The Severity of Car Accidents - How Can we Predict Danger?

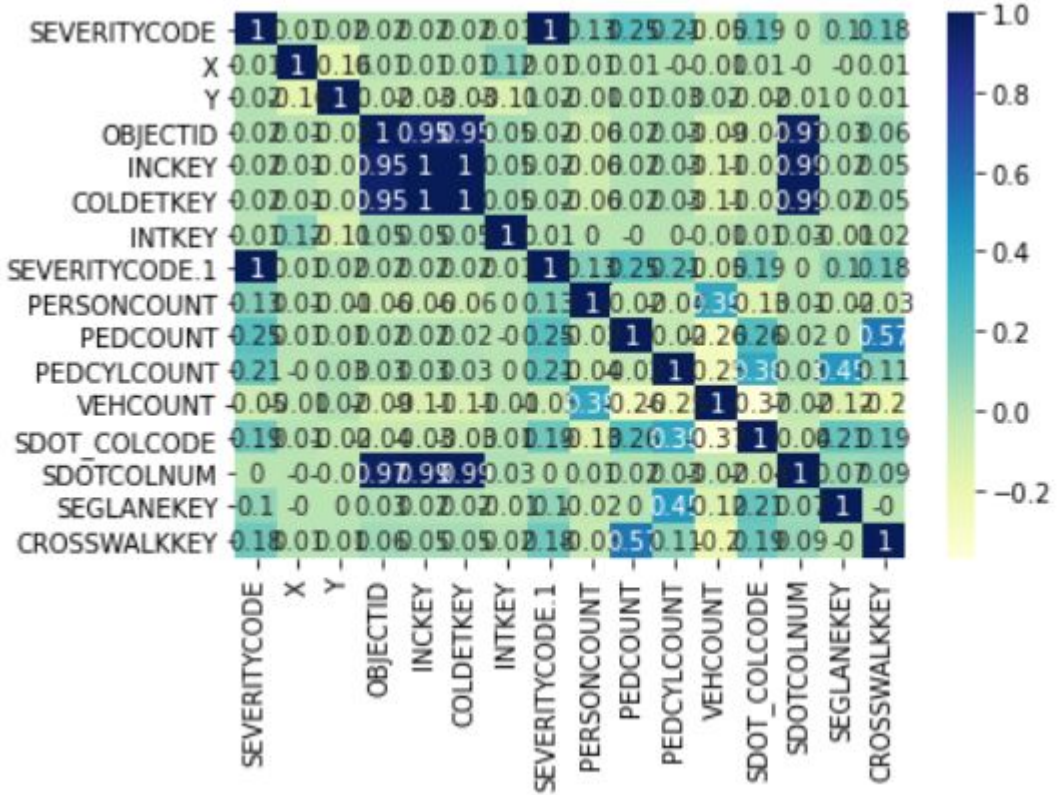
By Marcus Chiang
October 5, 2020

Car Accident Severity

- Car accidents happen frequently, but there are certain factors in these incidents that could help us predict the severity of any given accident
  - First check the correlation of different variables such as location and weather
- Using the Seattle Department of Transportation's (SDOT) data, we can determine how severe an accident could be
- This information would be useful for any driver

Data Acquisition and Cleaning

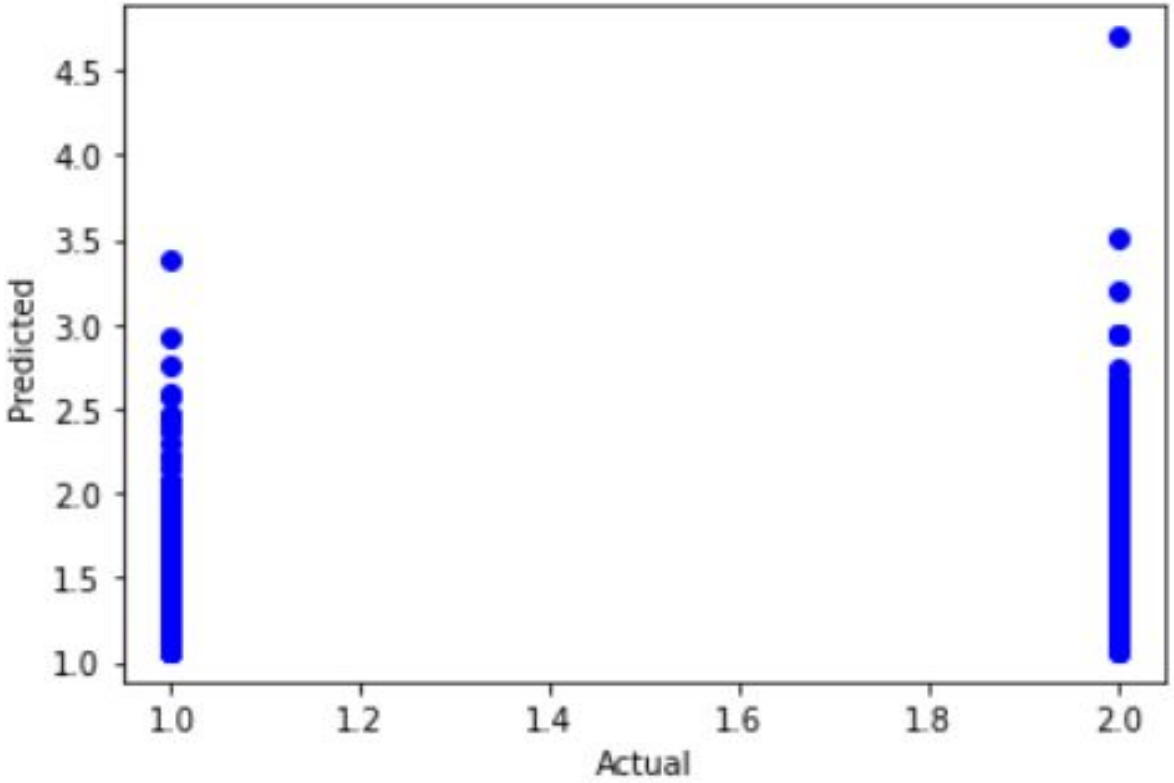
- The base data was proved by the Applied Capstone course
  - Data Collisions.csv
- Almost 200,000 rows of data and 39 columns for a total of 7,592,286 raw cells of data
- Dropped any NaN values by replacing them with the mean of the data column
- Only five columns were chosen in this model for a base accuracy - PEDCOUNT, PDECYLCOUNT, SDOT_COLCODE, PERSONCOUNT, CROSSWALKKEY



Correlation Chart Used to Determine Relevant Data Columns

Methodology

- Fit and defined X and Y to train the model
  - Initial RMSE and R2 scores were troubled and had to be fixed
- Proceeded to test X and Y models
  - Slightly improved scores
- Tested the model with several accidents documented in the Data Collisions dataset
  - Varying levels of accuracy
  - Due to an error, was forced to divide final prediction by 3 for better test scores
- Tested with several of my own accident scenarios similar to those that happened



Predicted and Actual Scores of my RMSE and R2 Graph

Conclusion

- The model was somewhat successful and could be improved and perfected
  - With more time, one could add the rest of the data columns for improved accuracy
    - Possibly even add other data types
- With enough data, the model could extend from just Seattle to entire states and possibly countries