

Examining the NYPD Arrests Dataset

Varun Somani, Mengjia Xia, Binxin Liu

ORIE4741 Big Messy Data
Fall 2019

The Dataset we will be using is the NYPD[1] arrests dataset. The dataset ranges from 2009 to 2017. It is a line-by-line record of each arrest the NYPD conducted during those years. We will use the dataset to try and answer the following question:

- How has the crime type changed in a certain region over the 8 years the dataset presents?

This question is important because the dataset ranges from time the world was reeling from the massive financial collapse which was centered in NYC. We can examine how an increase in prosperity from that time to now has affected the type of crime committed.

The dataset has exact dates for the crime, and we can also examine the race and gender. Thus, it will allow us a broad view of how crime has changed in the various regions, and whether certain racial or gender groups have fared better over the time than others.

References

1. OpenData, NYC. “NYPD Arrests Data (Historic).” Data.gov, Data.cityofnewyork.us, 21 May 2019, catalog.data.gov/dataset/nypd-arrests-data-historic.