

# 数理统计上机课第一次作业

2021 年 3 月 31 日

## 1 第一题：数据处理与可视化

数据为数据集 `crawfish.xlsx`，记录了七座城市共计 2527 家经营小龙虾的餐馆的相关数据。请使用你认为合适的方法完成：

- (1) 初步汇总数据，并简要描述小龙虾销售的基本情况。
- (2) 删除数据中没有提供人均价格 `price` 的店铺，并增加新的一列总销售额 `total` 为销量 `sale` 与人均价格 `price` 的乘积。
- (3) 使用你认为合适的统计图表展示七座城市的总销售额，解读你绘制的图表并推断可能的原因。
- (4) 分别绘制合适的图表展现分城市有无 `wifi`(`wifi`: 有 =1)、有无车位 (`park`: 有 =1) 以及不同折扣力度 (`distype`) 的总销售额的平均值，判断这些因素的不同对总销量平均值是否存在影响？陈述你的观点并适当分析原因。
- (5) 王多鱼是一位投资人，他的资金规模相对不大，近期希望投资小龙虾餐饮业。根据以上的分析，你推荐他投资哪一座城市的小龙虾餐饮业？简单陈述理由。

## 2 第二题：循环与函数

完成以下实验：

- (1) 针对特定函数  $y = \cos(\sin(x))$ ，使用循环编写梯度下降法的程序，要求：当迭代值  $|x^{(n+1)} - x^{(n)}| < 10^{-4}$  时迭代停止，自己选择一个合适的迭代起始点，并且选择一个较为合适的学习率。
- (2) 基于 (1) 中的循环，考虑循环中哪些变量条件可以事先给定，并编写梯度下降法的函数，封装为独立函数后自己选择一个具有多个极小值点的函数 (如  $\sin x$ )，绘制函数的图像，并比较不同的迭代起始位置会对迭代过程产生什么影响。
- (3) 经过以上实验，你觉得梯度下降法在实际使用中应该注意什么？