In [66]:
```
#TASK 1
import pandas  as  pd
import numpy as np
import matplotlib.pyplot as plt
import seaborn as sns
%matplotlib inline
import scipy.stats as stats
import datetime
import statsmodels.api as sm
from statsmodels.formula.api import ols
```

In [2]:
```
df = pd.read_csv(r'C:\Users\Bhavuk\Desktop\311_Service_Requests_from_2010_to_Present.csv
```

In [3]:
```
pd.set_option('display.max_columns',30)
pd.set_option('display.max_rows',800)
```

In [4]:
```
df.head()
```

Out[4]:

	Unique Key	Created Date	Closed Date	Agency	Agency Name	Complaint Type	Descriptor	Location Type	Incident Zip	
0	32310363	12/31/2015 11:59:45 PM	01-01-16 0:55	NYPD	New York City Police Department	Noise - Street/Sidewalk	Loud Music/Party	Street/Sidewalk	10034.0	VEI
1	32309934	12/31/2015 11:59:44 PM	01-01-16 1:26	NYPD	New York City Police Department	Blocked Driveway	No Access	Street/Sidewalk	11105.0	
2	32309159	12/31/2015 11:59:29 PM	01-01-16 4:51	NYPD	New York City Police Department	Blocked Driveway	No Access	Street/Sidewalk	10458.0	VA
3	32305098	12/31/2015 11:57:46 PM	01-01-16 7:43	NYPD	New York City Police Department	Illegal Parking	Commercial Overnight Parking	Street/Sidewalk	10461.0	
4	32306529	12/31/2015 11:56:58 PM	01-01-16 3:24	NYPD	New York City Police Department	Illegal Parking	Blocked Sidewalk	Street/Sidewalk	11373.0	

5 rows × 53 columns

In [5]:
```
df.shape
```

Out[5]:
```
(300698, 53)
```

In [6]:
```
df.columns
```

Out[6]:
```
Index(['Unique Key', 'Created Date', 'Closed Date', 'Agency', 'Agency Name',
       'Complaint Type', 'Descriptor', 'Location Type', 'Incident Zip',
       'Incident Address', 'Street Name', 'Cross Street 1', 'Cross Street 2',
       'Intersection Street 1', 'Intersection Street 2', 'Address Type',
       'City', 'Landmark', 'Facility Type', 'Status', 'Due Date',
       'Resolution Description', 'Resolution Action Updated Date',
       'Community Board', 'Borough', 'X Coordinate (State Plane)',
       'Y Coordinate (State Plane)', 'Park Facility Name', 'Park Borough',
       'School Name', 'School Number', 'School Region', 'School Code',
       'School Phone Number', 'School Address', 'School City', 'School State',
       'School Zip', 'School Not Found', 'School or Citywide Complaint',
       'Vehicle Type', 'Taxi Company Borough', 'Taxi Pick Up Location',
       'Bridge Highway Name', 'Bridge Highway Direction', 'Road Ramp',
       'Bridge Highway Segment', 'Garage Lot Name', 'Ferry Direction',
       'Ferry Terminal Name', 'Latitude', 'Longitude', 'Location'],
      dtype='object')
```

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

Out[7]:

	Unique Key	Incident Zip	X Coordinate (State Plane)	Y Coordinate (State Plane)	School or Citywide Complaint	Vehicle Type	Taxi Company Borough	Taxi Pick Up Location	Ga N
count	3.006980e+05	298083.000000	2.971580e+05	297158.000000	0.0	0.0	0.0	0.0	
mean	3.130054e+07	10848.888645	1.004854e+06	203754.534416	NaN	NaN	NaN	NaN	
std	5.738547e+05	583.182081	2.175338e+04	29880.183529	NaN	NaN	NaN	NaN	
min	3.027948e+07	83.000000	9.133570e+05	121219.000000	NaN	NaN	NaN	NaN	
25%	3.080118e+07	10310.000000	9.919752e+05	183343.000000	NaN	NaN	NaN	NaN	
50%	3.130436e+07	11208.000000	1.003158e+06	201110.500000	NaN	NaN	NaN	NaN	
75%	3.178446e+07	11238.000000	1.018372e+06	224125.250000	NaN	NaN	NaN	NaN	
max	3.231065e+07	11697.000000	1.067173e+06	271876.000000	NaN	NaN	NaN	NaN	

In [8]:
```
df['Resolution Description'].value_counts().head()
```

Out[8]:
```
The Police Department responded to the complaint and with the information available obse
rved no evidence of the violation at that time.    98490
The Police Department responded to the complaint and took action to fix the condition.
61624
The Police Department responded and upon arrival those responsible for the condition wer
e gone.                                            58831
The Police Department responded to the complaint and determined that police action was n
ot necessary.                                      38211
The Police Department issued a summons in response to the complaint.
28246
Name: Resolution Description, dtype: int64
```

In [9]:
```
df = df.drop(columns=['School Name', 'School Number', 'School Region', 'School Code',
                      'School Phone Number', 'School Address', 'School City', 'S
                        'School Zip', 'School Not Found', 'School or Citywide C
```

In [10]:
```
df = df.drop(columns=['Vehicle Type','Taxi Company Borough','Taxi Pick Up Location'],axi
```

In [11]:
```
df= df.drop(columns=['Garage Lot Name','Ferry Direction','Ferry Terminal Name'],axis=1)
```

In [12]:
```
df.info()
```

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 300698 entries, 0 to 300697
Data columns (total 36 columns):
 #   Column                          Non-Null Count   Dtype  
---  ------                          --------------   -----  
 0   Unique Key                      300698 non-null  int64  
 1   Created Date                    300698 non-null  object 
 2   Closed Date                     298534 non-null  object 
 3   Agency                          300698 non-null  object 
 4   Agency Name                     300698 non-null  object 
 5   Complaint Type                  300698 non-null  object 
 6   Descriptor                      294784 non-null  object 
 7   Location Type                   300567 non-null  object 
 8   Incident Zip                    298083 non-null  float64
 9   Incident Address                256288 non-null  object 
 10  Street Name                     256288 non-null  object 
 11  Cross Street 1                  251419 non-null  object 
 12  Cross Street 2                  250919 non-null  object 
 13  Intersection Street 1           43858 non-null   object 
 14  Intersection Street 2           43362 non-null   object 
 15  Address Type                    297883 non-null  object 
 16  City                            298084 non-null  object 
 17  Landmark                        349 non-null     object 
 18  Facility Type                   298527 non-null  object 
 19  Status                          300698 non-null  object 
 20  Due Date                        300695 non-null  object 
 21  Resolution Description          300698 non-null  object 
 22  Resolution Action Updated Date  298511 non-null  object 
 23  Community Board                 300698 non-null  object 
 24  Borough                         300698 non-null  object 
 25  X Coordinate (State Plane)      297158 non-null  float64
 26  Y Coordinate (State Plane)      297158 non-null  float64
 27  Park Facility Name              300698 non-null  object 
 28  Park Borough                    300698 non-null  object 
 29  Bridge Highway Name             243 non-null     object 
 30  Bridge Highway Direction        243 non-null     object 
 31  Road Ramp                       213 non-null     object 
 32  Bridge Highway Segment          213 non-null     object 
 33  Latitude                        297158 non-null  float64
 34  Longitude                       297158 non-null  float64
 35  Location                        297158 non-null  object 
dtypes: float64(5), int64(1), object(30)
memory usage: 82.6+ MB
```

In [13]:
```
df.isnull().sum()/len(df)*100
```

Out[13]:

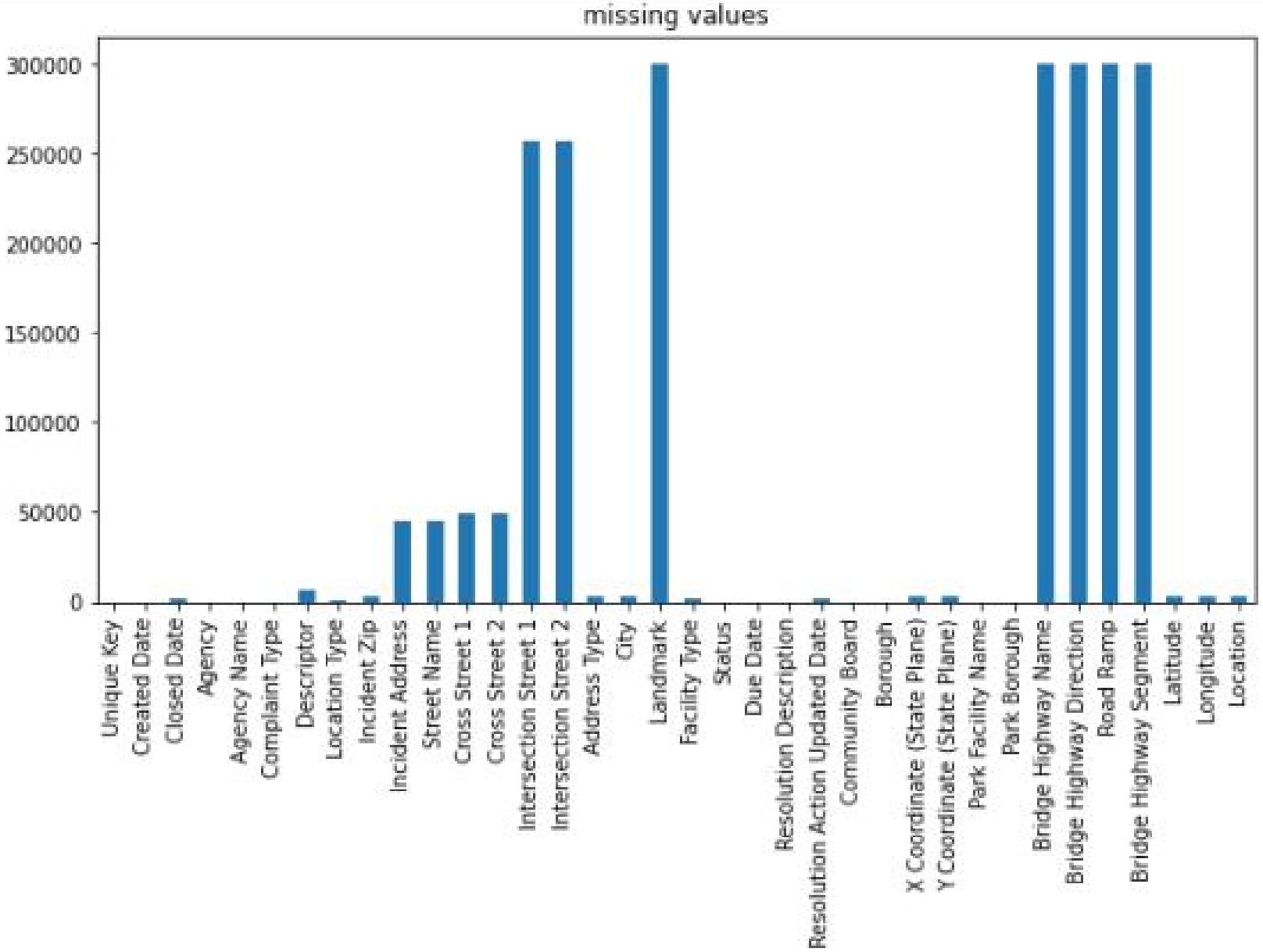
```
Unique Key                        0.000000
Created Date                      0.000000
Closed Date                       0.719659
Agency                            0.000000
Agency Name                       0.000000
Complaint Type                    0.000000
Descriptor                        1.966757
Location Type                     0.043565
Incident Zip                      0.869643
Incident Address                 14.768971
Street Name                      14.768971
Cross Street 1                   16.388203
Cross Street 2                   16.554483
Intersection Street 1            85.414602
Intersection Street 2            85.579552
Address Type                      0.936155
City                              0.869311
Landmark                         99.883937
Facility Type                     0.721987
Status                            0.000000
Due Date                          0.000998
Resolution Description            0.000000
Resolution Action Updated Date    0.727308
Community Board                   0.000000
Borough                           0.000000
X Coordinate (State Plane)        1.177261
Y Coordinate (State Plane)        1.177261
Park Facility Name                0.000000
Park Borough                      0.000000
Bridge Highway Name              99.919188
Bridge Highway Direction         99.919188
Road Ramp                        99.929165
Bridge Highway Segment           99.929165
Latitude                          1.177261
Longitude                         1.177261
Location                          1.177261
dtype: float64
```

In [14]:

```
df.isnull().sum().plot(kind='bar', figsize=(10,5),title = 'missing values')
```

Out[14]:

```
<AxesSubplot:title={'center':'missing values'}>
```



In [15]:
```
df.groupby(['Borough','Complaint Type','Descriptor']).size()
```

Out[15]:

Borough	Complaint Type	Descriptor	
BRONX	Animal Abuse	Chained	132
		In Car	36
		Neglected	673
		No Shelter	71
		Other (complaint details)	311
		Tortured	192
	Blocked Driveway	No Access	9884
		Partial Access	2871
	Derelict Vehicle	With License Plate	1953
	Disorderly Youth	Nuisance/Truant	7
		Playing in Unsuitable Place	56
	Drinking	After Hours - Licensed Est	10
		In Public	142
		Underage - Licensed Est	36
	Graffiti	Police Report Not Requested	1
		Police Report Requested	8
	Illegal Parking	Blocked Hydrant	1413
		Blocked Sidewalk	1206
		Commercial Overnight Parking	1151
		Detached Trailer	73
		Double Parked Blocking Traffic	593
		Double Parked Blocking Vehicle	830
		Overnight Commercial Storage	282
		Posted Parking Sign Violation	2342
		Unauthorized Bus Layover	49
	Noise - Commercial	Banging/Pounding	162
		Car/Truck Horn	28
		Car/Truck Music	90
		Loud Music/Party	1490
		Loud Talking	662
		Loud Television	2
	Noise - House of Worship	Banging/Pounding	7
		Loud Music/Party	45
		Loud Talking	27
	Noise - Park	Loud Music/Party	450
		Loud Talking	97
	Noise - Street/Sidewalk	Loud Music/Party	6173
		Loud Talking	2718
	Noise - Vehicle	Car/Truck Horn	177
		Car/Truck Music	2881
		Engine Idling	338
	Posting Advertisement	Building	4
		Vehicle	13
	Traffic	Chronic Speeding	32
		Chronic Stoplight Violation	30
		Congestion/Gridlock	250
		Drag Racing	24
		Truck Route Violation	21
	Vending	In Prohibited Area	107
		Unlicensed	272
BROOKLYN	Animal Abuse	Chained	165
		In Car	61
		Neglected	1196
		No Shelter	102
		Other (complaint details)	582
		Tortured	288
	Blocked Driveway	No Access	21422
		Partial Access	6726
	Derelict Vehicle	With License Plate	5181
	Disorderly Youth	Nuisance/Truant	14
		Playing in Unsuitable Place	58
	Drinking	After Hours - Licensed Est	19

		Underage - Licensed Est	48
	Graffiti	Police Report Not Requested	7
		Police Report Requested	36
	Illegal Parking	Blocked Hydrant	6697
		Blocked Sidewalk	4017
		Commercial Overnight Parking	4346
		Detached Trailer	132
		Double Parked Blocking Traffic	1958
		Double Parked Blocking Vehicle	1564
		Overnight Commercial Storage	558
		Posted Parking Sign Violation	7696
		Unauthorized Bus Layover	494
	Noise - Commercial	Banging/Pounding	1409
		Car/Truck Horn	338
		Car/Truck Music	339
		Loud Music/Party	8322
		Loud Talking	1027
		Loud Television	36
	Noise - House of Worship	Banging/Pounding	24
		Loud Music/Party	208
		Loud Talking	107
		Loud Television	1
	Noise - Park	Loud Music/Party	1245
		Loud Talking	310
	Noise - Street/Sidewalk	Loud Music/Party	8586
		Loud Talking	4769
	Noise - Vehicle	Car/Truck Horn	643
		Car/Truck Music	3127
		Engine Idling	1407
	Posting Advertisement	Building	16
		Vehicle	29
	Traffic	Chronic Speeding	57
		Chronic Stoplight Violation	66
		Congestion/Gridlock	618
		Drag Racing	41
		Truck Route Violation	304
	Vending	In Prohibited Area	178
		Unlicensed	337
MANHATTAN	Animal Abuse	Chained	69
		In Car	73
		Neglected	676
		No Shelter	72
		Other (complaint details)	488
		Tortured	147
	Blocked Driveway	No Access	1588
		Partial Access	485
	Derelict Vehicle	With License Plate	537
	Disorderly Youth	Nuisance/Truant	8
		Playing in Unsuitable Place	61
	Drinking	After Hours - Licensed Est	28
		In Public	191
		Underage - Licensed Est	76
	Graffiti	Police Report Not Requested	10
		Police Report Requested	12
	Illegal Parking	Blocked Hydrant	1701
		Blocked Sidewalk	1520
		Commercial Overnight Parking	467
		Detached Trailer	33
		Double Parked Blocking Traffic	1624
		Double Parked Blocking Vehicle	1253
		Overnight Commercial Storage	105
		Posted Parking Sign Violation	4933
		Unauthorized Bus Layover	496
	Noise - Commercial	Banging/Pounding	1481

		Car/Truck Horn	585
		Car/Truck Music	384
		Loud Music/Party	10669
		Loud Talking	1564
		Loud Television	37
	Noise - House of Worship	Banging/Pounding	11
		Loud Music/Party	157
		Loud Talking	26
	Noise - Park	Loud Music/Party	999
		Loud Talking	212
	Noise - Street/Sidewalk	Loud Music/Party	13798
		Loud Talking	6760
	Noise - Vehicle	Car/Truck Horn	1167
		Car/Truck Music	2862
		Engine Idling	1456
	Posting Advertisement	Building	23
		Vehicle	18
	Traffic	Chronic Speeding	34
		Chronic Stoplight Violation	84
		Congestion/Gridlock	1320
		Drag Racing	25
		Truck Route Violation	88
	Vending	In Prohibited Area	1526
		Unlicensed	873
QUEENS	Animal Abuse	Chained	133
		In Car	64
		Neglected	939
		No Shelter	112
		Other (complaint details)	458
		Tortured	178
	Blocked Driveway	No Access	22552
		Partial Access	9892
	Derelict Vehicle	With License Plate	8110
	Disorderly Youth	Nuisance/Truant	12
		Playing in Unsuitable Place	47
	Drinking	After Hours - Licensed Est	16
		In Public	245
		Underage - Licensed Est	96
	Graffiti	Police Report Not Requested	5
		Police Report Requested	32
	Illegal Parking	Blocked Hydrant	5230
		Blocked Sidewalk	3486
		Commercial Overnight Parking	4646
		Detached Trailer	156
		Double Parked Blocking Traffic	1154
		Double Parked Blocking Vehicle	469
		Overnight Commercial Storage	740
		Posted Parking Sign Violation	5867
		Unauthorized Bus Layover	234
	Noise - Commercial	Banging/Pounding	722
		Car/Truck Horn	110
		Car/Truck Music	134
		Loud Music/Party	4557
		Loud Talking	535
		Loud Television	17
	Noise - House of Worship	Banging/Pounding	202
		Loud Music/Party	83
		Loud Talking	14
	Noise - Park	Loud Music/Party	353
		Loud Talking	283
	Noise - Street/Sidewalk	Loud Music/Party	2613
		Loud Talking	1794
	Noise - Vehicle	Car/Truck Horn	477
		Car/Truck Music	1327

		Engine Idling	811
	Posting Advertisement	Building	16
		Vehicle	14
	Traffic	Chronic Speeding	104
		Chronic Stoplight Violation	88
		Congestion/Gridlock	468
		Drag Racing	67
		Truck Route Violation	586
	Vending	In Prohibited Area	207
		Unlicensed	270
STATEN ISLAND	Animal Abuse	Chained	36
		In Car	16
		Neglected	298
		No Shelter	25
		Other (complaint details)	136
		Tortured	46
	Blocked Driveway	No Access	1346
		Partial Access	796
	Derelict Vehicle	With License Plate	1766
	Disorderly Youth	Playing in Unsuitable Place	23
	Drinking	After Hours - Licensed Est	4
		In Public	157
		Underage - Licensed Est	14
	Graffiti	Police Report Requested	2
	Illegal Parking	Blocked Hydrant	818
		Blocked Sidewalk	725
		Commercial Overnight Parking	1319
		Detached Trailer	67
		Double Parked Blocking Traffic	288
		Double Parked Blocking Vehicle	92
		Overnight Commercial Storage	150
		Posted Parking Sign Violation	1373
		Unauthorized Bus Layover	62
	Noise - Commercial	Banging/Pounding	76
		Car/Truck Horn	4
		Car/Truck Music	15
		Loud Music/Party	534
		Loud Talking	50
	Noise - House of Worship	Banging/Pounding	4
		Loud Music/Party	10
		Loud Talking	3
	Noise - Park	Loud Music/Party	41
		Loud Talking	26
	Noise - Street/Sidewalk	Loud Music/Party	452
		Loud Talking	368
	Noise - Vehicle	Car/Truck Horn	51
		Car/Truck Music	139
		Engine Idling	166
	Posting Advertisement	Building	1
		Vehicle	515
	Traffic	Chronic Speeding	41
		Chronic Stoplight Violation	12
		Congestion/Gridlock	113
		Drag Racing	18
		Truck Route Violation	14
	Vending	In Prohibited Area	5
		Unlicensed	20
Unspecified	Agency Issues	Language Access Complaint	6
	Animal Abuse	In Car	1
		Neglected	5
		Other (complaint details)	2
		Tortured	3
	Animal in a Park	Animal Waste	1
	Blocked Driveway	No Access	184

```
                                Partial Access                  98
                Derelict Vehicle With License Plate            171
                Drinking        In Public                        7
                                Underage - Licensed Est          1
                Ferry Complaint Disruptive Passenger             1
                                Homeless Issue                   1
                Illegal Parking Blocked Hydrant                222
                                Blocked Sidewalk               167
                                Commercial Overnight Parking   260
                                Detached Trailer                 3
                                Double Parked Blocking Traffic 122
                                Double Parked Blocking Vehicle   3
                                Overnight Commercial Storage     2
                                Posted Parking Sign Violation  229
                                Unauthorized Bus Layover        32
                Noise - Commercial Banging/Pounding             67
                                Car/Truck Horn                  14
                                Car/Truck Music                 17
                                Loud Music/Party               215
                                Loud Talking                    53
                Noise - House of Worship Loud Music/Party        1
                                Loud Talking                     1
                Noise - Park    Loud Music/Party                21
                                Loud Talking                     5
                Noise - Street/Sidewalk Loud Music/Party       416
                                Loud Talking                   173
                Noise - Vehicle Car/Truck Horn                   5
                                Car/Truck Music                 38
                                Engine Idling                   11
                Posting Advertisement Vehicle                    1
                Traffic         Truck Route Violation            1
                Vending         In Prohibited Area               2
                                Unlicensed                       5
dtype: int64
```

In [ ]:

In [43]:
```
#TASK 2
df["Created Date"]=pd.to_datetime(df["Created Date"])
df["Closed Date"]=pd.to_datetime(df["Closed Date"])
df['Request_closing_time'] = (df['Closed Date'] - df['Created Date']).dt.total_seconds()
```

In [44]:
```
df.head()
```

Out[44]:

	Unique Key	Created Date	Closed Date	Agency	Agency Name	Complaint Type	Descriptor	Location Type	Incident Zip	I A
0	32310363	2015-12-31 23:59:45	2016-01-01 00:55:00	NYPD	New York City Police Department	Noise - Street/Sidewalk	Loud Music/Party	Street/Sidewalk	10034.0	VERI A
1	32309934	2015-12-31 23:59:44	2016-01-01 01:26:00	NYPD	New York City Police Department	Blocked Driveway	No Access	Street/Sidewalk	11105.0	2 A
2	32309159	2015-12-31 23:59:29	2016-01-01 04:51:00	NYPD	New York City Police Department	Blocked Driveway	No Access	Street/Sidewalk	10458.0	VALI A
3	32305098	2015-12-31 23:57:46	2016-01-01 07:43:00	NYPD	New York City Police Department	Illegal Parking	Commercial Overnight Parking	Street/Sidewalk	10461.0	B A
4	32306529	2015-12-31 23:56:58	2016-01-01 03:24:00	NYPD	New York City Police Department	Illegal Parking	Blocked Sidewalk	Street/Sidewalk	11373.0	8

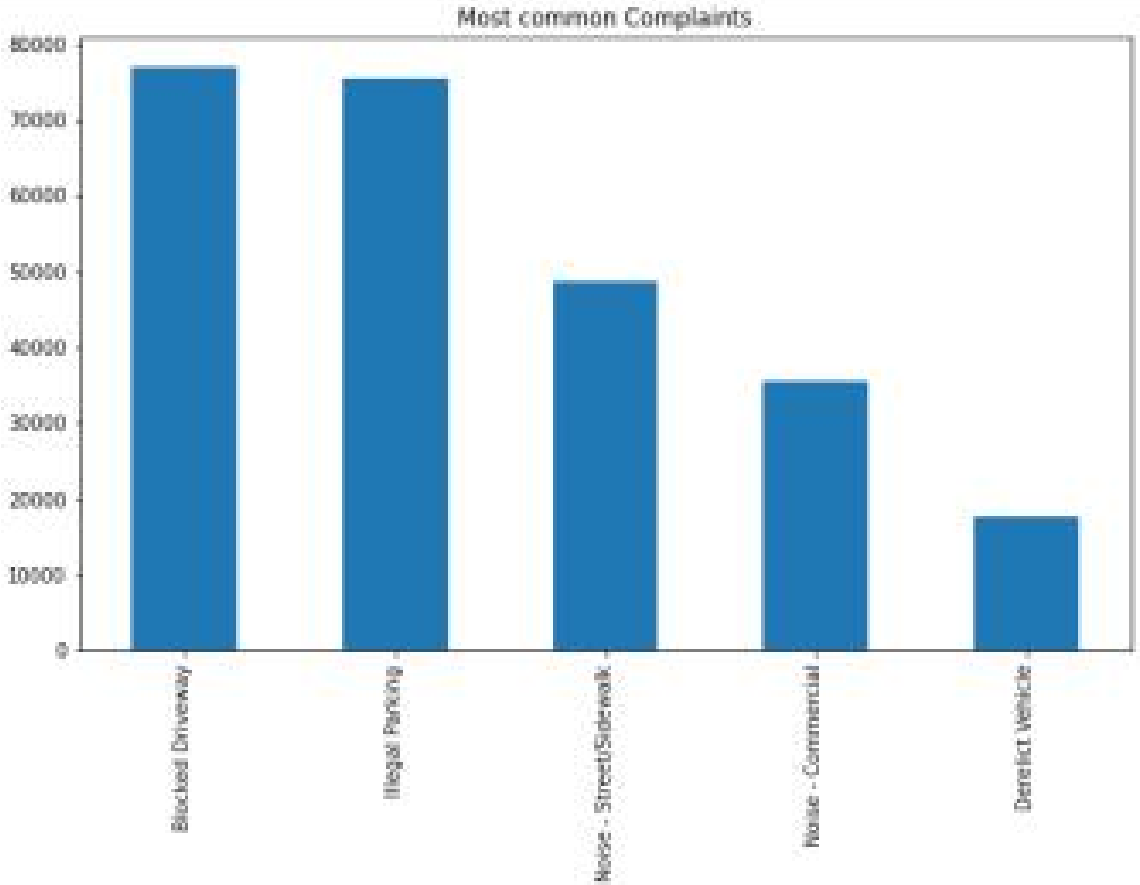
5 rows × 38 columns

In [45]:
```
#Task 3
#Prepare data
def prepareData(df):
    df['R'] = (df['Closed Date'] - df['Created Date']).dt.total_seconds()/3600 #in hours
    df_clean=df[df['Request_closing_time'].notnull()]
    df_perfect = df_clean[df_clean['Closed Date'] >= df_clean['Created Date']]
    df_perfect['Day of Week'] = df_perfect['Created Date'].dt.dayofweek
    df_perfect['Day of Month'] = df_perfect['Created Date'].dt.day
    df_perfect['Month'] = df_perfect['Created Date'].dt.month
    df_perfect['Year'] = df_perfect['Created Date'].dt.year
    df_perfect=df_perfect[df_perfect.Borough!='Unspecified']
    return df_perfect
#shape
df_perfect = prepareData(df)
df_perfect.shape
```

Out[45]:
```
(298068, 43)
```

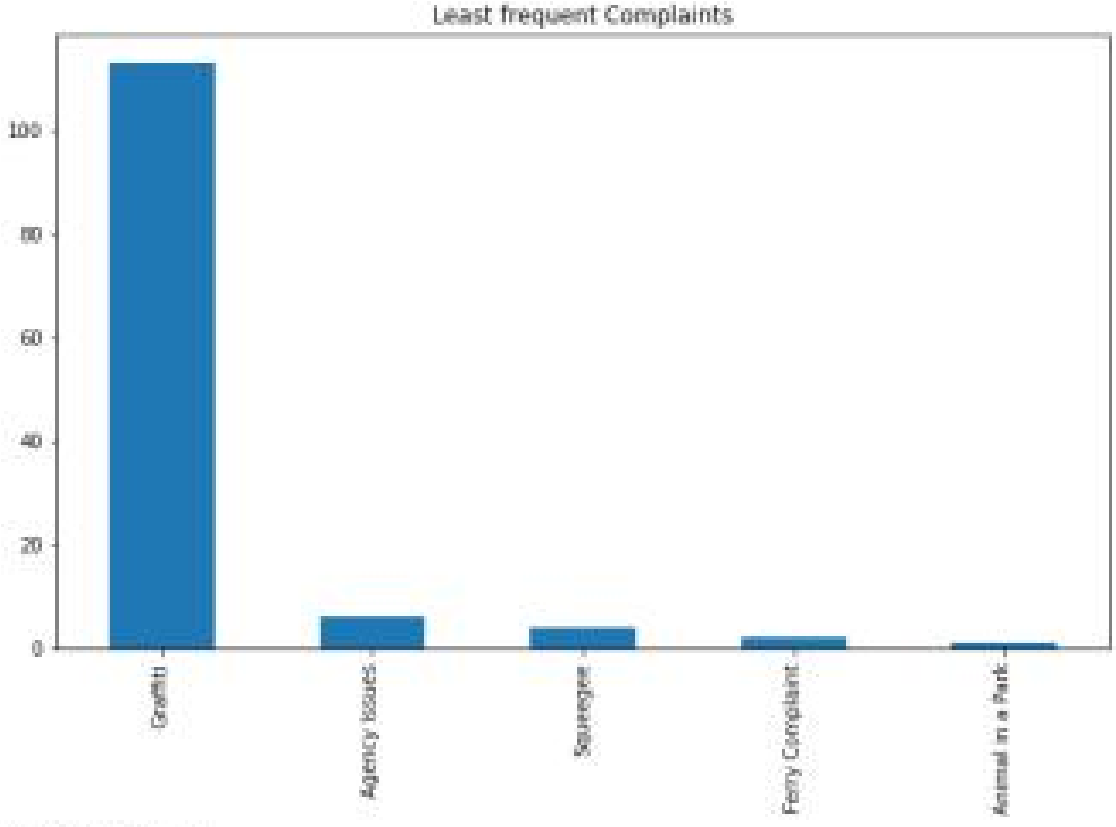
In [23]:
```
(df['Complaint Type'].value_counts()).head().plot(kind='bar',
                    figsize=(10,6), title = 'Most common Complaints')
```

Out[23]:
```
<AxesSubplot:title={'center':'Most common Complaints'}>
```

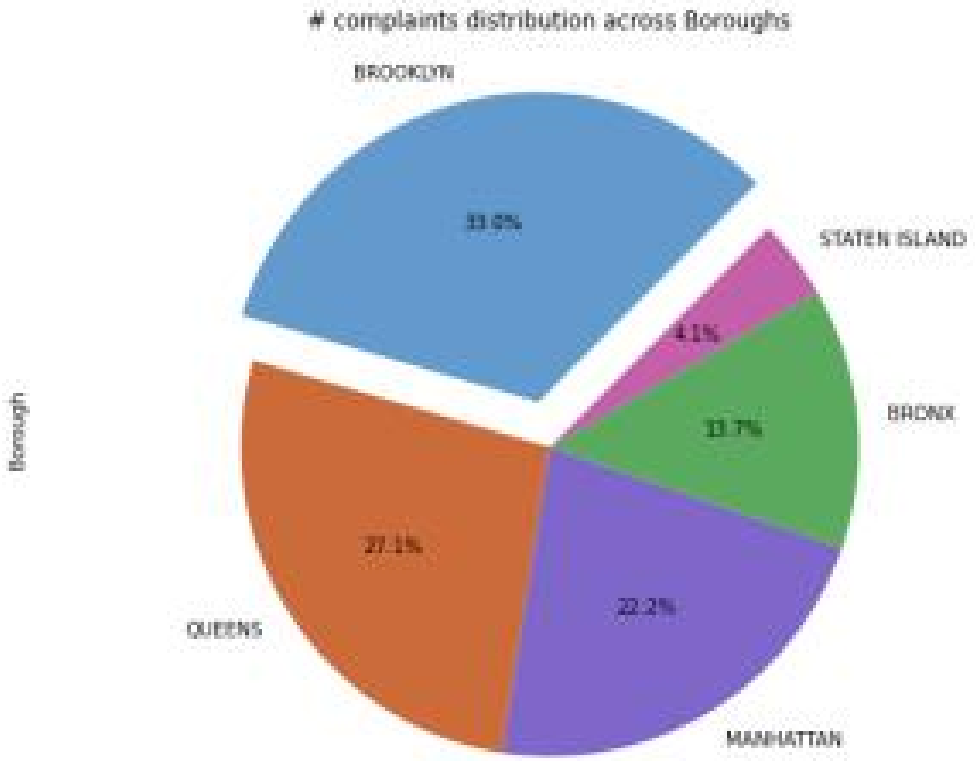


In [24]:
```
#Least frequent Complaints
(df['Complaint Type'].value_counts()).tail().plot(kind='bar',
                        figsize=(10,6), title = 'Least frequent Complaints')
```

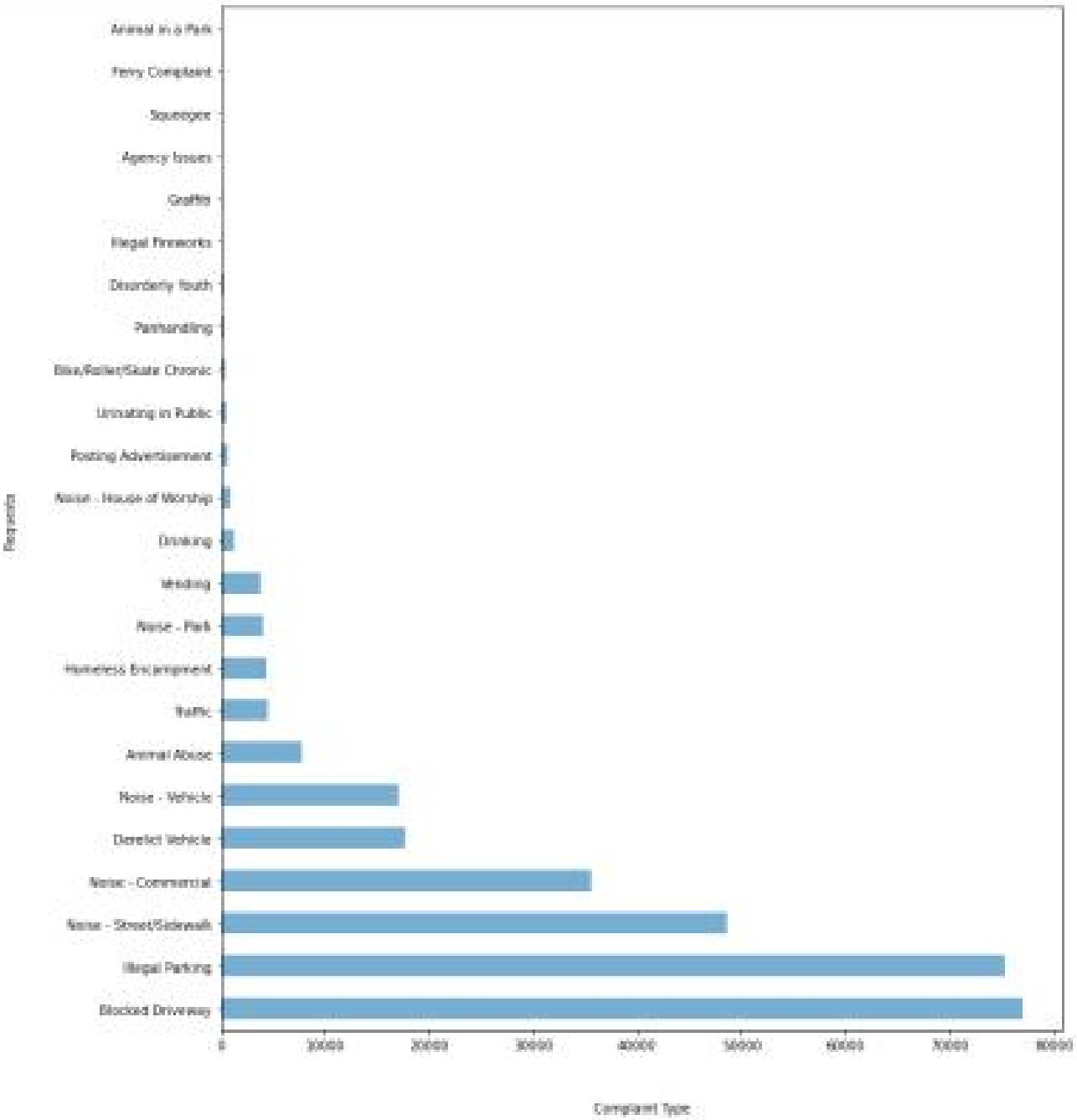
Out[24]:
```
<AxesSubplot:title={'center':'Least frequent Complaints'}>
```



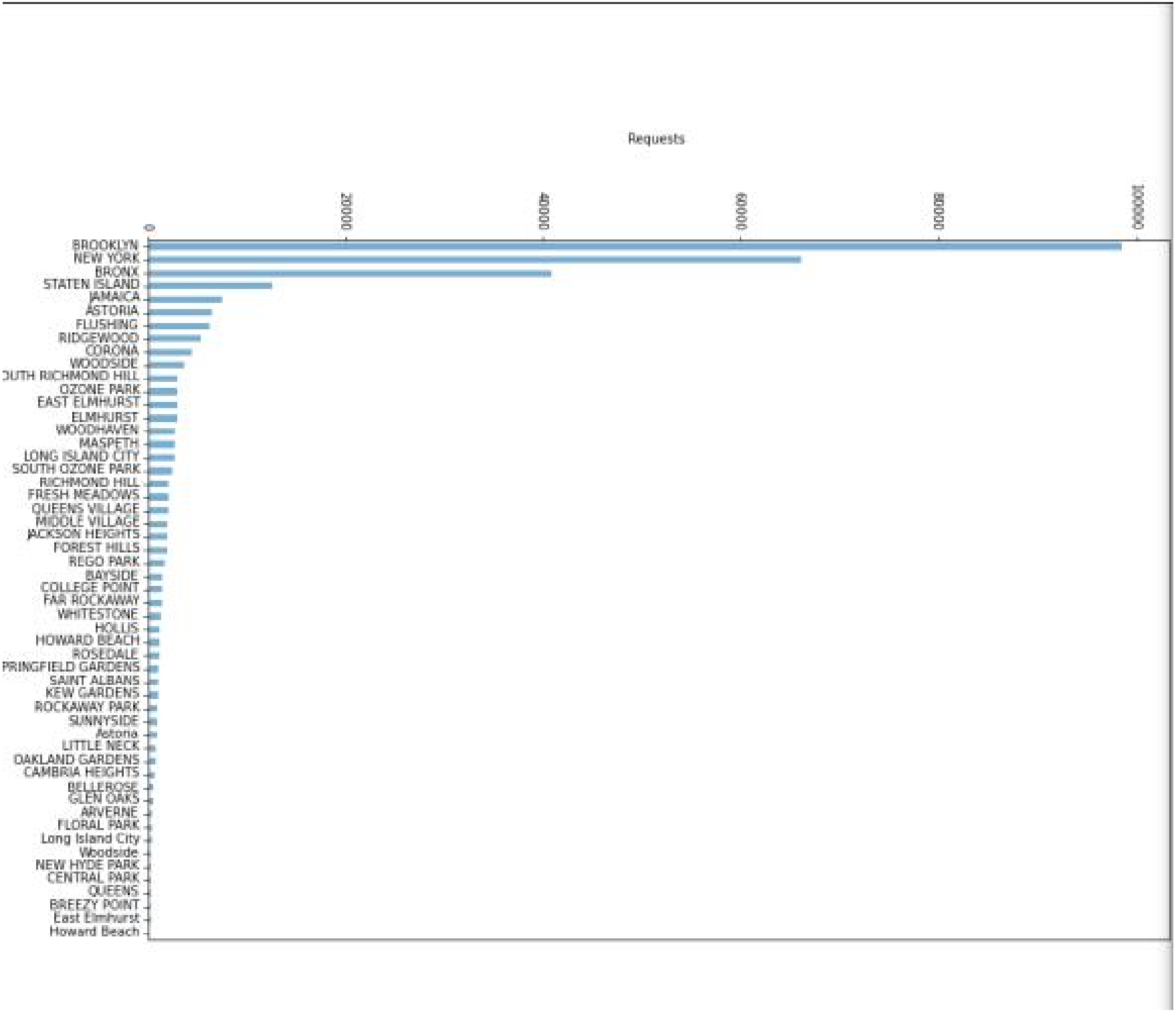
In [28]:
```
# complaints distribution across Boroughs
#conclusion 1
colors = ['#639ace','#ca6b39','#7f67ca','#5ba85f','#c368aa','#a7993f','#cc566a']
df_perfect['Borough'].value_counts().plot(kind='pie',autopct='%1.1f%%',
                          explode = (0.15, 0, 0, 0, 0), startangle=45, shadow=False, color
                          figsize = (8,6))
#plt.legend(title='BOROUGH', loc='upper right', bbox_to_anchor=(1.5,1))
plt.axis('equal')
plt.title('# complaints distribution across Boroughs\n')
plt.tight_layout()
plt.show()
```



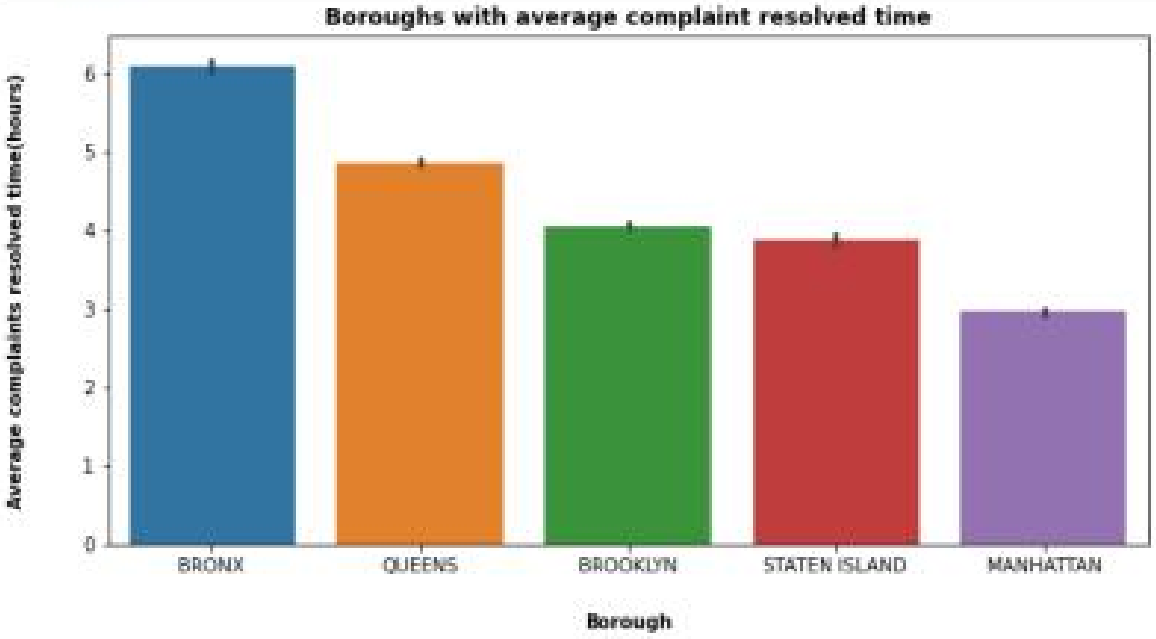
In [30]:
```
#Conclusion 2
#Most number of complaint request received is for  Blocked  Driveway
df['Complaint Type'].value_counts().plot(kind='barh',alpha=0.6, figsize=(12,15))
plt.xlabel('Complaint Type',labelpad=40)
plt.ylabel('Requests',labelpad=30)
plt.show()
```



In [33]:
```
#conclusion 3
##Most number of complaint requests received are from BROOKLYN city
df['City'].value_counts().plot(kind='bar', alpha=0.6, figsize=(10,15))
plt.xlabel('City',labelpad=20)
plt.ylabel('Requests',labelpad=30)
plt.show()
```

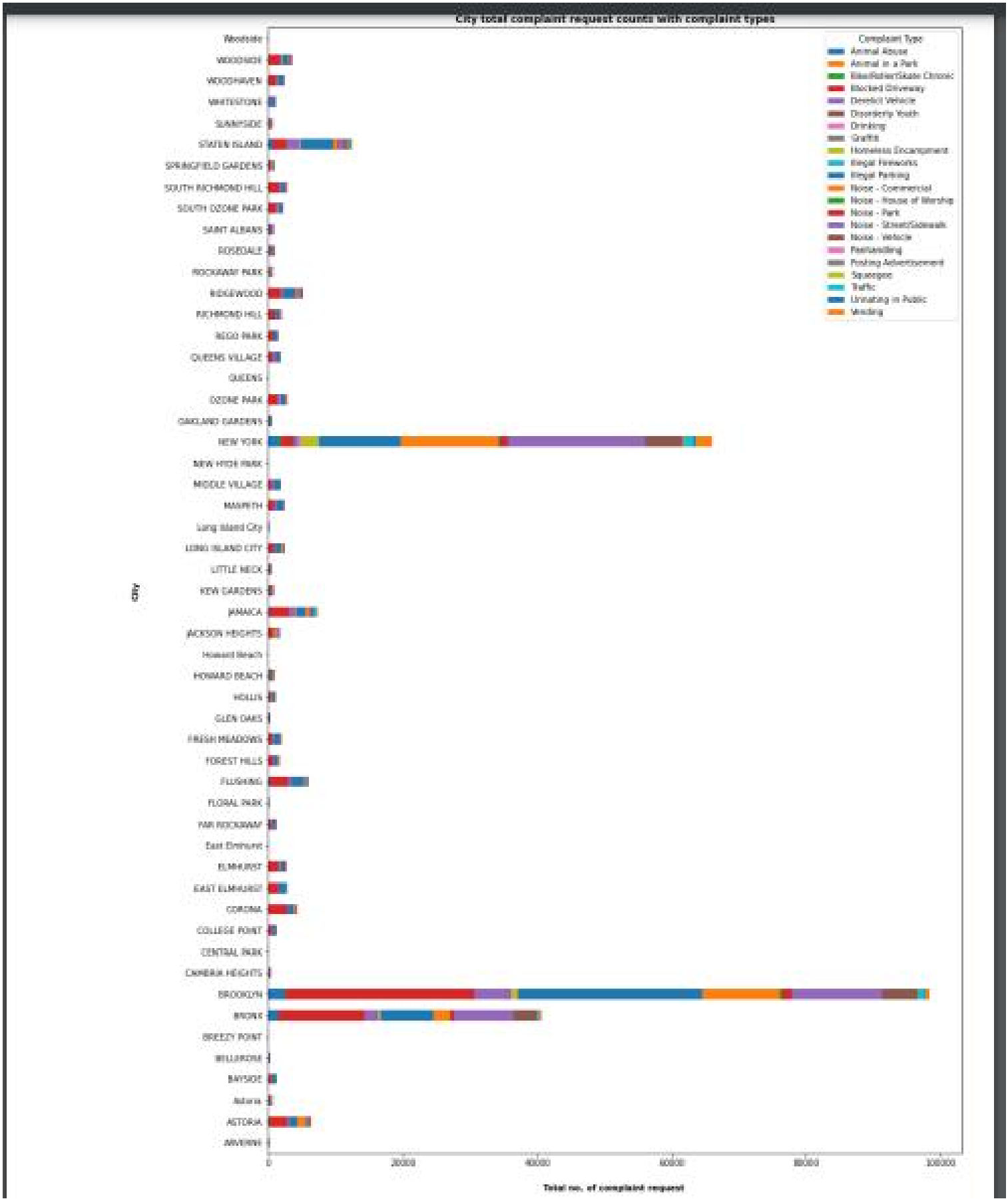

Requests
0
20000
40000
60000
80000
100000
BROOKLYN
NEW YORK
BRONX
STATEN ISLAND
JAMAICA
ASTORIA
FLUSHING
RIDGEWOOD
CORONA
WOODSIDE
SOUTH RICHMOND HILL
OZONE PARK
EAST ELMHURST
ELMHURST
WOODHAVEN
MASPETH
LONG ISLAND CITY
SOUTH OZONE PARK
RICHMOND HILL
FRESH MEADOWS
QUEENS VILLAGE
MIDDLE VILLAGE
JACKSON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

In [41]:
```
#Conclusion 4
#visulaizing  Boroughs with average complaint types
txt={'weight':'bold'}
plt.figure(figsize=(10,5))
sns.barplot(x='Borough',y='Resolution_Time',data=df,order=['BRONX','QUEENS','BROOKLYN','
plt.title("Boroughs with average complaint resolved time",fontdict=txt)
plt.xlabel("Borough",fontdict=txt,labelpad=20)
plt.ylabel("Average complaints resolved time(hours)",fontdict=txt,labelpad=30)
plt.show()
```

In [42]:
```
#Conclusion 5
# visualizing Cities with number of complaint requests received and its complaint types
City_Complaint_Types= pd.crosstab(index=df['City'],columns=df['Complaint Type'])
txt={'weight':'bold'}
plt.figure(figsize=(20,10))
City_Complaint_Types.plot(kind='barh',figsize=(15,25),stacked=True)
plt.title("City total complaint request counts with complaint types",fontdict=txt)
plt.xlabel("Total no. of complaint request ",fontdict=txt,labelpad=20)
plt.ylabel("City",fontdict=txt,labelpad=30)
plt.show()
```

```
<Figure size 1440x720 with 0 Axes>
```


City total complaint request counts with complaint types
Complaint Type
Animal Abuse
Animal in a Park
Bike/Roller/Skate Chronic
Blocked Driveway
Derelict Vehicle
Disorderly Youth
Drinking
Graffiti
Homeless Encampment
Illegal Fireworks
Illegal Parking
Noise - Commercial
Noise - House of Worship
Noise - Park
Noise - Street/Sidewalk
Noise - Vehicle
Panhandling
Posting Advertisement
Squeegee
Traffic
Urinating in Public
Vending
City
Total no. of complaint request
0
20000
40000
60000
80000
100000

In [46]:
```
#Task  4
grouped_data = df.groupby(['Complaint Type','Borough'])[['Request_closing_time']].mean()
```


In [47]:
```
grouped_data.head()
```

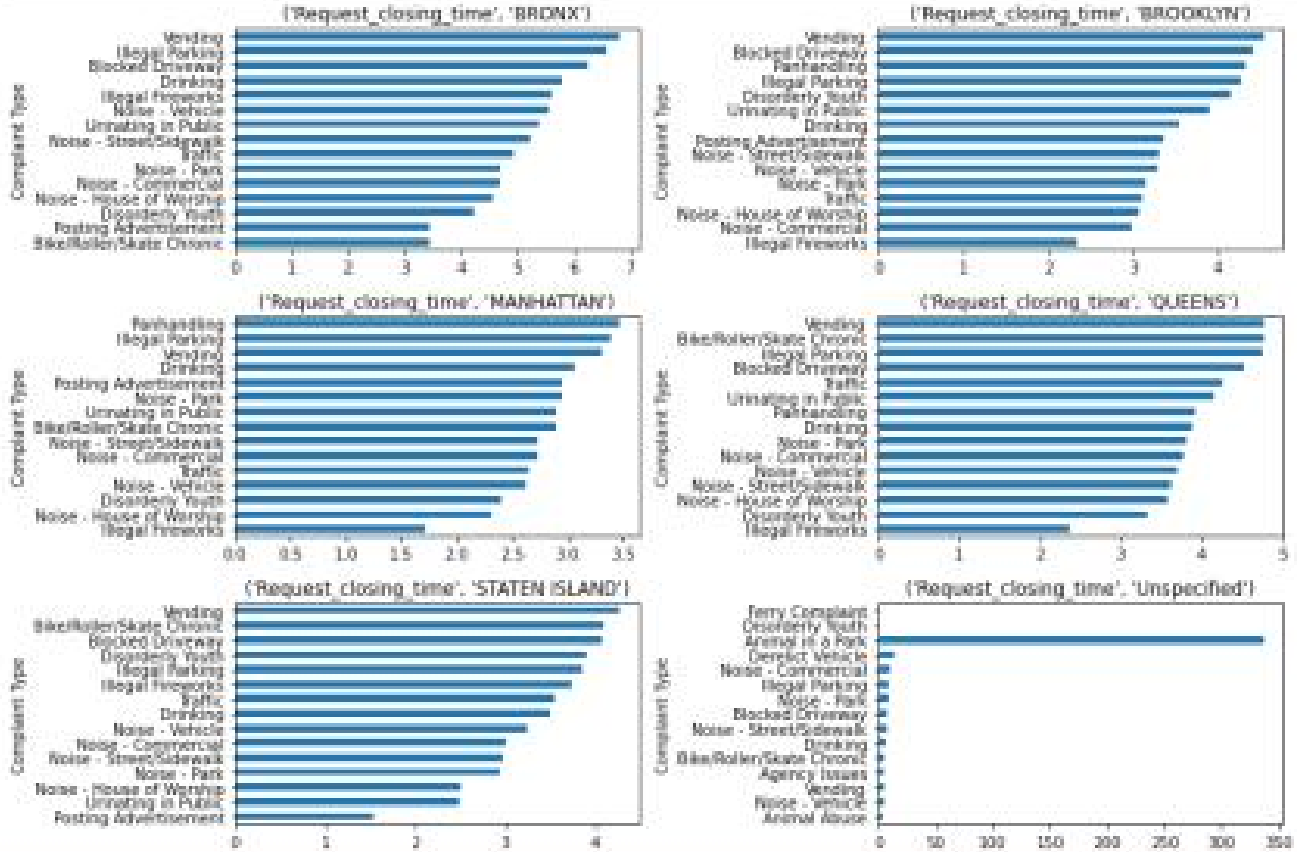
Out[47]:

	Request_closing_time					
Borough	**BRONX**	**BROOKLYN**	**MANHATTAN**	**QUEENS**	**STATEN ISLAND**	**Unspecified**
Complaint Type						
Agency Issues	NaN	NaN	NaN	NaN	NaN	5.260324
Animal Abuse	7.335495	4.832506	3.685151	5.413888	4.969506	3.409722
Animal in a Park	NaN	NaN	NaN	NaN	NaN	336.834722
Bike/Roller/Skate Chronic	3.458972	5.004705	2.892225	4.767222	4.077381	5.860000
Blocked Driveway	6.261764	4.410784	3.557437	4.537609	4.071254	8.466335

In [48]:
```
col_number = 2
row_number = 3
fig, axes = plt.subplots(row_number,col_number, figsize=(12,8))

for i, (label,col) in enumerate(grouped_data.iteritems()):
    ax = axes[int(i/col_number), i%col_number]
    col = col.sort_values(ascending=True)[:15]
    col.plot(kind='barh', ax=ax)
    ax.set_title(label)

plt.tight_layout()
```



In [ ]:
```
REQUEST CLOSING TIME IN HOURS FOR DIFFERENT LOCATION, ORDER BY COMPLAINTS
```

In [51]:
```
#TASK 5
df_avg_res_time_city = df_perfect.groupby(['City','Complaint Type']).Request_closing_tim
#df_perfect.sort_values('Complaint Type').groupby('City')
df_avg_res_time_city.head(300)
```


Out[51]:

City	Complaint Type	
ARVERNE	Animal Abuse	2.153626
	Blocked Driveway	2.525968
	Derelict Vehicle	2.968220
	Disorderly Youth	3.591250
	Drinking	0.238611
	Graffiti	1.533333
	Homeless Encampment	1.814792
	Illegal Parking	2.316365
	Noise - Commercial	2.287222
	Noise - House of Worship	1.562197
	Noise - Park	1.283333
	Noise - Street/Sidewalk	1.992395
	Noise - Vehicle	1.859881
	Panhandling	1.033333
	Urinating in Public	0.691944
	Vending	0.483333
ASTORIA	Animal Abuse	5.080169
	Bike/Roller/Skate Chronic	1.739315
	Blocked Driveway	4.816118
	Derelict Vehicle	9.689250
	Disorderly Youth	2.902593
	Drinking	4.722421
	Graffiti	14.095869
	Homeless Encampment	4.917587
	Illegal Fireworks	2.773472
	Illegal Parking	4.833391
	Noise - Commercial	3.133829
	Noise - House of Worship	2.022222
	Noise - Park	2.995109
	Noise - Street/Sidewalk	3.451164
	Noise - Vehicle	3.589176
	Panhandling	1.150000
	Posting Advertisement	5.870000
	Traffic	5.410183
	Urinating in Public	4.625710
	Vending	4.935417
Astoria	Blocked Driveway	4.914629
	Derelict Vehicle	6.236157
	Illegal Parking	4.711290
	Noise - Commercial	3.541890
	Noise - Street/Sidewalk	3.713231
BAYSIDE	Animal Abuse	3.274512
	Blocked Driveway	2.563872
	Derelict Vehicle	3.359780
	Disorderly Youth	2.973856
	Drinking	1.900000
	Graffiti	4.551574
	Homeless Encampment	2.875556
	Illegal Parking	2.562667
	Noise - Commercial	2.233667
	Noise - House of Worship	3.534861
	Noise - Park	3.272569

	Illegal Parking	2.562667
	Noise - Commercial	2.233667
	Noise - House of Worship	3.534861
	Noise - Park	3.272569
	Noise - Street/Sidewalk	1.531800
	Noise - Vehicle	1.708420
	Traffic	1.526944
	Vending	1.877917
BELLEROSE	Animal Abuse	12.724683
	Bike/Roller/Skate Chronic	4.908888
	Blocked Driveway	10.099412
	Derelict Vehicle	17.168143
	Disorderly Youth	1.850000
	Drinking	3.918856
	Homeless Encampment	39.133333
	Illegal Fireworks	6.666667
	Illegal Parking	8.283126
	Noise - Commercial	6.748773
	Noise - House of Worship	2.196944
	Noise - Park	1.414167
	Noise - Street/Sidewalk	9.068226
	Noise - Vehicle	2.583167
	Panhandling	7.483333
	Posting Advertisement	2.260000
	Traffic	5.758849
	Urinating in Public	7.541389
BREEZY POINT	Animal Abuse	2.613472
	Blocked Driveway	1.327685
	Derelict Vehicle	7.145185
	Drinking	2.633333
	Illegal Parking	3.942222
	Noise - Commercial	2.539514
	Noise - Street/Sidewalk	1.000000
	Noise - Vehicle	1.324444
BRONX	Animal Abuse	7.335495
	Bike/Roller/Skate Chronic	3.458972
	Blocked Driveway	6.261764
	Derelict Vehicle	9.227574
	Disorderly Youth	4.238408
	Drinking	5.793582
	Graffiti	8.899843
	Homeless Encampment	7.441901
	Illegal Fireworks	5.688194
	Illegal Parking	6.588843
	Noise - Commercial	4.697351
	Noise - House of Worship	4.559546
	Noise - Park	4.698525
	Noise - Street/Sidewalk	5.226309
	Noise - Vehicle	5.568793
	Panhandling	14.213728
	Posting Advertisement	3.461993
	Traffic	4.923685
	Urinating in Public	5.398103
	Vending	6.825697
BROOKLYN	Animal Abuse	4.832586
	Bike/Roller/Skate Chronic	5.004705
	Blocked Driveway	4.418784
	Derelict Vehicle	5.947490
	Disorderly Youth	4.158517
	Drinking	3.548561

	Vending	0.829897
BROOKLYN	Animal Abuse	4.832586
	Bike/Roller/Skate Chronic	5.004705
	Blocked Driveway	4.418784
	Derelict Vehicle	5.947498
	Disorderly Youth	4.158517
	Drinking	3.540561
	Graffiti	8.243284
	Homeless Encampment	4.694820
	Illegal Fireworks	2.339586
	Illegal Parking	4.272241
	Noise - Commercial	2.986892
	Noise - House of Worship	3.069860
	Noise - Park	3.150531
	Noise - Street/Sidewalk	3.296002
	Noise - Vehicle	3.283443
	Panhandling	4.316480
	Posting Advertisement	3.360951
	Traffic	3.112931
	Urinating in Public	3.899314
	Vending	4.528105
CAMBRIA HEIGHTS	Animal Abuse	11.355833
	Blocked Driveway	7.695979
	Derelict Vehicle	16.118778
	Homeless Encampment	22.789444
	Illegal Fireworks	1.538556
	Illegal Parking	11.243436
	Noise - Commercial	3.818823
	Noise - House of Worship	2.640972
	Noise - Street/Sidewalk	4.608044
	Noise - Vehicle	6.917092
	Traffic	8.546713
CENTRAL PARK	Illegal Parking	4.594386
	Noise - Street/Sidewalk	3.266942
COLLEGE POINT	Animal Abuse	4.645159
	Blocked Driveway	3.380623
	Derelict Vehicle	3.518946
	Disorderly Youth	0.600000
	Graffiti	11.864167
	Homeless Encampment	1.443981
	Illegal Parking	3.125850
	Noise - Commercial	3.972143
	Noise - Park	3.178750
	Noise - Street/Sidewalk	3.016010
	Noise - Vehicle	2.650594
	Traffic	2.292639
	Vending	4.658889
CORONA	Animal Abuse	3.668251
	Blocked Driveway	3.320945
	Derelict Vehicle	4.584795
	Disorderly Youth	2.878796
	Drinking	3.647929
	Graffiti	0.733333
	Homeless Encampment	3.669883
	Illegal Parking	3.361137
	Noise - Commercial	2.622118
	Noise - House of Worship	3.753611
	Noise - Park	2.320891
	Noise - Street/Sidewalk	2.364379
	Noise - Vehicle	2.660775

	Panhandling	1.174722
	Posting Advertisement	1.540000
	Traffic	2.285856
	Urinating in Public	1.952262
	Vending	3.181577
EAST ELMHURST	Animal Abuse	4.055861
	Bike/Roller/Skate Chronic	0.252222
	Blocked Driveway	3.733274
	Derelict Vehicle	5.886735
	Disorderly Youth	6.900000
	Drinking	1.861389
	Graffiti	7.647315
	Homeless Encampment	6.382917
	Illegal Parking	3.378084
	Noise - Commercial	2.352984
	Noise - House of Worship	1.438858
	Noise - Park	2.652833
	Noise - Street/Sidewalk	2.474237
	Noise - Vehicle	2.187741
	Posting Advertisement	1.858611
	Traffic	2.888861
	Urinating in Public	2.097167
	Vending	4.829831
ELMHURST	Animal Abuse	3.871988
	Bike/Roller/Skate Chronic	4.825694
	Blocked Driveway	3.433319
	Derelict Vehicle	4.818305
	Disorderly Youth	0.858194
	Drinking	3.862158
	Homeless Encampment	3.874705

	Illegal Fireworks	0.983333
	Illegal Parking	3.278206
	Noise - Commercial	2.790147
	Noise - House of Worship	1.884558
	Noise - Park	2.314551
	Noise - Street/Sidewalk	2.589251
	Noise - Vehicle	2.826909
	Panhandling	3.302315
	Posting Advertisement	0.733333
	Traffic	2.628532
	Urinating in Public	2.537833
	Vending	3.994378
East Elmhurst	Derelict Vehicle	0.485833
	Illegal Parking	5.783333
FAR ROCKAWAY	Animal Abuse	2.718200
	Blocked Driveway	2.634521
	Derelict Vehicle	3.665881
	Disorderly Youth	3.365833
	Drinking	2.418264
	Homeless Encampment	3.819147
	Illegal Parking	2.786477
	Noise - Commercial	1.926171
	Noise - House of Worship	1.133333
	Noise - Park	1.584686
	Noise - Street/Sidewalk	3.038487
	Noise - Vehicle	2.188743
	Traffic	2.385750
	Urinating in Public	1.585556
	Vending	2.804846
FLORAL PARK	Animal Abuse	26.582381
	Blocked Driveway	7.701278
	Derelict Vehicle	16.654871
	Disorderly Youth	3.258333
	Drinking	8.233333
	Illegal Parking	9.254314
	Noise - Commercial	4.388241
	Noise - Street/Sidewalk	6.988981
	Noise - Vehicle	1.952381
FLUSHING	Animal Abuse	3.557939
	Bike/Roller/Skate Chronic	5.733856
	Blocked Driveway	2.950928
	Derelict Vehicle	3.692664
	Disorderly Youth	2.144167
	Drinking	3.852410
	Graffiti	1.722569
	Homeless Encampment	3.886421
	Illegal Fireworks	0.691667
	Illegal Parking	2.982854
	Noise - Commercial	2.845192
	Noise - House of Worship	3.648333
	Noise - Park	2.888948
	Noise - Street/Sidewalk	2.838795
	Noise - Vehicle	3.359858
	Panhandling	1.146806
	Posting Advertisement	0.133333
	Traffic	2.863678
	Urinating in Public	1.225208
	Vending	2.449958
FOREST HILLS	Animal Abuse	3.283204
	Bike/Roller/Skate Chronic	4.823722
	Blocked Driveway	3.715294
	Derelict Vehicle	3.725887
	Disorderly Youth	4.151944
	Drinking	4.100000

```
                        Graffiti                      3.272222
                        Homeless Encampment           3.510877
                        Illegal Fireworks             0.666667
                        Illegal Parking               3.324728
                        Noise - Commercial            1.888749
                        Noise - House of Worship      3.903889
                        Noise - Park                  1.539389
                        Noise - Street/Sidewalk       2.320825
                        Noise - Vehicle               2.383696
                        Panhandling                   5.814167
                        Posting Advertisement         5.888056
                        Traffic                       2.607287
                        Urinating in Public           1.388194
                        Vending                       2.744778
FRESH MEADOWS           Animal Abuse                  3.279173
                        Blocked Driveway              3.973748
                        Derelict Vehicle              4.511535
                        Drinking                      3.738944
                        Homeless Encampment           5.899222
                        Illegal Parking               2.623412
                        Noise - Commercial            2.488869
                        Noise - Park                  3.416878
                        Noise - Street/Sidewalk       2.488247
                        Noise - Vehicle               2.145489
                        Panhandling                   1.578278
                        Traffic                       1.631517
                        Urinating in Public           1.688889
                        Vending                       2.631111
GLEN OAKS               Animal Abuse                 12.992944
                        Blocked Driveway             11.299241
                        Derelict Vehicle             15.244989
                        Illegal Parking               8.967151
                        Noise - Commercial            6.117639
                        Noise - Park                  4.482740
                        Noise - Street/Sidewalk      11.914213
                        Noise - Vehicle              11.421736
                        Traffic                       5.547963
                        Urinating in Public          11.688333
                        Vending                       4.843796
HOLLIS                  Animal Abuse                  4.371911
                        Blocked Driveway              4.795284
                        Derelict Vehicle             11.564689
                        Disorderly Youth              3.866667
                        Drinking                      7.298981
                        Homeless Encampment           4.284383
Name: Request_closing_time, dtype: float64
```

In [ ]:

In [ ]: AVERAGE RESPONSE TIME ACROSS COMPLAINT TYPES IN HOURS

In [52]:
```
df_avg_res_time = df_perfect.groupby('Complaint Type').Request_closing_time.mean().sort_
df_avg_res_time.head(30)
```

Out[52]:
```
Complaint Type
Posting Advertisement        1.975883
Illegal Fireworks            2.761139
Noise - Commercial           3.136565
Noise - House of Worship     3.193298
Noise - Park                 3.401711
Noise - Street/Sidewalk      3.438261
Traffic                      3.448681
Disorderly Youth             3.558584
Noise - Vehicle              3.588587
Urinating in Public          3.626664
Bike/Roller/Skate Chronic    3.758540
Drinking                     3.855364
Vending                      4.013628
Squeegee                     4.045625
Homeless Encampment          4.365570
Panhandling                  4.372768
Illegal Parking              4.485067
Blocked Driveway             4.738139
Animal Abuse                 5.213476
Graffiti                     7.151251
Derelict Vehicle             7.346087
Name: Request_closing_time, dtype: float64
```

In [ ]:
```
From the above data null hypothesis can be rejected . Since the average response time ac
Null Hypothesis : Average response time across complaint type are equal.
Alternate Hypothesis : Average response time across complaint type are equal
Following complains have resolution times which are very close.
Disorderly Youth 12818.982098 Noise - Vehicle 12018.914439
One group can be formed for these complaints and one way Anova for this can be performed
```

In [54]:
```
df_dis_youth = df_perfect[df_perfect['Complaint Type']=='Disorderly Youth']
df_dis_youth = df_dis_youth.loc[:,['Request_closing_time']]
df_dis_youth.head()
```

Out[54]:

	Request_closing_time
4870	0.198056
9034	1.279167
12527	0.851389
12178	5.393056
17181	1.902500

In [55]:
```
df_noise_veh = df_perfect[df_perfect['Complaint Type']=='Noise - Vehicle']
df_noise_veh = df_noise_veh.loc[:,['Request_closing_time']]
df_noise_veh.head()
```

Out[55]:

	Request_closing_time
87	6.374722
156	2.015000
172	3.144167
221	3.038056
319	0.726389

In [56]:
```
df_type_res = df_perfect.loc[:, ['Complaint Type','Resolution_Time']]
df_type_res.head()
```


olumns

Out[56]:
```
Index(['Complaint Type', 'Resolution_Time'], dtype='object')
```

In [57]:
```
fvalue, pvalue = stats.f_oneway(df_dis_youth, df_noise_veh)
pvalue
```

Out[57]:
```
array([0.91269878])
```

In [ ]:
```
Null Hypothesis to be accepted for Disorderly Youth and Noise - Vehicle p-value close to
```

In [ ]:
```

```

In [ ]:
```
One Way Anova for Posting Advertisement and Derelict Vehicle
```

In [59]:
```
df_post_ad = df_perfect[df_perfect['Complaint Type']=='Posting Advertisement']
df_post_ad = df_post_ad.loc[:,['Resolution_Time']]
df_post_ad.head()
```

Out[59]:

	Resolution_Time
20	2.110000
42	2.151389
46	2.176111
49	2.233889
51	2.260278

In [60]:
```
df_der_veh = df_perfect[df_perfect['Complaint Type']=='Derelict Vehicle']
df_der_veh = df_der_veh.loc[:,['Resolution_Time']]
df_der_veh.head()
```

Out[60]:

	Resolution_Time
14	10.489722
151	3.950278
255	1.364722
256	4.133056
285	0.753333

In [61]:
```
fvalue, pvalue = stats.f_oneway(df_post_ad, df_der_veh)
pvalue
```

Out[61]:
```
array([7.28776953e-35])
```

In [ ]:
```
Null Hypothesis for Posting Advertisement and Derelict Vehicle to be rejected p-value <
```

In [ ]:
```

```

In [ ]:
```
Anova table for complain type and resolution time
```

In [67]:
```
df_perfect['Complaint_Type']=df_perfect['Complaint Type']
df_type_res = df_perfect.loc[:, ['Complaint_Type','Request_closing_time']]
model = ols('Request_closing_time ~ Complaint_Type', data=df_type_res).fit()
anova_table = sm.stats.anova_lm(model, typ=2)
anova_table
```

Out[67]:

	sum_sq	df	F	PR(>F)
Complaint_Type	2.920400e+05	20.0	410.258598	0.0
Residual	1.060814e+07	298047.0	NaN	NaN

In [ ]:
```
Null Hypothesis to be rejected since p-value < 0.05. Thus we accept alternate hypothesis
```

In [ ]:
```

```

In [ ]:
```
Crosstab and Chi Square test for Location and Complaint type
```

In [68]:
```
df_city_type = pd.crosstab(df_perfect.City , df_perfect.Complaint_Type)
```

In [69]:
```
from scipy.stats import chi2_contingency
from scipy.stats import chi2
table = df_city_type
stat, p, dof, expected = chi2_contingency(table)
print('dof=%d' % dof)
print(expected)
prob = 0.95
critical = chi2.ppf(prob, dof)
print('probability=%.3f, critical=%.3f, stat=%.3f' % (prob, critical, stat))
if abs(stat) >= critical:
    print('Dependent (reject H0)')
else:
    print('Independent (fail to reject H0)')
alpha = 1.0 - prob
print('significance=%.3f, p=%.3f' % (alpha, p))
if p <= alpha:
    print('Dependent (reject H0)')
else:
    print('Independent (fail to reject H0)')
```

```
dof=1040
[[5.73350737e+00 3.11515400e-01 5.66574169e+01 ... 3.31741755e+00
  4.37807385e-01 2.80068584e+00]
 [1.64968644e+02 8.96314763e+00 1.63018841e+03 ... 9.54511504e+01
  1.25738943e+01 8.05833700e+01]
 [1.86599603e+01 1.01384103e+00 1.84394139e+02 ... 1.07966862e+01
  1.42226840e+00 9.11495938e+00]
 ...
 [6.41892211e+01 3.48755650e+00 6.34385536e+02 ... 3.71399974e+01
  4.89249632e+00 3.13549511e+01]
 [9.23615914e+01 5.01822989e+00 9.12699488e+02 ... 5.34485809e+01
  7.03979170e+00 4.51165829e+01]
 [3.12736765e+00 1.69917491e-01 3.09040456e+01 ... 1.80950848e+00
  2.38367665e-01 1.52764682e+00]]
probability=0.950, critical=1116.137, stat=110425.867
Dependent (reject H0)
significance=0.050, p=0.000
Dependent (reject H0)
```

In [ ]:
```
Since p value for the chi square test is less than 0.05(LOS) we can conclude that Compla
i.e specific type of complaint is raised from specific places.
```

In [ ]:
```
FINISH
```