

**CSE 240 Data Science with R**

**STUDENT WORK BOOK**

|  |  |  |
| --- | --- | --- |
| **Name** | **:** | Aparna.S |
| **Unique ID** | **:** | E0119020 |
| **Year** | **:** | II |
| **Quarter** | **:** | Q6 |
| **Department** | **:** | B.Tech CSE (CyS & IoT or AI &ML) |
| **Faculty Name** | **:** | Prof.B.Nirmala or Prof.N.Chiranjeevi |
| **Academic Year** | **:** | 2020-2021 |

**Date: 02.11.2020**

**Q1: Consider 2 vectors c(9,10,11,12) and c(13,14,15,16). Create a 4 by 2 matrix from these two vectors.**

**Program:**

a<-c(9,10,11,12)

b<-c(13,14,15,16)

#matrix 4x2

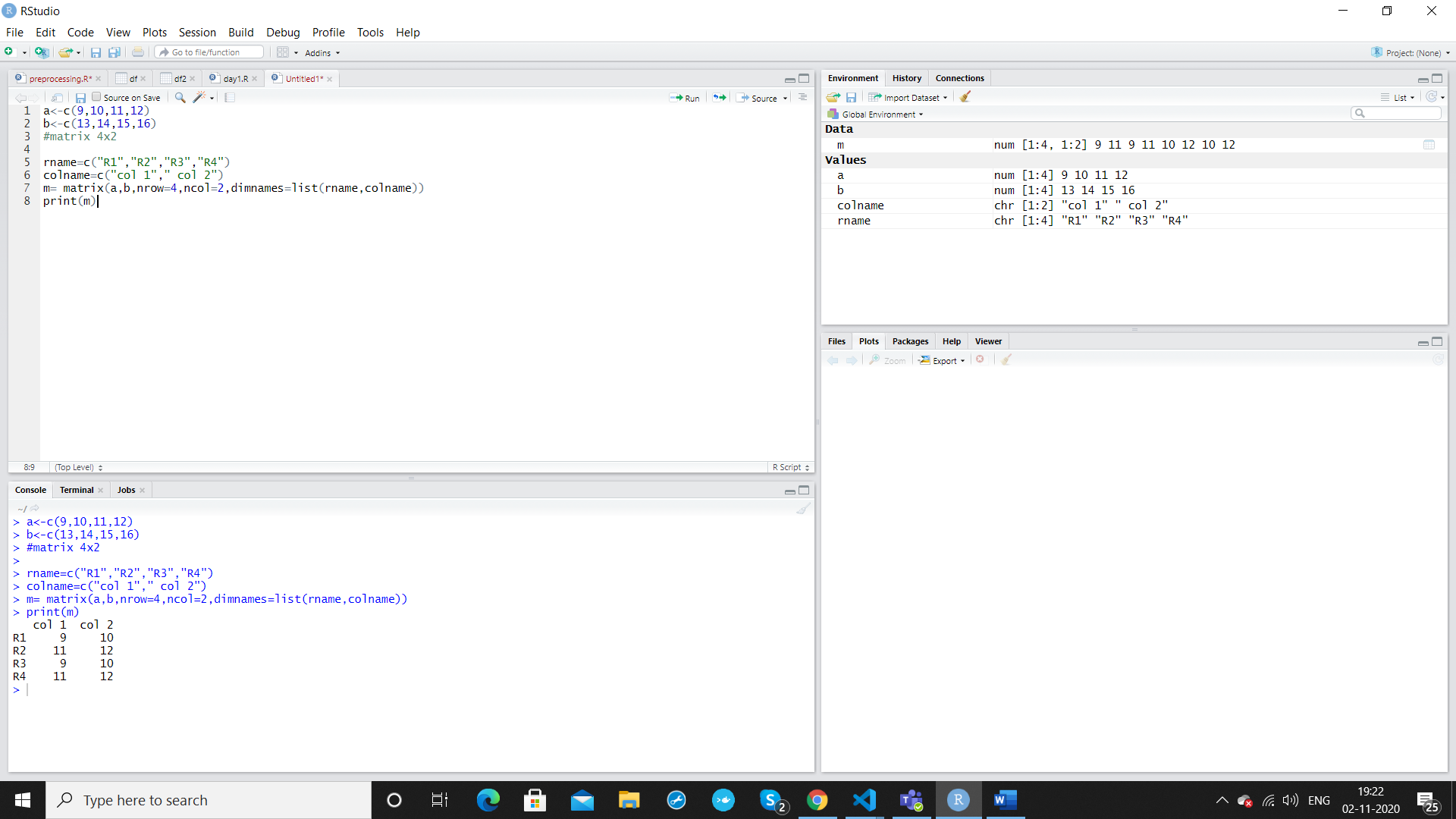
rname=c("R1","R2","R3","R4") # row labels

colname=c("col 1"," col 2") # column labels

m= matrix(a,b,nrow=4,ncol=2,dimnames=list(rname,colname))

print(m)

**Output:**



**Explanation: Concept or Program**

Vectors can be represented as any of the 6 types of atomic datatypes . It is declared as:

Eg: Vector <- c(1,2,3,4) (= can also be used instead of <-)

charV<-c(“a”,”b”,”c”)

Matrices are data objects in R and can be represented using the matrix() function. It contains following attributes : matrix(data, nrow=, ncol=, dimnames=,byrow=TRUE) (NOTE: by default the values are filled column wise ncol=TRUE) data – input vector, nrow- no of rows, ncol- no of columns, dimnames- labels or names for the rows and columns specified),byrow=TRUE, Boolean if true, the elements are arranged by row, otherwise by column.

**Q2:** **Write an R program to take input from the user (userID and Group/Branch) and display the values.**

**Program:**

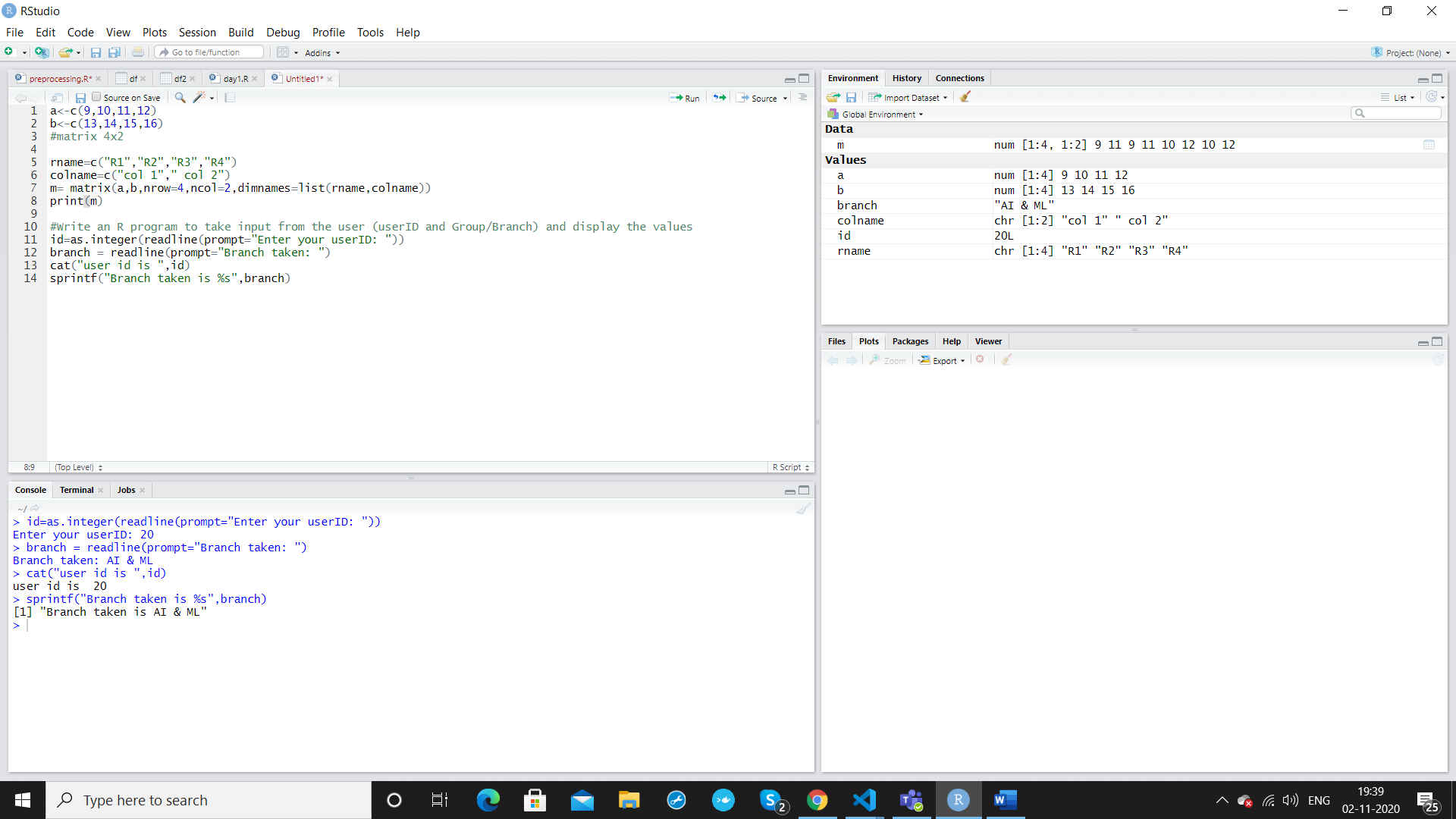
id=as.integer(readline(prompt="Enter your userID: "))

branch = readline(prompt="Branch taken: ")

cat("user id is ",id)

sprintf("Branch taken is %s",branch)

**Output:**



**Explanation: Concept or Program**

as.integer converts the input value as integer type .

To get user input, readline() with prompt attribute can be used

To print the details along with a string message cat() function or sprint() can be used.

Cat() contenates the string with the value passed and sprintf() prints string combined with the variable formatted with %d ,%s , %f .

**Q2:** **Create a data frame Write a R program to create a data frame from four given vectors. a name b. Subject C. Score d. Rank**

name=c("Sam","John","Cooper")

subject= c("CSE","Maths","English")

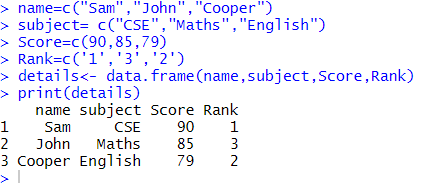
Score=c(90,85,79)

Rank=c('A','B','B')

details<- data.frame(name,subject,Score,Rank)

print(details)

**Output:**



**Explanation:**

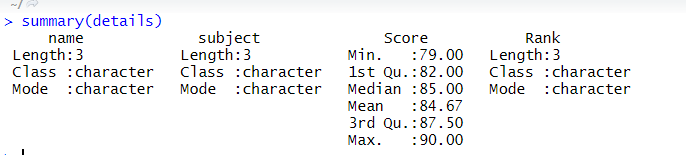
Dataframes are like 2-dimensional structure in the form of table, containing set of data in each columns. No column must be empty in the data.

**Q.4: Write a R program to get the statistical summary and nature of the data of a given data frame. ( use 3rd Question dataframe)**

**Program:**

summary(details)

**Output:**



**Explanation:**

Summary() function gives the overall statistical summary for numerical data and the length and class for characters in dataframe.

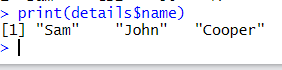
**Taking the above dataframe from Q3 as example.**

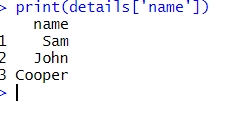
**Q.5:** **Write a R program to extract specific column from a data frame using column name.**

**Program:**

Print(details$names) or print(details[‘name’])

**Output:**





**Explanation:**

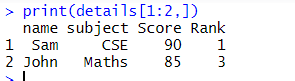
Column name can be specified with dataframe name followed by $ and the name of the column. ( df$colname)

**Q.6: Write a R program to extract first two rows from a given data frame**

**Program:**

print(details[1:2,])

**Output:**



**Explanation:**

The slicing operator used prints the 1st two rows 1:2 row wise, and all columns

**Date:**

**Q.NO : Question**

**Program:**

**Output:**

**Explanation:**