ISyE 3030 Project: Analyzing Player Performance and Team Success in the WNBA: Draft Impact and Seasonal Contributions

Team members' names: Sydney Bice, Johanna Lumbantobing

I.   Introduction

The goal of this project is to analyze how draft choices impact player performance and team success in the WNBA. Specifically, the project aims to identify if the draft round and pick correlate with key performance indicators for players and consequently, the overall team success in subsequent seasons. This project was inspired by the recent WNBA draft for the 2024 season.

II.   Data Collection and Preprocessing

This project utilizes publicly available data sourced from two main websites, across the timeline and basketball-reference. There were four types of data collected: draft data, player awards data, player performance data, and playoff performance data.
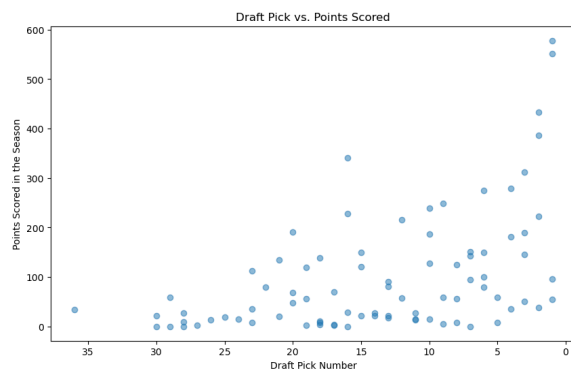
1. Draft Data (2019-2023): This dataset includes comprehensive details of WNBA drafts from 2019 to 2023, such as draft round picks, player, and team. Citation: Across the Timeline. (n.d.). WNBA Drafts. Retrieved April 20, 2024, from https://www.acrossthetimeline.com/wnba/drafts.html#season=2020
2. Player Award Data (2020-2023): This dataset includes Major WNBA awards from 2020 to 2023, with entries labeled by year. Citation: Across the Timeline. (n.d.). WNBA Awards. Retrieved April 10, 2024, from the specific URL listed above.
3. Player Performance Data (2020-2023) from Basketball Reference: This dataset includes detailed player statistics for WNBA seasons, with each year's data in separate sheets. Citation: Basketball Reference. (2023). WNBA Player Totals. Retrieved April 10, 2024, from https://www.basketball-reference.com/wnba/years/2023_totals.html
4. Playoff Performance Data (2020-2023) from Basketball Reference: This dataset includes detailed playoff performance stats for WNBA teams. Citation: Basketball Reference. (2021). WNBA Playoffs. Retrieved April 10, 2024, from https://www.basketball-reference.com/wnba/playoffs/2021.html

The bulk of the preprocessing involved merging datasets that were originally separated into a single database. The CSV files of data were loaded into pandas DataFrames,and each DataFrame was individually cleaned before being merged together on similar data columns. Additionally, for each player in this dataset, their playoff participation during their draft year was checked and recorded. Awards data was then added by matching players and their draft year awards, appending this information to the dataset. Duplicate entries and unnecessary columns were removed to finalize the dataset.

III. Descriptive Statistics

The table below presents the descriptive statistics for various metrics for drafted WNBA players. The measures include the mean, standard deviation, minimum, first quartile (Q1), median, third quartile (Q3), and maximum value.

	Mean	Standard Deviation	Minimum	Q1	Q2 (Median)	Q3	Maximum
Games Played	19.809524	11.494867	1	10	20	30	40
Round	1.654762	0.702732	1	1	2	2	3
Number Pick	13.66667	8.637734	1	6.75	13	19.25	36
Age	22.202381	0.9666671	19	22	22	23	24
Games Started	5.595238	10.043423	0	0	0	6	40
Minutes Played	294.392857	281.500199	2	62.5	193.5	439.5	1249
Field Goal	36.928571	45.184437	0	5	21	56	233
Field Goal Attempts	90.464286	101.547834	0	16.5	57.5	115.5	509
Free Throw Attempts %	0.74	0.185262	0	0.625	0.75	0.8527	1
Offensive Rebounds	15.595238	21.962794	0	1	6.5	18.75	125
Defensive Rebounds	37.440476	47.494667	0	5	19.5	40.25	210
Personal Fouls	30.83333	29.413477	0	6.75	22.5	47.75	125
Points	100.869048	121.289237	0	15	57	143.75	578
Team Rank	9.357143	3.743727	1	7	9	12	19

Visual 1: Scatter plot of draft pick number against season points for drafted players, displayed a moderate negative correlation of -0.485. This indicates that players drafted earlier (lower pick numbers, on the left due to the inverted axis) generally score more points than those drafted later, suggesting that factors beyond draft order also significantly impact scoring.

Visual 2: Heatmap of the correlation matrix between draft pick number and key performance metrics (points, total rebounds, assists), revealed varying degrees of correlation. This showed a moderate correlation between points and total rebounds, points and assists, and total rebounds and assists. This indicated that recently drafted players that score more points also tend to collect more rebounds and assists.
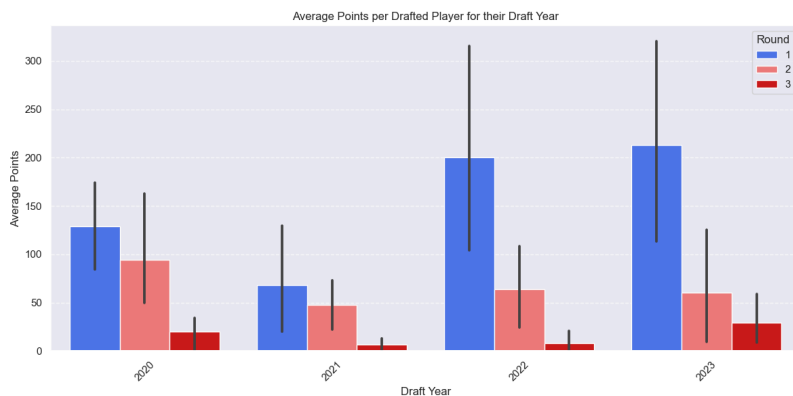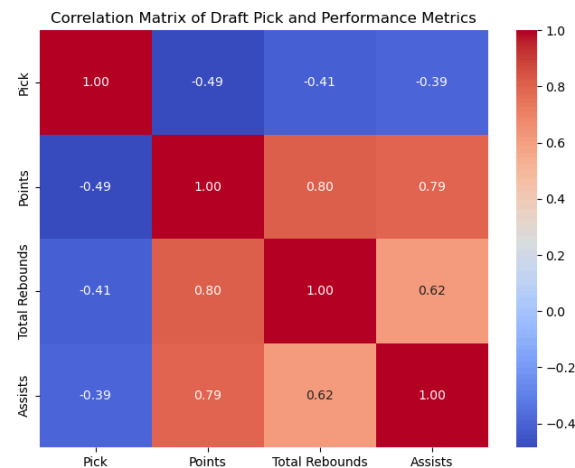
Visual 3:
Bar graph showing the average points scored by drafted players for their respective draft years, segmented by draft round.There were 24 players drafted in 2020, 19 drafted in 2021, 22 drafted in 2022, and 19 drafted in 2023. This shows that players drafted in round 1 were more likely to score more points than players drafted in later rounds. It also showed that in 2022 and 2023 there was a greater division between number of points scored for players drafted from round 1 compared to rounds 2 and 3 than in 2020 and 2021.This disparity indicates potential shifts in drafting strategies or player performances post-pandemic.

IV. Statistical Inference

- Null Hypothesis (H0): The average points scored by players from top college programs[1] are equal to the average points scored by players from other college programs.
- Alternative Hypothesis (H1): The average points scored by players from top college programs are not equal to the average points scored by players from other college programs.

▼	Top College Programs	Other College Programs
Sample Standard Deviation	142.426	83.44907
Sample Mean	142.2	63.2954

Degrees of Freedom	82
T-Value	3.05886
Critical Value	1.96
P-Value	0.003284

We chose to use a t-test because we are comparing the means of two independent samples (points scored by players from top college programs and points scored by players from other college programs). From the t-test, it was found that the p-value is 0.003284, which is less than the alpha value of 0.05. Because of this, we can reject the null hypothesis and say that there is a statistically significant difference between the mean points scored by players from top college programs and other college programs.

V. Regression Analysis

We performed a linear regression analysis to model the relationship between a team's average draft pick number and the team's rank for that season. A lower ranking correlates with better team performance. To the right is a graphical representation of the data. The final regression equation is ŷ = 8.495 - 0.098x. Additionally, R-squared = 0.046. The 8.495 represents the y-intercept which suggests that teams that have an average draft pick number of 0 with rank at 8.495 in that same season. The -0.098 is the slope of the equation. This means that for every increase of 1 in average draft pick number, team rank will decrease by 0.098, or perform better, that season. The

Team's Season Rank vs. Average Draft Pick Number

regression line generally fits the observed data points, but our residuals are quite large. Because our R-squared value is close to 0, it suggests that the average draft pick number is not strongly related to the team's season rank. Overall, there are many other factors that are affecting a team's season rank.
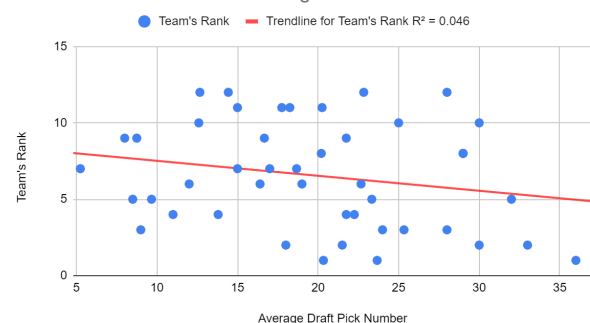
IV. Conclusion

This project investigates the impact of draft choices on player performance and team success in the WNBA, particularly examining correlations between draft round/pick and performance metrics. Inspired by the 2024 WNBA draft, our analysis consistently shows that players selected earlier tend to exhibit stronger performance, especially in scoring. Regression analysis suggests that a team's average draft pick isn't strongly linked to its season rank, prompting further exploration into factors like the number of draft picks and their performance. However, our study is limited by its focus on specific variables and lacks consideration of contextual factors such as coaching and team dynamics. Future research could explore these aspects to offer a more comprehensive understanding of team success and inform nuanced recruitment and performance strategies.

[1] According to the most recent AP Poll rankings from https://sportsdata.usatoday.com/basketball/wncaab/ap-poll, the top 25 women's college basketball programs are as follows: South Carolina, Iowa, UConn, North Carolina State, USC, LSU, Texas, Oregon State, Stanford, UCLA, Notre Dame, Indiana, Baylor, Gonzaga, Colorado, Ohio State, Duke, Virginia Tech, Kansas State, Syracuse, Oklahoma, Utah, Creighton, West Virginia, and Iowa State.