

第七课 数据可视化

第二节 EDA的目的

主 讲: Robin

EDA的目的

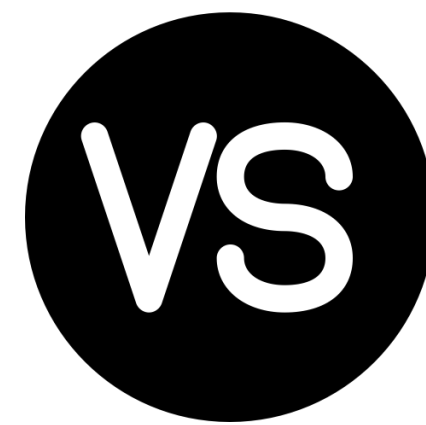
- 通过**可视化**对数据构建初始认知
- 对于数据集中的某些现象原因**提出假设**
- **论证**该使用何种统计推断进行分析
- 为后续使用适合的统计工具及技术**提供支持**
- 为进一步分析**提出问题**
- 对进一步的数据收集及调查**做基础**
- 通常情况，对于**未知的数据集**先进行EDA

EDA vs CDA

- 验证性数据分析 (Confirmatory Data Analysis, CDA) 注重对数据模型和假设的验证

EDA

- 开始没有假设
- 侧重于探索和生成假设
- 通常使用可视化的方法
- 没有统一的步骤、形式



CDA

- 从假设开始
- 侧重于对假设的验证
- 使用统计模型
- 有既定的策略、方法

EDA vs CDA

EDA: “自由爵士”



VS

CDA: “古典音乐”



只为遇见明天更优秀的你！