**Data Programming in R: Group Proposal**

Group Members: Norean Gardner, Melanie Adams, Jodie Carlson, Sam Etten, Brian Rolf

**Situation:** Bicycle Fatalities on a National Scale

**Data:** 2017 Crash Data from National Highway Transportation Safety Administration (NHTSA).

The NHTSA utilizes data from the Fatality Analysis Reporting System (FARS) to help in understanding the drivers of motor vehicle crashes to be able to improve traffic safety. The data includes information on accidents involving pedestrians, bicyclists, and people. It includes pre-crash actions and location of accidents relative to crosswalks and school zones.

Data is published in a zipped file located on the following website:

General website: <https://www-fars.nhtsa.dot.gov/Main/index.aspx>

File Transfer access: <https://www.nhtsa.gov/node/97996/251>

The data from 2017 (the most recent available) for the US will be downloaded as a csv file and loaded into R. We will filter the data to focus specifically on fatalities for bicyclists. We will utilize the user guide located at the following link to help in understanding our dataset so that we leverage the data appropriately in our analysis: <https://www.nhtsa.gov/filebrowser/download/119221>.

**Question:** What attributes such as time of day, location etc. lead to biking fatalities? Such as travelling with or against traffic? Was the appropriate gear such as a helmet being worn?

**Hypothesis:** If bicyclists are travelling during rush hour in urban areas in the evening then they are at a higher risk of fatal accidents.