Title of your report

Leave author name blank

Your submission date

Preliminaries

This is an R Markdown document. For more details on using R Markdown see http://rmarkdown.rstudio.com.

This document serves a multiple purpose:

- It provides you with some initial code which is needed in order to read the data which are required to complete this assignment;
- It contains instructions on what is expected from your assignment;
- You produce your actual assignment by adding text and code to this document, and eventually knitting it when finished.

Your finally submitted document will be a report, written in R Markdown, which may contain text, equations, code, and images, and which should be **1500** words long[1]. It is admitted that "word count" may be an unclearly defined concept for such a heterogeneous document. We therefore recommend to use the `wordcountaddin` to establish your wordcount (see code at the end of this document). To avoid the document becoming excessively long, we think of an image as representing 50 words, and of R code contributing another 20 (or so) words per chunk. So, bottom-line, assuming that you have a reasonable number of such elements included into your report, then a word count of **1000** according to the `wordcountaddin` will be appropriate for this assignment[2]. You *don't* need to evidence how this would add up to 1500 when accounting for code and/or figures.

The final document needs to be a knitted PDF version of this Markdown document, with the `Preliminaries` section removed, and of course your content added. This is non-negotiable. MS Office documents will not be accepted.

Your report needs to have four sections: Introduction, Methods, Analysis and Discussion of results. The framework of these is already set up below, with some indications of envisaged content. Within and between these four sections, you have some flexibility to move content around (in particular; these indications should not be understood as prescribing or limiting the scope of your work), but you should provide an overall sound story which addresses the overall question of this assignment.

R Packages The following are some R packages that you are likely to need for your analysis. So, it is recommended to load these now.

```
require(lme4)
require(lmerTest)
require(ggplot2)
require(sjPlot)
```

[1] In the summative assignment this would be 3000
[2] In the summative assignment this would be 2000

Data We now introduce the data to be used in this assignment. You can modify this part and move it into one of the subsections of your assignment as you deem appropriate.

We consider data resulting from a 3-level cluster randomized educational trial. That is, each of 20 schools were randomly designated as either a control or to receive and educational intervention. Test scores were calculated for pupils from two classes in each school both before and after the time at which the intervention took place. The data can be read into a data frame `CRT` as follows:

```
CRT <- read.csv("https://andygolightly.github.io/teaching/MATH43515/CRT.csv", header=TRUE)
head(CRT)
```

```
##   Pupil School Posttest Intervention Pretest FSM class
## 1     1      1       16            1       1   0     1
## 2     2      1       13            1       4   1     1
## 3     3      1       18            1       5   1     1
## 4     4      1       14            1       4   1     1
## 5     5      1       25            1       5   1     1
## 6     6      1       13            1       2   0     1
```

```
dim(CRT)
```

```
## [1] 260   7
```

We see that the data set possesses 260 rows (=pupils) and 7 columns. These columns represent the following variables:

- `Pupil`: anonymized student ID;
- `School`: anonymized school ID $\{1, \ldots, 20\}$;
- `Class`: anonymized class ID $\{1, 2\}$;
- `Intervention`: Intervention indicator (0=control; 1=treatment);
- `FSM`: pupil eligible for free school meal (0=not eligible; 1 =eligible)
- `Pretest`: A pre-test score for each pupil;
- `Posttest`: A post-test score for each pupil (**response**).

Question of interest Does the educational 'Intervention' have a significant impact on student attainment (post-test)?

Here your actual report will start. Remove all prior parts and the presentation section below before submitting.

Introduction (20 marks)

Give some brief background on randomized controlled trials and cluster randomized trials, and their relevance for assessing the effectiveness of an educational intervention. This subsection could also introduce the data, and give some simple illustrative exploratory analysis, without fitting any models.

Methods (20 marks)

Explain what multilevel models do and why they are relevant for the modelling of randomized controlled trials. Explain how the 3-level cluster randomized trial discussed here fits into this framework. This subsection could also give details on how to decompose variance (ICC etc) and brief details of the methodology used to test for the absence or presence of variance components.

Analysis (35 marks)

This section should contain the actual modelling and data analysis. You are welcome to fit several potential models, and apply suitable tests or diagnostic tools to compare them. Possible questions to address are:

1. What are the intra-class-correlations (ICCs) and variance partition coefficients (VPCs)?
2. What is the intervention effect and its confidence interval estimate; is the intervention effect significant?
3. How does the intervention effect change by including pre-test in the model? What role is played by FSM?

Discussion of results (15 marks)

In this section you will discuss and summarize the results from your analysis in the light of the original problem: do these data give evidence that the intervention works? You can also discuss other aspects of your analysis e.g. limitations of the approach taken, and interpret the final fitted model.

Presentation (10 marks)

Marks for correct mathematical typesetting (where required), writing style, presentation and clarity of figures etc.

Remove this section from your final report.

Word count

```
#install.packages("devtools")
#devtools::install_github("benmarwick/wordcountaddin", type = "source", dependencies = TRUE)
require(wordcountaddin)
```

```
## Loading required package: wordcountaddin
```

```
word_count()
```

```
## For information on available language packages for 'koRpus', run
##
##   available.koRpus.lang()
##
## and see ?install.koRpus.lang()
```

```
##
## Attaching package: 'koRpus'

## The following object is masked from 'package:wordcountaddin':
##
##     readability

## [1] 852
```

```
text_stats()
```

Method	koRpus	stringi
Word count	852	822
Character count	5415	5401
Sentence count	71	Not available
Reading time	4.3 minutes	4.1 minutes