

# Clustering Professional NBA Player Based on Performance



Hongfei Mei

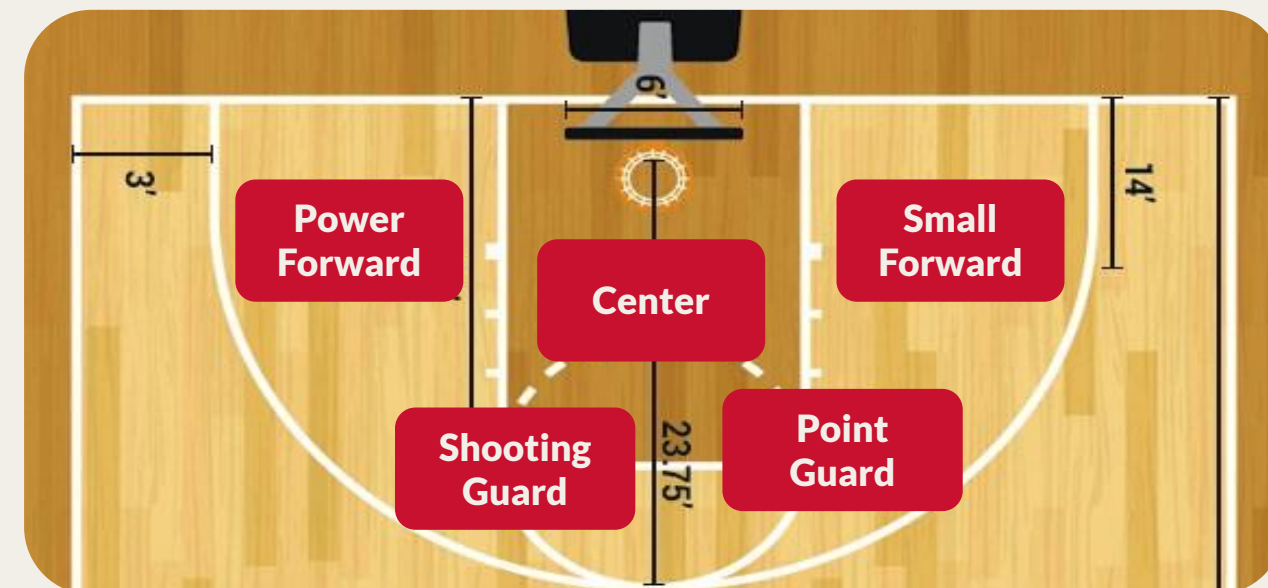
hongfei@umich.edu

SI 671 Data Mining Project Poster

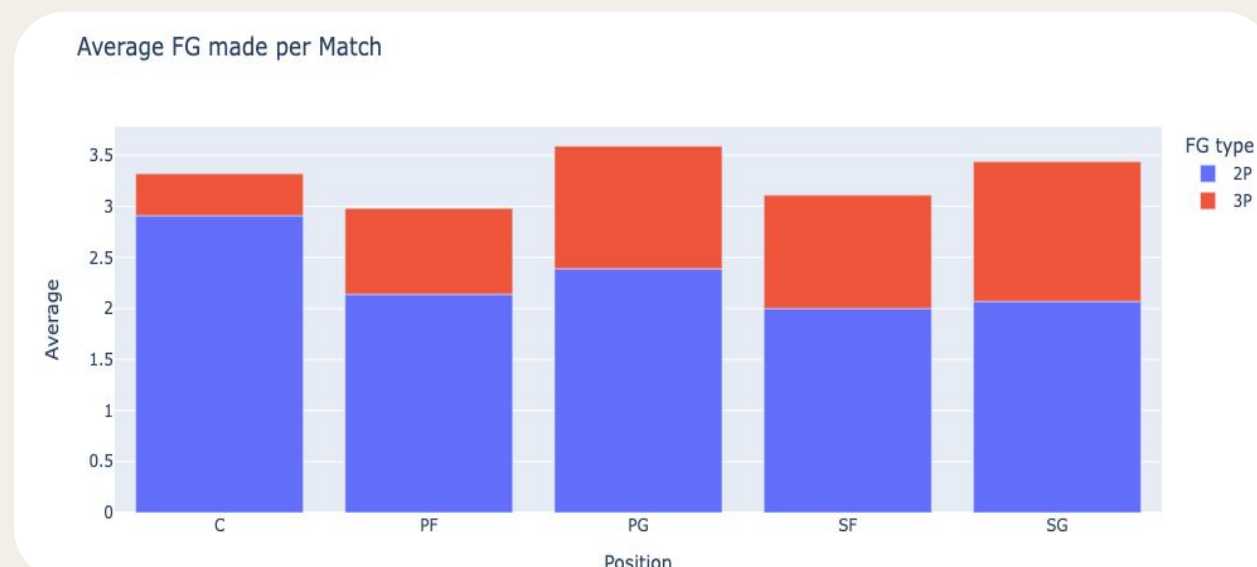
## INTRODUCTION

Basketball players are traditionally grouped into five distinct positions, but these designations are quickly becoming outdated. This study is an attempt to reclassify players into new groups based on personal performance in the 2021-2022 NBA season.

## Background



In Basketball, every player has a role on court. Each team has 5 players on the court at any given time. And each player often has a unique skillset to help their team win. Five distinct positions on court are:



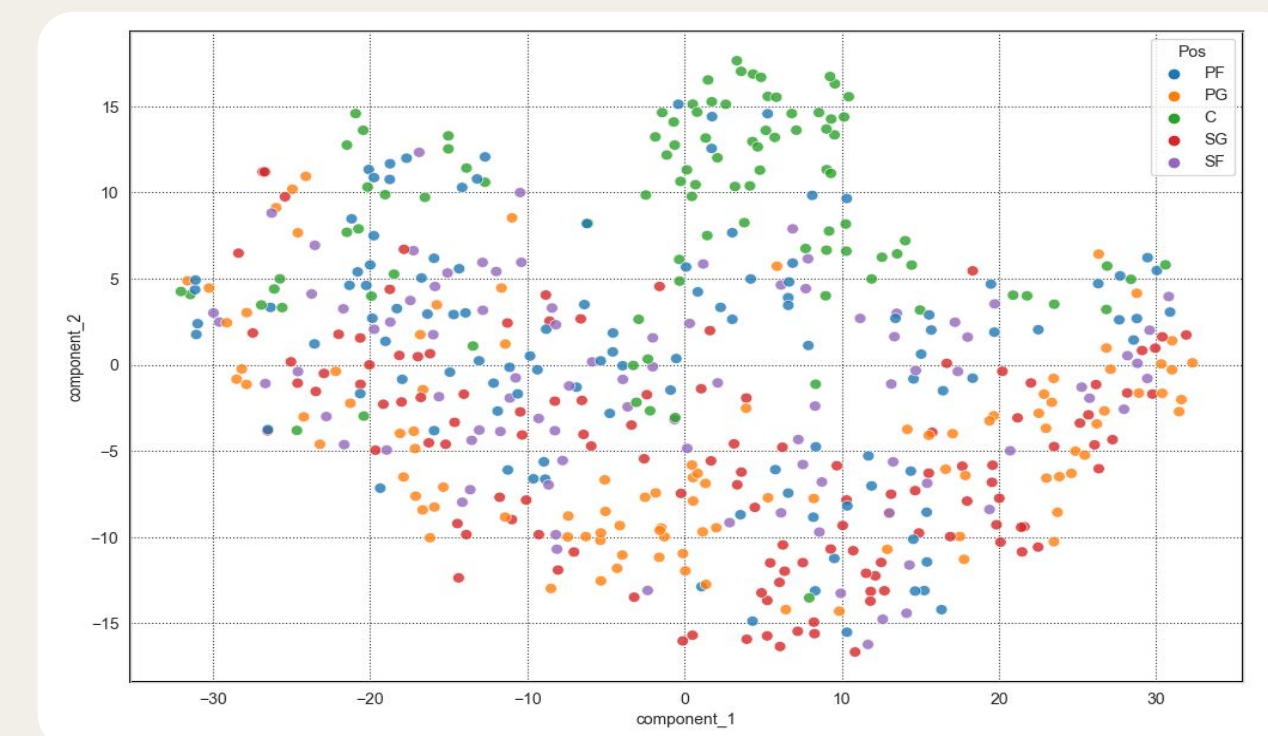
## Research Questions

- Is it possible that a new NBA player designation can be determined based on the actual performance of modern basketball players?
- If such a designation is found, could it be used to build a strong basketball team?

## RESEARCH

### 1 Cluster the Players

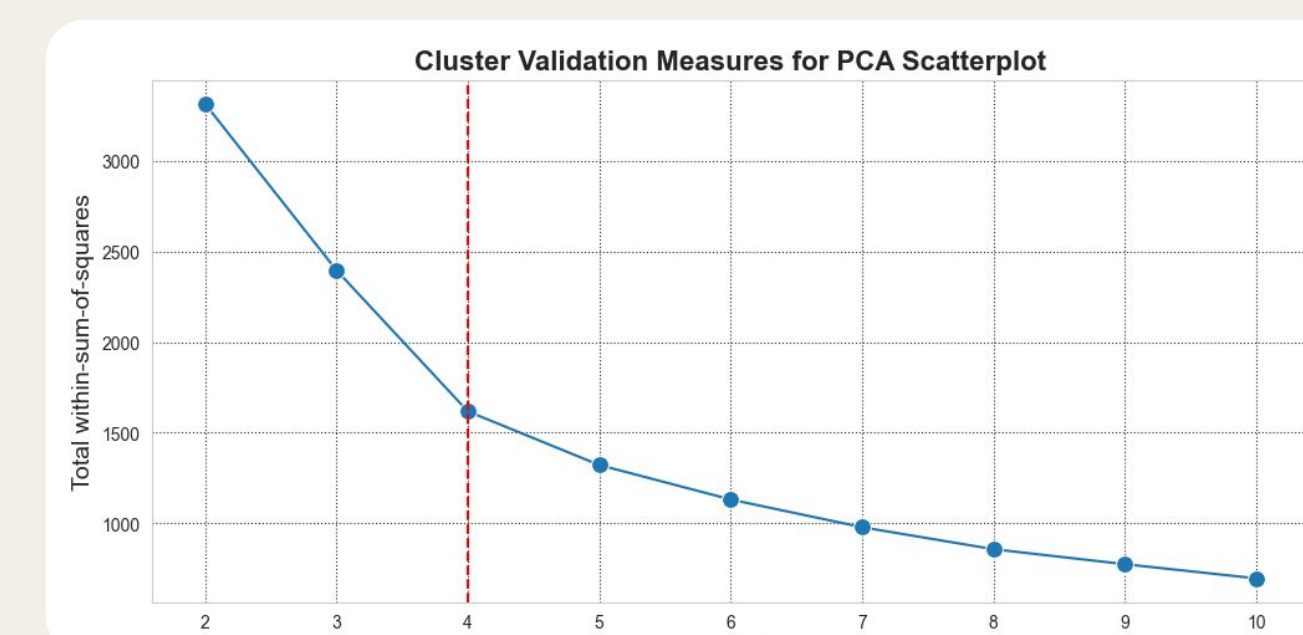
**t-SNE:** reduces dimensions non-linearly by preserving the local relationships of data points in a lower-dimensional space.



Centers are highly grouped and PG are also loosely grouped while the other three positions are relatively mixed among each other. There is some visible grouping but rest of it is fuzzy.

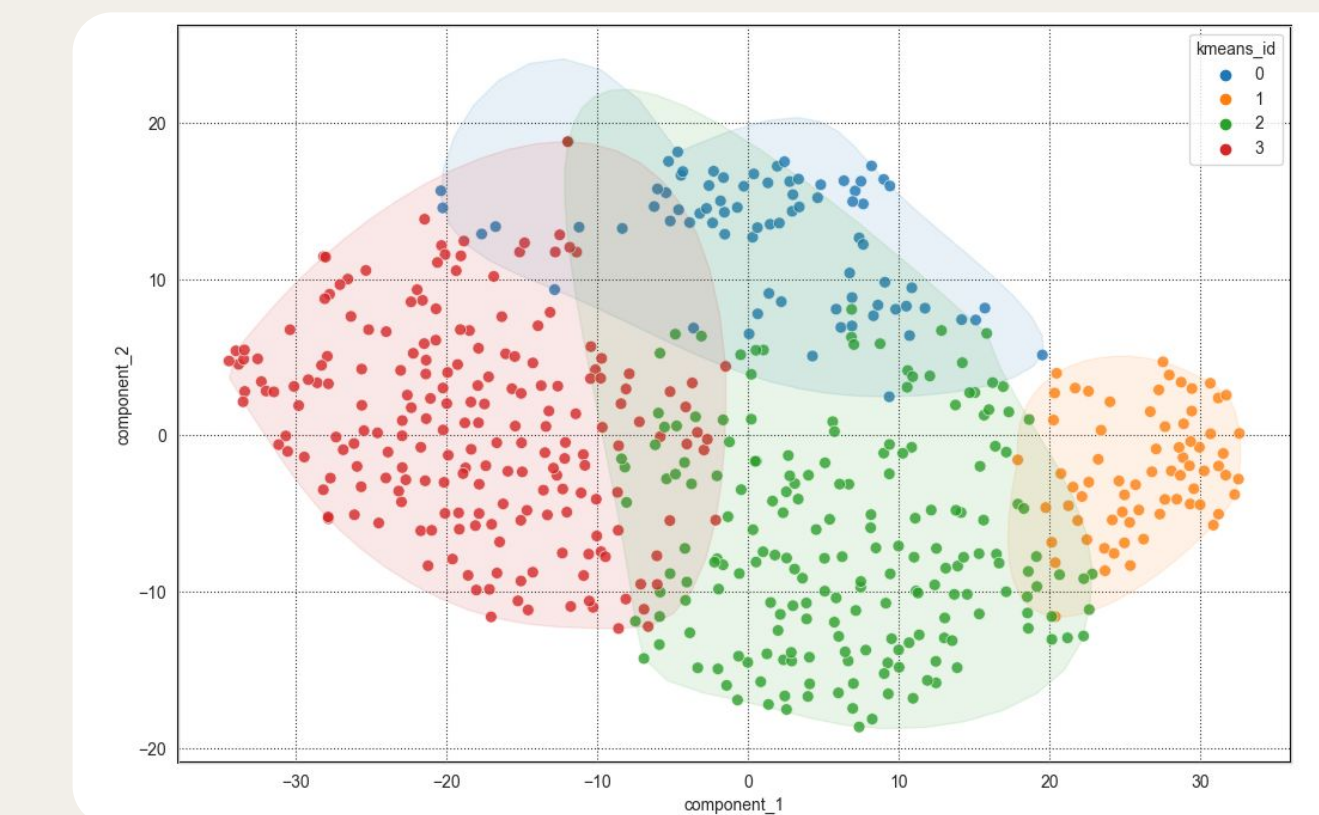
### 2 Finding Optimal number of Clusters (k)

**Elbow Method:** calculated as the sum of the squared distances between each member of a cluster and its cluster centroid[2].



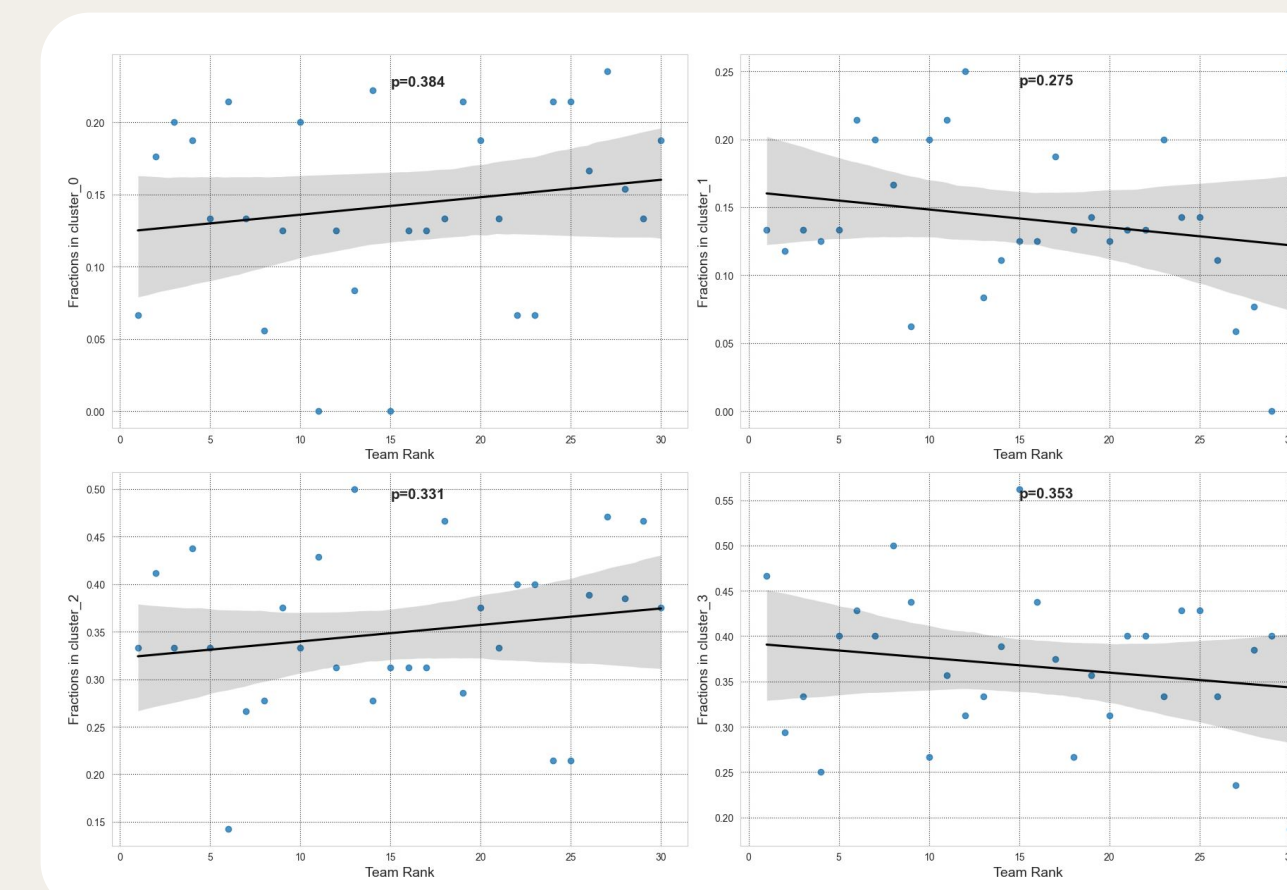
4 clusters is the optimal choice since it is the point where Elbow method shows an elbow

### 3 Cluster Players Again



### 4 Relationship of Team Rank with new clusters

#### Team rank vs. each cluster



There is no significance when applying linear regression to all four plots. This suggests that there is no relationship between how good a team is and a particular cluster. All clusters are equally important, suggesting there is no one type of player that dominates the NBA.

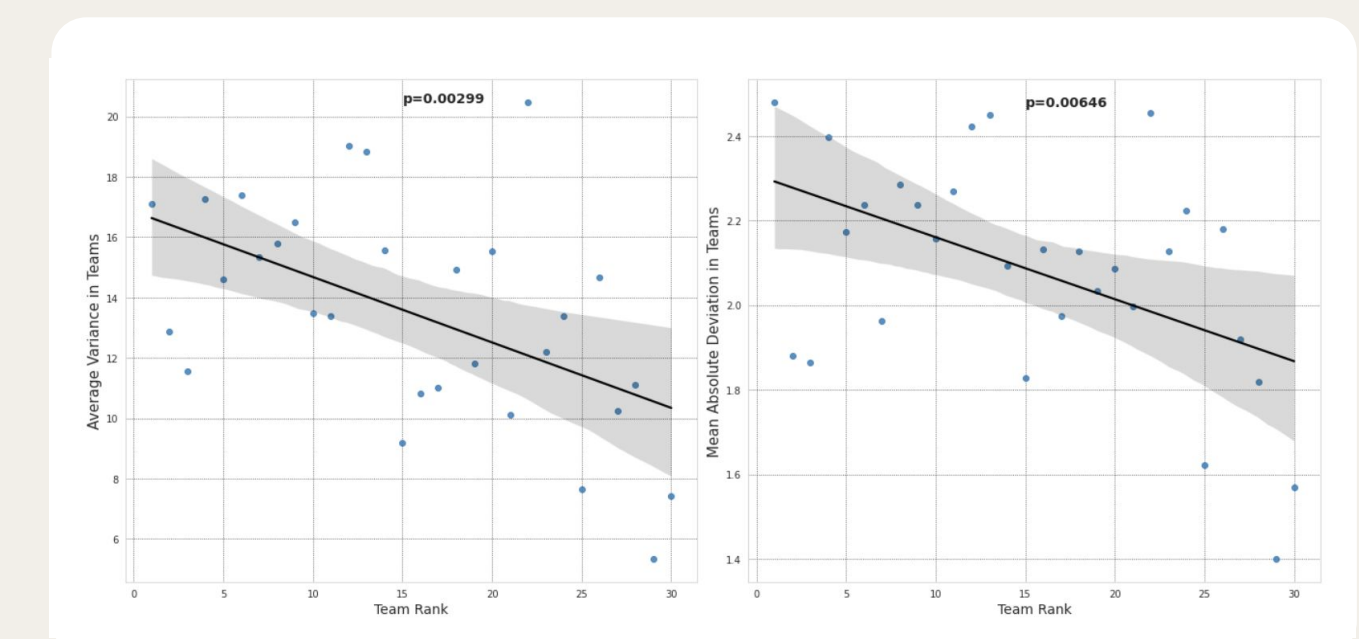
## CONCLUSION

The study reevaluates NBA player classification by clustering based on performance data, determining four optimal performance-based groups through t-SNE and the Elbow Method. The analysis reveals no direct correlation between team success and specific clusters, suggesting that diverse player skills contribute more to team performance than traditional positions. This highlights the importance of varied skill sets over conventional roles in team dynamics and success in the NBA

## REFERENCE

- [1] Data source: NBA Players data (1950 to 2022)
- [2] Eisen, Michael B., et al. "Cluster Analysis and Display of Genome-Wide Expression Patterns." Proceedings of the National Academy of Sciences of the United States of America, vol. 95, no. 25, 1998, pp. 14863-14868.

#### Team rank vs. cluster distribution



The results indicate that strong teams have players whose success cannot be attributed to fundamentals alone. Players being more spread apart on the performance metrics space means that the team has more diversity and diversity in turn ensures that the players' skillsets complement each other.