Using Advanced Predictive Modeling in a Nontraditional Discipline - The NFL Draft

Walker, Christopher (School: Henry W. Grady High School)
Rachwalski, Nathan (School: Henry W. Grady High School)

The purpose of this project is to determine if it is possible to effectively use traditional regression models to predict an event whose outcome has historically been unpredictable. In our case, we are analyzing the NFL Draft, but we hope our project could influence the way people attempt to analyze outcomes of all seemingly ambiguous decisions, not just those in sports. If we utilize historical NFL Combine and NFL Draft data we will be able to create many multiple linear regression models for each NFL position that demonstrate a correlation between performance in certain combine drills and where a player is taken in the NFL Draft. We started off by collecting data for all of the players using StatHead, a company geared towards recording statistics for American Professional Athletes, and used this data to develop spreadsheets. We used these spreadsheets as our data frames in R and were able to run multiple regression models and use packages such as the MuMIn package which allowed us to sort the multiple linear regression models by AIC. After selecting the combine events that correspond with our use of AIC, we ran multiple regression to generate results that can clearly link a player's performance in the NFL Combine to where they are taken in the draft. One example of this was our testing on the set of tight ends drafted in the last five years, a very small subset (about 40 players) of our larger set of data (about 1500 players).