

# 基于多源空间数据的广州市餐饮业选址

## 本科毕业论文答辩

答辩人: 钟沛  
指导老师: 陈艳男

华南师范大学数学科学学院

2023 年 5 月 11 日



## ① 选题背景及意义

## ② 研究思路

## ③ 实验结果分析

## ④ 总结与展望

# ① 选题背景及意义

## ② 研究思路

## ③ 实验结果分析

## ④ 总结与展望

# 选题背景及意义

- 随着国内餐饮行业市场规模的不断扩大，餐饮行业的竞争也变得更加激烈，门店选址的重要性日益凸显

# 选题背景及意义

- 随着国内餐饮行业市场规模的不断扩大，餐饮行业的竞争也变得更加激烈，门店选址的重要性日益凸显
- 选址要考虑的因素日益复杂，传统的选址方法不再受用，需要发掘更多基于大数据的选址方法

# 选题背景及意义

- 随着国内餐饮行业市场规模的不断扩大，餐饮行业的竞争也变得更加激烈，门店选址的重要性日益凸显
- 选址要考虑的因素日益复杂，传统的选址方法不再受用，需要发掘更多基于大数据的选址方法
- 当今基于大数据的选址模型
  - 移动信令数据、GPS 数据——在现实生活较难获取
  - 有监督机器学习算法——容易受到数据标签的影响

# 选题背景及意义

- 随着国内餐饮行业市场规模的不断扩大，餐饮行业的竞争也变得更加激烈，门店选址的重要性日益凸显
- 选址要考虑的因素日益复杂，传统的选址方法不再受用，需要发掘更多基于大数据的选址方法
- 当今基于大数据的选址模型
  - 移动信令数据、GPS 数据——在现实生活较难获取
  - 有监督机器学习算法——容易受到数据标签的影响

本文提出一种基于开源数据，将传统评价算法和机器学习算法结合的门店选址模型，为餐饮企业提供选址建议

## ① 选题背景及意义

## ② 研究思路

## ③ 实验结果分析

## ④ 总结与展望



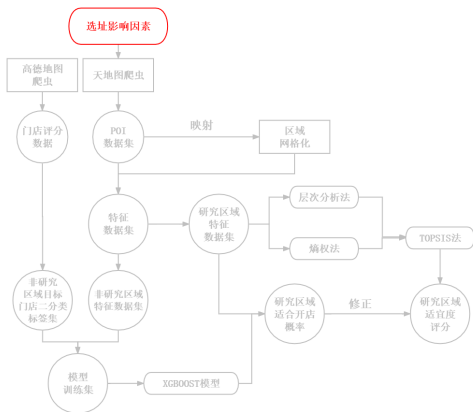
# 研究思路

- 研究区域  
广州市天河区
- 目标选址门店  
冷饮店
- 选址分析  
借助适宜度评分

# 研究思路

## 影响因素

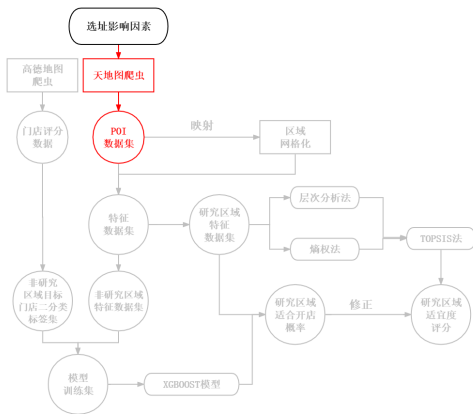
- 人口因素
  - 潜在消费者密度
  - 人口密度
- 交通因素
  - 交通流量
- 竞争因素
  - 竞争强度



# 研究思路

## 影响因素衡量

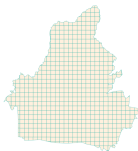
- 潜在消费者密度  
(商城、商务中心、中餐馆、快餐店)
- 人口密度  
(大厦、小区、学校)
- 交通流量  
(交叉路口、地铁站)
- 竞争强度  
(冷饮店、咖啡店、酒吧)



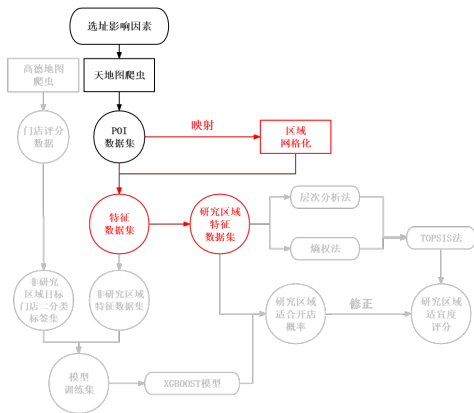
# 研究思路

## 数据预处理

### ① 地理图层网格划分



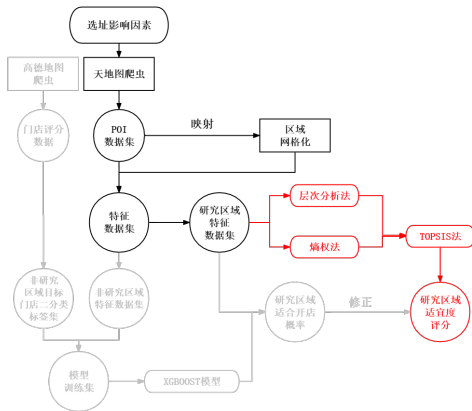
### ② 进行经纬度数据映射得到特征数据集



# 研究思路

## 适宜度评分计算

- 层次分析法
- 熵权法
- TOPSIS 法

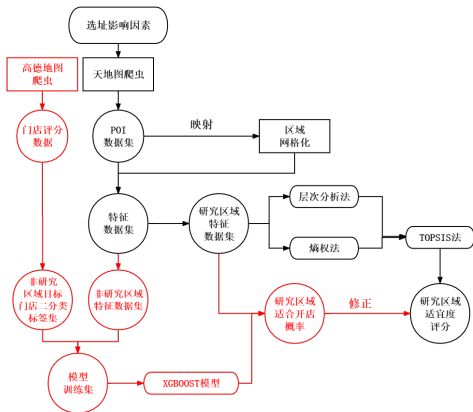


# 研究思路

## 适宜度评分修正

### ● 修正原因

- 在评价类算法中，竞争强度被指定为负向指标
- 竞争强度具有两面性，适度竞争可吸引消费者，但过度竞争会导致恶性竞争，影响经济收益

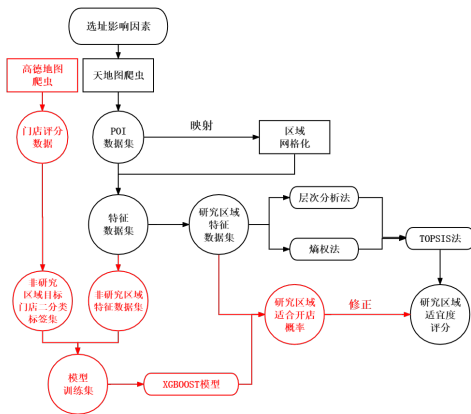


# 研究思路

## 适宜度评分修正

### 修正步骤

- ① 依据选址单元的高德地图门店评分值生成二分类标签集
- ② 利用非研究区域（广州市内，研究区域外）的数据集构建 XGBOOST 模型
- ③ 利用模型预测研究区域的选址单元的适合开店概率对评分进行修正



## ① 选题背景及意义

## ② 研究思路

## ③ 实验结果分析

## ④ 总结与展望

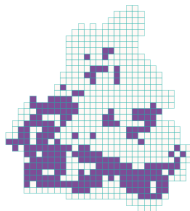


# 实验结果分析

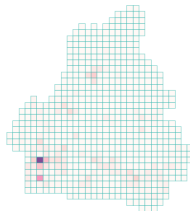
适宜度评分修正前后对比图（颜色越深，适宜度评分越高）

- 未修正的适宜度评分在冷饮店数量较多的区域较低，修正后有所提高，说明模型考虑了竞争的相对强度
- 适宜度评分较好的区域中都有目标选址的存在，表明模型很好地拟合了现有门店布局的空间特征

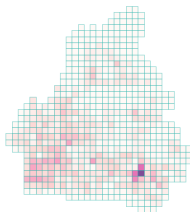
选址单元是否有目标门店存在



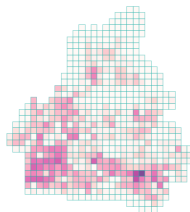
选址单元目标门店数量



未修正适宜度评分



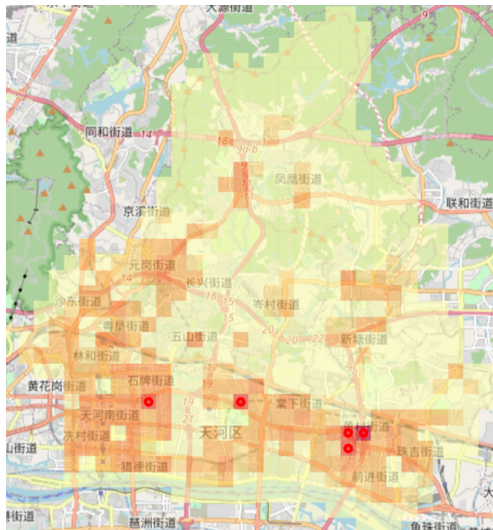
修正后适宜度评分





100%

综上所述,选址模型能够较好地体现研究区域门店选址的适宜度与空间分布的异质性,可以为餐饮门店的选址提供决策依据。



## ① 选题背景及意义

## ② 研究思路

## ③ 实验结果分析

## ④ 总结与展望

# 总结与展望

- 论文创新点
  - 本文的所有数据均来源于开源数据平台，数据可获取性较强
  - 选址算法融合传统选址评价算法和机器学习算法，可以更加全面地对数据进行分析
- 扩展方向
  - 考虑更多的开源数据
  - 考虑更多的指标，如空间统计量等

谢谢观看  
请各位老师批评指正