# NYC Crime Analysis

Greg Kimatov, Mehedi Shohag, Zeal Patel

# Introduction

# Motivation

Has it become more or less dangerous to live in NYC over time? Our goal is to predict the NYC crime rate in the near future by analyzing the crime rate trend over a recent 14-year time period.

# About the data

# NYC Crime Stats

- Collected from NYC Open Data
- Consists of arrests made in NYC between 2016 and 2019
- Consists of information like arrest date, offense description, offense level, age group the perp falls into, sex of the perp, and borough tthe arrest was made.
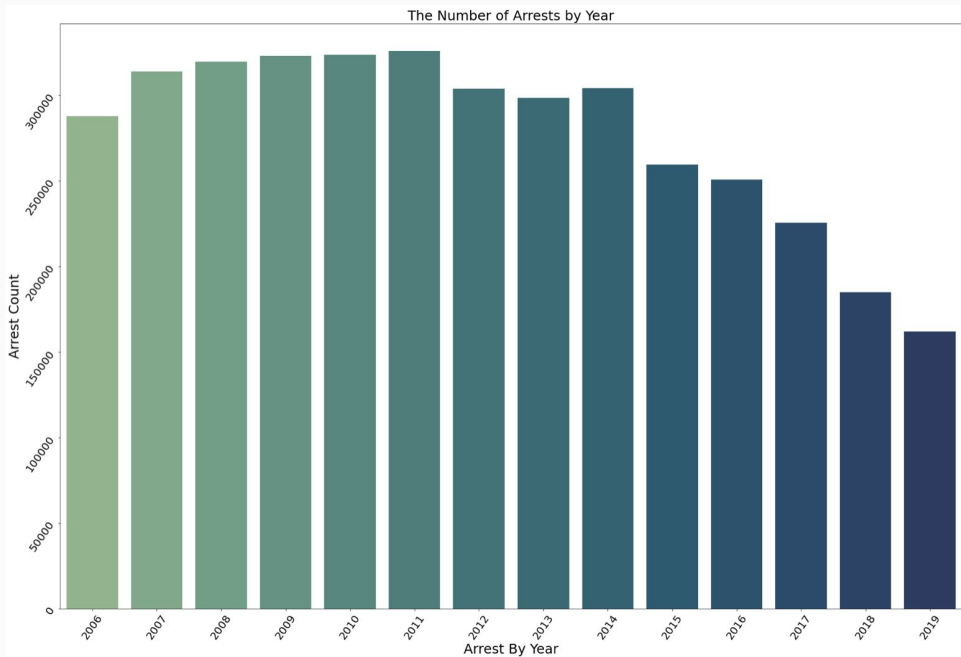
# Cleaning the data

# Cleaning the data

- Dropped null values
- Converted to string date to datetime objects
- Rename column names to be more descriptive
- Rename categorical cell values to be more descriptive
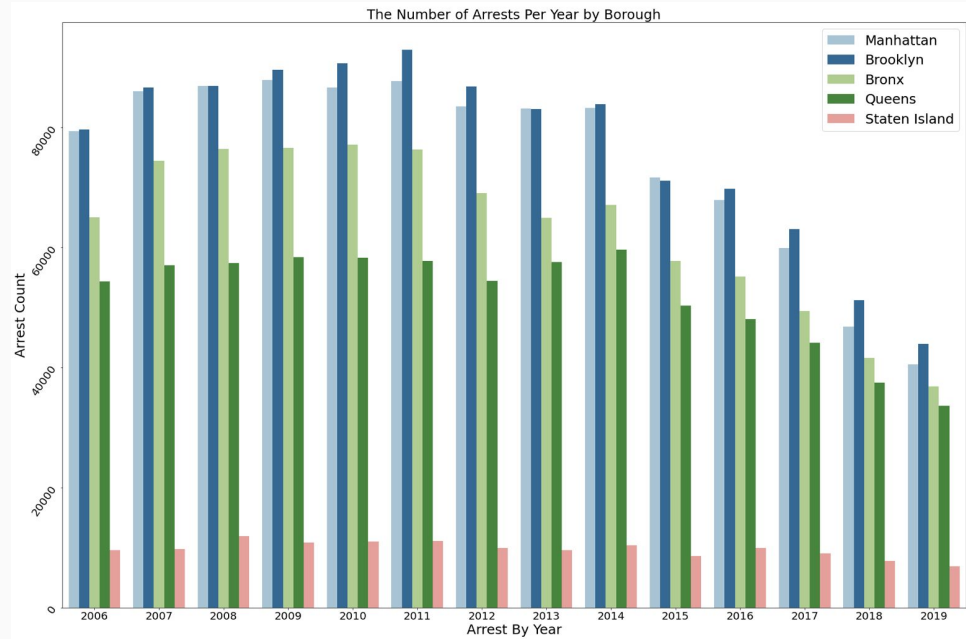
# Visualizing the data

# Number of arrests per year in the city

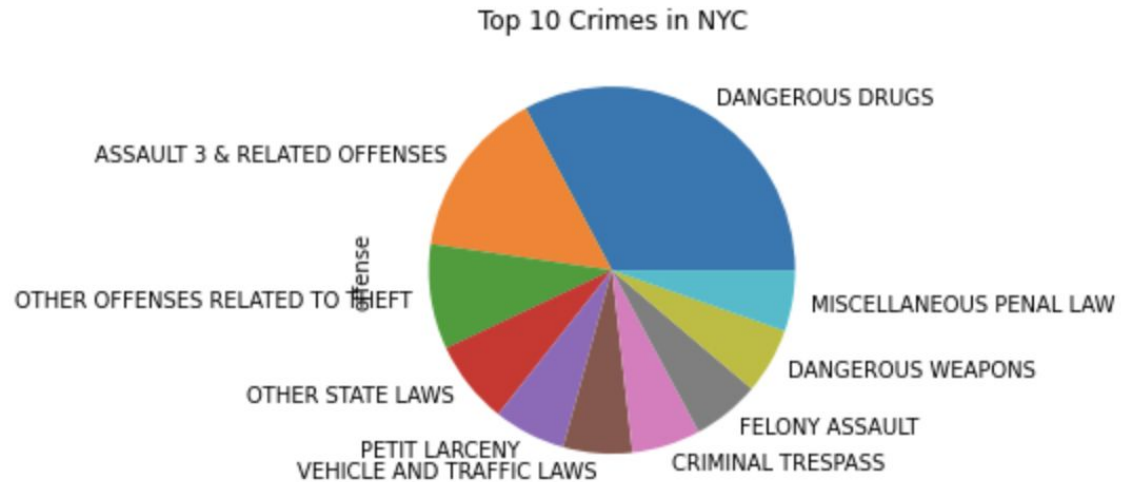The graph shows the total number of arrests across all five boroughs of NYC per year.

# Number of arrests per year by borough

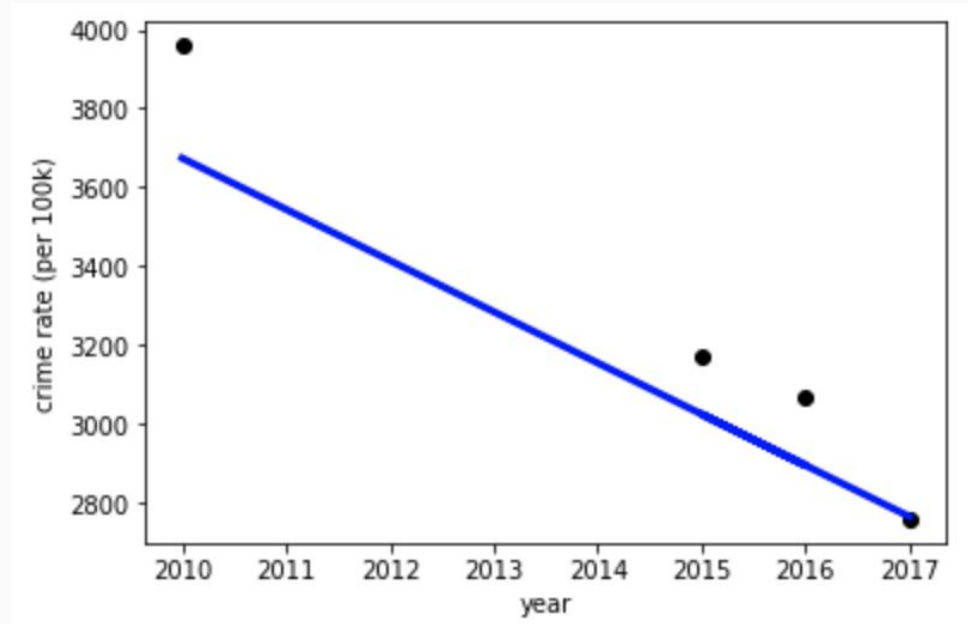The graph shows the number of arrests in each borough between 2006 and 2019



The Number of Arrests Per Year by Borough

# Top 10 types of offenses in NYC

The pie chart visualization shows the top 10 crimes committed in all of NYC



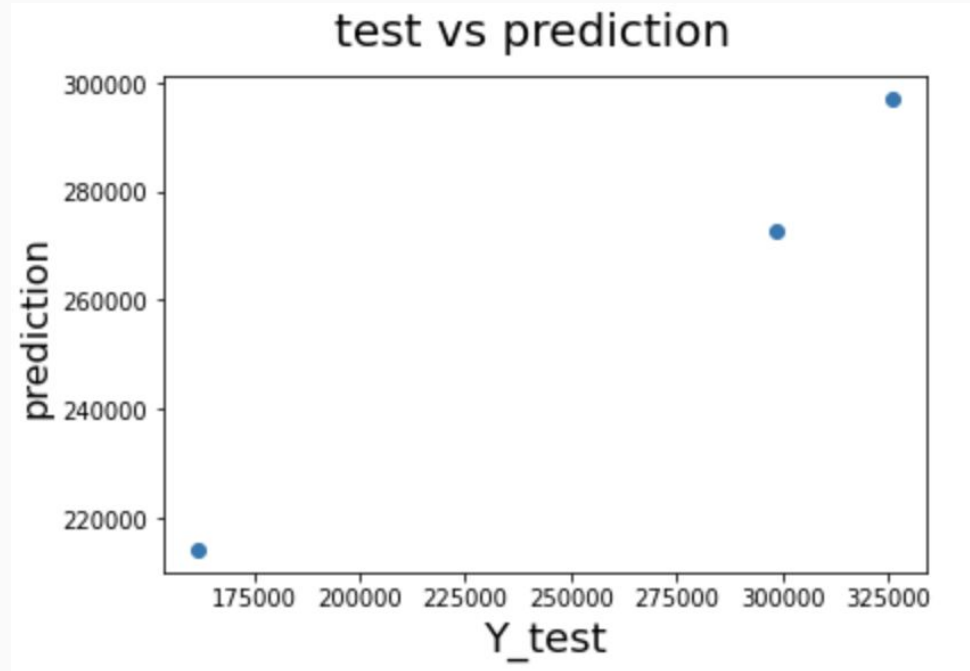Top 10 Crimes in NYC

# Making predictions

# Predicting the future yearly crime rate

- Feature: Arrest year
- Target: Crime rate per 100,000 people
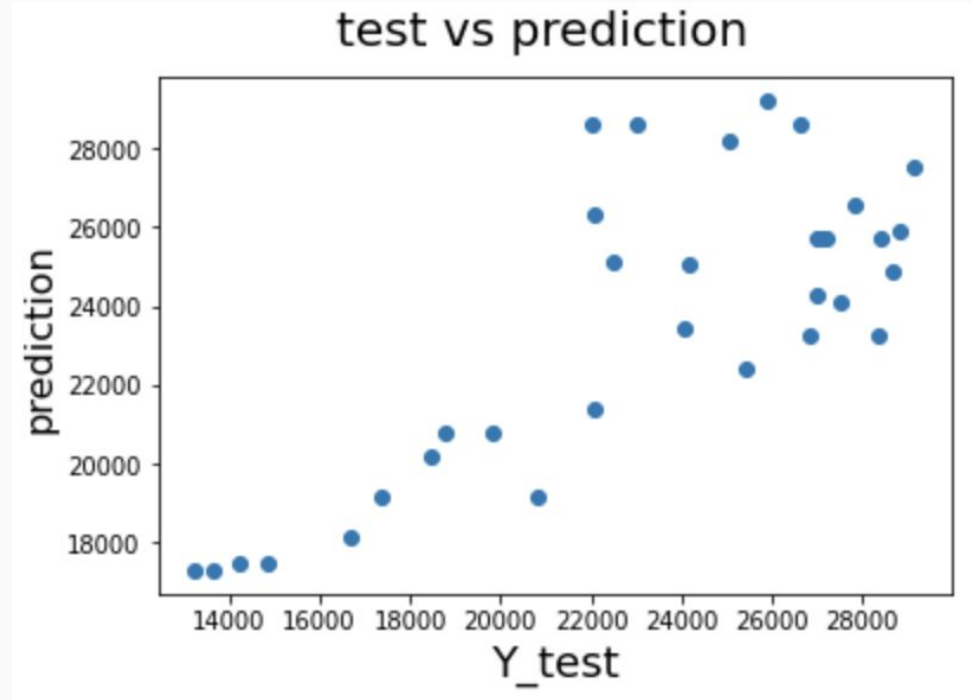- Accuracy score: 83%
- 2030 prediction: 1072

# Impact of annual GDP on crime rate

- Features: Arrest year, Annual percent change in GDP
- Target: Total number of arrests per year
- Accuracy: 67%
- Prediction if GDP percent change in NYC is +40% in 2045: 45376
- Prediction if GDP percent change in NYC is +1% in 2045: 473368



test vs prediction

# Impact of unemployment rate on crime rate

- Features: Arrest year, Unemployment rate
- Target: Total number of arrests per year
- Accuracy: 62%
- Prediction if unemployment rate is 88% in 2045: 154,920
- Prediction if GDP percent change in NYC is +40% in 2045: 55668

# Future work

# Future work

- We can collect additional data to keep our datasets up to date since we were only able to collected data until 2019.
- We can collect more detailed supplementary datasets to account for more economic features that can impact the crime rate.
- Create different types of models rather than just linear regression.

# Thanks!