

Title*

Subtitle

Joseph Burks

September 23, 2025

Abstract here.

Introduction

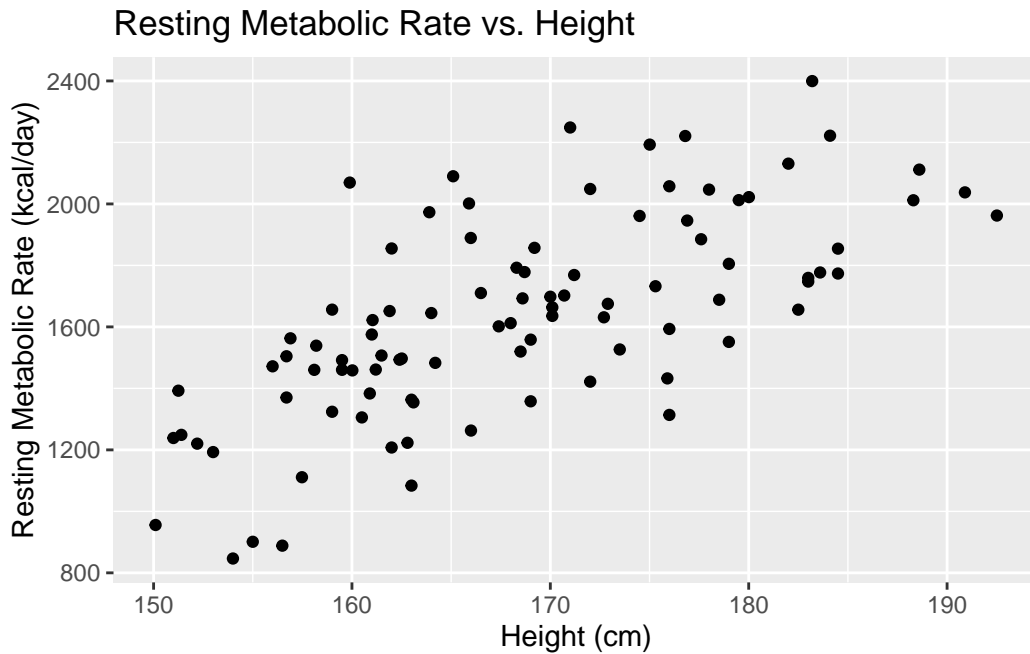
Resting metabolic rate is the amount of energy the body requires to maintain basic life functions while at rest. Another way to think of it, is as the amount of calories one burns through rest. Resting metabolic rate accounts for roughly 60-70% of the calories one burns in a day. While not super important for health diagnostics, knowing one's resting metabolic rate can help tailor nutrition and diet plans to an individual's needs. Calculating resting metabolic rate requires lab testing. Gas analysis is used to calculate resting metabolic rate. By measuring the amount of oxygen one consumes and the amount of carbon dioxide expelled from the body while at rest, one can obtain a decently accurate estimate of resting metabolic rate. There are formulas to estimate the rate. The most widely used is the Miffling-St Jeor Equation: $10(\text{weight}) + 6.25H - 5A + 5$ for men and $10W + 6.26\text{height} - 5A - 161$ for women. While considered the most accurate formula, the equation still has around 26% unknown variance. On top of having different equations for men and women, they have three variables. The question of this report is is height alone a good predictor of resting metabolic rate. This report fits a simple linear regression model, with height as the predictor and resting metabolic rate as the response variable. The fitted model is $Y_i = -2082.267 + 22.075$. This means that there is on average a 22.075 increase in the kilocal per day of the resting metabolic rate for a 1 centimeter increase in height.

Data

The data used in this report comes from the Harvard Dataverse. It was used in a study titled A Gender-agnostic Inclusive Estimation for Resting Metabolic Rate. (Navalta 2025) Height is

*Project repository available at: <https://github.com/peteragao/MATH261A-project-template>.

measured in centimeters and has a mean of 168.399 cm and a standard deviation of 10.18457 cm. Resting metabolic rate is measured in kcal per day with a mean 1635.097 kcal/day and standard deviation 328.4276 kcal/day.



There appears to be a strong positive linear relationship between height and resting metabolic rate. Linear regression will likely be a good model for this data.

Methods

Simple linear regression is the model used in this report. The general model is as followed: $Y_i = \text{Beta}_0 + \text{Beta}_1 x_i + \text{epsilon}$. Y_i are observed responses. In this case the observed resting metabolic rates. Beta_0 is the intercept. In this case it does not contain much meaning since the predictor variable is strictly positive. Beta_1 is the slope of the line of best fit. It represents the average increase in resting metabolic rates for a one unit increase in height. X_i are the observed predictor values. In this case, these are the heights. Epsilon is the error term and represents the variability inherent in all data. There are several assumptions for simple linear regression. One is that the relationship between the response and predictor must be linear. The others relate to the error terms. The errors must be independently and identically distributed with a normal distribution with constant We fit the simple linear regression model using the `lm()` function in R.

Results

```
#| warning: false
#| message: false

lm_fit <- lm(RMR..kcal.day. ~ Height..cm., data = rmr_data)
summary(lm_fit)
```

Call:

```
lm(formula = RMR..kcal.day. ~ Height..cm., data = rmr_data)
```

Residuals:

Min	1Q	Median	3Q	Max
-489.17	-161.49	-9.25	135.03	622.17

Coefficients:

	Estimate	Std. Error	t value	Pr(> t)
(Intercept)	-2082.267	413.468	-5.036	2.36e-06 ***
Height..cm.	22.075	2.451	9.007	2.81e-14 ***

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 240.7 on 92 degrees of freedom

Multiple R-squared: 0.4686, Adjusted R-squared: 0.4628

F-statistic: 81.13 on 1 and 92 DF, p-value: 2.806e-14

The estimated slope parameter is $b_1 = 22.075$. This means that for each one centimeter increase in height the resting metabolic rate increase on average by 22.075

The estimated intercept parameter is $b_0 = -2082.267$. This represents the average resting metabolic rate for some one who has a height of 0cm. Since height is strictly positive, there intercept does not contain much meaning.

This simple linear regression has an R-squared of 0.469. This means the model explains around 46.9% of the variation in the data.

We use a two-sided t-test with null hypothesis $\beta_1 = 0$ and alternative hypothesis $\beta_1 \neq 0$. For this test to be valid, we require that the error terms be independently and normally distributed with mean 0 and constant variance. Figure 1 seems to indicate that there is constant variance and independence, although the extremes start to diverge from what is expected a bit. This

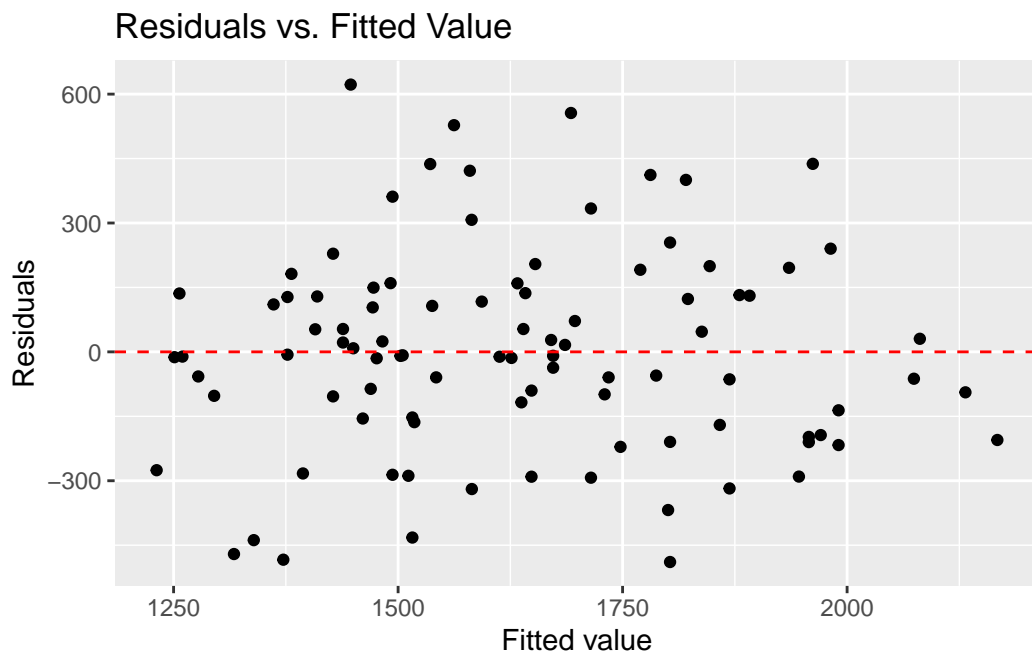


Figure 1: Residuals (y-axis) vs fitted values (x-axis) for a simple linear regression model with resting metabolic rate (kcal/day) as the response and height (cm) as the predictor

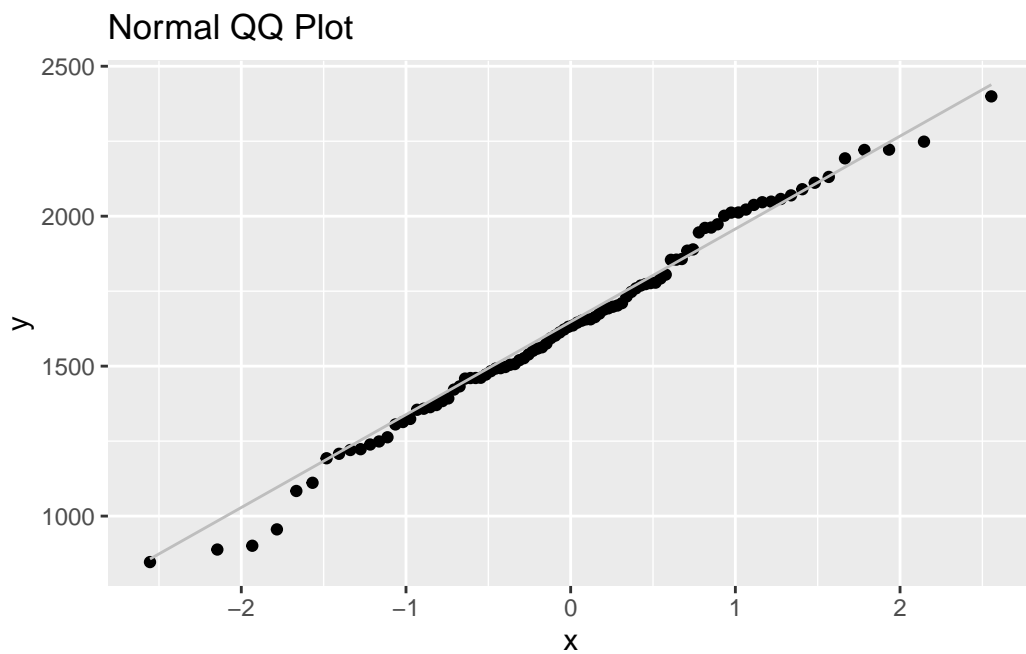


Figure 2: Plot of sample quantiles vs the theoretical normal distribution quantiles

is likely from the lack of data, rather than a violation of assumptions. Figure 2 indicates that the errors are normal, as there is very little deviation from the line.

The assumptions seemed to be satisfied, thus we can continue with the two-side t-test. The associated p-value for β_1 is $2.8057667 \times 10^{-14}$ which is much smaller than the classical significance of $\alpha = 0.05$. Hence we reject the null hypothesis. There is evidence to suggest that $\beta_1 \neq 0$.

Discussion {sec-discussion}

Is height a good predictor for resting metabolic rate. The R-squared offers is a decent size at 0.469. Meaning height is a decent predictor, although probably not good enough to completely move away from the more common formulas for resting metabolic rate.

The model is a good fit and all the necessary assumptions for simple linear regression are satisfied. Since resting metabolic rate is not a good metric for health, not everyone may want or need a super accurate measurement. In this case the simplicity of this model may be advantageous to some. However, the Miffling-St Jeor Equation is not super complex, so it is likely better in most situations.

Height may not be the best predictor. Height can often be a stand in for weight, as taller people on average will weigh more than shorter people. One question with this model, is is height influencing resting metabolic rate in a significant way or is it really just weight.

In the future we could use weight as a predictor to see if this is the case, or we could calculate the BMI to account for the expected increase in weight for an increase in height.

Fitting a multiple linear regression using more body composition variables as predictors is a good next step for this research. We could explore if it is really necessary to have different equations for male and female as the Miffling-St Jeor Equation does.

References

Navalta, James. 2025. "Data for: A Gender-agnostic Inclusive Estimation for Resting Metabolic Rate." Harvard Dataverse. <https://doi.org/10.7910/DVN/MIOESN>.