Intro to R

Why use R?

- Open source and free
  - R is available for free, making it accessible to everyone
  - A large, active community contributes to continuous improvement and knowledge sharing
  - Strong in statistical analysis
- R has a wide array of packages dedicated to statistical analysis
  - R is great at producing complex plots
  - R has specialised tools for various fields, such as econometrics and randomisation for research and evaluation
- Reproducibility
  - Script based workflow: R uses scripts for analysis, which can be shared and re-executed to reproduce results
  - R works well with Git for version control, helping track changes and collaborate efficiently
  - Tools like Quarto allow for the integration of code, results, and text into a single document, ensuring analyses are transparent and reproducible

Workflow

- RStudio: An IDE (integrated development environment) for R
  - RStudio provides a user friendly interface
- Working directories

  - The folder where R looks for files and saves output
  - Avoid using relative paths - use R projects

- Creating an R project
  - R Projects help organise all related files, scripts, and data in a single directory
  -

- Scripts
  - Scripts are written in `.R` files, where you can save commands and code for later use
  - You can run line-by-line or in chunks, making it easy to test and debug your scripts
  - Use comments within scripts to explain the purpose of code sections, improving readability and collaboration.

- Packages
  - Packages can be installed using `install.packages()`
  - Once installed packages are loaded into the R session using `library()`
  - Make sure to install and load tidyverse
    * `install.packages("tidyverse")`
    * `library(tidyverse)`

Reporting with Quarto

- Case study: TASO Technical Guide

Resources for further learning

- R for Data Science
- TASO's coding good practice
- TASO data visualisation style guide
- Productive R Workflow (paid for course)