{  
 "cells": [  
 {  
 "cell\_type": "code",  
 "execution\_count": 12,  
 "id": "4e5dccd1",  
 "metadata": {},  
 "outputs": [],  
 "source": [  
 "#Importing libraries\n",  
 "import numpy as np\n",  
 "import pandas as pd\n",  
 "import matplotlib.pyplot as plt\n",  
 "import seaborn as sns\n",  
 "sns.set()"  
 ]  
 },  
 {  
 "cell\_type": "code",  
 "execution\_count": 13,  
 "id": "09a76c77",  
 "metadata": {},  
 "outputs": [],  
 "source": [  
 "from scipy import stats\n",  
 "from scipy.stats import chi2\_contingency\n",  
 "import statsmodels.api as sm\n",  
 "from statsmodels.formula.api import ols"  
 ]  
 },  
 {  
 "cell\_type": "code",  
 "execution\_count": 14,  
 "id": "243e0a16",  
 "metadata": {},  
 "outputs": [  
 {  
 "name": "stderr",  
 "output\_type": "stream",  
 "text": [  
 "C:\\Users\\lrnem\\AppData\\Local\\Temp\\ipykernel\_8728\\2134542247.py:2: DtypeWarning: Columns (48,49) have mixed types. Specify dtype option on import or set low\_memory=False.\n",  
 " df=pd.read\_csv(\"311\_Service\_Requests\_from\_2010\_to\_Present.csv\")\n"  
 ]  
 }  
 ],  
 "source": [  
 "#Importing the dataset\n",  
 "df=pd.read\_csv(\"311\_Service\_Requests\_from\_2010\_to\_Present.csv\")"  
 ]  
 },  
 {  
 "cell\_type": "code",  
 "execution\_count": 15,  
 "id": "854fd89a",  
 "metadata": {},  
 "outputs": [  
 {  
 "data": {  
 "text/html": [  
 "<div>\n",  
 "<style scoped>\n",  
 " .dataframe tbody tr th:only-of-type {\n",  
 " vertical-align: middle;\n",  
 " }\n",  
 "\n",  
 " .dataframe tbody tr th {\n",  
 " vertical-align: top;\n",  
 " }\n",  
 "\n",  
 " .dataframe thead th {\n",  
 " text-align: right;\n",  
 " }\n",  
 "</style>\n",  
 "<table border=\"1\" class=\"dataframe\">\n",  
 " <thead>\n",  
 " <tr style=\"text-align: right;\">\n",  
 " <th></th>\n",  
 " <th>Unique Key</th>\n",  
 " <th>Created Date</th>\n",  
 " <th>Closed Date</th>\n",  
 " <th>Agency</th>\n",  
 " <th>Agency Name</th>\n",  
 " <th>Complaint Type</th>\n",  
 " <th>Descriptor</th>\n",  
 " <th>Location Type</th>\n",  
 " <th>Incident Zip</th>\n",  
 " <th>Incident Address</th>\n",  
 " <th>...</th>\n",  
 " <th>Bridge Highway Name</th>\n",  
 " <th>Bridge Highway Direction</th>\n",  
 " <th>Road Ramp</th>\n",  
 " <th>Bridge Highway Segment</th>\n",  
 " <th>Garage Lot Name</th>\n",  
 " <th>Ferry Direction</th>\n",  
 " <th>Ferry Terminal Name</th>\n",  
 " <th>Latitude</th>\n",  
 " <th>Longitude</th>\n",  
 " <th>Location</th>\n",  
 " </tr>\n",  
 " </thead>\n",  
 " <tbody>\n",  
 " <tr>\n",  
 " <th>0</th>\n",  
 " <td>32310363</td>\n",  
 " <td>12/31/2015 11:59:45 PM</td>\n",  
 " <td>01-01-16 0:55</td>\n",  
 " <td>NYPD</td>\n",  
 " <td>New York City Police Department</td>\n",  
 " <td>Noise - Street/Sidewalk</td>\n",  
 " <td>Loud Music/Party</td>\n",  
 " <td>Street/Sidewalk</td>\n",  
 " <td>10034.0</td>\n",  
 " <td>71 VERMILYEA AVENUE</td>\n",  
 " <td>...</td>\n",  
 " <td>NaN</td>\n",  
 " <td>NaN</td>\n",  
 " <td>NaN</td>\n",  
 " <td>NaN</td>\n",  
 " <td>NaN</td>\n",  
 " <td>NaN</td>\n",  
 " <td>NaN</td>\n",  
 " <td>40.865682</td>\n",  
 " <td>-73.923501</td>\n",  
 " <td>(40.86568153633767, -73.92350095571744)</td>\n",  
 " </tr>\n",  
 " <tr>\n",  
 " <th>1</th>\n",  
 " <td>32309934</td>\n",  
 " <td>12/31/2015 11:59:44 PM</td>\n",  
 " <td>01-01-16 1:26</td>\n",  
 " <td>NYPD</td>\n",  
 " <td>New York City Police Department</td>\n",  
 " <td>Blocked Driveway</td>\n",  
 " <td>No Access</td>\n",  
 " <td>Street/Sidewalk</td>\n",  
 " <td>11105.0</td>\n",  
 " <td>27-07 23 AVENUE</td>\n",  
 " <td>...</td>\n",  
 " <td>NaN</td>\n",  
 " <td>NaN</td>\n",  
 " <td>NaN</td>\n",  
 " <td>NaN</td>\n",  
 " <td>NaN</td>\n",  
 " <td>NaN</td>\n",  
 " <td>NaN</td>\n",  
 " <td>40.775945</td>\n",  
 " <td>-73.915094</td>\n",  
 " <td>(40.775945312321085, -73.91509393898605)</td>\n",  
 " </tr>\n",  
 " <tr>\n",  
 " <th>2</th>\n",  
 " <td>32309159</td>\n",  
 " <td>12/31/2015 11:59:29 PM</td>\n",  
 " <td>01-01-16 4:51</td>\n",  
 " <td>NYPD</td>\n",  
 " <td>New York City Police Department</td>\n",  
 " <td>Blocked Driveway</td>\n",  
 " <td>No Access</td>\n",  
 " <td>Street/Sidewalk</td>\n",  
 " <td>10458.0</td>\n",  
 " <td>2897 VALENTINE AVENUE</td>\n",  
 " <td>...</td>\n",  
 " <td>NaN</td>\n",  
 " <td>NaN</td>\n",  
 " <td>NaN</td>\n",  
 " <td>NaN</td>\n",  
 " <td>NaN</td>\n",  
 " <td>NaN</td>\n",  
 " <td>NaN</td>\n",  
 " <td>40.870325</td>\n",  
 " <td>-73.888525</td>\n",  
 " <td>(40.870324522111424, -73.88852464418646)</td>\n",  
 " </tr>\n",  
 " <tr>\n",  
 " <th>3</th>\n",  
 " <td>32305098</td>\n",  
 " <td>12/31/2015 11:57:46 PM</td>\n",  
 " <td>01-01-16 7:43</td>\n",  
 " <td>NYPD</td>\n",  
 " <td>New York City Police Department</td>\n",  
 " <td>Illegal Parking</td>\n",  
 " <td>Commercial Overnight Parking</td>\n",  
 " <td>Street/Sidewalk</td>\n",  
 " <td>10461.0</td>\n",  
 " <td>2940 BAISLEY AVENUE</td>\n",  
 " <td>...</td>\n",  
 " <td>NaN</td>\n",  
 " <td>NaN</td>\n",  
 " <td>NaN</td>\n",  
 " <td>NaN</td>\n",  
 " <td>NaN</td>\n",  
 " <td>NaN</td>\n",  
 " <td>NaN</td>\n",  
 " <td>40.835994</td>\n",  
 " <td>-73.828379</td>\n",  
 " <td>(40.83599404683083, -73.82837939584206)</td>\n",  
 " </tr>\n",  
 " <tr>\n",  
 " <th>4</th>\n",  
 " <td>32306529</td>\n",  
 " <td>12/31/2015 11:56:58 PM</td>\n",  
 " <td>01-01-16 3:24</td>\n",  
 " <td>NYPD</td>\n",  
 " <td>New York City Police Department</td>\n",  
 " <td>Illegal Parking</td>\n",  
 " <td>Blocked Sidewalk</td>\n",  
 " <td>Street/Sidewalk</td>\n",  
 " <td>11373.0</td>\n",  
 " <td>87-14 57 ROAD</td>\n",  
 " <td>...</td>\n",  
 " <td>NaN</td>\n",  
 " <td>NaN</td>\n",  
 " <td>NaN</td>\n",  
 " <td>NaN</td>\n",  
 " <td>NaN</td>\n",  
 " <td>NaN</td>\n",  
 " <td>NaN</td>\n",  
 " <td>40.733060</td>\n",  
 " <td>-73.874170</td>\n",  
 " <td>(40.733059618956815, -73.87416975810375)</td>\n",  
 " </tr>\n",  
 " </tbody>\n",  
 "</table>\n",  
 "<p>5 rows × 53 columns</p>\n",  
 "</div>"  
 ],  
 "text/plain": [  
 " Unique Key Created Date Closed Date Agency \\\n",  
 "0 32310363 12/31/2015 11:59:45 PM 01-01-16 0:55 NYPD \n",  
 "1 32309934 12/31/2015 11:59:44 PM 01-01-16 1:26 NYPD \n",  
 "2 32309159 12/31/2015 11:59:29 PM 01-01-16 4:51 NYPD \n",  
 "3 32305098 12/31/2015 11:57:46 PM 01-01-16 7:43 NYPD \n",  
 "4 32306529 12/31/2015 11:56:58 PM 01-01-16 3:24 NYPD \n",  
 "\n",  
 " Agency Name Complaint Type \\\n",  
 "0 New York City Police Department Noise - Street/Sidewalk \n",  
 "1 New York City Police Department Blocked Driveway \n",  
 "2 New York City Police Department Blocked Driveway \n",  
 "3 New York City Police Department Illegal Parking \n",  
 "4 New York City Police Department Illegal Parking \n",  
 "\n",  
 " Descriptor Location Type Incident Zip \\\n",  
 "0 Loud Music/Party Street/Sidewalk 10034.0 \n",  
 "1 No Access Street/Sidewalk 11105.0 \n",  
 "2 No Access Street/Sidewalk 10458.0 \n",  
 "3 Commercial Overnight Parking Street/Sidewalk 10461.0 \n",  
 "4 Blocked Sidewalk Street/Sidewalk 11373.0 \n",  
 "\n",  
 " Incident Address ... Bridge Highway Name Bridge Highway Direction \\\n",  
 "0 71 VERMILYEA AVENUE ... NaN NaN \n",  
 "1 27-07 23 AVENUE ... NaN NaN \n",  
 "2 2897 VALENTINE AVENUE ... NaN NaN \n",  
 "3 2940 BAISLEY AVENUE ... NaN NaN \n",  
 "4 87-14 57 ROAD ... NaN NaN \n",  
 "\n",  
 " Road Ramp Bridge Highway Segment Garage Lot Name Ferry Direction \\\n",  
 "0 NaN NaN NaN NaN \n",  
 "1 NaN NaN NaN NaN \n",  
 "2 NaN NaN NaN NaN \n",  
 "3 NaN NaN NaN NaN \n",  
 "4 NaN NaN NaN NaN

\n",  
 "\n",  
 " Ferry Terminal Name Latitude Longitude \\\n",  
 "0 NaN 40.865682 -73.923501 \n",  
 "1 NaN 40.775945 -73.915094 \n",  
 "2 NaN 40.870325 -73.888525 \n",  
 "3 NaN 40.835994 -73.828379 \n",  
 "4 NaN 40.733060 -73.874170 \n",  
 "\n",  
 " Location \n",  
 "0 (40.86568153633767, -73.92350095571744) \n",  
 "1 (40.775945312321085, -73.91509393898605) \n",  
 "2 (40.870324522111424, -73.88852464418646) \n",  
 "3 (40.83599404683083, -73.82837939584206) \n",  
 "4 (40.733059618956815, -73.87416975810375) \n",  
 "\n",  
 "[5 rows x 53 columns]"  
 ]  
 },  
 "execution\_count": 15,  
 "metadata": {},  
 "output\_type": "execute\_result"  
 }  
 ],  
 "source": [  
 "df.head()"  
 ]  
 },  
 {  
 "cell\_type": "code",  
 "execution\_count": 16,  
 "id": "33353669",  
 "metadata": {},  
 "outputs": [  
 {  
 "data": {  
 "text/html": [  
 "<div>\n",  
 "<style scoped>\n",  
 " .dataframe tbody tr th:only-of-type {\n",  
 " vertical-align: middle;\n",  
 " }\n",  
 "\n",  
 " .dataframe tbody tr th {\n",  
 " vertical-align: top;\n",  
 " }\n",  
 "\n",  
 " .dataframe thead th {\n",  
 " text-align: right;\n",  
 " }\n",  
 "</style>\n",  
 "<table border=\"1\" class=\"dataframe\">\n",  
 " <thead>\n",  
 " <tr style=\"text-align: right;\">\n",  
 " <th></th>\n",  
 " <th>Unique Key</th>\n",  
 " <th>Incident Zip</th>\n",  
 " <th>X Coordinate (State Plane)</th>\n",  
 " <th>Y Coordinate (State Plane)</th>\n",  
 " <th>School or Citywide Complaint</th>\n",  
 " <th>Vehicle Type</th>\n",  
 " <th>Taxi Company Borough</th>\n",  
 " <th>Taxi Pick Up Location</th>\n",  
 " <th>Garage Lot Name</th>\n",  
 " <th>Latitude</th>\n",  
 " <th>Longitude</th>\n",  
 " </tr>\n",  
 " </thead>\n",  
 " <tbody>\n",  
 " <tr>\n",  
 " <th>count</th>\n",  
 " <td>3.006980e+05</td>\n",  
 " <td>298083.000000</td>\n",  
 " <td>2.971580e+05</td>\n",  
 " <td>297158.000000</td>\n",  
 " <td>0.0</td>\n",  
 " <td>0.0</td>\n",  
 " <td>0.0</td>\n",  
 " <td>0.0</td>\n",  
 " <td>0.0</td>\n",  
 " <td>297158.000000</td>\n",  
 " <td>297158.000000</td>\n",  
 " </tr>\n",  
 " <tr>\n",  
 " <th>mean</th>\n",  
 " <td>3.130054e+07</td>\n",  
 " <td>10848.888645</td>\n",  
 " <td>1.004854e+06</td>\n",  
 " <td>203754.534416</td>\n",  
 " <td>NaN</td>\n",  
 " <td>NaN</td>\n",  
 " <td>NaN</td>\n",  
 " <td>NaN</td>\n",  
 " <td>NaN</td>\n",  
 " <td>40.725885</td>\n",  
 " <td>-73.925630</td>\n",  
 " </tr>\n",  
 " <tr>\n",  
 " <th>std</th>\n",  
 " <td>5.738547e+05</td>\n",  
 " <td>583.182081</td>\n",  
 " <td>2.175338e+04</td>\n",  
 " <td>29880.183529</td>\n",  
 " <td>NaN</td>\n",  
 " <td>NaN</td>\n",  
 " <td>NaN</td>\n",  
 " <td>NaN</td>\n",  
 " <td>NaN</td>\n",  
 " <td>0.082012</td>\n",  
 " <td>0.078454</td>\n",  
 " </tr>\n",  
 " <tr>\n",  
 " <th>min</th>\n",  
 " <td>3.027948e+07</td>\n",  
 " <td>83.000000</td>\n",  
 " <td>9.133570e+05</td>\n",  
 " <td>121219.000000</td>\n",  
 " <td>NaN</td>\n",  
 " <td>NaN</td>\n",  
 " <td>NaN</td>\n",  
 " <td>NaN</td>\n",  
 " <td>NaN</td>\n",  
 " <td>40.499135</td>\n",  
 " <td>-74.254937</td>\n",  
 " </tr>\n",  
 " <tr>\n",  
 " <th>25%</th>\n",  
 " <td>3.080118e+07</td>\n",  
 " <td>10310.000000</td>\n",  
 " <td>9.919752e+05</td>\n",  
 " <td>183343.000000</td>\n",  
 " <td>NaN</td>\n",  
 " <td>NaN</td>\n",  
 " <td>NaN</td>\n",  
 " <td>NaN</td>\n",  
 " <td>NaN</td>\n",  
 " <td>40.669796</td>\n",  
 " <td>-73.972142</td>\n",  
 " </tr>\n",  
 " <tr>\n",  
 " <th>50%</th>\n",  
 " <td>3.130436e+07</td>\n",  
 " <td>11208.000000</td>\n",  
 " <td>1.003158e+06</td>\n",  
 " <td>201110.500000</td>\n",  
 " <td>NaN</td>\n",  
 " <td>NaN</td>\n",  
 " <td>NaN</td>\n",  
 " <td>NaN</td>\n",  
 " <td>NaN</td>\n",  
 " <td>40.718661</td>\n",  
 " <td>-73.931781</td>\n",  
 " </tr>\n",  
 " <tr>\n",  
 " <th>75%</th>\n",  
 " <td>3.178446e+07</td>\n",  
 " <td>11238.000000</td>\n",  
 " <td>1.018372e+06</td>\n",  
 " <td>224125.250000</td>\n",  
 " <td>NaN</td>\n",  
 " <td>NaN</td>\n",  
 " <td>NaN</td>\n",  
 " <td>NaN</td>\n",  
 " <td>NaN</td>\n",  
 " <td>40.781840</td>\n",  
 " <td>-73.876805</td>\n",  
 " </tr>\n",  
 " <tr>\n",  
 " <th>max</th>\n",  
 " <td>3.231065e+07</td>\n",  
 " <td>11697.000000</td>\n",  
 " <td>1.067173e+06</td>\n",  
 " <td>271876.000000</td>\n",  
 " <td>NaN</td>\n",  
 " <td>NaN</td>\n",  
 " <td>NaN</td>\n",  
 " <td>NaN</td>\n",  
 " <td>NaN</td>\n",  
 " <td>40.912869</td>\n",  
 " <td>-73.700760</td>\n",  
 " </tr>\n",  
 " </tbody>\n",  
 "</table>\n",  
 "</div>"  
 ],  
 "text/plain": [  
 " Unique Key Incident Zip X Coordinate (State Plane) \\\n",  
 "count 3.006980e+05 298083.000000 2.971580e+05 \n",  
 "mean 3.130054e+07 10848.888645 1.004854e+06 \n",  
 "std 5.738547e+05 583.182081 2.175338e+04 \n",  
 "min 3.027948e+07 83.000000 9.133570e+05 \n",  
 "25% 3.080118e+07 10310.000000 9.919752e+05 \n",  
 "50% 3.130436e+07 11208.000000 1.003158e+06 \n",  
 "75% 3.178446e+07 11238.000000 1.018372e+06 \n",  
 "max 3.231065e+07 11697.000000 1.067173e+06 \n",  
 "\n",  
 " Y Coordinate (State Plane) School or Citywide Complaint Vehicle Type \\\n",  
 "count 297158.000000 0.0 0.0 \n",  
 "mean 203754.534416 NaN NaN \n",  
 "std 29880.183529 NaN NaN \n",  
 "min 121219.000000 NaN NaN \n",  
 "25% 183343.000000 NaN NaN \n",  
 "50% 201110.500000 NaN NaN \n",  
 "75% 224125.250000 NaN NaN \n",  
 "max 271876.000000 NaN NaN \n",  
 "\n",  
 " Taxi Company Borough Taxi Pick Up Location Garage Lot Name \\\n",  
 "count 0.0 0.0 0.0 \n",  
 "mean NaN NaN NaN \n",  
 "std NaN NaN NaN \n",  
 "min NaN NaN NaN \n",  
 "25% NaN NaN NaN \n",  
 "50% NaN NaN NaN \n",  
 "75% NaN NaN NaN \n",  
 "max NaN NaN NaN \n",  
 "\n",  
 " Latitude Longitude \n",  
 "count 297158.000000 297158.000000 \n",  
 "mean 40.725885 -73.925630 \n",  
 "std 0.082012 0.078454 \n",  
 "min 40.499135 -74.254937 \n",  
 "25% 40.669796 -73.972142 \n",  
 "50% 40.718661 -73.931781 \n",  
 "75% 40.781840 -73.876805 \n",  
 "max 40.912869 -73.700760 "  
 ]  
 },  
 "execution\_count": 16,  
 "metadata": {},  
 "output\_type": "execute\_result"  
 }  
 ],  
 "source": [  
 "#Understanding the data\n",  
 "df.describe()"  
 ]  
 },  
 {  
 "cell\_type": "code",  
 "execution\_count": 17,  
 "id": "09128c12",  
 "metadata": {},  
 "outputs": [  
 {  
 "data": {  
 "text/plain": [  
 "(300698, 53)"  
 ]  
 },  
 "execution\_count": 17,  
 "metadata": {},  
 "output\_type": "execute\_result"  
 }  
 ],  
 "source": [  
 "df.shape"  
 ]  
 },  
 {  
 "cell\_type": "code",  
 "execution\_count": 18,  
 "id": "41e101bb",  
 "metadata": {},  
 "outputs": [],  
 "source": [  
 "#Conversion to datetime format\n",  
 "df[\"Created Date\"]=pd.to\_datetime(df[\"Created Date\"])\n",  
 "df[\"Closed Date\"]=pd.to\_datetime(df[\"Closed Date\"])"  
 ]  
 },  
 {  
 "cell\_type": "code",  
 "execution\_count": 19,  
 "id": "dccdb2e8",  
 "metadata": {},  
 "outputs": [],  
 "source": [  
 "#Creating a new column ‘Request\_Closing\_Time’ as the time elapsed between request creation and request closing\n",  
 "df[\"Request\_Closing\_Time\"]=(df[\"Closed

Date\"]-df[\"Created Date\"])\n",  
 "Request\_Closing\_Time=[]\n",  
 "for x in (df[\"Closed Date\"]-df[\"Created Date\"]):\n",  
 " close=x.total\_seconds()/60\n",  
 " Request\_Closing\_Time.append(close)\n",  
 "df[\"Request\_Closing\_Time\"]=Request\_Closing\_Time"  
 ]  
 },  
 {  
 "cell\_type": "code",  
 "execution\_count": 20,  
 "id": "53017186",  
 "metadata": {},  
 "outputs": [  
 {  
 "data": {  
 "text/html": [  
 "<div>\n",  
 "<style scoped>\n",  
 " .dataframe tbody tr th:only-of-type {\n",  
 " vertical-align: middle;\n",  
 " }\n",  
 "\n",  
 " .dataframe tbody tr th {\n",  
 " vertical-align: top;\n",  
 " }\n",  
 "\n",  
 " .dataframe thead th {\n",  
 " text-align: right;\n",  
 " }\n",  
 "</style>\n",  
 "<table border=\"1\" class=\"dataframe\">\n",  
 " <thead>\n",  
 " <tr style=\"text-align: right;\">\n",  
 " <th></th>\n",  
 " <th>Unique Key</th>\n",  
 " <th>Created Date</th>\n",  
 " <th>Closed Date</th>\n",  
 " <th>Agency</th>\n",  
 " <th>Agency Name</th>\n",  
 " <th>Complaint Type</th>\n",  
 " <th>Descriptor</th>\n",  
 " <th>Location Type</th>\n",  
 " <th>Incident Zip</th>\n",  
 " <th>Incident Address</th>\n",  
 " <th>...</th>\n",  
 " <th>Bridge Highway Direction</th>\n",  
 " <th>Road Ramp</th>\n",  
 " <th>Bridge Highway Segment</th>\n",  
 " <th>Garage Lot Name</th>\n",  
 " <th>Ferry Direction</th>\n",  
 " <th>Ferry Terminal Name</th>\n",  
 " <th>Latitude</th>\n",  
 " <th>Longitude</th>\n",  
 " <th>Location</th>\n",  
 " <th>Request\_Closing\_Time</th>\n",  
 " </tr>\n",  
 " </thead>\n",  
 " <tbody>\n",  
 " <tr>\n",  
 " <th>0</th>\n",  
 " <td>32310363</td>\n",  
 " <td>2015-12-31 23:59:45</td>\n",  
 " <td>2016-01-01 00:55:00</td>\n",  
 " <td>NYPD</td>\n",  
 " <td>New York City Police Department</td>\n",  
 " <td>Noise - Street/Sidewalk</td>\n",  
 " <td>Loud Music/Party</td>\n",  
 " <td>Street/Sidewalk</td>\n",  
 " <td>10034.0</td>\n",  
 " <td>71 VERMILYEA AVENUE</td>\n",  
 " <td>...</td>\n",  
 " <td>NaN</td>\n",  
 " <td>NaN</td>\n",  
 " <td>NaN</td>\n",  
 " <td>NaN</td>\n",  
 " <td>NaN</td>\n",  
 " <td>NaN</td>\n",  
 " <td>40.865682</td>\n",  
 " <td>-73.923501</td>\n",  
 " <td>(40.86568153633767, -73.92350095571744)</td>\n",  
 " <td>55.250000</td>\n",  
 " </tr>\n",  
 " <tr>\n",  
 " <th>1</th>\n",  
 " <td>32309934</td>\n",  
 " <td>2015-12-31 23:59:44</td>\n",  
 " <td>2016-01-01 01:26:00</td>\n",  
 " <td>NYPD</td>\n",  
 " <td>New York City Police Department</td>\n",  
 " <td>Blocked Driveway</td>\n",  
 " <td>No Access</td>\n",  
 " <td>Street/Sidewalk</td>\n",  
 " <td>11105.0</td>\n",  
 " <td>27-07 23 AVENUE</td>\n",  
 " <td>...</td>\n",  
 " <td>NaN</td>\n",  
 " <td>NaN</td>\n",  
 " <td>NaN</td>\n",  
 " <td>NaN</td>\n",  
 " <td>NaN</td>\n",  
 " <td>NaN</td>\n",  
 " <td>40.775945</td>\n",  
 " <td>-73.915094</td>\n",  
 " <td>(40.775945312321085, -73.91509393898605)</td>\n",  
 " <td>86.266667</td>\n",  
 " </tr>\n",  
 " <tr>\n",  
 " <th>2</th>\n",  
 " <td>32309159</td>\n",  
 " <td>2015-12-31 23:59:29</td>\n",  
 " <td>2016-01-01 04:51:00</td>\n",  
 " <td>NYPD</td>\n",  
 " <td>New York City Police Department</td>\n",  
 " <td>Blocked Driveway</td>\n",  
 " <td>No Access</td>\n",  
 " <td>Street/Sidewalk</td>\n",  
 " <td>10458.0</td>\n",  
 " <td>2897 VALENTINE AVENUE</td>\n",  
 " <td>...</td>\n",  
 " <td>NaN</td>\n",  
 " <td>NaN</td>\n",  
 " <td>NaN</td>\n",  
 " <td>NaN</td>\n",  
 " <td>NaN</td>\n",  
 " <td>NaN</td>\n",  
 " <td>40.870325</td>\n",  
 " <td>-73.888525</td>\n",  
 " <td>(40.870324522111424, -73.88852464418646)</td>\n",  
 " <td>291.516667</td>\n",  
 " </tr>\n",  
 " <tr>\n",  
 " <th>3</th>\n",  
 " <td>32305098</td>\n",  
 " <td>2015-12-31 23:57:46</td>\n",  
 " <td>2016-01-01 07:43:00</td>\n",  
 " <td>NYPD</td>\n",  
 " <td>New York City Police Department</td>\n",  
 " <td>Illegal Parking</td>\n",  
 " <td>Commercial Overnight Parking</td>\n",  
 " <td>Street/Sidewalk</td>\n",  
 " <td>10461.0</td>\n",  
 " <td>2940 BAISLEY AVENUE</td>\n",  
 " <td>...</td>\n",  
 " <td>NaN</td>\n",  
 " <td>NaN</td>\n",  
 " <td>NaN</td>\n",  
 " <td>NaN</td>\n",  
 " <td>NaN</td>\n",  
 " <td>NaN</td>\n",  
 " <td>40.835994</td>\n",  
 " <td>-73.828379</td>\n",  
 " <td>(40.83599404683083, -73.82837939584206)</td>\n",  
 " <td>465.233333</td>\n",  
 " </tr>\n",  
 " <tr>\n",  
 " <th>4</th>\n",  
 " <td>32306529</td>\n",  
 " <td>2015-12-31 23:56:58</td>\n",  
 " <td>2016-01-01 03:24:00</td>\n",  
 " <td>NYPD</td>\n",  
 " <td>New York City Police Department</td>\n",  
 " <td>Illegal Parking</td>\n",  
 " <td>Blocked Sidewalk</td>\n",  
 " <td>Street/Sidewalk</td>\n",  
 " <td>11373.0</td>\n",  
 " <td>87-14 57 ROAD</td>\n",  
 " <td>...</td>\n",  
 " <td>NaN</td>\n",  
 " <td>NaN</td>\n",  
 " <td>NaN</td>\n",  
 " <td>NaN</td>\n",  
 " <td>NaN</td>\n",  
 " <td>NaN</td>\n",  
 " <td>40.733060</td>\n",  
 " <td>-73.874170</td>\n",  
 " <td>(40.733059618956815, -73.87416975810375)</td>\n",  
 " <td>207.033333</td>\n",  
 " </tr>\n",  
 " </tbody>\n",  
 "</table>\n",  
 "<p>5 rows × 54 columns</p>\n",  
 "</div>"  
 ],  
 "text/plain": [  
 " Unique Key Created Date Closed Date Agency \\\n",  
 "0 32310363 2015-12-31 23:59:45 2016-01-01 00:55:00 NYPD \n",  
 "1 32309934 2015-12-31 23:59:44 2016-01-01 01:26:00 NYPD \n",  
 "2 32309159 2015-12-31 23:59:29 2016-01-01 04:51:00 NYPD \n",  
 "3 32305098 2015-12-31 23:57:46 2016-01-01 07:43:00 NYPD \n",  
 "4 32306529 2015-12-31 23:56:58 2016-01-01 03:24:00 NYPD \n",  
 "\n",  
 " Agency Name Complaint Type \\\n",  
 "0 New York City Police Department Noise - Street/Sidewalk \n",  
 "1 New York City Police Department Blocked Driveway \n",  
 "2 New York City Police Department Blocked Driveway \n",  
 "3 New York City Police Department Illegal Parking \n",  
 "4 New York City Police Department Illegal Parking \n",  
 "\n",  
 " Descriptor Location Type Incident Zip \\\n",  
 "0 Loud Music/Party Street/Sidewalk 10034.0 \n",  
 "1 No Access Street/Sidewalk 11105.0 \n",  
 "2 No Access Street/Sidewalk 10458.0 \n",  
 "3 Commercial Overnight Parking Street/Sidewalk 10461.0 \n",  
 "4 Blocked Sidewalk Street/Sidewalk 11373.0 \n",  
 "\n",  
 " Incident Address ... Bridge Highway Direction Road Ramp \\\n",  
 "0 71 VERMILYEA AVENUE ... NaN NaN \n",  
 "1 27-07 23 AVENUE ... NaN NaN \n",  
 "2 2897 VALENTINE AVENUE ... NaN NaN \n",  
 "3 2940 BAISLEY AVENUE ... NaN NaN \n",  
 "4 87-14 57 ROAD ... NaN NaN \n",  
 "\n",  
 " Bridge Highway Segment Garage Lot Name Ferry Direction Ferry Terminal Name \\\n",  
 "0 NaN NaN NaN NaN \n",  
 "1 NaN NaN NaN NaN \n",  
 "2 NaN NaN NaN NaN \n",  
 "3 NaN NaN NaN NaN \n",  
 "4 NaN NaN NaN NaN \n",  
 "\n",  
 " Latitude Longitude Location \\\n",  
 "0 40.865682 -73.923501 (40.86568153633767, -73.92350095571744) \n",  
 "1 40.775945 -73.915094 (40.775945312321085, -73.91509393898605) \n",  
 "2 40.870325 -73.888525 (40.870324522111424, -73.88852464418646) \n",  
 "3 40.835994 -73.828379 (40.83599404683083, -73.82837939584206) \n",  
 "4 40.733060 -73.874170 (40.733059618956815, -73.87416975810375) \n",  
 "\n",  
 " Request\_Closing\_Time \n",  
 "0 55.250000 \n",  
 "1 86.266667 \n",  
 "2 291.516667 \n",  
 "3 465.233333 \n",  
 "4 207.033333 \n",  
 "\n",  
 "[5 rows x 54 columns]"  
 ]  
 },  
 "execution\_count": 20,  
 "metadata": {},  
 "output\_type": "execute\_result"

}  
 ],  
 "source": [  
 "df.head()"  
 ]  
 },  
 {  
 "cell\_type": "code",  
 "execution\_count": 21,  
 "id": "349008d2",  
 "metadata": {},  
 "outputs": [  
 {  
 "data": {  
 "text/plain": [  
 "array(['NYPD'], dtype=object)"  
 ]  
 },  
 "execution\_count": 21,  
 "metadata": {},  
 "output\_type": "execute\_result"  
 }  
 ],  
 "source": [  
 "#EDA\n",  
 "df[\"Agency\"].unique()"  
 ]  
 },  
 {  
 "cell\_type": "markdown",  
 "id": "46d7c127",  
 "metadata": {},  
 "source": [  
 "We can see the above data belongs to the NYPD."  
 ]  
 },  
 {  
 "cell\_type": "code",  
 "execution\_count": 22,  
 "id": "2a233b88",  
 "metadata": {},  
 "outputs": [  
 {  
 "name": "stderr",  
 "output\_type": "stream",  
 "text": [  
 "C:\\Users\\lrnem\\AppData\\Local\\Temp\\ipykernel\_8728\\5426915.py:1: UserWarning: \n",  
 "\n",  
 "`distplot` is a deprecated function and will be removed in seaborn v0.14.0.\n",  
 "\n",  
 "Please adapt your code to use either `displot` (a figure-level function with\n",  
 "similar flexibility) or `histplot` (an axes-level function for histograms).\n",  
 "\n",  
 "For a guide to updating your code to use the new functions, please see\n",  
 "https://gist.github.com/mwaskom/de44147ed2974457ad6372750bbe5751\n",  
 "\n",  
 " sns.distplot(df[\"Request\_Closing\_Time\"])\n"  
 ]  
 },  
 {  
 "data": {  
 "text/plain": [  
 "<function matplotlib.pyplot.show(close=None, block=None)>"  
 ]  
 },  
 "execution\_count": 22,  
 "metadata": {},  
 "output\_type": "execute\_result"  
 },  
 {  
 "data": {  
 "image/png": "





SqXCCy+8gOHDh+Phhx+GWq3GoUOH8H//939YuXIlAPd/D4jaAwYqInJb1pmypjz00EO3PKm8PevSpQs2bdqEN954AzNnzkR1dTXCw8OxevVqPPDAAx7xHhC1B1zyIyK3ZTKZ8P333zfZLigoqMUnpLd3fA+I5MFARURERNRKXDAnIiIiaiUGKiIiIqJWYqAiIiIiaiUGKiIiIqJWYqAiIiIiaiUGKiIiIqJWYqAiIiIiaqX/D7pBNqZwe2WbAAAAAElFTkSuQmCC",  
 "text/plain": [  
 "<Figure size 640x480 with 1 Axes>"  
 ]  
 },  
 "metadata": {},  
 "output\_type": "display\_data"  
 }  
 ],  
 "source": [  
 "sns.distplot(df[\"Request\_Closing\_Time\"])\n",  
 "plt.show"  
 ]  
 },  
 {  
 "cell\_type": "code",  
 "execution\_count": 23,  
 "id": "a7ed8f0c",  
 "metadata": {},  
 "outputs": [  
 {  
 "name": "stdout",  
 "output\_type": "stream",  
 "text": [  
 "Total Number of Complaints : 300698 \n",  
 "\n",  
 "Percentage of complaints that took 99 hours or less : 32.88 %\n",  
 "Percentage of complaints that took 999 hours or less : 97.18 %\n"  
 ]  
 }  
 ],  
 "source": [  
 "print(\"Total Number of Complaints : \",len(df),\"\\n\")\n",  
 "print(\"Percentage of complaints that took 99 hours or less : \",round((len(df)-(df[\"Request\_Closing\_Time\"]>=99).sum())/len(df)\*100,2),\"%\")\n",  
 "print(\"Percentage of complaints that took 999 hours or less : \",round((len(df)-(df[\"Request\_Closing\_Time\"]>=999).sum())/len(df)\*100,2),\"%\")"  
 ]  
 },  
 {  
 "cell\_type": "markdown",  
 "id": "85d5304f",  
 "metadata": {},  
 "source": [  
 "From the above data we can see that majority of the complaints needed more than 99 hours to be dealt with."  
 ]  
 },  
 {  
 "cell\_type": "code",  
 "execution\_count": 24,  
 "id": "f9ecfdde",  
 "metadata": {  
 "scrolled": true  
 },  
 "outputs": [  
 {  
 "name": "stderr",  
 "output\_type": "stream",  
 "text": [  
 "C:\\Users\\lrnem\\AppData\\Local\\Temp\\ipykernel\_8728\\2810770643.py:1: UserWarning: \n",  
 "\n",  
 "`distplot` is a deprecated function and will be removed in seaborn v0.14.0.\n",  
 "\n",  
 "Please adapt your code to use either `displot` (a figure-level function with\n",  
 "similar flexibility) or `histplot` (an axes-level function for histograms).\n",  
 "\n",  
 "For a guide to updating your code to use the new functions, please see\n",  
 "https://gist.github.com/mwaskom/de44147ed2974457ad6372750bbe5751\n",  
 "\n",  
 " sns.distplot(df[\"Request\_Closing\_Time\"])\n"  
 ]  
 },  
 {  
 "data": {  
 "image/png": "





",  
 "text/plain": [  
 "<Figure size 640x480 with 1 Axes>"  
 ]  
 },  
 "metadata": {},  
 "output\_type": "display\_data"  
 }  
 ],  
 "source": [  
 "sns.distplot(df[\"Request\_Closing\_Time\"])\n",  
 "plt.xlim((0,5000))\n",  
 "plt.ylim((0,0.0003))\n",  
 "plt.show()"  
 ]  
 },  
 {  
 "cell\_type": "markdown",  
 "id": "6541e3f8",  
 "metadata": {},  
 "source": [  
 "Now let us try to understand the major complaint types"  
 ]  
 },  
 {  
 "cell\_type": "code",  
 "execution\_count": 25,  
 "id": "9257bbce",  
 "metadata": {},  
 "outputs": [  
 {  
 "data": {  
 "image/png": "









",  
 "text/plain": [  
 "<Figure size 1500x1000 with 1 Axes>"  
 ]  
 },  
 "metadata": {},  
 "output\_type": "display\_data"  
 }  
 ],  
 "source": [  
 "df['Complaint Type'].value\_counts()[:10].plot(kind='barh',alpha=0.6,figsize=(15,10))\n",  
 "plt.show()"  
 ]  
 },  
 {  
 "cell\_type": "markdown",  
 "id": "ce4a03d0",  
 "metadata": {},  
 "source": [  
 "From the above graph we can see that majority of complaints are related to that of transporttion and parking and as such it needed more time to be resolved."  
 ]  
 },  
 {  
 "cell\_type": "code",  
 "execution\_count": 26,  
 "id": "423c7dd5",  
 "metadata": {},  
 "outputs": [  
 {  
 "data": {  
 "image/png": "





























",  
 "text/plain": [  
 "<Figure size 1500x700 with 1 Axes>"  
 ]  
 },  
 "metadata": {},  
 "output\_type": "display\_data"  
 }  
 ],  
 "source": [  
 "crt = sns.catplot(x='Complaint Type', y=\"Request\_Closing\_Time\",data=df)\n",  
 "crt.fig.set\_figwidth(15)\n",  
 "crt.fig.set\_figheight(7)\n",  
 "plt.xticks(rotation=90)\n",  
 "plt.ylim((0,5000))\n",  
 "plt.show()"  
 ]  
 },  
 {  
 "cell\_type": "markdown",  
 "id": "a0858fee",  
 "metadata": {},  
 "source": [  
 "We can see from the above lot that major complaints arises from transport and take up huge time to be resolved and as such the government should take steps to improve the transport system and implement stricter vehicle laws."  
 ]  
 },  
 {  
 "cell\_type": "code",  
 "execution\_count": 27,  
 "id": "aa3760f6",  
 "metadata": {},  
 "outputs": [  
 {  
 "data": {  
 "image/png": "



",  
 "text/plain": [  
 "<Figure size 1500x700 with 1 Axes>"  
 ]  
 },  
 "metadata": {},  
 "output\_type": "display\_data"  
 }  
 ],  
 "source": [  
 "df['Status'].value\_counts().plot(kind='bar',alpha=0.6,figsize=(15,7))\n",  
 "plt.show()"  
 ]  
 },  
 {  
 "cell\_type": "markdown",  
 "id": "3722c24a",  
 "metadata": {},  
 "source": [  
 "From the above plot we can see all complaints are now closed that is they are resolved."  
 ]  
 },  
 {  
 "cell\_type": "code",  
 "execution\_count": 28,  
 "id": "2126b693",  
 "metadata": {},  
 "outputs": [  
 {  
 "data": {  
 "image/png": "





",  
 "text/plain": [  
 "<Figure size 1200x700 with 1 Axes>"  
 ]  
 },  
 "metadata": {},  
 "output\_type": "display\_data"  
 }  
 ],  
 "source": [  
 "plt.figure(figsize=(12,7))\n",  
 "df['Borough'].value\_counts().plot(kind='bar',alpha=0.7)\n",  
 "plt.show()"  
 ]  
 },  
 {  
 "cell\_type": "code",  
 "execution\_count": 29,  
 "id": "0fcd5266",  
 "metadata": {},  
 "outputs": [  
 {  
 "name": "stdout",  
 "output\_type": "stream",  
 "text": [  
 "Percentage of complaints from MANHATTAN : 21.99\n",  
 "Percentage of complaints from QUEENS : 26.82\n",  
 "Percentage of complaints from

BRONX : 13.54\n",  
 "Percentage of complaints from BROOKLYN : 32.69\n",  
 "Percentage of complaints from Unspecified : 0.86\n",  
 "Percentage of complaints from STATEN ISLAND : 4.1\n"  
 ]  
 }  
 ],  
 "source": [  
 "for x in df[\"Borough\"].unique():\n",  
 " print(\"Percentage of complaints from \",x,\" : \",round((df[\"Borough\"]==x).sum()/len(df)\*100,2))"  
 ]  
 },  
 {  
 "cell\_type": "code",  
 "execution\_count": 30,  
 "id": "abbd2c19",  
 "metadata": {},  
 "outputs": [  
 {  
 "data": {  
 "text/plain": [  
 "array(['Street/Sidewalk', 'Club/Bar/Restaurant', 'Store/Commercial',\n",  
 " 'House of Worship', 'Residential Building/House',\n",  
 " 'Residential Building', 'Park/Playground', 'Vacant Lot',\n",  
 " 'House and Store', 'Highway', 'Commercial', 'Roadway Tunnel',\n",  
 " 'Subway Station', 'Parking Lot', 'Bridge', 'Terminal', nan,\n",  
 " 'Ferry', 'Park'], dtype=object)"  
 ]  
 },  
 "execution\_count": 30,  
 "metadata": {},  
 "output\_type": "execute\_result"  
 }  
 ],  
 "source": [  
 "#All unique locations\n",  
 "df[\"Location Type\"].unique()"  
 ]  
 },  
 {  
 "cell\_type": "code",  
 "execution\_count": 31,  
 "id": "cf6f2139",  
 "metadata": {},  
 "outputs": [  
 {  
 "data": {  
 "text/html": [  
 "<div>\n",  
 "<style scoped>\n",  
 " .dataframe tbody tr th:only-of-type {\n",  
 " vertical-align: middle;\n",  
 " }\n",  
 "\n",  
 " .dataframe tbody tr th {\n",  
 " vertical-align: top;\n",  
 " }\n",  
 "\n",  
 " .dataframe thead th {\n",  
 " text-align: right;\n",  
 " }\n",  
 "</style>\n",  
 "<table border=\"1\" class=\"dataframe\">\n",  
 " <thead>\n",  
 " <tr style=\"text-align: right;\">\n",  
 " <th></th>\n",  
 " <th>Request\_Closing\_Time</th>\n",  
 " </tr>\n",  
 " <tr>\n",  
 " <th>Location Type</th>\n",  
 " <th></th>\n",  
 " </tr>\n",  
 " </thead>\n",  
 " <tbody>\n",  
 " <tr>\n",  
 " <th>Subway Station</th>\n",  
 " <td>142.250980</td>\n",  
 " </tr>\n",  
 " <tr>\n",  
 " <th>Club/Bar/Restaurant</th>\n",  
 " <td>186.074330</td>\n",  
 " </tr>\n",  
 " <tr>\n",  
 " <th>House of Worship</th>\n",  
 " <td>191.833279</td>\n",  
 " </tr>\n",  
 " <tr>\n",  
 " <th>Store/Commercial</th>\n",  
 " <td>198.089073</td>\n",  
 " </tr>\n",  
 " <tr>\n",  
 " <th>Park/Playground</th>\n",  
 " <td>207.137129</td>\n",  
 " </tr>\n",  
 " <tr>\n",  
 " <th>Highway</th>\n",  
 " <td>223.424221</td>\n",  
 " </tr>\n",  
 " <tr>\n",  
 " <th>Bridge</th>\n",  
 " <td>229.158333</td>\n",  
 " </tr>\n",  
 " <tr>\n",  
 " <th>Roadway Tunnel</th>\n",  
 " <td>266.525714</td>\n",  
 " </tr>\n",  
 " <tr>\n",  
 " <th>Street/Sidewalk</th>\n",  
 " <td>268.515306</td>\n",  
 " </tr>\n",  
 " <tr>\n",  
 " <th>Residential Building</th>\n",  
 " <td>289.089941</td>\n",  
 " </tr>\n",  
 " <tr>\n",  
 " <th>House and Store</th>\n",  
 " <td>300.795699</td>\n",  
 " </tr>\n",  
 " <tr>\n",  
 " <th>Residential Building/House</th>\n",  
 " <td>309.505679</td>\n",  
 " </tr>\n",  
 " <tr>\n",  
 " <th>Parking Lot</th>\n",  
 " <td>320.130342</td>\n",  
 " </tr>\n",  
 " <tr>\n",  
 " <th>Commercial</th>\n",  
 " <td>320.566129</td>\n",  
 " </tr>\n",  
 " <tr>\n",  
 " <th>Vacant Lot</th>\n",  
 " <td>448.435498</td>\n",  
 " </tr>\n",  
 " <tr>\n",  
 " <th>Park</th>\n",  
 " <td>20210.083333</td>\n",  
 " </tr>\n",  
 " <tr>\n",  
 " <th>Ferry</th>\n",  
 " <td>NaN</td>\n",  
 " </tr>\n",  
 " <tr>\n",  
 " <th>Terminal</th>\n",  
 " <td>NaN</td>\n",  
 " </tr>\n",  
 " </tbody>\n",  
 "</table>\n",  
 "</div>"  
 ],  
 "text/plain": [  
 " Request\_Closing\_Time\n",  
 "Location Type \n",  
 "Subway Station 142.250980\n",  
 "Club/Bar/Restaurant 186.074330\n",  
 "House of Worship 191.833279\n",  
 "Store/Commercial 198.089073\n",  
 "Park/Playground 207.137129\n",  
 "Highway 223.424221\n",  
 "Bridge 229.158333\n",  
 "Roadway Tunnel 266.525714\n",  
 "Street/Sidewalk 268.515306\n",  
 "Residential Building 289.089941\n",  
 "House and Store 300.795699\n",  
 "Residential Building/House 309.505679\n",  
 "Parking Lot 320.130342\n",  
 "Commercial 320.566129\n",  
 "Vacant Lot 448.435498\n",  
 "Park 20210.083333\n",  
 "Ferry NaN\n",  
 "Terminal NaN"  
 ]  
 },  
 "execution\_count": 31,  
 "metadata": {},  
 "output\_type": "execute\_result"  
 }  
 ],  
 "source": [  
 "pd.DataFrame(df.groupby(\"Location Type\")[\"Request\_Closing\_Time\"].mean()).sort\_values(\"Request\_Closing\_Time\")"  
 ]  
 },  
 {  
 "cell\_type": "markdown",  
 "id": "d7fe1eb0",  
 "metadata": {},  
 "source": [  
 "Conclusion : Maximum time taken to resolved a complaint is in park and vacant lots whereas complaints from subway or club/bar/restaurant take the lowest."  
 ]  
 },  
 {  
 "cell\_type": "code",  
 "execution\_count": 32,  
 "id": "7754f982",  
 "metadata": {},  
 "outputs": [  
 {  
 "data": {  
 "text/html": [  
 "<div>\n",  
 "<style scoped>\n",  
 " .dataframe tbody tr th:only-of-type {\n",  
 " vertical-align: middle;\n",  
 " }\n",  
 "\n",  
 " .dataframe tbody tr th {\n",  
 " vertical-align: top;\n",  
 " }\n",  
 "\n",  
 " .dataframe thead th {\n",  
 " text-align: right;\n",  
 " }\n",  
 "</style>\n",  
 "<table border=\"1\" class=\"dataframe\">\n",  
 " <thead>\n",  
 " <tr style=\"text-align: right;\">\n",  
 " <th></th>\n",  
 " <th>Request\_Closing\_Time</th>\n",  
 " </tr>\n",  
 " <tr>\n",  
 " <th>City</th>\n",  
 " <th></th>\n",  
 " </tr>\n",  
 " </thead>\n",  
 " <tbody>\n",  
 " <tr>\n",  
 " <th>ARVERNE</th>\n",  
 " <td>135.895606</td>\n",  
 " </tr>\n",  
 " <tr>\n",  
 " <th>ROCKAWAY PARK</th>\n",  
 " <td>139.133736</td>\n",  
 " </tr>\n",  
 " <tr>\n",  
 " <th>LITTLE NECK</th>\n",  
 " <td>154.660316</td>\n",  
 " </tr>\n",  
 " <tr>\n",  
 " <th>OAKLAND GARDENS</th>\n",  
 " <td>157.853146</td>\n",  
 " </tr>\n",  
 " <tr>\n",  
 " <th>BAYSIDE</th>\n",  
 " <td>160.759992</td>\n",  
 " </tr>\n",  
 " <tr>\n",  
 " <th>FAR ROCKAWAY</th>\n",  
 " <td>167.399774</td>\n",  
 " </tr>\n",  
 " <tr>\n",  
 " <th>NEW YORK</th>\n",  
 " <td>178.357371</td>\n",  
 " </tr>\n",  
 " <tr>\n",  
 " <th>FLUSHING</th>\n",  
 " <td>181.081826</td>\n",  
 " </tr>\n",  
 " <tr>\n",  
 " <th>FOREST HILLS</th>\n",  
 " <td>193.449032</td>\n",  
 " </tr>\n",  
 " <tr>\n",  
 " <th>CORONA</th>\n",  
 " <td>193.670512</td>\n",  
 " </tr>\n",  
 " <tr>\n",  
 " <th>WHITESTONE</th>\n",  
 " <td>194.688843</td>\n",  
 " </tr>\n",  
 " <tr>\n",  
 " <th>FRESH MEADOWS</th>\n",  
 " <td>195.843207</td>\n",  
 " </tr>\n",  
 " <tr>\n",  
 " <th>COLLEGE POINT</th>\n",  
 " <td>196.417842</td>\n",  
 " </tr>\n",  
 " <tr>\n",  
 " <th>JACKSON HEIGHTS</th>\n",  
 " <td>196.419964</td>\n",  
 " </tr>\n",  
 " <tr>\n",  
 " <th>CENTRAL PARK</th>\n",  
 " <td>197.658591</td>\n",  
 " </tr>\n",  
 " <tr>\n",  
 " <th>ELMHURST</th>\n",  
 " <td>198.631095</td>\n",  
 " </tr>\n",  
 " <tr>\n",  
 " <th>REGO PARK</th>\n",  
 " <td>207.665668</td>\n",  
 " </tr>\n",  
 " <tr>\n",  
 " <th>BREEZY POINT</th>\n",  
 " <td>209.789444</td>\n",  
 " </tr>\n",  
 " <tr>\n",  
 " <th>EAST ELMHURST</th>\n",  
 " <td>214.659709</td>\n",  
 " </tr>\n",  
 " <tr>\n",  
 " <th>STATEN ISLAND</th>\n",  
 " <td>232.796699</td>\n",  
 " </tr>\n",  
 " <tr>\n",  
 " <th>Howard Beach</th>\n",  
 " <td>241.750000</td>\n",  
 " </tr>\n",  
 " <tr>\n",  
 " <th>BROOKLYN</th>\n",  
 " <td>242.878848</td>\n",  
 " </tr>\n",  
 " <tr>\n",  
 " <th>Long Island City</th>\n",  
 " <td>246.045522</td>\n",  
 " </tr>\n",  
 " <tr>\n",  
 " <th>Astoria</th>\n",  
 " <td>251.076304</td>\n",

" </tr>\n",  
 " <tr>\n",  
 " <th>RIDGEWOOD</th>\n",  
 " <td>266.507613</td>\n",  
 " </tr>\n",  
 " <tr>\n",  
 " <th>ASTORIA</th>\n",  
 " <td>275.934779</td>\n",  
 " </tr>\n",  
 " <tr>\n",  
 " <th>SAINT ALBANS</th>\n",  
 " <td>283.252098</td>\n",  
 " </tr>\n",  
 " <tr>\n",  
 " <th>KEW GARDENS</th>\n",  
 " <td>302.578556</td>\n",  
 " </tr>\n",  
 " <tr>\n",  
 " <th>Woodside</th>\n",  
 " <td>312.083333</td>\n",  
 " </tr>\n",  
 " <tr>\n",  
 " <th>JAMAICA</th>\n",  
 " <td>312.606051</td>\n",  
 " </tr>\n",  
 " <tr>\n",  
 " <th>SOUTH OZONE PARK</th>\n",  
 " <td>319.678662</td>\n",  
 " </tr>\n",  
 " <tr>\n",  
 " <th>MIDDLE VILLAGE</th>\n",  
 " <td>323.097583</td>\n",  
 " </tr>\n",  
 " <tr>\n",  
 " <th>RICHMOND HILL</th>\n",  
 " <td>329.658614</td>\n",  
 " </tr>\n",  
 " <tr>\n",  
 " <th>WOODHAVEN</th>\n",  
 " <td>335.728705</td>\n",  
 " </tr>\n",  
 " <tr>\n",  
 " <th>MASPETH</th>\n",  
 " <td>335.985805</td>\n",  
 " </tr>\n",  
 " <tr>\n",  
 " <th>SOUTH RICHMOND HILL</th>\n",  
 " <td>337.049201</td>\n",  
 " </tr>\n",  
 " <tr>\n",  
 " <th>OZONE PARK</th>\n",  
 " <td>340.863702</td>\n",  
 " </tr>\n",  
 " <tr>\n",  
 " <th>HOLLIS</th>\n",  
 " <td>345.610161</td>\n",  
 " </tr>\n",  
 " <tr>\n",  
 " <th>East Elmhurst</th>\n",  
 " <td>362.867857</td>\n",  
 " </tr>\n",  
 " <tr>\n",  
 " <th>BRONX</th>\n",  
 " <td>365.769723</td>\n",  
 " </tr>\n",  
 " <tr>\n",  
 " <th>HOWARD BEACH</th>\n",  
 " <td>369.652291</td>\n",  
 " </tr>\n",  
 " <tr>\n",  
 " <th>LONG ISLAND CITY</th>\n",  
 " <td>392.351457</td>\n",  
 " </tr>\n",  
 " <tr>\n",  
 " <th>SUNNYSIDE</th>\n",  
 " <td>411.120332</td>\n",  
 " </tr>\n",  
 " <tr>\n",  
 " <th>WOODSIDE</th>\n",  
 " <td>413.606029</td>\n",  
 " </tr>\n",  
 " <tr>\n",  
 " <th>NEW HYDE PARK</th>\n",  
 " <td>453.365646</td>\n",  
 " </tr>\n",  
 " <tr>\n",  
 " <th>GLEN OAKS</th>\n",  
 " <td>528.943900</td>\n",  
 " </tr>\n",  
 " <tr>\n",  
 " <th>SPRINGFIELD GARDENS</th>\n",  
 " <td>551.145130</td>\n",  
 " </tr>\n",  
 " <tr>\n",  
 " <th>ROSEDALE</th>\n",  
 " <td>601.867552</td>\n",  
 " </tr>\n",  
 " <tr>\n",  
 " <th>CAMBRIA HEIGHTS</th>\n",  
 " <td>607.426555</td>\n",  
 " </tr>\n",  
 " <tr>\n",  
 " <th>BELLEROSE</th>\n",  
 " <td>633.386578</td>\n",  
 " </tr>\n",  
 " <tr>\n",  
 " <th>QUEENS VILLAGE</th>\n",  
 " <td>654.411273</td>\n",  
 " </tr>\n",  
 " <tr>\n",  
 " <th>FLORAL PARK</th>\n",  
 " <td>703.171272</td>\n",  
 " </tr>\n",  
 " <tr>\n",  
 " <th>QUEENS</th>\n",  
 " <td>815.586458</td>\n",  
 " </tr>\n",  
 " </tbody>\n",  
 "</table>\n",  
 "</div>"  
 ],  
 "text/plain": [  
 " Request\_Closing\_Time\n",  
 "City \n",  
 "ARVERNE 135.895606\n",  
 "ROCKAWAY PARK 139.133736\n",  
 "LITTLE NECK 154.660316\n",  
 "OAKLAND GARDENS 157.853146\n",  
 "BAYSIDE 160.759992\n",  
 "FAR ROCKAWAY 167.399774\n",  
 "NEW YORK 178.357371\n",  
 "FLUSHING 181.081826\n",  
 "FOREST HILLS 193.449032\n",  
 "CORONA 193.670512\n",  
 "WHITESTONE 194.688843\n",  
 "FRESH MEADOWS 195.843207\n",  
 "COLLEGE POINT 196.417842\n",  
 "JACKSON HEIGHTS 196.419964\n",  
 "CENTRAL PARK 197.658591\n",  
 "ELMHURST 198.631095\n",  
 "REGO PARK 207.665668\n",  
 "BREEZY POINT 209.789444\n",  
 "EAST ELMHURST 214.659709\n",  
 "STATEN ISLAND 232.796699\n",  
 "Howard Beach 241.750000\n",  
 "BROOKLYN 242.878848\n",  
 "Long Island City 246.045522\n",  
 "Astoria 251.076304\n",  
 "RIDGEWOOD 266.507613\n",  
 "ASTORIA 275.934779\n",  
 "SAINT ALBANS 283.252098\n",  
 "KEW GARDENS 302.578556\n",  
 "Woodside 312.083333\n",  
 "JAMAICA 312.606051\n",  
 "SOUTH OZONE PARK 319.678662\n",  
 "MIDDLE VILLAGE 323.097583\n",  
 "RICHMOND HILL 329.658614\n",  
 "WOODHAVEN 335.728705\n",  
 "MASPETH 335.985805\n",  
 "SOUTH RICHMOND HILL 337.049201\n",  
 "OZONE PARK 340.863702\n",  
 "HOLLIS 345.610161\n",  
 "East Elmhurst 362.867857\n",  
 "BRONX 365.769723\n",  
 "HOWARD BEACH 369.652291\n",  
 "LONG ISLAND CITY 392.351457\n",  
 "SUNNYSIDE 411.120332\n",  
 "WOODSIDE 413.606029\n",  
 "NEW HYDE PARK 453.365646\n",  
 "GLEN OAKS 528.943900\n",  
 "SPRINGFIELD GARDENS 551.145130\n",  
 "ROSEDALE 601.867552\n",  
 "CAMBRIA HEIGHTS 607.426555\n",  
 "BELLEROSE 633.386578\n",  
 "QUEENS VILLAGE 654.411273\n",  
 "FLORAL PARK 703.171272\n",  
 "QUEENS 815.586458"  
 ]  
 },  
 "execution\_count": 32,  
 "metadata": {},  
 "output\_type": "execute\_result"  
 }  
 ],  
 "source": [  
 "#losing time of complaints with respect to city\n",  
 "pd.DataFrame(df.groupby(\"City\")[\"Request\_Closing\_Time\"].mean()).sort\_values(\"Request\_Closing\_Time\")"  
 ]  
 },  
 {  
 "cell\_type": "code",  
 "execution\_count": 33,  
 "id": "5dc8b1b2",  
 "metadata": {},  
 "outputs": [  
 {  
 "data": {  
 "text/html": [  
 "<div>\n",  
 "<style scoped>\n",  
 " .dataframe tbody tr th:only-of-type {\n",  
 " vertical-align: middle;\n",  
 " }\n",  
 "\n",  
 " .dataframe tbody tr th {\n",  
 " vertical-align: top;\n",  
 " }\n",  
 "\n",  
 " .dataframe thead th {\n",  
 " text-align: right;\n",  
 " }\n",  
 "</style>\n",  
 "<table border=\"1\" class=\"dataframe\">\n",  
 " <thead>\n",  
 " <tr style=\"text-align: right;\">\n",  
 " <th></th>\n",  
 " <th>0</th>\n",  
 " </tr>\n",  
 " </thead>\n",  
 " <tbody>\n",  
 " <tr>\n",  
 " <th>School or Citywide Complaint</th>\n",  
 " <td>100.000000</td>\n",  
 " </tr>\n",  
 " <tr>\n",  
 " <th>Garage Lot Name</th>\n",  
 " <td>100.000000</td>\n",  
 " </tr>\n",  
 " <tr>\n",  
 " <th>Vehicle Type</th>\n",  
 " <td>100.000000</td>\n",  
 " </tr>\n",  
 " <tr>\n",  
 " <th>Taxi Pick Up Location</th>\n",  
 " <td>100.000000</td>\n",  
 " </tr>\n",  
 " <tr>\n",  
 " <th>Taxi Company Borough</th>\n",  
 " <td>100.000000</td>\n",  
 " </tr>\n",  
 " <tr>\n",  
 " <th>Ferry Direction</th>\n",  
 " <td>99.999667</td>\n",  
 " </tr>\n",  
 " <tr>\n",  
 " <th>Ferry Terminal Name</th>\n",  
 " <td>99.999335</td>\n",  
 " </tr>\n",  
 " <tr>\n",  
 " <th>Road Ramp</th>\n",  
 " <td>99.929165</td>\n",  
 " </tr>\n",  
 " <tr>\n",  
 " <th>Bridge Highway Segment</th>\n",  
 " <td>99.929165</td>\n",  
 " </tr>\n",  
 " <tr>\n",  
 " <th>Bridge Highway Direction</th>\n",  
 " <td>99.919188</td>\n",  
 " </tr>\n",  
 " <tr>\n",  
 " <th>Bridge Highway Name</th>\n",  
 " <td>99.919188</td>\n",  
 " </tr>\n",  
 " <tr>\n",  
 " <th>Landmark</th>\n",  
 " <td>99.883937</td>\n",  
 " </tr>\n",  
 " <tr>\n",  
 " <th>Intersection Street 2</th>\n",  
 " <td>85.579552</td>\n",  
 " </tr>\n",  
 " <tr>\n",  
 " <th>Intersection Street 1</th>\n",  
 " <td>85.414602</td>\n",  
 " </tr>\n",  
 " <tr>\n",  
 " <th>Cross Street 2</th>\n",  
 " <td>16.554483</td>\n",  
 " </tr>\n",  
 " <tr>\n",  
 " <th>Cross Street 1</th>\n",  
 " <td>16.388203</td>\n",  
 " </tr>\n",  
 " <tr>\n",  
 " <th>Street Name</th>\n",  
 " <td>14.768971</td>\n",  
 " </tr>\n",  
 " <tr>\n",  
 " <th>Incident Address</th>\n",  
 " <td>14.768971</td>\n",

" </tr>\n",  
 " <tr>\n",  
 " <th>Descriptor</th>\n",  
 " <td>1.966757</td>\n",  
 " </tr>\n",  
 " <tr>\n",  
 " <th>Latitude</th>\n",  
 " <td>1.177261</td>\n",  
 " </tr>\n",  
 " </tbody>\n",  
 "</table>\n",  
 "</div>"  
 ],  
 "text/plain": [  
 " 0\n",  
 "School or Citywide Complaint 100.000000\n",  
 "Garage Lot Name 100.000000\n",  
 "Vehicle Type 100.000000\n",  
 "Taxi Pick Up Location 100.000000\n",  
 "Taxi Company Borough 100.000000\n",  
 "Ferry Direction 99.999667\n",  
 "Ferry Terminal Name 99.999335\n",  
 "Road Ramp 99.929165\n",  
 "Bridge Highway Segment 99.929165\n",  
 "Bridge Highway Direction 99.919188\n",  
 "Bridge Highway Name 99.919188\n",  
 "Landmark 99.883937\n",  
 "Intersection Street 2 85.579552\n",  
 "Intersection Street 1 85.414602\n",  
 "Cross Street 2 16.554483\n",  
 "Cross Street 1 16.388203\n",  
 "Street Name 14.768971\n",  
 "Incident Address 14.768971\n",  
 "Descriptor 1.966757\n",  
 "Latitude 1.177261"  
 ]  
 },  
 "execution\_count": 33,  
 "metadata": {},  
 "output\_type": "execute\_result"  
 }  
 ],  
 "source": [  
 "#Percentage of missing values\n",  
 "pd.DataFrame((df.isnull().sum()/df.shape[0]\*100)).sort\_values(0,ascending=False)[:20]"  
 ]  
 },  
 {  
 "cell\_type": "markdown",  
 "id": "96ac85c6",  
 "metadata": {},  
 "source": [  
 "We can see that school or city wide complaint, garage lot name, vehicle type, taxi pickup location, taxi company borough have 100% missing values, that could mean there are no complaints from those sectors."  
 ]  
 },  
 {  
 "cell\_type": "code",  
 "execution\_count": 34,  
 "id": "0a166e1b",  
 "metadata": {},  
 "outputs": [],  
 "source": [  
 "#We will now drop the missing values\n",  
 "dfn=df.loc[:,(df.isnull().sum()/df.shape[0]\*100)<=50]"  
 ]  
 },  
 {  
 "cell\_type": "code",  
 "execution\_count": 35,  
 "id": "98854be4",  
 "metadata": {},  
 "outputs": [  
 {  
 "name": "stdout",  
 "output\_type": "stream",  
 "text": [  
 "Old df shape : (300698, 54)\n",  
 "New df shape: (300698, 40)\n"  
 ]  
 }  
 ],  
 "source": [  
 "print(\"Old df shape :\",df.shape)\n",  
 "print(\"New df shape: \",dfn.shape)"  
 ]  
 },  
 {  
 "cell\_type": "code",  
 "execution\_count": 38,  
 "id": "950c98af",  
 "metadata": {},  
 "outputs": [  
 {  
 "name": "stdout",  
 "output\_type": "stream",  
 "text": [  
 "Agency : ['NYPD']\n",  
 "Agency Name : ['New York City Police Department' 'NYPD' 'Internal Affairs Bureau']\n",  
 "Facility Type : ['Precinct' nan]\n",  
 "Park Facility Name : ['Unspecified' 'Alley Pond Park - Nature Center']\n",  
 "School Name : ['Unspecified' 'Alley Pond Park - Nature Center']\n",  
 "School Number : ['Unspecified' 'Q001']\n",  
 "School Region : ['Unspecified' nan]\n",  
 "School Code : ['Unspecified' nan]\n",  
 "School Phone Number : ['Unspecified' '7182176034']\n",  
 "School Address : ['Unspecified' 'Grand Central Parkway, near the soccer field']\n",  
 "School City : ['Unspecified' 'QUEENS']\n",  
 "School State : ['Unspecified' 'NY']\n",  
 "School Zip : ['Unspecified' nan]\n",  
 "School Not Found : ['N']\n"  
 ]  
 }  
 ],  
 "source": [  
 "rem=[]\n",  
 "for x in dfn.columns.tolist():\n",  
 " if dfn[x].nunique()<=3:\n",  
 " print(x+ \" \"\*10+\" : \",dfn[x].unique())\n",  
 " rem.append(x)"  
 ]  
 },  
 {  
 "cell\_type": "markdown",  
 "id": "8a809023",  
 "metadata": {},  
 "source": [  
 "We can remove the unspecified data"  
 ]  
 },  
 {  
 "cell\_type": "code",  
 "execution\_count": 39,  
 "id": "a9618902",  
 "metadata": {},  
 "outputs": [  
 {  
 "name": "stderr",  
 "output\_type": "stream",  
 "text": [  
 "C:\\Users\\lrnem\\AppData\\Local\\Temp\\ipykernel\_8728\\3503437274.py:1: SettingWithCopyWarning: \n",  
 "A value is trying to be set on a copy of a slice from a DataFrame\n",  
 "\n",  
 "See the caveats in the documentation: https://pandas.pydata.org/pandas-docs/stable/user\_guide/indexing.html#returning-a-view-versus-a-copy\n",  
 " dfn.drop(rem,axis=1,inplace=True)\n"  
 ]  
 }  
 ],  
 "source": [  
 "dfn.drop(rem,axis=1,inplace=True)"  
 ]  
 },  
 {  
 "cell\_type": "code",  
 "execution\_count": 40,  
 "id": "9a1e6bac",  
 "metadata": {},  
 "outputs": [  
 {  
 "data": {  
 "text/plain": [  
 "(300698, 26)"  
 ]  
 },  
 "execution\_count": 40,  
 "metadata": {},  
 "output\_type": "execute\_result"  
 }  
 ],  
 "source": [  
 "dfn.shape"  
 ]  
 },  
 {  
 "cell\_type": "code",  
 "execution\_count": 41,  
 "id": "5d1df891",  
 "metadata": {},  
 "outputs": [  
 {  
 "data": {  
 "text/html": [  
 "<div>\n",  
 "<style scoped>\n",  
 " .dataframe tbody tr th:only-of-type {\n",  
 " vertical-align: middle;\n",  
 " }\n",  
 "\n",  
 " .dataframe tbody tr th {\n",  
 " vertical-align: top;\n",  
 " }\n",  
 "\n",  
 " .dataframe thead th {\n",  
 " text-align: right;\n",  
 " }\n",  
 "</style>\n",  
 "<table border=\"1\" class=\"dataframe\">\n",  
 " <thead>\n",  
 " <tr style=\"text-align: right;\">\n",  
 " <th></th>\n",  
 " <th>Unique Key</th>\n",  
 " <th>Created Date</th>\n",  
 " <th>Closed Date</th>\n",  
 " <th>Complaint Type</th>\n",  
 " <th>Descriptor</th>\n",  
 " <th>Location Type</th>\n",  
 " <th>Incident Zip</th>\n",  
 " <th>Incident Address</th>\n",  
 " <th>Street Name</th>\n",  
 " <th>Cross Street 1</th>\n",  
 " <th>...</th>\n",  
 " <th>Resolution Action Updated Date</th>\n",  
 " <th>Community Board</th>\n",  
 " <th>Borough</th>\n",  
 " <th>X Coordinate (State Plane)</th>\n",  
 " <th>Y Coordinate (State Plane)</th>\n",  
 " <th>Park Borough</th>\n",  
 " <th>Latitude</th>\n",  
 " <th>Longitude</th>\n",  
 " <th>Location</th>\n",  
 " <th>Request\_Closing\_Time</th>\n",  
 " </tr>\n",  
 " </thead>\n",  
 " <tbody>\n",  
 " <tr>\n",  
 " <th>0</th>\n",  
 " <td>32310363</td>\n",  
 " <td>2015-12-31 23:59:45</td>\n",  
 " <td>2016-01-01 00:55:00</td>\n",  
 " <td>Noise - Street/Sidewalk</td>\n",  
 " <td>Loud Music/Party</td>\n",  
 " <td>Street/Sidewalk</td>\n",  
 " <td>10034.0</td>\n",  
 " <td>71 VERMILYEA AVENUE</td>\n",  
 " <td>VERMILYEA AVENUE</td>\n",  
 " <td>ACADEMY STREET</td>\n",  
 " <td>...</td>\n",  
 " <td>01-01-16 0:55</td>\n",  
 " <td>12 MANHATTAN</td>\n",  
 " <td>MANHATTAN</td>\n",  
 " <td>1005409.0</td>\n",  
 " <td>254678.0</td>\n",  
 " <td>MANHATTAN</td>\n",  
 " <td>40.865682</td>\n",  
 " <td>-73.923501</td>\n",  
 " <td>(40.86568153633767, -73.92350095571744)</td>\n",  
 " <td>55.250000</td>\n",  
 " </tr>\n",  
 " <tr>\n",  
 " <th>1</th>\n",  
 " <td>32309934</td>\n",  
 " <td>2015-12-31 23:59:44</td>\n",  
 " <td>2016-01-01 01:26:00</td>\n",  
 " <td>Blocked Driveway</td>\n",  
 " <td>No Access</td>\n",  
 " <td>Street/Sidewalk</td>\n",  
 " <td>11105.0</td>\n",  
 " <td>27-07 23 AVENUE</td>\n",  
 " <td>23 AVENUE</td>\n",  
 " <td>27 STREET</td>\n",  
 " <td>...</td>\n",  
 " <td>01-01-16 1:26</td>\n",  
 " <td>01 QUEENS</td>\n",  
 " <td>QUEENS</td>\n",  
 " <td>1007766.0</td>\n",  
 " <td>221986.0</td>\n",  
 " <td>QUEENS</td>\n",  
 " <td>40.775945</td>\n",  
 " <td>-73.915094</td>\n",  
 " <td>(40.775945312321085, -73.91509393898605)</td>\n",  
 " <td>86.266667</td>\n",  
 " </tr>\n",  
 " <tr>\n",  
 " <th>2</th>\n",  
 " <td>32309159</td>\n",  
 " <td>2015-12-31 23:59:29</td>\n",  
 " <td>2016-01-01 04:51:00</td>\n",  
 " <td>Blocked Driveway</td>\n",  
 " <td>No Access</td>\n",  
 " <td>Street/Sidewalk</td>\n",  
 " <td>10458.0</td>\n",  
 " <td>2897 VALENTINE AVENUE</td>\n",  
 " <td>VALENTINE AVENUE</td>\n",  
 " <td>EAST 198 STREET</td>\n",  
 " <td>...</td>\n",  
 " <td>01-01-16 4:51</td>\n",  
 " <td>07 BRONX</td>\n",  
 " <td>BRONX</td>\n",  
 " <td>1015081.0</td>\n",  
 " <td>256380.0</td>\n",  
 " <td>BRONX</td>\n",  
 " <td>40.870325</td>\n",  
 " <td>-73.888525</td>\n",  
 " <td>(40.870324522111424, -73.88852464418646)</td>\n",  
 " <td>291.516667</td>\n",  
 " </tr>\n",  
 " <tr>\n",  
 " <th>3</th>\n",  
 " <td>32305098</td>\n",  
 " <td>2015-12-31 23:57:46</td>\n",  
 " <td>2016-01-01 07:43:00</td>\n",  
 " <td>Illegal Parking</td>\n",  
 " <td>Commercial Overnight Parking</td>\n",  
 " <td>Street/Sidewalk</td>\n",

" <td>10461.0</td>\n",  
 " <td>2940 BAISLEY AVENUE</td>\n",  
 " <td>BAISLEY AVENUE</td>\n",  
 " <td>EDISON AVENUE</td>\n",  
 " <td>...</td>\n",  
 " <td>01-01-16 7:43</td>\n",  
 " <td>10 BRONX</td>\n",  
 " <td>BRONX</td>\n",  
 " <td>1031740.0</td>\n",  
 " <td>243899.0</td>\n",  
 " <td>BRONX</td>\n",  
 " <td>40.835994</td>\n",  
 " <td>-73.828379</td>\n",  
 " <td>(40.83599404683083, -73.82837939584206)</td>\n",  
 " <td>465.233333</td>\n",  
 " </tr>\n",  
 " <tr>\n",  
 " <th>4</th>\n",  
 " <td>32306529</td>\n",  
 " <td>2015-12-31 23:56:58</td>\n",  
 " <td>2016-01-01 03:24:00</td>\n",  
 " <td>Illegal Parking</td>\n",  
 " <td>Blocked Sidewalk</td>\n",  
 " <td>Street/Sidewalk</td>\n",  
 " <td>11373.0</td>\n",  
 " <td>87-14 57 ROAD</td>\n",  
 " <td>57 ROAD</td>\n",  
 " <td>SEABURY STREET</td>\n",  
 " <td>...</td>\n",  
 " <td>01-01-16 3:24</td>\n",  
 " <td>04 QUEENS</td>\n",  
 " <td>QUEENS</td>\n",  
 " <td>1019123.0</td>\n",  
 " <td>206375.0</td>\n",  
 " <td>QUEENS</td>\n",  
 " <td>40.733060</td>\n",  
 " <td>-73.874170</td>\n",  
 " <td>(40.733059618956815, -73.87416975810375)</td>\n",  
 " <td>207.033333</td>\n",  
 " </tr>\n",  
 " </tbody>\n",  
 "</table>\n",  
 "<p>5 rows × 26 columns</p>\n",  
 "</div>"  
 ],  
 "text/plain": [  
 " Unique Key Created Date Closed Date \\\n",  
 "0 32310363 2015-12-31 23:59:45 2016-01-01 00:55:00 \n",  
 "1 32309934 2015-12-31 23:59:44 2016-01-01 01:26:00 \n",  
 "2 32309159 2015-12-31 23:59:29 2016-01-01 04:51:00 \n",  
 "3 32305098 2015-12-31 23:57:46 2016-01-01 07:43:00 \n",  
 "4 32306529 2015-12-31 23:56:58 2016-01-01 03:24:00 \n",  
 "\n",  
 " Complaint Type Descriptor Location Type \\\n",  
 "0 Noise - Street/Sidewalk Loud Music/Party Street/Sidewalk \n",  
 "1 Blocked Driveway No Access Street/Sidewalk \n",  
 "2 Blocked Driveway No Access Street/Sidewalk \n",  
 "3 Illegal Parking Commercial Overnight Parking Street/Sidewalk \n",  
 "4 Illegal Parking Blocked Sidewalk Street/Sidewalk \n",  
 "\n",  
 " Incident Zip Incident Address Street Name Cross Street 1 \\\n",  
 "0 10034.0 71 VERMILYEA AVENUE VERMILYEA AVENUE ACADEMY STREET \n",  
 "1 11105.0 27-07 23 AVENUE 23 AVENUE 27 STREET \n",  
 "2 10458.0 2897 VALENTINE AVENUE VALENTINE AVENUE EAST 198 STREET \n",  
 "3 10461.0 2940 BAISLEY AVENUE BAISLEY AVENUE EDISON AVENUE \n",  
 "4 11373.0 87-14 57 ROAD 57 ROAD SEABURY STREET \n",  
 "\n",  
 " ... Resolution Action Updated Date Community Board Borough \\\n",  
 "0 ... 01-01-16 0:55 12 MANHATTAN MANHATTAN \n",  
 "1 ... 01-01-16 1:26 01 QUEENS QUEENS \n",  
 "2 ... 01-01-16 4:51 07 BRONX BRONX \n",  
 "3 ... 01-01-16 7:43 10 BRONX BRONX \n",  
 "4 ... 01-01-16 3:24 04 QUEENS QUEENS \n",  
 "\n",  
 " X Coordinate (State Plane) Y Coordinate (State Plane) Park Borough \\\n",  
 "0 1005409.0 254678.0 MANHATTAN \n",  
 "1 1007766.0 221986.0 QUEENS \n",  
 "2 1015081.0 256380.0 BRONX \n",  
 "3 1031740.0 243899.0 BRONX \n",  
 "4 1019123.0 206375.0 QUEENS \n",  
 "\n",  
 " Latitude Longitude Location \\\n",  
 "0 40.865682 -73.923501 (40.86568153633767, -73.92350095571744) \n",  
 "1 40.775945 -73.915094 (40.775945312321085, -73.91509393898605) \n",  
 "2 40.870325 -73.888525 (40.870324522111424, -73.88852464418646) \n",  
 "3 40.835994 -73.828379 (40.83599404683083, -73.82837939584206) \n",  
 "4 40.733060 -73.874170 (40.733059618956815, -73.87416975810375) \n",  
 "\n",  
 " Request\_Closing\_Time \n",  
 "0 55.250000 \n",  
 "1 86.266667 \n",  
 "2 291.516667 \n",  
 "3 465.233333 \n",  
 "4 207.033333 \n",  
 "\n",  
 "[5 rows x 26 columns]"  
 ]  
 },  
 "execution\_count": 41,  
 "metadata": {},  
 "output\_type": "execute\_result"  
 }  
 ],  
 "source": [  
 "dfn.head()"  
 ]  
 },  
 {  
 "cell\_type": "code",  
 "execution\_count": 43,  
 "id": "150705bf",  
 "metadata": {},  
 "outputs": [  
 {  
 "data": {  
 "text/plain": [  
 "(0.0, 2000.0)"  
 ]  
 },  
 "execution\_count": 43,  
 "metadata": {},  
 "output\_type": "execute\_result"  
 },  
 {  
 "data": {  
 "image/png": "

































",  
 "text/plain": [  
 "<Figure size 1500x800 with 1 Axes>"  
 ]  
 },  
 "metadata": {},  
 "output\_type": "display\_data"  
 }  
 ],  
 "source": [  
 "#Hypothesis testing\n",  
 "crt=sns.catplot(x=\"Complaint Type\",y=\"Request\_Closing\_Time\",kind=\"box\",data=dfn)\n",  
 "crt.fig.set\_figheight(8)\n",  
 "crt.fig.set\_figwidth(15)\n",  
 "plt.xticks(rotation=90)\n",  
 "plt.ylim((0,2000))"  
 ]  
 },  
 {  
 "cell\_type": "markdown",  
 "id": "09c8a74b",  
 "metadata": {},  
 "source": [  
 "H0: There is no significant different in mean of Request\_Closing\_Time for different Complaint\n",  
 "\n",  
 "\n",  
 "H1: There is signficant different in mean of Request\_Closing\_Time for different Complaint"  
 ]  
 },  
 {  
 "cell\_type": "code",  
 "execution\_count": 46,  
 "id": "e914deec",  
 "metadata": {},  
 "outputs": [  
 {  
 "data": {  
 "text/html": [  
 "<div>\n",  
 "<style scoped>\n",  
 " .dataframe tbody tr th:only-of-type {\n",  
 " vertical-align: middle;\n",  
 " }\n",  
 "\n",  
 " .dataframe tbody tr th {\n",  
 " vertical-align: top;\n",  
 " }\n",  
 "\n",  
 " .dataframe thead th {\n",  
 " text-align: right;\n",  
 " }\n",  
 "</style>\n",  
 "<table border=\"1\" class=\"dataframe\">\n",  
 " <thead>\n",  
 " <tr style=\"text-align: right;\">\n",  
 " <th></th>\n",  
 " <th>Request\_Closing\_Time</th>\n",  
 " <th>Complaint</th>\n",  
 " </tr>\n",  
 " </thead>\n",  
 " <tbody>\n",  
 " <tr>\n",  
 " <th>0</th>\n",  
 " <td>55.250000</td>\n",  
 " <td>Noise - Street/Sidewalk</td>\n",  
 " </tr>\n",  
 " <tr>\n",  
 " <th>1</th>\n",  
 " <td>86.266667</td>\n",  
 " <td>Blocked Driveway</td>\n",  
 " </tr>\n",  
 " <tr>\n",  
 " <th>2</th>\n",  
 " <td>291.516667</td>\n",  
 " <td>Blocked Driveway</td>\n",  
 " </tr>\n",  
 " <tr>\n",  
 " <th>3</th>\n",  
 " <td>465.233333</td>\n",  
 " <td>Illegal Parking</td>\n",  
 " </tr>\n",  
 " <tr>\n",  
 " <th>4</th>\n",  
 " <td>207.033333</td>\n",  
 " <td>Illegal Parking</td>\n",  
 " </tr>\n",  
 " </tbody>\n",  
 "</table>\n",  
 "</div>"  
 ],  
 "text/plain": [  
 " Request\_Closing\_Time Complaint\n",  
 "0 55.250000 Noise - Street/Sidewalk\n",  
 "1 86.266667 Blocked Driveway\n",  
 "2 291.516667 Blocked Driveway\n",  
 "3 465.233333 Illegal Parking\n",  
 "4 207.033333 Illegal Parking"  
 ]  
 },  
 "execution\_count": 46,  
 "metadata": {},  
 "output\_type": "execute\_result"  
 }  
 ],  
 "source": [  
 "anova\_df=pd.DataFrame()\n",  
 "anova\_df[\"Request\_Closing\_Time\"]=dfn[\"Request\_Closing\_Time\"]\n",  
 "anova\_df[\"Complaint\"]=dfn[\"Complaint Type\"]\n",  
 "\n",  
 "anova\_df.dropna(inplace=True)\n",  
 "anova\_df.head()"  
 ]  
 },  
 {  
 "cell\_type": "code",  
 "execution\_count": 47,  
 "id": "3672536b",  
 "metadata": {},  
 "outputs": [  
 {  
 "data": {  
 "text/html": [  
 "<div>\n",  
 "<style scoped>\n",  
 " .dataframe tbody tr th:only-of-type {\n",  
 " vertical-align: middle;\n",  
 " }\n",  
 "\n",  
 " .dataframe tbody tr th {\n",  
 " vertical-align: top;\n",  
 " }\n",  
 "\n",  
 " .dataframe thead th {\n",  
 " text-align: right;\n",  
 " }\n",  
 "</style>\n",  
 "<table border=\"1\" class=\"dataframe\">\n",  
 " <thead>\n",  
 " <tr style=\"text-align: right;\">\n",  
 " <th></th>\n",  
 " <th>df</th>\n",  
 " <th>sum\_sq</th>\n",  
 " <th>mean\_sq</th>\n",  
 " <th>F</th>\n",  
 " <th>PR(&gt;F)</th>\n",  
 " </tr>\n",  
 " </thead>\n",  
 " <tbody>\n",  
 " <tr>\n",  
 " <th>Complaint</th>\n",  
 " <td>22.0</td>\n",  
 " <td>1.455049e+09</td>\n",  
 " <td>6.613860e+07</td>\n",  
 " <td>514.177089</td>\n",  
 " <td>0.0</td>\n",  
 " </tr>\n",  
 " <tr>\n",  
 " <th>Residual</th>\n",  
 " <td>298511.0</td>\n",  
 " <td>3.839747e+10</td>\n",  
 " <td>1.286300e+05</td>\n",  
 " <td>NaN</td>\n",  
 " <td>NaN</td>\n",  
 " </tr>\n",  
 " </tbody>\n",  
 "</table>\n",  
 "</div>"  
 ],  
 "text/plain": [  
 " df sum\_sq mean\_sq F PR(>F)\n",  
 "Complaint 22.0 1.455049e+09 6.613860e+07 514.177089 0.0\n",  
 "Residual 298511.0 3.839747e+10 1.286300e+05 NaN NaN"  
 ]  
 },  
 "execution\_count": 47,  
 "metadata": {},  
 "output\_type": "execute\_result"  
 }  
 ],  
 "source": [  
 "lm=ols(\"Request\_Closing\_Time~Complaint\",data=anova\_df).fit()\n",  
 "table=sm.stats.anova\_lm(lm)\n",  
 "table"  
 ]  
 },  
 {  
 "cell\_type": "code",  
 "execution\_count": 48,  
 "id": "cf9f869b",  
 "metadata": {},  
 "outputs": [],  
 "source": [  
 "chi\_sq=pd.DataFrame()\n",  
 "chi\_sq[\"Location Type\"]=dfn[\"Location Type\"]\n",  
 "chi\_sq[\"Complaint Type\"]=dfn[\"Complaint Type\"]\n",  
 "chi\_sq.dropna(inplace=True)"  
 ]  
 },  
 {  
 "cell\_type": "code",  
 "execution\_count": 49,  
 "id": "e9f033de",  
 "metadata": {},  
 "outputs": [],  
 "source": [  
 "data\_crosstab = pd.crosstab( chi\_sq[\"Location Type\"],chi\_sq[\"Complaint Type\"])"  
 ]  
 },  
 {  
 "cell\_type": "code",  
 "execution\_count": 50,  
 "id": "d932ba9a",  
 "metadata": {},  
 "outputs": [  
 {  
 "name": "stdout",  
 "output\_type": "stream",  
 "text": [  
 "Dependent (reject H0)\n"  
 ]  
 }  
 ],  
 "source": [  
 "stat, p, dof, expected = chi2\_contingency(data\_crosstab) \n",  
 "\n",  
 "alpha = 0.05\n",  
 "if p <= alpha: \n",  
 " print('Dependent (reject H0)') \n",  
 "else: \n",  
 " print('Independent (H0 holds true)')"  
 ]  
 },  
 {  
 "cell\_type": "markdown",  
 "id": "6dbf6631",  
 "metadata": {},  
 "source": [  
 "Conclusions :\n",  
 " 1. Different complaints last for different duration.\n",  
 " 2. Complaints are different in different locations.\n",  
 " 3. Majority complaints are from transport sector.\n",  
 " 4. School sector has the lowest number of complaints(next to none)."  
 ]  
 },  
 {  
 "cell\_type": "code",  
 "execution\_count": null,  
 "id": "ca6b12c6",  
 "metadata": {},  
 "outputs": [],  
 "source": []  
 }  
 ],  
 "metadata": {  
 "kernelspec": {  
 "display\_name": "Python 3 (ipykernel)",  
 "language": "python",  
 "name": "python3"  
 },  
 "language\_info": {  
 "codemirror\_mode": {  
 "name": "ipython",  
 "version": 3  
 },  
 "file\_extension": ".py",  
 "mimetype": "text/x-python",  
 "name": "python",  
 "nbconvert\_exporter": "python",  
 "pygments\_lexer": "ipython3",  
 "version": "3.11.3"  
 }  
 },  
 "nbformat": 4,  
 "nbformat\_minor": 5  
}