problem statement

In this project, the rookies are the player who is still in his first shoes contract and these players normally played less than 3 years in NBA. This project will focus on the player who joined the NBA in 2015 and 2016.Because the rookies in 2017 didn't play enough games.

The first aim is to find who has potential to be an all-star player. We will collect the list of NBA player and their game data from basketball-reference. we will find the game data of rookie players in the last ten years. Those players who had ever been an all-star player will be seen as a successful player and labelled as one. Then we can build a model to predict which rookie will be an all-star player in the future. Finally, the fan map of these rookies will be produced

challenges from a data science point of view

The number and quality of data:
Lots of communities ban the permission which used to obtain information about users. Therefore, obtaining valid data in the legal pathway is a big challenge.

How to filtrate the valuable rookies in NBA:
There are 60 players join the NBA draft every year, however, this project should focus on the potential rookies, therefore to research their data and pick the target players in another challenge.

explain the pipeline

1. Obtain rookies list in past 3 years
2. Obtain their game data
3. Build a model to pick potential rookies.
4. Obtain the social account of these potential rookies.
5. Obtain the followers' location.
6. Draw the fan map
7. Analysis the business value

details on data sets

1. List of NBA rookies in the past 3 years
2. The game data of players will include PER(The Player Efficiency Rating) and their win-share per 36 minutes. These two advanced data are the most useful data to estimate the value of one player.
3. The location of followers.

NBA draft

https://simple.wikipedia.org/wiki/NBA_draft