Estimating Body Fat Using Circumference Measurements*

Abdomen as the Strongest Predictor With Contributions From Wrist, Height, and Age For Affordable Health Monitoring

Wendy Yuan

November 29, 2024

This study uses a multi-linear regression model to estimate body fat percentage based on circumference measurements. Abdominal size was the strongest predictor, followed by wrist size, height, and age, which showed minimal contributions. These results demonstrate that accessible methods can effectively estimate body fat, offering methods suitable for non-clinical settings. This approach is important for identifying and managing health risks associated with extremely low or high body fat levels in a practical and affordable way.

Table of contents

*Code and data are available at: https://github.com/kiwindyy/Body-Fat

1 Introduction

Monitoring body fat is important because of the major role it plays in overall health. Current methods, like DXA scans and underwater weighing, are very reliable but often expensive, difficult to access, or only available in specialized facilities. More affordable methods, like using a measuring tape for body size, can provide reasonable estimates of body fat and are easier for more people to use (Tinsley 2023). A lack of easy and reliable ways to measure body fat can lead to missed chances to detect health risks. Excess body fat is strongly linked to conditions such as heart disease, type 2 diabetes, high blood pressure, and even certain cancers. It can also increase strain on the joints, leading to issues like osteoarthritis, and contribute to breathing problems like sleep apnea (Diabetes, Digestive, and Diseases 2023). Extremely low body fat can interfere with key body functions, affecting hormone levels, immune health, and even heart rhythm. For example, not enough body fat can lead to irregular heartbeats, extreme fatigue, and reproductive issues (Fetters 2023). Without proper tools to measure and manage body fat, individuals may unknowingly face these risks, making it important to research methods that are both reliable and affordable.

This study examines the reliability of using circumference measurements to estimate body fat. Starting with a dataset containing many measurements, we used backward selection to identify the most relevant variables impacting body fat. The four variables are: Age, Height, Abdomen, and Wrist circumference. After making sure these variables did not display collinearity, we tested the linear regression assumptions to confirm the model's validity. The selected variables were analyzed in a multiple linear regression model and individually in simple linear regression plots against body fat. This approach gives a clear understanding of how each variable relates to body fat. Finally, we evaluated the coefficients in the multiple regression model to make sure they aligned logically with the model, supporting its reliability.

The study found that different body measurements had different effects on body fat percentage. Abdominal size had the strongest influence, with larger measurements strongly linked to higher body fat. Wrist size and height were associated with lower body fat, meaning that people with bigger wrists or greater height tend to have less fat overall. Age played a smaller role, with older individuals showing a slight tendency to have more body fat, but the effect was weaker than the other factors. Together, these measurements offered a practical way to estimate body fat, with abdominal size standing out as the most important factor.

The remainder of this paper is structured as follows: Section 2 explains the dataset source, how it was collected, and describes the variables used. Section 3 outlines the multi-linear regression setup and the rationale for its use. Section 4 presents the regression outcomes, showing how the predictors explain body fat percentage. Section 5 interprets the findings, notes limitations, and suggests improvements. Section A explores survey, sampling, and observational data methods, including measurement techniques, simulations.

1.1 Estimand

The estimand of this study is body fat percentage, which we aim to estimate based on measurements from The Data And Story Library (Ricard n.d.). Body fat percentage is an important health metric that cannot be measured directly for every individual, especially outside of specialized settings. To address this, we used a multi-linear regression model, combining predictors such as age, height, abdomen, and wrist circumference to estimate body fat. By analyzing these variables, this study seeks to understand their effects on body fat. The model not only provides a framework for estimation making sure our approach is both structured and practical.

2 Data

2.1 Measurement

This study uses the statistical programming language R (R Core Team 2024) and the packages: tidyverse (Wickham et al. 2019), Arrow (Richardson et al. 2024), MASS (Venables and Ripley 2002), GGally (Schloerke et al. 2024), gridExtra (Auguie 2017), Knitr (Xie 2024), kableExtra (Zhu 2024), Car (Fox and Weisberg 2019). The dataset used in this study, sourced from The Data And Story Library (Ricard n.d.), contains measurements on body fat percentage alongside various physical metrics for 250 male participants. The dataset was collected to analyze body fat as an important health measure, providing variables such as age, weight, height, and circumferences of specific body parts. These measurements were obtained through direct assessments to ensure accuracy and uniform input into the dataset. The dataset allows researchers to consider accessible ways to estimate body fat without relying on costly or specialized methods.

Before analysis, the raw dataset was cleaned to standardize and prepare the data for use. Since the original dataset included measurements in both inches and centimeters, all measurements were converted into centimeters to ensure uniformity, as detailed in Table 1 & Table 2. Additional cleaning steps included removing duplicates, addressing outliers, and removing missing observations to improve dataset's usability. This cleaned data makes sure that the variables in the study are accurate and consistent. For reference, the original dataset before cleaning is provided in Table 4 & Table 5, showing the changes made during this process.

Table 1: Cleaned Data of Body Fat Variables Part 1

Density	Pct.BF	Age	Weight	Height	Neck	Chest	Abdomen
1.0708	12.3	23	69.96	172.08	36.2	93.1	85.2
1.0853	6.1	22	78.58	183.51	38.5	93.6	83.0
1.0414	25.3	22	69.85	168.27	34.0	95.8	87.9
1.0751	10.4	26	83.80	183.51	37.4	101.8	86.4

1.0340	28.7	24	83.57	180.97	34.4	97.3	100.0
1.0502	20.9	24	95.36	189.86	39.0	104.5	94.4
1.0549	19.2	26	82.10	177.16	36.4	105.1	90.7
1.0704	12.4	25	79.83	184.15	37.8	99.6	88.5
1.0900	4.1	25	86.63	187.96	38.1	100.9	82.5
1.0722	11.7	23	89.92	186.69	42.1	99.6	88.6

Table 2: Cleaned Data of Body Fat Variables Part 2

Waist	Hip	Thigh	Knee	Ankle	Bicep	Forearm	Wrist
85.2	94.5	59.0	37.3	21.9	32.0	27.4	17.1
83.0	98.7	58.7	37.3	23.4	30.5	28.9	18.2
87.9	99.2	59.6	38.9	24.0	28.8	25.2	16.6
86.4	101.2	60.1	37.3	22.8	32.4	29.4	18.2
100.0	101.9	63.2	42.2	24.0	32.2	27.7	17.7
94.4	107.8	66.0	42.0	25.6	35.7	30.6	18.8
90.7	100.3	58.4	38.3	22.9	31.9	27.8	17.7
88.5	97.1	60.0	39.4	23.2	30.5	29.0	18.8
82.5	99.9	62.9	38.3	23.8	35.9	31.1	18.2
88.6	104.1	63.1	41.7	25.0	35.6	30.0	19.2

2.2 Outcome variables

The dependent variable in this study is percent body fat, which represents the proportion of an individual's total body mass that is composed of fat. This variable is a continuous measure and serves as the target outcome we aim to estimate using predictor variables. Percent body fat is an important indicator of health, as both excessive and extremely low levels can lead to various medical issues. In our analysis, understanding the distribution of this variable is essential for ensuring that it aligns with the assumptions required for linear regression modeling.

As shown in Figure 1, the histogram of percent body fat has a shape that is close to bell-like, though it is not perfectly symmetrical. Most individuals in the dataset have body fat levels clustered within a common range, with fewer individuals having much higher or much lower levels. There is a slight tendency for more values to stretch toward higher body fat percentages, creating a small imbalance in the shape of the distribution. The spread of values covers a wide range, reflecting the diversity in body fat levels across the individuals studied. At the extremes, a few outliers stand out - some with very low body fat and others with notably high levels.

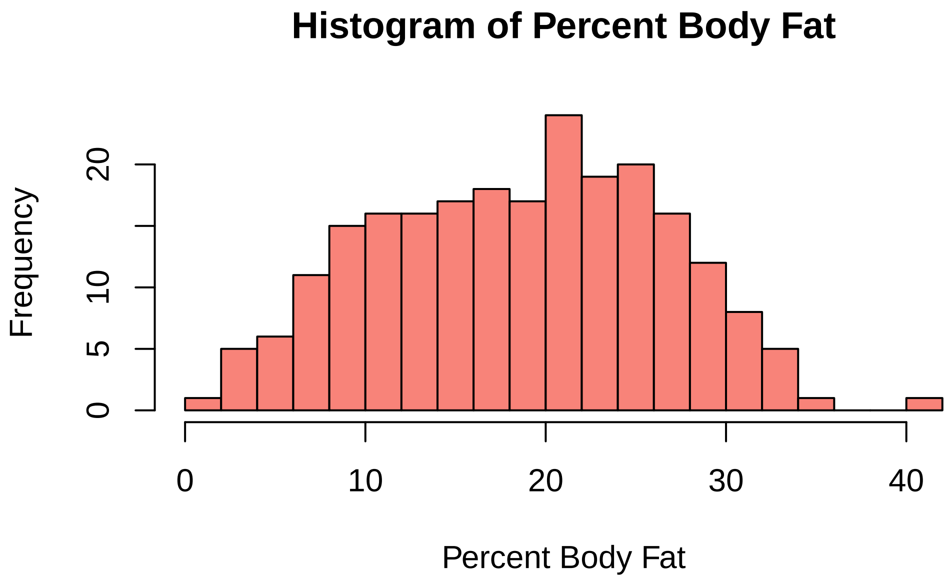

Figure 1: Distribution of Percent Body Fat

2.3 Predictor variables

The original uncleaned dataset contained the variables below:

- Density: The measure of body density, used as a key variable in estimating body fat percentage. It combines body weight and volume to provide insights into overall composition.
- Age: The age of the individual in years, providing a measure of the person's stage of life, which can influence body composition.
- Weight: The total body mass of the individual, reflecting the combined weight of bones, muscles, fat, and other tissues.
- Height: The standing height of the individual, often used alongside weight to calculate proportions and indices like BMI.
- Neck: The circumference of the neck, offering a measurement of fat and muscle distribution in the upper body.
- Chest: The circumference of the chest, reflecting the size and structure of the upper torso, including muscle and fat.
- Abdomen: The circumference around the abdomen, a key indicator of central fat distribution and a significant predictor of body fat percentage.
- Waist: The measurement around the waist, often used with hip circumference to assess body shape and fat distribution.
- Hip: The circumference of the hips, providing a measure of lower-body proportions and often paired with waist circumference to calculate ratios.

- Thigh: The circumference of the thigh, measured to understand muscle and fat distribution in the upper leg.
- Knee: The circumference of the knee, offering additional detail about lower-body structure and proportions.
- Ankle: The measurement around the ankle, reflecting skeletal and soft tissue composition in the lower leg.
- Bicep: The circumference of the bicep, measured with the arm flexed, highlighting upper-arm muscle and fat distribution.
- Forearm: The circumference of the forearm, providing information about the composition of the lower arm.
- Wrist: The circumference of the wrist, a useful proxy for skeletal size and overall body frame.

To focus on the most important variables, this study used backward selection to find the four that gave the best model fit, with the lowest Akaike Information Criterion (AIC). Backward selection works by starting with all the variables and gradually removing those that add the least to explaining body fat percentage, based on statistical rules. This process ensures the final set of variables is strongly related to the outcome. As shown in Figure 2, this approach narrowed the variables down to Age, Height, Abdomen, and Wrist circumference. These four were chosen because they worked together to provide the best explanation of the differences in body fat levels while keeping the model simple and effective.

x
Age
Height
Abdomen
Wrist

Figure 2: AIC Optimizing Best 4 Variables

The histograms of Age, Height, Abdomen, and Wrist shown in Figure 3 explain their role in predicting body fat. Most participants are between 30 and 60 years old, with fewer older individuals, making the sample mainly middle-aged. Height is evenly spread and follows a balanced shape, making it a reliable variable for the analysis. Abdomen size is mostly within a common range, with a few larger measurements that could influence the results due to their link to higher body fat. Wrist size has less variation but is consistent and still important because of its connection to lower body fat. Overall, these variables work well for modeling body fat, although the uneven spread of Age might need adjustments for better results.

The pairwise plot, Figure 4, shows how Age, Height, Abdomen, and Wrist are related to each other. Abdomen and Wrist have the strongest relationship. Larger abdominal circumferences are generally linked to larger wrists, showing overall body size. Age has a weak connection

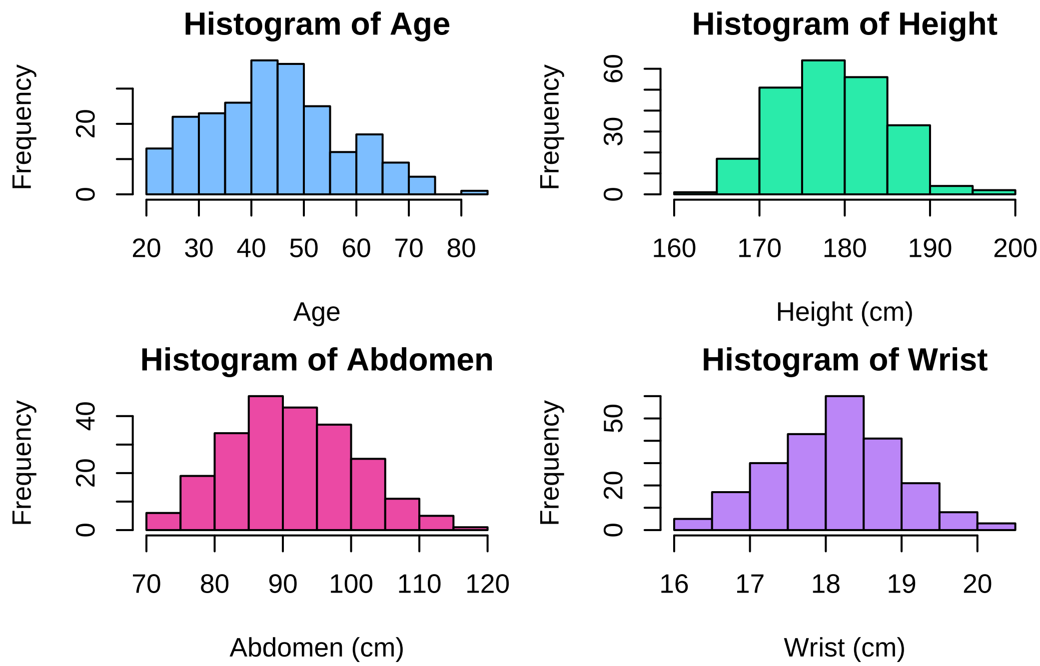

Figure 3: Distributions of Age, Height, Abdomen, Wrist

to Abdomen. Older individuals tend to have slightly larger abdominal sizes. However, Age shows no clear relationship with Height or Wrist. Height has small links to both Abdomen and Wrist. Taller individuals tend to have slightly smaller abdominal sizes and larger wrists. Abdomen and Wrist are the most important variables for predicting body fat. Height and Age provide smaller, additional contributions. The lack of strong correlations between these variables means they work well together in a regression model.

The variance inflation factor (VIF) is used to check if the predictors in a regression model are too closely related, a problem called multicollinearity. VIF shows how much a predictor's contribution is influenced by its relationship with other variables. A VIF of 1 means no overlap, while values over 5 or 10 might cause issues and need fixing.

Figure 5 shows the VIF values for Age, Height, Abdomen, and Wrist are all below 2 - meaning there is very little overlap between them. Wrist has the highest value (1.724), but this is still low, meaning it is only slightly related to other variables. The scatterplots showed some small relationships between variables, and the VIF confirms that these are not strong enough to cause problems. This means each variable adds its own useful information to the model.

3 Model

The goal of our model is to estimate body fat percentage using four predictors: Age, Height, Abdomen, and Wrist circumference. This helps us understand how each variable affects body fat while making sure the model gives accurate estimates.

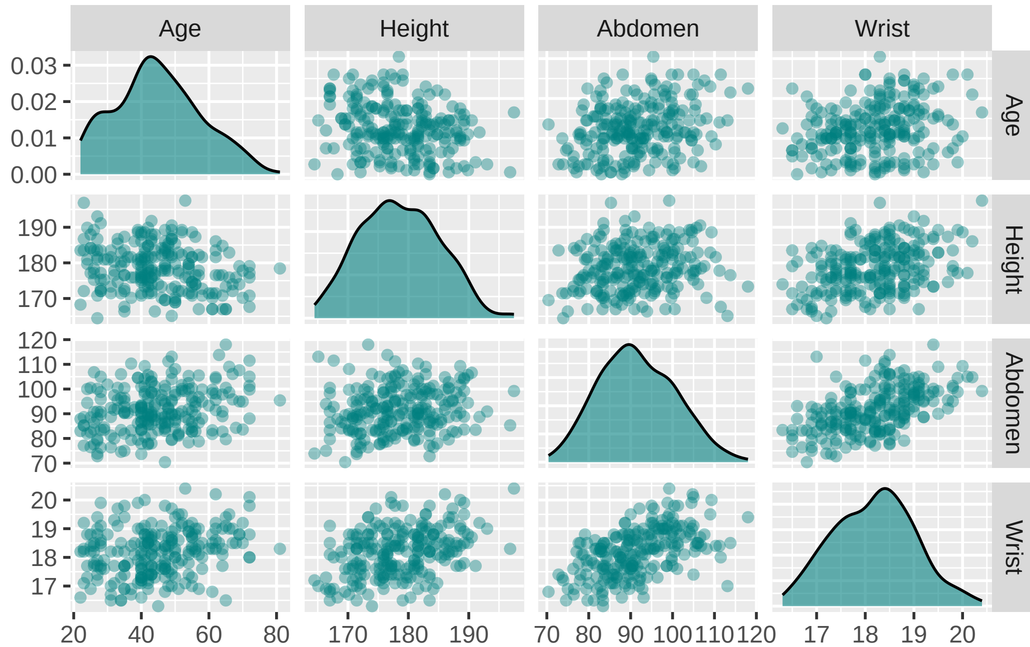

Figure 4: Pairwise Plot Compairing the Relationship Between Age, Heigh, Abdomen, Wrist

	x
Age	1.266260
Height	1.346920
Abdomen	1.450437
Wrist	1.695736

Figure 5: VIF Values of Age, Height, Abdomen, Wrist

3.1 Model set-up

Define y_i as the body fat percentage for the i-th individual. The predictors in the model are:

- x_{1i}: Age, representing the individual's age in years.
- x_{2i}: Height, measured in centimeters.
- x_{3i}: Abdomen, the abdominal circumference in centimeters.
- x_{4i}: Wrist, the wrist circumference in centimeters.

The model is expressed as:

$y_i = \beta_0 + \beta_1 x_{1i} + \beta_2 x_{2i} + \beta_3 x_{3i} + \beta_4 x_{4i}$

where β_0 is the intercept, $\beta_1, \beta_2, \beta_3, \beta_4$ are the coefficients for each predictor, and ϵ_i represents the error term, assumed to be normally distributed with mean 0. Diagnostic plots and residual checks were performed to assess whether the model passed linear regression assumptions

3.1.1 Model justification

Linear regression has four key assumptions to ensure the model works well. Linearity means the relationship between the predictors and the outcome should be straight, so the model captures the correct pattern. Constance variance means the residuals should have the same spread across all predictor variables. Normality requires the errors to follow a normal distribution. Independence ensures the errors are not related to each other, avoiding bias. Checking these assumptions helps make sure the model is accurate and trustworthy.

The residual plot, shown in Figure 6, helps assess both the linearity and independence assumptions in the linear regression model. There is no strong curve or trend in the residuals, confirming that the linearity assumption holds well. The residuals are randomly scattered around the red horizontal line at zero, with no noticeable patterns or clustering, indicating that the model correctly captures the relationships between the predictors and the dependent variable. This randomness also suggests that the residuals are independent, as they are not systematically related to each other or to the fitted values. Together, these results confirm that the model satisfies both the linearity and independence assumptions, providing a solid basis for interpreting the results.

The Scale-Location plot, Figure 7, is used to assess the assumption of constant variance in a linear regression model. The standardized residuals are plotted against the fitted values, with the red line representing the trend. For the assumption of constant variance to hold, the points should be scattered randomly around the red line without forming a pattern or a funnel shape. In this plot, the points seem to be spread out evenly across the predicted values, with no clear signs of the spread getting bigger or smaller. This means the errors have a consistent variance confirming the linear regression assumption.

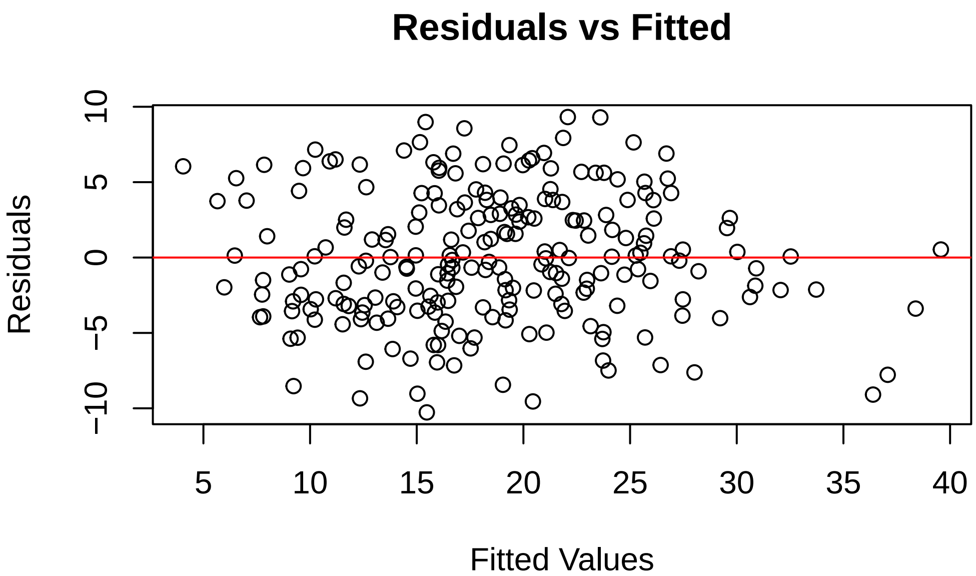

Figure 6: Residual Plot Verifying the Linearity Assumption

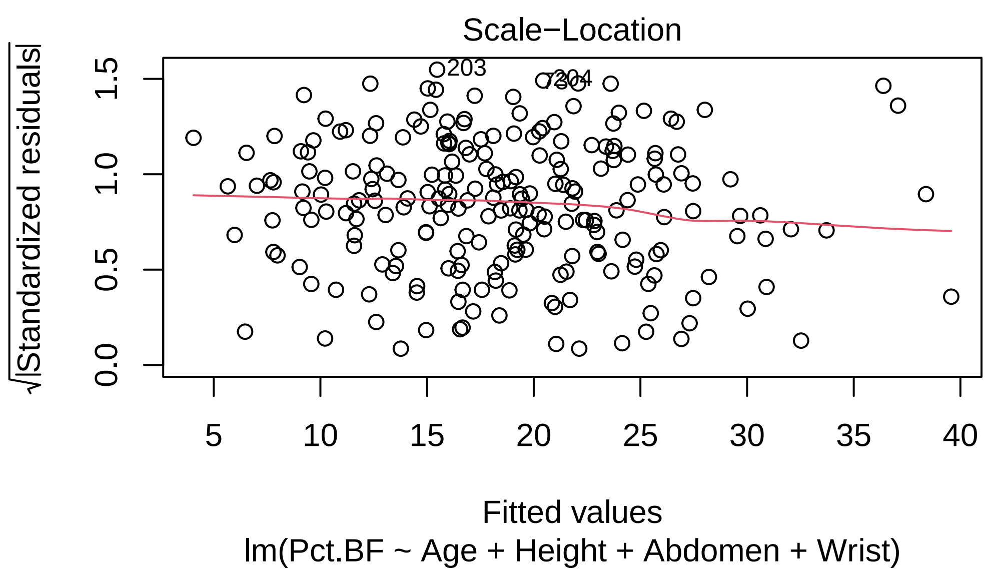

Figure 7: Scale Location Verifying Constance Variance Assumption

The Normal Q-Q plot, Figure 8 checks if the residuals (errors) in the model follow a normal distribution. If the errors are normal, the points should stay close to the red line. Here, most points are close to the line, meaning the errors are roughly normal. At the ends, some points are farther from the line, which might mean there are some extreme values or outliers. Overall, the errors seem mostly normal satisfying the linear regression assumption.

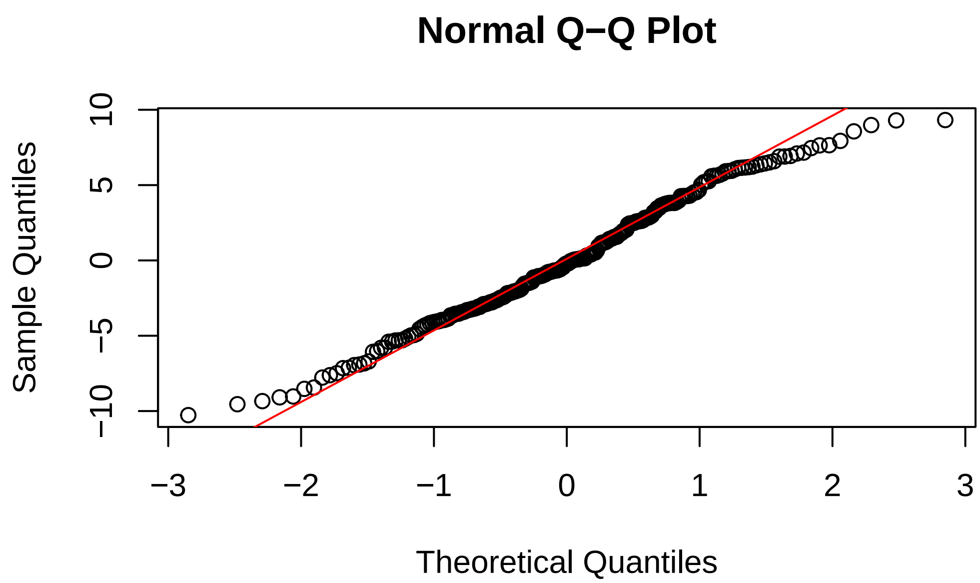

Figure 8: QQ-Plot Verifying the Normality of Errors Assumption

We expect larger body measurements to be linked to higher body fat percentages. Specifically, people with bigger abdominal or wrist circumferences are likely to have more body fat. Abdomen size often shows how much fat is stored in the center of the body, which contributes a lot to total body fat. Wrist circumference, while less variable, can reflect overall body size and fat distribution. In contrast, we expect taller individuals to have less body fat because taller people usually have leaner body compositions.

We also expect older individuals to have slightly more body fat. As people age, changes in metabolism and lifestyle can lead to gradual increases in fat. By using these variables in a multi-linear regression model, we can measure how each one affects body fat percentage. For example, the model will estimate how much body fat percentage increases when abdomen size grows by one unit, while keeping the other variables unchanged. This helps us understand the role of each measurement in predicting body fat.

4 Results

The scatterplots from Figure 9 provide an initial look at the relationships between body fat percentage and the variables: Age, Height, Abdomen, and Wrist circumference. It looks at whether trends exist between the dependent variable and the predictors, giving a sense of how each variable may contribute to the model.

- Pct.BF vs Age: There is a slight positive trend, indicating that body fat percentage tends to increase as Age increases. While this relationship is not very strong, it suggests age could have a modest impact on body fat.
- Pct.BF vs Height: The plot shows no clear trend, with a nearly flat line indicating that Height likely has little influence on body fat percentage.
- Pct.BF vs Abdomen: A strong positive relationship is visible, as body fat percentage increases significantly with Abdomen circumference. This suggests Abdomen size is a key predictor of body fat due to its role in central fat storage.
- Pct.BF vs Wrist: A weak negative trend is observed, suggesting that individuals with larger Wrist circumferences tend to have slightly lower body fat percentages.

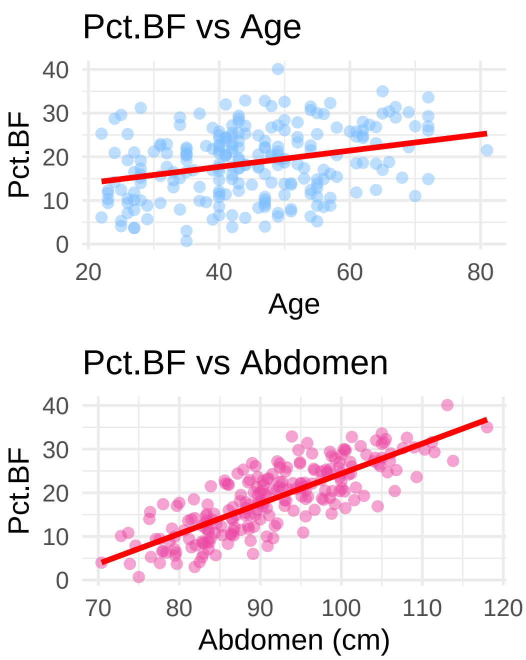

Figure 9: Simple Linear Regression Plots Against Body Fat Percent for Age, Height, Abdomen, Wrist

The regression results are summarized in Table 3. The intercept, which represents the estimated body fat percentage when all predictors are zero, is 2.01. However, this value doesn't have much real-world meaning because variables like height and abdomen size cannot actually be zero. Among the predictors, Abdomen circumference has the strongest positive effect on body fat percentage. For every 1 cm increase in abdomen size, body fat percentage increases by about 0.78%, while keeping other factors unchanged.

Wrist circumference shows a significant negative relationship with body fat percentage. This

Table 3: Multi-linear Regression Model

```
Call:
lm(formula = Pct.BF ~ Age + Height + Abdomen + Wrist, data = data)

Residuals:
     Min       1Q   Median       3Q      Max 
-10.2708  -3.0998  -0.2117   3.3125   9.3195 

Coefficients:
            Estimate Std. Error t value Pr(>|t|)    
(Intercept)  2.79763    8.86682   0.316   0.7527    
Age          0.04491    0.02559   1.755   0.0806 .  
Height      -0.12353    0.05130  -2.408   0.0168 *  
Abdomen      0.78091    0.03816  20.463  < 2e-16 ***
Wrist       -1.96214    0.45097  -4.351 2.06e-05 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 4.33 on 223 degrees of freedom
Multiple R-squared:  0.6982,	Adjusted R-squared:  0.6927 
F-statistic: 128.9 on 4 and 223 DF,  p-value: < 2.2e-16
```

means that people with larger wrists tend to have lower body fat percentages. Height also has a small negative effect, indicating that taller individuals are likely to have slightly lower body fat. Age has a weak positive effect, meaning body fat percentage may slightly increase with age, but this result is less certain.

The model explains about 70.3% of the variation in body fat percentage, showing it fits the data well. The residual standard error is 4.33, meaning the model's predictions are fairly accurate on average. Overall, Abdomen and Wrist are the most important predictors, with Height and Age playing smaller roles. Abdomen circumference stands out as the strongest factor for estimating body fat percentage.

4.1 Model Validation

The plot, Figure 10 verifies the reliability of the multi-linear regression model by showing the estimated effects of each predictor and their 90% confidence intervals. It confirms that the model's results are consistent and aligns with key assumptions of linear regression.

- Abdomen: The strong positive coefficient and confidence interval far from zero confirm that Abdomen is a significant predictor of body fat percentage. This shows the model correctly identifies meaningful relationships between predictors and the dependent variable.
- Wrist: The negative coefficient and confidence interval that does not cross zero confirm that the model captures the relationship between Wrist size and lower body fat. While the interval for Wrist is slightly wider, it still supports the model's reliability.
- Height: The negative coefficient with a confidence interval just avoiding zero confirms the model captures a small but reliable relationship between Height and body fat percentage.
- Age: The confidence interval for Age includes zero, confirming the model's result that Age has a weaker and statistically insignificant effect on body fat percentage.

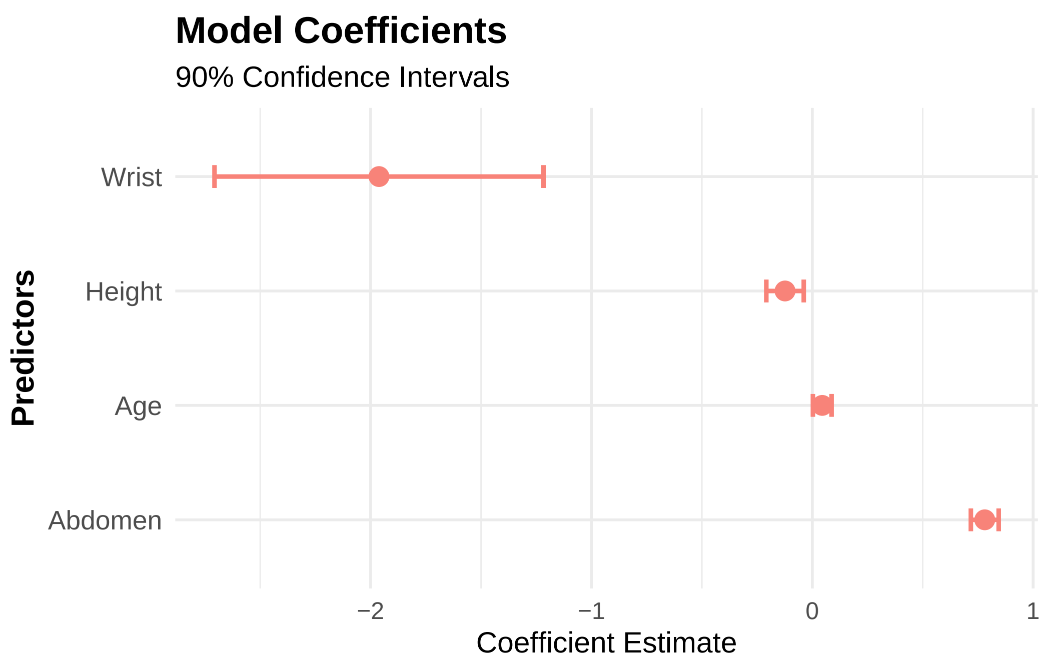

Figure 10: Model Coefficients Confidence Intervals

5 Discussion

5.1 Best Predictor: Abdomen

The regression results confirm that abdomen circumference is the strongest predictor of body fat percentage, with a coefficient of 0.784 and a highly significant p-value (<2e-16). This means that for every 1 cm increase in abdomen size, body fat percentage increases by approximately

0.78%, holding all other variables constant. Abdomen size is directly related to fat accumulation around the central part of the body, which is often linked to overall body fat levels. This finding aligns with the understanding that abdominal fat is a strong indicator of general adiposity and health risks, such as heart disease and metabolic disorders. The importance of abdomen as a predictor highlights its role in practical applications of body fat estimation models, particularly for tracking health outcomes and identifying individuals at risk for fat-related conditions.

5.2 Contributions of Wrist and Height

The results also show meaningful contributions from wrist circumference and height, though their relationships with body fat differ. Wrist circumference has a significant negative coefficient (-1.942, p = 2.38e-05), indicating that individuals with larger wrists tend to have lower body fat percentages. This could reflect differences in body structure, where a larger wrist size often suggests a higher proportion of bone mass compared to fat mass. Height, on the other hand, has a smaller negative coefficient (-0.123, p = 0.0169), showing that taller individuals generally have slightly lower body fat percentages. This relationship may stem from taller individuals having a larger lean body mass proportion relative to their overall size. These results suggest that, while Abdomen plays the main role, Wrist and Height provide important context in understanding body fat distribution. They highlight the need to consider body structure, not just size, when evaluating fat levels.

5.3 Model Fit and Predictive Strength

The model demonstrates strong predictive ability, with an R^2 value of 0.7033 and an adjusted R^2 of 0.698. This means that approximately 70% of the variation in body fat percentage is explained by the four predictors in the model: Age, Height, Abdomen, and Wrist. The residual standard error of 4.33 indicates that the model's predictions are reasonably close to the actual values, with only a small degree of error. While Abdomen is a primary predictor, and Wrist and Height also contribute, Age appears to be a weaker predictor. Its coefficient (0.046) has a p-value of 0.071, indicating it is not statistically significant at the 5% level. This suggests that while body fat percentage might slightly increase with age, the relationship is not strong enough to rely on for predictive purposes. Including Age adds only marginal value to the model, and its role is less important compared to the other predictors. These results validate the model's usefulness in estimating body fat and emphasize its role in applications such as health monitoring, where accessible and reliable measurement tools are needed.

6 Weaknesses and Next Steps

This paper creates a multi-linear regression model to estimate body fat percentage using four predictors: Age, Height, Abdomen, and Wrist circumference. The dataset includes male participants, and the study ensures the model follows linear regression assumptions. The results show how each predictor contributes to body fat percentage, with Abdomen being the most important factor.

One finding is that abdominal size is a strong predictor of body fat percentage. This supports the idea that fat stored in the abdominal area plays a main role in overall body fat levels. Abdomen circumference serves as a useful and accessible measure for monitoring body fat, making it helpful for health-related purposes.

We also learn that structural factors like wrist size are connected to body fat percentage. Larger wrists are linked to lower body fat, possibly because of a higher proportion of bone mass compared to fat mass. This finding shows the need to consider body structure, not just size, when assessing body fat.

6.0.1 Weaknesses

One issue with this study is that the data only includes men, making it difficult to apply the findings to other groups, such as women or people from different backgrounds. Body fat distribution and its predictors vary by gender and other factors, which limits how useful the model is for the general population. The dataset also focuses mostly on people aged 30 to 60, with very few younger or older individuals. As a result, the model may not work well for teenagers, young adults, or seniors, who might have different patterns of body fat.

Another limitation is the number of predictors used. The model includes Age, Height, Abdomen, and Wrist but leaves out others like weight, hip size, and thigh size, which could improve predictions. These variables were excluded during backward selection to simplify the model, but some of them might still be important. Excluding these predictors could make the model less accurate in explaining body fat differences across individuals.

The residual analysis also showed minor issues. While the model works well overall, there are some outliers and slight deviations from normality in the residuals. These could affect how well the model predicts body fat for people with very high or very low values. Additionally, the residual patterns suggest there might be other predictors not included in the model that could explain some of the remaining variability.

6.0.2 Future Steps

Future research should focus on expanding the dataset to include more diverse groups. Adding data for women and people from different backgrounds would make the model more useful for a wider population. Including people of all ages, especially teenagers and seniors, would help the model predict body fat for everyone, not just middle-aged adults. A more balanced dataset would also make it easier to detect patterns that apply to specific groups.

Researchers should also consider adding more predictors to the model. Including variables like weight, hip size, and thigh size could give a fuller picture of body fat distribution and improve the model's accuracy. Looking at how important each variable is could help decide which ones to add while keeping the model manageable.

Finally, using different methods to model the data could address some of the limitations. For example, machine learning techniques or non-linear models could find patterns that linear regression might miss, especially for outliers. Future studies could also use data collected over time to see how body fat changes as people age and whether the same predictors still apply. This would help create a better understanding of body fat and how it changes throughout life.

A Appendix

A.1 Appendix A

Table 4: Raw Data of Body Fat Variables Part 1

V1	V2	V3	V4	V5	V6	V7	V8
Density	Pct.BF	Age	Weight	Height	Neck	Chest	Abdomen
1.0708	12.3	23	154.25	67.75	36.2	93.1	85.2
1.0853	6.1	22	173.25	72.25	38.5	93.6	83
1.0414	25.3	22	154	66.25	34	95.8	87.9
1.0751	10.4	26	184.75	72.25	37.4	101.8	86.4
1.034	28.7	24	184.25	71.25	34.4	97.3	100
1.0502	20.9	24	210.25	74.75	39	104.5	94.4
1.0549	19.2	26	181	69.75	36.4	105.1	90.7
1.0704	12.4	25	176	72.5	37.8	99.6	88.5
1.09	4.1	25	191	74	38.1	100.9	82.5

Table 5: Raw Data of Body Fat Variables Part 2

V9	V10	V11	V12	V13	V14	V15	V16
Waist	Hip	Thigh	Knee	Ankle	Bicep	Forearm	Wrist
33.543307	94.5	59	37.3	21.9	32	27.4	17.1
32.677165	98.7	58.7	37.3	23.4	30.5	28.9	18.2
34.606299	99.2	59.6	38.9	24	28.8	25.2	16.6
34.015748	101.2	60.1	37.3	22.8	32.4	29.4	18.2
39.370079	101.9	63.2	42.2	24	32.2	27.7	17.7
37.165354	107.8	66	42	25.6	35.7	30.6	18.8
35.708661	100.3	58.4	38.3	22.9	31.9	27.8	17.7
34.842520	97.1	60	39.4	23.2	30.5	29	18.8
32.480315	99.9	62.9	38.3	23.8	35.9	31.1	18.2

A.2 Appendix B

References

Auguie, Baptiste. 2017. *gridExtra: Miscellaneous Functions for "Grid" Graphics*.

Diabetes, National Institute of, Digestive, and Kidney Diseases. 2023. "Health Risks of Overweight & Obesity." National Institutes of Health. https://www.niddk.nih.gov/health-information/weight-management/adult-overweight-obesity/health-risks-overweight-obesity.

Fetters, K. Aleisha. 2023. "15 Negative Effects of a Low Body Fat Percentage." *Men's Journal*. https://www.mensjournal.com/health-fitness/15-negative-effects-having-low-body-fat-percentage.

Fox, John, and Sanford Weisberg. 2019. *An R Companion to Applied Regression*. Third. Thousand Oaks CA: Sage. https://www.john-fox.ca/Companion/.

R Core Team. 2024. *R: A Language and Environment for Statistical Computing*. Vienna, Austria: R Foundation for Statistical Computing. https://www.R-project.org/.

Ricard, Mark. n.d. "Bodyfat." Brigham Young University; BYU Human Performance Research Center. https://dasl.datadescription.com/datafile/bodyfat/?_sf_s=body+fat&_sfm_cases=4+59943.

Richardson, Neal, Ian Cook, Nic Crane, Dewey Dunnington, Romain François, Jonathan Keane, Dragoș Moldovan-Grünfeld, Jeroen Ooms, Jacob Wujciak-Jens, and Apache Arrow. 2024. *Arrow: Integration to 'Apache' 'Arrow'*. https://github.com/apache/arrow/.

Schloerke, Barret, Di Cook, Joseph Larmarange, Francois Briatte, Moritz Marbach, Edwin Thoen, Amos Elberg, and Jason Crowley. 2024. *GGally: Extension to 'Ggplot2'*. https://ggobi.github.io/ggally/.

Tinsley, Grant. 2023. "The 10 Best Ways to Measure Your Body Fat Percentage." *Healthline*. https://www.healthline.com/nutrition/ways-to-measure-body-fat.

Venables, W. N., and B. D. Ripley. 2002. *Modern Applied Statistics with s*. Fourth. New York: Springer. https://www.stats.ox.ac.uk/pub/MASS4/.

Wickham, Hadley, Mara Averick, Jennifer Bryan, Winston Chang, Lucy D'Agostino McGowan, Romain François, Garrett Grolemund, et al. 2019. "Welcome to the tidyverse." *Journal of Open Source Software* 4 (43): 1686. https://doi.org/10.21105/joss.01686.

Xie, Yihui. 2024. *Knitr: A General-Purpose Package for Dynamic Report Generation in r*. https://yihui.org/knitr/.

Zhu, Hao. 2024. *kableExtra: Construct Complex Table with 'Kable' and Pipe Syntax*. http://haozhu233.github.io/kableExtra/.