



Programming: R

报告人：张诏月



标准数据类型

- 向量 `eg. x <- c('s1', 's3', 's5', 's2', 's4', 's6') x[1]`
- 因子 `eg. b <- factor(x,levels = c('s1', 's3', 's5', 's2', 's4', 's6'))`
- 列表 `eg. list('abcd', 786, 2.23, 'runoob', 70.2)`
- 矩阵 `eg. matrix(data, nrow=, ncol=, byrow = F) + - * /`
- 数组 `eg. array(1:16,dim=(2,4,2))`
- 数据框 `eg. data.frame(x = , y =)`



基本语句

- 赋值 `<- / =`
- 输出 `print`
- 条件判断语句 `if-else if-else / switch`
- 循环语句 `for-in / while / repeat`
- 异常处理语句

```
re1 = tryCatch({  
  code},warning = function(w){  
  code},error = function(e){  
  code})
```

文件读写

```
f <- file("filename.txt", open="rt")
```

```
readLines(f)
```

```
writeLines ( )
```

```
close(f)
```

```
read.table()/read.csv()
```

```
> install.packages("GEOquery")
```

```
WARNING: Rtools is required to build R packages but is not currently installed. Please  
download and install the appropriate version of Rtools before proceeding:
```

```
https://cran.rstudio.com/bin/windows/Rtools/
```

```
Installing package into 'C:/Users/zzdfr/Documents/R/win-library/3.6'  
(as 'lib' is unspecified)
```

```
Warning in install.packages :  
package 'GEOquery' is not available (for R version 3.6.0)
```

函数

```
f <- function_name(x)
```

调用包

```
install.packages()
```

```
library()
```

```
suppressMessages(library())
```

```
if(!requireNamespace("BiocManager",quietly = TRUE))
```

```
  install.packages("BiocManager")
```

```
BiocManager::install("GEOquery")
```

正则表达式

```
grep() grepl() sub() gsub()...
```



1. 读取exercise2.fasta (<https://github.com/ZhaoyueZhang/Lin-Group-Bioinformatics-seminar>)
计算每条序列的长度
按组别绘制序列长度boxplot
使用ggplot2中的geom_histogram绘制序列长度概率密度分布图

色彩搭配网站: COLRD, Color Hunt

zyzhang@uestc.edu.cn

姓名+exercise2

