



# CRIME PREDICTOR



## Team Members:

Amit Panda, Murtaza Dhuliawala, Niranjana Uprethi, Pulkit Sachdeva, Vivek George



# THE SYSTEM

- Analyze the crime data for the city of San Francisco over the last 12 years.
- Visualize the trends in crime by rate of incidents, type of crimes in the different locations.
- Build a model which, given time and place of crime, can predict the category of crime that has occurred as also the probability of a crime occurring in a given location at a given time.



# THE BIG QUESTION !!!



- This system is being built with an aim to provide a tool to the police departments, to better prepare for a criminal incident.
- Officers can ask more relevant and pointed questions on receiving a 911 call if they have an idea of what crime may have occurred rather than relying on their gut feeling.
- The interactive visualization would help administrators and police departments explore and understand trends in crime over the last twelve years.

# MAKE A DIFFERENCE

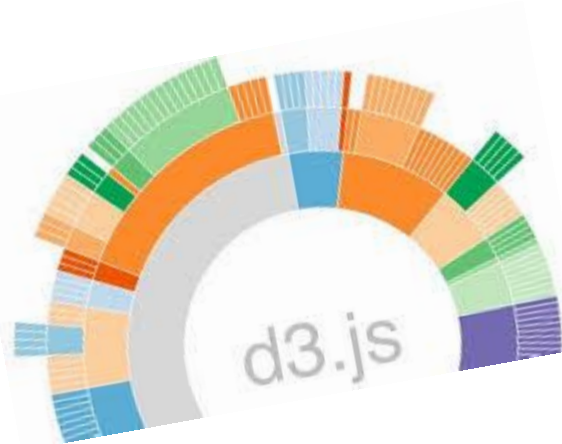


- Predicting type of crime would help SFPD to be better prepared. To plan and have the resources available to deal with the crime.
- The system could be used by civilians to get an understanding of areas that would be safer to live in, or areas that are unsafe to visit or pass-by at certain times.
- Given a source of data, the system could be expanded to cover any major city in the world.
- Other metrics like population density, income of people in an area, office buildings in a location could be used to draw more inferences on crime data.



# THE RISKS AND THE COSTS

- The risk is that there might be no relation between the area a crime happens in, the time it happens at and the kind of crime that happens, and the prediction accuracy might be very low.
- The system will be built on open source technologies like D3, python, R. The data is clean and freely available on the SF Open Data and Kaggle website.



kaggle™

# THE PLAN — GOALS FOR THE NEXT MONTH

- Approximately two months completion time
- Design visualizations and start implementation
- Exploring different classification techniques
- Start learning about various validation techniques
- Test if other data sources (e.g. median income statistics from census data) improve classification accuracy
- Design an application that works with real time data



