How Age Affects Physical Attributes
Exploring An Answered Question

Christopher Sarno *Washington State University*

In this article we compare how *age* correlates to and affects a variety of variables, including *height*, *floor to hip length*, and *hip to head length*. The goal is to reject or fail to reject a null hypothesis that age affects these physical attributes equally.

Keywords: multiple comparisons to control; correlation strength; nonlinear growth curves; critical points

December 11, 2020

1 Introduction

Age affects everyone differently. The question that stood out to me when looking at the measure data was how age correlates with height, whether different parts of the body grew quicker than others, and if gender was a factor.

I decided to narrow down the comparisons to the few attributes listed above for their close relation to each other. Hip to floor is the height of the lower half of the body and we can find the height of the upper half of the body by subtracting the height of the lower half from the full height. This measurement is later referred to as "hip to head".

Figure 1: Conceptual Model

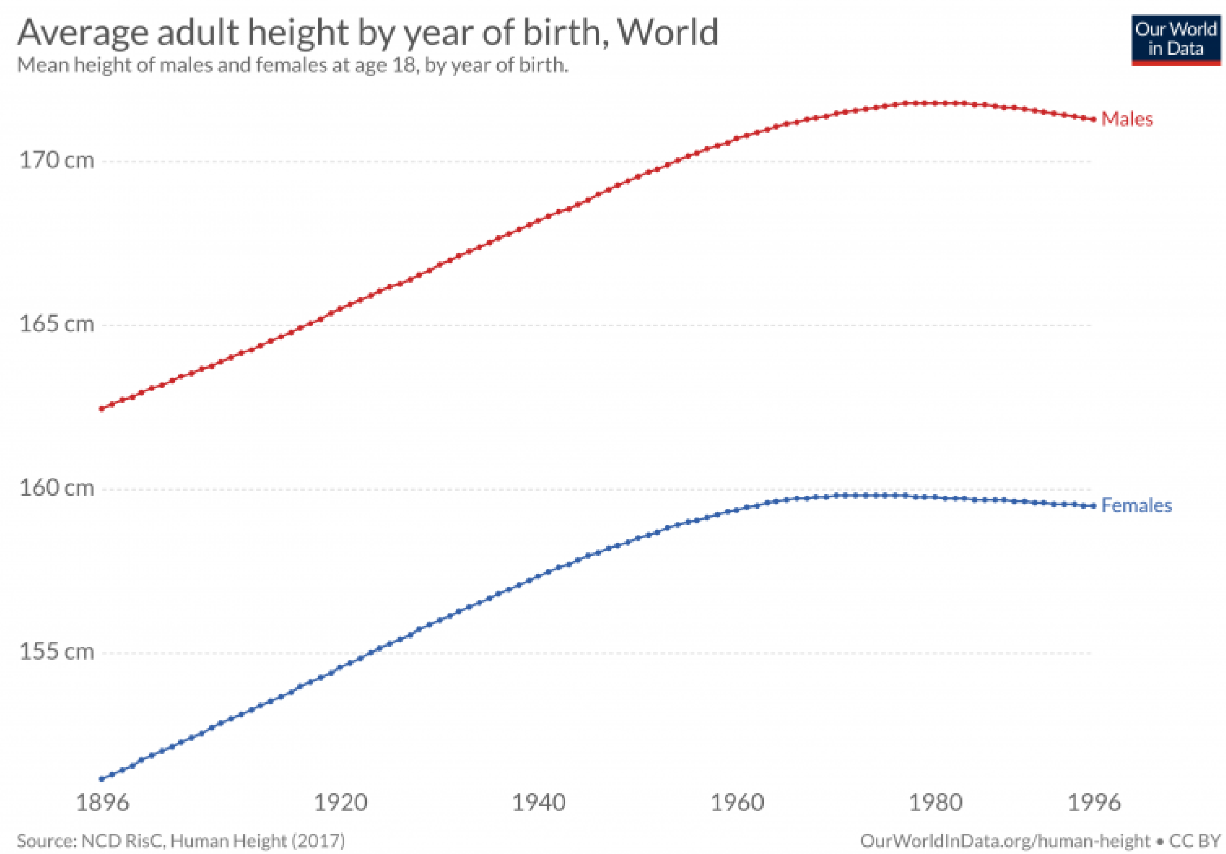
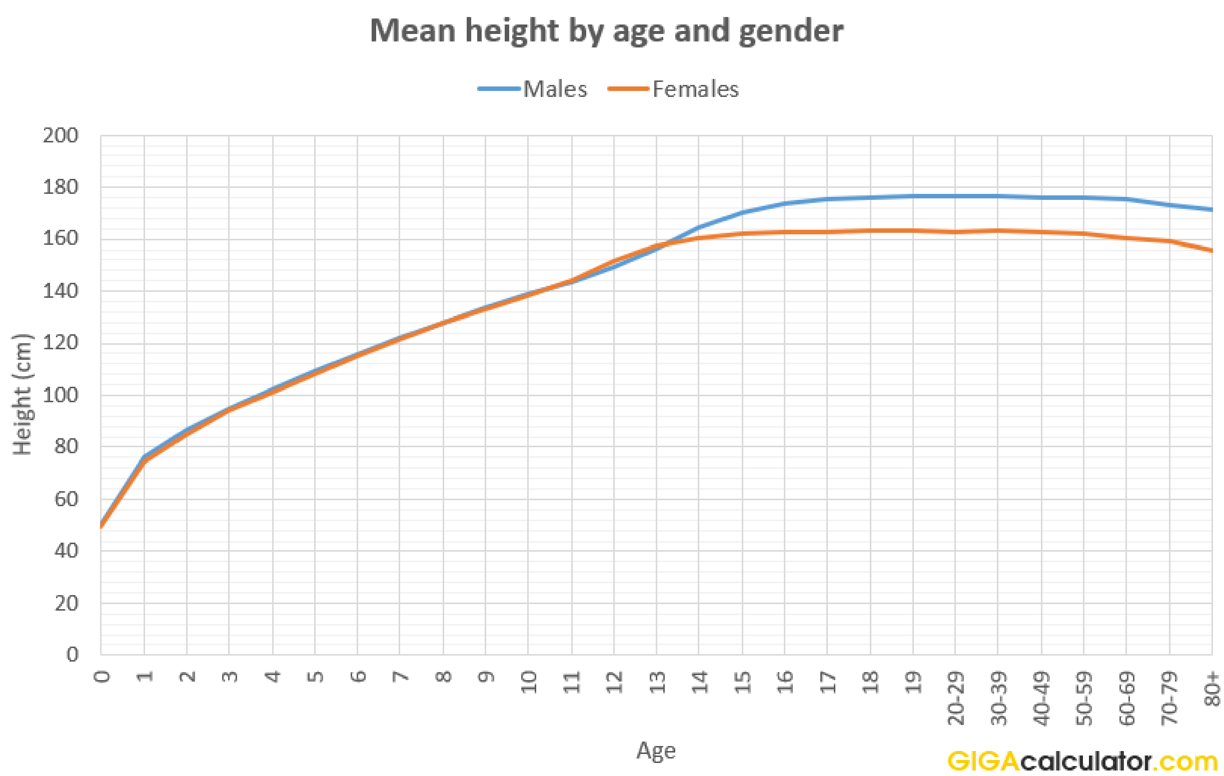


Figure 2: A look at average adult height throughout the past century

See Figure 3.

Figure 3: Conceptual Model



2 Research Question: The primary research question for this article is as follows: What is the age people start and stop growing?

2.1 Secondary question: Do different attributes scale differently with age than others?

2.2 Tertiary question: How strong is the correlation between age and height, and between those different attributes?

3 Data Description

The circumstances surrounding the gathering of this data was quite abnormal. Due to the Covid-19 pandemic, in-person data collection was just about off the table. A data collection sheet was instead drafted up, asking the participants to enter a number of personal information including covariates such as age, ethnicity, and gender, and a plethora of physical measurements.

The data collection handout was distributed virtually to ten participants per analyst and the data was compiled into one large database. Names of the participants were run through a md5 text converted, assuring anonymity.

3.1 Summary of Sample

Figure 4: Summary Of Data

age	height	floor.hip
Min. : 3.00	Min. :36.61	Min. :13.78
1st Qu.:22.00	1st Qu.:63.00	1st Qu.:35.98
Median :27.00	Median :66.50	Median :37.80
Mean :34.89	Mean :66.07	Mean :37.51
3rd Qu.:52.00	3rd Qu.:70.00	3rd Qu.:39.76
Max. :94.00	Max. :75.00	Max. :44.49

3.2 Summary Statistics of Data

Figure 5: Conceptual Model

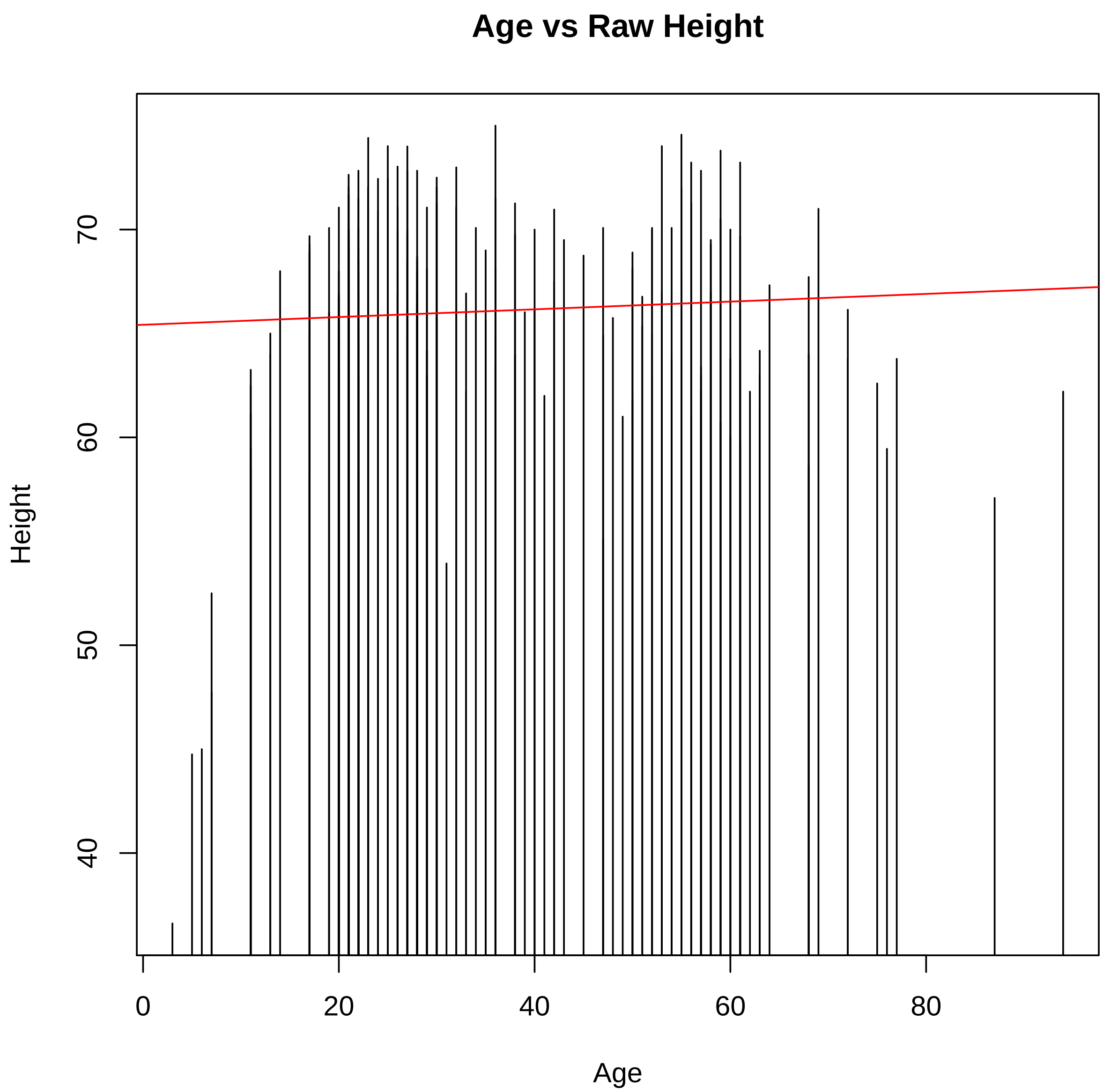


Figure 6: A histogram-like age vs height plot

Figure 7: Lower Half

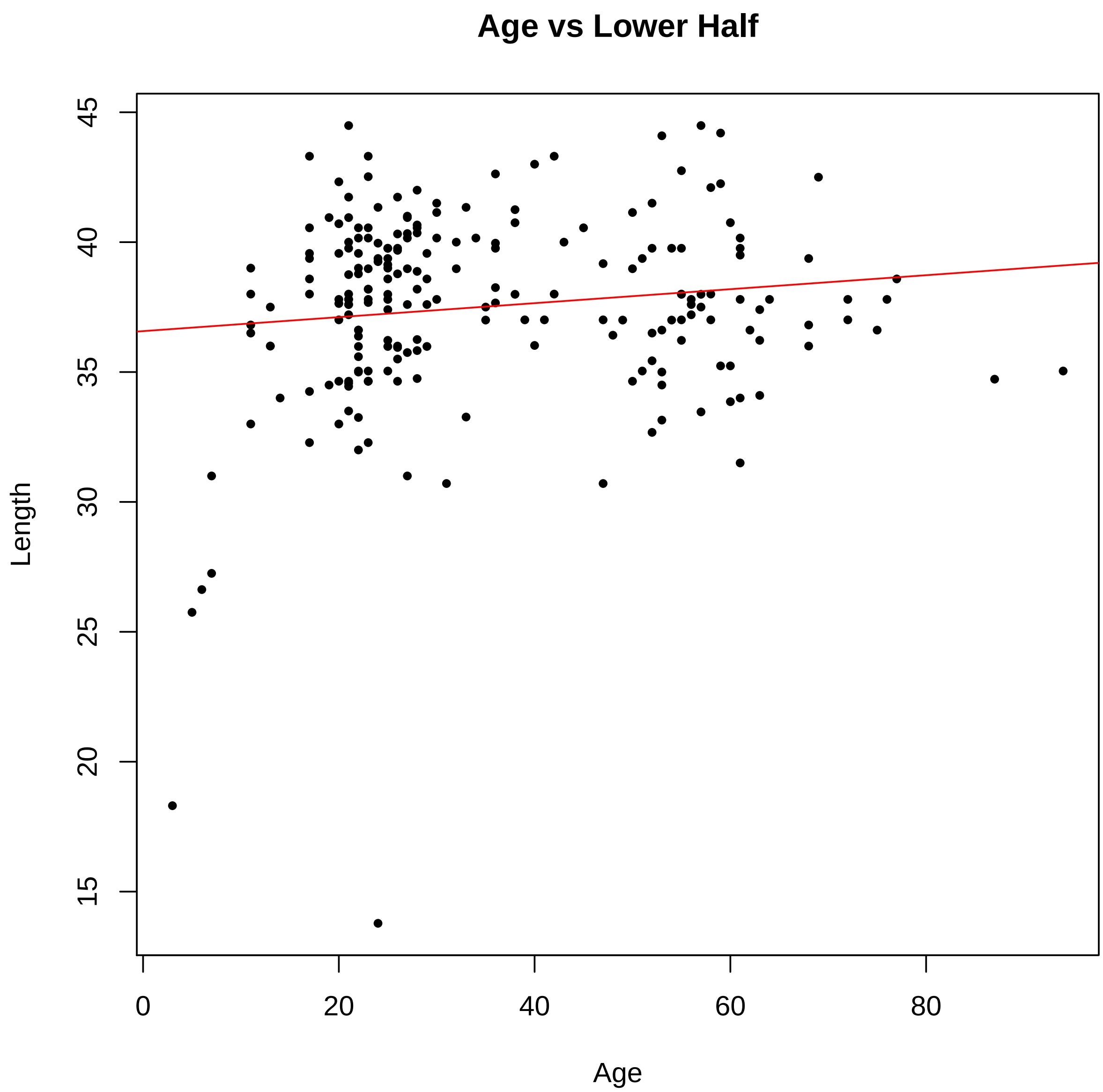


Figure 8: Plotting Age Vs Hip To Floor Measurement

Figure 9: Upper Half

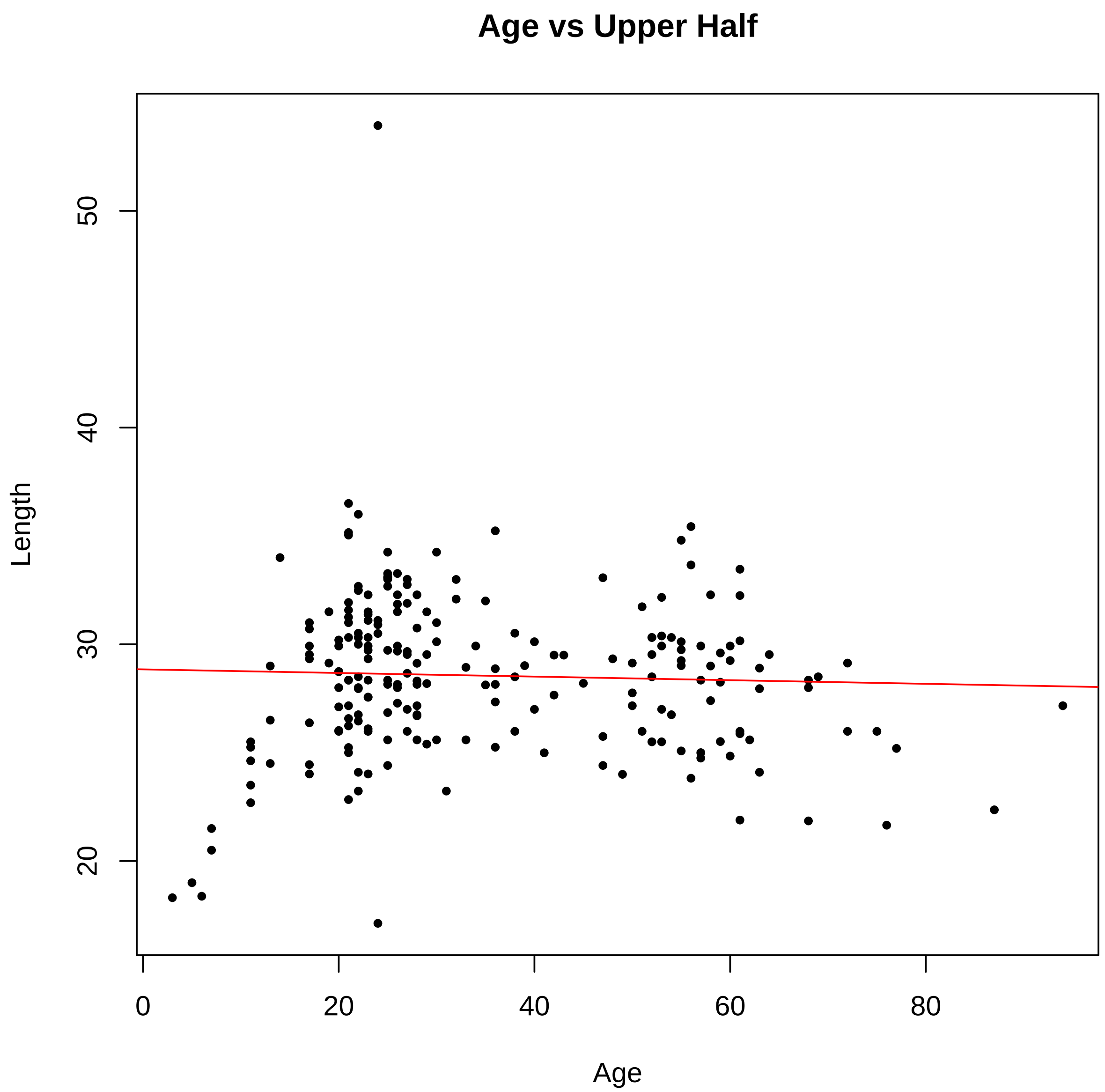


Figure 10: Plotting Age Vs Hip To Head Measurement

4 Key Findings

Beginning with the primary research question, Figure 5 indicates that growth begins very rapidly almost immediately after birth, and begins to plateau soon after age 20. Interestingly, the few datapoints we have for the height of those over the age of 60 indicates that height seems to regress a bit after age 60. However, due to the extremes at both ends, the correlation between age and height is very weak at 0.06.

To answer the secondary and tertiary questions regarding how different physical attributes scale with height, if they scale independently of one another, and their correlations, we will refer to the scatterplots, their regression lines, and correlation values. Figure 7 shows a constant, positive trend between age and the length from hip to floor. The scatterplot loosely follows that of the height graph, but with a bit more variation. However, the correlation between age and lower half height is quite weak with a value of only 0.128. While this is

We can compare this to Figure 9, which shows a constant, slightly negative trend between age and the length from hip to head. Again, however, the correlation between age and the length from hip to head is extremely weak and surprisingly negative at -0.04, indicating that the length between a person's hip and their head decreases with age.

While all referenced correlations are weak, it is significant to point out the difference between the rate at which age scales with the lengths of the upper half and lower half of the body. Ignoring intercepts, the length from floor to hip scales positively with 0.0269 times age yet the length from hip to head scales negatively with -0.008323 times age. This is a 3.5% difference in age scaling between the upper half and lower half of the body, which is not an insignificant amount.

5 Conclusion

To summarize the answers to our research questions, the critical points for growth are located at birth (x=0) and around age 20. Growth begins immediately and very rapidly starting at the first critical point, then tapers off and stagnates at the second critical point. While the scatterplots show that the upper half and the lower half of the body scale inversely and independently with age, the correlation between the three factors is too low to make a statement with any real confidence. Therefore, we fail to reject the null hypothesis that age affects the selected physical attributes equally.

Refer to the Appendices in section~6 where I am going to cite John (Tukey 1962, pp. 2-3).
Here is a quote by Tukey (1962, pp. 2-3):

> For a long time I have thought I was a statistician, interested in inferences from the particular to the general. But as I have watched mathematical statistics evolve, I have had to cause to wonder and to doubt. [...] All in all, I have come to feel that my central interest is in *data analysis*, which I take to include among other things: procedures for analyzing data, techniques for interpreting the results of such procedures, ways of planning the gathering of data to make its analysis easier, more precise or more accurate, and all the machinery and results of (mathematical) statistics which apply to analyzing the data.
>
> Large parts of data analysis are inferential in the sample-to-population sense, but these are only parts, not the whole. Large parts of data analysis are incisive, laying bare indications which we could not perceive by simple and direct examination of the raw data, but these too are only parts, not the whole. Some parts of data analysis, as the term is her stretch beyond its philology, are allocation, in the sense that they guide us in the distribution of effort and other valuable considerations in observation, experimentation, or analysis. Data analysis is a larger and more varied field than inference, or incisive procedures, or allocation.
>
> Statistics has contributed much to data analysis. In the future it can, and in my view should, contribute more. For such contributions to exist, and be valuable, it is not necessary that they be direct. They need not provide new techniques, or better tables for old techniques, in order to influence the practice of data analysis.

Table 1: Descriptive Statistics and Correlation Analysis

		M	SD	1	2	3
1	**age**	34.9	17.83	1		
2	**height**	66.1	5.38	.06	1	
3	**floor.hip**	37.5	3.74	.13†	.71***	1
4	**NA**	28.6	3.82	-.04	.72***	.01

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

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

6 APPENDICES

6.1 Data Provenance

As stated above, gathering the data from the source was difficult due to complications from the global pandemic. Since in-person data collection was impossible, it was impossible to assure impeccable quality of data.

Quality scores for the data was estimated by the analysts but we cannot be sure of its accuracy. As such, further studies and testing could be performed only on data that meet certain quality requirements, which perhaps could influence our analysis and findings presented in this article.

6.1.1 Data Collection Handout

Figure 11: Handout

Project Data Collection Handout

Participant Name :

Dominant Writing Hand (left/right) :

Dominant Eye (left/right) :

Eye Color :

Dominant Hand For Swinging :

Age :

Gender :

Ethnicity :

Quality of measurements :

Minutes taken to perform measurements :

Data Collector's Notes :

Side	Measurement	Value	Side	Measurement	Value
NA	Height		Right	Floor To Extended Arm	
NA	Head Height		NA	Arm Spam	
NA	Head Circumference		Left	Foot Length	
Left	Hand Length		Right	Foot Length	
Right	Hand length		Left	Floor to Knee Pit	
Left	Hand Width		Right	Floor to Knee Pit	
Right	Hand Width		Left	Floor to Hip	
Left	Hand to Elbow		Right	Floor to Hip	
Right	Hand to Elbow		NA	Floor to Navel	
Left	Elbow to Armpit		Left	Floor to Armpit	
Right	Elbow to Armpit		Right	Floor to Armpit	
Left	Floor To Extended Arm				

Note for participant: If only doing half of measurements, please do all left side or all right side.

Chris Sarno

Christopher.sarno@wsu.edu

11517322

6.2 Preparing the Report Workspace as a subsection

6.2.1 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.

```
library(devtools);       # required for source_url

path.humanVerseWSU = "https://raw.githubusercontent.com/MonteShaffer/humanVerseWSU/"
source_url( paste0(path.humanVerseWSU,"master/misc/functions-project-measure.R") );
```

```
## Warning: package 'Hmisc' was built under R version 4.0.3
```

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

```
path.project = "E:/Users/Mcshuggets/Documents/School/Math/Statistics/Stat 419 (Multivariate Statistics)/
path.figures = "E:/Users/Mcshuggets/Documents/School/Math/Statistics/Stat 419 (Multivariate Statistics)/

path.to.secret = "E:/Users/Mcshuggets/Documents/School/Math/Statistics/Secret/";

measure = utils::read.csv(paste0(path.to.secret,"measure-students.txt"), header=TRUE, quote="", sep="|")

#path.github = "https://raw.githubusercontent.com/this-IS-YOUR-PATH-TO-GITHUB/";
#source_url( paste0(path.github,"master/functions/functions-project-measure.R") );

# this is your function
# put in the same "units"
# merge left/right
# build proportion data
# and so on ...
prepareMeasureData<- function(file)
{
  df <- subset(file, select=c("age", "height","floor.hip" ))
  df <- na.omit(df)
  return(df)
}
measure.df = prepareMeasureData(measure);
measure.df$hip.head = measure.df$height-measure.df$floor.hip
path.tables = paste0(path.project,"tables/");
  createDirRecursive(path.tables);
file.correlation = paste0(path.tables,"age-height-table.tex");

measure.mrtx <- as.matrix(measure.df)
buildLatexCorrelationTable(measure.mrtx,
  rotateTable = FALSE,
  width.table = 0.75,
  myFile = file.correlation,
  myNames = c("age", "height", "floor.hip") );
```

Below is the code to generate the summary statistics and save them as a table that you see in Section 3.1.

```r
summary(measure.df)
```

```
##       age            height         floor.hip        hip.head    
##  Min.   : 3.00   Min.   :36.61   Min.   :13.78   Min.   :17.13  
##  1st Qu.:22.00   1st Qu.:63.00   1st Qu.:35.98   1st Qu.:25.98  
##  Median :27.00   Median :66.50   Median :37.80   Median :28.66  
##  Mean   :34.89   Mean   :66.07   Mean   :37.51   Mean   :28.55  
##  3rd Qu.:52.00   3rd Qu.:70.00   3rd Qu.:39.76   3rd Qu.:30.51  
##  Max.   :94.00   Max.   :75.00   Max.   :44.49   Max.   :53.94
```

```r
sink("summary.txt")
print(summary(measure.df))
```

```
##       age            height         floor.hip        hip.head    
##  Min.   : 3.00   Min.   :36.61   Min.   :13.78   Min.   :17.13  
##  1st Qu.:22.00   1st Qu.:63.00   1st Qu.:35.98   1st Qu.:25.98  
##  Median :27.00   Median :66.50   Median :37.80   Median :28.66  
##  Mean   :34.89   Mean   :66.07   Mean   :37.51   Mean   :28.55  
##  3rd Qu.:52.00   3rd Qu.:70.00   3rd Qu.:39.76   3rd Qu.:30.51  
##  Max.   :94.00   Max.   :75.00   Max.   :44.49   Max.   :53.94
```

```r
sink()
```

Below is the code to generate the graphs and correlation coefficients referenced in Section 3.2 and Section 4.

```r
# Building the age vs height graph
pdf(file=paste0(path.figures, "height-graph.pdf"))
plot(measure.df$age, measure.df$height,
     pch = 16, cex = 0.7,
     main="Age vs Raw Height",
     xlab="Age",
     ylab="Height",
     type="h")
abline(lm(measure.df$height~measure.df$age), col="red")
dev.off()
```

```
## pdf 
##   2
```

```r
lm(measure.df$height~measure.df$age)
```

```
## 
## Call:
## lm(formula = measure.df$height ~ measure.df$age)
## 
## Coefficients:
##    (Intercept)  measure.df$age  
##       65.41827         0.01858
```

```r
cor(measure.df$height, measure.df$age)
```

```
## [1] 0.06154268
```

```r
# Building the age vs hip to floor graph
pdf(file=paste0(path.figures, "hip-floor-graph.pdf"))
plot(measure.df$age, measure.df$floor.hip,
     pch = 16, cex = 0.7,
     main="Age vs Lower Half",
     xlab="Age",
     ylab="Length",
     type="p")
abline(lm(measure.df$floor.hip~measure.df$age), col="red")
dev.off()
```

```
## pdf
##   2
```

```r
lm(measure.df$floor.hip~measure.df$age)
```

```
##
## Call:
## lm(formula = measure.df$floor.hip ~ measure.df$age)
##
## Coefficients:
##    (Intercept)  measure.df$age
##        36.5750          0.0269
```

```r
cor(measure.df$floor.hip, measure.df$age)
```

```
## [1] 0.1281182
```

```r
# Building the age vs hip to head graph
pdf(file=paste0(path.figures, "hip-head-graph.pdf"))
plot(measure.df$age, measure.df$hip.head,
     pch = 16, cex = 0.7,
     main="Age vs Upper Half",
     xlab="Age",
     ylab="Length",
     type="p")
abline(lm(measure.df$hip.head~measure.df$age), col="red")
dev.off()
```

```
## pdf
##   2
```

```r
lm(measure.df$hip.head~measure.df$age)
```

```
##
## Call:
## lm(formula = measure.df$hip.head ~ measure.df$age)
##
## Coefficients:
##    (Intercept)  measure.df$age
##      28.843232       -0.008323
```

```r
cor(measure.df$hip.head, measure.df$age)
```

```
## [1] -0.03887501
```

```r
# Combining the lower half and upper half graphs
pdf(file=paste0(path.figures, "combined-graph.pdf"))
plot(measure.df$age, measure.df$hip.head,
     pch = 16, cex = 0.7,
     main="Lower vs Upper Half",
     xlab="Age",
     ylab="Length",
     type="p")
abline(lm(measure.df$hip.head~measure.df$age), col="red")
abline(lm(measure.df$floor.hip~measure.df$age), col="blue")
text(50, 40, "Lower Half")
text(80, 30, "Upper Half")
dev.off()
```

```
## pdf
##   2
```

```r
lm(measure.df$height~measure.df$age)
```

```
##
## Call:
## lm(formula = measure.df$height ~ measure.df$age)
##
## Coefficients:
##    (Intercept)  measure.df$age
##       65.41827         0.01858
```

```r
cor(measure.df$age, measure.df$height)
```

```
## [1] 0.06154268
```

ENDNOTES

REFERENCES

Tukey, John W. 1962. The Future of Data Analysis. *The Annals of Mathematical Statistics* **33**(1), 1–67.

TABLE OF CONTENTS