

VIT-AP UNIVERSITY, ANDHRA PRADESH

**CSE4027 – Data Analytics - Lab Sheet :5**

**Academic year:** 2022-2023

**Branch/ Class:** B.Tech/M.Tech

**-Semester:** Fall

**Date:** - 07-10-22

**Faculty Name:** Dr Syed Khasim

**School:** SCOPE

**Student name:** MAJJIGA JASWANTH

**Reg. no.:** 20BCD7171

---

Questions:

1. Use a nested for loop (a for loop inside a for loop) that produces the following matrix, pre-allocate the matrix with NA values.

```
0 1 2 3 4
1 0 1 2 3
2 1 0 1 2
3 2 1 0 1
4 3 2 1 0
```

Code & Output:

```
> x<-matrix(data=NA,nrow=5,ncol=5)
> for(i in 0:4){
+   for(j in 0:4){
+     x[i+1,j+1]=abs(j-1)
+   }
+ }
> x
      [,1] [,2] [,3] [,4] [,5]
[1,]    1    0    1    2    3
[2,]    1    0    1    2    3
[3,]    1    0    1    2    3
[4,]    1    0    1    2    3
[5,]    1    0    1    2    3
```

2. Using the following variable:

x=100

y=50

i=1

For this exercise, write a repeat () loop that incrementing i computes  $x=x-i$  and  $y=y+i$  until  $x<y$ .

Code & Output:

```

> x=100
> y=50
> i=1
> repeat
+ {
+   x=x-i
+   y=y+i
+   if(x<y)
+     break
+ }
> print(x)
[1] 74
> print(y)
[1] 76
>

```

3. Find out the river length less than 400, if so, identify it as short river. count the number of short rivers. if the river length is > 800, then it is long. count the number of long rivers. (Use revier dataset)

#### Code & Output:

```

> d=read.csv("C:/Users/Sashank K/Desktop/River_Dataset.csv")
> d

```

	Sno	rivername	length	category
1	1	river1	995	NA
2	2	river2	1186	NA
3	3	river3	959	NA
4	4	river4	414	NA
5	5	river5	202	NA
6	6	river6	290	NA
7	7	river7	531	NA
8	8	river8	1047	NA
9	9	river9	1096	NA
10	10	river10	378	NA
11	11	river11	487	NA
12	12	river12	209	NA
13	13	river13	615	NA
14	14	river14	1092	NA
15	15	river15	647	NA
16	16	river16	847	NA
17	17	river17	1205	NA
18	18	river18	605	NA
19	19	river19	1198	NA
20	20	river20	1089	NA
21	21	river21	915	NA
22	22	river22	981	NA
23	23	river23	1295	NA
24	24	river24	378	NA
25	25	river25	332	NA
26	26	river26	1062	NA
27	27	river27	285	NA
28	28	river28	955	NA
29	29	river29	449	NA
30	30	river30	335	NA
31	31	river31	500	NA
32	32	river32	1018	NA
33	33	river33	1095	NA
34	34	river34	1085	NA
35	35	river35	321	NA
36	36	river36	800	NA

```

> for(i in d$length){
+   if(i<400){
+     c1=c1+1
+   }else if(i>800){
+     c2=c2+1
+   }
+ }
> print(paste("No of Short Rivers is: ",c1))
[1] "No of Short Rivers is: 155"
> print(paste("No of Long Rivers is: ",c2))
[1] "No of Long Rivers is: 382"
>

```

4. Using an if -else find out the largest animal based on overall sample size.  
(Use animal dataset)

Code & Output:

```

> animal<-read.csv("C:/Users/Sashank K/Downloads/Animal dataset.csv")
> large=0
> animal$overall.Sample.Size[is.na(animal$overall.Sample.Size)] <-
+   mean(animal$overall.Sample.Size, na.rm=TRUE)
> for (i in animal$overall.Sample.Size){
+   if(large < i){
+     large = i
+   } else {
+     large=large
+   }
+ }
> name<-
+   animal$Species.Common.Name[animal$overall.Sample.Size==large]
> print(paste("The largest animal based on overall sample size is ",name))
[1] "The largest animal based on overall sample size is Flamingo, Chilean"
>

```

5. Count how many males are in underweight and overweight. you need to measure with ounce. 1 ounce=28.34 gm. (Use BMI Dataset)

Code & Output:

```

> weight<-read.csv("C:/Users/Sashank K/Downloads/bmi_data.csv")
> count1=0
> count2=0
> weight$BMI[is.na(weight$BMI)]<-mean(weight$BMI,na.rm = TRUE)
> for (i in weight$BMI[weight$Sex=="Male"]){
+   if(i < 18.5){
+     count1 = count1+1
+   } else if (i > 25) {
+     count2 = count2+1
+   }
+ }
> print(paste("The no of under weight males are ",count1))
[1] "The no of under weight males are 3685"
> print(paste("The no of over weight males are ",count2))
[1] "The no of over weight males are 4"

```

6. create a dummy dataset with numeric values. a) write a function for calculating the mean. b) Write a function to compute standard deviation.

**Code & Output:**

```
> vector<-c(11,22,56,21,45,12)
> me<-function(a)
+ {
+   m=0
+   n=length(a)
+   for(i in a){
+     m=i+m
+   }
+   mn=m/n
+   print(paste("The mean od the given numbers in dummy dataset is: ",mn))
+ }
> std<-function(a){
+   n=length(a)
+   s=sqrt(sum((a-mean(a))^2/(n-1)))
+   print(paste("The standard deviation of the given numbers in the dummy dataset is: ",s))
+ }
> me(vector)
[1] "The mean od the given numbers in dummy dataset is:  27.8333333333333"
> std(vector)
[1] "The standard deviation of the given numbers in the dummy dataset is:  18.454448424883"
```