TBA2102: Introduction to Business Analytics
AY 2020/21 Semester 2
Tutorial 2 Assignment: R Basics
Tutorial session: Tue 2 Feb 2021 8.30pm

Instructions:
- Prepare the answers for Section 1 for discussion during Tutorial 2 in Week 4.
- The answers for questions in Section B will be discussed during Tutorial 3 in week 5.

In this tutorial, we will review and apply concepts related to the Basics of R.

Section 1: R Basics (Answers to be discussed in class)
R objects are assigned for each question 1a, 1b… 2a, 2b, etc… so their values are independent across questions.

1. What is the output for each of the following sets of codes?

a. i.
```
x<-"2"
class(x)
```

ii.
```
x<-4
y<-10
z<-y/x
class(z)
```

b. i.
```
y<-c(20,36,10)
sort(y, decreasing = TRUE)
```

ii.
```
y<-c(20,36,10,10,10)
order(y, decreasing = FALSE)
```

iii.
```
y<-c(20,36,10,10,10)
x<-c(2,3,1,4,5)
order(y, x, decreasing = FALSE)
```

iv.
```
y<-c(20,36,10,10,10)
x<-c(2,3,1,4,5)
z<-data.frame(cbind(y,x))
z[order(z$x, decreasing = TRUE, )
```

c. i.
```
size<-c("medium", "small", "big", "big")
size_fac<- factor(size, levels=c("small","medium","big"), ordered=TRUE)
size_fac[1] < size_fac[3]
```

d. i.
```
C<-c(1,3,6,8,0,10)
C[2:4]
```

ii.
```
class(C)
```

e. i.
```
df<-data.frame(x=c("a","b","c","d","e"), y=c(1,4,6,8,10), stringsAsFactors=FALSE)
class(df$x)
```

ii.
```
class(df$y)
```

iii.
```
df[c(3:5),"y"]
```

iv.
```
df$y<-as.integer(df$y)
class(df$y)
```

v.
```
subset(df, y>6, select=x)
```

2. For each question part below, what is the missing code ("?") required to return the output? If you think "?" can be blank, then type "blank" for your answer.

a. i.
```
> vol<- c(109, 59, 56, 97, 86, 40, 39)
>  ?  (vol, decreasing = TRUE)
[1] 1 4 5 2 3 6 7
```

ii.
```
> vol<- ? *vol
> vol
[1] 218 118 112 194 172  80  78
```

b. i.
```
> shop1<-list(c("A", "B","C"), c(30,50), c(500, 1000))
>   ?  (shop1) <- c("Product","Cost", "Qty")
> shop1[["Qty"]]
[1]  500 1000
```

ii.
```
> shop1$ ?
[1] "A" "B" "C"
```

iii.
```
> shop1$ ?
[1] 30
```

c. i.
```
> x<- c("w","w","e","w")
> y<-factor(x)
>   ?   (y)<-c("east","west")
> y
[1] west west east west
Levels: east west
```

ii.
```
> x<- c("west","west","east","west")
> xfac<-factor(x, levels = c(      ?      ))
> xfac
[1] west west east west
Levels: east west
```

d. i.
```
> Candidates <- c("Mary","Natalie","James","Pete")
> Vote <- c(23, 44, 5, 66)
> Vote[ ? ]
[1] 5
```

ii.
```
> Candidates[   ?   ]
[1] "Mary" "Pete"
```

iii.
```
> dfvoting <-   ?   (Candidates,Vote)
> dfvoting
     Candidates Vote
[1,] "Mary"     "23"
[2,] "Natalie"  "44"
[3,] "James"    "5"
[4,] "Pete"     "66"
```

e.

i.
```
> df<-data.frame(Name=c("Henry", "Mary","Natalie","James","Pete"), Age=c(16,
 23, 44, 5, 66), Gender=c("M","F","F","M","M"), stringsAsFactors=FALSE )
> df[      ?       ]
   Name Age
1 Henry  16
2  Mary  23
```

ii.
```
>    ?    (df,Age>50)
  Name Age Gender
5 Pete  66      M
```

iii.
```
> subset(df,Gender=="M",       ?        )
   Name
1 Henry
4 James
5  Pete
```

iv.
```
> df$ ?
[1] "Henry"   "Mary"    "Natalie" "James"   "Pete"
```

3. A variable *rain_vol* contains the following values (which is the rain volume for each day): 100, 150, 140, 125, 20, 30, 55

a. What is the code to create the *rain_vol* vector?

b. What is the code to assign the first 3 letters of the days of the week (from "Mon", "Tue"… "Sun") as names of the *rain_vol* vector?

c. What is the code to sort *rain_vol* in increasing volume?

d. There was an error in the measuring gauge. Could you subtract 10 from each of the values in the *rain_vol*? What is the code to do this?

Section 2: (Answers to be submitted) [1 mark per question; total 25 marks]

1. What is the output for each of the following sets of codes?

a.
```
x<-c(4,2,2,1)
y<-c(2,1,2,1)
z<-x/y
z
```

b.

i.
```
height<- c(110,120,125,100)
order(height,decreasing=TRUE)
```

ii.
```
sort(height,decreasing=FALSE)
```

c.
```
grade <- c("good","bad","good","bad")
factor(grade, levels=c("good","bad"), ordered = FALSE)
```

d. i.
```
s<-c(11,13,21,15,9,"false")
class(s)
```

ii.
```
s[c(2,6)]
```

e. i.
```
df<-data.frame(candidate=c("Andy","Bob","Dylan","Elyse","Fay"), score=c(4,8,5,8,7))
class(df$score)
```

ii.
```
df[c(2,4),2]
```

iii.
```
df$candidate <-as.character(df$candidate)
df[4,"candidate"]
```

iv.
```
subset(df,score>7,select= candidate)
```

2. For each question part below, what is the missing code ("?") required to return the output?

a.
```
> x<-c(1,3,10, 8)
> y<-c("Mon","Tue","Wed","Thu")
>   ?   (x)<-y
> x
Mon Tue Wed Thu
  1   3  10   8
```

b. i.
```
> Satisfaction<-c("good","excellent", "poor","fair")
> Satisfac<-factor(Satisfaction, levels=c(              ?              ), ordered=TRUE)
> Satisfac
[1] good      excellent poor      fair
Levels: poor < fair < good < excellent
```

ii.
```
>     ?     >Satisfac[3]
[1] TRUE
```

c. i.

```
> recipe<-list(c("Pancake","Egg","Cereal","Bread"),          ?          ), c(2,3,1))
> names(recipe) <- c("Breakfast","Snacks","Qty")
> recipe
$Breakfast
[1] "Pancake" "Egg"     "Cereal"  "Bread"

$Snacks
[1] "Cookie"  "Pretzel"

$Qty
[1] 2 3 1
```

ii.

```
> recipe$ ?
[1] "Egg"
```

iii.

```
> recipe[[    ?    ]]
[1] "Cookie"  "Pretzel"
```

d. i.

```
> petallen<- c(4.5,5.5,2,3,4)
>   ?  (petallen)
[1] 2.0 3.0 4.0 4.5 5.5
```

ii.

```
> petal2len<- petallen ?
> petal2len
[1] 6.5 7.5 4.0 5.0 6.0
```

e. i.

```
> df2<-data.frame(Name=c("Henry", "Mary","James","Pete"), Age=c(16, 44, 5, 66),
 Gender=c("M","F","M","M"), stringsAsFactors = FALSE)
> df2[      ?      ]
[1] "Henry" "Mary"
```

ii.

```
> subset(df2,Age>40,       ?       ')
  Name
2 Mary
4 Pete
```

iii.

```
> subset(df2,Name    ?    select =Age)
  Age
2  44
```

3. Mary planted 5 seeds. At the end of week 2, she measured the height of each seeding (A, B, C, D, E) and recorded them in the variable *ht2* in the respective order (i.e. A, B,..,E).

Height (cm) measurements taken for seedlings A, B… E at the end of week 2 were: 2, 2.5, 4, 3, 3.5

a. What is the code to assign the height measurements to *ht2*?

b. What is the code to assign the values "A", "B",… "E" as *names* for *ht2*?

c. What is the code to sort *ht2* in decreasing value?

d. May recorded the height of plant B incorrectly. What code would you write, to change the value to 3?