Introduction

Why R and Installation Procedure

Kanahaiya Kumar [Data Mining Central]

What is R?

- R is an environment for data manipulation, statistical computing, graphics display and data analysis

- One can effectively handle the data and storage of outputs

- Simple as well as complicated calculations could be done

- Graphical display on screen and hard copy could be done

- Programming language is effective which includes all possibilities just like any other good programming language

- The R language is very similar in appearance to "S" Language based on a software s-plus

Switching to R?

- Initially R was developed by statistician for statistician and today diverse communities are using for different purposes

- The R language is widely used for research & development, analytics, etc

- The R language is open source and free and many statistical packages are freely available

- R has statistical computing environment and great support of communities

- It has a computer language which is convenient to use for statistical and graphical applications

- The commands can be saved, run and stored in script files

Switching to R?

- R is available for Windows, Unix, Linux, and Macintosh platforms

- Initially R was developed to compete with S-Plus

- Built in and contributed packages are available, and users are provided tools to make packages. It is possible to contribute your own packages.

- This language provides the logical control of branching and looping, and modular programming using functions

- Error messages are provided while executing a programme or a function which are helpful in finding the errors

Switching to R?

- R is interpreted computer language [Not compiler]

- When we use command line interface, each command or expression to be evaluated is typed at the command prompt, and immediately evaluated when the enter key is pressed at the end of a syntactically complete statement

- Tip
  - Press the up-arrow key to recall commands and edit them
  - Use the Esc [Escape] key to cancel a command

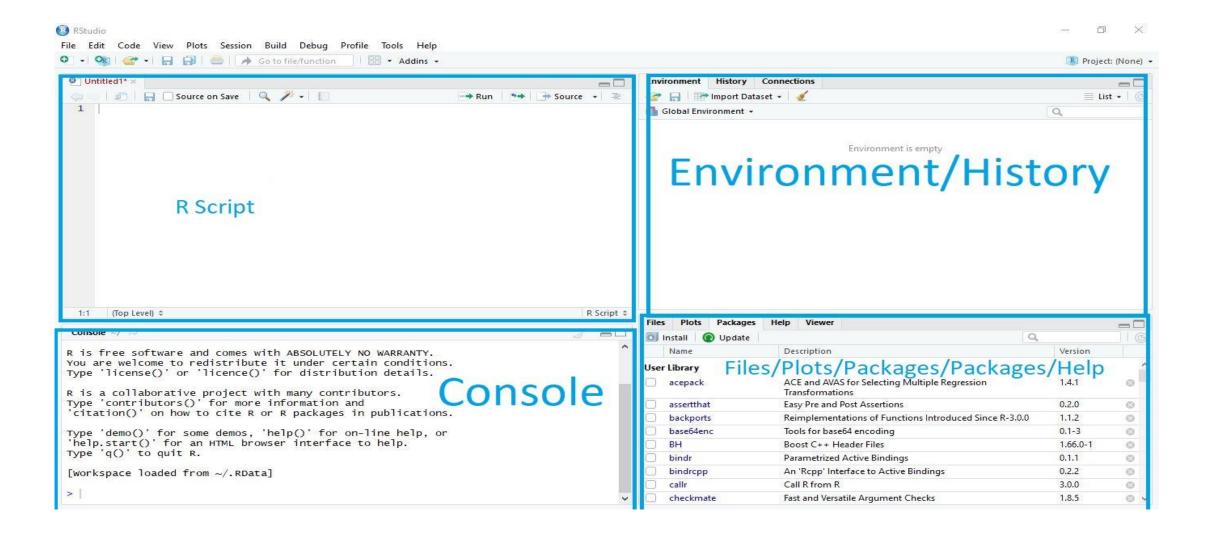- Graphics can be directly saved in a postscript or pdf format

Installing R

- You may install R by downloading it from https://www.r-project.org
- Else download it from the Comprehensive R Archive Network [CRAN] website: http://cran.r-project.org/
- Installing R Studio
  - R Studio is a software which helps in running the R Software
  - Several such editors are available , e.g. Tinn R [https://sourceforge.net/projects/tinn-r]
  - R Studio written in C++ programming language
  - R Studio is a free and open source integrated development environment [IDE] for R
  - Download R Studio software from website https://www.rstudio.com

How to

- Set the working directory
- Create an R file and save it
- Execute an R file
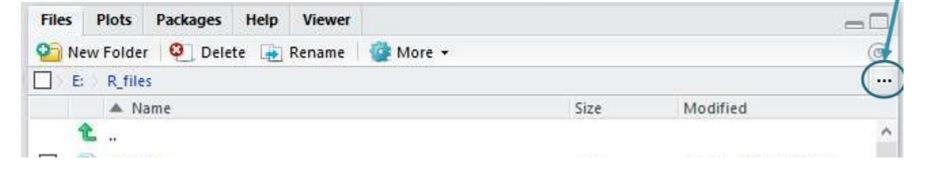- Execute pieces of code

A first look of R Studio

Set the working directory

Enter in console: setwd("directorypath")

Or

Step 1: Choose a suitable location by clicking on the indicated button *To choose working directory*

Set the working directory

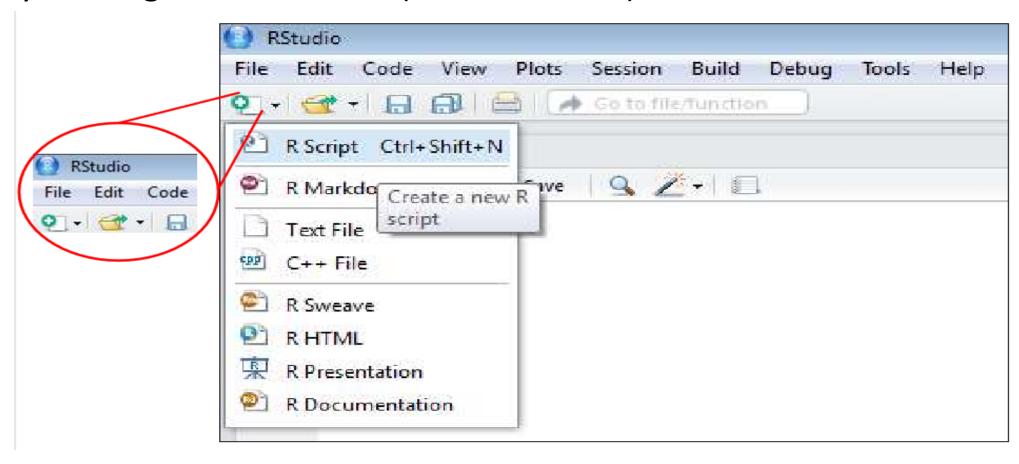Step 2: Once directory is chosen, select the more icon and choose "set as working directory"

Ctrl + Shift + H ➔ Set the working directory

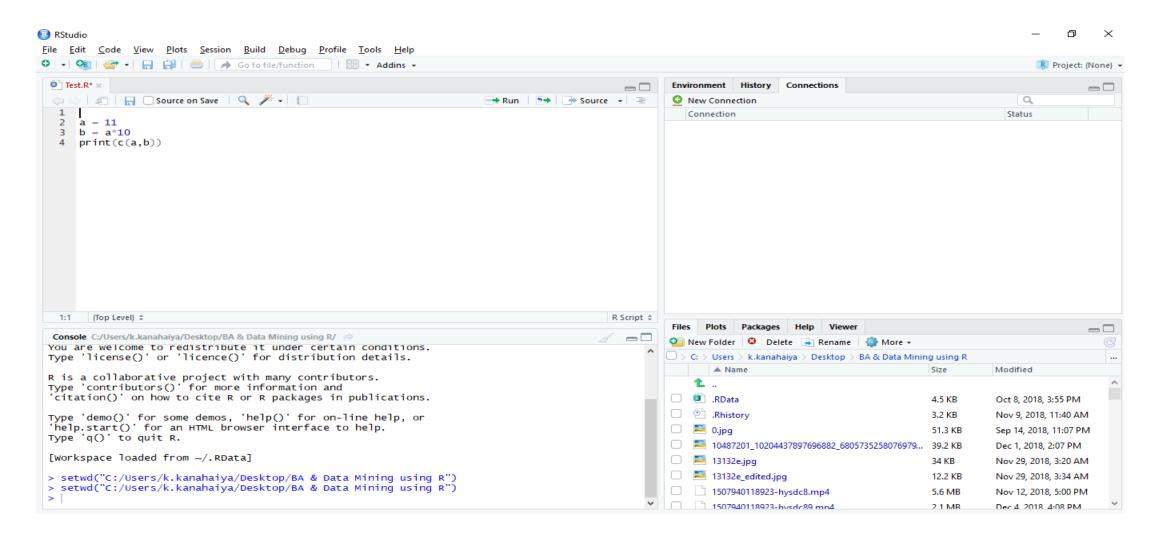Or Click on session tab ➔ Set working directory ➔ Choose directory

Self Study➔ Explore all the tabs/buttons to be familiar with R Studio IDE
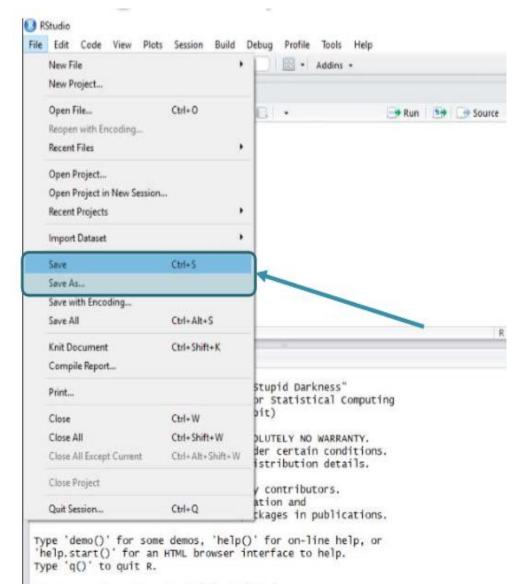
Create an R file
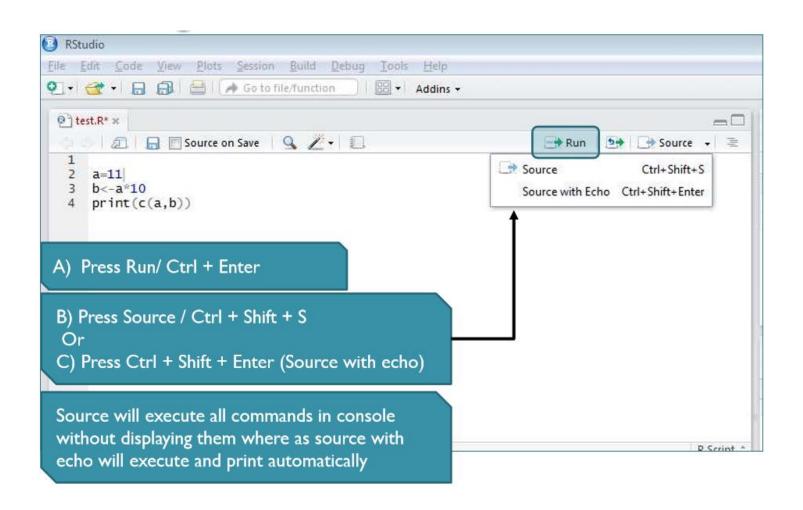
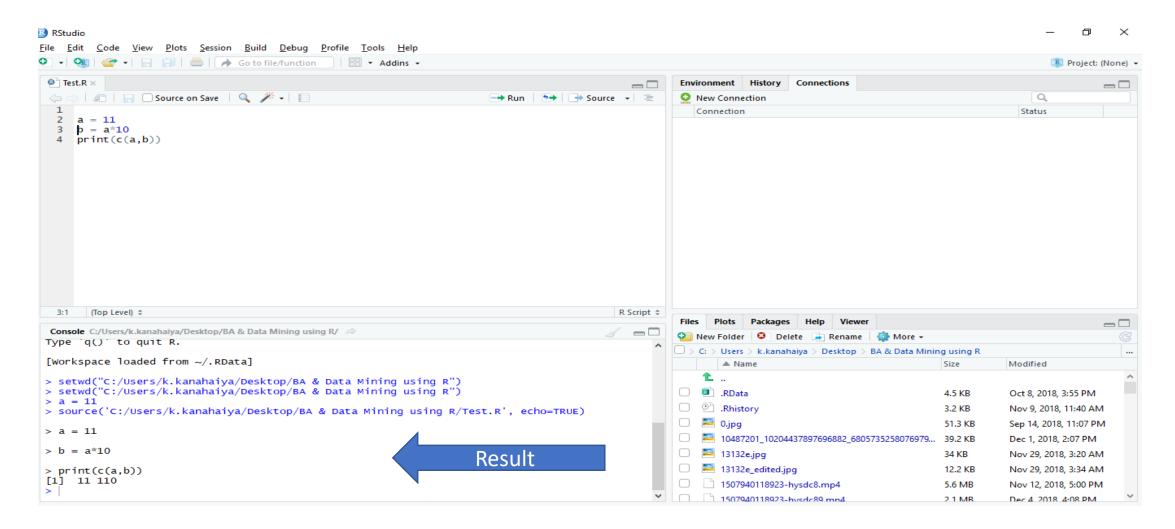By Clicking the "+" Button (Below File Tab)
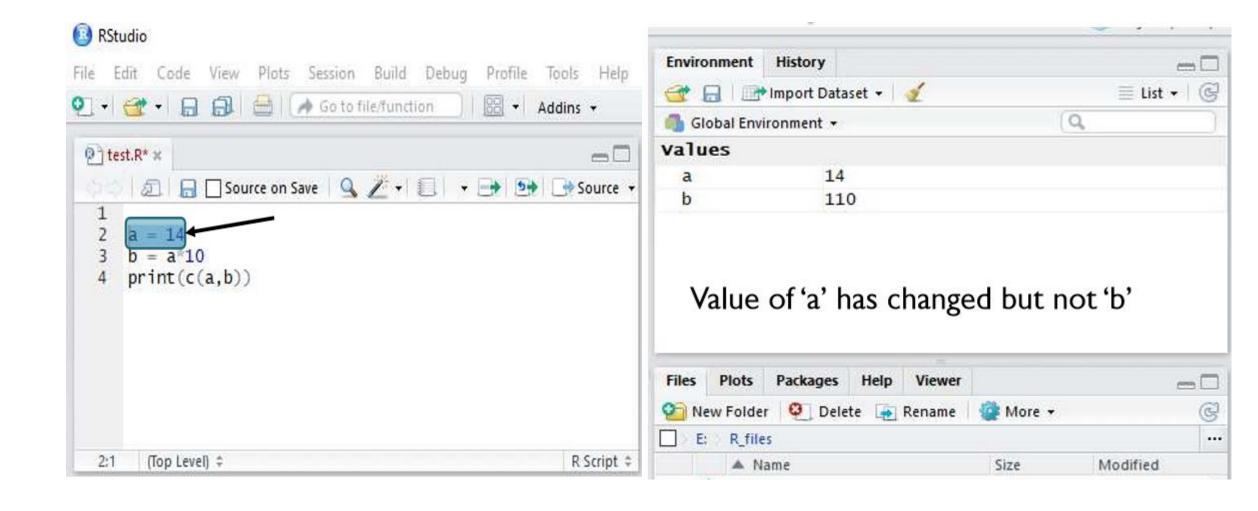
Writing Scripts

Saving R Script File

Executing an R File

Executing an R File: Example

Execute pieces of code - Run

Executing R Files – Summary

- Run can be used to execute selected lines

- Source / Source with echo is for a whole file

- Advantages – Using Run:
  - Troubleshooting /debugging

- Disadvantages – Using Run:
  - For large section, console will be over populated and messy

How to

- Add Comments
  - For single line comment, insert '#' at the start of the line
  - For multiple line comment, select multiple lines using cursor, then press CTRL+Shift+C
  - Or Select multiple lines using cursor, click on "Code" in menu and select "Comment/Uncomment lines"
- Clear the environment
  - To clear console, CTRL+L
  - To clear environment, Single variable: Enter in console/R Script: rm(variable)
  - All variables: Enter in console/R Script: rm(list=ls()

How to

- Saving the workspace
  - Workspace data: Workspace information is temporary
  - Workspace data is not retained after the session
    - If you close the R-Session
    - If you restart the computer