**Bayesian Workshop (Barry CA)**

Deliverables: 1 zipped file (Code + binned data file) + 1 report

Are smaller, lighter cars really putting passengers at substantially greater risk of injury or death? One could hypothesize that so many other factors influence the probability and severity of injuries, including highly advanced restraint systems, that vehicle size may ultimately not determine life or death.

**Data prep**

Target variable 🡪 OA\_MAIS

<https://support.bayesfusion.com/docs/GeNIe/learning_cleaningdata.html>

Explore (EDA)

1. Check missing values (volume) 🡪 DONE
2. Check unusual states 🡪 DONE
3. See distribution (histogram) 🡪 # of bins (DONE)

Currently defining the # of bins based on its normality. But unknown if it works that way.

Some cases have outliers

Act

1. Remove / fill in missing values 🡪 DONE
2. Discretize into bins 🡪 DONE

Notes

1. Delta V is the direction of the impact angle

**Models**

1. Naïve bayes 🡪 DONE
2. Tree bayes 🡪 DONE

**Results for each Network**

1. Default 🡪 DONE
2. Leave One Out
3. K Fold (10) 🡪 DONE
4. K Fold (20) 🡪 DONE

**Research**

<https://www.statisticshowto.datasciencecentral.com/choose-bin-sizes-statistics/>

<https://www.analyticsvidhya.com/blog/2016/03/tutorial-powerful-packages-imputing-missing-values/>