

Nova TCNJ Harrisburg Group

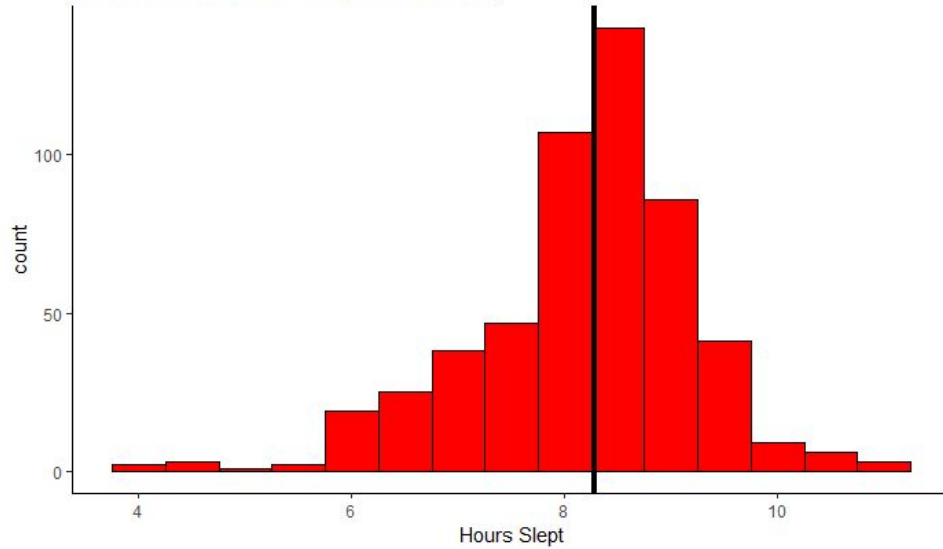
DataFest Philly 2019



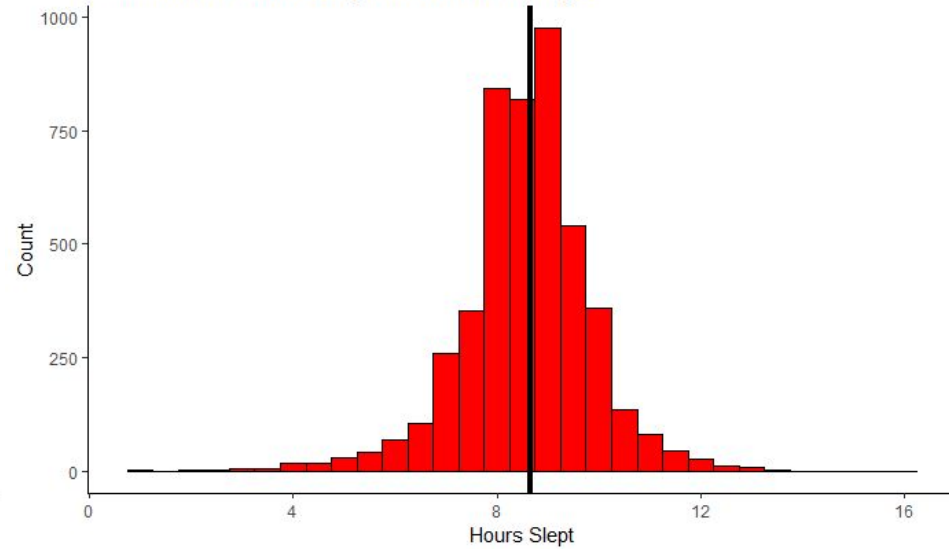
Daniel Brace, Thomas Holland, Emily Miller, and Alex Nguyen

Hours Slept On and Off Gamedays

Distribution of Hours Slept on Gameday



Distribution of Hours Slept on Non-Gamedays



Predicting Fatigue from Soreness and Sleep Quality

- Explored the correlations of 21 different variables with Fatigue.
- Sleep Quality and Soreness were the best predictors.

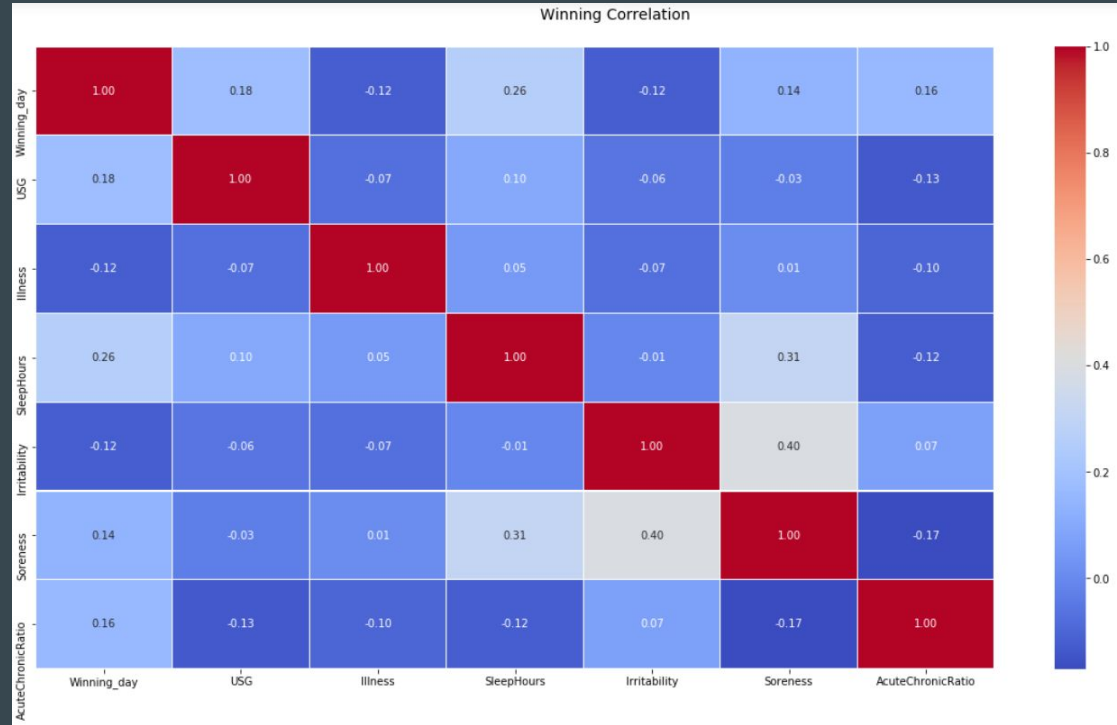
Linear Regression:	
MSE	0.579
GaussianNB:	
MSE	0.902
Cross Value Scores	
	0.609
	0.773
	0.545
	0.545
	0.545
RandomForest:	
MSE	0.563



Predicting a Win from Correlated Variables

- Wrangle data by taking game day average and apply this to wellness/RPE data.

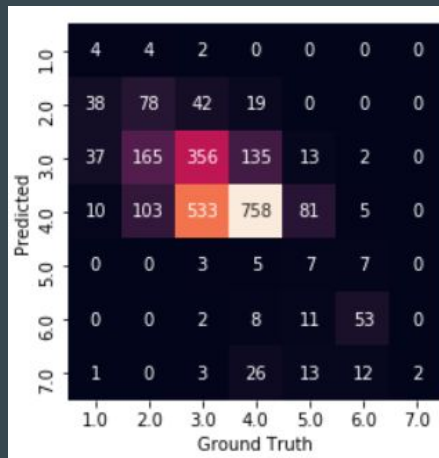
Linear Regression:	
MSE	0.174
GaussianNB:	
MSE	0.065
Cross Value Scores	0.344
	0.678
	0.629
	0.518
RandomForest:	
MSE	0.065



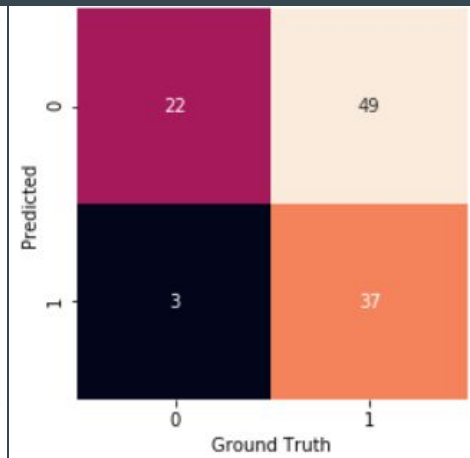
Future Research

- Record more accurate data for Winning games.
- More individual player reports through better estimations.
- Round Y_prediction matrix from decimals to 0 or 1.
- Compare results of other machine learning algorithms.
- Take into account opponent and score difference and explore any relation to fatigue.
 - ex: By how much did they win or lose? Was there a significant, average, or small difference in scores?

Predicting Fatigue from
Soreness and Sleep Quality



Predicting Winning from
Correlated Values



Thank you James and Spencer for talking to us about Rugby, It helped us to better understand the data.

Thank you Jacob, Timothy, and Kevin. Your guidance regarding data science was invaluable.