**Capstone Project**

**Introduction problem**

Say you are driving to another city for work or to visit some friends. It is rainy and windy, and on the way, you come across a terrible traffic jam on the other side of the highway. Long lines of cars barely moving. As you keep driving, police car start appearing from afar shutting down the highway. Oh, it is an accident and there's a helicopter transporting the ones involved in the crash to the nearest hospital. They must be in critical condition for all of this to be happening. Now, wouldn't it be great if there is something in place that could warn you, given the weather and the road conditions about the possibility of you getting into a car accident and how severe it would be, so that you would drive more carefully or even change your travel if you are able to.

People who like travelling by car, bus drivers, truckers and even ordinary persons who hate being stuck in traffic jams will be happy to know about possibility of getting into a car accident. It can save a lot of time and you will always be happy about your journey.

**Data problem**

I will use data about collisions. All collisions provided by SPD and recorded by Traffic Records. This includes all types of collisions. Collisions will display at the intersection or mid-block of a segment. Firstly, I am going to preprocess the data, dealing with all the missing data. Secondly, I will change type of columns to numeric. After that I will decide which features to take by correlation. Then I am going to scale the data with StandardScale and use SVM as machine learning algorithm to predict severity of collision.

Example of final dataset:

