

talk05 练习与作业

目录

练习和作业说明	1
Talk05 内容回顾	1
练习与作业：用户验证	1
练习与作业 1: dplyr 练习	2

练习和作业说明

将相关代码填写入以 “{r}” 标志的代码框中，运行并看到正确的结果；

完成后，用工具栏里的 “Knit” 按键生成 PDF 文档；

将 PDF 文档改为：姓名-学号-talk05 作业.pdf，并提交到老师指定的平台/钉群。

Talk05 内容回顾

- dplyr 、tidyr (超级强大的数据处理) part 1
 - 长宽数据转换
 - dplyr 几个重要函数

练习与作业：用户验证

请运行以下命令，验证你的用户名。

如你当前用户名不能体现你的真实姓名，请改为拼音后再运行本作业！

```
Sys.getenv("USER")
```

```
## [1] "wchen"
```

```
Sys.getenv("HOME")
```

```
## [1] "/Users/wchen"
```

练习与作业 1: dplyr 练习

使用 mouse.tibble 变量做统计

- 每个染色体上每种基因类型的数量、平均长度、最大和最小长度，挑出最长和最短的基因
- 去掉含有 500 以下基因的染色体，按染色体、数量高 -> 低进行排序

```
## 代码写这里，并运行；
```

使用 grades 变量做统计

首先，用下面命令生成 grades 变量：

```
grades <- tibble( "Name" = c("Weihua Chen", "Mm Hu", "John Doe", "Jane Doe",  
                             "Warren Buffet", "Elon Musk", "Jack Ma"),  
                 "Occupation" = c("Teacher", "Student", "Teacher", "Student",  
                                   rep( "Entrepreneur", 3 ) ),  
                 "English" = sample( 60:100, 7 ),  
                 "ComputerScience" = sample(80:90, 7),  
                 "Biology" = sample( 50:100, 7),  
                 "Bioinformatics" = sample( 40:90, 7)  
               );
```

然后统计：1. 每个人最差的学科和成绩分别是什么？2. 哪个职业的平均成绩最好？3. 每个职业的最佳学科分别是什么（按平均分排序）???

```
## 代码写这里，并运行；
```

使用 `starwars` 变量做计算

1. 计算每个人的 BMI；
2. 挑选出肥胖 ($\text{BMI} \geq 30$) 的人类，并且只显示其 `name`, `sex` 和 `homeworld`；

```
## 代码写这里，并运行；
```

3. 挑选出金发碧眼的人类；
4. 按 BMI 将他们分为三组， <18 , $18\sim25$, >25 ，统计每组的人数，并用 `barplot` 进行展示；注意：展示时三组的按 BMI 从小到大排序；
5. 改变排序方式，按每组人数从小到大排序；

```
## 代码写这里，并运行；
```