

VIT-AP UNIVERSITY, ANDHRA PRADESH

CSE2047 – Data Analytics - Lab Sheet : 3

Academic year: 2020-2021

Branch/ Class: B.Tech/M.Tech

Semester: Fall

Date:

Faculty Name: Prof. S.Gopikrishnan

School: SCOPE

Student name: Valiveti Manikanta Bhuvanesh

Reg. no.: 19BCD7088

LAB 3

Questions:

1. Find out mean of all columns of BMI dataset using for loop.

```
df<-read.csv("bmi_data.csv")
df
```

```
sum =0
count=0
for (i in df$Age) {
  sum = sum+i
  count=count+1
}
sum/count
```

```
sum =0
count=0
for (i in na.omit(df$Height)) {
  sum = sum+i
  count=count+1
}
sum/count
```

```
sum =0
count=0
for (i in na.omit(df$Weight)) {
  sum = sum+i
  count=count+1
}
sum/count
```

```
sum =0
```

```

count=0
for (i in na.omit(df$BMI)) {
  sum = sum+i
  count=count+1
}
sum/count
>
> sum =0
> count=0
> for (i in df$Age) {
+   sum = sum+i
+   count=count+1
+ }
> sum/count
[1] 26.4996
>
> sum =0
> count=0
> for (i in na.omit(df$Height)) {
+   sum = sum+i
+   count=count+1
+ }
> sum/count
[1] 67.99297
>
> sum =0
> count=0
> for (i in na.omit(df$Weight)) {
+   sum = sum+i
+   count=count+1
+ }
> sum/count
[1] 127.0782
>
> sum =0
> count=0
> for (i in na.omit(df$BMI)) {
+   sum = sum+i
+   count=count+1
+ }
> sum/count
[1] 19.32137
.

```

2. Create two vectors - height and weight of 20 students and convert it to a matrix

```
height = c(23,34,54,34,67,98,56,76,23,56,87,56,45,34,32,54,76,98,89,90)
```

```
weight = c(23,54,65,76,87,78,56,45,34,23,24,35,36,67,56,87,56,47,76,73)
```

```
m=cbind(height,weight)
```

```
m
```

```
> height = c(23,34,54,34,67,98,56,76,23,56,87,56,45,34,32,54,76,98,89,90)
```

```
> weight = c(23,54,65,76,87,78,56,45,34,23,24,35,36,67,56,87,56,47,76,73)
```

```
> m=cbind(height,weight)
```

```
> m
```

	height	weight
[1,]	23	23
[2,]	34	54
[3,]	54	65
[4,]	34	76
[5,]	67	87
[6,]	98	78
[7,]	56	56
[8,]	76	45
[9,]	23	34
[10,]	56	23
[11,]	87	24
[12,]	56	35
[13,]	45	36
[14,]	34	67
[15,]	32	56
[16,]	54	87
[17,]	76	56
[18,]	98	47
[19,]	89	76
[20,]	90	73

3. Convert matrix into data frame and find the weight of 12th student

```
df1= as.data.frame(m)
```

```
df1
```

```
df1$weight[12]
```

```
> df1= as.data.frame(m)
```

```
> df1
```

	height	weight
1	23	23
2	34	54
3	54	65
4	34	76
5	67	87
6	98	78
7	56	56
8	76	45
9	23	34
10	56	23
11	87	24
12	56	35
13	45	36
14	34	67
15	32	56
16	54	87
17	76	56
18	98	47
19	89	76
20	90	73

```
> df1$weight[12]
```

```
[1] 35
```

4. Categorize high and low based on bmi factor using if condition

```
v=c(1:dim(df)[1])
```

V

```
for (i in v) {
```

```
if(!is.na(df$BMI[i])){
```

```
if(df$BMI[i]>=19){
```

```
df$bmifactor[i]="High"
```

}

```
else if(df$BMI[i]< 19){
```

```
df$bmifactor[i]="Low"
```

}

```
else{
```

```
df$bmifactor[i]="NA"
```

}

}

}

```
print(df)
```

```
> for (i in v) {
```

```
+   if(!is.na(df$BMI[i])){
```

```
+   if(df$BMI[i]>=19){
```

```
+ df$bmifactor[1]="High"
```

+	j
+	p

```
+ df$bmifactor[i]="Low"
```

$$+ \quad \}$$

```
+ else{
```

```
+ df$bmi.factor[i]="NA"
```

$$\left\{ \begin{array}{l} + \\ - \end{array} \right\}$$
$$\begin{array}{l} + \\ + \end{array} \quad \}$$
$$+ \}$$

```
> print(df)
```

	Sex	Age	Height	Weight	BMI	bmfactor
1	Female	21	65.78331	112.99250	18.35765	Low
2	Female	35	71.51521	136.48730	18.76265	Low
3	Female	27	69.39874	153.02690	22.33898	High
4	Male	24	68.21660	142.33540	21.50461	High
5	Female	18	67.78781	144.29710	22.07767	High
6	Female	22	68.69784	123.30240	18.36894	Low

5. Using switch case categorize into

Category

BMI range - weight/height

Severe Thinness < 16

Moderate Thinness 16 - 17

Mild Thinness 17 - 18.5

Normal 18.5 - 25

Overweight 25 - 30

Obese Class I 30 - 35

Obese Class II 35 - 40

Obese Class III > 40

```
flag=0
for (i in v) {
  n=df$BMI[i]
  if(!is.na(n)){
    if((n<16)){
      flag=1
    }
    else if(n>=16 & n<17){
      flag=2
    }
    else if(n>=17 & n<18.5){
      flag=3
    }
    else if(n>=18.5 & n<25){
      flag=4
    }
    else if(n>=25 & n<30){
      flag=5
    }
    else if(n>=30 & n<35){
      flag=6
    }
    else if(n>=35 & n<40){
      flag=7
    }
    else{
      flag=8
    }
  }
  x<- switch (flag,
    "Severe Thinness",
    "Moderate Thinness",
    "Mild Thinness",
    "Normal",
    "Overweight",
    "Obese Class I",
    "Obese Class II",
    "Obese Class III"
  )
  df$bmifactor1[i]=x
}
print(df)
```

```

> flag=0
> for (i in v) {
+   n=dfsBMI[i]
+   if(!is.na(n)){
+     if((n<16)){
+       flag=1
+     }
+     else if(n>=16 & n<17){
+       flag=2
+     }
+     else if(n>=17 & n<18.5){
+       flag=3
+     }
+     else if(n>=18.5 & n<25){
+       flag=4
+     }
+     else if(n>=25 & n<30){
+       flag=5
+     }
+     else if(n>=30 & n<35){
+       flag=6
+     }
+     else if(n>=35 & n<40){
+       flag=7
+     }
+     else{
+       flag=8
+     }
+     x<- switch (flag,
+       "Severe Thinness",
+       "Moderate Thinness",
+       "Mild Thinness",
+       "Normal",
+       "Overweight",
+       "Obese Class I",
+       "Obese Class II",
+       "Obese Class III"
+     )
+   }
+   dfsbmifactor1[i]=x
+ }
> print(df)

```

	Sex	Age	Height.Inches.	Weight.Pounds.	BMI	bmifactor	bmifactor1
1	Female	21	65.78331	112.99250	18.35765	Low	Mild Thinness
2	Female	35	71.51521	136.48730	18.76265	Low	Normal
3	Female	27	69.39874	153.02690	22.33898	High	Normal
4	Male	24	68.21660	142.33540	21.50461	High	Normal
5	Female	18	67.78781	144.29710	22.07767	High	Normal
6	Female	22	68.69784	123.30240	18.36894	Low	Mild Thinness
7	Male	35	69.80204	141.49470	NA	Low	Mild Thinness
8	Male	19	70.01472	136.46230	19.57189	High	Normal
9	Female	28	67.90265	112.37230	17.13502	Low	Mild Thinness
10	Male	25	66.78236	120.66720	19.02237	High	Normal
11	Male	34	NA	127.45160	NA	Low	Normal
12	Female	28	67.62333	114.14300	17.54911	Low	Mild Thinness
13	Male	21	68.30248	125.61070	18.93008	Low	Normal
14	Male	25	67.11656	122.46180	19.11349	High	Normal
15	Female	27	68.27967	116.08660	17.50645	Low	Mild Thinness
16	Male	23	71.09160	139.99750	19.47523	High	Normal
17	Male	20	66.46100	129.50230	20.61306	High	Normal
18	Female	31	68.64927	142.97330	21.32956	High	Normal
19	Male	19	71.23033	137.90250	19.10914	High	Normal
20	Female	32	67.13118	124.04490	19.35215	High	Normal
21	Male	22	67.83379	141.28070	21.58686	High	Normal
22	Female	32	68.87881	143.53920	21.27150	High	Normal
23	Female	31	63.48115	97.90191	17.08049	Low	Mild Thinness