Statistical Visualization using R Lab

By

L. ATRI DATTA RAVI TEZ
Asst. Professor

Course Outcomes

CO1: Employ math and simulation in R [K2]

CO2: Demonstrate various types of data structures in R [K3]

CO3: Apply appropriate control structures to solve a particular Programming problem [K3]

CO4: Use R to graphically visualize data and results of statistical calculations [K3]

List of Experiments

1. Demonstrate the basic math functions in R

2. Demonstrate Vector operations in R

3. Demonstrate Matrix operations in R

4. Demonstrate Array operations in R

5. Demonstrate Dataframes in R

6. Demonstrate Lists in R

List of Experiments

7. Illustrate the following controls statements in R
   a) if and else
   b) ifelse
   c) switch

8. Demonstrate for and while loops in R

9. Demonstrate importing and exporting data using R

10. Illustrate the descriptive statistics using summary() in R

List of Experiments

11. Demonstrate the following statistical distribution functions in R:

   a) Normal Distribution
   b) Binomial Distribution
   c) Poisson Distribution
   d) Chi Square Distribution

12. Illustrate the following basic graphics in R:
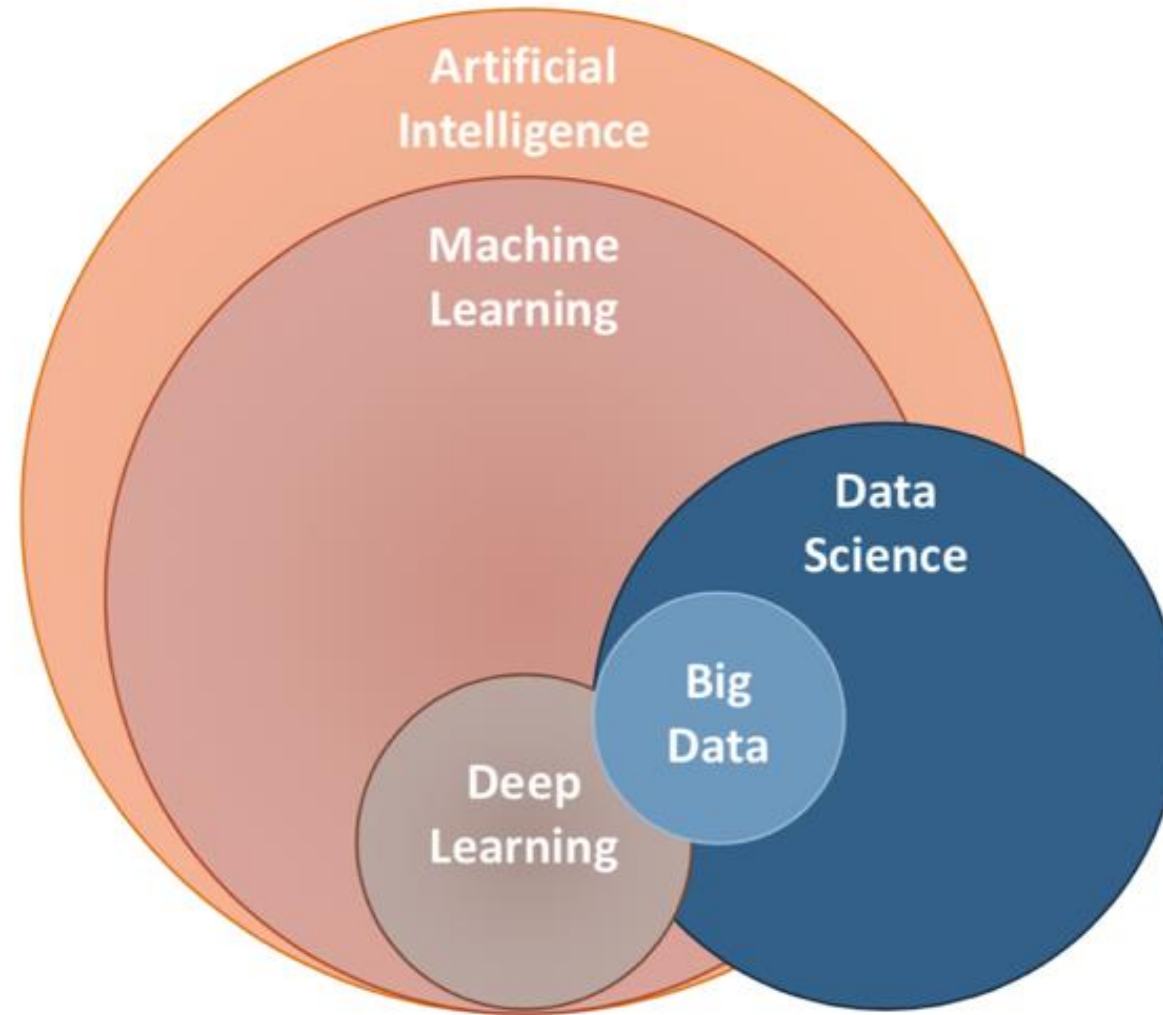
   a) Bar plots
   b) Pie Charts
   c) Histograms
   d) Kernel density plots
   e) Boxplots
   f) Dotplots

List of Experiments

13. Illustrate the Correlation and Covariance analysis using R

14. Illustrate the different types of t-tests using R

15. Illustrate the ANOVA test using R

Introduction to R
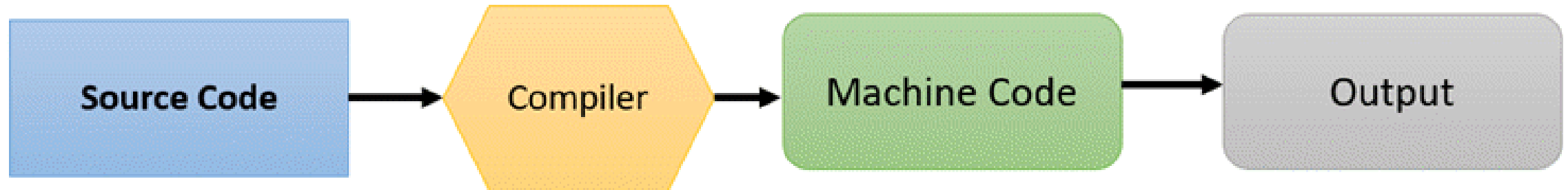
AI vs. ML vs. DL vs. DS vs. Big Data

R Language

- R is an open source programming language and software environment for statistical computing and graphics.
- The R language is widely used among statisticians and data miners for developing statistical software and data analytics tools
- Modelled after S & S-plus, developed at AT&T labs in late 1980s.
- R project was started by Robert Gentleman and Ross Ihaka Department of Statistics, University of Auckland (1995).
- Currently maintained by R core development team – an international team of volunteer developers (since 1997).

Features of R

- ➢ R is an interpreted language; users typically access it through a command-line interpreter.
- ➢ R is a well-developed, simple and effective programming language which includes conditionals, loops, user defined recursive functions and input and output facilities.
- ➢ R has an effective data handling and storage facility.
- ➢ R provides a suite of operators for calculations on arrays, lists, vectors and matrices.
- ➢ R provides a large, coherent and integrated collection of tools for data analysis.
- ➢ R provides graphical facilities for data analysis and display either directly at the computer or printing at the papers.

How Compiler Works

Source Code	→	Compiler	→	Machine Code	→	Output

How Interpreter Works

Source Code	→	Interpreter	→	Output

R Resources

- http://www.r-project.org/
- http://cran.r-project.org/doc/contrib/Verzani-SimpleR.pdf
- Download R :

    http://cran.r-project.org/bin/
- Download RStudio :

    http://www.rstudio.com/ide/download/desktop

R Installation

Installing R on windows PC :

➢ Use internet browser to point to : http://mirror.aarnet.edu.au/pub/CRAN

➢ Under the heading Precompiled Binary Distributions, choose the link Windows.

➢ Next heading is R for Windows; choose the link base.

➢ Click on download option(R 3.4.1 for windows).

➢ Save this to the folder C:\R on your PC.

➢ When downloading is complete, close or minimize the Internet browser.
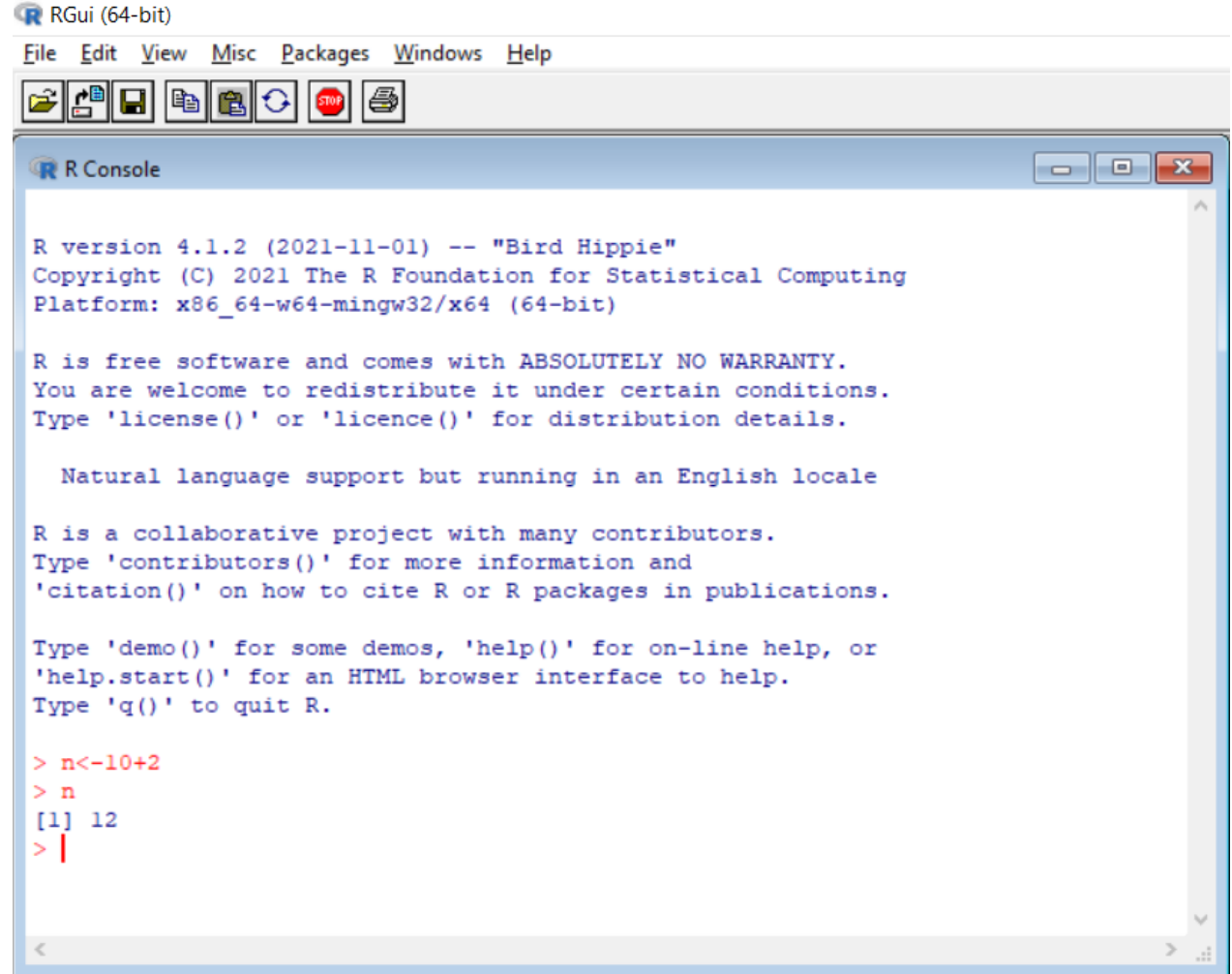
➢ Double click on R 3.4.1-win32.exe in C:\R to install.

Installing R on Linux:

➢   sudo apt-get install r-base-core

R-Studio Installation

➤ Go to www.rstudio.com and click on the "Download RStudio" button.

➤ Click on "Download RStudio Desktop."

➤ Click on the version recommended for your system, or the latest Windows version, and save the executable file. Run the .exe file and follow the installation instructions.

A test run with R in Windows

▸ Double click the R icon on the Desktop and the R Console will open.

▸ Wait while the program loads. You observe something like this.

▸ You can type your own program at the prompt line >

```
R RGui (64-bit)
File  Edit  View  Misc  Packages  Windows  Help

R R Console

R version 4.1.2 (2021-11-01) -- "Bird Hippie"
Copyright (C) 2021 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.

> n<-10+2
> n
[1] 12
>
```
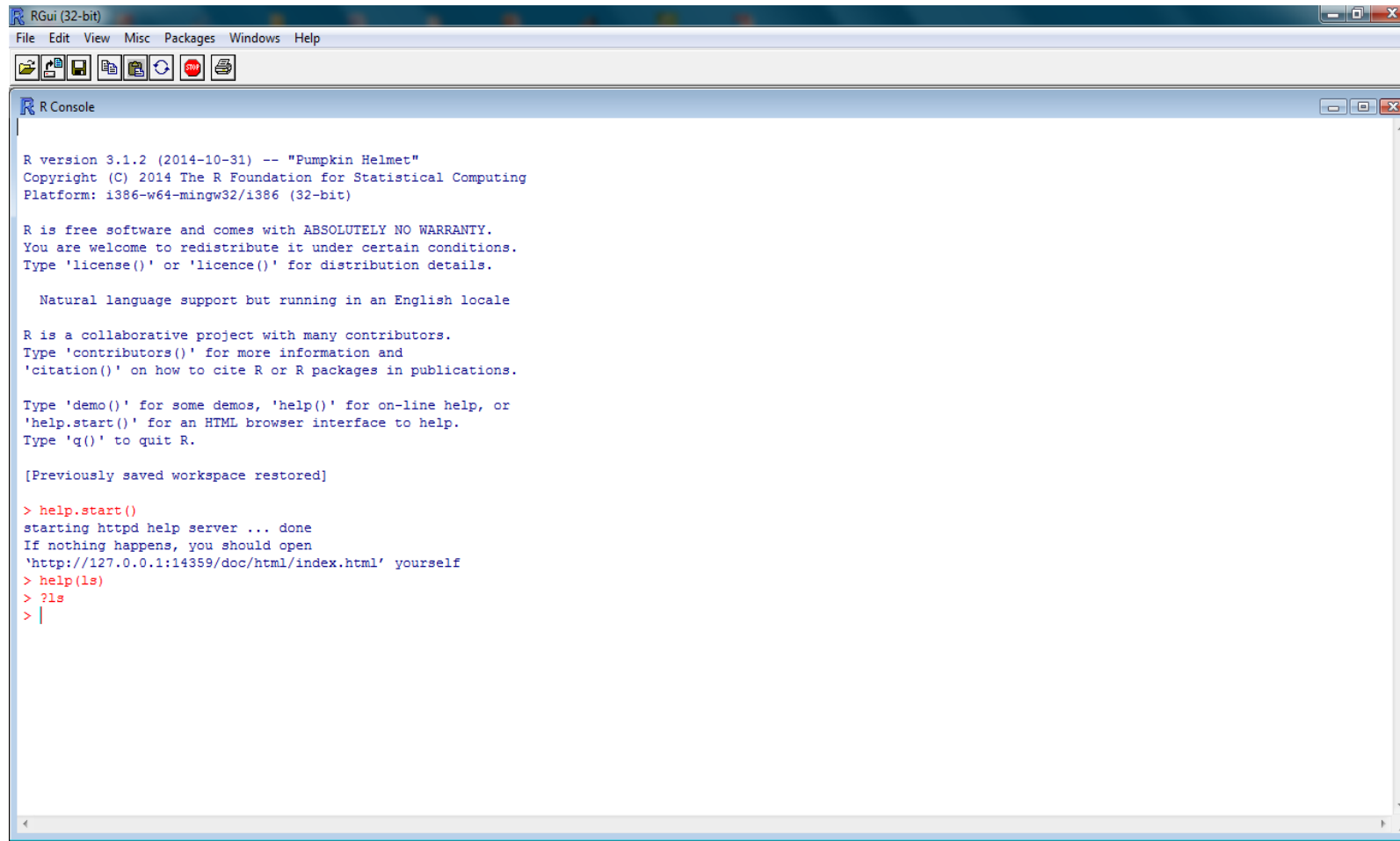
Getting help from R console

- help.start()
- help(topic)
- ?topic
- ??topic

16

R commands in integrated environment

Exp. 1: Demonstrate the basic math functions in R

Aim: To understand about operators, variables, built-in constants, built-in functions in R and how to execute basic mathematical operations in R.

▶ **Operators:** An operator is a symbol that tells the compiler to perform specific mathematical or logical manipulations. R language is rich in built-in operators and provides following types of operators.

Exp. 1: Demonstrate the basic math functions in R

- # Addition
  > 7+3
  [1]   10
- # Subtraction
  > 7-3
  [1]   4
- # Multiplication
  > 8*7
  [1]   56

- #Division
  > 100/50
  [1]   2
- # Square root
  > sqrt(81)
  [1]   9
- # Exponents
  > 9^2
  [1]   81

Exp. 1: Demonstrate the basic math functions in R

Basic R Operators

Operation	Description
x + y	Addition
x - y	Subtraction
x * y	Multiplication
x / y	Division
x ^ y	Exponentiation
x %% y	Modular arithmetic
x %/% y	Integer division

Exp. 1: Demonstrate the basic math functions in R

Basic R Operators

Operation	Description
x == y	Test for equality
x <= y	Test for less than or equal to
x >= y	Test for greater than or equal to
x && y	Boolean AND for scalars
x || y	Boolean OR for scalars
x & y	Boolean AND for vectors
x | y	Boolean OR for vectors
!x	Boolean negation

Exp. 1: Demonstrate the basic math functions in R

Operator Type	Operators
Arithmetic Operators	+, -, *, /, %%, %/%, ^
Relational Operators	<, >, ==, <=, >=, !=
Logical Operators	&, |, !, &&, ||
Assignment Operators	= or <- or <<- & ->, ->>
Miscellaneous Operators	:, %in%, %*%

Exp. 1: Demonstrate the basic math functions in R

▸ Arithmetic Operators:

> 3+5   ##8

> 12 + 3 / 4 − 5 + 3*8   ##31.75

> (12 + 3)/ 4 − 5 + 3*8   ##22.75

> ((12+3)/(4 − 5) +3)*8   ##-96

Exp. 1: Demonstrate the basic math functions in R

◦ Order of Basic Mathematical Operations:

1. Brackets	4. Multiplication
2. Orders (Powers, Roots)	5. Addition
3. Division	6. Subtraction

> 256%%13   ##9

> 256%/%13   ##19

>6^3   ##216

Exp. 1: Demonstrate the basic math functions in R

- Relational Operators:

> 3<5	##TRUE
> 6>9	##FALSE
> 5==6	##FALSE
> 12!=52	##TRUE
>5<=6	##TRUE
>125>=50	##TRUE

Exp. 1: Demonstrate the basic math functions in R

- Logical Operators:
  ```
  > TRUE&TRUE                    ##TRUE
  > TRUE&FALSE                   ##FALSE
  > FALSE&TRUE                   ##FALSE
  > FALSE&FALSE                  ##FALSE
  > TRUE|TRUE                    ##TRUE
  > TRUE|FALSE                   ## TRUE
  > FALSE|TRUE                   ## TRUE
  > FALSE|FALSE                  ##FALSE
  >!TRUE                         ##FALSE
  ```

Exp. 1: Demonstrate the basic math functions in R

Variables

▶ A variable provides us with named storage that our programs can manipulate.

▶ A variable in R can store an atomic vector, group of atomic vectors or a combination of many R objects.

▶ A valid variable name consists of letters, numbers and the dot or underline characters

▶ The variable name starts with a letter or the dot not followed by a number.

Exp. 1: Demonstrate the basic math functions in R

- Assignment Operators:
  > x=10
  >y<-15
  >18->z

Note: Miscellaneous Operators will be discussed later

Exp. 1: Demonstrate the basic math functions in R

▸ **Built-in Constants**

>LETTERS   ##the 26 upper-case letters of the Roman alphabet;

>letters   ##the 26 lower-case letters of the Roman alphabet;

>month.abb   ##the three-letter abbreviations for the English month names;

>month.name   ##the English names for the months of the year;

>pi   ##the ratio of the circumference of a circle to its diameter.

Exp. 1: Demonstrate the basic math functions in R

▸ **<u>Basic Math Functions</u>**

  ◦ R includes an extensive set of built-in math functions.

Function	Description
abs(x)	absolute value
sqrt(x)	square root
ceiling(x)	ceiling(3.475) is 4
floor(x)	floor(3.475) is 3
trunc(x)	trunc(5.99) is 5
round(x, digits=n)	round(3.475, digits=2) is 3.48
signif(x, digits=n)	signif(3.475, digits=2) is 3.5
cos(x), sin(x), tan(x)	also acos(x), cosh(x), acosh(x), etc.
log(x)	natural logarithm
log10(x)	common logarithm
exp(x)	e^x

▸ Built-in Functions (Math)

```
>abs(-4)            ##4
>sqrt(25)           ##5
>ceiling(5.7)       ##6
>ceiling(4.2)       ##5
```

Exp. 1: Demonstrate the basic math functions in R

>floor(3.3)   ##3
>floor(10.75)   ##10
>trunc(2.5)   ##2
>round(3.475, digits=2)   ##3.48

Exp. 1: Demonstrate the basic math functions in R

> log(2,10)   ##0.30103
> log(2, base=10)   ##0.30103
> log10(2)   ##0.30103
> log(2)   ##0.6931472

Exp. 1: Demonstrate the basic math functions in R

>exp(4)   ##54.59815
>cos(pi)   ##-1
>sin(pi)   ##1.224606e-16~0
>tan(pi/4)   ##1

TASK

▸ Find the area and circumference of a circle given the radius=6

Practice Tasks

Solve:

- $\sqrt{78}$
- |-56| =56
- 3.49 should be printed as 4
- What are two ways of printing 5.14 as 5?
- 3.12321 should be printed as 3.12
- Log(2.3) with base 2
- e^5
- cos(90)
- Display Month names viz. January, February
- Calculate and display Area of a rectangle given length=8 and breadth=3