

STATS 201 Data Analysis Assignment 3, NEFU, 2021

Instructions concerning this assignment:

We are providing you an R Markdown document called **STATS201_2021_NEFU_A3.Rmd** which will have some answers already filled in. You will need to fill in and complete the rest of the document. The data files you will be using for the assignment are described in the questions and are available online. Make sure you put these datasets in the same place you put the R markdown document because it is going to look for them there. The first change you need to make to the markdown document is put your name and ID number at the top.

Question 1. [13 Marks]

A musical college wished to study the effectiveness of various practice methods for its trombonists. To this end 30 trombonists were randomly assigned to one of 5 practice methods. Each student's ability with the trombone was initially assessed. To assess improvement, after a set period of time using the assigned practice method only, they reassessed each student and recorded the difference between the student's final and initial scores.

The dataset is in the file `trombone.csv` and includes these variables:

<code>diff</code>	The difference in trombone assessment scores (after – before)
<code>method</code>	The practice method assigned to the student: NP - no practice control, MP - mental practice only, MPS - mental practice with simulated slide movement, PP - physical practice only, CP - combination of mental and physical practice

We are interested in whether any of the methods show significant improvement? If so, which method shows the most significant improvement and by how much?

- Comment on what the plot of the data and summary statistics reveal.
- Fit an appropriate model to the data and check the assumptions.
- Write a **Methods and Assumption Checks** section.
- Write an **Executive Summary**.

Question 2. [19 Marks]

A researcher was interested in how much males tend to be taller than females and whether this height difference depended on the ethnicity of people. They randomly sampled data from a large cross-sectional study called `xs.nz` that had collected various information on people in New Zealand. This data can be found in the file "**height.txt**", with variables:

<code>height</code>	the person's height (in metres).
<code>sex</code>	a factor with levels M for male and F for female.
<code>ethnicity</code>	a factor with the levels European, Maori, Polynesian and Other.

- Outline the questions of interest or the goal of the analysis.
- Produce and comment on appropriate interaction plot(s).
- Fit an appropriate two-way ANOVA. Determine whether the model can be simplified and determine an appropriate final model.
- Write a **Methods and Assumption Checks** section.
- Write an **Executive Summary**.