```
import numpy as np
import pandas as pd
import matplotlib.pyplot as plt
import datetime
```

```
dataset = pd.read_csv('311_Service_Requests_from_2010_to_Present.csv')
dataset
```

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

	Unique Key	Created Date	Closed Date	Agency	\
0	32310363	12/31/2015 11:59:45 PM	01-01-16 0:55	NYPD	
1	32309934	12/31/2015 11:59:44 PM	01-01-16 1:26	NYPD	
2	32309159	12/31/2015 11:59:29 PM	01-01-16 4:51	NYPD	
3	32305098	12/31/2015 11:57:46 PM	01-01-16 7:43	NYPD	
4	32306529	12/31/2015 11:56:58 PM	01-01-16 3:24	NYPD	
...	...	...	...	...	
300693	30281872	03/29/2015 12:33:41 AM	NaN	NYPD	
300694	30281230	03/29/2015 12:33:28 AM	03/29/2015 02:33:59 AM	NYPD	
300695	30283424	03/29/2015 12:33:03 AM	03/29/2015 03:40:20 AM	NYPD	
300696	30280004	03/29/2015 12:33:02 AM	03/29/2015 04:38:35 AM	NYPD	
300697	30281825	03/29/2015 12:33:01 AM	03/29/2015 04:41:50 AM	NYPD	

	Agency Name	Complaint Type	\
0	New York City Police Department	Noise - Street/Sidewalk	
1	New York City Police Department	Blocked Driveway	
2	New York City Police Department	Blocked Driveway	
3	New York City Police Department	Illegal Parking	
4	New York City Police Department	Illegal Parking	
...	...	...	
300693	New York City Police Department	Noise - Commercial	
300694	New York City Police Department	Blocked Driveway	
300695	New York City Police Department	Noise - Commercial	
300696	New York City Police Department	Noise - Commercial	
300697	New York City Police Department	Noise - Commercial	

```
                         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
...                               ...                  ...           ..
.
300693             Loud Music/Party  Club/Bar/Restaurant
NaN
300694               Partial Access      Street/Sidewalk
11418.0
300695             Loud Music/Party  Club/Bar/Restaurant
11206.0
300696             Loud Music/Party  Club/Bar/Restaurant
10461.0
300697             Loud Music/Party     Store/Commercial
10036.0

               Incident Address  ... Bridge Highway Name  \
0            71 VERMILYEA AVENUE  ...                 NaN
1                27-07 23 AVENUE  ...                 NaN
2           2897 VALENTINE AVENUE  ...                 NaN
3             2940 BAISLEY AVENUE  ...                 NaN
4                   87-14 57 ROAD  ...                 NaN
...                           ...  ...                 ...
300693            CRESCENT AVENUE  ...                 NaN
300694           100-17 87 AVENUE  ...                 NaN
300695          162 THROOP AVENUE  ...                 NaN
300696  3151 EAST TREMONT AVENUE  ...                 NaN
300697         251 WEST 48 STREET  ...                 NaN

       Bridge Highway Direction Road Ramp Bridge Highway Segment  \
0                           NaN       NaN                    NaN
1                           NaN       NaN                    NaN
2                           NaN       NaN                    NaN
3                           NaN       NaN                    NaN
4                           NaN       NaN                    NaN
...                         ...       ...                    ...
300693                      NaN       NaN                    NaN
300694                      NaN       NaN                    NaN
300695                      NaN       NaN                    NaN
300696                      NaN       NaN                    NaN
300697                      NaN       NaN                    NaN
```

```
       Garage Lot Name Ferry Direction Ferry Terminal Name   Latitude
\
0                  NaN            NaN                  NaN  40.865682

1                  NaN            NaN                  NaN  40.775945

2                  NaN            NaN                  NaN  40.870325

3                  NaN            NaN                  NaN  40.835994

4                  NaN            NaN                  NaN  40.733060

...                ...            ...                  ...        ...

300693             NaN            NaN                  NaN        NaN

300694             NaN            NaN                  NaN  40.694077

300695             NaN            NaN                  NaN  40.699590

300696             NaN            NaN                  NaN  40.837708

300697             NaN            NaN                  NaN  40.760583


        Longitude                                   Location
0      -73.923501   (40.86568153633767, -73.92350095571744)
1      -73.915094  (40.775945312321085, -73.91509393898605)
2      -73.888525  (40.870324522111424, -73.88852464418646)
3      -73.828379   (40.83599404683083, -73.82837939584206)
4      -73.874170  (40.733059618956815, -73.87416975810375)
...           ...                                       ...
300693        NaN                                       NaN
300694 -73.846087    (40.69407728322387, -73.8460866160573)
300695 -73.944234   (40.69959035300927, -73.94423377144169)
300696 -73.834587    (40.8377075854206, -73.83458731019586)
300697 -73.985922   (40.76058322950115, -73.98592204392392)

[300698 rows x 53 columns]

dataset.head()

   Unique Key            Created Date    Closed Date Agency  \
0    32310363  12/31/2015 11:59:45 PM  01-01-16 0:55   NYPD
1    32309934  12/31/2015 11:59:44 PM  01-01-16 1:26   NYPD
2    32309159  12/31/2015 11:59:29 PM  01-01-16 4:51   NYPD
3    32305098  12/31/2015 11:57:46 PM  01-01-16 7:43   NYPD
4    32306529  12/31/2015 11:56:58 PM  01-01-16 3:24   NYPD
```

```
                        Agency Name           Complaint Type  \
0  New York City Police Department  Noise - Street/Sidewalk
1  New York City Police Department         Blocked Driveway
2  New York City Police Department         Blocked Driveway
3  New York City Police Department          Illegal Parking
4  New York City Police Department          Illegal Parking

                     Descriptor    Location Type  Incident Zip  \
0             Loud Music/Party  Street/Sidewalk       10034.0
1                    No Access  Street/Sidewalk       11105.0
2                    No Access  Street/Sidewalk       10458.0
3  Commercial Overnight Parking  Street/Sidewalk       10461.0
4             Blocked Sidewalk  Street/Sidewalk       11373.0

        Incident Address  ... Bridge Highway Name Bridge Highway
Direction  \
0    71 VERMILYEA AVENUE  ...                 NaN
NaN
1        27-07 23 AVENUE  ...                 NaN
NaN
2  2897 VALENTINE AVENUE  ...                 NaN
NaN
3    2940 BAISLEY AVENUE  ...                 NaN
NaN
4          87-14 57 ROAD  ...                 NaN
NaN

  Road Ramp Bridge Highway Segment Garage Lot Name Ferry Direction  \
0       NaN                     NaN             NaN             NaN
1       NaN                     NaN             NaN             NaN
2       NaN                     NaN             NaN             NaN
3       NaN                     NaN             NaN             NaN
4       NaN                     NaN             NaN             NaN

  Ferry Terminal Name   Latitude  Longitude  \
0                 NaN  40.865682 -73.923501
1                 NaN  40.775945 -73.915094
2                 NaN  40.870325 -73.888525
3                 NaN  40.835994 -73.828379
4                 NaN  40.733060 -73.874170

                                  Location
0   (40.86568153633767, -73.92350095571744)
1  (40.775945312321085, -73.91509393898605)
2  (40.870324522111424, -73.88852464418646)
3   (40.83599404683083, -73.82837939584206)
4  (40.733059618956815, -73.87416975810375)
```

[5 rows x 53 columns]

```
dataset.shape
```

```
(300698, 53)
```

```
dataset.info()
```

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 300698 entries, 0 to 300697
Data columns (total 53 columns):
 #   Column                          Non-Null Count   Dtype  
---  ------                          --------------   -----  
 0   Unique Key                      300698 non-null  int64  
 1   Created Date                    300698 non-null  object 
 2   Closed Date                     298534 non-null  object 
 3   Agency                          300698 non-null  object 
 4   Agency Name                     300698 non-null  object 
 5   Complaint Type                  300698 non-null  object 
 6   Descriptor                      294784 non-null  object 
 7   Location Type                   300567 non-null  object 
 8   Incident Zip                    298083 non-null  float64
 9   Incident Address                256288 non-null  object 
 10  Street Name                     256288 non-null  object 
 11  Cross Street 1                  251419 non-null  object 
 12  Cross Street 2                  250919 non-null  object 
 13  Intersection Street 1           43858 non-null   object 
 14  Intersection Street 2           43362 non-null   object 
 15  Address Type                    297883 non-null  object 
 16  City                            298084 non-null  object 
 17  Landmark                        349 non-null     object 
 18  Facility Type                   298527 non-null  object 
 19  Status                          300698 non-null  object 
 20  Due Date                        300695 non-null  object 
 21  Resolution Description          300698 non-null  object 
 22  Resolution Action Updated Date  298511 non-null  object 
 23  Community Board                 300698 non-null  object 
 24  Borough                         300698 non-null  object 
 25  X Coordinate (State Plane)      297158 non-null  float64
 26  Y Coordinate (State Plane)      297158 non-null  float64
 27  Park Facility Name              300698 non-null  object 
 28  Park Borough                    300698 non-null  object 
 29  School Name                     300698 non-null  object 
 30  School Number                   300698 non-null  object 
 31  School Region                   300697 non-null  object 
 32  School Code                     300697 non-null  object 
 33  School Phone Number             300698 non-null  object 
 34  School Address                  300698 non-null  object 
 35  School City                     300698 non-null  object 
 36  School State                    300698 non-null  object 
```

```
 37  School Zip                      300697 non-null  object
 38  School Not Found                300698 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             243 non-null     object
 44  Bridge Highway Direction        243 non-null     object
 45  Road Ramp                       213 non-null     object
 46  Bridge Highway Segment          213 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                        297158 non-null  float64
 51  Longitude                       297158 non-null  float64
 52  Location                        297158 non-null  object
dtypes: float64(10), int64(1), object(42)
memory usage: 121.6+ MB

dataset.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')

#dropping null,unnecessary,same type of columns
drop_columns = ['Agency Name','Incident Address','Street Name','Cross
Street 1','Cross Street 2','Intersection Street 1',
'Intersection Street 2','Address Type','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','Landmark',
'X Coordinate (State Plane)','Y Coordinate (State Plane)','Due
Date','Resolution Action Updated Date','Community Board','Facility
Type',
'Location']
dataset = dataset.drop(drop_columns,axis=1)

dataset
```

```
        Unique Key            Created Date              Closed Date
Agency  \
0         32310363  12/31/2015 11:59:45 PM           01-01-16 0:55
NYPD
1         32309934  12/31/2015 11:59:44 PM           01-01-16 1:26
NYPD
2         32309159  12/31/2015 11:59:29 PM           01-01-16 4:51
NYPD
3         32305098  12/31/2015 11:57:46 PM           01-01-16 7:43
NYPD
4         32306529  12/31/2015 11:56:58 PM           01-01-16 3:24
NYPD
...            ...                     ...                    ...
...
300693    30281872  03/29/2015 12:33:41 AM                    NaN
NYPD
300694    30281230  03/29/2015 12:33:28 AM  03/29/2015 02:33:59 AM
NYPD
300695    30283424  03/29/2015 12:33:03 AM  03/29/2015 03:40:20 AM
NYPD
300696    30280004  03/29/2015 12:33:02 AM  03/29/2015 04:38:35 AM
NYPD
300697    30281825  03/29/2015 12:33:01 AM  03/29/2015 04:41:50 AM
NYPD

                 Complaint Type                   Descriptor  \
0       Noise - Street/Sidewalk              Loud Music/Party
1              Blocked Driveway                     No Access
2              Blocked Driveway                     No Access
3               Illegal Parking  Commercial Overnight Parking
4               Illegal Parking              Blocked Sidewalk
...                         ...                           ...
300693       Noise - Commercial              Loud Music/Party
300694         Blocked Driveway                Partial Access
```

```
300695       Noise - Commercial             Loud Music/Party
300696       Noise - Commercial             Loud Music/Party
300697       Noise - Commercial             Loud Music/Party

              Location Type  Incident Zip           City  Status  \
0           Street/Sidewalk       10034.0       NEW YORK  Closed
1           Street/Sidewalk       11105.0        ASTORIA  Closed
2           Street/Sidewalk       10458.0          BRONX  Closed
3           Street/Sidewalk       10461.0          BRONX  Closed
4           Street/Sidewalk       11373.0       ELMHURST  Closed
...                     ...           ...            ...     ...
300693  Club/Bar/Restaurant           NaN            NaN    Open
300694      Street/Sidewalk       11418.0  RICHMOND HILL  Closed
300695  Club/Bar/Restaurant       11206.0       BROOKLYN  Closed
300696  Club/Bar/Restaurant       10461.0          BRONX  Closed
300697     Store/Commercial       10036.0       NEW YORK  Closed

                                    Resolution Description      Borough
\
0       The Police Department responded and upon arriv...    MANHATTAN

1       The Police Department responded to the complai...       QUEENS

2       The Police Department responded and upon arriv...        BRONX

3       The Police Department responded to the complai...        BRONX

4       The Police Department responded and upon arriv...       QUEENS

...                                                   ...          ...

300693  Your complaint has been forwarded to the New Y...  Unspecified

300694  The Police Department responded and upon arriv...       QUEENS

300695  The Police Department responded to the complai...     BROOKLYN

300696  The Police Department responded to the complai...        BRONX

300697  The Police Department responded to the complai...    MANHATTAN


         Latitude  Longitude
0       40.865682 -73.923501
1       40.775945 -73.915094
2       40.870325 -73.888525
3       40.835994 -73.828379
4       40.733060 -73.874170
...           ...        ...
```

```
300693        NaN        NaN
300694  40.694077 -73.846087
300695  40.699590 -73.944234
300696  40.837708 -73.834587
300697  40.760583 -73.985922

[300698 rows x 14 columns]
```

```
dataset.shape
```

```
(300698, 14)
```

```
dataset.isnull().sum()
```

```
Unique Key                   0
Created Date                 0
Closed Date               2164
Agency                       0
Complaint Type               0
Descriptor                5914
Location Type              131
Incident Zip              2615
City                      2614
Status                       0
Resolution Description       0
Borough                      0
Latitude                  3540
Longitude                 3540
dtype: int64
```

```
# only work on closed cases
dataset = dataset[dataset['Status']== 'Closed']
```

```
dataset
```

```
        Unique Key            Created Date              Closed Date
Agency  \
0         32310363  12/31/2015 11:59:45 PM          01-01-16 0:55
NYPD
1         32309934  12/31/2015 11:59:44 PM          01-01-16 1:26
NYPD
2         32309159  12/31/2015 11:59:29 PM          01-01-16 4:51
NYPD
3         32305098  12/31/2015 11:57:46 PM          01-01-16 7:43
NYPD
4         32306529  12/31/2015 11:56:58 PM          01-01-16 3:24
NYPD
...            ...                      ...                   ...
...
300692    30281370  03/29/2015 12:34:32 AM  03/29/2015 01:13:01 AM
NYPD
300694    30281230  03/29/2015 12:33:28 AM  03/29/2015 02:33:59 AM
```

```
NYPD
300695    30283424  03/29/2015 12:33:03 AM  03/29/2015 03:40:20 AM
NYPD
300696    30280004  03/29/2015 12:33:02 AM  03/29/2015 04:38:35 AM
NYPD
300697    30281825  03/29/2015 12:33:01 AM  03/29/2015 04:41:50 AM
NYPD

                Complaint Type                   Descriptor  \
0       Noise - Street/Sidewalk              Loud Music/Party
1              Blocked Driveway                     No Access
2              Blocked Driveway                     No Access
3               Illegal Parking  Commercial Overnight Parking
4               Illegal Parking               Blocked Sidewalk
...                         ...                            ...
300692       Noise - Commercial              Loud Music/Party
300694         Blocked Driveway                Partial Access
300695       Noise - Commercial              Loud Music/Party
300696       Noise - Commercial              Loud Music/Party
300697       Noise - Commercial              Loud Music/Party

              Location Type  Incident Zip           City  Status  \
0           Street/Sidewalk       10034.0       NEW YORK  Closed
1           Street/Sidewalk       11105.0        ASTORIA  Closed
2           Street/Sidewalk       10458.0          BRONX  Closed
3           Street/Sidewalk       10461.0          BRONX  Closed
4           Street/Sidewalk       11373.0       ELMHURST  Closed
...                     ...           ...            ...     ...
300692     Store/Commercial       10002.0       NEW YORK  Closed
300694      Street/Sidewalk       11418.0  RICHMOND HILL  Closed
300695  Club/Bar/Restaurant       11206.0       BROOKLYN  Closed
300696  Club/Bar/Restaurant       10461.0          BRONX  Closed
300697     Store/Commercial       10036.0       NEW YORK  Closed

                                    Resolution Description
Borough  \
0       The Police Department responded and upon arriv...  MANHATTAN

1       The Police Department responded to the complai...     QUEENS

2       The Police Department responded and upon arriv...      BRONX

3       The Police Department responded to the complai...      BRONX

4       The Police Department responded and upon arriv...     QUEENS

...                                                   ...        ...

300692  The Police Department responded to the complai...  MANHATTAN
```

```
300694  The Police Department responded and upon arriv...     QUEENS

300695  The Police Department responded to the complai...   BROOKLYN

300696  The Police Department responded to the complai...      BRONX

300697  The Police Department responded to the complai...  MANHATTAN


         Latitude  Longitude
0       40.865682 -73.923501
1       40.775945 -73.915094
2       40.870325 -73.888525
3       40.835994 -73.828379
4       40.733060 -73.874170
...           ...        ...
300692  40.716053 -73.991378
300694  40.694077 -73.846087
300695  40.699590 -73.944234
300696  40.837708 -73.834587
300697  40.760583 -73.985922

[298471 rows x 14 columns]
```

```
dataset.isnull().sum()
```

```
Unique Key                  0
Created Date                0
Closed Date                 0
Agency                      0
Complaint Type              0
Descriptor               5903
Location Type              65
Incident Zip              507
City                      506
Status                      0
Resolution Description      0
Borough                     0
Latitude                 1432
Longitude                1432
dtype: int64
```

```
dataset = dataset.drop('Status',axis=1)
dataset
```

```
        Unique Key            Created Date             Closed Date
Agency  \
0         32310363  12/31/2015 11:59:45 PM          01-01-16 0:55
NYPD
```

```
1         32309934  12/31/2015 11:59:44 PM           01-01-16 1:26
NYPD
2         32309159  12/31/2015 11:59:29 PM           01-01-16 4:51
NYPD
3         32305098  12/31/2015 11:57:46 PM           01-01-16 7:43
NYPD
4         32306529  12/31/2015 11:56:58 PM           01-01-16 3:24
NYPD
...            ...                     ...                     ...
...
300692    30281370  03/29/2015 12:34:32 AM  03/29/2015 01:13:01 AM
NYPD
300694    30281230  03/29/2015 12:33:28 AM  03/29/2015 02:33:59 AM
NYPD
300695    30283424  03/29/2015 12:33:03 AM  03/29/2015 03:40:20 AM
NYPD
300696    30280004  03/29/2015 12:33:02 AM  03/29/2015 04:38:35 AM
NYPD
300697    30281825  03/29/2015 12:33:01 AM  03/29/2015 04:41:50 AM
NYPD

                Complaint Type                   Descriptor  \
0       Noise - Street/Sidewalk              Loud Music/Party
1              Blocked Driveway                     No Access
2              Blocked Driveway                     No Access
3               Illegal Parking  Commercial Overnight Parking
4               Illegal Parking               Blocked Sidewalk
...                         ...                            ...
300692       Noise - Commercial              Loud Music/Party
300694         Blocked Driveway                Partial Access
300695       Noise - Commercial              Loud Music/Party
300696       Noise - Commercial              Loud Music/Party
300697       Noise - Commercial              Loud Music/Party

              Location Type  Incident Zip           City  \
0           Street/Sidewalk       10034.0       NEW YORK
1           Street/Sidewalk       11105.0        ASTORIA
2           Street/Sidewalk       10458.0          BRONX
3           Street/Sidewalk       10461.0          BRONX
4           Street/Sidewalk       11373.0       ELMHURST
...                     ...           ...            ...
300692     Store/Commercial       10002.0       NEW YORK
300694      Street/Sidewalk       11418.0  RICHMOND HILL
300695  Club/Bar/Restaurant       11206.0       BROOKLYN
300696  Club/Bar/Restaurant       10461.0          BRONX
300697     Store/Commercial       10036.0       NEW YORK

                                  Resolution Description
Borough  \
0       The Police Department responded and upon arriv...  MANHATTAN
```

```
1       The Police Department responded to the complai...     QUEENS

2       The Police Department responded and upon arriv...      BRONX

3       The Police Department responded to the complai...      BRONX

4       The Police Department responded and upon arriv...     QUEENS

...                                                   ...        ...

300692  The Police Department responded to the complai...  MANHATTAN

300694  The Police Department responded and upon arriv...     QUEENS

300695  The Police Department responded to the complai...   BROOKLYN

300696  The Police Department responded to the complai...      BRONX

300697  The Police Department responded to the complai...  MANHATTAN


         Latitude  Longitude
0       40.865682 -73.923501
1       40.775945 -73.915094
2       40.870325 -73.888525
3       40.835994 -73.828379
4       40.733060 -73.874170
...           ...        ...
300692  40.716053 -73.991378
300694  40.694077 -73.846087
300695  40.699590 -73.944234
300696  40.837708 -73.834587
300697  40.760583 -73.985922

[298471 rows x 13 columns]

dataset.shape

(298471, 13)

dataset.isnull().sum()

Unique Key                   0
Created Date                 0
Closed Date                  0
Agency                       0
Complaint Type               0
Descriptor                5903
Location Type               65
```

```
Incident Zip               507
City                       506
Resolution Description       0
Borough                      0
Latitude                  1432
Longitude                 1432
dtype: int64
```

```
#removing all null values
dataset.dropna(subset=['Descriptor','Location Type','Incident
Zip','City','Latitude','Longitude'],inplace = True)
```

```
dataset.info()
```

```
<class 'pandas.core.frame.DataFrame'>
Int64Index: 291107 entries, 0 to 300697
Data columns (total 13 columns):
 #   Column                  Non-Null Count   Dtype
---  ------                  --------------   -----
 0   Unique Key              291107 non-null  int64
 1   Created Date            291107 non-null  object
 2   Closed Date             291107 non-null  object
 3   Agency                  291107 non-null  object
 4   Complaint Type          291107 non-null  object
 5   Descriptor              291107 non-null  object
 6   Location Type           291107 non-null  object
 7   Incident Zip            291107 non-null  float64
 8   City                    291107 non-null  object
 9   Resolution Description  291107 non-null  object
 10  Borough                 291107 non-null  object
 11  Latitude                291107 non-null  float64
 12  Longitude               291107 non-null  float64
dtypes: float64(3), int64(1), object(9)
memory usage: 31.1+ MB
```

```
dataset.shape
```

```
(291107, 13)
```

Converting date columns to date type and creating new column

```
dataset.dtypes
```

```
Unique Key                 int64
Created Date              object
Closed Date               object
Agency                    object
Complaint Type            object
Descriptor                object
Location Type             object
Incident Zip             float64
City                      object
Resolution Description    object
```

```
Borough                    object
Latitude                  float64
Longitude                 float64
dtype: object
```

```
cols = ['Created Date','Closed Date']
for col in cols:
 dataset[col] =
pd.to_datetime(dataset[col],infer_datetime_format=True)
```

```
dataset.dtypes
```

```
Unique Key                          int64
Created Date              datetime64[ns]
Closed Date               datetime64[ns]
Agency                            object
Complaint Type                    object
Descriptor                        object
Location Type                     object
Incident Zip                     float64
City                              object
Resolution Description            object
Borough                           object
Latitude                         float64
Longitude                        float64
dtype: object
```

```
#creating new column
dataset['Request_Closing_Time'] = dataset[cols[1]] - dataset[cols[0]]
```

```
dataset.info()
```

```
<class 'pandas.core.frame.DataFrame'>
Int64Index: 291107 entries, 0 to 300697
Data columns (total 14 columns):
 #   Column                  Non-Null Count   Dtype
---  ------                  --------------   -----
 0   Unique Key              291107 non-null  int64
 1   Created Date            291107 non-null  datetime64[ns]
 2   Closed Date             291107 non-null  datetime64[ns]
 3   Agency                  291107 non-null  object
 4   Complaint Type          291107 non-null  object
 5   Descriptor              291107 non-null  object
 6   Location Type           291107 non-null  object
 7   Incident Zip            291107 non-null  float64
 8   City                    291107 non-null  object
 9   Resolution Description  291107 non-null  object
 10  Borough                 291107 non-null  object
 11  Latitude                291107 non-null  float64
 12  Longitude               291107 non-null  float64
 13  Request_Closing_Time    291107 non-null  timedelta64[ns]
```

```
dtypes: datetime64[ns](2), float64(3), int64(1), object(7),
timedelta64[ns](1)
memory usage: 33.3+ MB
```

```
dataset.describe()
```

```
         Unique Key   Incident Zip       Latitude      Longitude  \
count  2.911070e+05  291107.000000  291107.000000  291107.000000
mean   3.130158e+07   10857.977349      40.725681     -73.925035
std    5.753777e+05     580.280774       0.082411       0.078654
min    3.027948e+07      83.000000      40.499135     -74.254937
25%    3.079934e+07   10314.000000      40.668926     -73.970957
50%    3.130675e+07   11209.000000      40.717782     -73.930774
75%    3.179091e+07   11238.000000      40.782973     -73.875788
max    3.231065e+07   11697.000000      40.912869     -73.700760

            Request_Closing_Time
count                     291107
mean   0 days 04:18:32.132665995
std    0 days 06:03:45.509089128
min              0 days 00:01:00
25%              0 days 01:16:30
50%              0 days 02:42:38
75%              0 days 05:20:24
max             24 days 16:52:22
```

```
dataset.columns
```

```
Index(['Unique Key', 'Created Date', 'Closed Date', 'Agency',
'Complaint Type',
       'Descriptor', 'Location Type', 'Incident Zip', 'City',
       'Resolution Description', 'Borough', 'Latitude', 'Longitude',
       'Request_Closing_Time'],
      dtype='object')
```

```
dataset['Agency'].value_counts()
```

```
NYPD    291107
Name: Agency, dtype: int64
```

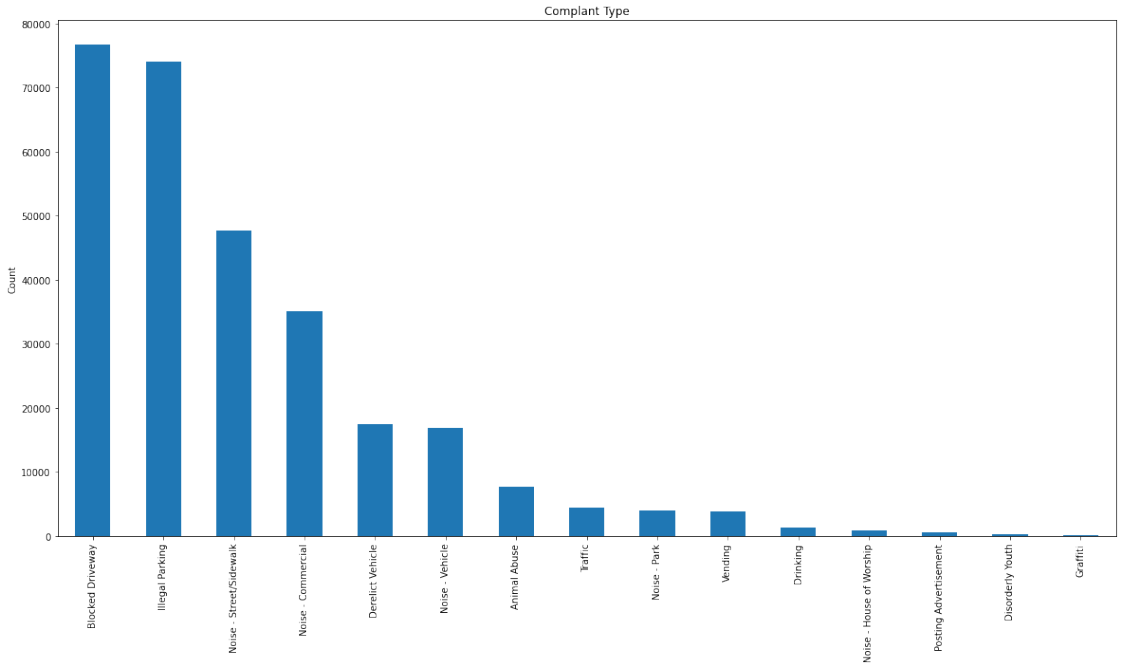
```
dataset['Complaint Type'].value_counts()
```

```
Blocked Driveway            76676
Illegal Parking             74021
Noise - Street/Sidewalk     47747
Noise - Commercial          35144
Derelict Vehicle            17506
Noise - Vehicle             16868
Animal Abuse                 7744
Traffic                      4466
Noise - Park                 3927
Vending                      3773
```

```
Drinking                      1270
Noise - House of Worship       920
Posting Advertisement          647
Disorderly Youth               285
Graffiti                       113
Name: Complaint Type, dtype: int64
```

```
dataset['Complaint Type'].value_counts().plot(kind =
'bar',figsize=(20,10),title='Complant Type',ylabel = 'Count')
```

```
<AxesSubplot:title={'center':'Complant Type'}, ylabel='Count'>
```




```
dataset['Descriptor'].value_counts()
```

```
Loud Music/Party                    60444
No Access                           56725
Posted Parking Sign Violation       22103
Loud Talking                        21254
Partial Access                      19951
With License Plate                  17506
Blocked Hydrant                     15837
Commercial Overnight Parking        11908
Car/Truck Music                     11114
Blocked Sidewalk                    10930
Double Parked Blocking Traffic       5558
Double Parked Blocking Vehicle       4147
Engine Idling                        4134
Banging/Pounding                     4090
Neglected                            3771
Car/Truck Horn                       3477
```

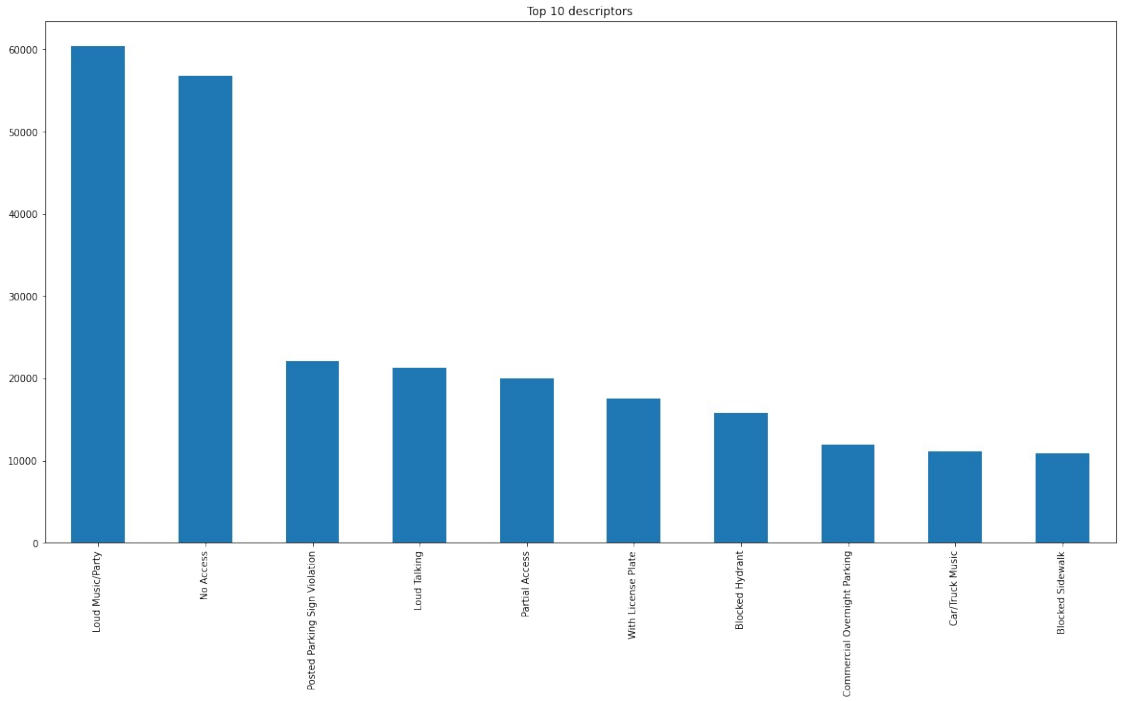
```
Congestion/Gridlock               2736
In Prohibited Area                2017
Other (complaint details)         1961
Unlicensed                        1756
Overnight Commercial Storage      1746
Unauthorized Bus Layover          1333
Truck Route Violation             1010
In Public                          923
Tortured                           849
Vehicle                            587
Chained                            534
Detached Trailer                   459
No Shelter                         381
Chronic Stoplight Violation        280
Underage - Licensed Est            270
Chronic Speeding                   266
In Car                             248
Playing in Unsuitable Place        245
Drag Racing                        174
Loud Television                     93
Police Report Requested             90
After Hours - Licensed Est          77
Building                            60
Nuisance/Truant                     40
Police Report Not Requested         23
Name: Descriptor, dtype: int64
```

```
dataset['Descriptor'].value_counts().head(10)
```

```
Loud Music/Party                 60444
No Access                        56725
Posted Parking Sign Violation    22103
Loud Talking                     21254
Partial Access                   19951
With License Plate               17506
Blocked Hydrant                  15837
Commercial Overnight Parking     11908
Car/Truck Music                  11114
Blocked Sidewalk                 10930
Name: Descriptor, dtype: int64
```

```
#plotting top 10 descriptors
dataset['Descriptor'].value_counts().head(10).plot(kind =
'bar',figsize=(20,10),title='Top 10 descriptors')
```

```
<AxesSubplot:title={'center':'Top 10 descriptors'}>
```

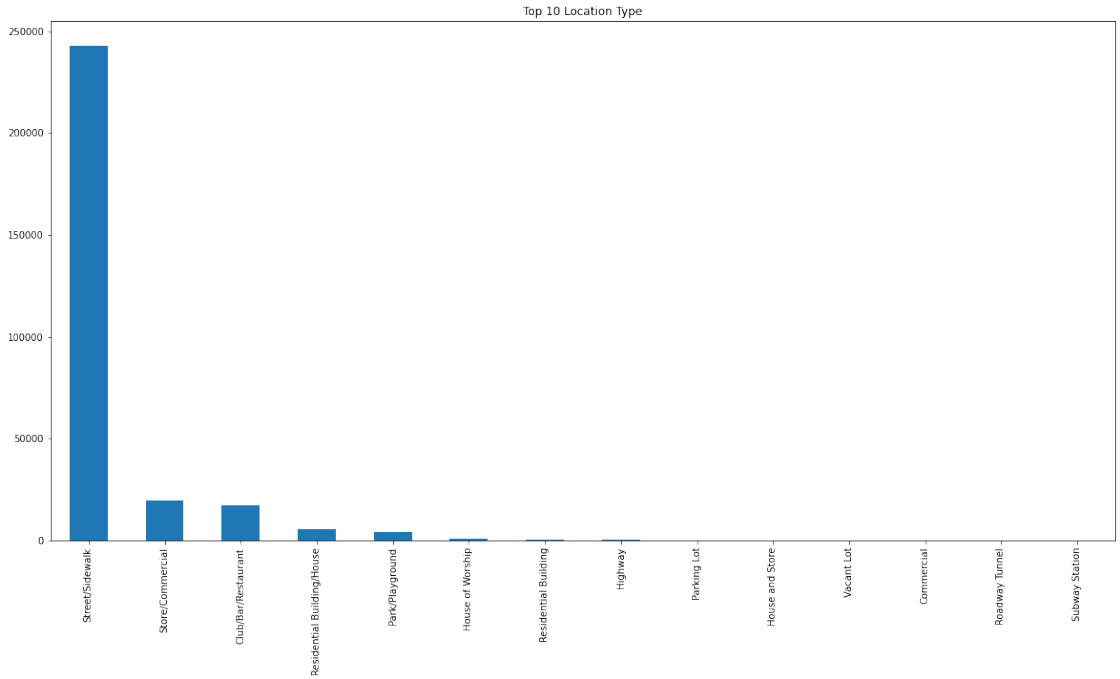



```
dataset['Location Type'].value_counts()
```

```
Street/Sidewalk                242765
Store/Commercial                19425
Club/Bar/Restaurant             17172
Residential Building/House       5753
Park/Playground                  4246
House of Worship                  920
Residential Building              226
Highway                           196
Parking Lot                       116
House and Store                    93
Vacant Lot                         77
Commercial                         62
Roadway Tunnel                     34
Subway Station                     22
Name: Location Type, dtype: int64
```

```
dataset['Location Type'].value_counts().plot(kind='bar', figsize=(20,
10), title='Top 10 Location Type')
```

```
<AxesSubplot:title={'center':'Top 10 Location Type'}>
```



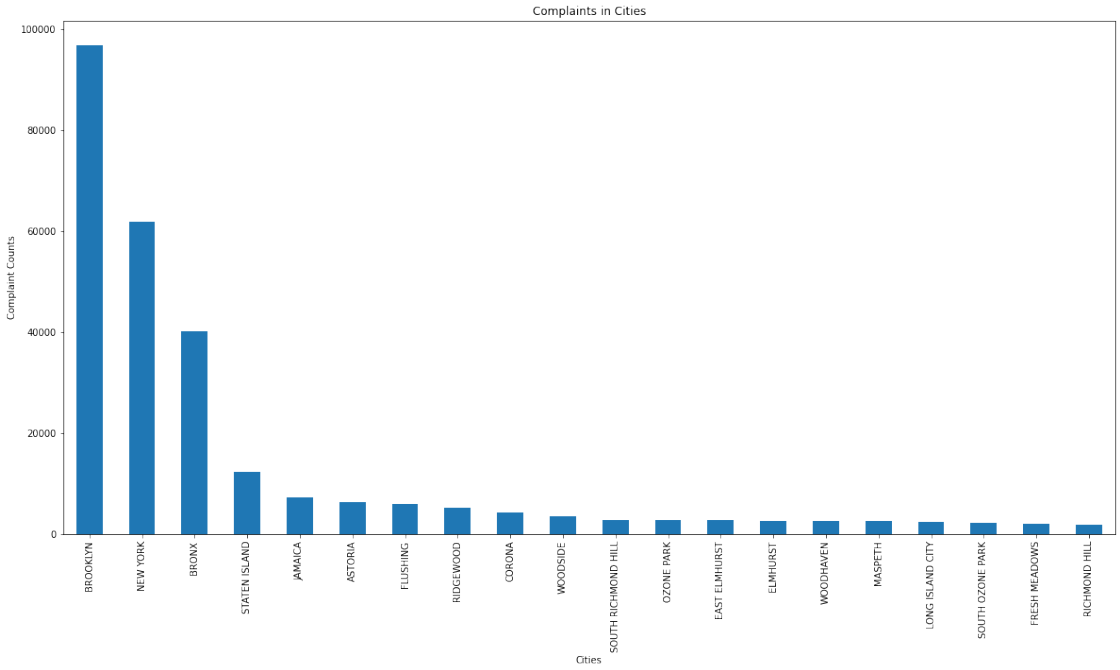

```
dataset['City'].value_counts()
```

```
BROOKLYN               96858
NEW YORK               61935
BRONX                  40216
STATEN ISLAND          12211
JAMAICA                 7155
ASTORIA                 6254
FLUSHING                5916
RIDGEWOOD               5124
CORONA                  4265
WOODSIDE                3493
SOUTH RICHMOND HILL     2759
OZONE PARK              2733
EAST ELMHURST           2724
ELMHURST                2624
WOODHAVEN               2448
MASPETH                 2445
LONG ISLAND CITY        2408
SOUTH OZONE PARK        2165
FRESH MEADOWS           1886
RICHMOND HILL           1865
QUEENS VILLAGE          1788
MIDDLE VILLAGE          1759
JACKSON HEIGHTS         1671
FOREST HILLS            1655
REGO PARK               1477
COLLEGE POINT           1216
BAYSIDE                 1214
FAR ROCKAWAY            1163
```

```
WHITESTONE              1093
HOLLIS                   998
HOWARD BEACH             921
ROSEDALE                 913
SPRINGFIELD GARDENS      871
SAINT ALBANS             825
KEW GARDENS              763
ROCKAWAY PARK            738
Astoria                  716
SUNNYSIDE                708
LITTLE NECK              558
OAKLAND GARDENS          546
CAMBRIA HEIGHTS          471
BELLEROSE                369
GLEN OAKS                304
ARVERNE                  214
FLORAL PARK              152
Long Island City         134
Woodside                 120
CENTRAL PARK              97
NEW HYDE PARK             96
BREEZY POINT              30
QUEENS                    28
East Elmhurst             14
Howard Beach               1
Name: City, dtype: int64
```

```
dataset['City'].value_counts().head(20).plot(kind='bar',figsize=(20,10
), title='Complaints in Cities', ylabel='Complaint Counts')
plt.xlabel('Cities')
```

```
Text(0.5, 0, 'Cities')
```

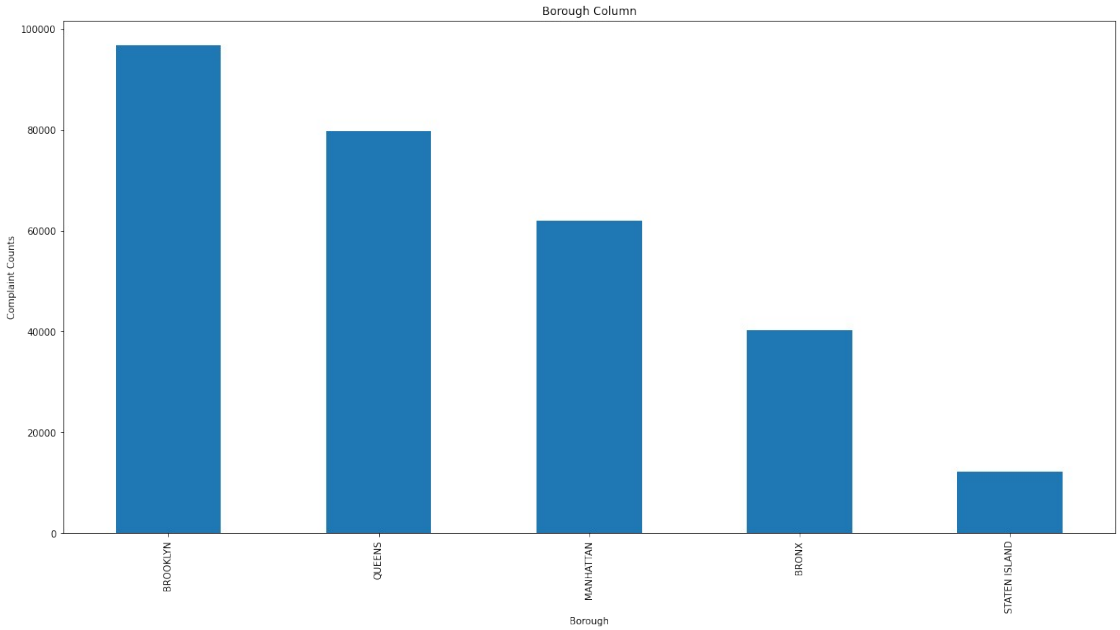



```
dataset['Borough'].value_counts()
```

```
BROOKLYN         96858
QUEENS           79790
MANHATTAN        62033
BRONX            40217
STATEN ISLAND    12209
Name: Borough, dtype: int64
```

```
dataset['Borough'].value_counts().plot(kind='bar',figsize=(20,10) ,tit
le='Borough Column', ylabel='Complaint Counts')
plt.xlabel('Borough')
```

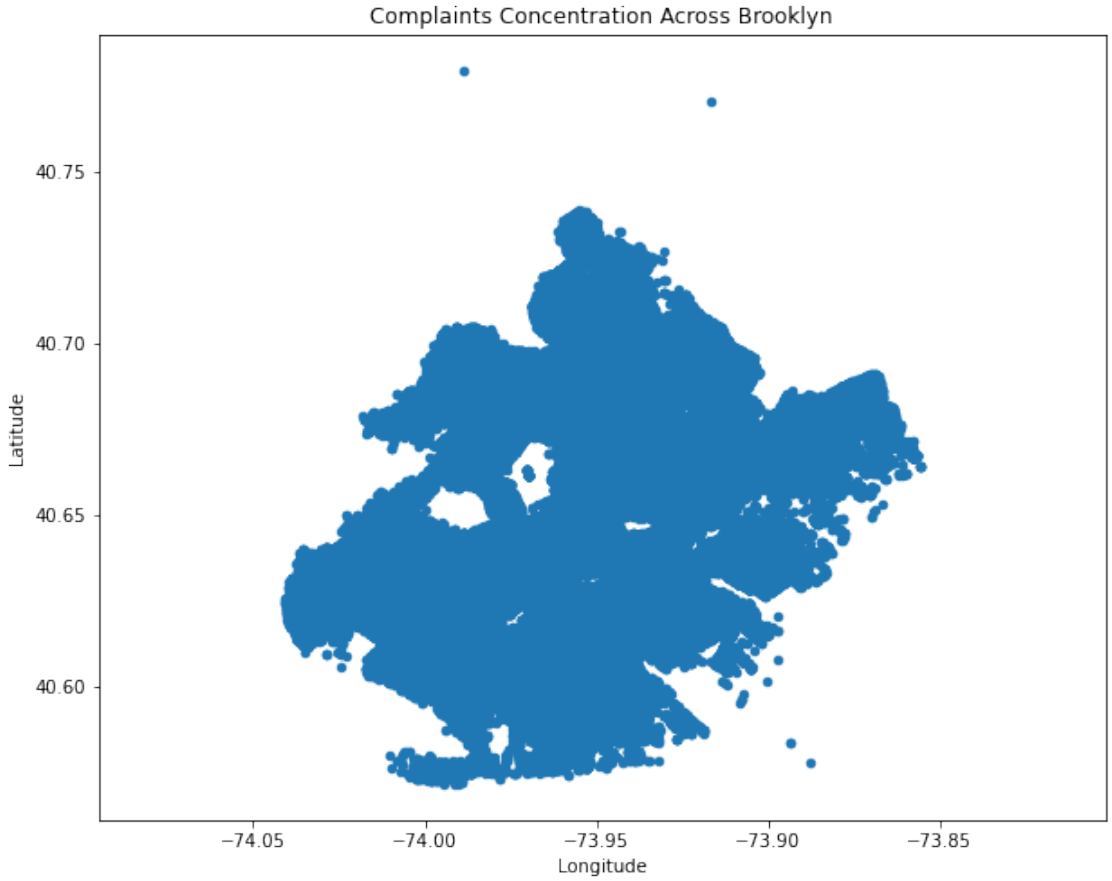
```
Text(0.5, 0, 'Borough')
```



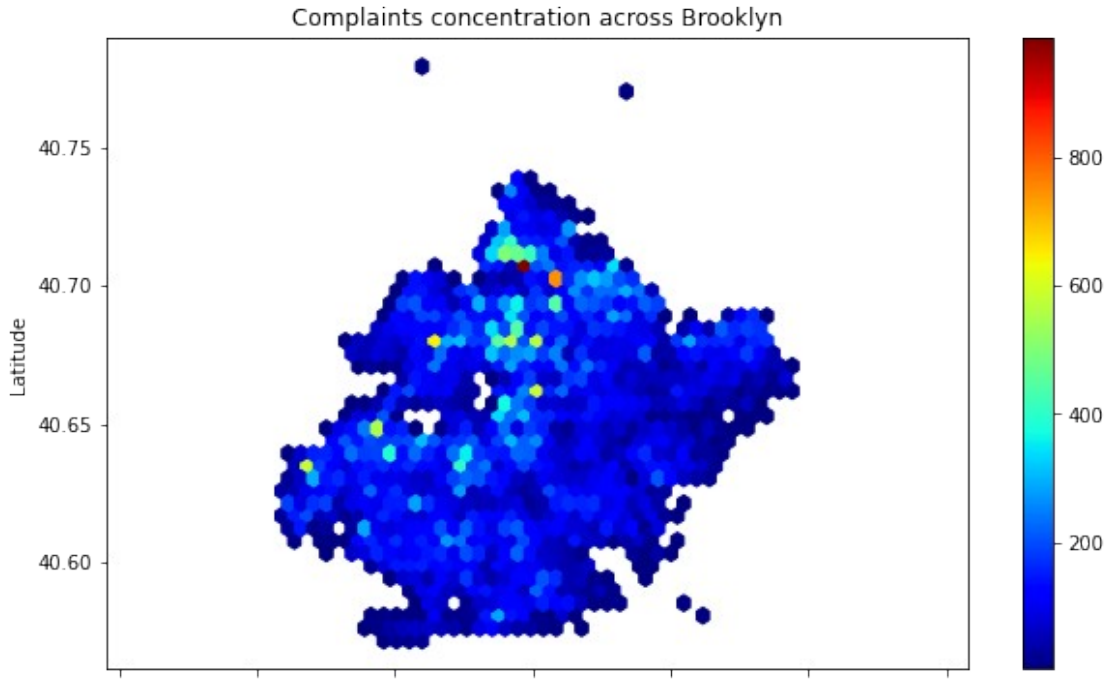

```
brooklyn_data = dataset[dataset['City']=='BROOKLYN']

brooklyn_data[['Longitude', 'Latitude']].plot(kind='scatter',
                                              x='Longitude',
                                              y='Latitude',
                                              figsize=(10,8),
                                              title = 'Complaints
Concentration Across Brooklyn'
                                             ).axis('equal')

plt.xlabel('Longitude')
plt.ylabel('Latitude')
plt.show()
```




```
brooklyn_data.plot(kind='hexbin',
                   x='Longitude',
                   y='Latitude',
                   gridsize=40,
                   colormap = 'jet',
                   mincnt=1,
                   title = 'Complaints concentration across Brooklyn',
                   figsize=(10,6)
                  ).axis('equal')
plt.xlabel('Longitude')
plt.ylabel('Latitude')
plt.show()
```




```
#borough on the basis of complaints
top_5_complaints = dataset['Complaint Type'].value_counts()[:5].keys()
top_5_complaints
```

```
Index(['Blocked Driveway', 'Illegal Parking', 'Noise -
Street/Sidewalk',
       'Noise - Commercial', 'Derelict Vehicle'],
      dtype='object')
```

```
borough_complaints = dataset.groupby(['Borough', 'Complaint
Type']).size().unstack()
borough_complaints = borough_complaints[top_5_complaints]
borough_complaints
```

```
Complaint Type  Blocked Driveway  Illegal Parking  Noise -
Street/Sidewalk  \
Borough

BRONX                      12740             7829
8864
BROOKLYN                   28119            27386
13315
MANHATTAN                   2055            11981
20362
QUEENS                     31621            21944
4391
STATEN ISLAND               2141             4881
815

Complaint Type  Noise - Commercial  Derelict Vehicle
```

```
Borough
BRONX                          2431             1948
BROOKLYN                      11451             5164
MANHATTAN                     14528              530
QUEENS                         6057             8102
STATEN ISLAND                   677             1762

#complaints on the basis of borough
top_borough = dataset['Borough'].value_counts().keys()

complaint_per_borough = dataset.groupby(['Complaint Type',
'Borough']).size().unstack()
complaint_per_borough = complaint_per_borough[top_borough]
complaint_per_borough
```

Borough	BROOKLYN	QUEENS	MANHATTAN	BRONX	STATEN ISLAND
Complaint Type					
Animal Abuse	2390	1874	1511	1412	557
Blocked Driveway	28119	31621	2055	12740	2141
Derelict Vehicle	5164	8102	530	1948	1762
Disorderly Youth	72	59	68	63	23
Drinking	257	357	294	187	175
Graffiti	43	37	22	9	2
Illegal Parking	27386	21944	11981	7829	4881
Noise - Commercial	11451	6057	14528	2431	677
Noise - House of Worship	338	297	189	79	17
Noise - Park	1537	634	1167	522	67
Noise - Street/Sidewalk	13315	4391	20362	8864	815
Noise - Vehicle	5145	2608	5374	3385	356
Posting Advertisement	45	30	41	16	515
Traffic	1082	1302	1531	355	196
Vending	514	477	2380	377	25

```
dataset.info()
```

```
<class 'pandas.core.frame.DataFrame'>
Int64Index: 291107 entries, 0 to 300697
Data columns (total 14 columns):
 #   Column                  Non-Null Count   Dtype
---  ------                  --------------   -----
 0   Unique Key              291107 non-null  int64
 1   Created Date            291107 non-null  datetime64[ns]
 2   Closed Date             291107 non-null  datetime64[ns]
 3   Agency                  291107 non-null  object
 4   Complaint Type          291107 non-null  object
 5   Descriptor              291107 non-null  object
 6   Location Type           291107 non-null  object
 7   Incident Zip            291107 non-null  float64
 8   City                    291107 non-null  object
 9   Resolution Description  291107 non-null  object
 10  Borough                 291107 non-null  object
 11  Latitude                291107 non-null  float64
 12  Longitude               291107 non-null  float64
 13  Request_Closing_Time    291107 non-null  timedelta64[ns]
dtypes: datetime64[ns](2), float64(3), int64(1), object(7),
timedelta64[ns](1)
memory usage: 33.3+ MB

dataset.head(5)

   Unique Key        Created Date         Closed Date Agency  \
0    32310363 2015-12-31 23:59:45 2016-01-01 00:55:00   NYPD
1    32309934 2015-12-31 23:59:44 2016-01-01 01:26:00   NYPD
2    32309159 2015-12-31 23:59:29 2016-01-01 04:51:00   NYPD
3    32305098 2015-12-31 23:57:46 2016-01-01 07:43:00   NYPD
4    32306529 2015-12-31 23:56:58 2016-01-01 03:24:00   NYPD

            Complaint Type                     Descriptor    Location
Type  \
0  Noise - Street/Sidewalk              Loud Music/Party
Street/Sidewalk
1         Blocked Driveway                      No Access
Street/Sidewalk
2         Blocked Driveway                      No Access
Street/Sidewalk
3          Illegal Parking  Commercial Overnight Parking
Street/Sidewalk
4          Illegal Parking              Blocked Sidewalk
Street/Sidewalk

   Incident Zip      City                           Resolution
Description  \
0       10034.0  NEW YORK  The Police Department responded and upon
arriv...
1       11105.0   ASTORIA  The Police Department responded to the
```

```
complai...
2       10458.0     BRONX  The Police Department responded and upon
arriv...
3       10461.0     BRONX  The Police Department responded to the
complai...
4       11373.0  ELMHURST  The Police Department responded and upon
arriv...
```

```
     Borough   Latitude  Longitude Request_Closing_Time
0  MANHATTAN  40.865682 -73.923501      0 days 00:55:15
1     QUEENS  40.775945 -73.915094      0 days 01:26:16
2      BRONX  40.870325 -73.888525      0 days 04:51:31
3      BRONX  40.835994 -73.828379      0 days 07:45:14
4     QUEENS  40.733060 -73.874170      0 days 03:27:02
```

```
#converting Request_Closing_time into hours for a new column
dataset['Request_Closing_Time_in_Hours'] =
dataset['Request_Closing_Time'].astype('timedelta64[h]')+1

dataset[['Request_Closing_Time', 'Request_Closing_Time_in_Hours']]
```

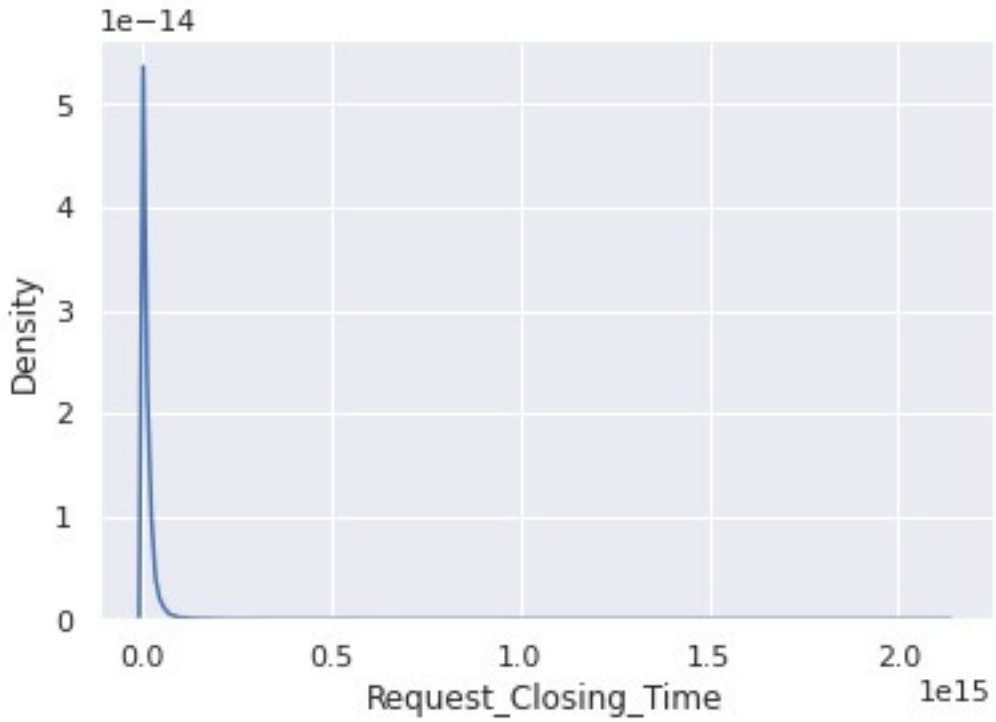
```
       Request_Closing_Time  Request_Closing_Time_in_Hours
0           0 days 00:55:15                            1.0
1           0 days 01:26:16                            2.0
2           0 days 04:51:31                            5.0
3           0 days 07:45:14                            8.0
4           0 days 03:27:02                            4.0
...                     ...                            ...
300692      0 days 00:38:29                            1.0
300694      0 days 02:00:31                            3.0
300695      0 days 03:07:17                            4.0
300696      0 days 04:05:33                            5.0
300697      0 days 04:08:49                            5.0

[291107 rows x 2 columns]
```

```
import seaborn as sns

sns.set()
sns.distplot(dataset.Request_Closing_Time, hist= False)
plt.show()
```

```
/usr/local/lib/python3.7/site-packages/seaborn/distributions.py:2619:
FutureWarning: `distplot` is a deprecated function and will be removed
in a future version. Please adapt your code to use either `displot` (a
figure-level function with similar flexibility) or `kdeplot` (an axes-
level function for kernel density plots).
  warnings.warn(msg, FutureWarning)
```




```
data_avg_time_in_hrs = dataset.groupby(['City','Complaint Type'])
['Request_Closing_Time_in_Hours'].mean()

data_avg_time_in_hrs.head(10)

City     Complaint Type
ARVERNE  Animal Abuse                2.631579
         Blocked Driveway            3.028571
         Derelict Vehicle            3.407407
         Disorderly Youth            4.000000
         Drinking                    1.000000
         Graffiti                    2.000000
         Illegal Parking             2.827586
         Noise - Commercial          3.000000
         Noise - House of Worship    2.090909
         Noise - Park                1.500000
Name: Request_Closing_Time_in_Hours, dtype: float64

dataset['Request_Closing_Time'].describe()

count                         291107
mean     0 days 04:18:32.132665995
std      0 days 06:03:45.509089128
min                0 days 00:01:00
25%                0 days 01:16:30
50%                0 days 02:42:38
75%                0 days 05:20:24
max               24 days 16:52:22
Name: Request_Closing_Time, dtype: object
```

```
#Finding the mean and standard hours
mean_hrs = dataset['Request_Closing_Time_in_Hours'].mean()
std_hrs = dataset['Request_Closing_Time_in_Hours'].std()

mean_hrs
```

```
4.81868179054437
```

```
std_hrs
```

```
6.06442731490891
```

```
#Check for nomality of Age
# Shapiro- Wilk Test
from scipy.stats import shapiro
stat, p= shapiro(dataset.Request_Closing_Time)
print(stat, p)
```

```
0.5173379182815552 0.0

/usr/local/lib/python3.7/site-packages/scipy/stats/morestats.py:1676:
UserWarning: p-value may not be accurate for N > 5000.
  warnings.warn("p-value may not be accurate for N > 5000.")
```