**Introduction**

*Types of data*

Body fat (%) – Response variable (continues quantitative)

Age (years) - Explanatory variable (discrete quantitative)

Chest circumference (cm) - Explanatory variable (continues quantitative)

Density (g/cm3) - Explanatory variable (continues quantitative)

Knee circumference (cm) - Explanatory variable (continues quantitative)

Weight (lbs) - Explanatory variable (continues quantitative)

*Type of study*

Observation study

*Sample size*

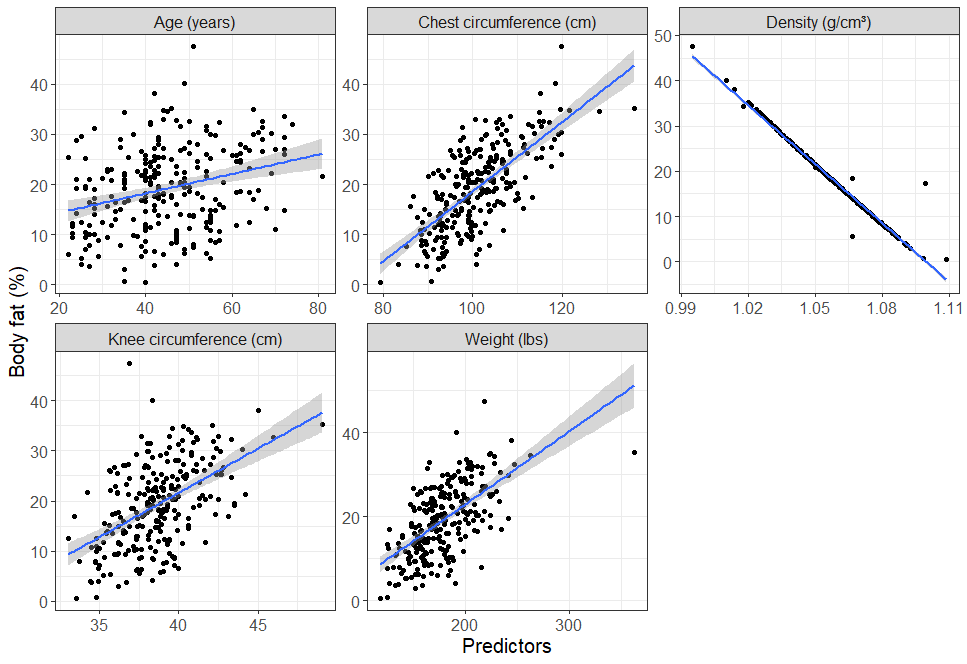
Sample size is 252 for both response and explanatory variables.

**Descriptive Statistics**

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
|  | Measurement | | | | | |
| Variable | Age (years) | Body fat (%) | Chest circumference (cm) | Density (g/cm³) | Knee circumference (cm) | Weight (lbs) |
| Sample Size | 252 | 252 | 252 | 252 | 252 | 252 |
| Mean | 44.88 | 19.15 | 100.82 | 1.06 | 38.59 | 178.92 |
| Standard deviation | 12.602 | 8.364 | 8.430 | 0.019 | 2.412 | 29.389 |
| Median | 43.00 | 19.20 | 99.65 | 1.05 | 38.50 | 176.50 |
| 1st quartile | 35.75 | 12.47 | 94.35 | 1.04 | 36.98 | 159.00 |
| 3rd quartile | 54.00 | 25.30 | 105.38 | 1.07 | 39.92 | 197.00 |
| Min | 22.0 | 0.5 | 79.3 | 1.0 | 33.0 | 118.5 |
| Max | 81.00 | 47.50 | 136.20 | 1.11 | 49.10 | 363.15 |

*Normality*

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| Variable | Age (years) | Body fat (%) | Chest circumference (cm) | Density (g/cm³) | Knee circumference (cm) | Weight (lbs) |
| Sample Size | 252 | 252 | 252 | 252 | 252 | 252 |
| Mean | 44.88 | 19.15 | 100.82 | 1.06 | 38.59 | 178.92 |
| Median | 43.00 | 19.20 | 99.65 | 1.05 | 38.50 | 176.50 |
| Skewness | 0.280 | 0.148 | 0.673 | -0.020 | 0.511 | 1.191 |
| Normally Distributed | No | Yes | No | Yes | No | No |
| p-value | 0.0010 | 0.1410 | 0.0001 | 0.6571 | 0.0033 | 0.0000 |

**Scatter Plot**

Correlation coefficients

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
|  | Age | Chest | Density | Knee | Weight |
| Pearson Correlation Coefficient | 0.2915 | 0.7024 | -0.9876 | 0.5084 | 0.6123 |
| P Value | 0 | 0 | 0 | 0 | 0 |

* There’s a to be a weak positive linear relationship between body fat and age.
* There’s a strong positive linear relationship between body fat and chest circumference.
* There’s a strong negative linear relationship between body fat and density.
* There’s a medium positive linear relationship between body fat and knee circumference.
* There’s a medium positive linear relationship between body fat and weight.

**Correlation Matrix**

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
|  | Age | Chest | Density | Knee | Weight |
| Age | 1 | . | . | . | . |
| Chest | 0.1764 | 1 | . | . | . |
| Density | -0.2776 | -0.6826 | 1 | . | . |
| Knee | 0.0175 | 0.7195 | -0.495 | 1 | . |
| Weight | -0.0127 | 0.8942 | -0.5941 | 0.8532 | 1 |

A multi-collinearity is present between following variables:

There’s a strong positive correlation between Weight and Chest. r>0.5

There’s a strong positive correlation between Knee and Chest. r>0.5

There’s a strong negative correlation between Knee and Chest. r>0.5

**Model selection**

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| R squared | Variables | Regression Coefficients | Standardised  Regression Coefficients | p values | VIF |
| 97.69% | Intercept | 453.2336 |  |  |  |
|  | Age | 0.0154 | 0.0232 | 0.035 | 1.2989 |
|  | Chest Circumference | 0.0352 | 0.0355 | 0.1668 | 7.1104 |
|  | Density | -415.572 | -0.9456 | <0.0005 | 1.9705 |
|  | Knee Circumference | -0.024 | -0.0069 | 0.7148 | 3.8947 |
|  | Weight | 0.0071 | 0.025 | 0.4271 | 10.7358 |
| 97.69% | Intercept | 451.7504 |  |  |  |
|  | Chest Circumference | 0.0132 | 0.0199 | 0.0504 | 1.1145 |
|  | Density | 0.0506 | 0.051 | 0.0022 | 2.9526 |
|  | Knee Circumference | -415.628 | -0.9457 | <0.0005 | 1.9702 |
|  | Weight | 0.0111 | 0.0032 | 0.8191 | 2.132 |
| 97.70% | Intercept | 452.026 |  |  |  |
|  | Chest Circumference | 0.013 | 0.0195 | 0.0512 | 1.0839 |
|  | Density | 0.0529 | 0.0533 | <0.0005 | 1.8731 |
|  | Knee Circumference | -415.688 | -0.9458 | <0.0005 | 1.9664 |

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| Labels | Adjusted R² | Regression coefficient | 95% CI (Lower) | 95% CI (Upper) | Standardised Regression coefficient | p-value |
| Overall | 97.7% |  |  |  |  | <0.0005 |
| Intercept |  | 452.026 |  |  |  |  |
| Age | Age | 0.013 | -0.0001 | 0.026 | 0.0195 | 0.0512 |
| Chest | Chest Circumference | 0.0529 | 0.0273 | 0.0785 | 0.0533 | <0.0005 |
| Density | Density | -415.688 | -427.304 | -404.072 | -0.9458 | <0.0005 |

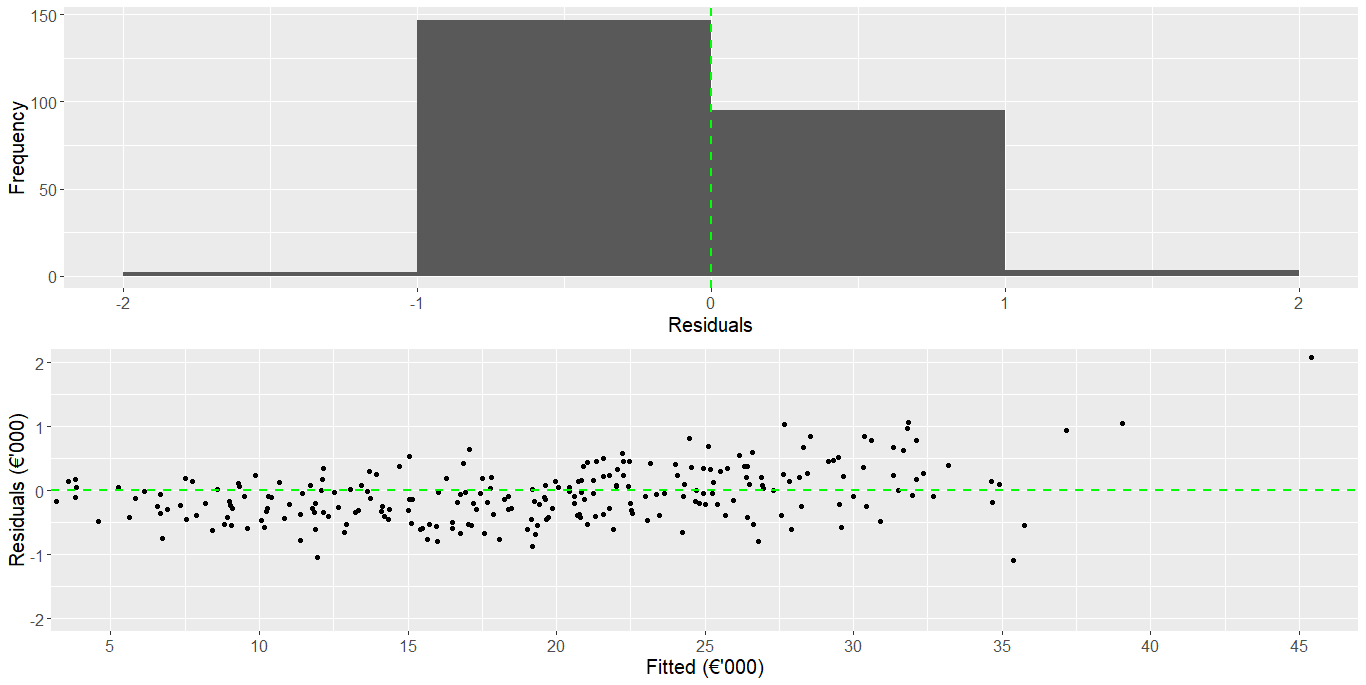
**The variation inflation factor (VIF)**

A value of 1 means that the predictor is not correlated with other variables

Values of more than 4 or 5 are sometimes regarded as being moderate to high, with values of 10 or more being regarded as very high.

**Final Model**

**Residuals**



|  |  |  |
| --- | --- | --- |
| Statistics | Outcome | p\_value |
| Normally Distributed | No | 0 |
| Zero mean | Yes | 1 |
| Random | Yes | 0.6944 |

Distribution of the residual is normally distributed.

Difference exists between measurements about zero.

There is auto-correlation among the residuals.