

```
In [1]: import pandas as pd
```

```
In [2]: data = pd.read_csv('311_Service_Requests_from_2010_to_Present.csv')
```

```
C:\Users\mha0125\Anaconda3\lib\site-packages\IPython\core\interactiveshell.py:3146: DtypeWarning: Columns (48,49) have mixed types.Specify dtype option on import or set low_memory=False.
  has_raised = await self.run_ast_nodes(code_ast.body, cell_name,
```

```
In [3]: data
```

1	32309934	12/31/2015 23:59	1/1/2016 1:26	NYPD	New York City Police Department	Blocked Driveway	No Access	Street/Sidewalk
2	32309159	12/31/2015 23:59	1/1/2016 4:51	NYPD	New York City Police Department	Blocked Driveway	No Access	Street/Sidewalk
3	32305098	12/31/2015 23:57	1/1/2016 7:43	NYPD	New York City Police Department	Illegal Parking	Commercial Overnight Parking	Street/Sidewalk
4	32306529	12/31/2015 23:56	1/1/2016 3:24	NYPD	New York City Police Department	Illegal Parking	Blocked Sidewalk	Street/Sidewalk
...
300693	30281872	3/29/2015 0:33	NaN	NYPD	New York City Police Department	Noise - Commercial	Loud Music/Party	Club/Bar/Restaurant
300694	30281230	3/29/2015 0:33	3/29/2015 2:33	NYPD	New York City Police Department	Blocked Driveway	Partial Access	Street/Sidewalk
300695	30283424	3/29/2015 0:33	3/29/2015 3:40	NYPD	New York City Police Department	Noise - Commercial	Loud Music/Party	Club/Bar/Restaurant
300696	30280004	3/29/2015 0:33	3/29/2015 4:38	NYPD	New York City Police Department	Noise - Commercial	Loud Music/Party	Club/Bar/Restaurant
300697	30281825	3/29/2015 0:33	3/29/2015 4:41	NYPD	New York City Police Department	Noise - Commercial	Loud Music/Party	Store/Commercial

300698 rows x 53 columns

```
data.tail()
```

	Unique Key	Created Date	Closed Date	Agency	Agency Name	Complaint Type	Descriptor	Location Type	Incident Type
300693	30281872	3/29/2015 0:33	NaN	NYPD	New York City Police Department	Noise - Commercial	Loud Music/Party	Club/Bar/Restaurant	
300694	30281230	3/29/2015 0:33	3/29/2015 2:33	NYPD	New York City Police Department	Blocked Driveway	Partial Access	Street/Sidewalk	11
300695	30283424	3/29/2015 0:33	3/29/2015 3:40	NYPD	New York City Police Department	Noise - Commercial	Loud Music/Party	Club/Bar/Restaurant	11
300696	30280004	3/29/2015 0:33	3/29/2015 4:38	NYPD	New York City Police Department	Noise - Commercial	Loud Music/Party	Club/Bar/Restaurant	10

300698 rows x 53 columns

```
In [4]: data.tail()
```

	Unique Key	Created Date	Closed Date	Agency	Agency Name	Complaint Type	Descriptor	Location Type	Incident Zip
	300693	30281872	3/29/2015 0:33	NaN	NYPD	New York City Police Department	Noise - Commercial	Loud Music/Party	Club/Bar/Restaurant
	300694	30281230	3/29/2015 0:33	3/29/2015 2:33	NYPD	New York City Police Department	Blocked Driveway	Partial Access	Street/Sidewalk
	300695	30283424	3/29/2015 0:33	3/29/2015 3:40	NYPD	New York City Police Department	Noise - Commercial	Loud Music/Party	Club/Bar/Restaurant
	300696	30280004	3/29/2015 0:33	3/29/2015 4:38	NYPD	New York City Police Department	Noise - Commercial	Loud Music/Party	Club/Bar/Restaurant
	300697	30281825	3/29/2015 0:33	3/29/2015 4:41	NYPD	New York City Police Department	Noise - Commercial	Loud Music/Party	Store/Commercial

5 rows x 53 columns

```
In [5]: import datetime as dt
```

```
In [7]: nyc311 = pd.read_csv('311_Service_Requests_from_2010_to_Present.csv', parse_dates=['Closed Date', 'Created Date'])
```

```
C:\Users\mha0125\Anaconda3\lib\site-packages\IPython\core\interactiveshell.py:3146: DtypeWarning: Columns (48,49) have mixed types.Specify dtype option on import or set low_memory=False.
  has_raised = await self.run_ast_nodes(code_ast.body, cell_name,
```

```
In [8]: nyc311
```

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

	300693	30281872	2015-03-29 00:33:00	NaT	NYPD	New York City Police Department	Noise - Commercial	Loud Music/Party	Club/Bar/Restaurant
	300694	30281230	2015-03-29 00:33:00	2015-03-29 02:33:00	NYPD	New York City Police Department	Blocked Driveway	Partial Access	Street/Sidewalk
	300695	30283424	2015-03-29 00:33:00	2015-03-29 03:40:00	NYPD	New York City Police Department	Noise - Commercial	Loud Music/Party	Club/Bar/Restaurant
	300696	30280004	2015-03-29 00:33:00	2015-03-29 04:38:00	NYPD	New York City Police Department	Noise - Commercial	Loud Music/Party	Club/Bar/Restaurant
	300697	30281825	2015-03-29 00:33:00	2015-03-29 04:41:00	NYPD	New York City Police Department	Noise - Commercial	Loud Music/Party	Store/Commercial

300698 rows x 53 columns

```
In [9]: completion_time = nyc311['Closed Date'] - nyc311['Created Date']
```

```
In [10]: type(completion_time)
```

```
Out[10]: pandas.core.series.Series
```

```
In [11]: nyc311['Completion Time'] = completion_time
```

```
In [12]: nyc311
```

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

	300693	30281872	2015-03-29 00:33:00	NaT	NYPD	New York City Police Department	Noise - Commercial	Loud Music/Party	Club/Bar/Restaurant
	300694	30281230	2015-03-29 00:33:00	2015-03-29 02:33:00	NYPD	New York City Police Department	Blocked Driveway	Partial Access	Street/Sidewalk
	300695	30283424	2015-03-29 00:33:00	2015-03-29 03:40:00	NYPD	New York City Police Department	Noise - Commercial	Loud Music/Party	Club/Bar/Restaurant
	300696	30280004	2015-03-29 00:33:00	2015-03-29 04:38:00	NYPD	New York City Police Department	Noise - Commercial	Loud Music/Party	Club/Bar/Restaurant
	300697	30281825	2015-03-29 00:33:00	2015-03-29 04:41:00	NYPD	New York City Police Department	Noise - Commercial	Loud Music/Party	Store/Commercial

300698 rows x 54 columns

```
In [13]: nyc311.columns
```

```
Out[13]: 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', 'Completion Time'],
      dtype='object')
```

```
In [14]: nyc311.count()
```

Unique Key	300698
Created Date	300698
Closed Date	298534
Agency	300698
Agency Name	300698
Complaint Type	300698
Descriptor	294784
Location Type	300567
Incident Zip	298083
Incident Address	256288
Street Name	256288
Cross Street 1	251419
Cross Street 2	250919
Intersection Street 1	43858
Intersection Street 2	43362
Address Type	297883
City	298084
Landmark	349
Facility Type	298527
Status	300698
Due Date	300695
Resolution Description	300698
Resolution Action Updated Date	298511
Community Board	300698
Borough	300698
X Coordinate (State Plane)	297158
Y Coordinate (State Plane)	297158
Park Facility Name	300698
Park Borough	300698
School Name	300698
School Number	300698
School Region	300697
School Code	300697
School Phone Number	300698
School Address	300698
School City	300698
School State	300698
School Zip	300697
School Not Found	300698
School or Citywide Complaint	0
Vehicle Type	0
Taxi Company Borough	0
Taxi Pick Up Location	0
Bridge Highway Name	243
Bridge Highway Direction	243
Road Ramp	213
Bridge Highway Segment	213
Garage Lot Name	0
Ferry Direction	1
Ferry Terminal Name	2
Latitude	297158
Longitude	297158
Location	297158
Completion Time	298534
dtype: int64	

```
In [15]: nyc311.index
```

```
Out[15]: RangeIndex(start=0, stop=300698, step=1)
```

```
In [16]: nyc311.describe()
```

	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
count	3.006980e+05	298083.000000	2.971580e+05	2.97158.000000	0.0	0.0	0.0	0.0	
mean	3.130054e+07	10848.886645	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 [17]: import matplotlib.pyplot as plt
```

```
In [18]: group_by_borough = nyc311.groupby('Borough')
```

```
In [19]: group_by_borough.count()
```

	Unique Key	Created Date	Closed Date	Agency	Agency Name	Complaint Type	Descriptor	Location Type	Incident Zip	Incident Address
Borough										
BRONX	40702	40702	40697	40702	40702	40702	40342	40688	40701	35317
BROOKLYN	98307	98307	98299	98307	98307	98307	97093	98263	98307	86261
MANHATTAN	66131	66131	66109	66131	66131	66131	62644	66116	66095	52162
QUEENS	80641	80641	80629	80641	80641	80641	79909	80598	80639	69470
STATEN ISLAND	12343	12343	12338	12343	12343	12343	12229	12337	12341	10513
Unspecified	2574	2574	466	2574	2574	2574	2567	2565	0	2565

6 rows x 53 columns

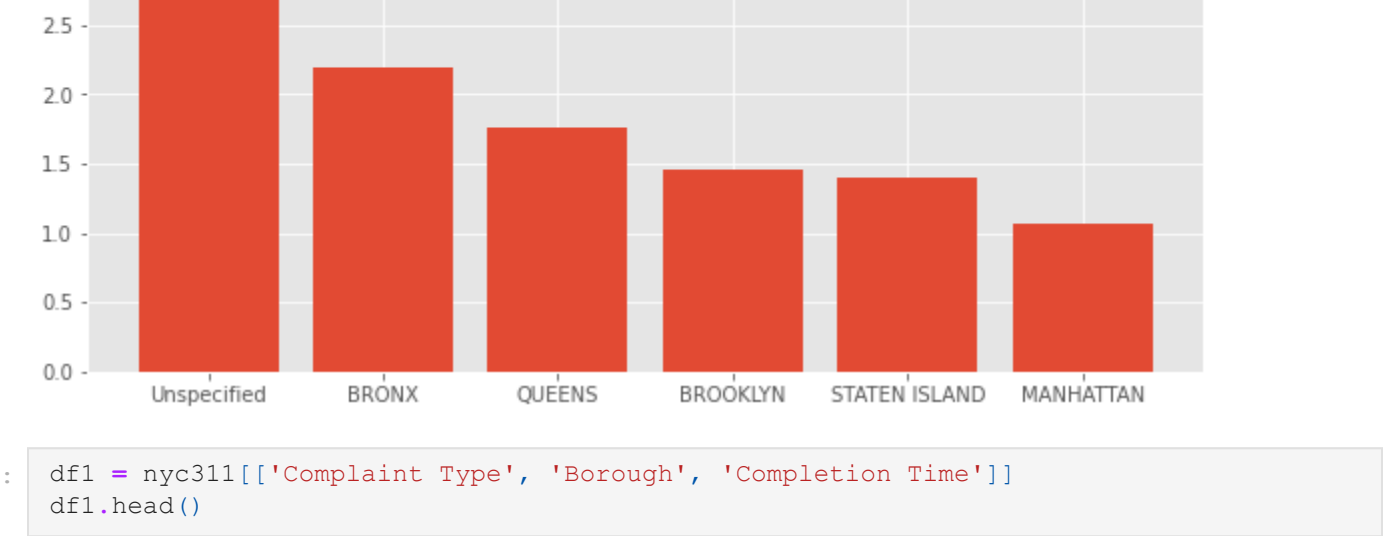
```
In [20]: import numpy as np
```

```
In [21]: borough_size = group_by_borough.size()
```

```
In [22]: borough_size.sort_values(inplace=True, ascending=False)
```

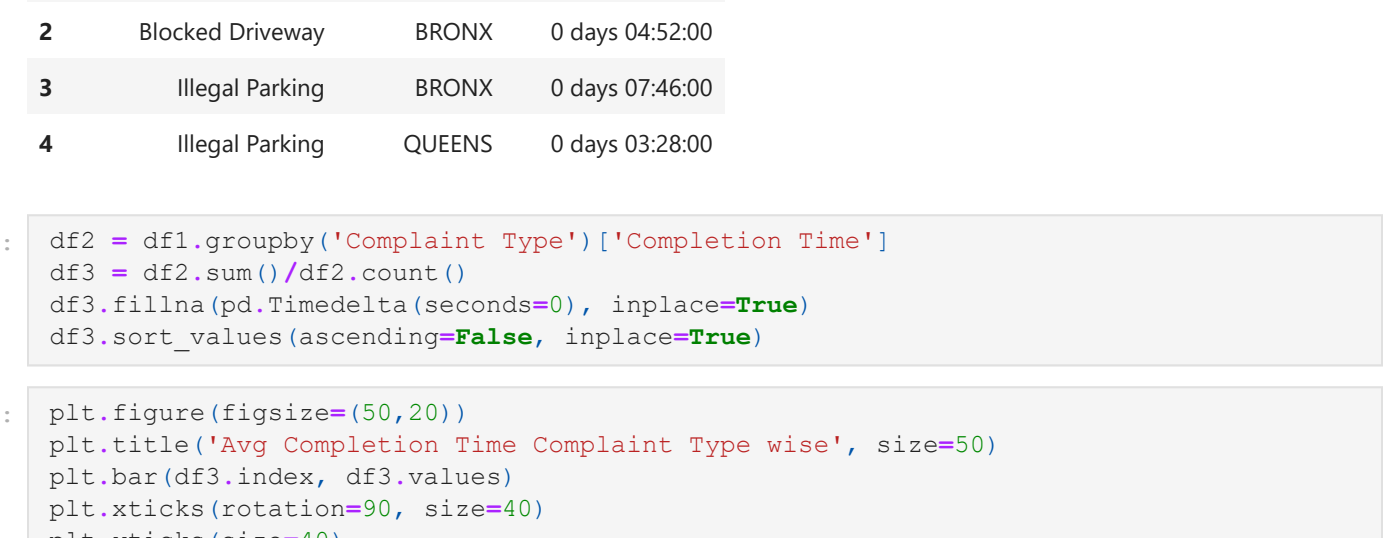
```
In [23]: borough_index = borough_size.index
```

```
In [24]: borough_values = borough_size.values
```



```
In [26]: groupby_complainttype = nyc311.groupby('Complaint Type')
```

```
In [27]: complainttype_size = groupby_complainttype.size().sort_values(ascending=False)
```



```
In [29]: groupby_locationtype = nyc311.groupby('Location Type')
```

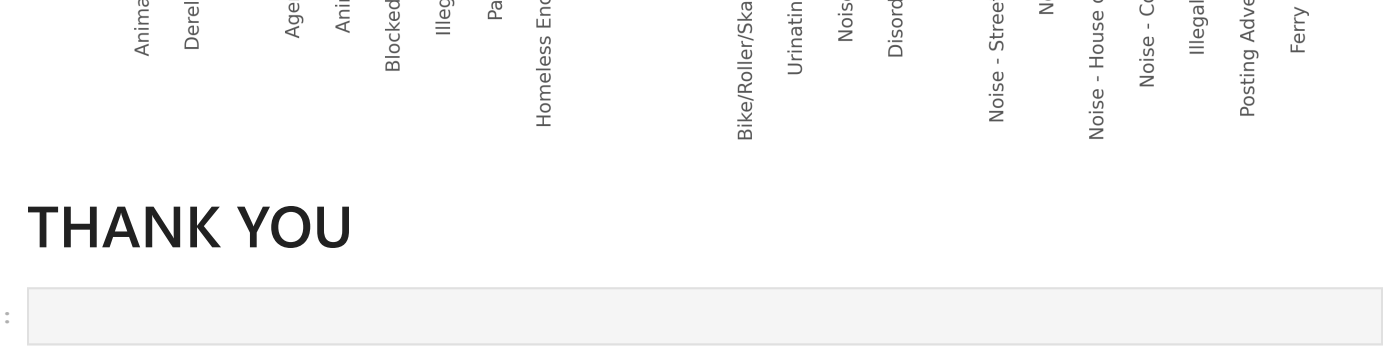
```
In [30]: location_type = groupby_locationtype.size().sort_values(ascending=False).head()
```

```
location_type.count()
```

```
Out[30]: 5
```

```
In [31]: explode = [0.1,0.1,0.1,0.1,0.1]
```

```
In [32]: plt.figure(figsize=(30,10))
plt.pie(location_type.values, labels=location_type.index, explode=explode, autopct='%')
plt.title('Complaint Count by Location Type', size=10)
plt.show()
```



```
In [33]: avg_comptime = group_by_borough['Completion Time'].sum()/group_by_borough['Completion Time'].count()
avg_comptime.sort_values(ascending=False, inplace=True)
avg_comptime
```

Borough	Unspecified	0 days 10:23:35.150214592
BRONX	0 days 06:05:46.538074059	
QUEENS	0 days 04:52:16.440486673	
BROOKLYN	0 days 04:02:52.260033572	
STATEN ISLAND	0 days 03:52:46.432160804	
MANHATTAN	0 days 02:58:22.599040977	
Name: Completion Time, dtype: timedelta64[ns]		

```
In [34]: plt.figure(figsize=(10,5))
plt.title('Avg Completion Time Boroughwise')
plt.bar(avg_comptime.index, avg_comptime.values)
plt.show()
```



```
In [35]: df1 = nyc311[['Complaint Type', 'Borough', 'Completion Time']]
df1.head()
```

	Complaint Type	Borough	Completion Time
0	Noise - Street/Sidewalk	MANHATTAN	0 days 00:56:00
1	Blocked Driveway	QUEENS	0 days 01:27:00
2	Blocked Driveway	BRONX	0 days 04:52:00
3	Illegal Parking	BRONX	0 days 07:46:00
4	Illegal Parking	QUEENS	0 days 03:28:00

```
In [36]: df2 = df1.groupby('Complaint Type')['Completion Time']
df3 = df2.sum()/df2.count()
df3.fillna(pd.Timedelta(seconds=0), inplace=True)
df3.sort_values(ascending=False, inplace=True)
```

```
In [37]: plt.figure(figsize=(50,20))
plt.title('Avg Completion Time Complaint Type wise', size=50)
plt.bar(df3.index, df3.values)
plt.xticks(rotation=90, size=40)
plt.yticks(size=40)
plt.show()
```

