# Predicting Severity of Car Accidents IBM Data Science Capstone

David Wang

# Intro

- Every year approximately 1.35 million people die in road crashes
- On average 3,700 people lose their lives every day

# Business Problem

- This project will involve working on a case study where we will predict car accident severity. The stakeholders will be the drivers of ridesharing companies. This is important for employee safety and their customers. Drivers usually don't pay attention to how carefully they should be driving given certain weather and road conditions. Developing a model to predict the severity of an accident it would be can help drivers make better decisions on how cautious to drive and assess their travel plan.

# Data

The dataset is provided by IBM

- CSV file
  - A target label ‘severity code’. 1 represents property damage only and 2 for injury collision.
  - Different features that describe the accident
- PDF file
  - Metadata for the csv file

# Methodolgy

1. Hot encode the data so machine learning algorithms can use the data
2. Clean the data by removing all the unknown and normalizing the feature data
3. Plot charts to visualize the data
4. Select the features to be used by the KNN
5. Build the KNN model to predict the severity of a crash

# Results

- Using k = 14, the number of neighbors, to create a KNN model, I was able to predict accident severity with 73% accuracy.

# Reference


- https://www.asirt.org/safe-travel/road-safety-facts/#:~:text=Annual%20Global%20Road%20Crash%20Statistics,resulting%20in%20long%2Dterm%20disabilities.