

R编程结构



课堂测试时间

- 1、填写下表：

功能	函数	功能	函数
1:4-1		sep(1.3, 5, 1)	
rep(1:2, 3)		x <- c(1,4,6,7); x[-2]; x*x;	
z <- c(1,5,3,8,1) z[z>3], order(z)		y <- matrix(1:10, nrow=2) y[2, c(4,5)]	
	sort/sort.list()/order		which.min()/which.max()

- 2、请列举R中常见的数据类型数据结构，描述矩阵和数据框的区别？
- 3、使用rep()函数构建一个向量x。向量是由5个1，3个2，4个3和2个4组成的。
- 4、写出命令：1) 创建一个矩阵mat,矩阵的值为1-20之间的整数，四行五列，按行排列
2) 获取矩阵的第三行第四列的元素值 3) 获取矩阵第一行的所有值。
- 5、写出命令：1) 创建字符串向量names,元素值为"zhang san"、"li si"、"wang wu",创建数值型向量scores,元素为70、80、90，创建字符串向量levels,元素值为"C""B""A"; 2)将levels转换成因子类型 3) 创建数据框exam,其列向量为names、scoes、levels; 4) 根据列名称，提出第一列和第三列。

- 数据结构定义: `c()`; `matrix()`; `array()`; `data.frame()`; `factor()`; `list()`;
- 数据结构访问: 下标; 下标向量; 逻辑向量; 负下标;
- 向量: `::`; `seq()`; `rep()`;
- 算术运算符: `+`; `-`; `*`; `/`; `**`; `^`; `%%`; `%/`;
- 逻辑运算: `>`; `<`; `>=`; `<=`; `==`; `!=`; `!`; `|`; `&`; `isTRUE()`; `identical()`; `any()`; `all()`;
- 属性函数: `length()`; `dim()`; `class()`; `names()`; `head()`; `tail()`;
- 排序函数: `order()`; `sort()`; `sort.list()`; `which()`; `which.max()`; `which.min()`;
- 运算函数: `max()`; `min()`; `range()`; `sum()`; `prod()`; `sqrt()`; `abs()`;
- 类型函数: `is.numeric()`; `is.integer()`; `is.logical()`; `is.character()`; `as.xxxx()`;
- 其余函数: `attach()`; `detach()`; `with()`; `$`; `t()`; `diag()`; `solve()`; `eigen()`;

- 矩阵运算: `t()`; `det()`; `array()`; `crossprod()`; `tcrossprod()`; `diag()`; `solve()`; `eigen()`;
- 缺失值: `NA`; `is.na()`; `na.rm = TRUE`; `na.omit()`;
- 类型函数: `is.numeric()`; `is.integer()`; `is.logical()`; `is.character()`; `as.xxxx()`
- 字符处理: `nchar()`; `substr()`; `strsplit()`; `toupper()`; `tolower()`; `paste()`;
- 日期和时间: `Sys.Date()`; `date()`; `difftime()`; `format()`; `as.Date()`; `%d`, `%a`, `%A`, `%m`, `%b`, `%B`, `%y`, `%Y`;
- 统计函数: `mean()`; `median()`; `sd()`; `var()`; `max()`; `min()`; `range()`; `sum()`; `quantile()`; `diff()`; `scale()`;
- 数据集合合并: `rbind()`; `cbind()`;

- 完成大作业0001
- 按照要求提交

- 流程控制
- 循环控制
- 输入输出
- 自写函数

if-else	<i>if(cond) statement</i> <i>if(cond) statement1 else statement2</i>
ifelse	<i>if(cond, statement1, statement2)</i>

```
> if(FALSE)
+ {
+   message("This won't execute...")
+ }else
+ {
+   message("and you'll get an error before you reach this.")
+ }
and you'll get an error before you reach this.
```

一个表达式，不能是NA

```
> ifelse(rbinom(10, 1, 0.5), "Head", "Tail")
[1] "Tail" "Head" "Head" "Head" "Tail" "Head" "Head" "Tail"
[9] "Tail" "Tail"
```


switch

switch(expr, ...)

```
> feelings <- c("sad", "afraid")
> for (i in feelings)
+   print(
+     switch(i,
+       happy = "I am glad you are happy",
+       afraid = "There is nothing to fear",
+       sad = "Cheer up",
+       angry = "Calm down now"
+     )
+   )
[1] "Cheer up"
[1] "There is nothing to fear"
```

repeat	<i>repeat(statement)</i>
for	<i>for(var in seq) statement</i>
while	<i>while(cond) statement</i>

```

> repeat
+ {
+   message("Happy Groundhog Day!")
+   action <- sample(
+     c(
+       "Learn French",
+       "Make an ice statue",
+       "Rob a bank",
+       "Win heart of Andie McDowell"
+     ),
+     1
+   )
+   message("action = ", action)
+   if(action == "Win heart of Andie McDowell") break
+ }

```

```

-
Happy Groundhog Day!
action = Make an ice statue
Happy Groundhog Day!
action = Rob a bank
Happy Groundhog Day!
action = Win heart of Andie McDowell

```

break

next

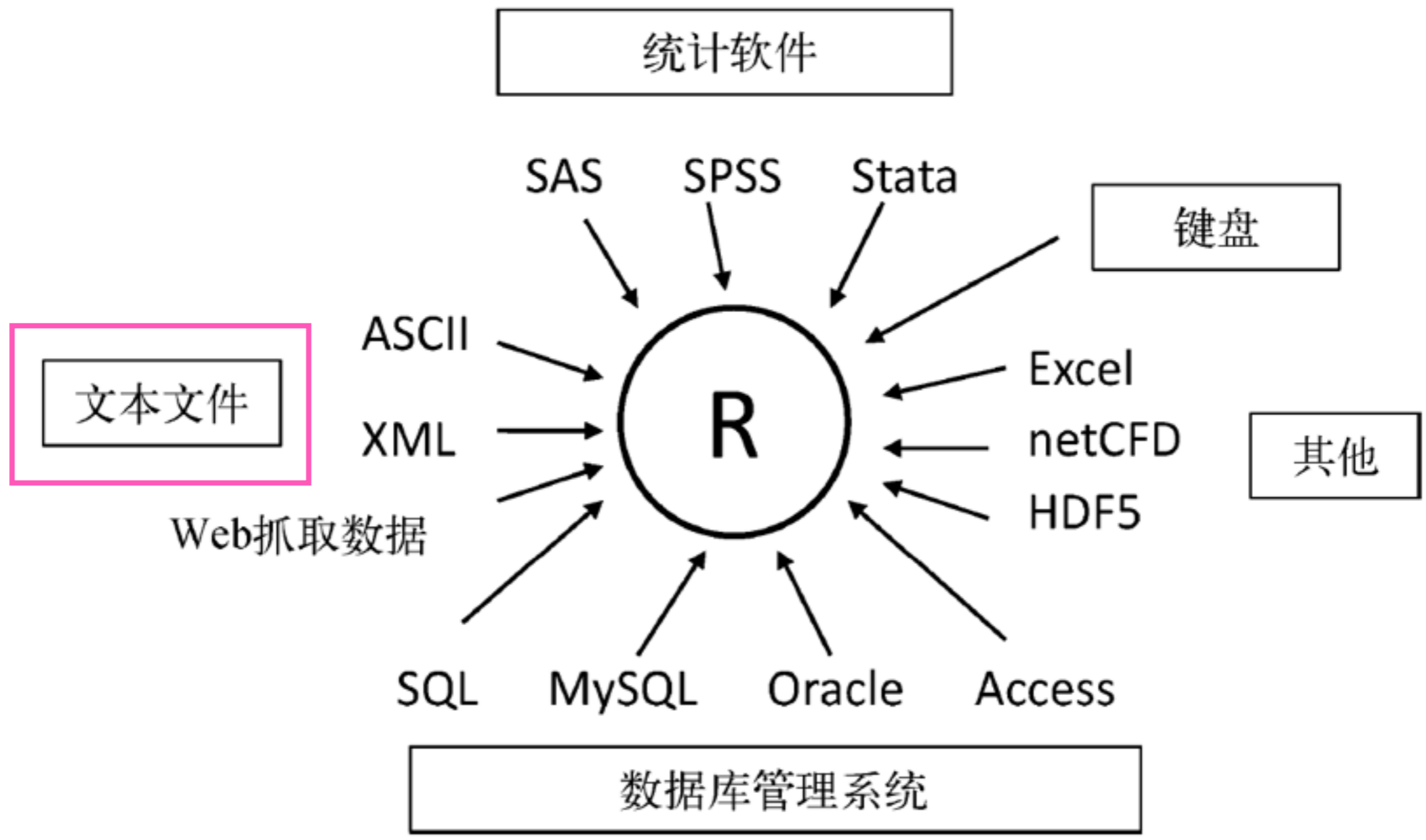


图2-2 可供R导入的数据源

```
read.table(file,  
           header=FALSE,  
           sep=";",  
           row.names="")
```

```
read.csv()  
write.csv()
```

```
write.table(file,  
            append=FALSE,  
            sep=";",  
            row.names=TRUE,  
            col.names=TRUE)
```

- `write.table(student.data, file="test.txt")`
- `a <- read.table("test.txt")`

```
function(arg1,arg2,...) {  
  statements  
  return (object)  
}
```

```
myfun <- function() {  
  print("hello world")  
  return ()  
}
```

```
> f <- function(x,y) x + y  
> f  
function(x,y) x + y  
> f(1,2)  
[1] 3
```

提问时间!

孙惠平

sunhp@ss.pku.edu.cn

练习

2.5 已知有 5 名学生的数据，如表 2.3 所示。用数据框的形式读入数据。

表 2.3: 学生数据

序号	姓名	性别	年龄	身高 (cm)	体重 (kg)
1	张三	女	14	156	42.0
2	李四	男	15	165	49.0
3	王五	女	16	157	41.5
4	赵六	男	14	162	52.0
5	丁一	女	15	159	45.5

2.6 将例 2.5 中的数据表 2.3 的数据写成一个纯文本文件，用函数 `read.table()` 读该文件，然后再用函数 `write.csv()` 写成一个能用 *Excel* 表能打开的文件，并用 *Excel* 表打开。

2.7 编写一个 *R* 程序 (函数). 输入一个整数 n , 如果 $n \leq 0$, 则中止运算, 并输出一句话: “要求输入一个正整数”; 否则, 如果 n 是偶数, 则将 n 除 2, 并赋给 n ; 否则, 将 $3n + 1$ 赋给 n . 不断循环, 只到 $n = 1$, 才停止计算, 并输出一句话: “运算成功”. 这个例子是为了检验数论中的一个简单的定理.

● 0010-1

例 2.4 编写一个用二分法求非线性方程根的函数, 并求方程

● 0010-2

$$x^3 - x - 1 = 0$$

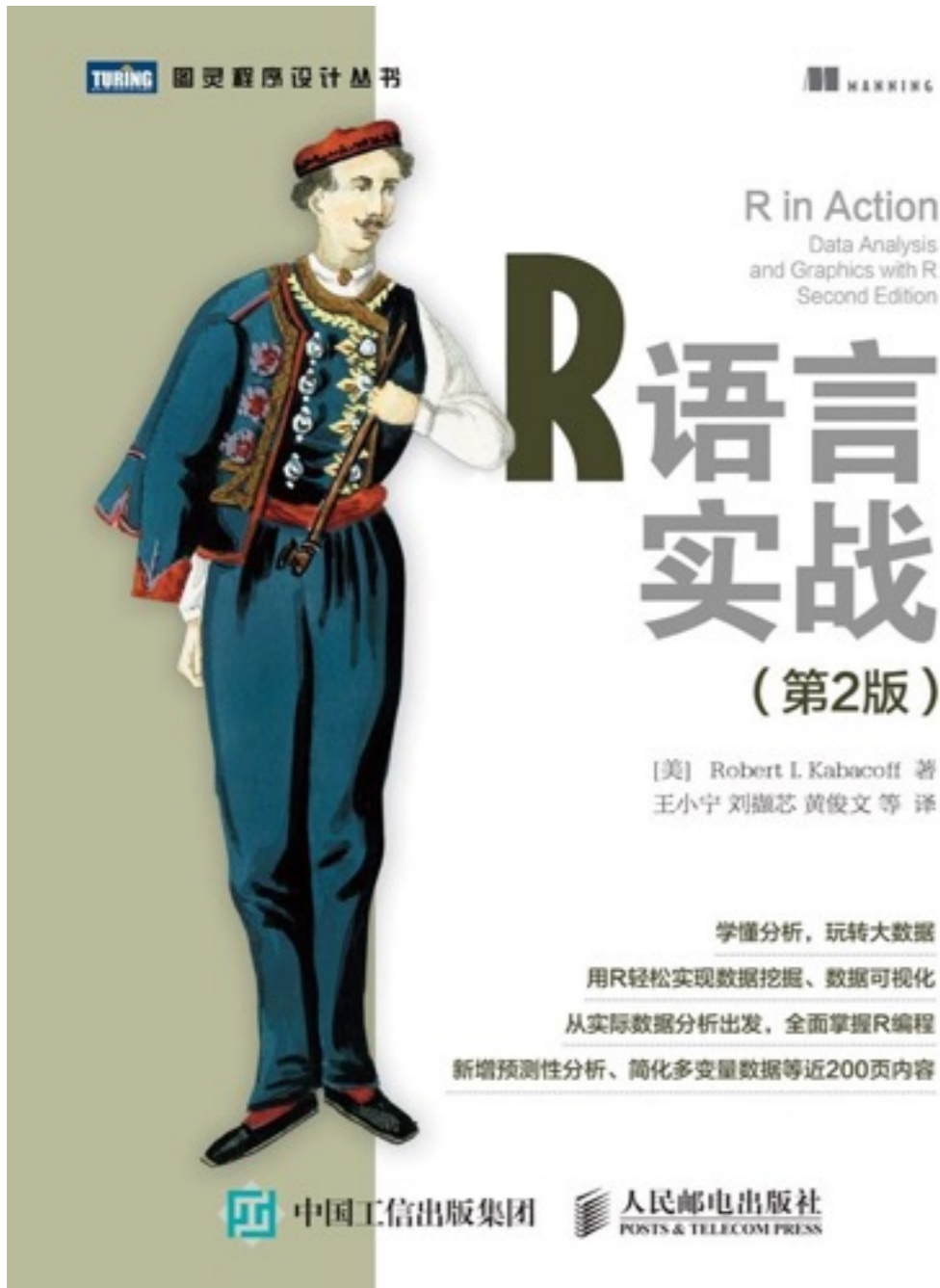
在区间 $[1, 2]$ 内的根, 精度要求 $\varepsilon = 10^{-6}$.

二分法计算过程如下: 取中点 $x = \frac{a+b}{2}$, 若 $f(a)$ 与 $f(x)$ 异号, 则置 $b = x$; 否则 $a = x$. 当区间长度小于指定要求时, 停止计算.

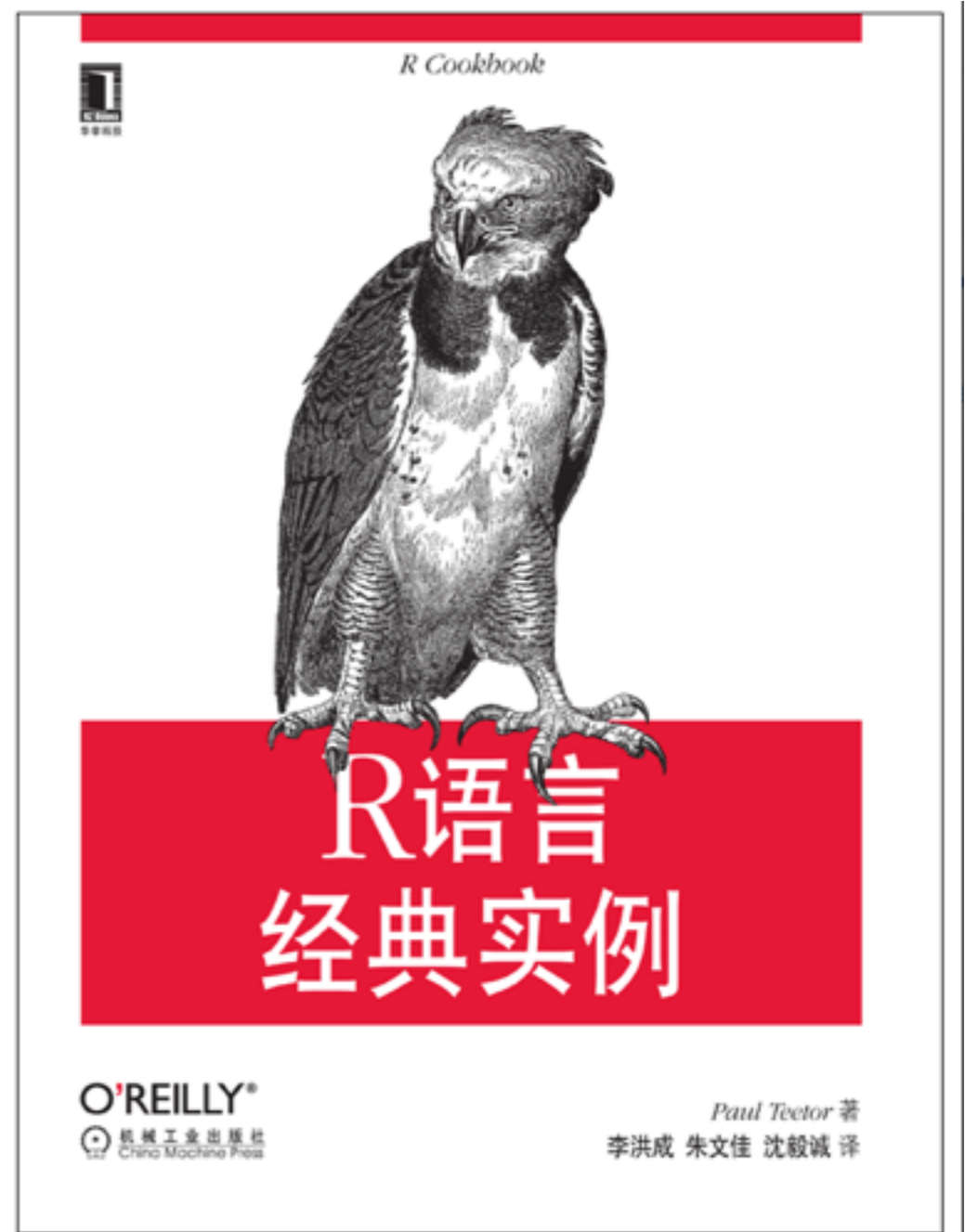
- 有48位应聘者应聘公司某职位，公司为这些应聘者的15个指标打分，分数从0到10，0最低，10最高，具体分数见0011_applicant.csv，公司要录用其中优秀的8名，写一个程序来选择：
 - 如果各个指标权重是同样的，怎么选
 - 随机产生15个指标的一个权重表，按照权重进行选择
 - 随机制定3个指标，这三个指标不能低于8，如何选择

- `0012_test.txt`文件中包含两个班学生的6门课程成绩，要求：
 - 在最后增加一列班级 (`class`)，前35个为1班，后边35个为3班，中间的为2班，班级形式为“01, 02, 03”
 - 在最前面增加一列学生序号 (`number`)，序号形式为“`student_class_number`”
 - 统计成绩缺失的信息，包括哪些学生成绩出现缺失，每个课程有多少学生成绩有缺失，每个班级有多少学生成绩有缺失，使用本班级该课程的平均分作为该缺失的值
 - 统计每个课程的平均分、每个班级的平均分，每个人的总分
 - 求1班每科成绩均大于80的学生集合，2班每科成绩均小于90的学生集合
 - 给每个学生评定优、良、中、差四个等级，按照每个班级划分，按照所有班级一起划分，找出两种划分中等级不同的学生

- 从 `0013_grade.csv` 中读取两班成绩
- 计算每个班级的均值和标准方差
- 计算每个人的标准化成绩，添加到数据中，写到 `0013grade.txt` 中



2.3、5.4、5.5
例子5-6、5-8



第四章

- 完成大作业0002
- 按照要求提交

谢谢!

孙惠平

sunhp@ss.pku.edu.cn