NAME – ANSH GOEL

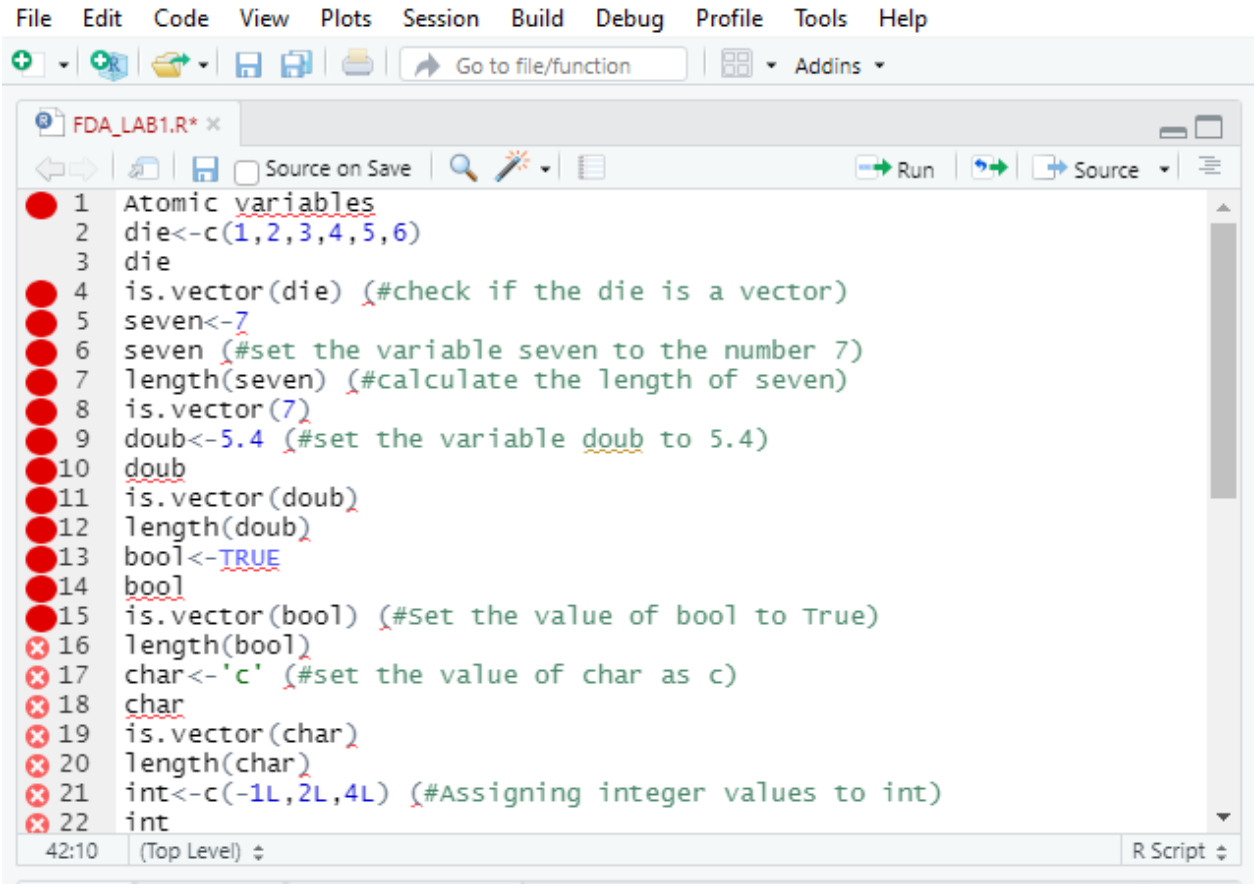
REG NO. – 20BCE1798

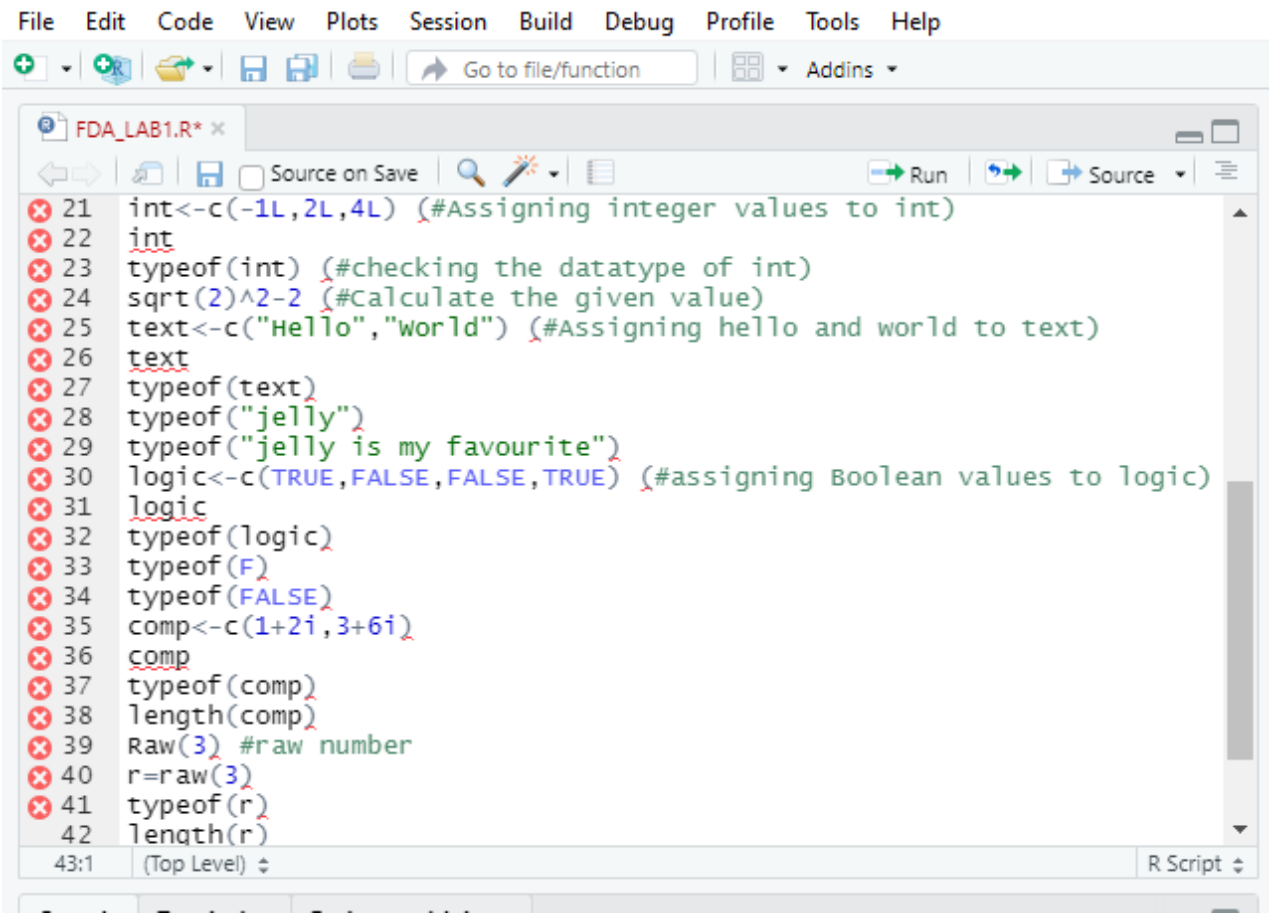
COURSE NAME – FOUNDATION OF DATA ANALYTICS (FDA)

COURSE CODE: 3505

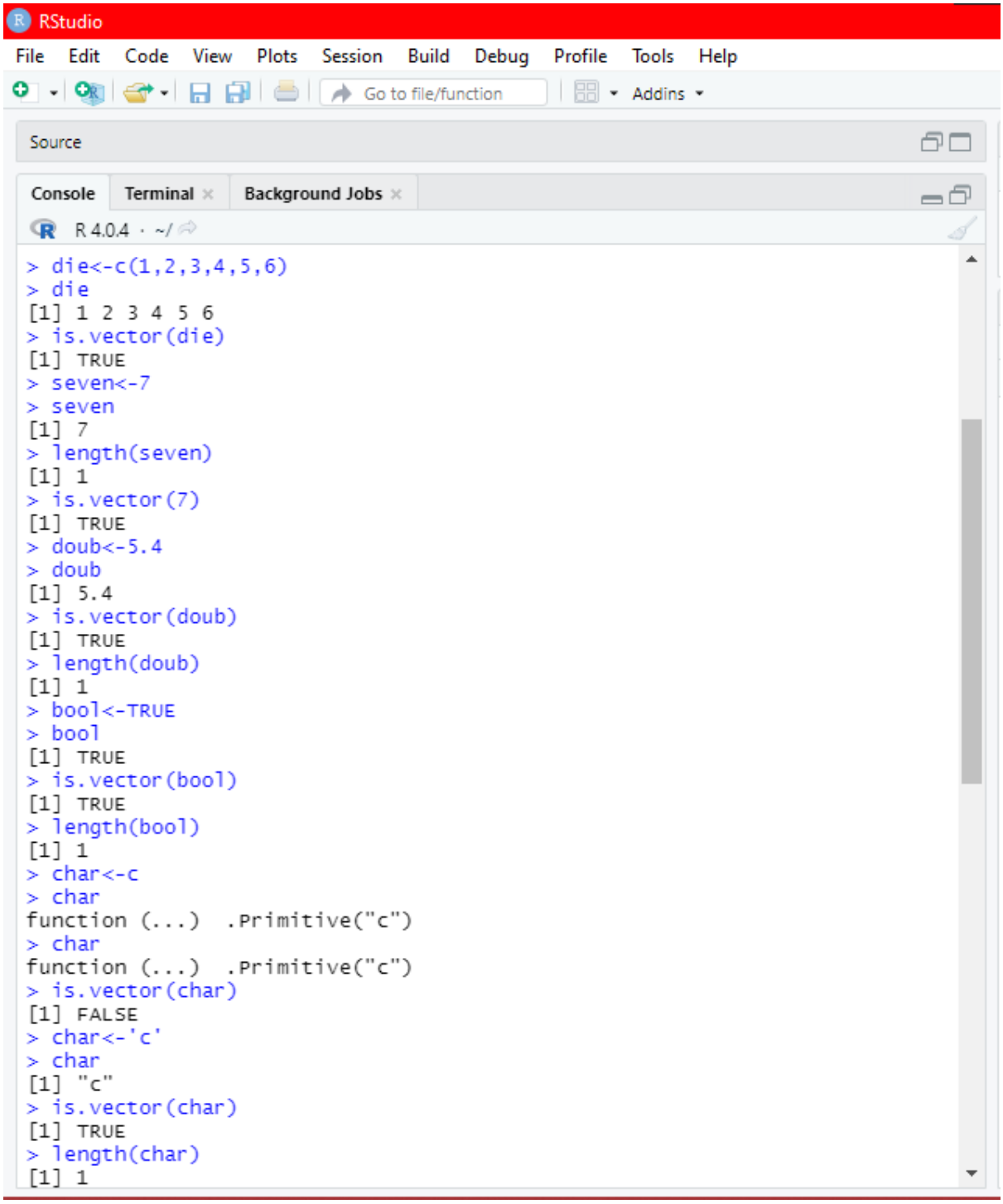
DATE – 21st JULY, 2022

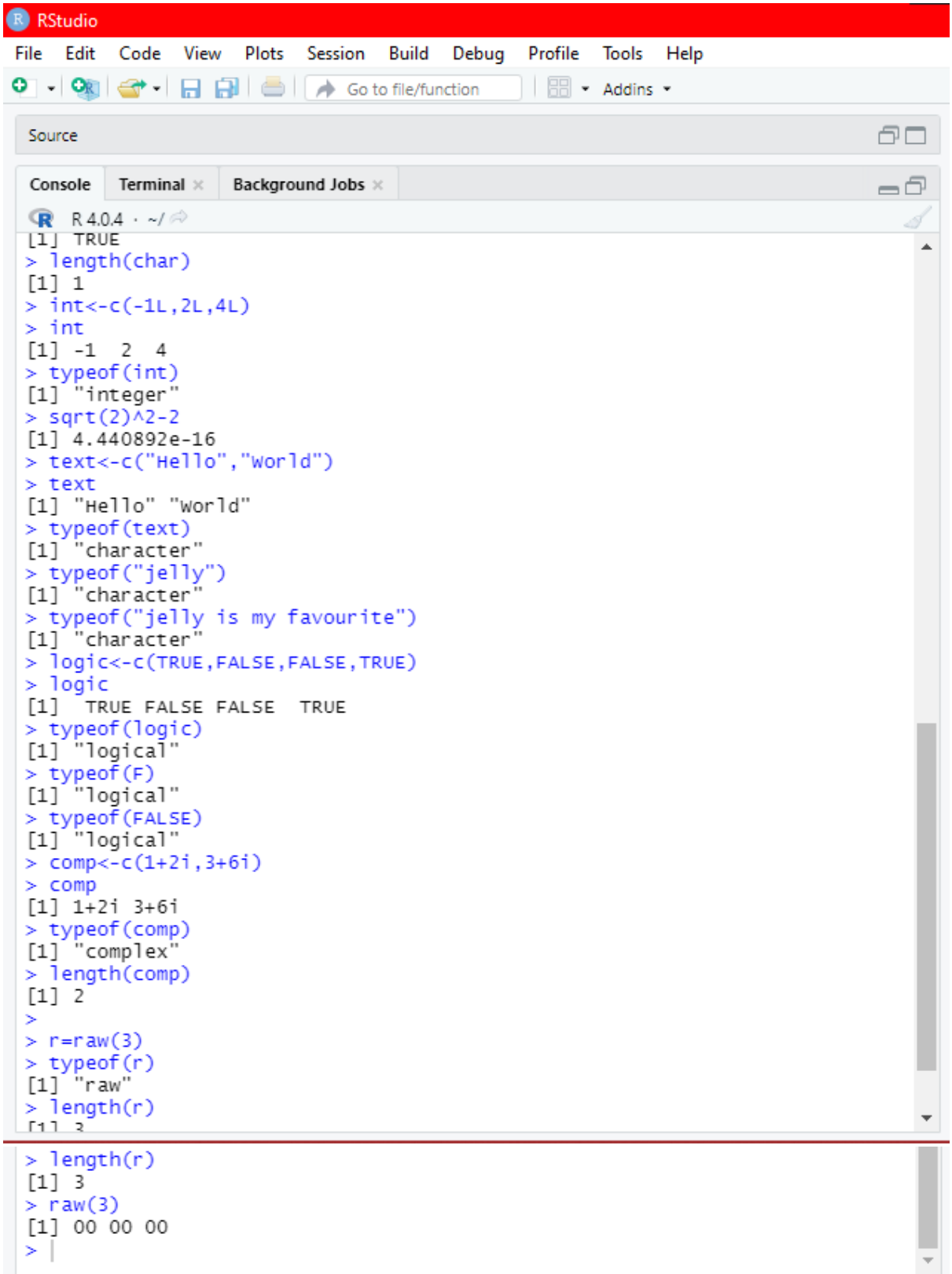
LAB 1 – BASIC INTRODUCTION TO R STUDIO




File Edit Code View Plots Session Build Debug Profile Tools Help
Go to file/function
Addins
FDA_LAB1.R*
Source on Save
Run
Source
int<-c(-1L,2L,4L) (#Assigning integer values to int)
int
typeof(int) (#checking the datatype of int)
sqrt(2)^2-2 (#Calculate the given value)
text<-c("Hello","World") (#Assigning hello and world to text)
text
typeof(text)
typeof("jelly")
typeof("jelly is my favourite")
logic<-c(TRUE,FALSE,FALSE,TRUE) (#assigning Boolean values to logic)
logic
typeof(logic)
typeof(F)
typeof(FALSE)
comp<-c(1+2i,3+6i)
comp
typeof(comp)
length(comp)
Raw(3) #raw number
r=raw(3)
typeof(r)
length(r)
43:1
(Top Level)
R Script

OUTPUT:

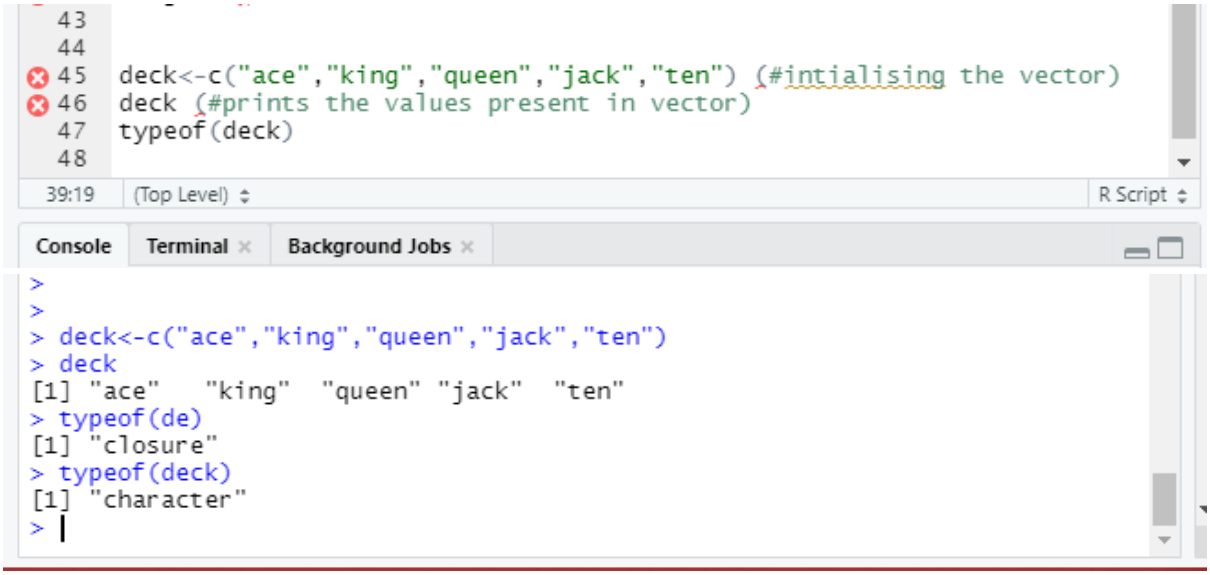

RStudio
File Edit Code View Plots Session Build Debug Profile Tools Help
Go to file/function
Addins
Source
Console
Terminal
Background Jobs
R 4.0.4 · ~/
> die<-c(1,2,3,4,5,6)
> die
[1] 1 2 3 4 5 6
> is.vector(die)
[1] TRUE
> seven<-7
> seven
[1] 7
> length(seven)
[1] 1
> is.vector(7)
[1] TRUE
> doub<-5.4
> doub
[1] 5.4
> is.vector(doub)
[1] TRUE
> length(doub)
[1] 1
> bool<-TRUE
> bool
[1] TRUE
> is.vector(bool)
[1] TRUE
> length(bool)
[1] 1
> char<-c
> char
function (...) .Primitive("c")
> char
function (...) .Primitive("c")
> is.vector(char)
[1] FALSE
> char<-'c'
> char
[1] "c"
> is.vector(char)
[1] TRUE
> length(char)
[1] 1


RStudio
File Edit Code View Plots Session Build Debug Profile Tools Help
Go to file/function
Addins
Source
Console
Terminal
Background Jobs
R 4.0.4 · ~/
[1] TRUE
> length(char)
[1] 1
> int<-c(-1L,2L,4L)
> int
[1] -1 2 4
> typeof(int)
[1] "integer"
> sqrt(2)^2-2
[1] 4.440892e-16
> text<-c("Hello","World")
> text
[1] "Hello" "World"
> typeof(text)
[1] "character"
> typeof("jelly")
[1] "character"
> typeof("jelly is my favourite")
[1] "character"
> logic<-c(TRUE,FALSE,FALSE,TRUE)
> logic
[1] TRUE FALSE FALSE TRUE
> typeof(logic)
[1] "logical"
> typeof(F)
[1] "logical"
> typeof(FALSE)
[1] "logical"
> comp<-c(1+2i,3+6i)
> comp
[1] 1+2i 3+6i
> typeof(comp)
[1] "complex"
> length(comp)
[1] 2
>
> r=raw(3)
> typeof(r)
[1] "raw"
> length(r)
[1] 3
> length(r)
[1] 3
> raw(3)
[1] 00 00 00
>

Q1)

Create an atomic vector that stores the name of the face of cards in a royal flush. For example, the ace of spades, king of spades, queen of spades, jack of spades and ten of spades. The face name of the ace of spades would be ace and spades is the suit.

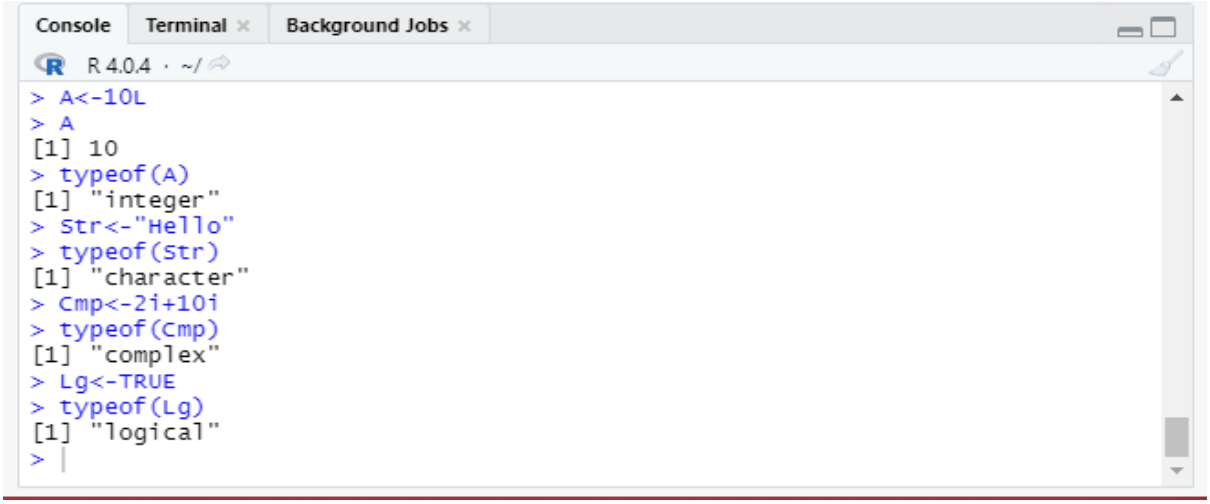
Answer:



Q2)

An R program is provided below. Identify and print the types of the R objects in the given code.

Ans.



Q3)
What is the difference between vector and list in R? Give examples for each.
Ans.

Data of all kinds, including numeric, character, logical, integer, double, and more, are stored in lists.

The same type of element is stored in a vector.

A vector age example is c (20, 25, 30) # A vector cannot contain more than one datatype.

List(name="ansh" gender=”M” age="18," company="abc") as an example. Lists may contain a variety of datatypes.

Q4)

Write R code to perform arithmetic operation, relational operation, logical operation and assignment operation using R.

Ans.

```
62
63  Arithmetic:
64  2+3
65  9*4
66  97-55
67  7865/5
68
```

Console | Terminal | Background Jobs

R 4.0.4 · ~/

```
> 2+3
[1] 5
> 9*4
[1] 36
> 97-55
[1] 42
> 7865/5
[1] 1573
```

```
69  Relational:
70  2>3 #greater than
71  3<2 #less than
72  4*1==5 #Checking Equlaity
73  6==9
74  8==8
75
```

```
> 2>3
[1] FALSE
> 3<2
[1] FALSE
> 4*1==5
[1] FALSE
> 6==9
[1] FALSE
> 8==8
[1] TRUE
>
```

```
76  Logical:
77  1&&0 #AND operator
78  1||1 #OR operator
79  !FALSE #NOT operator
80
```

```
>
> 1&&0
[1] FALSE
> 1||1
[1] TRUE
> !FALSE
[1] TRUE
>
```

```
81  Assignment:
82  x<-8 #left assignment
83  x
84  y=26 #argument assignment
85  y
86  w<<-5 #Left lexicographic assignment
87  w
88  5->>z #Right lexicographic assignment
89  z
```

```
>
>
> x<-8
> x
[1] 8
> y=26
> y
[1] 26
> w<<-5
> w
[1] 5
> 5->>z
> z
[1] 5
```

Q5)
List a minimum of 20 different packages and its usage in R.

Ans.

1. ggplot2

It is possible to create graphs with one, two, or three variables, as well as category and numerical data. Additionally, grouping can be accomplished using a symbol, size, colour, etc.

2. data.table

The fastest package for handling a large quantity of data while manipulating data is data.table. Predictive analytics is mostly employed in the health care industry for genomic data and other industries like business. Additionally, the data sizes range from 10GB to 100GB.

3. dplyr

dplyr is the package which is used for data manipulation by providing different sets of verbs like select(), arrange(), filter(), summarise(), and mutate().

4. tidyr

To make tidy data, use tidyr. The bulk of the labour is spent organising and cleaning up the data. Basically, datasets with clean data are ones in which every cell represents a single value, every row represents an observation, and every column represents a variable.

5. Shiny

Shiny can be used to build web applications without requiring JavaScript. It can be used together with html widgets, JavaScript actions, and CSS themes to have extended features. It can also be used to build dashboards along with the standalone web applications.

6. plotly

plotly is the graphing library used to create graphs that are interactive and can also be used with JavaScript known as plotly.js.

7. knitr

knitr is the package mostly used for research. It is reproducible, used for report creation, and integrates with various types of code structures like LaTeX, HTML, Markdown, LyX, etc.
8. mlr3

mlr3 package is created for doing Machine Learning. Additionally, it supports Object-Oriented programming and provides "R6" objects together with machine learning workflow, making it efficient. It is also regarded as one of the expandable frameworks for survival analysis, clustering, regression, and classification.

9. XGBoost

XGBoost is an implementation of the gradient boosting framework. Additionally, it offers an interface for R that includes the model from R's caret package. Its implementation in H20, Spark, and Python is slower than its speed and performance. Machine learning tasks like classification, ranking issues, and regression are the main use cases for this software.

10. Caret

A caret package is a short form of Classification And Regression Training used for predictive modeling

11 XLConnect, xlsx - These packages help you read and write Micorsoft Excel files from R. You can also just export your spreadsheets from Excel as .csv's.

12. randomForest - Random forest methods from machine learning

13. foreign - Foreign provides functions that help you load data files from other programs into R.

14. haven - Enables R to read and write data from SAS, SPSS, and Stata.

15. tidyverse - An opinionated collection of R packages designed for data science that share an underlying design philosophy, grammar, and data structures. This collection includes all the packages in this section, plus many more for data import, tidying, and visualization listed here.

16. dplyr - Essential shortcuts for subsetting, summarizing, rearranging, and joining together data sets. It is the go to package for fast data manipulation.

17. stringr - Easy to learn tools for regular expressions and character strings.

18. lubridate - Tools that make working with dates and times easier.

19. ggvis - Interactive, web based graphics built with the grammar of graphics.

20. rgl - Interactive 3D visualizations with R