



Predicting Severity of Car Collisions

Ariella Goldman
IBM Data Science Coursera Capstone Project



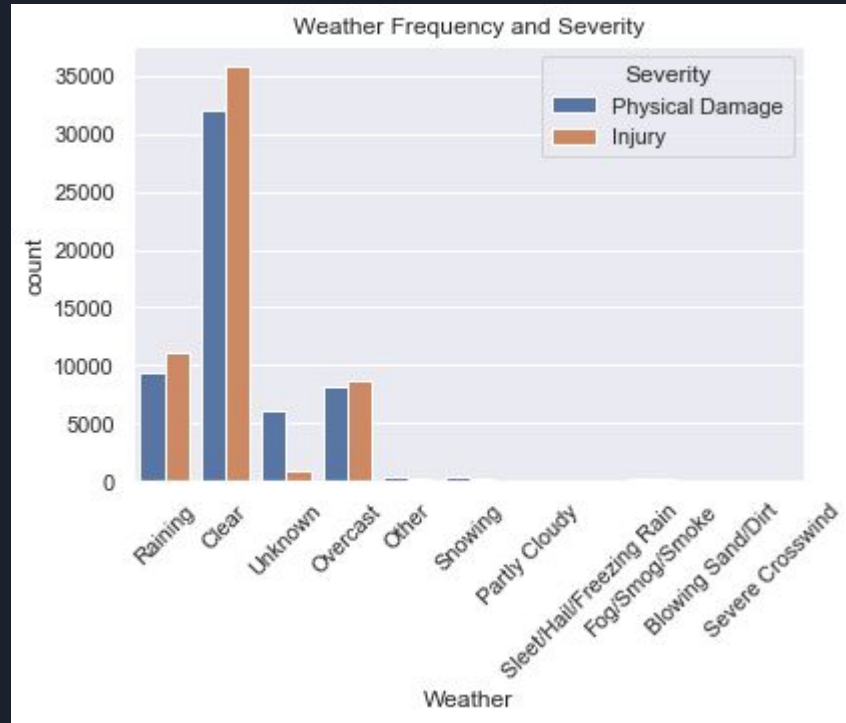
Predicting Collision Severity is Important to Private Insurance Companies

- In 2019, there were 16,000 car crashes per day in the United States
 - In 2013, the average auto liability claim for property damage was \$3,231.
 - The average auto liability claim for bodily injury was \$15,433
 - Private insurers pay approximately 50% of all motor vehicle crash costs.
-
- If something could warn a driver, given for instance, the weather and the road conditions, about how severe a potential accident would be, a driver might drive more carefully or perhaps change travel plans, saving insurance companies money

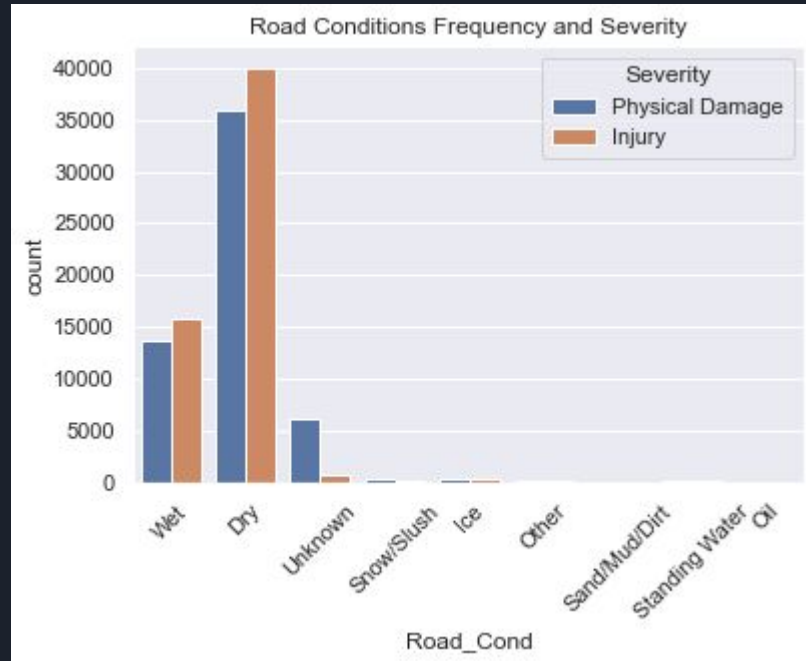


Data Acquisition and Cleaning

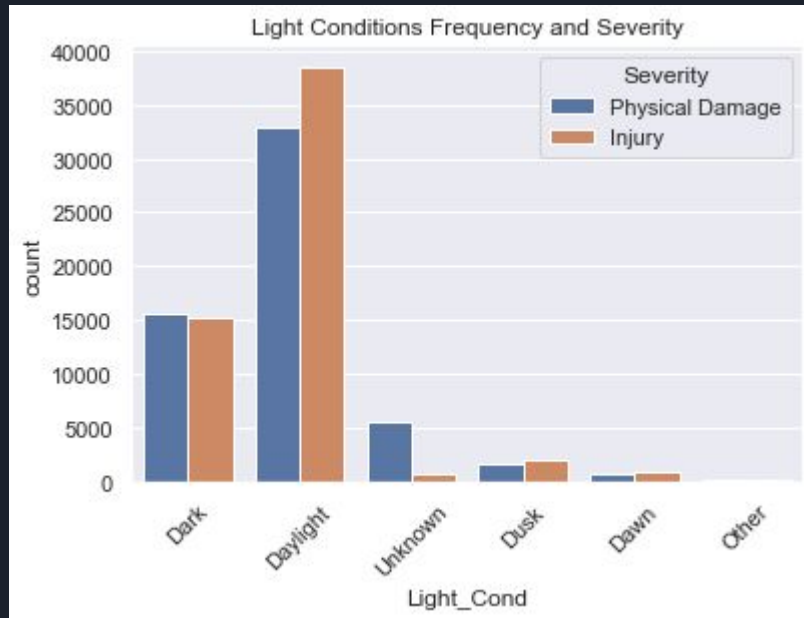
- The dataset regarding car collisions was provided by Coursera
- The data dates weekly from 2004 to present
- The data has been collected from the Seattle Department of Transportation
- This data set contains 194,673 rows and 37 attributes
- The data set is labeled with the severity of the accident - 1 or 2 indicating property damage or injury, respectively
- 3 features were used for classification, after One Hot Encoding to convert categorical data to numeric data



Weather Value Counts and Severity



Road Conditions Value Counts and Severity



Light Conditions Value Counts and Severity



Classification Models

	Jaccard	F1-Score	Log-loss
KNN	0.46	0.54	N/A
Decision Tree	0.47	0.53	N/A
Random Forest	0.48	0.54	N/A
Logistic Regression	0.47	0.54	0.67



Conclusion and Future Directions

- Built useful models to predict severity label of car collisions based on weather, road conditions, and light conditions
- Classification models have room for improvement
- Further analysis on this data set should be done to determine the impact of other independent variables on severity level
 - Collision type effect on severity level
 - Day of week
 - Hour of day
 - X, Y location of collision
- Machine Learning models take a while to run on Jupyter