**Task 2- Crime Data Analysis**

7.

from pyspark.sql.functions import year, to\_timestamp, to\_date, col

# Convert 'REPORTDATETIME' to timestamp format

data = data.withColumn("Date", to\_date(col("REPORTDATETIME"), "MM-dd-yyyy").alias("to\_date"))

# Extract the year from 'DATE'

data = data.withColumn("Year", year("Date"))

# Filter records for the year 2013

offenses\_2013 = data.filter(data["Year"] == 2013)

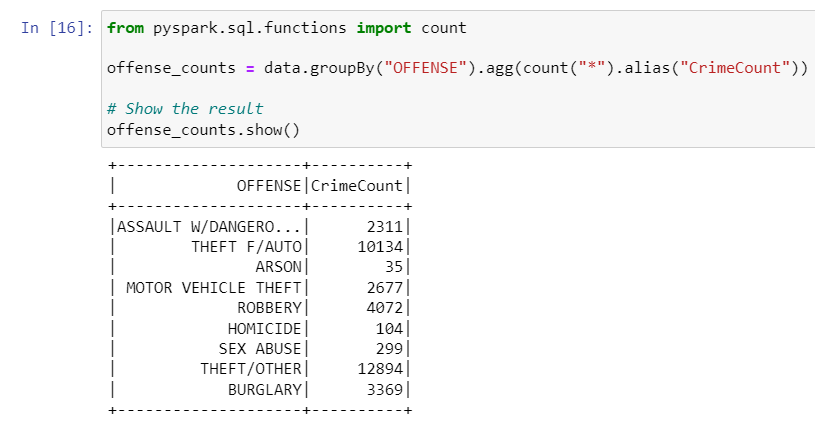
# Count the number of offenses in 2013

num\_offenses\_2013 = offenses\_2013.count()

# Print the result

print(f"Number of offenses in 2013: {num\_offenses\_2013}")

8.



9.

A computer screen shot of a person

Description automatically generated

10.

A screenshot of a computer code

Description automatically generated