In [1]:
```
import pandas as pd
import numpy as np
import matplotlib.pyplot as plt
%matplotlib inline
```

In [2]:
```
import seaborn as sns
```

In [3]:
```
pwd
```

Out(3):
```
'C:\Users\HP'
```

In [4]:
```
df = pd.read_csv("C:\Users\HP\Desktop\project\311_Service_Requests_from_2010_to_Present.csv")
```

```
C:\ANACONDA\lib\site-packages\IPython\core\interactiveshell.py:3165: 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 [5]:
```
df.head()
```

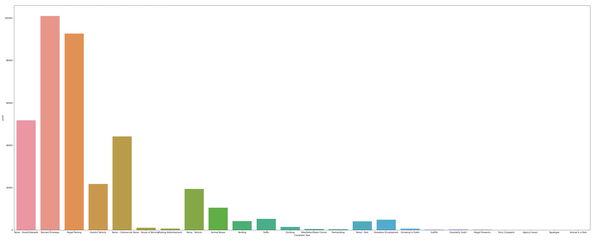
Out[5]:

	Unique Key	Created Date	Closed Date	Agency	Agency Name	Complaint Type	Descriptor	Location Type	Incident Zip	Incident Address	...	Bridge Highway Name	Bridge Highway Direction	Road Ramp	Bridge Highway Segment	Garage Lot Name	Ferry Direction	Ferry Terminal Name	Latitude	Longitude
0	32310363	12/31/2015 11:59:45 PM	01/01/2016 12:55:15 AM	NYPD	New York City Police Department	Noise - Street/Sidewalk	Loud Music/Party	Street/Sidewalk	10034.0	71 VERMILYEA AVENUE	...	NaN	NaN	NaN	NaN	NaN	NaN	NaN	40.865682	-73.923501
1	32309934	12/31/2015 11:59:44 PM	01/01/2016 01:26:57 AM	NYPD	New York City Police Department	Blocked Driveway	No Access	Street/Sidewalk	11105.0	27-07 23 AVENUE	...	NaN	NaN	NaN	NaN	NaN	NaN	NaN	40.775945	-73.915094
2	32309159	12/31/2015 11:59:29 PM	01/01/2016 04:51:03 AM	NYPD	New York City Police Department	Blocked Driveway	No Access	Street/Sidewalk	10458.0	2897 VALENTINE AVENUE	...	NaN	NaN	NaN	NaN	NaN	NaN	NaN	40.870325	-73.888525
3	32305098	12/31/2015 11:57:46 PM	01/01/2016 07:43:13 AM	NYPD	New York City Police Department	Illegal Parking	Commercial Overnight Parking	Street/Sidewalk	10461.0	2940 BAISLEY AVENUE	...	NaN	NaN	NaN	NaN	NaN	NaN	NaN	40.835994	-73.828379
4	32306529	12/31/2015 11:56:58 PM	01/01/2016 03:24:42 AM	NYPD	New York City Police Department	Illegal Parking	Blocked Sidewalk	Street/Sidewalk	11373.0	87-14 57 ROAD	...	NaN	NaN	NaN	NaN	NaN	NaN	NaN	40.733060	-73.874170

5 rows × 53 columns

In [6]:
```
plt.figure(figsize=(50,20))
sns.countplot(x= "Complaint Type", data=df)
```

Out[6]:
```
<AxesSubplot:xlabel='Complaint Type', ylabel='count'>
```



In [7]:
```
import datetime
```

In [11]:
```
df['Created Date'] = pd.to_datetime(df['Created Date'])
df['Created Date']
```

Out[11]:
```
0        2015-12-31 23:59:45
1        2015-12-31 23:59:44
2        2015-12-31 23:59:29
3        2015-12-31 23:57:46
4        2015-12-31 23:56:58
                   ...
364553   2015-01-01 00:04:44
364554   2015-01-01 00:04:28
364555   2015-01-01 00:01:30
364556   2015-01-01 00:01:29
364557   2015-01-01 00:00:50
Name: Created Date, Length: 364558, dtype: datetime64[ns]
```

In [14]:
```
df2=df[['Created Date']]
df2
```

Out[14]:

	Created Date
0	2015-12-31 23:59:45
1	2015-12-31 23:59:44
2	2015-12-31 23:59:29
3	2015-12-31 23:57:46
4	2015-12-31 23:56:58
...	...
364553	2015-01-01 00:04:44
364554	2015-01-01 00:04:28
364555	2015-01-01 00:01:30
364556	2015-01-01 00:01:29
364557	2015-01-01 00:00:50

364558 rows × 1 columns

In [15]:
```
df2['hour'] = df2['Created Date'].dt.hour
df2['minute'] = df2['Created Date'].dt.minute
df2['second'] = df2['Created Date'].dt.second
df2
```

```
<ipython-input-15-fabc94044f2db>:1: SettingWithCopyWarning:
A value is trying to be set on a copy of a slice from a DataFrame.
Try using .loc[row_indexer,col_indexer] = value instead

See the caveats in the documentation: https://pandas.pydata.org/pandas-docs/stable/user_guide/indexing.html#returning-a-view-versus-a-copy
  df2['hour'] = df2['Created Date'].dt.hour
<ipython-input-15-fabc94044f2db>:2: SettingWithCopyWarning:
A value is trying to be set on a copy of a slice from a DataFrame.
Try using .loc[row_indexer,col_indexer] = value instead

See the caveats in the documentation: https://pandas.pydata.org/pandas-docs/stable/user_guide/indexing.html#returning-a-view-versus-a-copy
  df2['minute'] = df2['Created Date'].dt.minute
<ipython-input-15-fabc94044f2db>:3: SettingWithCopyWarning:
A value is trying to be set on a copy of a slice from a DataFrame.
Try using .loc[row_indexer,col_indexer] = value instead

See the caveats in the documentation: https://pandas.pydata.org/pandas-docs/stable/user_guide/indexing.html#returning-a-view-versus-a-copy
  df2['second'] = df2['Created Date'].dt.second
```

Out[15]:

	Created Date	hour	minute	second
0	2015-12-31 23:59:45	23	59	45
1	2015-12-31 23:59:44	23	59	44
2	2015-12-31 23:59:29	23	59	29
3	2015-12-31 23:57:46	23	57	46
4	2015-12-31 23:56:58	23	56	58
...	...	...	...	...
364553	2015-01-01 00:04:44	0	4	44
364554	2015-01-01 00:04:28	0	4	28
364555	2015-01-01 00:01:30	0	1	30
364556	2015-01-01 00:01:29	0	1	29
364557	2015-01-01 00:00:50	0	0	50

364558 rows × 4 columns

In [25]:
```
df['Closed Date'] = pd.to_datetime(df['Closed Date'])
df['Closed Date']
```

Out[25]:
```
0        2016-01-01 00:55:15
1        2016-01-01 01:26:57
2        2016-01-01 04:51:03
3        2016-01-01 07:43:13
4        2016-01-01 03:24:42
                   ...
364553   2015-01-01 10:22:31
364554   2015-01-01 02:25:02
364555   2015-01-01 00:20:33
364556   2015-01-01 02:42:22
364557   2015-01-01 02:47:50
Name: Closed Date, Length: 364558, dtype: datetime64[ns]
```

In [26]:
```
df3=df[['Closed Date']]
df3
```

Out[26]:

	Closed Date
0	2016-01-01 00:55:15
1	2016-01-01 01:26:57
2	2016-01-01 04:51:03
3	2016-01-01 07:43:13
4	2016-01-01 03:24:42
...	...
364553	2015-01-01 10:22:31
364554	2015-01-01 02:25:02
364555	2015-01-01 00:20:33
364556	2015-01-01 02:42:22
364557	2015-01-01 02:47:50

364558 rows × 1 columns

In [27]:
```
df3['hour'] = df3['Closed Date'].dt.hour
df3['minute'] = df3['Closed Date'].dt.minute
df3['second'] = df3['Closed Date'].dt.second
df3
```

```
<ipython-input-27-d9b1c96e87c4>:1: SettingWithCopyWarning:
A value is trying to be set on a copy of a slice from a DataFrame.
Try using .loc[row_indexer,col_indexer] = value instead

See the caveats in the documentation: https://pandas.pydata.org/pandas-docs/stable/user_guide/indexing.html#returning-a-view-versus-a-copy
  df3['hour'] = df3['Closed Date'].dt.hour
<ipython-input-27-d9b1c96e87c4>:2: SettingWithCopyWarning:
A value is trying to be set on a copy of a slice from a DataFrame.
Try using .loc[row_indexer,col_indexer] = value instead

See the caveats in the documentation: https://pandas.pydata.org/pandas-docs/stable/user_guide/indexing.html#returning-a-view-versus-a-copy
  df3['minute'] = df3['Closed Date'].dt.minute
<ipython-input-27-d9b1c96e87c4>:3: SettingWithCopyWarning:
A value is trying to be set on a copy of a slice from a DataFrame.
Try using .loc[row_indexer,col_indexer] = value instead

See the caveats in the documentation: https://pandas.pydata.org/pandas-docs/stable/user_guide/indexing.html#returning-a-view-versus-a-copy
  df3['second'] = df3['Closed Date'].dt.second
```

Out[27]:

	Closed Date	hour	minute	second
0	2016-01-01 00:55:15	0.0	55.0	15.0
1	2016-01-01 01:26:57	1.0	26.0	57.0
2	2016-01-01 04:51:03	4.0	51.0	3.0
3	2016-01-01 07:43:13	7.0	43.0	13.0
4	2016-01-01 03:24:42	3.0	24.0	42.0
...	...	...	...	...
364553	2015-01-01 10:22:31	10.0	22.0	31.0
364554	2015-01-01 02:25:02	2.0	25.0	2.0
364555	2015-01-01 00:20:33	0.0	20.0	33.0
364556	2015-01-01 02:42:22	2.0	42.0	22.0
364557	2015-01-01 02:47:50	2.0	47.0	50.0

364558 rows × 4 columns

In [34]:
```
df4=df[['Created Date','Closed Date']]
df4
```

Out[34]:

	Created Date	Closed Date
0	2015-12-31 23:59:45	2016-01-01 00:55:15
1	2015-12-31 23:59:44	2016-01-01 01:26:57
2	2015-12-31 23:59:29	2016-01-01 04:51:03
3	2015-12-31 23:57:46	2016-01-01 07:43:13
4	2015-12-31 23:56:58	2016-01-01 03:24:42
...	...	...
364553	2015-01-01 00:04:44	2015-01-01 10:22:31
364554	2015-01-01 00:04:28	2015-01-01 02:25:02
364555	2015-01-01 00:01:30	2015-01-01 00:20:33
364556	2015-01-01 00:01:29	2015-01-01 02:42:22
364557	2015-01-01 00:00:50	2015-01-01 02:47:50

364558 rows × 2 columns

In [36]:
```
df4['hourdiff']= df3['hour']-df2['hour']
df4['minutediff']= df3['minute']-df2['minute']
df4['seconddiff']= df3['second']-df2['second']
df4
```

```
<ipython-input-36-000e5a3818c4>:1: SettingWithCopyWarning:
A value is trying to be set on a copy of a slice from a DataFrame.
Try using .loc[row_indexer,col_indexer] = value instead

See the caveats in the documentation: https://pandas.pydata.org/pandas-docs/stable/user_guide/indexing.html#returning-a-view-versus-a-copy
  df4['hourdiff']= df3['hour']-df2['hour']
<ipython-input-36-000e5a3818c4>:2: SettingWithCopyWarning:
A value is trying to be set on a copy of a slice from a DataFrame.
Try using .loc[row_indexer,col_indexer] = value instead

See the caveats in the documentation: https://pandas.pydata.org/pandas-docs/stable/user_guide/indexing.html#returning-a-view-versus-a-copy
  df4['minutediff']= df3['minute']-df2['minute']
<ipython-input-36-000e5a3818c4>:3: SettingWithCopyWarning:
A value is trying to be set on a copy of a slice from a DataFrame.
Try using .loc[row_indexer,col_indexer] = value instead

See the caveats in the documentation: https://pandas.pydata.org/pandas-docs/stable/user_guide/indexing.html#returning-a-view-versus-a-copy
  df4['seconddiff']= df3['second']-df2['second']
```

Out[36]:

	Created Date	Closed Date	hourdiff	minutediff	seconddiff
0	2015-12-31 23:59:45	2016-01-01 00:55:15	-23.0	-4.0	30.0
1	2015-12-31 23:59:44	2016-01-01 01:26:57	-22.0	-33.0	13.0
2	2015-12-31 23:59:29	2016-01-01 04:51:03	-19.0	-8.0	-26.0
3	2015-12-31 23:57:46	2016-01-01 07:43:13	-16.0	-14.0	-33.0
4	2015-12-31 23:56:58	2016-01-01 03:24:42	-20.0	-32.0	-16.0
...	...	...	...	...	...
364553	2015-01-01 00:04:44	2015-01-01 10:22:31	10.0	18.0	-13.0
364554	2015-01-01 00:04:28	2015-01-01 02:25:02	2.0	21.0	-26.0
364555	2015-01-01 00:01:30	2015-01-01 00:20:33	0.0	19.0	3.0
364556	2015-01-01 00:01:29	2015-01-01 02:42:22	2.0	41.0	-7.0
364557	2015-01-01 00:00:50	2015-01-01 02:47:50	2.0	47.0	0.0

364558 rows × 5 columns

In [39]:
```
df4.iloc[:,2:5].apply(np.mean)
```

Out[39]:
```
hourdiff     -1.201840
minutediff   -0.290783
seconddiff    0.035400
dtype: float64
```

In [ ]:

In [ ]:

In [ ]:

In [ ]:

In [ ]:

In [ ]: