Probability and Statistics Assignment 1

Harshad Kumar Elangovan - 19200349

01/10/2019

The table with exam marks of each student is created using the below commands.

```
studentname=c("John","Anne","Terry","Fred","Maria")
exam1=c(92,75,98,62,79)
exam2=c(82,96,60,55,72)
results=data.frame(studentname,exam1,exam2)
print(results)
```

```
##   studentname exam1 exam2
## 1        John    92    82
## 2        Anne    75    96
## 3       Terry    98    60
## 4        Fred    62    55
## 5       Maria    79    72
```

Part A

What code creates the list of exam2 marks?

The code for creating the list of exam2 marks is

```
exam2=c(82,96,60,55,72)
```

Part B

Who has a better grade, Maria or Terry?

From the table, we can see that Maria and Terry have the same grade ‘B’. But the average of Terry is more than that of Maria. So, **Terry has better grade than Maria**.

```
results$avg = ((results$exam1 + results$exam2)/2)
results$grade = cut(results$avg, breaks = c(0,49,69,84,100),labels=c("D","C","B","A"))
results = results[order(-results$avg),]
print(results)
```

```
##   studentname exam1 exam2  avg grade
## 1        John    92    82 87.0     A
## 2        Anne    75    96 85.5     A
## 3       Terry    98    60 79.0     B
## 5       Maria    79    72 75.5     B
## 4        Fred    62    55 58.5     C
```

Part C

Suppose you wanted to print out only the names and exam 1 marks of the students. What code would you type?

We can display the names and exam 1 of the students separately by creating another dataframe with only those two columns in it. This can be done using the below code,

```
results1 = data.frame(results$studentname,results$exam1)
print(results1)
```

```
##   results.studentname results.exam1
## 1                John            92
## 2                Anne            75
## 3               Terry            98
## 4               Maria            79
## 5                Fred            62
```

Part D

Mention the mean and median of Pulse1 and Pulse2 and which do you think is pulse after exercise?

```
pulse = read.table("C:\\Users\\harshie\\Documents\\UCD\\Probability\\Lab1\\pulse.txt", header
=TRUE)
pulse = data.frame(pulse)
```

```
summary(pulse$pulse1)
```

```
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   48.00   64.00   71.00   72.87   80.00  100.00
```

```
summary(pulse$pulse2)
```

```
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##      50      68      76      80      85     140
```

From the above table, the mean and median of both the pulses are as follows,

Pulse1

Mean:72.87

Median:71.00

Pulse2

Mean:80

Median:76

The pulse 2 with mean and median value higher is the pulse after exercise.

Part E

Mention the mean and median of Ran = 1 and Ran = 2 and Which group do you think ran and which walked?

The mean and median of the pulses with Ran value 1 and 2 can be found using the below mentioned function in markdown. This will produce the output based on the conditions.

```
summary(pulse$pulse2[pulse$ran ==1])
```

```
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   58.00   76.00   88.00   92.51  105.00  140.00
```

```
summary(pulse$pulse2[pulse$ran ==2])
```

```
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   50.00   66.00   70.00   72.32   78.00   94.00
```

From the above table, the mean and median of both the pulses are as follows,

Ran = 1

Mean:92.51

Median:88.00

Ran = 2

Mean:72.32

Median:70.00

So, from the above mean and median values, the pulse with Ran =1 ran while exercising and pulse with Ran=2 walked.

Part F

What is the most common (modal) pulse1 measurement?

From the output below, the most common measurement in pulse1 is 68.

```
table(pulse$pulse1)
```

```
##
##  48  54  58  60  61  62  64  66  68  70  72  74  76  78  80  82  84  86
##   1   2   3   4   1   9   4   5  11   6   6   5   5   5   3   3   4   1
##  87  88  90  92  94  96 100
##   1   3   4   2   1   2   1
```

What is the most common (modal) pulse2 measurement?

From the output below, the most common measurement in pulse2 is 76.

```
table(pulse$pulse2)
```

```
##
##  50  54  56  58  60  62  64  66  68  70  72  74  75  76  78  80  82  84
##   1   1   2   2   1   4   1   6   6   7   3   5   1  12   2   5   1   9
##  88  90  92  94  96  98 100 102 104 106 110 112 115 116 118 128 140
##   2   1   2   3   1   1   2   1   1   1   1   1   1   1   2   1   1
```

Part G

Sketch the histogram of the weights of female volunteers here:

The code for producing the histogram of the weights of female volunteers is as follows,

```
hist(pulse$weight[pulse$sex==2],xlab="Weight",main="Histogram of Female Volunteers based on w
eight")
```