Zachary Tankersley

Southern New Hampshire University

DAT-375-T2962

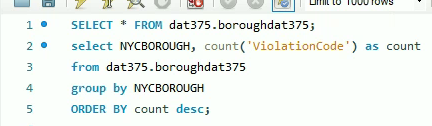
November 15, 2022

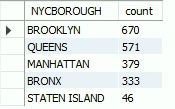
**Module Three Assignment**

For this report, I have been tasked with finding the three New York City boroughs with the highest accident rates, regardless of accident type, to calculate insurance rate changes for customers that are located in the area. The first order of business will be defining the query parameters to be used. The data set being analyzed contains ID numbers, crash dates, the borough location of the incident, the description of the violations, and the violation codes. The parameters of interest for this analysis are going to be:

1. Borough location or “NYCBOROUGH”
2. Violation codes or “ViolationCode”

Using the MySQL Workbench analysis tool on this data set I will write a simple search query where I will select *all* data points from the data set. To make the query results easier to read I will select *all* boroughs and then while selected, count all violations, then finally group the results by borough and order the counts in descending order. The following screenshot will present the full query.



 According to the query performed, the five boroughs in New York City have the following violations per borough:

1. Brooklyn: 670 violations
2. Queens: 571 violations
3. Manhattan: 379 violations
4. Bronx: 333 violations
5. Staten Island: 46 violations

The results of this query would indicate clearly that Brooklyn, Queens, and Manhattan have the top three incidence rates, respectively.