Oxygen intake rates - a measure of aerobic fitness

180029941

06/11/2018

..

Executive Summary

Introduction

The present report aims to build a model that predicts Oxygen intake rates (a measure of aerobic fitness) supported on a series of measurements. The fitness dataset from Rawlings (1998) contains measurements of the following seven variables obtained from 31 men:

- Age: Age in years;
- Weight: Weight in kg;
- Oxygen: Oxygen intake rate, ml per kg body weight per minute;
- RunTime: time to run 1.5 miles in minutes;
- RestPulse: heart rate while resting;
- RunPulse: heart rate at end of run;
- MaxPulse: maximum heart rate recorded while running;

From the data set fitness.csv a linear model (predicting Oxygen) will be developed. The bootstrapping function used to provide confidence intervals came from an original function provided by Donovan (2018), which was improved at a later stage.

The current report uses R 3.5.1 software (R Core Team, 2018). It was produced a linear model which was fitted in each analysis and the bootstrap used to generate confidence intervals for each of the covariates of interest. We aim to exclude variables that present, essentially, the same information about response avoiding this way collinearity.

The reasonability of the assumptions on which the model is based were assessed:

1. Linearity
2. Homoscedasticity
3. Independence
4. Normality

Bootstrap methods were used in order draw conclusion to hypothesis tests in regards to the significance of the relationships between the response and the parameter estimates can be drawn using bootstrap methods. If the confidence interval contains zero, one fails to reject the null hypothesis, and if it does not contain zero, one can reject the null hypothesis.

Findings

Based on our fitness data set, we are going to implement a Linear Regression Model which explanatory variables (e.g. Age, Weight, RunTime, RestPulse, RunPulse and MaxPulse) will help explain or predict the behaviour of the response variable (Oxygen). The model is specified as follows:

..

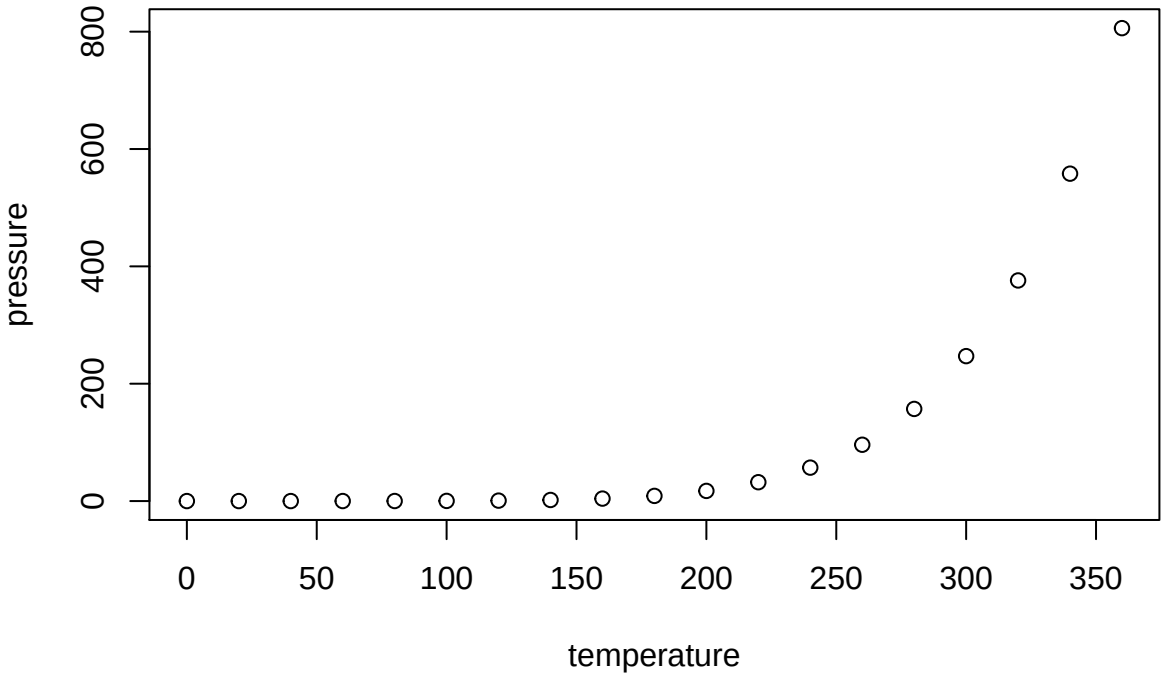
```
##   Age Weight Oxygen RunTime RestPulse RunPulse MaxPulse
## 1  44  89.47 44.609   11.37        62      178      182
## 2  40  75.07 45.313   10.07        62      185      185
## 3  44  85.84 54.297    8.65        45      156      168
## 4  42  68.15 59.571    8.17        40      166      172
## 5  38  89.02 49.874    9.22        55      178      180
## 6  47  77.45 44.811   11.63        58      176      176
```

```
##
## Call:
## lm(formula = Oxygen ~ Age + Weight + RunTime + RestPulse + RunPulse +
##     MaxPulse, data = fitness)
##
## Coefficients:
## (Intercept)          Age       Weight      RunTime    RestPulse
##   102.93448     -0.22697     -0.07418     -2.62865     -0.02153
##    RunPulse     MaxPulse
##    -0.36963      0.30322
```

..

Including Plots

You can also embed plots, for example:



Note that the `echo = FALSE` parameter was added to the code chunk to prevent printing of the R code that generated the plot.