UNIVERSITI TEKNOLOGI MARA
MID TERM TEST

COURSE	:	**STATISTICAL COMPUTING**
COURSE CODE	:	**STA705**
EXAMINATION	:	**OCTOBER 2019**
TIME	:	**2 HOURS (Total: 65 marks)**

Scanned with CamScanner

QUESTION 1

a) What is the output from the following command?

```
rep(seq(10,40,8),5:2)
```

```
[1] 10 10 10 10 10 18 18 18 18 26 26 26 34 34
```

(3 marks)

b) Write the R command to produce a vector of four 8's, seven 6's and three 5's.

```
q1b = c(rep(8,4), rep(6,7), rep(5,3))
```

(3 marks)

c) Use the **seq** command to produce the following output.

```
400 380 360 340 320 300 280 260 240 220 200 180 160 140
120 100  80  60
```

```
q1c = rev(seq(60,400,by=20))
```

(2 marks)

d) Produce the following output by using the data in part c).

```
     [,1] [,2] [,3] [,4] [,5] [,6]
[1,]  400  380  360  340  320  300
[2,]  280  260  240  220  200  180
[3,]  160  140  120  100   80   60
```

```
q1d = matrix(q1c, nrow=3, byrow=T)
```

(3 marks)

e) Based on the data in part d), obtain the following output.

```
, , 1
     [,1] [,2] [,3]
[1,]  400  160  260
[2,]  280  380  140

, , 2
     [,1] [,2] [,3]
[1,]  360  120  220
[2,]  240  340  100
```

© HakCiptaUniversitiTeknologi MARA

Scanned with CamScanner

```
, , 3
     [,1] [,2] [,3]
[1,]  320   80  180
[2,]  200  300   60
```

qie = array(qrd, c(2,3,3))

2

(2 marks)

f) Write a suitable command to replace the value 80 in part d) with 888.

qld[row = 3, col = 5] = 888

1½

(2 marks)

> qie[1,2,3] = 8

QUESTION 2

a) Given that:

59 48 25 99 10 34 67 83

Obtain the R command for the following results.

Results	R Command
10 25 34 48 59 67 83 99	sort(c(59,48,25, 99,10,34,67,83))
5 3 6 2 1 7 8 4	order(c(59,48,25, 99,10,34,67,83))
5 4 2 8 1 3 6 7	rank(c(59,48,25, 99,10,34,67,83))

6

(6 marks)

© HakCiptaUniversitiTeknologi MARA

Scanned with CamScanner

b) Merge the data in part (a), X=A:H and Y={1:3,M,N,O,P}. The results should be obtained as below.

```
> Data
  xx X Y
1 59 A 1
2 48 B 2
3 25 C 3
4 99 D 4
5 10 E M
6 34 F N
7 67 G O
8 83 H P
```

```
q2a = c(59,48,25,99,10,34,67,83)
X = c("A","B","C","D","E","F","G","H")
Y = c(1:4,"M","N","O","P")
q2b = data.frame(q2a, X, Y)
```

3

(3 marks)

QUESTION 3

a) Given that A = 10, 30, 50, 29, 23, 45, 67, 35, B= {Group 1, Group 2, Group 3} and D={Bag X, Bag Y}. By using the data in part (a) together with A, B and D, produce the following dataset.

```
> data1
  xt  A       B     D
1 59 10 Group 1 Bag X
2 48 30 Group 3 Bag Y
3 25 50 Group 2 Bag X
4 99 29 Group 1 Bag Y
5 10 23 Group 2 Bag X
6 34 45 Group 3 Bag Y
7 67 67 Group 1 Bag X
8 83 35 Group 3 Bag Y
```

```
> data2
  xt  A       B     D
1 59 10 Group 1 Bag X
2 48 30 Group 3 Bag Y
3 25 50 Group 2 Bag X
4 99 29 Group 1 Bag Y
5 10 23 Group 2 Bag X
6 34 45 Group 3 Bag Y
7 67 67 Group 1 Bag X
8 83 35 Group 3 Bag Y
```

© HakCiptaUniversitiTeknologi MARA

Scanned with CamScanner

```
xt = c(59, 18, 25, 99, 10, 34, 67, 83)
A = c(10, 30, 50, 29, 23, 45, 67, 35)
B = c("Group 1", "Group 3", "Group 2", "Group 1", "Group 2",
      "Group 3", "Group 1", "Group 3")
D = gl(2, 1, 8, labels = c("Bag X", "Bag Y"))
library(data.table)
data1 = data.frame(xt, A, B, D)
data2 = data.table(xt, A, B, D)
```

(6 marks)

b) Compute the dataset above to obtain the following results.

Output 1:

```
Group 1 Group 2 Group 3
    106      73     110
```

Output 2:

```
  Group.1 Group.2   x
1 Group 1   Bag X  77
2 Group 2   Bag X  73
3 Group 1   Bag Y  29
4 Group 3   Bag Y 110
```

Output 3:

```
          B  V1
1: Group 1 106
2: Group 3 110
3: Group 2  73
```

Output 4:

```
       D  V1
1: Bag X 161
2: Bag Y 264
```

© HakCiptaUniversitiTeknologi MARA

Scanned with CamScanner

Output 5:

```
          B     D  V1
1: Group 1 Bag X  77
2: Group 3 Bag Y 110
3: Group 2 Bag X  73
4: Group 1 Bag Y  29
```

Output 1: tapply(A, B, sum) 1

Output 2: aggregate(A, by=list(B, D), sum) 2

Output 3: aggregate(A, by=list(B), sum) 1/2

library (data-table)

Output 4: data.table(aggregate(x+, by=list(D), sum)) 1/2

1/2

Output 5: data.table(aggregate(A, by=list(B, D), sum)) 1/2

(10 marks)

QUESTION 4

The information on the ***Prestige*** dataset obtained from library(car) in R system is given as follows:

Education *Average education of occupational incumbents, years, in 1971.*
Income *Average income of incumbents, dollars, in 1971.*
Women *Percentage of incumbents who are women.*
Prestige *Pineo-Porter prestige score for occupation, from a social survey conducted in the mid-1960s.*
Census *Canadian Census occupational code.*
Type *Type of occupation. A factor with levels (note: out of order): bc, Blue Collar; prof, Professional, Managerial, and Technical; wc, White Collar.*

© HakCiptaUniversitiTeknologi MARA

Scanned with CamScanner

a) Give the R command to remove the missing values in the dataset.

```
q4q = na.omit(Prestige)
```

(2 marks)

b) State the dimension of the new dataset obtained in part (a)

rows = 98, column = 6

(1 mark)

c) By using the complete dataset obtained from part (a), write the appropriate command to determine the mean of ***income*** and ***prestige*** according to the three levels of ***type***. The output is shown as below.

```
                bc        prof          wc
income    5374.13636 10559.45161 5052.30435
prestige    35.52727    67.84839   42.24348
```

```
attach(q4q)
mean1 = tapply(income, type, mean)
mean2 = tapply(prestige, type, mean)
data1 = rbind(income = mean1, prestige = mean2)
```

(5 marks)

d) Write the R command to categorize the variable ***income*** into three groups A (high), B (middle) and C (low) as shown below. Hint: use the R command "***cut***".

Output:

```
> Pres2
   income women prestige type  Group
1   12351 11.16     68.8 prof   high
2   25879  4.02     69.1 prof   high
3    9271 15.70     63.4 prof   high
4    8865  9.11     56.8 prof   high
5    8403 11.68     73.5 prof   high
6   11030  5.13     77.6 prof   high
7    8258 25.65     72.6 prof   high
```

row = 1656.
median = 6035.5
max = 25879

0, 1656,

© HakCiptaUniversitiTeknologi MARA

Scanned with CamScanner

8	14163	2.69	78.1	prof	high
9	11377	1.03	73.1	prof	high
10	11023	0.94	68.8	prof	high
11	5902	1.91	62.0	prof	middle
12	7059	7.83	60.0	prof	middle
13	8425	15.33	53.8	prof	high
14	8049	57.31	62.2	prof	middle
15	7405	48.28	74.9	prof	middle
16	6336	54.77	55.1	prof	middle
17	19263	5.13	82.3	prof	high
18	6112	77.10	58.1	prof	middle
19	9593	34.89	58.3	prof	high
....					
87	5299	0.56	38.9	bc	middle
88	5959	0.52	36.2	bc	middle
89	4549	2.46	29.9	bc	middle
90	6928	0.61	42.9	bc	middle
91	3910	1.09	26.5	bc	low
92	14032	0.58	66.1	prof	high
93	8845	0.00	48.9	bc	high
94	5562	9.47	35.9	bc	middle
95	4224	3.59	25.1	bc	middle
96	4753	0.00	26.1	bc	middle
97	6462	13.58	42.2	bc	middle
98	3617	70.87	35.2	bc	low

low = min 1656, max 3910 (4199)

middle = min 4224, max 8049 (8403)

high = min 8425, max 25879.

```
Group= cut (q4a $income, break = c(1650, 4199, 8200, 25879), labels =
      c("Low", "Middle", "High"))

Pres2 = data.frame (q4a $income, q4a $women, q4a$ prestige,
      q4a $type, Group)
```

4

(5 marks)

© HakCiptaUniversitiTeknologi MARA

Scanned with CamScanner

e) Write the R command to aggregate the mean of ***Group.1(type)*** and ***Group.2(Group)*** as shown below.

	Group.1	Group.2	x
1	bc	low	3164.429
2	wc	low	3235.333
3	bc	middle	5933.320
4	prof	middle	6261.167
5	wc	middle	6023.462
6	bc	high	8765.400
7	prof	high	13274.158
8	wc	high	8780.000

Qe = aggregate(income, by=list(type, Group), mean).

3½

(4 marks)

f) Write the R command to produce the following plots:

Figure 1

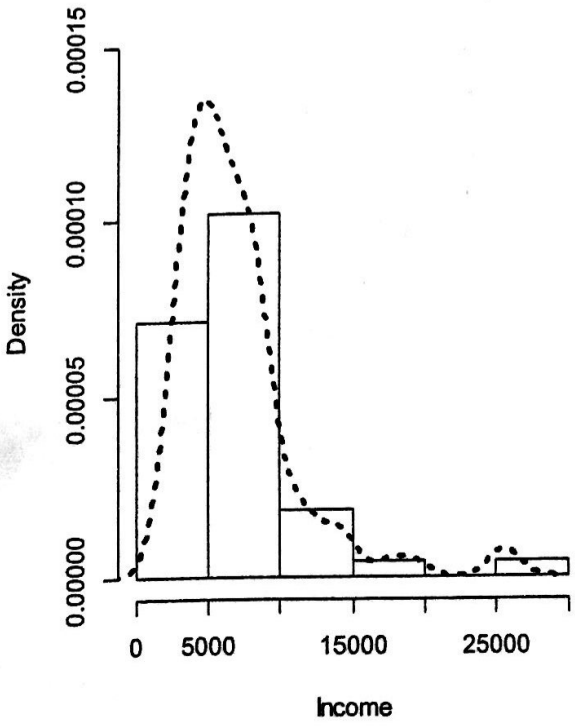
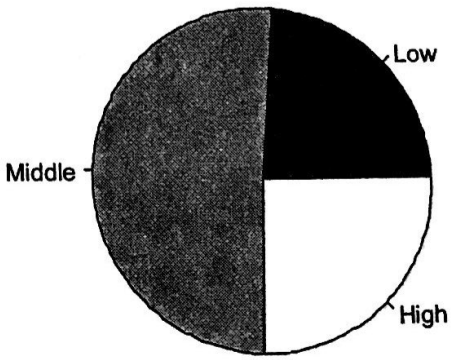


Figure 2: Level of Income



© HakCiptaUniversitiTeknologi MARA

Scanned with CamScanner

```
> par (mfrow = c(1,2))
#histogram
hist (income, xlab = "income", ylab = "Density",
 main = expression("Figure 1"), ylim = c(0.00000, 0.00015))
lines (density(income), lty = "dotted", lwd = 3)

#pie chart
pieee = table (Group)
pie (Pieee, main = expression ("Figure 2 : Level of Income"))
```

(6 marks)

g) Write the R command to produce the following tabulation.

type	Group low	middle	high
bc	14	25	5
prof	0	12	19
wc	9	13	1

```
ftable (type, Group)
```

(2 marks)

END OF TEST PAPER

© HakCiptaUniversitiTeknologi MARA

Scanned with CamScanner