**INTRODUCTION / BUSINESS PROBLEM**

This essay details the data preparation and modeling work that uses a publicly available dataset on vehicle accidents registered by the Seattle City Department of Transportation. This work benefits the American public by identifying those factors conducive to various types of traffic incidents and using this information to minimize this risk.

**INTRODUCTION / BUSINESS PROBLEM**

In predicting the target variable of accident severity, there are 37 key input variables to be assessed for their value. Those variables and value-adding descriptors are summarized below:

1. X and Y appear to be coordinates, focused on one particular area
2. ObjectID may suggest what is involved in incident - it's a unique identifier.
3. Inckey is unique key for incident. LIKELY UNINSIGHTFUL, SINCE IT"S A CLASSIFIER.
4. Intkey is unique key for intersection where incident happens.
5. Coldetkey corresponds to collision detail
6. Seglane and Crosswalk have keys reflecting where incident happened
7. Hit Parked Car reflects whether this happened
8. Collison code and description given by State Dept of Transportation
9. Whether speeding was factor and whether Pedestrian Right Of Way Was Not Granted
10. Lighting and road conditions
11. Type of Junction, whether or not driver was DUI
12. Whether inattention or DUI was a leading factor
13. Date and time of incident wether collision was due to lack of attention