

Find the ideal player for a team

A study of changing style in NBA games and
a new approach of classifying player functions

Yiding Weng

Introduction

We want to explore:

- How have NBA basketball games and players changed over the years?
- How to improve basketball team roster during off-seasons?
- How to define player profile, in order to find cheaper or better substitute ?

Approach

- Observe and verify the changing playing style in NBA games and players
- Find a way to segment NBA players according to their performance statistics, using clustering.
- Discuss the profiles found through clustering

Data preparation

Historic NBA team records

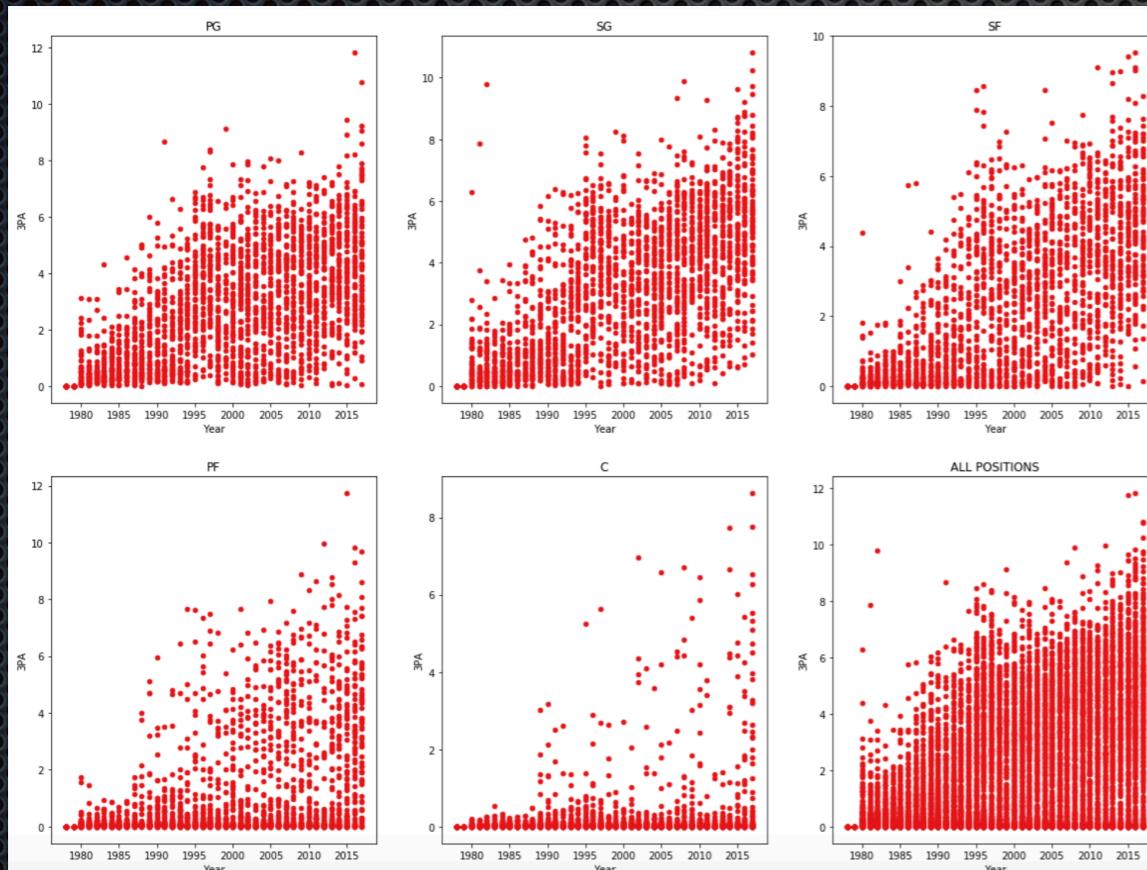
Data Location: <https://www.kaggle.com/drgilermo/nba-players-stats/data>

- **Data cleaning**, use only PG, SG, PG, PF, C as the 5 designated positions
- **Normalization**, transform player data as they have played equal amount of time
- **Deal with missing values**, remove the data before 1978
- **Handle outliers**, keep players who have played more than 400 minutes per season

Exploratory analysis

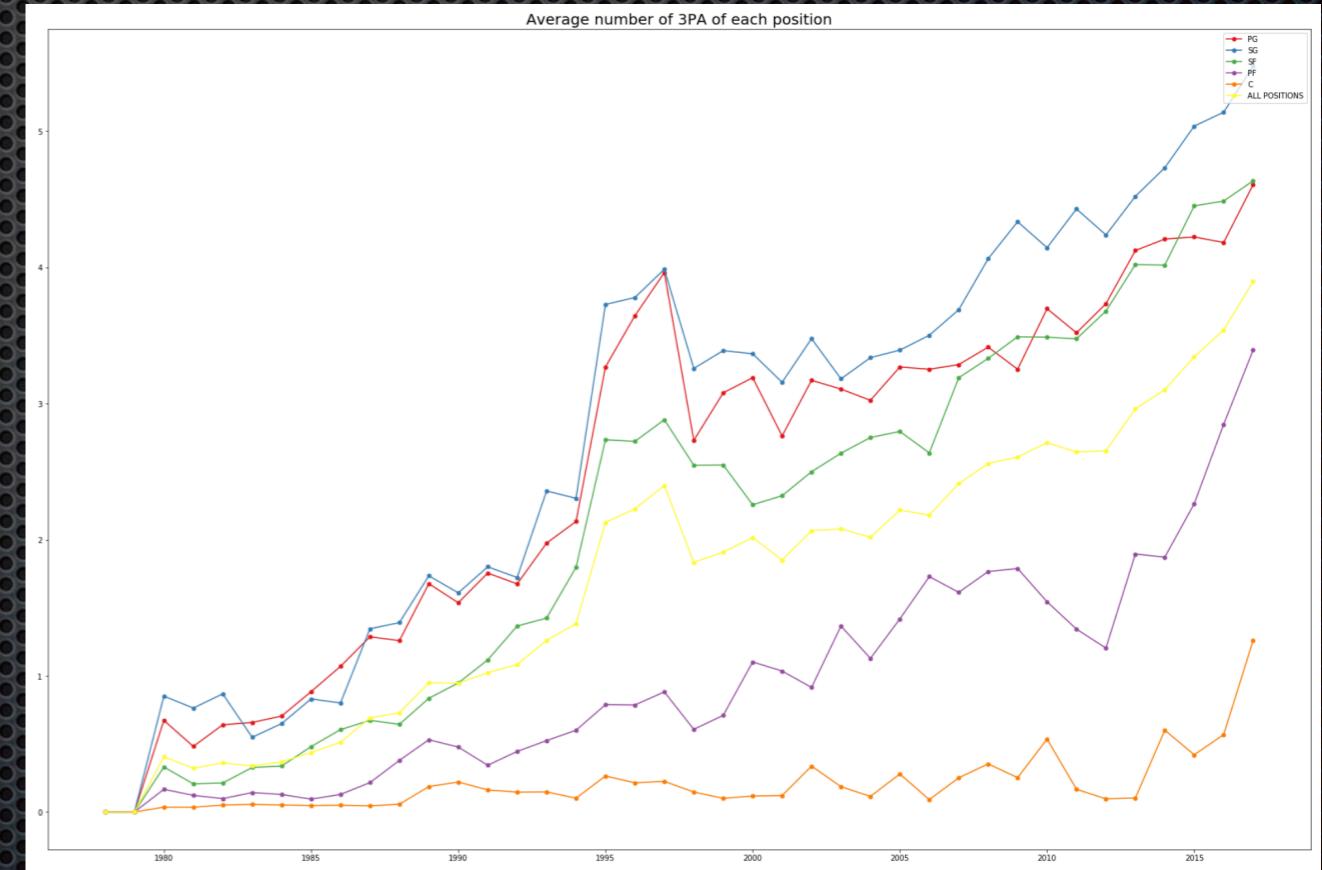
Scatter plot

- 3PA (number of 3-Pointer attempt) of individual player in each position from 1978-2017



Line plot

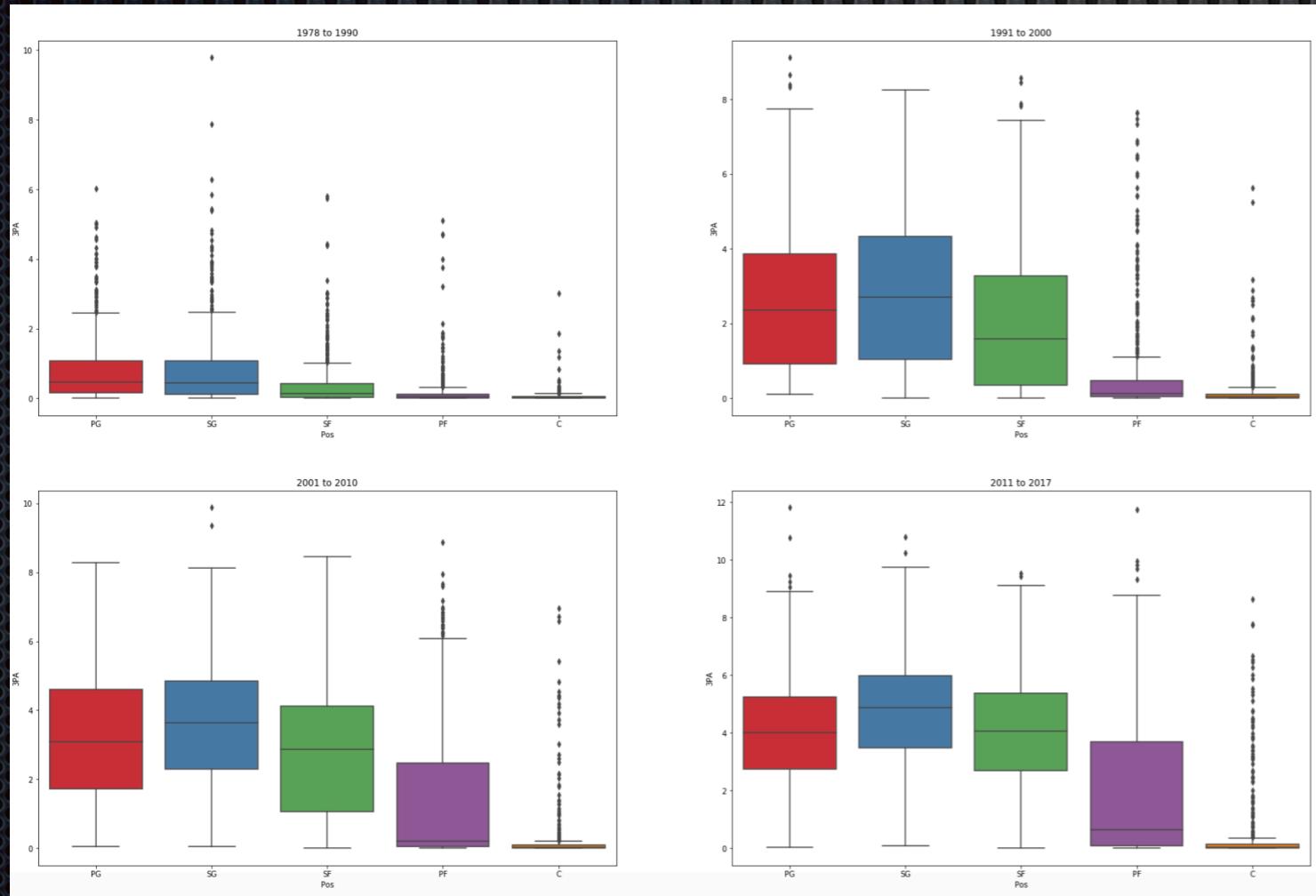
- average number of 3-pointer attempts per game of each players in different positio



Exploratory analysis

Box plot

- number of 3-pointer attempts for player distribution in each time period



Quick Summary

PTS - Center (C) and point guard (PG) took small forward (SF) and shooting guard (SG)'s places in leading score positions;

3PA - Overall increase, especially in point guard (PG), shooting guard (SG) and small forward (SF);

2PA - Overall decrease;

AST - Overall stable, point guard (PG) becomes less dominant in leading the number;

TRB - Overall stable, Center (C) becomes more dominant;

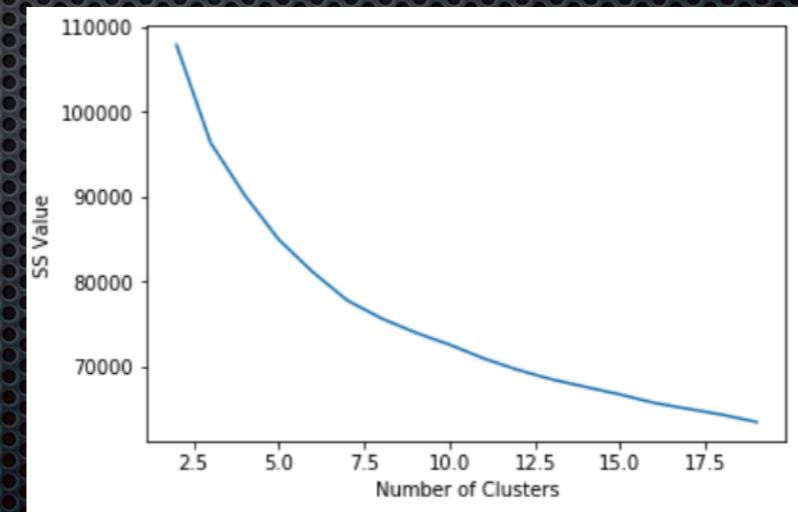
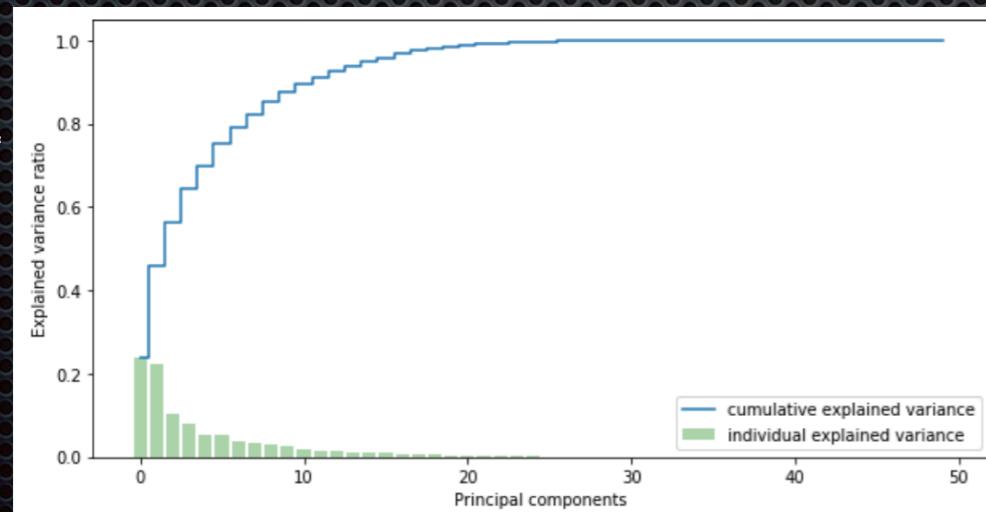
PF - Overall decrease;

FT - Overall decrease in FT, Center (C) and point guard (PG) took small forward (SF) and shooting guard (SG)'s places in leading number of free throws;

BLK - Overall stable, Center (C) becomes more dominant;

Cluster modeling

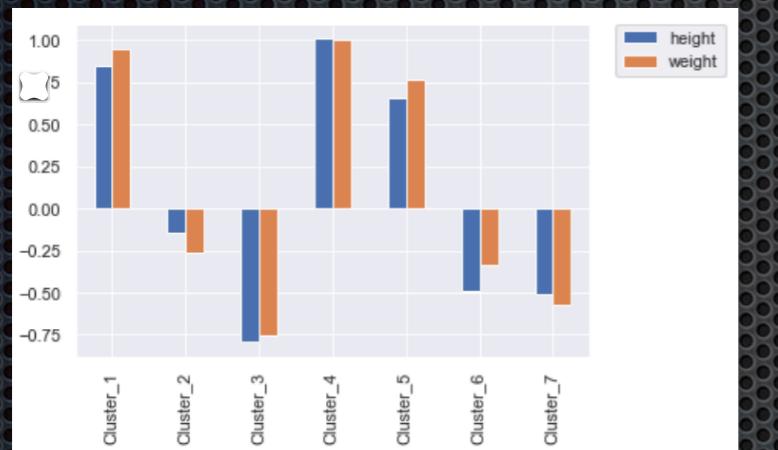
- **Select and modify data for modeling**, in order to observe the most obvious change in limited time, I chose two subsets have the largest time gap in between, which is pre-90 (players data before 1990) and 2000s (players data since 2010)
- **Feature analysis with PCA**, PCA stands for principal component analysis, It reduces the dimension of data with the aim of retaining as much information as possible.
- **Find the optimal number for clusters**, with the two reduced dimension datasets, I explored the different choices of number for clustering, with the options from 2 to 20 clusters, and decided to move on with 7 as the number of clusters for both datasets.



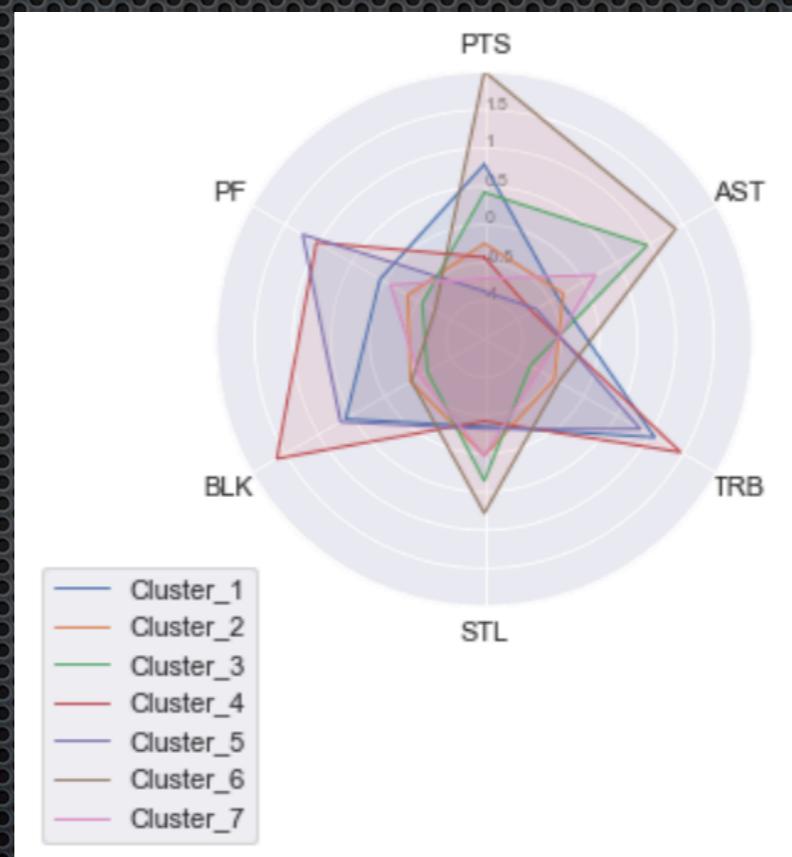
Cluster analysis

For pre-1990 data

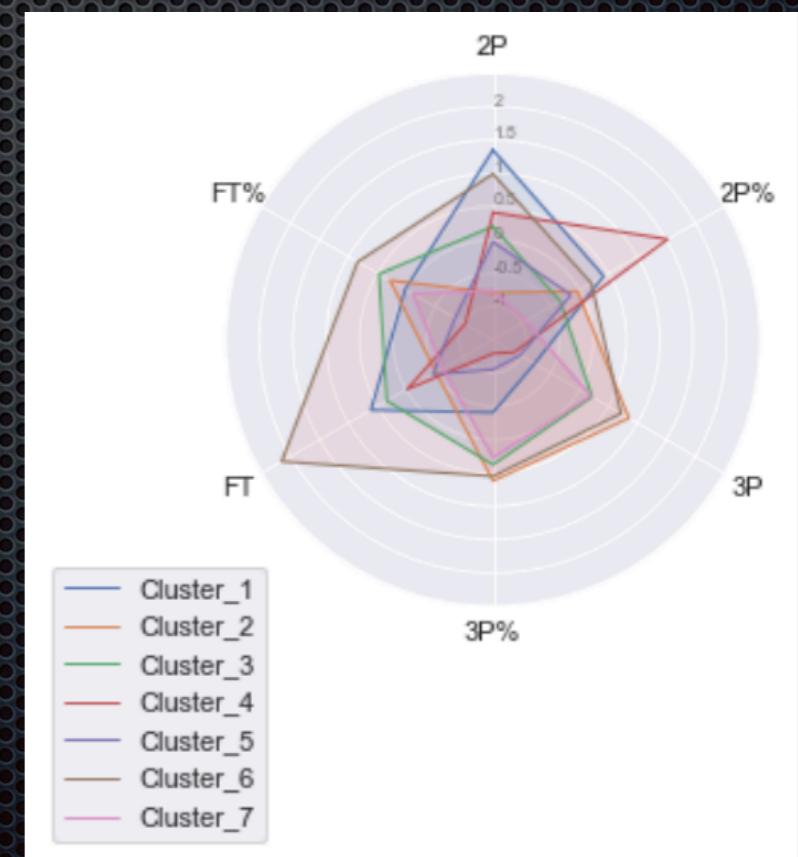
Physical feature
assessment



Function
assessment



Scoring
assessment



Cluster analysis

For pre-1990 data

1. Middle man that carried the team - SG, SF

Isiah Thomas, George Gervin, Julius Erving, Magic Johnson.

2. Well-rounded small players with limited playing time - PG, SG, SF

Nick Weatherspoon, Jim Paxson, Charlie Scott, Armond Hill

3. Defensive big man who does the physical work - C, PF

George Johnson, Clifford Ray, Paul Silas, Benoit Benjamin

4. Distant range shooter - PG, SG

Michael Adams, Reggie Miller, Johnny Newman, Jeff Hornacek

5. Elite big man with all traits - PF, C

Larry Bird, Moses Malone, Karl Malone, Hakeem Olajuwon

6. Well-rounded big man - SF, PF and C

Charles Oakley, Sam Perkins, Chuck Person, Alex English

7. Small players with all traits - PG, SG

John Stockton, John Havlicek, Bob Wilkerson, Kevin Porter

Cluster analysis

For pre-2010 data

1. Big man with all traits - PF, C

Marc Gasol, Blake Griffin, Kevin Love, Dirk Nowitzki

2. Three-point shooter, outside defender - SG, SF, PF

Klay Thompson, J.J. Redick, Ray Allen, Kyle Korver.

3. Outside assistant - PG, SG

Ricky Rubio, Rajon Rondo, John Wall, Andre Miller

4. Under basket offender, ring protector - C, PF

Zaza Pachulia, Andrew Bogut, Andre Drummond, Rudy Gobert.

5. In between big men - PF, C

Lamar Odom, Tony Allen, Taj Gibson, Yi Jianlian

6. Elite player who carries the team - PG, SG and SF

LeBron James, Kobe Bryant, James Harden, Kevin Durant

7. Secondary ball handler, 3-point shooter - PG, SG and SF

Marco Belinelli, Corey Brewer, Matthew Dellavedova, Derek Fisher

Conclusions

- The league players are increasingly more dominated by 'smaller' players, which means the game is moving away from near-basket towards outside range
- The increased freedom for players to switch between positions, this implies the speed and range of ball movement have increased, rather than simply move towards the big man inside or close to basket.
- Contemporary games are becoming less physical, fast-paced, and covering wider space. Therefore, we can state the NBA basketball playing style is moving towards 'small ball' after 1990s to 2010s.

Future work

- **Feature selection**, use silhouette score to compare the performance of all the clustering based on different feature selection options
- **Cluster modeling for data from 1990 - 2000**
- **Import Salary Data**, if the salary is incorporated into the player information, it can provide extra perspective while comes to player trade consideration.
- **Team roaster analysis**, study the combination of players from different clusters

Recommendations

Contemporary NBA players have evolved their functions beyond what their positions normally implies. Signing and trading players during off season involves careful consideration of the player functions. This project has provided general function to profile NBA players.

A relatively successful team seeks to cut down their salary cap can consider trading their existing player with new player in the same function cluster with lower salary expectation.

And for the team intends to improve their roaster performance should avoid trading player for the one within the similar function cluster, but for player from different ones to change the team chemistry.

.