期末考核问题

JohnLu

今天是2021-06-24

（把代码和结果粘贴下边，不要截图！！）

1.R语言 最早始于哪一年?

1993

2.R语言为什么命名为R语言?

最开始的两个开发者名首字母是R

3.R语言是基于什么语言开发的?

S语言

4.R语言的两大基本功能是?

统计计算，图形输出

5.Rstudio和R的关系是?

统计计算，图形输出

6.以tidyverse包为例，列举R两种安装包的方式（截图语言描述都可以）?

在Rstudio中使用Install Package/直接用install函数

7.R语言有哪些数据类型?

逻辑，浮点，整数，字符，复数

8.R语言的数据结构类型有哪些?

向量，矩阵，数据框

9.R中的向量是什么?

同一种数据的有序集合

10.各列举三条R语言的优势和劣势

优势：方便调库，语法为方便数学家使用(如下标从1开始)，方便绘图(样式多，美观)

劣势：慢，中文支持不佳(编码问题)，开源导致代码质量参差不齐

11.实际操作–rivers是R自带的北美河流长度信息，请以rivrs为例，求出哪一个河流（序号）最长?

```
which.max(rivers)
```

```
## [1] 68
```

12.实际操作–以mtcars为例，除am、cyl和vs之外求所有量的平均值

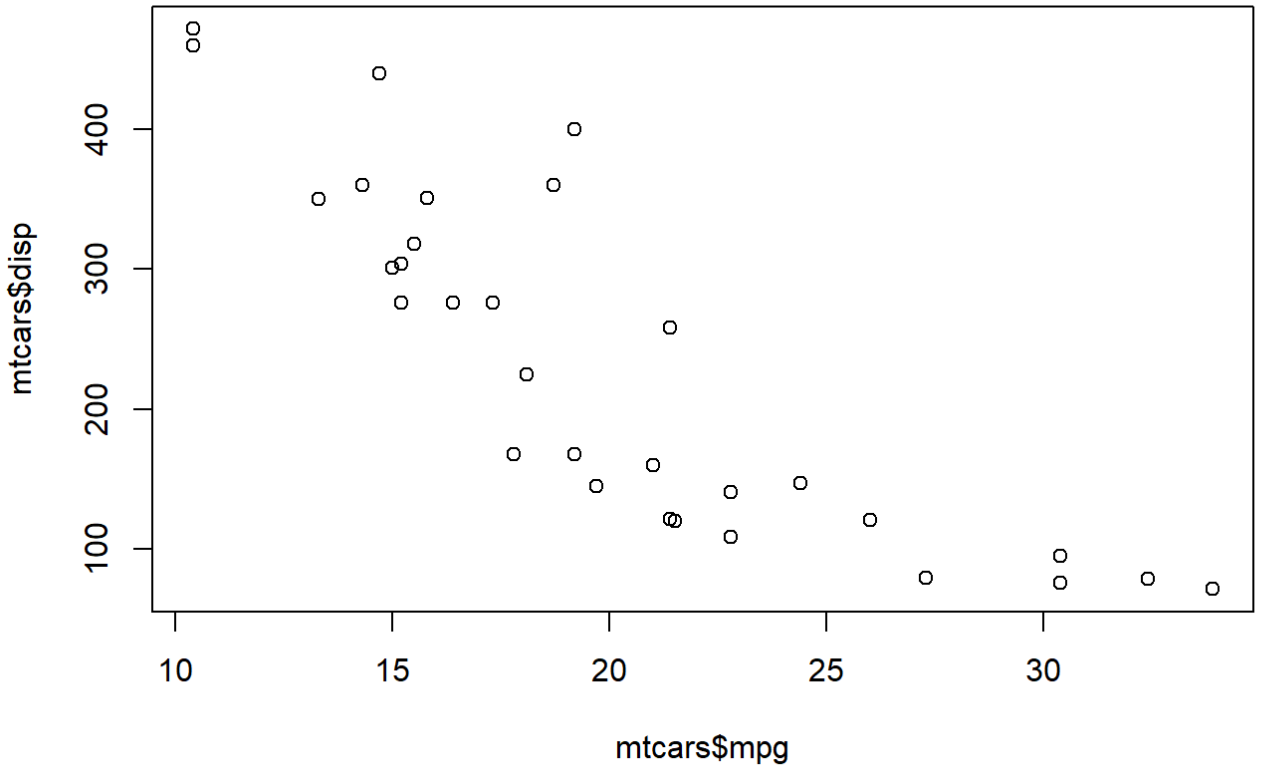
```
colnames(mtcars)
apply(mtcars[-c(2,8,9)],2,mean)
```

```
##  [1] "mpg"  "cyl"  "disp" "hp"   "drat" "wt"   "qsec" "vs"   "am"   "gear" "car
b"
```

```
##        mpg       disp         hp       drat         wt       qsec       gear
carb
##  20.090625 230.721875 146.687500   3.596563   3.217250  17.848750   3.687500
2.812500
```

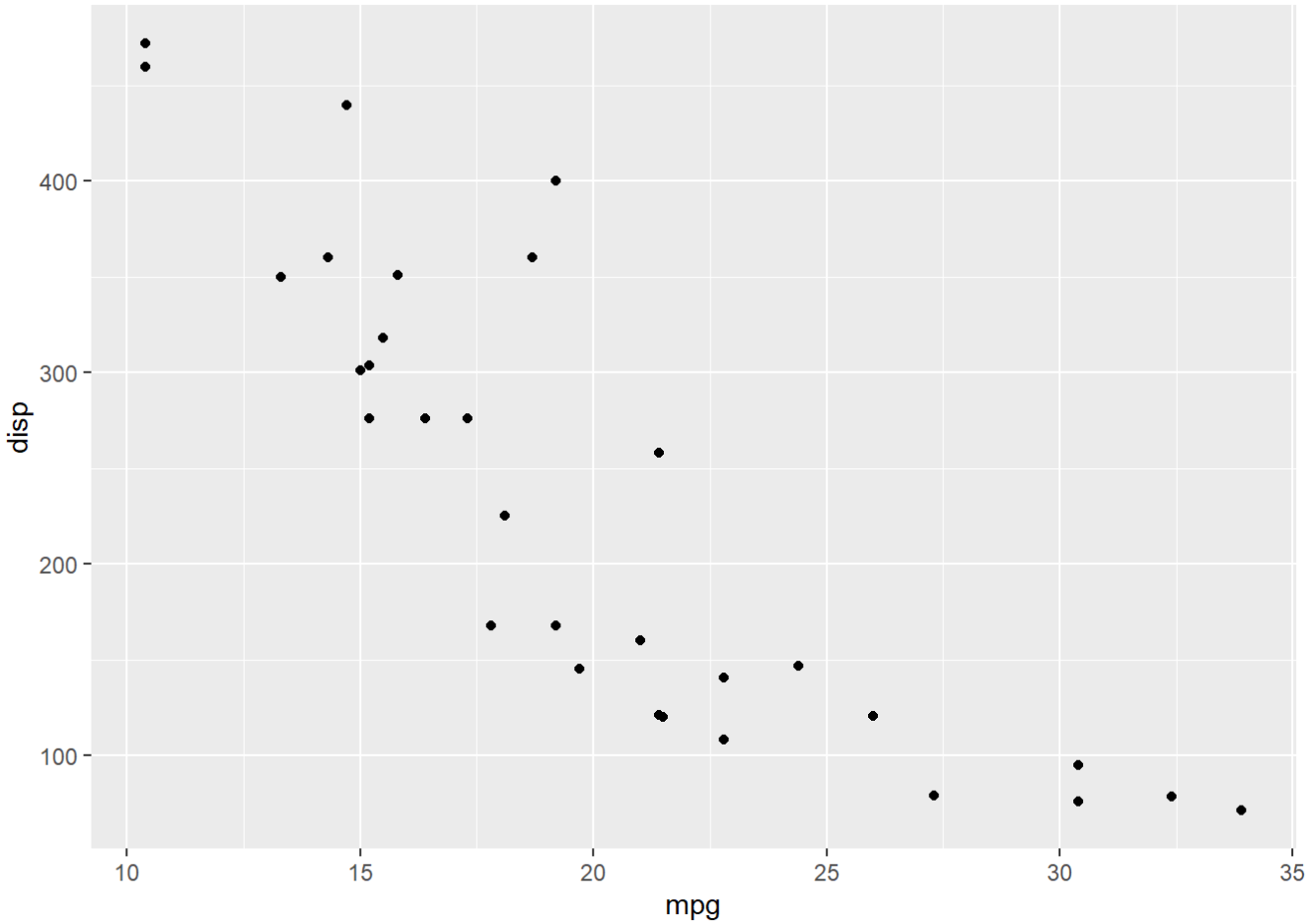
13.实际操作–以mtcars为例，分别用plot和ggplot函数绘mpg与disp关系的散点图

```
plot(mtcars$mpg, mtcars$disp)
```




```
library(ggplot2)
ggplot(mtcars,aes(mpg,disp))+geom_point()
```

```
## Learn more about the underlying theory at https://ggplot2-book.org/
```



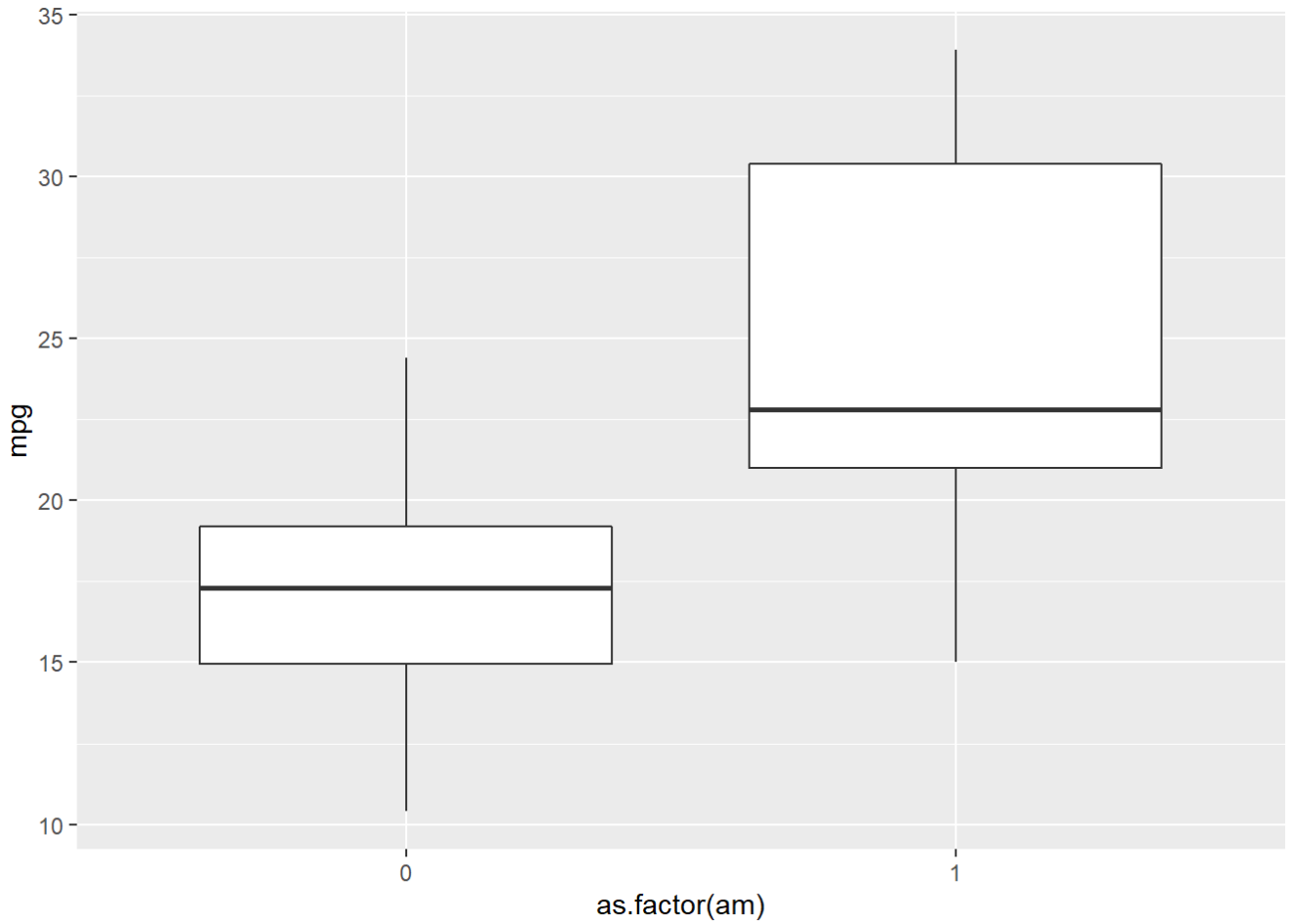
14.实际操作–随机生成100个均值为1，标准差为2的正态随数

```
rnorm(100,1,2)
```

```
##   [1] -5.867492945 -0.005839735  2.844053113  3.453403332  0.348225763 -0.367725
286 -0.565543886  1.097821774 -0.947164609  3.654288767  0.533973646
##  [12] -0.623420320  1.860283937  1.495117257  2.842524101 -0.302458595  2.597103
172  3.772143711 -1.346393137 -0.018548991  0.061115898  2.020602302
##  [23]  1.585364137 -0.350968391 -0.580414033  2.240953566  3.389156429  0.581695
489  2.819044532  0.855041788 -0.269662121  0.304686084  1.357658484
##  [34]  2.203948760  1.896712731  1.145288561 -2.282885701  1.122882553  1.411354
785  1.003490506  3.078973202 -0.618506058 -0.654855756  1.270629312
##  [45] -0.766261866  0.366714666 -0.987377101 -2.601498809 -0.433638570 -1.661756
606  4.094689911  2.763594713  3.797362885  1.840056328 -2.993422585
##  [56]  0.435001669 -1.552729101 -0.962979914  2.269660410  1.041922050  2.065464
568 -0.701844873  0.668874167  3.133542688  2.467379301 -0.577184855
##  [67]  2.350828177  5.672768750  0.728474906  1.730468137  1.390387814 -0.674577
558  1.404850876 -0.474438508  3.730438385  4.083220922  0.906856798
##  [78] -0.928542059 -3.013978552  3.802265134  1.104071163  0.603600672 -0.807901
329  1.424392424  1.464790608 -0.161508299  0.211034804  0.575141253
##  [89]  2.975206913  1.598739335  0.962186420  0.437059391  2.012882835  0.427667
622  1.022693824  0.460253636 -0.030883736 -0.568913215  2.744036787
## [100]  0.816548967
```

15.实际操作–利用ggplot并以mtcars数据集为例，演示如生成箱式图

```
library(ggplot2)
ggplot(mtcars,aes(as.factor(am),mpg))+geom_boxplot()
```



16.实际操作–利用for循环实现从1开始到100的累加运算

```
s=0
for(i in 1:100){
    s <- s+i
}
s
```

```
## [1] 5050
```

17.实际操作–利用t检验检验mtcars中手动挡和自动挡的车mg是否有显著性差异

```
t.test(mtcars$mpg~mtcars$am)
```

```
##
##  Welch Two Sample t-test
##
## data:  mtcars$mpg by mtcars$am
## t = -3.7671, df = 18.332, p-value = 0.001374
## alternative hypothesis: true difference in means between group 0 and group 1 is
not equal to 0
## 95 percent confidence interval:
##  -11.280194  -3.209684
## sample estimates:
## mean in group 0 mean in group 1
##        17.14737        24.39231
```

18.实际操作–以diamonds数据为例，按照color和claity分类计算平均价格

```
library(tidyverse)
diamonds %>% group_by(color,clarity) %>% summarise(mean(price))
```

```
## -- Attaching packages
--------------------------------------------------------------------------------
------------------------- tidyverse 1.3.1 --
```

```
## v tibble  3.1.2     v dplyr   1.0.6
## v tidyr   1.1.3     v stringr 1.4.0
## v readr   1.4.0     v forcats 0.5.1
## v purrr   0.3.4
```

```
## -- Conflicts
-------------------------------------------------------------------------------
---------------------------- tidyverse_conflicts() --
## x dplyr::filter() masks stats::filter()
## x dplyr::lag()    masks stats::lag()
```

```
## `summarise()` has grouped output by 'color'. You can override using the `.groups
` argument.
```

```
## # A tibble: 56 x 3
## # Groups:   color [7]
##    color clarity `mean(price)`
##    <ord> <ord>           <dbl>
##  1 D     I1              3863.
##  2 D     SI2             3931.
##  3 D     SI1             2976.
##  4 D     VS2             2587.
##  5 D     VS1             3030.
##  6 D     VVS2            3351.
##  7 D     VVS1            2948.
##  8 D     IF              8307.
##  9 E     I1              3488.
## 10 E     SI2             4174.
## # ... with 46 more rows
```

19.实际操作–利用R求下面矩阵行列式的值，并计算该矩阵逆\(\begin{bmatrix} 6&2\\ 2&5 \end{bmatrix}\)

```
m <- matrix(c(6,2,2,5),2)
det(m)
solve(m)
```

```
## [1] 26
```

```
##             [,1]        [,2]
## [1,]  0.19230769 -0.07692308
## [2,] -0.07692308  0.23076923
```

20.实际操作–利用R求下面矩阵的单位特征向量和特征根并验证两个特征值对应的特征向量是正交的\(\begin{bmatrix} 6&2\\ 2&5 \end{bmatrix}\)

```
m <- matrix(c(6,2,2,5),2)
eigen(m)
a <- eigen(m)$vectors
a[1, ] %*% a[2, ]
```

```
## eigen() decomposition
## $values
## [1] 7.561553 3.438447
##
## $vectors
##            [,1]       [,2]
## [1,] -0.7882054  0.6154122
## [2,] -0.6154122 -0.7882054
```

```
##      [,1]
## [1,]    0
```