SAN FRANCISCO CRIME PREDICTOR

INTRODUCTION:
- Analyse San Francisco crime data for last 12 years
- Visualize trends in crime
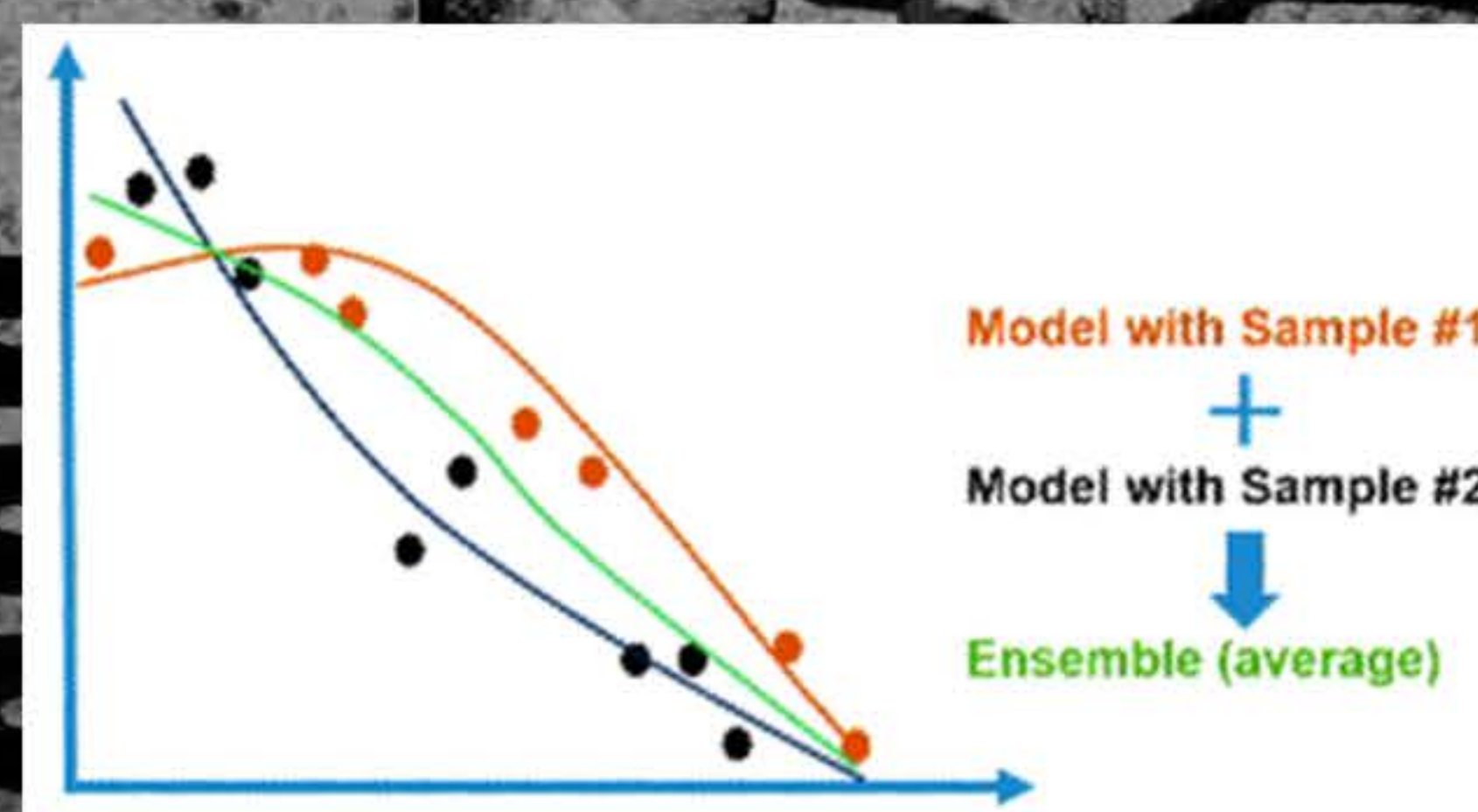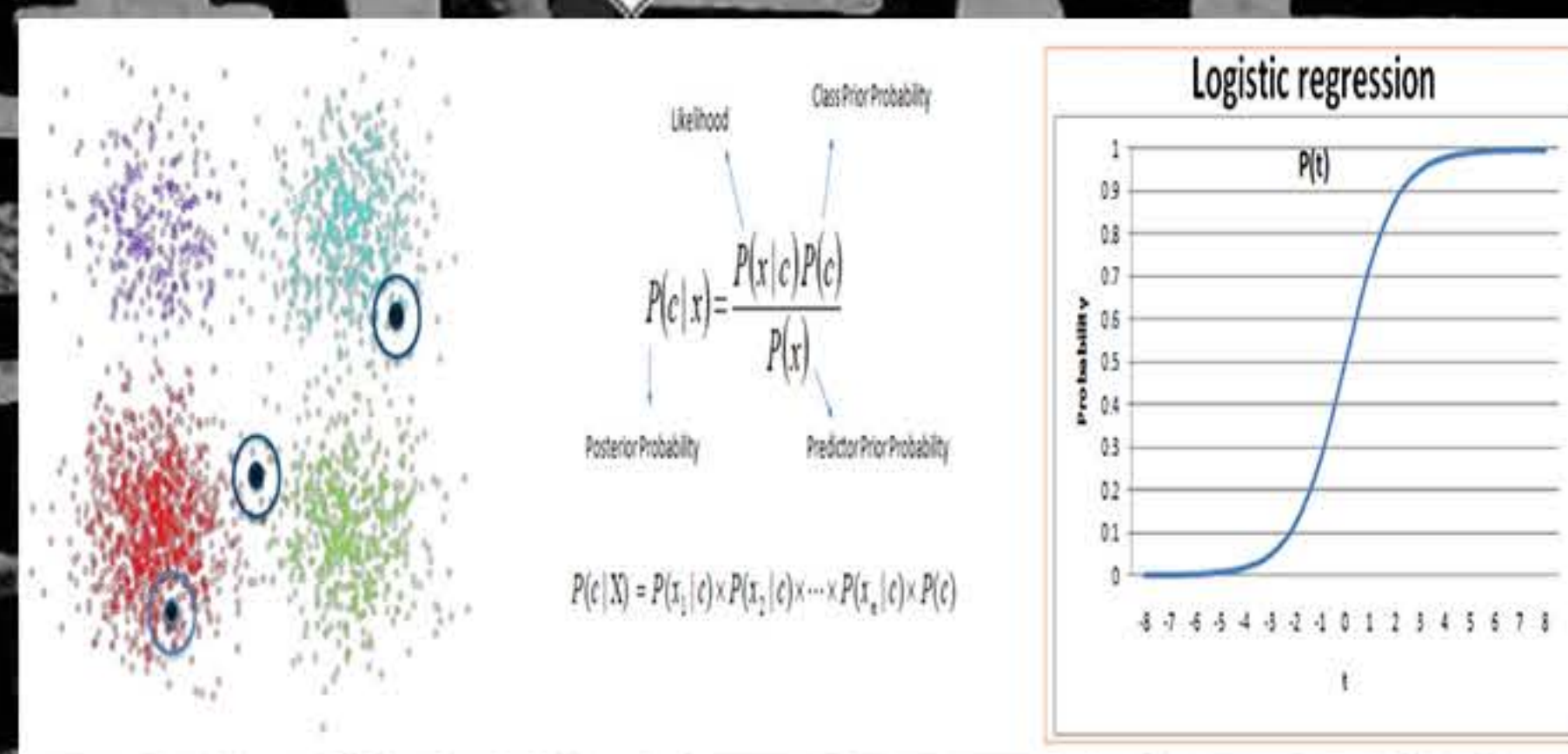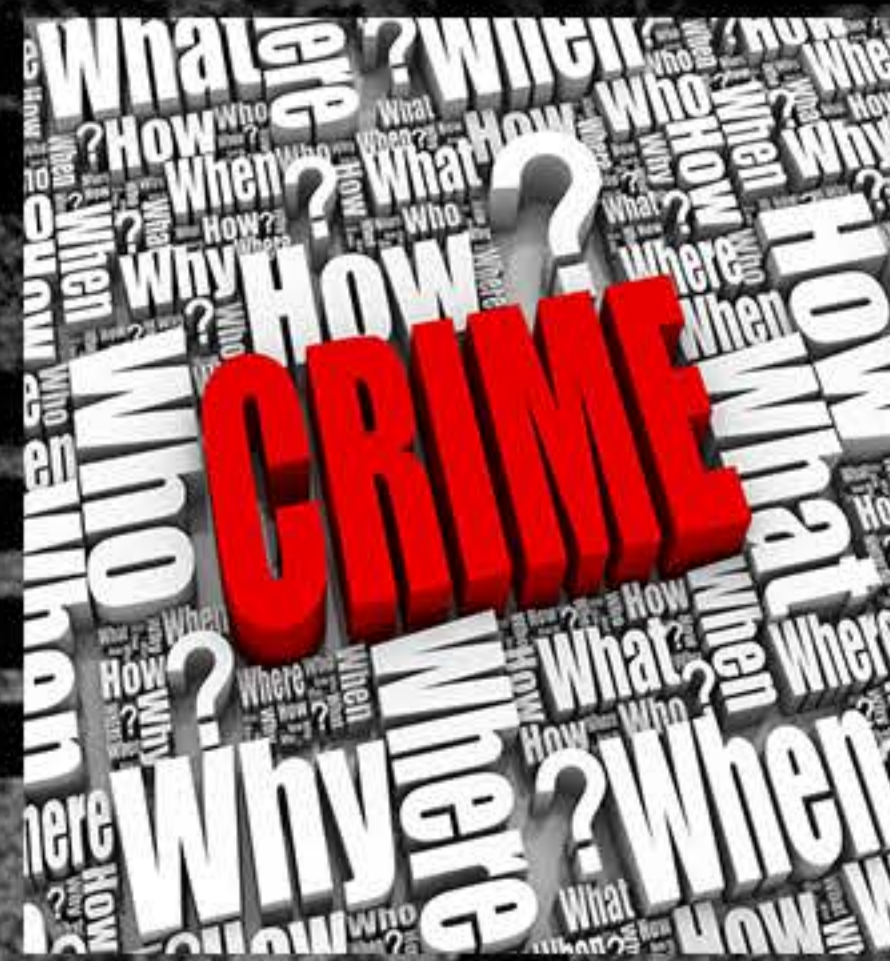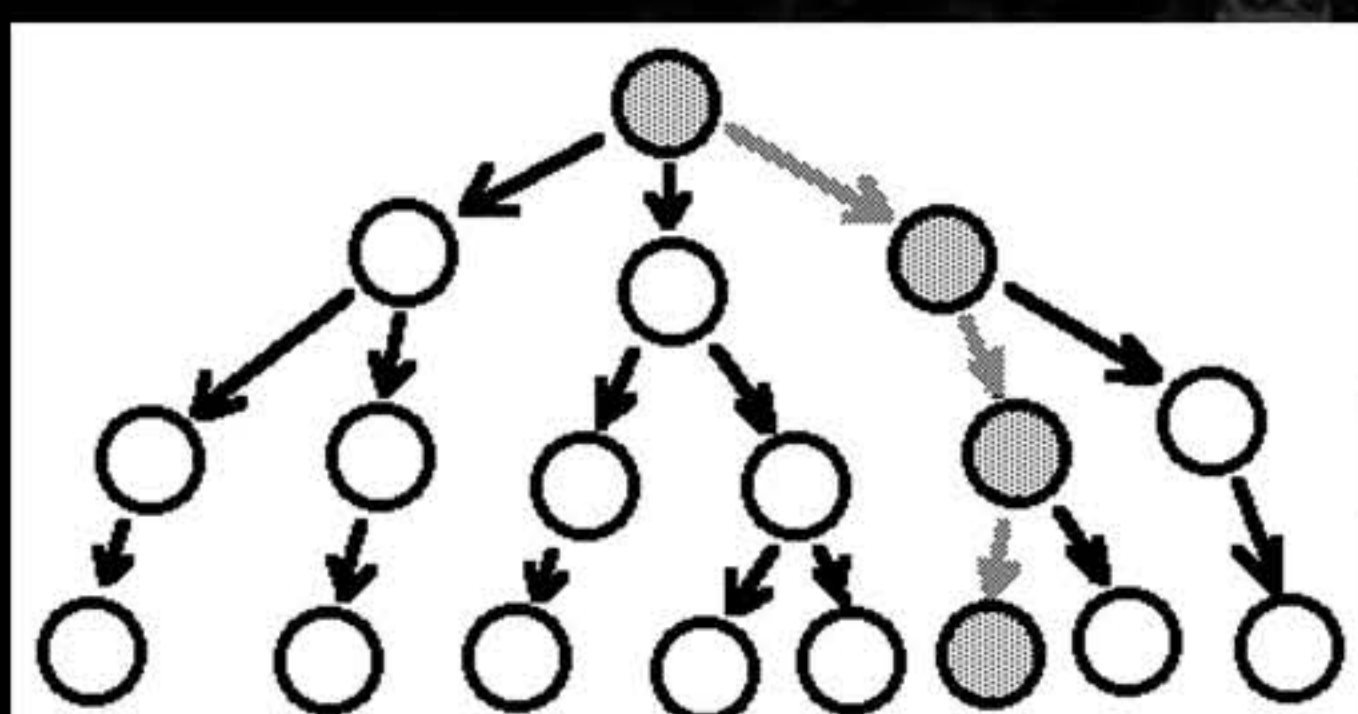- Predict category and probablity of crime

MOTIVATION:
- Help SFPD to be better prepared with resources
- Make civilians aware of safe areas in the city

DATA:
- Kaggle Competition, SF Open Data website
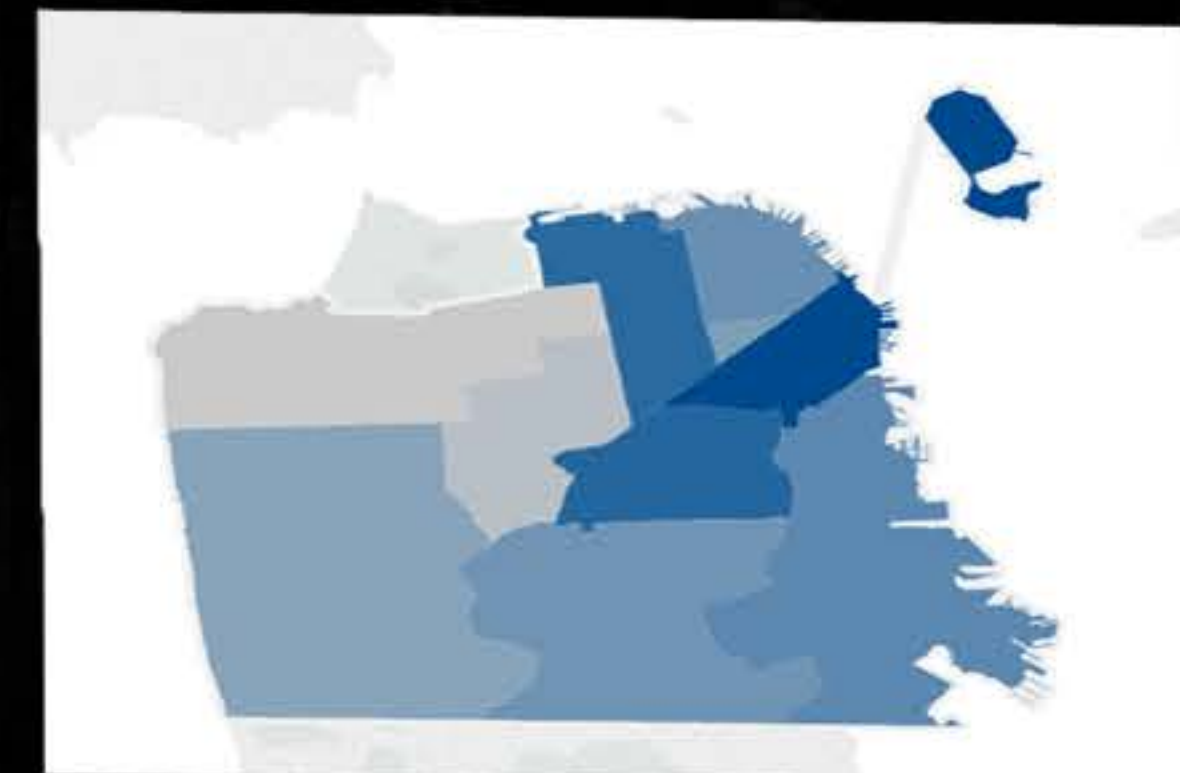- Over 800k rows of crime data from past 12 years

PREDICTION:
- Given the time and location of crime, the model predicts the type o f crime committed.

- Tested the dataset on 11 different models including Binary Naive Bayes (with & without adaboost), decision trees, Random Forest and K-Nearest Neighbor.

- We finally decided to go for Ensemble Modelling as it produced better results comapred to all indivudual predictive modeling algorithms.
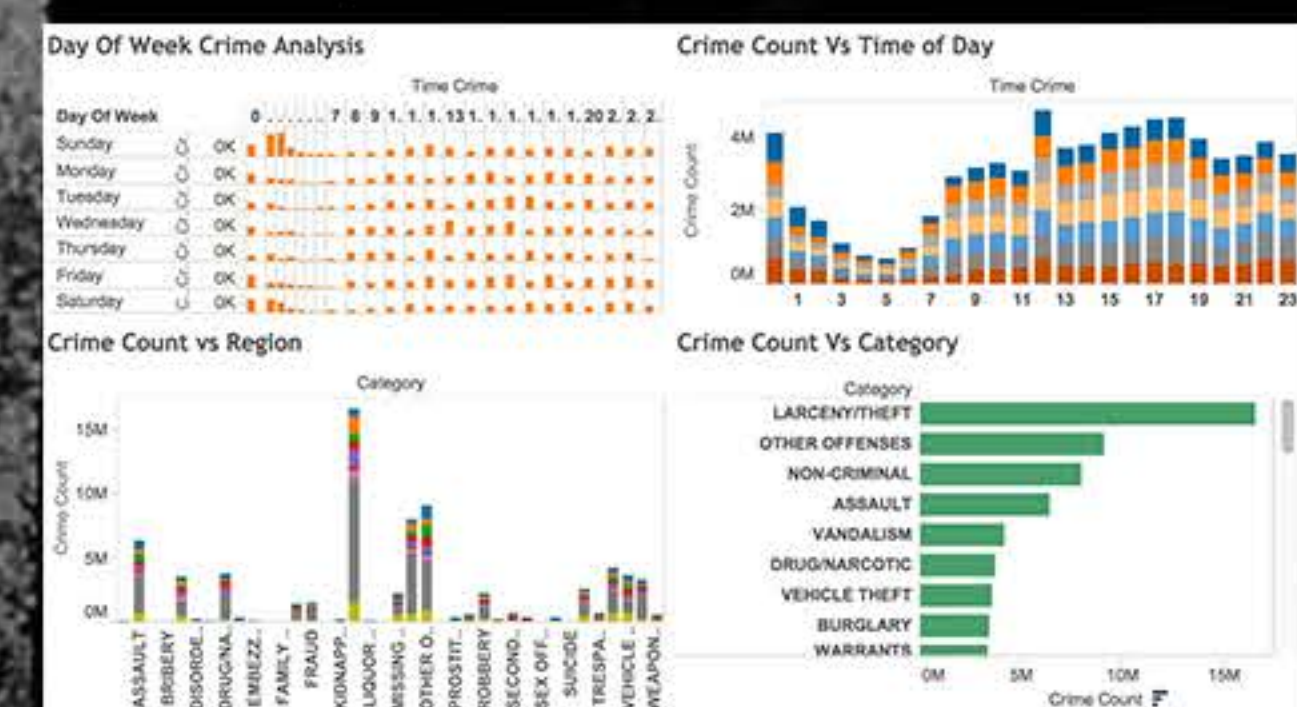
VISUALIZATION:
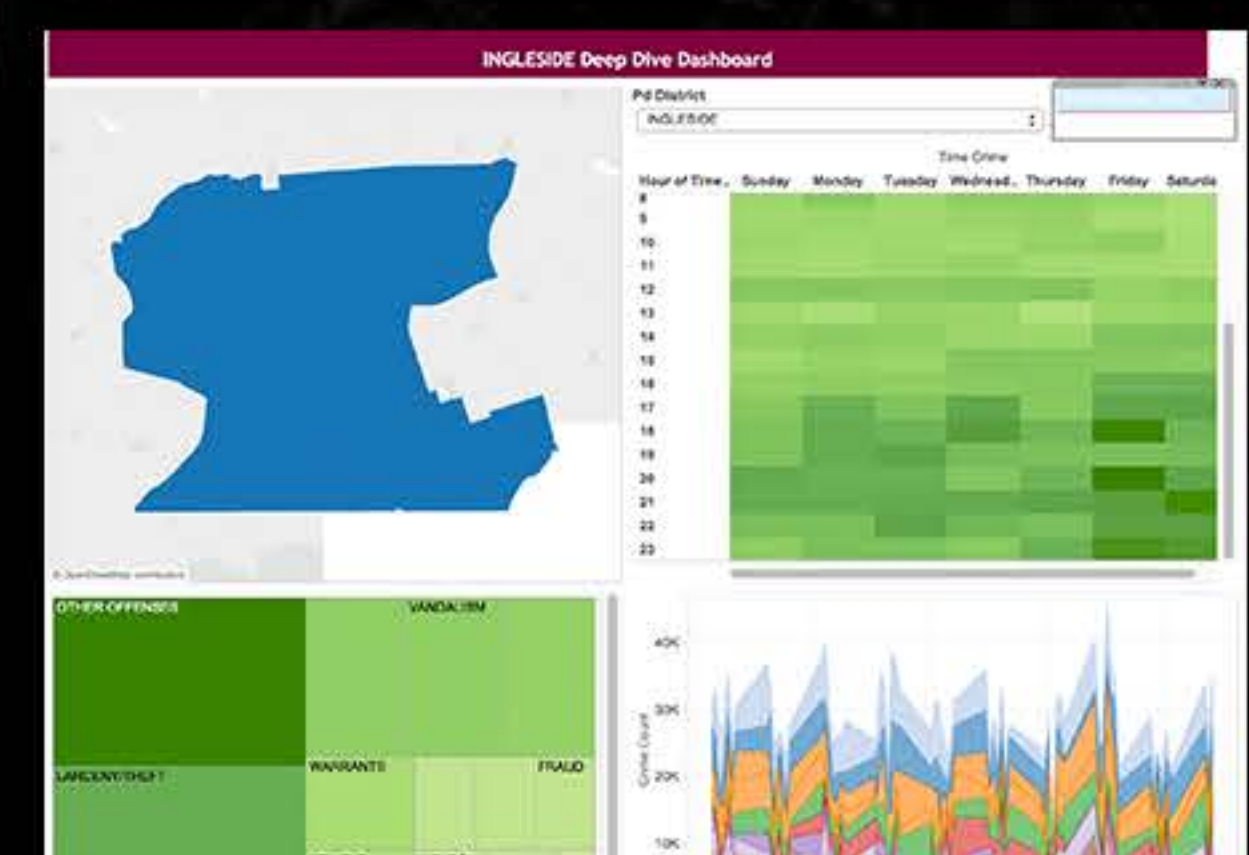- Interactive visualization using Tableau to help SFPD explore trends in crime

- The deep-dive system can be used to track evolving crime trends at District Level.

- SFPD can use the dashboard to derive meaningful insights for taking corrective measures to prevent future crimes

- List of visualizations include bar charts, tree maps, area charts and chloropleths with filters like type, place and time of crime.

Logistic regression

$$P(c|x) = \frac{P(x|c)P(c)}{P(x)}$$

$$P(c|X) = P(x_1|c) \times P(x_2|c) \times \cdots \times P(x_n|c) \times P(c)$$

Model with Sample #1
+
Model with Sample #2
→
Ensemble (average)

PRESENTED BY: Amit Panda, Murtaza Dhuliawala, Pulkit Sachdeva, Niranjan Upreti & Vivek George