Casey Martin, Domas Budrys, and Faith Sang - Project Proposal

For our project, we propose to analyze the following dataset:

<https://www.kaggle.com/sohier/baltimore-911-calls>

This dataset is large, with 2.8 million entries, and is well structured but its information is rather raw. There isn’t any pre-processing done to this dataset, such as averages or standard deviations. The dataset details 911 calls to operators in Baltimore and contains the following information:

Date, Time, District, Reason, Street Address, and GPS Location

We intend to answer questions that might inform authorities what areas they should focus on to either improve their 911 services or reduce wasted time for 911 operators. For example, preliminary looks at the data set shows many non-emergency calls go to 911 that could be referred to the local non-emergency police number. We also want to know what the hot-spots of 911 calls are so we can better inform our officers where they should be patrolling and where police precincts should be located. Finally, we would like to know if there is a any seasonal trends in 911 calls to determine if there any strategy changes are in order during different seasons. We have also included a question about weird and quirky things 911 operators respond to in case we find some strange items in our datasets during our research.

We have broken these into four Big Questions and have three to five smaller supporting questions for each one. We think this approach is best because the Big Questions represent what we really want to know, but the “Possibilities” of how to answer the big questions actually break the questions down into more actionable pieces.

Big Question 1:

How are the 911 calls geographically distributed?

Possibilities: Plotting GPS coordinates on a map overlay of Baltimore (heatmap?)

Determine deviation from central average (how spread out)

Relate call density to population density (external data required)

What are the most dangerous streets and why?

Big Question 2:

How is 911 generally being used by the populace?

Possibilities: What are the most common reasons for calling 911?

How many non-emergency/no voice calls?

Whats the distribution of priority? (Non-emergency, low, medium, high)

How many man hours are wasted by non-emergency 911 calls?

Mention that no-voice calls are medium priority and why that is a problem.

Big Question 3:

How does the use of 911 change over time? (throughout the year/throughout the day/throughout the week)

Possibilities: Is the frequency of 911 calls seasonal? (more during the summer than winter?)

Has the frequency changed over the past three years?

Has the type of call changed over the years?

Big Question 4:

What are some of the weird and interesting things that 911 operators deal with daily?

Possibilities: Top 5 non-emergency calls

Frequency of low priority calls by children

Other stupid things people call 911 for