Metrics Impacting NBA Player Contract Values

Executive Summary

Background:

An NBA agent hypothetically hired our team to analyze the factors impacting contract values. The agent will earn a 5% commission for each contract; higher contract values result in higher commissions for the agent. Both the agent and the players stand to gain from a growing NBA market.

Objectives:

- A specific list of factors likely to increase or decrease contract values.
- Expected contract values of newly signed players based on college performance.
- A list of schools to focus recruitment efforts.

Data Sources:

Explain the sources

Data Manipulation Methods:

A blurb data cleaning steps

Analysis:

A blurb on insights that includes visuals

- Two levels (EDA – Descriptive)
- Regression coef blurb 1
- Regression coef blub 2
- Comparison of regression output

Limitations:

A blurb that explains limitations

Inspiration

To Help Players and Agents Capitalize On A Growing NBA Market

Background:

For this project, we have created a hypothetical scenario in which an agent of an NBA player management firm hired us to deliver a report on the leading factors impacting an NBA player's guaranteed contract value. The agent receives a 5% commission on each contract he negotiates; the higher the value, the more money he earns. The agent will use the analysis to advise players and prioritize performance improvement metrics.

Context:

Like most other sports, the NBA sets a salary cap that determines how much a team can spend on salaries. Since some franchises reside in more significant, popular cities, they generate more revenue than smaller market teams. The cap aims to keep the game competitive by limiting the amount of money big market teams can spend on their roster of players. The cap is partly determined by the league's total revenue, which has doubled in the last decade. As the NBA makes more money, the players and their respective agents stand to gain financially. Figure 1. on the right shows salary projections reaching more than 80 million annually by 2029.

Similar Studies:

An article titled "NBA Player Salary Analysis based on Multivariate Regression Analysis" was published by *Highlights in Science, Engineering and Technology*. The authors conducted a similar study but focused on ways to balance disproportionate salaries between the players. This study gave us an idea of what variables to investigate for our project. Additionally, we referred to a prior study from Koki Ando, who also conducted a regression analysis, but his goal was to predict a player's future salary; from that study, we got an idea of how to structure our experiment.

Objectives:

For this project we aim to deliver the following to the agent:

- A specific list of factors likely to increase or decrease contract values.
- Expected contract values of newly signed players based on college performance.
- A list of schools to focus recruitment efforts.

NBA Supermax Salary Projections

SEASON	SALARY CAP PROJECTION	MAX SALARY
2023-24	$136.021 mil	$47.6 mil
2024-25	$149.623 mil	$52.37 mil
2025-26	$164.585 mil	$57.6 mil
2026-27	$181.044 mil	$63.37 mil
2027-28	$199.148 mil	$69.7 mil
2028-29	$219.063 mil	$76.67 mil
2029-30	$240.969 mil	$84.34 mil

Figure 1. Table of supermax salary projections. *NBA salaries keep going up. Prepare to have your mind blown in the future* by Mike Vorkunov, 2023, The Athletic. https://theathletic.com/4740069/2023/08/03/nba-salary-cap-rise-jaylen-brown/

Name	Description & Important Variables	Size	Format	Links
Player Career Stats		419 KB		https://github.com/swar/nba_api/blob/master/docs/nba_api/stats/endpoints/playercareerstats.md
Player IDs		99 KB	Pandas DF	https://github.com/swar/nba_api/blob/master/docs/nba_api/stats/static/players.md
Draft History				https://github.com/swar/nba_api/blob/master/docs/nba_api/stats/endpoints/drafthistory.md
NBA Teams		312 B		https://github.com/swar/nba_api/blob/master/docs/nba_api/stats/static/teams.md
Contract Values		42 KB 476 rows		2023-24 NBA Player Contracts
College Stats		~12.4 MB		SR-CBB

Data Manipulation Methods

Clean and Prepare NBA Data for Regression Analysis

How did we need to manipulate the data:

The default file type for the player contract data is .xls, and this generates a value error because pandas cannot determine the file format. In this case, we opened the downloaded file, renamed it, and saved it as a .xlsx file. Additionally, we executed the following manipulations:

- Convert player names to lowercase characters.
- Replace accented characters with their English versions.
- Split the player column into first and last name columns.
- Rename columns to be more representative.
- Drop duplicate rows.

The player ID dataset needed the following manipulations:

- Convert first and last name columns to lowercase.
- Convert the player ID column to string datatype.
- Update specific player names to allow for joining.

The career stats needed the following manipulations:

- Convert the season column to datetime truncated to year.
- Drop rows corresponding to the 2024 season.
- Limit career stats to the last five years.

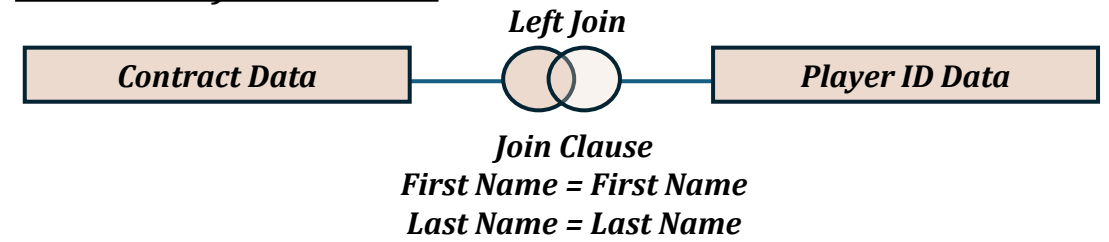
Challenges:

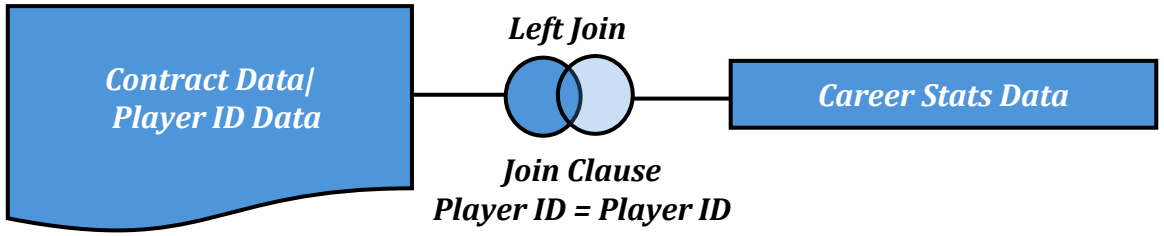
Our data sources contained accented and non-accented names.

How did we address data quality concerns and challenges:

The 2024 season is in progress and contains a partial season's metrics. We removed rows corresponding to the year 2024 from the dataset. Our dataset did contain null values because one player did not have a contract value, and several players are rookies competing in their first season this year; these players were dropped from the dataset because they did not fit the scope of our data goals. Several players needed name corrections because our data sources recorded names differently. The basketball reference website stores a player's name with accented characters, while the NBA_API stores the name without the accents. This was the most challenging data quality issue because it prevented the accurate joining of the two datasets. We addressed this using code that corrected specific names.

How did we join the data:





COLLEGE STATS MANIPULATION

- DATA RETREIVAL
- FILTERING
- CLEANING
- MERGING
- QUALITY ASSURANCE

- * link notebook

NBA Exploratory Analysis

How did we need to manipulate the data:

Refer to .py file docstring

How did we data quality concerns:

- Refer to .ipynb mardown cells
- Missing
- Incomplete
- Incorrect

How did we perform conversion steps:

Refer to .ipyb notebook

How did we join the data:

Discuss the variables and steps used to join the two data sources to perform the analysis

Challenges:

Describe the challenges we encountered and how we solved them.

NBA REGRESSION ANALYSIS WITH VISUAL(S)

- * link notebook

COLLEGE REGRESSION ANALYSIS WITH VISUAL(S)

- * link notebook

COMPARISION

- A
- * link notebook

CONSIDERATIONS & LIMITATIONS

- D

References & Statement of Work

placeholder text

References

- Vorkunov,M. (2023). Vorkunov: NBA salaries keep going up. Prepare to have your mind blown in the future. *The Athletic*. https://theathletic.com/4740069/2023/08/03/nba-salary-cap-rise-jaylen-brown/
- Feng, X., Wang, Y., & Xiong, T. (2023). NBA Player Salary Analysis based on Multivariate Regression Analysis. *Highlights in Science, Engineering and Technology*, *49*, 157-166. https://doi.org/10.54097/hset.v49i.8498
- Ando, K. (2018). NBA Players' Salary Prediction Using Linear Regression Model. https://rstudio-pubs-static.s3.amazonaws.com/371407_e21330910f3c4bd2b6e19440013ea793.html#