Optimizing Baltimore Police Dispatch Locations
Tom Hervieu, Jason Kolbush, Chris Kwan

Introduction:

The city of Baltimore has one of the highest crime rates in the country. It is well above the national average with over 300 murders in 2015. About 0.4% of the entire population is directly involved or impacted by a violent crime every year. Many of the crimes occur in drug clusters located around poor areas of the city. We wonder if there is a way to help prevent crime from happening in a faster span of time. Even seconds matter when it comes to violent crime. Thus, the goal of our project is to determine the best dispatch locations for police officers to be in order to arrive at the crime locations as quickly as possible.

Data Set: https://catalog.data.gov/dataset/calls-for-service

Through data.gov, we were able to find a "911 Calls for Service" from data.baltimorecity.gov, which contains many different data collections for the city of Baltimore. This dataset contains 2016 data for every 911 call made prior to its location (i.e. latitude and longitude coordinates). It also indicates the level of urgency of the call, which could potentially benefit the police in the future. The entire dataset contains two million rows for 2015-2017. The data set is contained in a CSV file containing the date, priority, description, and location in three different forms: district, street #, and longitude/latitude coordinates. We will most likely filter out the "Non-Emergency" and "Low" Priority data points, and we will also eliminate data points entailing accidents in order to focus on more strictly on crime-prevention.

Machine Learning Problem:

This is a K-Clustering problem. We will be displaying the ideal dispatch locations according to our analysis of the 2016 data on a map of Baltimore. The optimal dispatch locations will be represented by the cluster centers that we create. An issue that may arise is that we may not be able to create as many cluster centers as there are dispatch locations in the actual city of Baltimore. However, we can pinpoint optimal locations for some dispatch units, as well as police stations.

Algorithm/Method Applied:

We will apply the K-Means Algorithm and create as many clusters as possible to create a reasonable, specific distribution of cluster centers (dispatch locations) across Baltimore- relative to the distribution of crime. Due to the fact that dispatch units are driving cars on the streets, we can view the path of travel as a grid. Therefore, we will be using Manhattan distance when allocating each data point to its designated cluster. Due to the size of the dataset, we may choose to run our algorithm until our cluster centers are near convergence with a certain margin of error. Our visual will comprise of a map of Baltimore, with markers at the center of a specific cluster of crimes.