The Difference Between You and An NBA Player.
Comparing the Ape Indexes of NBA Players and Ordinary Males.

Nahom Debela   *Washington State University*
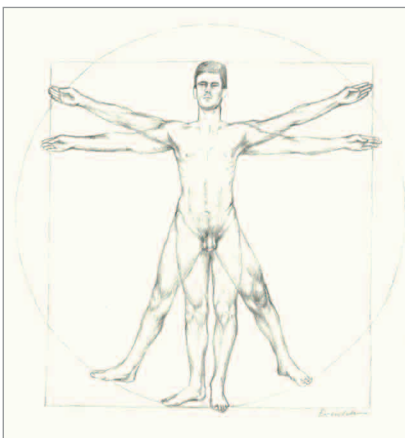
In this report I walk through my steps in answering my research questions. I seek to find relationships between NBA measurement data and measurements from ordinary people. I perform statistical data analysis to uncover trends and solve my questions.
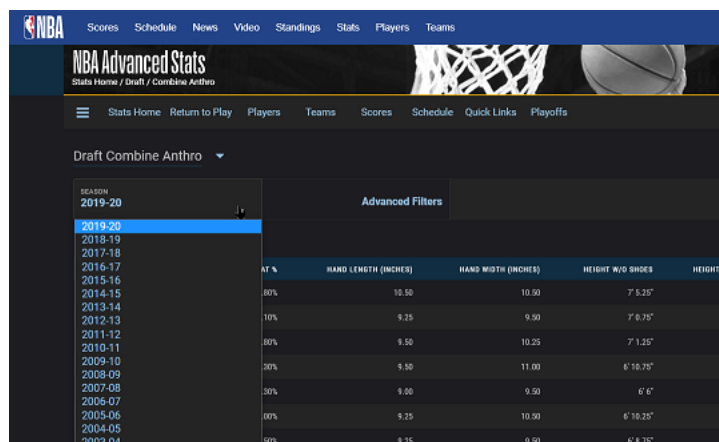
December 26, 2020

1   Introduction

(a) Figure 1: Male Vitruvian Figure

(b) Figure 2: NBA Measurements

Figure 1:   The Vitruvian drawing was made by Italian polymath Leonardo da Vinci in about 1490. It is based upon the theory that the human body can be symmetrically inscribed within both a circle and a square

Figure 2: NBA Data of Measurements that were taken from the NBA website.

2   Research Question: How does the ratio between height and arm span compare between NBA players and normal males?

2.1   *Is there a strong relationship between height and arm span?*

2.2   *Has the ratio between height and arm span changed over time for NBA Players?*

Figure 2: Wingpsan

---

3   Data Description

I've used 3 separate data sets. One data set consists of measurement data from participants who have filled out the handout. From this data I have used "Arm Span","Height", and I have calculated "Ape Index". This measure is simply the ratio between one's height and arm span. In my analysis, I used ape index to observe if there is a relationship between the average NBA players ape index and the ordinary male's.

The other 2 data sets are NBA measurent data from "Kaggle". This dataset came as one file. I filtered the NBA data and split it into 2 separate data sets. One included previous NBA players that played in 2002 and before. The other contains recent NBA players who have played in 2016 and after. For my analysis, I used: "Arm Span", "Height", and I have calculated "Ape Index" for both time frames. These datasets are helpful in answering: Does the ratio between height and arm span change over time?

3.1　Summary of Sample

Gender: This is a factor. Possible values include, male, female, other. 100% of my sample is "male."

Units: All my observations are in cm.
Age: All my observations are samples where age is at least 18.
The rest of my data includes Height, Wingspan and Ape index for all 3 datasets.

Figure 3: Summary Table

```
   height.NA        arm.span.NA        Measure Ape Scale
 Min.   :154.0    Min.    :156.5     Min.    :0.8324
 1st Qu.:170.0    1st Qu.:171.5      1st Qu.:1.0000
 Median :176.5    Median :177.0      Median :1.0141
 Mean   :174.4    Mean    :176.3     Mean    :1.0112
 3rd Qu.:179.1    3rd Qu.:182.5      3rd Qu.:1.0327
 Max.   :188.0    Max.    :191.1     Max.    :1.0606
 Height (in cm) 2002 Wingspan (in cm) 2002 NBA Ape Scale 2002
 Min.   :177.8      Min.    :182.9       Min.    :0.9639
 1st Qu.:190.5      1st Qu.:203.2        1st Qu.:1.0242
 Median :200.7      Median :207.0        Median :1.0438
 Mean   :199.5      Mean    :208.4       Mean    :1.0449
 3rd Qu.:205.7      3rd Qu.:215.9        3rd Qu.:1.0627
 Max.   :215.9      Max.    :229.9       Max.    :1.1165
 Height (in cm) 2015 Wingspan (in cm) 2015 NBA Ape Scale 2015
 Min.   :177.8      Min.    :177.8       Min.    :0.9878
 1st Qu.:198.1      1st Qu.:207.0        1st Qu.:1.0288
 Median :203.2      Median :212.1        Median :1.0427
 Mean   :202.0      Mean    :212.0       Mean    :1.0495
 3rd Qu.:208.3      3rd Qu.:219.1        3rd Qu.:1.0737
 Max.   :213.4      Max.    :229.2       Max.    :1.1033
```

From all 3 datasets the largest mean in ape index is in the 2015 NBA Data set.
Ape index seems to be the largest in the 2015 NBA dataset (M = 1.0495), followed by 2002 NBA Dataset (M=1.0449cm) and finally the Normal measurement data (M=1.0112cm)

Wingspan seems to be the largest in the 2015 NBA dataset (M = 212.0), followed by 2002 NBA Dataset (M=208.4cm) and finally the Normal measurement data (M=176.3cm)
Wingpsan of NBA players in both datasets is on average larger than the normal males wingspan by 33.9cm.

Height seems to be the largest in the 2015 NBA dataset (M = 202.0), followed by 2002 NBA Dataset (M=199.5cm) and finally the Normal measurement data (M=174.4cm)
Height of NBA players in both datasets is on average larger than the normal males height by 26.4cm.

Figure 4: Height vs Wingspan Plot
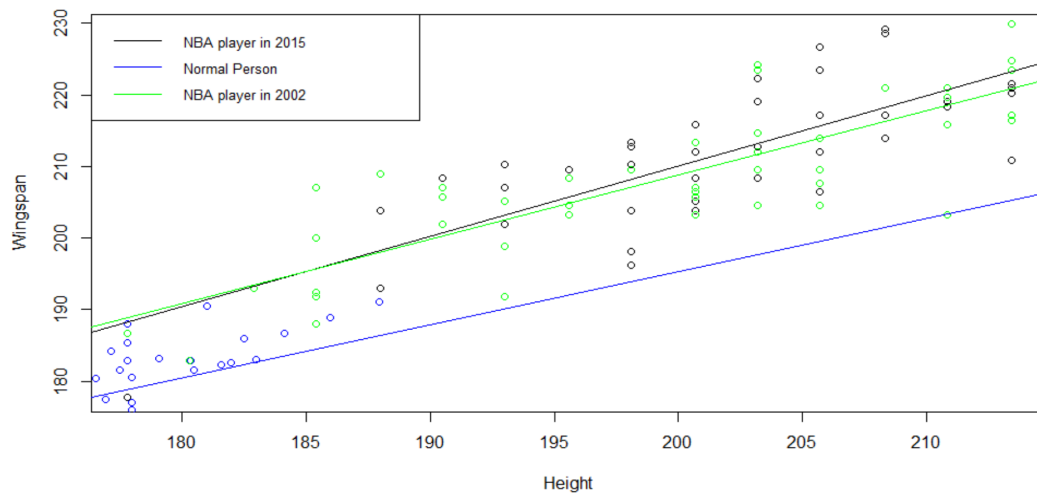
Clearly, we can see from the plot that NBA Height and Wingspan in both 2015 and 2002 are significantly larger than the normal males height and weight.

The lines of NBA players in 2015 and 2012 are very close and it is difficult to make any assumptions.

3.2 Summary Statistics of Data

Correlation Table is on Next Page

Table 1: NBA Descriptive Statistics and Correlation Analysis

	M	SD	1	2	3	4	5	6	7	8
1 height.NA	174.4	7.75	1							
2 arm.span.NA	176.3	8.59	.67***	1						
3 Measure Ape Scale	1.0	.04	-.32*	.49***	1					
4 Height (in cm) 2002	201.1	9.89	-.17	-.13	.05	1				
5 Wingspan (in cm) 2002	211.0	10.08	-.10	-.07	.03	.86***	1			
6 NBA Ape Scale 2002	1.0	.03	.14	.11	-.03	-.31*	.21	1		
7 Height (in cm) 2015	202.0	7.55	-.11	-.35*	-.31*	-.12	-.17	-.09	1	
8 Wingspan (in cm) 2015	212.0	9.60	-.03	-.26†	-.31*	.08	-.04	-.25†	.77***	1
9 NBA Ape Scale 2015	1.0	.03	.11	.04	-.09	.28*	.15	-.27†	-.07	.58***

Notes: Pearson pairwise correlations are reported;
a two-side test was performed to report correlation significance.

$^\dagger p < .10$ $^* p < .05$ $^{**} p < .01$ $^{***} p < .001$

4 Key Findings

From my correlation table I have several strong relationships:
For Normal Males: Wingspan correlates with height corr=.67
For NBA players in 2015: Wingspan correlates with height corr=.86
For NBA players in 2002: Wingspan correlates with height corr=.77

Is there a strong relationship between height and arm span?

These three correlations tell me that Wingspan and Height have a strong positive relationship. Additionally, if we take a look at the mean of wingspan in any of the datasets, it is not very different than the height. I can conclude that the relationship Wingspan and Height is positively strong. We can assume that as a males height increases, so does their wingspan.

Has the ratio between height and arm span changed over time for NBA Players?

I found the ape index of both NBA datasets (2002, 2015):
NBA Ape Index 2002: 1.0449
NBA Ape Index 2015: 1.0495

There is no significant difference between ape index of NBA players who played before or during the 2002 season and from NBA players who played after the 2015 season.

How does the ratio between height and arm span compare between NBA players and normal males?

Since ape index is extremely close in the NBA data I can just average it out as one NBA ape index to compare to the normal males ape index.

I found the average ape index of both NBA datasets (2002, 2015):
Average NBA ape index (2002,2015): 1.0472
Average Male ape index: 1.0112
As we can see, the ape index of NBA players is slightly higher than the normal male's: with a total difference of 0.036.

5   Conclusion

I set out to answer 3 questions:

Is there a strong relationship between height and arm span?
Has the ratio between height and arm span changed over time for NBA Players?
How does the ratio between height and arm span compare between NBA players and normal males?

From my research and analysis I can conclude that there is in fact a strong relationship between height and arm span. They seem to be highly correlated. It seems as height increases, arm span also increases. I am not very surprised with these results because typically the human body's wingspan is the same as their height. This results in an expected ape index of 1 for the average human. This theory is taken from Leonardo da Vinci.

To answer my next two questions I found the proportion between ape index and height which is known as "ape index". I compared this ape index score between my two NBA measurement samples and they turned out to be almost identical in ape index. From this I can conclude: there is no significant change over time between NBA player proportions of height to wingspan.

Finally, I wanted to know how the proportion of height and wingspan differed between normal males and NBA players. From my samples of 49 observations, I was not able to find a significant difference. Although ape index was larger for NBA players than normal males, I cannot safely conclude that NBA players have larger indexes.

6 APPENDICES

6.1 Data Provenance

My measurement data which came from those that have filled out the handout consists of 428 rows and 37 columns. This data contains both numerical and categorical data. I filtered the data to only include Male samples of at least 18 years old. It is my belief this is an appropriate age to select since the current minimum age to play in the NBA is 18. After data cleansing, data wrangling and filtering, I was left with 49 observations to work with. I went from 126 rows to 49 after removing all outliers, which points to the much of the data being inaccurate.

My NBA dataset that came from Kaggle is titled "NBA Players - Measurements (1947-2017)." The data set included 4551 rows and 18 columns. I was able to find a popular dataset that was publicly shared by Fernando Blanco on Kaggle. The Data represents numerous body measurements of players who have were in the NBA from 1947, all the way to 2017. I split my data into two seperate data sets: NBA players who played before or during the 2002 season and NBA players who played after the 2015 season. From these two data pools, I took random samples of 49 observations to match the number of observations from my measurement data. I also made a final merged dataset, In which I merged my 3 datasets and used in my analysis.

6.2 Correlation Table

The correlation table was made by running the function BuildLatexCorrelationTable(). This function was made by instructor: Monte Shaffer. The code to run the correlation table is in another file, however the output of the function which is a .tex file is already stored in a tables folder.

6.3 Kaggle Reference

Kaggle is a data science community website that offers thousands of #datasets.
https://www.kaggle.com/whitefero/nba-players-measurements-19472017
 Vitruvian Drawing by Leonardo da Vinci

https://en.wikipedia.org/wiki/Vitruvian_Man

6.3.1   Data Collection Handout

Figure 5: Handout Page 1

Please measure as accurately as possible and fill out the fields below that you can. If for some reason you cannot, please write "NA." Measure both left and right if possible, if not then fill out one side for all measurements. (L/R) indicates to measure both left and right sides. Please also fill out the quality of the measurements (Q) and the time it took (M) on the back page.

Person: _____

Data Collector: _____

Units: _____   Eye Color: _____

Dominant Hand: _____   Dominant Eye:_____

Dominant Swinging Arm: _____   Ethnicity: _____

Gender: _____   Age: _____

Height: _____   Head Circumference: _____

Head Height: _____   Hand Length (L/R): _____/_____

Hand to Elbow (L/R): _____/_____   Hand Width (L/R): _____/_____

Elbow to Armpit (L/R): _____/_____   Foot Length (L/R): _____/_____

Arm Reach(L/R): _____/_____   Arm Span: _____

Floor to Knee Pit (L/R): _____/_____   Floor to Armpit (L/R): _____/_____

Floor to Hip (L/R): _____/_____   Floor to Navel : _____

Nahom Debela
nahom.debela@wsu.edu
011510499
STATS 419
Instructor: Monte J. Shaffer
Date: 9/4/2020

Figure 6: Handout Page 2

.

TO BE COMPLETED LATER

Q: _____

M: _____

Notes:

References

Arm Reach: **Arm Span:**

Head Circumference: Distance around head, measured right about ears and eyes,

Floor to Navel: Floor to belly button.

Head Height: From chin to top of head.

Hand Length: Length of hand from middle finger to wrist

Hand Width: Width of hand from pinky finger to thumb fully stretched.

Nahom Debela
nahom.debela@wsu.edu
011510499
STATS

Instructor: Monte J. Shaffer
Date: 9/4/2020

419

6.4   Preparing the Report Workspace as a subsection

6.4.1   Preparing the Report Workspace as a subsubsection

Preparing the Report Workspace as a paragraph

Preparing the Report Workspace as a subparagrah   Below is the necessary functions and libraries required to run the code referenced in this document.

```r
library(devtools) # required for source_url
my.source = 'local';
local.path = "C:/Users/nahom/_git_/WSU_STATS419_FALL2020/";
# C:\Users\nahom\Desktop\STATS 419\secret\datasets
local.data.path.to.secret = "C:/Users/nahom/Desktop/STATS 419/secret/datasets/measure/";
local.data.path.to.secret.NBA = "C:/Users/nahom/Desktop/STATS 419/secret/datasets/measure-NBA/"
source( paste0(local.path, "functions/libraries.R"), local = T);

library(humanVerseWSU);
path.github = "https://raw.githubusercontent.com/MonteShaffer/humanVerseWSU/master/";
source_url("https://raw.githubusercontent.com/MonteShaffer/humanVerseWSU/master/humanVerseWSU/R/functio
# EDA functions
source_url( paste0(path.github,"humanVerseWSU/R/functions-EDA.R") );  #
source( paste0(local.path, "functions/functions-project-measure.R"), local = T);
```

Below is the code to load the data and prepare it for analysis.

```r
measure.uncleaned = readData("measure-students.txt")
measure.cleaned = dataCleaning(measure.uncleaned)

nbaData.uncleaned = readDataNBA("Player - Bio Stats (1947-2017).csv")
nbaData.cleaned.2015 = cleanNbaData2015(nbaData.uncleaned)
nbaData.cleaned.2002 = cleanNbaData2002(nbaData.uncleaned)
```

Below is the code I used in my analysis and plots

```r
# For summary tables
# I called get_summary() function
# get_summary(measure.cleaned.df,nbaData.2002,nbaData.2015)

# For height vs wingspan plot
# I called displayPlotHeightWing
# displayPlotHeightWing(measure.cleaned.df,nbaData.2015,nbaData.2002)
```

ENDNOTES

TABLE OF CONTENTS