

# 第五讲作业

---

生信 2001 张子栋 2020317210101

GitHub 地址: [MarkdownNotes/R at main · Bluuur/MarkdownNotes \(github.com\)](https://github.com/Bluuur/MarkdownNotes).

1. 有一个外部文件为 `class.txt`, 存储的是某个班级学生的姓名, 年龄, 身高, 体重和性别的信息

(1) 把此文件读入到 R 中并把他记录为数据框 `class`

```
1 | class <- read.table('/home/ubuntu/R_course/R_data/class.txt')
2 | class
```

A data.frame: 40 × 5

V1	V2	V3	V4	V5
<chr>	<int>	<int>	<dbl>	<chr>
Lawrence	17	172	78.1	M
Jeffery	14	169	51.3	M
Edward	14	167	50.8	M
Phillip	16	167	58.1	M
Kirk	17	167	60.8	M
Robert	15	164	58.1	M
Jaclyn	12	162	65.8	F
Danny	15	162	48.1	M
Clay	15	162	47.7	M
Henry	14	159	54.0	M
Leslie	14	159	64.5	F
John	13	159	44.5	M
William	15	159	50.4	M
Martha	16	159	50.8	F
Lewis	14	157	41.8	M
Amy	15	157	50.8	F
Alfred	14	157	44.9	M
Chris	14	157	44.9	M
Fredrick	14	154	42.2	M
Carol	14	154	38.1	F
Joe	13	154	47.7	M
Mary	15	152	41.8	F
Linda	17	152	52.7	F
Mark	15	152	47.2	M
Patty	14	152	38.6	F
Elizabet	14	152	41.3	F
Judy	14	149	36.8	F
Louis	12	149	55.8	F
Alice	13	149	48.6	F

V1	V2	V3	V4	V5
<chr>	<int>	<int>	<dbl>	<chr>
James	12	149	58.1	M
Marian	16	147	52.2	F
Tim	12	147	38.1	M
Barbara	13	147	50.8	F
David	13	145	35.9	M
Katie	12	145	43.1	F
Michael	13	142	43.1	M
Susan	13	137	30.4	F
Jane	12	135	33.6	F
Lillie	12	127	29.1	F
Robert	12	125	35.9	M

(2) 给数据框 class 的列命名为 name, age, height, weight, sex

```

1 class <- read.table('/home/ubuntu/R_course/R_data/class.txt')
2 names(class)[1:5] <- c('name', 'age', 'height', 'weight', 'sex')
3 class

```

A data.frame: 40 × 5

name	age	height	weight	sex
<chr>	<int>	<int>	<dbl>	<chr>
Lawrence	17	172	78.1	M
Jeffery	14	169	51.3	M
Edward	14	167	50.8	M
Phillip	16	167	58.1	M
Kirk	17	167	60.8	M
Robert	15	164	58.1	M
Jaclyn	12	162	65.8	F
Danny	15	162	48.1	M
Clay	15	162	47.7	M
Henry	14	159	54.0	M
Leslie	14	159	64.5	F
John	13	159	44.5	M
William	15	159	50.4	M
Martha	16	159	50.8	F
Lewis	14	157	41.8	M
Amy	15	157	50.8	F
Alfred	14	157	44.9	M
Chris	14	157	44.9	M
Fredrick	14	154	42.2	M
Carol	14	154	38.1	F
Joe	13	154	47.7	M
Mary	15	152	41.8	F
Linda	17	152	52.7	F
Mark	15	152	47.2	M
Patty	14	152	38.6	F
Elizabet	14	152	41.3	F
Judy	14	149	36.8	F
Louis	12	149	55.8	F
Alice	13	149	48.6	F

name	age	height	weight	sex
<chr>	<int>	<int>	<dbl>	<chr>
James	12	149	58.1	M
Marian	16	147	52.2	F
Tim	12	147	38.1	M
Barbara	13	147	50.8	F
David	13	145	35.9	M
Katie	12	145	43.1	F
Michael	13	142	43.1	M
Susan	13	137	30.4	F
Jane	12	135	33.6	F
Lillie	12	127	29.1	F
Robert	12	125	35.9	M

(3) 添加一个根据年龄分组的变量 rank, 分成两组:

- low ( $\leq 14$  岁)
- high ( $\geq 15$  岁)

```

1 class <- read.table('/home/ubuntu/R_course/R_data/class.txt')
2 names(class)[1:5] <- c('name', 'age', 'height', 'weight', 'sex')
3 class$rank <- ifelse(class$age > 14, 'high', 'low')
4 class

```

A data.frame: 40 × 6

name	age	height	weight	sex	rank
<chr>	<int>	<int>	<dbl>	<chr>	<chr>
Lawrence	17	172	78.1	M	high
Jeffery	14	169	51.3	M	low
Edward	14	167	50.8	M	low
Phillip	16	167	58.1	M	high
Kirk	17	167	60.8	M	high
Robert	15	164	58.1	M	high
Jaclyn	12	162	65.8	F	low
Danny	15	162	48.1	M	high
Clay	15	162	47.7	M	high
Henry	14	159	54.0	M	low
Leslie	14	159	64.5	F	low
John	13	159	44.5	M	low
William	15	159	50.4	M	high
Martha	16	159	50.8	F	high
Lewis	14	157	41.8	M	low
Amy	15	157	50.8	F	high
Alfred	14	157	44.9	M	low
Chris	14	157	44.9	M	low
Fredrick	14	154	42.2	M	low
Carol	14	154	38.1	F	low
Joe	13	154	47.7	M	low
Mary	15	152	41.8	F	high
Linda	17	152	52.7	F	high
Mark	15	152	47.2	M	high
Patty	14	152	38.6	F	low
Elizabet	14	152	41.3	F	low
Judy	14	149	36.8	F	low
Louis	12	149	55.8	F	low
Alice	13	149	48.6	F	low

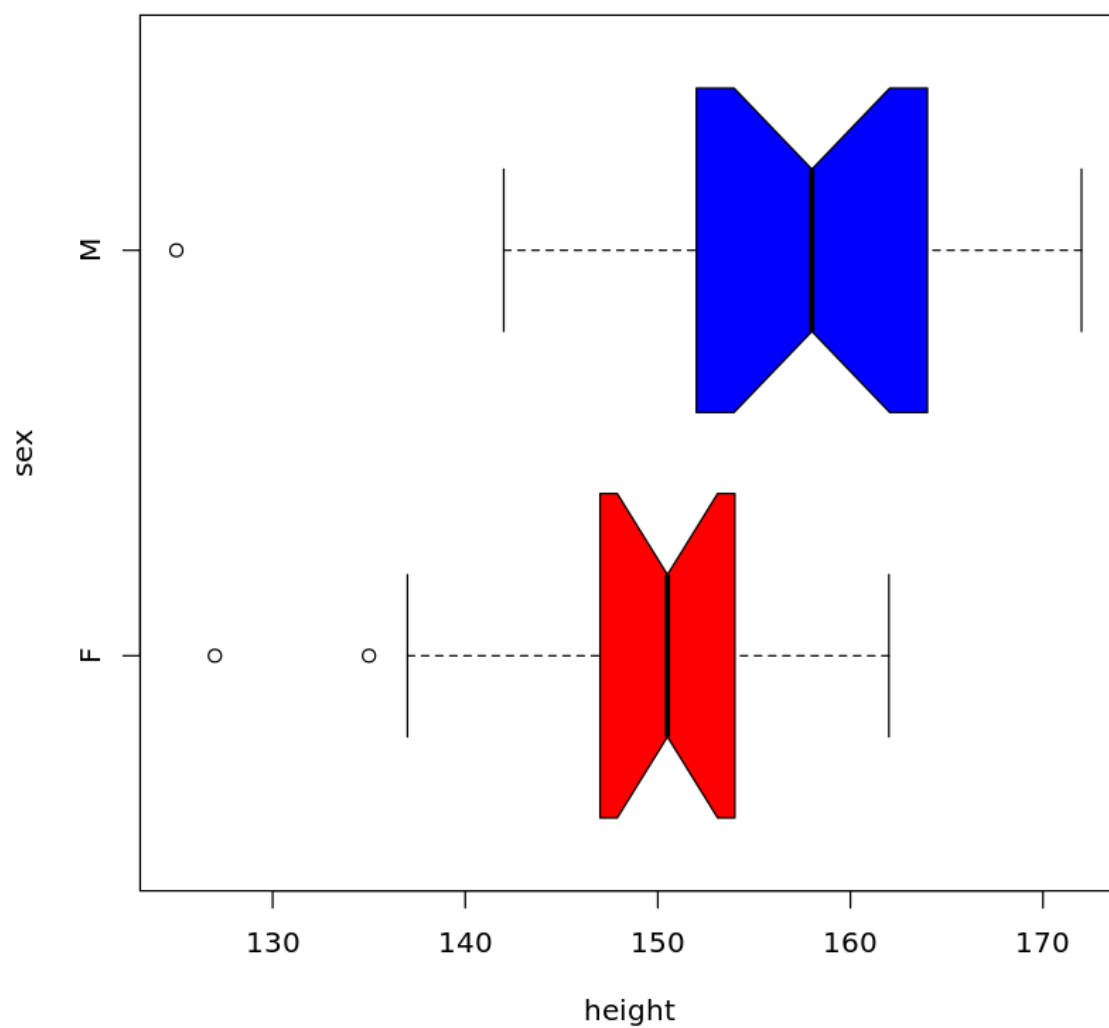
name	age	height	weight	sex	rank
<chr>	<int>	<int>	<dbl>	<chr>	<chr>
James	12	149	58.1	M	low
Marian	16	147	52.2	F	high
Tim	12	147	38.1	M	low
Barbara	13	147	50.8	F	low
David	13	145	35.9	M	low
Katie	12	145	43.1	F	low
Michael	13	142	43.1	M	low
Susan	13	137	30.4	F	low
Jane	12	135	33.6	F	low
Lillie	12	127	29.1	F	low
Robert	12	125	35.9	M	low

(4) 按性别画出体重的箱型图(水平, 凹槽, 颜色控制等) (第 1 图)

```

1 class <- read.table('/home/ubuntu/R_course/R_data/class.txt')
2 names(class)[1:5] <- c('name', 'age', 'height', 'weight', 'sex')
3 class$rank <- ifelse(class$age > 14, 'high', 'low')
4 with(class, boxplot(height ~ sex,
5                     horizontal = T,
6                     notch = T,
7                     col = c('red', 'blue'))))

```



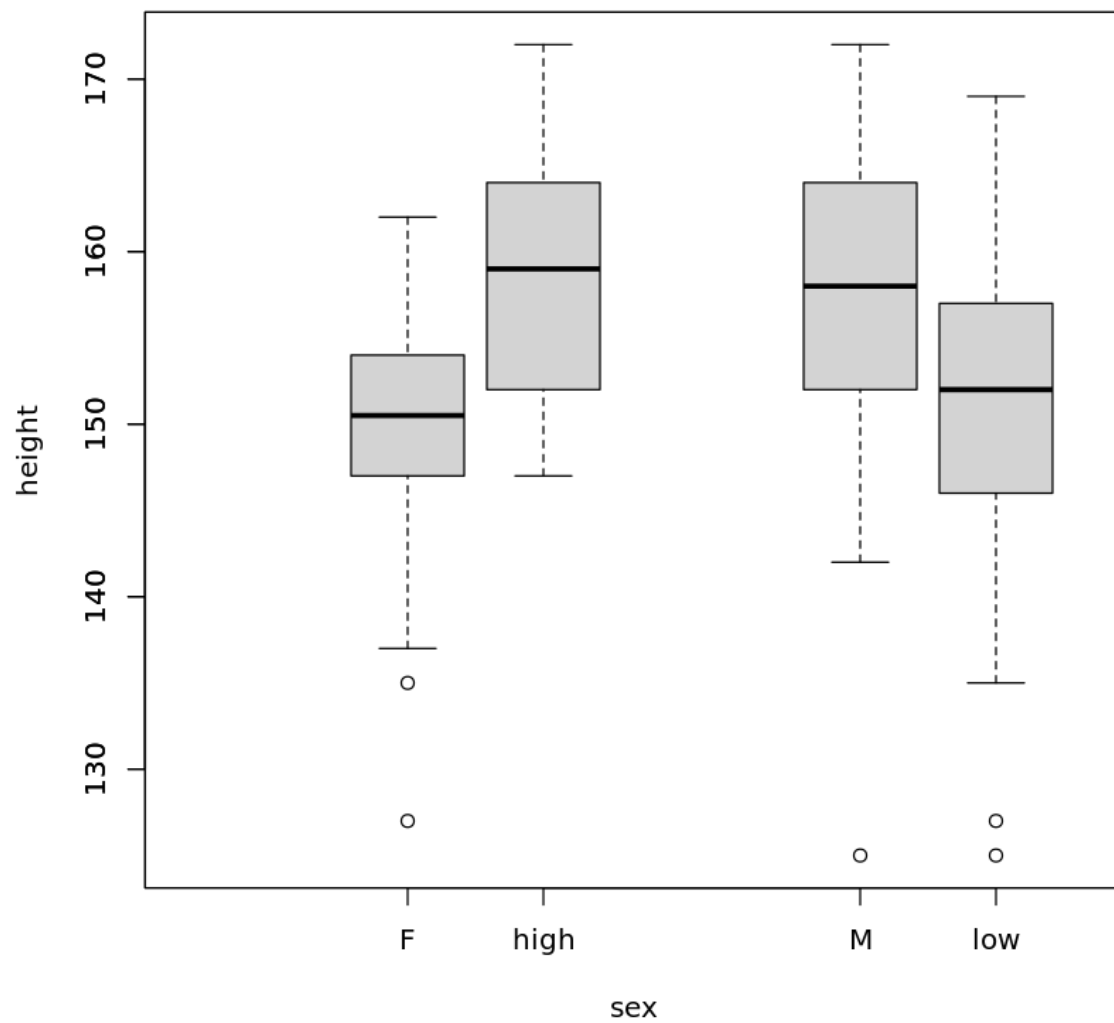
(5) 根据性别及年龄分组画出箱型图 (第二图)

```

1 class <- read.table('/home/ubuntu/R_course/R_data/class.txt')
2 names(class)[1:5] <- c('name', 'age', 'height', 'weight', 'sex')
3 class$rank <- ifelse(class$age > 14, 'high', 'low')
4 with(class, boxplot(height ~ sex, boxwex = 0.25, at = 1:2 - 0.1))
5 with(class, boxplot(height ~ rank, boxwex = 0.25, at = 1:2 + 0.2, add = T))

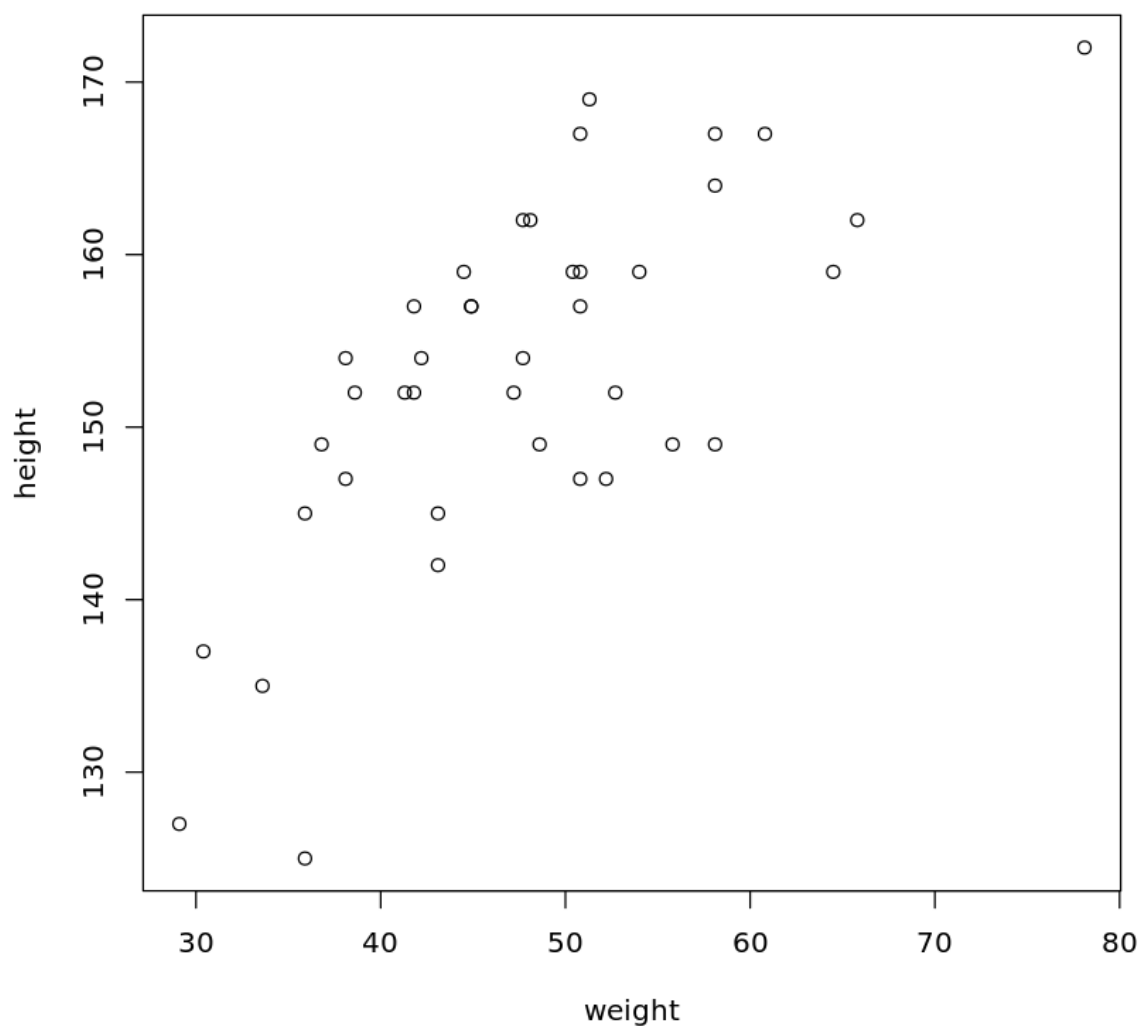
```





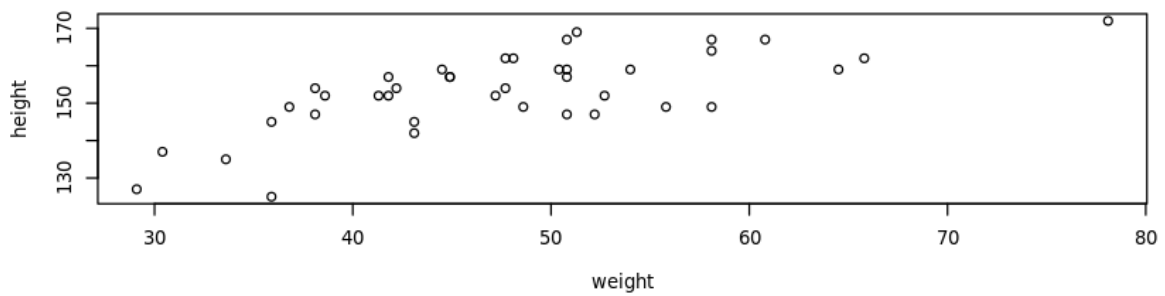
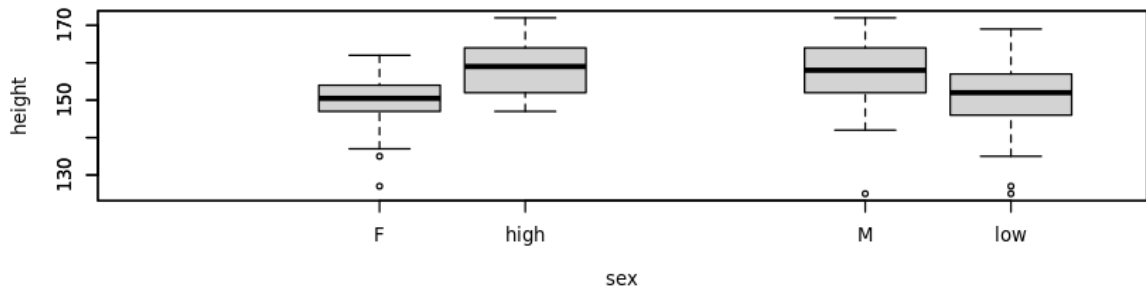
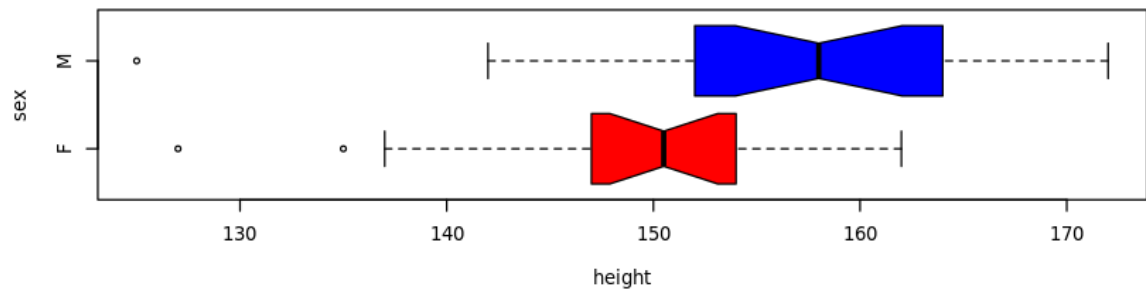
(6) 画出体重与身高关系的散点图

```
1 class <- read.table('/home/ubuntu/R_course/R_data/class.txt')
2 names(class)[1:5] <- c('name', 'age', 'height', 'weight', 'sex')
3 plot(class$weight, class$height, xlab = 'weight', ylab = 'height')
```



(7) 通过设置参数 `mfrow` 将所有图放到同一页中并保存为一个 `png` 文件

```
1 par(mfrow = c(3, 1))
2 class <- read.table('/home/ubuntu/R_course/R_data/class.txt')
3 names(class)[1:5] <- c('name', 'age', 'height', 'weight', 'sex')
4 class$rank <- ifelse(class$age > 14, 'high', 'low')
5 with(class, boxplot(height ~ sex,
6                     horizontal = T,
7                     notch = T,
8                     col = c('red', 'blue'))))
9 with(class, boxplot(height ~ sex, boxwex = 0.25, at = 1:2 - 0.1))
10 with(class, boxplot(height ~ rank, boxwex = 0.25, at = 1:2 + 0.2, add = T))
11 plot(class$weight, class$height, xlab = 'weight', ylab = 'height')
```

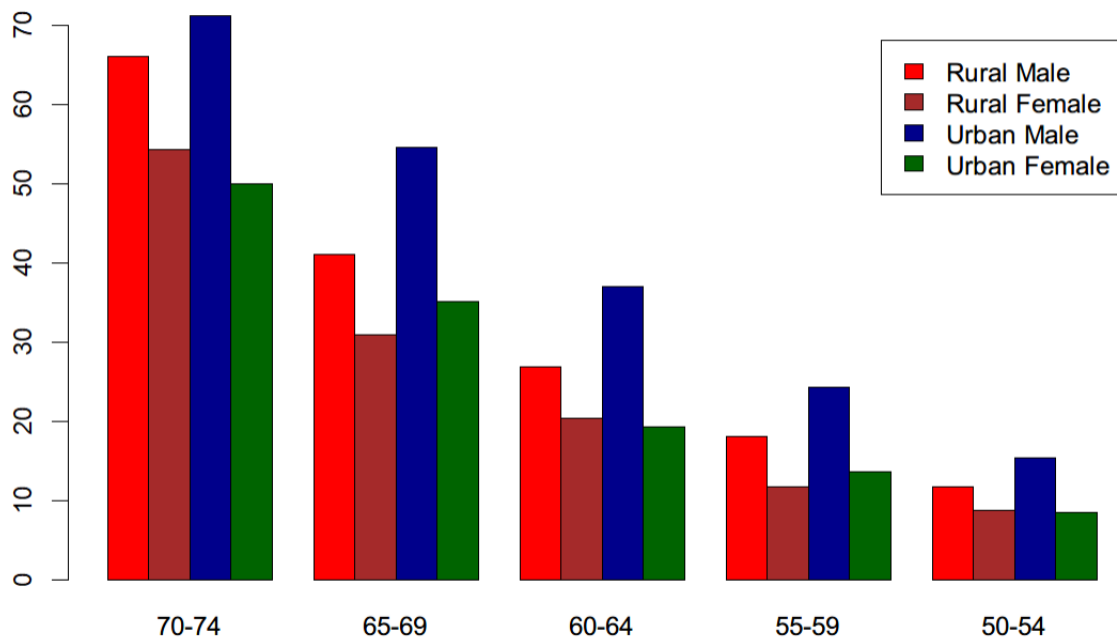


2. `VADeaths` 是 1940 年弗吉尼亚州分年龄组, 分地区和分性别死亡率数据. 画出按年龄进行分组的不同性别及地区的条形图.

```
1 data <- VADeaths
2 data <- t(data)
3 data
4 barplot(data[, 5:1], beside = T, legend = T, col = c('red', 'brown',
  'darkblue', 'darkgreen'), ylim = c(0, 70))
```

A matrix: 4 × 5 of type dbl

	50-54	55-59	60-64	65-69	70-74
<b>Rural Male</b>	11.7	18.1	26.9	41.0	66.0
<b>Rural Female</b>	8.7	11.7	20.3	30.9	54.3
<b>Urban Male</b>	15.4	24.3	37.0	54.6	71.1
<b>Urban Female</b>	8.4	13.6	19.3	35.1	50.0



3.下面左右两组图是同一套数据的不同呈现方式

(1) 左右两组图的类别名称各是什么?

左: 饼图

右: 条形图

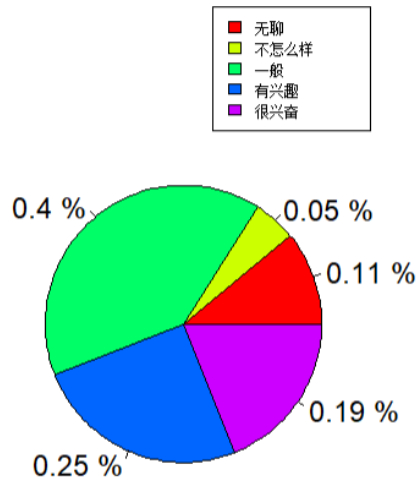
(2) 用 R 复现这两组图形(不要求颜色和文字完全一致)

```

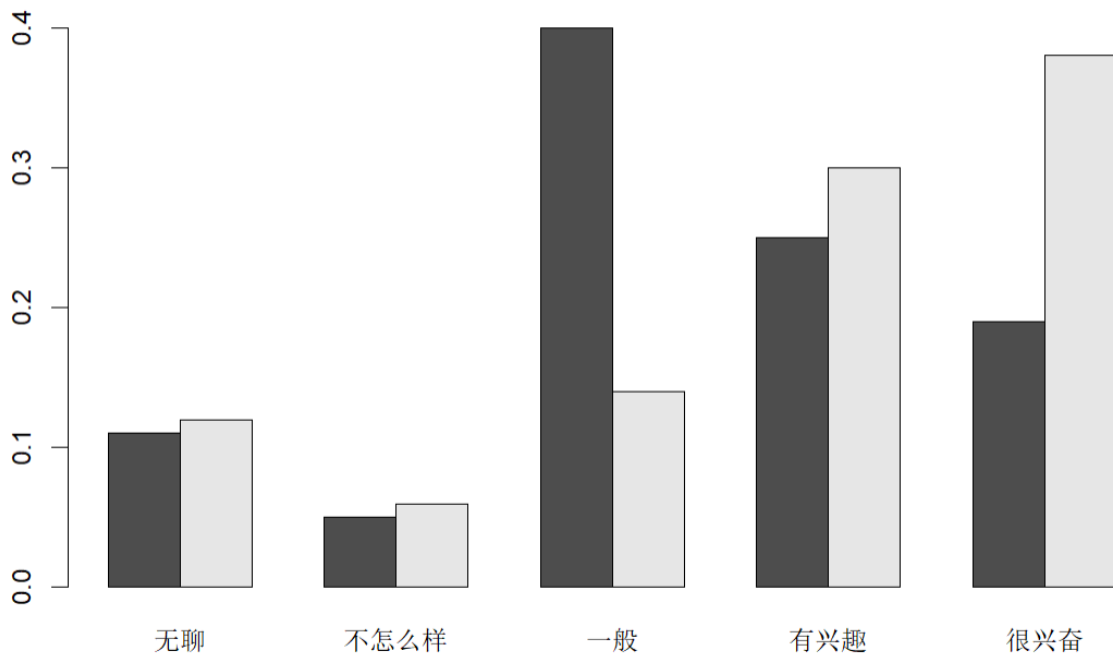
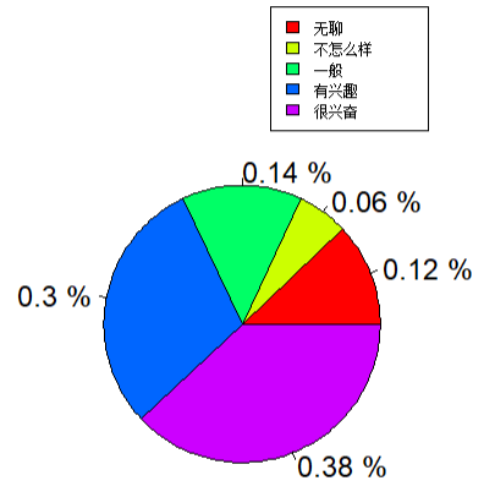
1 data1 <- c(0.11, 0.05, 0.4, 0.25, 0.19)
2 data2 <- c(0.12, 0.06, 0.14, 0.30, 0.38)
3
4 percent1 <- paste(data1, '%')
5 percent2 <- paste(data2, '%')
6 atti <- c('无聊', '不怎么样', '一般', '有兴趣', '很兴奋')
7 inOne <- rbind(data1, data2)
8 colnames(inOne) <- atti
9 par(mfrow = c(1, 2))
10 pie(inOne[1,], labels = percent1, main = 'before', col =
  rainbow(length(data1)))
11 legend("topright", atti, cex = 0.5, fill = rainbow(length(data1)))
12 pie(inOne[2,], labels = percent2, main = 'after', col =
  rainbow(length(data1)))
13 legend("topright", atti, cex = 0.5, fill = rainbow(length(data1)))
14 par(mfrow = c(1, 1))
15 barplot(inOne, beside = T)

```

before



after



(3) 在这里, 从讲故事和传达信息的角度看, 那种图是更适合的呈现方式? 为什么?

- 柱状图更适合
  - 这里关注项目前后对科学的态度变化
  - 前后相邻, 对比更明显

4.访问网站 重复动图

```

1 library(gapminder)
2 library(ggplot2)
3 library(gganimate)
4
5 ggplot(gapminder, aes(gdpPercap, lifeExp, size = pop, color = continent)) +
6   geom_point() +
7   scale_x_log10() +
8   theme_bw() +
9   labs(title = 'Year: {frame_time}', x = 'GDP per capita', y = 'life
expectancy') +
10  transition_time(year) +
11  ease_aes('linear')
12 anim_save("271-ggplot2-animated-gif-chart-with-gganimate1.gif")

```

