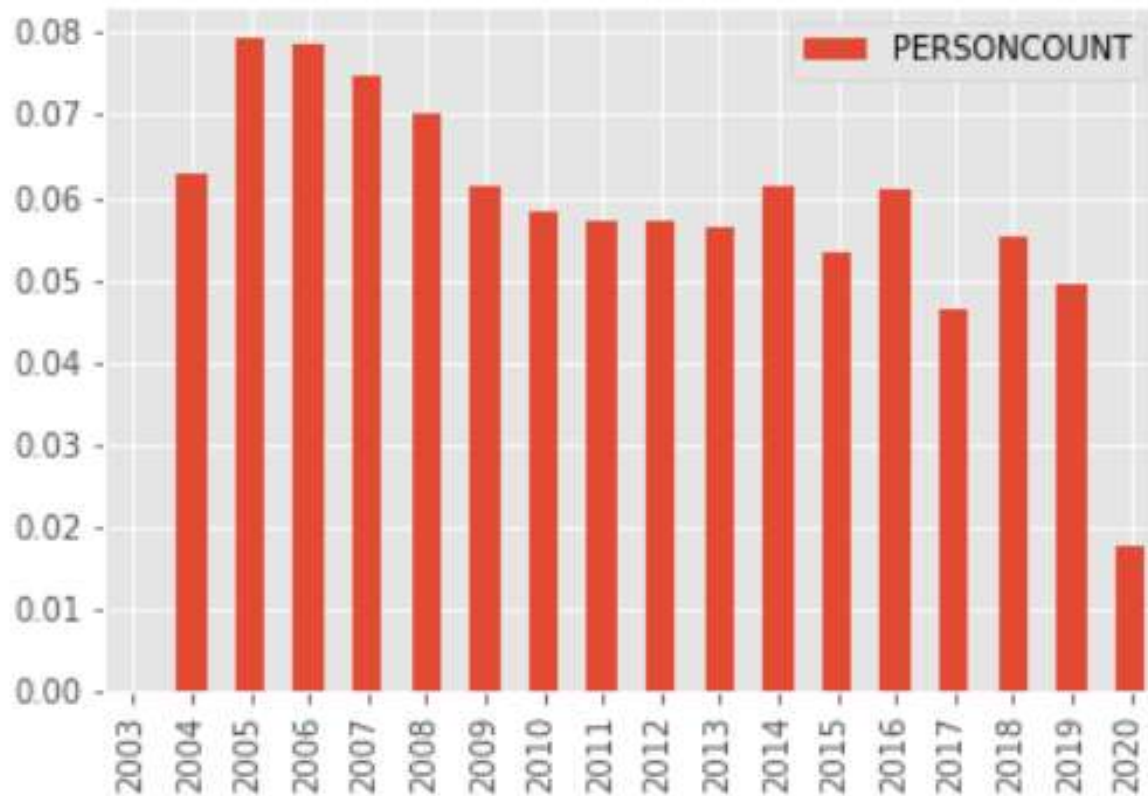


Seattle Collisions severity recognition using Machine Learning

Data exploratory analysis

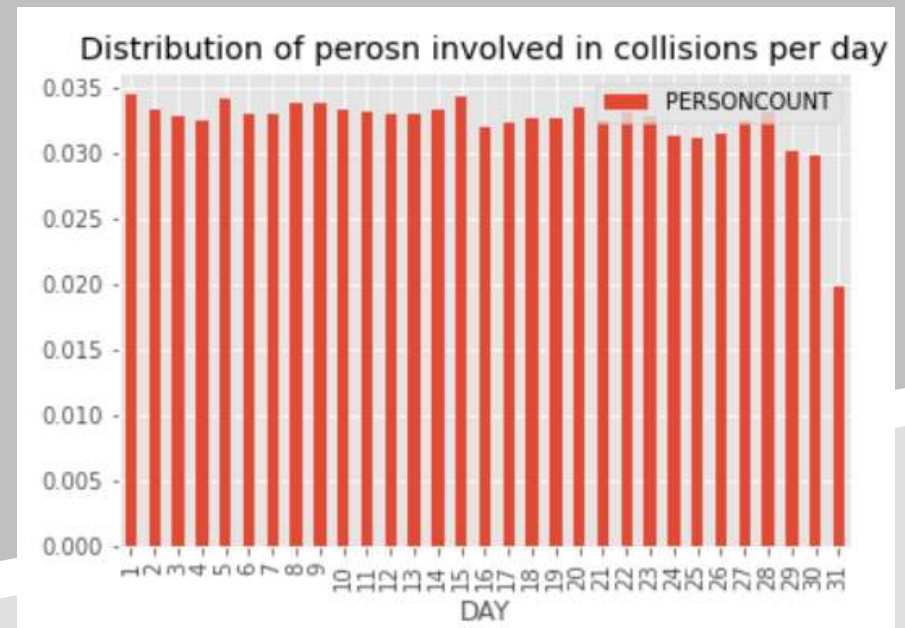
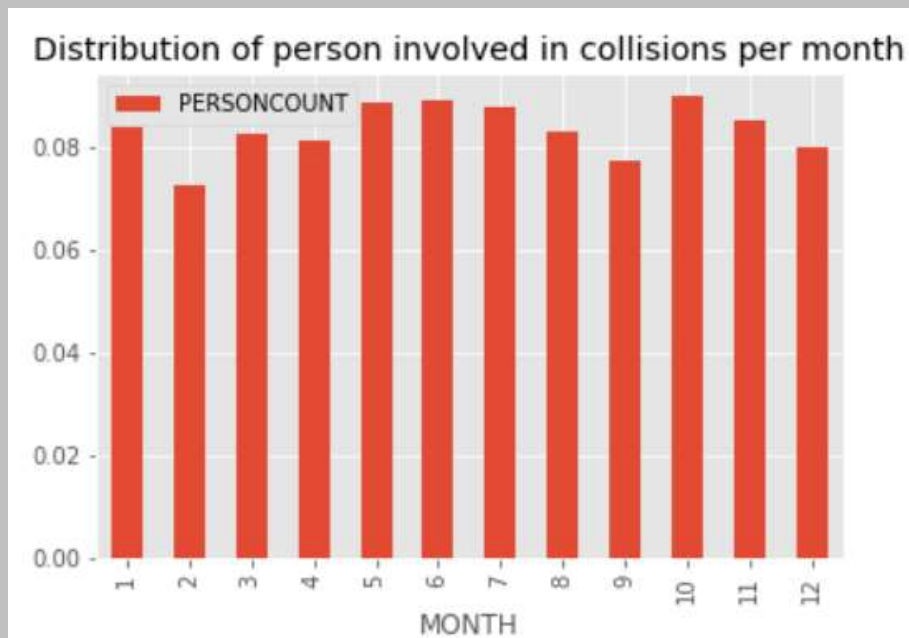
- Incidents have decreased over the years

Person involved in collisions count from 2004 to present



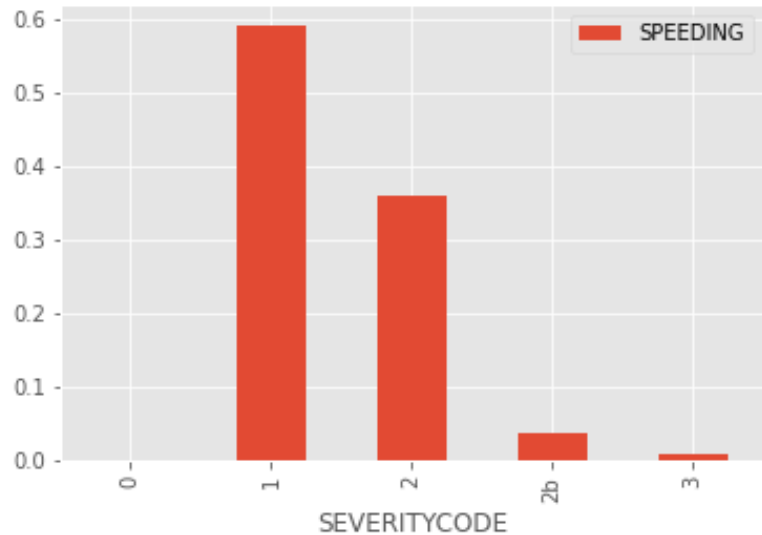
Data exploratory analysis

- There is no preferred month of day
- February and September have the least number of incidents

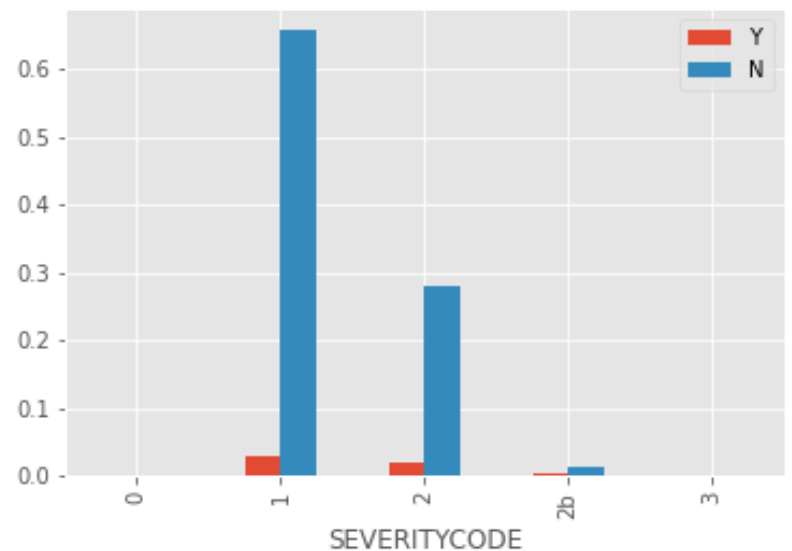


Data exploratory analysis: Driver-Pedestrian

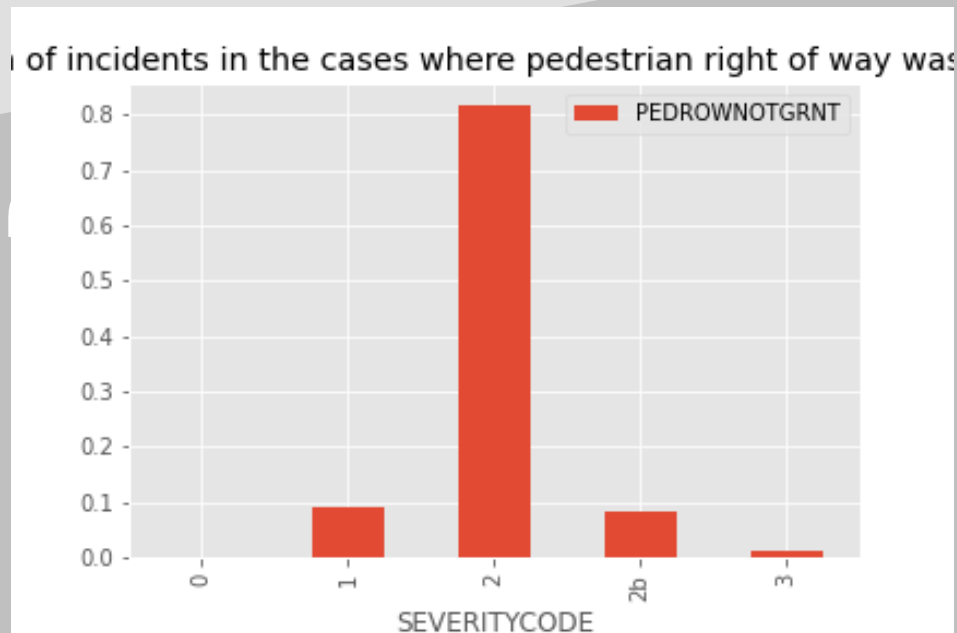
Distribution of incidents where the driver was speeding



Distribution of incidents whether the driver was under alcohol or drug

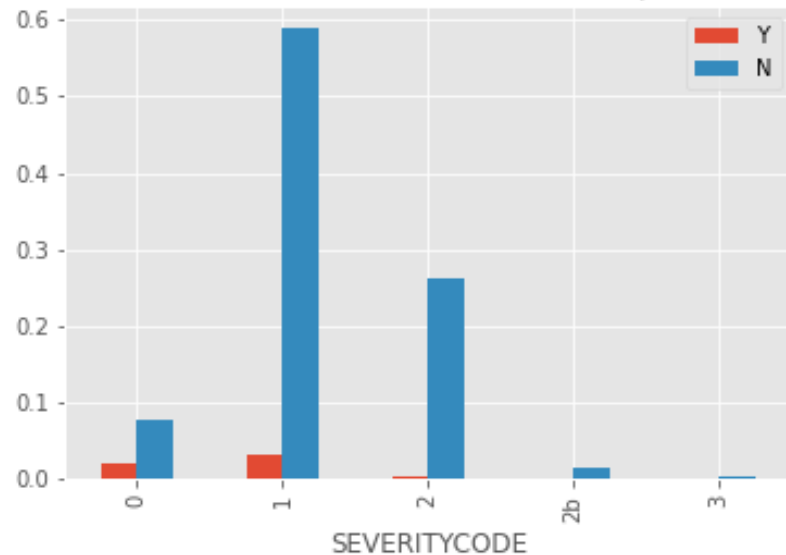


Data exploratory analysis: Driver-Pedestrian

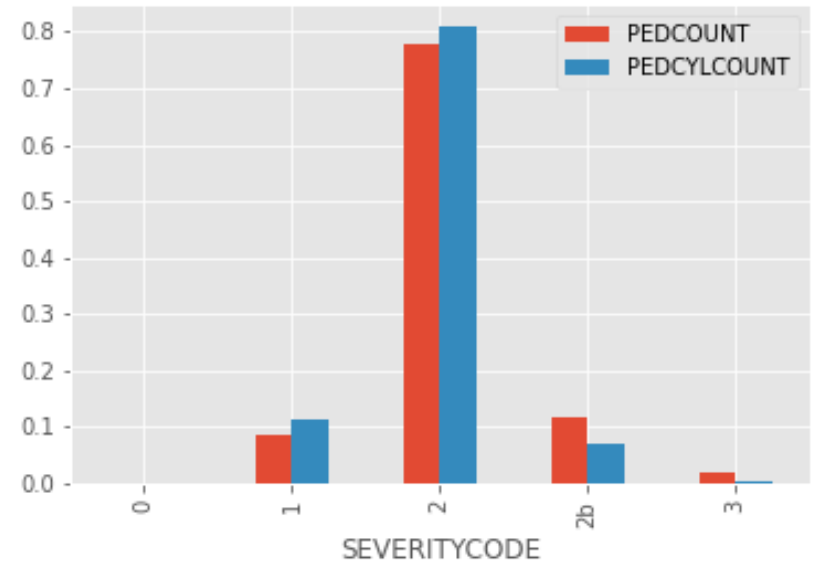


Data exploratory analysis: Driver-Pedestrian

Distribution of incidents whether or not was a parked car involved

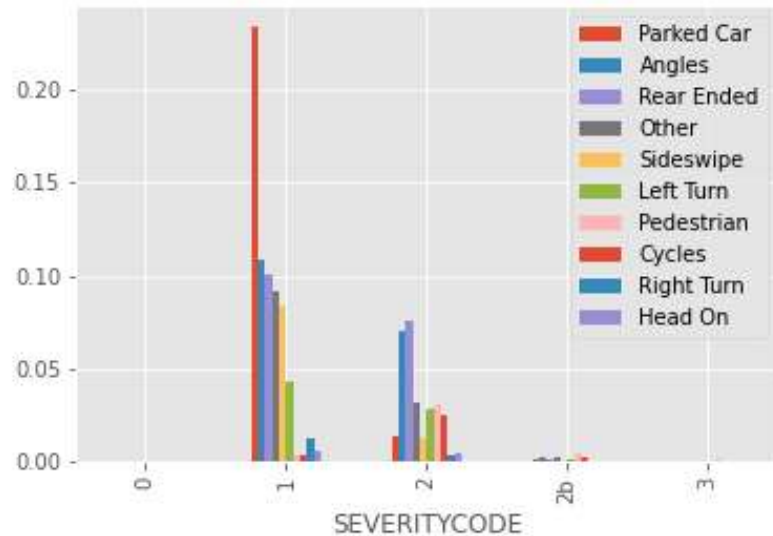


Distribution of incidents according to the number of pedestrians and bicyclists involved

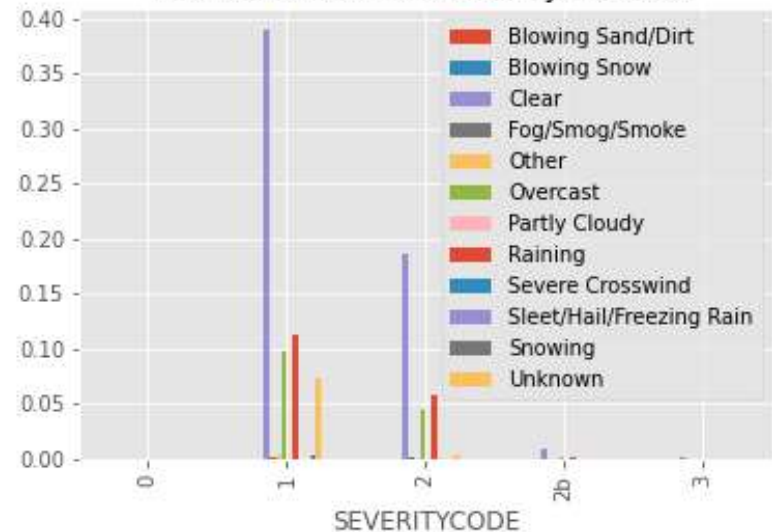


Data exploratory analysis: Environment

Distribution of incidents according to the collision type



Distribution of incidents by weather

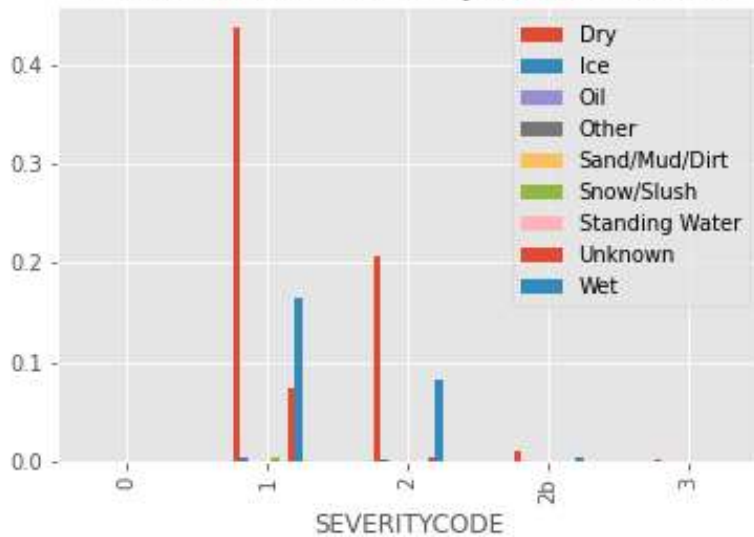


Data exploratory analysis: Environment

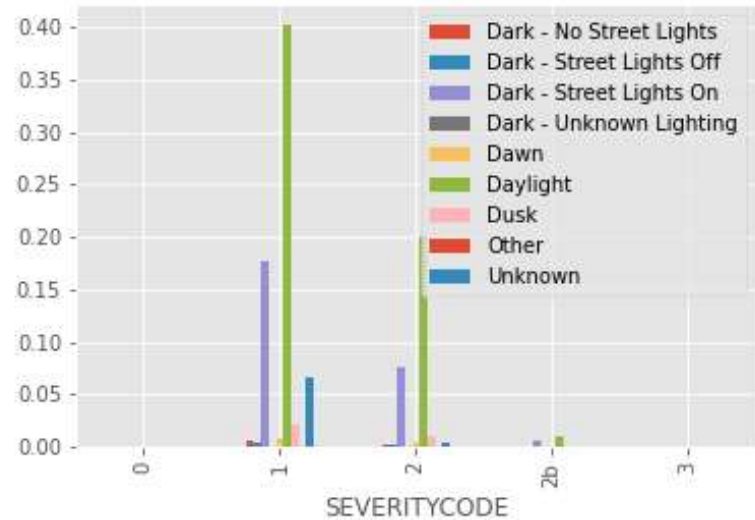
- There is a small fatalitie percentage
- “• Most accidents implies property damage
- Just a small percetage involves serious injuries although the injuries percentage is quite elevated.”

Data exploratory analysis: Environment

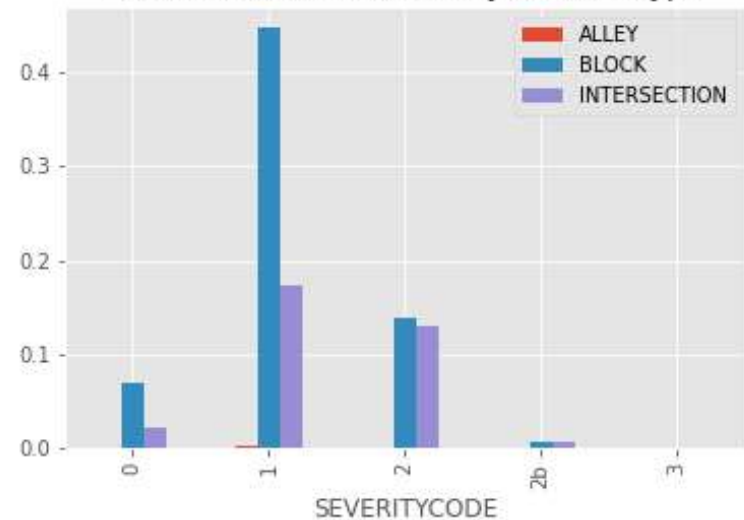
Distribution of incidents by Road conditions



Distribution of incidents by light conditions



Distribution of incidents by address type



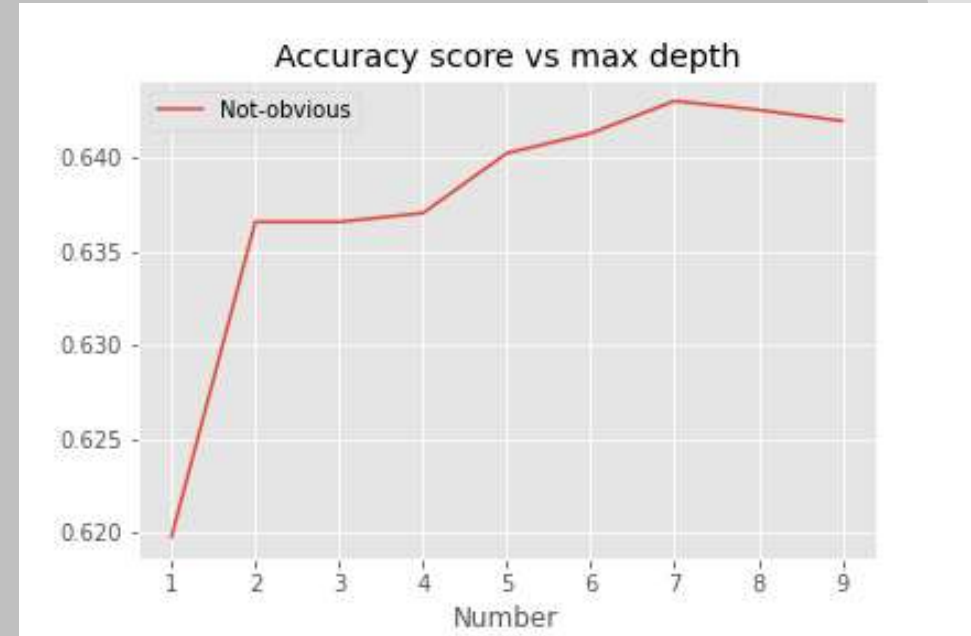
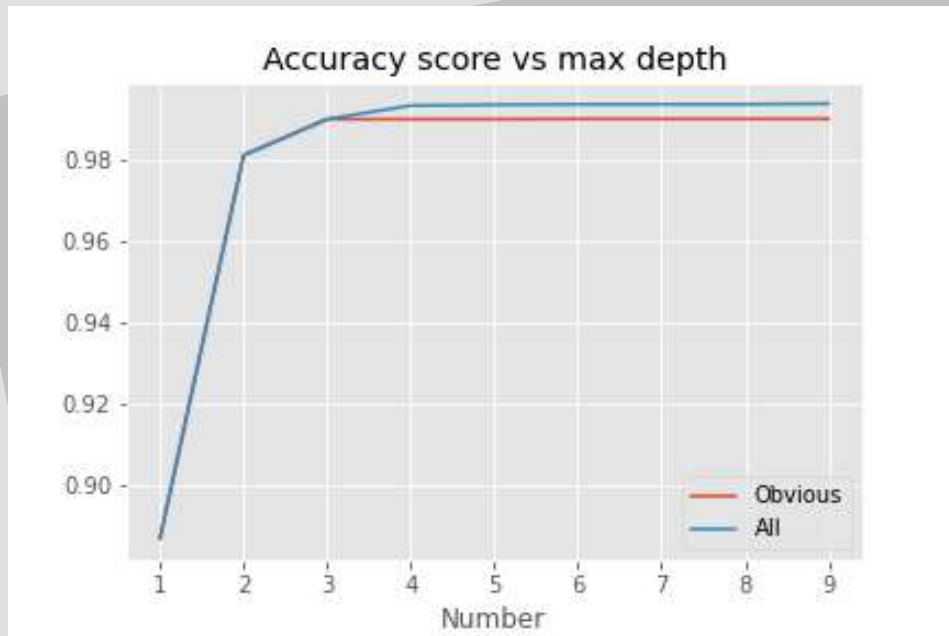
Data exploratory analysis: Environment

**“IT IS ALL ABOUT DRIVER
RESPONSIBILITY!!”**

Modeling

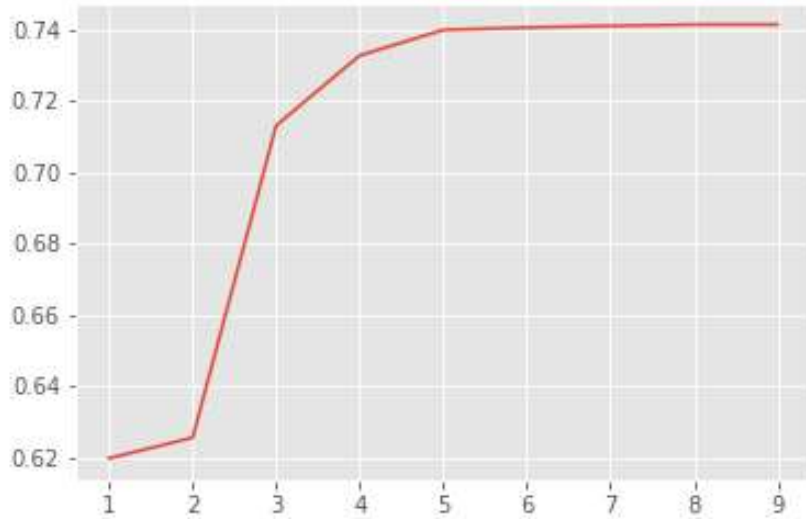
- Decision tree algorithm was accurate
- Other algorithms take high computing times

Modeling: Numerical Attributes

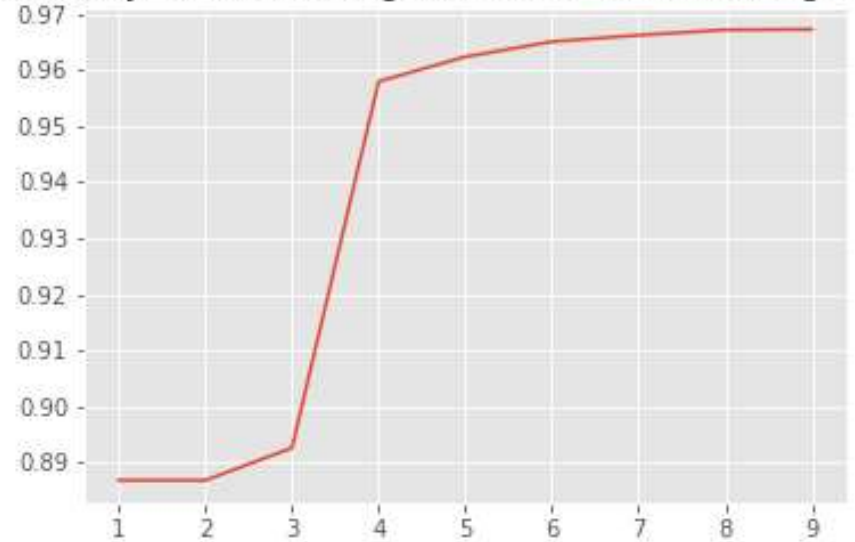


Modeling: Categorical Attributes

Accuracy Score for categorical attributes implementation



Accuracy score for categorical attributes including Injuries



Including injuries highly improve accuracy!

Conclusions

- Incidents have decreased over the years
 - There is no preferred month of day
 - February and September have the least number of incidents
 - There is a small fatality percentage
 - Most accidents imply property damage
 - Just a small percentage involves serious injuries although the injury percentage is quite elevated.
 - Most collisions occur in good environmental conditions, so they are a result of a lack of responsibility
- “
- Decision tree algorithm was accurate
 - One can either model with numerical or categorical attributes with high accuracies.
 - Modeling with attributes improves its accuracy by including injuries rather than fatalities
 - A tree depth between 5 and 7 is optimal
- ”