

定量分析：数据思维与 商业统计

陈文波

cwb@whu.edu.cn

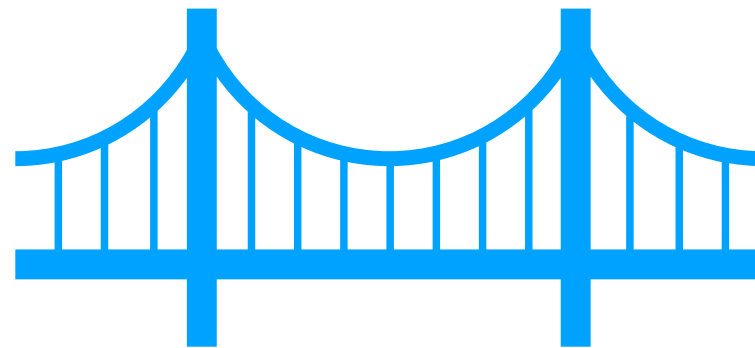
2021年10月

课程梗概

描述性分析与EDA

- 数值方法：三类指标
- 图形方法：五种图形

R



- 抽样分布、CLT
- 正态分布
- T分布
- 卡方分布
- F分布

推断性分析

- 区间估计
- 假设检验
- 应用：
 - 分类 vs. 分类
 - 分类 vs. 数值
 - 数值 vs. 数值

课程梗概

- Session 1 : 内容简介、基本概念、描述性统计1
- Session 2: 描述性统计2
- Session 3: 抽样、中心极限定理、正态分布
- Session 4: 其他三个分布, 区间估计
- Session 5: 假设检验, 两个总体均值和比例的推断
- Session 6: 总体方差和方差比的推断、分类数据 Vs. 数值型数据: 卡方检验与列联表分析
- Session 7: 分类数据 Vs. 数值型数据: 方差分析; 数值型数据: 回归分析
- Session 8: 课程总结

S4-重要分布、区间估计、假设检验

- Chi-squared Distribution
- t-distribution
- F-distribution

点估计与区间估计

- 点估计
- 区间估计: 基本过程和解释

假设检验

- 假设的提法
- 检验的方向
- 2种不同的方法：临界值方法与p_value
- R演示

Wrap-up

- 抽样
- 四种概率抽样
- 抽样分布
- 正态分布

两天总结

- S1：数据与统计，数据分类
- S2：分清类别，三类指标，五种图形PLUS
- S3：抽样与抽样分布，正态分布
- S4：3种分布，推断初步

课后练习

- 请使用你掌握的数据分析方法，对I数据集进行探索性数据分析（描述性方法为主）。可以讨论的问题包括（但不限于）：价格特点、年代特点、区位特点；价格与年代、年代与房产供应，等等。希望大家发现更有趣的现象。
- 要求：使用markdown工具（例如Rmarkdown, Typora)撰写数据分析报告，并尽量导出为pdf在学习通里提交。
 - 提交时间：10月28日中午12点之前。
 - 要有完整的形式：
 - 主要结论
 - 数据分析思路与方法
 - 分析过程