

Data Science with R Exam Questions

[Time: 2 hrs]
[Total Marks: 100]

Marks

- Q1. Write an R program to create a sequence of numbers from 20 to 50 and find the mean of numbers from 20 to 60 and the sum of numbers from 51 to 91. [3]
- Q2. A student scored 70 marks in English, 95 marks in Science, 80 marks in Maths and 74 marks in History. Write an R program to plot a simple bar chart displaying the scores of the given subjects. [5]
- Q3. Write a R program to create a data frame to store the following details of 5 employees. [3]

Name	Gender	Age	Designation	SSN
Anastasia S	M	23	Clerk	123-34-2346
Dima R	M	22	Manager	123-44-779
Katherine S	F	25	Executive	556-24-433
JAMES A	F	26	CEO	123-98-987
LAURA MARTIN	M	32	ASSISTANT	679-77-576

- Q4. Write an R program to create a list of heterogeneous data, which includes character, numeric and logical vectors. Print the list. [3]
- Q.5 Write an R program to convert a given matrix to a 1-dimensional array. [2]

	[,1]	[,2]	[,3]	[,4]
--	------	------	------	------

PG Program in Analytics

[1,]	1	4	7	10
[2,]	2	5	8	11
[3,]	3	6	9	12

- Q.6 Write a R program to create a list containing a given vector, a matrix, and a list and add an element at the end of the list [5]

Vector: c("Red", "Green", "Black")

Matrix:

	[,1]	[,2]	[,3]
[1,]	1	5	9
[2,]	3	7	11

List: ("Python", "PHP", "Java")

New Element: 4

- Q.7 Write an R program to merge two given lists into one list. [5]

List1= list(1,2,3)

List2 = list("Red", "Green", "Black")

- Q.8 Write an R program to convert a given data frame to a list by rows. [8]

Name	Score	attempts	qualify
Anastasia	12.5	1	yes
Dima	9.0	3	no
Katherine	16.5	2	yes
James	12.0	3	no
Emily	9.0	2	no
Michael	20.0	3	yes
Matthew	14.5	1	yes
Laura	13.5	1	no
Kevin	8.0	2	no

- Q.9 Write an R program to create a correlation matrix from a data frame of the same data type. [2]

```
d = data.frame(x1=rnorm(5),  
              x2=rnorm(5),  
              x3=rnorm(5))
```

- Q.10 Write an R program to rotate a given matrix 90 degrees clockwise. [8]

	[,1]	[,2]	[,3]
[1,]	1	4	7
[2,]	2	5	8
[3,]	3	6	9

- Q.11 Check for missing values in the 'mtcars' data set. [3]
- Q.12 Check which attributes are important to determine the mpg of a car in the 'mtcars' data set. [8]
- Q.13 Build a simple linear model to predict the mpg of a car in the 'mtcars' data set. [8]
- Q.14 Build a logistic regression model using the glm function to know the effect of admission into graduate school. The target variable, admit/don't admit, is a binary variable [8]
Use the given "binary.csv" dataset
- Q.15 Use the given variables from the titanic dataset and build the decision tree on train data. [5]
Variables from dataset: survived, embarked, sex, sibsp, parch, fare
- Q.16 Create a plot to display the result of decision tree. [5]
- Q.17 Create the confusion matrix for the above model. [3]
- Q.18 Perform k-means clustering on USArrest dataset. Scale the data before performing clustering. [8]
Use the below code to load the data

```
data("USArrests")
```

- Q.19 Print the cluster number for each observation and cluster size **[3]**
for the above k-means model.
- Q.20 Plot the result of the k-means cluster. **[5]**