Service Request Final

January 22, 2020

0.1 Service Request - NYC 311

Project NYC 311 Service Request - To analyse the dataset and obtain insights and conclusions.

0.1.1 1.Importing Libraries and Dataset

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

        import matplotlib.pyplot as plt
        import seaborn as sns

        import datetime

        %matplotlib inline

        from sklearn.preprocessing import StandardScaler

        # importing dataset
        service_request = pd.read_csv('311_Service_Requests_from_2010_to_Present.csv')

        # shape of the dataset
        print("shape" , service_request.shape)
```

```
shape (300698, 53)


C:\ProgramData\Anaconda3\lib\site-packages\IPython\core\interactiveshell.py:3020: DtypeWarning
  interactivity=interactivity, compiler=compiler, result=result)
```

```
In [2]: # printing first five records
        service_request.head()
```

```
Out[2]:    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                   ...                     \
0    71 VERMILYEA AVENUE                   ...
1        27-07 23 AVENUE                   ...
2  2897 VALENTINE AVENUE                   ...
3    2940 BAISLEY AVENUE                   ...
4          87-14 57 ROAD                   ...

  Bridge Highway Name Bridge Highway Direction Road Ramp  \
0                 NaN                      NaN       NaN
1                 NaN                      NaN       NaN
2                 NaN                      NaN       NaN
3                 NaN                      NaN       NaN
4                 NaN                      NaN       NaN

  Bridge Highway Segment Garage Lot Name Ferry Direction Ferry Terminal Name  \
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

    Latitude  Longitude                                 Location
0  40.865682 -73.923501   (40.86568153633767, -73.92350095571744)
1  40.775945 -73.915094  (40.775945312321085, -73.91509393898605)
2  40.870325 -73.888525  (40.870324522111424, -73.88852464418646)
3  40.835994 -73.828379   (40.83599404683083, -73.82837939584206)
4  40.733060 -73.874170  (40.733059618956815, -73.87416975810375)

[5 rows x 53 columns]
```

0.1.2 2. Making Column 'Request_Closing_Time'

Procedure :

- convert the closed date to datetime format
- convert the created date to datetime format
- Take the difference and store it into a new column in the dataframe **NB :** Since there are Nan values in closed date field and the percentage of nan values in closed date field is less than 0.8% , so we drop those rows without losing much informations.

```
In [3]: # Drop the nan values
        service_request.dropna(axis=0 , subset=['Closed Date'] , inplace=True)
```

```
In [4]: service_request['Created Date'].apply(len).unique()
```

```
Out[4]: array([22, 14, 13], dtype=int64)
```

```
In [5]: service_request['Closed Date'].apply(len).unique()
```

```
Out[5]: array([13, 14, 22], dtype=int64)
```

Therefore , there are two distinct formats in Created date (all are of same data type) field. They are of the form :

- '%m/%d/%Y %I:%M:%S %p'
- '%m-%d-%y %H:%M'

```
In [6]: # custom function for converting created date field to datetime :
        def date_converter_(value) :
            try :
                return datetime.datetime.strptime(value , '%m/%d/%Y %I:%M:%S %p')
            except ValueError :
                return datetime.datetime.strptime(value , '%m-%d-%y %H:%M')
```

```
In [7]: service_request['Closed Date'] = service_request['Closed Date'].apply(date_converter_)
        service_request['Created Date'] = service_request['Created Date'].apply(date_converter
```

```
In [8]: # apply the custom function to the Created Date Field
        service_request['Request_Closing_Time'] = service_request['Closed Date'] - service_requ
```

```
In [9]: service_request.head()
```

```
Out[9]:    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

                               Agency Name           Complaint Type  \
        0  New York City Police Department  Noise - Street/Sidewalk
```

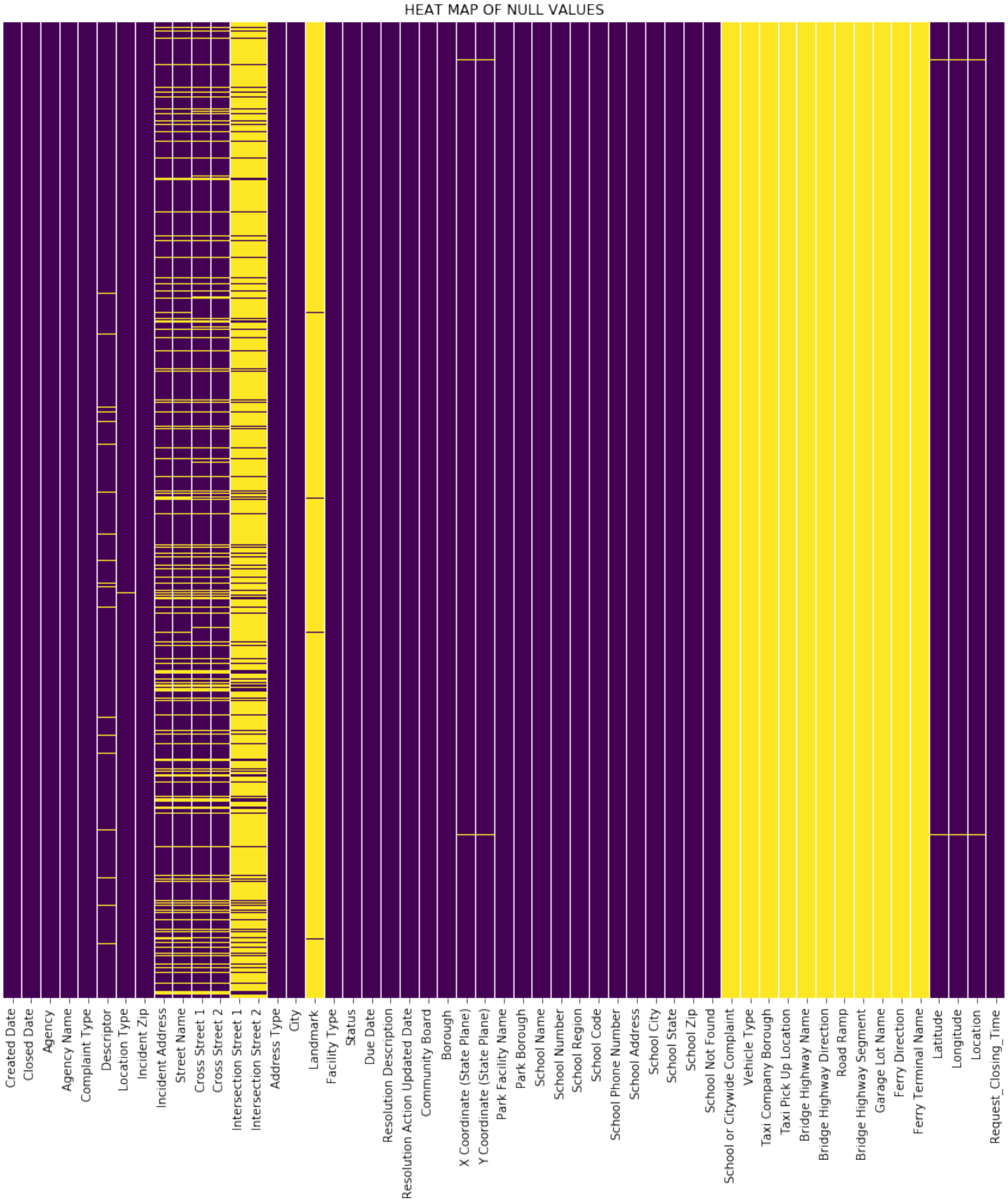
```
1  New York City Police Department         Blocked Driveway
2  New York City Police Department         Blocked Driveway
3  New York City Police Department          Illegal Parking
4  New York City Police Department          Illegal Parking

                     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 Direction  \
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

  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 Request_Closing_Time
0   (40.86568153633767, -73.92350095571744)             00:55:15
1  (40.775945312321085, -73.91509393898605)             01:26:16
2  (40.870324522111424, -73.88852464418646)             04:51:31
3   (40.83599404683083, -73.82837939584206)             07:45:14
4  (40.733059618956815, -73.87416975810375)             03:27:02

[5 rows x 54 columns]
```

```
In [10]: service_request['Agency Name'].unique()
```

```
Out[10]: array(['New York City Police Department', 'NYPD',
                'Internal Affairs Bureau'], dtype=object)
```

0.1.3 3.Data Insights

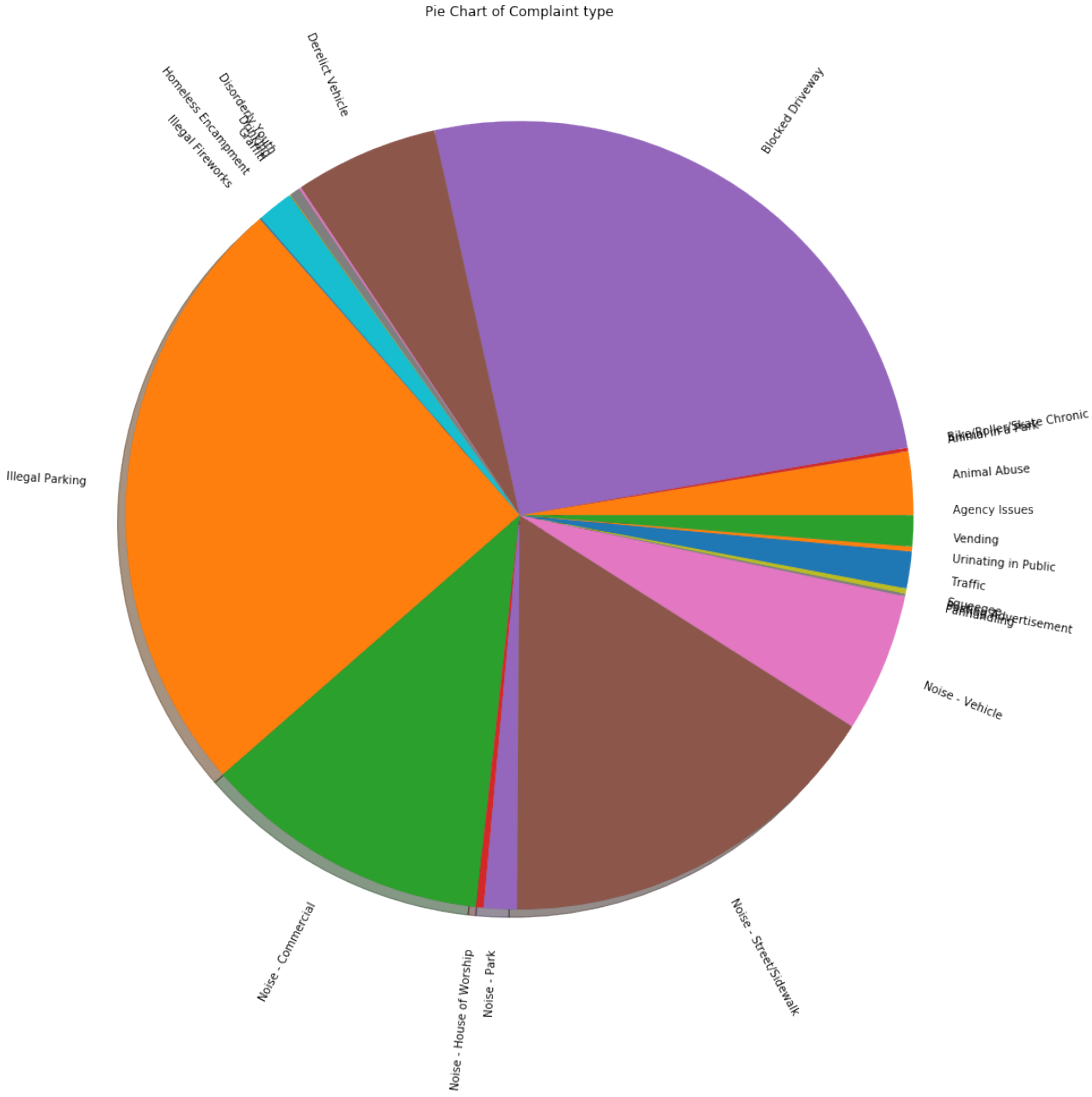
```
In [11]: # Checking missing Values
         plt.figure(figsize=(15,15))
         plt.title("HEAT MAP OF NULL VALUES")
         sns.heatmap(service_request.iloc[:,1:].isnull() , yticklabels=False ,cbar=False, cmap=
         plt.show()
```



So from above analysis we can see that **'Intersection Street 1','Intersection Street 2', 'Landmark','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'** are not possible for analysis as they have too much null values

Complaint Type Analysis

```
In [12]: freq_complaint = service_request.groupby('Complaint Type').agg('count')['Unique Key']
         # explode = np.zeros(23)
         # explode[0] = 2
         plt.figure(figsize=(16,16))
         plt.title('Pie Chart of Complaint type')
         plt.pie(x=freq_complaint.values.astype('float64') ,shadow=True, labels=freq_complaint
         plt.show()

         print("STATS\n----------------------\n" , service_request['Complaint Type'].describe()

         print('----------------------')
         print(service_request.groupby('Complaint Type').agg('count')['Unique Key'].sort_values
```

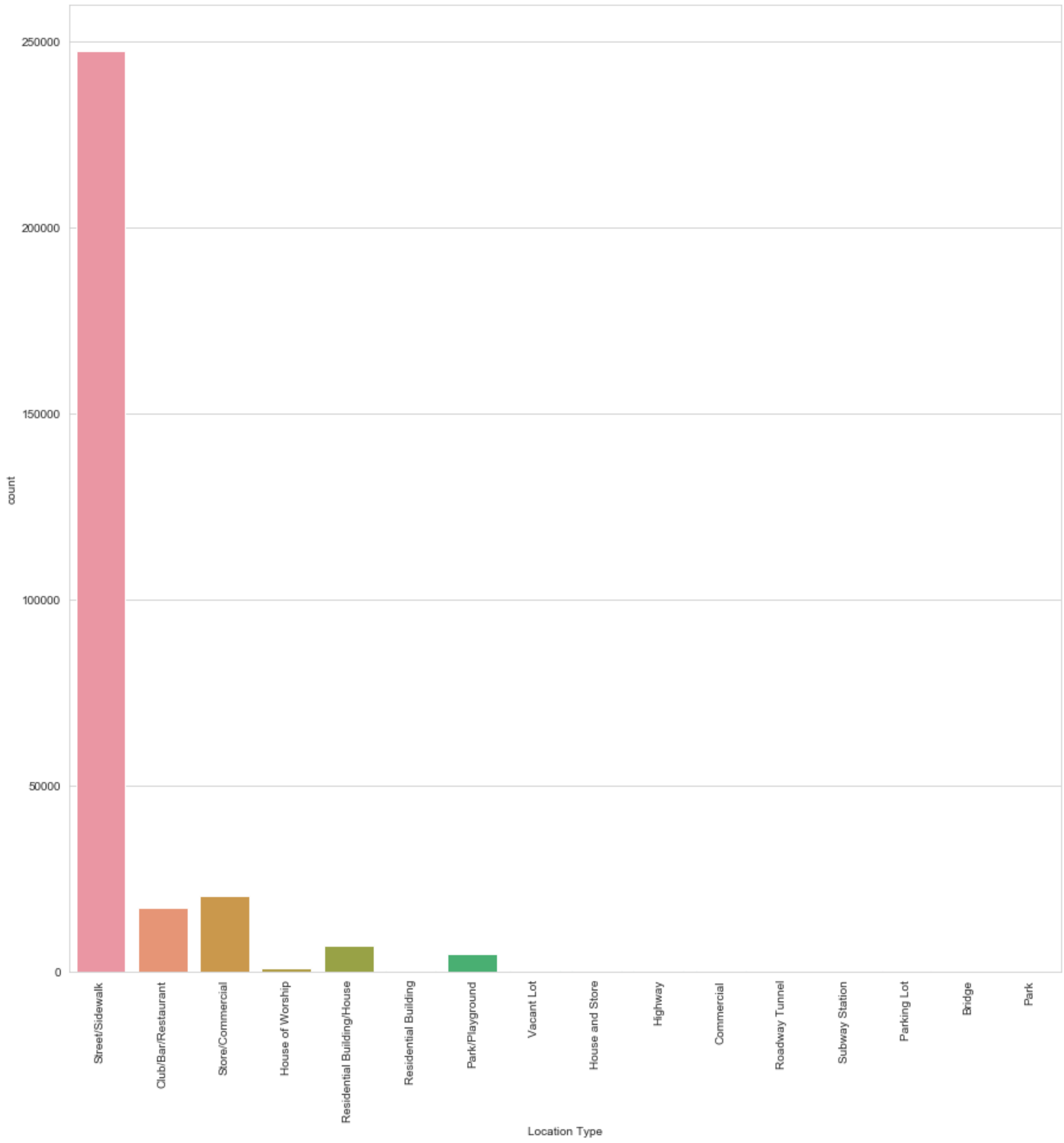


```
STATS
----------------------
count                 298534
unique                    23
top         Blocked Driveway
freq                   76810
Name: Complaint Type, dtype: object
----------------------
Complaint Type
Animal in a Park                  1
Squeegee                          4
Agency Issues                     6
Graffiti                        113
```

```
Illegal Fireworks              168
Disorderly Youth               286
Panhandling                    305
Bike/Roller/Skate Chronic      424
Urinating in Public            592
Posting Advertisement          648
Noise - House of Worship       929
Drinking                      1275
Vending                       3795
Noise - Park                  4022
Homeless Encampment           4416
Traffic                       4496
Animal Abuse                  7768
Noise - Vehicle              17033
Derelict Vehicle             17588
Noise - Commercial           35247
Noise - Street/Sidewalk      48076
Illegal Parking              74532
Blocked Driveway             76810
Name: Unique Key, dtype: int64
```

Location Type Analysis

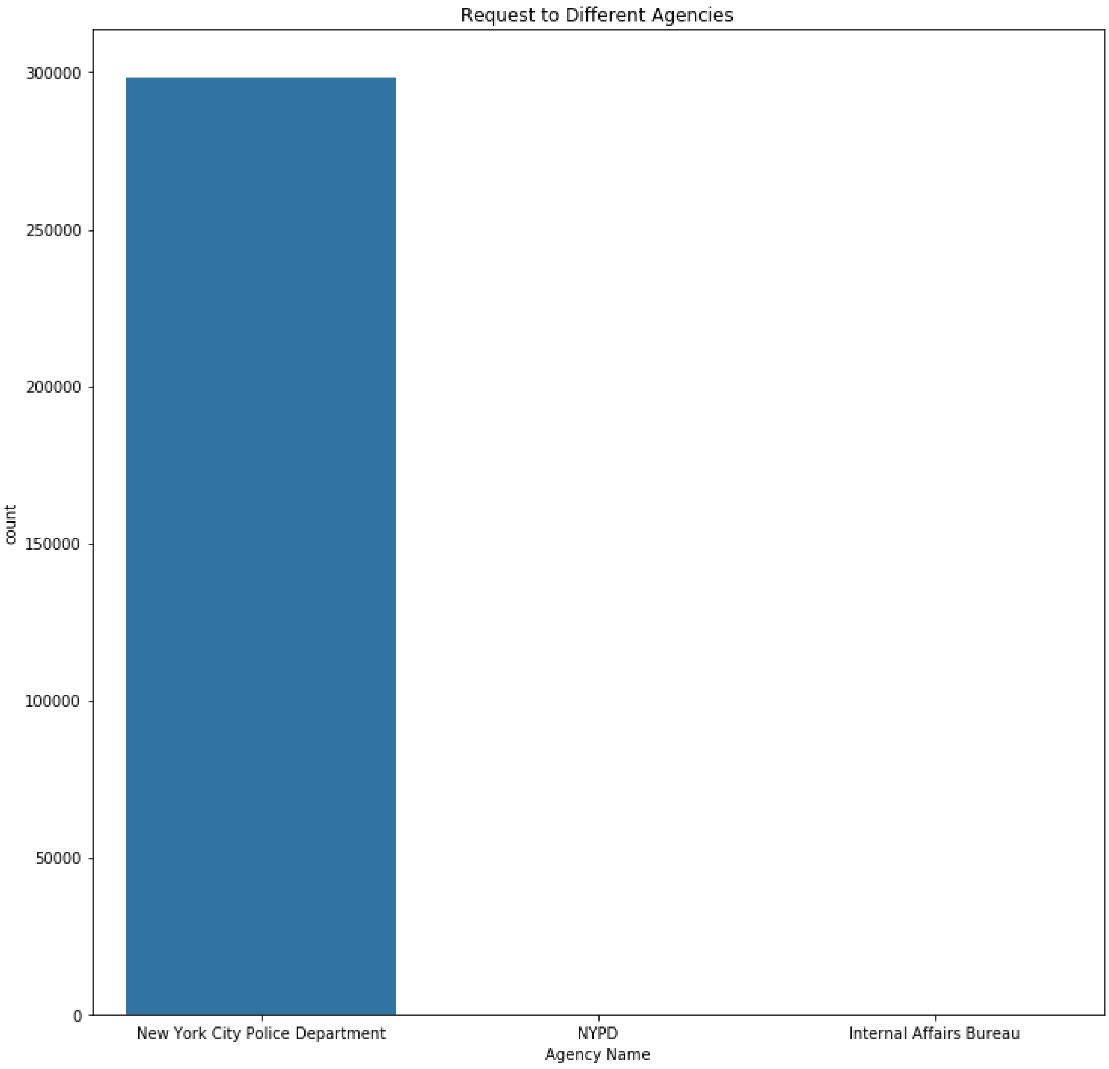
```
In [88]: plt.figure(figsize=(15,15))
         sns.countplot(x='Location Type' , data=service_request )
         plt.xticks(rotation='vertical')
         plt.show()
         print("STATS\n----------------------------\n" , service_request['Location Type'].desc
```



```
STATS
----------------------------
count                 298406
unique                    16
top          Street/Sidewalk
freq                  247503
Name: Location Type, dtype: object
```

```
In [14]: plt.figure(figsize=(12,12))
         plt.title("Request to Different Agencies")
```

```
sns.countplot(x = 'Agency Name' , data=service_request)

plt.show()

print("STATS\n-------------------------\n" , service_request['Agency Name'].describe()
print("-------------------------")
print(service_request.groupby('Agency Name').agg('count')['Unique Key'])
```



```
STATS
-------------------------
count                               298534
unique                                   3
top          New York City Police Department
```

```
freq                               298527
Name: Agency Name, dtype: object
-------------------------
Agency Name
Internal Affairs Bureau                 6
NYPD                                    1
New York City Police Department    298527
Name: Unique Key, dtype: int64
```

0.1.4 4.Mean of request closing time based on complaint type grouped under different location

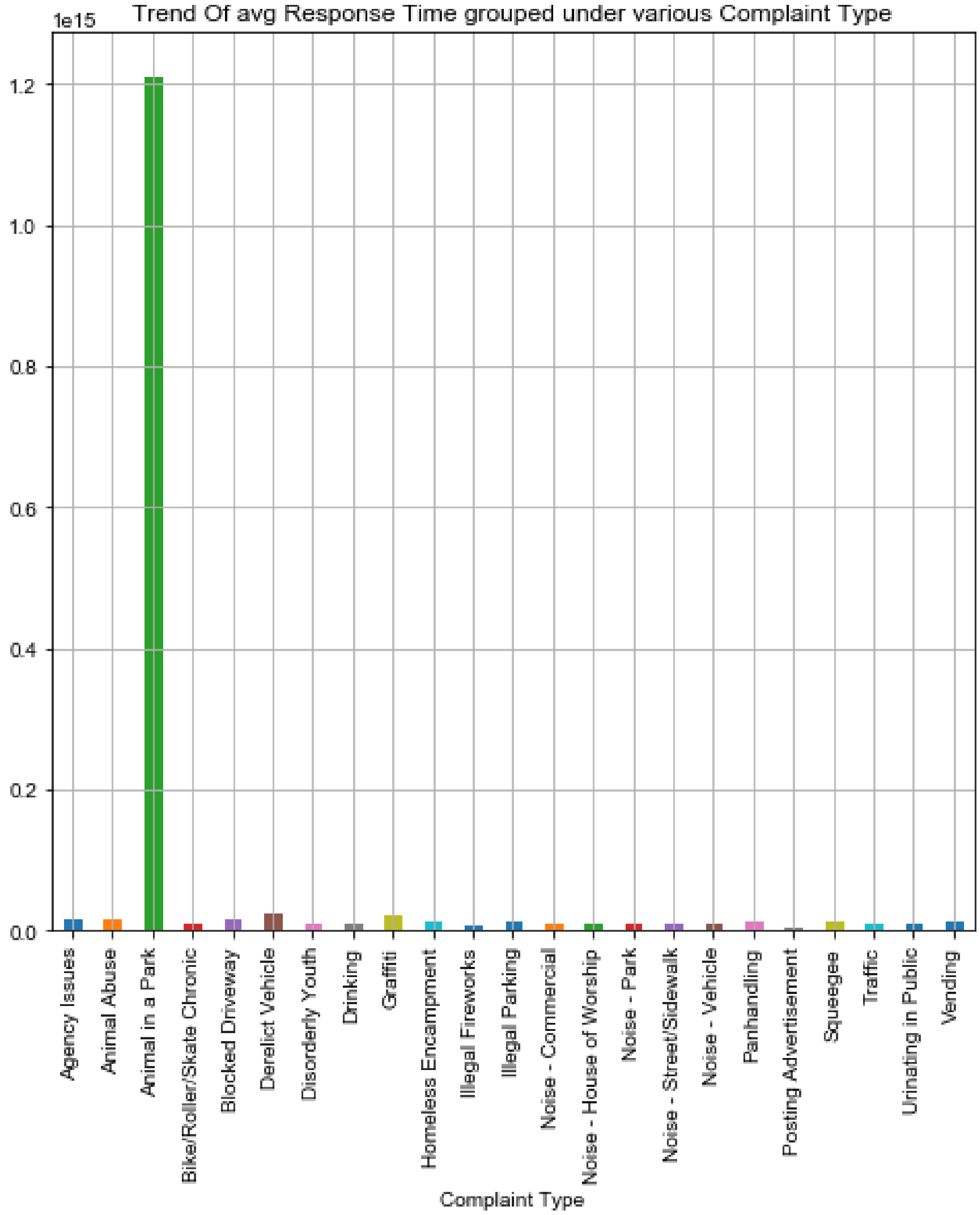
```
In [15]: service_request['new'] = service_request['Request_Closing_Time'].values.astype(np.int6

         newss = service_request.groupby(['Location Type' , 'Complaint Type']).mean()

         newss.new.apply(pd.to_timedelta)
```

```
Out[15]: Location Type               Complaint Type
         Bridge                      Homeless Encampment         0 days 03:49:09.500000
         Club/Bar/Restaurant         Drinking                    0 days 04:32:44.923287
                                     Noise - Commercial          0 days 03:03:50.116976
                                     Urinating in Public                0 days 07:55:12
         Commercial                  Animal Abuse                0 days 05:20:33.967741
         Highway                     Derelict Vehicle            0 days 08:11:44.538461
                                     Homeless Encampment         0 days 03:16:16.200000
                                     Traffic                     0 days 03:26:51.639784
         House and Store             Animal Abuse                0 days 05:00:47.741935
         House of Worship            Noise - House of Worship    0 days 03:11:49.996763
         Park                        Animal in a Park                  14 days 00:50:05
         Park/Playground             Animal Abuse                0 days 03:42:08.975409
                                     Drinking                    0 days 03:26:28.785714
                                     Homeless Encampment         0 days 03:55:25.787535
                                     Illegal Fireworks                  0 days 05:00:11
                                     Noise - Park                0 days 03:24:32.394429
                                     Panhandling                 0 days 01:13:07.166666
                                     Urinating in Public         0 days 02:51:44.973684
                                     Vending                     0 days 03:28:01.647619
         Parking Lot                 Animal Abuse                0 days 05:32:25.436363
                                     Posting Advertisement       0 days 02:06:56.714285
         Residential Building        Animal Abuse                0 days 04:49:05.396475
         Residential Building/House  Animal Abuse                0 days 05:23:58.263924
                                     Bike/Roller/Skate Chronic   0 days 03:36:40.680000
                                     Disorderly Youth            0 days 03:51:16.116883
                                     Drinking                    0 days 03:40:41.813148
                                     Graffiti                    0 days 05:52:43.678571
                                     Homeless Encampment         0 days 04:55:26.776195
```

```
                            Illegal Fireworks           0 days 03:05:58.969696
                            Panhandling                        0 days 05:09:56
                                                                 ...
Store/Commercial            Drinking                    0 days 03:15:09.344444
                            Graffiti                    0 days 05:33:38.312500
                            Homeless Encampment         0 days 04:14:14.476562
                            Illegal Fireworks                  0 days 01:55:27
                            Noise - Commercial          0 days 03:13:25.207597
                            Panhandling                 0 days 04:20:46.833333
                            Posting Advertisement              0 days 02:22:09
                            Urinating in Public         0 days 03:26:23.075757
                            Vending                     0 days 03:58:25.817129
Street/Sidewalk             Animal Abuse                0 days 05:06:38.977734
                            Bike/Roller/Skate Chronic   0 days 03:45:59.072254
                            Blocked Driveway            0 days 04:44:27.674677
                            Derelict Vehicle            0 days 07:21:38.166885
                            Disorderly Youth            0 days 03:28:25.870646
                            Drinking                    0 days 03:37:51.009259
                            Graffiti                    0 days 12:02:15.080000
                            Homeless Encampment         0 days 04:13:47.726879
                            Illegal Fireworks           0 days 02:32:29.992000
                            Illegal Parking             0 days 04:30:02.851416
                            Noise - Street/Sidewalk     0 days 03:26:43.506813
                            Noise - Vehicle             0 days 03:35:20.531826
                            Panhandling                 0 days 04:24:28.295964
                            Posting Advertisement       0 days 01:49:08.375862
                            Squeegee                    0 days 04:02:44.250000
                            Traffic                     0 days 03:27:35.132834
                            Urinating in Public         0 days 03:17:06.835443
                            Vending                     0 days 04:01:36.091980
Subway Station              Animal Abuse                0 days 03:02:08.181818
                            Urinating in Public         0 days 01:09:07.666666
Vacant Lot                  Derelict Vehicle            0 days 07:28:26.129870
Name: new, Length: 69, dtype: timedelta64[ns]
```

0.1.5 5.1.Avg Response time for various complaint type

```
In [16]: que_5a = service_request.groupby('Complaint Type').mean()
         plt.figure(figsize=(8,8))
         plt.title('Trend Of avg Response Time grouped under various Complaint Type')
         sns.set_style("whitegrid")
         que_5a.new.apply(pd.to_timedelta).plot(kind='bar')
         # que_5a.new.apply(pd.to_timedelta).plot(kind='line' , cmap = 'viridis' , x=None)
         plt.show()

         print("------------------------------------------------------------------------")

         mean_exc_animal_in_park = pd.to_timedelta(que_5a['new'].drop(index='Animal in a Park'
         std_exc_animal_in_park = pd.to_timedelta(que_5a['new'].drop(index='Animal in a Park')
         mean_ = que_5a['new'].apply(pd.to_timedelta).mean()
         print("Conclusion :\n")
         print("Mean =" , mean_)
         print("Mean Excluding Animal in a park Location Type  =" , mean_exc_animal_in_park)
         print("Std Excluding Animal in a park  Location Type =" , std_exc_animal_in_park)
```



Trend Of avg Response Time grouped under various Complaint Type

```
----------------------------------------------------------------------
Conclusion :

Mean = 0 days 18:35:36.238295
Mean Excluding Animal in a park row  = 0 days 04:07:40.385490
Std Excluding Animal in a park row = 0 days 01:15:51.604037
```

We except Animal in park (freq = 1) Complaint type row , other complaint types have almost similar avg response time.

5.2.The Relation between Complaint Type and Location In this section we make a Logistics Regression model to predict the Location of inputted Complaint Type.

```
In [17]: service_request.head()
```

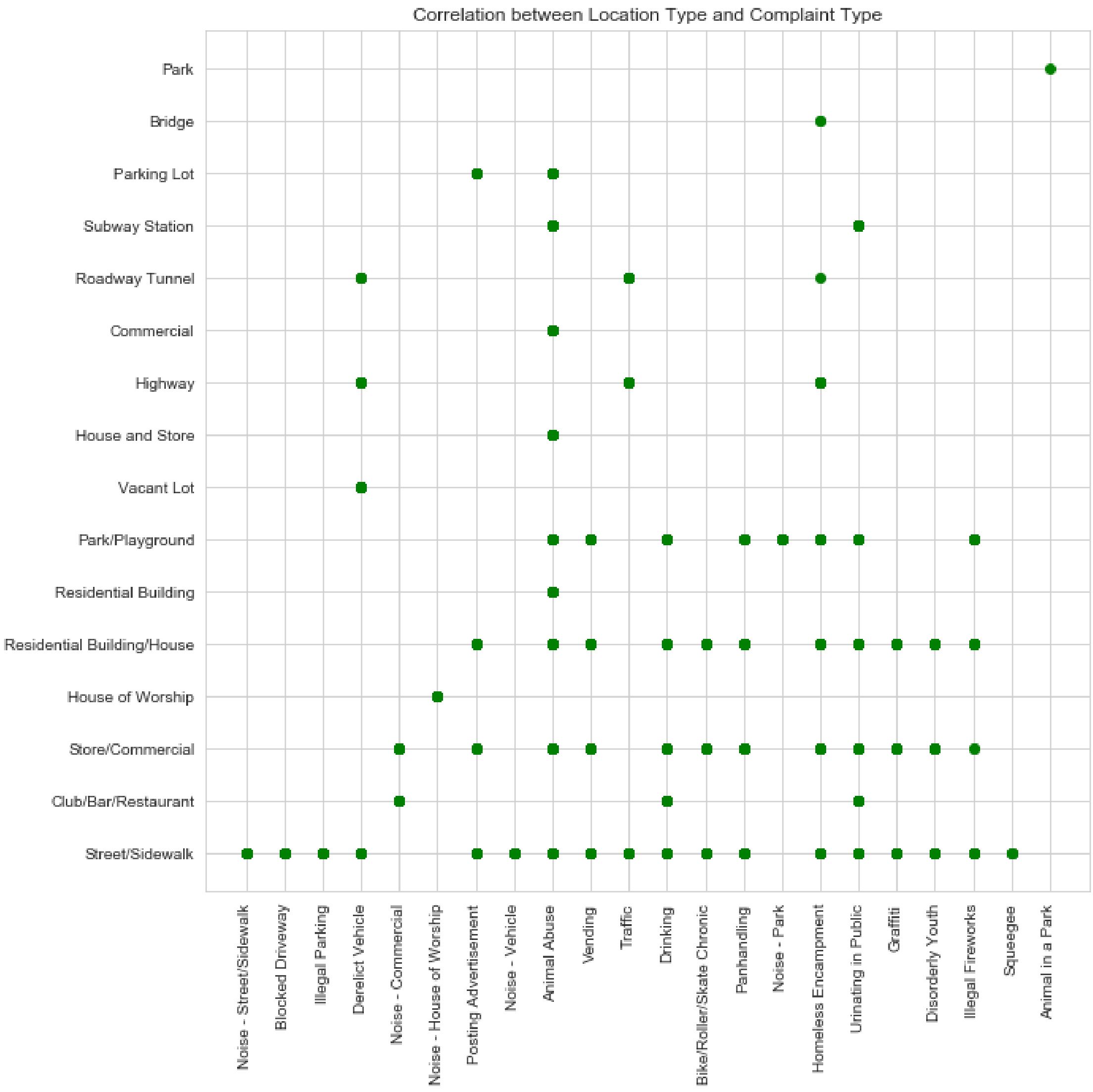
```
Out[17]:    Unique Key        Created Date         Closed Date Agency  \
         0    32310363 2015-12-31 23:59:45 2016-01-01 00:55:00   NYPD
         1    32309934 2015-12-31 23:59:44 2016-01-01 01:26:00   NYPD
         2    32309159 2015-12-31 23:59:29 2016-01-01 04:51:00   NYPD
         3    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

                                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       ...       Road Ramp Bridge Highway Segment  \
         0    71 VERMILYEA AVENUE       ...             NaN                    NaN
         1        27-07 23 AVENUE       ...             NaN                    NaN
         2  2897 VALENTINE AVENUE       ...             NaN                    NaN
         3    2940 BAISLEY AVENUE       ...             NaN                    NaN
         4            87-14 57 ROAD       ...             NaN                    NaN

           Garage Lot Name Ferry Direction Ferry Terminal Name   Latitude  Longitude  \
         0             NaN             NaN                 NaN  40.865682 -73.923501
         1             NaN             NaN                 NaN  40.775945 -73.915094
         2             NaN             NaN                 NaN  40.870325 -73.888525
```

```
3             NaN             NaN                 NaN  40.835994 -73.828379
4             NaN             NaN                 NaN  40.733060 -73.874170

                                   Location Request_Closing_Time  \
0   (40.86568153633767, -73.92350095571744)             00:55:15
1  (40.775945312321085, -73.91509393898605)             01:26:16
2  (40.870324522111424, -73.88852464418646)             04:51:31
3   (40.83599404683083, -73.82837939584206)             07:45:14
4  (40.733059618956815, -73.87416975810375)             03:27:02

              new
0   3315000000000
1   5176000000000
2  17491000000000
3  27914000000000
4  12422000000000

[5 rows x 55 columns]
```

```
In [18]: service_request.groupby('Location Type' ).agg('count')['Complaint Type'].reset_index(
```

```
Out[18]:                  Location Type  Complaint Type
         0                       Bridge               2
         1          Club/Bar/Restaurant           17227
         2                   Commercial              62
         3                      Highway             214
         4              House and Store              93
         5             House of Worship             927
         6                         Park               1
         7              Park/Playground            4751
         8                  Parking Lot             117
         9         Residential Building             227
         10  Residential Building/House            6953
         11              Roadway Tunnel              35
         12            Store/Commercial           20183
         13             Street/Sidewalk          247503
         14              Subway Station              34
         15                  Vacant Lot              77
```

```
In [20]: plt.figure(figsize=(10,10))
         plt.title('Correlation between Location Type and Complaint Type')
         plt.scatter(data=service_request.dropna(subset=['Complaint Type' , 'Location Type'])
         plt.xticks(rotation='vertical')
         plt.show()
```



```
In [47]: que_5b = service_request[['Complaint Type' , 'Location Type']].dropna(subset=['Locatio
```

```
In [48]: que_5b.describe()
```

```
Out[48]:            Complaint Type    Location Type
         count              298406           298406
         unique                 22               16
         top      Blocked Driveway  Street/Sidewalk
         freq                76773           247503
```

```
In [49]: que_5b_test = pd.get_dummies(data=que_5b , columns=['Complaint Type'] )
```

```
In [66]: from sklearn.preprocessing import LabelEncoder
```

```
         label_encoder = LabelEncoder()

         label_encoder.fit(que_5b_test['Location Type'])

         y = label_encoder.transform(que_5b_test['Location Type'])
         X = que_5b_test.iloc[: , 1:]

In [67]: X.head()

Out[67]:    Complaint Type_Animal Abuse  Complaint Type_Animal in a Park  \
         0                            0                                0
         1                            0                                0
         2                            0                                0
         3                            0                                0
         4                            0                                0

            Complaint Type_Bike/Roller/Skate Chronic  Complaint Type_Blocked Driveway  \
         0                                         0                                0
         1                                         0                                1
         2                                         0                                1
         3                                         0                                0
         4                                         0                                0

            Complaint Type_Derelict Vehicle  Complaint Type_Disorderly Youth  \
         0                                0                                0
         1                                0                                0
         2                                0                                0
         3                                0                                0
         4                                0                                0

            Complaint Type_Drinking  Complaint Type_Graffiti  \
         0                        0                        0
         1                        0                        0
         2                        0                        0
         3                        0                        0
         4                        0                        0

            Complaint Type_Homeless Encampment  Complaint Type_Illegal Fireworks  \
         0                                   0                                 0
         1                                   0                                 0
         2                                   0                                 0
         3                                   0                                 0
         4                                   0                                 0

                     ...            Complaint Type_Noise - House of Worship  \
         0           ...                                                  0
         1           ...                                                  0
         2           ...                                                  0
```

```
3           ...                                                  0
4           ...                                                  0

   Complaint Type_Noise - Park  Complaint Type_Noise - Street/Sidewalk  \
0                            0                                       1
1                            0                                       0
2                            0                                       0
3                            0                                       0
4                            0                                       0

   Complaint Type_Noise - Vehicle  Complaint Type_Panhandling  \
0                               0                           0
1                               0                           0
2                               0                           0
3                               0                           0
4                               0                           0

   Complaint Type_Posting Advertisement  Complaint Type_Squeegee  \
0                                     0                        0
1                                     0                        0
2                                     0                        0
3                                     0                        0
4                                     0                        0

   Complaint Type_Traffic  Complaint Type_Urinating in Public  \
0                       0                                   0
1                       0                                   0
2                       0                                   0
3                       0                                   0
4                       0                                   0

   Complaint Type_Vending
0                       0
1                       0
2                       0
3                       0
4                       0

[5 rows x 22 columns]
```

In [70]:
```
from sklearn.linear_model import LogisticRegression
from sklearn.model_selection import train_test_split

model = LogisticRegression()

X_train , X_test , y_train , y_test = train_test_split(X , y , test_size=0.3 , random_
```

In [71]:
```
X_train.head()
```

```
Out[71]:         Complaint Type_Animal Abuse  Complaint Type_Animal in a Park  \
         191109                            0                                0
         201522                            0                                0
         219279                            0                                0
         55782                             0                                0
         115588                            0                                0

                 Complaint Type_Bike/Roller/Skate Chronic  \
         191109                                         0
         201522                                         0
         219279                                         0
         55782                                          0
         115588                                         0

                 Complaint Type_Blocked Driveway  Complaint Type_Derelict Vehicle  \
         191109                                0                                1
         201522                                1                                0
         219279                                1                                0
         55782                                 0                                0
         115588                                1                                0

                 Complaint Type_Disorderly Youth  Complaint Type_Drinking  \
         191109                                0                        0
         201522                                0                        0
         219279                                0                        0
         55782                                 0                        0
         115588                                0                        0

                 Complaint Type_Graffiti  Complaint Type_Homeless Encampment  \
         191109                        0                                   0
         201522                        0                                   0
         219279                        0                                   0
         55782                         0                                   0
         115588                        0                                   0

                 Complaint Type_Illegal Fireworks           ...            \
         191109                                 0           ...
         201522                                 0           ...
         219279                                 0           ...
         55782                                  0           ...
         115588                                 0           ...

                 Complaint Type_Noise - House of Worship  Complaint Type_Noise - Park  \
         191109                                        0                            0
         201522                                        0                            0
         219279                                        0                            0
         55782                                         0                            0
         115588                                        0                            0
```

```
        Complaint Type_Noise - Street/Sidewalk  \
191109                                       0
201522                                       0
219279                                       0
55782                                        0
115588                                       0

        Complaint Type_Noise - Vehicle  Complaint Type_Panhandling  \
191109                               0                           0
201522                               0                           0
219279                               0                           0
55782                                0                           0
115588                               0                           0

        Complaint Type_Posting Advertisement  Complaint Type_Squeegee  \
191109                                     0                        0
201522                                     0                        0
219279                                     0                        0
55782                                      0                        0
115588                                     0                        0

        Complaint Type_Traffic  Complaint Type_Urinating in Public  \
191109                       0                                   0
201522                       0                                   0
219279                       0                                   0
55782                        0                                   0
115588                       0                                   0

        Complaint Type_Vending
191109                       0
201522                       0
219279                       0
55782                        0
115588                       0

[5 rows x 22 columns]
```

```
In [72]: model.fit(X_train , y_train)
```

```
C:\ProgramData\Anaconda3\lib\site-packages\sklearn\linear_model\logistic.py:433: FutureWarning
  FutureWarning)
C:\ProgramData\Anaconda3\lib\site-packages\sklearn\linear_model\logistic.py:460: FutureWarning
  "this warning.", FutureWarning)
```

```
Out[72]: LogisticRegression(C=1.0, class_weight=None, dual=False, fit_intercept=True,
                   intercept_scaling=1, max_iter=100, multi_class='warn',
                   n_jobs=None, penalty='l2', random_state=None, solver='warn',
                   tol=0.0001, verbose=0, warm_start=False)
```

```
In [73]: predictions = model.predict(X_test)
```

```
In [86]: from sklearn.metrics import confusion_matrix,classification_report

         # print(predictions[:15])
         # print(y_test[:15])
```

```
[12 12 13 12 13 13 13 12 10 13 13 12 13 13 13]
[ 1 12 13  1 13 13 13 12 10 13 13  1 13 13 13]
```

```
In [91]: print(classification_report(label_encoder.inverse_transform(y_test) , label_encoder.i

         # print(confusion_matrix(y_test , predictions))
```

```
C:\ProgramData\Anaconda3\lib\site-packages\sklearn\metrics\classification.py:1143: UndefinedMe
  'precision', 'predicted', average, warn_for)
C:\ProgramData\Anaconda3\lib\site-packages\sklearn\metrics\classification.py:1143: UndefinedMe
  'precision', 'predicted', average, warn_for)


                           precision    recall  f1-score   support

       Club/Bar/Restaurant      0.00      0.00      0.00      5220
                Commercial      0.00      0.00      0.00        21
                   Highway      0.00      0.00      0.00        76
           House and Store      0.00      0.00      0.00        32
          House of Worship      1.00      1.00      1.00       261
                      Park      0.00      0.00      0.00         1
           Park/Playground      1.00      0.84      0.92      1359
               Parking Lot      0.00      0.00      0.00        31
      Residential Building      0.00      0.00      0.00        70
Residential Building/House      0.65      0.74      0.69      2045
            Roadway Tunnel      0.00      0.00      0.00        13
          Store/Commercial      0.52      0.91      0.66      6006
           Street/Sidewalk      0.98      0.99      0.99     74363
            Subway Station      0.00      0.00      0.00         6
                Vacant Lot      0.00      0.00      0.00        18

                 micro avg      0.92      0.92      0.92     89522
                 macro avg      0.28      0.30      0.28     89522
              weighted avg      0.88      0.92      0.90     89522



C:\ProgramData\Anaconda3\lib\site-packages\sklearn\metrics\classification.py:1143: UndefinedMe
  'precision', 'predicted', average, warn_for)
```