Capstone Project - Car accident severity

**Problem Statement**

Seattle, a city on Puget Sound in the Pacific Northwest, is surrounded by water, mountains and evergreen forests, and contains thousands of acres of parkland. Washington State’s largest city, its home to a large tech industry, with Microsoft and Amazon headquartered in its metropolitan area. The futuristic Space Needle, a 1962 World’s Fair legacy, is its most iconic landmark. The state has an excellent connection of highways. It also has a dense population of suburbs, and once you get past many kilometers of tightly knit older towns, there are suburban, pastoral and even rural areas. With that said it increases the changes of collisions due to heavy traffic and high commute. The collision dataset given in this course is what we will explore in this project. The data provides a number of factors that, collectively, or individually, can be evaluated to find out the chances of collision in a particular location, weather or road condition.

**Audience**

People who travel on a daily basis or even if someone new commuting to the highway will have an idea of the road condition and how time, weather and other scenarios while driving may cause collision. They will get awareness to take precautions

**Benefits**

it will reduce the number of collision since they will have better idea of what to expect

**Data**

The data we will be using is the collision data from the example, as well as the pdf describing the data, for reference

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

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

**Methodology**

#### Initial data will be cleaned in Excel and then using Python for additional cleaning, analyzing and visualizing state traffic accident data. The idea is to extract data to evaluate its finding, combine factors leading to collision together, and building linear regression models to evaluate the chances of collision.

To do this, we will be using the following modules/libraries

* Pandas
* Numpy
* Folium
* Geocode
* Matplotlib
* Scikit-learn