第二十三章 统计与数据科学

- 第二十三章 统计与数据科学
  - 一、统计学
    - 描述统计：研究数据收集、整理和描述
    - 推断统计：利用样本数据来推断总体特征
      - 参数估计、假设检验
  - 二、变量和数据
    - 变量：定量变量、分类变量、顺序变量；定性变量
    - 数据：数值型数据、分类数据、顺序数据
  - 三、数据的来源
    - 按其收集方法：观测数据实验数据
    - 从使用者的角度：一手数据、二手数据
  - 四、统计调查
    - 概念与分类
      - 按调查对象的范围
        - 全面调查：全面统计报表、普查
        - 非全面调查
          - 非全面统计报表
          - 抽样调查、重点调查、典型调查
      - 按调查登记的时间是否连续
        - 连续调查
        - 不连续调查
    - 统计调查的方式
      - 统计报表
      - 普查
        - 1、通常一次性/周期性
          - 经济普查：尾数逢3、8的年份
          - 人口普查：逢0的年份
          - 农业普查：逢6的年份
        - 2、一般需规定统一的标准调查时间
        - 3、数据一般较准确，规范化程度较高
        - 4、使用范围较窄
      - 抽样调查：经济性、时效性强、适应面广、准确性高
      - 重点调查
        - 一种非全面调查、适用范围很广
        - 目的：了解基本状况和发展趋势，不要求掌握全面数据
      - 典型调查
        - 作用：弥补全面调查的不足、可验证其数据真实性
        - 优点：灵活机动　不足：大程度上受到观认识的影响
    - 统计质量评价标准：真实性、准确性、完整性、及时性、适用性、经济性、可比性、协调性、可获得性。
  - 五、数据科学与大数据
    - 数据科学
    - 大数据：数据量大、数据多样性、价值密度低、数据的产生和处理速度快　“4V”特性
    - 数据挖掘

Presented with xmind