San Francisco Crime Data Report

Coursera Capstone Project for Data Science

Introduction

- This report will analyze San Francisco crime data from 2018/1 to now
- The report will be useful for government, politician, public and tourists
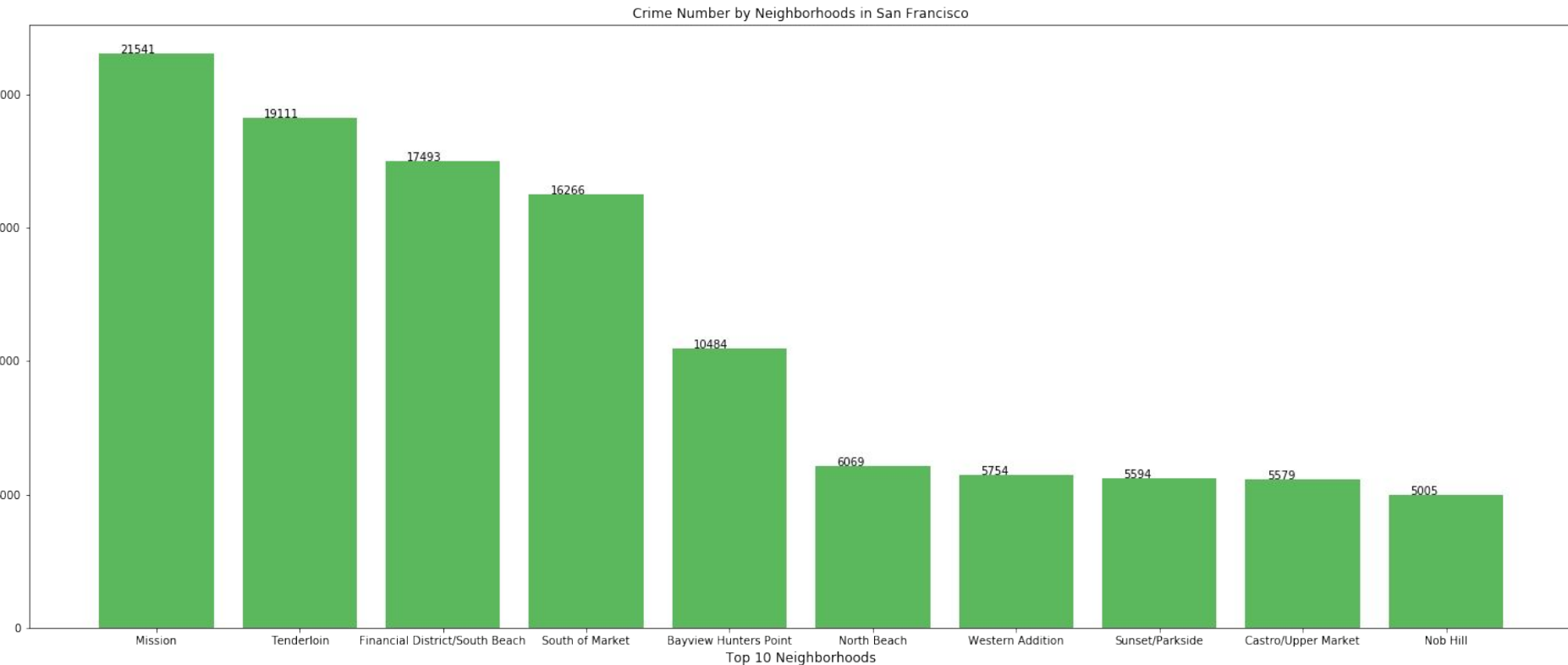- We will have exploratory data analysis on the dataset

Data Preparation

- The data is from publicly available website from sfgov website. The link is here https://data.sfgov.org/Public-Safety/Police-Department-Incident-Reports-2018-to-Present/wg3w-h783
- Rows with NaN values will be dropped from the dataset since we have around 200,000 rows in the dataset. The valid data will be adequate for us to generate the report
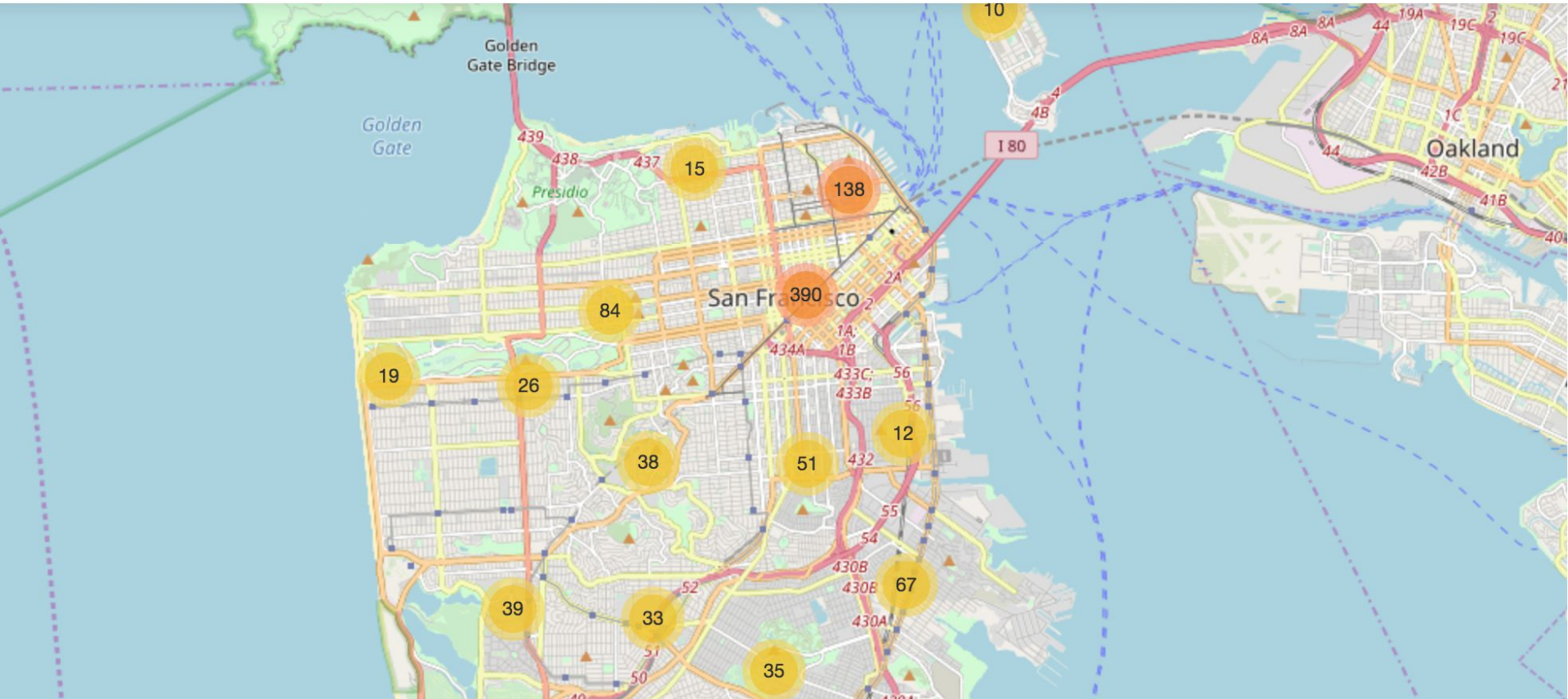
Exploratory Data Analysis

For exploratory section, we will categorize the data on different features, such as neighborhoods, date of the crime, day of the crime and time of the crime.

- Categorize the data by neighborhoods
- Categorize the data by incident category
- Categorize the data by month
- Categorize the data by day of the week
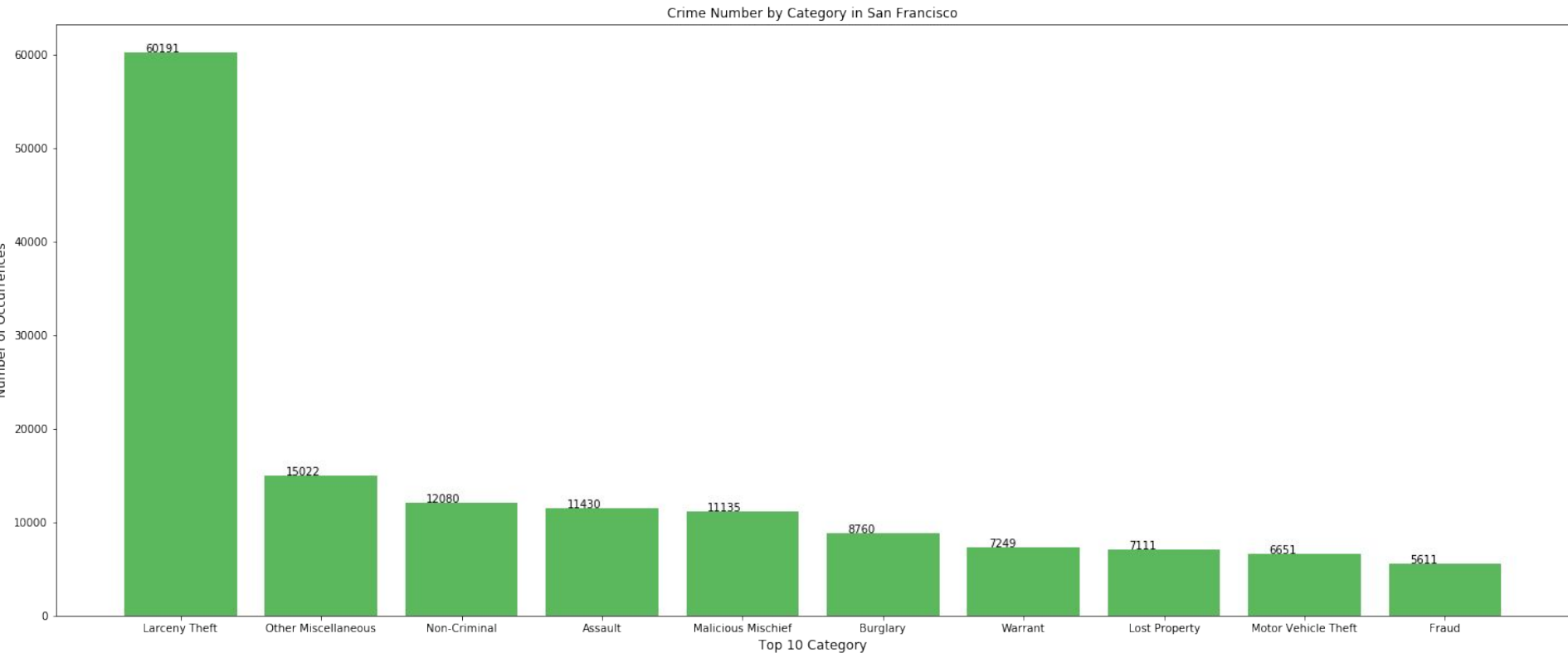- Categorize the data by time of the crime
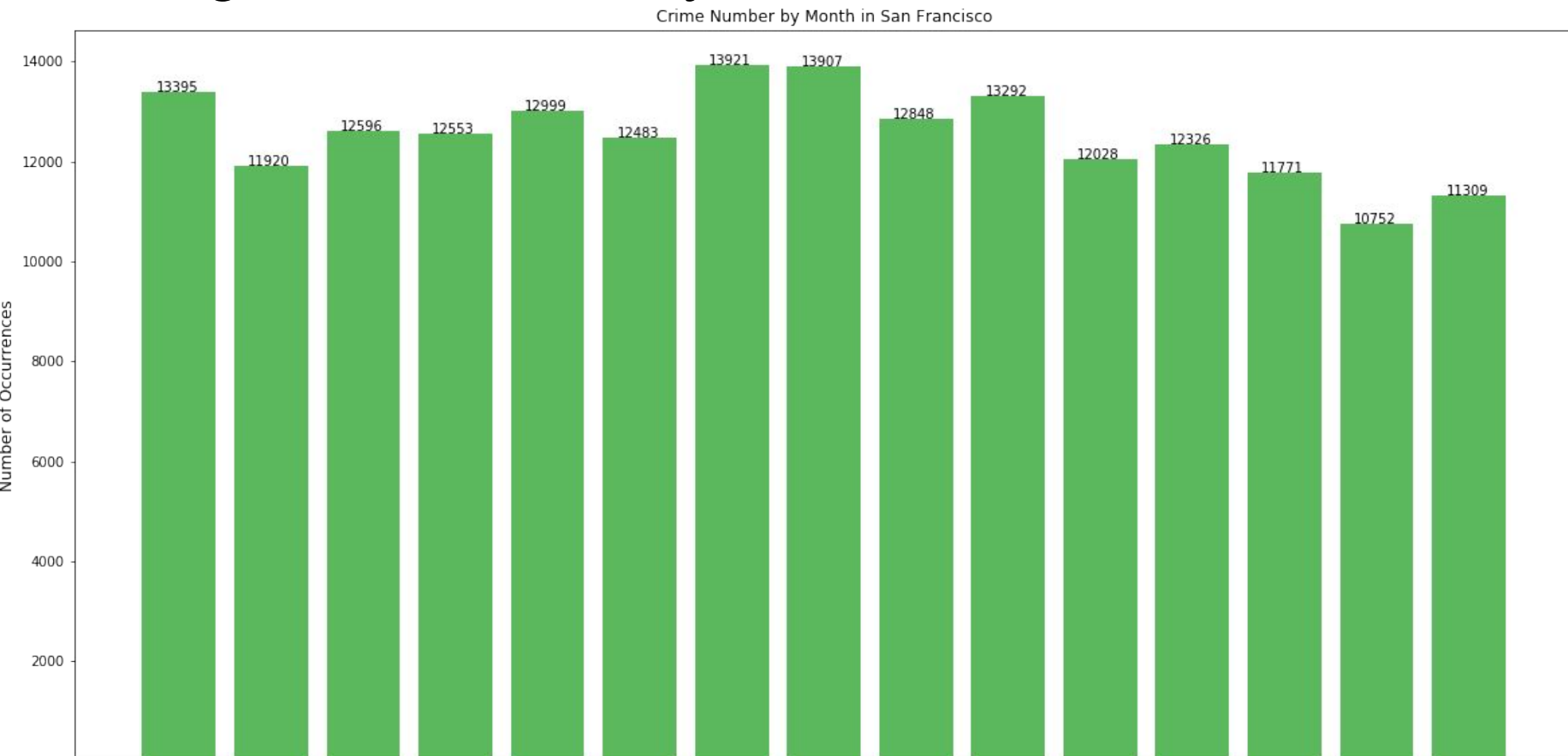
Categorize the data by neighborhoods

Crime Number by Neighborhoods in San Francisco

Top 10 Neighborhoods

Sample 1000 crime distribution on map

Categorize the data by incident category

Crime Number by Category in San Francisco
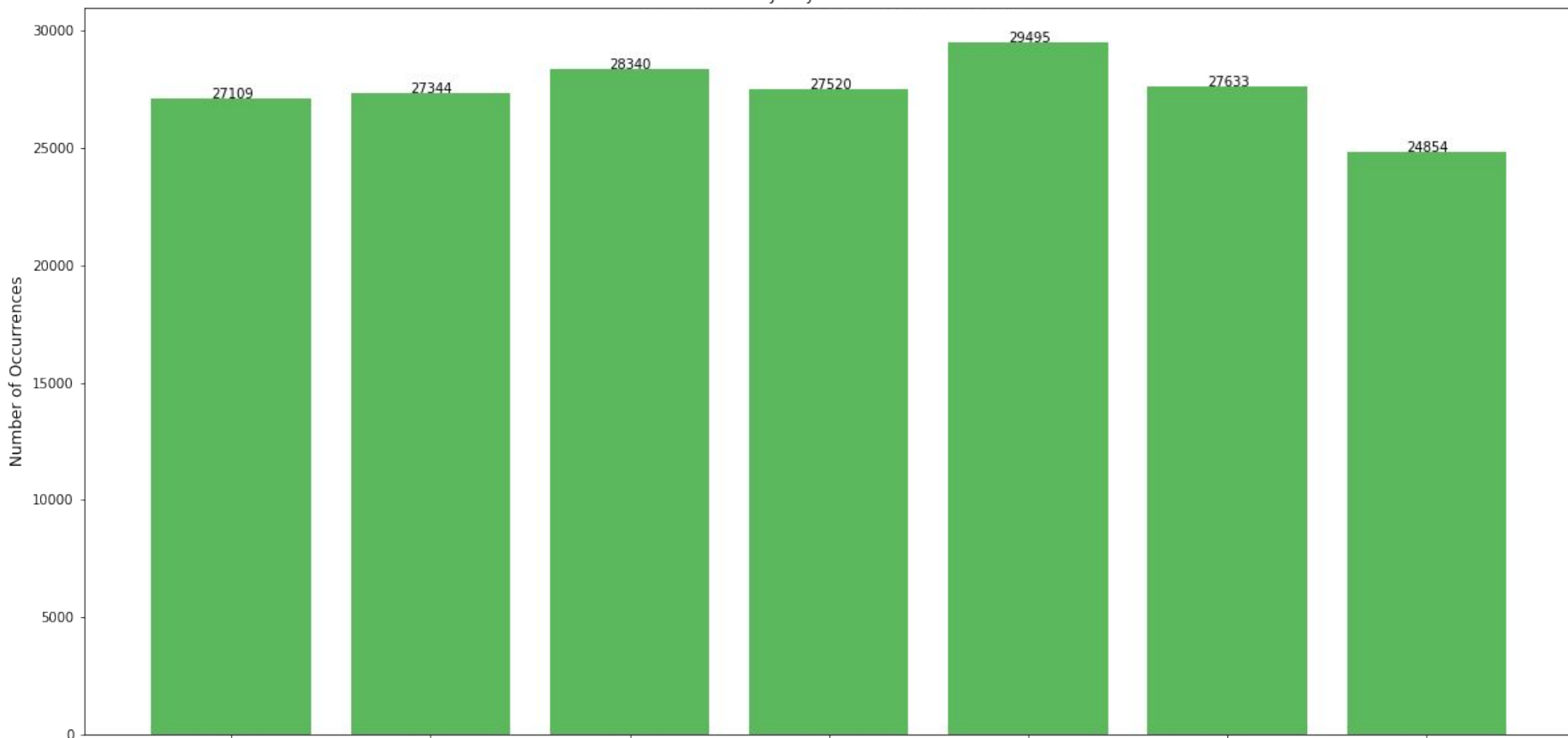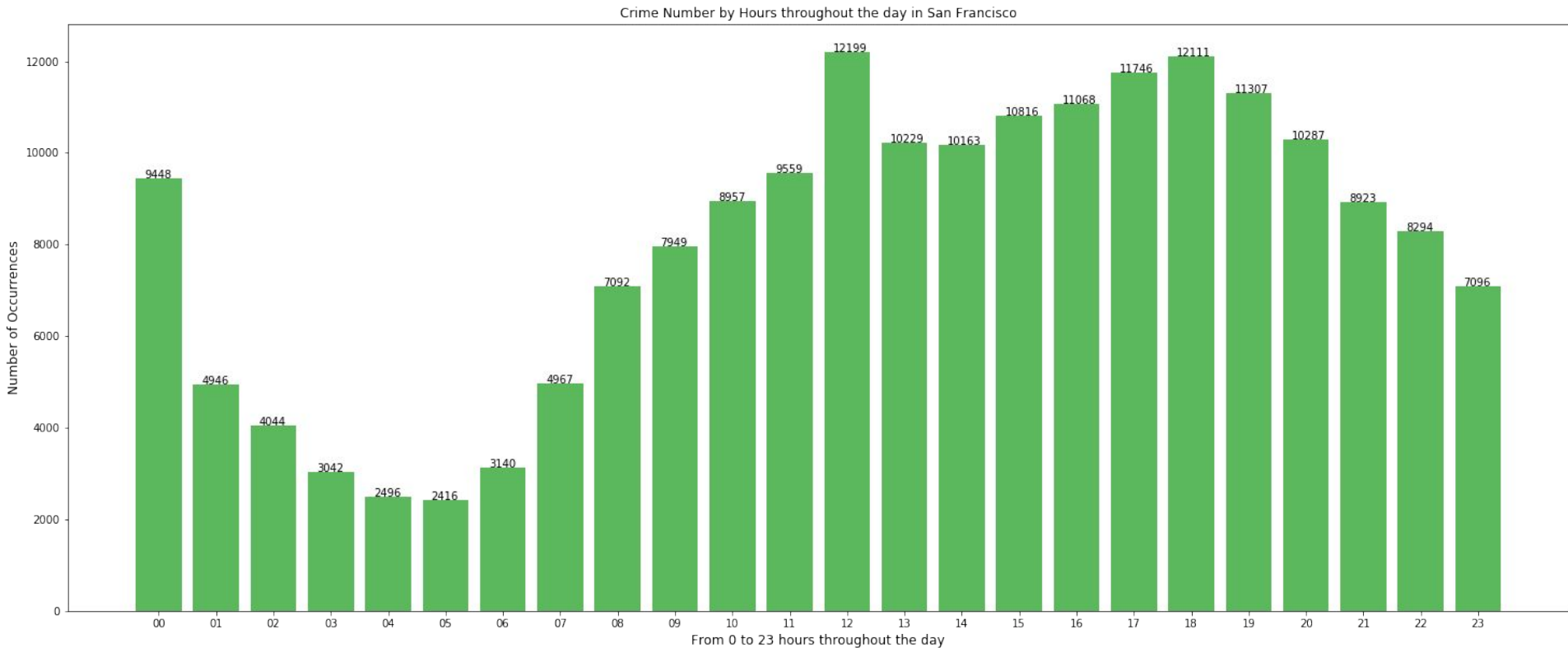
Categorize the data by month

Crime Number by Month in San Francisco

Categorize the data by week of day

Crime Number by Day of Week in San Francisco

Categorize the data by time of day

Crime Number by Hours throughout the day in San Francisco

Conclusion

- Mission neighborhood has the highest crime number
- Larceny Theft has the highest number of crime
- July and August have the highest crime number throughout the year
- Friday has the highest crime number throughout the week
- Crime number throughout the day has a clear trend