

Problem

According to the Insurance Institute for Highway Safety (IIHS), motor vehicle collisions in the United States are responsible for roughly 35,000 deaths each year. In order to determine improved vehicle safety guidelines, we analyzed data to answer these business questions: can we predict the severity of traffic accidents based on the light conditions, and is there a statistically significant correlation between the level of light at the time of an accident and the severity level of the accident? With this information, we can make recommendations pertaining to the optimal amount of public lighting infrastructure in a neighborhood.

Data

For this business case we will be utilizing the Seattle vehicle collisions data. We will incorporate the collision severity data per crash, the light conditions at the time of the crash, as well as generalized data on where the crash occurred: intersection, block, or alley. These will allow us to determine the correlation between light conditions and total number of collisions, as well as determine which part of a street is most heavily impacted by the current lighting conditions. We will then determine the impact of light conditions and street location on the severity of the crash. We will assess the data using common data science techniques and statistical modeling.

Methodology

Firstly, we cleaned the data by removing any rows which had columns without values. This reduced the sample size yet maintained a large enough sample for statistical accuracy. Second, the rows containing data which did not pertain to the business questions was removed. This left a data set containing street locations, number of collisions, and the light condition at the time of the collision.

Results

Firstly, the data was analyzed to determine the where on a street the majority of vehicle accidents occur. As shown in the table below, a vast majority of collisions occurred in the middle of the city block, roughly 66% of collisions. This was followed by nearly 34% at road intersections, and less than .5% in alleys.

Street Location	Number of Collisions	Percent of Collisions
Alley	744	0.40%
Block	123400	65.77%
Intersection	63486	33.84%