaN	
3	2940 BAISLEY AVENUE	…	NaN	
4	87-14 57 ROAD	…	NaN	
…	…	…	…	
364553	84-25 85 ROAD	…	NaN	

```
364554   2555 SEDGWICK AVENUE  …                NaN
364555    508 WEST 139 STREET  …                NaN
364556    931 EAST 226 STREET  …                NaN
364557      123-19 135 STREET  …                NaN

       Bridge Highway Direction Road Ramp Bridge Highway Segment  \
0                           NaN       NaN                    NaN
1                           NaN       NaN                    NaN
2                           NaN       NaN                    NaN
3                           NaN       NaN                    NaN
4                           NaN       NaN                    NaN
…                         …       …                      …
364553                      NaN       NaN                    NaN
364554                      NaN       NaN                    NaN
364555                      NaN       NaN                    NaN
364556                      NaN       NaN                    NaN
364557                      NaN       NaN                    NaN

       Garage Lot Name 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
…                …             …                    …        …
364553             NaN             NaN                  NaN  40.695145
364554             NaN             NaN                  NaN  40.867830
364555             NaN             NaN                  NaN  40.821647
364556             NaN             NaN                  NaN  40.886361
364557             NaN             NaN                  NaN  40.674212

        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)
…           …                                       …
364553 -73.860949   (40.69514470265117, -73.86094888534394)
364554 -73.907178   (40.86782963689454, -73.90717786644662)
364555 -73.950873  (40.821646626438095, -73.95087342885292)
364556 -73.853290   (40.88636077906953, -73.85329048666742)
364557 -73.803585  (40.674211762243935, -73.80358548685278)

[364558 rows x 53 columns]
```

[5]:
```
# 1.3 Print the columns of the DataFrame
service_request_nyc.columns
```

[5]:
```
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')
```

[6]:
```
# 1.4 Identify the shape of the dataset
service_request_nyc.shape
```

[6]:
```
(364558, 53)
```

[7]:
```
## 1.5 Identify the variables with null value
service_request_nyc_null= service_request_nyc.isnull().sum()
service_request_nyc_null
```

[7]:
```
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
```

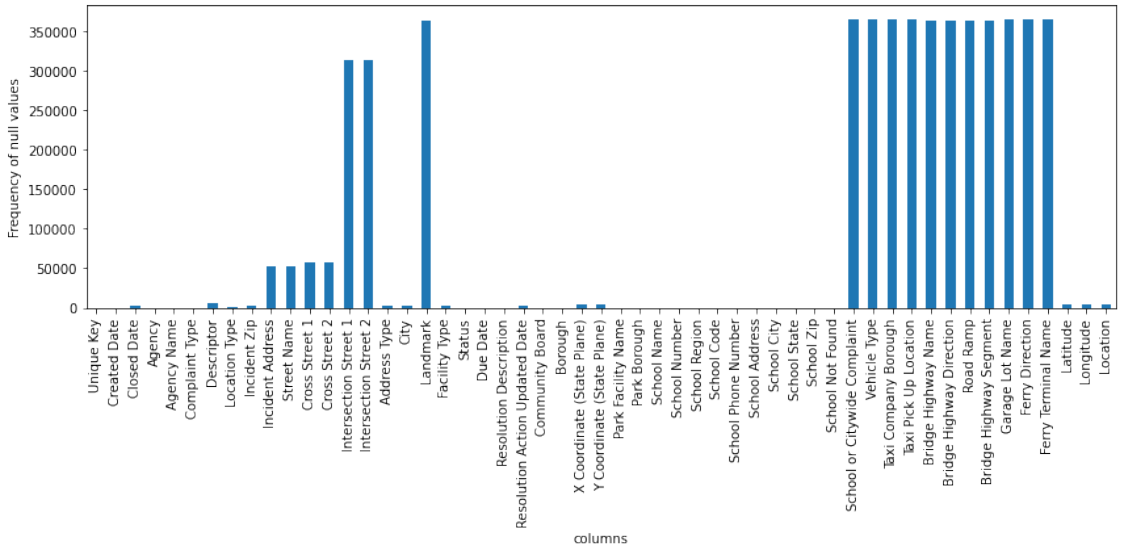
```
Facility Type                         2389
Status                                   0
Due Date                                 3
Resolution Description                   0
Resolution Action Updated Date        2402
Community Board                          0
Borough                                  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
```

0.2 2.Perform basic data exploratory analysis

[8]:
```
# 2.1 Draw a frequency plot to show the number of null values in each column of␣
 ↪the DataFrame

plt.figure(figsize=(12,6))
service_request_nyc_null.plot(kind= 'bar')
plt.xlabel('columns')
plt.ylabel('Frequency of null values')
```

```
plt.xticks(rotation=90)
plt.tight_layout()
plt.show();
```



[9]:
```
## 2.2 2.2 Missing value treatment
## 2.2.1 Remove the records whose Closed Date values are null

service_request_nyc.dropna(subset= ['Closed Date'], inplace= True)
```

[10]:
```
service_request_nyc.info()
```

```
<class 'pandas.core.frame.DataFrame'>
Int64Index: 362177 entries, 0 to 364557
Data columns (total 53 columns):
 #   Column                          Non-Null Count   Dtype
---  ------                          --------------   -----
 0   Unique Key                      362177 non-null  int64
 1   Created Date                    362177 non-null  object
 2   Closed Date                     362177 non-null  object
 3   Agency                          362177 non-null  object
 4   Agency Name                     362177 non-null  object
 5   Complaint Type                  362177 non-null  object
 6   Descriptor                      355681 non-null  object
 7   Location Type                   362047 non-null  object
 8   Incident Zip                    361502 non-null  float64
 9   Incident Address                310491 non-null  object
 10  Street Name                     310491 non-null  object
 11  Cross Street 1                  306846 non-null  object
 12  Cross Street 2                  306713 non-null  object
```

```
 13  Intersection Street 1           50628 non-null   object
 14  Intersection Street 2           50504 non-null   object
 15  Address Type                    361248 non-null  object
 16  City                            361503 non-null  object
 17  Landmark                        375 non-null     object
 18  Facility Type                   362159 non-null  object
 19  Status                          362177 non-null  object
 20  Due Date                        362176 non-null  object
 21  Resolution Description          362177 non-null  object
 22  Resolution Action Updated Date  362138 non-null  object
 23  Community Board                 362177 non-null  object
 24  Borough                         362177 non-null  object
 25  X Coordinate (State Plane)      360470 non-null  float64
 26  Y Coordinate (State Plane)      360470 non-null  float64
 27  Park Facility Name              362177 non-null  object
 28  Park Borough                    362177 non-null  object
 29  School Name                     362177 non-null  object
 30  School Number                   362177 non-null  object
 31  School Region                   362176 non-null  object
 32  School Code                     362176 non-null  object
 33  School Phone Number             362177 non-null  object
 34  School Address                  362177 non-null  object
 35  School City                     362177 non-null  object
 36  School State                    362177 non-null  object
 37  School Zip                      362176 non-null  object
 38  School Not Found                362177 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             297 non-null     object
 44  Bridge Highway Direction        297 non-null     object
 45  Road Ramp                       262 non-null     object
 46  Bridge Highway Segment          262 non-null     object
 47  Garage Lot Name                 0 non-null       float64
 48  Ferry Direction                 0 non-null       object
 49  Ferry Terminal Name             0 non-null       object
 50  Latitude                        360470 non-null  float64
 51  Longitude                       360470 non-null  float64
 52  Location                        360470 non-null  object
dtypes: float64(10), int64(1), object(42)
memory usage: 149.2+ MB
```

[11]:
```
service_request_nyc.head()
```

[11]:
```
   Unique Key            Created Date            Closed Date Agency  \
0    32310363  12/31/2015 11:59:45 PM  01/01/2016 12:55:15 AM   NYPD
```

```
1    32309934  12/31/2015 11:59:44 PM  01/01/2016 01:26:57 AM   NYPD
2    32309159  12/31/2015 11:59:29 PM  01/01/2016 04:51:03 AM   NYPD
3    32305098  12/31/2015 11:57:46 PM  01/01/2016 07:43:13 AM   NYPD
4    32306529  12/31/2015 11:56:58 PM  01/01/2016 03:24:42 AM   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 Garage Lot Name 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 53 columns]
```

0.3 2.3 Analyze the date column, and remove entries that have an incorrect timeline

[12]:
```
## 2.3.1 Calculate the time elapsed in closed and creation date

service_request_nyc['Created Date']= pd.
 ↪to_datetime(service_request_nyc['Created Date'])
service_request_nyc['Created Date']
```

[12]:
```
0         2015-12-31 23:59:45
1         2015-12-31 23:59:44
2         2015-12-31 23:59:29
3         2015-12-31 23:57:46
4         2015-12-31 23:56:58
                  …
364553    2015-01-01 00:04:44
364554    2015-01-01 00:04:28
364555    2015-01-01 00:01:30
364556    2015-01-01 00:01:29
364557    2015-01-01 00:00:50
Name: Created Date, Length: 362177, dtype: datetime64[ns]
```

[13]:
```
service_request_nyc['Closed Date']= pd.to_datetime(service_request_nyc['Closed␣
 ↪Date'])
service_request_nyc['Closed Date']
```

[13]:
```
0         2016-01-01 00:55:15
1         2016-01-01 01:26:57
2         2016-01-01 04:51:03
3         2016-01-01 07:43:13
4         2016-01-01 03:24:42
                  …
364553    2015-01-01 10:22:31
364554    2015-01-01 02:25:02
364555    2015-01-01 00:20:33
364556    2015-01-01 02:42:22
364557    2015-01-01 02:47:50
Name: Closed Date, Length: 362177, dtype: datetime64[ns]
```

[14]:
```
service_request_nyc['Elapsed_Time']= service_request_nyc['Closed Date']-␣
 ↪service_request_nyc['Created Date']
service_request_nyc['Elapsed_Time']
```

[14]:
```
0        0 days 00:55:30
1        0 days 01:27:13
2        0 days 04:51:34
3        0 days 07:45:27
4        0 days 03:27:44
               …
364553   0 days 10:17:47
364554   0 days 02:20:34
364555   0 days 00:19:03
364556   0 days 02:40:53
364557   0 days 02:47:00
Name: Elapsed_Time, Length: 362177, dtype: timedelta64[ns]
```

[15]:
```
## 2.3.2 Convert the calculated date to seconds to get a better representation

(service_request_nyc['Elapsed_Time'].astype(int)/10**9).astype(int)
```

[15]:
```
0          3330
1          5233
2         17494
3         27927
4         12464
          …
364553    37067
364554     8434
364555     1143
364556     9653
364557    10020
Name: Elapsed_Time, Length: 362177, dtype: int64
```

[16]:
```
service_request_nyc['Elapsed_Time'].isnull().sum()
```

[16]:
```
0
```

[17]:
```
service_request_nyc.info()
```

```
<class 'pandas.core.frame.DataFrame'>
Int64Index: 362177 entries, 0 to 364557
Data columns (total 54 columns):
 #   Column                          Non-Null Count   Dtype
---  ------                          --------------   -----
 0   Unique Key                      362177 non-null  int64
 1   Created Date                    362177 non-null  datetime64[ns]
 2   Closed Date                     362177 non-null  datetime64[ns]
 3   Agency                          362177 non-null  object
 4   Agency Name                     362177 non-null  object
 5   Complaint Type                  362177 non-null  object
```

```
6   Descriptor                      355681 non-null  object
7   Location Type                   362047 non-null  object
8   Incident Zip                    361502 non-null  float64
9   Incident Address                310491 non-null  object
10  Street Name                     310491 non-null  object
11  Cross Street 1                  306846 non-null  object
12  Cross Street 2                  306713 non-null  object
13  Intersection Street 1           50628 non-null   object
14  Intersection Street 2           50504 non-null   object
15  Address Type                    361248 non-null  object
16  City                            361503 non-null  object
17  Landmark                        375 non-null     object
18  Facility Type                   362159 non-null  object
19  Status                          362177 non-null  object
20  Due Date                        362176 non-null  object
21  Resolution Description          362177 non-null  object
22  Resolution Action Updated Date  362138 non-null  object
23  Community Board                 362177 non-null  object
24  Borough                         362177 non-null  object
25  X Coordinate (State Plane)      360470 non-null  float64
26  Y Coordinate (State Plane)      360470 non-null  float64
27  Park Facility Name              362177 non-null  object
28  Park Borough                    362177 non-null  object
29  School Name                     362177 non-null  object
30  School Number                   362177 non-null  object
31  School Region                   362176 non-null  object
32  School Code                     362176 non-null  object
33  School Phone Number             362177 non-null  object
34  School Address                  362177 non-null  object
35  School City                     362177 non-null  object
36  School State                    362177 non-null  object
37  School Zip                      362176 non-null  object
38  School Not Found                362177 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             297 non-null     object
44  Bridge Highway Direction        297 non-null     object
45  Road Ramp                       262 non-null     object
46  Bridge Highway Segment          262 non-null     object
47  Garage Lot Name                 0 non-null       float64
48  Ferry Direction                 0 non-null       object
49  Ferry Terminal Name             0 non-null       object
50  Latitude                        360470 non-null  float64
51  Longitude                       360470 non-null  float64
52  Location                        360470 non-null  object
53  Elapsed_Time                    362177 non-null  timedelta64[ns]
```

```
dtypes: datetime64[ns](2), float64(10), int64(1), object(40), timedelta64[ns](1)
memory usage: 152.0+ MB
```

[53]:
```
# 2.3.3 View the descriptive statistics for the newly created column

service_request_nyc['Elapsed_Time'].describe()
```

[53]:
```
count    3.621770e+05
mean     1.511330e+13
std      2.110255e+13
min      6.100000e+10
25%      4.533000e+12
50%      9.616000e+12
75%      1.887800e+13
max      2.134342e+15
Name: Elapsed_Time, dtype: float64
```

[19]:
```
# 2.3.4 Check the number of null values in the Complaint_Type and City columns
service_request_nyc[['Complaint Type', 'City']].isnull().sum()
```

[19]:
```
Complaint Type      0
City              674
dtype: int64
```

[20]:
```
# 2.3.5 Impute the NA value with Unknown City
service_request_nyc['City'].fillna('Unknown City', inplace= True)
service_request_nyc['City'].isnull().sum()
```

[20]:
```
0
```

[21]:
```
service_request_nyc['City']
```

[21]:
```
0                 NEW YORK
1                  ASTORIA
2                    BRONX
3                    BRONX
4                  ELMHURST
                   ...
364553           WOODHAVEN
364554               BRONX
364555            NEW YORK
364556               BRONX
364557    SOUTH OZONE PARK
Name: City, Length: 362177, dtype: object
```

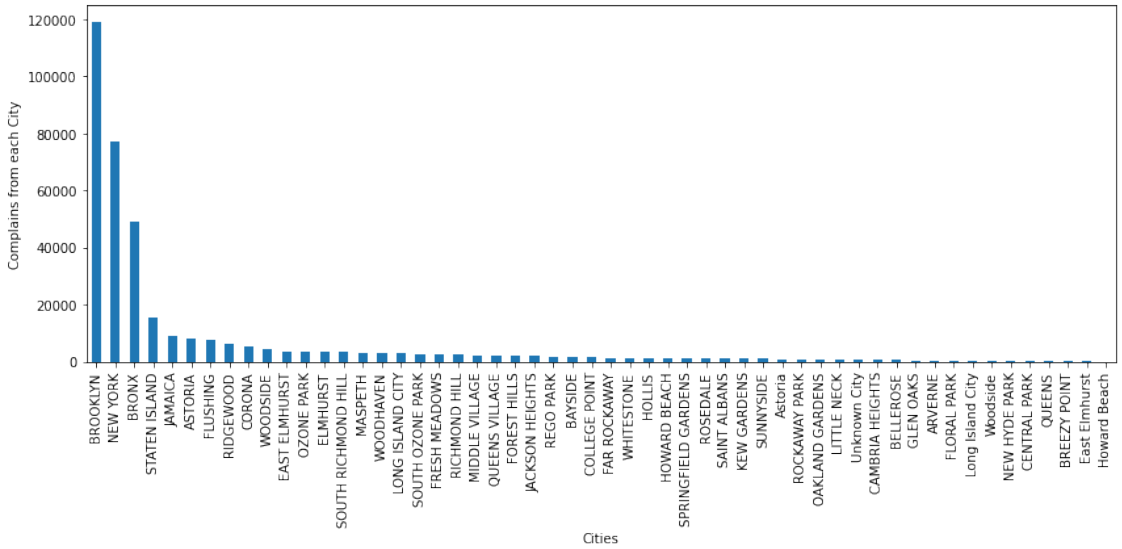
[22]:
```
#2.3.6 Draw a frequency plot for the complaints in each city
service_request_nyc['City'].value_counts()
```

[22]:

BROOKLYN	118849
NEW YORK	77289
BRONX	49166
STATEN ISLAND	15335
JAMAICA	8930
ASTORIA	7991
FLUSHING	7486
RIDGEWOOD	6391
CORONA	5383
WOODSIDE	4357
EAST ELMHURST	3557
OZONE PARK	3446
ELMHURST	3438
SOUTH RICHMOND HILL	3431
MASPETH	3117
WOODHAVEN	3102
LONG ISLAND CITY	3027
SOUTH OZONE PARK	2668
FRESH MEADOWS	2453
RICHMOND HILL	2333
MIDDLE VILLAGE	2291
QUEENS VILLAGE	2251
FOREST HILLS	2122
JACKSON HEIGHTS	2105
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	905
ROCKAWAY PARK	831
OAKLAND GARDENS	717
LITTLE NECK	712
Unknown City	674
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
```

[23]:
```
plt.figure(figsize=(12,6))
service_request_nyc['City'].value_counts().plot(kind= 'bar')
plt.xlabel('Cities')
plt.ylabel('Complains from each City')
plt.xticks(rotation=90)
plt.tight_layout()
plt.show();
```



[24]:
```
service_request_nyc['City'].value_counts().iloc[0]
```

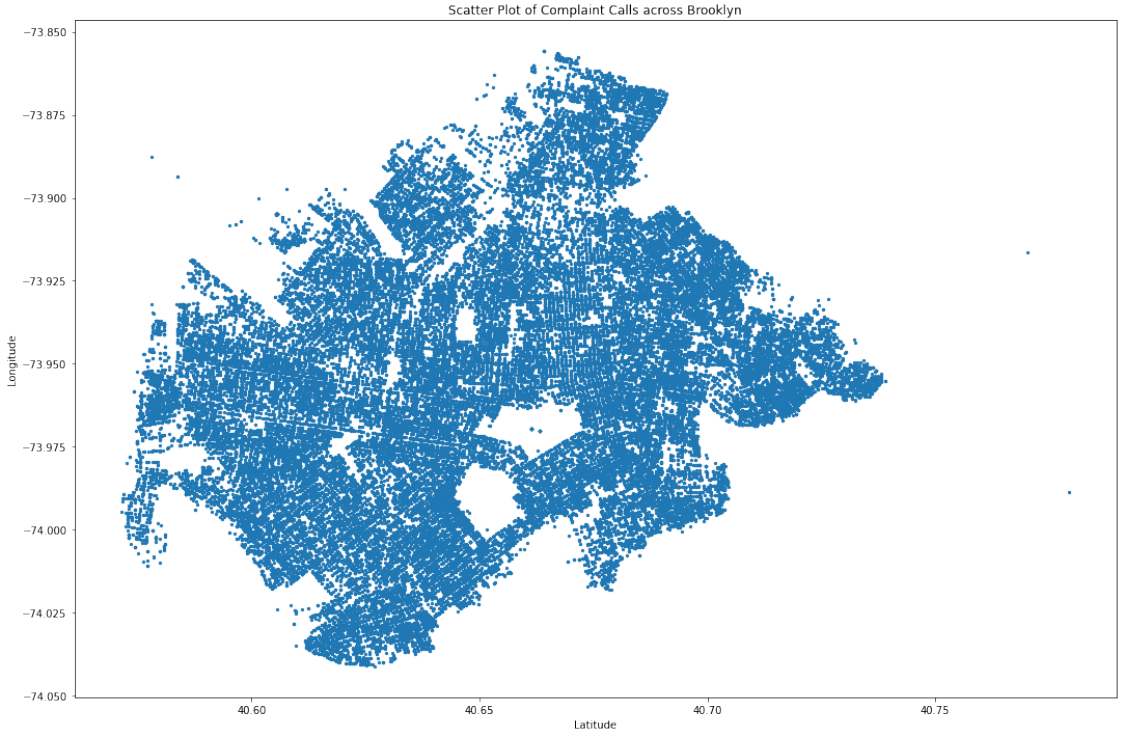
[24]:
```
118849
```

[25]:
```
## 2.3.7 Create a scatter and hexbin plot of the concentration of complaints␣
↪across Brooklyn
service_request_nyc_brooklyn =␣
↪service_request_nyc[service_request_nyc['City']== 'BROOKLYN']
```

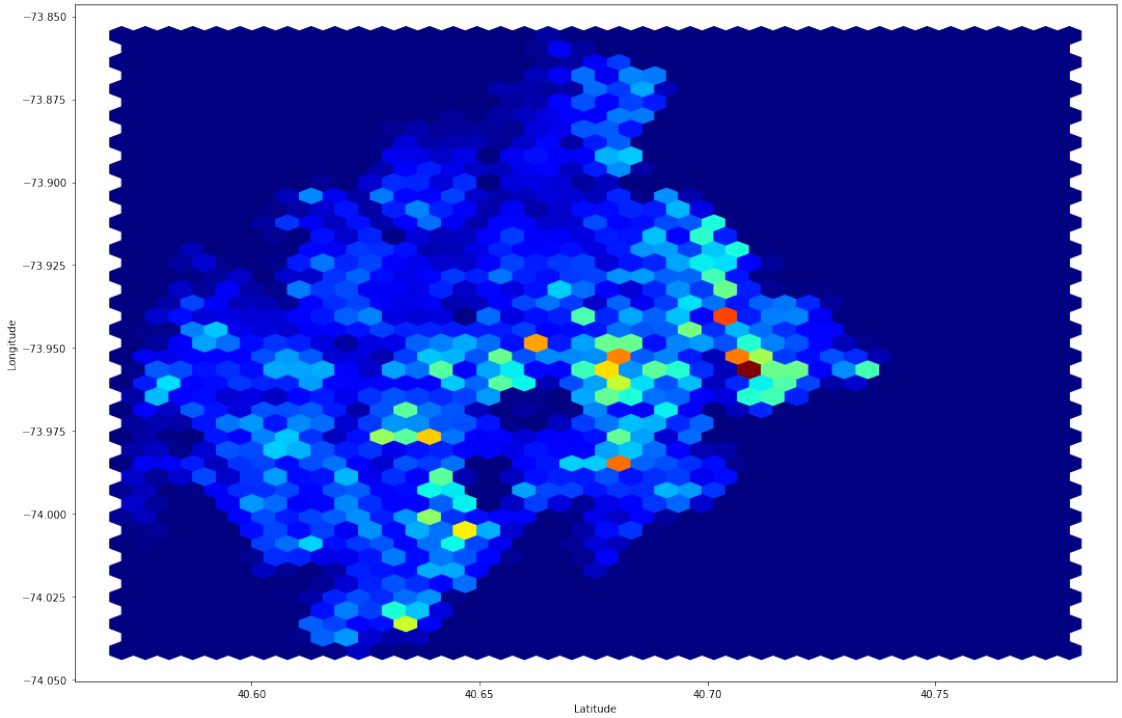
[26]:
```
plt.figure(figsize=(18,12))
```

```
plt.scatter(service_request_nyc_brooklyn['Latitude'],␣
 ↪service_request_nyc_brooklyn['Longitude'], s=5)
plt.xlabel('Latitude')
plt.ylabel('Longitude')
plt.xticks= 'o'
plt.title('Scatter Plot of Complaint Calls across Brooklyn')
plt.show()

plt.figure(figsize=(18,12))
plt.hexbin(service_request_nyc_brooklyn['Latitude'],␣
 ↪service_request_nyc_brooklyn['Longitude'], gridsize= 40,
          cmap= 'jet')
plt.xlabel('Latitude')
plt.ylabel('Longitude')
```

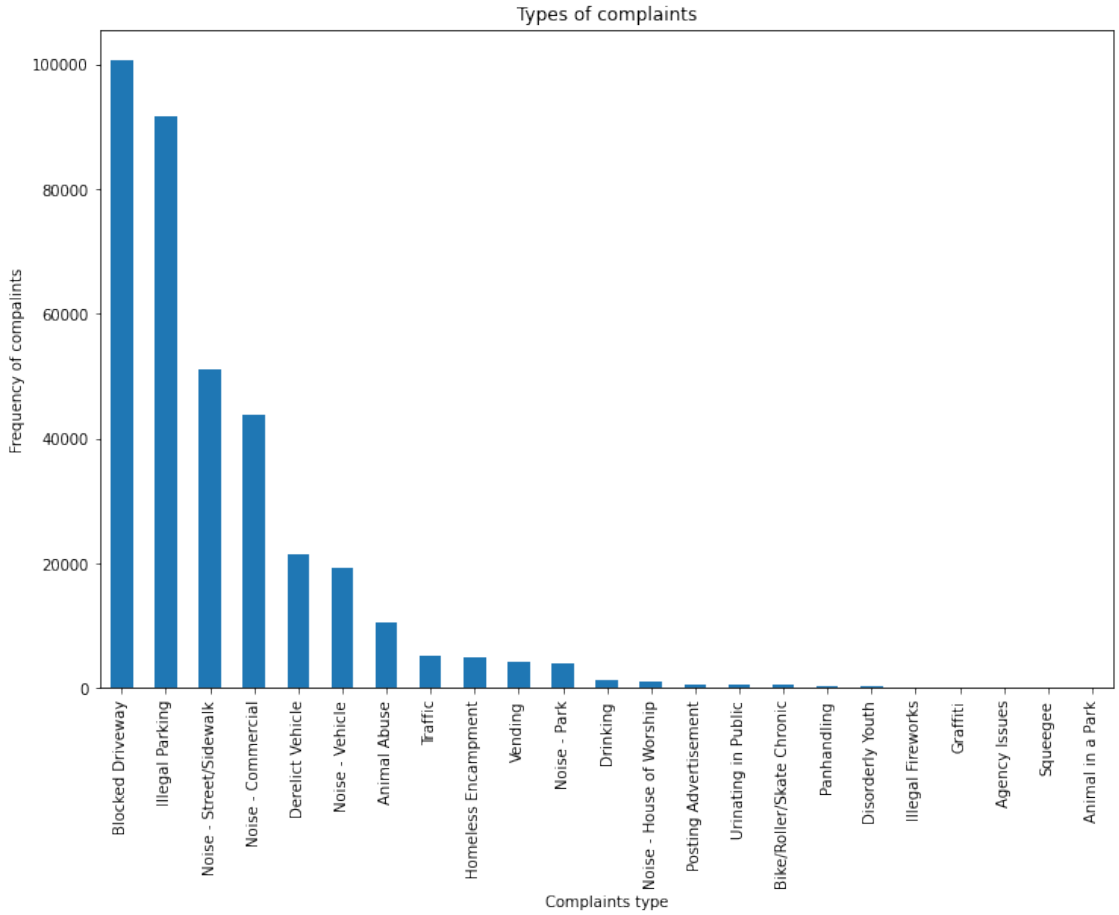


[26]: Text(0, 0.5, 'Longitude')



0.4 3. Finding majaor type of complaints

[27]:
```
## 3.1 Plot a bar graph to show the types of complaints
plt.figure(figsize= (12,8))
service_request_nyc['Complaint Type'].value_counts().plot(kind= 'bar')
plt.xlabel('Complaints type')
plt.ylabel('Frequency of compalints')
plt.title('Types of complaints')
plt.show();
```



[28]:
```
## 3.2Check the frequency of various types of complaints for New York City
service_request_nyc['City'].value_counts().iloc[1]
```

[28]:
```
77289
```

[29]:
```
service_request_nyc_newyork= service_request_nyc[service_request_nyc['City']==␣
↪'NEW YORK']
service_request_nyc_newyork['Complaint Type'].value_counts()
```

[29]:
```
Noise - Street/Sidewalk      22245
Noise - Commercial           18686
Illegal Parking              14549
Noise - Vehicle               6294
Homeless Encampment           3060
Blocked Driveway              2705
Vending                       2638
Animal Abuse                  1941
Traffic                       1769
```

```
Noise - Park                  1243
Derelict Vehicle               695
Drinking                       321
Urinating in Public            264
Bike/Roller/Skate Chronic      254
Noise - House of Worship       222
Panhandling                    206
Disorderly Youth                81
Posting Advertisement           49
Illegal Fireworks               38
Graffiti                        25
Squeegee                         4
Name: Complaint Type, dtype: int64
```

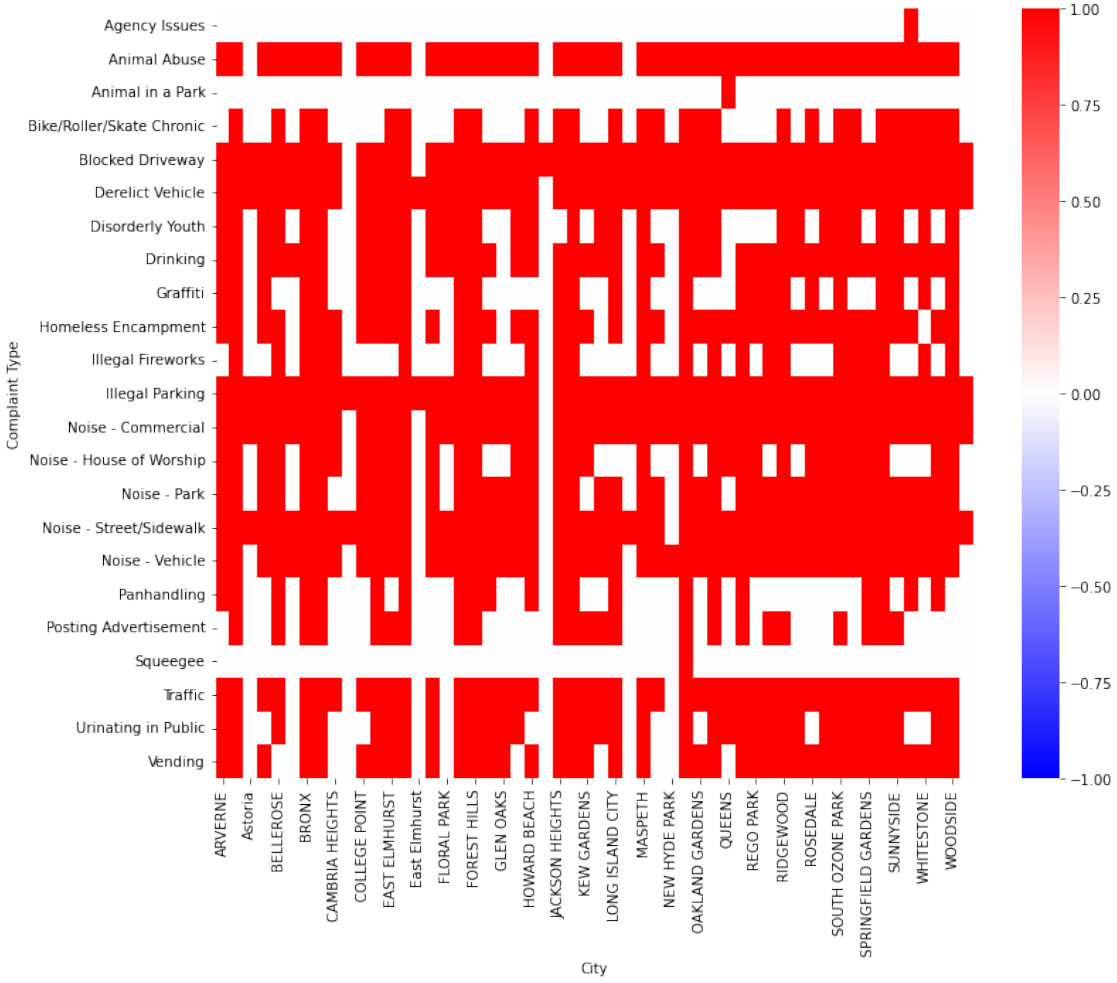
[30]:
```
## 3.3 Find the top 10 complaint types
service_request_nyc['Complaint Type'].value_counts().nlargest(10)
```

[30]:
```
Blocked Driveway           100624
Illegal Parking             91716
Noise - Street/Sidewalk     51139
Noise - Commercial          43751
Derelict Vehicle            21518
Noise - Vehicle             19301
Animal Abuse                10530
Traffic                      5196
Homeless Encampment          4879
Vending                      4185
Name: Complaint Type, dtype: int64
```

[31]:
```
## 3.4 Display the various types of complaints in each city
crosstab= pd.crosstab(index= service_request_nyc['Complaint Type'], columns=␣
 ↪service_request_nyc['City'])
```

[32]:
```
plt.figure(figsize=(12,10))
sns.heatmap(crosstab, cmap= 'bwr', vmin=-1, vmax=1, annot= False)
```

[32]:
```
<AxesSubplot:xlabel='City', ylabel='Complaint Type'>
```



[33]:
```
##  3.5 Create a DataFrame, df_new, which contains cities as columns and␣
↪complaint types in rows

df_new= pd.DataFrame(crosstab)
```

[34]:
```
df_new
```

[34]:

City	ARVERNE	ASTORIA	Astoria	BAYSIDE	BELLEROSE	\
Complaint Type						
Agency Issues	0	0	0	0	0	
Animal Abuse	46	170	0	53	15	
Animal in a Park	0	0	0	0	0	
Bike/Roller/Skate Chronic	0	16	0	0	1	
Blocked Driveway	50	3436	159	514	138	
Derelict Vehicle	32	426	14	231	120	
Disorderly Youth	2	5	0	2	2	

Drinking	1	43	0	1	1
Graffiti	1	4	0	3	0
Homeless Encampment	4	32	0	2	1
Illegal Fireworks	0	4	0	0	1
Illegal Parking	62	1340	277	638	132
Noise - Commercial	2	1653	310	47	38
Noise - House of Worship	14	21	0	3	1
Noise - Park	2	64	0	4	1
Noise - Street/Sidewalk	29	409	145	17	13
Noise - Vehicle	10	236	0	24	11
Panhandling	1	2	0	0	1
Posting Advertisement	0	3	0	0	1
Squeegee	0	0	0	0	0
Traffic	1	60	0	9	9
Urinating in Public	1	10	0	0	1
Vending	1	57	0	2	0

City	BREEZY POINT	BRONX	BROOKLYN	CAMBRIA HEIGHTS	\
Complaint Type					
Agency Issues	0	0	0	0	
Animal Abuse	2	1971	3191	15	
Animal in a Park	0	0	0	0	
Bike/Roller/Skate Chronic	0	22	124	0	
Blocked Driveway	3	17062	36445	177	
Derelict Vehicle	3	2402	6257	148	
Disorderly Youth	0	66	79	0	
Drinking	1	206	291	0	
Graffiti	0	15	60	0	
Homeless Encampment	0	275	948	6	
Illegal Fireworks	0	24	61	1	
Illegal Parking	16	9889	33532	113	
Noise - Commercial	4	2944	13855	19	
Noise - House of Worship	0	90	389	2	
Noise - Park	0	548	1575	0	
Noise - Street/Sidewalk	1	9144	13982	29	
Noise - Vehicle	1	3556	5965	100	
Panhandling	0	20	49	0	
Posting Advertisement	0	18	58	0	
Squeegee	0	0	0	0	
Traffic	0	427	1258	7	
Urinating in Public	0	54	155	0	
Vending	0	433	575	0	

City	CENTRAL PARK	…	SOUTH OZONE PARK	\
Complaint Type		…		
Agency Issues	0	…	0	
Animal Abuse	0	…	74	

Animal in a Park	0	…	0
Bike/Roller/Skate Chronic	0	…	1
Blocked Driveway	0	…	1202
Derelict Vehicle	0	…	425
Disorderly Youth	0	…	2
Drinking	0	…	14
Graffiti	0	…	2
Homeless Encampment	0	…	5
Illegal Fireworks	0	…	1
Illegal Parking	5	…	602
Noise - Commercial	0	…	82
Noise - House of Worship	0	…	5
Noise - Park	0	…	4
Noise - Street/Sidewalk	105	…	108
Noise - Vehicle	0	…	97
Panhandling	0	…	0
Posting Advertisement	0	…	1
Squeegee	0	…	0
Traffic	0	…	36
Urinating in Public	0	…	2
Vending	0	…	5

City	SOUTH RICHMOND HILL	SPRINGFIELD GARDENS	\
Complaint Type			
Agency Issues	0	0	
Animal Abuse	40	42	
Animal in a Park	0	0	
Bike/Roller/Skate Chronic	1	0	
Blocked Driveway	1946	330	
Derelict Vehicle	356	267	
Disorderly Youth	2	0	
Drinking	25	6	
Graffiti	0	0	
Homeless Encampment	12	7	
Illegal Fireworks	2	1	
Illegal Parking	596	291	
Noise - Commercial	223	38	
Noise - House of Worship	3	1	
Noise - Park	2	1	
Noise - Street/Sidewalk	93	42	
Noise - Vehicle	93	48	
Panhandling	0	2	
Posting Advertisement	0	2	
Squeegee	0	0	
Traffic	12	12	
Urinating in Public	1	3	
Vending	24	1	

City	STATEN ISLAND	SUNNYSIDE	Unknown City	WHITESTONE	\
Complaint Type					
Agency Issues	0	0	8	0	
Animal Abuse	786	40	1	43	
Animal in a Park	0	0	0	0	
Bike/Roller/Skate Chronic	10	2	2	4	
Blocked Driveway	2845	278	86	279	
Derelict Vehicle	2184	17	63	279	
Disorderly Youth	25	2	0	1	
Drinking	188	12	3	3	
Graffiti	6	1	0	1	
Homeless Encampment	77	12	1	0	
Illegal Fireworks	11	0	0	1	
Illegal Parking	6224	167	312	631	
Noise - Commercial	783	238	79	21	
Noise - House of Worship	18	0	0	0	
Noise - Park	67	16	7	7	
Noise - Street/Sidewalk	885	69	99	35	
Noise - Vehicle	424	53	9	31	
Panhandling	13	0	1	0	
Posting Advertisement	516	3	0	0	
Squeegee	0	0	0	0	
Traffic	229	17	2	32	
Urinating in Public	19	2	0	0	
Vending	25	15	1	1	

City	WOODHAVEN	WOODSIDE	Woodside
Complaint Type			
Agency Issues	0	0	0
Animal Abuse	57	111	0
Animal in a Park	0	0	0
Bike/Roller/Skate Chronic	2	5	0
Blocked Driveway	1363	2038	27
Derelict Vehicle	369	298	8
Disorderly Youth	0	1	0
Drinking	4	15	0
Graffiti	0	4	0
Homeless Encampment	10	38	0
Illegal Fireworks	0	1	0
Illegal Parking	896	1083	124
Noise - Commercial	209	256	2
Noise - House of Worship	3	4	0
Noise - Park	3	38	0
Noise - Street/Sidewalk	89	261	5
Noise - Vehicle	81	136	0
Panhandling	1	0	0

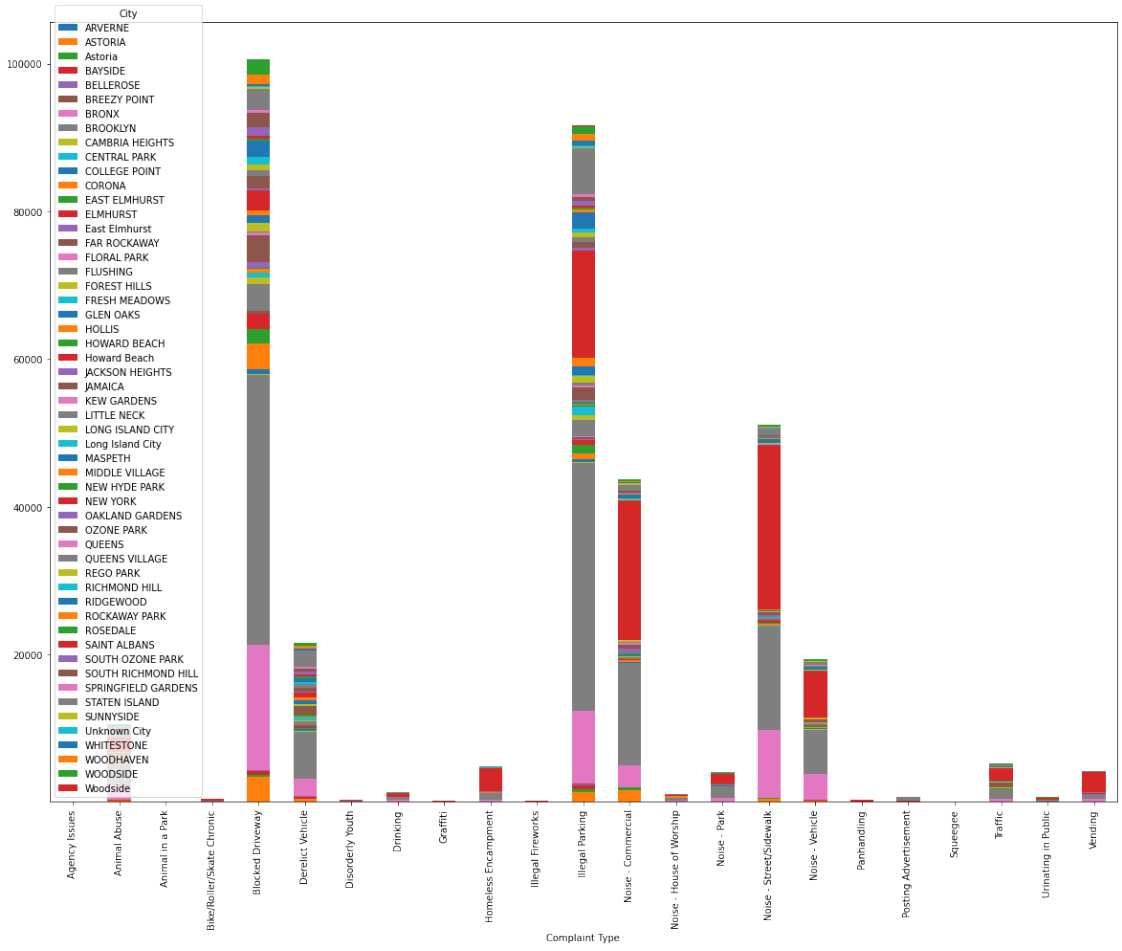
Posting Advertisement	0	0	0
Squeegee	0	0	0
Traffic	7	45	0
Urinating in Public	2	8	0
Vending	6	15	0

[23 rows x 54 columns]

0.5 4. Visualize the major types of complaints in each city

[35]:
```
## 4.1 Draw another chart that shows the types of complaints in each city in a␣
↪single chart,
##where different colors show the different types of complaints
crosstab.plot(kind= 'bar', stacked=True, figsize=(20,15))
```

[35]: <AxesSubplot:xlabel='Complaint Type'>



[37]:
```
## 4.2 Sort the complaint types based on the average Request_Closing_Time␣
↪grouping them for different
##locations
service_request_nyc['Elapsed_Time'] = service_request_nyc['Elapsed_Time'].
↪astype(int)

grouped_data= service_request_nyc.groupby(['City', 'Complaint␣
↪Type'])['Elapsed_Time'].mean()

grouped_data_sort= pd.DataFrame(grouped_data.sort_values())
grouped_data_sort
```

[37]:

City	Complaint Type	Elapsed_Time
ARVERNE	Drinking	8.590000e+11
RIDGEWOOD	Posting Advertisement	8.950000e+11
EAST ELMHURST	Bike/Roller/Skate Chronic	9.080000e+11
OZONE PARK	Illegal Fireworks	1.191000e+12
QUEENS	Urinating in Public	1.226000e+12
…		…
OAKLAND GARDENS	Homeless Encampment	1.031560e+14
SOUTH RICHMOND HILL	Bike/Roller/Skate Chronic	1.112770e+14
BELLEROSE	Homeless Encampment	1.409300e+14
QUEENS VILLAGE	Graffiti	1.919950e+14
QUEENS	Animal in a Park	1.212634e+15

[792 rows x 1 columns]

0.6 5. See whether the average response time across different complaint types is similar (overall)

[41]:
```
## 5.1 Visualize the average of Request_Closing_Time

grouped_data_compalint= service_request_nyc.groupby(['Complaint␣
↪Type'])['Elapsed_Time'].mean()
grouped_data_compalint_sort= pd.DataFrame(grouped_data_compalint.sort_values())
grouped_data_compalint_sort
```

[41]:

Complaint Type	Elapsed_Time
Posting Advertisement	7.286256e+12
Illegal Fireworks	1.011348e+13
Noise - Commercial	1.108576e+13
Noise - House of Worship	1.139109e+13
Noise - Park	1.222606e+13

```
Noise - Street/Sidewalk    1.223130e+13
Traffic                    1.230912e+13
Disorderly Youth           1.236375e+13
Noise - Vehicle            1.256180e+13
Urinating in Public        1.295929e+13
Bike/Roller/Skate Chronic  1.312369e+13
Drinking                   1.382130e+13
Vending                    1.436628e+13
Squeegee                   1.456025e+13
Homeless Encampment        1.545138e+13
Illegal Parking            1.565044e+13
Panhandling                1.585355e+13
Blocked Driveway           1.623252e+13
Animal Abuse               1.803256e+13
Agency Issues              1.828912e+13
Graffiti                   2.327634e+13
Derelict Vehicle           2.535960e+13
Animal in a Park           1.212634e+15
```

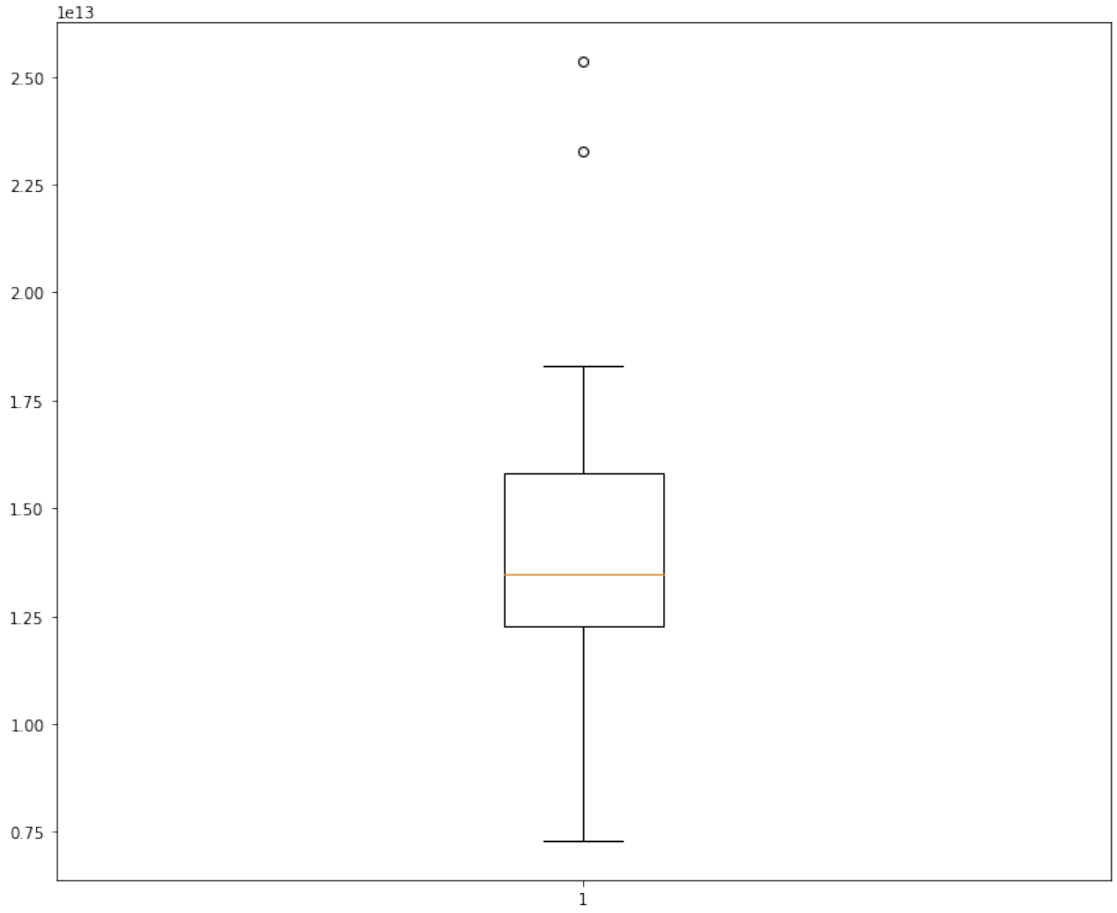
[42]:
```
grouped_data_compalint_sort.describe()
```

[42]:
```
       Elapsed_Time
count  2.300000e+01
mean   6.657300e+13
std    2.498643e+14
min    7.286256e+12
25%    1.227021e+13
50%    1.382130e+13
75%    1.604304e+13
max    1.212634e+15
```

[45]:
```
Elapsed_time_Range= grouped_data_compalint_sort.max()-␣
 ↪grouped_data_compalint_sort.min()
Elapsed_time_Range
```

[45]:
```
Elapsed_Time    1.205348e+15
dtype: float64
```

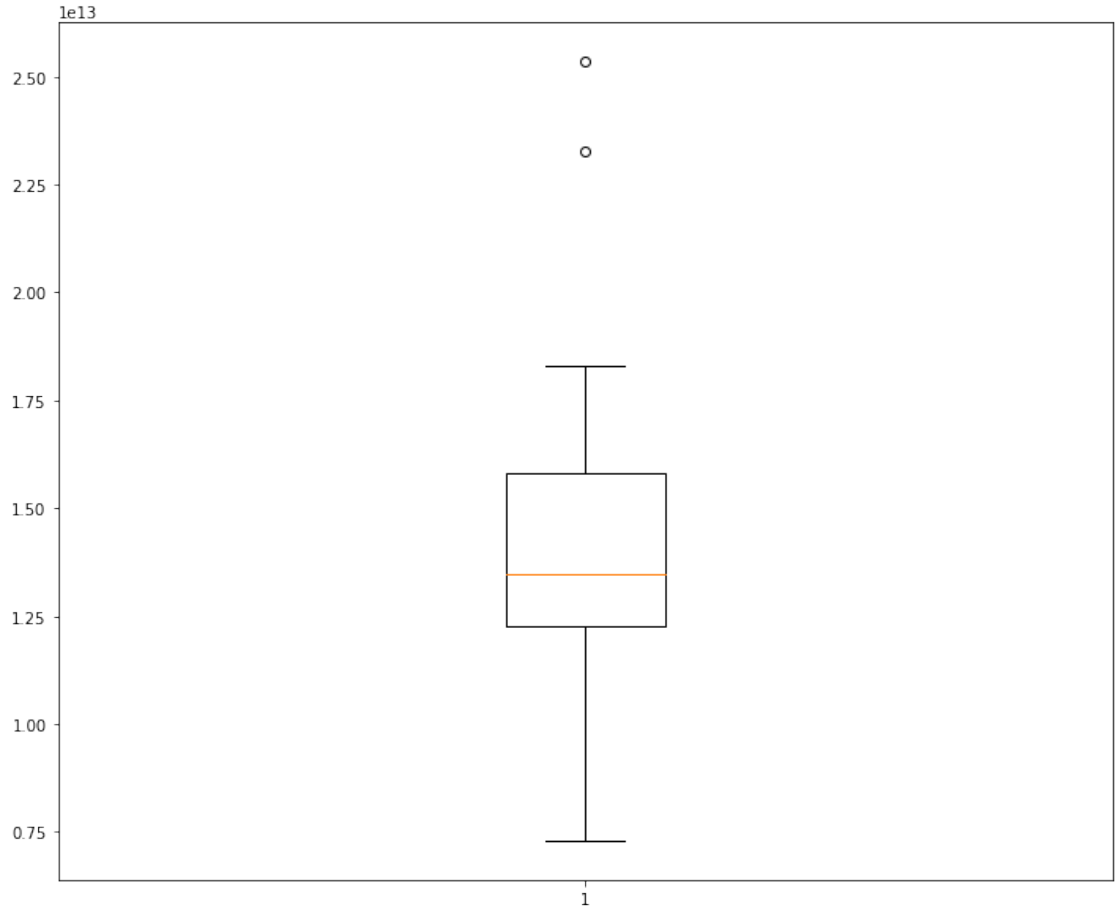
[52]:
```
plt.figure(figsize=(12,10))
plt.boxplot(grouped_data_compalint_sort);
```



[48]:
```
## Find the index of the maximum value
max_index= grouped_data_compalint_sort.idxmax()

# drop the row with maximum value
grouped_data_compalint_sort= grouped_data_compalint_sort.drop(max_index)
```

[50]:
```
plt.figure(figsize=(12,10))
plt.boxplot(grouped_data_compalint_sort);
```


1e13
2.50
2.25
2.00
1.75
1.50
1.25
1.00
0.75
1

[ ]: