This document contains settings that help exporting your analysis to a Word document. The expectation is that you have already gained familiarity with editing R Markdown documents from the [R for Data Science](https://r4ds.had.co.nz/r-markdown.html) book, which was on the module reading list, particularly Chapters 27, 29 and 30. It assumes, therefore, that you are familiar with the purpose of yaml headers (i.e. the settings between the two sets of - - - marks at the top of an .Rmd document). There are numerous settings that can be included in the yaml header, but for our purposes specifying the output format (Word document) is the only essential one.

Options relating to how code chunks (i.e. the commands included between two sets of ``` marks throughout) are rendered when written (i.e. “knitted”) to the output document format can be set both in the header of the individual code chunks, or as a global setting that applies to all chunks. The first code chunk below (in the .Rmd version) is a global setting chunk that applies some useful setting for the purpose of printing outputs only to Word documents.

To learn more about R Markdown documents, these two resources are extremely useful:

* Xie, Allaire, and Grolemund (2021) [R Markdown: The Definitive Guide](https://bookdown.org/yihui/rmarkdown/)
* Xie, Dervieux, and Riederer (2022) [R Markdown Cookbook](https://bookdown.org/yihui/rmarkdown-cookbook/)

These resources go far beyond what you need to know at this stage, but they can be helpful for searching particular settings/issues that come up in your own coding.

Below is a recommended structure and some example code for tabulations that seems to work well with Word documents.

# Your title

## Introduction

## Literature review (or a better title for this section)

### Any subtitles

## Data and methods

Somewhere in this section or early on in the Results section you may want to show a summary of your variables of interest. For example, a table such as this (using the descr() function from the summarytools package):

|  | Mean | Std.Dev | Median | Min | Max | N.Valid |
| --- | --- | --- | --- | --- | --- | --- |
| agea | 51.13 | 12.86 | 53.00 | 25.00 | 80.00 | 42079.00 |
| cntry | 4.65 | 2.25 | 4.00 | 1.00 | 8.00 | 42079.00 |
| eduyrs25 | 12.38 | 4.32 | 12.00 | 0.00 | 25.00 | 41626.00 |
| facntr | 0.96 | 0.19 | 1.00 | 0.00 | 1.00 | 42079.00 |
| female | 0.54 | 0.50 | 1.00 | 0.00 | 1.00 | 42079.00 |
| mocntr | 0.97 | 0.18 | 1.00 | 0.00 | 1.00 | 42079.00 |
| paredu\_a\_high | 0.33 | 0.47 | 0.00 | 0.00 | 1.00 | 39765.00 |
| pplfair | 6.44 | 2.21 | 6.00 | 1.00 | 11.00 | 41879.00 |
| ppltrst | 5.67 | 2.36 | 6.00 | 1.00 | 11.00 | 42031.00 |
| trustindex3 | 4.93 | 1.86 | 5.00 | 0.00 | 10.00 | 42079.00 |

The main aesthetic limitation is that the number of decimal points cannot be set separately for each statistic, which in this particular table is inconvenient. Decimals can be rounded down to 0, but then the Mean and Std. Dev also gets rounded down.Also, variables are re-ordered alphabetically.

Another option is using another function also called descr(), but from the sjmisc package:

| var | n | NA.prc | mean | sd | md | range |
| --- | --- | --- | --- | --- | --- | --- |
| trustindex3 | 42079 | 0.00 | 4.93 | 1.86 | 5 | 10 (0-10) |
| ppltrst | 42031 | 0.11 | 5.67 | 2.36 | 6 | 10 (1-11) |
| pplfair | 41879 | 0.48 | 6.44 | 2.21 | 6 | 10 (1-11) |
| cntry | 42079 | 0.00 | 4.65 | 2.25 | 4 | 7 (1-8) |
| facntr | 42079 | 0.00 | 0.96 | 0.19 | 1 | 1 (0-1) |
| mocntr | 42079 | 0.00 | 0.97 | 0.18 | 1 | 1 (0-1) |
| female | 42079 | 0.00 | 0.54 | 0.50 | 1 | 1 (0-1) |
| agea | 42079 | 0.00 | 51.13 | 12.86 | 53 | 55 (25-80) |
| eduyrs25 | 41626 | 1.08 | 12.38 | 4.32 | 12 | 25 (0-25) |
| paredu\_a\_high | 39765 | 5.50 | 0.33 | 0.47 | 0 | 1 (0-1) |

Here, on the other hand, what’s inconvenient is that the order of the statistics (column headings) cannot be changed. That’s probably easier to do manually in Word, however, if needed.

## Results

This is the section where you describe the results of your main analysis, which will be some regression model. In this example, I fit three nested models - M1, M2 and M3 - starting with a simple bivariate regression of trust on education years only (M1), then expanding the list of predictors with age and sex (M2), and finally including also parental education and parents born abroad. The *.Rmd* file fits and saves these models in the background.

In class, your aim was to print summary statistics from the fit model in a format that helped you to interpret the results. Apart from the standard call to summary(), we have tried summ() from the jtools package.

Here, your aim is to find the best way to present your model(s) to readers. The two summary tools we used in class are not ideal for directly knitting nicely formatted regression model summaries to Word documents in RMarkdown. The *.Rmd* file contains some better functions for this purpose.

Here is an example of a summary table of a single model using the parameters package):

| Parameter | Coefficient | SE | 95% CI | t(39381) | p |
| --- | --- | --- | --- | --- | --- |
| (Intercept) | 3.028 | 0.053 | (2.92, 3.13) | 57.373 | < .001 |
| eduyrs25 | 0.097 | 0.002 | (0.09, 0.10) | 42.015 | < .001 |
| agea | 0.012 | 7.294e-04 | (0.01, 0.01) | 16.751 | < .001 |
| female | 0.046 | 0.018 | (0.01, 0.08) | 2.543 | 0.011 |
| paredu a high | 0.179 | 0.021 | (0.14, 0.22) | 8.601 | < .001 |
| fmnoncntr | -0.010 | 0.040 | (-0.09, 0.07) | -0.245 | 0.807 |

A summary table of several models (using the parameters package):

| Parameter | M1 | M2 | M3 |
| --- | --- | --- | --- |
| (Intercept) | 3.75\*\*\* (0.03) | 3.04\*\*\* (0.05) | 3.03\*\*\* (0.05) |
| eduyrs25 | 0.10\*\*\* (2.06e-03) | 0.10\*\*\* (2.11e-03) | 0.10\*\*\* (2.32e-03) |
| agea |  | 0.01\*\*\* (7.10e-04) | 0.01\*\*\* (7.29e-04) |
| female |  | 0.04\* (0.02) | 0.05\* (0.02) |
| paredu a high |  |  | 0.18\*\*\* (0.02) |
| fmnoncntr |  |  | -9.74e-03 (0.04) |
|  |  |  |  |
| Observations | 41626 | 41626 | 39387 |

## Discussion

## Conclusions

## References