# Applied Data Science Capstone

IBM Data Science Capstone

# Lowering Accident Recurrence and Severity in Seattle, Washington

Kristopher Zapata

September 2020

# Table of Contents

# Introduction

Even though accidents happen every day and are bound to happen eventually, since it's the very nature of humans to commit mistakes/errors and some events are beyond our control, there are ways we can use Data Science to reduce accidents recurrence and the possible negative effects of an accident, such as fatalities. The Seattle Police Department, specifically the Seattle Department of Transportation, has been gathering data of all the accidents that have occurred, with a specific purpose in mind, which is improve the public's safety on the road. Therefore, the purpose of this Data Science Capstone Project is to improve the public's safety on the road in Seattle and ultimately, reduce the amount of accident's with high severity, which would lead to less fatalities.

## Business Problem

This project aim is to analyze the historic data regarding accidents in order to identify how each possible factor has an effect on accident severity and use it to develop a model that would be able to predict accident severity according to the features used to train the model in the first place. These factors/features include road conditions, light conditions, weather conditions, whether the person or people involved were under the influence of alcohol and/or other drugs, whether the driver was speeding, whether the driver was not paying attention, collision address type(crossway, intersection etc...), whether or not the pedestrian right of way was granted or not and other less influential features.

This project is destined to help the Seattle Department of Transportation in identifying certain patterns and features that lead to increased accident severity such as a dangerous crossway or certain weather conditions. After identifying these patterns or features that can have an effect on accident severity, transit authorities can draw out an action plan to accordingly address these issues to improve road safety for the public. It will help the Seattle Department of Transportation prevent accidents and lower accident severity, which in turn would lead to less fatalities which is the ultimate goal, improving road safety for the public.

## Target Audience for this Project

The main target audience for this project would be the Seattle Department of Transportation, given they're the entity responsible for improving on-the-road safety for the public in general. Additional audiences who would be interested in the results of this project would be:

- Transit authorities for other cities interested in implementing a similar project.
- General Public for their own safety
- Elected and non-elected officials in public office.

# Data

In order to provide an answer to the business problem explained above, we'll require the database to include the following features:

- Location of the collision.
- Collision Address Type.
- Severity of the Collision, according to the SDOT classification.
- The number of pedestrians involved in the collision.
- The number of bicycles involved in the collision.
- The number of vehicles involved in the collision.
- The date and time of the incident.
- Whether or not collision was due to inattention.
- Whether or not a driver involved was under the influence of drugs or alcohol.
- Weather Conditions
- Road Conditions
- Light Conditions
- Whether or not the pedestrian right of way was not granted.
- Whether or not speeding was a factor in the collision.
- A code provided by the state that describes the collision.
- A key for the lane segment in which the collision occurred.
- A key for the crosswalk at which the collision occurred.
- Whether or not the collision involved hitting a parked car.

## Data Source

The dataset employed for this project is named "Collisions-All Years" and is provided by the Seattle Police Department and recorded by Traffic Records. This database includes all types of collisions and it's updated on a weekly basis.

For more information regarding the dataset employed in this project, you may contact the SDOT (Seattle Department of Transportation) Traffic Management Division, specifically the Traffic Records Group at DOT_IT_GIS@seattle.gov.