Laura E. Barnes
Julianne Quinn APMA 6430

Homework 1 & 2
Describing the Current Rail Accident Situation

Assignment: In this homework, you will demonstrate that you can use evidence informed systems engineering to describe the current safety situation of rail accidents. Use at least four graphics to describe the current situation. After each graphic, provide a description of your observations. Use your graphics to generate preliminary hypotheses that you could potentially investigate to look at contributors to the severity of rail accidents using our data set. Turn in both the *RMD* and *PDF* version of your R Markdown Notebook.

Assignment Objective: In this assignment, you will demonstrate your ability to use data visualization techniques to describe the current raila ccident situation over the last 18 years and generate preliminary formulate actionable hypotheses about relationships between train accident characteristics and their severity that could be used to inform FRA policies.

Data: The data in RailAccidentsXX.csv provide information on historical train accidents between 2001-2019.

Instructions:

1. You may discuss this homework with other students in the class. However, the work you turn in must be your own, and you must reference any contributions from others. You must pledge your submission.

2. Use the assignment page on the Collab site for submission.

3. Perform the analyses described below using the techniques described in class to provide a visual picture of the current rail accident safety situation.

   visualize relationships and formulate hypotheses about the severity of train accidents. For this assignment, you will turn in detailed R code (that instructors should be able to run) with comments responding to the questions below. You can alternatively turn in a detailed R Markdown notebook in both **Rmd** and **PDF** formats with your analysis and responses to the below questions.

Assignment: Utilize the visualization methods learned in class to develop hypotheses about what actions the FRA could take to reduce the severity of rail accidents.

1. **Visualizing & Selecting Severity Metrics** (10 points)

   (a) Identify at least 6 severity metrics the FRA should consider in evaluating their safety regulations

   (b) From these 6 severity metrics, identify 2 metrics to describe accident severity. These can be chosen by down-selecting from the metrics and/or by combining 2 metrics to form a new one. Use principal component analysis to motivate your metric choices.

2. **Summarize the Current Safety Situation** (25 points)

   (a) Describe the current accident situation in terms of both frequency of accidents and severity of accidents from 2001-2019. Use the severity metrics identified in Part 1. Use visualization and summary statistics to explain why the analysis should focus on only the few extreme accidents.

   (b) Obtain an understanding of the contributors to accident severity in your extreme accidents data using the multivariate visualization techniques discussed in class. This analysis should explore both quantitative and categorical variables to describe the severe accidents, identify predictors, and inform actionable hypotheses.

3. **Generating Hypotheses** (15 points)

   (a) From your analysis in Part 2, generate at least 2 well-formed, actionable hypotheses for each severity metric. Explain why the hypotheses are actionable demonstrate how you arrived at each hypothesis. Write out your null and alternative hypotheses.

BONUS (5 points) if you also find references (i.e., news articles, reports from FRA) to support hypotheses