

R 暑假第一次作业

一、

1、张三、李四、王五、赵六的身高分别为 170cm、165cm、178cm、174cm，体重分别为 50kg、60kg、59kg、62kg，年龄分别为 13、18、45、32，创建相应的 dataframe，并将体重这一变量改为 factor 类型

2、将 factor 类型变量改为 numeric

3、求年龄为 20 以下人们的平均身高和年龄为 20 以上人们的平均体重

1

```
name=c("zhangsan","lisi","wangwu","zhaoliu")
```

```
height=c(170,165,178,174)
```

```
weight=c(50,60,59,62)
```

```
age=c(13,18,45,32)
```

```
df=data.frame(name,height,weight,age)
```

```
df$weight=as.factor(df$weight)
```

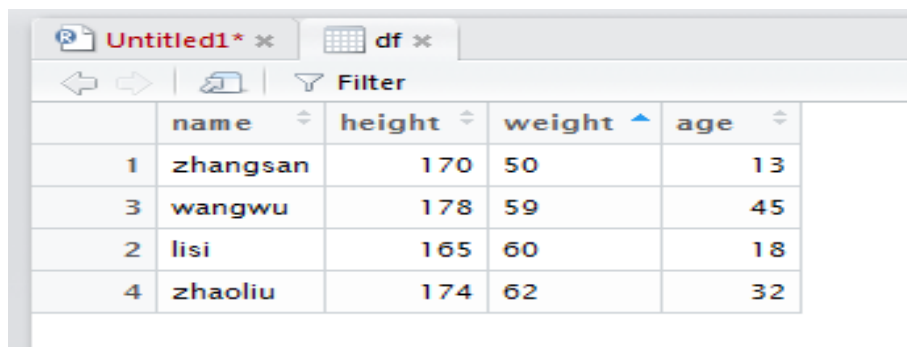
2

```
df$weight=as.numeric(df$weight)
```

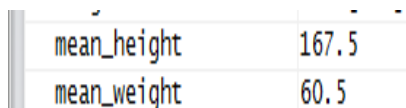
3

```
mean_height=mean(df$height[df$age<20])
```

```
mean_weight=mean(df$weight[df$age>20])
```



	name	height	weight	age
1	zhangsan	170	50	13
3	wangwu	178	59	45
2	lisi	165	60	18
4	zhaoliu	174	62	32



mean_height	167.5
mean_weight	60.5

二、

1、小张，小李，小王，小叶，四位同学的考试成绩分别为 90、87、54、51，低于 60 分为不及格，请筛选出哪几位同学不及格

2、小芬，小陆两门同学的考试成绩分别为 87、76 平时成绩分别为 70、85，总成绩为 60% 的考试成绩与 40% 的平时成绩之和，算出两门同学的总成绩

3、将 day1 赋值为 mon, tue, wen, day2 赋值为 day1+thu, fri

1

```
name1=c("xiao Zhang","xiao Li","xiao Wang","xiao Ye")
```

```
score=c(90,87,54,51)
```

```
df1=data.frame(name1,score)
```

```
df1$name1[df1$score<60]
```

2

```
xiaofen_score=87*0.6+70*0.4
```

```
xiaolu_score=76*0.6+85*0.4
```

```
> df1$name1[df1$score<60]
[1] xiao Wang xiao Ye
Levels: xiao Li xiao Wang xiao Ye xiao Zhang
> |
```

xiaofen_score	80.2
xiaolu_score	79.6

三、

小组中有三名同学，名字分别是：lieren，shushi，saman，他们的体重分别为 65，70，85，身高分别为 180，180，210。

1.构造数据框，包含三人的信息，三列分别为 name weight height。

2.现在小组又增加一名成员，名字叫 fashi，体重 45 身高 165，将该成员加入到数据框中。

3.现在 shushi 同学退出小组，将该成员的信息在数据框中删除。

1

```
name=c("lieren","shushi","saman")
```

```
weight=c(65,70,85)
```

```
height=c(180,180,210)
```

```
df=data.frame(name,weight,height)
```

2

```
name1=c(name,"fashi")
```

```
weight1=c(weight,45)
height1=c(height,165)
df1=data.frame(name1,weight1,height1)
```

	name1	weight1	height1
1	lieren	65	180
2	shushi	70	180
3	saman	85	210
4	fashi	45	165

```
3
df1=df1[-2,]
```

	name1	weight1	height1
1	lieren	65	180
3	saman	85	210
4	fashi	45	165

四、

1、形成一个以名字、性别、是否加分、分数为列的表格，现有 4 人名字分别是 A/B/C/D，性别分别为男、女、女、男，是否加分分别为是、是、否、否，分数分别为 58/59/85/90，现在将成绩进行调整，将成绩低于 60 分并且加分的人成绩加上 10 分，将成绩高于 80 分且性别为男的人分数减 5 分。

2、在已有表格的基础上，新增一名同学，名字是 E，性别女，不加分，分数 89

3、新增一列活动参与，如果是女生就填“参与”，男生就填“不参与”，让系统判断填写。

```
1
name=c("A","B","C","D")
gener=c("male","female","female","male")
add_score=c("yes","yes","no","no")
score=c(58,59,85,90)
df=data.frame(name,gener,add_score,score)
```

```
df$score[df$score<60&df$add_score=="yes"]=df$score[df$score<60&df$add_score=="yes"]+10
df$score[df$score>80&df$gener=="male"]=df$score[df$score>80&df$gener=="male"]-5
```

	name	gener	add_score	score
1	A	male	yes	68
2	B	female	yes	69
3	C	female	no	85
4	D	male	no	85

2

```
name1=c(name,"E")
gener1=c(gener,"female")
add_score1=c(add_score,"no")
score1=c(score,89)
df=data.frame(name=name1,gener=gener1,add_score=add_score1,score=score1)
df$score[df$score<60&df$add_score=="yes"]=df$score[df$score<60&df$add_score=="yes"]+10
df$score[df$score>80&df$gener=="male"]=df$score[df$score>80&df$gener=="male"]-5
```

	name	gener	add_score	score
1	A	male	yes	68
2	B	female	yes	69
3	C	female	no	85
4	D	male	no	85
5	E	female	no	89

3

```
df$attend=rep("yes")
df$attend[df$gener=="male"]=df$attend[df$gener=="male"]="no"
```

	name	gender	add_score	score	attend
1	A	male	yes	68	no
2	B	female	yes	69	yes
3	C	female	no	85	yes
4	D	male	no	85	no
5	E	female	no	89	yes

五、

1. 有一个班里 5 名学生的成绩分别为 61, 75, 90, 82, NA; 由于有一名学生成绩不翼而飞,老师决定用其他 4 名学生的平均分代替。最后将这 5 名学生的成绩保存到变量 `score` 中。
2. 老师突发奇想决定将这 5 名学生的成绩分别重复 4, 2, 1, 4, 2 次, 然后将它保存到 `score1` 变量中。(`rep()`, 使用 `help` 深入了解这个函数)
3. 老师觉得这 5 名学生成绩不真实, 于是决定调整成绩。最终成绩等于原始成绩的 70% 加上平时成绩, 他们的平时成绩分别为 30, 24, 33, 28, 25, 求这 5 名学生的最终成绩并保存到变量 `score2` 中。

1.

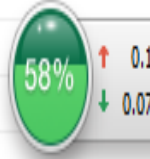
```
a=mean(61,75,90,82)
score =c(61,75,90,82,a)
```

2.

```
score1=rep(score,times=c(4,2,1,4,2))
```

3.

```
pingshi =c(30,24,33,28,25)
score2=score*0.7+pingshi
```

Global Environment		Q
Values		
a	61	
pingshi	num [1:5] 30 24 33 28 25	
score	num [1:5] 61 75 90 82 61	
score1	num [1:13] 61 61 61 61 75 75 90 82 82 82 ...	
score2	num [1:5] 72.7 76.5 96 85.4 67.7	

六，

1.现班里有四名同学 A、B、C、D 在进行体育测验。其中 A、B 为女生，C、D 为男生。测验项目有一分钟仰卧起坐、100 米跑步、跳远。经过测试，他们仰卧起坐的个数依次为 30、55、55、70，100 米跑步时间依次为 11 秒、13 秒、10 秒、9 秒，跳远的长度依次为 2 米、2.5 米、3.5 米、4 米。请构建适宜的 dataframe 来记录结果。

2.体育测试最终成绩统计方法如下：①女生仰卧起坐 60 个为 100 分，少一个减一分；男生仰卧起坐 70 个为满分，少一个减一分。②女生 100 米跑步 10 秒 100 分，多一秒减 10 分；男生 9 秒满分，多一秒减 10 分。③女生 3 米 100 分，少 0.5 米减 10 分，男生 4.5 米满分，少 0.5 米减 10 分。总成绩由三个项目的平均分组成，请新建一个适宜的 dataframe 来记录成绩，并新增一列记录总成绩。

3.B 嘲笑 C 做的仰卧起坐个数跟她一样，于是 C 气势汹汹地跑去再测一次，并打赌如果能做到 70 个，B 要自降 1 分，如果不能做到 70 个，他要在更新后成绩自降 5 分。因为 C 从来没有做到 70 个过，于是 B 答应了。结果 C 做了 69 个。请更新两个数据框里的数据。

1.

```
name=c("A","B","C","D")
gender=c("F","F","M","M")
yangwo=c(30,55,55,70)
run=c(11,13,10,9)
jump=c(2,2.5,3.5,4)
result=data.frame(name,gender,yangwo,run,jump)
```

	name	gender	yangwo	run	jump
1	A	F	30	11	2.0
2	B	F	55	13	2.5
3	C	M	55	10	3.5
4	D	M	70	9	4.0

2.

```
t1=c(70,95,85,100)
```

```
t2=c(90,70,90,100)
```

```
t3=c(80,90,80,90)
```

```
total=t1+t2+t3
```

```
result=cbind(result,total)
```

	name	gender	yangwo	run	jump	total
1	A	F	30	11	2.0	240
2	B	F	55	13	2.5	255
3	C	M	55	10	3.5	255
4	D	M	70	9	4.0	290

3

```
result$total[result$name=="C"]=result$total[result$name=="C"]-5
```

	name	gender	yangwo	run	jump	total
1	A	F	30	11	2.0	240
2	B	F	55	13	2.5	255
3	C	M	55	10	3.5	250
4	D	M	70	9	4.0	290

七，

1.应用题：小明去年身高 165cm，今年长高了 30%，请问：小明今年身高是多少？

```
height=165*1.3
```

2.将一个小组中四个人 aa，bb，cc，dd 的名字保存到向量类型 names 中，将四个人的体重 65，66，64，63 存储到变量 weight 中，并构成数据框 dataframe，此数据框由两列构成。

```
names=c("aa","bb","cc","dd")
```

```
weight=c(65,66,64,63)
```

```
df=data.frame(names,weight)
```

3.利用 seq（）函数，生成一系列数字[100，1000，10000，100000],并将结果保存到 result（）中。

```
a=10
```

```
i=seq(2,5,1)
```

```
result=a^i
```

八，

1.请使用函数生成下面几组数据

(1) 1,2,3……100

(2) 2,4,6,……100

(3) 1,1,1,2,2,3,3,3,3

```
seq(1,100,1)
```

```
seq(2,100,2)
```

```
rep(seq(1,3,1),times=c(3,2,4))
```

2.(1)生成数据框，name xiaoli xiaowang xiaohong;height 188 177 166;

(2)添加一行数据，score 66 77 88 生成新的数据框

(3) 将 height,score 中的数据类型改为 numeric

(4) 将小李的分数改为 99，更新数据框。

```
name=c("xiaoli","xiaowang ","xiaohong ")
```

```
height=c(188,177,166)
```

```
df=data.frame(name,height)
```

```
score=c(66,77,88)
```

```
df1=cbind(df,score)
```

```
height=as.numeric(height)
```



```
score=as.numeric(score)
df1$score[df1$name=="xiaoli"]=99
```

3. 在 2 题最后生成的数据框中，请找出（1）分数最低的同学姓名，（2）身高最高的同学姓名（3）身高超过 170，且成绩及格的同学姓名

```
X=df1$name[df1$score==min(df1$score)];
Y=df1$name[df1$height==max(df1$height)];
Z=df1$name[df1$height>170&df1$score>60];
```