**NYC Borough Analysis**

Camilo Hoyos

DAT-375 Data Analysis Techniques

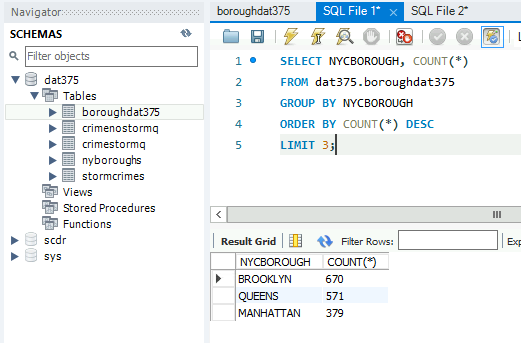
Jacqueline U. Clauss, M.S.

January 31, 2021

The data we will be reviewing is vehicle accident rates in New York City (NYC) boroughs. The attributes and attribute types of the data are as follows:

* ID - String
* Crash Date - DateTime
* NYCBorough - String
* Violation - String
* Violation Code - String

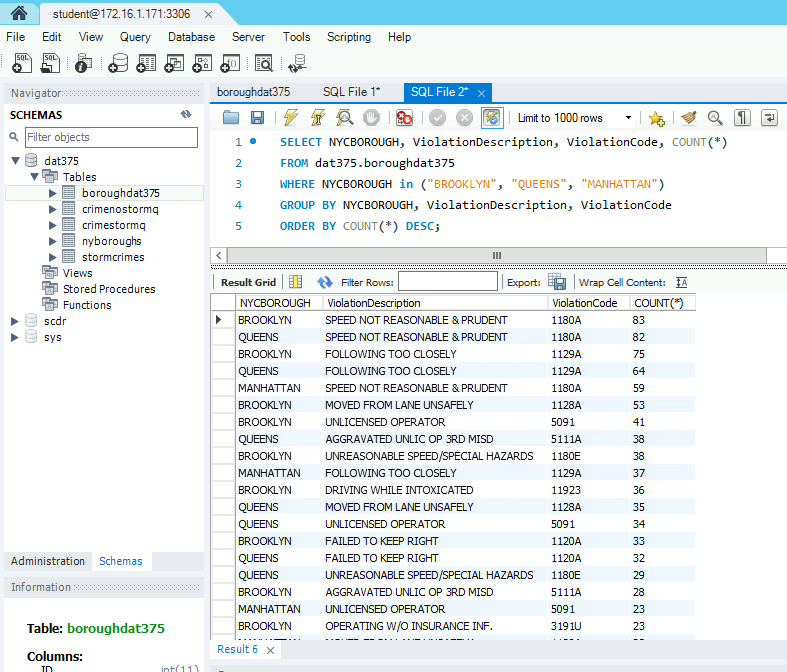
However, we will mainly be working with NYCBorough and a manually created column COUNT. The problem that is looking to be identified is “what are the top three boroughs with the highest number of accidents?”. The SQL statement below provides an exact response to the request.



“SELECT NYCBOROUGH, COUNT(\*) FROM dat375.boroughdat375 GROUP BY NYCBOROUGH ORDER BY COUNT(\*) DESC LIMIT 3;”

Here we can see that the results in descending order are Brooklyn, Queens, and Manhattan.

In spirit of the CRISP and reanalyzing for potential future analysis, I have included an additional query. (Larose & Larose, 2015) In our requirements there is a statement “The company will be using this report to calculate the insurance rates to charge customers living in those areas”. It is a fair argument to say that not every violation is equal in terms of risk to an insurance company. Perhaps instead of just applying an evaluation of insurance cost based on the number, the type of violation should be considered for a more granular cost. That being said, the new problem would be “Of the top three boroughs, what trend of violations do we see?”. It would be possible to take the query below, identify which violations cost more to the insurance company, and then create a price based on the distribution of violations. While the screenshot below does not provide the comprehensive results it provides the necessary context to calculate a better rate. This can be seen as a differentiator against other insurance companies.



References

Larose, D. T., &amp; Larose, C. D. (2015). Chapter 1.4. In Data mining and predictive analytics (pp. 5-7). Hoboken, NJ: John Wiley &amp; Sons. doi:https://mbsdirect.vitalsource.com/#/books/9781118991121/cfi/6/8!/4/14/6/2@0:100