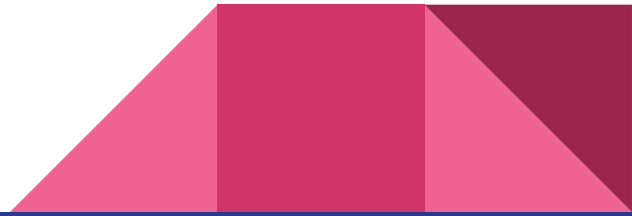
Chicago Car Crash Analysis

Avi Rubin

Outline

- Business Question
- Data
- Recommendations



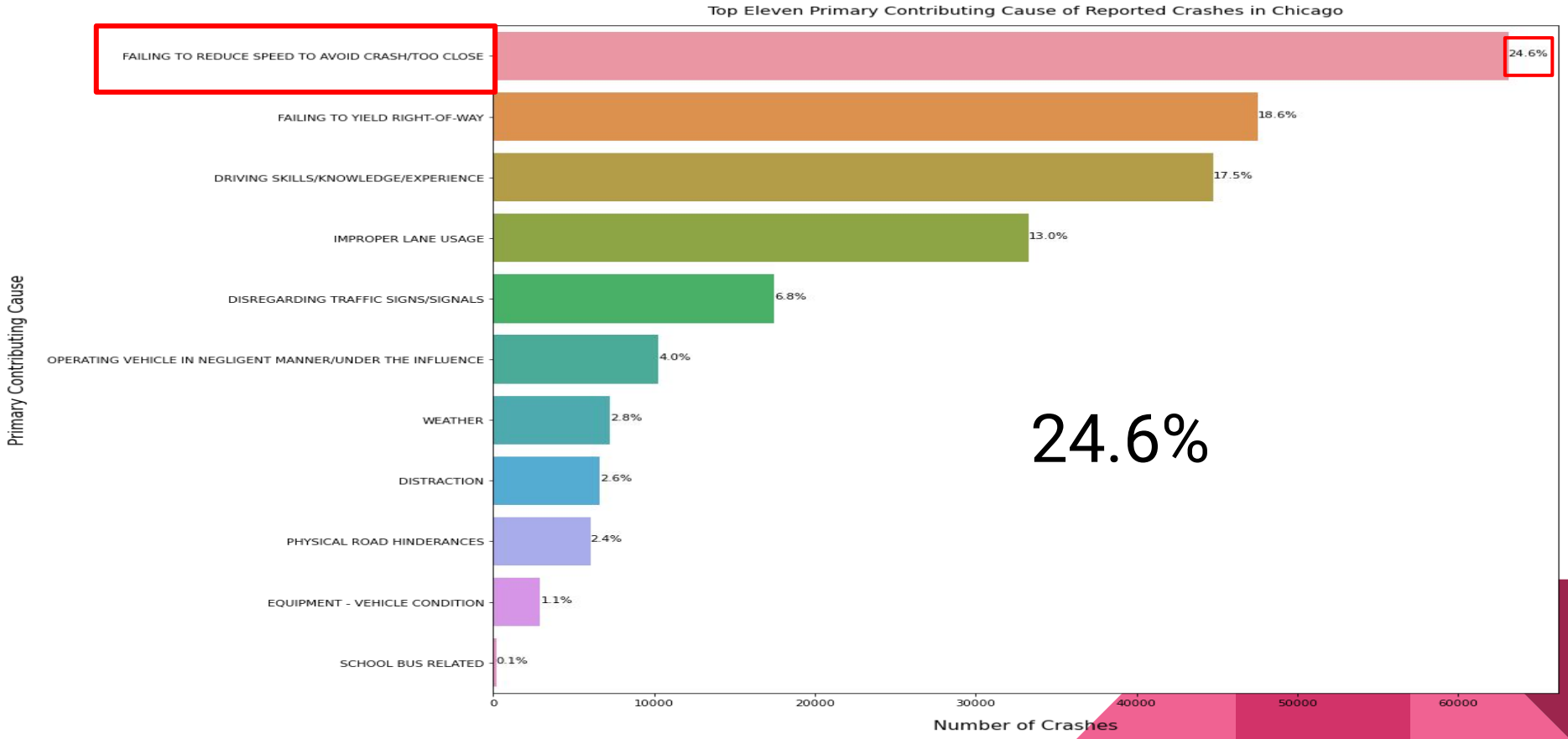
Business Question

1. Can we predict the primary cause of accidents to fill in the missing data?
2. Are we able to analyze that data to eliminate accidents and fatalities in the future?

Data

- Chicago, IL Car Crash Data
  - Obtained from the city of Chicago website
  - 3 data sets that work together and include:
    - Crash Information
    - Information on the People involved in the accident
    - Information on the Vehicles involved in the accident
  - Merge datasets, Cleaned and Manipulated data to work for our purposes
  - Created and tested models to predict classifications for the primary cause of the accidents
  - Obtain results/recommendations

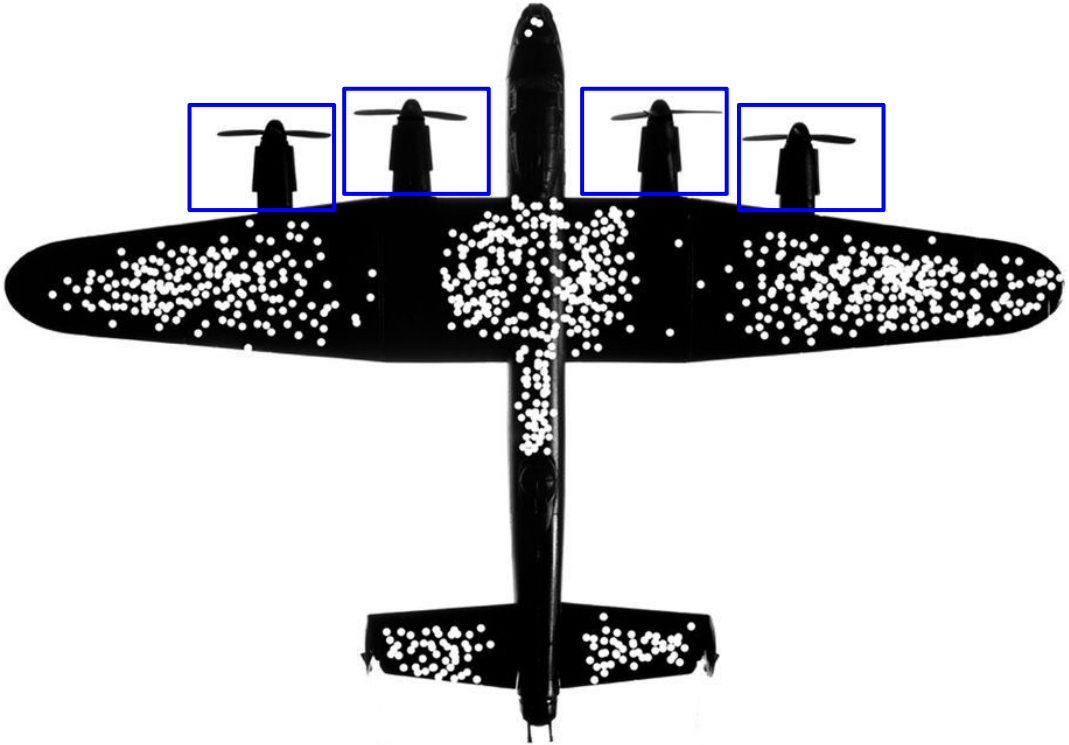
Top Eleven Primary Cause of Crashes in Chicago

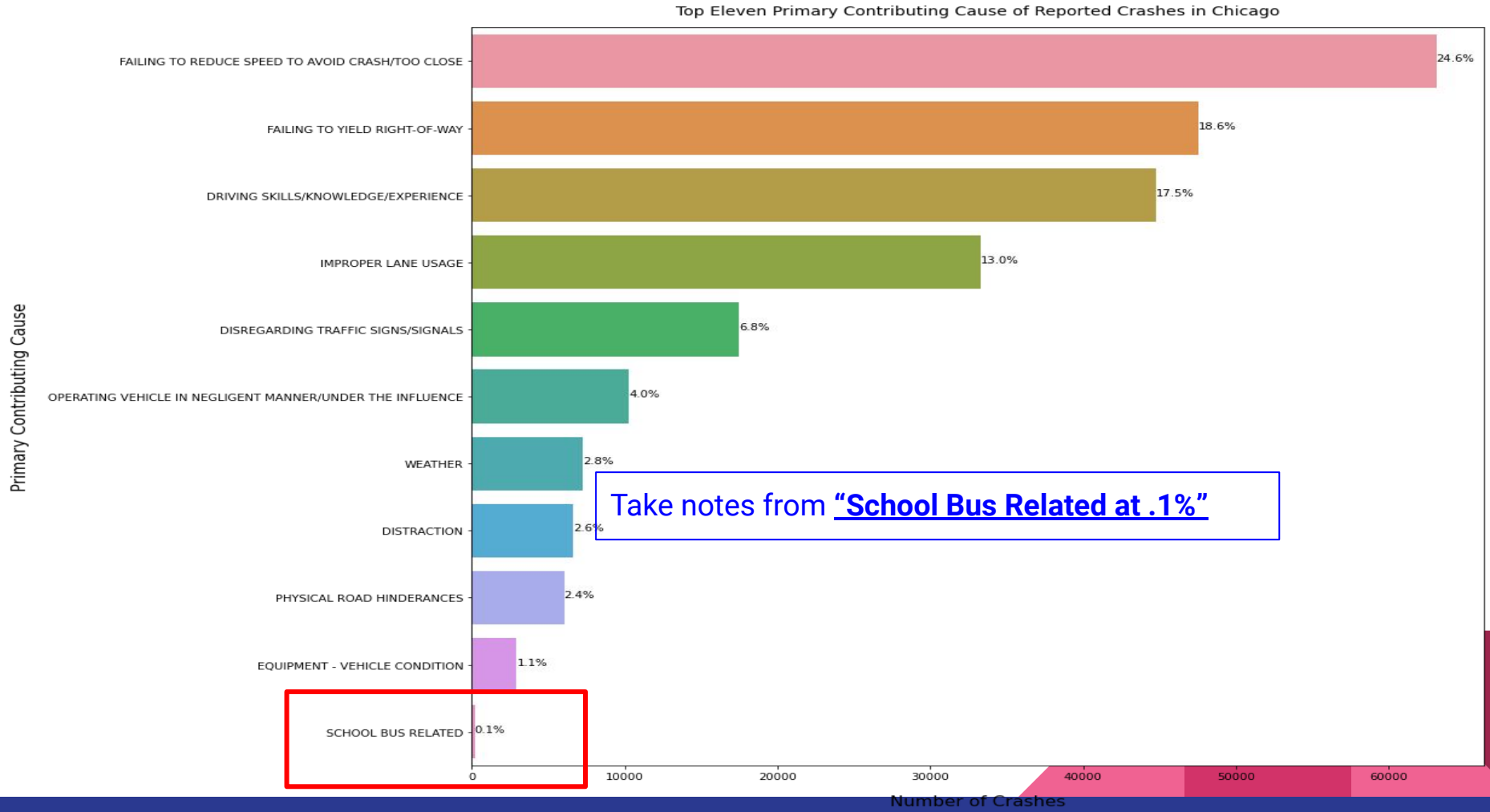


24.6%

WW2: Where to armor the planes?

Solution: Abraham Wald (Data Scientist?) said to Armor the planes where there were no bullet holes

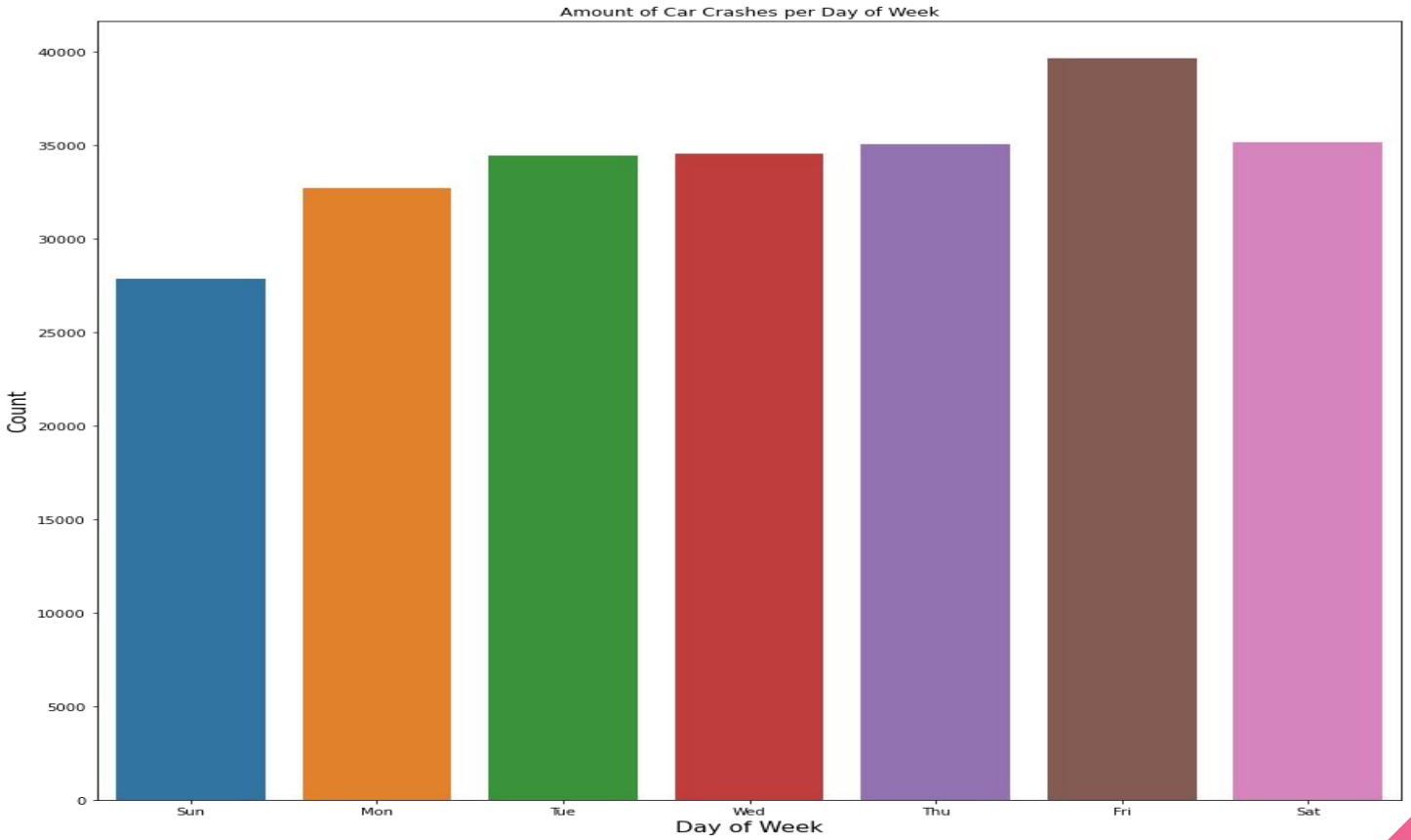




Take notes from **"School Bus Related at .1%"**

Recommended Solutions - Mimic School bus laws and characteristics

- Failed to yield to right of way was 18.6% of reported crashes
  - School busses are not allowed to make a right on red (included in the above category)
- School busses are big and painted bright yellow (maybe vehicle color could help?)
- Cannot overtake a school bus when their stop signs are out
- School busses typically travel at lower speeds so maybe lower the speed limit



Crashes are pretty evenly distributed except Sunday when school busses aren't around. So the school bus theory holds up.

Next Steps...

- See why Friday has the highest amount of crashes?
  - Driving while intoxicated?
- Explore which hours of the day have the most accidents
- Check what speeds accidents occur at most frequently

Models for Predicting Classifications

- Our best model for predicting all the classifications had an accuracy of 47%
- The best model was 56% accurate

- I would suggest in the future maybe combining some classes and we would be able to get a better model predictor and that will help us mitigate accidents in the future

Thank you

Avi Rubin - GitHub.com/ajr930