San-Francisco Venues Crime Analytics

San-Francisco Venues Crime Analytics Agenda

- Analyzing Crimes in San-Francisco Neighbourhood
- Data Acquisition and Cleaning
- Analysis of Top Crime in San-Francisco
- Peak Day for Larceny Crime in San-Francisco
- Distribution during Week day for Larceny Crime
- Distribution of Larceny Crime in San-Francisco
- Spread of Venues for Financial District/South Beach
- Non – Linear Relationship between Crime and Time
- Linear Regression & Polynomial Testing
- Conclusion and Future Direction

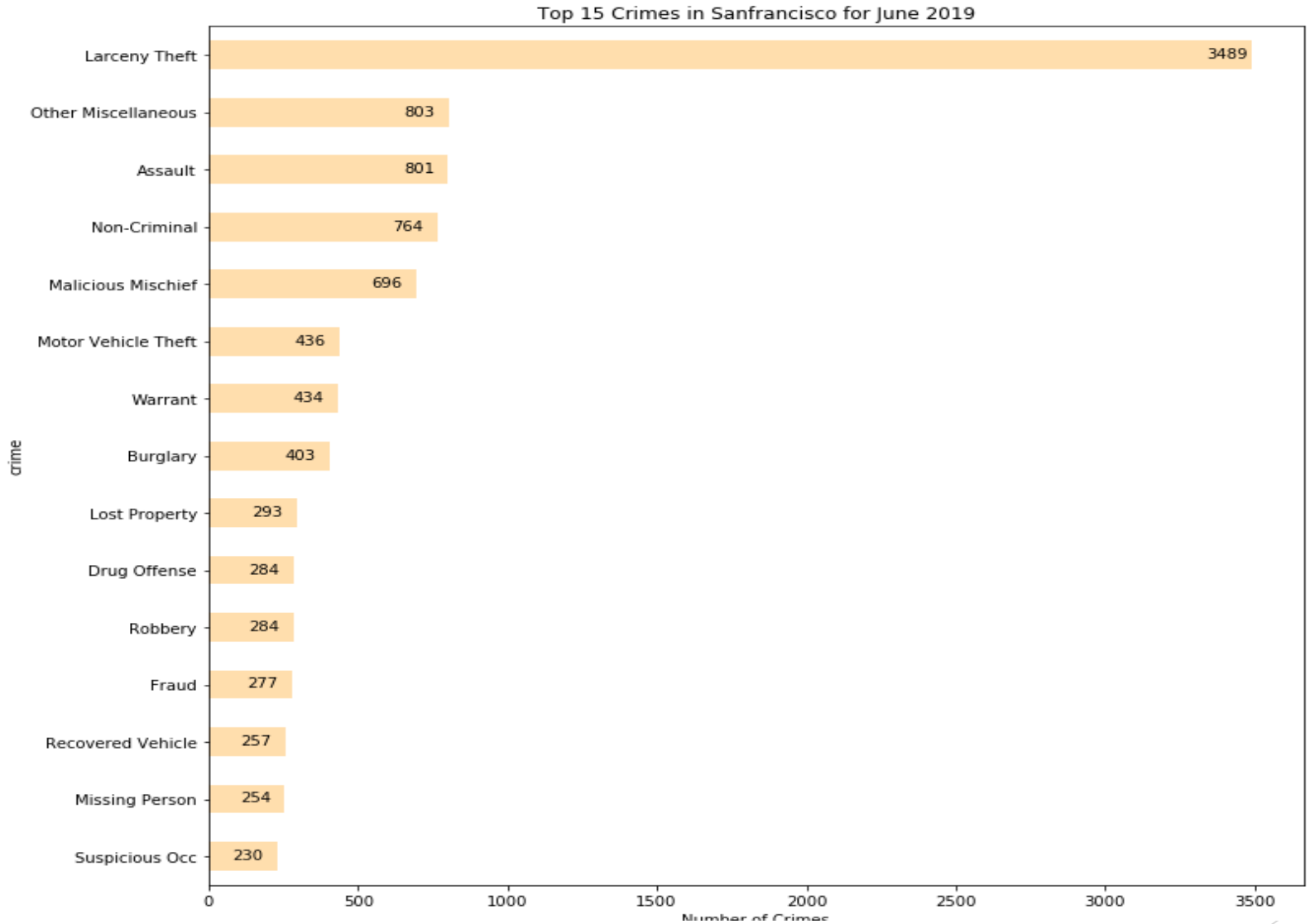
Analysing Crimes in San-Francisco Neighbourhood

- San Francisco is one of the 20 fastest growing cities in the United States
- San Francisco is the nation's leader in Burglary, larceny, shoplifting, and vandalism
- Crimes Effect growth of economy and hindrance to commerce
- Optimal resource deployment by SF Police Department
  - Control of Crimes by speedy response
  - Strategic deployment of force at Key Venues
  - Peak hours work load optimisation for SFPD

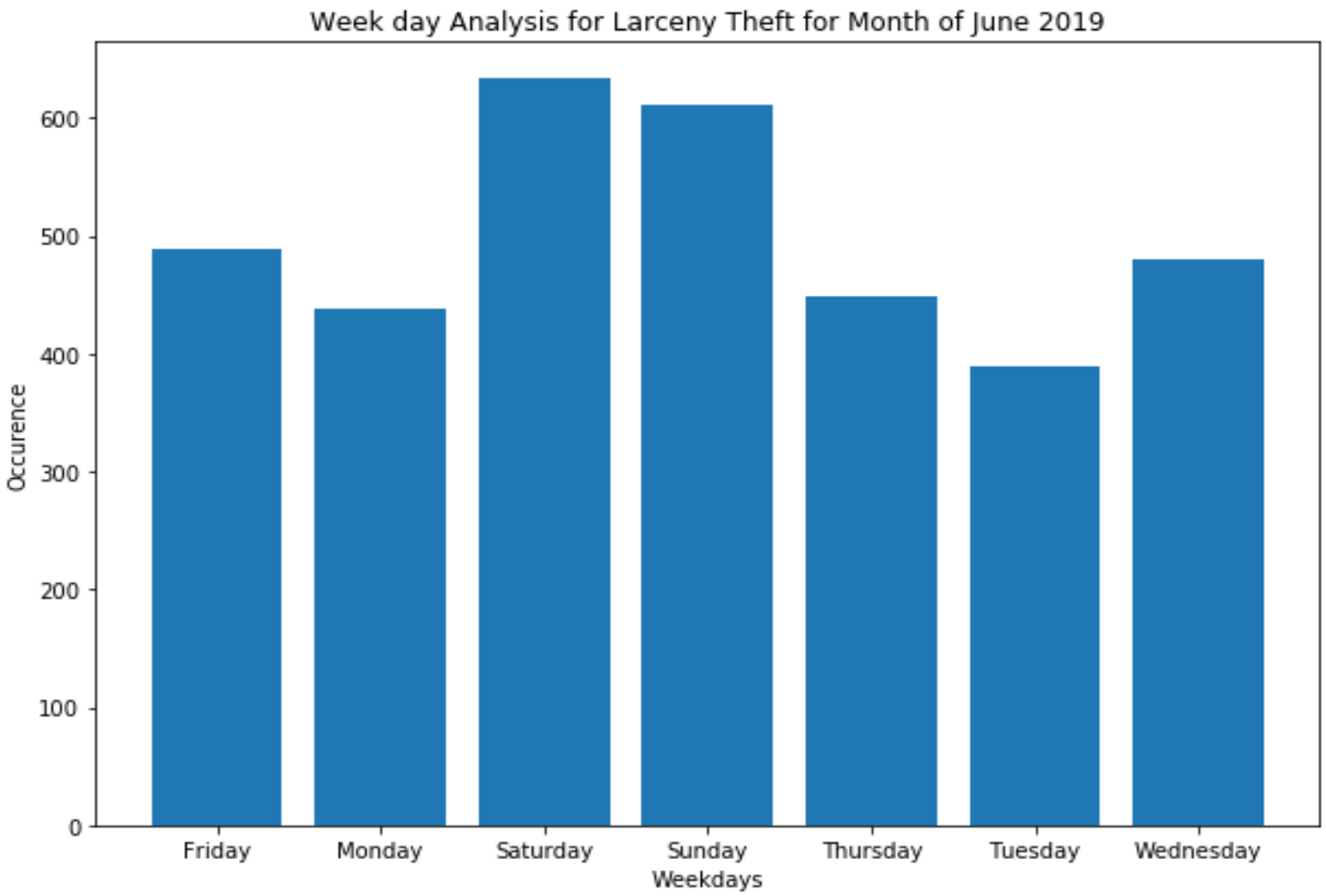
Data Acquisition and Cleaning

- San Francisco Crime report downloaded from SFPD Site
- Data for June 2019 selected to speed computation
- Formatting changed for Time and Week day for Modelling
- Focused study on Larceny Theft due to its highest count
- Venues Listing and Category sourced from Four Square
- Web scrapping used to format Json file to Dataframe

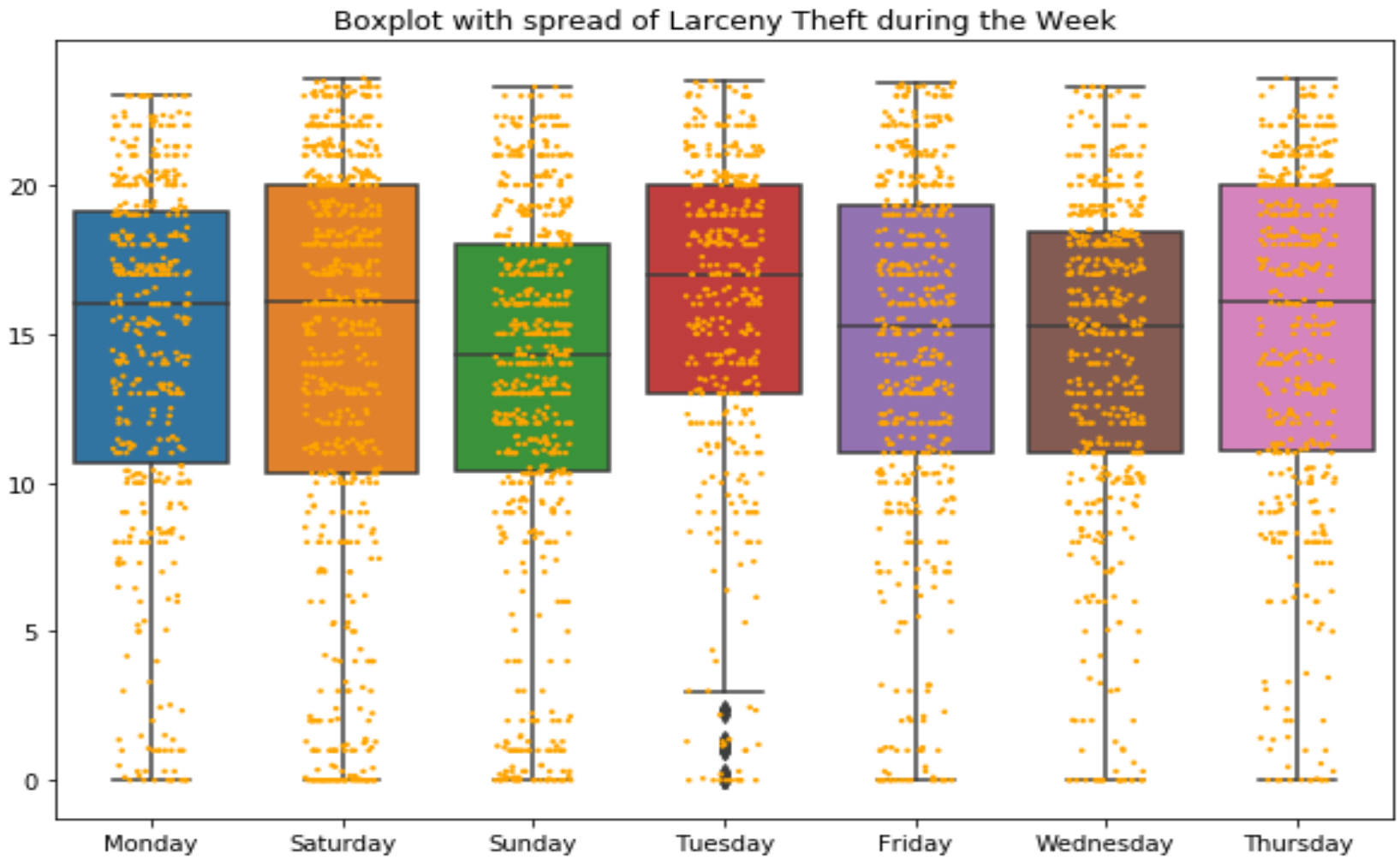
Analysis of Top Crime in San-Francisco



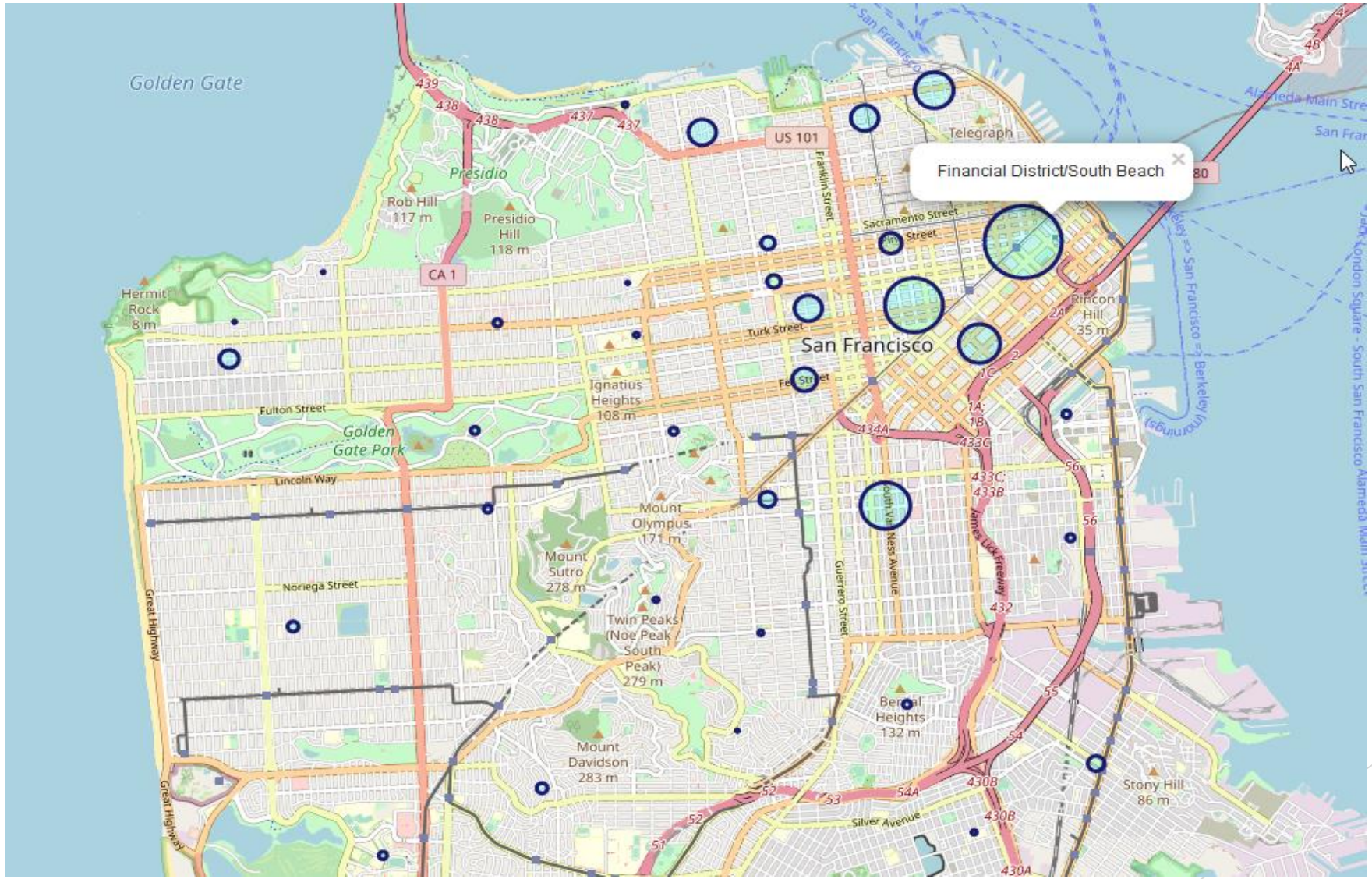
Peak Day for Larceny Crime in San-Francisco



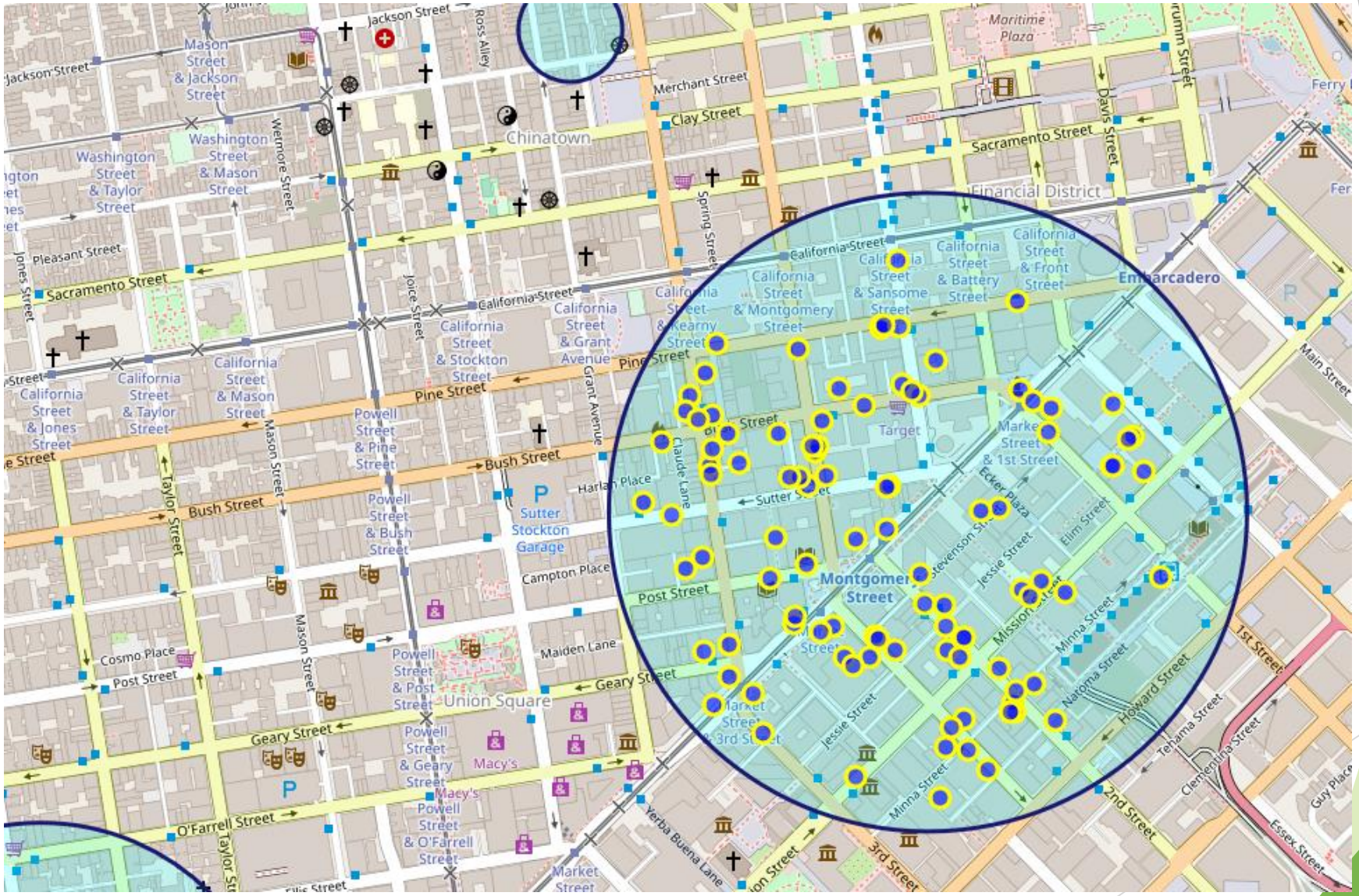
Distribution during Week day for Larceny Crime



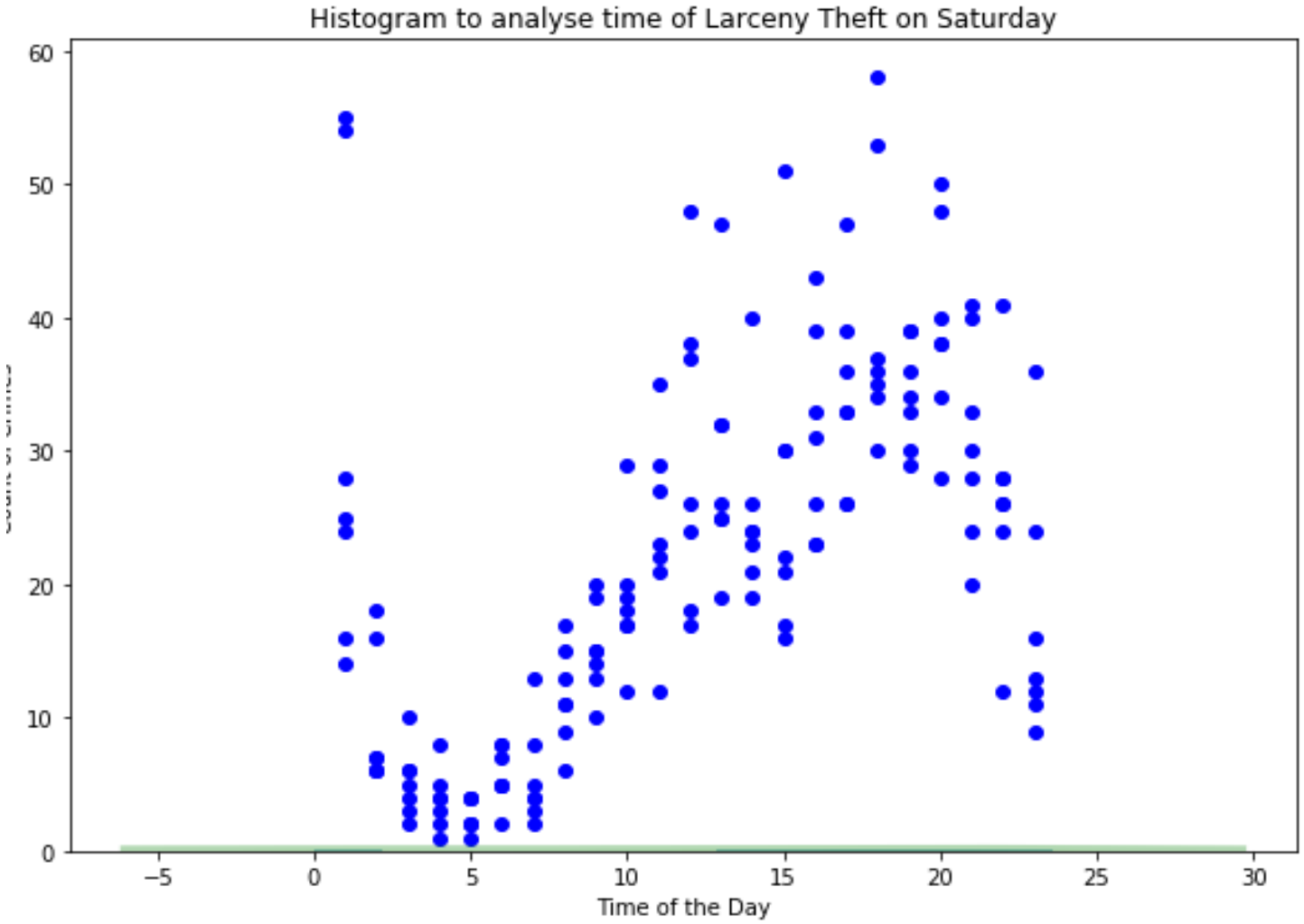
Distribution of Larceny Crime in San-Francisco



Spread of Venues for Financial District/South Beach

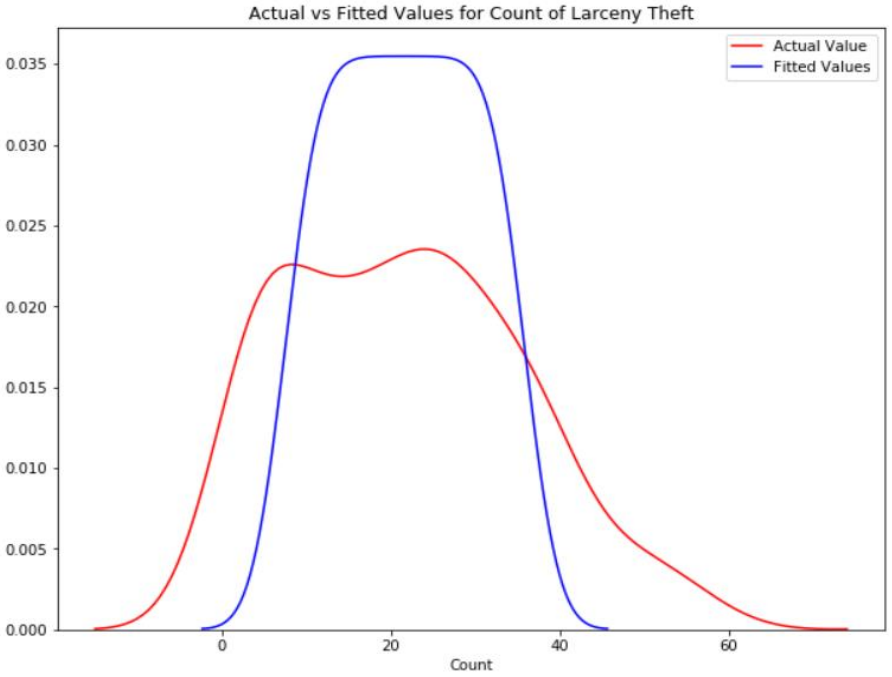


Non – Linear Relationship between Crime and Time



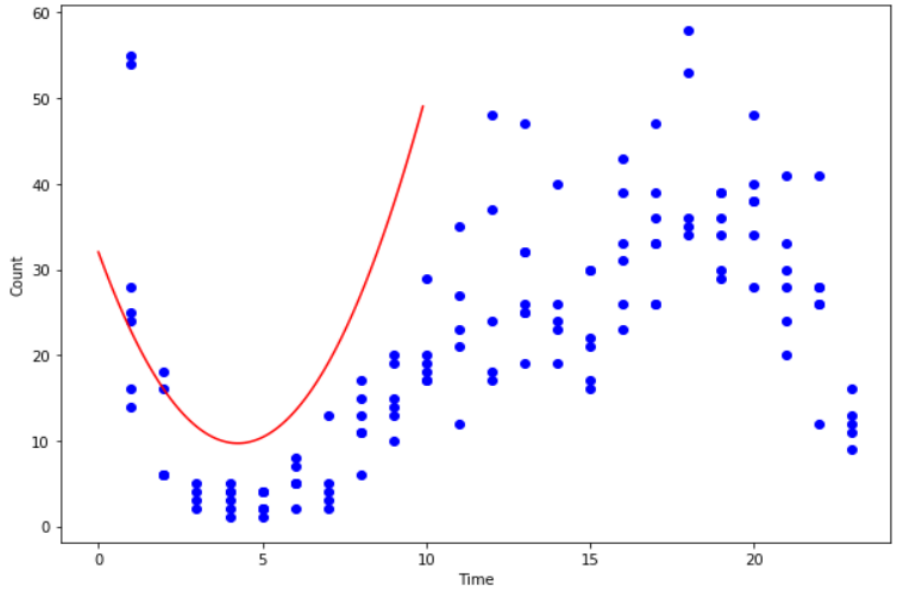
Linear Regression

Intercept is: 5.55
Coefficient is:[0.35403727 1.22543761]
R2-score: 0.35



Polynomial Testing

Intercept: [32.02728127]
Mean absolute error: 6.65
Residual sum of squares (MSE): 76.12
R2-score: 0.31



Conclusion and Future Direction

Conclusion

- Larceny Theft” highest recorded crime in San-Francisco
- Trend exist between Crime count on Saturday and Sunday
- Crimes happen towards late evening to early morning on Weekends.
- Non Linear Relationship exist between Crime and Time of Crime

Future

- Predictive model to be built Non Linear Modelling techniques
- Dashboard to analyse Venues most effected by the crime
- Clustering approach for type of venue and Count of crimes