DUE: \*Monday\* 10/2/23 Noon (Report and Slides)

Instructions:

Outline a modeling approach for the estimated travel time data. Discuss how to

model dependence of the estimated travel times on scenario and other covariates (e.g. summer season, hour of the day, system load), including event/call ID (as each scenario is assessed on each call) and describe how to use this model to compare the scenarios. As we have seen with the mock data, there will likely be numerous events where a load metric does not vary by scenario. Approach this by modeling

* ~~(1) the probability of there being variation/no variation across scenarios (logistic regression if there are covariates like season and time) and~~
* ~~(2) the dependence of the travel times on scenario given that there is variation (using a linear mixed model).~~
* ~~Add to Overleaf doc (latex)~~
* ~~Add a few (<=5) slides to the Google slides document summarizing your section of the report (see below). Link to Google Doc:~~ [~~https://docs.google.com/presentation/d/1yZhci99kUhmjBbrTIK3N\_VziEF0AQ9VllpOp4X3A\_kA/edit?usp=sharing~~](https://docs.google.com/presentation/d/1yZhci99kUhmjBbrTIK3N_VziEF0AQ9VllpOp4X3A_kA/edit?usp=sharing)

Note:

* Covariates: season, hour of the day, day of the week, priority, system load, etc.

Paragraph 1:

* ~~Discuss travel time differ-by-scenario-or-not binary outcome, calculate probability~~
* ~~Model binary outcome with logistic regression model~~
* ~~glm(differ ~ covariates)~~

Ambulance travel time, a significant component of response time, is the metric you aim to optimize by comparing the average travel time across various ambulance distribution scenarios. The initial step in modeling travel time is checking whether the estimated travel times exhibit variation across different scenarios, which enables the identification of instances where travel time is influenced by the scenario under consideration, and those where it remains the same. A binary outcome, hereby referred to as “differ,” is formulated, wherein “1” denotes a variation in travel time across scenarios, and “0” indicates an absence of variation. This differentiation is modeled using a logistic regression model, specifically utilizing the glm() function in R, due to the binary nature of the outcome. The response variable is the binary outcome “differ,” and the predictor variables, or covariates, encompass factors such as season, hour of the day, day of the week, priority, and system load. This model calculates the probability of observing a variation in travel times across scenarios, contingent upon the covariates. For instance, during peak summer seasons or specific hours of the day, the model may unveil a heightened probability of variation in travel times across various scenarios, potentially attributable to factors such as escalated traffic or road closures.

Paragraph 2:

* ~~Discuss travel time given differ-by-scenario, model with linear mixed model using R package lme4::lmer()~~
* ~~lmer(tt ~ scenario + (1 | eventID) + covariates)~~

Upon recognizing instances where travel times show discernible variations across different scenarios (i.e. when “differ” is predicted to be 1 in the logistic model mentioned in the last paragraph), it becomes vital to model these times, particularly focusing on their dependency on the respective scenarios, given the existence of such variations. The Linear Mixed Model (LMM), especially employing the lme4::lmer() function in R, stands out as a robust tool for this endeavor (Bates et al., 2015). In this model, travel time is designated as the response variable you want to predict; the scenario operates as a fixed effect, symbolizing the specific ambulance allocation strategy being examined and quantifying the impact of various deployment strategies on travel time; a term based on event ID should also incorporated to capture the random effect, allowing each event to have its unique baseline travel time, thereby addressing the unobserved heterogeneity and inherent correlations within each event’s measurements; and covariates, which include additional fixed effects like time of day, season, and system load, as previously mentioned, offer a mechanism to control and adjust for these variables in the model. Regarding the model’s interpretation, since the primary focus is on the impact of ambulance distribution on travel time, you should identify the baseline scenario and examine the offset term for each scenario in the model output. The scenario presenting the smallest or most negative offset term corresponds to the shortest travel time predicted by this linear mixed-effect model and would be the scenario of interest.