

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 3<sup>rd</sup> and 5<sup>th</sup> rows with 1<sup>st</sup> and 3<sup>rd</sup> columns from a given data frame
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")
```
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 inbuilt 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 grouped 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 variables by two categorical variables: Species and Sepal.Length.Cate:

(vi) Average mean value of numerical variables by Species and Sepal.Length.Cate

(vii)Create Pivot Table based on Species and Sepal.Length.Cate.