

# Peer-graded Assignment: Capstone Project - Car accident severity

## By Edward Scott

### Table of Contents

Introduction.....	1
The data.....	1
Methodology.....	1
Data.....	1

### Introduction

Car accidents are an inevitability of modern life, but wouldn't it be great if you could predict the severity of an accident on a given day? Factors such as weather, road condition, time of day and many others could have an impact on the severity of a collision. Knowing the likelihood of a severe accident happening could make individuals take a little extra care whilst driving, or they may even wish to seek alternative transport.

In this project I am going to try to predict the severity of a collision based on Seattle City collision data. I will create a Machine Learning model based on a training set of the cleaned data and will then will use the remaining data (test set) to evaluate the model.

### The data

The data I am using was collected between January 2004 and April 2020. There are 194,673 collision records with 1 row representing 1 collision. There are 38 columns included in the dataset, each representing a piece of information about the collision. I cannot verify the validity of the dataset as it was provided to me by Coursera.

A link to the original dataset can be found here:

<https://s3.us.cloud-object-storage.appdomain.cloud/cf-courses-data/CognitiveClass/DP0701EN/version-2/Data-Collisions.csv>

A link to the MetaData form which describes the raw data file can be found here:

<https://s3.us.cloud-object-storage.appdomain.cloud/cf-courses-data/CognitiveClass/DP0701EN/version-2/Metadata.pdf>