

心理学编程基础 · 第0课

PsychoR
COStudy 心理学
夏晓凯

课程介绍

为什么学 · 学什么 · 怎样学

导论



为什么

学什么

怎样学

心理学前沿研究中
不会编程就无法进行下去

编程的应用场景

- 实证研究中
 - 呈现复杂的实验刺激
 - 紧跟论文上最新的研究方法
 - 实现自己创想的各种idea
- 调查研究中
 - 实现较复杂的统计分析（IRT，软计算，潜变量模型）
- 国际协作（OSF）与发表（SCI）的需要

为什么

学什么

怎样学

如何挑选合适的编程语言



采集数据	<ul style="list-style-type: none">• E-Prime• Presentation	Psychtoolbox PsychoPy	Python
分析数据	<ul style="list-style-type: none">• R• Python	Matlab Julia	Mathematica
论文发表	<ul style="list-style-type: none">• R• LaTeX		

为什么

学什么

怎样学

本课程最终选择



为什么

学什么

怎样学

R：最适合心理学学生的编程语言



- 高级语言
 - 类似人类语言，更好理解
 - 容易迁移到其他编程语言
- 语法风格一致：函数式编程
- 语法特性较少
- 专门为数据分析所设计：心理学研究中随处可用
- 社区支持完善
 - 国内的组织：统计之都，WISE R CLUB，PsychoR
 - 开源社区：完整的统计学生态

为什么

学什么

怎样学

如何学习这门课

- 熟悉基本概念
- 掌握流程和思路
- 填充具体知识
- 解决心理学问题

为什么

学什么

怎样学

如何学习这门课

- 熟悉基本概念
- 掌握流程和思路 —— 串讲课
- 填充具体知识 —— 后续课程
- 解决实际问题 —— 案例课、具体研究工作

为什么

学什么

怎样学

获取R相关资源



CRAN

<http://cran.r-project.org>



Rstudio

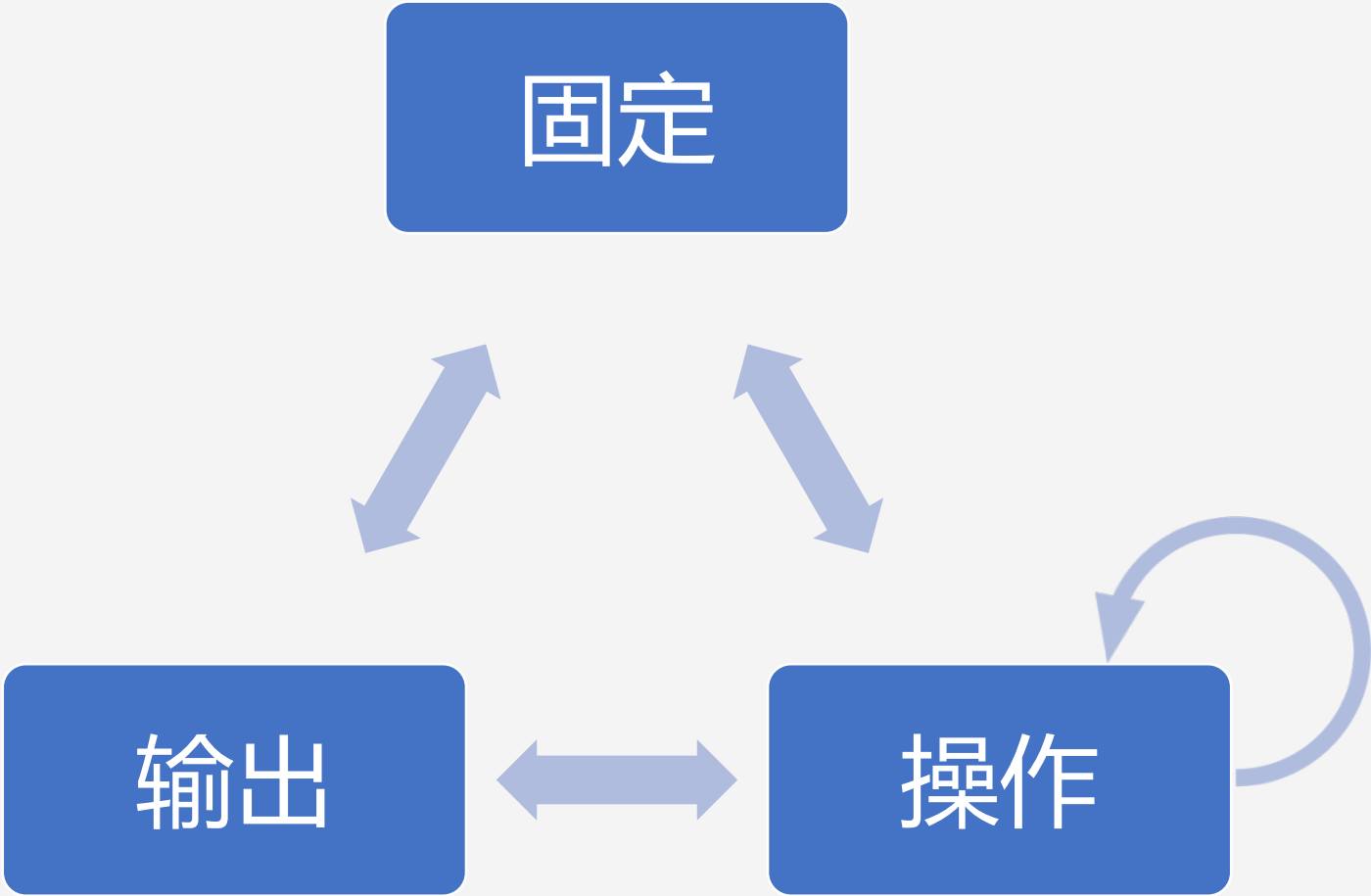
<http://www.rstudio.com>



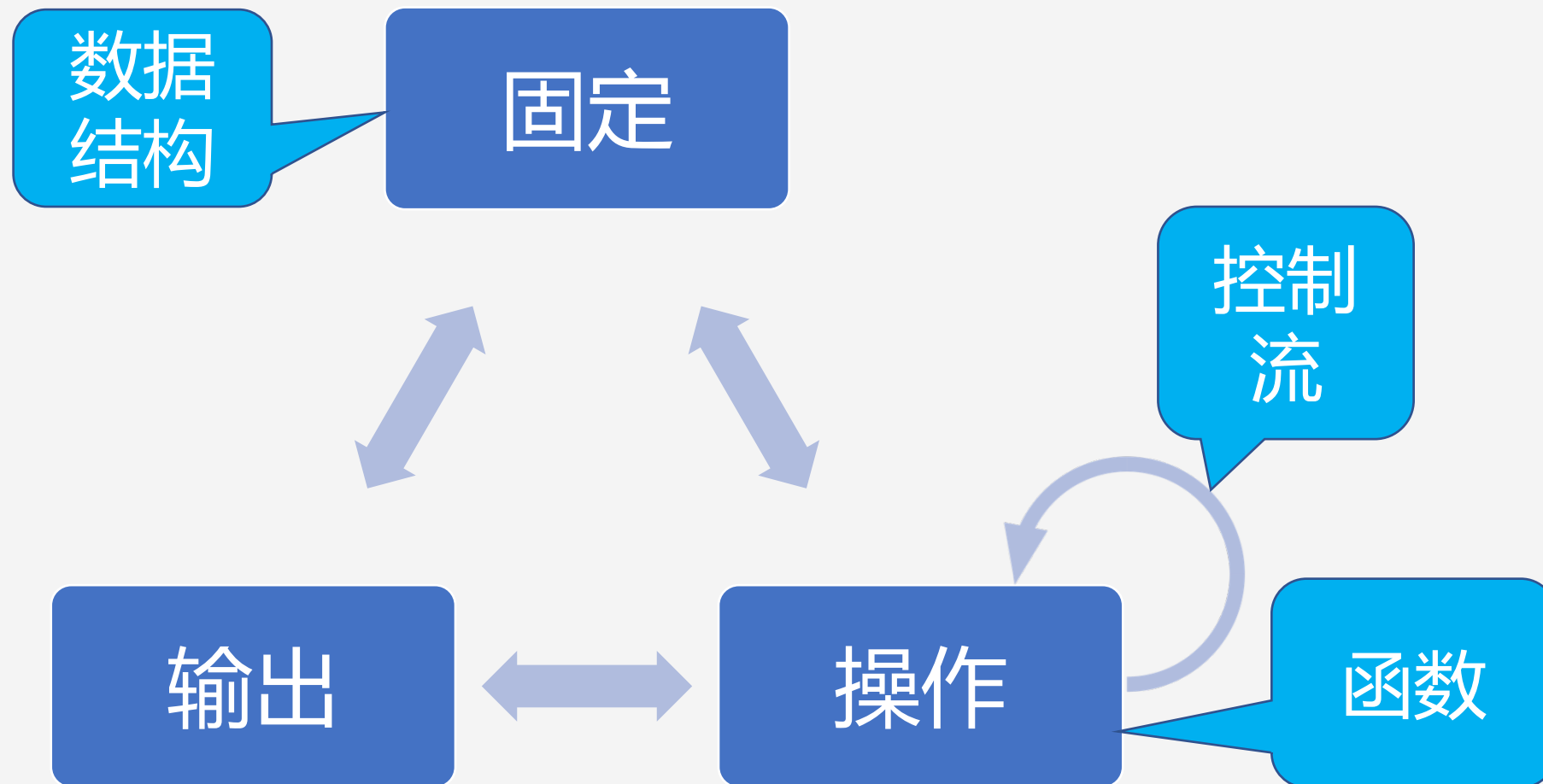
心理学编程基本范式

固定 – 操作 – 输出的 Workflow

编程基本流程



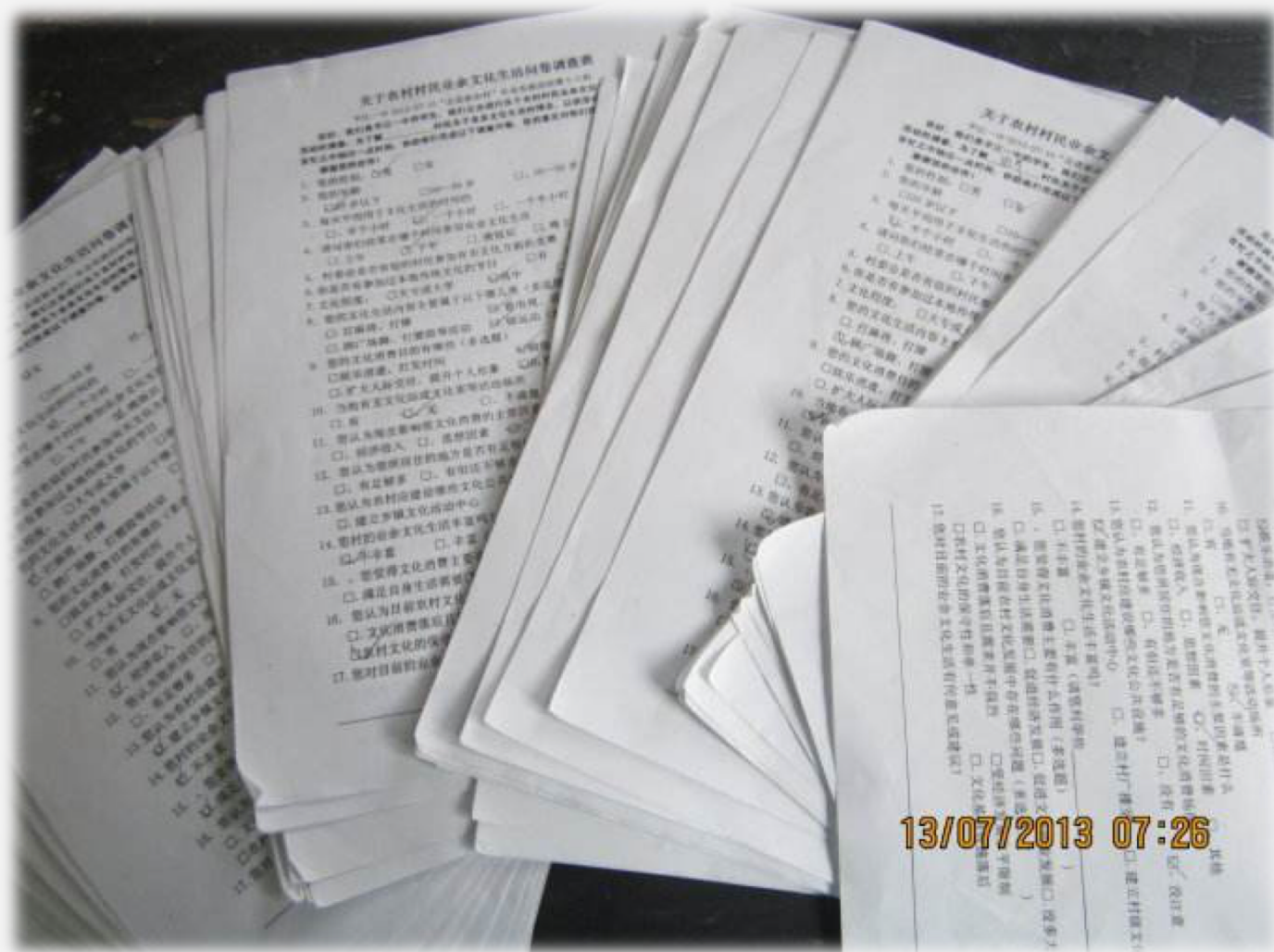
编程基本流程



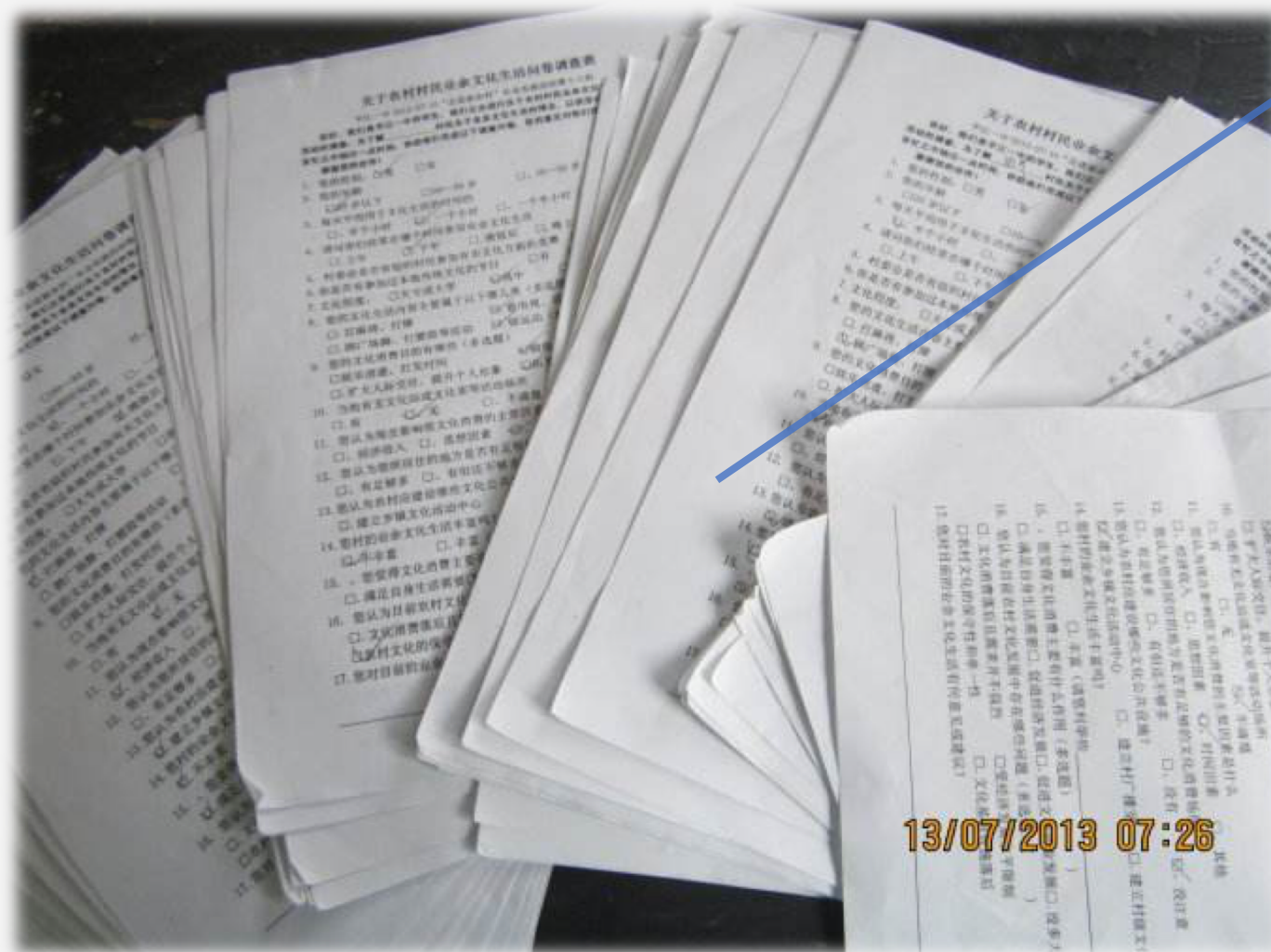
数据结构

赋予杂乱的数据以意义

生活中的数据



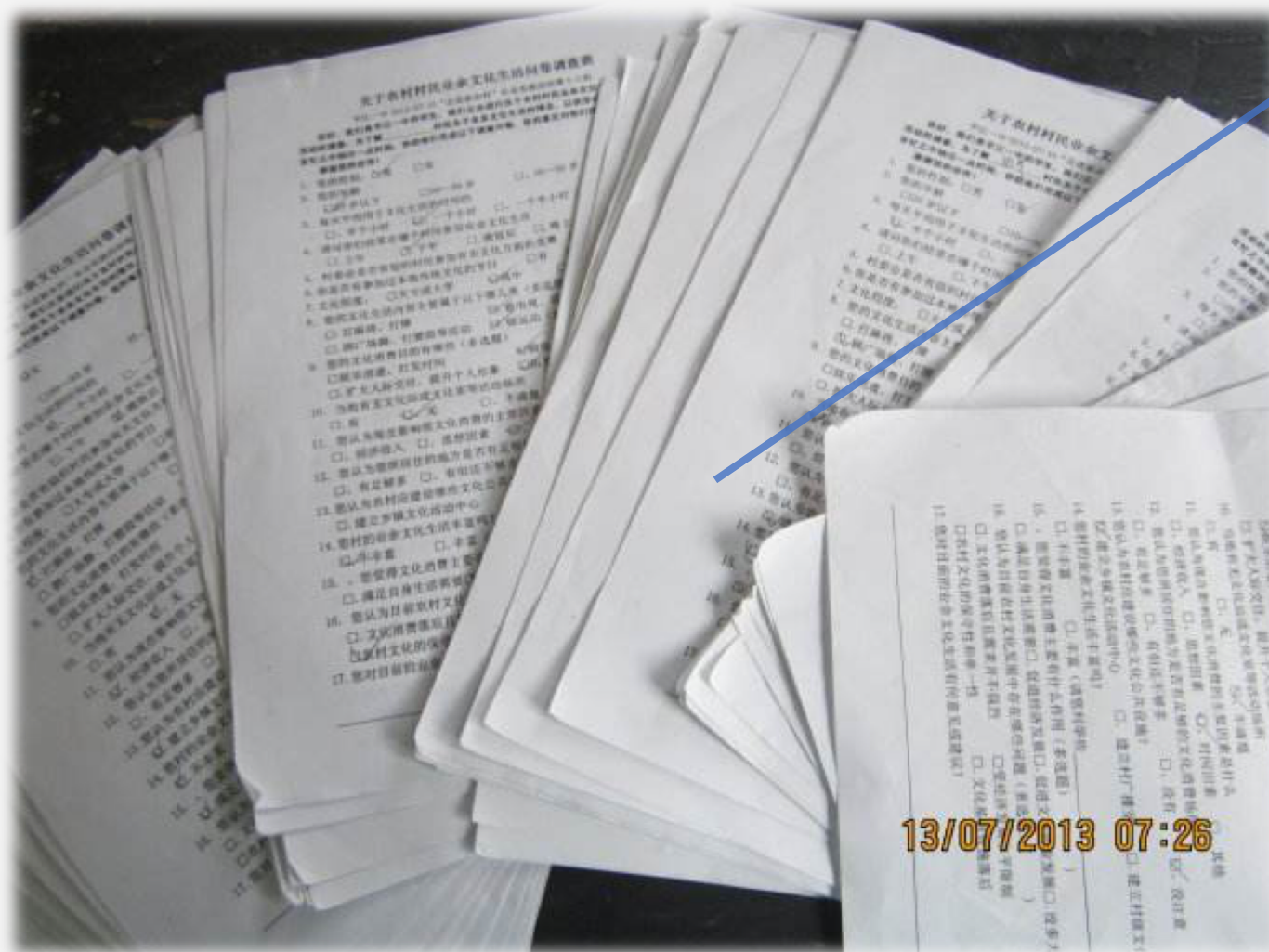
生活中的数据



- 难以索引
- 类型不分
- 编码多样

13/07/2013 07:26

生活中的数据



- 难以索引
- 类型不分
- 编码多样



- 建立索引
- 分开类型
- 统一编码

整理后的数据

被试编号	性别 (factor)	年龄 (dbl)	第一题 (factor)	第二题 (factor)
1	男	23	A	3
2	女	22	B	2
3	女	22	A	1

整理后的数据

被试编号	性别 (factor)	年龄 (dbl)	第一题 (factor)	第二题 (factor)
1	男			3
2			B	2
3	女	22	A	1

Data Frame

整理后的数据



被试编号	性别 (factor)	年龄 (dbl)	第一题 (factor)	第二题 (factor)
1	男	23	A	3
2	女	22	B	2
3	女	22	A	1

整理后的数据



被试编号	性别 (factor)	年龄 (dbl)	第一题 (factor)	第二题 (factor)
1	男	23	A	3
2	女	22	B	2
3	女	22	A	1

数据结构的知识



- 选择
- 建立
- 索引
- 操作
 - 增
 - 删
 - 改

如何选择数据结构

- 根据数据的属性与涵义选择
- 根据分析方法的要求选择

R中常见的数据类型与含义

- 向量（一个数、一列数、线性空间中的一个点）
- 矩阵（一组数、线性空间中的变换）
- 数据框（一个表格，每行是观测，每列是观测的属性）
- 因子（离散的数据、对数据的分类）
- 列表（装东西用的）

数据结构的建立

- 向量：使用 `c()` 函数
- 其他数据结构：数据结构名同名函数建立

数据结构的索引

- 基本索引符号
 - [] : 依据线性空间位置索引
 - \$: 依据数据对应名称索引
 - [[]] : 索引数据块
- 数据结构的索引思路
 - 查看数据结构
 - 依据结构提取相应数据

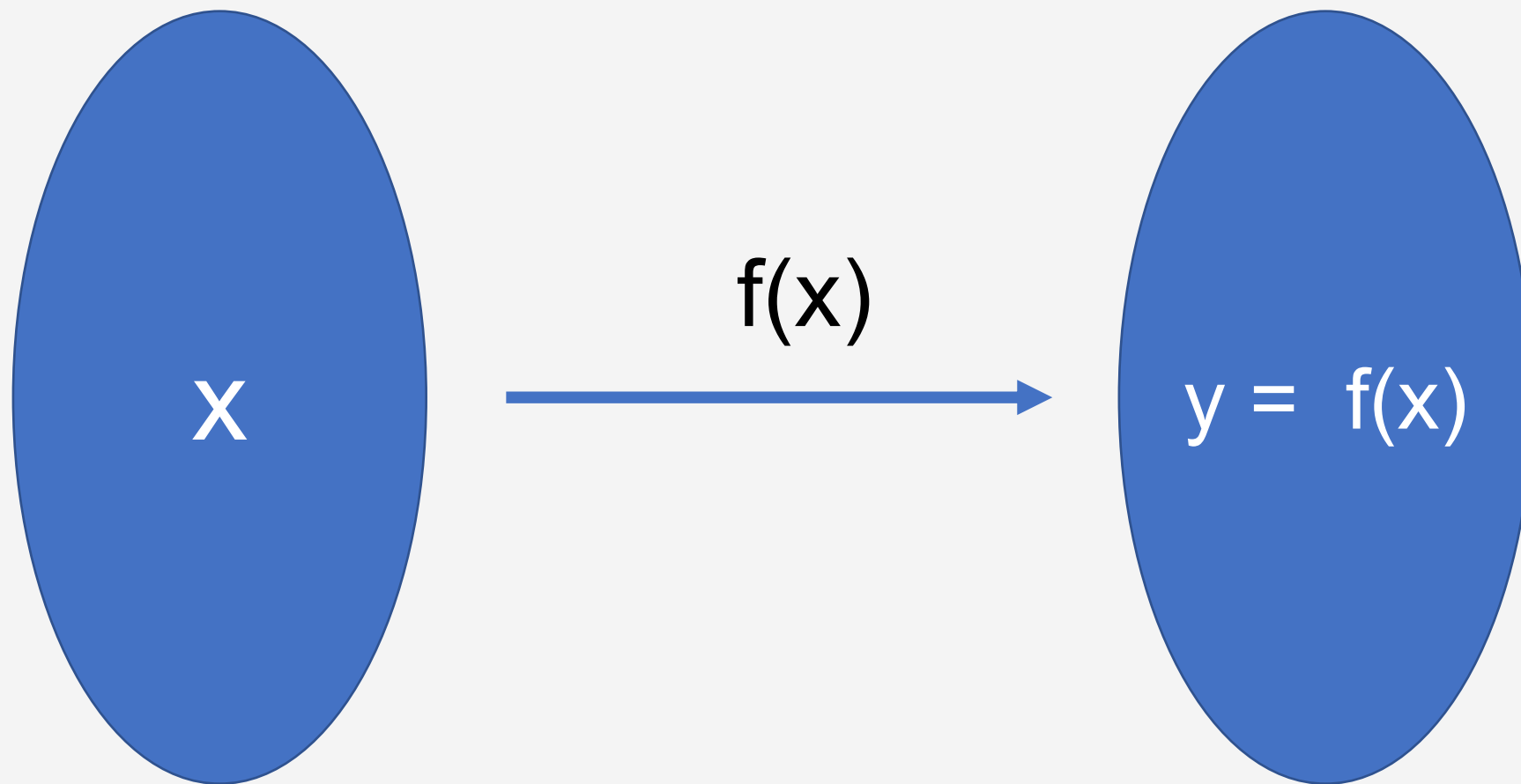
数据结构的操作

- 基础操作
 - 依靠索引 + 赋值操作实现
- 高级操作
 - 基于 tidyverse 包实现

函数/方法

对数据进行操作

为什么叫函数



函数的结构

t.test(x = a, y = b)

函数名

参数部分

函数的结构

t.test(x = a, y = b)

函数名

参数部分

组成部分	功能	格式
函数名	决定函数做什么	由作者确定
参数部分	决定函数如何去做	形式参数 = 实际参数

函数的使用



?t.test

什么是方法

- 对象的内置函数
- 对象是对一个操作客体的抽象
- 使用方法（Python）：对象.函数（ ）

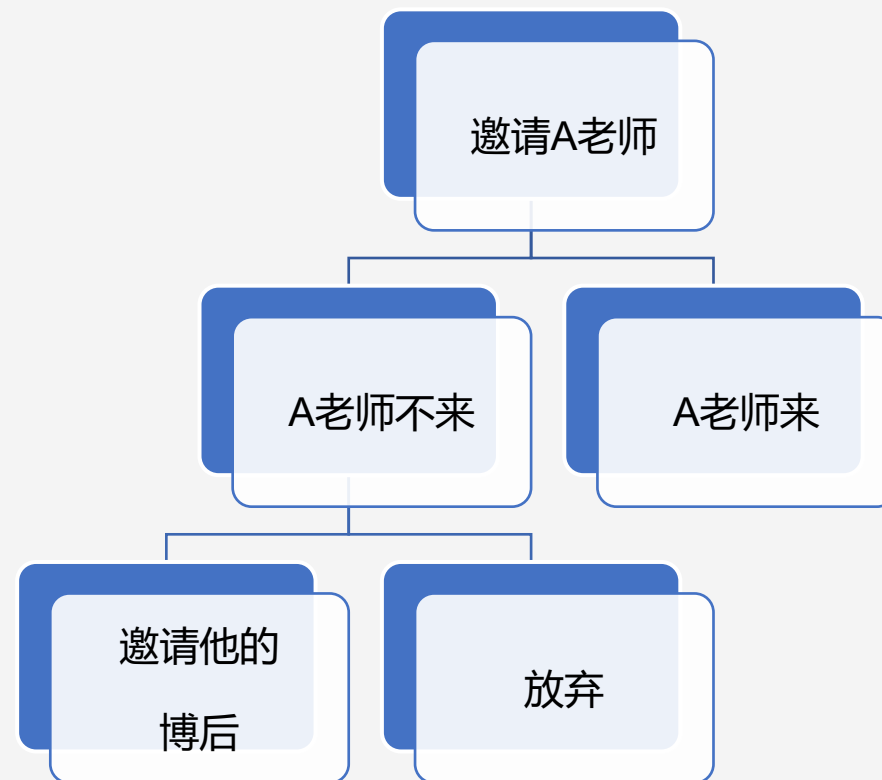
控制流

控制函数协同工作

两种控制流



循环



判断

到达图灵完备

练习

熟悉R语言环境

练习1：向世界问声好



- 输入Hello World!

结果输出

输入符：
可以输入代码

文字要加双引
号才能识别

```
> "Hello World!"  
[1] "Hello World!"  
> |
```

交互式编程：
一个输入一个反馈

索引符号：
标识元素的位置

光标：
开始输入下一段代码吧

赋值操作



```
> a <- "Hello World!"
```

```
> a
```

练习2：认识一下正态分布（可视化）



```
> data <- rnorm(1000)
```

```
> hist(data)
```

练习3：一个完整的数据分析流程



- 对鸢尾花数据不同种类的花瓣长度进行t检验
- 选取合适的数据结构
- 使用合适的函数进行分析
- 正确理解输出结果