

Objective

The objective of this Project is to apply some of the skills learned in the Course IBM Data Analytics using Python.

Business Understanding.

Several factors can contribute to the risk of collisions and can result in injury, death, property damage as well as financial cost to both society and individuals.

The dataset that will be used contains information about types of collisions from 2004 to date and is hosted by the Seattle Department of Transportation (SDOT).

The analysis that we are going to provide could be useful to the transit department and general public to prevent/avoid collisions and could be the trigger for plans to reduce this events.

The model aims to predict the severity of an accident based on variable "Severity Code" (see description on [Accident-Colissions](#)).

Source of information.

The dataset that will be used contains information about types of collisions from 2004 to date and is hosted by the Seattle Department of Transportation (SDOT) on the following link [Accident-Colissions](#).

Dataset description can be found on the following link [File Description Colissions](#)

The dataset description can be found on the following link: <https://www.kaggle.com/jonleon/seattle-sdot-collisions-data>.

Data Volume: Database contains 194,673 observations and 38 variables.