Summer weekends at night are the most dangerous in Boston.

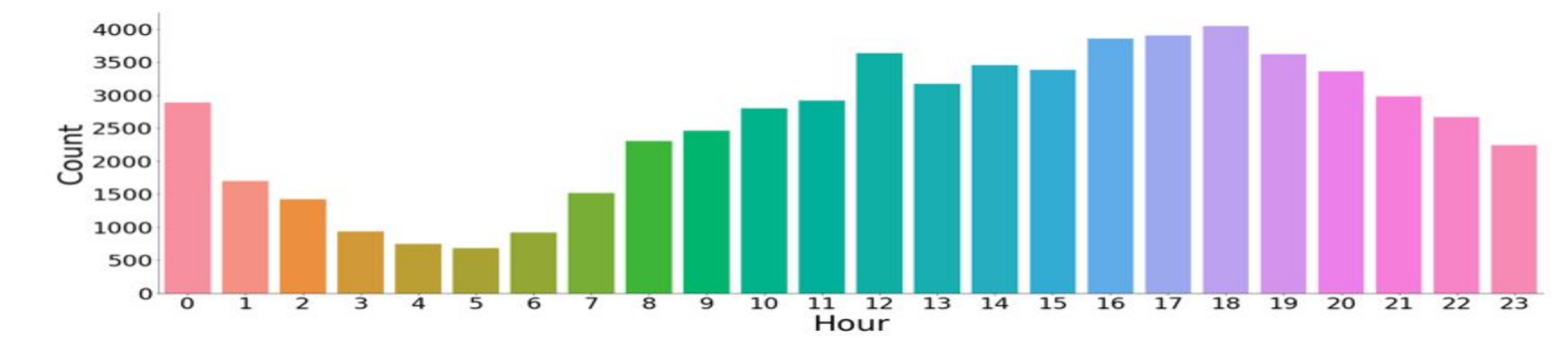
...

Luke Gruszka

The Variables

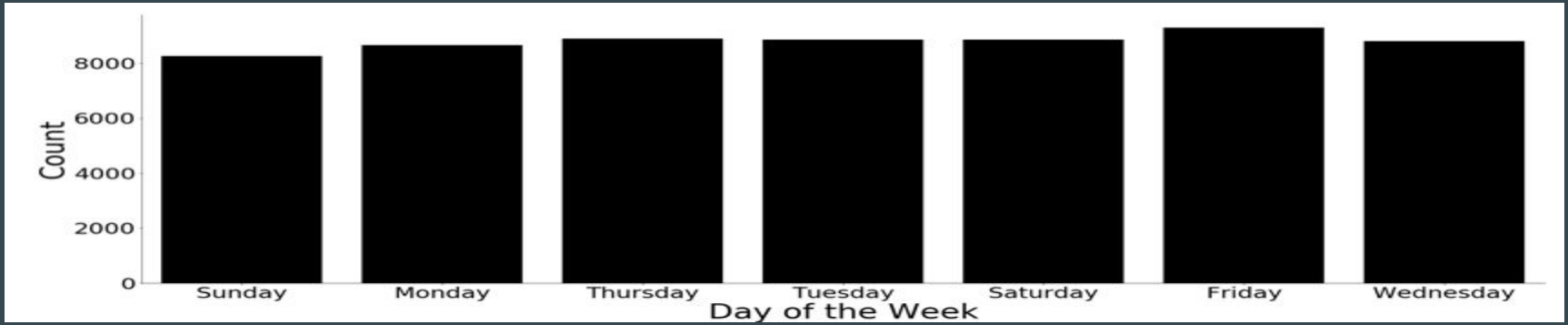
- Offense Description
  - To understand the most violent and reoccuring crimes
- Hour
  - Which part of the day is most likely to have a crime
- Month
  - Which month has the most amount of crime
- Day of Week
  - What day tends to have the most amount of crime
- Shooting
  - If the crimes were deadly
- UCR Part
  - How bad the crime was compared to the other crimes

Hour



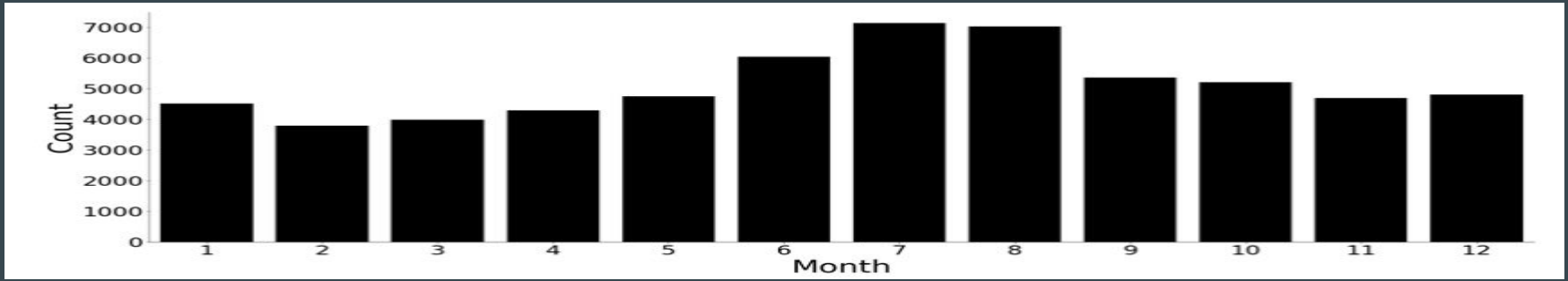
- No Outliers. All data stays in the flow of the data since there is a set number of hours in the day.
- It is skewed Left and shows that more crime happens in the late afternoon.

Day of the Week



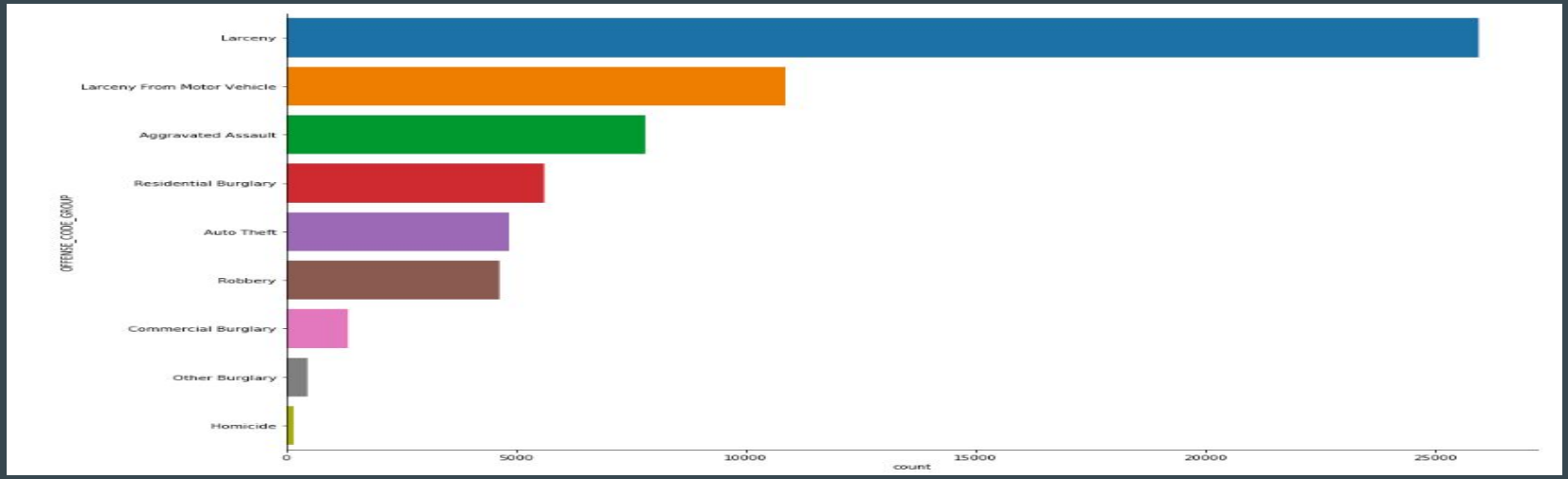
- No Outliers. There is a set number of values for days of week.
- The data is pretty similar although it looks to have some more in the Friday value.

Month



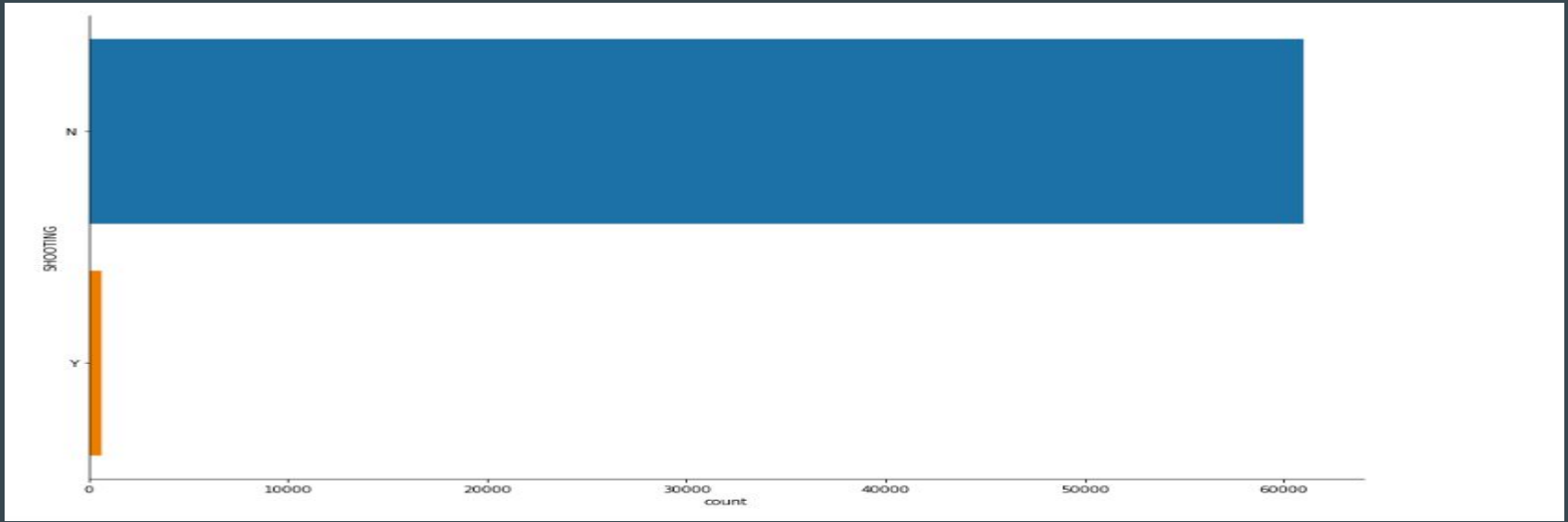
- No outliers because there is a set amount of values for the months.
- The data shows that there tends to be more crime when the weather is nicer particularly in the summer.

Type of Crime



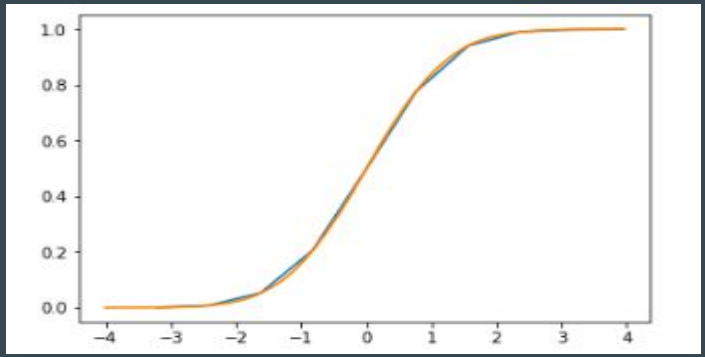
- The data was controlled to incorporate only the major crimes. This will help take away any outliers that are one offs and don't help provide to the major goal of this analysis.

Shooting



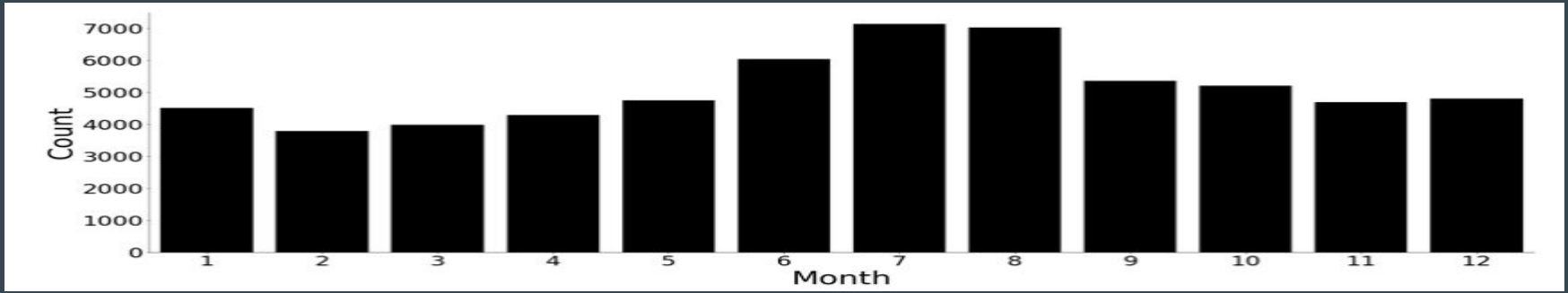
- Shooting were taken into account to see if the crimes turned deadly. These will show how bad the crimes were and if there are correlations to the time of day.

CDF on Month

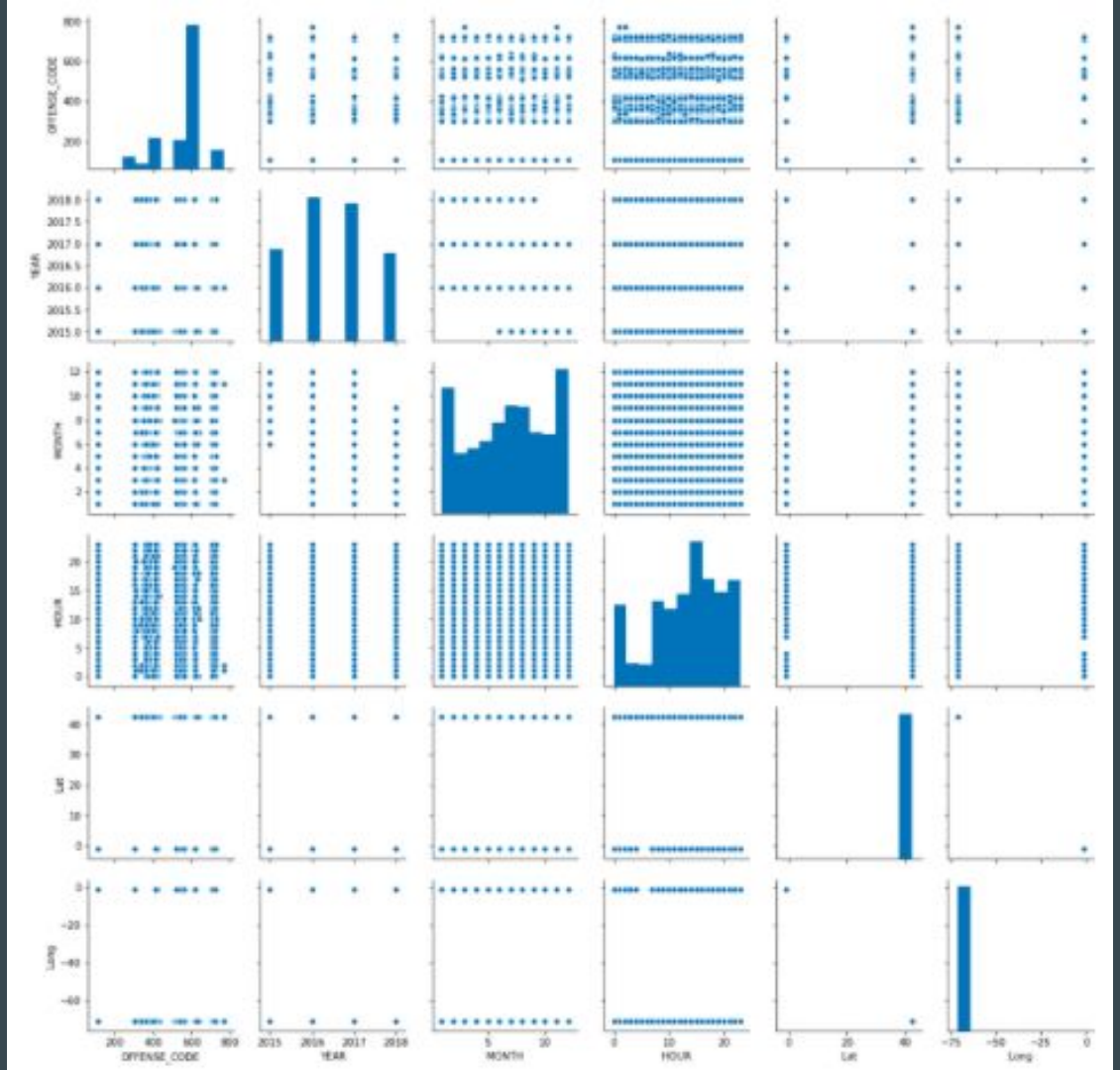


- This shows that the value that there are no outliers in the data.
- It also shows that the data in normally distributed and is not skewed.
- This tells me that crime is typically probably in all months and does not quite mean that it is heavily committed to one month. Also, this could become the values are pretty central and only 12 options.

Months Analytic Distribution



- The data shows that there is a correlation between the warmer months and the number of crimes.
- The unique topic would be January where it seems to be slightly higher would could resemble the holidays
- Also, in the summer days are longer with more sunlight so that means there is more time spent in the day where it could allows for a higher probability that a crime can occur.


OFFENSE_CODE
YEAR
MONTH
HOUR
Lat
Long

Analysis of the Scatterplots

- There is a correlation between the months and the number of crimes committed.
- This will then lead to having a more likely correlation between the type of crimes committed. It will have casaution on the likelihood of the more violence crimes being committed since the probability of the crimes are more likely.
- Also, that causes a causation for the Shooting to be more likely when there are more crimes taking place.
-

Hypothesis Testing

- I tested the Month, Day of Week, and hour is different tests.
- Month can be accepted. The summer month of July tends to have a high number of crimes committed.
- Day of Week was reject. Friday was the highest for the number of crimes committed. Although Saturday was the Second highest it can be approved in the 95% percentile.
- Hour was accepted. On the confidence interval of 95% it fell within the variance of the most probability of crimes being committed.

Regression Analysis

- The regression shows that causation is at work in the analysis. There is correlation between the time of day and the count of the crimes. But the weather is a causation on the month of which the crimes are being committed. The type or group of crime is also showing not correlation to the month or time of day. Also, the day of the week is also showing that it is static through all months of the year with friday being the most relevant. Also, shooting are showing correlation between the day, month, and hour. All the most popular in those variables show the highest number of shootings. This could be because the probability of having shooting have increases significantly.