Project Proposal

due October 11, 2021 by 11:59 PM

Maggie Lundberg, Isabella Kjaerulff, and Riya Mohan

Load Packages

```
library(tidyverse)
```

Load Data

Introduction and Data, including Research Questions

Our data is provided by NEMSIS, the National EMS Information System, and was collected as an effort to help standardize and improve EMS databases. From this data, we ask, what factors influence somebody's process of emergency care in an EMS call?

To answer our overarching research question, we plan to center our analysis around these questions: Is there a correlation between drug use and observed quality of care? Is there a correlation between race and observed quality of care? Are there biases present in the EMS quality of care sample size? What is the relationship between symptoms and vital sign ranges? Is race correlated to technicians' assessment of GCS?

Variables of interest include Primary Symptom, Other Associated Symptoms, Cause of Injury, Barriers to Patient Care, Alcohol/Drug Use Indicators, Protocols Used, Medication Given, Procedure, Procedure Successful/Response to Procedure, Suspected EMS Work Related Exposure/Injury/Death, Race, Initial Patient Acuity vs Final Patient Acuity, and Vital Signs–Blood Pressure, Pulse Oximetry, Respiratory Rate, O2 Saturation, Blood Glucose, Glasgow Coma Score, and Level of Responsiveness.

Glimpse

(Please use `glimpse` for your data, uploaded into the data folder, here.)

Data Analysis Plan

As a part of our data analysis, a few of the specific predictors we are looking at are race, symptoms, cause of injury, BGL, and Alcohol/Drug Use Indicators. As a result, the outcomes we will be looking at are primarily related to the quality of care, which we will be able to determine from Response Time, GCS, Barriers to Patient Care, and Vital Signs. Furthermore, in order to address our research questions we will be using race, cause of injury, and symptoms as comparison groups.

In addition, we have identified a few vital statistical methods needed for analysis. These include p-value/t-test (null hypothesis), standard deviation for the symptoms/conditions as related to different vital signs, and conditional probability. As a result, we expect to see a correlation between race and quality of care, blood glucose level, alcohol/drug use indicators, symptoms, and vital signs. We also expect that the null hypothesis can be rejected ($P < 0.05$) and a standard distribution of vitals will be related to their associated symptoms. Various conditional probabilities are expected to bolster our expected conclusion as well

(Specify the outcome (response, Y) and predictor (explanatory, X) variables you will use to answer your question, as well as the comparison groups you will use, if applicable. You may include very preliminary exploratory data analysis, including some summary statistics and visualizations, along with some explanation on how they help you learn more about your data. Note the statistical method(s) that you believe will be useful in answering your question(s). What results from these specific statistical methods are needed to support your hypothesized answer?)