**Applied Data Science Capstone Project**

**Prediction of Car Accidents Severity**

**Introduction/Business Problem**

The primary audience for this data science project is the unit responsible for safety analysis within Seattle Department of Transportation. They will be interested in achieving the highest levels of safety performance for the general public through monitoring and controlling the rate of accidents in Seattle. For this to be achieved, data analysis has to applied for the data set in hand, a predictive model for accidents severity to be developed. Through this model, factors leading to the most severe accidents, as a priority, can be identified and controlled through targeted mitigation actions at the root causes.

**Data**

The data that will be used for this project will be provided from Seattle Department of Transport, traffic management division. The data is a historical data set from 2004 to present on collisions, with the size of 194673 observations (rows) and 38 attributes (columns).

Among the 38 attributes within the data set is information about accidents severity (SEVERITYCODE attribute) which our model will predict with the highest possible level of accuracy as a function of the selected features from the data set.

A link providing description to the attributes of data set can be found [here](https://s3.us.cloud-object-storage.appdomain.cloud/cf-courses-data/CognitiveClass/DP0701EN/version-2/Metadata.pdf).