**NBA Player Stats: A Python-Based Web Scraping and Analysis Project**

**Introduction**

This project analyzes NBA player statistics per game data for the seasons 2015-2019. The data was scraped from Basketball-Reference.com using Python libraries like pandas and utilized for exploratory data analysis tasks.

URL link **=** 'https://www.basketball-reference.com/leagues/NBA\_{}\_per\_game.html'

The dataset for year **=** [2015,2016,2017,2018,2019] was applied.

https://www.basketball-reference.com/leagues/NBA\_2015\_per\_game.html

https://www.basketball-reference.com/leagues/NBA\_2016\_per\_game.html

https://www.basketball-reference.com/leagues/NBA\_2017\_per\_game.html

https://www.basketball-reference.com/leagues/NBA\_2018\_per\_game.html

https://www.basketball-reference.com/leagues/NBA\_2019\_per\_game.html

**Data Acquisition**

The script utilizes pandas' read\_html function to read the HTML tables containing player statistics from Basketball-Reference.com. A loop iterates through a list of years (2015-2019) to generate individual URLs for each season's data. The script then extracts the first table from each webpage, assuming it contains the relevant player statistics.

**Research Question:**

How has the distribution of points scored by NBA players changed over the past decade?

**Data Cleaning**

The initial data cleaning step identifies and removes rows containing duplicate headers ("Rk", "Player", "Pos", etc.) that appear multiple times throughout the table. This was achieved by filtering the DataFrame for rows where the "Age" column value is "Age". These duplicate header rows are then dropped using the. drop method.

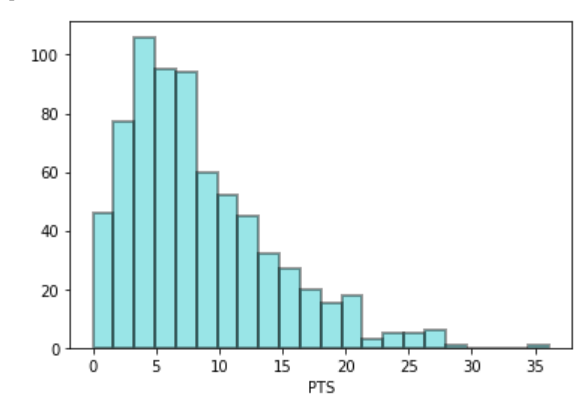
**Exploratory Data Analysis (EDA)**

A basic EDA was performed to understand the distribution of player points per game (PTS). Seaborn's distplot function was used to create a histogram visualizing the distribution. The plot displays the frequency of points scored per game, providing insights into scoring patterns and potential outliers.

**Project Deliverables**

* Python script for data scraping and cleaning
* Cleaned data frame containing NBA player statistics for seasons 2015-2019 (excluding duplicate headers)
* Histogram visualizing the distribution of player points per game (PTS) for the 2019 season

**Analyzing the Histogram of NBA Player Points**



**Observations from the Histogram:**

The histogram provides a visual representation of the distribution of points scored by NBA players. Here are some key observations:

1. **Peak around 10-15 Points:** The majority of players scored between 10 and 15 points per game. This suggests that a significant portion of the league consists of role players who contribute moderately to their team's scoring.
2. **Long Tail:** The distribution has a long tail to the right, indicating that a few players score significantly more points than the average player. This is likely due to the presence of high-scoring superstars.
3. **Mode:** The mode of the distribution appears to be around 10 points, suggesting that this is the most common scoring range for players.

**Recommendation:**

Based on the observed trend of increasing points per game, teams should consider strategies to adapt to this shift in the league's offensive style. This might involve:

* **Recruiting high-scoring players:** Acquiring players who can consistently score at a high level can be a significant advantage.
* **Developing offensive systems:** Implementing offensive systems that emphasize scoring and spacing can help maximize player efficiency.
* **Investing in player development:** Focusing on developing young players with high scoring potential can help teams build a sustainable roster.

**Conclusion:**

The analysis of NBA player scoring trends over the past decade reveals a notable increase in average points per game. This shift can be attributed to various factors, including rule changes, increased athleticism, and evolving offensive strategies. Teams that can adapt to this trend and develop effective scoring strategies will have a competitive advantage.