



北京邮电大学

BEIJING UNIVERSITY OF POSTS AND TELECOMMUNICATIONS



Data Analytics with R  
语言数据分析



气象万千，数以等观

艾新波 / 2018 • 北京



# 课程体系



## R语言数据分析



### 上部：论道



- 第1章 气象万千、数以等观
- 第2章 所谓学习、归类而已
- 第3章 格言联璧话学习
- 第4章 源于数学、归于工程



### 中部：执具



- 第5章 工欲善其事必先利其器
- 第6章 基础编程
- 第7章 数据对象



- 第8章 人人都爱tidyverse
- 第9章 最美不过数据框



### 下部：博术



- 第10章 观数以形
- 第11章 相随相伴、谓之关联
- 第12章 既是世间法、自当有分别
- 第13章 方以类聚、物以群分
- 第14章 庐山烟雨浙江潮

# 课程体系







## R语言数据分析



### 上部：论道






-  第1章 气象万千、数以等观
-  第2章 所谓学习、归类而已
-  第3章 格言联璧话学习
-  第4章 源于数学、归于工程




### 中部：执具



-  第5章 工欲善其事必先利其器
-  第6章 基础编程
-  第7章 数据对象








-  第8章 人人都爱tidyverse
-  第9章 最美不过数据框

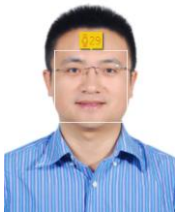


### 下部 博术

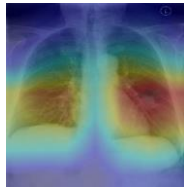


-  第10章 观数以形
-  第11章 相随相伴、谓之关联
-  第12章 既是世间法、自当有分别
-  第13章 方以类聚、物以群分
-  第14章 庐山烟雨浙江潮

# 数据分析：发现数据背后的规律



<https://www.how-old.net/>



现在开始语音转文字测试。

转换完成



猜你喜欢

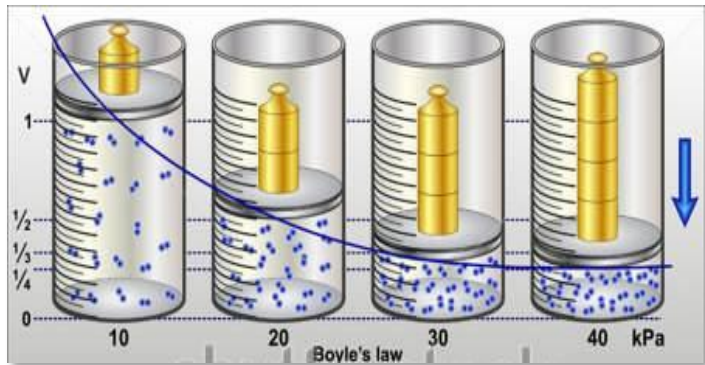


每日推荐



音乐圈

## 从人类历史上第一个被发现的定律说起

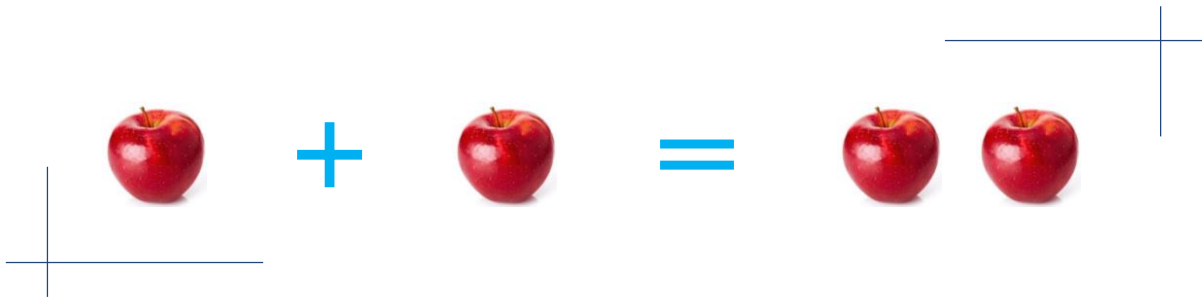


**波义耳定律（1662年）：密闭容器中的定量气体**

**在恒温下，其体积和压强成反比：**  $p = f(v) = \frac{k}{v}$

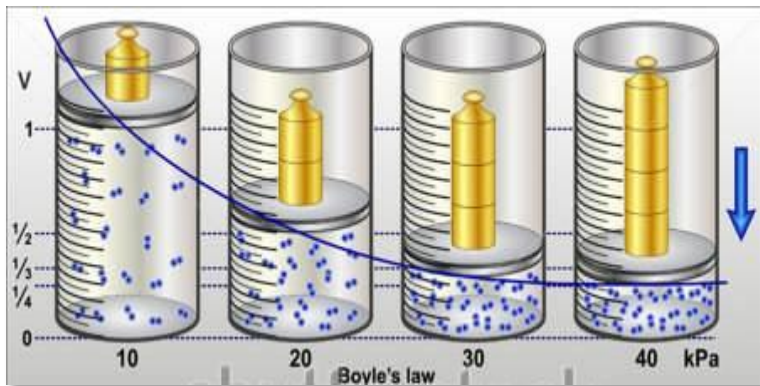
## 我们学过的第一个数学模型

也许是世界上最简单的数学模型：



$$1+1=2$$

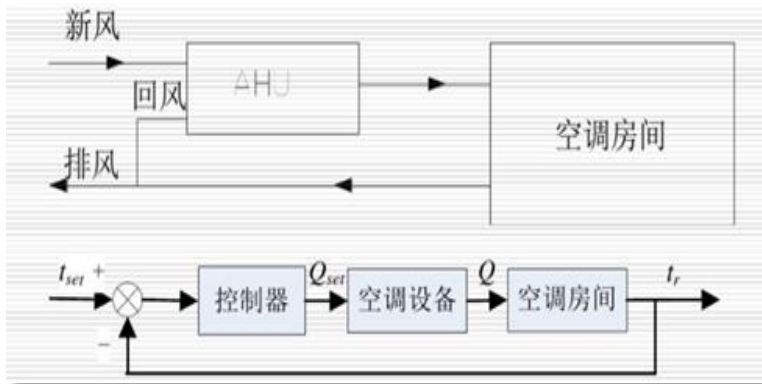
# 风马牛不相及



$$p = f(v) = \frac{k}{v}$$

压强 = 体积?

# 风马牛不相及



$$u = f(e)$$

送风量 = 温度 (温差)



# 风马牛不相及



$$PM2.5 = 0.29T + 0.53U + 0.2V$$

污染 = 温度 + 湿度 + 风速?

# 等号的含义

$$x = 5$$
$$1 + 1 = 2$$

equal to

决定

$$e^{\pi \cdot i} = -1$$

等于

相当于

correspond to

defined as

转化为

$$i^2 = -1$$

$$F = ma$$

定义为

transform

$$\begin{cases} y_1 = a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n, \\ y_2 = a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n, \\ \dots\dots\dots \\ y_m = a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{mn}x_n \end{cases}$$

$$y = f(x)$$

联系

依存

等同于

equivalent to

$$p = f(v) = k \cdot v^{-1}$$

Value as

取值为

映射

Depend on

$$u = f(e)$$

等价于

mapped to

$$y = 0.29t + 0.53u + 0.2v$$

制约

作用

## 用等号表征万物之间的联系



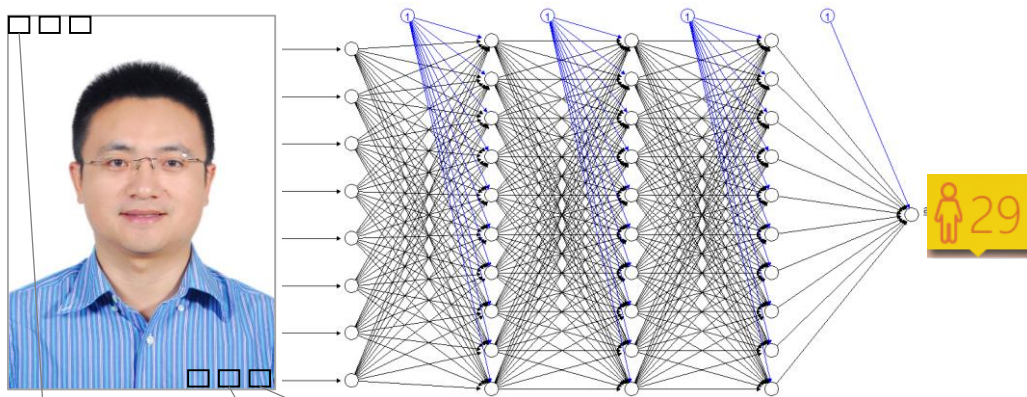
气象万千，数以**等**观

## 用等号表征万物之间的联系



气象万千，数以 = 观

气象万千，数以=观

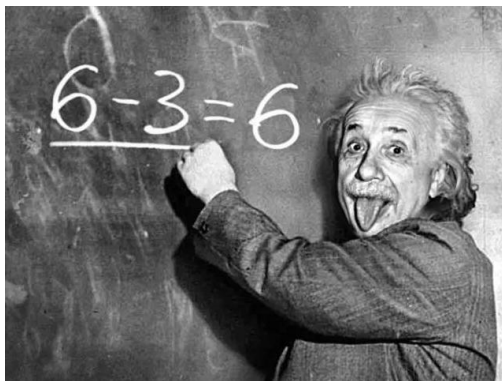


$$y = f(x_1 = 1.00, \dots, x_{314999} = 0.16, x_{315000} = 0.90) = 29$$

## 气象万千，数以=观

等号左侧 =	= 等号右侧	对客观规律的认识
邮件文本	1:是 2:否	垃圾邮件过滤
道路图像	道路、天空、行人、.....	语义分割
医疗图像	1:有病 2:正常	疾病诊疗
光纤信号	事件类别	声纹监测
交易记录	正常、异常	欺诈交易侦测
用户评论	情绪类别	情绪分析
温度、压力等	氯化石蜡粘度	生产质量控制
.....		

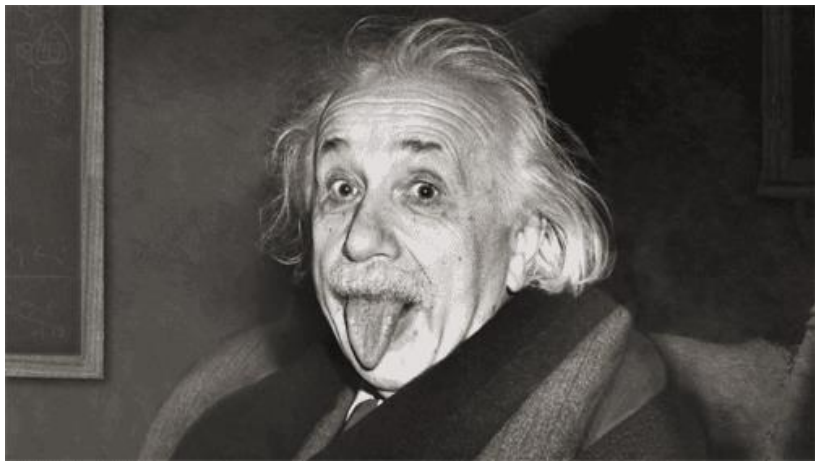
气象万千，数以=观



*I have no special talent, I am only passionately curious*

**数据科学最令人着迷的地方：**一旦进行量化，看似风马牛不相及的事物/属性，经过数学运算，居然可以画上等号，刻画各种各样的规律

气象万千，数以=观



**原动力：**永葆对这个世界的好奇心



A decorative blue border with rounded corners frames the entire slide. Two thin blue crosshair symbols are positioned on the right and left sides of the text.

**谢谢聆听**  
**Thank you**

# 教师个人联系方式

艾新波

手机: 13641159546

QQ: 23127789

微信: 13641159546

E-mail: [13641159546@126.com](mailto:13641159546@126.com)

[axb@bupt.edu.cn](mailto:axb@bupt.edu.cn)

地址: 北京邮电大学科研楼917室

课程网址: <https://github.com/byaxb/RDataAnalytics>

