PROGRESS REPORT NBA EDITION

Nirvan Ashish
University of Colorado
nias4854@colorado.edu

Yifei Niu
University of Colorado
yifei.niu@colorado.edu

1. ABSTRACT

We should be able to successfully predict how a rookie may perform in his rookie season based on physical measurements and previous statistics (may be college level). We should also be able to formulate and give the correct order of the current rookies completing their first season based on our work. We will also be using general information about the NBA from online sources such as ESPN and other sports services.. We know the NBA has so many different stars, for example Stephen Curry and Russa Westbrook. We want to find how to evaluate a good NBA star. Most NBA player styles are different in the NBA. We should follow a different situation to find a good player.

2. LITERARY SURVEY

1. Processing: we will need to convert our data into compatible forms and then merge

   them together. This will include creating NBA player physical measurement, each season statistics and how about NBA physical measurement changed in each season. We should clean another player in the NBA player stats. We will focus on those data sets getting results in the final.

2. Analysis using the datasets, we will form logistic regression with NBA physical measurements. This will form a baseline of our research, we will need to analyze the effect of each attribute on the low level rate NBA physical measurement star. After analyzing the relationships between all variables, we will create a regression model to generate the probability of NBA body situation and physical measurement .

3. DATA SETS

3.1 NBA physical measurement:

https://stats.nba.com/draft/combine-anthro/

NBA physical measurements has 9 parts, we will research which part is so important in the graph. The dataset contains information on the draft combine of the rookies that are going to make their first appearance in the NBA. This data set contains the physical measurements of the players and we plan on using the important data and finding a trend that helps us identify their ideal position with which they should enter the NBA.

3.2 Every team resource:

https://www.basketball-reference.com/
NBA team and player resources follows this graph.
We intend to see the shift in the NBA player performances every season and find trends that support our claim. In addition to this we also will look to build a model draft for the young rookies based on their information and trends that we can find. For example 2013 draft, the first pick was not Stephen Curry whereas he is arguably the greatest point guard in today's game.

3.3 Every player Stats :

https://www.kaggle.com/drgilermo/nba-players-stats

This dataset may qualify as important. We will be using this for reference as this contains all the player statistics from the previous years. Finding the general level in the NBA usin graphy. We could predict NBA physical measurements based on this dataset. When looking at different NBA superstars a big research group is required.

6. PROPOSED WORK :

1. Processing: we will need to convert our data into compatible forms and then merge them together. This will include creating NBA player physical measurement, each season statistics and how about NBA physical measurement changed in each season. We should clean another player in the NBA player stats. We will focus on those data sets getting results in the final.

2. Analysis using the datasets, we will form logistic regression with NBA physical measurements. This will form a baseline of our research, we will need to analyze the effect of each attribute on the low level rate NBA physical measurement star. After analysing the relationships between all variables, we will create a regression model to generate the probability of NBA body situation and physical measurement .

5. EVALUATION

We will evaluate our research by testing to see if it accurately predicts whether or not a NBA rookie can become a star. We hope to get an accuracy measure of 60%-70%. Some players who did not get the first draft pick but in turn become super stars. We should rethink how to make this error before pickup. Most players change so quickly. When they suffer big hurt and sometimes mental health issues, this may still influence his status. (For example: James Harden(Houston Rockets) has ever suffered these problems.). A comparison between our predicted best rookie and some of the greatest current players will be made in terms of physical measurements and see if there is any correlation between the two. A true order of the rookie season will be found by the end of this project.

6. TOOLS

Jupyter Notebooks:

From this part, we will build NBA physical measurement, we should build a graph, how about this research in the NBA physical measurement.

Python 3:

We will use this download basic datasets graphy, because they couldn't give us crv results directly. These things help us find data tools, because some data have protection levels.

NumPy:

We could get this graph from an NBA resource, this helps us get a good analsize result. Some websites need to use these tools.

Pandas:

We will research basic data, building graphy. How about research from this data graphy.

Excel:

We will be using excel as well to make a few graphs that are relevant to our project.

SciPy:

We will be using libraries such as SciPy in order to find correlation constants and attributes that link the different datasets

7. MILESTONES

To start with, we will need to clean the data. This would involve cleaning and pruning of the data so as to use only the important attributes. we will find a graph in this data. We will find some data we could not use in the graphy.

In addition, we will start to merge bthe relevant attributes of our datasets. We will also build a graph in this step to show any type of correlation between physical measurements and statistics.

After cleaning and merging we will be able to analyse the data looking for general patterns and trends between the players of the draft combine and points in their respective player seasons. Furthermore, we will need to build datasets to make this easily understanding.

We will be continuing to document our work till the end, we should build the next steps.

By the end we will have our final report ready with all our source code put on Github.

The project will end by making a short presentation explaining our work.

8. COMPLETED

The initial phase of the project involved cleaning up pf the data and pruning. We have been successful in doing so and this makes it easier to analyse the data without null values present in between.

We then merged the datasets in order to help analyzing the data easier. We then were able to bring out a trend in the heights, weights, wingspans of the rookie physical measurements datasets.

We were able to conclude that the heights involving the guard positions such as Pg and Sg do not differ much with a few outliers whereas positions such as forwards usually have more change in their physical measurements.

9. REMAINING

Using different algorithms and cluster methods we intend to group a couple of rookies based on their physical measurements based on their first season statistics. A true order of the first season rookies will be found out by the end after all the analyzing has been done.

While we look for any correlations between physical measurements and player performance .

We also wish to predict players that may go on to become superstars. A few comparisons will be made to current superstars of the game to enhance the accuracy of our project.

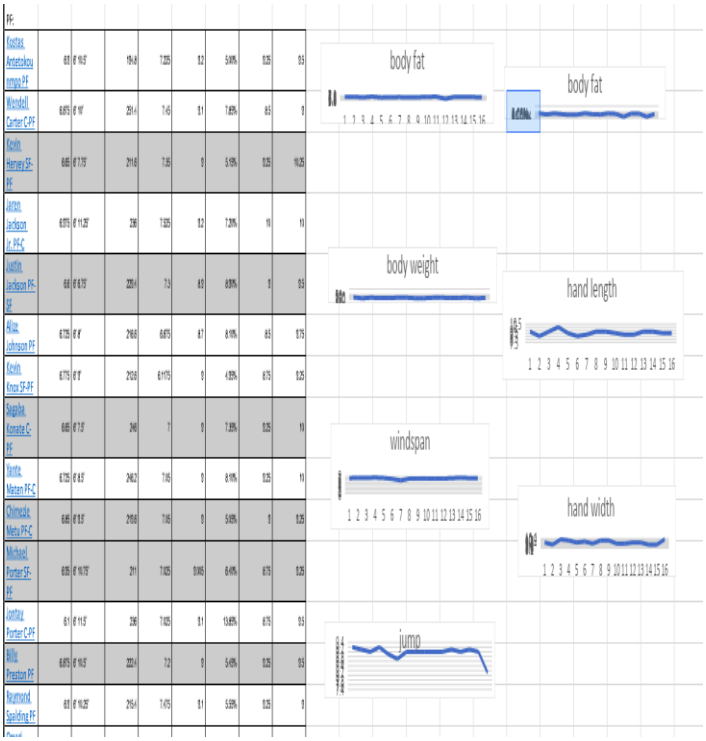
The final report involving everything the project has to offer will be submitted on github too.

The source code involving the procedure used to make the graphs shall be shared to the github account as well.

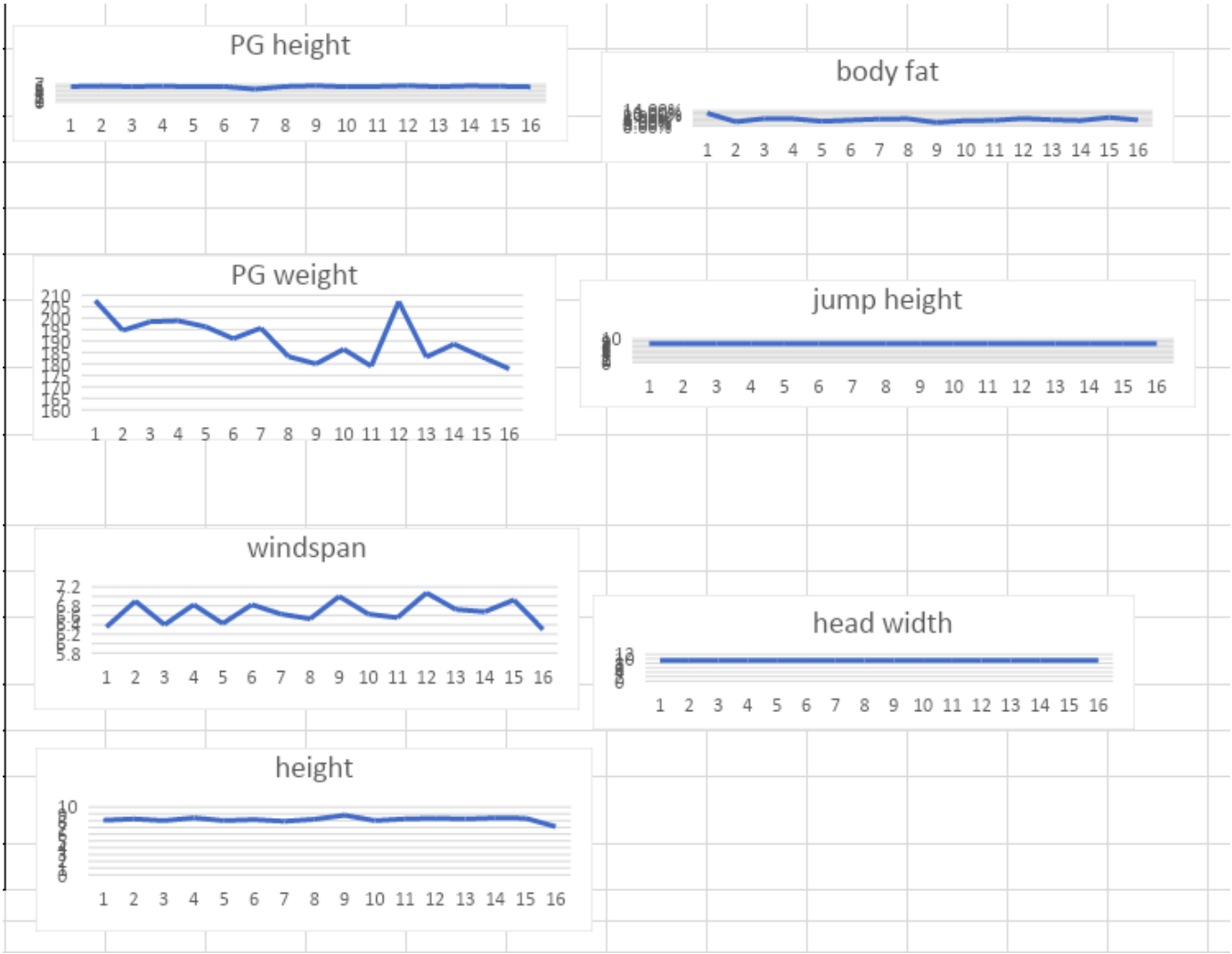
Finally the group presentation explaining our project will be shared with pur class.

10. RESULTS

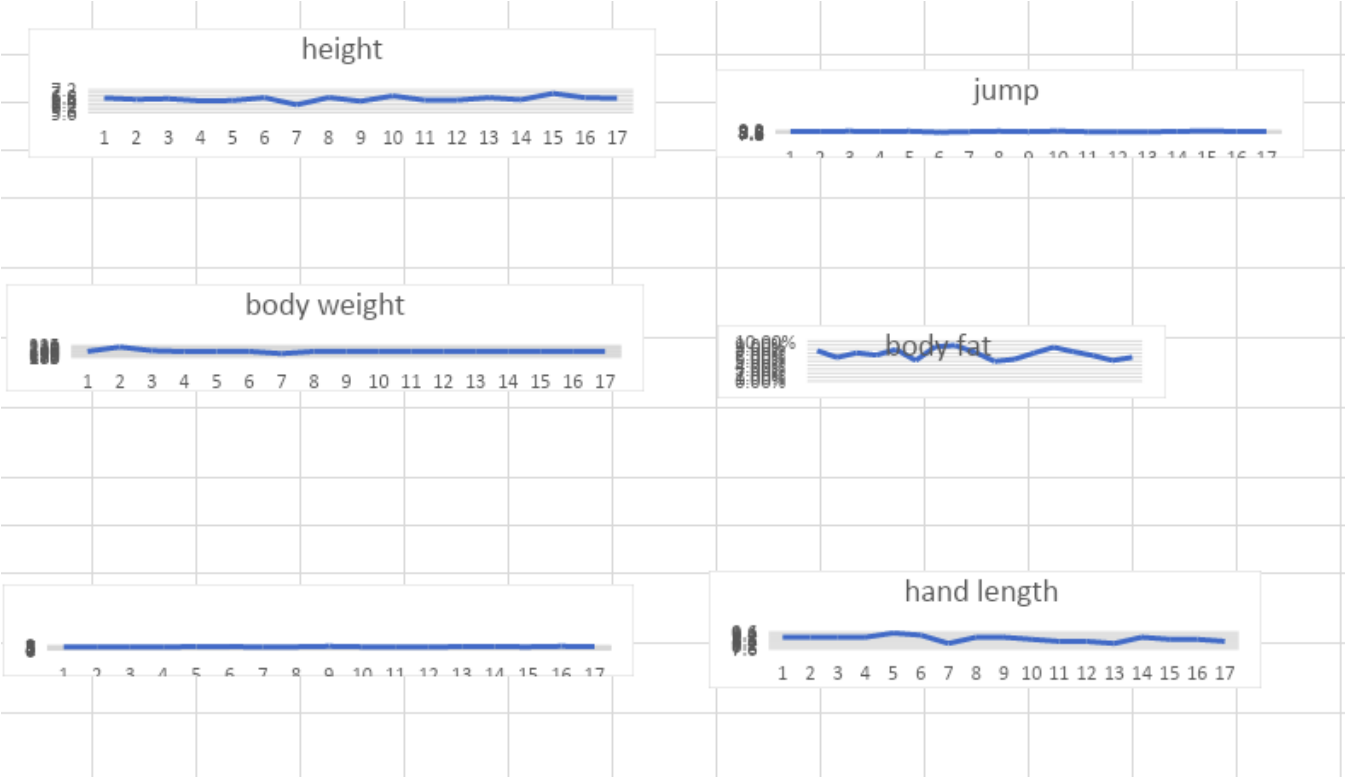
The PF position NBA rookie trends shows very less difference but there is a sudden dip in jump when the height increases. Wing span, Hand width and other physical attributes remain the same with no dips.



.



PG weights show a lot of change and therefore weight is not generally taken into account for a PG since strength of PG is not valued as the PG usually adds in the team for assists and shooting. There is a lot of change in wingspan too since most of the PG do not rebound.



The above graphs are that for a Centre. The graphs are usually all similar and show very less outliers since C are players that usually play the defensive role such as rebounding and blocking having shooting or finishing or playmaking as secondary abilities.