Project - Customer Service Requests Analysis

In [1]:
```
# import the required libraries
import numpy as np
import pandas as pd
import matplotlib.pyplot as plt
%matplotlib inline
```

In [2]:
```
# Set the working directory
import io
%cd "F:\Akshay\Simplilearn\Electives\PYTHON_DATA_SCIENCE\PROJECTS\Customer Service Requests Analysis"
```

```
F:\Akshay\Simplilearn\Electives\PYTHON_DATA_SCIENCE\PROJECTS\Customer Service Requests Analysis
```

Analysis Tasks to be performed:

1. Import a 311 NYC service request.

In [3]:
```
# import the data
NYC311 = pd.read_csv('311_Service_Requests_from_2010_to_Present.csv',low_memory=False)
```

In [4]:
```
NYC311.shape # Rows and columns
```

Out[4]:
```
(300698, 53)
```

In [5]:
```
NYC311.head() # first 5 records
```

Out[5]:

	Unique Key	Created Date	Closed Date	Agency	Agency Name	Complaint Type	Descriptor	Location Type	Incident Zip	Incident Address	...	Bridge Highway Name	Bridge Highway Direction	Road Ramp	Bridge Highway Segment	Garag L Nan
0	32310363	12/31/2015 11:59:45 PM	01-01-16 0:55	NYPD	New York City Police Department	Noise - Street/Sidewalk	Loud Music/Party	Street/Sidewalk	10034.0	71 VERMILYEA AVENUE	...	NaN	NaN	NaN	NaN	Na

	Unique Key	Created Date	Closed Date	Agency	Agency Name	Complaint Type	Descriptor	Location Type	Incident Zip	Incident Address	...	Bridge Highway Name	Bridge Highway Direction	Road Ramp	Bridge Highway Segment	Garag L Nam
1	32309934	12/31/2015 11:59:44 PM	01-01-16 1:26	NYPD	New York City Police Department	Blocked Driveway	No Access	Street/Sidewalk	11105.0	27-07 23 AVENUE	...	NaN	NaN	NaN	NaN	Na
2	32309159	12/31/2015 11:59:29 PM	01-01-16 4:51	NYPD	New York City Police Department	Blocked Driveway	No Access	Street/Sidewalk	10458.0	2897 VALENTINE AVENUE	...	NaN	NaN	NaN	NaN	Na
3	32305098	12/31/2015 11:57:46 PM	01-01-16 7:43	NYPD	New York City Police Department	Illegal Parking	Commercial Overnight Parking	Street/Sidewalk	10461.0	2940 BAISLEY AVENUE	...	NaN	NaN	NaN	NaN	Na
4	32306529	12/31/2015 11:56:58 PM	01-01-16 3:24	NYPD	New York City Police Department	Illegal Parking	Blocked Sidewalk	Street/Sidewalk	11373.0	87-14 57 ROAD	...	NaN	NaN	NaN	NaN	Na

5 rows × 53 columns

In [6]:
```
NYC311.describe() # More info on the data
```

Out[6]:

	Unique Key	Incident Zip	X Coordinate (State Plane)	Y Coordinate (State Plane)	School or Citywide Complaint	Vehicle Type	Taxi Company Borough	Taxi Pick Up Location	Garage Lot Name	Latitude	Longitude
count	3.006980e+05	298083.000000	2.971580e+05	297158.000000	0.0	0.0	0.0	0.0	0.0	297158.000000	297158.000000
mean	3.130054e+07	10848.888645	1.004854e+06	203754.534416	NaN	NaN	NaN	NaN	NaN	40.725885	-73.925630
std	5.738547e+05	583.182081	2.175338e+04	29880.183529	NaN	NaN	NaN	NaN	NaN	0.082012	0.078454
min	3.027948e+07	83.000000	9.133570e+05	121219.000000	NaN	NaN	NaN	NaN	NaN	40.499135	-74.254937
25%	3.080118e+07	10310.000000	9.919752e+05	183343.000000	NaN	NaN	NaN	NaN	NaN	40.669796	-73.972142
50%	3.130436e+07	11208.000000	1.003158e+06	201110.500000	NaN	NaN	NaN	NaN	NaN	40.718661	-73.931781
75%	3.178446e+07	11238.000000	1.018372e+06	224125.250000	NaN	NaN	NaN	NaN	NaN	40.781840	-73.876805
max	3.231065e+07	11697.000000	1.067173e+06	271876.000000	NaN	NaN	NaN	NaN	NaN	40.912869	-73.700760

In [7]:
```
NYC311.dtypes # check datatypes of data
```

Out[7]:

```
Unique Key                       int64
Created Date                    object
Closed Date                     object
Agency                          object
Agency Name                     object
Complaint Type                  object
Descriptor                      object
Location Type                   object
Incident Zip                   float64
Incident Address                object
Street Name                     object
Cross Street 1                  object
Cross Street 2                  object
Intersection Street 1           object
Intersection Street 2           object
Address Type                    object
City                            object
Landmark                        object
Facility Type                   object
Status                          object
Due Date                        object
Resolution Description          object
Resolution Action Updated Date  object
Community Board                 object
Borough                         object
X Coordinate (State Plane)     float64
Y Coordinate (State Plane)     float64
Park Facility Name              object
Park Borough                    object
School Name                     object
School Number                   object
School Region                   object
School Code                     object
School Phone Number             object
School Address                  object
School City                     object
School State                    object
School Zip                      object
School Not Found                object
School or Citywide Complaint   float64
Vehicle Type                   float64
Taxi Company Borough           float64
Taxi Pick Up Location          float64
Bridge Highway Name             object
Bridge Highway Direction        object
Road Ramp                       object
Bridge Highway Segment          object
Garage Lot Name                float64
Ferry Direction                 object
```

```
Ferry Terminal Name               object
Latitude                         float64
Longitude                        float64
Location                          object
dtype: object
```

In [8]:

```
# Check Null values/missing values in the dataset
NYC311.isnull().sum().sort_values(ascending=False)
```

Out[8]:

```
School or Citywide Complaint      300698
Vehicle Type                      300698
Taxi Company Borough              300698
Taxi Pick Up Location             300698
Garage Lot Name                   300698
Ferry Direction                   300697
Ferry Terminal Name               300696
Road Ramp                         300485
Bridge Highway Segment            300485
Bridge Highway Name               300455
Bridge Highway Direction          300455
Landmark                          300349
Intersection Street 2             257336
Intersection Street 1             256840
Cross Street 2                     49779
Cross Street 1                     49279
Street Name                        44410
Incident Address                   44410
Descriptor                          5914
X Coordinate (State Plane)          3540
Latitude                            3540
Longitude                           3540
Y Coordinate (State Plane)          3540
Location                            3540
Address Type                        2815
Incident Zip                        2615
City                                2614
Resolution Action Updated Date      2187
Facility Type                       2171
Closed Date                         2164
Location Type                        131
Due Date                               3
School Region                          1
School Code                            1
School Zip                             1
Borough                                0
Agency                                 0
Agency Name                            0
Complaint Type                         0
Status                                 0
```

```
School Not Found                    0
Resolution Description              0
Community Board                     0
School State                        0
School City                         0
School Address                      0
School Phone Number                 0
School Number                       0
School Name                         0
Park Borough                        0
Park Facility Name                  0
Created Date                        0
Unique Key                          0
dtype: int64
```

In [9]:
```
# Descriptor - we will fill null values in Descriptor by Unspecified Description instead of deleting that rows
NYC311.Descriptor.fillna('Unspecified Description',inplace=True)
```

In [10]:
```
# City - we will fill null values in City by Unknown City instead of deleting that rows
NYC311.City.fillna('Unknown City',inplace=True)
```

In [11]:
```
# Closed Date - we will keep this as it is as it indicates that a complaint is raised but it is not closed/resolved yet
```

In [12]:
```
NYC311.isnull().sum().sort_values(ascending=False)
```

Out[12]:
```
School or Citywide Complaint      300698
Vehicle Type                      300698
Taxi Company Borough              300698
Garage Lot Name                   300698
Taxi Pick Up Location             300698
Ferry Direction                   300697
Ferry Terminal Name               300696
Bridge Highway Segment            300485
Road Ramp                         300485
Bridge Highway Direction          300455
Bridge Highway Name               300455
Landmark                          300349
Intersection Street 2             257336
Intersection Street 1             256840
Cross Street 2                     49779
Cross Street 1                     49279
Street Name                        44410
Incident Address                   44410
X Coordinate (State Plane)          3540
```

```
Latitude                          3540
Longitude                         3540
Y Coordinate (State Plane)        3540
Location                          3540
Address Type                      2815
Incident Zip                      2615
Resolution Action Updated Date    2187
Facility Type                     2171
Closed Date                       2164
Location Type                      131
Due Date                             3
School Region                        1
School Zip                           1
School Code                          1
City                                 0
Community Board                      0
Descriptor                           0
Complaint Type                       0
Agency Name                          0
Agency                               0
Status                               0
School Not Found                     0
Resolution Description               0
Borough                              0
School State                         0
School City                          0
School Address                       0
School Phone Number                  0
School Number                        0
School Name                          0
Park Borough                         0
Park Facility Name                   0
Created Date                         0
Unique Key                           0
dtype: int64
```

2. Read or convert the columns 'Created Date' and Closed Date' to datetime datatype and create a new column 'Request_Closing_Time' as the time elapsed between request creation and request closing.

In [13]:
```
# Converting Created Date, Closed Date to datetime format
NYC311['Created Date'] = pd.to_datetime(NYC311['Created Date'])
NYC311['Closed Date'] = pd.to_datetime(NYC311['Closed Date'])
```

In [14]:

```
# Creating a new column Request_Closing_Time
NYC311['Request_Closing_Time'] = NYC311['Closed Date'] - NYC311['Created Date']
```

In [15]:

```
NYC311.head(10) # first 10 records
```

Out[15]:

	Unique Key	Created Date	Closed Date	Agency	Agency Name	Complaint Type	Descriptor	Location Type	Incident Zip	Incident Address	...	Bridge Highway Direction	Road Ramp	Bridge Highway Segment	Garage Lot Name	Ferr Directio
0	32310363	2015-12-31 23:59:45	2016-01-01 00:55:00	NYPD	New York City Police Department	Noise - Street/Sidewalk	Loud Music/Party	Street/Sidewalk	10034.0	71 VERMILYEA AVENUE	...	NaN	NaN	NaN	NaN	Na
1	32309934	2015-12-31 23:59:44	2016-01-01 01:26:00	NYPD	New York City Police Department	Blocked Driveway	No Access	Street/Sidewalk	11105.0	27-07 23 AVENUE	...	NaN	NaN	NaN	NaN	Na
2	32309159	2015-12-31 23:59:29	2016-01-01 04:51:00	NYPD	New York City Police Department	Blocked Driveway	No Access	Street/Sidewalk	10458.0	2897 VALENTINE AVENUE	...	NaN	NaN	NaN	NaN	Na
3	32305098	2015-12-31 23:57:46	2016-01-01 07:43:00	NYPD	New York City Police Department	Illegal Parking	Commercial Overnight Parking	Street/Sidewalk	10461.0	2940 BAISLEY AVENUE	...	NaN	NaN	NaN	NaN	Na
4	32306529	2015-12-31 23:56:58	2016-01-01 03:24:00	NYPD	New York City Police Department	Illegal Parking	Blocked Sidewalk	Street/Sidewalk	11373.0	87-14 57 ROAD	...	NaN	NaN	NaN	NaN	Na
5	32306554	2015-12-31 23:56:30	2016-01-01 01:50:00	NYPD	New York City Police Department	Illegal Parking	Posted Parking Sign Violation	Street/Sidewalk	11215.0	260 21 STREET	...	NaN	NaN	NaN	NaN	Na
6	32306559	2015-12-31 23:55:32	2016-01-01 01:53:00	NYPD	New York City Police Department	Illegal Parking	Blocked Hydrant	Street/Sidewalk	10032.0	524 WEST 169 STREET	...	NaN	NaN	NaN	NaN	Na
7	32307009	2015-12-31 23:54:05	2016-01-01 01:42:00	NYPD	New York City Police Department	Blocked Driveway	No Access	Street/Sidewalk	10457.0	501 EAST 171 STREET	...	NaN	NaN	NaN	NaN	Na
8	32308581	2015-12-31 23:53:58	2016-01-01 08:27:00	NYPD	New York City Police Department	Illegal Parking	Posted Parking Sign Violation	Street/Sidewalk	11415.0	83-44 LEFFERTS BOULEVARD	...	NaN	NaN	NaN	NaN	Na

	Unique Key	Created Date	Closed Date	Agency	Agency Name	Complaint Type	Descriptor	Location Type	Incident Zip	Incident Address	...	Bridge Highway Direction	Road Ramp	Bridge Highway Segment	Garage Lot Name	Ferr Directio
9	32308391	2015-12-31 23:53:58	2016-01-01 01:17:00	NYPD	New York City Police Department	Blocked Driveway	No Access	Street/Sidewalk	11219.0	1408 66 STREET	...	NaN	NaN	NaN	NaN	Na

10 rows × 54 columns

In [16]:
```
# View of the Request_Closing_Time
NYC311.groupby(['Created Date','Closed Date','Request_Closing_Time']).size().reset_index()
```

Out[16]:

	Created Date	Closed Date	Request_Closing_Time	0
0	2015-03-29 00:33:01	2015-03-29 04:41:50	0 days 04:08:49	1
1	2015-03-29 00:33:02	2015-03-29 04:38:35	0 days 04:05:33	1
2	2015-03-29 00:33:03	2015-03-29 03:40:20	0 days 03:07:17	1
3	2015-03-29 00:33:28	2015-03-29 02:33:59	0 days 02:00:31	1
4	2015-03-29 00:34:32	2015-03-29 01:13:01	0 days 00:38:29	1
...	...	...	...	...
297811	2015-12-31 23:56:58	2016-01-01 03:24:00	0 days 03:27:02	1
297812	2015-12-31 23:57:46	2016-01-01 07:43:00	0 days 07:45:14	1
297813	2015-12-31 23:59:29	2016-01-01 04:51:00	0 days 04:51:31	1
297814	2015-12-31 23:59:44	2016-01-01 01:26:00	0 days 01:26:16	1
297815	2015-12-31 23:59:45	2016-01-01 00:55:00	0 days 00:55:15	1

297816 rows × 4 columns

3. Provide major insights/patterns that you can offer in a visual format (graphs or tables); at least 4 major conclusions that you can come up with after generic data mining.

In [17]:

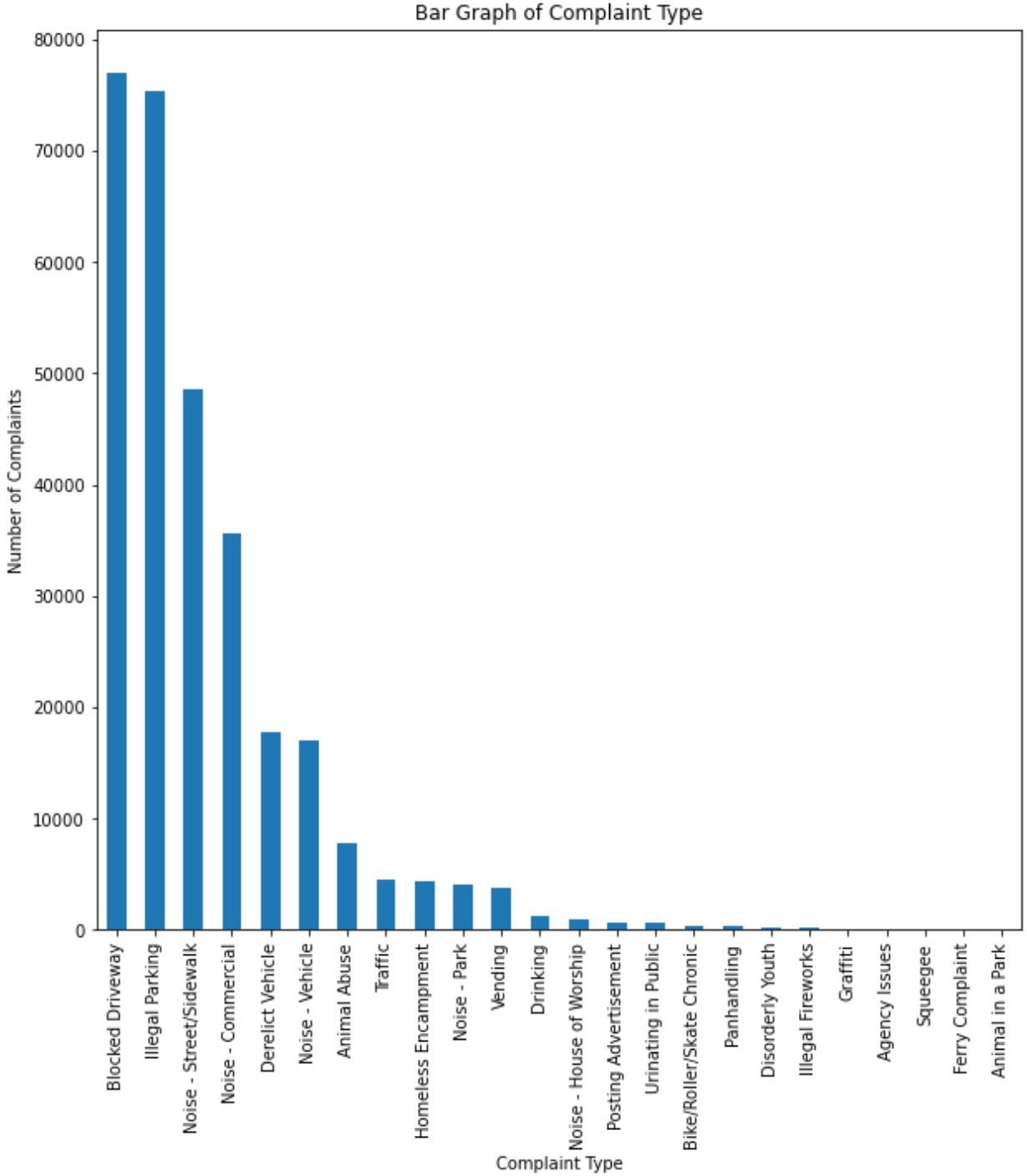
```
# Complaint Type Analysis
ComplaintType = NYC311['Complaint Type']
ComplaintType.value_counts()
```

Out[17]:

```
Blocked Driveway             77044
Illegal Parking              75361
Noise - Street/Sidewalk      48612
Noise - Commercial           35577
Derelict Vehicle             17718
Noise - Vehicle              17083
Animal Abuse                  7778
Traffic                       4498
Homeless Encampment           4416
Noise - Park                  4042
Vending                       3802
Drinking                      1280
Noise - House of Worship       931
Posting Advertisement          650
Urinating in Public            592
Bike/Roller/Skate Chronic      427
Panhandling                    307
Disorderly Youth               286
Illegal Fireworks              168
Graffiti                       113
Agency Issues                    6
Squeegee                         4
Ferry Complaint                  2
Animal in a Park                 1
Name: Complaint Type, dtype: int64
```

In [50]:

```
# Lets create a Bar Graph of Complaint Type
plt.figure(figsize=(10,10))
ComplaintType.value_counts().plot(kind='bar')
plt.xlabel('Complaint Type')
plt.ylabel('Number of Complaints')
plt.title('Bar Graph of Complaint Type')
plt.show()
```


Bar Graph of Complaint Type
Number of Complaints
80000
70000
60000
50000
40000
30000
20000
10000
0
Blocked Driveway
Illegal Parking
Noise - Street/Sidewalk
Noise - Commercial
Derelict Vehicle
Noise - Vehicle
Animal Abuse
Traffic
Homeless Encampment
Noise - Park
Vending
Drinking
Noise - House of Worship
Posting Advertisement
Urinating in Public
Bike/Roller/Skate Chronic
Panhandling
Disorderly Youth
Illegal Fireworks
Graffiti
Agency Issues
Squeegee
Ferry Complaint
Animal in a Park
Complaint Type

From above analysis we found that the maximum number of complaints are of Blocked Driveway

In [19]:
```
# City with Highest Number of Complaints
Citywise_complaints = NYC311['City'].value_counts()
Citywise_complaints
```

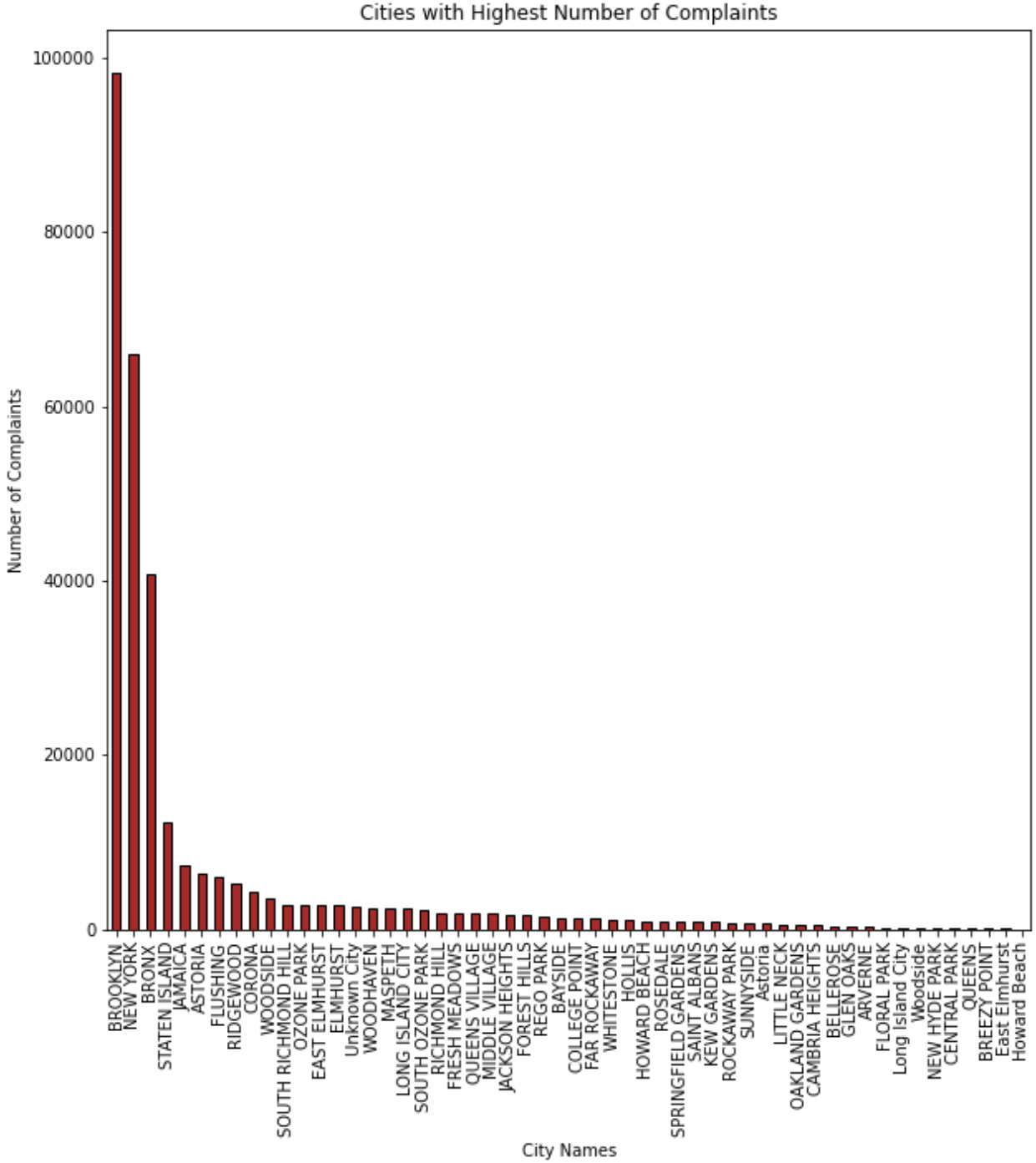
Out[19]:

BROOKLYN	98307
NEW YORK	65994
BRONX	40702
STATEN ISLAND	12343
JAMAICA	7296
ASTORIA	6330
FLUSHING	5971
RIDGEWOOD	5163
CORONA	4295
WOODSIDE	3544
SOUTH RICHMOND HILL	2774
OZONE PARK	2755
EAST ELMHURST	2734
ELMHURST	2673
Unknown City	2614
WOODHAVEN	2464
MASPETH	2462
LONG ISLAND CITY	2437
SOUTH OZONE PARK	2173
RICHMOND HILL	1904
FRESH MEADOWS	1899
QUEENS VILLAGE	1814
MIDDLE VILLAGE	1765
JACKSON HEIGHTS	1689
FOREST HILLS	1688
REGO PARK	1486
BAYSIDE	1221
COLLEGE POINT	1220
FAR ROCKAWAY	1179
WHITESTONE	1098
HOLLIS	1012
HOWARD BEACH	931
ROSEDALE	922
SPRINGFIELD GARDENS	883
SAINT ALBANS	834
KEW GARDENS	771
ROCKAWAY PARK	745
SUNNYSIDE	723
Astoria	717
LITTLE NECK	559

```
OAKLAND GARDENS          551
CAMBRIA HEIGHTS          477
BELLEROSE                375
GLEN OAKS                306
ARVERNE                  220
FLORAL PARK              152
Long Island City         134
Woodside                 120
NEW HYDE PARK             98
CENTRAL PARK              97
QUEENS                    32
BREEZY POINT              30
East Elmhurst             14
Howard Beach               1
Name: City, dtype: int64
```

In [51]:

```
# Lets create a Bar Graph of Cities with Highest Number of Complaints
plt.figure(figsize=(10,10))
Citywise_complaints.plot(kind='bar',color='brown',edgecolor='black')
plt.xlabel('City Names')
plt.ylabel('Number of Complaints')
plt.title('Cities with Highest Number of Complaints')
plt.show()
```


Cities with Highest Number of Complaints
Number of Complaints
100000
80000
60000
40000
20000
0
BROOKLYN
NEW YORK
BRONX
STATEN ISLAND
JAMAICA
ASTORIA
FLUSHING
RIDGEWOOD
CORONA
WOODSIDE
SOUTH RICHMOND HILL
OZONE PARK
EAST ELMHURST
ELMHURST
Unknown City
WOODHAVEN
MASPETH
LONG ISLAND CITY
SOUTH OZONE PARK
RICHMOND HILL
FRESH MEADOWS
QUEENS VILLAGE
MIDDLE VILLAGE
JACKSON HEIGHTS
FOREST HILLS
REGO PARK
BAYSIDE
COLLEGE POINT
FAR ROCKAWAY
WHITESTONE
HOLLIS
HOWARD BEACH
ROSEDALE
SPRINGFIELD GARDENS
SAINT ALBANS
KEW GARDENS
ROCKAWAY PARK
SUNNYSIDE
Astoria
LITTLE NECK
OAKLAND GARDENS
CAMBRIA HEIGHTS
BELLEROSE
GLEN OAKS
ARVERNE
FLORAL PARK
Long Island City
Woodside
NEW HYDE PARK
CENTRAL PARK
QUEENS
BREEZY POINT
East Elmhurst
Howard Beach
City Names

From above analysis we found that the highest number of complaints are from BROOKLYN and lowest are from Howard Beach

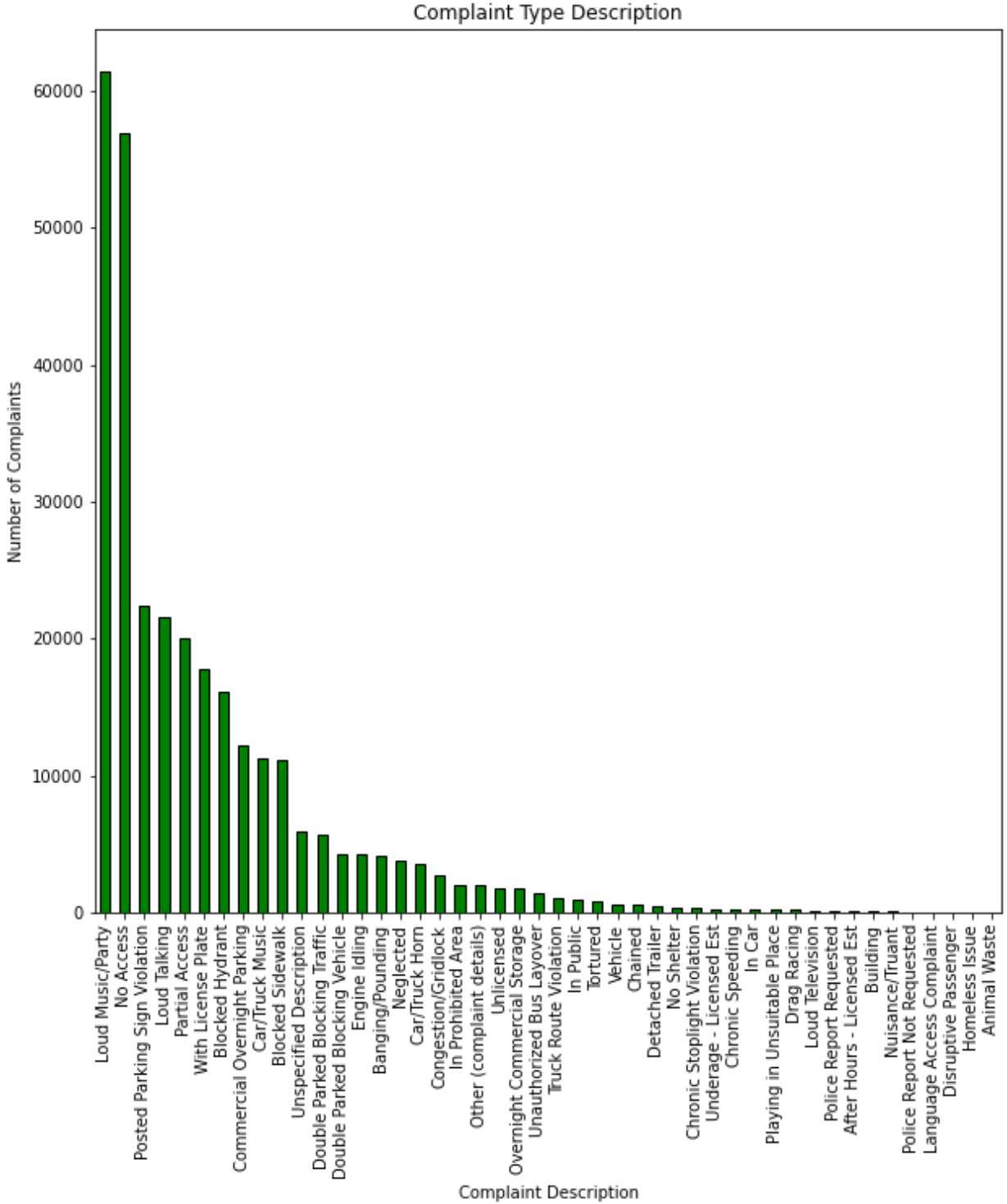
In [21]:
```
# Description of the type of complaints
Complaint_Description = NYC311.Descriptor.value_counts()
Complaint_Description
```

Out[21]:
```
Loud Music/Party                  61430
No Access                         56976
Posted Parking Sign Violation     22440
Loud Talking                      21584
Partial Access                    20068
With License Plate                17718
Blocked Hydrant                   16081
Commercial Overnight Parking      12189
Car/Truck Music                   11273
Blocked Sidewalk                  11121
Unspecified Description            5914
Double Parked Blocking Traffic     5731
Double Parked Blocking Vehicle     4211
Engine Idling                      4189
Banging/Pounding                   4165
Neglected                          3787
Car/Truck Horn                     3511
Congestion/Gridlock                2761
In Prohibited Area                 2025
Other (complaint details)          1969
Unlicensed                         1777
Overnight Commercial Storage       1757
Unauthorized Bus Layover           1367
Truck Route Violation              1014
In Public                           932
Tortured                            854
Vehicle                             590
Chained                             535
Detached Trailer                    464
No Shelter                          382
Chronic Stoplight Violation         280
Underage - Licensed Est             271
Chronic Speeding                    268
In Car                              251
Playing in Unsuitable Place         245
Drag Racing                         175
Loud Television                      93
Police Report Requested              90
After Hours - Licensed Est           77
Building                             60
```

```
Nuisance/Truant                     41
Police Report Not Requested         23
Language Access Complaint            6
Disruptive Passenger                 1
Homeless Issue                       1
Animal Waste                         1
Name: Descriptor, dtype: int64
```

In [52]:

```
# Lets create a Bar Graph of Complaint Type Description
plt.figure(figsize=(10,10))
Complaint_Description.plot(kind='bar',color='green',edgecolor='black')
plt.xlabel('Complaint Description')
plt.ylabel('Number of Complaints')
plt.title('Complaint Type Description')
plt.show()
```


Complaint Type Description
Number of Complaints
0
10000
20000
30000
40000
50000
60000
Loud Music/Party
No Access
Posted Parking Sign Violation
Loud Talking
Partial Access
With License Plate
Blocked Hydrant
Commercial Overnight Parking
Car/Truck Music
Blocked Sidewalk
Unspecified Description
Double Parked Blocking Traffic
Double Parked Blocking Vehicle
Engine Idling
Banging/Pounding
Neglected
Car/Truck Horn
Congestion/Gridlock
In Prohibited Area
Other (complaint details)
Unlicensed
Overnight Commercial Storage
Unauthorized Bus Layover
Truck Route Violation
In Public
Tortured
Vehicle
Chained
Detached Trailer
No Shelter
Chronic Stoplight Violation
Underage - Licensed Est
Chronic Speeding
In Car
Playing in Unsuitable Place
Drag Racing
Loud Television
Police Report Requested
After Hours - Licensed Est
Building
Nuisance/Truant
Police Report Not Requested
Language Access Complaint
Disruptive Passenger
Homeless Issue
Animal Waste
Complaint Description

From above analysis we found that the maximum number of complaints are for Loud Music/Party

In [23]:
```
# Now we will create a data for city and complaint type
City_Complaints = NYC311.groupby(['Complaint Type','City']).size().reset_index(name='Count')
City_Complaints
```

Out[23]:

	Complaint Type	City	Count
0	Agency Issues	Unknown City	6
1	Animal Abuse	ARVERNE	38
2	Animal Abuse	ASTORIA	125
3	Animal Abuse	BAYSIDE	37
4	Animal Abuse	BELLEROSE	7
...	...	...	...
777	Vending	SUNNYSIDE	15
778	Vending	Unknown City	7
779	Vending	WHITESTONE	1
780	Vending	WOODHAVEN	6
781	Vending	WOODSIDE	15

782 rows × 3 columns

In [24]:
```
# view of the complaint type and city for which it is maximum
City_Complaints.loc[City_Complaints.groupby('Complaint Type')['Count'].idxmax()]
```

Out[24]:

	Complaint Type	City	Count
0	Agency Issues	Unknown City	6
7	Animal Abuse	BROOKLYN	2394
48	Animal in a Park	QUEENS	1
63	Bike/Roller/Skate Chronic	NEW YORK	225

	Complaint Type	City	Count
83	Blocked Driveway	BROOKLYN	28148
135	Derelict Vehicle	BROOKLYN	5181
185	Disorderly Youth	BROOKLYN	72
236	Drinking	NEW YORK	295
255	Ferry Complaint	Unknown City	2
260	Graffiti	BROOKLYN	43
301	Homeless Encampment	NEW YORK	2775
323	Illegal Fireworks	BROOKLYN	61
350	Illegal Parking	BROOKLYN	27462
425	Noise - Commercial	NEW YORK	14550
450	Noise - House of Worship	BROOKLYN	340
484	Noise - Park	BROOKLYN	1555
551	Noise - Street/Sidewalk	NEW YORK	20433
600	Noise - Vehicle	NEW YORK	5485
634	Panhandling	NEW YORK	193
660	Posting Advertisement	STATEN ISLAND	516
663	Squeegee	NEW YORK	4
688	Traffic	NEW YORK	1549
728	Urinating in Public	NEW YORK	251
763	Vending	NEW YORK	2399

From above table we can get information on which complaint is maximum by city-wise and its count. For e.g. Animal Abuse is maximum in BROOKLYN with count = 2394, Bike/Roller/Skate Chronic is maximum in NEW YORK with count = 225 and so on

In [25]:
```
# Here we will combine the Complaint Type, Descriptor and City
completedata = NYC311.groupby(['Complaint Type','Descriptor','City']).size().reset_index(name='Count')
completedata.head(50)
```

Out[25]:

	Complaint Type	Descriptor	City	Count
0	Agency Issues	Language Access Complaint	Unknown City	6
1	Animal Abuse	Chained	ARVERNE	3
2	Animal Abuse	Chained	ASTORIA	16
3	Animal Abuse	Chained	BAYSIDE	6
4	Animal Abuse	Chained	BREEZY POINT	1
5	Animal Abuse	Chained	BRONX	132
6	Animal Abuse	Chained	BROOKLYN	165
7	Animal Abuse	Chained	CAMBRIA HEIGHTS	2
8	Animal Abuse	Chained	CORONA	3
9	Animal Abuse	Chained	EAST ELMHURST	5
10	Animal Abuse	Chained	ELMHURST	1
11	Animal Abuse	Chained	FAR ROCKAWAY	3
12	Animal Abuse	Chained	FLUSHING	7
13	Animal Abuse	Chained	FRESH MEADOWS	2
14	Animal Abuse	Chained	HOLLIS	1
15	Animal Abuse	Chained	HOWARD BEACH	1
16	Animal Abuse	Chained	JACKSON HEIGHTS	1
17	Animal Abuse	Chained	JAMAICA	29
18	Animal Abuse	Chained	LITTLE NECK	1
19	Animal Abuse	Chained	MASPETH	1
20	Animal Abuse	Chained	NEW YORK	69
21	Animal Abuse	Chained	OAKLAND GARDENS	1
22	Animal Abuse	Chained	OZONE PARK	3

	Complaint Type	Descriptor	City	Count
23	Animal Abuse	Chained	QUEENS VILLAGE	3
24	Animal Abuse	Chained	REGO PARK	2
25	Animal Abuse	Chained	RICHMOND HILL	2
26	Animal Abuse	Chained	RIDGEWOOD	4
27	Animal Abuse	Chained	ROCKAWAY PARK	4
28	Animal Abuse	Chained	ROSEDALE	5
29	Animal Abuse	Chained	SAINT ALBANS	6
30	Animal Abuse	Chained	SOUTH OZONE PARK	6
31	Animal Abuse	Chained	SOUTH RICHMOND HILL	3
32	Animal Abuse	Chained	SPRINGFIELD GARDENS	4
33	Animal Abuse	Chained	STATEN ISLAND	36
34	Animal Abuse	Chained	WHITESTONE	3
35	Animal Abuse	Chained	WOODHAVEN	3
36	Animal Abuse	Chained	WOODSIDE	1
37	Animal Abuse	In Car	ASTORIA	9
38	Animal Abuse	In Car	BAYSIDE	4
39	Animal Abuse	In Car	BELLEROSE	1
40	Animal Abuse	In Car	BRONX	36
41	Animal Abuse	In Car	BROOKLYN	61
42	Animal Abuse	In Car	COLLEGE POINT	3
43	Animal Abuse	In Car	CORONA	1
44	Animal Abuse	In Car	EAST ELMHURST	3
45	Animal Abuse	In Car	ELMHURST	4
46	Animal Abuse	In Car	FAR ROCKAWAY	3
47	Animal Abuse	In Car	FLUSHING	4
48	Animal Abuse	In Car	FOREST HILLS	3

	Complaint Type	Descriptor	City	Count
49	Animal Abuse	In Car	FRESH MEADOWS	2

In [26]:
```
# View of the Complaint Type, Descriptor and City for which it is maximum
completedata.loc[completedata.groupby('Descriptor')['Count'].idxmax()]
```

Out[26]:

	Complaint Type	Descriptor	City	Count
463	Drinking	After Hours - Licensed Est	NEW YORK	28
234	Animal in a Park	Animal Waste	QUEENS	1
1085	Noise - Commercial	Banging/Pounding	NEW YORK	1481
642	Illegal Parking	Blocked Hydrant	BROOKLYN	6697
693	Illegal Parking	Blocked Sidewalk	BROOKLYN	4017
1663	Posting Advertisement	Building	NEW YORK	23
1526	Noise - Vehicle	Car/Truck Horn	NEW YORK	1167
1550	Noise - Vehicle	Car/Truck Music	BROOKLYN	3127
6	Animal Abuse	Chained	BROOKLYN	165
1690	Traffic	Chronic Speeding	BROOKLYN	57
1744	Traffic	Chronic Stoplight Violation	NEW YORK	84
742	Illegal Parking	Commercial Overnight Parking	BROOKLYN	4346
1781	Traffic	Congestion/Gridlock	NEW YORK	1317
791	Illegal Parking	Detached Trailer	BROOKLYN	132
541	Ferry Complaint	Disruptive Passenger	Unknown City	1
833	Illegal Parking	Double Parked Blocking Traffic	BROOKLYN	1958
880	Illegal Parking	Double Parked Blocking Vehicle	BROOKLYN	1564
1803	Traffic	Drag Racing	BROOKLYN	41
1616	Noise - Vehicle	Engine Idling	NEW YORK	1456
542	Ferry Complaint	Homeless Issue	Unknown City	1
54	Animal Abuse	In Car	NEW YORK	73

	Complaint Type	Descriptor	City	Count
1917	Vending	In Prohibited Area	NEW YORK	1526
491	Drinking	In Public	NEW YORK	191
0	Agency Issues	Language Access Complaint	Unknown City	6
1430	Noise - Street/Sidewalk	Loud Music/Party	NEW YORK	13674
1479	Noise - Street/Sidewalk	Loud Talking	NEW YORK	6759
1264	Noise - Commercial	Loud Television	NEW YORK	37
72	Animal Abuse	Neglected	BROOKLYN	1196
269	Blocked Driveway	No Access	BROOKLYN	21422
116	Animal Abuse	No Shelter	BROOKLYN	102
416	Disorderly Youth	Nuisance/Truant	BROOKLYN	14
153	Animal Abuse	Other (complaint details)	BROOKLYN	582
927	Illegal Parking	Overnight Commercial Storage	BROOKLYN	558
320	Blocked Driveway	Partial Access	BROOKLYN	6726
444	Disorderly Youth	Playing in Unsuitable Place	NEW YORK	61
548	Graffiti	Police Report Not Requested	NEW YORK	10
555	Graffiti	Police Report Requested	BROOKLYN	36
973	Illegal Parking	Posted Parking Sign Violation	BROOKLYN	7696
196	Animal Abuse	Tortured	BROOKLYN	288
1844	Traffic	Truck Route Violation	JAMAICA	484
1042	Illegal Parking	Unauthorized Bus Layover	NEW YORK	496
526	Drinking	Underage - Licensed Est	NEW YORK	76
1952	Vending	Unlicensed	NEW YORK	873
594	Homeless Encampment	Unspecified Description	NEW YORK	2775
1683	Posting Advertisement	Vehicle	STATEN ISLAND	515
370	Derelict Vehicle	With License Plate	BROOKLYN	5181

From above table we can get information like Complaint Type - Drinking with Descriptor - After Hours - Licensed Est is maximum in the city NEW YORK with count = 28, Complaint Type - Illegal Parking with Descriptor - Blocked Hydrant is maximum in the city BROOKLYN with count = 6697

4. Order the complaint types based on the average 'Request_Closing_Time', grouping them for different locations.

In [27]:
```
# Creating a new data which has columns Complaint Type, Request_Closing_Time, City
Request_Time_City = NYC311.groupby(['Complaint Type','Request_Closing_Time','City']).size().reset_index(name='Count')
Request_Time_City.head(50)
```

Out[27]:

	Complaint Type	Request_Closing_Time	City	Count
0	Agency Issues	0 days 01:07:53	Unknown City	1
1	Agency Issues	0 days 02:42:16	Unknown City	1
2	Agency Issues	0 days 02:57:08	Unknown City	1
3	Agency Issues	0 days 06:51:26	Unknown City	1
4	Agency Issues	0 days 07:32:00	Unknown City	1
5	Agency Issues	0 days 10:23:00	Unknown City	1
6	Animal Abuse	0 days 00:03:53	NEW YORK	1
7	Animal Abuse	0 days 00:03:54	BROOKLYN	1
8	Animal Abuse	0 days 00:04:06	NEW YORK	1
9	Animal Abuse	0 days 00:04:46	NEW YORK	1
10	Animal Abuse	0 days 00:05:25	BROOKLYN	1
11	Animal Abuse	0 days 00:05:36	ELMHURST	1
12	Animal Abuse	0 days 00:05:41	BROOKLYN	1
13	Animal Abuse	0 days 00:06:00	BROOKLYN	2
14	Animal Abuse	0 days 00:06:00	NEW YORK	1

	Complaint Type	Request_Closing_Time	City	Count
15	Animal Abuse	0 days 00:06:08	BROOKLYN	1
16	Animal Abuse	0 days 00:06:38	BROOKLYN	1
17	Animal Abuse	0 days 00:06:47	BROOKLYN	1
18	Animal Abuse	0 days 00:07:00	BROOKLYN	2
19	Animal Abuse	0 days 00:07:00	ROCKAWAY PARK	1
20	Animal Abuse	0 days 00:07:00	STATEN ISLAND	1
21	Animal Abuse	0 days 00:07:19	BROOKLYN	1
22	Animal Abuse	0 days 00:07:38	BROOKLYN	1
23	Animal Abuse	0 days 00:07:45	NEW YORK	1
24	Animal Abuse	0 days 00:07:51	JAMAICA	1
25	Animal Abuse	0 days 00:08:00	BROOKLYN	2
26	Animal Abuse	0 days 00:08:00	WOODHAVEN	1
27	Animal Abuse	0 days 00:08:03	NEW YORK	1
28	Animal Abuse	0 days 00:08:11	BROOKLYN	1
29	Animal Abuse	0 days 00:08:12	BROOKLYN	1
30	Animal Abuse	0 days 00:08:15	FLUSHING	1
31	Animal Abuse	0 days 00:08:17	FRESH MEADOWS	1
32	Animal Abuse	0 days 00:08:19	BRONX	1
33	Animal Abuse	0 days 00:08:36	BROOKLYN	1
34	Animal Abuse	0 days 00:08:40	BROOKLYN	1
35	Animal Abuse	0 days 00:08:55	NEW YORK	1
36	Animal Abuse	0 days 00:09:00	BROOKLYN	4
37	Animal Abuse	0 days 00:09:00	NEW YORK	3
38	Animal Abuse	0 days 00:09:00	STATEN ISLAND	1
39	Animal Abuse	0 days 00:09:22	BROOKLYN	1
40	Animal Abuse	0 days 00:09:27	BRONX	1

	Complaint Type	Request_Closing_Time	City	Count
41	Animal Abuse	0 days 00:09:32	BROOKLYN	1
42	Animal Abuse	0 days 00:09:37	OAKLAND GARDENS	1
43	Animal Abuse	0 days 00:09:43	RIDGEWOOD	1
44	Animal Abuse	0 days 00:10:00	BROOKLYN	1
45	Animal Abuse	0 days 00:10:00	NEW YORK	2
46	Animal Abuse	0 days 00:10:23	NEW YORK	1
47	Animal Abuse	0 days 00:10:37	BROOKLYN	1
48	Animal Abuse	0 days 00:10:43	FRESH MEADOWS	1
49	Animal Abuse	0 days 00:10:53	BROOKLYN	1

In [28]:

```
# Minimum Time to close/complete the request by city-wise data
Request_Time_City.loc[Request_Time_City.groupby('Complaint Type')['Request_Closing_Time'].idxmin()]
```

Out[28]:

	Complaint Type	Request_Closing_Time	City	Count
0	Agency Issues	0 days 01:07:53	Unknown City	1
6	Animal Abuse	0 days 00:03:53	NEW YORK	1
6799	Animal in a Park	14 days 00:50:05	QUEENS	1
6800	Bike/Roller/Skate Chronic	0 days 00:04:00	NEW YORK	1
7213	Blocked Driveway	0 days 00:02:51	STATEN ISLAND	1
59786	Derelict Vehicle	0 days 00:03:00	WOODHAVEN	1
75200	Disorderly Youth	0 days 00:06:03	BRONX	1
75484	Drinking	0 days 00:04:56	SOUTH OZONE PARK	1
76717	Graffiti	0 days 00:09:23	ROSEDALE	1
76830	Homeless Encampment	0 days 00:05:00	NEW YORK	1
80442	Illegal Fireworks	0 days 00:08:06	BROOKLYN	1
80604	Illegal Parking	0 days 00:02:37	FRESH MEADOWS	1
130664	Noise - Commercial	0 days 00:01:00	NEW YORK	1

	Complaint Type	Request_Closing_Time	City	Count
153218	Noise - House of Worship	0 days 00:04:20	BROOKLYN	1
154090	Noise - Park	0 days 00:04:17	NEW YORK	1
157549	Noise - Street/Sidewalk	0 days 00:02:00	STATEN ISLAND	1
185685	Noise - Vehicle	0 days 00:02:43	ASTORIA	1
198106	Panhandling	0 days 00:08:58	OZONE PARK	1
198404	Posting Advertisement	0 days 00:02:00	STATEN ISLAND	1
198975	Squeegee	0 days 01:10:45	NEW YORK	1
198979	Traffic	0 days 00:04:40	BROOKLYN	1
202877	Urinating in Public	0 days 00:08:00	BROOKLYN	1
203449	Vending	0 days 00:03:09	NEW YORK	1

Above table gives information of minimum time taken to resolve a complaint type according to city. for e.g. Request_Closing_Time for the complaint type Animal Abuse is 0 days 00:03:53 in the city NEW YORK

In [29]:
```
# Maximum Time to close/complete the request city-wise data
Request_Time_City.loc[Request_Time_City.groupby('Complaint Type')['Request_Closing_Time'].idxmax()]
```

Out[29]:

	Complaint Type	Request_Closing_Time	City	Count
5	Agency Issues	0 days 10:23:00	Unknown City	1
6798	Animal Abuse	21 days 15:16:01	BROOKLYN	1
6799	Animal in a Park	14 days 00:50:05	QUEENS	1
7212	Bike/Roller/Skate Chronic	1 days 09:54:52	BROOKLYN	1
59785	Blocked Driveway	6 days 04:17:00	BRONX	1
75199	Derelict Vehicle	9 days 07:22:12	BROOKLYN	1
75483	Disorderly Youth	1 days 04:03:27	BROOKLYN	1

	Complaint Type	Request_Closing_Time	City	Count
76716	Drinking	3 days 22:46:00	BRONX	1
76829	Graffiti	2 days 06:36:43	BROOKLYN	1
80441	Homeless Encampment	3 days 19:18:44	RICHMOND HILL	1
80603	Illegal Fireworks	1 days 03:51:10	BROOKLYN	1
130663	Illegal Parking	24 days 01:21:36	BROOKLYN	1
153217	Noise - Commercial	5 days 06:24:00	NEW YORK	1
154089	Noise - House of Worship	2 days 01:06:00	HOLLIS	1
157548	Noise - Park	2 days 09:41:00	BRONX	1
185684	Noise - Street/Sidewalk	24 days 16:52:22	BROOKLYN	1
198105	Noise - Vehicle	6 days 03:27:00	BRONX	1
198403	Panhandling	6 days 01:05:00	BRONX	1
198974	Posting Advertisement	1 days 01:05:13	SPRINGFIELD GARDENS	1
198978	Squeegee	0 days 06:47:12	NEW YORK	1
202876	Traffic	2 days 12:08:00	JAMAICA	1
203448	Urinating in Public	3 days 09:12:01	JAMAICA	1
206676	Vending	3 days 04:55:28	BRONX	1

Above table gives information of maximum time taken to resolve a complaint type according to city. for e.g. Request_Closing_Time for the complaint type Animal Abuse is 21 days 15:16:01 in the city BROOKLYN

5. Perform a statistical test for the following :

1. Whether the average response time across complaint types is similar or not (overall)

Null Hypothesis - The average response time across complaint types is similar

Alternate Hypothesis - The average response time across complaint types is different

In [30]:
```
# Creating a new column Time_in_Seconds which gives the total seconds for column Request_Closing_Time for analysis
NYC311['Time_in_Seconds'] = NYC311['Request_Closing_Time'].dt.total_seconds()
```

In [31]:
```
NYC311.head() # first 5 records
```

Out[31]:

	Unique Key	Created Date	Closed Date	Agency	Agency Name	Complaint Type	Descriptor	Location Type	Incident Zip	Incident Address	...	Road Ramp	Bridge Highway Segment	Garage Lot Name	Ferry Direction	Ferry Terminal Name
0	32310363	2015-12-31 23:59:45	2016-01-01 00:55:00	NYPD	New York City Police Department	Noise - Street/Sidewalk	Loud Music/Party	Street/Sidewalk	10034.0	71 VERMILYEA AVENUE	...	NaN	NaN	NaN	NaN	NaN
1	32309934	2015-12-31 23:59:44	2016-01-01 01:26:00	NYPD	New York City Police Department	Blocked Driveway	No Access	Street/Sidewalk	11105.0	27-07 23 AVENUE	...	NaN	NaN	NaN	NaN	NaN
2	32309159	2015-12-31 23:59:29	2016-01-01 04:51:00	NYPD	New York City Police Department	Blocked Driveway	No Access	Street/Sidewalk	10458.0	2897 VALENTINE AVENUE	...	NaN	NaN	NaN	NaN	NaN
3	32305098	2015-12-31 23:57:46	2016-01-01 07:43:00	NYPD	New York City Police Department	Illegal Parking	Commercial Overnight Parking	Street/Sidewalk	10461.0	2940 BAISLEY AVENUE	...	NaN	NaN	NaN	NaN	NaN
4	32306529	2015-12-31 23:56:58	2016-01-01 03:24:00	NYPD	New York City Police Department	Illegal Parking	Blocked Sidewalk	Street/Sidewalk	11373.0	87-14 57 ROAD	...	NaN	NaN	NaN	NaN	NaN

5 rows × 55 columns

In [32]:
```
# Lets create a new data with columns Complaint Type and Time in Seconds
```

```
Complaint_Time = NYC311[['Complaint Type','Time_in_Seconds']]
Complaint_Time.head() # first 5 records
```

Out[32]:

	Complaint Type	Time_in_Seconds
0	Noise - Street/Sidewalk	3315.0
1	Blocked Driveway	5176.0
2	Blocked Driveway	17491.0
3	Illegal Parking	27914.0
4	Illegal Parking	12422.0

In [33]:
```
Complaint_Time.dtypes # Check datatypes
```

Out[33]:
```
Complaint Type      object
Time_in_Seconds    float64
dtype: object
```

In [34]:
```
Complaint_Time.shape # rows and columns
```

Out[34]:
```
(300698, 2)
```

In [35]:
```
# Check null/missing values
Complaint_Time.isnull().sum()
```

Out[35]:
```
Complaint Type        0
Time_in_Seconds    2164
dtype: int64
```

In [36]:
```
# we will drop na values from data as the complaint associated with it has not closed/completed
Complaint_Time_Final = Complaint_Time.dropna()
```

In [46]:
```
# missing values are dropped
Complaint_Time_Final.isnull().sum()
```

Out[46]:
```
Complaint Type     0
Time_in_Seconds    0
dtype: int64
```

In [37]:
```
Complaint_Time_Final.shape # rows and columns
```

Out[37]:
```
(298534, 2)
```

In [38]:
```
# Calculating the average of time taken to resolve a complaint type in seconds
Complaint_Time_Final.Time_in_Seconds.groupby(Complaint_Time_Final['Complaint Type']).mean()
```

Out[38]:
```
Complaint Type
Agency Issues                1.893717e+04
Animal Abuse                 1.876768e+04
Animal in a Park             1.212605e+06
Bike/Roller/Skate Chronic    1.355926e+04
Blocked Driveway             1.706726e+04
Derelict Vehicle             2.651090e+04
Disorderly Youth             1.281090e+04
Drinking                     1.390260e+04
Graffiti                     2.574450e+04
Homeless Encampment          1.571605e+04
Illegal Fireworks            9.940101e+03
Illegal Parking              1.620415e+04
Noise - Commercial           1.132976e+04
Noise - House of Worship     1.149587e+04
Noise - Park                 1.227864e+04
Noise - Street/Sidewalk      1.240281e+04
Noise - Vehicle              1.292038e+04
Panhandling                  1.574196e+04
Posting Advertisement        7.112892e+03
Squeegee                     1.456425e+04
Traffic                      1.241525e+04
Urinating in Public          1.305599e+04
Vending                      1.445011e+04
Name: Time_in_Seconds, dtype: float64
```

In [39]:
```
# Creating separate data for each complaint type
Agency_Issues = Complaint_Time_Final[Complaint_Time_Final['Complaint Type']=='Agency Issues']
Animal_Abuse = Complaint_Time_Final[Complaint_Time_Final['Complaint Type']=='Animal Abuse']
Animal_in_a_Park = Complaint_Time_Final[Complaint_Time_Final['Complaint Type']=='Animal in a Park']
Bike_Roller_Skate_Chronic = Complaint_Time_Final[Complaint_Time_Final['Complaint Type']=='Bike/Roller/Skate Chronic']
Blocked_Driveway = Complaint_Time_Final[Complaint_Time_Final['Complaint Type']=='Blocked Driveway']
Derelict_Vehicle = Complaint_Time_Final[Complaint_Time_Final['Complaint Type']=='Derelict Vehicle']
Disorderly_Youth = Complaint_Time_Final[Complaint_Time_Final['Complaint Type']=='Disorderly Youth']
Drinking = Complaint_Time_Final[Complaint_Time_Final['Complaint Type']=='Drinking']
Graffiti = Complaint_Time_Final[Complaint_Time_Final['Complaint Type']=='Graffiti']
Homeless_Encampment = Complaint_Time_Final[Complaint_Time_Final['Complaint Type']=='Homeless Encampment']
Illegal_Fireworks = Complaint_Time_Final[Complaint_Time_Final['Complaint Type']=='Illegal Fireworks']
Illegal_Parking = Complaint_Time_Final[Complaint_Time_Final['Complaint Type']=='Illegal Parking']
```

```
Noise_Commercial = Complaint_Time_Final[Complaint_Time_Final['Complaint Type']=='Noise - Commercial']
Noise_House_of_Worship = Complaint_Time_Final[Complaint_Time_Final['Complaint Type']=='Noise - House of Worship']
Noise_Park = Complaint_Time_Final[Complaint_Time_Final['Complaint Type']=='Noise - Park']
Noise_Street_Sidewalk = Complaint_Time_Final[Complaint_Time_Final['Complaint Type']=='Noise - Street/Sidewalk']
Noise_Vehicle = Complaint_Time_Final[Complaint_Time_Final['Complaint Type']=='Noise - Vehicle']
Panhandling = Complaint_Time_Final[Complaint_Time_Final['Complaint Type']=='Panhandling']
Posting_Advertisement = Complaint_Time_Final[Complaint_Time_Final['Complaint Type']=='Posting Advertisement']
Squeegee = Complaint_Time_Final[Complaint_Time_Final['Complaint Type']=='Squeegee']
Traffic = Complaint_Time_Final[Complaint_Time_Final['Complaint Type']=='Traffic']
Urinating_in_Public = Complaint_Time_Final[Complaint_Time_Final['Complaint Type']=='Urinating in Public']
Vending = Complaint_Time_Final[Complaint_Time_Final['Complaint Type']=='Vending']
```

In [40]:

```
# Anova Single Factor
from scipy.stats import f_oneway
```

In [41]:

```
f_oneway(Agency_Issues.Time_in_Seconds,Animal_Abuse.Time_in_Seconds,Animal_in_a_Park.Time_in_Seconds,
        Bike_Roller_Skate_Chronic.Time_in_Seconds,Blocked_Driveway.Time_in_Seconds,Derelict_Vehicle.Time_in_Seconds,
        Disorderly_Youth.Time_in_Seconds,Drinking.Time_in_Seconds,Graffiti.Time_in_Seconds,Homeless_Encampment.Time_in_Seconds,
        Illegal_Fireworks.Time_in_Seconds,Illegal_Parking.Time_in_Seconds,Noise_Commercial.Time_in_Seconds,
        Noise_House_of_Worship.Time_in_Seconds,Noise_Park.Time_in_Seconds,Noise_Street_Sidewalk.Time_in_Seconds,
        Noise_Vehicle.Time_in_Seconds,Panhandling.Time_in_Seconds,Posting_Advertisement.Time_in_Seconds,
        Squeegee.Time_in_Seconds,Traffic.Time_in_Seconds,Urinating_in_Public.Time_in_Seconds,Vending.Time_in_Seconds)
```

Out[41]:

```
F_onewayResult(statistic=514.1770889253739, pvalue=0.0)
```

Here the p-value is 0, hence we reject the Null Hypothesis

So the average response time across complaint types is different

2. Are the type of complaint or service requested and location related

Null Hypothesis = The type of complaint or service requested and location is related

Alternate Hypothesis = The type of complaint or service requested and location are not related

In [42]:
```
# we will use pd.crosstab function to get a view of complaint type by city
pd.crosstab(NYC311['Complaint Type'],NYC311.City)
```

Out[42]:

City	ARVERNE	ASTORIA	Astoria	BAYSIDE	BELLEROSE	BREEZY POINT	BRONX	BROOKLYN	CAMBRIA HEIGHTS	CENTRAL PARK	...	SOUTH OZONE PARK	SOUTH RICHMOND HILL	SPRINGFIELD GARDENS	STAT ISLAI
Complaint Type															
Agency Issues	0	0	0	0	0	0	0	0	0	0	...	0	0	0	
Animal Abuse	38	125	0	37	7	2	1415	2394	11	0	...	55	26	24	5
Animal in a Park	0	0	0	0	0	0	0	0	0	0	...	0	0	0	
Bike/Roller/Skate Chronic	0	15	0	0	1	0	20	111	0	0	...	1	1	0	
Blocked Driveway	35	2618	116	377	95	3	12755	28148	147	0	...	942	1548	262	21
Derelict Vehicle	27	351	12	198	89	3	1953	5181	115	0	...	358	289	210	17
Disorderly Youth	2	3	0	1	2	0	63	72	0	0	...	2	2	0	
Drinking	1	35	0	1	1	1	188	257	0	0	...	13	23	6	1
Ferry Complaint	0	0	0	0	0	0	0	0	0	0	...	0	0	0	
Graffiti	1	4	0	3	0	0	9	43	0	0	...	0	0	0	
Homeless Encampment	4	32	0	2	1	0	247	857	5	0	...	4	11	5	
Illegal Fireworks	0	4	0	0	1	0	24	61	1	0	...	1	2	1	
Illegal Parking	58	1068	213	514	106	15	7859	27462	76	2	...	494	462	238	48
Noise - Commercial	2	1293	262	40	37	4	2434	11463	12	0	...	70	198	36	6
Noise - House of Worship	11	19	0	2	1	0	79	340	2	0	...	3	3	1	
Noise - Park	2	61	0	4	1	0	547	1555	0	0	...	4	2	1	

City	ARVERNE	ASTORIA	Astoria	BAYSIDE	BELLEROSE	BREEZY POINT	BRONX	BROOKLYN	CAMBRIA HEIGHTS	CENTRAL PARK	...	SOUTH OZONE PARK	SOUTH RICHMOND HILL	SPRINGFIELD GARDENS	STAT ISLAI
Complaint Type															
Noise - Street/Sidewalk	29	386	114	15	13	1	8892	13356	25	95	...	105	91	38	8
Noise - Vehicle	7	204	0	16	10	1	3396	5177	77	0	...	85	81	42	3
Panhandling	1	1	0	0	1	0	19	49	0	0	...	0	0	2	
Posting Advertisement	0	1	0	0	1	0	17	45	0	0	...	1	0	2	5
Squeegee	0	0	0	0	0	0	0	0	0	0	...	0	0	0	
Traffic	0	47	0	9	7	0	355	1085	6	0	...	28	11	11	2
Urinating in Public	1	9	0	0	1	0	51	136	0	0	...	2	0	3	
Vending	1	54	0	2	0	0	379	515	0	0	...	5	24	1	

24 rows × 54 columns

In [43]:
```
# Chi square test
from scipy.stats import chi2_contingency
```

In [44]:
```
chi2_contingency(pd.crosstab(NYC311['Complaint Type'],NYC311.City))
```

Out[44]:
```
(121788.33122567515,
 0.0,
 1219,
 array([[4.38978643e-03, 1.26306128e-01, 1.43067130e-02, ...,
         4.91656080e-02, 7.07154687e-02, 2.39442896e-03],
        [5.69062648e+00, 1.63734844e+02, 1.85462690e+01, ...,
         6.37350165e+01, 9.16708192e+01, 3.10397808e+00],
        [7.31631072e-04, 2.10510213e-02, 2.38445217e-03, ...,
         8.19426800e-03, 1.17859114e-02, 3.99071494e-04],
        ...,
        [3.29087656e+00, 9.46874938e+01, 1.07252659e+01, ...,
         3.68578175e+01, 5.30130297e+01, 1.79502358e+00],
        [4.33125594e-01, 1.24622046e+01, 1.41159569e+00, ...,
```

```
 4.85100666e+00, 6.97725958e+00, 2.36250324e-01],
[2.78166133e+00, 8.00359829e+01, 9.06568717e+00, ...,
 3.11546069e+01, 4.48100353e+01, 1.51726982e+00]]))
```

Here the p-value is 0, hence we reject the Null Hypothesis

So the type of complaint or service requested and location are not related

Summary :

1. We created a new column ‘Request_Closing_Time’ as the time elapsed between request creation and request closing

2. The maximum number of complaints are of Blocked Driveway

3. The highest number of complaints are from BROOKLYN and lowest are from Howard Beach

4. The maximum number of complaints are for Loud Music/Party

5. We created a table that gives information on which complaint is maximum by city-wise and its count. For e.g. Animal Abuse is maximum in BROOKLYN with count = 2394, Bike/Roller/Skate Chronic is maximum in NEW YORK with count = 225 and so on

6. We created a table that gives information like Complaint Type - Drinking with Descriptor - After Hours - Licensed Est is maximum in the city NEW YORK with count = 28, Complaint Type - Illegal Parking with Descriptor - Blocked Hydrant is maximum in the city BROOKLYN with count = 6697

7. We created a table that gives information of minimum time taken to resolve a complaint type according to city. for e.g. Request_Closing_Time for the complaint type Animal Abuse is 0 days 00:03:53 in the city NEW YORK

8. We created a table that gives information of maximum time taken to resolve a complaint type according to city. for e.g. Request_Closing_Time for the complaint type Animal Abuse is 21 days 15:16:01 in the city BROOKLYN

9. We performed statistical tests that shows : 1. the average response time across complaint types is different, 2. the type of complaint or service requested and location are not related