

# **Evaluating First-Round NBA Draft Picks: Bust or Success?**

## **DSC 405 (001) Web Scraping Project**

### **Mehul Suji and Landon Rao**

#### **Project Summary**

Annually, young players enter the NBA Draft, with a select few drafted in the first round, leading to higher pay and fame. These players are expected to excel, but many are deemed busts. This project aims to determine the success rate of first-round picks by comparing their game stats to league averages. Success is defined as stats equal to or greater than the average, while lower stats indicate a bust. This analysis, conducted via Python web scraping, provides valuable insights for fans and teams about the potential value of their rookies.

#### **Source**

Website: Basketball Reference

Url: [https://www.basketball-reference.com/leagues/NBA\\_2024\\_totals.html](https://www.basketball-reference.com/leagues/NBA_2024_totals.html)

Limitations: N/A

Explanatory Variable: Statistics of the first-round picks, which include Points per Game, Rebounds per Game, Assists per Game, Steals per Game, and Turnovers per Game.

Response Variable: Status of the first-round picks, which is determined based on whether their statistics are greater than equal to, or lower than the league's first-round draft pick's averages. This status is categorized as either 'successful' or 'bust'.

Information Status: HTML

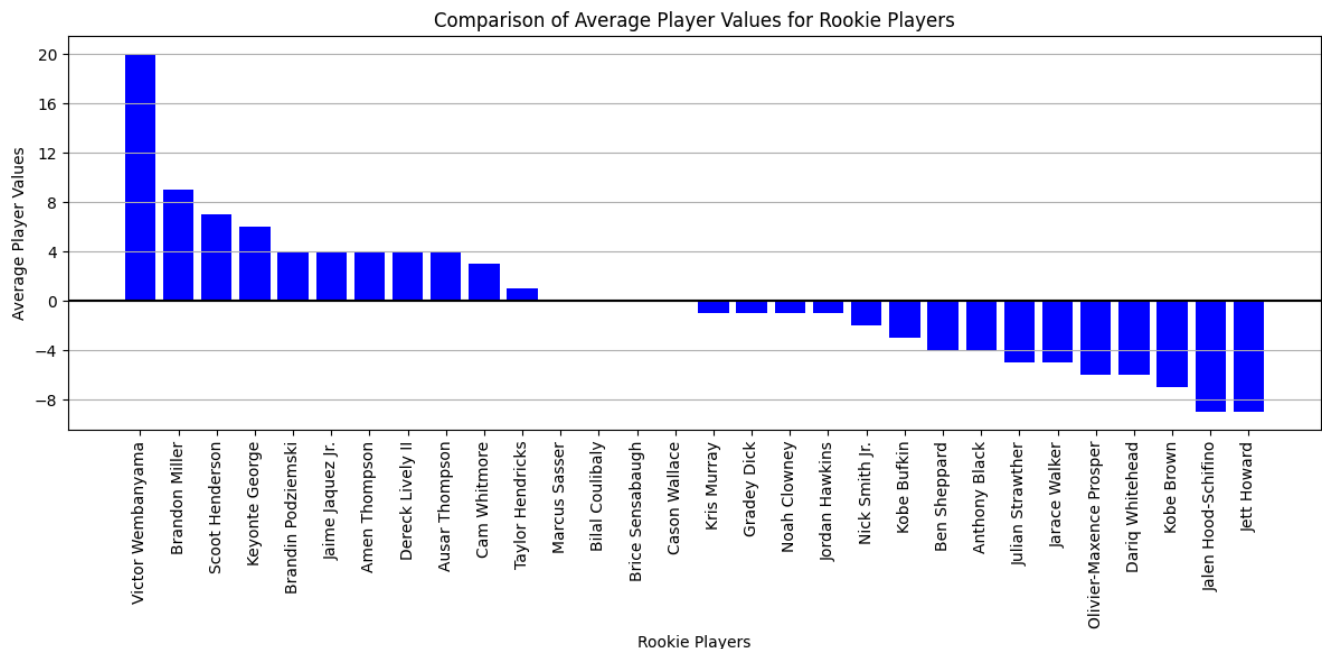
#### **Data Moves:**

1. Import necessary libraries and load the website's HTML data. This is the initial setup where you bring in the tools you need and the data you'll be working with.
2. Transform the HTML data into a usable data frame. This involves finding the appropriate HTML tags and classes that contain the data you're interested in, and converting that data into a structured format.
3. Select the columns needed for analysis. The columns "Player", "G", "PTS", "TRB", "AST", "STL", and "TOV".
4. Combine rows with the same player names. This step consolidates data for each player if they appear in multiple rows.
5. Calculate per-game statistics. Using the columns "PTS", "TRB", "AST", "STL", and "TOV", create new columns 'Points per Game', 'Rebounds per Game', 'Assists per Game', 'Steals per Game', 'Turnovers per Game' by dividing the values into these columns by the number of games played ('G' column).
6. Calculate average statistics. Create new columns 'Average Points per Game', 'Average

Rebounds per Game’, ‘Average Assists per Game’, ‘Average Steals per Game’, and ‘Average Turnovers per Game’ by finding the means of the ‘Points per Game’, ‘Rebounds per Game’, ‘Assists per Game’, ‘Steals per Game’, ‘Turnovers per Game’ columns.

- Visualize the data. Create bar charts comparing the rookie players and their ‘Points per Game’, ‘Rebounds per Game’, ‘Assists per Game’, ‘Steals per Game’, and ‘Turnovers per Game’.
- Calculate and visualize differences. Calculate the difference between the players’ per-game statistics and the average statistics, and plot these differences on a bar chart. This will give you a visual representation of how each player compares to the average.

## Findings:



From the findings, we can conclude that Victor Wembanyama, Brandon Miller, Scout Henderson, Keyonte George, Brandin Podziemski, Jamie Jaquez Jr., Amen Thompson, Dereck Lively II, Ausar Thompson, Taylor Hendricks, Marcus Sasser, Billy Coulibaly, and Brice Sensabaugh are all playing at a level equal or above the league average hailing them as a ‘Success’ according to our parameters. On the other hand we found that Cason Wallace, Kris Murray, Gradey Dick, Noah Clowney, Nick Smith Jr., Kobe Bufkin, Ben Sheppard, Anthony Black, Julian Strawther, Jarace Walker, Oliver-Maxence Prosper, Dariq Whitehead, Kobe Brown, Jalen Hood-Schifino, and Jett Howard are all playing at a level below league average hailing them a ‘Bust’ according to our parameters.