In [1]:
```
import pandas as pd
import numpy as np
```

In [2]:
```
import matplotlib.pyplot as plt
%matplotlib inline
```

In [5]:
```
#Import Dataset
customers = pd.read_csv('311_Service_Requests_from_2010_to_Present.csv')
customers.head()
```

```
C:\Users\bhavna\AppData\Local\Temp\ipykernel_18716\1356398677.py:2: DtypeWarning:
Columns (48,49) have mixed types. Specify dtype option on import or set low_memory
=False.
  customers = pd.read_csv('311_Service_Requests_from_2010_to_Present.csv')
```

Out[5]:

	Unique Key	Created Date	Closed Date	Agency	Agency Name	Complaint Type	Descriptor	Location T
0	32310363	12/31/2015 11:59:45 PM	01/01/2016 12:55:15 AM	NYPD	New York City Police Department	Noise - Street/Sidewalk	Loud Music/Party	Street/Sidev
1	32309934	12/31/2015 11:59:44 PM	01/01/2016 01:26:57 AM	NYPD	New York City Police Department	Blocked Driveway	No Access	Street/Sidev
2	32309159	12/31/2015 11:59:29 PM	01/01/2016 04:51:03 AM	NYPD	New York City Police Department	Blocked Driveway	No Access	Street/Sidev
3	32305098	12/31/2015 11:57:46 PM	01/01/2016 07:43:13 AM	NYPD	New York City Police Department	Illegal Parking	Commercial Overnight Parking	Street/Sidev
4	32306529	12/31/2015 11:56:58 PM	01/01/2016 03:24:42 AM	NYPD	New York City Police Department	Illegal Parking	Blocked Sidewalk	Street/Sidev

5 rows × 53 columns

In [7]:
```
customers.describe()
```

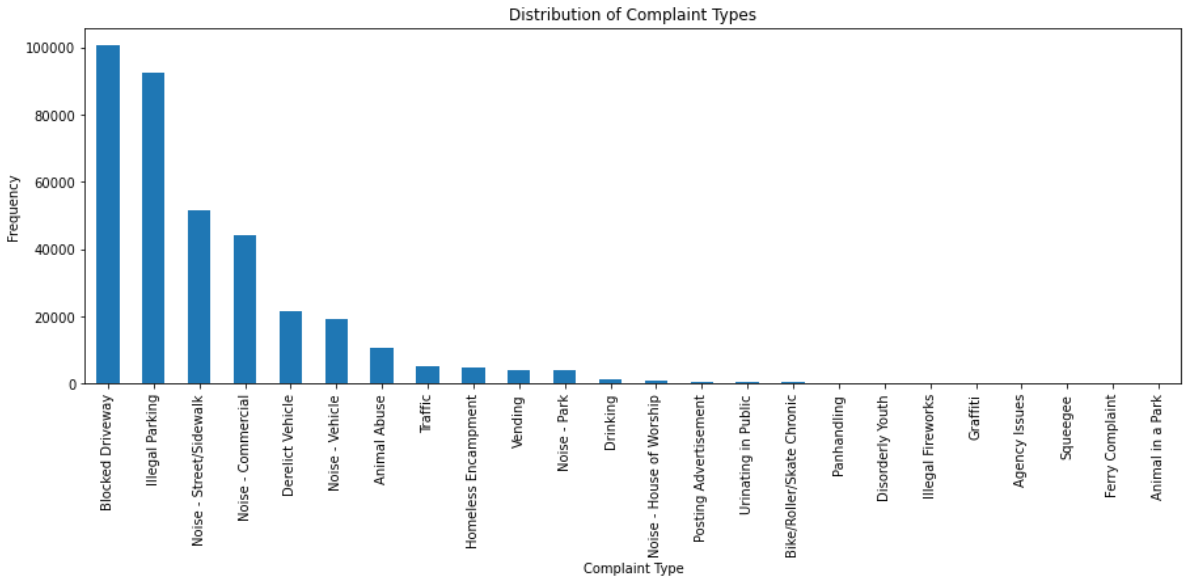
Out[7]:

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

In [8]:
```
customers.shape
```

Out[8]:
```
(364558, 53)
```

In [9]:
```
#visualize the dataset
complaint_types = customers['Complaint Type'].value_counts()
complaint_types.plot(kind='bar', figsize=(15,5))
plt.title("Distribution of Complaint Types")
plt.xlabel("Complaint Type")
plt.ylabel("Frequency")
plt.show()
```



In [12]:
```
#Print the columns of the DataFrame
print(customers.columns)
```

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

In [13]:
```
#Identify the shape of the dataset
customers.shape
```

Out[13]:
```
(364558, 53)
```

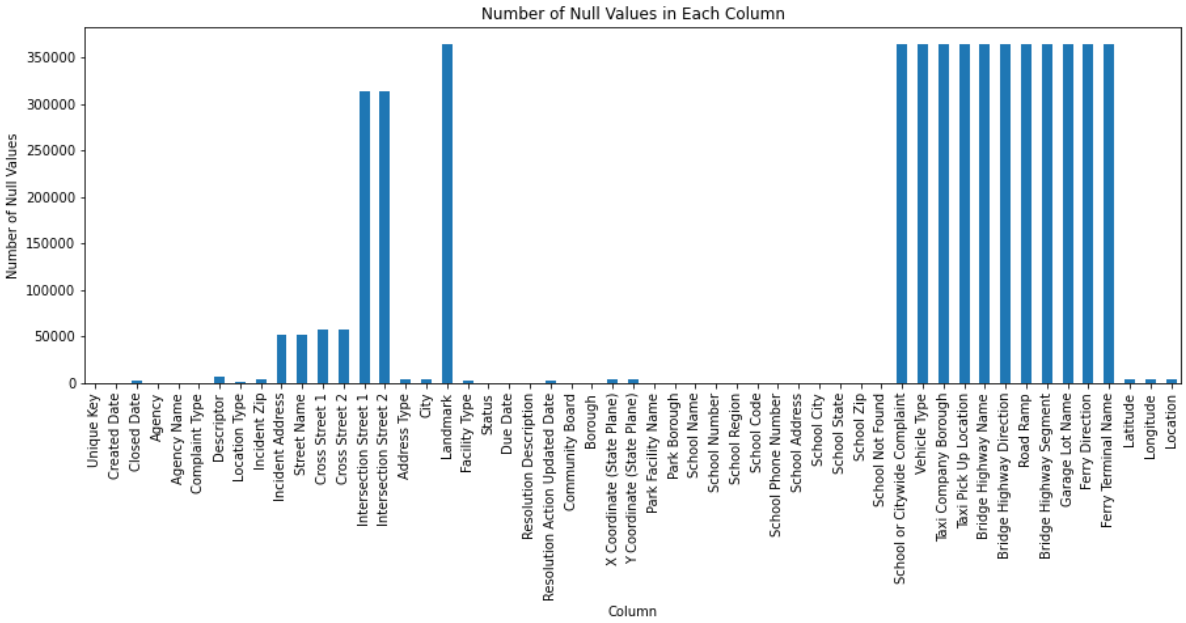
In [10]:
```
#Identify the variables with null values
customers.isnull()
```

Out[10]:

	Unique Key	Created Date	Closed Date	Agency	Agency Name	Complaint Type	Descriptor	Location Type	Incident Zip	In A
0	False	False	False	False	False	False	False	False	False	
1	False	False	False	False	False	False	False	False	False	
2	False	False	False	False	False	False	False	False	False	
3	False	False	False	False	False	False	False	False	False	
4	False	False	False	False	False	False	False	False	False	
...	...	...	...	...	...	...	...	...	...	
364553	False	False	False	False	False	False	False	False	False	
364554	False	False	False	False	False	False	False	False	False	
364555	False	False	False	False	False	False	False	False	False	
364556	False	False	False	False	False	False	False	False	False	
364557	False	False	False	False	False	False	False	False	False	

364558 rows × 53 columns

In [14]:
```
#Draw a frequency plot to show the number of null values in each column of the Data
null_values = customers.isnull().sum()
null_values.plot(kind='bar', figsize=(15,5))
plt.title("Number of Null Values in Each Column")
plt.xlabel("Column")
plt.ylabel("Number of Null Values")
plt.show()
```



In [15]:
```
#Remove the records whose Closed Date values are null
customers.dropna(subset = ['Closed Date'])
```

Out[15]:

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

362177 rows × 53 columns

In [20]:
```
customers.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  datetime64[ns]
 2   Closed Date                     362177 non-null  datetime64[ns]
 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: datetime64[ns](2), float64(10), int64(1), object(40)
memory usage: 147.4+ MB
```

In [19]:
```
#Calculate the time elapsed in closed and creation date
```

```
customers["Time Elapsed"] = customers["Closed Date"] - customers["Created Date"]
print(customers["Time Elapsed"])
```

```
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: Time Elapsed, Length: 364558, dtype: timedelta64[ns]
```

In [22]:
```
#Convert the calculated date to seconds to get a better representation
customers["Time Elapsed (seconds)"] = customers["Time Elapsed"].apply(lambda x: x.
print(customers["Time Elapsed (seconds)"])
```

```
0          3330.0
1          5233.0
2         17494.0
3         27927.0
4         12464.0
           ...
364553    37067.0
364554     8434.0
364555     1143.0
364556     9653.0
364557    10020.0
Name: Time Elapsed (seconds), Length: 364558, dtype: float64
```

In [24]:
```
# View the descriptive statistics for the newly created column
print(customers["Time Elapsed (seconds)"].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 (seconds), dtype: float64
```

In [25]:
```
#Check the number of null values in the Complaint_Type and City columns
print("Number of null values in Complaint Type column:", customers["Complaint Type
print("Number of null values in City column:", customers["City"].isna().sum())
```

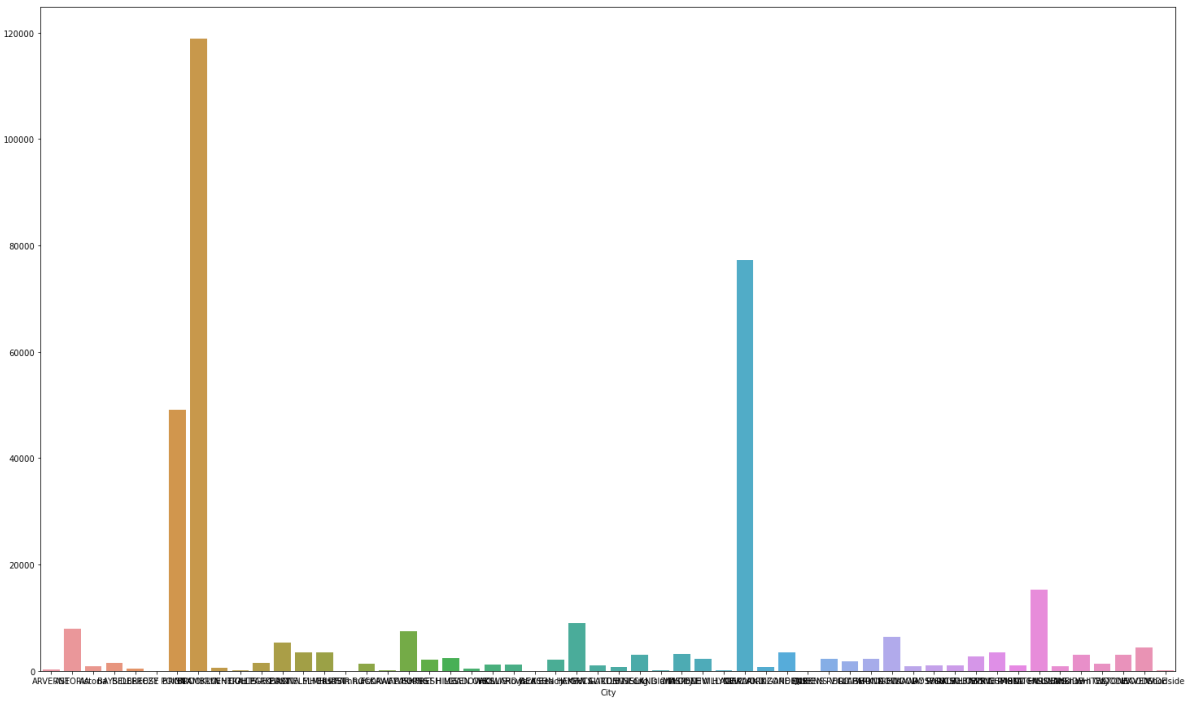
```
Number of null values in Complaint Type column: 0
Number of null values in City column: 2997
```

In [28]:
```
# Impute the NA value with Unknown City
customers["City"] = customers["City"].fillna("Unknown City")
print(customers)
```

```
        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  ... Road Ramp Bridge Highway Segment  \
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

       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
```

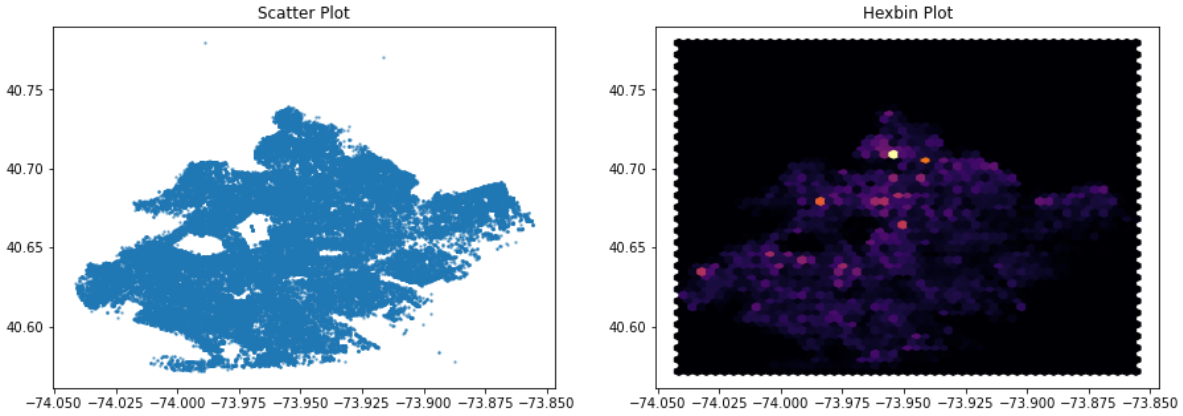
```
        Longitude                                  Location    Time Elapsed  \
0      -73.923501   (40.86568153633767, -73.92350095571744) 0 days 00:55:30
1      -73.915094  (40.775945312321085, -73.91509393898605) 0 days 01:27:13
2      -73.888525  (40.870324522111424, -73.88852464418646) 0 days 04:51:34
3      -73.828379   (40.83599404683083, -73.82837939584206) 0 days 07:45:27
4      -73.874170  (40.733059618956815, -73.87416975810375) 0 days 03:27:44
...           ...                                       ...             ...
364553 -73.860949   (40.69514470265117, -73.86094888534394) 0 days 10:17:47
364554 -73.907178   (40.86782963689454, -73.90717786644662) 0 days 02:20:34
364555 -73.950873  (40.821646626438095, -73.95087342885292) 0 days 00:19:03
364556 -73.853290   (40.88636077906953, -73.85329048666742) 0 days 02:40:53
364557 -73.803585  (40.674211762243935, -73.80358548685278) 0 days 02:47:00

       Time Elapsed (seconds)
0                      3330.0
1                      5233.0
2                     17494.0
3                     27927.0
4                     12464.0
...                       ...
364553                37067.0
364554                 8434.0
364555                 1143.0
364556                 9653.0
364557                10020.0

[364558 rows x 55 columns]
```

In [35]:
```
# Draw a frequency plot for the complaints in each city
import seaborn as sns


city_complaints = customers.groupby("City")["Complaint Type"].count()

plt.figure(figsize=(25, 15))
sns.barplot(x=city_complaints.index, y=city_complaints.values)
plt.show()
```



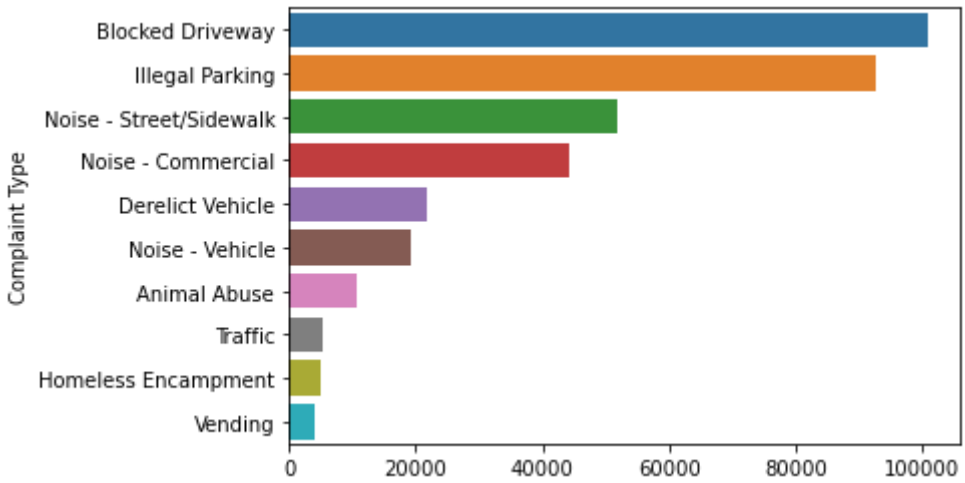
In [36]:
```
# Create a scatter and hexbin plot of the concentration of complaints across Brookl
```

```
brooklyn_df = customers[customers["City"] == "BROOKLYN"]

plt.figure(figsize=(15, 5))

plt.subplot(121)
plt.scatter(brooklyn_df["Longitude"], brooklyn_df["Latitude"], s=2, alpha=0.5)
plt.title("Scatter Plot")

plt.subplot(122)
plt.hexbin(brooklyn_df["Longitude"], brooklyn_df["Latitude"], gridsize=50, cmap="i
plt.title("Hexbin Plot")

plt.show()
```



In [37]:
```
#Plot a bar graph to show the types of complaints

complaint_types = customers.groupby("Complaint Type")["City"].count().sort_values(
sns.barplot(x=complaint_types.values[:10], y=complaint_types.index[:10])
```

Out[37]:
```
<AxesSubplot:ylabel='Complaint Type'>
```



In [39]:
```
# Check the frequency of various types of complaints for New York City
nyc_df = customers[customers["City"] == "NEW YORK"]
complaint_types = nyc_df["Complaint Type"].value_counts()
print(complaint_types)
```

```
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
```

In [40]:
```
#Find the top 10 complaint types
complaint_types = customers["Complaint Type"].value_counts()
top_10_complaint_types = complaint_types[:10]
print(top_10_complaint_types)
```

```
Blocked Driveway           100881
Illegal Parking             92679
Noise - Street/Sidewalk     51692
Noise - Commercial          44109
Derelict Vehicle            21661
Noise - Vehicle             19352
Animal Abuse                10541
Traffic                      5198
Homeless Encampment          4879
Vending                      4192
Name: Complaint Type, dtype: int64
```

In [41]:
```
#Display the various types of complaints in each city
city_complaint_types = customers.groupby("City")["Complaint Type"].value_counts()
print(city_complaint_types)
```

```
City      Complaint Type
ARVERNE   Illegal Parking            62
          Blocked Driveway           50
          Animal Abuse               46
          Derelict Vehicle           32
          Noise - Street/Sidewalk    29
                                    ...
Woodside  Illegal Parking           124
          Blocked Driveway           27
          Derelict Vehicle            8
          Noise - Street/Sidewalk     5
          Noise - Commercial          2
Name: Complaint Type, Length: 795, dtype: int64
```

In [43]:
```
#Create a DataFrame, df_new, which contains cities as columns and complaint types
df_new = customers.pivot_table(index="Complaint Type", columns="City", aggfunc=len
print(df_new)
```

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	...	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	

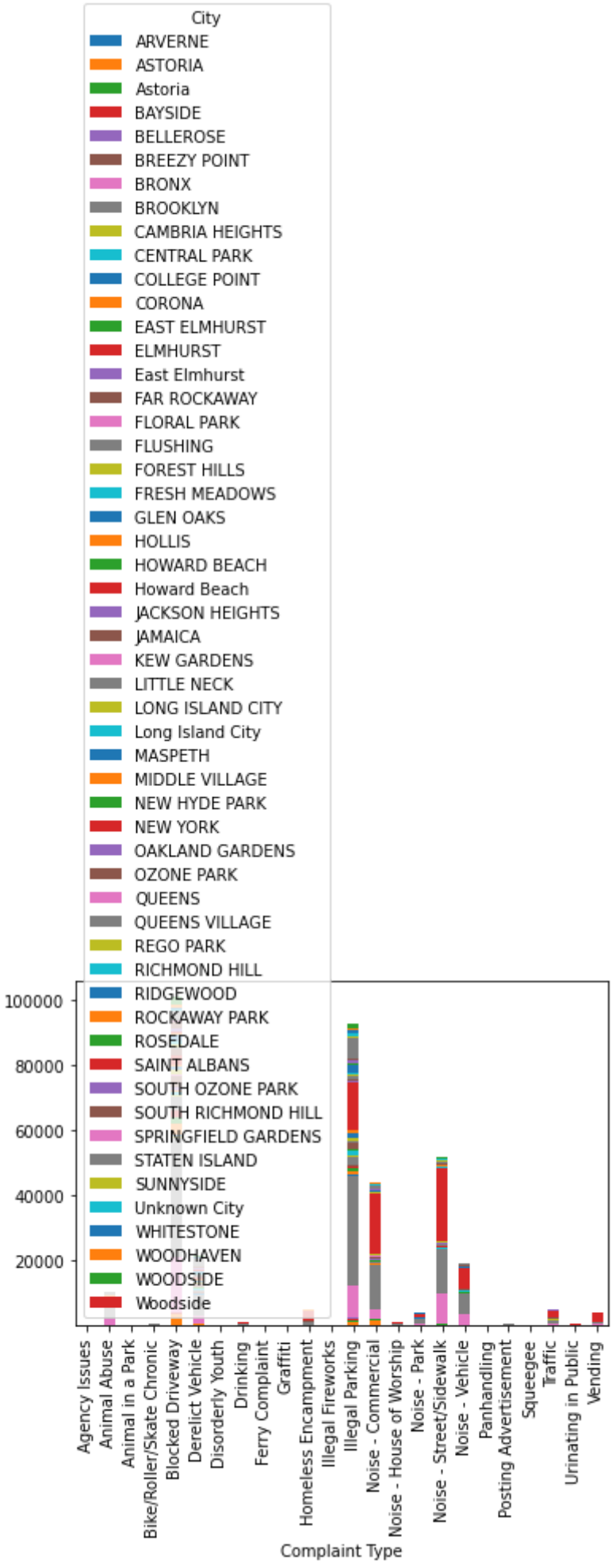
City	STATEN ISLAND	SUNNYSIDE	Unknown City	WHITESTONE	\
Complaint Type					
Agency Issues	NaN	NaN	8.0	NaN	
Animal Abuse	786.0	40.0	12.0	43.0	
Animal in a Park	NaN	NaN	NaN	NaN	
Bike/Roller/Skate Chronic	10.0	2.0	5.0	4.0	
Blocked Driveway	2845.0	278.0	333.0	279.0	
Derelict Vehicle	2184.0	17.0	201.0	279.0	
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	

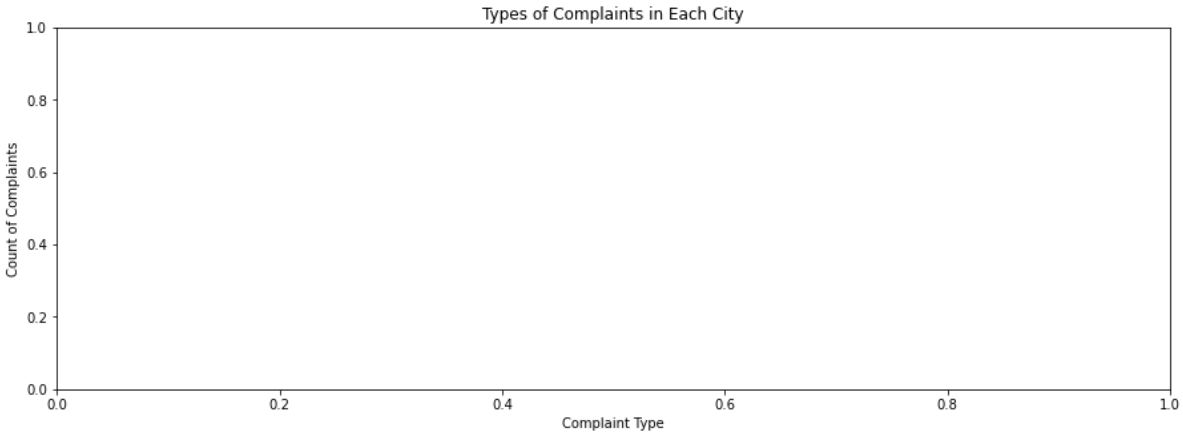
Panhandling	13.0	NaN	3.0	NaN
Posting Advertisement	517.0	3.0	1.0	NaN
Squeegee	NaN	NaN	NaN	NaN
Traffic	229.0	17.0	3.0	32.0
Urinating in Public	19.0	2.0	NaN	NaN
Vending	25.0	15.0	7.0	1.0

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 54 columns]

In [48]:
```
#Draw another chart that shows the types of complaints in each city in a single cha
df_new.plot(kind='bar', stacked=True)
plt.figure(figsize=(15, 5))
plt.xlabel("Complaint Type")
plt.ylabel("Count of Complaints")
plt.title("Types of Complaints in Each City")
plt.show()
```


City
ARVERNE
ASTORIA
Astoria
BAYSIDE
BELLEROSE
BREEZY POINT
BRONX
BROOKLYN
CAMBRIA HEIGHTS
CENTRAL PARK
COLLEGE POINT
CORONA
EAST ELMHURST
ELMHURST
East Elmhurst
FAR ROCKAWAY
FLORAL PARK
FLUSHING
FOREST HILLS
FRESH MEADOWS
GLEN OAKS
HOLLIS
HOWARD BEACH
Howard Beach
JACKSON HEIGHTS
JAMAICA
KEW GARDENS
LITTLE NECK
LONG ISLAND CITY
Long Island City
MASPETH
MIDDLE VILLAGE
NEW HYDE PARK
NEW YORK
OAKLAND GARDENS
OZONE PARK
QUEENS
QUEENS VILLAGE
REGO PARK
RICHMOND HILL
RIDGEWOOD
ROCKAWAY PARK
ROSEDALE
SAINT ALBANS
SOUTH OZONE PARK
SOUTH RICHMOND HILL
SPRINGFIELD GARDENS
STATEN ISLAND
SUNNYSIDE
Unknown City
WHITESTONE
WOODHAVEN
WOODSIDE
Woodside
100000
80000
60000
40000
20000
Agency Issues
Animal Abuse
Animal in a Park
Bike/Roller/Skate Chronic
Blocked Driveway
Derelict Vehicle
Disorderly Youth
Drinking
Ferry Complaint
Graffiti
Homeless Encampment
Illegal Fireworks
Illegal Parking
Noise - Commercial
Noise - House of Worship
Noise - Park
Noise - Street/Sidewalk
Noise - Vehicle
Panhandling
Posting Advertisement
Squeegee
Traffic
Urinating in Public
Vending
Complaint Type



In [50]:
```
#Sort the complaint types based on the average Request_Closing_Time grouping them j
customers["Request_Closing_Time"] = customers["Time Elapsed (seconds)"] / 3600
grouped = customers.groupby(["City", "Complaint Type"]).mean().reset_index()
grouped = grouped.sort_values("Request_Closing_Time", ascending=False)
print(grouped[["City", "Complaint Type", "Request_Closing_Time"]])
```

```
                   City            Complaint Type  Request_Closing_Time
510              QUEENS          Animal in a Park            336.842778
525       QUEENS VILLAGE                  Graffiti             53.331944
63             BELLEROSE       Homeless Encampment             39.147222
653  SOUTH RICHMOND HILL  Bike/Roller/Skate Chronic             30.910278
482      OAKLAND GARDENS       Homeless Encampment             28.654444
..                   ...                       ...                   ...
585            RIDGEWOOD     Posting Advertisement              0.248611
4                ARVERNE                  Drinking              0.238611
729        Unknown City           Ferry Complaint                   NaN
733        Unknown City  Noise - House of Worship                   NaN
738        Unknown City     Posting Advertisement                   NaN

[795 rows x 3 columns]
```

In [52]:
```
#Visualize the average of Request_Closing_Time Identify the significant variables b
import statsmodels.api as sm

customers["Request_Closing_Time"] = customers["Time Elapsed (seconds)"] / 3600
df_ols = customers[["Request_Closing_Time", "City", "Complaint Type"]]
df_ols = pd.get_dummies(df_ols, columns=["City", "Complaint Type"])

model = sm.OLS(df_ols["Request_Closing_Time"], df_ols.drop("Request_Closing_Time",
result = model.fit()

print(result.summary())
```

```
                            OLS Regression Results
==============================================================================
Dep. Variable:     Request_Closing_Time   R-squared:                       nan
Model:                              OLS   Adj. R-squared:                  nan
Method:                   Least Squares   F-statistic:                     nan
Date:                  Sun, 12 Feb 2023   Prob (F-statistic):              nan
Time:                          19:49:25   Log-Likelihood:                  nan
No. Observations:                364558   AIC:                             nan
Df Residuals:                    364481   BIC:                             nan
Df Model:                            76
Covariance Type:               nonrobust
==============================================================================
                                                coef    std err          t      P>|
t|      [0.025      0.975]
------------------------------------------------------------------------------
-------------------------
City_ARVERNE                                     nan        nan        nan        n
an         nan         nan
City_ASTORIA                                     nan        nan        nan        n
an         nan         nan
City_Astoria                                     nan        nan        nan        n
an         nan         nan
City_BAYSIDE                                     nan        nan        nan        n
an         nan         nan
City_BELLEROSE                                   nan        nan        nan        n
an         nan         nan
City_BREEZY POINT                                nan        nan        nan        n
an         nan         nan
City_BRONX                                       nan        nan        nan        n
an         nan         nan
City_BROOKLYN                                    nan        nan        nan        n
an         nan         nan
City_CAMBRIA HEIGHTS                             nan        nan        nan        n
an         nan         nan
City_CENTRAL PARK                                nan        nan        nan        n
an         nan         nan
City_COLLEGE POINT                               nan        nan        nan        n
an         nan         nan
City_CORONA                                      nan        nan        nan        n
an         nan         nan
City_EAST ELMHURST                               nan        nan        nan        n
an         nan         nan
City_ELMHURST                                    nan        nan        nan        n
an         nan         nan
City_East Elmhurst                               nan        nan        nan        n
an         nan         nan
City_FAR ROCKAWAY                                nan        nan        nan        n
an         nan         nan
City_FLORAL PARK                                 nan        nan        nan        n
an         nan         nan
City_FLUSHING                                    nan        nan        nan        n
an         nan         nan
City_FOREST HILLS                                nan        nan        nan        n
an         nan         nan
City_FRESH MEADOWS                               nan        nan        nan        n
an         nan         nan
City_GLEN OAKS                                   nan        nan        nan        n
an         nan         nan
City_HOLLIS                                      nan        nan        nan        n
an         nan         nan
City_HOWARD BEACH                                nan        nan        nan        n
an         nan         nan
City_Howard Beach                                nan        nan        nan        n
```

```
an         nan         nan
City_JACKSON HEIGHTS                    nan        nan        nan        n
an         nan         nan
City_JAMAICA                            nan        nan        nan        n
an         nan         nan
City_KEW GARDENS                        nan        nan        nan        n
an         nan         nan
City_LITTLE NECK                        nan        nan        nan        n
an         nan         nan
City_LONG ISLAND CITY                   nan        nan        nan        n
an         nan         nan
City_Long Island City                   nan        nan        nan        n
an         nan         nan
City_MASPETH                            nan        nan        nan        n
an         nan         nan
City_MIDDLE VILLAGE                     nan        nan        nan        n
an         nan         nan
City_NEW HYDE PARK                      nan        nan        nan        n
an         nan         nan
City_NEW YORK                           nan        nan        nan        n
an         nan         nan
City_OAKLAND GARDENS                    nan        nan        nan        n
an         nan         nan
City_OZONE PARK                         nan        nan        nan        n
an         nan         nan
City_QUEENS                             nan        nan        nan        n
an         nan         nan
City_QUEENS VILLAGE                     nan        nan        nan        n
an         nan         nan
City_REGO PARK                          nan        nan        nan        n
an         nan         nan
City_RICHMOND HILL                      nan        nan        nan        n
an         nan         nan
City_RIDGEWOOD                          nan        nan        nan        n
an         nan         nan
City_ROCKAWAY PARK                      nan        nan        nan        n
an         nan         nan
City_ROSEDALE                           nan        nan        nan        n
an         nan         nan
City_SAINT ALBANS                       nan        nan        nan        n
an         nan         nan
City_SOUTH OZONE PARK                   nan        nan        nan        n
an         nan         nan
City_SOUTH RICHMOND HILL                nan        nan        nan        n
an         nan         nan
City_SPRINGFIELD GARDENS                nan        nan        nan        n
an         nan         nan
City_STATEN ISLAND                      nan        nan        nan        n
an         nan         nan
City_SUNNYSIDE                          nan        nan        nan        n
an         nan         nan
City_Unknown City                       nan        nan        nan        n
an         nan         nan
City_WHITESTONE                         nan        nan        nan        n
an         nan         nan
City_WOODHAVEN                          nan        nan        nan        n
an         nan         nan
City_WOODSIDE                           nan        nan        nan        n
an         nan         nan
City_Woodside                           nan        nan        nan        n
an         nan         nan
Complaint Type_Agency Issues            nan        nan        nan        n
an         nan         nan
Complaint Type_Animal Abuse             nan        nan        nan        n
```

```
an         nan         nan
Complaint Type_Animal in a Park                nan        nan        nan        n
an         nan         nan
Complaint Type_Bike/Roller/Skate Chronic       nan        nan        nan        n
an         nan         nan
Complaint Type_Blocked Driveway                nan        nan        nan        n
an         nan         nan
Complaint Type_Derelict Vehicle                nan        nan        nan        n
an         nan         nan
Complaint Type_Disorderly Youth                nan        nan        nan        n
an         nan         nan
Complaint Type_Drinking                        nan        nan        nan        n
an         nan         nan
Complaint Type_Ferry Complaint                 nan        nan        nan        n
an         nan         nan
Complaint Type_Graffiti                        nan        nan        nan        n
an         nan         nan
Complaint Type_Homeless Encampment             nan        nan        nan        n
an         nan         nan
Complaint Type_Illegal Fireworks               nan        nan        nan        n
an         nan         nan
Complaint Type_Illegal Parking                 nan        nan        nan        n
an         nan         nan
Complaint Type_Noise - Commercial              nan        nan        nan        n
an         nan         nan
Complaint Type_Noise - House of Worship        nan        nan        nan        n
an         nan         nan
Complaint Type_Noise - Park                    nan        nan        nan        n
an         nan         nan
Complaint Type_Noise - Street/Sidewalk         nan        nan        nan        n
an         nan         nan
Complaint Type_Noise - Vehicle                 nan        nan        nan        n
an         nan         nan
Complaint Type_Panhandling                     nan        nan        nan        n
an         nan         nan
Complaint Type_Posting Advertisement           nan        nan        nan        n
an         nan         nan
Complaint Type_Squeegee                        nan        nan        nan        n
an         nan         nan
Complaint Type_Traffic                         nan        nan        nan        n
an         nan         nan
Complaint Type_Urinating in Public             nan        nan        nan        n
an         nan         nan
Complaint Type_Vending                         nan        nan        nan        n
an         nan         nan
==============================================================================
Omnibus:                        nan   Durbin-Watson:                     nan
Prob(Omnibus):                  nan   Jarque-Bera (JB):                  nan
Skew:                           nan   Prob(JB):                          nan
Kurtosis:                       nan   Cond. No.                      1.77e+14
==============================================================================

Notes:
[1] Standard Errors assume that the covariance matrix of the errors is correctly s
pecified.
[2] The smallest eigenvalue is 5.25e-24. This might indicate that there are
strong multicollinearity problems or that the design matrix is singular.
```

In [55]:

```
# Perform a Kruskal-Wallis H test
import scipy.stats as stats

# Your data
data = [customers[customers['City'] == city]['Request_Closing_Time'] for city in cu
```

```
# Perform the Kruskal-Wallis H test
statistic, pvalue = stats.kruskal(*data)
```

In [56]:
```
if pvalue < 0.05:
    print("The difference in median Request Closing Time between the different loca
else:
    print("The difference in median Request Closing Time between the different loca
```

```
The difference in median Request Closing Time between the different locations is n
ot statistically significant (p-value = nan)
```

In [59]:
```
print(customers)
```

```
        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 Segment Garage Lot Name  \
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

       Ferry Direction Ferry Terminal Name   Latitude  Longitude  \
0                  NaN                NaN  40.865682 -73.923501
1                  NaN                NaN  40.775945 -73.915094
2                  NaN                NaN  40.870325 -73.888525
3                  NaN                NaN  40.835994 -73.828379
4                  NaN                NaN  40.733060 -73.874170
...                ...                ...        ...        ...
364553             NaN                NaN  40.695145 -73.860949
364554             NaN                NaN  40.867830 -73.907178
364555             NaN                NaN  40.821647 -73.950873
364556             NaN                NaN  40.886361 -73.853290
364557             NaN                NaN  40.674212 -73.803585
```

```
                                        Location    Time Elapsed  \
0        (40.86568153633767, -73.92350095571744) 0 days 00:55:30
1       (40.775945312321085, -73.91509393898605) 0 days 01:27:13
2       (40.870324522111424, -73.88852464418646) 0 days 04:51:34
3        (40.83599404683083, -73.82837939584206) 0 days 07:45:27
4       (40.733059618956815, -73.87416975810375) 0 days 03:27:44
...                                          ...             ...
364553   (40.69514470265117, -73.86094888534394) 0 days 10:17:47
364554   (40.86782963689454, -73.90717786644662) 0 days 02:20:34
364555  (40.821646626438095, -73.95087342885292) 0 days 00:19:03
364556   (40.88636077906953, -73.85329048666742) 0 days 02:40:53
364557  (40.674211762243935, -73.80358548685278) 0 days 02:47:00

       Time Elapsed (seconds) Request_Closing_Time
0                      3330.0             0.925000
1                      5233.0             1.453611
2                     17494.0             4.859444
3                     27927.0             7.757500
4                     12464.0             3.462222
...                       ...                  ...
364553                37067.0            10.296389
364554                 8434.0             2.342778
364555                 1143.0             0.317500
364556                 9653.0             2.681389
364557                10020.0             2.783333

[364558 rows x 56 columns]
```

In [ ]: