Univariate Statistics and Methodology in R  
Coursework report

Department of Psychology, \

Academic year 2020-2021

* [Key Dates](https://uoepsy.github.io/usmr/report/usmr_2021_assignment.html#key-dates)
* [What you need to submit](https://uoepsy.github.io/usmr/report/usmr_2021_assignment.html#what-you-need-to-submit)
* [A template .Rmd file is provided for you](https://uoepsy.github.io/usmr/report/usmr_2021_assignment.html#a-template-.rmd-file-is-provided-for-you)
* [How to approach the questions](https://uoepsy.github.io/usmr/report/usmr_2021_assignment.html#how-to-approach-the-questions)
  + [Helpful tip:](https://uoepsy.github.io/usmr/report/usmr_2021_assignment.html#helpful-tip)
* [Report Formatting](https://uoepsy.github.io/usmr/report/usmr_2021_assignment.html#report-formatting)
* [Grading](https://uoepsy.github.io/usmr/report/usmr_2021_assignment.html#grading)
  + [Additionally:](https://uoepsy.github.io/usmr/report/usmr_2021_assignment.html#additionally)
* [Queries concerning the task](https://uoepsy.github.io/usmr/report/usmr_2021_assignment.html#queries-concerning-the-task)
* [Submitting your files](https://uoepsy.github.io/usmr/report/usmr_2021_assignment.html#submitting-your-files)
  + [Filenames](https://uoepsy.github.io/usmr/report/usmr_2021_assignment.html#filenames)
  + [Where to submit](https://uoepsy.github.io/usmr/report/usmr_2021_assignment.html#where-to-submit)
* [COURSEWORK TASK](https://uoepsy.github.io/usmr/report/usmr_2021_assignment.html#coursework-task)
  + [Questions](https://uoepsy.github.io/usmr/report/usmr_2021_assignment.html#questions)

Your task is to describe and analyse some data in order to provide answers to a set of research questions. Analyses will draw on the methodologies we have discussed in lectures and labs.

Key Dates

**Coursework set**: 12 noon, Wednesday 18th November 2020  
**Coursework due**: 12 noon, Thursday **17th December** 2020

What you need to submit

You are required to submit 2 documents.

1. Your final compiled report, detailing your analyses, results, interpretation and conclusions. This can be **.docx** (or equivalent), **.pdf**, or **.html**.  
   [https://rstudio.ppls.ed.ac.uk](https://rstudio.ppls.ed.ac.uk/) now supports knitting to **.pdf** and word (**.docx**); choose by clicking the drop-down arrow next to “Knit”.  
   Graphical user interface, text, application

   Description automatically generated
2. A **.Rmd** document which reproduces the results you give in the report.

A template .Rmd file is provided for you

We have provided a template **.Rmd** file for you to use, which can be downloaded from Learn or from <http://uoepsy.github.io/usmr/report/usmr_2021_assignment_template.Rmd>.  
In the template, you will find empty code chunks for you to add R code to. However, you should feel free to create more if you wish: This is just a template, and is mainly designed to make your life easier.  
The template sets echo = FALSE for all code-chunks. This means that your R code (but not the output) is hidden in the compiled document.

**If you use RStudio on the PPLS Server, you will need to upload the template .Rmd file to your space on the server in order to use it.**

How to approach the questions

For each of the questions below (repeated in the template), the *compiled* report (final .pdf, .html, or word document) is expected to include:

1. **Clear written details of the analysis conducted** in order to answer the question, including transparency with regards to decisions made about the data prior to and during analysis.
2. **Results**, in appropriate detail (for instance, a test statistic, standard error and p-value, not just one of these).
3. **Presentation** of results where appropriate (in the form of tables or plots).
4. **Interpretation** (in the form of a written paragraph(s) referencing relevant parts of your results) leading to a conclusion regarding the question.

The code you write in your submitted .Rmd file should successfully undertake the analysis described in A), which returns B). You should also include the code to produce C).

Helpful tip:

As the compiled document will not contain visible R code, a large part of the challenge comes in clearly describing all aspects of the analysis procedure.  
A reader of your compiled document should be able to more or less replicate your analyses **without** referring to your R code.

Report Formatting

Beyond assessing your ability to create reproducible results, the focus of this assignment is not on the formatting of RMarkdown files; it is on implementing and interpreting analyses in R.

This means that you are more than welcome, for instance, to knit to word, and then to subsequently make edits to your text and formatting in the word document itself. *The important thing to remember is that the data analysis and modelling results should match those produced in your RMarkdown file.*

If you do wish to do your formatting in RMarkdown, then we suggest the following readings for help:

* [RMarkdown CheatSheet](https://rstudio.com/wp-content/uploads/2015/02/rmarkdown-cheatsheet.pdf)
* [Writing Math in Rmd](https://rpruim.github.io/s341/S19/from-class/MathinRmd.html#:~:text=Math%20inside%20RMarkdown,10n%3D1n2.)
* For linear regression tables to .html and .pdf, the [**stargazer** package](https://cran.r-project.org/web/packages/stargazer/vignettes/stargazer.pdf)
* For linear regression tables to .html, the [tab\_model() from the](https://strengejacke.github.io/sjPlot/articles/tab_model_estimates.html)**[sjPlot](https://strengejacke.github.io/sjPlot/articles/tab_model_estimates.html)**[package](https://strengejacke.github.io/sjPlot/articles/tab_model_estimates.html)

Feel free also to post formatting questions on the discussion forum on Learn.

Grading

Marks will be awarded for providing evidence of the ability to understand and execute appropriate statistical methods to answer each of the questions, and for clear presentation of results and conclusions.

Additionally:

* Ten points will be deducted from your final grade if your marker cannot determine how the results in your report were generated in your .Rmd document (for example, if the .Rmd document produces errors during knitting, or produces values different from those reported). This means that an 82 out of 100 becomes 72 out of 100.
* **Up to** ten points will be added for good use of R and RMarkdown (for example, where code is elegant, or when inline constructs are used to report results in the text). For example, a 72 might be raised up to 82.

Queries concerning the task

This document contains a basic overview of the task and of how to submit it. If you have any questions concerning the coursework report, we ask that you post them on the designated section of the on-line discussion forum on Learn. If you have a question, it is likely your classmates may have the same question. Before posting a question, please check the on-line board in case it has already been answered.

Submitting your files

Before submitting, we strongly advise you to check that your code runs. The easiest way to check this is by checking that your .Rmd compiles (i.e., can you knit your Rmarkdown document into .html/.pdf/.docx without error?)

**If you use RStudio on the PPLS Server, you will need to export the file to your computer in order to upload it to Turnitin.**

Filenames

For both files which you submit, the filename should be your Exam Number with the appropriate extension, and nothing else.  
For example, a student with exam number **B047847** would submit two files:

* **B047847.Rmd**
* one of: **B047847.docx** / **B047847.pdf** / **B047847.html**

Where to submit

Go to the Assessments page on Learn, and look for “Assessment Submission”. There you will find an own-work declaration which requires marking as reviewed, before two submission boxes will be visible (one for each file).

COURSEWORK TASK

**Sleep quality in academia**  
The exponential increase of electronic devices involved in every day life is often claimed to have a detrimental influence on sleep quality. Due to advances in technology, it is now the norm for many of these devices to log user engagement, and there are specific products aimed entirely at logging daily activities (for instance, smart-watches).

In the present study, data was collected from 150 participants from the University of Edinburgh community (students, staff, and in some cases friends and relatives). Each participant was given a smart-watch for a period of 24 hours and was asked to keep it on their wrist at all times until returning it the following day. At the end of this period, researchers downloaded data from the device capturing two measures: sleep duration (hours) and sleep quality (score 0-100, a metric taking into account heart rate, movement, and respiration rate).  
Additionally, participants filled in a short questionnaire. Details of the data collected can be found in the table below.

| Data Dictionary | |
| --- | --- |
|  | |
| variable | description |
| Study | |
| id | Unique participant identifier |
| Smart Watch | |
| sleep\_qual | Sleep Quality Score (0-100; metric taking into account heart rate, movement, and respiration rate, collected via smartwatch device) |
| sleep\_dur | Sleep Duration based on movement and heart rate, collected via smartwatch device (hours) |
| Questionnaire | |
| age | Age (years) |
| exercise | Did you partake in >20 minutes rigorous (or >40 minutes moderate) exercise during the day before sleep was recorded? (No = 0, Yes = 1) |
| role | Are you currently: Undergraduate Student (UG); Postgraduate Student (PG); Staff (S); Other (O) |
| hrs\_caff | Approximately how many hours before sleep did you last consume caffeine, and what was the drink? |
| screentime | What was your phone usage in the last 60 minutes before sleep (determined through smartphone screen time usage data & time of sleep according to smartwatch) (minutes) |
| light | For the last hour before going to sleep, was the light in your room on or off? (Off = 0, On = 1) |
| sr\_qual | How would you rate your overall sleep last night? ("below average", "above average") |

Questions

1. Assess the data for any unusual values (given the context of the data description). Report any decisions and actions taken regarding any such values, and use R to provide some descriptive statistics for the dataset.
2. Is self-reported sleep quality independent of academic role? Perform an appropriate statistical test, and report and interpret the results in text.
3. Do smart-watch sleep quality scores relate to individuals’ self-reports of perceived sleep quality? Does this relationship differ depending on academic role?
4. Is sleep quality (as measured by smart-watch) influenced by the duration of screen usage in the hour prior to sleep?
5. Is sleep quality (as measured by smart-watch) influenced by the duration of screen usage in the hour prior to sleep **after accounting for variation explained by sleep duration, age, and exercise**?
6. Is the amount to which screen usage influences sleep quality different depending on whether screens are used in a light vs. a dark environment?
7. What factors increase the likelihood of reporting above-average sleep quality? Here you should perform a more exploratory analysis using self-reported sleep quality as the outcome (dependent) variable.
   * Extra:  
     Can you plot the predicted probabilities of an individual reporting above average sleep quality for 20 year old undergraduates who range from 0 to 12 hours of sleep?