# ETC5513

## Reproducible and Collaborative Practices

Assignment 2

# Contents

# Assignment objectives

Demonstrate that you are capable to create a reproducible report using R, Rstudio, Git and GitHub employing all the tools that we have learned in Weeks 1 to 4. In particular:

- Working on a reproducible Rstudio project

- Produce a pdf report using the template discuss in Week 7.

- Practice markdown syntax

- Practice R coding

- Use R chunk options to customize your report template

- Create tables, add table captions and refer to them in the report text as described in Week 4

- Create simple figures to visualize the data, add figure captions and refer to the figures in the report text as you learned in Week 4

- Demonstrate that you are able to work in groups to clone a GitHub repository locally and synchronize the changes between your local and remote repositories while using individual branches.

- Show that you can create and merge branches as well as resolve conflicts.

- Show that you can create meaningful Git commits so that the changes and the history of the project can be recorded and tracked.

==This assignment contains both INDIVIDUAL and GROUP components.==

==The printed final PDF report cannot exceed 16 pages for teams of four or 20 pages for teams of five (not including appendix pages). Additionally, you cannot use a dataset that has been previously used for other assignments or in other units this year or previous years. Please read the instructions below carefully.==

# Assignment description

I have created a GitHub repository for you. The assignment repo is inside GitHub classroom in the link below (**please log in into your GitHub account before accessing the link**

below). <mark>Please nominate one of the team members to first clone the repository and create a new team with the team name you got in Moodle. Then the rest of the team can join that team using the link below too:</mark>

[Link to GitHub Assignment 2 repo]()

- Once you access the link below, please follow the instructions (as you did in assignment 1) so that you can access the GitHub repo that contains the assignment repo. Remember that you first need to log in into your GitHub account!

<mark>The entire history must be recorded in this repository since the begining. If that is not the case the assignment will be grade with a 0.</mark>

In this assignment, you need to create an Rstudio project that will be able to produce a pdf business report using the template provided in Week 7 (please find Rmd temple in the Moodle site under Week 7 and modify it as you see fit). Discuss with your team what is the best way to do that. Probably the best way to do this is to nominate a team mate to add the template from Moodle in the main branch of the repository and then you all can create your branches from there and modify the template as a team.

Please use the template provided in this repo and read here about how the report should be structure and the elements that should have. <mark>Please make sure you read these resources:</mark>

- https://www.monash.edu/learnhq/excel-at-writing/annotated-assessment-samples/business-and-economics/buseco-report-writing and

- https://www.monash.edu/learnhq/excel-at-writing/annotated-assessment-samples/business-and-economics/buseco-report-writing/writing-the-business-report].

The structure of the report must have the following components:

- Title and authors
- Table of contents
- Executive summary
- Introduction
- The body of the report should contain the following sections:
  - Methodology
  - Results
  - Discussion, conclusion and recommendations
- Reference section.

You may choose whether to include the R code output in your report or not. However, please bear in mind that we expect a professional business report. Therefore, make sure not to include R messages or warnings in the report output. <mark>You may choose to include relevant</mark>

==code parts that support your report==. When selecting a data set or data sets for your report, please choose ones that align with your group's interests from reputable sources. Please note that **Kaggle** data sets are **not** permitted.

==The size of the data set is not relevant, however each data file cannot be larger than 50MB.==

Your **reproducible** report should be created using an Rmd file in Rstudio and must knit into a pdf file **without errors**.

The reproducible project files structure should look like this. Please add additional folders if needed:
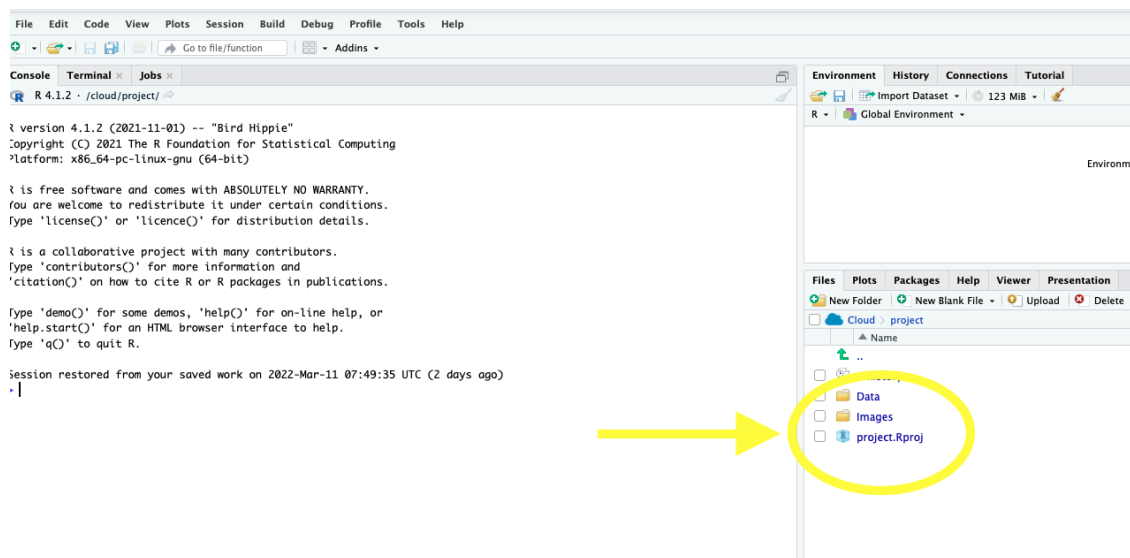


# Instruction for completing the assignment

As a group, you need to agree on a topic with several subtopics for each individual in the team and select a data set or data sets for the assignment.

## Part A

Each team member will work on their own Git branch named after their student **username** and produce a report with the following components for their assigned subtopic:

- Title and authors (1pt)
- Table of contents: Defining using adequate YAML syntax (1pt)
- Executive summary: Maximum of 4 sentences (3pts)
- Introduction: Maximum 10 sentences (5pts)
- The body of the report should contain the following sections:
    - Methodology: Maximum 300 words (5pts). Should include a figure and a table and those must be referenced in the text and have adequate captions (4pts).

– Results: Maximum 200 words (3pts). Should include either a figure **or** a table.
– Discussion, conclusion and recommendations
– Reference section: Include at least 1 reference. (1pt)

## Part B

As a team each of you will need to:

1. Merge the branches: Start by merging all the branches into the main branch. You will need to get organized and do this one by one so that you can resolve the conflicts that will appear.

2. Resolve conflicts: Since you will be working on the same sections, conflicts may arise when the same lines of code are modified in different branches. To resolve conflicts, edit the affected files and choose which version to keep. Once you have resolved the conflicts, commit the changes.

3. Make sure that the report knits into a pdf with errors each time you do a merge. If errors arise please solve them.

4. Review and Revise: Carefully review the final report to ensure that it meets all the requirements of a cohesive report and is free of errors. Make any necessary revisions and edits before submitting the report. For this, you will all probably need to work on the main branch. Please ensure that any changes made are in terms of the requirements and in a coordinated fashion to avoid any problems.

**The final business report must look professional, coherent, and connected**

**Students who do not participate in Part B will receive a grade of 0 for their assignments. Please note that teams will not be penalized if individual students do not contribute to the work on their respective branches.**

# Marking rubric

In addition to the points for each section as described above, you will be graded on your:

- Report template (5pts). For example the default Rmd template will be marked with 1pt, if you use more sophisticated templates that will add points into this component of the assignment (to a max of 5pts)
- Issues with spelling and grammar (up to -5pts)
- R code style (5pts)
- Report quality: Sections in the report are connected and aligned with the research question in a coherent way. (10pts)

- Messages in commits are clear and informative (5pts)
- Work on individual branches of the git repository (10pts)
- Work on the main branch of the git repository (10pts)
- Git tree structure (5pts)

# The usage of AI (chatGPT for example)

- You may use chatGPT to **correct your English** or to help you with your R code (you can use chatGPT to query **your code** and for example to find code bugs or to ask for hints to enhance **your existing code**). However, if you use chatGPT you must declare it by adding a section in your Rmd report called **Appendix** where you display screen shots of your chatGPT queries and all the interactions that you have had with chatGPT related to this assignment.

- You **cannot** use chatGPT to generate content for the assignment from scratch or to create code for you from scratch. You must always provide chatGPT with an input and you must screenshot all the those interactions, and report that usage in the appendix of your reproducible report.

- Monash University supports the responsible and ethical use of generative AI. For more info please refer to Monash Policy and practise guidance around acceptable and responsible use of AI technologies.

Remember you can be better than chatGPT. If you just use chatGPT to create content for you, where is your value and why are your skills special?!

# Assignment submission: Moodle

The report must be rendered into a PDF format and be reproducible. Any reports that are not reproducible will be marked as a 0.

Each team member must submit in Moodle the following items to be submitted in Moodle:

1. pdf file of the report
2. Paste the http of your GitHub repo in the text box and your team name.
3. A zip file of your entire repository.

VERY IMPORTANT: The printed final PDF report cannot exceed 16 pages for teams of four or 20 pages for teams of five (not including appendix pages)! You cannot use a data set that has been used for other assignment or in another units.

# Submission deadline

Friday May 5, 4:30pm (Melbourne time). **Late submissions will not be accepted.**

# Plagiarism

Monash University is committed to honesty and academic integrity. There are serious consequences for plagiarism and collusion. If plagiarism and/or collusion is detected further actions will be taken according to Monash University policy and procedures. More info here:

https://www.monash.edu/students/admin/policies/academic-integrity

- You cannot re-use assignments that have been submitted or used in other units!