**Car Accident Causes and Severity in the Seattle Area**

**Introduction and Business Problem**

The problem to be addressed in this course is centered around both the causes and the severity of car accidents. These are important concerns because in 2019, it was recorded that almost 227,754,100 licensed drivers were using the various roadways in the United States. That number is more than just a statistic, it is human lives that are at risk of being lost every day. For this reason, I think that the causes and severity of car accidents should be analyzed in the effort to benefit the general population, civil and road engineers, the Bureau of Motor Vehicles, and Emergency Medical Services. This applied study could help the general population by making them aware of which driving conditions have the highest risk factors for their safety. With this information, commuters and travelers could drive more carefully or take a different route. Additionally, this study can offer insight to professionals and consultants who design the roadways and bridges in the United States to ensure safer design strategies upon implementation. In a similar fashion, the Bureau of Motor Vehicles could use this information to enforce design regulations on automobile manufacturing companies to verify that new models have a higher degree of safety. Finally, Emergency Medical Services could benefit from this study because they could allocate more resources and personnel during high risk time periods. For example, during times of day that have a higher frequency of reported accidents or during periods of inclement weather.

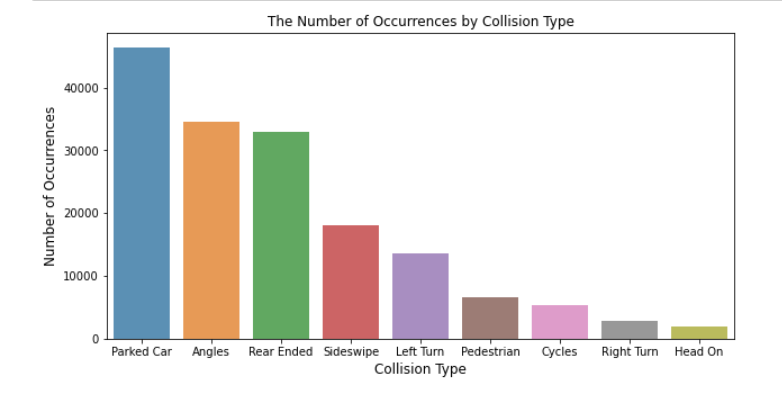
**Data**

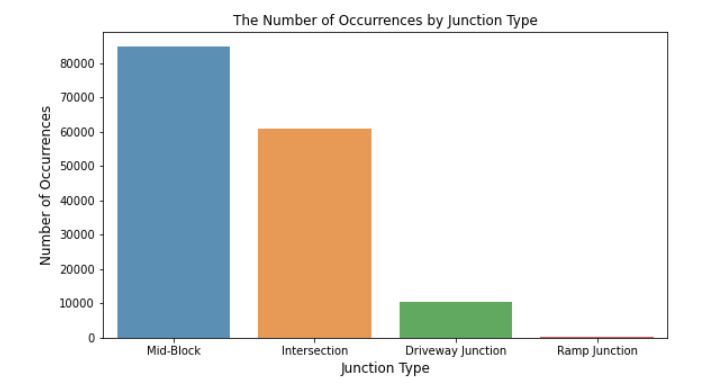
The data that will be used to find the causes and severity of car accidents will be the Data-Collisons.csv file that was provided in this course. More specifically, this data pertains to the city of Seattle, Washington. After further inspection of the data set, it is apparent that it contains a severity code, on the scale of 0-3, where 0 is unknown and 3 is a fatality. A location is provided, given in x,y coordinates. Along with these two items is a description of the severity of the collision as well as the collision type. The number of people involved in the accident is identified and is further categorized as a person, pedestrian or cycling pedestrian. Various conditions such as weather and time of day are given. For example, based on the first row of data, it is evident that the severity code was 2 and that an injury occurred. Furthermore, it is clear that two people were involved both of which were drivers of motor vehicles. While it is apparent that inattention or the influence of alcohol were not contributing factors to this accident, wet road conditions were listed in the case description. Based on just this particular accident, it would be beneficial to explore the effects of weather and the number of people on accident probability.

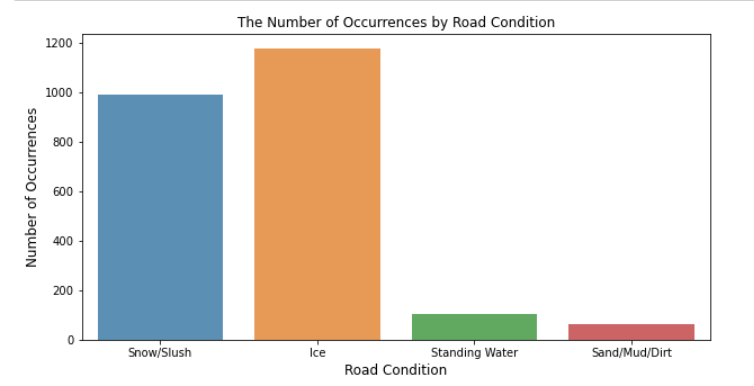
**Methodology**

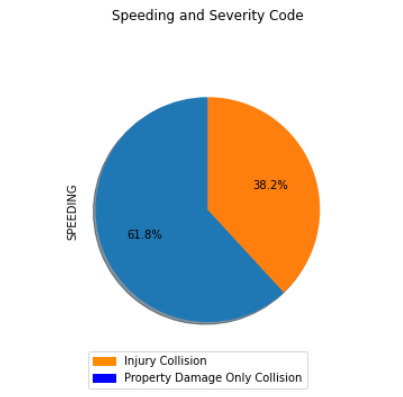
A variety of methods were used to understand and analyze the data for this project. Pandas functions such as df.describe(), df.value\_counts() and df.shape were utilized to become familiar with the data set on a surface level. To visualize the data, bar charts, pie charts, and maps were developed. Groupby functions were also used to group the speeding and severity code data together in order to understand their categorical relationship. No machine learning techniques were used on this data set, all relations were discovered by means of manual interpretation.

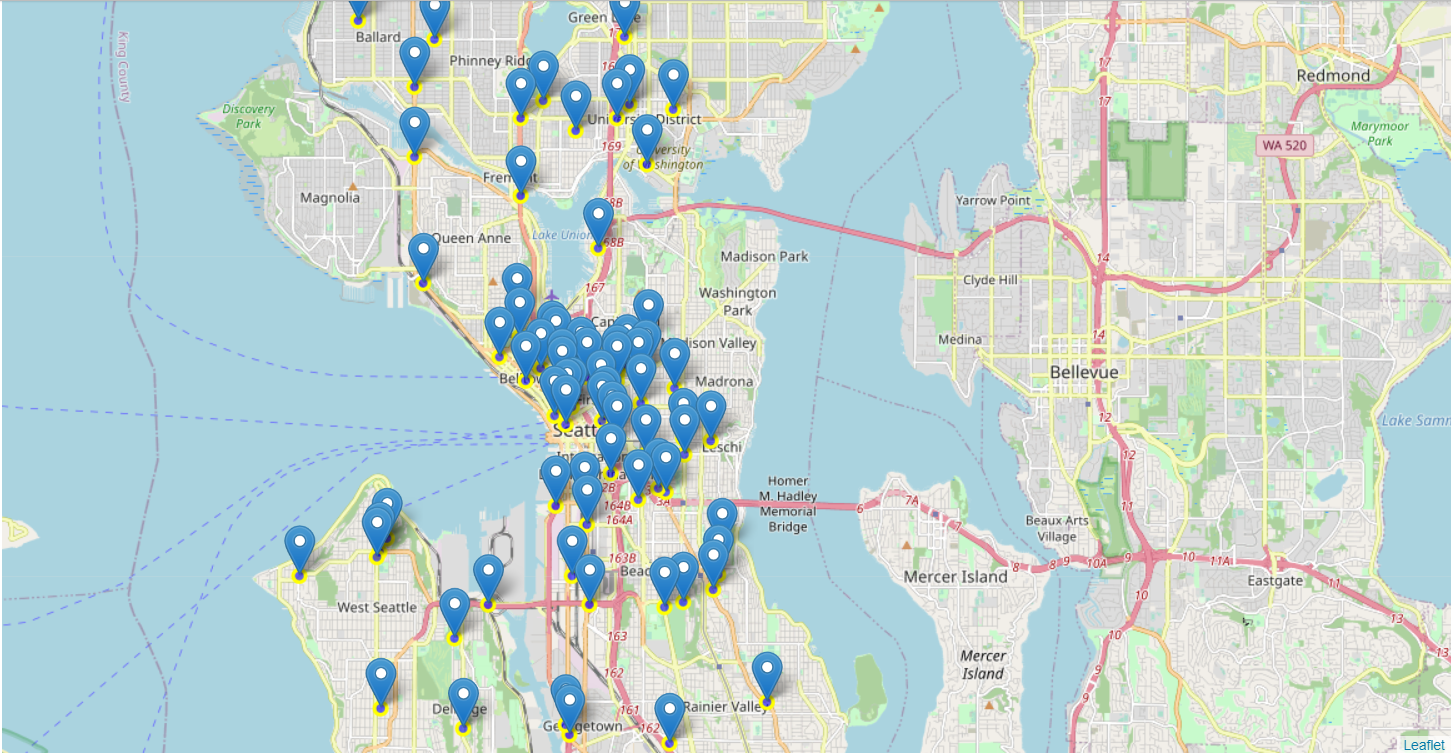
**Results**











**Discussion**

The above results showcase a wide variety of observations. From the first data set, it was apparent that head on collisions were the least frequent type of collision while parked car collisions were the most frequent. This is important to note, for it highlights that it is more likely to see property damage accidents than to see accidents that result in injuries. The severity code pie chart emphasizes this theme, for only 38% of speeding incidents resulted in an injury to a person. For pedestrians and commuters, the above map demonstrates that it is more likely to have an accident closer to Seattle’s city center than in the outer belt areas. With more pedestrian activity and density in the city center, it is advisable that drivers should exhibit a greater degree of caution when driving downtown. Icy conditions were found to be the most dangerous of road conditions while dirt and sand on the road only contributed a small fraction to accidents frequency.

**Conclusions and Insights**

It was found that the presence of ice on the road was the worst road condition possible. Additionally, parked car collisions were the most frequently occurring accident type. In terms of the most common junction for an accident to take place, it was concluded that the middle of a block was the most common. Furthermore, the majority of speeding accidents only resulted in property damage instead of in injury. Finally, accident density was greater near the city center than in the northern peninsula or other suburbs of Seattle.