**Practical 5**

**Aim:**

Implementation of the basics of R and data acquisition, Install packages, Loading packages Data types, checking the type of variable, printing variable and objects (Vector, Matrix, List, Factor, Data frame, Table) cbind-ing and rbind-ing, Reading and Writing data. setwd(), getwd(), data(), rm(), Attaching and Detaching data. Reading data from the console. Loading data from different Data sources. (CSV, Excel).

**Theory:**

**Basic R commands**:

R is an open-source programming language that is widely used as a statistical software and data analysis tool. R generally comes with the Command-line interface. R is available across widely used platforms like Windows, Linux, and macOS. Also, the R programming language is the latest cutting-edge tool. It was designed by Ross Ihaka and Robert Gentleman at the University of Auckland, New Zealand, and is currently developed by the R Development Core Team. R programming language is an implementation of the S programming language. It also combines with lexical scoping semantics inspired by Scheme. Moreover, the project conceives in 1992, with an initial version released in 1995 and a stable beta version in 2000.

**Features of R:**

Statistical features:

a. Basic Statistics: The most common basic statistics terms are the mean, mode, and median. These are all known as “Measures of Central Tendency.” So using the R language we can measure central tendency very easily.

b. Static graphics: R is rich with facilities for creating and developing interesting static graphics. R contains functionality for many plot types including graphic maps, mosaic plots, biplots, and the list goes on.

c. Probability distributions: Probability distributions play a vital role in statistics and by using R we can easily handle various types of probability distribution such as Binomial Distribution, Normal Distribution, Chi-squared Distribution, and many more.

d. Data analysis: It provides a large, coherent, and integrated collection of tools for data analysis. Programming features:

a. R Packages: One of the major features of R is it has a wide availability of libraries. R has CRAN(Comprehensive R Archive Network), which is a repository holding more than 10, 0000 packages.

b. Distributed Computing: Distributed computing is a model in which components of a software system are shared among multiple computers to improve efficiency and performance. Two new packages ddR and multiplier used for distributed programming in R were released in November 2015.

1. print() prints out code.

2. To assign a value to a variable, use the <- sign.



setwd(dir) is used to set the working directory to dir.

getwd(): It returns an absolute filepath representing the current working directory of

the R process.

dir() function lists all the files in a directory.

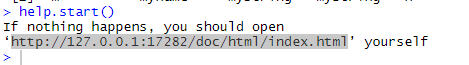


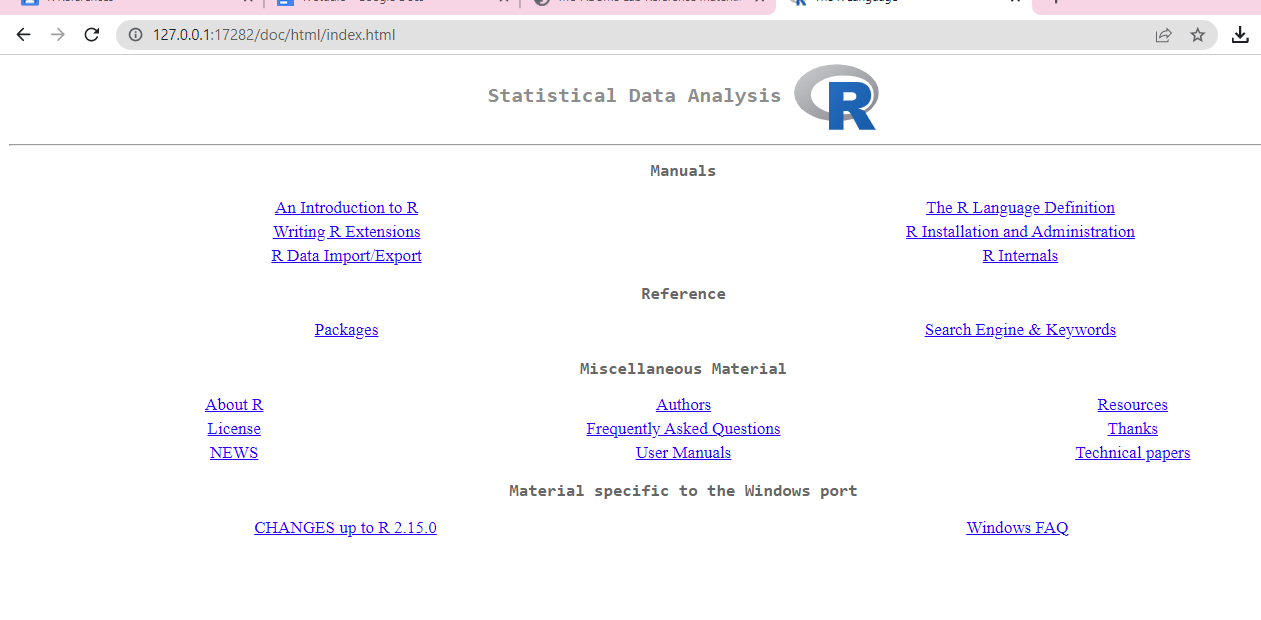


ls() function is used to list the names of all the objects that are present in the working

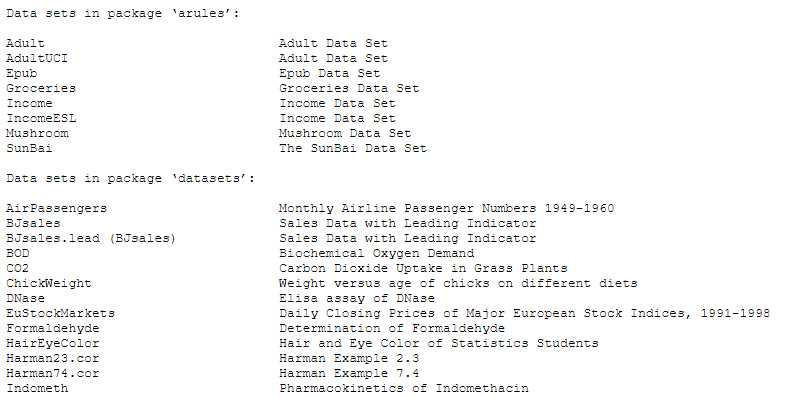
directory.



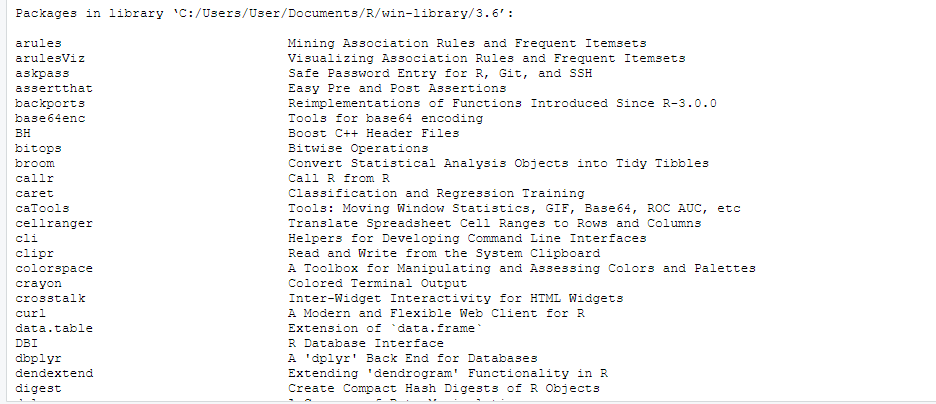


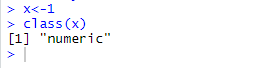






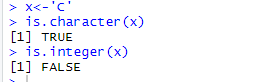








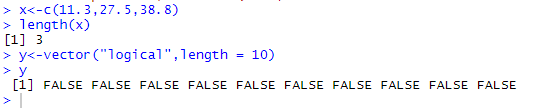






c() function is used to create a vector or list of items.

vector() is used to produce a vector of the given length and mode.

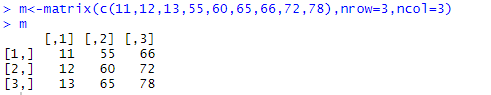






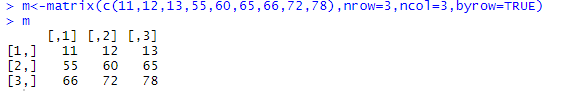


matrix() function used to create matrix of 2-D array having elements of same class.





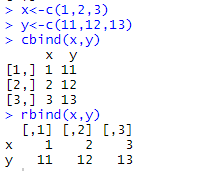


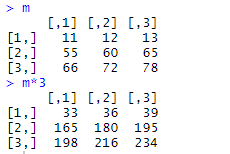


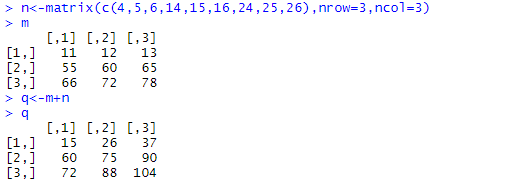
Use the cbind() function to add additional columns in a Matrix.

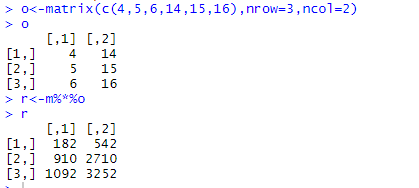
Use the rbind() function to add additional rows in a Matrix.

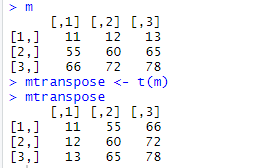
Matrix operations/functions:





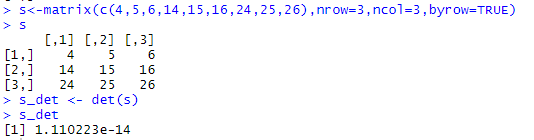




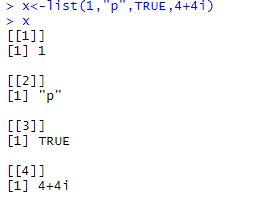


det() function is used to calculate the determinant of the specified matrix.

t() functions is used to transpose the rows and columns.

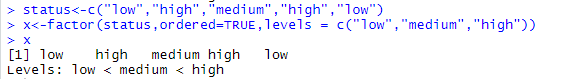


list() is used to create a list.



Factors are used to categorize data.

To create a factor, use the factor() function and add a vector as an argument.



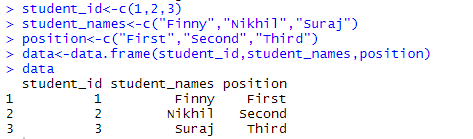
Data Frames are data displayed in a format as a table. Data Frames can have

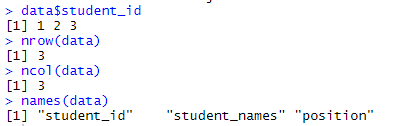
different types of data inside them. While the first column can be numeric, the second

and third can be character or logical. However, each column should have the same

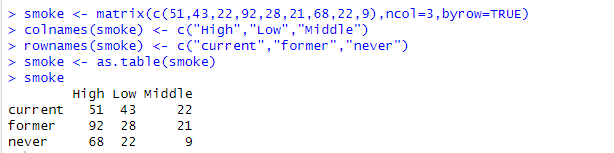
type of data.

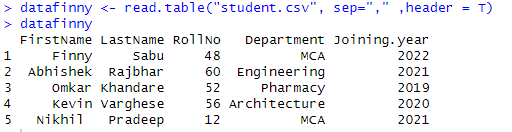
data.frame() is used to create a data frame.

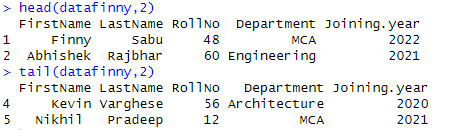




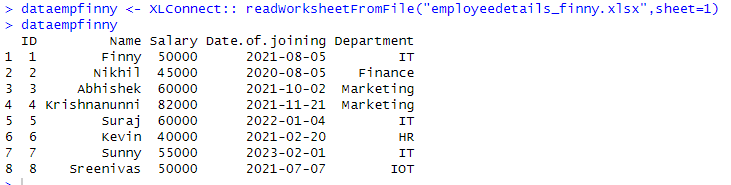
as.table() function is used to convert an object into a table.

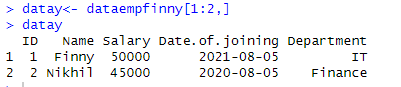






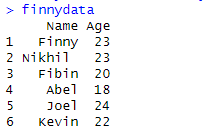
readWorksheetFromFile() is used to reads data from worksheets in an Excel file.











Conclusion:

Successfully implemented the basics R language and data acquisition in R studio