

DATA ANALYSIS USING R

CSAEC49

Credits: 1:0:0

Dr.Parkavi.A

Associate Professor,

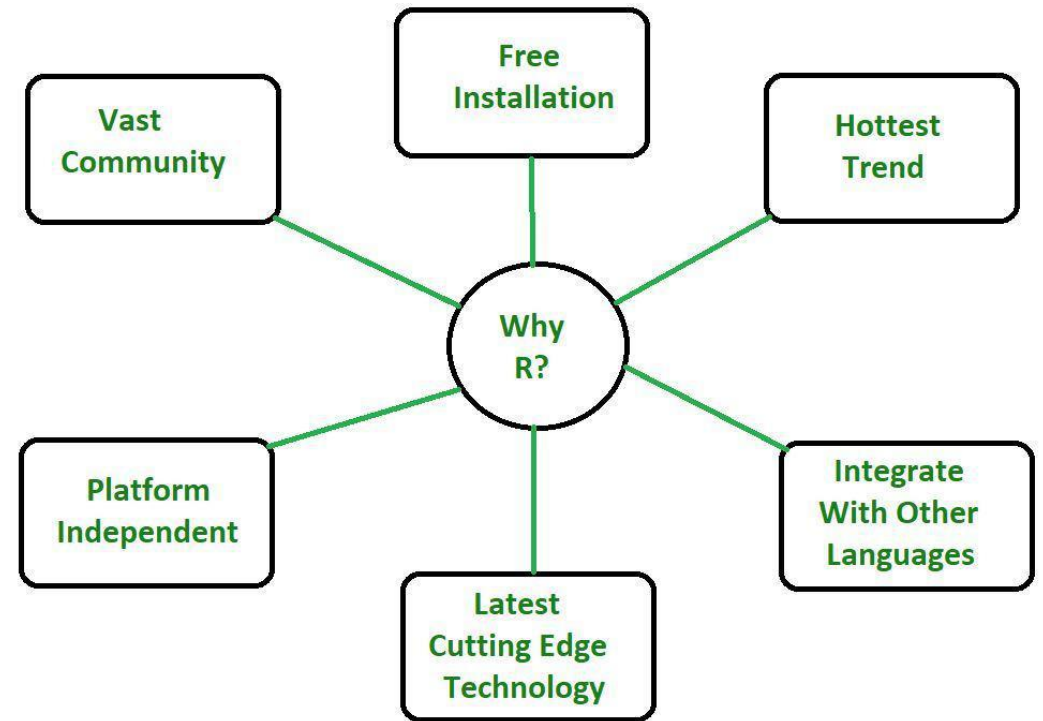
Dept of CSE, RIT

20 Mark component

- Individual practical assignment
- One research article to be studied
- Explain to Faculty ->5 marks (after CIE-1)
- Implement it->10 marks (After CIE-1 take 2 weeks to finish)
- Report->5 marks
- Ieee, acm, Elsevier, science direct, mdpi, springer, sage publishers, igi global (2023,2024)

Introduction to R programming

- Open-source programming language
- Statistical software and data analysis tool
- Command-line interface
- Windows, Linux, and macOS
- Tool for machine learning, statistics, and data analysis
- Objects, functions, and packages
- Integrate with other languages
- Ross Ihaka and Robert Gentleman at the University of Auckland, New Zealand,



What is R?

- Platform Independent
- Latest Tech
- Statistical Analysis, Machine Learning, Data Analytics
- Open Source and free
- Data Visualization and Manipulation
- Command – Line Interface
- Easy integration with other languages
- Community and Packages
- Distributed Computing
- IDE like R Studio
- Memory and Speed

How to Install R and R Studio?

- Text, and images ,doc,html
- Shiny
- <https://posit.co/download/rstudio-desktop/>

1: Install R

RStudio requires R 3.3.0+. Choose a version of R that matches your computer's operating system.

DOWNLOAD AND INSTALL R

2: Install RStudio

DOWNLOAD RSTUDIO DESKTOP FOR WINDOWS

Size: 214.34 MB | SHA-256: FE62B784 | Version:
2023.09.1+494 | Released: 2023-10-17

Install R

Download and Install R

Precompiled binary distributions of the base system and contributed packages, **Windows and Mac** users most likely want one of these versions of R:

- [Download R for Linux \(Debian, Fedora/Redhat, Ubuntu\)](#)
- [Download R for macOS](#)
- [Download R for Windows](#)

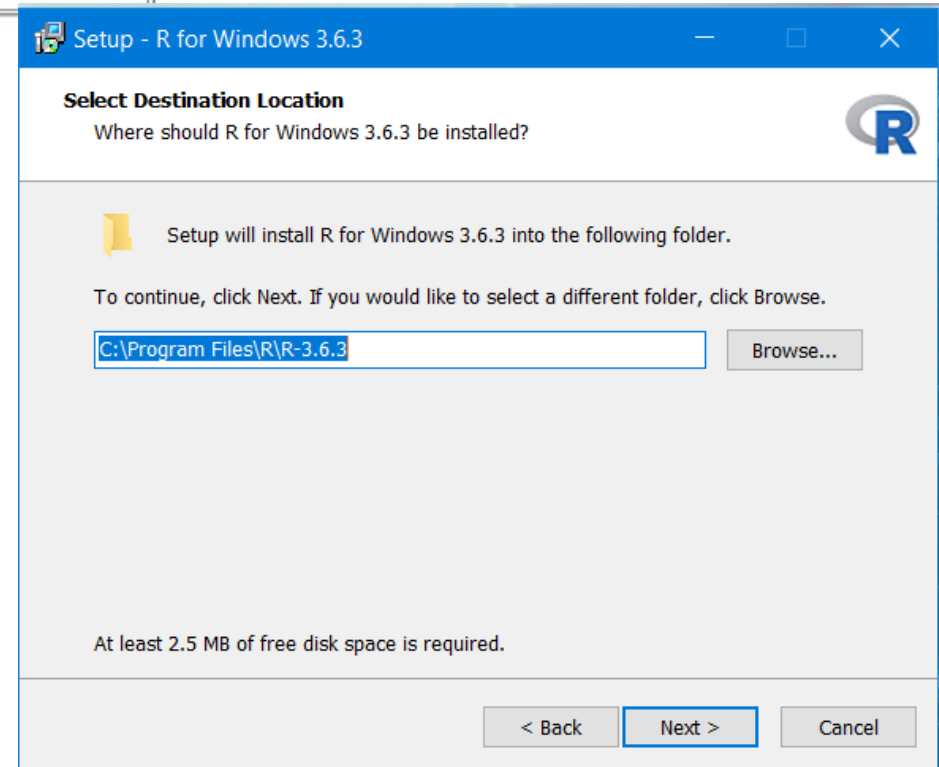
R is part of many Linux distributions, you should check with your Linux package management system in addition to the link above.

Source Code for all Platforms

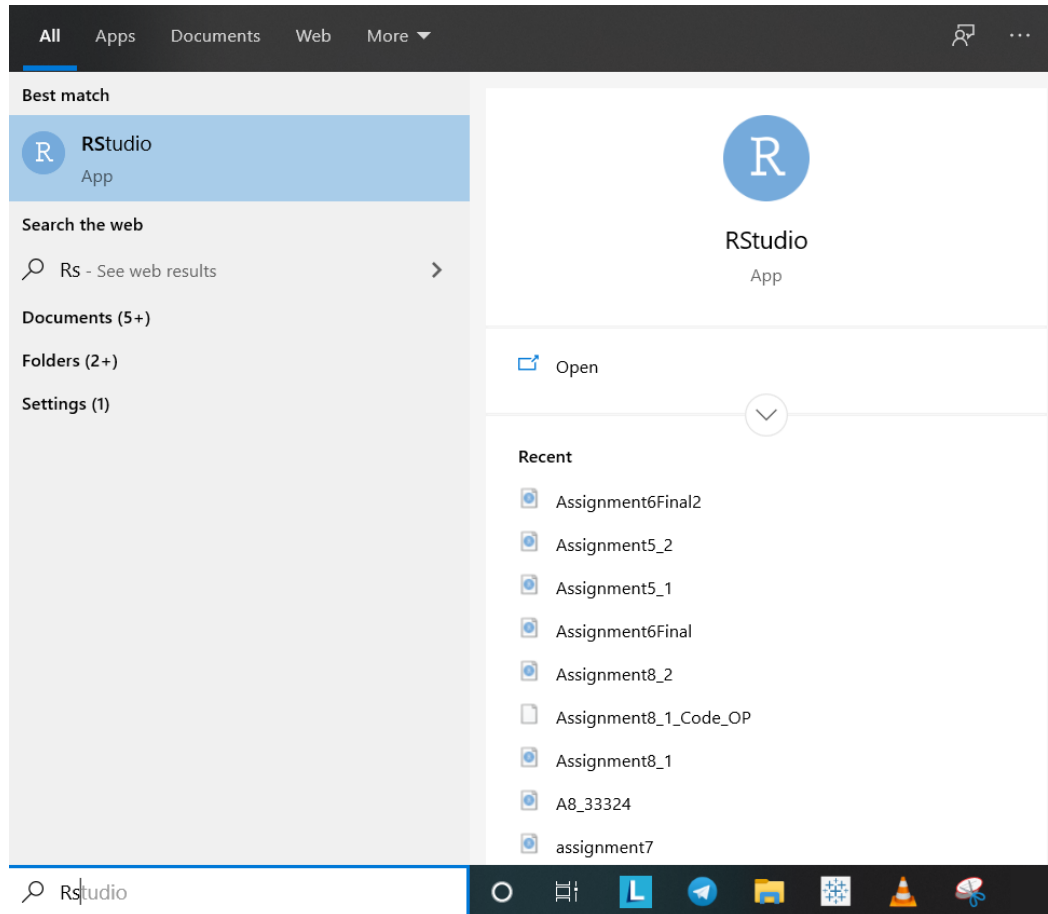
[Download R-4.3.2 for Windows](#) (79 megabytes, 64 bit)

[README on the Windows binary distribution](#)

[New features in this version](#)



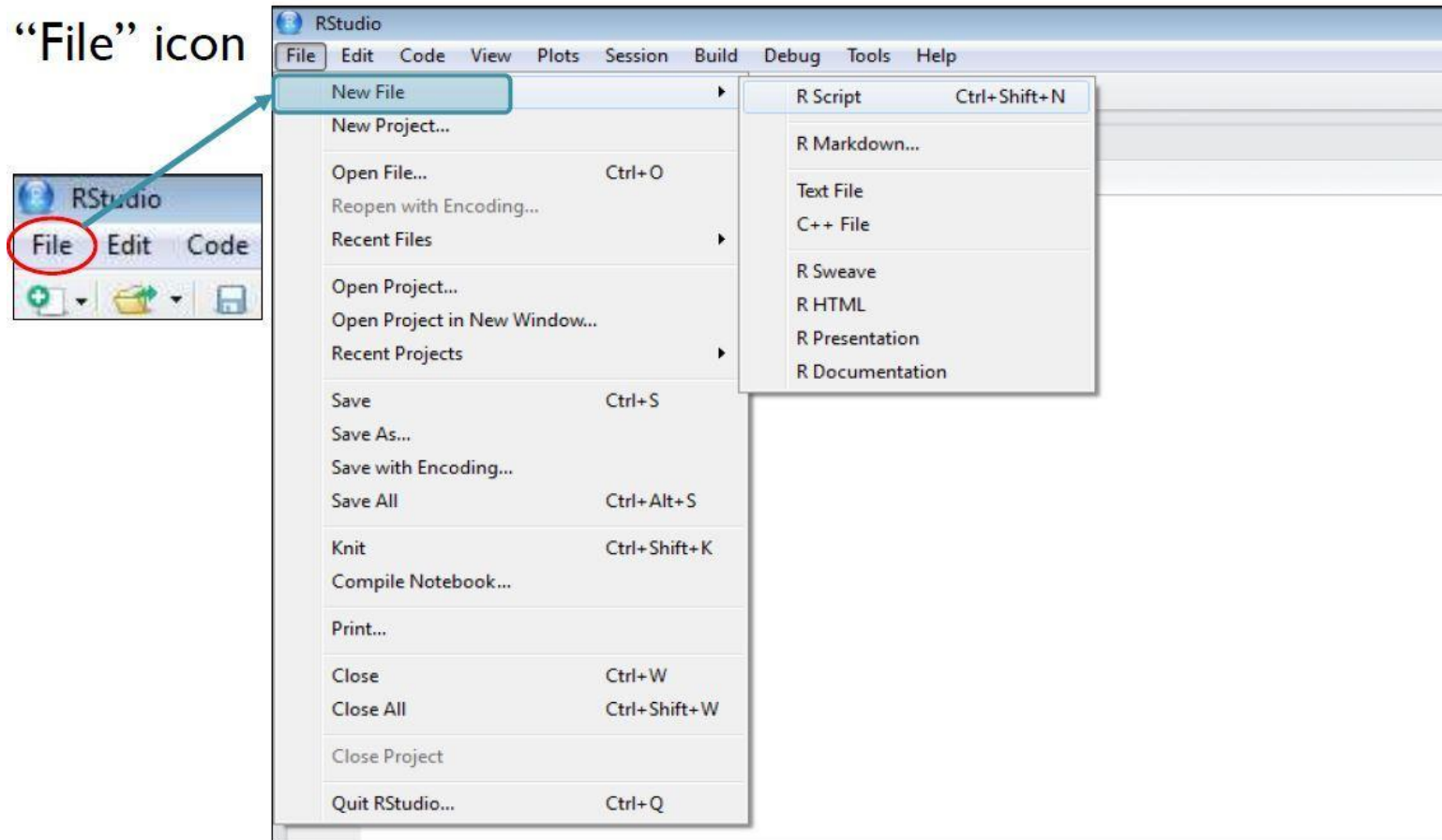
Install Rstudio



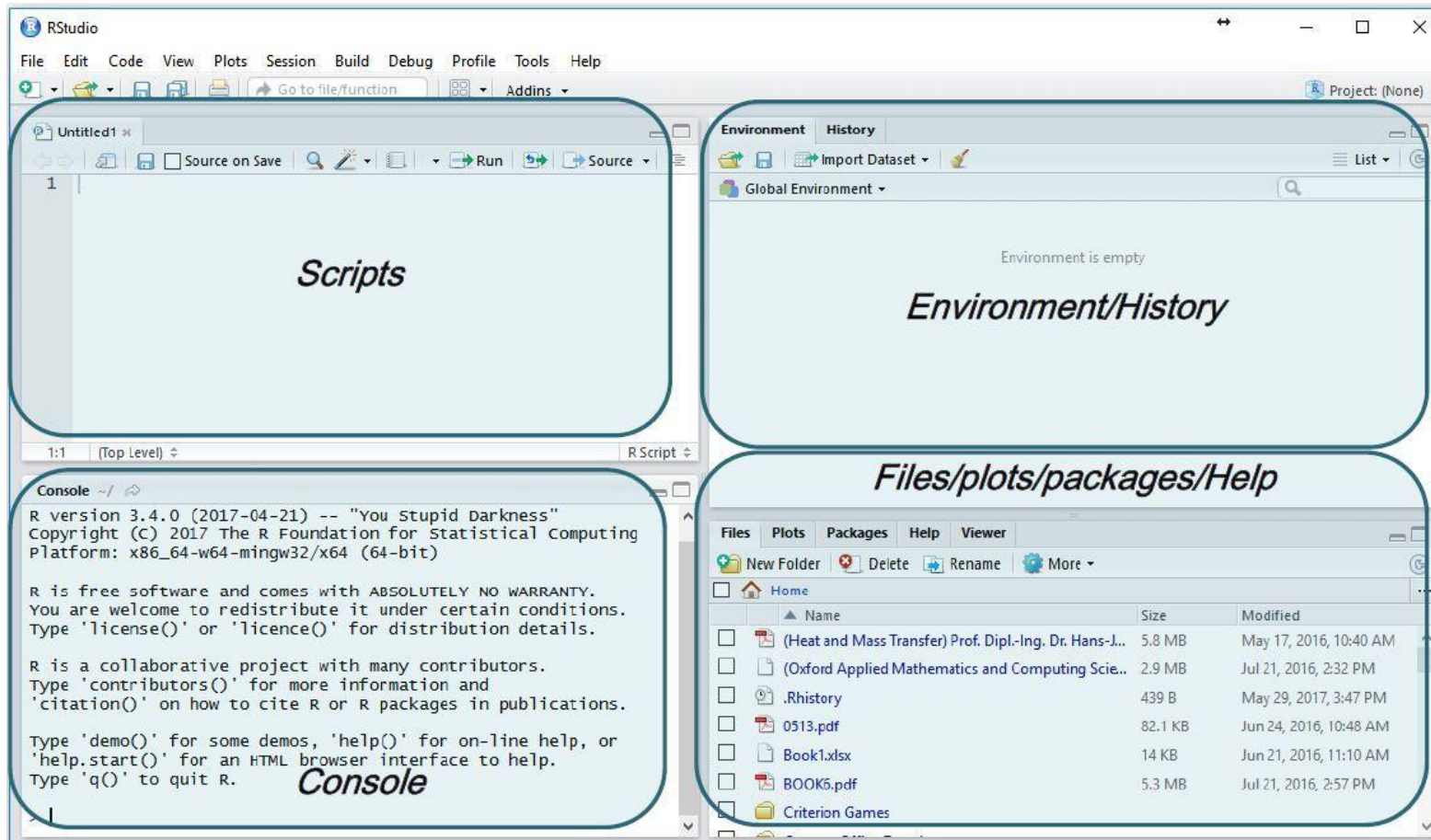
R Studio Overview

Creating an R file

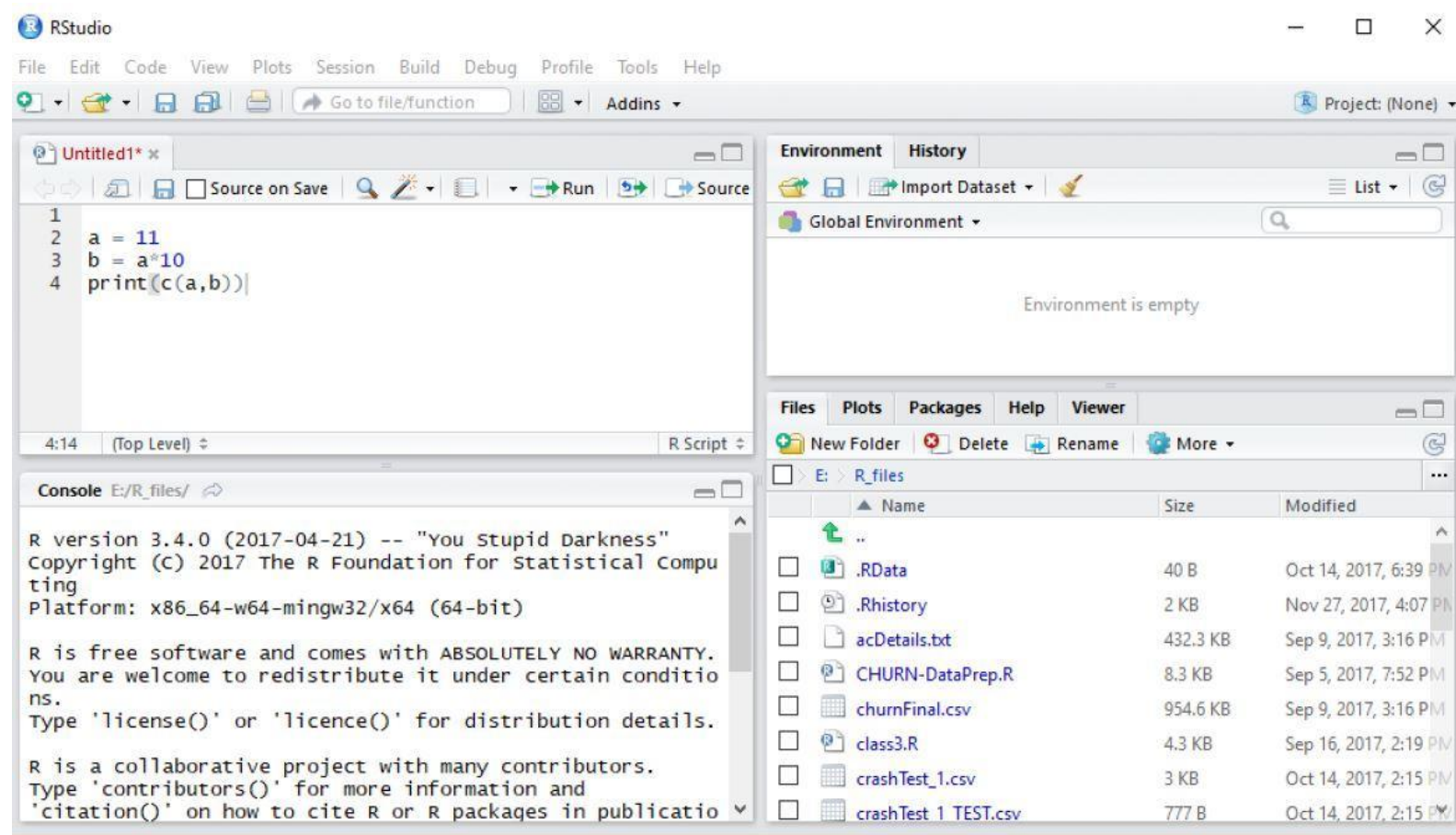
“File” icon



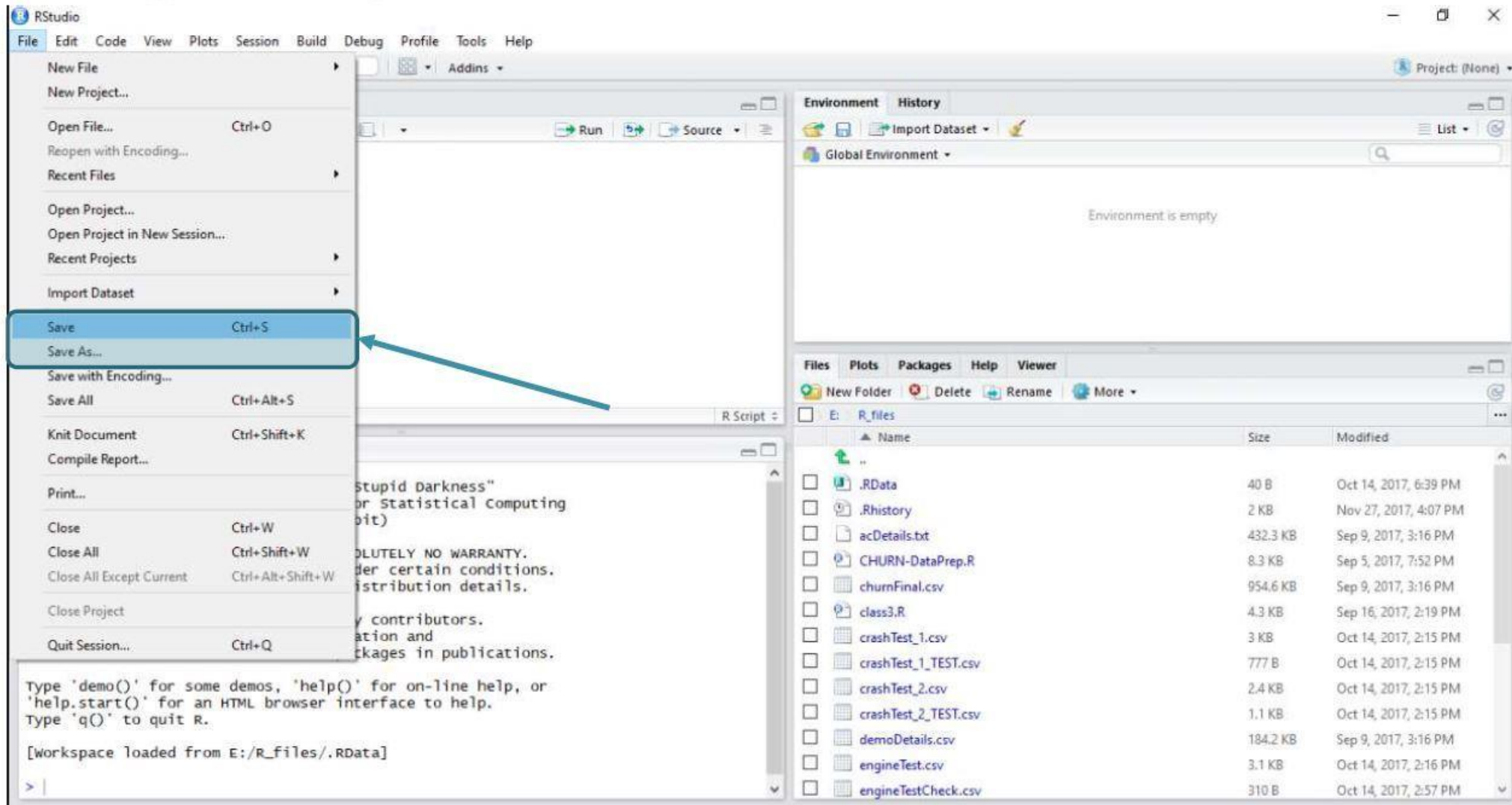
R Studio with the script file



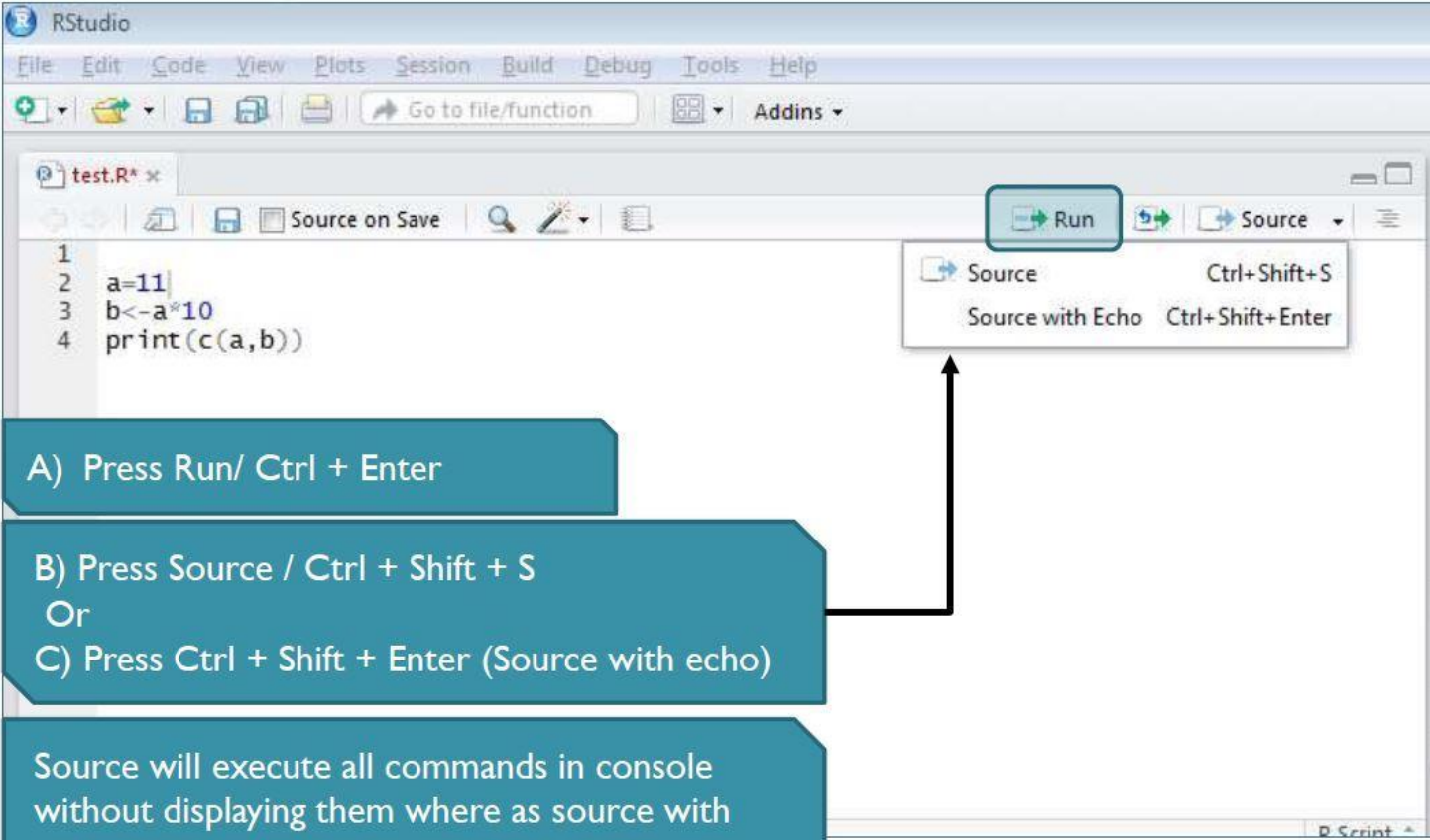
Writing Scripts in an R File



Saving an R File



Execution of an R file



The screenshot shows the RStudio interface with a script file named 'test.R' open. The script contains the following code:

```
1  
2 a=11  
3 b<-a*10  
4 print(c(a,b))
```

The 'Run' button is highlighted in the top toolbar. A context menu is open, showing the following options:

- Run (highlighted)
- Source (Ctrl+Shift+S)
- Source with Echo (Ctrl+Shift+Enter)

Three blue callout boxes provide instructions for executing the code:

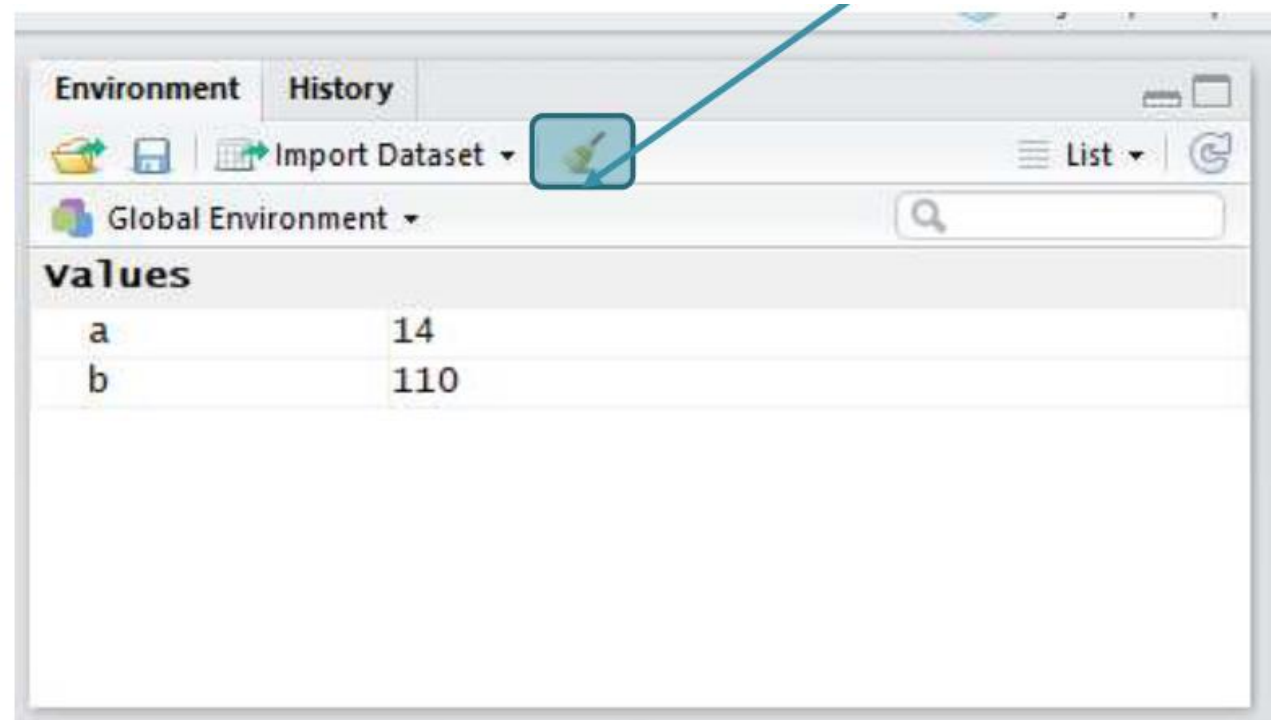
- A) Press Run/ Ctrl + Enter
- B) Press Source / Ctrl + Shift + S
- C) Press Ctrl + Shift + Enter (Source with echo)

A final blue callout box explains the difference between the two execution methods:

Source will execute all commands in console without displaying them where as source with echo will execute and print automatically

Working in the Console

- **Clearing the Console: ctrl + L**
- **Clearing the Environment:**
 - *rm(variable)*



Arithmetic Operators



```
# R program to illustrate  
# the use of Arithmetic operators  
vec1 <- c(0, 2)  
vec2 <- c(2, 3)  
  
# Performing operations on Operands  
cat ("Addition of vectors :", vec1 + vec2, "\n")  
cat ("Subtraction of vectors :", vec1 - vec2, "\n")  
cat ("Multiplication of vectors :", vec1 * vec2, "\n")  
cat ("Division of vectors :", vec1 / vec2, "\n")  
cat ("Modulo of vectors :", vec1 %% vec2, "\n")  
cat ("Power operator :", vec1 ^ vec2)
```

Output

```
Addition of vectors : 2 5  
Subtraction of vectors : -2 -1  
Multiplication of vectors : 0 6  
Division of vectors : 0 0.6666667  
Modulo of vectors : 0 2  
Power operator : 0 8
```


Logical operators



```
# R program to illustrate
# the use of Logical operators
vec1 <- c(0,2)
vec2 <- c(TRUE,FALSE)

# Performing operations on Operands
cat ("Element wise AND :", vec1 & vec2, "\n")
cat ("Element wise OR :", vec1 | vec2, "\n")
cat ("Logical AND :", vec1[1] && vec2[1], "\n")
cat ("Logical OR :", vec1[1] || vec2[1], "\n")
cat ("Negation :", !vec1)
```

Output

```
Element wise AND : FALSE FALSE
Element wise OR : TRUE TRUE
Logical AND : FALSE
Logical OR : TRUE
Negation : TRUE FALSE
```

Relational operators



```
# R program to illustrate  
# the use of Relational operators  
vec1 <- c(0, 2)  
vec2 <- c(2, 3)  
  
# Performing operations on Operands  
cat ("Vector1 less than Vector2 :", vec1 < vec2, "\n")  
cat ("Vector1 less than equal to Vector2 :", vec1 <= vec2, "\n")  
cat ("Vector1 greater than Vector2 :", vec1 > vec2, "\n")  
cat ("Vector1 greater than equal to Vector2 :", vec1 >= vec2, "\n")  
cat ("Vector1 not equal to Vector2 :", vec1 != vec2, "\n")
```

Output

```
Vector1 less than Vector2 : TRUE TRUE  
Vector1 less than equal to Vector2 : TRUE TRUE  
Vector1 greater than Vector2 : FALSE FALSE  
Vector1 greater than equal to Vector2 : FALSE FALSE  
Vector1 not equal to Vector2 : TRUE TRUE
```


Problems

- Write a R script to store student details(usn,name, 6 subjects marks) using variables.
 - Find out their total marks and average.
 - Check whether they are pass or fail in the all subjects using logical & relational operator.
- Write a R script to store faculty details(name, fid,salary, no. of papers published, no of books written, no of patents published, no. of consultancy works ,no of funded projects)using variables.
 - Give weightage for their contributions(eg. For each papers published 5 points)
 - Find out the faculty total points for their contributions.
 - If they score >75 display that “Appraisal is good” else “not satisfactory”

```
# Student details

# Define the student variables
usn <- "1RV20CS001"
name <- "John Doe"
marks <- c(85, 78, 92, 88, 76, 81) # Marks for 6
subjects

# Calculate total marks and average
total_marks <- sum(marks)
average_marks <- total_marks / length(marks)

# Check pass or fail in all subjects (assuming pass
mark is 40)

# pass_status <- all(marks >= 40)
pass_status <- TRUE
```

```
for (mark in marks){
    if (mark < 40) {
        pass_status <- FALSE
        break
    }
}

# Display student details and results
cat("Student Details:\n")
cat("USN:", usn, "\n")
cat("Name:", name, "\n")
cat("Total Marks:", total_marks, "\n")
cat("Average Marks:", average_marks, "\n")
cat("Pass Status:", ifelse(pass_status, "Pass",
"Fail"), "\n")
```

```
# Faculty details
# Define the faculty variables
faculty_name <- "Dr. Jane Smith"
fid <- "F001"
salary <- 95000
num_papers <- 15
num_books <- 3
num_patents <- 2
num_consultancy <- 4
num_funded_projects <- 5
# Define the weightage points
points_paper <- 5
points_book <- 10
points_patent <- 8
points_consultancy <- 6
points_funded_project <- 12
```

```
# Calculate the total points for contributions
total_points <- (num_papers * points_paper) +
  (num_books * points_book) +
  (num_patents * points_patent) +
  (num_consultancy * points_consultancy) +
  (num_funded_projects *
points_funded_project)

# Evaluate faculty appraisal
appraisal_status <- ifelse(total_points > 75, "Appraisal is
good", "Not satisfactory")

# Display faculty details and appraisal status
cat("Faculty Details:\n")
cat("Name:", faculty_name, "\n")
cat("FID:", fid, "\n")
cat("Salary:", salary, "\n")
cat("Total Points for Contributions:", total_points, "\n")
cat("Appraisal Status:", appraisal_status, "\n")
```

References

- <https://posit.co/download/rstudio-desktop/>
- <https://www.w3schools.com/r/default.asp>
- <https://www.geeksforgeeks.org/r-programming-language-introduction/>