

Introduction

Maintaining a healthy weight can help improving the overall lifestyle and reducing the risk of hearth diseases, high blood pressure, stroke, and diabetes, and hence help maintaining good health overall.

Through this case study, we aims at studying association between weights and eating habits as well as physical conditions. We will study a dataset relating to 2111 individuals ages 14 to 61 from Mexico, Peru and Columbia. The data includes atributes such as gender, weight, height, attributes related to eating habits, and attributes related to physical condition. We will use this data to answer the following question

- Whether there is an association between frequency of exercise and weight? Does weight has a significant relationship with weight, and how the relationship between exercise and weight is when controlled for height.
- What is the relationship between weight and gender+ the use of technological devices.

Data Description

The data is downloaded from UCLA machine learning database <https://archive-beta.ics.uci.edu/dataset/544/estimation+of+obesity+levels+based+on+eating+habits+and+physical+condition>. The data is available on our SAS server as `mydata.Obesity`. The variables of interest for our case study are

Variable	Description
weight	weight of the individual
Gender	gender (female/ male)
Age	Age of the individual
height	Height of the individual
Family	has a family member suffered from overweight? (yes/no)
FAVC	Frequent consumption of high caloric food? (yes/no)
FCVC	Frequency consumption of vegetables (1=never, 2=sometimes, 3=always)
NCP	Number of main meals (1,2, 3,4)
CAEC	Do you eat food between meals? (Always/Frequently/Sometimes/No)
SMOKE	Do you smoke? (yes/no)
CH20	How much do you drink daily? (1=less than 1 litre, 2=1 to 2 litres, 3=more than 2 litres)
SCC	Calory sonsumption monitoring? (yes/no)
FAF	Physical activity frequency per week (0=none, 1=1 to 2 days, 2= 2 to 4 days, 3= 4 to 5 days)
TUE	How much time do you use technological devices per day (0=0 to 2 hours, 1=3 to 5 hours, 2= 5 or more hours)
CALC	Consumption of alcohol (never, sometimes, frequently, always)
MTRANS	which transportation do you use? (Automobile, Motorbike, bike, public transport, walking)
NObeyesdad	Weight class (Insufficient_weight, normal_weight, overweight_level_I, overweight_level_II, Obesity_Type_I, Obesity_Type_II, Obesity_Type_III)

Assignment Tasks

Q1 (55 marks)

- (a) **(20 marks)** Carry out a one-way analysis of variance (ANOVA) relating `weight` to frequency of exercise `FAF`. Use contrasts to test at least one a-priori hypothesis of your choice. Examine and comment on residuals. Also carry out appropriate post-hoc comparisons and discuss your results.
- (b) **(10 marks)** If the assumptions for ANOVA is not satisfied, use a non-parametric method to validate the results in question (a).
- (c) **(25 marks)** Use SAS to perform a one-way ANCOVA relating `weight` to `FAF` and `height` with `height` as a covariate, including appropriate post-hoc comparisons:
- Confirm that there is a linear relationship between the response variable and the covariate (a scatterplot and correlation coefficient plus a comment will suffice);
 - Check the two additional ANCOVA considerations (report and comments only on the parts of the output most directly relevant to condition checking):
 - Independence of the covariate and the treatment effect (perform a one-way ANOVA test);
 - Equality of slopes (add and check significance of the interaction term);
 - Report and briefly discuss your results.

Technical note: Make sure you obtain and examine Type III Sum of Squares (ss3). Also obtain estimates of 'least squares means' (lsmeans) which are means by treatment adjusted for the covariate.

Q2 (30 marks)

Perform and analyse a factorial ANOVA model to determine whether there is statistically significant difference in `weight` by `Gender` and the use of technological devices `TUE`. Carry out to test whether there is evidence of interaction between `Gender` and `TUE`. Examine and comment on residuals. Carry out appropriate follow-up analysis and discuss your results. Test the following a-priori hypotheses:

- mean `weight` for male with `TUE`=1 is the same as mean `weight` for female with `TUE`=1.

State and test two other a-priori hypotheses of your choice.

Q3 (15 marks)

Write a summary of your findings from Questions 1–3. Keep the technical details of the analyses that led you to these conclusions to the absolute minimum. Rather,

focus on practical significance and present your findings in non-specialist terms. One to two paragraphs (up to a page) will be sufficient.