appliedpython

March 10, 2023

[39]:
```
# 1. Understand the dataset:
# 1.1 Import the dataset
# 1.2 Visualize the dataset
# 1.3 Print the columns of the DataFrame
# 1.4 Identify the shape of the dataset
# 1.5 Identify the variables with null values
# 2. Perform basic data exploratory analysis:
# 2.1 Draw a frequency plot to show the number of null values in
# each column of the DataFrame
# 2.2 Missing value treatment
# 2.2.1 Remove the records whose Closed Date values are
# 2.3 Analyze the date column, and remove entries that have an
# incorrect timeline
# 2.3.1 Calculate the time elapsed in closed and creation date
# 2.3.2 Convert the calculated date to seconds to get a better
# representation
# 2.3.3 View the descriptive statistics for the newly created
# column
# 2.3.4 Check the number of null values in the Complaint_Type
# and City columns
# 2.3.5 Impute the NA value with Unknown City
# 2.3.6 Draw a frequency plot for the complaints in each city
# 2.3.7 Create a scatter and hexbin plot of the concentration of
# complaints across Brooklyn
# 3. Find major types of complaints:
# 3.1 Plot a bar graph to show the types of complaints
# 3.2 Check the frequency of various types of complaints for New
# York City
# 3.3 Find the top 10 complaint types
# 3.4 Display the various types of complaints in each city
# 3.5 Create a DataFrame, df_new, which contains cities as
# columns and complaint types in rows
# 4. Visualize the major types of complaints in each city
# 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
# 4.2 Sort the complaint types based on the average
```

```
# Request_Closing_Time grouping them for different
# locations
# 5. See whether the average response time across different
# complaint types is similar (overall)
# 5.1 Visualize the average of Request_Closing_Time
# 6. Identify the significant variables by performing statistical
# analysis using p-values
# 7. Perform a Kruskal-Wallis H test
# 7.1 Fail to reject H0: All sample distributions are equal
# 7.2 Reject H0: One or more sample distributions are not
# equal
# 8. Present your observations
# null
```

[1]:
```
import pandas as pd
import seaborn as sns
import matplotlib.pyplot as plt
import missingno as msno
data_filename = 'pythonproject1/311_Service_Requests_from_2010_to_Present.csv'
df = pd.read_csv(data_filename)
df
```

```
/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)
```

[1]:

	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]
```

[2]:
```
df.columns
```

[2]:
```
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')
```

[3]:
```
print(df.shape)
```

```
(364558, 53)
```

[4]:
```
df.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
```

[5]:
```
null_cols = df.columns[df.isnull().any()]
print(null_cols)
```

```
Index(['Closed Date', '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', 'Due Date',
       'Resolution Action Updated Date', 'X Coordinate (State Plane)',
       'Y Coordinate (State Plane)', 'School Region', 'School Code',
       'School Zip', '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')
```

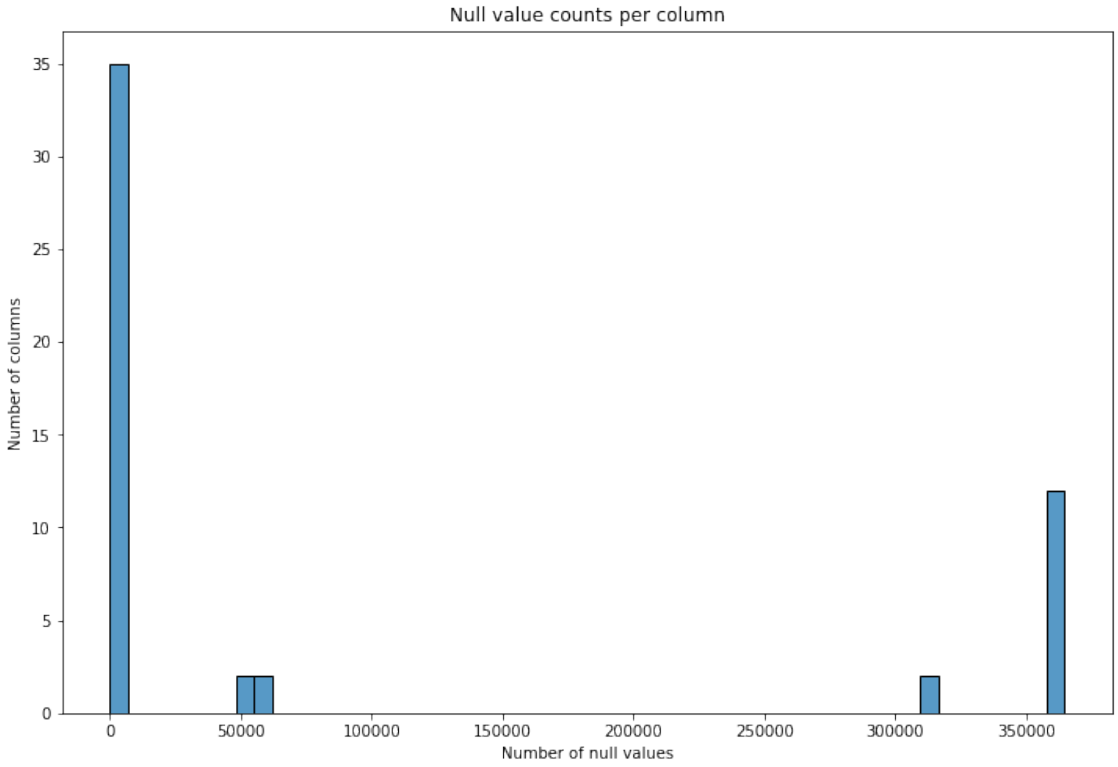
[45]:
```
# Count the number of null values in each column
null_counts = df.isnull().sum()
null_counts
```

[45]:
```
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
```

[46]:
```
# Set figure size
plt.figure(figsize=(12, 8))
```

```
# Plot a bar chart of null value counts
sns.histplot(df.isnull().sum(), bins=df.shape[1])

# Set x-axis label
plt.xlabel('Number of null values')

# Set y-axis label
plt.ylabel('Number of columns')

# Set plot title
plt.title('Null value counts per column')

# Show the plot
plt.show()
```



[47]:
```
# remove the rows where the 'Closed Date' value is null
df = df.dropna(subset=['Closed Date'])

# show the updated dataframe
print(df)
```

```
      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)

[362177 rows x 53 columns]
```

[6]:
```
# convert the Closed Date and Created Date columns to pandas datetime objects
df['Closed Date'] = pd.to_datetime(df['Closed Date'])
df['Created Date'] = pd.to_datetime(df['Created Date'])
```

```
# calculate the time elapsed between the Closed Date and Created Date columns
df['Time Elapsed'] = (df['Closed Date'] - df['Created Date']).dt.total_seconds()

# show the updated dataframe
print(df)
```

```
        Unique Key        Created Date         Closed Date Agency  \
0         32310363 2015-12-31 23:59:45 2016-01-01 00:55:15   NYPD
1         32309934 2015-12-31 23:59:44 2016-01-01 01:26:57   NYPD
2         32309159 2015-12-31 23:59:29 2016-01-01 04:51:03   NYPD
3         32305098 2015-12-31 23:57:46 2016-01-01 07:43:13   NYPD
4         32306529 2015-12-31 23:56:58 2016-01-01 03:24:42   NYPD
…              …                 …                   …    …
364553    29609918 2015-01-01 00:04:44 2015-01-01 10:22:31   NYPD
364554    29608392 2015-01-01 00:04:28 2015-01-01 02:25:02   NYPD
364555    29607589 2015-01-01 00:01:30 2015-01-01 00:20:33   NYPD
364556    29610889 2015-01-01 00:01:29 2015-01-01 02:42:22   NYPD
364557    29611816 2015-01-01 00:00:50 2015-01-01 02:47:50   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 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
4               87-14 57 ROAD  …                      NaN       NaN
…                         …  …                      …         …
364553          84-25 85 ROAD  …                      NaN       NaN
364554   2555 SEDGWICK AVENUE  …                      NaN       NaN
364555    508 WEST 139 STREET  …                      NaN       NaN
364556    931 EAST 226 STREET  …                      NaN       NaN
364557      123-19 135 STREET  …                      NaN       NaN

       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
…                         …               …               …
364553                    NaN             NaN             NaN
364554                    NaN             NaN             NaN
364555                    NaN             NaN             NaN
364556                    NaN             NaN             NaN
364557                    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
…                      …         …         …
364553                 NaN  40.695145 -73.860949
364554                 NaN  40.867830 -73.907178
364555                 NaN  40.821647 -73.950873
364556                 NaN  40.886361 -73.853290
364557                 NaN  40.674212 -73.803585

                                        Location Time Elapsed
0        (40.86568153633767, -73.92350095571744)       3330.0
1       (40.775945312321085, -73.91509393898605)       5233.0
2       (40.870324522111424, -73.88852464418646)      17494.0
3        (40.83599404683083, -73.82837939584206)      27927.0
4       (40.733059618956815, -73.87416975810375)      12464.0
…                                             …           …
364553   (40.69514470265117, -73.86094888534394)      37067.0
364554   (40.86782963689454, -73.90717786644662)       8434.0
364555  (40.821646626438095, -73.95087342885292)       1143.0
```

```
364556   (40.88636077906953, -73.85329048666742)       9653.0
364557  (40.674211762243935, -73.80358548685278)      10020.0

[364558 rows x 54 columns]
```

[49]:
```
# view the descriptive statistics for a column
print(df['Time Elapsed'].describe())
```

```
count    3.621770e+05
mean     1.511330e+04
std      2.110255e+04
min      6.100000e+01
25%      4.533000e+03
50%      9.616000e+03
75%      1.887800e+04
max      2.134342e+06
Name: Time Elapsed, dtype: float64
```

[50]:
```
# # count the number of null values in the Complaint_Type column
print('Number of null values in Complaint_Type column:', df['Complaint Type'].
 ↪isnull().sum())

# count the number of null values in the City column
print('Number of null values in City column:', df['City'].isnull().sum())
```

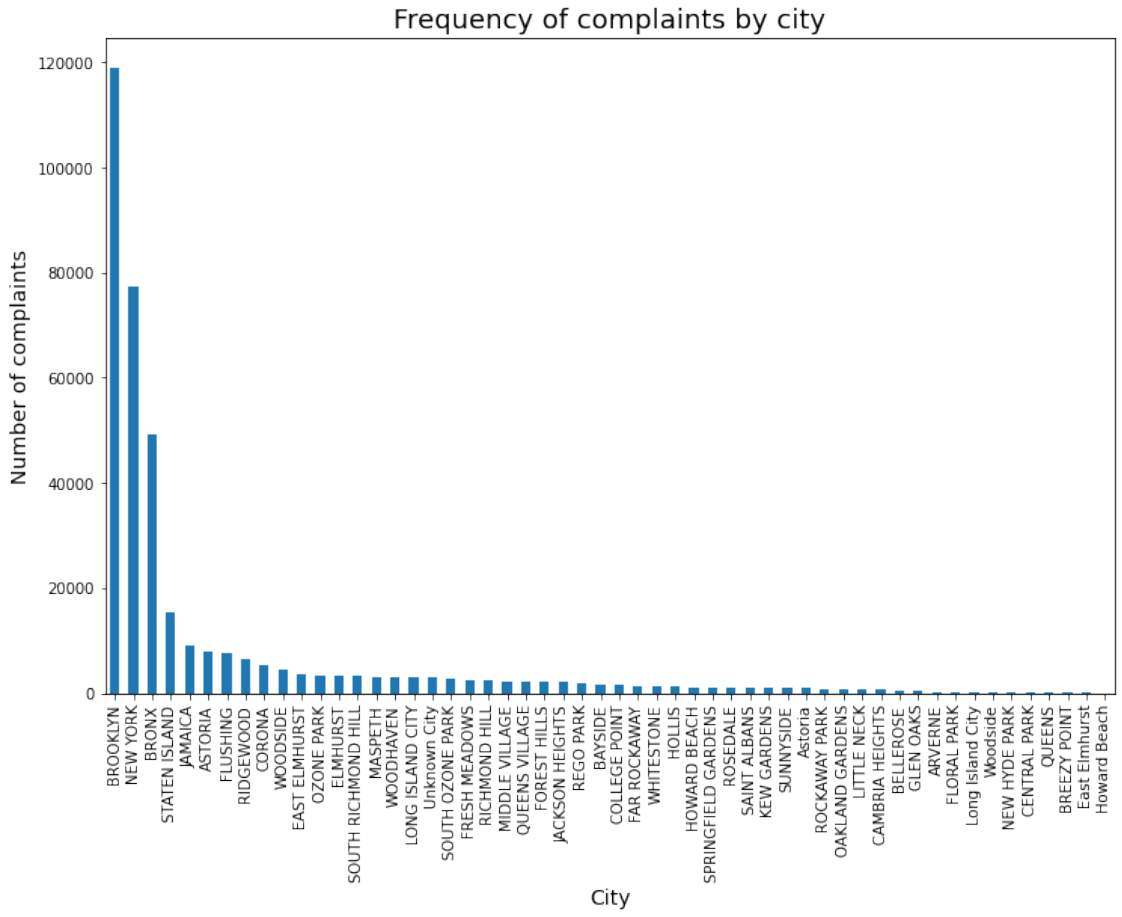
```
Number of null values in Complaint_Type column: 0
Number of null values in City column: 674
```

[7]:
```
#to Impute the NA value with Unknown City
df['City'] = df['City'].fillna('Unknown City')

# Save the updated dataset
df.to_csv('updated_dataset.csv', index=False)
single_column = df['City']

# Print the column
print(single_column)
```

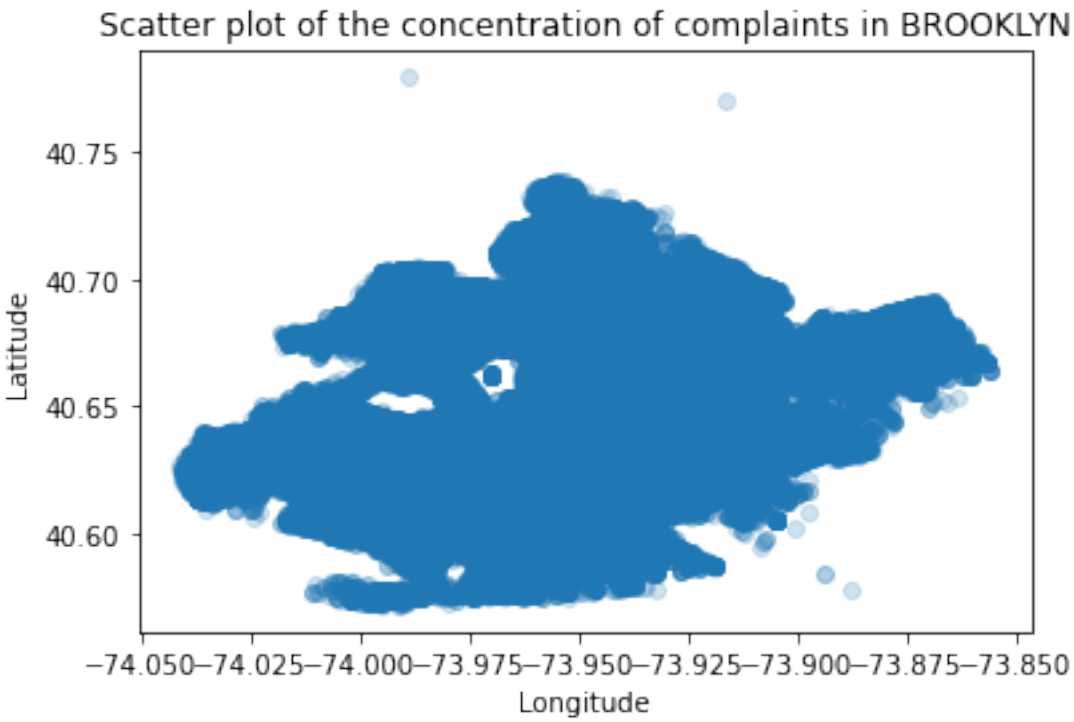
```
0                 NEW YORK
1                  ASTORIA
2                    BRONX
3                    BRONX
4                  ELMHURST
                   …
364553            WOODHAVEN
364554                BRONX
364555             NEW YORK
364556                BRONX
```

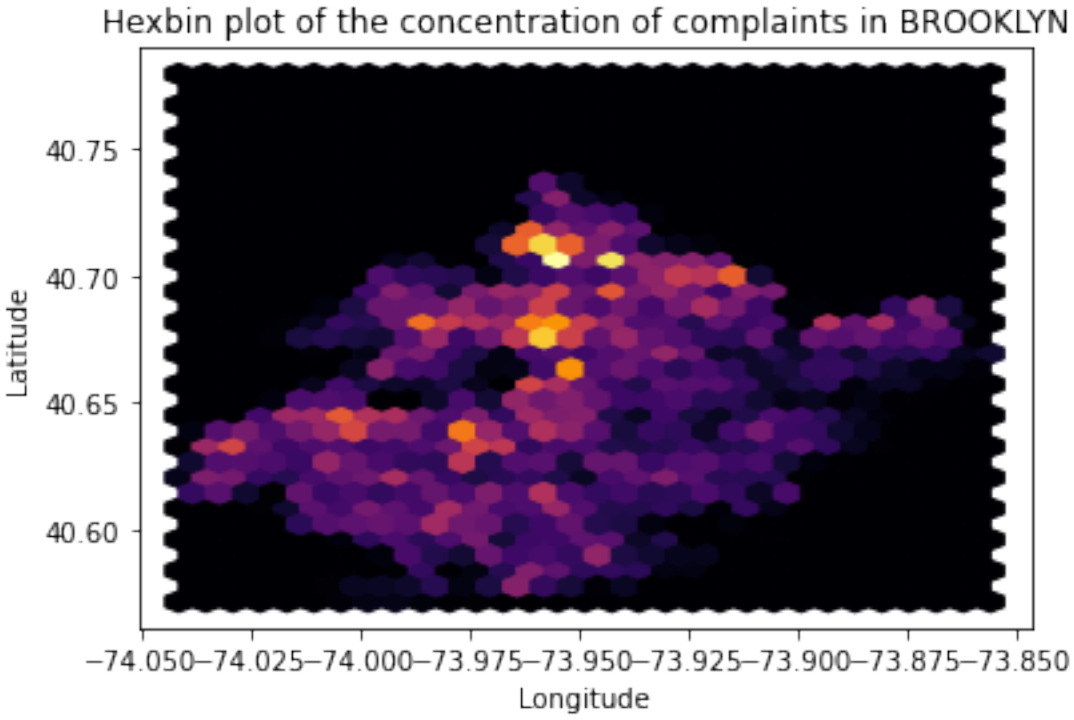
```
364557    SOUTH OZONE PARK
Name: City, Length: 364558, dtype: object
```

[17]:
```
# Count the number of complaints in each city
complaints_by_city = df['City'].value_counts()

# Set the figure size to 12x8 inches
fig, ax = plt.subplots(figsize=(12,8))

# Plot the frequency of complaints by city
complaints_by_city.plot(kind='bar', ax=ax)

# Set the plot title and axis labels
ax.set_title('Frequency of complaints by city', fontsize=18)
ax.set_xlabel('City', fontsize=14)
ax.set_ylabel('Number of complaints', fontsize=14)

# Show the plot
plt.show()
```

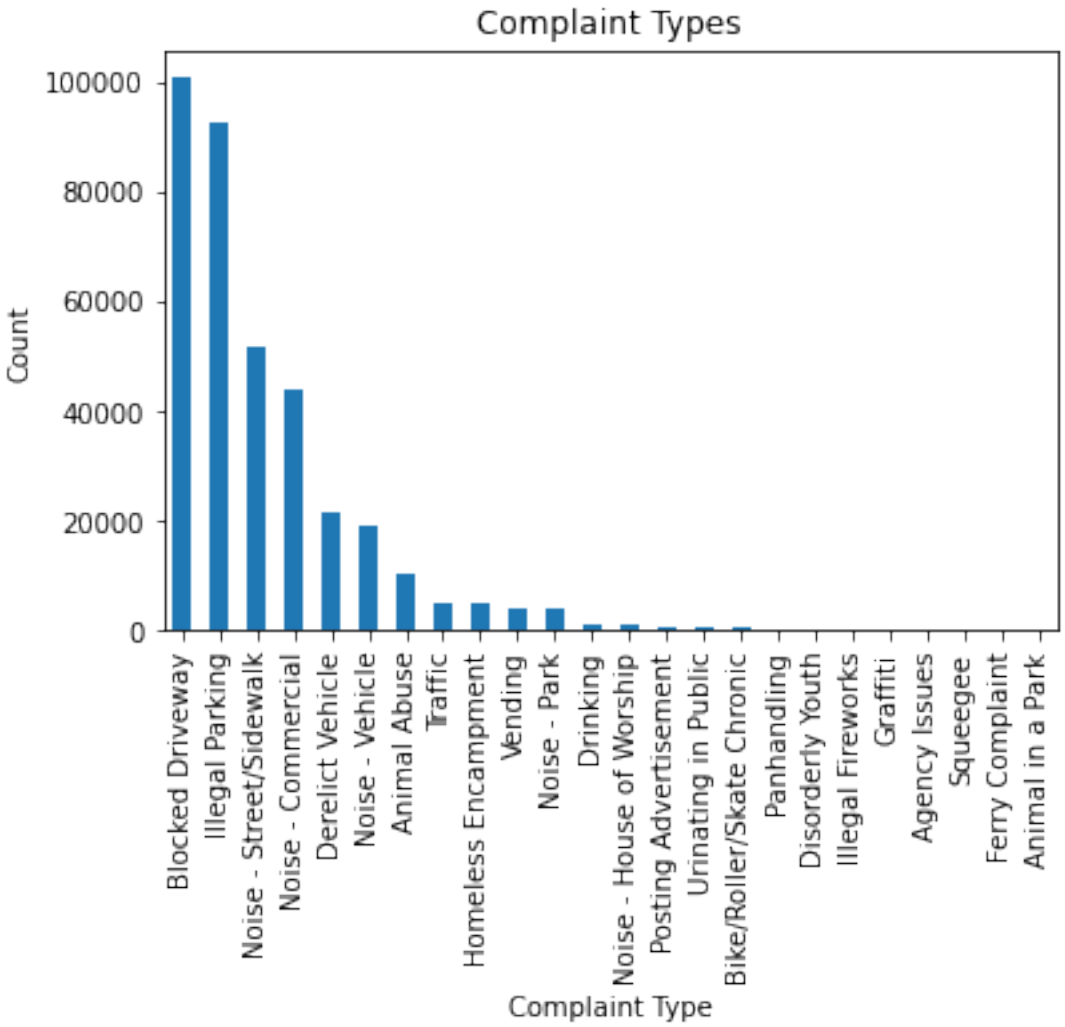


[19]:
```
# Filter the dataset to only include BROOKLYN
brooklyn_df = df[df['City'].str.upper() == 'BROOKLYN']

# Create a scatter plot of the concentration of complaints in BROOKLYN
plt.scatter(brooklyn_df['Longitude'], brooklyn_df['Latitude'], alpha=0.2)



# Set the plot title and axis labels
plt.title('Scatter plot of the concentration of complaints in BROOKLYN')
plt.xlabel('Longitude')
plt.ylabel('Latitude')

# Show the plot
plt.show()

# Create a hexbin plot of the concentration of complaints in BROOKLYN
plt.hexbin(brooklyn_df['Longitude'], brooklyn_df['Latitude'], gridsize=30,␣
 ↪cmap='inferno')

# Set the plot title and axis labels
plt.title('Hexbin plot of the concentration of complaints in BROOKLYN')
plt.xlabel('Longitude')
plt.ylabel('Latitude')

# Show the plot
plt.show()
```


Scatter plot of the concentration of complaints in BROOKLYN
40.75
40.70
40.65
40.60
Latitude
−74.050
−74.025
−74.000
−73.975
−73.950
−73.925
−73.900
−73.875
−73.850
Longitude


Hexbin plot of the concentration of complaints in BROOKLYN
40.75
40.70
40.65
40.60
Latitude
−74.050
−74.025
−74.000
−73.975
−73.950
−73.925
−73.900
−73.875
−73.850
Longitude

[20]:
```
# Count the occurrences of each complaint type
complaint_counts = df['Complaint Type'].value_counts()

# Plot a bar graph of the complaint types
complaint_counts.plot(kind='bar')

# Set the plot title and axis labels
plt.title('Complaint Types')
plt.xlabel('Complaint Type')
plt.ylabel('Count')

# Show the plot
plt.show()
```



[21]:
```
# Filter the dataset to only include New York City
nyc_df = df[df['City'] == 'NEW YORK']

# Count the occurrences of each complaint type
complaint_counts = nyc_df['Complaint Type'].value_counts()

# Print the frequency of each complaint type
print(complaint_counts)
```

```
Noise - Street/Sidewalk      22252
Noise - Commercial           18692
Illegal Parking              14553
Noise - Vehicle               6295
Homeless Encampment           3060
Blocked Driveway              2707
Vending                       2639
Animal Abuse                  1941
Traffic                       1770
Noise - Park                  1244
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
```

[52]:
```
# Count the occurrences of each complaint type and get the top 10
top_10_complaints = df['Complaint Type'].value_counts().head(10)

# Print the top 10 complaint types
print(top_10_complaints)
```

```
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
```

[23]:
```
# Group the data by City and Complaint Type and count the occurrences
complaint_counts = df.groupby(['City', 'Complaint Type']).size()

# Print the result
print(complaint_counts)
```

```
City      Complaint Type
ARVERNE   Animal Abuse                46
          Blocked Driveway            50
          Derelict Vehicle            32
          Disorderly Youth             2
          Drinking                     1
                                      ..
Woodside  Blocked Driveway            27
          Derelict Vehicle             8
          Illegal Parking            124
          Noise - Commercial           2
          Noise - Street/Sidewalk      5
Length: 795, dtype: int64
```

[24]:
```
# Create a pivot table with cities as columns and complaint types as rows
df_new = pd.pivot_table(df, index='Complaint Type', columns='City', aggfunc=len)

# Print the result
print(df_new)
```

```
                          Address Type                                    \
City                           ARVERNE ASTORIA Astoria BAYSIDE BELLEROSE
Complaint Type
Agency Issues                      NaN     NaN     NaN     NaN       NaN
Animal Abuse                      46.0   170.0     NaN    53.0      15.0
Animal in a Park                   NaN     NaN     NaN     NaN       NaN
Bike/Roller/Skate Chronic          NaN    16.0     NaN     NaN       1.0
Blocked Driveway                  50.0  3436.0   159.0   514.0     138.0
Derelict Vehicle                  32.0   426.0    14.0   231.0     120.0
Disorderly Youth                   2.0     5.0     NaN     2.0       2.0
Drinking                           1.0    43.0     NaN     1.0       1.0
Ferry Complaint                    NaN     NaN     NaN     NaN       NaN
Graffiti                           1.0     4.0     NaN     3.0       NaN
Homeless Encampment                4.0    32.0     NaN     2.0       1.0
Illegal Fireworks                  NaN     4.0     NaN     NaN       1.0
Illegal Parking                   62.0  1340.0   277.0   638.0     132.0
Noise - Commercial                 2.0  1653.0   311.0    47.0      38.0
Noise - House of Worship          14.0    21.0     NaN     3.0       1.0
Noise - Park                       2.0    64.0     NaN     4.0       1.0
Noise - Street/Sidewalk           29.0   409.0   145.0    17.0      13.0
```

Noise - Vehicle	10.0	236.0	NaN	24.0	11.0
Panhandling	1.0	2.0	NaN	NaN	1.0
Posting Advertisement	NaN	3.0	NaN	NaN	1.0
Squeegee	NaN	NaN	NaN	NaN	NaN
Traffic	1.0	60.0	NaN	9.0	9.0
Urinating in Public	1.0	10.0	NaN	NaN	1.0
Vending	1.0	57.0	NaN	2.0	NaN

City	BREEZY POINT	BRONX	BROOKLYN	CAMBRIA HEIGHTS
Complaint Type				
Agency Issues	NaN	NaN	NaN	NaN
Animal Abuse	2.0	1971.0	3191.0	15.0
Animal in a Park	NaN	NaN	NaN	NaN
Bike/Roller/Skate Chronic	NaN	22.0	124.0	NaN
Blocked Driveway	3.0	17063.0	36447.0	177.0
Derelict Vehicle	3.0	2403.0	6259.0	148.0
Disorderly Youth	NaN	66.0	79.0	NaN
Drinking	1.0	206.0	291.0	NaN
Ferry Complaint	NaN	NaN	NaN	NaN
Graffiti	NaN	15.0	60.0	NaN
Homeless Encampment	NaN	275.0	948.0	6.0
Illegal Fireworks	NaN	24.0	61.0	1.0
Illegal Parking	16.0	9889.0	33533.0	113.0
Noise - Commercial	4.0	2945.0	13860.0	19.0
Noise - House of Worship	NaN	90.0	389.0	2.0
Noise - Park	NaN	548.0	1575.0	NaN
Noise - Street/Sidewalk	1.0	9146.0	13984.0	29.0
Noise - Vehicle	1.0	3556.0	5966.0	100.0
Panhandling	NaN	20.0	49.0	NaN
Posting Advertisement	NaN	18.0	58.0	NaN
Squeegee	NaN	NaN	NaN	NaN
Traffic	NaN	427.0	1258.0	7.0
Urinating in Public	NaN	54.0	155.0	NaN
Vending	NaN	433.0	575.0	NaN

City	CENTRAL PARK	…	Y Coordinate (State Plane) SOUTH OZONE PARK
Complaint Type		…	
Agency Issues	NaN	…	NaN
Animal Abuse	NaN	…	74.0
Animal in a Park	NaN	…	NaN
Bike/Roller/Skate Chronic	NaN	…	1.0
Blocked Driveway	NaN	…	1202.0
Derelict Vehicle	NaN	…	425.0
Disorderly Youth	NaN	…	2.0
Drinking	NaN	…	14.0
Ferry Complaint	NaN	…	NaN

Graffiti	NaN	…	2.0
Homeless Encampment	NaN	…	5.0
Illegal Fireworks	NaN	…	1.0
Illegal Parking	5.0	…	602.0
Noise - Commercial	NaN	…	82.0
Noise - House of Worship	NaN	…	5.0
Noise - Park	NaN	…	4.0
Noise - Street/Sidewalk	105.0	…	108.0
Noise - Vehicle	NaN	…	97.0
Panhandling	NaN	…	NaN
Posting Advertisement	NaN	…	1.0
Squeegee	NaN	…	NaN
Traffic	NaN	…	36.0
Urinating in Public	NaN	…	2.0
Vending	NaN	…	5.0

City	SOUTH RICHMOND HILL	SPRINGFIELD GARDENS
Complaint Type		
Agency Issues	NaN	NaN
Animal Abuse	40.0	42.0
Animal in a Park	NaN	NaN
Bike/Roller/Skate Chronic	1.0	NaN
Blocked Driveway	1946.0	330.0
Derelict Vehicle	356.0	267.0
Disorderly Youth	2.0	NaN
Drinking	25.0	6.0
Ferry Complaint	NaN	NaN
Graffiti	NaN	NaN
Homeless Encampment	12.0	7.0
Illegal Fireworks	2.0	1.0
Illegal Parking	596.0	291.0
Noise - Commercial	223.0	38.0
Noise - House of Worship	3.0	1.0
Noise - Park	2.0	1.0
Noise - Street/Sidewalk	93.0	42.0
Noise - Vehicle	93.0	48.0
Panhandling	NaN	2.0
Posting Advertisement	NaN	2.0
Squeegee	NaN	NaN
Traffic	12.0	12.0
Urinating in Public	1.0	3.0
Vending	24.0	1.0

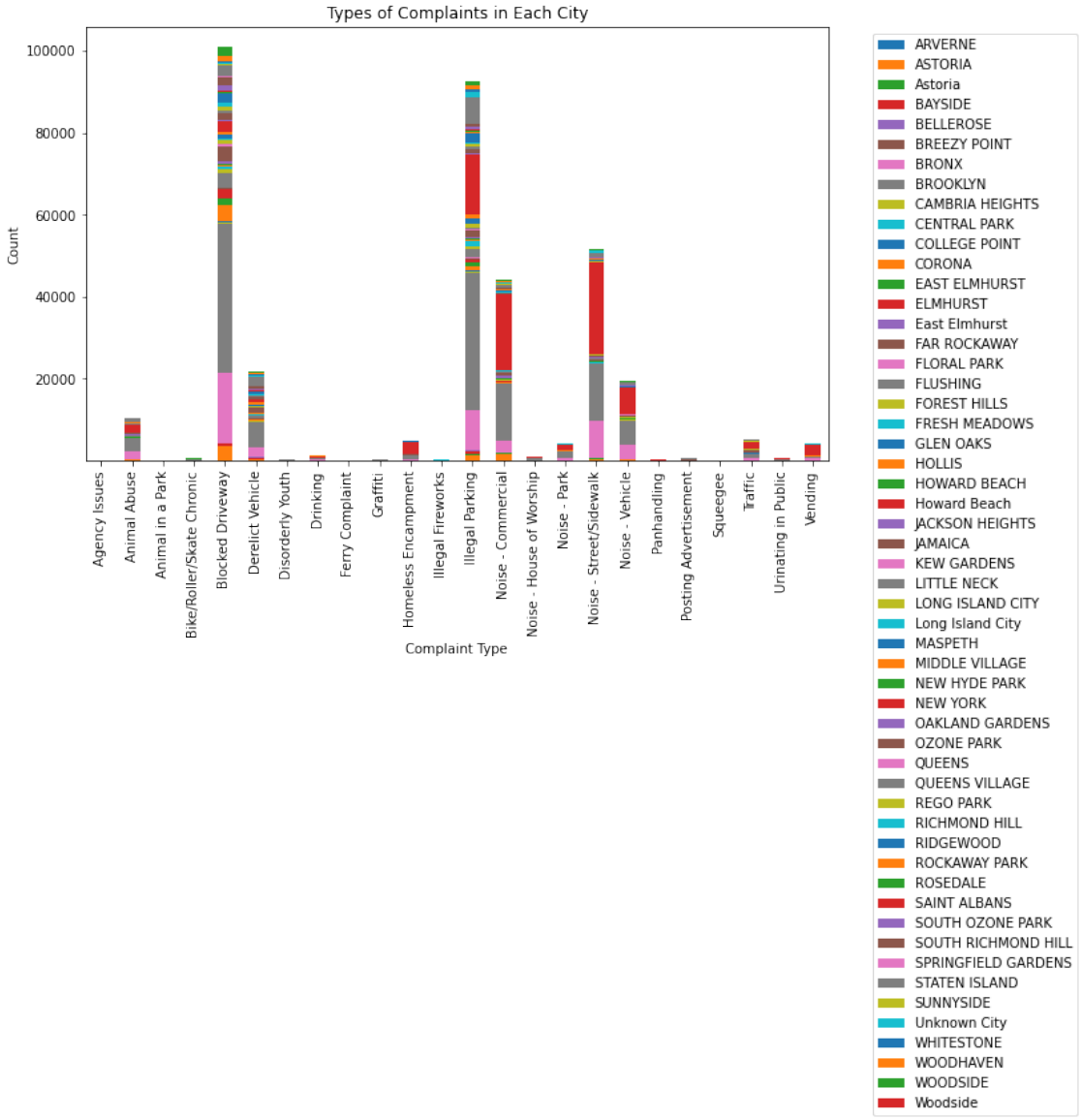
City	STATEN ISLAND	SUNNYSIDE	Unknown City	WHITESTONE
Complaint Type				
Agency Issues	NaN	NaN	8.0	NaN

Animal Abuse	786.0	40.0	12.0	43.0
Animal in a Park	NaN	NaN	NaN	NaN
Bike/Roller/Skate Chronic	10.0	2.0	5.0	4.0
Blocked Driveway	2845.0	278.0	333.0	279.0
Derelict Vehicle	2184.0	17.0	201.0	279.0
Disorderly Youth	25.0	2.0	NaN	1.0
Drinking	188.0	12.0	8.0	3.0
Ferry Complaint	NaN	NaN	2.0	NaN
Graffiti	6.0	1.0	NaN	1.0
Homeless Encampment	77.0	12.0	1.0	NaN
Illegal Fireworks	11.0	NaN	NaN	1.0
Illegal Parking	6224.0	167.0	1267.0	631.0
Noise - Commercial	784.0	238.0	422.0	21.0
Noise - House of Worship	18.0	NaN	2.0	NaN
Noise - Park	67.0	16.0	26.0	7.0
Noise - Street/Sidewalk	888.0	69.0	638.0	35.0
Noise - Vehicle	424.0	53.0	58.0	31.0
Panhandling	13.0	NaN	3.0	NaN
Posting Advertisement	517.0	3.0	1.0	NaN
Squeegee	NaN	NaN	NaN	NaN
Traffic	229.0	17.0	3.0	32.0
Urinating in Public	19.0	2.0	NaN	NaN
Vending	25.0	15.0	7.0	1.0

City	WOODHAVEN	WOODSIDE	Woodside
Complaint Type			
Agency Issues	NaN	NaN	NaN
Animal Abuse	57.0	111.0	NaN
Animal in a Park	NaN	NaN	NaN
Bike/Roller/Skate Chronic	2.0	5.0	NaN
Blocked Driveway	1364.0	2038.0	27.0
Derelict Vehicle	369.0	298.0	8.0
Disorderly Youth	NaN	1.0	NaN
Drinking	4.0	15.0	NaN
Ferry Complaint	NaN	NaN	NaN
Graffiti	NaN	4.0	NaN
Homeless Encampment	10.0	38.0	NaN
Illegal Fireworks	NaN	1.0	NaN
Illegal Parking	896.0	1083.0	124.0
Noise - Commercial	209.0	256.0	2.0
Noise - House of Worship	3.0	4.0	NaN
Noise - Park	3.0	38.0	NaN
Noise - Street/Sidewalk	89.0	261.0	5.0
Noise - Vehicle	81.0	136.0	NaN
Panhandling	1.0	NaN	NaN
Posting Advertisement	NaN	NaN	NaN
Squeegee	NaN	NaN	NaN

```
Traffic                          7.0     45.0      NaN
Urinating in Public              2.0      8.0      NaN
Vending                          6.0     15.0      NaN

[24 rows x 2808 columns]
```

[9]:
```python
# Group the data by City and Complaint Type and count the occurrences
complaint_counts = df.groupby(['City', 'Complaint Type']).size()

# Reshape the data into a pivot table
complaint_counts = complaint_counts.unstack(level=0)

# Create a stacked bar chart with a larger figure size
ax = complaint_counts.plot(kind='bar', stacked=True, figsize=(10, 6))

# Set the plot title and axis labels
ax.set_title('Types of Complaints in Each City')
ax.set_xlabel('Complaint Type')
ax.set_ylabel('Count')

# Move the legend outside the plot
ax.legend(bbox_to_anchor=(1.05, 1), loc='upper left')

# Show the plot
plt.show()
```



[58]:
```
# Convert the 'Created Date' and 'Closed Date' columns to datetime format
df['Created Date'] = pd.to_datetime(df['Created Date'])
df['Closed Date'] = pd.to_datetime(df['Closed Date'])

# Calculate the time difference in hours
df['Request_Closing_Time'] = (df['Closed Date'] - df['Created Date']).dt.
 ↪total_seconds() / 3600

# Group the data by 'City' and 'Complaint Type' and calculate the mean␣
 ↪'Request_Closing_Time'
complaint_times = df.groupby(['City', 'Complaint␣
 ↪Type'])['Request_Closing_Time'].mean()
```

```
# Reshape the data into a pivot table
complaint_times = complaint_times.unstack(level=0)

# Sort the columns of the pivot table based on the mean 'Request_Closing_Time'
complaint_times = complaint_times[complaint_times.mean().sort_values().index]

# Display the sorted pivot table
print(complaint_times)
```

```
City                          ARVERNE  ROCKAWAY PARK  FAR ROCKAWAY  LITTLE NECK  \
Complaint Type
Agency Issues                     NaN            NaN           NaN          NaN
Animal Abuse                2.333110       2.376852      2.553488     2.712500
Animal in a Park                  NaN            NaN           NaN          NaN
Bike/Roller/Skate Chronic         NaN            NaN           NaN          NaN
Blocked Driveway            2.310789       2.383299      2.537480     2.487754
Derelict Vehicle            3.165000       2.949284      3.614658     3.670510
Disorderly Youth            3.591250       3.426806      3.365833     4.352917
Drinking                    0.238611       2.330785      2.420625     1.582500
Ferry Complaint                   NaN            NaN           NaN          NaN
Graffiti                    1.530000            NaN           NaN          NaN
Homeless Encampment         1.817014       1.715069      2.720694          NaN
Illegal Fireworks                 NaN            NaN           NaN          NaN
Illegal Parking             2.335022       2.620457      2.560960     2.742772
Noise - Commercial          2.287222       2.041902      1.786634     1.339859
Noise - House of Worship    1.848175            NaN      1.147778          NaN
Noise - Park                1.288333       1.028750      1.502174     1.073194
Noise - Street/Sidewalk     1.992557       1.870510      3.006875     2.066222
Noise - Vehicle             1.576000       2.162567      2.169428     1.861667
Panhandling                 1.020278            NaN           NaN          NaN
Posting Advertisement             NaN            NaN           NaN     2.230556
Squeegee                          NaN            NaN           NaN          NaN
Traffic                     1.115000       3.560675      2.376919     2.114458
Urinating in Public         0.691944       0.720833      1.505556     3.075833
Vending                     0.481944       2.944444      2.613778          NaN

City                          BAYSIDE  ELMHURST  BREEZY POINT    CORONA  \
Complaint Type
Agency Issues                     NaN       NaN           NaN       NaN
Animal Abuse                3.112096  3.481106      2.613472  3.342642
Animal in a Park                  NaN       NaN           NaN       NaN
Bike/Roller/Skate Chronic         NaN  4.622778           NaN       NaN
Blocked Driveway            2.596096  3.297199      1.326481  3.210154
Derelict Vehicle            3.406008  4.814406      7.148056  4.313410
Disorderly Youth            2.594028  0.863889           NaN  2.879769
Drinking                    1.902222  3.063184      2.637500  3.576904
```

Ferry Complaint	NaN	NaN	NaN	NaN
Graffiti	4.550370	1.132500	NaN	2.593542
Homeless Encampment	2.872917	3.553309	NaN	3.535011
Illegal Fireworks	NaN	0.975000	NaN	NaN
Illegal Parking	2.519456	3.240384	3.765122	3.281186
Noise - Commercial	2.302512	2.965150	2.542083	2.612403
Noise - House of Worship	4.482407	3.020185	NaN	3.753889
Noise - Park	3.273264	2.369127	NaN	2.321644
Noise - Street/Sidewalk	1.481928	2.516650	1.003333	2.379040
Noise - Vehicle	1.561620	2.089428	1.317222	2.597811
Panhandling	NaN	3.302315	NaN	1.174722
Posting Advertisement	NaN	0.733333	NaN	1.540000
Squeegee	NaN	NaN	NaN	NaN
Traffic	1.528025	2.508920	NaN	2.605853
Urinating in Public	NaN	2.536639	NaN	1.952500
Vending	1.877917	4.313333	NaN	3.008453

City	FLUSHING	FRESH MEADOWS	…	SOUTH RICHMOND HILL	\
Complaint Type			…		
Agency Issues	NaN	NaN	…	NaN	
Animal Abuse	3.219258	3.364074	…	4.183007	
Animal in a Park	NaN	NaN	…	NaN	
Bike/Roller/Skate Chronic	5.733056	NaN	…	30.910278	
Blocked Driveway	2.892248	3.943145	…	4.573347	
Derelict Vehicle	3.784891	4.735897	…	10.951176	
Disorderly Youth	2.144167	NaN	…	3.065000	
Drinking	2.860697	3.736944	…	5.494789	
Ferry Complaint	NaN	NaN	…	NaN	
Graffiti	1.864676	NaN	…	NaN	
Homeless Encampment	3.006720	5.239028	…	5.037199	
Illegal Fireworks	0.687917	NaN	…	0.475694	
Illegal Parking	2.892264	2.729423	…	5.398858	
Noise - Commercial	2.742076	2.166058	…	3.719600	
Noise - House of Worship	3.641556	NaN	…	1.987222	
Noise - Park	2.858142	3.414097	…	3.602083	
Noise - Street/Sidewalk	2.826192	2.518646	…	4.083527	
Noise - Vehicle	3.354520	2.385344	…	3.702136	
Panhandling	1.154028	1.570278	…	NaN	
Posting Advertisement	6.121667	NaN	…	NaN	
Squeegee	NaN	NaN	…	NaN	
Traffic	2.275664	1.459852	…	5.398796	
Urinating in Public	1.227662	1.600000	…	8.309444	
Vending	2.593521	2.631111	…	5.269664	

City	SPRINGFIELD GARDENS	FLORAL PARK	Unknown City	\
Complaint Type				
Agency Issues	NaN	NaN	5.080312	
Animal Abuse	12.497130	11.300595	3.409722	

Animal in a Park	NaN	NaN	NaN
Bike/Roller/Skate Chronic	NaN	NaN	5.858750
Blocked Driveway	8.166322	5.990202	8.956686
Derelict Vehicle	10.614604	13.846761	14.587037
Disorderly Youth	NaN	3.258333	NaN
Drinking	4.845787	8.227500	6.602037
Ferry Complaint	NaN	NaN	NaN
Graffiti	NaN	NaN	NaN
Homeless Encampment	5.783968	NaN	2.534444
Illegal Fireworks	5.556389	NaN	NaN
Illegal Parking	8.372372	8.908032	9.228882
Noise - Commercial	3.686382	4.368241	9.825355
Noise - House of Worship	3.710556	NaN	NaN
Noise - Park	1.666944	NaN	8.587024
Noise - Street/Sidewalk	4.065073	6.908981	6.863791
Noise - Vehicle	4.794774	1.958056	6.186821
Panhandling	6.348056	NaN	12.316389
Posting Advertisement	19.502917	NaN	NaN
Squeegee	NaN	NaN	NaN
Traffic	9.704884	NaN	8.726944
Urinating in Public	4.788796	NaN	NaN
Vending	6.851667	NaN	5.117222

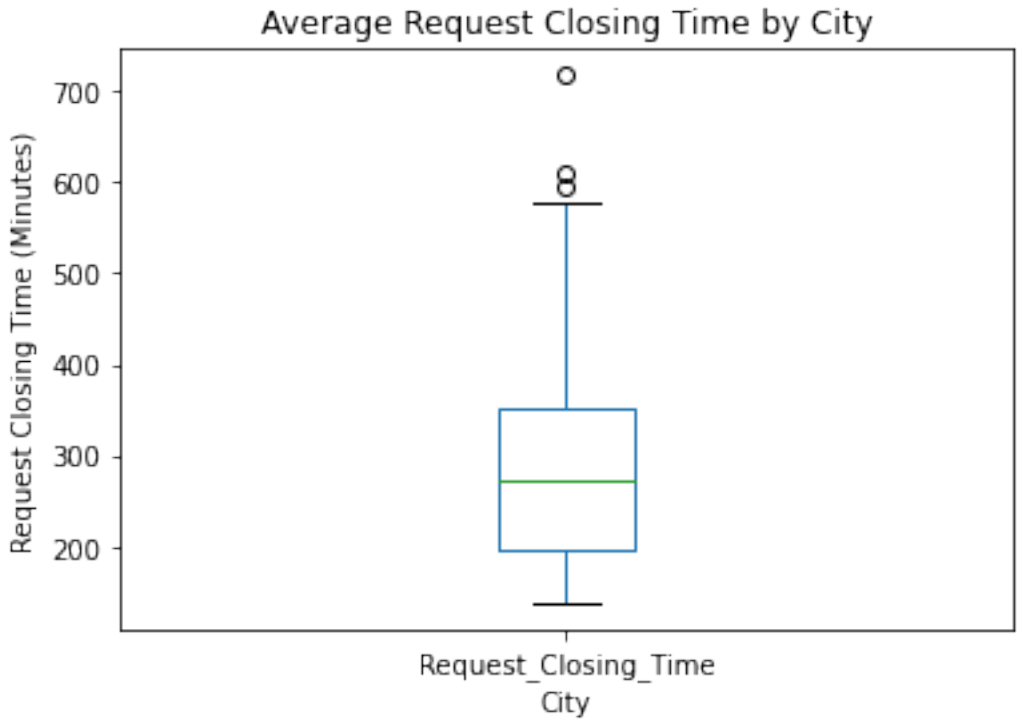
City	CAMBRIA HEIGHTS	BELLEROSE	ROSEDALE	GLEN OAKS	\
Complaint Type					
Agency Issues	NaN	NaN	NaN	NaN	
Animal Abuse	9.734167	8.880333	12.793535	12.991278	
Animal in a Park	NaN	NaN	NaN	NaN	
Bike/Roller/Skate Chronic	NaN	4.893333	8.520694	NaN	
Blocked Driveway	7.575615	8.365091	6.886983	9.469531	
Derelict Vehicle	13.978598	15.346356	13.920744	14.904274	
Disorderly Youth	NaN	1.842083	NaN	NaN	
Drinking	NaN	3.918056	5.500139	NaN	
Ferry Complaint	NaN	NaN	NaN	NaN	
Graffiti	NaN	NaN	1.624861	NaN	
Homeless Encampment	19.474954	39.147222	18.860764	NaN	
Illegal Fireworks	1.530556	6.663056	NaN	NaN	
Illegal Parking	8.943373	7.641894	9.432899	7.752140	
Noise - Commercial	4.042076	6.575278	7.445923	5.982794	
Noise - House of Worship	2.636944	2.196944	4.504444	NaN	
Noise - Park	NaN	1.414167	5.778164	4.367529	
Noise - Street/Sidewalk	4.539205	9.068953	5.881987	11.909954	
Noise - Vehicle	6.323172	3.757803	5.846678	11.423681	
Panhandling	NaN	7.490556	NaN	NaN	
Posting Advertisement	NaN	2.260000	NaN	NaN	
Squeegee	NaN	NaN	NaN	NaN	
Traffic	7.415754	5.603025	10.449611	5.547963	
Urinating in Public	NaN	7.541389	NaN	11.012500	

```
Vending                                NaN        NaN   6.876140   4.655088
```

```
City                       QUEENS VILLAGE      QUEENS
Complaint Type
Agency Issues                         NaN         NaN
Animal Abuse                    11.179960    2.374444
Animal in a Park                      NaN  336.842778
Bike/Roller/Skate Chronic             NaN         NaN
Blocked Driveway                 8.522350    3.063148
Derelict Vehicle                13.942425    4.533889
Disorderly Youth                      NaN         NaN
Drinking                         4.680889         NaN
Ferry Complaint                       NaN         NaN
Graffiti                        53.331944         NaN
Homeless Encampment              8.468713    7.205000
Illegal Fireworks                3.285167         NaN
Illegal Parking                  9.268298    3.068889
Noise - Commercial               9.468628    1.326296
Noise - House of Worship         2.513750    2.589722
Noise - Park                     3.526111         NaN
Noise - Street/Sidewalk          5.788313    3.651852
Noise - Vehicle                  8.114053    1.324722
Panhandling                      9.058889         NaN
Posting Advertisement            3.061667         NaN
Squeegee                              NaN         NaN
Traffic                          7.068467    2.116806
Urinating in Public              6.332778    0.340556
Vending                         14.599722         NaN

[24 rows x 54 columns]
```

[27]:
```
# Convert the Closed Date column to datetime format
df['Closed Date'] = pd.to_datetime(df['Closed Date'])

# Calculate the Request_Closing_Time in minutes
df['Request_Closing_Time'] = (df['Closed Date'] - df['Created Date']).dt.
 ↪total_seconds() / 60

# Group the data by City and calculate the mean Request_Closing_Time
mean_closing_time = df.groupby('City')['Request_Closing_Time'].mean()

# Create a boxplot to visualize the distribution of mean Request_Closing_Time␣
 ↪by City
mean_closing_time.plot(kind='box')

# Set the plot title and axis labels
plt.title('Average Request Closing Time by City')
```

```
plt.xlabel('City')
plt.ylabel('Request Closing Time (Minutes)')

# Show the plot
plt.show()
```



[28]:
```
import pandas as pd
from scipy.stats import kruskal


# Subset to relevant columns
df = df[['City', 'Request_Closing_Time']]

# Drop missing values in Request_Closing_Time
df = df.dropna(subset=['Request_Closing_Time'])

# Group data by City
grouped_data = [group['Request_Closing_Time'] for name, group in df.
 ↪groupby('City')]

# Perform Kruskal-Wallis H test
statistic, pvalue = kruskal(*grouped_data)
```

```
# Print results
print('Kruskal-Wallis H test results:')
print('Test statistic:', statistic)
print('p-value:', pvalue)
```

```
Kruskal-Wallis H test results:
Test statistic: 21612.302431410746
p-value: 0.0
```

[ ]:
```
df['City'].value_counts().head(1)
```

[ ]:
```
BROOKLYN    118849
Name: City, dtype: int64
```