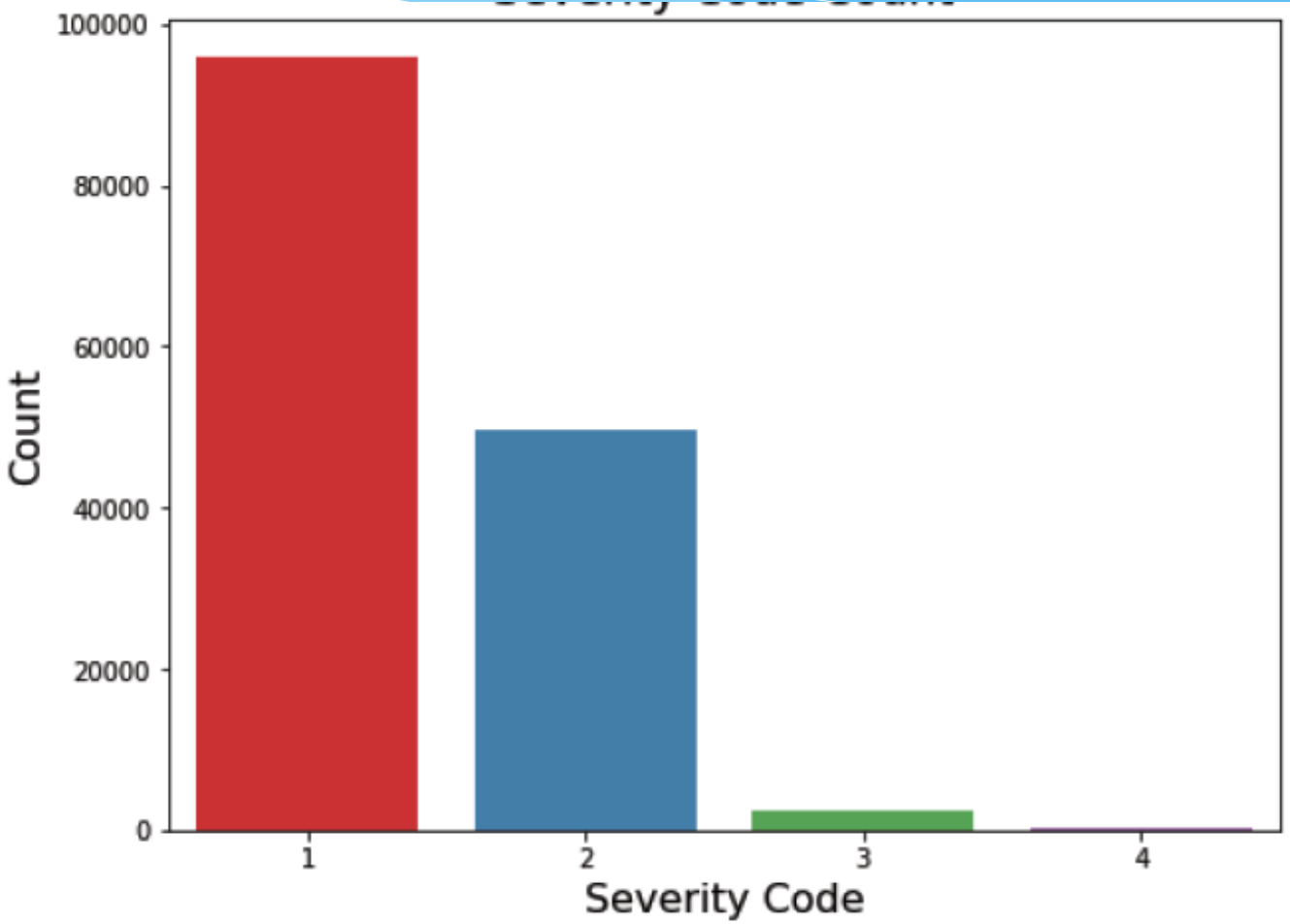
Car Collision Severity: A Case study of Seattle

- Mohd Shabbiruddin Khan Miftahi

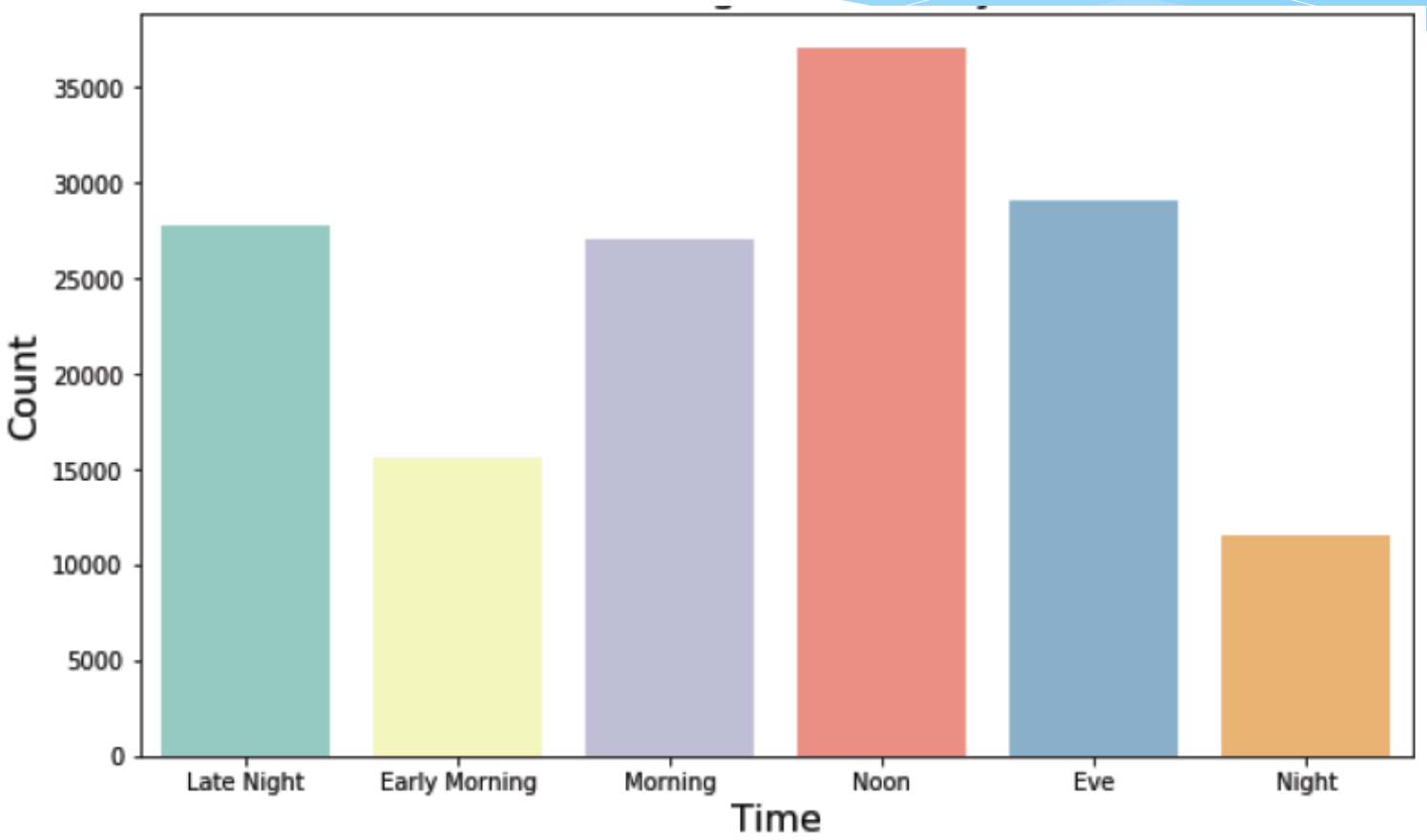
Introduction

* Road accidents are an important issue of our modern societies, responsible for millions of deaths and injuries every year in the world.
* Seattle, a city on Puget Sound in the Pacific Northwest., Washington State's largest city, is a home to a large tech industry, with Microsoft and Amazon headquartered in its metropolitan area.
* The rapid economic growth that the city has developed in the past has resulted in improving the lifestyle of the people living over there.
* With the raised standard of living in the city, and greater purchasing power of its inhabitants, the number of automobiles on the roads has increased rapidly.
* With so many people owning and operating vehicles on the roads, problems like traffic jams, congestion on roads, and road accidents are becoming a common sight.
* The aim of this work is to create a model that could warn the road users, given the weather and the road conditions about the possibility of them getting into a car accident and how severe it would be.
* The data for this work was taken from open data platform hosted by the city of Seattle.

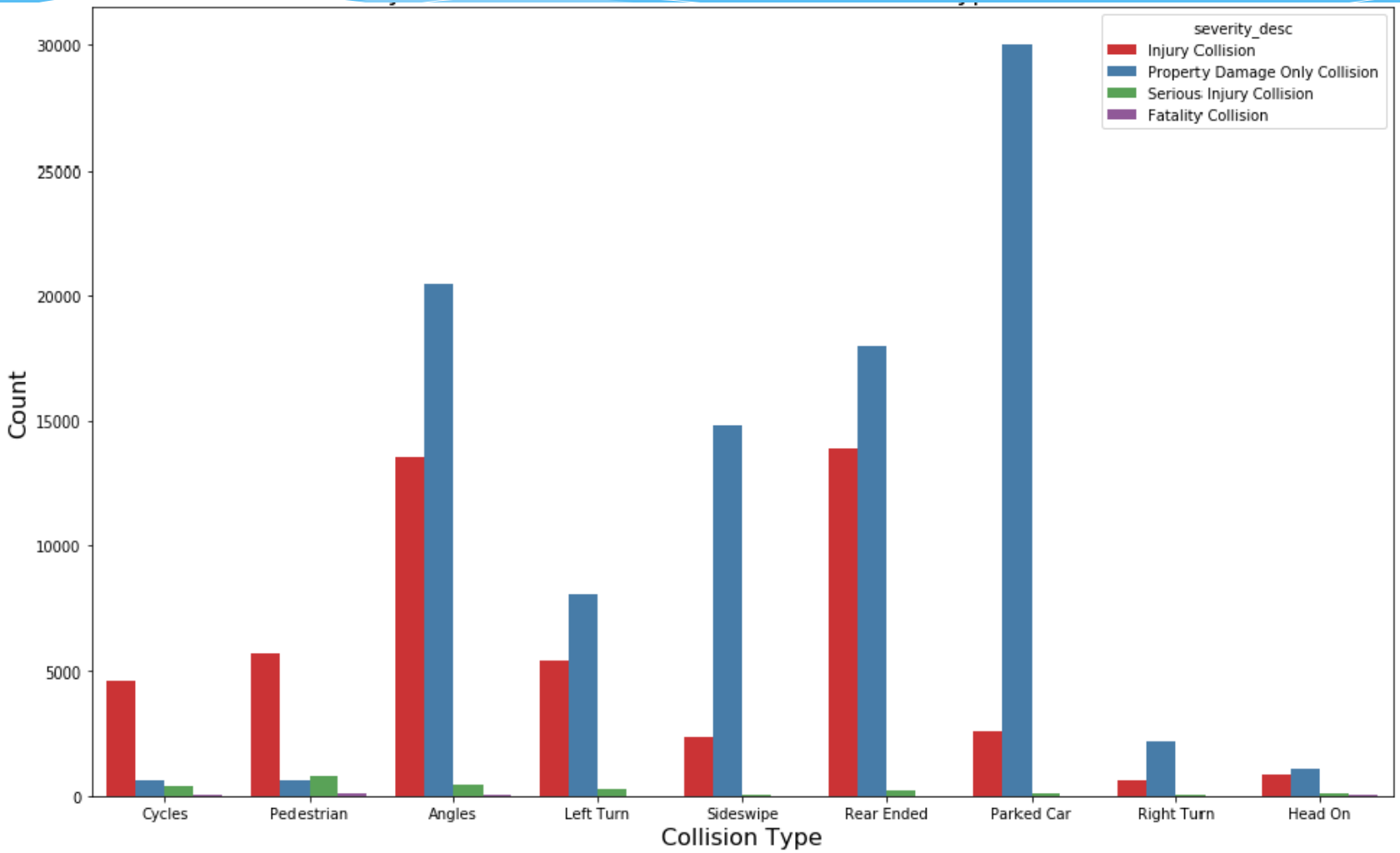
Severity Codes



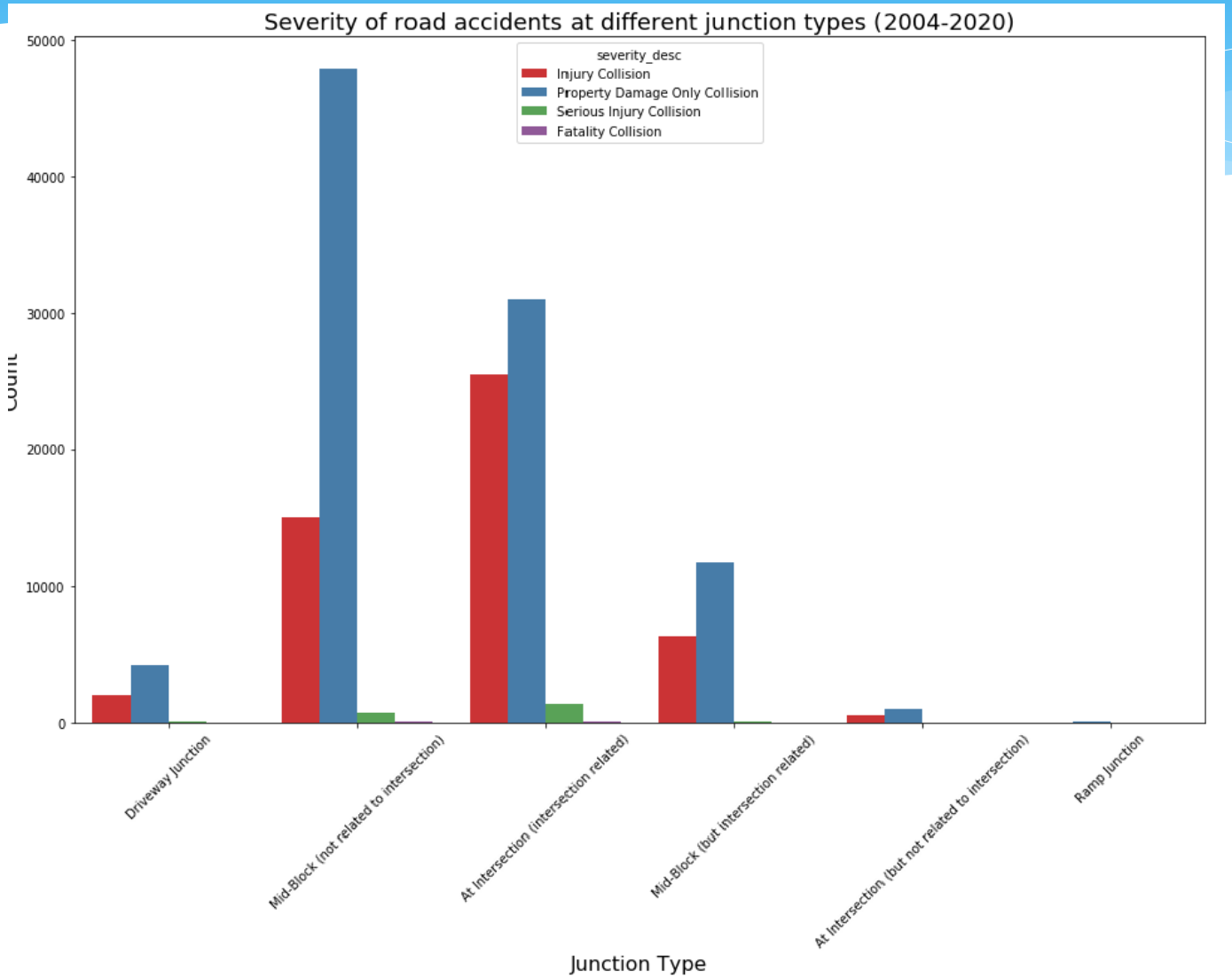
Accident Trend throughout the day



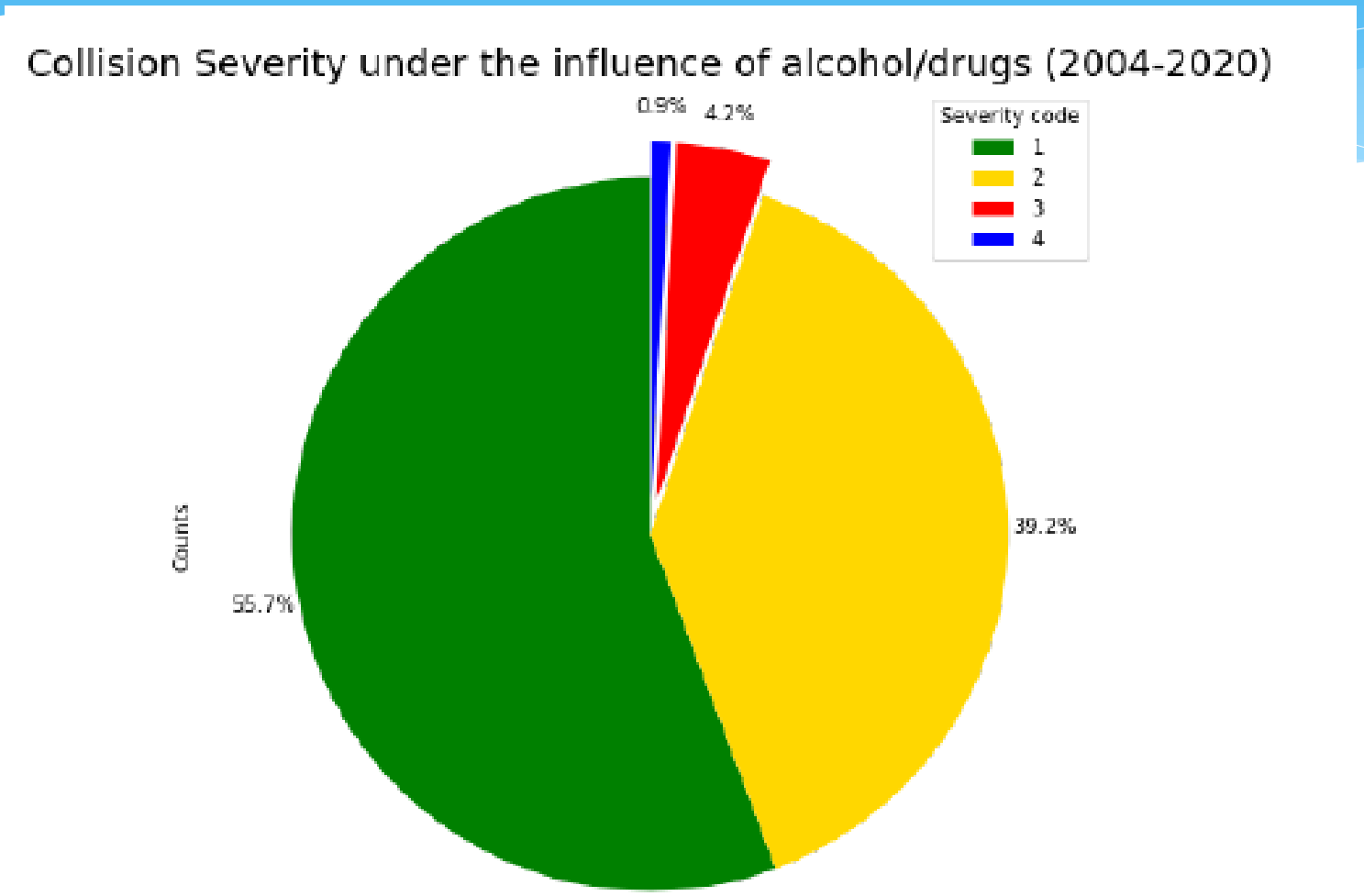
Severity of different Collision types



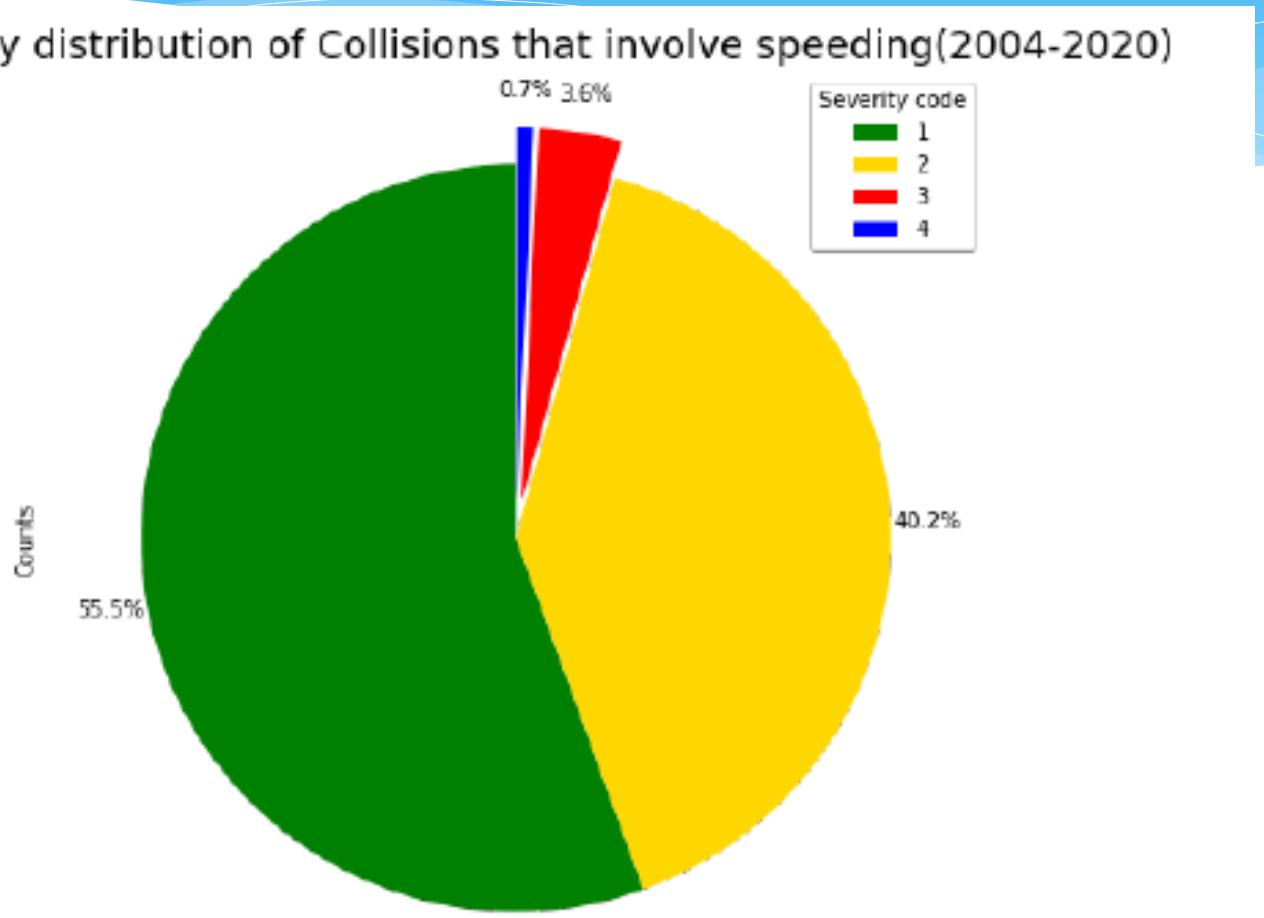
Severity of road accidents at different junction types



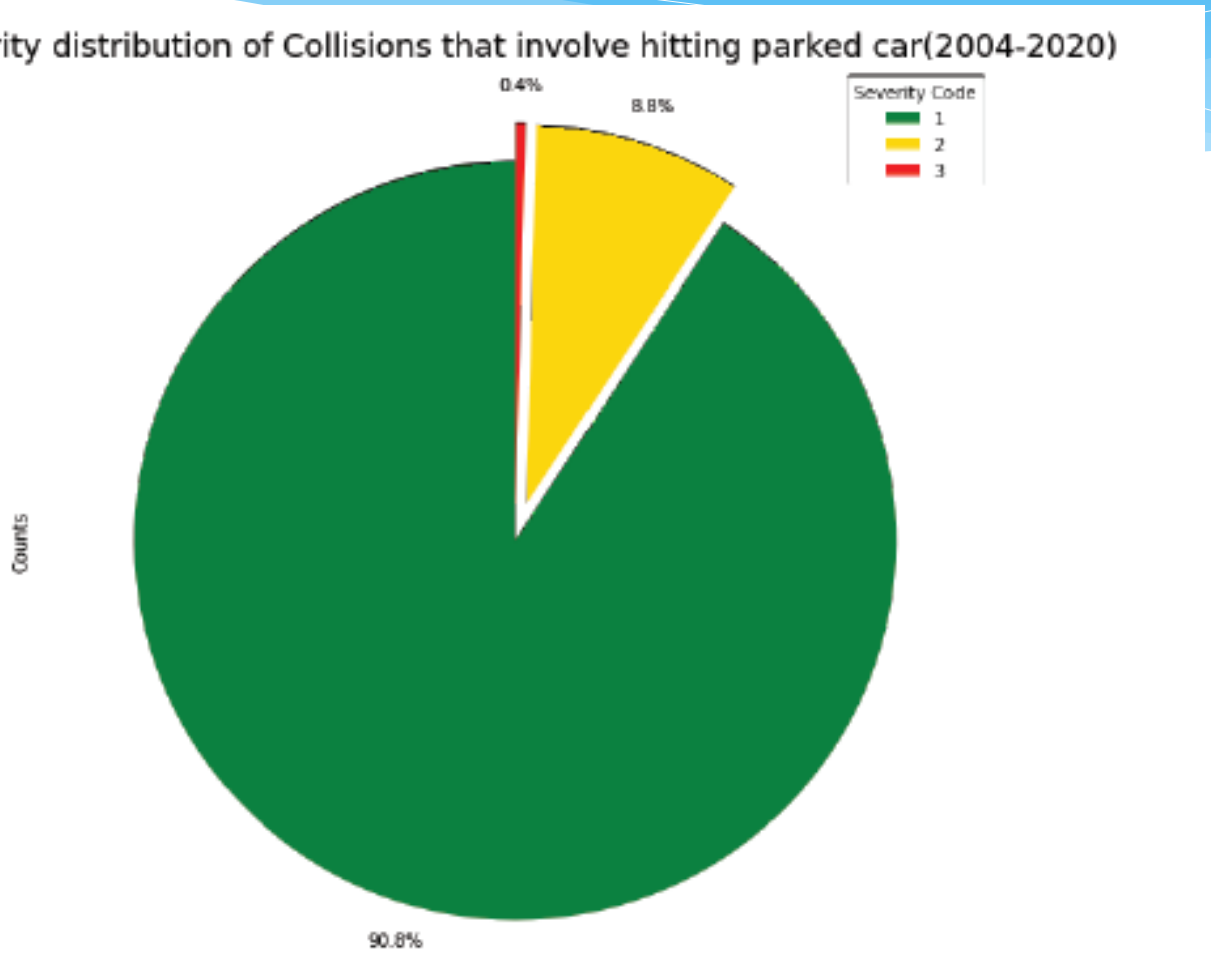
Influence of Drugs/Alcohol



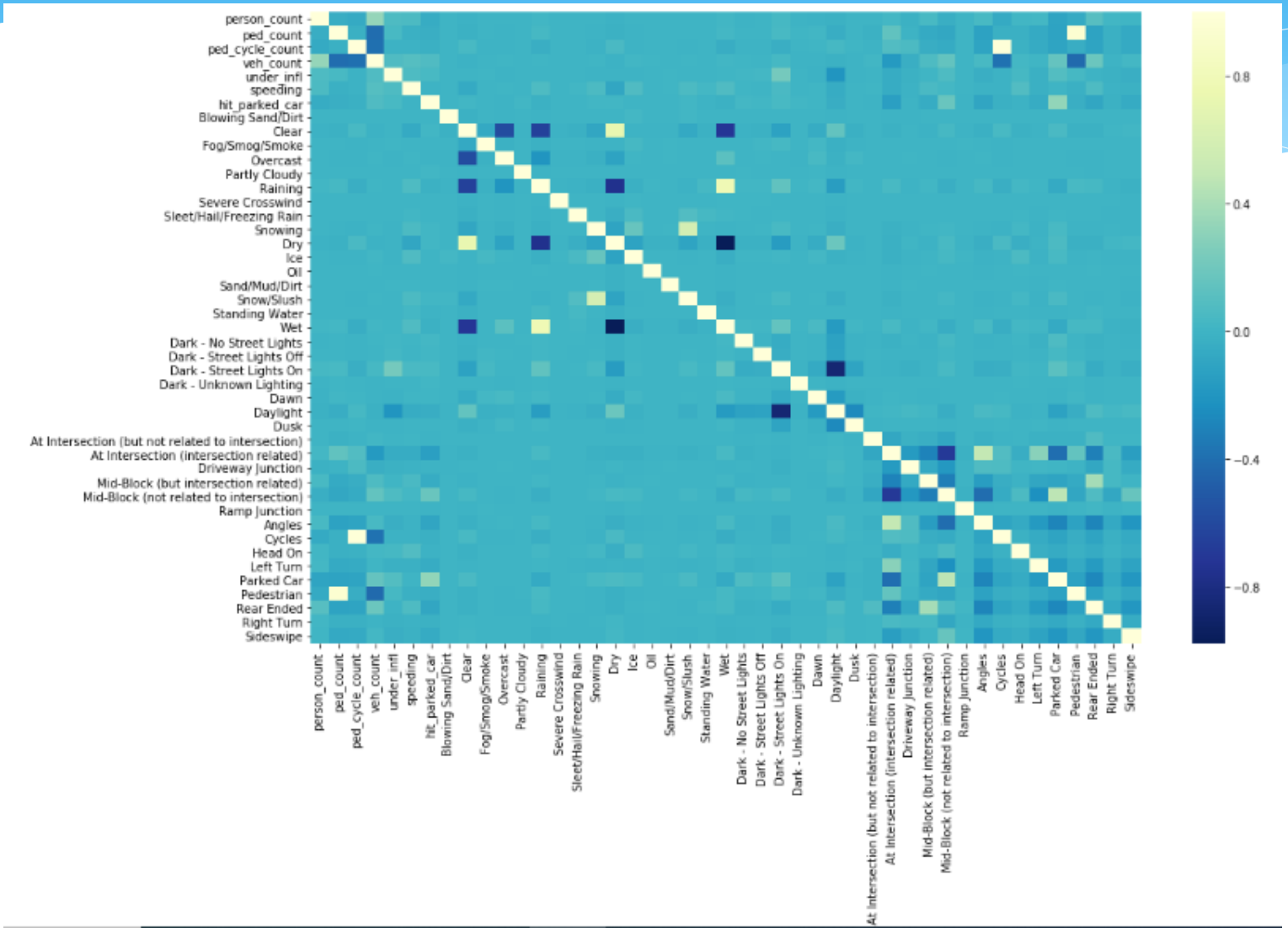
Influence of Speeding on Collision Severity



Severity of cases involving hitting of a parked car



Correlation between Independent features selected



Conclusion

* We can conclude that the Decision Tree is the best model in this scenario an KNN is the second best.
* Logistic Regression appears to be the worst of the four techniques.
* The initial data had a lot of missing values and was unbalanced, so we had to ensure proper cleaning and balancing of the data in order to prevent a skewed model.
* From the exploratory analysis, we see that most accidents cause only property damage or minor injuries.
* Most accidents are angled crashes followed by parked car and rear end crashes.
* Most accidents happen during clear days on dry roads in daylight.
* Most property damage only collisions occur at midblocks while most injury collisions happen at intersections.
* Most serious injuries and fatalities are caused by collisions with pedestrians. As such special attantion should be paid towards pedestrian safety.
* Further, we find that most property damage only collisions occur at midblocks while most injury collisions happen at intersections.
* Finally, influence of alcohol/drugs and speeding appear to increase the severity of accidents.