

Homework 2

1. (练习 `dataframe` 的添加, 删除, 修改, 查询操作) 案例小组有 4 名成员, 名字分别是: `zhangsan`, `lisi`, `wangwu`, `xiaoming`, 身高分别是 175, 170, 180, 168, 性别分别是: F, M, F, M (F 表示女, M 代表男)。(5 分)
 - a. 请构造数据框, 包含 4 位同学的信息, 三列的名称分别为, `name`, `height`, `gender`。
 - b. 现在该小组又增加了一名女生, 名字叫 `xiaohong`, 身高 165, 请将该学生加入到数据框中 (提示: 用 `rbind` 函数)
 - c. 经过考试, 上次小组中 5 位同学的成绩分别是 60, 40, 55, 55, 40, 将成绩加入到数据框中, 列名是 `score` (提示: 用 `cbind` 函数)
 - d. 请查找身高超过 170 的同学信息, 保存到数据框变量 `data170` 中
 - e. 请查找考试分数超过 50 分的女生的信息, 保存到数据框变量 `femaledata50` 中
2. 请下载第三次课大家的考试成绩, 由于大家系统不一样, 数据有多种格式供大家使用。(10 分)

CSV: `score_data = read.csv(file = "Exam1-score.csv");`
Txt: `score_data = read.table(file = "Exam1-score.txt", sep = ",", header = TRUE);`

 - a. 请将数据读入 R 软件中进行分析, 通过浏览整个数据看看数据中每一列的特点, 第一列表示提交的顺序 ID, 第二列是提交时间, 第三列是考试用时, 第四列表示成绩, 第五列姓名, 第六列是专业
 - b. 结合第二次课中大家交作业的例子, 找出是否同学重复提交考试(可能没有), 如果有的话请删除最后一次作答成绩。
 - c. 使用 `plot` 和 `hist` 函数观察各位同学的考试成绩, 对大家的考试成绩特点进行简要的分析
 - d. 考试过程中有一些同学意外关闭考试网页, 然后又重新进行了作答, 请仔细观察数据, 提供一些思路能够找出这些意外关闭考试的同学。
 - e. 编写自己的函数, `get_score_description`, 该函数的输入是 `scores`, 返回的结果是成绩的最小值, 最大值, 中位数和均值, 保存到向量中返回。函数定义如下:

```
get_score_description = function(scores)
{

}

}
```
 - f. 现在想得到每个专业的考试情况, 请利用上面的函数对每个专业同学的考试成绩进行统计, 保存到大数据框变量 `major_score_description` (提示: 先按照专业的名称将数据保存到临时的数据框变量中, 然后和上一步 d 中方法一样, 将该专业的成绩作为参数传入函数 `get_score_description` 中得到统计结果, 接着将专业名称和返回的结果构成一行加入到数据框变量 `major_score_description` 中。)

3. 上课我们介绍了 dailyshow 的数据，请完成下面工作：
- 首先用 `summary` 来查看数据的总体情况，回答你观察到的数据存在哪些问题
 - 数据清理，提取含有 NA 的数据，回答 NA 数据的特点
 - 观察 `Occupation` 和 `Group` 的数据，将由于大小写原因造成的不一致修正
 - 编写函数 `get_top5` 将 `Occupation, Group`, 频次最高的前 5 位的名字返回（提示，首先用函数 `table` 得到频次，然后 `sort` 之后得到大小顺序，如果要从高到低排序，请查看 `sort` 的帮助文档，将参数 `decreasing` 设置为 `TRUE`。但是这里的数据是频次，所以要得到名字，需要用函数 `names()` 从刚才 `sort` 得到的结果中得到从高到低排序的名字，返回前五个即可）。（10 分）

作业要求：

- 写清楚题号和需要回答的内容
- 保证作业具有很高的可读性
- 报告和程序需要进行必要的说明，能够让没有背景知识的同学也能看懂
- 提交作业的时间是 2016 年 5 月 20 日星期五晚上 12 点，过时将认为是缺作业。
- 如果对作业题目有疑问的地方请在群里问，不要私信及问如果解答题目的细节问题，希望大家能够锻炼自己解决问题的能力！
- 所有作业用 word 完成后另存为 pdf 格式（如果不会请百度如何将 word 另存为 pdf），邮件主题和作业名字一样，格式为：DSJJYB_姓名_专业_HW2，所有作业提交至邮箱 jasonyaopku@gmail.com