```
In [1]: sc
```

Out[1]: **SparkContext**

Spark UI (http://DESKTOP-2BUG72N:4041)
Version
 v2.4.1
Master
 local[2]
AppName
 PySparkShell

```
In [67]: from pyspark import SparkFiles
         url="https://ahmaditustorage.blob.core.windows.net/datasciencetask/crimes_sample.
         spark.sparkContext.addFile(url)

         df=spark.read.csv(SparkFiles.get("crimes_sample.csv"), header=True)
```

```
In [69]: df
```

Out[69]: DataFrame[ID: string, Case Number: string, Date: string, Block: string, IUCR: s
         tring, Primary Type: string, Description: string, Location Description: string,
         Arrest: string, Domestic: string, Beat: string, District: string, Ward: string,
         Community Area: string, FBI Code: string, X Coordinate: string, Y Coordinate: s
         tring, Year: string, Updated On: string, Latitude: string, Longitude: string, L
         ocation: string]

```
In [12]: type(df)
```

Out[12]: pyspark.sql.dataframe.DataFrame

```
In [70]: df2=df.filter(df["Year"] == "2007")
```

```
In [71]: df3=df2.select("Primary Type","Year")
```

```
In [72]: from pyspark.sql.functions import count
```

```
In [73]: df4=df3.groupBy("Primary Type").count()
```

```
In [74]: import pyodbc
```

In [75]:
```python
server = 'cloudahmaddb.database.windows.net'
database = 'SampleCrimeDB'
username = 'ahmad'
password = '******'
driver= '{ODBC Driver 17 for SQL Server}'
cnxn = pyodbc.connect('DRIVER='+driver+';SERVER='+server+';PORT=1433;DATABASE='+d
cursor = cnxn.cursor()
cursor.execute("select * from CrimeCount")
row = cursor.fetchone()
print("executed")
while row:
    print (str(row[0]) + " " + str(row[1]))
    row = cursor.fetchone()
```

executed

In [76]:
```python
for row in df4.collect():
    query="insert into CrimeCount(Crime,Count) values ('{}',{})"
    query= query.format(row[0],row[1])
    print(query)
    print(row[0],row[1])
    cursor.execute(query)
```

```
insert into CrimeCount(Crime,Count) values ('OFFENSE INVOLVING CHILDREN',42)
OFFENSE INVOLVING CHILDREN 42
insert into CrimeCount(Crime,Count) values ('STALKING',7)
STALKING 7
insert into CrimeCount(Crime,Count) values ('PUBLIC PEACE VIOLATION',34)
PUBLIC PEACE VIOLATION 34
insert into CrimeCount(Crime,Count) values ('ARSON',4)
ARSON 4
insert into CrimeCount(Crime,Count) values ('GAMBLING',7)
GAMBLING 7
insert into CrimeCount(Crime,Count) values ('CRIMINAL TRESPASS',119)
CRIMINAL TRESPASS 119
insert into CrimeCount(Crime,Count) values ('ASSAULT',225)
ASSAULT 225
insert into CrimeCount(Crime,Count) values ('LIQUOR LAW VIOLATION',11)
LIQUOR LAW VIOLATION 11
insert into CrimeCount(Crime,Count) values ('MOTOR VEHICLE THEFT',186)
MOTOR VEHICLE THEFT 186
insert into CrimeCount(Crime,Count) values ('THEFT',918)
THEFT 918
insert into CrimeCount(Crime,Count) values ('BATTERY',793)
BATTERY 793
insert into CrimeCount(Crime,Count) values ('ROBBERY',142)
ROBBERY 142
insert into CrimeCount(Crime,Count) values ('CRIM SEXUAL ASSAULT',14)
CRIM SEXUAL ASSAULT 14
insert into CrimeCount(Crime,Count) values ('PROSTITUTION',94)
PROSTITUTION 94
insert into CrimeCount(Crime,Count) values ('DECEPTIVE PRACTICE',215)
DECEPTIVE PRACTICE 215
insert into CrimeCount(Crime,Count) values ('SEX OFFENSE',9)
SEX OFFENSE 9
insert into CrimeCount(Crime,Count) values ('CRIMINAL DAMAGE',522)
CRIMINAL DAMAGE 522
insert into CrimeCount(Crime,Count) values ('NARCOTICS',732)
NARCOTICS 732
insert into CrimeCount(Crime,Count) values ('OTHER OFFENSE',282)
OTHER OFFENSE 282
insert into CrimeCount(Crime,Count) values ('KIDNAPPING',1)
KIDNAPPING 1
insert into CrimeCount(Crime,Count) values ('BURGLARY',230)
BURGLARY 230
insert into CrimeCount(Crime,Count) values ('WEAPONS VIOLATION',42)
WEAPONS VIOLATION 42
insert into CrimeCount(Crime,Count) values ('INTERFERENCE WITH PUBLIC OFFICER',8)
INTERFERENCE WITH PUBLIC OFFICER 8
```

In [77]: `cnxn.commit()`

In [ ]: