



Boom or Bust: An Investigation of Rookie Running Back Performance Through the NFL Draft



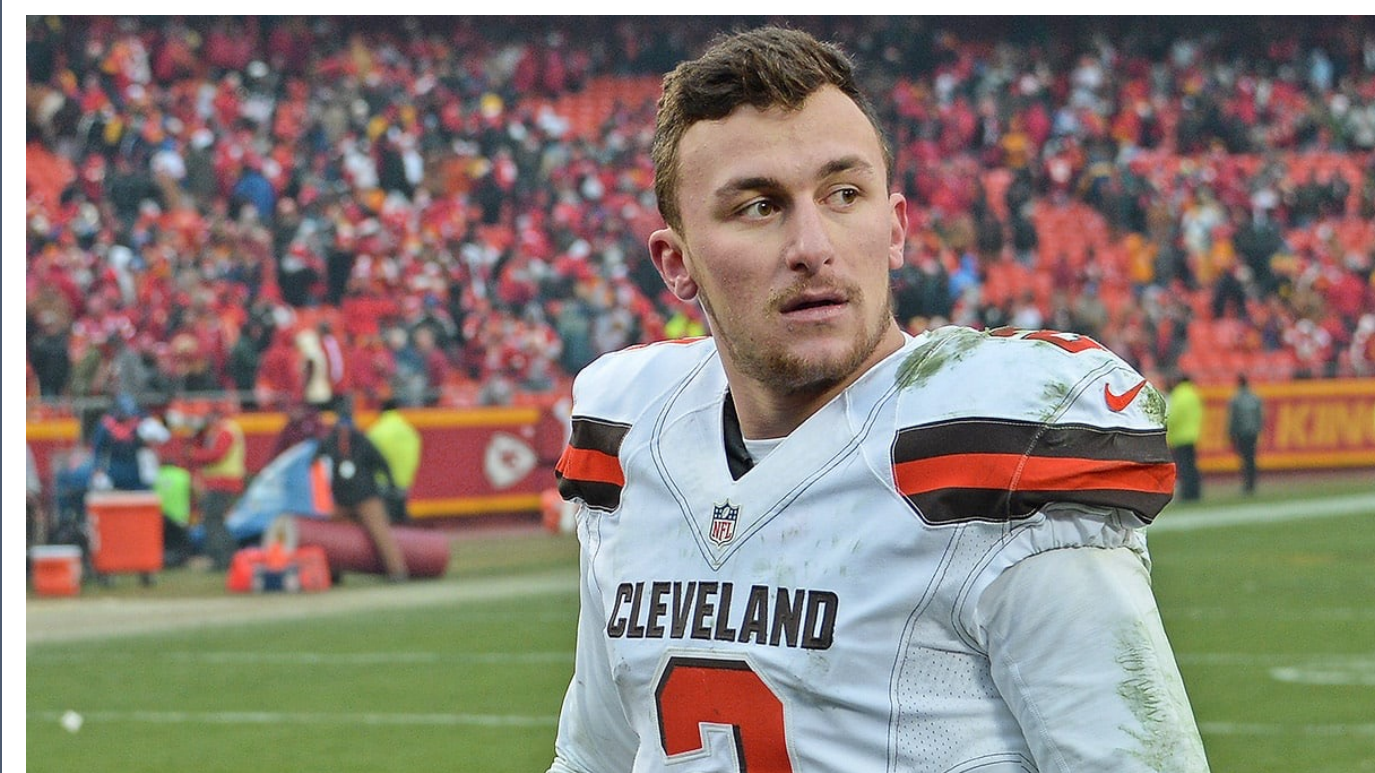
Shane Epstein-Petrullo Advised by Dr. Jillian Morrison
Statistics and Data Science

ABSTRACT

This study aims to investigate rookie running back performance in the NFL based on a variety of factors. The NFL draft is an event in which NFL executives and team decision-makers can assess and select players that they think can add value to their organization. However, it is often the case that an exceptional college player who is selected early in the draft and paid a significant amount of money underperforms in the NFL. The player would then likely be considered a "bust". The reason for this drop in playing quality is not solely due to athletic ability or football intelligence though. Each player is subject to random and unique circumstances, which has historically made assessing talent via the NFL draft so difficult for executives leading to "busts" time and time again. Some positions are much more subject to being "busts" and are considered volatile in this study. As such, the least volatile position is deemed to be the running back, who carries the football for as many yards as possible. Running backs also have the shortest career average of about 2.57 years. This study aims to investigate how random factors as well as performance statistics can be used to account for differences in rookie running back quality throughout the NFL draft.

Establishing Significance

- The National Football League (NFL) is the highest-grossing sports league in the U.S. Moreover, roughly 74.5% of Americans follow football.
- It is important to note that college football, specifically at the Division I level in the NCAA, also holds significant cultural and economic value in the U.S., and the NFL aims to garner new players from the top talent in collegiate competitions. The way in which these college players make it to the NFL is through the NFL draft.
- The NFL draft is an event that draws a significant amount of attention. Players from around the country are assessed and analyzed by a plethora of experts, team management leaders, coaches, and team owners. Based on collegiate athletic performance, team-specific workouts, and NFL combine performance, players' "draft stock" can rise, fall, or stay stagnant.
- **Boom:** A player who exceeds expectations despite the position in which they were drafted in.
- **Bust:** A player who underperforms and fails to meet expectations given the position they were drafted in.



Bust: Johnny Manziel

What the Data Shows: Prolific college career, speed, dynamic, playmaker.
What the Data Doesn't Show: Alcoholism, Depression, Drug Abuse, Unsustainable Lifestyle.

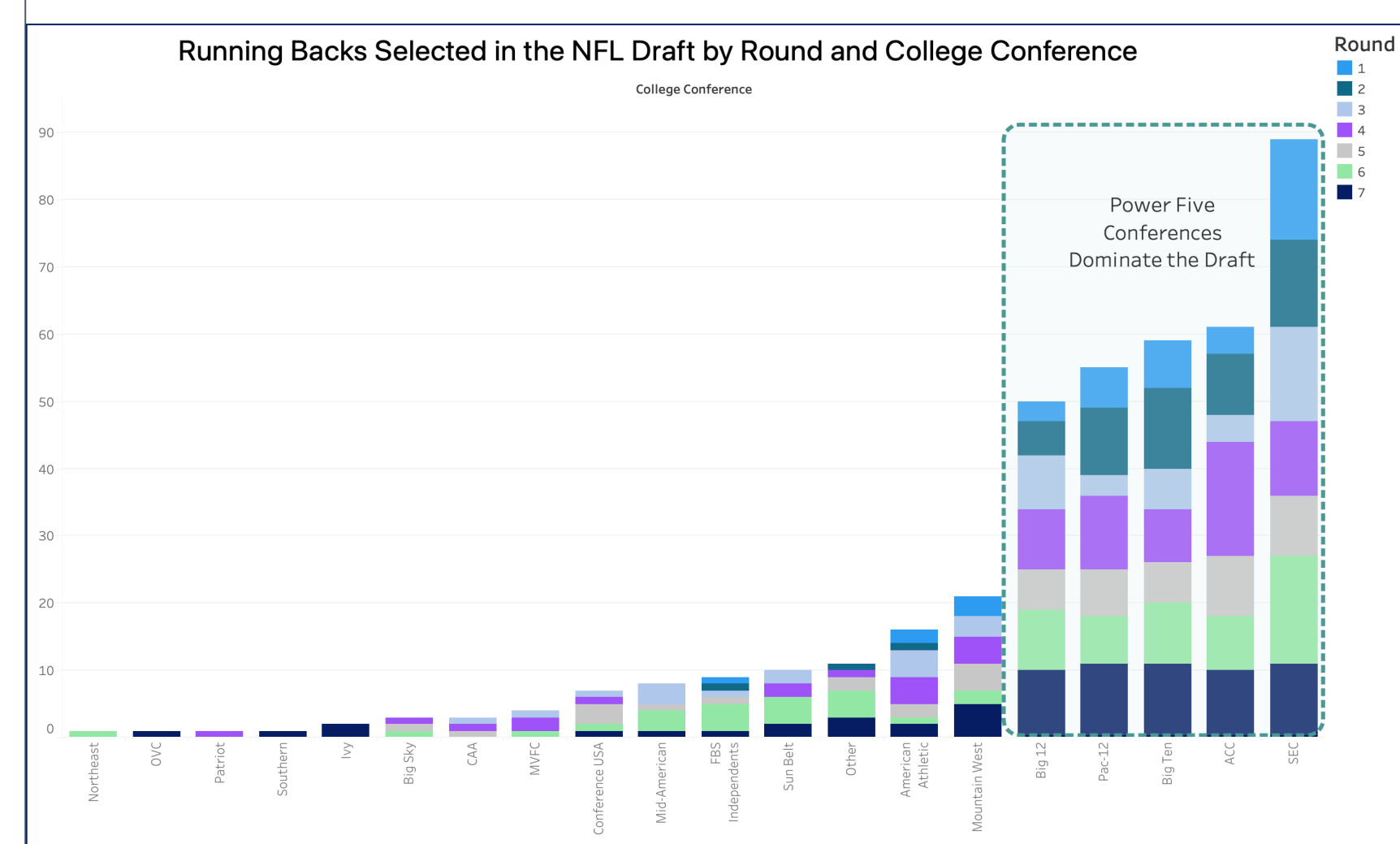


Boom: Trevon Diggs

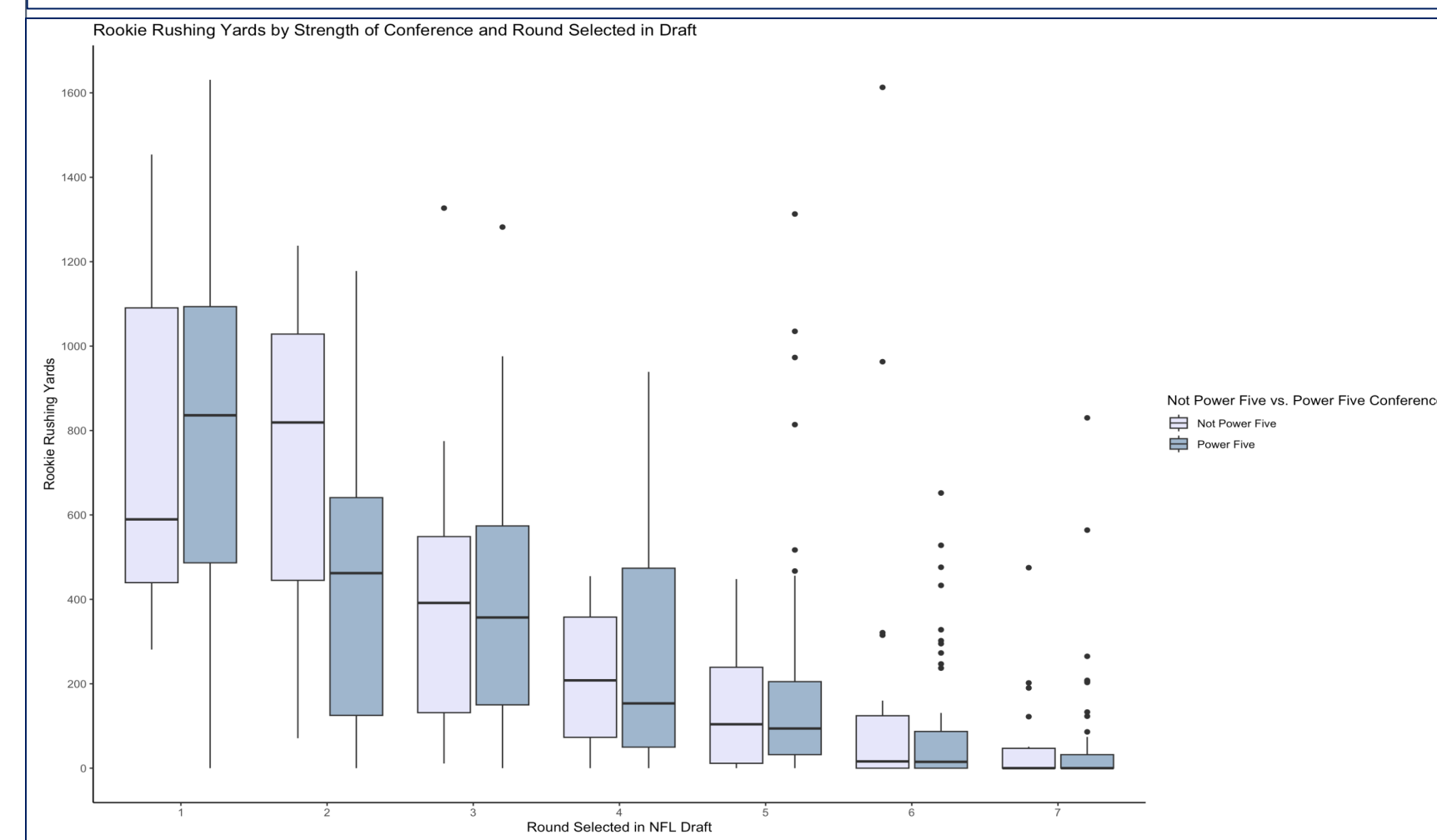
What the Data Shows: Not a significant contributor until senior year of college, position switch, Power 5 school.
What the Data Does Not Show: Work Ethic, Fatherhood, Motivation

Data Collection and Exploratory Data Analysis

The data set used for this study was collected via a variety of methods from various sources. Every running back drafted over the past 20 years was featured in this data set and variables regarding rookie year rushing statistics, college rushing statistics, NFL draft position, and contract variables were collected for each of the over 400 players included. The *cfbfastr* package was used to obtain the names of players over the course of the last 20 drafts. ProFootball Reference was used to obtain NFL statistics while each player's college statistics were manually input. Contract variables came from Overthecap and Spotrac.

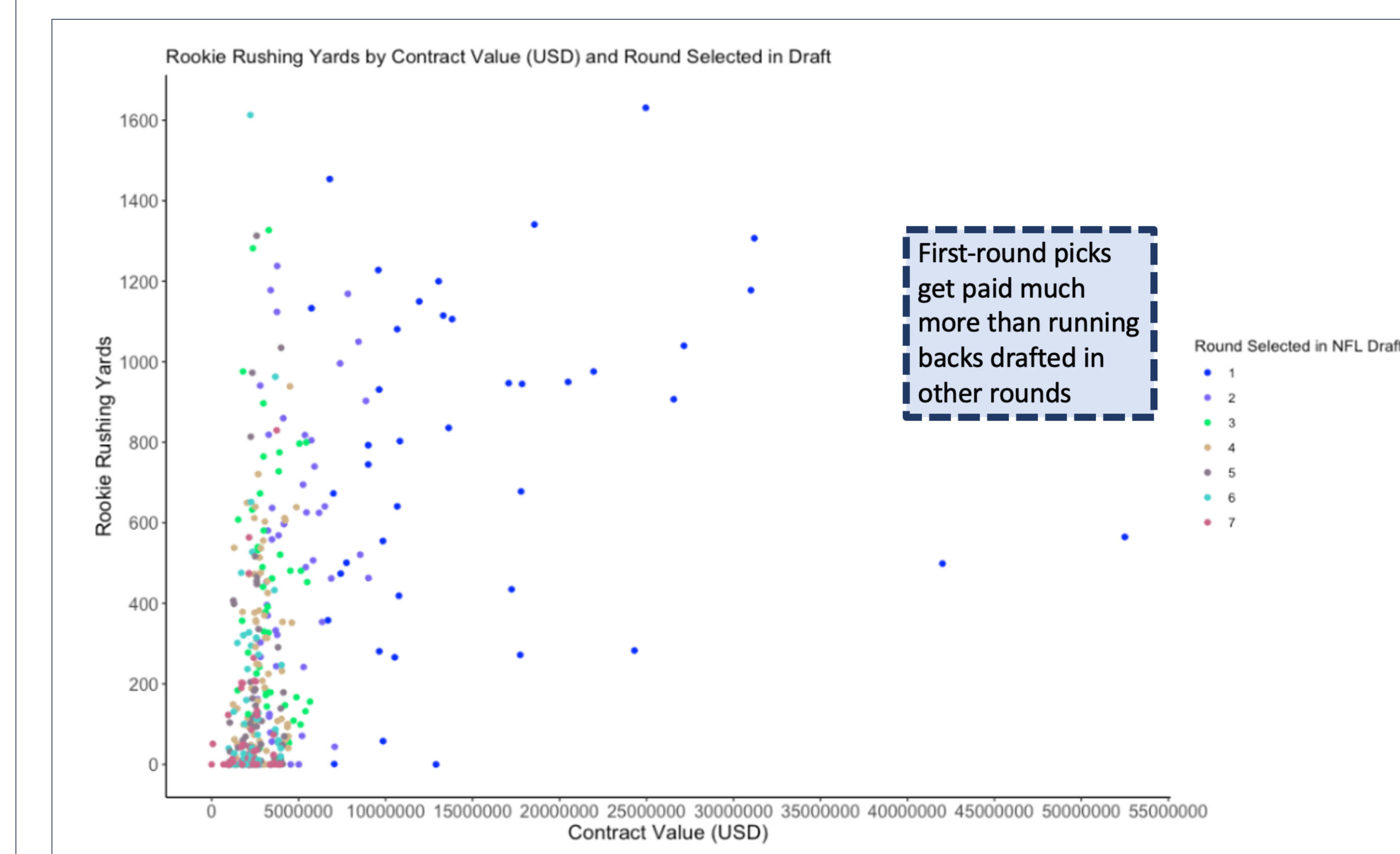


Running backs from the Power 5 tend to dominate the NFL draft, and NFL executives are largely the decision-makers when it comes to drafting new talent.



However, from the plot to the left, it can be observed that there does not seem to be a significant difference in rookie running back performance between running backs from the Power 5 versus running backs who do not come from the Power 5.

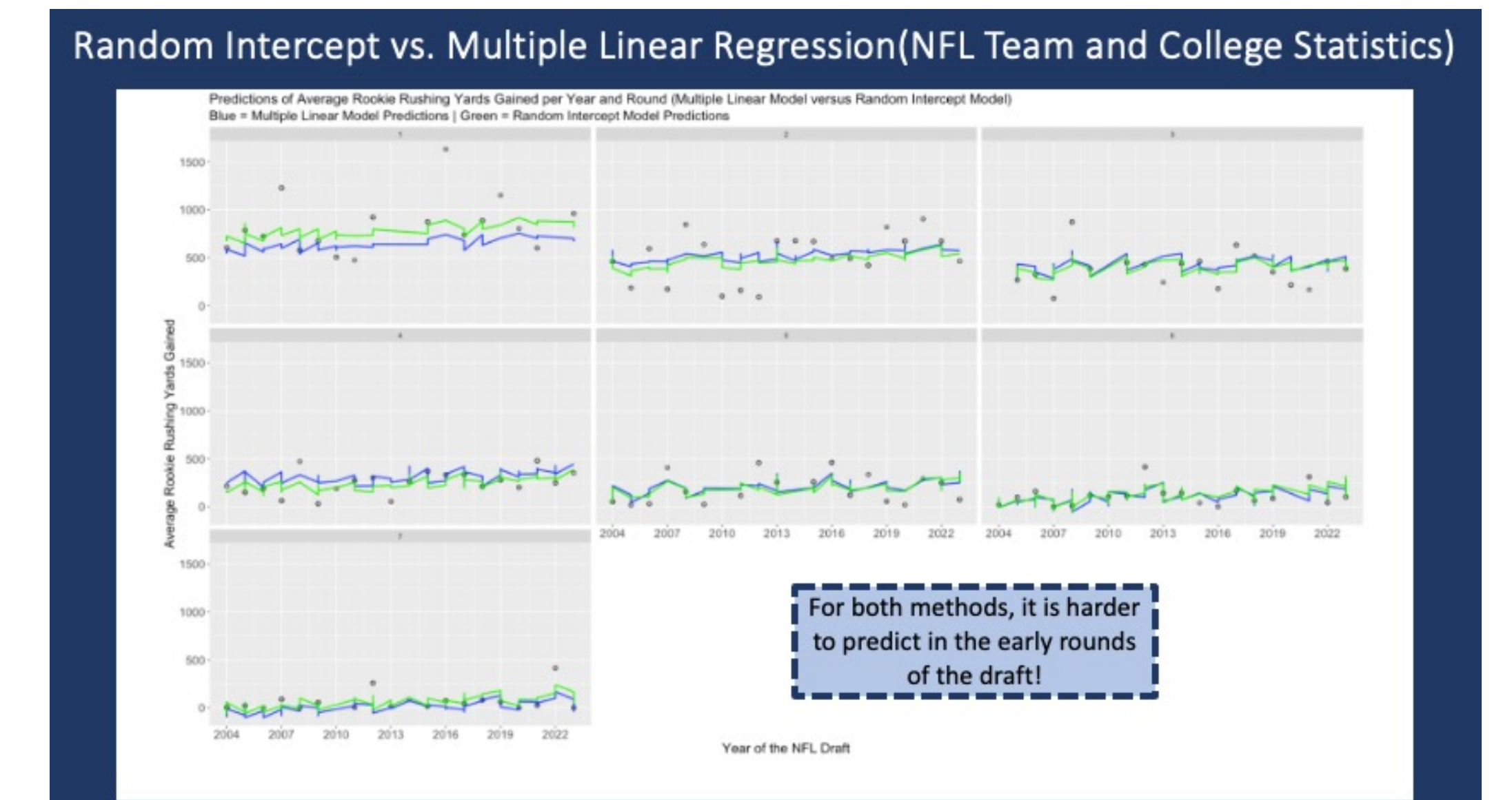
Exploratory Data Analysis



Despite being paid much more than any selection in the other rounds, first-round draft picks largely performed similarly to backs drafted in later rounds who were paid much less. These findings are consistent with those of Massey and Thaler 2013, who argued that executives overvalue high picks in the draft as the money spent on these high picks is not conducive to their on-field performance returns.

RESULTS

- Comparative Analysis between Fixed Effects Models and Random Effects Models
- Amount of Variance in Rookie Running Back Performance Explained by . . .
 - Fixed Effects Models: 39%-40%
 - Random Effects Models: 71%-72%
- The Numbers DON'T always say it all. Running backs are humans and subject to random factors such as personal life, injuries, motivation variances, and work ethic changes, which NFL executives cannot see from performance statistics alone.
- The Fixed Effects that had the most significant effect on performance were college statistics and the team a running back was drafted to. A random intercept was placed on the round of the NFL draft.
- There is variance in performance between rounds of the NFL draft, so coupling this information round by round is important!
- It is harder to predict performance in earlier rounds of the draft.



CONCLUSION, IMPLICATIONS, AND FUTURE DIRECTIONS

- **Conclusion:**
 - Spending a lot of money on an early-round running back is unwise.
 - NFL executives gather performance data well but fail to consider the random nature of human subjects (too much emphasis on prior performance and not enough on personal factors).
 - The team that a running back is drafted to as well as their college statistics are influential fixed effects.
- **Implications:**
 - Personal information about a player is not easily quantifiable data!
 - There is a lack of available data regarding contract variables and NFL combine stats before 2003.
 - The NFL has become increasingly more pass-dominant.
- **Future Directions:**
 - In the future, .txt files with investigations and reports of players' work ethic and attitude over the years can be useful.
 - Using more advanced machine learning models and AI can help to make predictions that will give the probability that a running back will be a bust or not.

ACKNOWLEDGEMENTS

Thank you to the entire Statistical and Data Sciences Department at the College of Wooster. I would like to acknowledge Dr. Jillian Morrison, who has been a monumental figure not just through the Independent Study process, but also as a mentor and person to me throughout my time at Wooster. She is always willing to navigate any problem with a smile on her face and the determination to help others. I would further like to acknowledge Dr. Drew Pasteur, whose guidance in the field of sports analytics was extremely helpful throughout this process. Dr. Pasteur has been an extraordinarily valued mentor to me who is there to provide guidance and wisdom from the chalkboard to the real world. Further, Dr. Pamela Pierce has instilled a love of learning and tackling difficult problems within me. She is always ready to help any student with a very impressive arsenal of jokes to help the learning process. Dr. Colby Long has been an engaging and enthusiastic professor as well as an interested second reader. Thank you, Dr. Long.

I would like to further thank Mr. Brian Fischer and Ms. Jenna Welsh who first introduced statistical research to me.

I am forever grateful to the Wooster Men's Lacrosse Team and Nate Huston, who have been my constant for the past four years. The support from each and every one of my teammates and coaches is overwhelming, and I am proud of the team we have become.

Thank you to my parents and those closest to me who have helped me and supported me throughout this process. Thank you Grace for the overwhelming support and reassurance. Thank you Taemour for the long nights of hard work. Thank you Ben, Vinny, Henry, Oliver, Rowan, Casey, Abby, Jaden, Alvaro, and Luke. I cannot imagine this process without all of you. Thank you Fungai Jani, who helped me find the idea for this Independent Study project.