DAY 2 PROGRAMS

1.Write a R program to create an array of two 3x3 matrices each with 3 rows and 3 columns from two given two vectors. Print the second row of the second matrix of the array and the element in the 3rd row and 3rd column of the 1st matrix.

2. Write a R program to combine three arrays so that the first row of the first array is followed by the first row of the second array and then first row of the third array.

3. Write a R program to create an array using four given columns, three given rows, and two given tables and display the content of the array.

4. Write a R program to create a two-dimensional 5x3 array of sequence of even integers greater than 50. Use Below Data frame from question 5 to 9 exam\_data = data.frame( name = c('Anastasia', 'Dima', 'Katherine', 'James', 'Emily', 'Michael', 'Matthew', 'Laura', 'Kevin', 'Jonas'), score = c(12.5, 9, 16.5, 12, 9, 20, 14.5, 13.5, 8, 19), attempts = c(1, 3, 2, 3, 2, 3, 1, 1, 2, 1), qualify = c('yes', 'no', 'yes', 'no', 'no', 'yes', 'yes', 'no','no','yes') )

5. Write a R program to extract 3rd and 5th rows with 1st and 3rd columns from a given data frame

6. Write a R program to add a new column named country in a given data frameCountry<-c("USA","USA","USA","USA","UK","USA","USA","India","USA","USA")

7. Write a R program to add new row(s) to an existing data frame new\_exam\_data = data.frame(name = c('Robert', 'Sophia'),score =c(10.5, 9), attempts = c(1, 3),qualify = c('yes', 'no'))

8. Write a R program to sort a given data frame by name and score

9. Write a R program to save the information of a data frame in a file and display the information of the file.

10. Write a R program to call the (built-in) dataset airquality. Check whether it is a data frame or not? Order the entire data frame by the first and second column. remove the variables 'Solar.R' and 'Wind' and display the data frame.

11. Write a R program to create a factor corresponding to height of women data set , which inbuild in R, contains height and weights for a sample of women.

12. Write a R program to extract the five of the levels of factor created from a random sample from the LETTERS (Part of the base R distribution.)

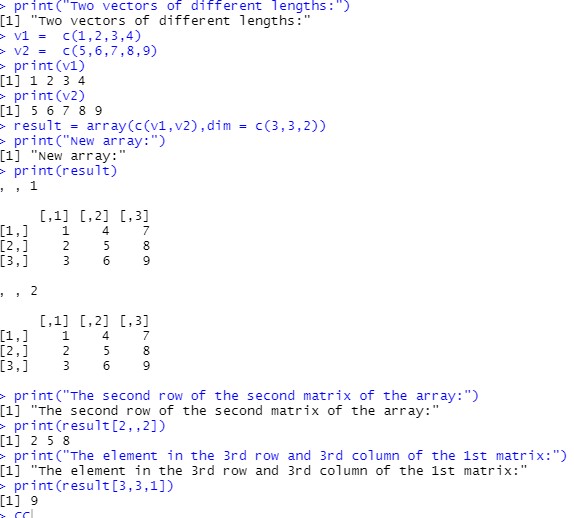
13. Iris dataset is a very famous dataset in almost all data mining, machine learning courses, and it has been an R build-in dataset. The dataset consists of 50 samples from each of three species of Iris flowers (Iris setosa, Iris virginica and Iris versicolor). Four features(variables) were measured from each sample, they are the length and the width of sepal and petal, in centimetres. Perform the following EDA steps . (i)Find dimension, Structure, Summary statistics, Standard Deviation of all features. (ii)Find mean and standard deviation of features groped by three species of Iris flowers (Iris setosa, Iris virginica and Iris versicolor) (iii)Find quantile value of sepal width and length (iV)create new data frame named iris1 which have a new column name Sepal.Length.Cate that categorizes “Sepal.Length” by quantile (V) Average value of numerical varialbes by two categorical variables: Species and Sepal.Length.Cate: (vi) Average mean value of numerical varialbes by Species and Sepal.Length.Cate (vii)Create Pivot Table based on Species and Sepal.Length.Cate.

ANSWERS

1.Write a R program to create an array of two 3x3 matrices each with 3 rows and 3 columns from two given two vectors. Print the second row of the second matrix of the array and the element in the 3rd row and 3rd column of the 1st matrix.

a) print("Two vectors of different lengths:") v1 = c(1,2,3,4) v2 = c(5,6,7,8,9) print(v1) print(v2) result = array(c(v1,v2),dim = c(3,3,2)) print("New array:") print(result) print("The second row of the second matrix of the array:") print(result[2,,2]) print("The element in the 3rd row and 3rd column of the 1st matrix:") print(result[3,3,1])

OUTPUT



2. Write a R program to combine three arrays so that the first row of the first array is followed by the first row of the second array and then first row of the third array.

A) num1 = rbind(rep("D",2), rep("E",2), rep("F",2)) print("num1") print(num1) num2 = rbind(rep("S",2), rep("T",2), rep("U",2)) print("num2") print(num2) num3 = rbind(rep("V",2), rep("W",2), rep("X",2)) print("num3") print(num3) a = matrix(t(cbind(num1,num2,num3)),ncol=2, byrow=T) print("Combine three arrays, taking one row from each one by one:") print(a)

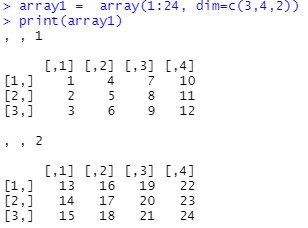
OUTPUT



3. Write a R program to create an array using four given columns, three given rows, and two given tables and display the content of the array.

A) array1 = array(1:24, dim=c(3,4,2)) print(array1)

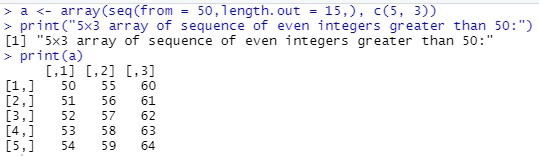
OUTPUT



4. Write a R program to create a two-dimensional 5x3 array of sequence of even integers greater than 50.

A) a <- array(seq(from = 50,length.out = 15,), c(5, 3)) print("5×3 array of sequence of even integers greater than 50:") print(a)

OUTPUT



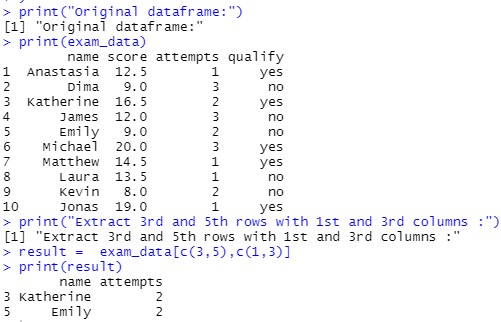
5. Write a R program to extract 3 rd and 5 th rows with 1 st and 3 rd columns from a given data frame

A) exam\_data = data.frame( name = c('Anastasia', 'Dima', 'Katherine', 'James', 'Emily', 'Michael', 'Matthew', 'Laura', 'Kevin', 'Jonas'), score = c(12.5, 9, 16.5, 12, 9, 20, 14.5, 13.5, 8, 19), attempts = c(1, 3, 2, 3, 2, 3, 1, 1, 2, 1), qualify = c('yes', 'no', 'yes', 'no', 'no', 'yes', 'yes', 'no', 'no', 'yes')

)

print("Original dataframe:") print(exam\_data) print("Extract 3rd and 5th rows with 1st and 3rd columns :") result = exam\_data[c(3,5),c(1,3)] print(result)

OUTPUT



6. Write a R program to add a new column named country in a given data frame

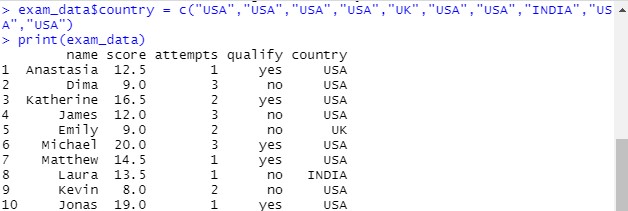
Country<-C (“USA”,”USA”,”USA”,”USA”,”UK”,”USA”,”USA”,”India”,”USA”,”USA”) A) exam\_data = data.frame( name = c('Anastasia', 'Dima', 'Katherine', 'James', 'Emily', 'Michael', 'Matthew', 'Laura', 'Kevin', 'Jonas'), score = c(12.5, 9, 16.5, 12, 9, 20, 14.5, 13.5, 8, 19),

attempts = c(1, 3, 2, 3, 2, 3, 1, 1, 2, 1), qualify = c('yes', 'no', 'yes', 'no', 'no', 'yes', 'yes', 'no', 'no', 'yes')

)

print("Original dataframe:") print(exam\_data) print("New data frame after adding the 'country' column:") exam\_data$country = c("USA","USA","USA","USA","UK","USA","USA","INDIA","USA","USA") print(exam\_data)

OUTPUT



7. Write a R program to add new row(s) to an existing data frame new\_exam\_data = data.frame(name = c(&#39;Robert&#39;, &#39;Sophia&#39;),score = c(10.5, 9), attempts = c(1, 3),qualify = c(&#39;yes&#39;, &#39;no&#39;)) A) exam\_data = data.frame( name = c('Anastasia', 'Dima', 'Katherine', 'James', 'Emily', 'Michael', 'Matthew', 'Laura', 'Kevin', 'Jonas'), score = c(12.5, 9, 16.5, 12, 9, 20, 14.5, 13.5, 8, 19), attempts = c(1, 3, 2, 3, 2, 3, 1, 1, 2, 1), qualify = c('yes', 'no', 'yes', 'no', 'no', 'yes', 'yes', 'no', 'no', 'yes')

)

print("Original dataframe:")

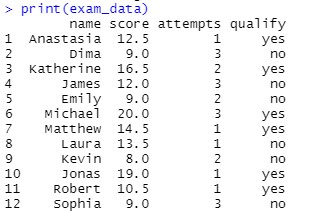
print(exam\_data) new\_exam\_data = data.frame(

name = c('Robert', 'Sophia'), score = c(10.5, 9), attempts = c(1, 3), qualify = c('yes', 'no')

)

exam\_data = rbind(exam\_data, new\_exam\_data) print("After adding new row(s) to an existing data frame:") print(exam\_data)

OUTPUT



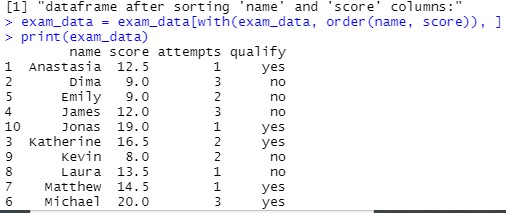
1. Write a R program to sort a given data frame by name and scoreA) exam\_data = data.frame( name = c('Anastasia', 'Dima', 'Katherine', 'James', 'Emily', 'Michael', 'Matthew', 'Laura', 'Kevin', 'Jonas'), score = c(12.5, 9, 16.5, 12, 9, 20, 14.5, 13.5, 8, 19), attempts = c(1, 3, 2, 3, 2, 3, 1, 1, 2, 1), qualify = c('yes', 'no', 'yes', 'no', 'no', 'yes', 'yes', 'no', 'no', 'yes')

)

print("Original dataframe:")

print(exam\_data) print("dataframe after sorting 'name' and 'score' columns:") exam\_data = exam\_data[with(exam\_data, order(name, score)), ] print(exam\_data)

OUTPUT

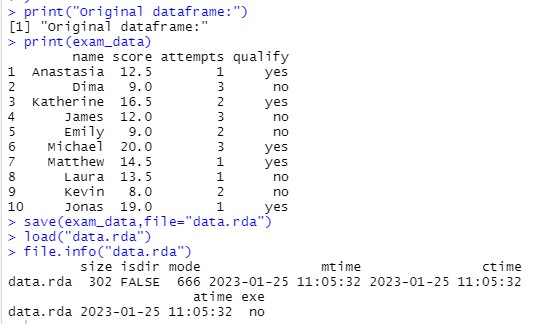


1. Write a R program to save the information of a data frame in a file and display the information of the file. A) exam\_data = data.frame( name = c('Anastasia', 'Dima', 'Katherine', 'James', 'Emily', 'Michael', 'Matthew', 'Laura', 'Kevin', 'Jonas'), score = c(12.5, 9, 16.5, 12, 9, 20, 14.5, 13.5, 8, 19), attempts = c(1, 3, 2, 3, 2, 3, 1, 1, 2, 1), qualify = c('yes', 'no', 'yes', 'no', 'no', 'yes', 'yes', 'no', 'no', 'yes')

)

print("Original dataframe:") print(exam\_data) save(exam\_data,file="data.rda") load("data.rda") file.info("data.rda")

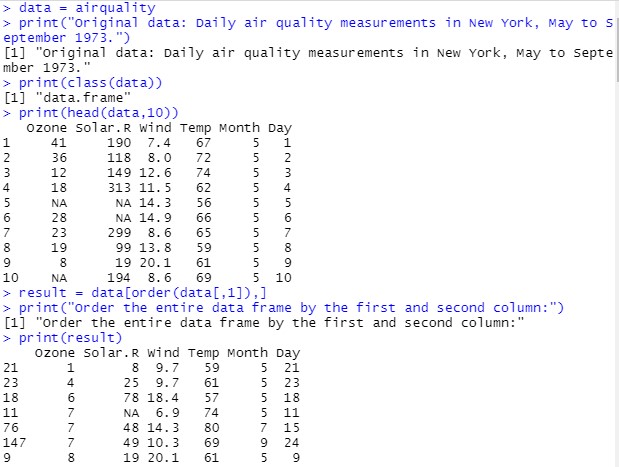
OUTPUT



1. Write a R program to call the (built-in) dataset airquality. Check whether it is a data frame or not? Order the entire data frame by the first and second column.

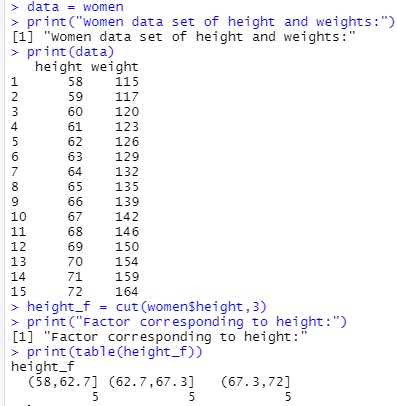
remove the variables &#39;Solar.R&#39; and &#39;Wind&#39; and display the data frame. A) data = airquality print("Original data: Daily air quality measurements in New York, May to September 1973.") print(class(data)) print(head(data,10)) result = data[order(data[,1]),] print("Order the entire data frame by the first and second column:") print(result) data[,c("Solar.R")]=NULL data[,c("Wind")]=NULL print("data.frame after removing 'Solar.R' and 'Wind' variables:") print(data)

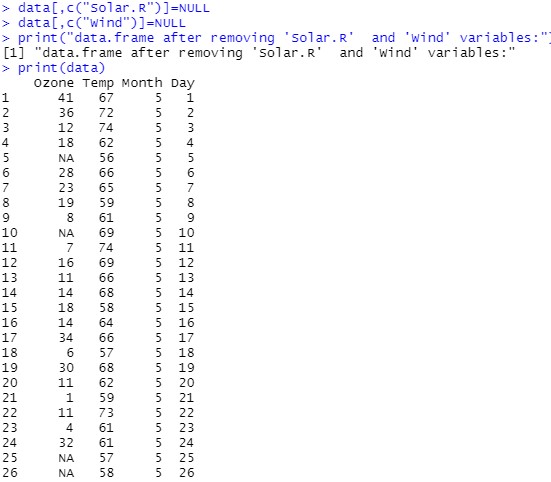
OUTPUT



1. Write a R program to create a factor corresponding to height of women data set , which inbuild in R, contains height and weights for a sample of women. A) data = women print("Women data set of height and weights:") print(data) height\_f = cut(women$height,3) print("Factor corresponding to height:") print(table(height\_f))

OUTPUT





1. Write a R program to extract the five of the levels of factor created from a random sample from the LETTERS (Part of the base R distribution.) A) L = sample(LETTERS,size=50,replace=TRUE) print("Original data:") print(L) f = factor(L) print("Original factors:") print(f) print("Only five of the levels") print(table(L[1:5]))

OUTPUT

