Introduction Practical

Kursus R: Pengenalan dan Praktikal

Mohd Azmi Bin Suliman

azmi.suliman@moh.gov.my

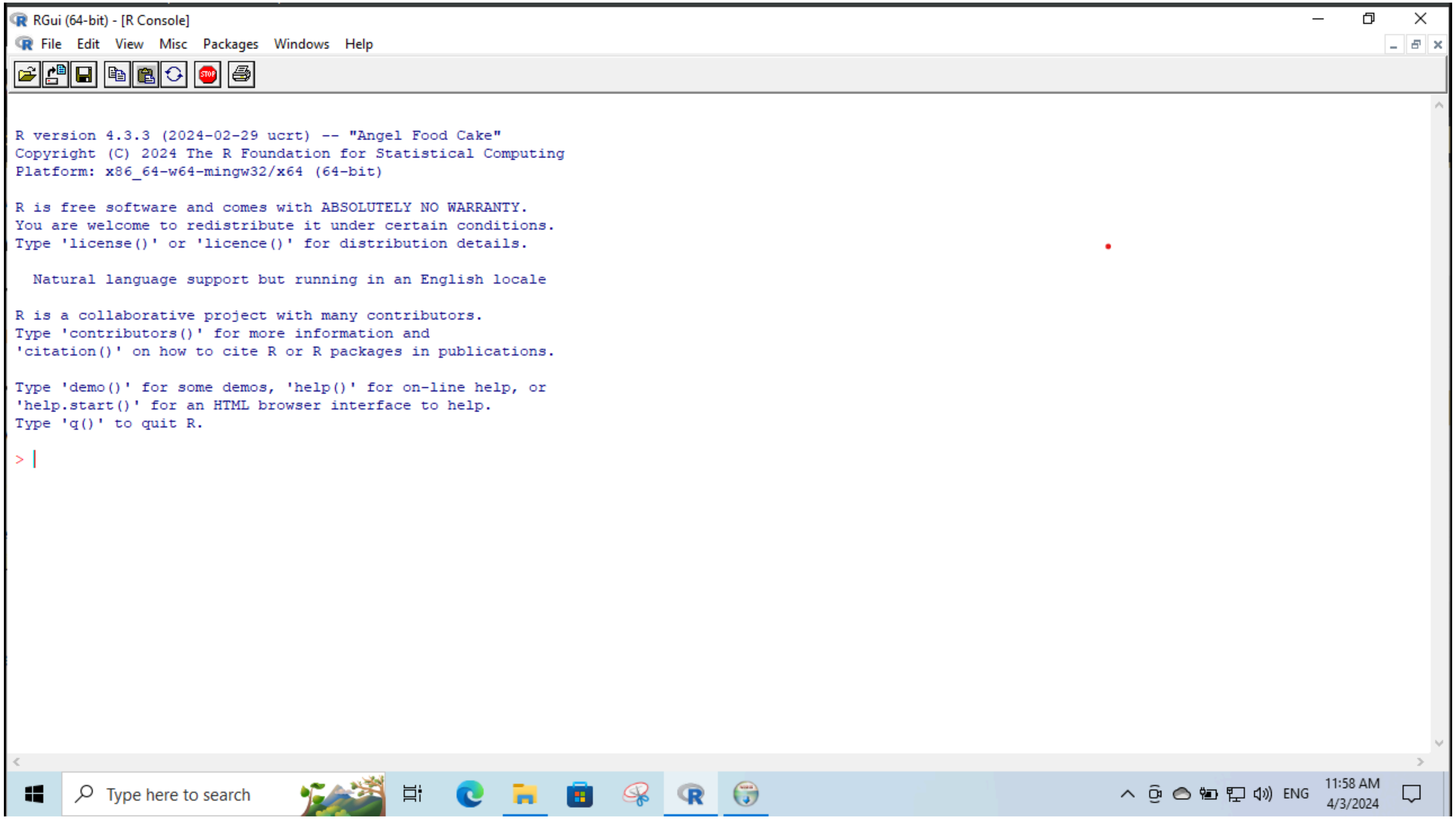
Pusat Penyelidikan Penyakit Tak Berjangkit, Institut Kesihatan Umum

Wednesday, 02 October 2024

Getting to know R

Typical R Session

- Open your R console.

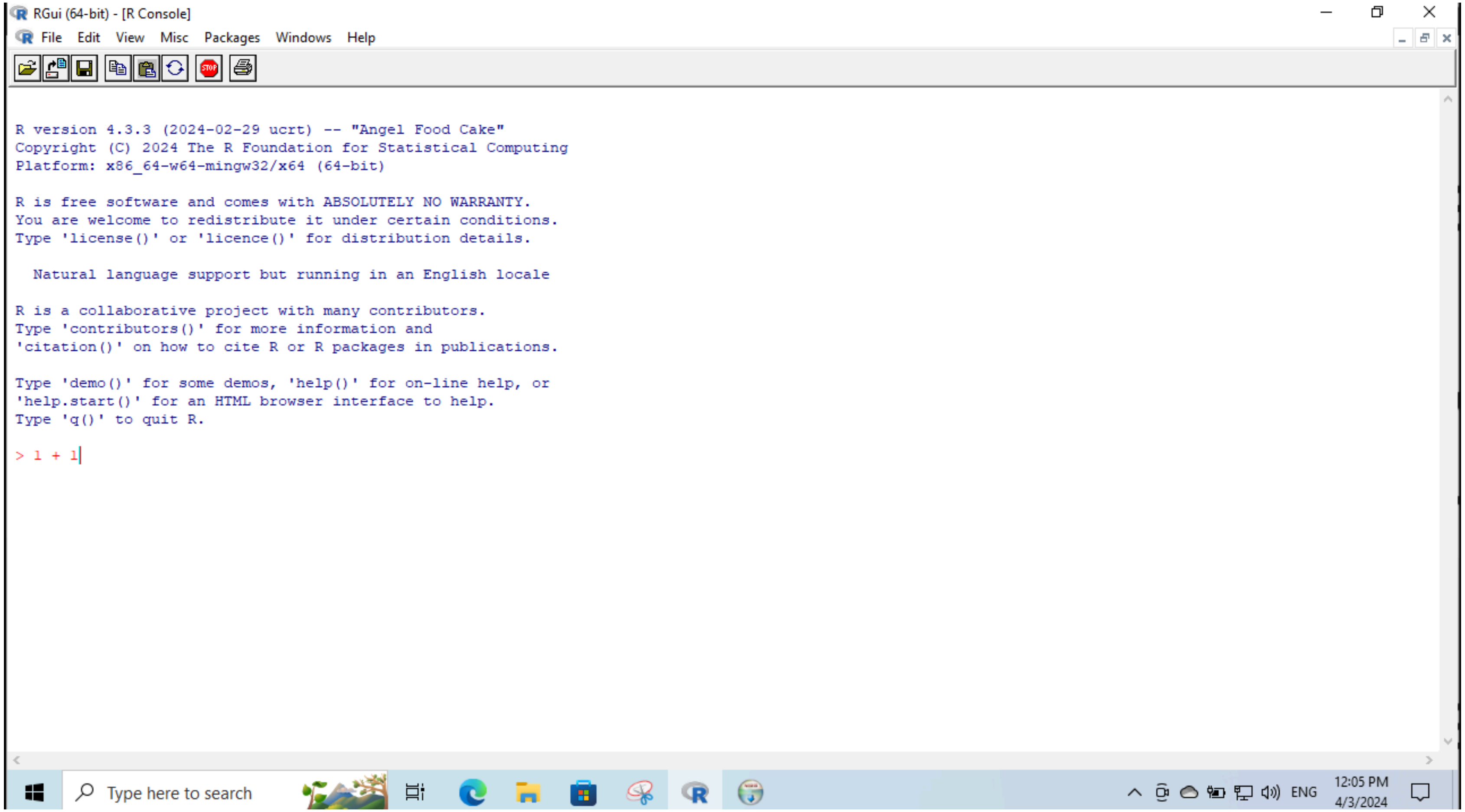


RGui (64-bit) - [R Console]
File Edit View Misc Packages Windows Help
Type here to search
ENG
11:58 AM
4/3/2024

```
R version 4.3.3 (2024-02-29 ucrt) -- "Angel Food Cake"
Copyright (C) 2024 The R Foundation for Statistical Computing
Platform: x86_64-w64-mingw32/x64 (64-bit)

R is free software and comes with ABSOLUTELY NO WARRANTY.
You are welcome to redistribute it under certain conditions.
Type 'license()' or 'licence()' for distribution details.

  Natural language support but running in an English locale

R is a collaborative project with many contributors.
Type 'contributors()' for more information and
'citation()' on how to cite R or R packages in publications.

Type 'demo()' for some demos, 'help()' for on-line help, or
'help.start()' for an HTML browser interface to help.
Type 'q()' to quit R.

> 
```

Typical R Session

- R console is where you can type in the R command/code.

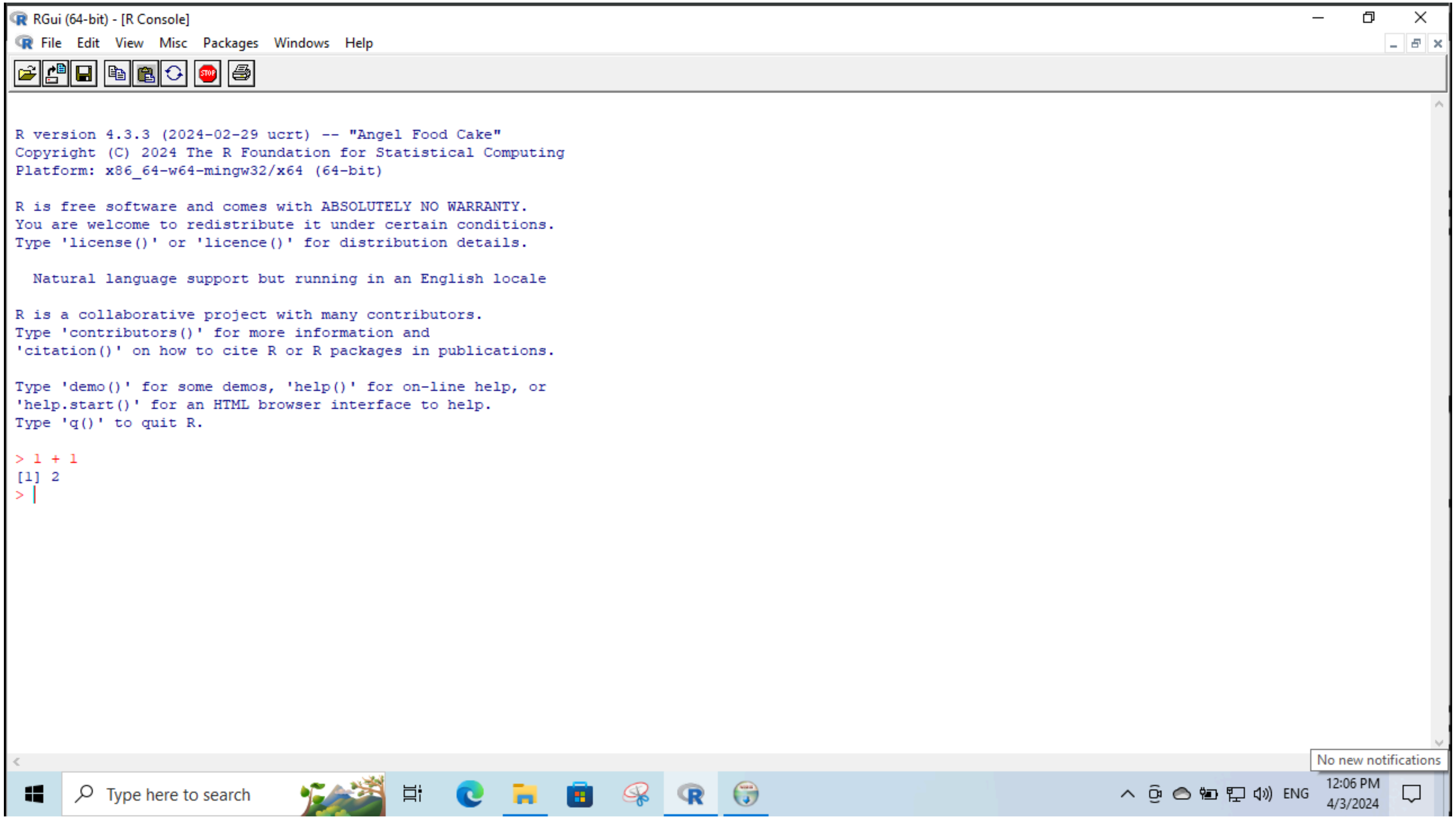


RGui (64-bit) - [R Console]
File Edit View Misc Packages Windows Help
Type here to search
ENG
12:05 PM
4/3/2024

```
R version 4.3.3 (2024-02-29 ucrt) -- "Angel Food Cake"
Copyright (C) 2024 The R Foundation for Statistical Computing
Platform: x86_64-w64-mingw32/x64 (64-bit)

R is free software and comes with ABSOLUTELY NO WARRANTY.
You are welcome to redistribute it under certain conditions.
Type 'license()' or 'licence()' for distribution details.

  Natural language support but running in an English locale

R is a collaborative project with many contributors.
Type 'contributors()' for more information and
'citation()' on how to cite R or R packages in publications.

Type 'demo()' for some demos, 'help()' for on-line help, or
'help.start()' for an HTML browser interface to help.
Type 'q()' to quit R.

> 1 + 1
```

Typical R Session

- The output of the command is shown below the command.



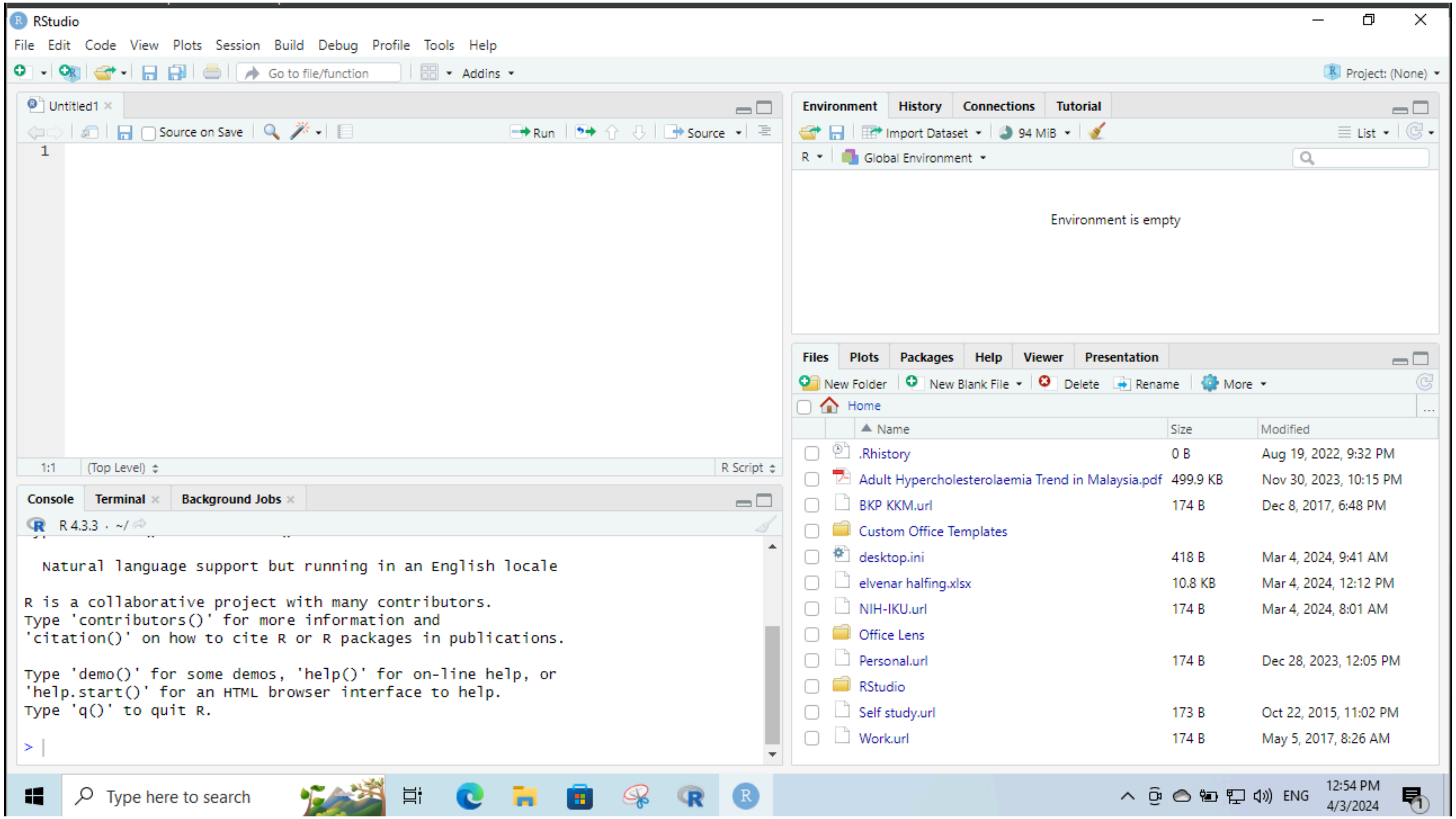
RGui (64-bit) - [R Console]
File Edit View Misc Packages Windows Help
No new notifications
Type here to search
ENG
12:06 PM
4/3/2024

```
R version 4.3.3 (2024-02-29 ucrt) -- "Angel Food Cake"
Copyright (C) 2024 The R Foundation for Statistical Computing
Platform: x86_64-w64-mingw32/x64 (64-bit)

R is free software and comes with ABSOLUTELY NO WARRANTY.
You are welcome to redistribute it under certain conditions.
Type 'license()' or 'licence()' for distribution details.

  Natural language support but running in an English locale

R is a collaborative project with many contributors.
Type 'contributors()' for more information and
'citation()' on how to cite R or R packages in publications.

Type 'demo()' for some demos, 'help()' for on-line help, or
'help.start()' for an HTML browser interface to help.
Type 'q()' to quit R.

> 1 + 1
[1] 2
> |
```

RStudio, the IDE for R

RStudio Layout

- When you first open RStudio, you will usually see a layout like this.



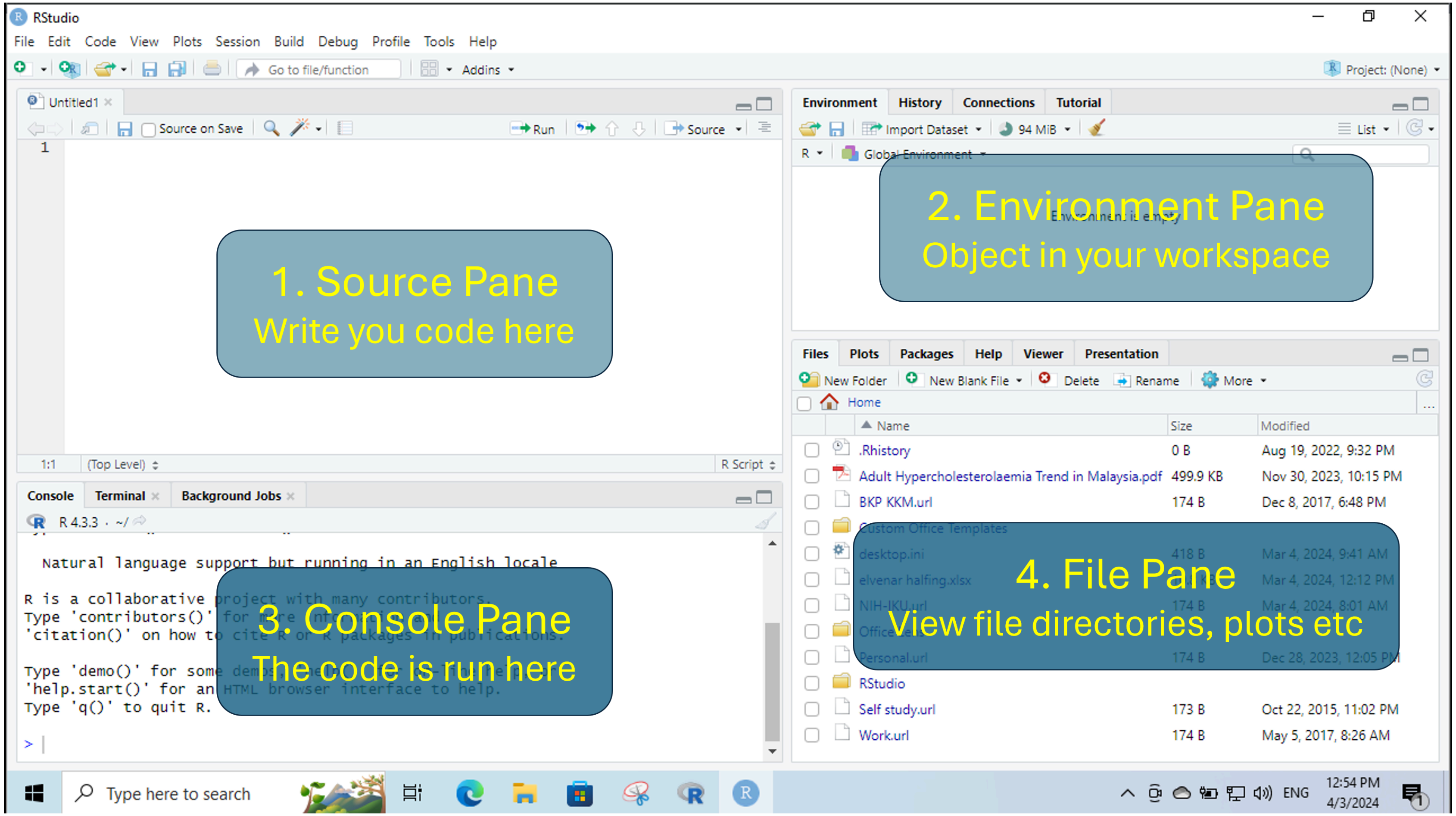
Environment is empty

```
  Natural language support but running in an English locale

R is a collaborative project with many contributors.
Type 'contributors()' for more information and
'citation()' on how to cite R or R packages in publications.

Type 'demo()' for some demos, 'help()' for on-line help, or
'help.start()' for an HTML browser interface to help.
Type 'q()' to quit R.

> |
```

Name	Size	Modified
.Rhistory	0 B	Aug 19, 2022, 9:32 PM
Adult Hypercholesterolaemia Trend in Malaysia.pdf	499.9 KB	Nov 30, 2023, 10:15 PM
BKP KKM.url	174 B	Dec 8, 2017, 6:48 PM
Custom Office Templates		
desktop.ini	418 B	Mar 4, 2024, 9:41 AM
elvenar halfing.xlsx	10.8 KB	Mar 4, 2024, 12:12 PM
NIH-IKU.url	174 B	Mar 4, 2024, 8:01 AM
Office Lens		
Personal.url	174 B	Dec 28, 2023, 12:05 PM
RStudio		
Self study.url	173 B	Oct 22, 2015, 11:02 PM
Work.url	174 B	May 5, 2017, 8:26 AM

RStudio Layout

- There are four (4) panes in RStudio.



1. Source Pane
Write you code here

2. Environment Pane
Object in your workspace

3. Console Pane
The code is run here

4. File Pane
View file directories, plots etc

```
  Natural language support but running in an English locale

R is a collaborative project with many contributors.
Type 'contributors()' for more information and
'citation()' on how to cite R or R packages in publications.

Type 'demo()' for some demos, 'help()' for on-line help, or
'help.start()' for an HTML browser interface to help.
Type 'q()' to quit R.

>
```

Name	Size	Modified
.Rhistory	0 B	Aug 19, 2022, 9:32 PM
Adult Hypercholesterolaemia Trend in Malaysia.pdf	499.9 KB	Nov 30, 2023, 10:15 PM
BKP KKM.url	174 B	Dec 8, 2017, 6:48 PM
Custom Office Templates		
desktop.ini	418 B	Mar 4, 2024, 9:41 AM
elvenar halfing.xlsx	[illegible]	Mar 4, 2024, 12:12 PM
NIH-IKU.url	174 B	Mar 4, 2024, 8:01 AM
Office Lens		
Personal.url	174 B	Dec 28, 2023, 12:05 PM
RStudio		
Self study.url	173 B	Oct 22, 2015, 11:02 PM
Work.url	174 B	May 5, 2017, 8:26 AM

Setting up your RStudio

- Different people may use R and RStudio differently.
- Having a sensible workflow can improve your work.
  1. Treat individual R process and the associated workspace as disposable.
  2. Use IDE & Project Management
  3. **Specified your default working directory**
  4. Save your work in a script (or Quarto)
  5. Always start with a blank slate.
     - → When quit R, **DO NOT SAVE WORKSPACE.**
     - → When starting R, **DO NOT RESTORE WORKSPACE.**
- Further reading: What They Forgot to Teach You About R

Setting up your RStudio

1. Open your RStudio.
2. Go to `Tools` > `Global Options` > `General`

Setting up your RStudio

1. Open your RStudio.
2. Go to `Tools` > `Global Options` > `General`

Default Working Directory

3. Set your default working directory.
    - Set working directory inside RStudio folder in your Document folder
    - Click `Browse...` > Create New Folder > name it `RStudio` > Select Folder > OK

Setting up your RStudio

1. Open your RStudio.
2. Go to `Tools` > `Global Options` > `General`

Blank slate setting

4. Unchecked these
    - `Restore most recently opened project at startup`
    - `Restore previously open source documents at startup`
    - `Restore .RData into workspace at startup`
    - Save workspace to .RData on exit: `Never`
    - `Always save history (even when not saving .RData)`

Creating Project (and have sensible workflow)

RStudio Project

- RStudio have built in project management.
- RStudio project act as a container for your work.
- All the codes, data, scripts and outputs are organized in one place
- Allow for easy sharing and collaboration
- You might forget what you did last time, but having a project will help you to revisit your old work.

Create a new project

- Go to `File` > `New Project...`
- `New Directory` > `New Project`
- Name your project `R_Practical`
- Don't forget to check your working directory
- Click `Browse...` > Select `RStudio` folder > Select `Open`
- Click `Create Project`

Quarto, Notebook for R

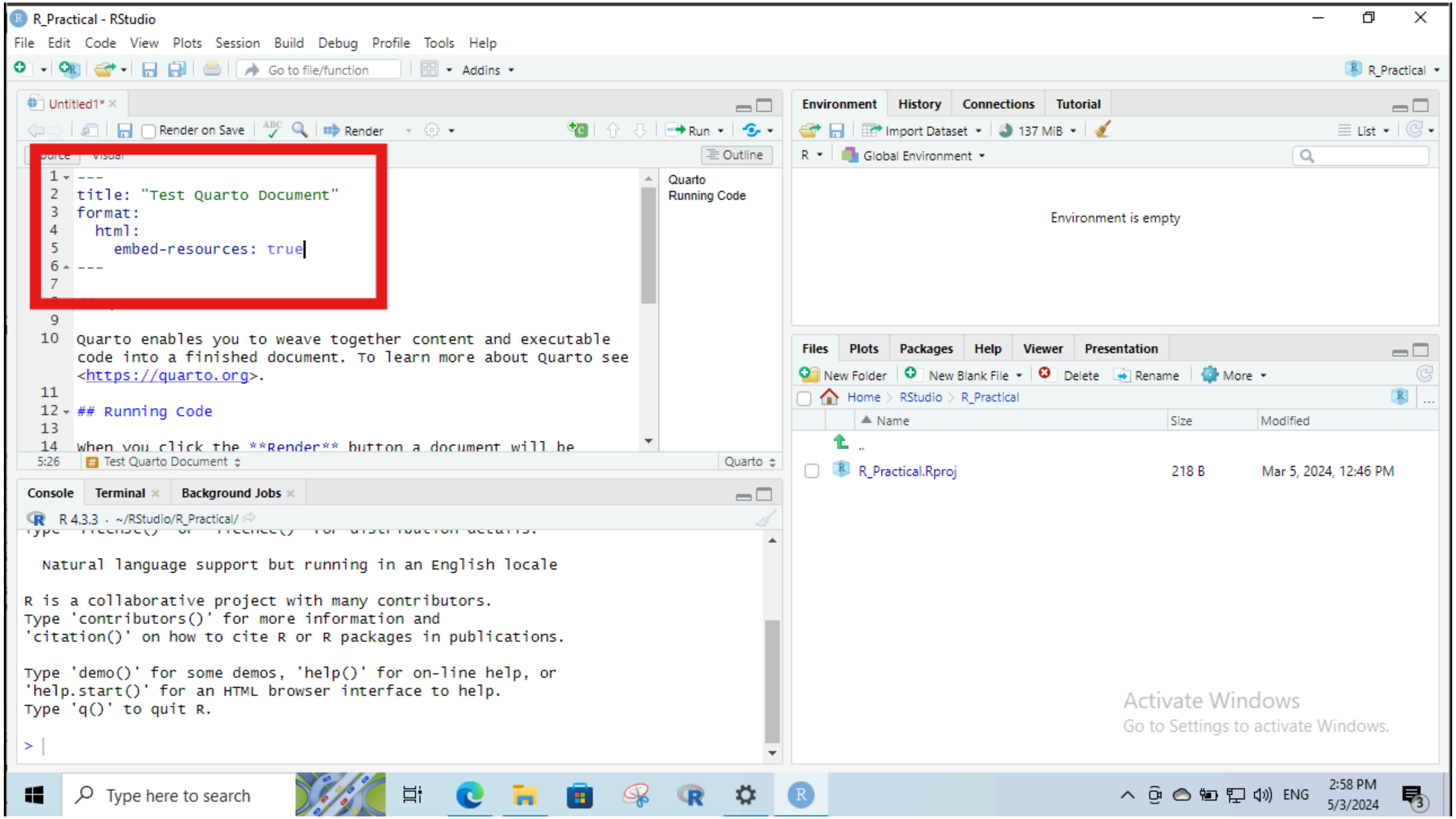
Setting up your Quarto

- Go to `File` > `New File` > `Quarto Document...`
- Add your title `Test Quarto Document`
- Add the author (*optional*)
- Select `HTML`
- Set engine as `Knitr`
- **UNTICK** `Use visual markdown editor`
- Click `Create`

Setting up your Quarto

- **WAIT!!**, we need to ensure that it is standalone.

  → Add `embed-resources: true` to the YAML header.



```
---
title: "Test Quarto Document"
format:
  html:
    embed-resources: true
---


Quarto enables you to weave together content and executable
code into a finished document. To learn more about Quarto see
<https://quarto.org>.

## Running Code

When you click the **Render** button a document will be
```

Environment is empty

Name	Size	Modified
..		
R_Practical.Rproj	218 B	Mar 5, 2024, 12:46 PM

```
  Natural language support but running in an English locale

R is a collaborative project with many contributors.
Type 'contributors()' for more information and
'citation()' on how to cite R or R packages in publications.

Type 'demo()' for some demos, 'help()' for on-line help, or
'help.start()' for an HTML browser interface to help.
Type 'q()' to quit R.

> |
```

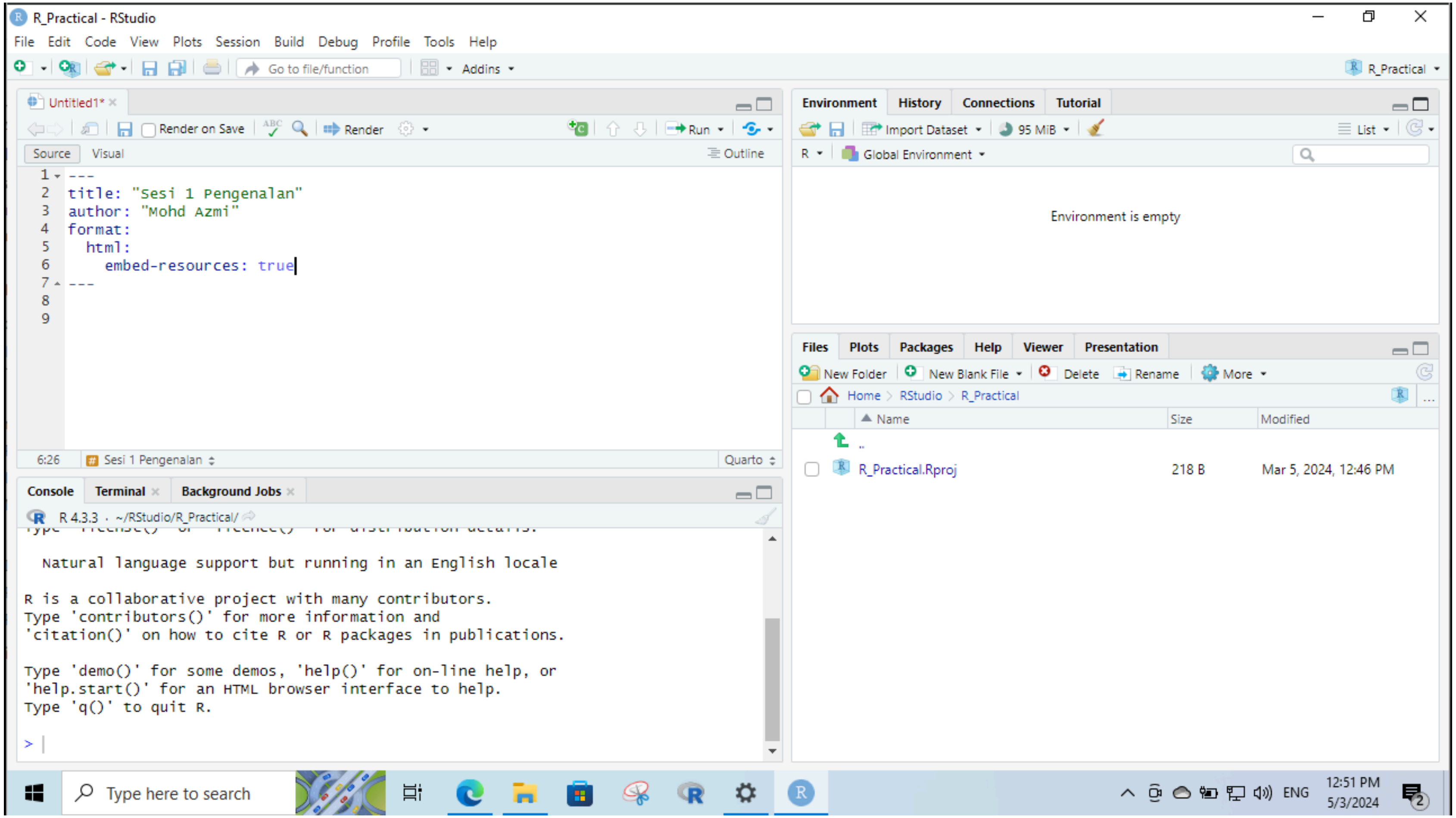
Now let get familiar with Quarto Notebook

Key Concepts: Objects and Data Types

TODO: Setting up blank Quarto Document

- Go to `File` > `New File` > `Quarto Document...`
- Add your title `Sesi 1 Pengenalan`
- Add the author (*optional*)
- Select `HTML`
- Set engine as `Knitr`
- **UNTICK** `Use visual markdown editor`
- Click `Create Blank Document`
- Add `embed-resources: true` to the YAML header.

TODO: Setting up blank Quarto Document



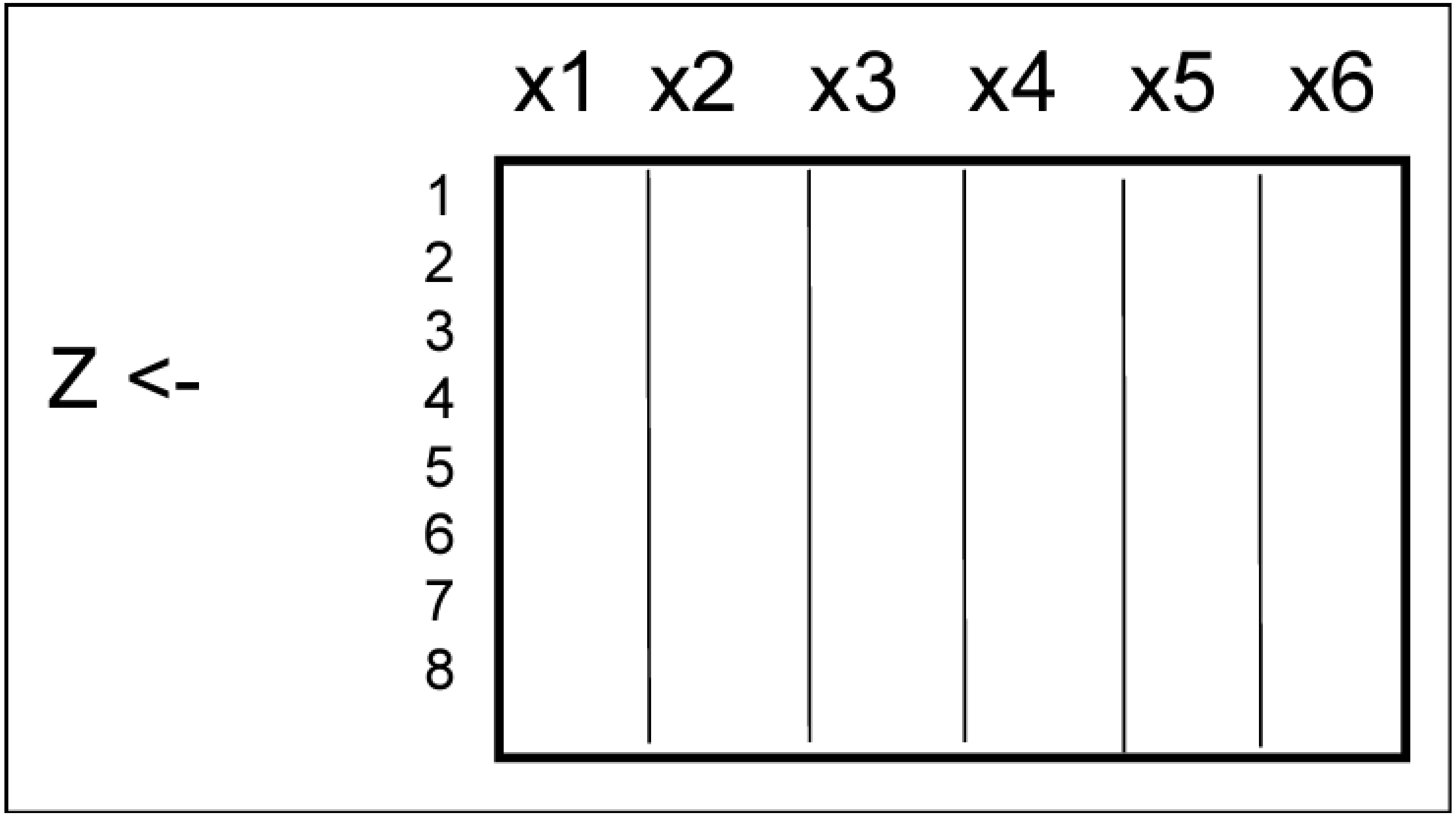
```
---
title: "Sesi 1 Pengenalan"
author: "Mohd Azmi"
format:
  html:
    embed-resources: true
---
```

```
Type 'license()' or 'licence()' for distribution details.

  Natural language support but running in an English locale

R is a collaborative project with many contributors.
Type 'contributors()' for more information and
'citation()' on how to cite R or R packages in publications.

Type 'demo()' for some demos, 'help()' for on-line help, or
'help.start()' for an HTML browser interface to help.
Type 'q()' to quit R.

>
```

R Objects

- Almost everything in R is objects: functions, datasets, results, and models.
- Script can be thought of as a way to make objects.
- Your goal is usually to write a script that, by its end, has created the objects that you need.
  - → statistical results
  - → graphics

R Objects



R modes

- Modes of an object in R refers to the basic type of its elements.
  - → lower-level descriptions
- R has several modes of objects

Modes	**Examples**
Numeric	1, 2, 3
Character	a, b, c
Logical	TRUE
Complex	1+0i
Raw	
List	
Function	mean(age)
NULL	

Data Structure

- In R, the data can be in various forms.
  - → Arrays
  - → Vectors
  - → Lists
  - → Factors
  - → Matrices
  - → Data Frames

Vectors (Logical)

- A vector is a sequence of data elements of the same basic type
- There are five types of vectors in R: Logical, Numeric, Integer, Complex, Character
- Logical Vector

```r
logical_vector <- c(TRUE, FALSE, TRUE)
logical_vector
```

```
[1]  TRUE FALSE  TRUE
```

Vectors (Numeric)

- A vector is a sequence of data elements of the same basic type
- There are five types of vectors in R: Logical, Numeric, Integer, Complex, Character
- Numeric Vector

```r
numeric_vector <- c(1, 2, 3)
numeric_vector
```

```
[1] 1 2 3
```

Vectors (Character)

- A vector is a sequence of data elements of the same basic type
- There are five types of vectors in R: Logical, Numeric, Integer, Complex, Character
- Character Vector

```r
character_vector <- c("a", "b", "c")
character_vector
```

```
[1] "a" "b" "c"
```

Lists

- A list is a special type of vector that can contain elements of different types.

```r
list_vector <- list(1, "a", TRUE)
list_vector
```

```
[[1]]
[1] 1

[[2]]
[1] "a"

[[3]]
[1] TRUE
```

Factors

- A factor is a vector that can contain only predefined values.
- Used to store categorical data.

```r
education_levels <- c("MSc", "BSc", "Dip", "PhD")
education_factor <- factor(education_levels, ordered = TRUE,
                            levels = c("Dip", "BSc", "MSc", "PhD"))
education_factor
```

```
[1] MSc BSc Dip PhD
Levels: Dip < BSc < MSc < PhD
```

Matrix

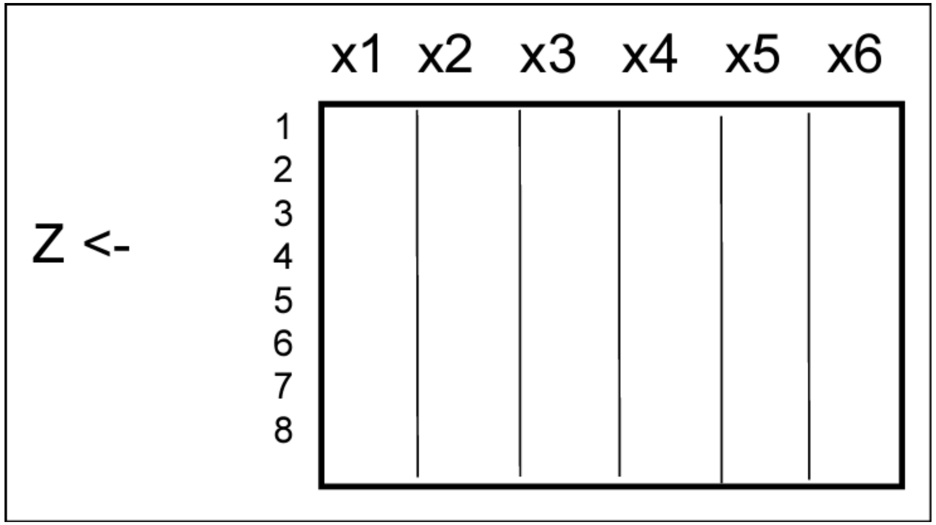
- A matrix is a 2D array-like structure.
- Unlike a list, a matrix only holds single basic types.

```r
example_matrix <- matrix(1:9, nrow = 3, ncol = 3)
example_matrix
```

```
     [,1] [,2] [,3]
[1,]    1    4    7
[2,]    2    5    8
[3,]    3    6    9
```

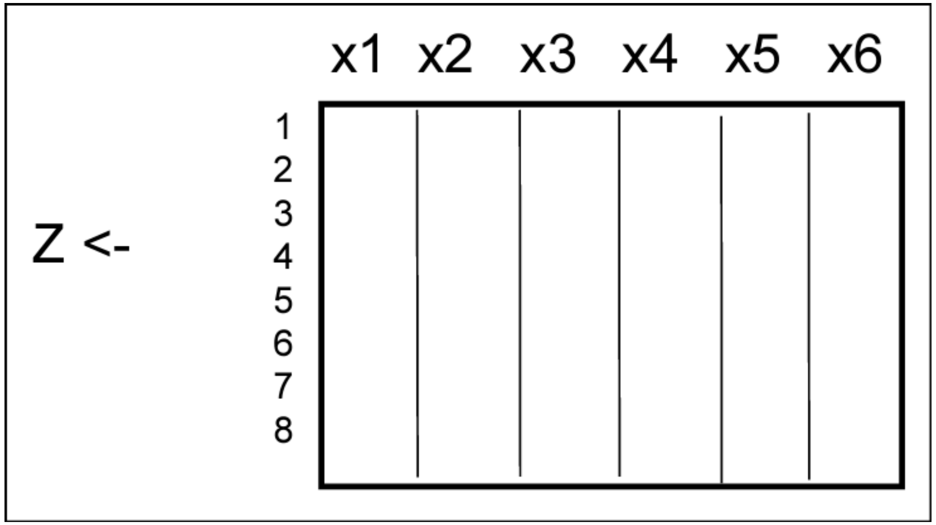
Data Frame

- Most familiar to Excel and SPSS user
- Unlike a matrix, a data frame allows for different data types in each column.
- A data frame is a table or a 2D array-like structure.
  - → Each column contains values of one variable.
  - → Each row contains one set of values from each column



Data Frame

- Data frame is the most common way of storing data in R.
- Allow user to perform row-wise, column-wise and cell-wise functions



Operators and Functions

Operators

- R has several operators.
  - → Arithmetic
  - → Assignment
  - → Logical
  - → Relational
  - → Special

Arithmetic Operators

- Addition

```r
1 + 2
```

```
[1] 3
```

- Subtraction

```r
2-1
```

```
[1] 1
```

- Multiplication

```r
3*4
```

```
[1] 12
```

- Division

```r
4/2
```

```
[1] 2
```

Assignment Operators

- Assignment operators are used to assign values to variables.
- R accept four types of assignment operators: `<-`, `<<-`, `=`, `->`

```
a <- 5
a
```

```
[1] 5
```

```
b = 7
b
```

```
[1] 7
```

```
9 -> c
c
```

```
[1] 9
```

Logical Operators

- Logical operators are used to combine or compare logical values

- Basic logical operators
  - → !: NOT
  - → &: AND
  - → |: OR

```
TRUE | FALSE
```

```
[1] TRUE
```

```
TRUE & FALSE
```

```
[1] FALSE
```

```
!TRUE
```

```
[1] FALSE
```

```
!FALSE
```

```
[1] TRUE
```

Logical Operators

- Logical operators are used to combine or compare logical values

- Equality and relational operators
  - → ==: equal
  - → !=: not equal
  - → >: greater than
  - → <: less than
  - → >=: greater than or equal to
  - → <=: less than or equal to

```r
1 == 1
```

```
[1] TRUE
```

```r
1 != 1
```

```
[1] FALSE
```

```r
1 > 1
```

```
[1] FALSE
```

```r
1 < 1
```

```
[1] FALSE
```

```r
1 >= 1
```

```
[1] TRUE
```

```r
1 <= 1
```

```
[1] TRUE
```

Special Operators

- Special operators are used to perform special operations.
  - → `:`: sequence
  - → `%in%`: match
  - → `%*%`: matrix multiplication
  - → `%/%`: integer division
  - → `%%`: modulus

```r
1:5
```

```
[1] 1 2 3 4 5
```

```r
5 %in% 1:10
```

```
[1] TRUE
```

```r
2 %*% 3
```

```
     [,1]
[1,]    6
```

```r
5 %/% 2
```

```
[1] 2
```

```r
5 %% 2
```

```
[1] 1
```

Functions

- Functions are a set of commands grouped together to perform a specific task.
- R has many built-in functions.
- Functions are called by their name followed by parentheses ()

```r
sqrt(4)
```

```
[1] 2
```

- R also allow the user to create their own function (advance)