# NBA Correlation Analysis

This repository contains R scripts designed to analyze correlations and relationships between various player and positional statistics in NBA games. The analyses focus on understanding how different stats (like points, assists, steals, etc.) relate to each other, both at the player level and the positional level. These insights can be particularly valuable for making informed decisions in sports betting, especially in over/under betting scenarios.

## Files

### NBA Correlation Analysis by Player.R

This script performs a detailed correlation analysis between individual NBA players. It calculates standardized Spearman correlations and adjusted \( R^2 \) values between various stats (e.g., points, rebounds, assists) for all possible pairs of players. The analysis includes:

- \*\*Standardized Correlations\*\*: Standardizing stats to Z-scores to account for different levels of variability across metrics.

- \*\*Adjusted \( R^2 \)\*\*: Provides insight into how well one player's performance in a particular stat (e.g., points) predicts another player's performance in a different or the same stat (e.g., assists).

\*\*Use Case\*\*: This analysis can help identify player pairs whose performances are closely linked. For instance, if a player tends to score high points when another player assists a lot, betting on the over for both players in games where they play together might be advantageous.

### NBA Correlation Analysis by Position.R

This script extends the correlation analysis to the positional level, aggregating player stats by position (e.g., guards, forwards, centers). It calculates both correlations and adjusted \( R^2 \) values between different positional stats across all games in the dataset. The analysis includes:

- \*\*Position-Based Correlations\*\*: Evaluates how the performance of one position (e.g., guards) correlates with another position (e.g., centers).

- \*\*Adjusted \( R^2 \)\*\*: Determines how well the performance of one position explains the variance in another position’s performance.

\*\*Use Case\*\*: This analysis can be used to understand how different positions interact during games. For example, if a team's guards' performance in assists is highly correlated with the points scored by their forwards, this insight can guide betting strategies focusing on team dynamics.

## Application in Sports Betting

These analyses are especially useful for sports bettors who focus on over/under betting markets. By understanding which player or positional stats are most closely linked, bettors can make more informed decisions about where to place their bets. For instance, if the analysis shows a strong correlation between a specific player's assists and another player's points, bettors might consider placing bets on both players exceeding their expected performance in a game where they both play significant minutes.

Overall, these scripts provide a robust statistical foundation for making data-driven decisions in NBA sports betting.

## Limitations

While the analyses provided in these scripts offer valuable insights, there are several important limitations to consider:

1. \*\*Data Range\*\*: The dataset used in these analyses is limited to NBA games played between October 2023 and March 2024. This relatively short time frame means that the correlations and relationships identified may not fully capture long-term trends or player performance variability across entire seasons or multiple seasons.

2. \*\*Sample Size\*\*: Due to the limited time frame, the number of games and data points available for analysis is constrained. This can impact the robustness of the correlations and adjusted \( R^2 \) values, particularly for players or positions that may have played fewer games during this period.

3. \*\*Applicability to Other Sports\*\*: The analysis is specific to NBA basketball, where the dynamics of player interactions and positional play are unique. While similar analytical methods could be applied to other sports (e.g., football, baseball, soccer), the nature of these sports could lead to different types of correlations or relationships. It could be beneficial to explore these methods in other sports to identify whether stronger or more useful correlations exist.

4. \*\*Contextual Factors\*\*: The analysis does not account for contextual factors such as injuries, changes in team dynamics, or specific matchups that might influence player or positional performance on a game-by-game basis. These factors can significantly impact the predictive power of the correlations identified in the data.

Given these limitations, while the insights derived from this analysis can be useful for guiding sports betting decisions, they should be considered as part of a broader strategy that includes other data sources and considerations. Expanding the dataset to cover more extensive time periods and exploring similar analyses in other sports could provide more comprehensive and potentially more accurate insights.