
title: "实验四 R 语言函数和程序包" author: "郑骏明" date: "2019/11/9" output: html_document

实验四 R 语言函数和程序包

实验说明

(一) 实验类型：综合性 (二) 实验目的：1.了解R 语言函数的定义和使用。 2.了解R 语言程序包的安装和使用。

(三) 实验内容：1.R 语言用户自定义函数的语法结构。 2.使用read.table 和write.table 函数进行文本处理。 3.安装并加载oligo 程序包。 4.安装并加载limma 程序包。

(四) 要求：必开

实验讲解

apply

- [文档下方的apply部分](#)
- [读取和写入外部数据](#)
- [实验报告电子文档收集问卷](#)

实验项目

1. 将下面的代码改写，不要使用for循环，而是使用sapply

```
# 使用mtcars数据集，将miles per gallon转化为每升公里数1 mile per gallon = 0.425公里/升
kpl = c()
for (i in 1:dim(mtcars)[1]){
  kpl[i] = mtcars[i, "'mpg']*0.425
}
mtcars$kpl = kpl
# 将以上代码用sapply的形式重新书写
```

2. 读取和写入外部数据

读取外部数据是数据分析的基本操作。常用的外部数据格式包括csv以及excel文件。csv文件可以使用R自带的read.csv完成. 要读取文件首先要知道相对的位置。所以应该使用getwd()命令获取当前所在的路径。

建立工作文件夹，然后把代码、数据、分析结果还有输出图片分门别类的放好，这样才不会混乱

```
# 读取potatoes.csv
df_potatoes = read.csv(__)
# 使用read.csv2读取potatoes.csv，观察结果。为何会出错？
```

```

df_potatoes2 = read.csv2(__)
# 使用read.csv读取potatoes_headerless.csv
# 该csv没有表头, 请使用下面提供的v_header_pota作为表头
v_header_pota = c("area", "temp", "size", "storage", "method", "texture",
"flavor", "moistness")
df_potatoes_hdless = read.csv(__,header=__,col.names =__)
# 使用read.delim读取gse71868_short.txt这个文件
# 阅读该文件, 可以看到!series_matrix_table_begin在第61行, 实际的表格在第62行, 可以使用
skip属性跳过前面的61行
# 总共读取15行
df_gse = read.delim(__, skip=__, nrows=__)
# 使用readxl读取urbanpop.xlsx
# 列出urbanpop中的工作表

—
# 读取urbanpop中的1960-1966这个工作表

—
# 将mtcars这个数据框写入mtcars.csv

—
# 将mtcars以及iris这两个数据框写入combine.xlsx, 分别写入该xlsx文件的mtcars数据表和iris
数据表
library(__)
list_comb = list(__,__)
write_xlsx(__,__)

```

3. 使用readxl包读取pccr.xlsx文件, 这个文件有那两个数据表? 在下次实验课我们会利用tidyr和dplyr这两个工具包进行数据处理. 建议课后预习这两个包
4. 使用for loop以及if...else将pccr.xlsx文件中的两个数据表进行整合. 在两个表中查询sampleid一样的行, 然后将actin表中的ct数据放在cyclin表的新列Ct_actin中

```

library(readxl)
df_cyclin = read_excel('__',sheet = '__')
df_actin = read_excel('__',sheet = '__')
# 因为这里用的是非常规方法, 先转换成传统dataframe
df_cyclin = as.data.frame(df_cyclin)
df_actin = as.data.frame(df_actin)
# 建立一个向量储存接下来循环返回的结果
vect_ct_actin = c()
for (i in __:dim(__)[__]){
  id = df_cyclin[__,'__']
  ct_actin = df_actin[__==$__=='__']
  vect_ct_actin[i] = ct_actin
}
df_cyclin = cbind(df_cyclin,vect_ct_actin)

```

5. [选做]在自己的电脑安装并加载oligo以及limma程序包, 截图上传至[实验报告电子文档收集问卷](#)
6. 在纸质实验报告中举例说明如何使用sapply函数
7. 在纸质实验报告中总结陈述读取外部csv文件的流程