



主讲老师：徐娜

中级经济师  
经济基础知识  
课程精讲班



## 第二十三章

# 统计与数据科学



# 目录

## 第二十三章 统计与数据科学

- 第一节统计学
  - 1. 统计学的概念
  - 2. 描述统计和推断统计
- 第二节变量和数据
  - 1. 变量的定义和分类
  - 2. 数据的含义和分类
- 第三节数据的来源
  - 1. 收集方法的不同：观测数据和实验数据
  - 2. 数据本身的来源：一手数据和二手数据
- 第四节统计调查
  - 1. 统计调查的分类、方式
  - 2. 统计质量评价标准
- 第五节数据科学与大数据
  - 1. 数据科学
  - 2. 大数据
  - 3. 数据挖掘



## 第一节 统计学

定义	统计学是关于收集、整理、分析数据和从数据中得出结论的科学
分支	描述统计 描述统计是研究数据收集、整理和描述的统计学方法。其内容包括如何取得所需要的数据，如何用图表或数学方法对数据进行整理和展示，如何描述数据的一般性特征
	推断统计 推断统计是研究如何利用样本数据来推断总体特征的统计学方法，其内容包括参数估计和假设检验两大类。参数估计是利用样本信息推断总体特征；假设检验是利用样本信息判断对总体的假设是否成立



## 典型真题

【真题·2023单选】下列统计处理中，属于推断统计的是（ ）。

- A. 采用散点图展示城镇居民人均可支配收入和人均消费支出间的相关关系
- B. 采用圆形图展示能源消费构成
- C. 利用统计表展示2023年10月份新建商品住宅零售价格变化情况
- D. 利用样本客户调查数据估计某品牌消费者忠诚度



## 典型真题

答案：D

解析：推断统计是研究如何利用样本数据来推断总体特征的统计学方法，其内容包括参数估计和假设检验两大类。描述统计是研究数据收集、整理和描述的统计学方法。



## 第二节 变量和数据

### 一、变量

定义	研究对象的属性或特征，相对于常数而言可以有两个或更多个可能的取值		
分类	定量变量	当变量的取值是数量时，如企业销售额，注册员工数量等	
	定性变量	分类变量	当变量的取值表现为类别时，如企业所属行业
		顺序变量	当变量的取值表现为类别且具有一定顺序时，如员工受教育水平



## 第二节 变量和数据

### 二、数据

含义	对变量进行测量、观测的结果。根据需要可以是数值、文字或者图像等形式	
分类	分类数据	分类变量的观测结果，表现为类别，一般用文字来表述，也可用数值代码表示，如用1表示“男性”，2表示“女性”
	顺序数据	顺序变量的观测结果，表现为类别，一般用文字表述，也可用数值代码表示，如用1表示“硕士及以上”，2表示“本科”，3表示“大专及以下”
	定量数据	即数值型数据，是对定量变量的观测结果，其取值表现为具体的数值；可以进行数学运算。





## 第二节 变量和数据

收集方法	观测数据	通过直接调查或测量而收集到的数据。在没有对事物加任何人为控制因素的条件下得到的，几乎所有与社会经济现象有关的统计数据都是观测数据，如GDP、CPI、房价等数据
	实验数据	通过在实验中控制实验对象以及其所处的实验环境收集到的数据。比如，一种新产品使用寿命的数据，一种新药疗效的数据，自然科学领域的数据大多都是。
来源	一手数据	直接的调查和科学实验
	二手数据	别人的调查或实验的数据



## 典型真题

【真题·2023单选】下列统计产量中，属于分类变量的是

( )。

- A. 职业类别
- B. 国内生产总值
- C. 营业收入
- D. 工人工资



## 典型真题

答案：A

解析：分类变量的观测结果称为分类数据，表现为类别，一般用文字来表述，也可用数值代码表示。比如用1表示“男性”，用2表示“女性”。



## 典型真题

【真题·2023单选】下列统计数据中，属于实验数据的是

( )。

- A. 国内生产总值
- B. 某新型产品使用寿命
- C. 住房价格指数
- D. 居民消费价格指数



## 典型真题

答案：B

解析：实验数据是通过在实验中控制实验对象以及其所处的实验环境收集到的数据，称为实验数据。比如，一种新产品使用寿命的数据，一种新药疗效的数据。自然科学领域的数据大多都是实验数据。