**Introduction/Business Problem:**

The Seattle administration has been reviewing accident cases as a means to deploy preventive measures to avoid damage to individuals and property. An ability to convert historic accident reports and use that to help state police and authorities to create targeted awareness measures to prevent accidents will be helpful. Further, insurance companies would also benefit by preventive and pro-active counter-measures that will reduce third-party property damage.

Using existing fields from records such as environmental factors, the state can use this model to preempt accidents and plan its response accordingly. The target audience for this model are the state police department, emergency services, town planners and insurance companies that can use this model to plan better and reduce accidents.

**Data**

The dataset consists of approx 195 thousand accidents in the Seattle area from January 2004 ranked by severity. Severity 1 indicates property damage whereas severity 2 indicates injury. The data includes 37 independent features that have been recorded at the time of the incident. The features include record details of ID, environmental factors such as weather, physical condition of roads, street lights and time of accident.

From the dataset, there were 136,485 cases of property damage and 58,188 cases of injury. On further analysis of the data, the number of incidents was highest on Friday and lowest on Sunday. Other features that will be used to determine severity are Weather, Road Condition, Lighting and Junction type. The highest number of injuries (Severity 2 incident) happened at intersections.

Using these features, I plan to build a classification model that will determine the probability of the severity of accidents.