The dataset includes statistics for players who participated in the NFL Combine from 1999 to 2015. The combine is an event where players are measured at various strength and speed exercises such as the forty yard dash and bench press, along with measuring their weight, height, and more. The dataset also includes players' positions and alma mater which was used as categories for analysis. The dataset is from nflsavant.com which is an NFL data site and was created by the current Director of Baseball Research and Development for Major League Baseball (MLB) Daren William.

The final project analyzes relationships in the data such as the correlation between high vertical and fast forty yard dashes. A linear regression line is calculated with a steep slope confirming that players who jump high tend to run quickly. Players' strength is often exaggerated in the media and the actual population mean of NFL prospects bench press score is determined with a T confidence interval. Bias exists for SEC players because they have won the most bowl games lately and the percent of NFL Combine participants who are from the SEC is determined. The differences among positions is also analyzed in terms of weight and draft status. All offensive lineman do not have the same dimensions and weights for centers are illustrated and compared to weights for guards. Quarterbacks are typically taken in the first round of the draft but other positions are more likely to be chosen in the first 32 picks as well. If someone played multiple positions in college, it can be useful for them to declare as a certain position that has a better chance of getting drafted. Also as with most datasets, the empirical distributions can be modeled as known distributions. Vertical jumps fits well with the Weibull distribution and bench scores can be modeled with the normal distribution. Overall the analysis shows that a scout, coach, or NFL prospect can have an advantage or at least be better informed by statistically analyzing the NFL Combine dataset.