

Research Project Proposal

Rishika Katamneni, Nabil Salehiyan, Leslie Malave
EPPS 6323 Knowledge Mining Project Proposal
February 21st, 2023

Area/Topic:

We want to analyze San Francisco 911 calls for service to determine if there is a discrepancy between priority given to neighborhoods based on median household income, despite an important emergency.

Data:

We will be using the dataset “Law Enforcement Dispatched Calls for Service Real Time” found at data.gov. This is a continuously updating dataset from the City of San Francisco which began on December 3rd, 2021. To obtain demographic and income data for specific San Francisco neighborhoods, we will be using city-data.com from 2019.

Research Statement:

America has an apparent problem with its criminal justice system. Our law enforcement has marginalized certain groups of people based on socioeconomic status not just regarding proper situation handling but how fast the first responders show up to the scene. The question we want to answer using 911 emergency data from San Francisco is if there is a pattern or trend surrounding these calls and their responses. Does socioeconomic status determine the speed at which first responders show up at the scene?

Being able to answer these questions will shed further light on the disparities that marginalized groups face in America, in hopes to find better solutions. The bias that officers have is not as inherent as people perceive it to be, research rather shows that these biases are more implicit and not self-actualized by law enforcement. If we are able to show this bias through knowledge-mining methods, these analyses could help bring more attention to the problems faced within our justice system.

Method:

We will use a supervised logistic regression method with our output variable being onscene_datetime. Our input variables will be received_datetime, dispatch_datetime, enroute_datetime, intersection_name, analysis_neighborhood, and police_district. We will also add input variables of median household income and majority race for each neighborhood. The data from data.gov consists of recorded logs of these 911 calls from San Francisco. The goal of the knowledge mining method will be to extract the hidden pattern behind the data to see if there is a correlation between these variables.

References:

City of San Francisco. 2021-2023. *Law Enforcement Dispatched Calls for Service Real Time*.
<https://catalog.data.gov/dataset/law-enforcement-dispatched-calls-for-service-real-time>