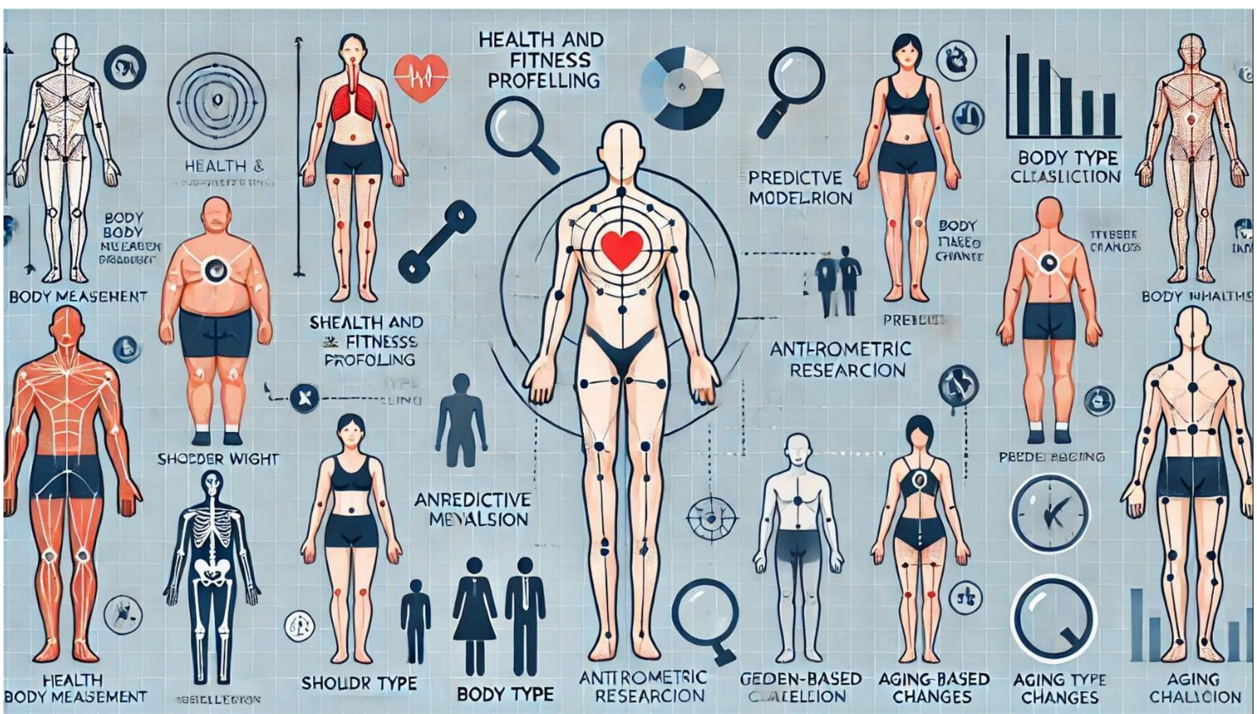
Fall 2024 TTU Multivariate Analysis

(ISQS-6350)

Body Measurement Analysis for Health Research



RAWLS COLLEGE - MSDS

Oreoluwa Coker - Coco Choo - Siavash Arbabi

Introduction

This section is authored by Siavash Arbabi.

This dataset contains detailed anthropometric (body measurement) information collected from 507 respondents. The dataset includes various body dimensions, girth measurements, demographic data, and physiological measurements that provide a comprehensive picture of the physical profiles of individuals. For more information about the breakdown of the columns, please visit the last page.

We found this dataset at Kaggle.
(https://www.kaggle.com/datasets/mexwell/body-measurements)

Motivation

This section is authored by Siavash Arbabi.

The primary motivation for studying this body measurement data is to gain insights into human body structure and understand how various measurements and demographic variables interrelate. Analyzing this dataset can help in several applied areas:

1. **Health and Fitness Profiling:**
   - By identifying patterns in body measurements, it's possible to classify different body types, which can aid in personalized health assessments, fitness plans, and nutritional recommendations.
   - Variables like waist girth, hip girth, and weight are associated with health metrics like BMI and body fat distribution, which can be risk factors for various health conditions.
2. **Classification of Body Types:**
   - Cluster analysis on body measurements could reveal natural groupings in the data, which could relate to established body types. This information is valuable for industries involved in fitness, apparel, and body composition analysis.

3. **Predictive Modeling for Health Metrics:**
    - The dataset allows predictive modeling, such as estimating weight based on body girths and diameters, which can be valuable in contexts where weight measurements are not feasible.
    - Linear regression or machine learning models could be built to predict weight, height, or other physical attributes from related measurements.
4. **Gender-Based Analysis:**
    - By analyzing body measurements by gender, we can identify dimensions that show significant differences. This is useful for understanding sex-based anatomical distinctions, aiding in tailored approaches for medical, ergonomic, or design applications.
    - Discriminant analysis may classify individuals by sex based on measurements, highlighting physical characteristics that differ between males and females.
5. **Aging and Physical Changes:**
    - Studying how body measurements correlate with age can provide insights into physical changes over time, such as increases in girth measurements or decreases in height, which are relevant for aging research and healthcare.

Data Preparation

This section is authored by Ore Coker and Coco Choo.

To prepare the body measurement data to be used in classification, prediction and analysis, preliminary data cleaning was conducted. This includes renaming columns to enhance clarity and comprehension, for example, changing column name wri_d to wrist_d.

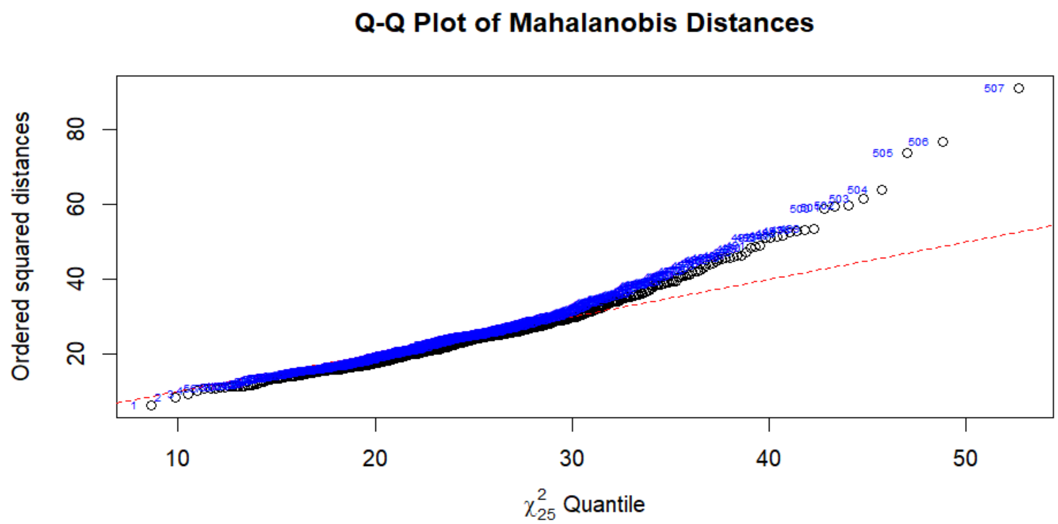


Figure 1: Mahalanobis Distance Plot to identify outliers

Initially, seven outliers were identified using a bivariate boxplot analysis between biacromial diameter and biiliac diameter. However, additional outliers were detected through a bivariate boxplot comparing weight and height. To incorporate all variables in the outlier identification process, a Mahalanobis distance plot was utilized. Based on Figure 1, seven respondents were identified as outliers, 501 through 507, and excluded from the dataset, leaving the body measurement data for a total of 500 respondents.

We then performed correlation analysis on both the original dataset, which includes all observations, and the cleansed dataset, with the identified outliers removed. A comparison of the correlation coefficients between the two datasets revealed minimal differences. This is likely due to the large sample size of the dataset, where the removal of seven outliers had a negligible impact on the overall correlations. Below is a partial correlation matrix of the cleansed dataset, with values rounded to two decimal places.

	biacromial_d	biiliac_d	bitrochanteric_d	chest_dep	chest_d
biacromial_d	1.00	0.30	0.48	0.59	0.77
biiliac_d	0.30	1.00	0.68	0.35	0.33
bitrochanteric_d	0.48	0.68	1.00	0.47	0.52
chest_dep	0.59	0.35	0.47	1.00	0.67
chest_d	0.77	0.33	0.52	0.67	1.00
elbow_d	0.76	0.31	0.52	0.66	0.76
wrist_d	0.72	0.27	0.46	0.61	0.73
knee_d	0.63	0.41	0.60	0.55	0.65
ankle_d	0.66	0.35	0.48	0.60	0.67
shoulder_gi	0.79	0.26	0.47	0.74	0.87
chest_gi	0.72	0.31	0.48	0.81	0.87
waist_gi	0.64	0.42	0.56	0.80	0.79
navel_gi	0.29	0.57	0.61	0.62	0.49
hip_gi	0.33	0.56	0.74	0.55	0.51
thigh_gi	0.11	0.42	0.52	0.35	0.30
bicep_gi	0.69	0.27	0.47	0.73	0.79
forearm_gi	0.75	0.27	0.47	0.72	0.81
knee_gi	0.50	0.45	0.61	0.56	0.59

Figure 2: Portion of the Correlation Matrix

The correlation matrix for the dataset, comprising 25 distinct variables, has dimensions of 25x25. As anticipated in body measurement data, many variables have strong positive correlations with one another, reflecting the inherent relationships between these measurements.

Visualizations

This section is authored by Coco Choo.

To further analyze the distribution of body measurement variables and uncover patterns within the data, kernel density plots were generated to provide a continuous estimate of the probability density function. Specifically, the distributions of body weight and height were examined using a Gaussian kernel to ensure smooth and accurate representation of the data.

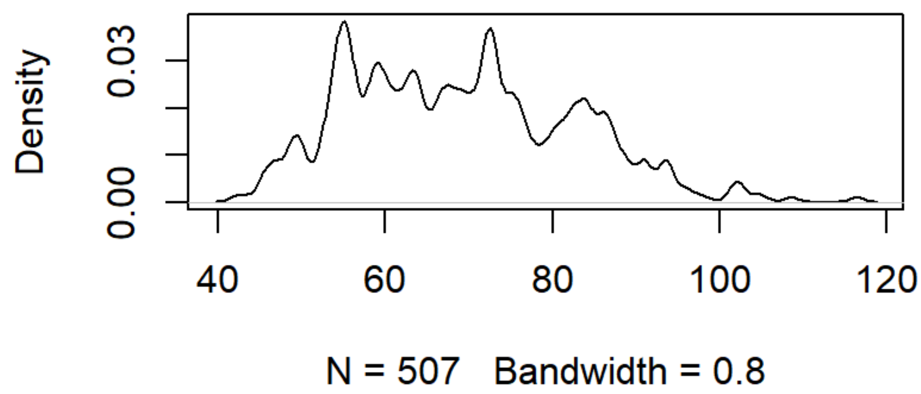


Figure 3: Kernel Density Plot of Weight (kg)

Figure 3 reveals that the density curve has multiple peaks, indicating a multimodal distribution. This observation aligns with the presence of two distinct subgroups in the data: males and females. The primary peaks are centered around 60 and 70, with additional smaller peaks observed. Based on the above plot, it is reasonable to infer that the peak near 60 corresponds to female weights, while the peak near 70 represents male weights.

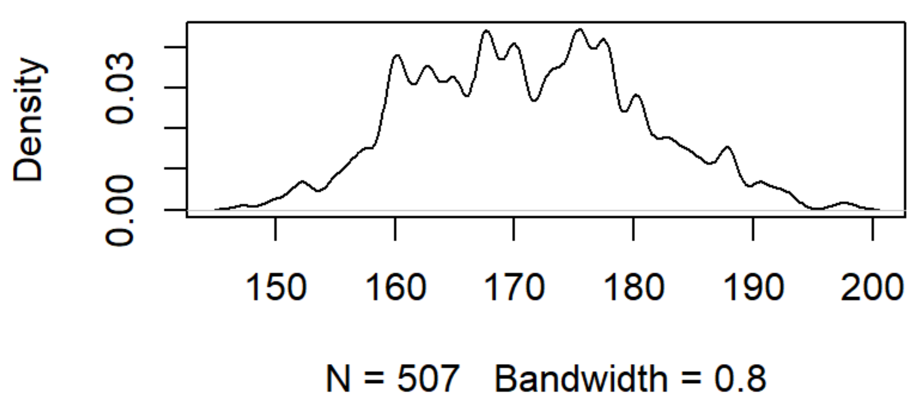


Figure 4: Kernel Density Plot of Height (cm)

Similarly, Figure 4 demonstrates that the density curve exhibits multiple peaks, further indicating a multimodal distribution. The primary peaks are observed around 165 and 175, with additional smaller peaks present. It is reasonable to infer that the peak near 165 corresponds to female heights, while the peak around 175 represents male heights.

Building on this understanding of the individual distributions of weight and height among respondents, a bivariate density plot was created to provide a more comprehensive analysis of the relationship between these variables, categorized by sex.

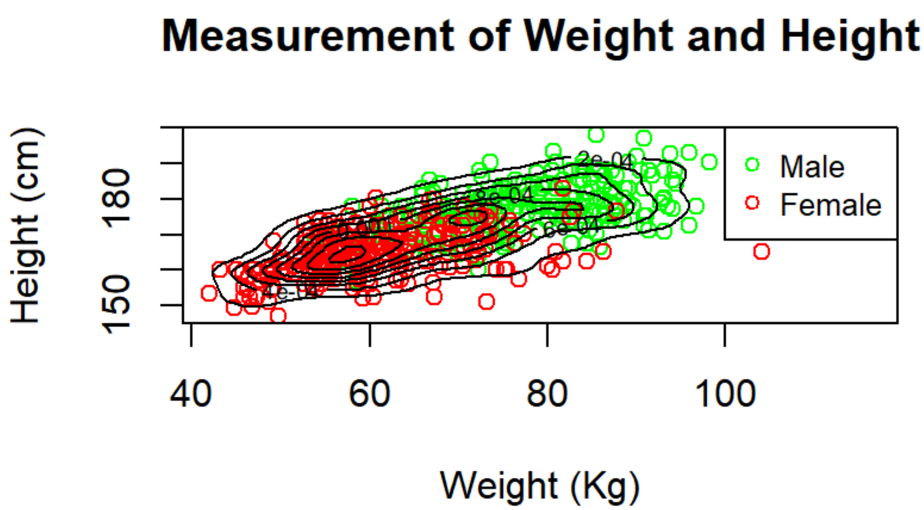


Figure 5: Bivariate Density Plot of Weight and Height by Sex

Figure 5 highlights potential biological differences, as evidenced by distinct high density contour regions for each sex. Males, which are represented by green circles, generally show higher densities in both weight and height compared to females, represented by red circles. Notably, a concentrated overlapping area with multiple data points clustered is observed, indicating shared ranges of weight and height between the two groups.

Dimension Reduction Analysis

This section is authored by Ore Coker.

With 25 distinct variables, the dataset is likely affected by the curse of dimensionality, which complicates analysis. To address this, Principal Component Analysis (PCA) was conducted to reduce the dataset's dimensionality while retaining as much of the original variation as possible.

	Comp.1	Comp.2	Comp.3	Comp.4	Comp.5
biacromial_d	0.20148176	0.203039806	0.0613983815	0.175888857	0.22194407
biiliac_d	0.12228632	-0.293354140	-0.1564792670	0.489223578	0.33407482
bitrochanteric_d	0.17156755	-0.262308029	-0.0407612554	0.365078612	0.11264573
chest_dep	0.20342225	-0.004366424	-0.2053329057	-0.202736160	0.13178322
chest_d	0.21952186	0.087473538	-0.0258467520	-0.089656956	0.17549800
elbow_d	0.22244090	0.162308467	0.0016827942	0.121035586	-0.05767279
wrist_d	0.21360060	0.162927725	0.0181450721	0.092468650	-0.18722062
knee_d	0.20694909	-0.032542214	0.1454169802	0.209184359	-0.24321536
ankle_d	0.20465261	0.123303600	-0.0341393694	0.241807213	-0.20863333
shoulder_gi	0.23128974	0.121081538	-0.0009033164	-0.179983356	0.12508684
chest_gi	0.23304140	0.073412041	-0.1037835150	-0.239974322	0.13826969
waist_gi	0.22511159	-0.049814221	-0.2152106531	-0.199202389	0.12029065
navel_gi	0.16996430	-0.320682659	-0.2817867751	-0.128803981	0.08543423
hip_gi	0.17786008	-0.382801754	0.0164100642	-0.090571504	0.06908473
thigh_gi	0.11962460	-0.431511514	0.2657559995	-0.220298256	0.04008916
bicep_gi	0.22857474	0.077050389	0.0079557799	-0.238029713	0.05206857
forearm_gi	0.23301796	0.124473553	0.0676071015	-0.140890087	-0.02004103
knee_gi	0.19935879	-0.202402402	0.2238192867	0.036842467	-0.20272783
calf_max_gi	0.19515904	-0.169999844	0.2660155322	-0.030831701	-0.29736705
ankle_min_gi	0.20418704	-0.039103728	0.1767613346	0.026747431	-0.33657881
wrist_min_gi	0.22657073	0.156370663	0.0558641963	-0.009211643	-0.17817941
age	0.06544149	-0.083355980	-0.7259859871	0.066490003	-0.47092893
wgt	0.24622156	-0.064396384	0.0220473992	-0.067095354	0.10882868
hgt	0.18693155	0.159205942	0.0676770284	0.360801136	0.23605930
sex	0.18726039	0.339502346	-0.0724032690	-0.012921686	0.08539778

Figure 6: Loadings for the First 5 Principal Components

Through PCA, five components were identified, representing 85% of the original dataset's variation. As a result, the dataset was reduced to these five components. A further review of the loadings for each component provided the following insights:

Component 1 represents overall body measurements.

Records with smaller Component 2 scores correspond to individuals with smaller biiliac diameter, bitrochanteric diameter, navel girth, hip girth, thigh girth, knee girth. These individuals are generally smaller in size.

Records with larger Component 3 scores indicate individuals with larger knee girth, thigh girth, calf max girth, and ankle min girth. These individuals tend to have wider legs.

Records with larger Component 4 scores represent individuals with larger ankle diameter and height but smaller chest girth, thigh girth, and bicep girth. These are taller individuals with a smaller upper body.

Records with larger Component 5 scores are associated with individuals having a larger biacromial diameter and height but smaller wrist diameter, ankle min girth, wrist min girth. These individuals are taller with broad shoulders but smaller hinge joints.

Cluster Analysis

This section is authored by Ore Coker.

To better understand and analyze our robust dataset, Cluster analysis with K-Means was conducted. To determine the optimal number of clusters, a scree plot was generated using the scaled data. Based on the plot, three clusters were identified as optimal, and a K-Means model was developed accordingly.

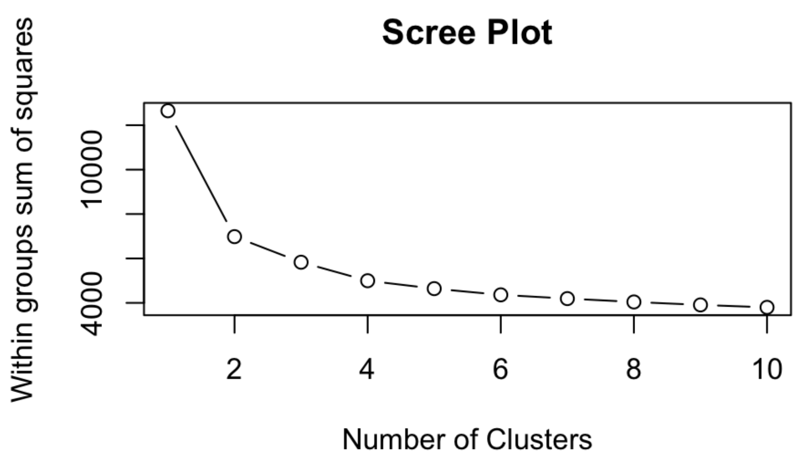


Figure 7: Scree Plot to Visualize the Optimal Number of Clusters

After clustering the data into three groups, summary tables were created to display the distribution of the groups and the column means for each cluster.

1	2	3
131	177	199

Figure 8: Classification Table for K-Means Clustering

Column Means for each Cluster

	biacromial_d	biiliac_d	bitrochanteric_d	chest_dep	chest_d	elbow_d	wrist_d	knee_d	ankle_d	shoulder_gi	chest_gi	waist_gi
Cluster 1	0.990	0.634	0.865	1.117	1.128	1.144	1.061	1.025	1.051	1.152	1.183	1.183
Cluster 2	0.256	0.002	0.073	0.084	0.184	0.185	0.209	0.200	0.170	0.240	0.209	0.150
Cluster 3	-0.879	-0.419	-0.635	-0.810	-0.906	-0.918	-0.884	-0.852	-0.843	-0.972	-0.964	-0.912

	navel_gi	hip_gi	thigh_gi	bicep_gi	forearm_gi	knee_gi	calf_max_gi	ankle_min_gi	wrist_min_gi	age	wgt	hgt	sex
Cluster 1	0.828	0.911	0.600	1.155	1.174	1.001	1.018	1.103	1.163	0.324	1.299	0.991	0.933
Cluster 2	0.118	0.035	-0.013	0.221	0.238	0.104	0.074	0.054	0.192	0.068	0.129	0.139	0.291
Cluster 3	-0.650	-0.631	-0.383	-0.957	-0.984	-0.752	-0.736	-0.774	-0.936	-0.274	-0.970	-0.776	-0.873

Figure 9: Cluster Column Means Table

The Cluster Column Means Table reveals the following insights:

Cluster 1 represents observations with larger measurements across all variables, corresponding to larger individuals. Cluster 2 represents observations with average measurements, with many means close to 0. These are average-sized individuals. Cluster 3 represents observations with smaller measurements across all variables, corresponding to smaller individuals.

This cluster analysis suggests that the data likely provides an accurate representation of the real world, reflecting three broad categories of body sizes and grouping observations accordingly.

Confirmatory Factor Analysis

This section is authored by Siavash Arbabi.

We used factor analysis for dimensionality reduction, identifying latent (hidden) variables that explain the underlying relationships among observed variables. This method allows us to gain insights into human body structure and understand how various measurements and demographic variables interrelate.

Using a two-factor model and evaluating the p-value, the null hypothesis - that two factors are sufficient - should be rejected. However, the RMSE for this model is only 5%, suggesting that the two-factor model is likely sufficient to effectively explain the data.

Factor 1 (Dominates upper-body and general measurements):

Variables such as biacromial_d (shoulder width), elbow_d, wrist_d, shoulder_gi, forearm_gi, and wrist_min_gi have high loadings (>0.8).

Sex has an extremely high loading (0.922), suggesting this factor heavily differentiates by biological sex, which strongly influences upper-body dimensions.

General body size measurements such as wgt (weight) and hgt (height) also have moderate loadings (0.739 and 0.724, respectively).

Factor 2 (Dominates lower-body and girth measurements):

Variables like bitrochanteric_d (hip width), hip_gi, thigh_gi, navel_gi, and waist_gi have high loadings (>0.7), indicating this factor captures variation in lower-body girths and widths.

Weight (wgt) also has a strong loading here (0.647), suggesting a connection between body mass and lower-body girths.

Loadings:

	Factor1	Factor2
biacromial_d	0.822	
biiliac_d		0.579
bitrochanteric_d		0.692
chest_dep	0.661	
chest_d	0.791	
elbow_d	0.859	
wrist_d	0.825	
knee_d	0.625	
ankle_d	0.754	
shoulder_gi	0.868	
chest_gi	0.835	
waist_gi	0.694	0.575
navel_gi		0.799
hip_gi		0.943
thigh_gi		0.867
bicep_gi	0.830	
forearm_gi	0.882	
knee_gi		0.694
calf_max_gi		0.633
ankle_min_gi	0.611	
wrist_min_gi	0.873	
age		
wgt	0.739	0.647
hgt	0.724	
sex	0.922	

Figure 10: Loadings for Exploratory Factor Analysis

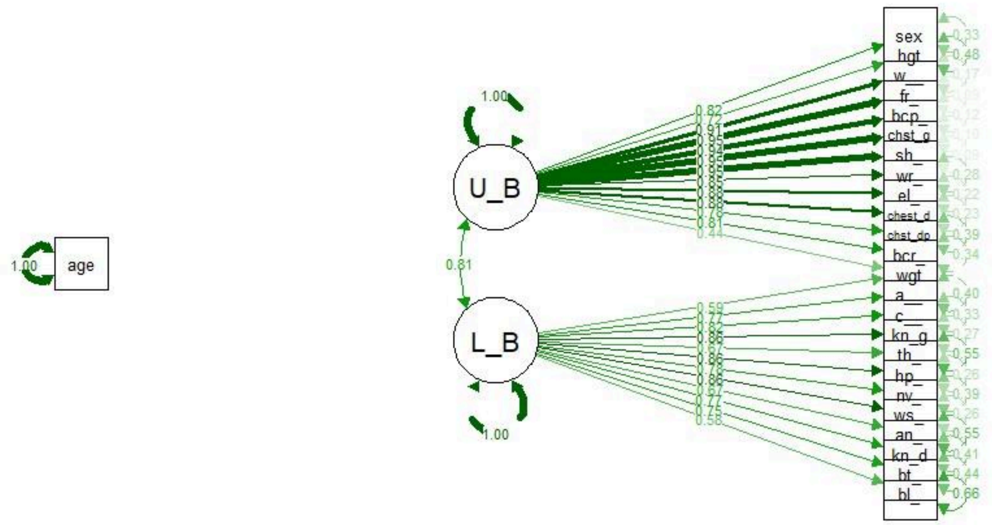


Figure 11: Path diagram for Confirmatory Factor Analysis

Based on the result of loadings, age has low loadings (less than 0.2) on both factors, indicating that it does not contribute substantially to either. Therefore, it is appropriate to treat age as an independent variable in the model. This can be done by specifying its variance without linking it to the latent factors.

Summary of Model

1. GFI (Goodness of Fit Index):

- Result: 0.4788
- The GFI is well below the recommended threshold of 0.90. Typically, a GFI value above 0.90 indicates a good fit, and values below 0.80 suggest a poor fit. A value of 0.4788 indicates a poor fit of the model to the data.

2. AGFI (Adjusted Goodness of Fit Index):

- Result: 0.3818
- Similar to GFI, AGFI is a more conservative measure. It also suggests a poor fit, as it is well below the acceptable threshold of 0.80. Values closer to 1 indicate a better fit.

3. SRMR (Standardized Root Mean Square Residual):

- Result: 0.1240
- The SRMR is a good measure of how well the model fits the data, with values below 0.1 indicating a good fit. A value of 0.1240 is above the recommended threshold, suggesting that the model does not fit the data well.

To enhance the existing model, we decided to exclude the four variables representing general size indicators - age, weight, height and sex - as they do not have a significant relationship with either the upper body or lower body factors. The model was subsequently reanalyzed using CFA.

Criterion	Value	Threshold	Condition Met
GFI	0.524909	> 0.90	FALSE
AGFI	0.416244	> 0.80	FALSE
SRMR	0.103581	< 0.1	FALSE

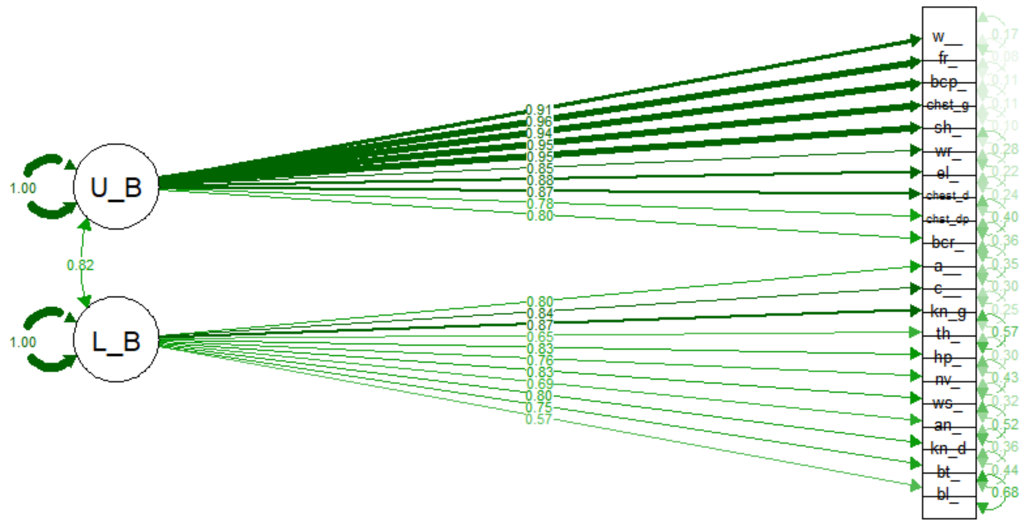


Figure 12: Path diagram for Confirmatory Factor Analysis(after dropping four columns)

The comparison between the original and revised models revealed minimal differences in overall fit indices. While the revised model showed slight improvements in metrics, the results were still unsatisfactory. This indicates that excluding the general size indicators had a limited impact on the model's ability to explain the underlying factor structure.

Findings

This section is authored by Coco Choo.

To sum up, our report provides a thorough analysis of a body measurement dataset, employing various statistical techniques to uncover meaning patterns and relationships. In the data preparation phase, outlier detection was conducted using Mahalanobis distance to ensure data quality, and the removal of seven outliers was shown to have minimal impact on overall correlations. Kernel density plots revealed multimodal distributions in weight (kg) and height (cm), suggesting distinct subgroups corresponding to sex differences. Further analysis through Principal Component Analysis reduced the dataset's dimensionality, capturing 85% of the variance with the first five components. This reduction provided insights into key body measurement variations. Cluster analysis with K-means identified three groupings of individuals: Larger, average and smaller body types. Lastly, Confirmatory Factor Analysis reinforced these findings and identified two primary factors representing the upper-body and the lower-body characteristics.

Likewise, the use of visualizations like density plots, bivariate contour plots, and scree plots effectively enhances the interpretability of the analysis and the insights derived from loadings and clusters reflect an effort to connect statistical findings to practical interpretations. However, the report could add more detailed explanations of the statistical methods and underlying assumptions used so it is accessible to readers without a strong statistical background. Additionally, while the results are well-presented and insightful, the report does not sufficiently address potential limitations. For example, it does not discuss how preprocessing choices like scaling or variable selection may have influenced the outcome of the analysis. This leaves questions about the broader applicability of the findings in real-world contexts. Addressing these aspects would significantly enhance thereport's overall credibility and depth. Suggestions for potential future work are outlined below.

Future Work

This section is authored by Siavash Arbabi.

1. Advanced Predictive Modeling

- **Objective**: Build machine learning models (e.g., decision trees, random forests, or neural networks) to predict health metrics such as BMI, fat percentage, or risk factors for diseases.
- **Benefits**: These models can provide actionable insights for healthcare providers and fitness planners.

2. Longitudinal Analysis

- **Objective**: Investigate how body measurements evolve over time if a similar dataset with repeated measures is available.
- **Benefits**: Insights into aging-related changes or the effects of lifestyle interventions.

3. Gender-Specific Analysis

- **Objective**: Perform more nuanced analyses that separate male and female data to study anatomical and physiological differences.
- **Benefits**: Useful for tailoring medical treatments, ergonomic designs, or fitness programs.

4. Custom Models for Specific Industries

- **Objective**: Develop models for industries such as fashion (e.g., clothing sizing), furniture design (e.g., ergonomic chairs)
- **Benefits**: Improves product design and market segmentation.

5. Development of Personalized Tools

- **Objective**: Use the findings to create user-friendly tools or applications that provide personalized health or fitness recommendations based on body measurements.
- **Benefits**: Encourages public adoption and practical utility of the research findings.

APPENDIX A

Column Breakdown:

1. **Biacromial Diameter (biacromial_d)**: The distance across the shoulders, measured in centimeters.
2. **Biiliac Diameter (biiliac_d)**: Also known as pelvic breadth, it is the width of the pelvis measured in centimeters.
3. **Bitrochanteric Diameter (bitrochanteric_d)**: The distance between the two trochanters of the femur, indicating hip width, measured in centimeters.
4. **Chest Depth (chest_dep)**: The depth of the chest measured between the spine and sternum at nipple level, mid-expiration, in centimeters.
5. **Chest Diameter (chest_d)**: The diameter of the chest at nipple level, mid-expiration, measured in centimeters.
6. **Elbow Diameter (elbow_d)**: The combined width of both elbows, measured in centimeters.
7. **Wrist Diameter (wrist_d)**: The combined width of both wrists, measured in centimeters.
8. **Knee Diameter (knee_d)**: The combined width of both knees, measured in centimeters.
9. **Ankle Diameter (ankle_d)**: The combined width of both ankles, measured in centimeters.
10. **Shoulder Girth (shoulder_gi)**: The circumference around the shoulders, measured over the deltoid muscles, in centimeters.
11. **Chest Girth (chest_gi)**: The circumference around the chest, measured at the nipple line for males and above breast tissue for females, mid-expiration, in centimeters.
12. **Waist Girth (waist_gi)**: The circumference at the narrowest part of the torso below the rib cage, taken as an average of contracted and relaxed measurements, in centimeters.

13. **Navel (Abdominal) Girth (navel_gi)**: The circumference around the abdomen measured at the level of the umbilicus and iliac crest, in centimeters.
14. **Hip Girth (hip_gi)**: The circumference around the hips at the level of the bitrochanteric diameter, in centimeters.
15. **Thigh Girth (thigh_gi)**: The circumference of the thigh measured below the gluteal fold, taken as an average of the right and left thighs, in centimeters.
16. **Bicep Girth (bicep_gi)**: The circumference of the biceps when flexed, taken as an average of the right and left arms, in centimeters.
17. **Forearm Girth (forearm_gi)**: The circumference of the forearm when extended, palm up, taken as an average of the right and left arms, in centimeters.
18. **Knee Girth (knee_gi)**: Repeated measurement of knee diameter, measured as the sum of both knees, in centimeters.
19. **Calf Maximum Girth (calf_max_gi)**: The maximum circumference of the calves, taken as an average of the right and left calves, in centimeters.
20. **Ankle Minimum Girth (ankle_min_gi)**: The minimum circumference around the ankles, taken as an average of the right and left ankles, in centimeters.
21. **Wrist Minimum Girth (wrist_min_gi)**: The minimum circumference around the wrists, taken as an average of the right and left wrists, in centimeters.
22. **Age (age)**: The age of the respondent in years.
23. **Weight (wgt)**: The weight of the respondent in kilograms.
24. **Height (hgt)**: The height of the respondent in centimeters.
25. **Sex (sex)**: A categorical variable where 1 indicates male and 0 indicates female.

Detecting Outliers:

```
x <- data

xbar <- colMeans(x)

S <- cov(x)

d2 <- mahalanobis(x, xbar, S)

numVariables <- ncol(x)

position <- (1:nrow(x) - 1/2) / nrow(x)

quantiles <- qchisq(position, df = numVariables)

plot(quantiles, sort(d2), xlab = expression(paste(chi[25]^2, " Quantile")),  ylab = "Ordered squared distances", main = "Q-Q Plot of Mahalanobis Distances")

abline(a = 0, b = 1, col = "red", lty = 2)

text(quantiles, sort(d2), labels = 1:nrow(x), col = "blue", pos = 2, cex = 0.5)
```

Visualizations:

```
library(MVA)

library("KernSmooth")
```

Figure 3: `plot(density(bodym$wgt, bw = 0.8, kernel = "gaussian"))`

Figure 4: `plot(density(bodym$hgt, bw = 0.8, kernel = "gaussian"))`

Figure 5:

```
bw <- c(dpik(bodym$wgt), dpik(bodym$hgt))

body_wh <- bodym[, c(23,24)]

density <- bkde2D(body_wh, bandwidth = bw)

col = ifelse(bodym$sex == 1, "green", "red")

plot(body_wh, xlab = "Weight (Kg)", ylab = "Height (cm)", main="Measurement of Weight and Height", col = col)
```

```
contour(x = density$x1, y = density$x2, z = density$fhat, add = TRUE)

legend("topright", legend = c("Male", "Female"), col = c("green", "red"), pch = 1, cex = .8)
```

Principal Component Analysis:

```
body.s <- scale(data)

body.pca <- princomp(body.s, cor = T)

body.pca$loading[, 1:5]
```

K-Means Clustering Analysis:

```
library(mclust)

library(knitr)

body.s <- scale(bodym)

plot.wgss = function(mydata, maxc = nrow(mydata) - 1) {

wgss = numeric(maxc)

for (i in 1:maxc) {

km <- kmeans(mydata, centers = i, nstart = 20)

wgss[i] = km$tot.withinss

}

plot(1:maxc, wgss, type = "b", xlab = "Number of Clusters", ylab = "Within groups sum of
squares",

main = "Scree Plot")

}

plot.wgss(body.s, maxc = 10)

km <- kmeans(body.s, centers = 3, nstart = 10)

km.t <- table(km$cluster)
```

km.t

clusterMeans <- km$centers

rownames(clusterMeans) <- c("Cluster 1", "Cluster 2", "Cluster 3")

kable(clusterMeans, caption = "Column Means for each Cluster")

Confirmatory Factor Analysis:

Figure 10:

data.fa<- factanal(bodym, factors = 2)

data.fa

corr<- cor(bodym)

f.loading<- data.fa$loadings[,1:2]

corHat<- (f.loading)%*%t(f.loading)+diag(data.fa$uniquenesses)

rsme<- sqrt(mean((corr-corHat)^2))

rsme

print(data.fa$loadings,cut=0.5)

Figure 11:

library("sem")

body_model <- specifyModel(text = "

Upper_Body -> biacromial_d, lambda1, NA

Upper_Body -> chest_dep, lambda2, NA

Upper_Body -> chest_d, lambda3, NA

Upper_Body -> elbow_d, lambda4, NA

Upper_Body -> wrist_d, lambda5, NA

```
Upper_Body     -> shoulder_gi, lambda6, NA
Upper_Body     -> chest_gi, lambda7, NA
Upper_Body     -> bicep_gi, lambda8, NA
Upper_Body     -> forearm_gi, lambda9, NA
Upper_Body     -> wrist_min_gi, lambda10, NA
Upper_Body     -> wgt, lambda11, NA
Upper_Body     -> hgt, lambda12, NA
Upper_Body     -> sex, lambda13, NA
Lower_Body     -> biiliac_d, lambda14, NA
Lower_Body     -> bitrochanteric_d, lambda15, NA
Lower_Body     -> waist_gi, lambda16, NA
Lower_Body     -> navel_gi, lambda17, NA
Lower_Body     -> hip_gi, lambda18, NA
Lower_Body     -> thigh_gi, lambda19, NA
Lower_Body     -> knee_d, lambda20, NA
Lower_Body     -> ankle_d, lambda21, NA
Lower_Body     -> knee_gi, lambda22, NA
Lower_Body     -> calf_max_gi, lambda23, NA
Lower_Body     -> ankle_min_gi, lambda24, NA
Lower_Body     -> wgt, lambda25, NA
Upper_Body     <-> Lower_Body, rho, NA
biacromial_d   <-> biacromial_d, theta1, NA
chest_dep      <-> chest_dep, theta2, NA
chest_d        <-> chest_d, theta3, NA
elbow_d        <-> elbow_d, theta4, NA
wrist_d        <-> wrist_d, theta5, NA
shoulder_gi    <-> shoulder_gi, theta6, NA
```

```
chest_gi       <-> chest_gi, theta7, NA
bicep_gi       <-> bicep_gi, theta8, NA
forearm_gi     <-> forearm_gi, theta9, NA
wrist_min_gi   <-> wrist_min_gi, theta10, NA
wgt            <-> wgt, theta11, NA
hgt            <-> hgt, theta12, NA
sex            <-> sex, theta13, NA
biiliac_d      <-> biiliac_d, theta14, NA
bitrochanteric_d <-> bitrochanteric_d, theta15, NA
waist_gi       <-> waist_gi, theta16, NA
navel_gi       <-> navel_gi, theta17, NA
hip_gi         <-> hip_gi, theta18, NA
thigh_gi       <-> thigh_gi, theta19, NA
knee_d         <-> knee_d, theta20, NA
ankle_d        <-> ankle_d, theta21, NA
knee_gi        <-> knee_gi, theta22, NA
calf_max_gi    <-> calf_max_gi, theta23, NA
ankle_min_gi   <-> ankle_min_gi, theta24, NA
age            <-> age, theta25, NA
Upper_Body     <-> Upper_Body, NA, 1
Lower_Body <-> Lower_Body, NA, 1
")


body_sem <- sem(body_model, corr, nrow(bodym))


library("semPlot")
```

```
semPlot::semPaths(body_sem, rotation = 2, 'std', 'est')

options(fit.indices = c("GFI", "AGFI", "SRMR"))

criteria=summary(body_sem)

criteria$GFI

criteria$AGFI

criteria$SRMR


criteria$GFI > 0.90

criteria$AGFI > 0.80

criteria$SRMR < 0.1
```

Figure 12:

```
data_new <- bodym[, -(22:25)]

data.fa<- factanal(data_new,factors = 2)

data.fa

corr<- cor(data_new)

f.loading<- data.fa$loadings[,1:2]

corHat<- (f.loading)%*%t(f.loading)+diag(data.fa$uniquenesses)

rsme<- sqrt(mean((corr-corHat)^2))

rsme

print(data.fa$loadings)
```

```
library("sem")

body_mode3 <- specifyModel(text = "
Upper_Body     -> biacromial_d, lambda1, NA
Upper_Body     -> chest_dep, lambda2, NA
Upper_Body     -> chest_d, lambda3, NA
Upper_Body     -> elbow_d, lambda4, NA
Upper_Body     -> wrist_d, lambda5, NA
Upper_Body     -> shoulder_gi, lambda6, NA
Upper_Body     -> chest_gi, lambda7, NA
Upper_Body     -> bicep_gi, lambda8, NA
Upper_Body     -> forearm_gi, lambda9, NA
Upper_Body     -> wrist_min_gi, lambda10, NA
Lower_Body     -> biiliac_d, lambda11, NA
Lower_Body     -> bitrochanteric_d, lambda12, NA
Lower_Body     -> knee_d, lambda13, NA
Lower_Body     -> ankle_d, lambda14, NA
Lower_Body     -> waist_gi, lambda15, NA
Lower_Body     -> navel_gi, lambda16, NA
Lower_Body     -> hip_gi, lambda17, NA
Lower_Body     -> thigh_gi, lambda18, NA
Lower_Body     -> knee_gi, lambda19, NA
Lower_Body     -> calf_max_gi, lambda20, NA
Lower_Body     -> ankle_min_gi, lambda21, NA
Upper_Body    <-> Lower_Body, rho, NA
biacromial_d  <-> biacromial_d, theta1, NA
chest_dep     <-> chest_dep, theta2, NA
chest_d       <-> chest_d, theta3, NA
```

```
elbow_d        <-> elbow_d, theta4, NA
wrist_d        <-> wrist_d, theta5, NA
shoulder_gi    <-> shoulder_gi, theta6, NA
chest_gi       <-> chest_gi, theta7, NA
bicep_gi       <-> bicep_gi, theta8, NA
forearm_gi     <-> forearm_gi, theta9, NA
wrist_min_gi   <-> wrist_min_gi, theta10, NA
biiliac_d      <-> biiliac_d, theta11, NA
bitrochanteric_d <-> bitrochanteric_d, theta12, NA
waist_gi       <-> waist_gi, theta13, NA
navel_gi       <-> navel_gi, theta14, NA
hip_gi         <-> hip_gi, theta15, NA
thigh_gi       <-> thigh_gi, theta16, NA
knee_d         <-> knee_d, theta17, NA
ankle_d        <-> ankle_d, theta18, NA
knee_gi        <-> knee_gi, theta19, NA
calf_max_gi    <-> calf_max_gi, theta20, NA
ankle_min_gi   <-> ankle_min_gi, theta21, NA
Upper_Body     <-> Upper_Body, NA, 1
Lower_Body <-> Lower_Body, NA, 1
")


body_sem <- sem(body_mode3, corr, nrow(data_new))


library("semPlot")

semPlot::semPaths(body_sem, rotation = 2, 'std', 'est')
```

```
options(fit.indices = c("GFI", "AGFI", "SRMR"))

criteria=summary(body_sem)

criteria$GFI

criteria$AGFI

criteria$SRMR


criteria$GFI > 0.90

criteria$AGFI > 0.80

criteria$SRMR < 0.1
```