

# Outline day 1\*

Tran Thai Hung

Last version: January 27, 2025

## Table of contents

<b>1. Set up</b>	<b>2</b>
1.1 Git vs Github . . . . .	2
1.2 Rproject / Project Environment . . . . .	2
<b>2. R Foundations</b>	<b>2</b>
2.1 Rscript . . . . .	2
2.2 Variable/Classes/Special Values . . . . .	2
2.3 Functions . . . . .	2
2.4 Scope . . . . .	2
2.5 Packages . . . . .	3
<b>3. In-class Exercise</b>	<b>3</b>
<b>4. Data Manipulation in R</b>	<b>3</b>
4.1 Data Import/Export . . . . .	3
4.2 Clean Column Names . . . . .	3
4.3 Simple Data Extraction . . . . .	3
4.4 Data Summary . . . . .	3

## List of Figures

---

\*Please do not circulate.

# 1. Set up

## 1.1 Git vs Github

- Difference between Git and GitHub.
- Why version control is important for collaboration.
- Hands-on: Creating and cloning repositories.

## 1.2 Rproject / Project Environment

- Overview of RStudio projects.
- Benefits of using a project environment.
- Setting up your first R project.

# 2. R Foundations

## 2.1 Rscript

- What is an R script?
- How to create and run scripts in RStudio.

## 2.2 Variable/Classes/Special Values

- Variables and data types (numeric, character, logical, etc.).
- Special values (e.g., NA, NaN, Inf, NULL).

## 2.3 Functions

- Creating and using functions.
- Built-in vs custom functions.

## 2.4 Scope

- Local vs global scope.
- The environment in R.

## 2.5 Packages

- Installing and loading packages.
- Exploring CRAN, Bioconductor, and GitHub.

## 3. In-class Exercise

- Practice using what we've learned so far.
- Tasks involving creating variables, using functions, and exploring packages.

## 4. Data Manipulation in R

### 4.1 Data Import/Export

- Reading data from CSV, Excel, and other formats.
- Writing data to files.

### 4.2 Clean Column Names

- Standardizing column names using `janitor::clean_names()`.

### 4.3 Simple Data Extraction

- Selecting rows and columns.

### 4.4 Data Summary

- Summarizing data with `summary` and `skim`.