

NBA Player Performance Analysis

Ian Sellers & Ryan Brown

1. Introduction and Data Description

This project focused on analyzing NBA player performance data from 2012 to the present, with the goal of uncovering key insights into player efficiency, team dynamics, and modern basketball trends. The primary data source for this analysis was an NBA player dataset obtained from Kaggle, which included important metrics such as.

- Field Goal Percentage (FG%)
- 3-Point Percentage (FG3%)
- Free Throw Percentage (FT%)
- Points (PTS)
- Rebounds (REB)
- Assists (AST)
- Steals (STL)
- Blocks (BLK)
- Minutes Played (MIN)

This Kaggle dataset provided a comprehensive and reliable view of player and team performance across multiple seasons.

Preprocessing Steps.

Using Python's Pandas library, we handled missing values and standardized data formats to ensure consistency across different seasons. We calculated new metrics such as Player Efficiency Rating (PER) and Offensive Contribution to provide deeper insights into performance. We filtered the data to focus on regular season performance, excluding playoff data to maintain consistency in the analysis.

2. Methods of Analysis

We approached the analysis with several key questions in mind.

1. Which teams have consistently led in points, rebounds, and assists over the past decade?

2. Who are the most efficient players based on a combination of shooting accuracy, points, and assists?
3. How do 3-point shooting trends correlate with overall team performance?
4. What factors contribute most to a player's overall efficiency (PER)?

Fields of Use.

The data was analyzed across several fields.

1. Investigating the balance between scoring and assists.
2. Exploring how different teams achieve high performance metrics like rebounds and assists.
3. Analyzing player consistency in shooting percentages across multiple seasons.
4. Ranking players by their efficiency using custom calculations of PER, incorporating metrics like points, assists, and turnovers.

Resulting Output.

The analysis produced several key visualizations and results.

- A breakdown of top-performing teams by points, rebounds, and assists.
- A ranked list of the most efficient players, such as Giannis Antetokounmpo and Nikola Jokic.
- A correlation matrix showing the relationship between different performance metrics (e.g., points and assists, field goals made and points).

3. Program Description

The program, written in Python, consisted of several key processes to ensure comprehensive analysis.

1. The first step involved importing the Kaggle dataset into a Pandas DataFrame, addressing missing values, converting columns to the correct data types, and normalizing key metrics for comparison across seasons.
2. We calculated custom metrics such as Offensive Contribution (points + assists) and Player Efficiency Rating (simplified version), which included factors like points, assists, rebounds, steals, blocks, and turnovers.
3. Using Matplotlib and Seaborn, we created multi-line and bar charts to visualize top team performance across multiple seasons. For example, bar charts were used to display the top 5 teams in terms of points, rebounds, and assists, while line graphs were employed to show player efficiency trends over time.

4. We computed the correlation between different performance metrics using Pandas' `corr()` function and visualized the relationships using heatmaps. This allowed us to see how strongly metrics like 3-point shooting correlated with overall points scored.

4. Results and Analysis

The results of the analysis provided several clear insights into player and team performance trends.

- Teams like the Golden State Warriors (GSW), Oklahoma City Thunder (OKC), and Houston Rockets (HOU) consistently ranked among the top performers in terms of points, rebounds, and assists. GSW, in particular, led in both points and assists, highlighting their offensive efficiency.
- Players like Giannis Antetokounmpo and Nikola Jokic consistently ranked at the top in terms of PER, thanks to their ability to contribute across multiple areas like scoring, rebounding, and playmaking.
- We identified DeAndre Jordan as a top performer in field goal percentage, leading the league in FG% for several consecutive seasons, reaching a high of 71.4% in the 2016-17 season.
- The correlation matrix revealed a strong positive correlation (~ 0.99) between field goals made and points scored, as well as a moderate positive correlation between assists and points (~ 0.79). This reinforces the importance of playmaking in modern basketball.
- We observed that while 3-point shooting has a strong correlation with points, not all top scorers are reliant on it. Players like Stephen Curry and James Harden are heavily reliant on 3-point shooting, while others like Shai Gilgeous-Alexander employ a more traditional scoring style.

5. Group Roles and Contributions

The project was completed by two team members, Ian and Ryan. Ian focused on the data preprocessing, ensuring data integrity and consistency across the entire dataset. Ian also took lead on developing custom metrics, such as Player Efficiency Rating (PER) and Offensive Contribution, using Python. Ryan focused on data visualization and statistical analysis. Ryan was responsible for creating the correlation matrices and heatmaps, as well as building the visual storytelling through Seaborn and Matplotlib charts. Both Ian and Ryan collaborated on the final analysis, ensuring that the insights derived from the data were consistent, accurate, and relevant to modern NBA trends.

6. Conclusion

Through our analysis of NBA player performance data from Kaggle, spanning from 2012 to the present, we were able to uncover valuable insights into how teams and players achieve success on the court. The data highlighted the growing importance of 3-point shooting in modern basketball and showcased the dominance of certain players in terms of efficiency and offensive contribution.

- Golden State Warriors, Houston Rockets, and Oklahoma City Thunder have consistently demonstrated their offensive prowess, ranking among the top teams in multiple metrics.
- Players like James Harden, Giannis Antetokounmpo, and Nikola Jokic continue to lead the league in both traditional and advanced metrics, proving their versatility and value in multiple aspects of the game.

This analysis, powered by Python, provided a deep dive into the evolving strategies and player performances shaping the modern NBA, offering a comprehensive view of the league's dynamics.