**Introduction**

Across the US, crashes each year result in thousands of lives lost, injured victims, and billions of dollars in property damage. The National Highway Traffic Safety Administration (NHTSA) uses data from many sources, and one of their goals is to reduce both human and property damage.

Many different factors are provided to the NHTSA, and not all factors are important in determining the severity of accidents. By narrowing the accident locality to a select area, we can identify certain factors contributing to accident severity. A model can be built and applied once these factors are identified, in hopes to apply the model to other locales.

**The Data**

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

We are focusing on the provided data by IBM from Seattle Department of Transportation (SDOT), timeframe from 2004 to present.

**Business**

Not all information provided to dispatch personnel is relevant to the severity of the accident. The goal is to identify and predict factors that contribute to injuries to people and the severity of the injuries. By identifying these factors, we aim to predict and minimize further serious accidents as well as giving emergency personnel the ability to prepare for the severity of the accident based on the information provided by dispatch personnel using our model.