

Taras Kaminsky, Eric Tran, Sid Rao, Numfor Tiapo

Database Systems Spring 2019 *Final Project Memo*

2/8/19

Datasets being used:

- 1) NYPD Motor Vehicle Collisions
<https://data.cityofnewyork.us/Public-Safety/NYPD-Motor-Vehicle-Collisions/h9gi-nx95/data>
- 2) 311 Service Requests from 2010 to Present
<https://data.cityofnewyork.us/Social-Services/311-Service-Requests-from-2010-to-Present/erm2-nwe9/data>

Both datasets are presented by NYC OpenData. No license is necessary to access and use the data. The two datasets can be joined using a street name and a date. Dataset 1 describes car collisions, where they took place, the type of the car, when they took place, and how many people were injured or killed. Dataset 2 describes all service requests reported to 311, such as potholes, illegal parking, blocked driveway, etc. These entries also have a street address and date to help join the two datasets together.

Dataset 1 has 1.4 million rows, and Dataset 2 has 19.7 million rows, but NYC OpenData has a tool that filters some things out before downloading. So, maybe a more appropriate scope is to view all entries from a whole calendar year, such as 2018.