**Capstone Project - Car accident severity**

**Introduction**

For the final Capstone Project in the IBM Certificate Course, Coursera has provided a set of traffic accident related data to base this final project on. The data was collected by the Seattle Police Department, recorded by Traffic Records, and provided by Coursera via a download link. The time period for this data is from 2004- present and contains such information as severity, location, collision type, weather conditions, road conditions, and light conditions, among others.

**Data**

The data was collected by the Seattle Police Department and Accident Traffic Records Department from 2004 to present. The data consists of 37 independent variables and 194,673 rows. The dependent variable, “SEVERITYCODE”, contains numbers that correspond to different levels of severity caused by an accident from 0 to 3.

A code that corresponds to the severity of the collision:

• 3—fatality

• 2b—serious injury

• 2—injury

• 1—prop damage

• 0—unknown

To accurately build a model to prevent future accidents and/or reduce their severity, we will use the following attributes — ADDRTYPE, WEATHER, ROADCOND, VEHCOUNT, PERSONCOUNT.