Predict 911 calls in San francisco city

By: Simantini Patil

Project Overview

This project analyzed **911 emergency call** data compiled from **San Francisco metropolitan area** and predicted the number of calls for violent and property crimes.

The project will help the authorities to know the crimes that have occurred over the time and help reduce the occurrence of them.

Project Objective

To help police authorities to know the crimes that have occurred over the time and to give them an ability to foresee the occurrences of crime/ emergencies so that they can implement proactive strategies to reduce the calls in the future.

Data Set

911 calls data published by City and County of San Francisco available at https://data.sfgov.org/

- Made available by Socrata Open Data API (SODA)
- Covers the period 03/31/2016-present
- Has over two million rows and 14 columns

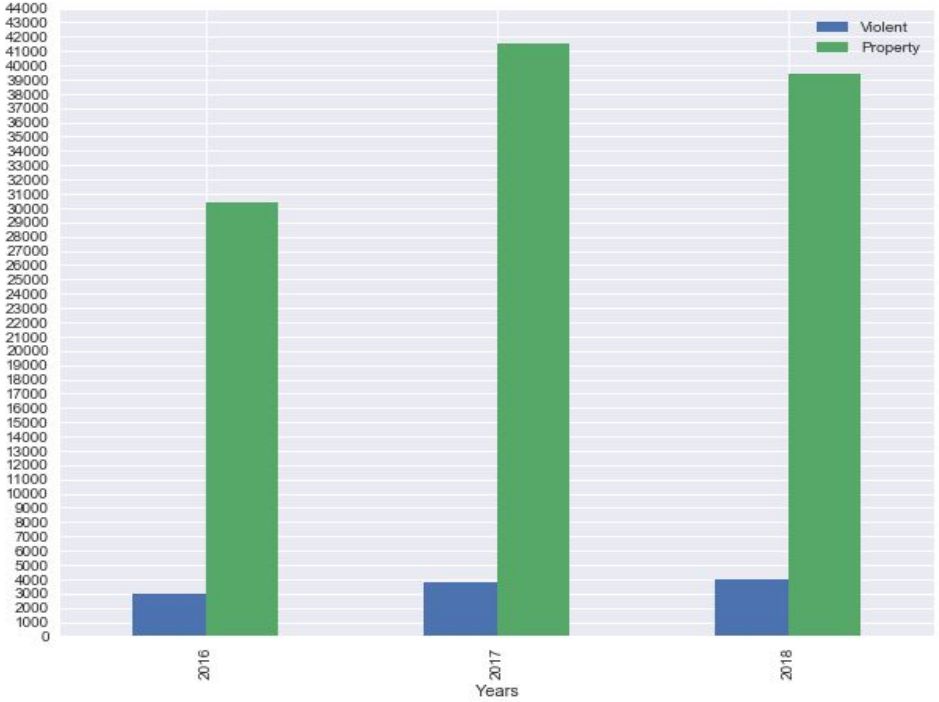
Approach

Dataset was divided into violent crimes calls and property crimes calls to make data manageable.
The problem was divided into several steps:

1. **Data Wrangling:** Clean and wrangle the data
2. **EDA:** Extensive data visualisation was used to extract insights and pattern from the dataset.
3. **Inferential Statistics**: Hypothesis tests were built to derive the statistical significance of the features.
4. **Machine Learning**: Various regressors were tested and their accuracy recorded. Hyperparamters were tuned. The model was then fit on the test data and the best model was obtained.

Exploratory Data Analysis

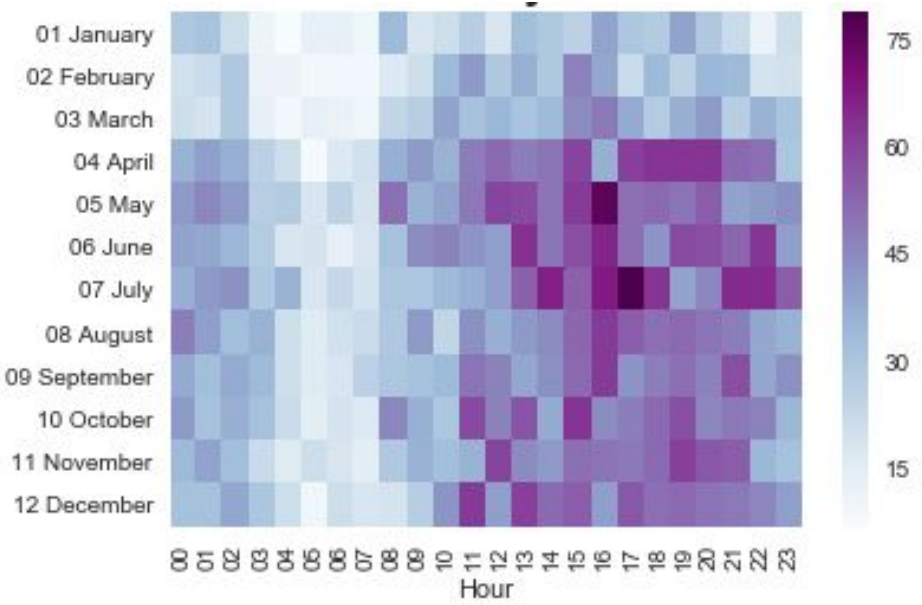
How number of calls has changed over the years? is it decreasing?



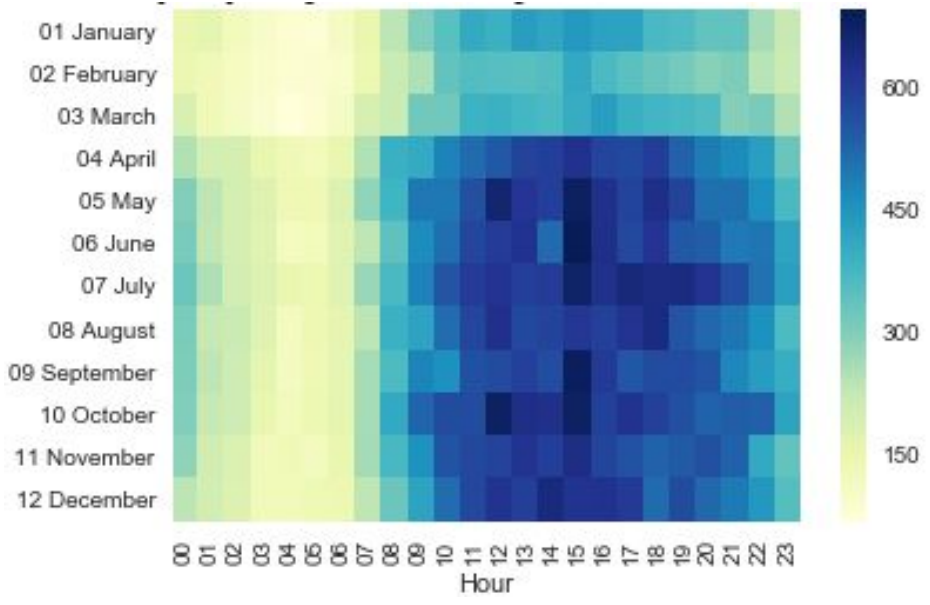
Exploratory Data Analysis

911 Calls for violent crimes and property by month and hour of the day

Violent Crime



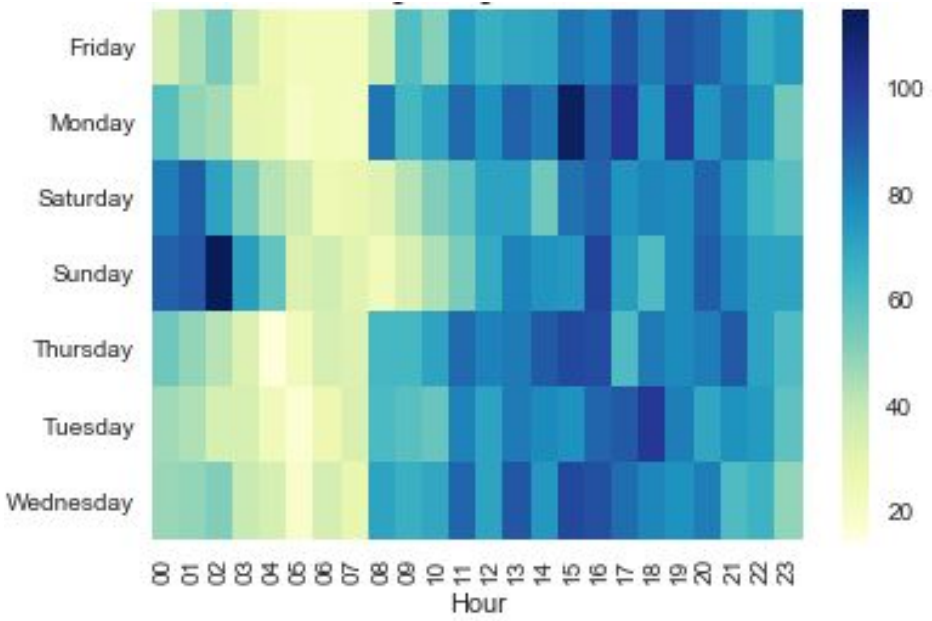
Property Crime



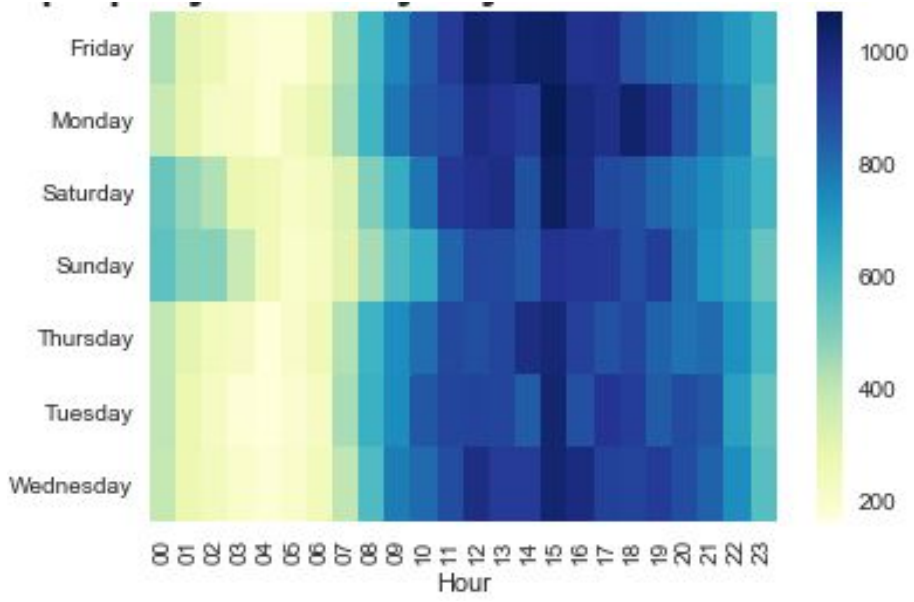
Exploratory Data Analysis

At what time of a day more calls are received?

Violent Crime



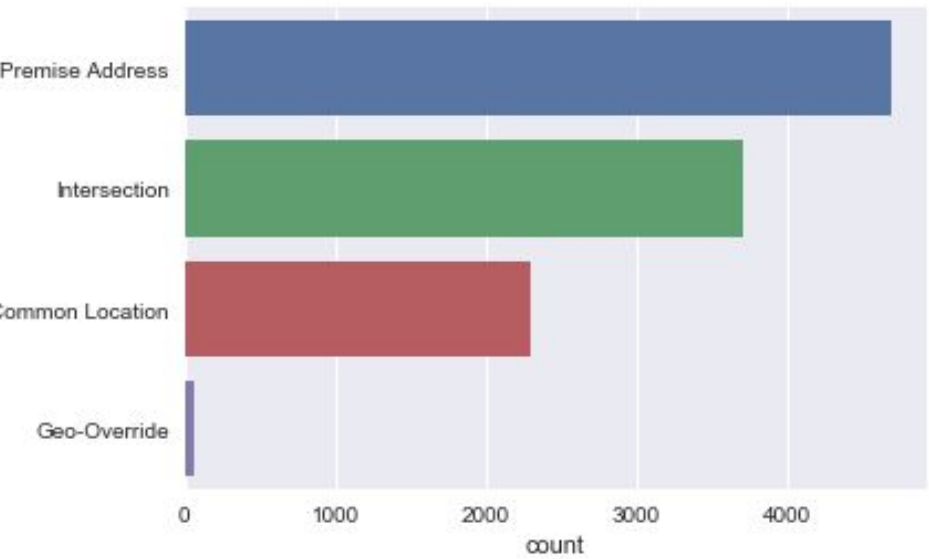
Property Crime



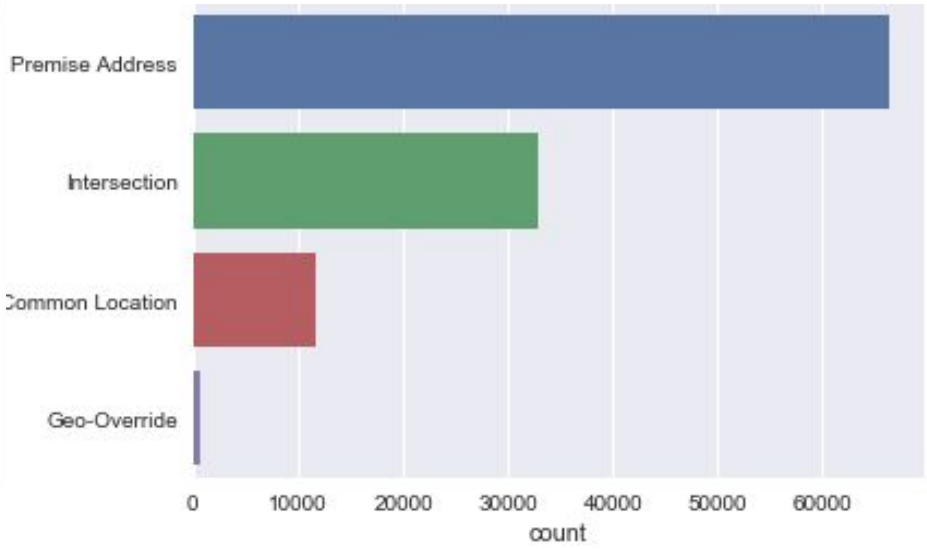
Exploratory Data Analysis

911 Calls for violent and property crimes reported on different address types

Violent Crime



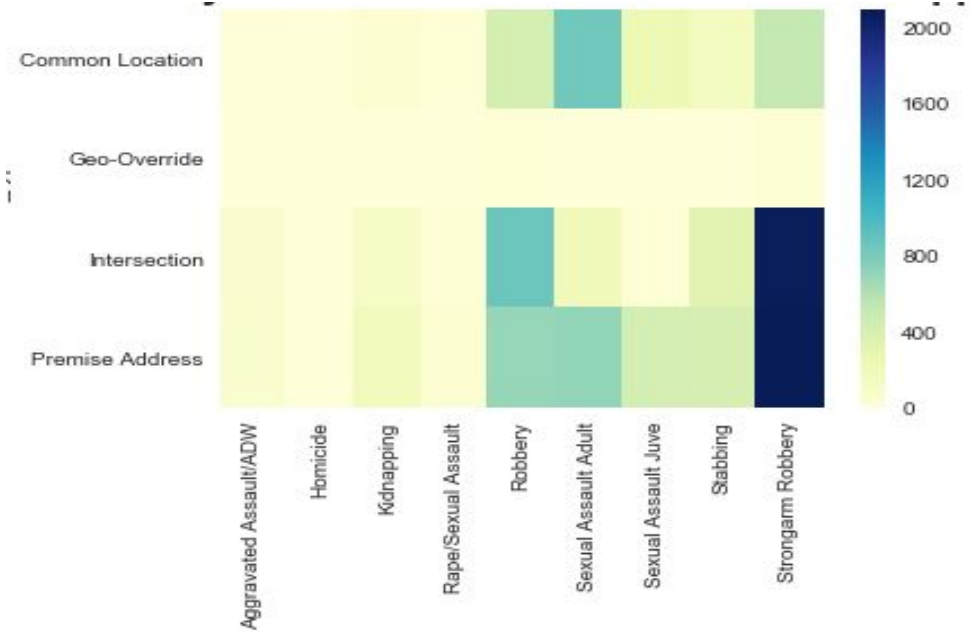
Property Crime



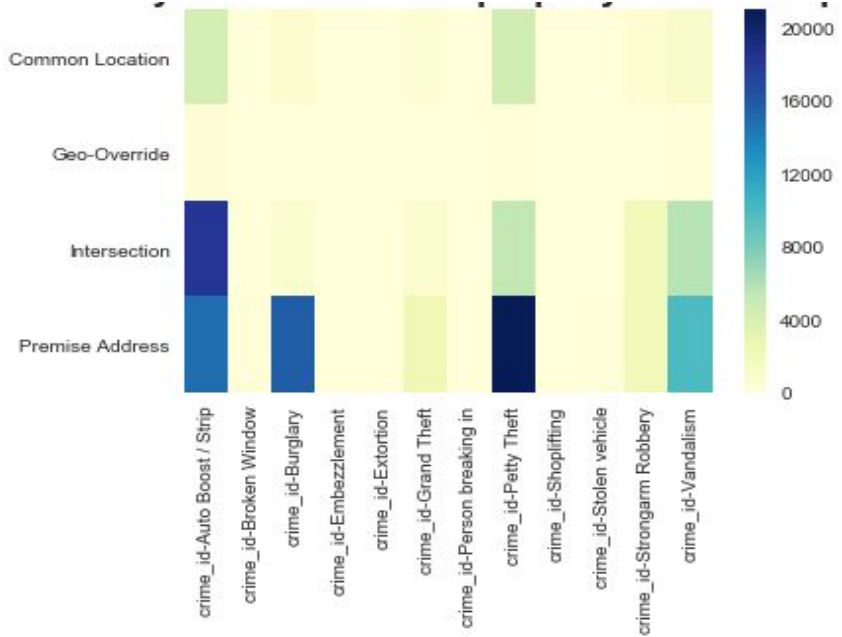
Exploratory Data Analysis

What is the most common location for a certain type of crime to happen?

Violent Crime



Property Crime



Feature Engineering

Below new features were created from original dataset

- event: no. of calls occured on a particular day
- Hour: extracted hour data from ‘call_dttm’ datetime column
- Year: extracted Year data from ‘call_dttm’ datetime column
- Month: extracted month data from ‘call_dttm’ datetime column
- Date : extracted date from ‘call_dttm’ datetime column
- Day: extracted day of the week data from ‘call_dttm’ datetime column

Hypothesis testing

From data visualization, it was observed that most calls were being made in summer but to prove hypothesis tests were performed.

1. The Null hypothesis : There is no difference in number of calls for service in summer and Spring
2. The Null hypothesis : There is no difference in number of calls for service in summer and winter

Both the Null hypothesis were rejected as p-value= 3.73e-06 and 3.69e-08 respectively.

Modeling

The goal was to predict the number of calls for the period June 2018 - December 2018. The data was split accordingly.

- The data from March 2016 - May 2018 was planned to be used for training

- The data from June 2018- December 2018 was our testing dataset.

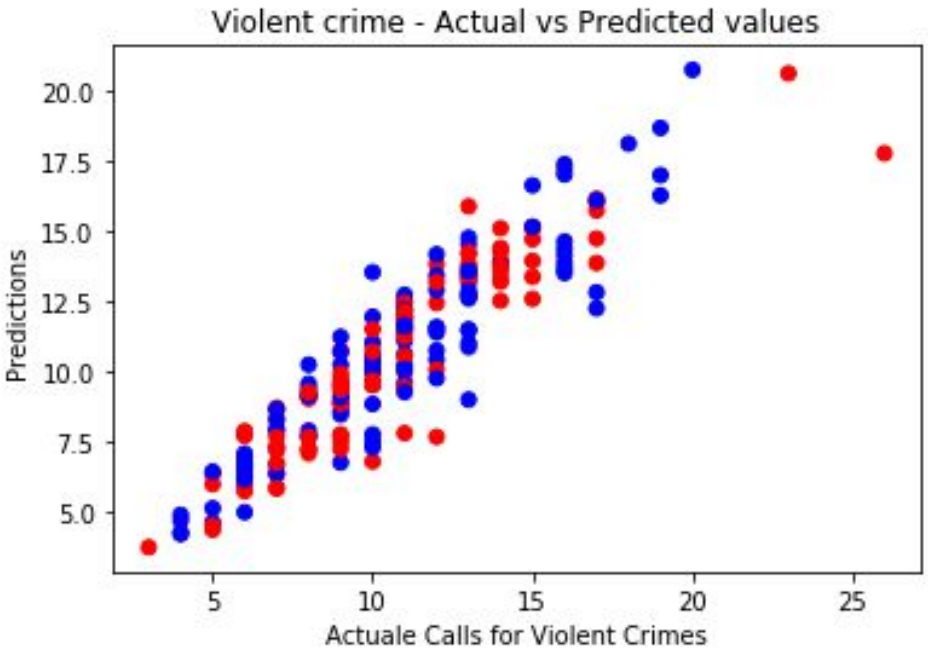
Linear Regression, Random Forest Regressor and Gradient Boosting Regressor models used on the data.

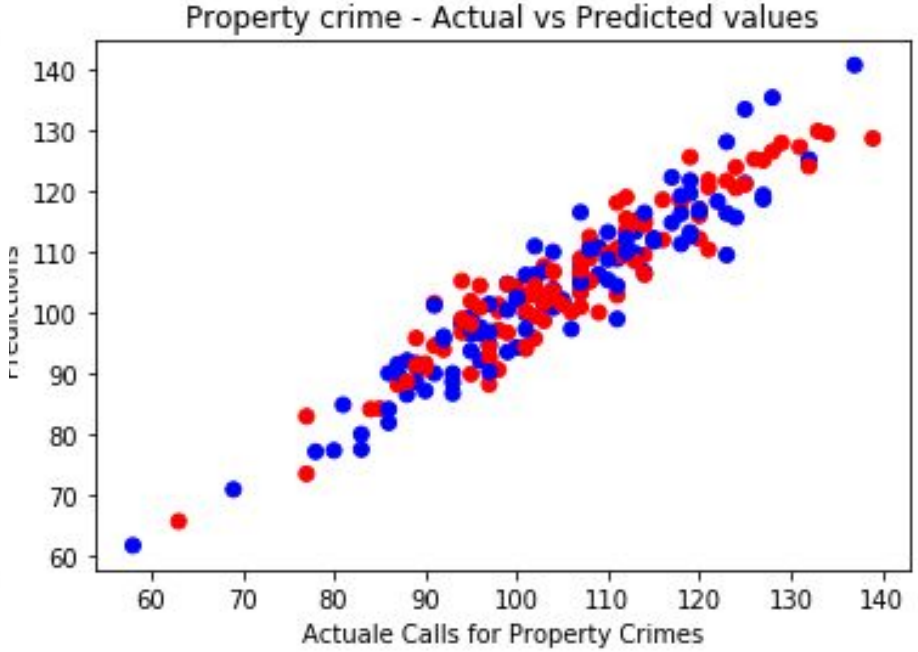
Model comparison

	Linear Regression		Random Forest Regressor		Gradient Boosting Regressor	
	MSE	r2_score	MSE	r2_score	MSE	r2_score
Calls for Violent crimes	2.60	0.811	2.54	0.815	2.16	0.843
Calls for Property crimes	20.46	0.890	23.38	0.874	21.71	0.883

The best model to predict Violent crime was Gradient Boosting Regressor and Linear Regression for property crime calls.

Model Predictions





Conclusions

The following Insights can be given.

- There is a spike observed in no. of calls for violent and property crimes during summer.
- Most number calls are expected to happen at 3.00 pm so authorities can implement proactive strategies to reduce the calls in the future at this time.
- Aggravated Assault is mostly reported from noon till midnight.In short, San Francisco is as safe as any big city neighborhood. But at night, it's wise for everyone to stick to streets with lots of foot traffic, or travel with an equally alert, sensible companion.