

HW3

Shuangshuang Xu

2019/9/17

Problem 4

(a)

```
#a
#import data
urla<-"https://www2.isye.gatech.edu/~jeffwu/wuhamadabook/data/Sensory.dat"
sensory_data<-fread(urla,sep=" ",fill=TRUE,skip=1)
colnames(sensory_data)<-c("item","x1","x2","x3","x4","x5")
#get data frame contains NA
temp<-sensory_data %>% filter(is.na(x5)==TRUE) %>% select(1:5) %>% cbind(rep(1:10,each=2))
colnames(temp)<-c("x1","x2","x3","x4","x5","item")
#get data frame without NA
temp1<-sensory_data %>% filter(is.na(x5)==FALSE)
#combine
sensory_data<-rbind(temp1,temp)
sensory_data<-sensory_data[order(sensory_data$item),]
rownames(sensory_data)<-NULL

kable(sensory_data,caption="table of sensory data")
```

Table 1: table of sensory data

item	x1	x2	x3	x4	x5
1	4.3	4.9	3.3	5.3	4.4
1	4.3	4.5	4.0	5.5	3.3
1	4.1	5.3	3.4	5.7	4.7
2	6.0	5.3	4.5	5.9	4.7
2	4.9	6.3	4.2	5.5	4.9
2	6.0	5.9	4.7	6.3	4.6
3	2.4	2.5	2.3	3.1	2.4
3	3.9	3.0	2.8	2.7	1.3
3	1.9	3.9	2.6	4.6	2.2
4	7.4	8.2	6.4	6.8	6.0
4	7.1	7.9	5.9	7.3	6.1
4	6.4	7.1	6.9	7.0	6.7
5	5.7	6.3	5.4	6.1	5.9
5	5.8	5.7	5.4	6.2	6.5
5	5.8	6.0	6.1	7.0	4.9
6	2.2	2.4	1.7	3.4	1.7
6	3.0	1.8	2.1	4.0	1.7
6	2.1	3.3	1.1	3.3	2.1
7	1.2	1.5	1.2	0.9	0.7
7	1.3	2.4	0.8	1.2	1.3
7	0.9	3.1	1.1	1.9	1.6

item	x1	x2	x3	x4	x5
8	4.2	4.8	4.5	4.6	3.2
8	3.0	4.5	4.7	4.9	4.6
8	4.8	4.8	4.7	4.8	4.3
9	8.0	8.6	9.0	9.4	8.8
9	9.0	7.7	6.7	9.0	7.9
9	8.9	9.2	8.1	9.1	7.6
10	5.0	4.8	3.9	5.5	3.8
10	5.4	5.0	3.4	4.9	4.6
10	2.8	5.2	4.1	3.9	5.5

(b)

```
#b
#import data
urlb<-"https://www2.isye.gatech.edu/~jeffwu/wuhamadabook/data/LongJumpData.dat"
temp<-fread(urlb,sep=" ",fill=TRUE,skip=1)
#combine
colnames(temp)<-rep(c("Year","Long_Jump"),4)
data<-rbind(temp[,1:2],temp[,3:4],temp[,5:6],temp[,7:8])
#remove NA
gold_medal<-data %>% filter(is.na(Year)==FALSE | is.na(Long_Jump)==FALSE)

kable(gold_medal,caption="table of gold medal")
```

Table 2: table of gold medal

Year	Long_Jump
-4	249.75
0	282.88
4	289.00
8	294.50
12	299.25
20	281.50
24	293.13
28	304.75
32	300.75
36	317.31
48	308.00
52	298.00
56	308.25
60	319.75
64	317.75
68	350.50
72	324.50
76	328.50
80	336.25
84	336.25
88	343.25
92	342.50

(c)

```
#c
#import data
urlc<-"https://www2.isye.gatech.edu/~jeffwu/wuhamadabook/data/BrainandBodyWeight.dat"
temp<-fread(urlc,sep=" ",fill=TRUE,skip=1)
#combine
colnames(temp)<-rep(c("body_wt","brain_wt"),3)
data<-rbind(temp[,1:2],temp[,3:4],temp[,5:6])
#remove NA
weight<-data %>% filter(is.na(body_wt)==FALSE | is.na(brain_wt)==FALSE)

kable(weight, caption="table of body weight and brain weight")
```

Table 3: table of body weight and brain weight

body_wt	brain_wt
3.385	44.50
0.480	15.50
1.350	8.10
465.000	423.00
36.330	119.50
27.660	115.00
14.830	98.20
1.040	5.50
4.190	58.00
0.425	6.40
0.101	4.00
0.920	5.70
1.000	6.60
0.005	0.10
0.060	1.00
3.500	10.80
2.000	12.30
1.700	6.30
2547.000	4603.00
0.023	0.30
187.100	419.00
521.000	655.00
0.785	3.50
10.000	115.00
3.300	25.60
0.200	5.00
1.410	17.50
529.000	680.00
207.000	406.00
85.000	325.00
0.750	12.30
62.000	1320.00
6654.000	5712.00
3.500	3.90
6.800	179.00
35.000	56.00

body_wt	brain_wt
4.050	17.00
0.120	1.00
0.023	0.40
0.010	0.30
1.400	12.50
250.000	490.00
2.500	12.10
55.500	175.00
100.000	157.00
52.160	440.00
10.550	179.50
0.550	2.40
60.000	81.00
3.600	21.00
4.288	39.20
0.280	1.90
0.075	1.20
0.122	3.00
0.048	0.33
192.000	180.00
3.000	25.00
160.000	169.00
0.900	2.60
1.620	11.40
0.104	2.50
4.235	50.40

(d)

```
#d
#import data
urld<-"https://www2.isye.gatech.edu/~jeffwu/wuhamadabook/data/tomato.dat"
temp<-as.data.frame(fread(urld,sep=" ",fill=TRUE,skip=1,header=TRUE))
#split
c<-c()
for(i in 1:2){
  for(j in 2:4){
    c<-c(c,unlist(strsplit(temp[i,j],",")))
  }
}
c<-as.numeric(c)
m<-t(matrix(c,nrow=9))
#data frame
tomato<-data.frame(m)
rownames(tomato)<-temp[,1]
colnames(tomato)<-rep(c("10000","20000","30000"),each=3)

kable(tomato, caption="table of tomato data")
```

Table 4: table of tomato data

	10000	10000	10000	20000	20000	20000	30000	30000	30000
Ife#1	16.1	15.3	17.5	16.6	19.2	18.5	20.8	18.0	21.0
PusaEarlyDwarf	8.1	8.6	10.1	12.7	13.7	11.5	14.4	15.4	13.7