
湖南城市学院

实验名称	实验三 POI 数据空间分析
姓 名	李灵慧
学 号	2202501-18
学 院	市政与测绘工程学院
专 业	地理空间信息工程
指导教师	汤淼

2025 年 3 月 30 日

1 数据描述

使用 ai 帮助分类后，POI 签到数据的一级类与二级类如下图所示

一级类	二级类	一级类	二级类	一级类	二级类	一级类	二级类	一级类	二级类	一级类	二级类
教育培训	幼儿园	餐饮	湖南菜	出行	火车票代售点	文化休闲娱乐	羽毛球场	商业服务	眼镜店	企事业单位	律师事务所
	小学		西餐		飞机票代售点		保龄球馆		个人用品/化妆品店		出版社
	中学		巴西菜		公交卡 / 月票代售点		乒乓球馆		干洗店		杂志社
	高等院校		浙江菜		溜冰场		植物园		古玩字画店		电视台
住宿	校园生活	小区	台湾菜	文化休闲娱乐	公园户外	银行	溜冰场	党政及福利机构	摄影器材店	其他	楼宇
	职业技术学校		云南菜		一般景点		海滨浴场		老字号		劳动保障类公司
	成人教育		意式菜品餐厅		电影院		售票处		特色商业街		工农贸易类公司
	培训机构		熟食		美容美发店		综合体育馆		宠物用品店		电子商务公司
餐饮	培训学校	医疗卫生	零食	文化休闲娱乐	文化宫	商业服务	国家级景点	企事业单位	旧货市场	其他	IT类公司
	教育出国类公司		西北菜		网吧		省级景点		快速		文艺团体
	学校大门		湖北菜		休闲场所		棋牌室		宠物用品店		报社
	高校中心点		中式餐馆		KTV		档案馆		旅游咨询		机构及其它类公司
餐饮	出行住宿	出行	江苏菜	文化休闲娱乐	健身中心	商业服务	美术馆	企事业单位	维修机构	其他	未知分类
	旅馆招待所		茶艺馆		公园		图书馆		化妆卫浴用品类公司		楼宇大门
	经济型连锁酒店		综合酒楼		公园景点售票处		会展中心		钟表店		无中心点楼宇大门
	三星级宾馆		外国餐厅		公园广场		银行		服装服饰类公司		一般地点标志
餐饮	四星级宾馆	住宅区	自助	文化休闲娱乐	洗浴推拿场所	商业服务	金融服务类公司	党政及福利机构	儿童用品店	其他	大门
	五星级宾馆		东北菜		滑雪场		便民商店/便利店		政府机构		
	度假疗养场所		上海菜		纪念馆		购物服务		警察局		
	餐饮美食		社区服务		展览馆		商场		税务局		
餐饮	中餐厅	医疗	三甲医院	文化休闲娱乐	博物馆	商业服务	家电维修卖场	企事业单位	消防	其他	
	河北菜		诊所		世界遗产		家居装饰类公司		丧葬		
	四川菜		药店		夜市		汽车交通类公司		教堂		
	甜品店		医疗健康类公司		城市广场		消费电子类公司		回教寺		
餐饮	快餐店	出行	加油站	文化休闲娱乐	剧场	商业服务	房地产类公司	企事业单位	寺庙道观	其他	
	火锅店		收费站		酒吧		家具建材市场		邮局		
	烧烤		火车站		游戏厅		珠宝首饰工艺品		公司		
	特色 / 地方风味餐厅		公交车站		迪厅		服装鞋帽皮具店		楼宇机构		
餐饮	日本料理	出行	长途汽车站	文化休闲娱乐	台球厅	商业服务	专卖店	企事业单位	工厂	其他	
	休闲餐饮场所		飞机场		动物园		体育用品店		通讯及服务类公司		
	糕饼店		汽车维修		游乐园		书店		产业园区		
	冷饮店		汽车租赁		垂钓园		烟酒专卖店		知名企业		
餐饮	清真餐馆	出行	二手车交易	文化休闲娱乐	运动场所	商业服务	文化用品店	企事业单位	商务服务类公司	其他	
	广东菜		汽车俱乐部		网球场		音像店		文化体育类公司		
	海鲜酒楼		4S店		游泳馆		小商品市场		食品饮料类公司		
	北京菜				高尔夫相关						
餐饮	韩国料理	出行		文化休闲娱乐	跆拳道馆	商业服务		企事业单位		其他	
	咖啡厅				科技馆						

图 1 一级类与二级类

2 数据采集

新浪微博签到 POI 数据通过新浪微博开放平台 API 获取数据获取，时间：2014 年 11 月，范围：河北省，共 322878 条数据，有效数据 321690 条，其中，石家庄 64184 条，餐饮 10082 条（清洗后），住宅区 2287 条（清洗后）。

石家庄市主城区行政边界数据来源于阿里云 DataV 数据可视化平台,获取时间为 2025 年 3 月，格式 geojson。

交通路网数据来源于 OpenStreetMap 网站(<https://www.openstreetmap.org>),获取时间为 2024 年 6 月，格式 osm。

2015 常住人口栅格数据（1km）,源自 WorldPop 网站(<https://hub.worldpop.org/>)。

2015 年夜间灯光和 GDP 栅格数据（1km 分辨率）来自中国科学院资源环境数据中心(<https://www.resdc.cn/>)

3 数据预处理

3.1 数据问题

（1）定位异常，经纬度坐标明显超出研究区范围

lon	lat
108.331384	22.83508056
115.63932	22.945308
120.228746	22.961889

图 2.1.1 定位异常

- (2) 种类繁多，category_name 字段有两百多种值，不利于分析
- (3) 只需要研究石家庄市的数据，但数据中包含其它城市
- (4) 获取的栅格数据范围过大，栅格数据值无法归一计算
- (5) ArcMap 不支持 OSM 格式
- (6) 地理探测器需要整理表格数据，各字段的数据尺度统一
- (7) 数据坐标系和单位不统一，栅格数据像元大小不统一

3.2 应对方法

(1) 查询研究区经纬度范围，使用 python 的 pandas 库进行筛选处理，或者在 ArcMap 中使用 clip 工具裁剪。

(2) 获取所有种类名称，使用 ai 将种类分成教育培训、住宿、餐饮、小区、医疗卫生、出行、文化休闲娱乐、银行、商业服务、党政及福利机构、企事业单位、其它这 12 个大类，因为我选取餐饮作为研究对象，又将餐饮分为中式快餐、国际餐饮、小吃快餐。在 ai 完成分类后，进行校对，制作成映射表，用 python 进行数据识别。最后筛选需要的数据并输出。

分类类型	原始类型	分类类型	原始类型	分类类型	原始类型	分类类型	原始类型
中式餐饮	中餐厅	中式餐饮	江苏菜	国际餐饮	日本料理	甜点饮品	甜品店
	河北菜		福建菜		韩国料理		糕饼店
	四川菜		东北菜		巴西菜		冷饮店
	广东菜		上海菜		意式菜品餐厅		茶艺馆
	北京菜		火锅店		外国餐厅		咖啡厅
	湖南菜		烧烤	小吃快餐	西餐厅	其它	餐饮美食
	浙江菜		海鲜酒楼		快餐厅		休闲餐饮场所
	台湾菜		清真餐馆		熟食		
	云贵菜		中式素餐馆		零食		
	西北菜		综合酒楼		自助		
湖北菜	特色 / 地方风味餐厅	甜点饮品	甜品店				

图 2.2.2 餐饮分类对照表

(3) 查询新浪微博开放平台中的城市代码表，可知石家庄代码为 0311，筛选 city 字段中等于 0311 的数据。

(4) 使用研究区 shp 文件进行裁剪、然后重分类，根据自然断点法分割成 10 个等级。

(5) 使用插件 ArcGIS editor for OSM 转化数据。

(6) 将道路网使用密度工具转成栅格数据，将所有栅格数据坐标系统一，重分类

后, 使用提取至点工具提取到点要素, 然后导出表格进行地理探测器分析。

3.3 处理结果

(2) 得到分类后的石家庄餐饮数据点

4 空间分布分析

部四类，其中中部餐饮点最为密集，其余呈零星分布

餐饮点分布K-means聚类分析

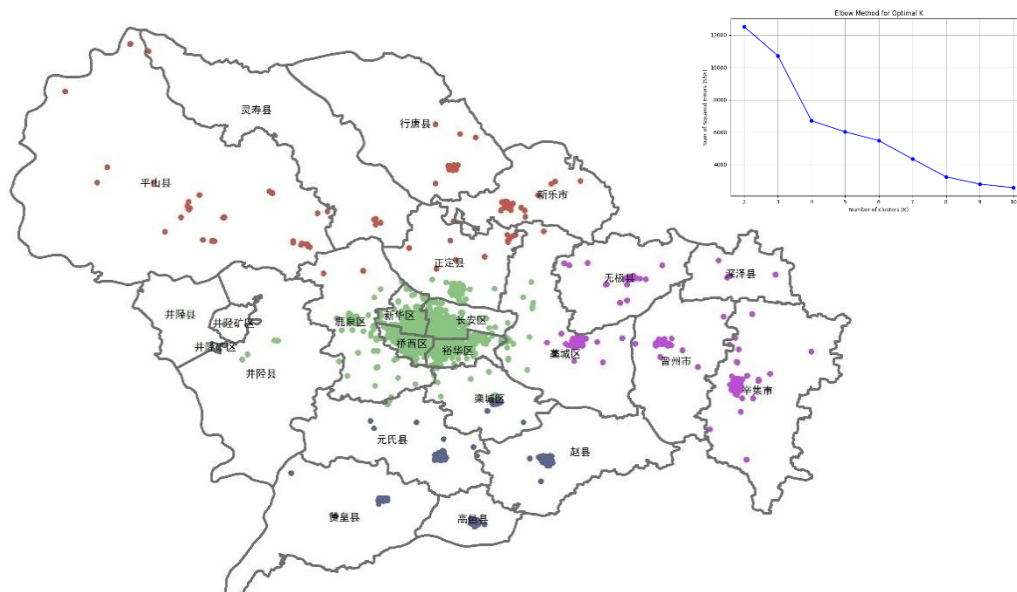


图 3.2 餐饮点分布聚类分析

4.3 近邻表。

通过近邻表计算要素之间的距离及其他邻近性信息,探索城市中各对象之间的空间距离。本实验利用近邻表统计每一餐饮网点与其最邻近餐饮网点之间的距离,分析石家庄市餐饮分布的空间距离特征。

由下图可知，四类餐饮绝大部分在 500m 以内,极少数散落在 1000m 之外。国际餐饮的竞争最为激烈，平均距离最短，而且由图 3.2 可知，国际餐饮主要分布于中心城区，这应该是由于国际餐饮消费水平高，只有经济发达地区可以支持其运营。

类型	总数	与最邻近餐饮点之间的距离/m					
		平均距离	最大距离	<200	<500	>1000	>2000
中式餐饮	5742	67.33	8714.16	5510	5660	48	36
国际餐饮	194	38.29	1256.96	187	190	1	0
小吃快餐	1302	73.15	17966.78	1252	1279	8	7
甜点饮品	1328	51.93	10313.62	1288	1318	4	2

图 3.3 邻近表

4.4 核密度。

核密度是研究区域差异的常用方法[33],通过识别空间中的热点区域反映空间数据点的集聚程度。本实验运用该方法分析石家庄市餐饮分布的空间集聚特征。

由下图可见，中心城区，尤其是新华区、长安区、桥西区、裕华区四区的交界处餐饮点最为集中，其余地区只有零星分布，其中国际餐饮的集中性最为明显

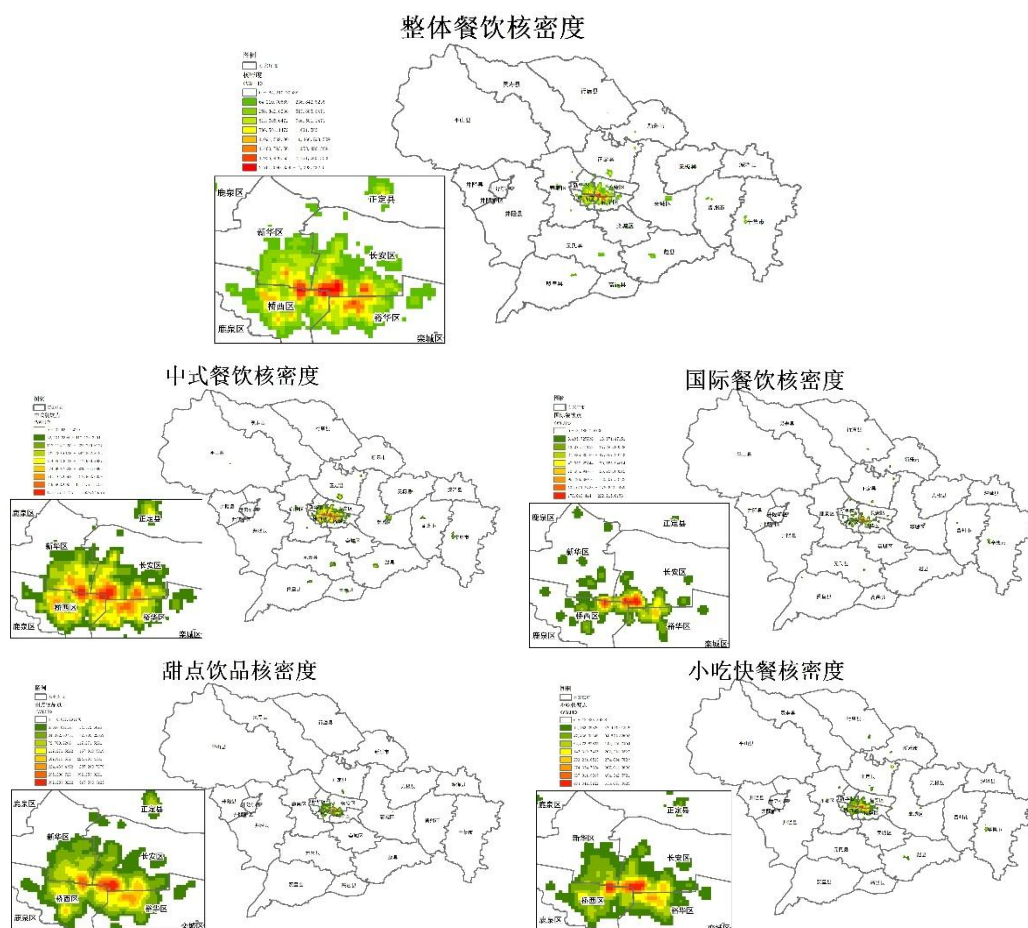


图 3.4 核密度分析

5 相关性分析

5.1 空间统计

逐个统计了各个住宅点 2000m 范围内的餐饮点个数，并将不同个数的住宅点分以不同颜色显示，反映了住宅点和餐饮点之间的相关性，计数越多的住宅点就餐越便利且选择更多样，在中心四区交界的住宅就餐是最方便的，而且便利程度以交界区为中心呈放射性衰减。

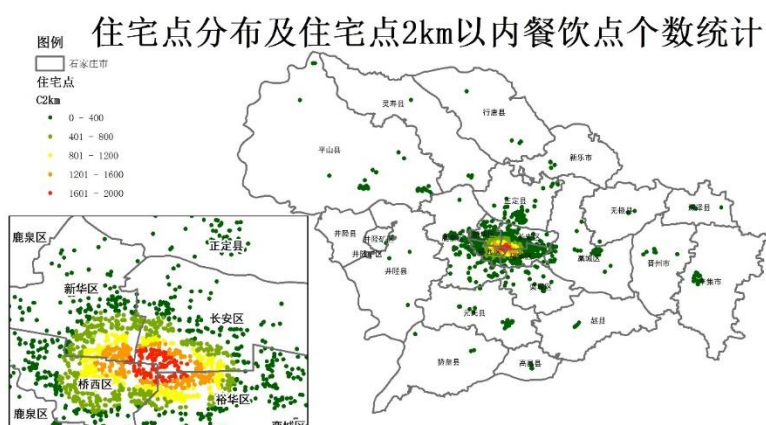


图 3.1 各个住宅点 2km 以内餐饮点个数统计

5.2 地理探测器

使用 GeoDetector 软件对就餐便利性（住宅点附近餐饮点个数分 10 类得到）和道路密度（RDD）、人口（POP）、夜间灯光（NPP）、经济（GDP）进行影响因子交互探测分析和单因子探测分析，得到结果如下，由图可知，道路密度的影响是最大的，其余的因素差不多，道路密度和夜间灯光以微弱的差距占据相关性最大值，由此可见对于餐饮来说，交通的便利性是首要考虑的因素。

	RDD	POP	NPP	GDP
RDD	0.5120			
POP	0.6086	0.4210		
NPP	0.6217	0.5625	0.4651	
GDP	0.6137	0.4606	0.5618	0.4015

	RDD	POP	NPP	GDP
q 值	0.5120	0.4210	0.4651	0.4015
p 值	0	0	0	0

图 3.2 地理探测器

6 总结

本次实验基于石家庄市主城区的 POI 数据，通过多源数据整合与空间分析方法，系统探究了餐饮点的空间分布特征及其与交通、人口、经济等要素的关联性。主要结论如下：

空间分布特征：

石家庄市餐饮点呈现显著的集聚性，集中分布于中心城区的交界地带（新华区、长安区、桥西区、裕华区），其中中式餐饮占比最高，国际餐饮则高度集中于经济活跃区域。

K-means 聚类与核密度分析表明，餐饮点可分为中部、北部、南部、东部四类集群，中部区域密度最高，其余区域分布相对稀疏。

近邻表分析显示，餐饮网点间平均距离多在 500 米以内，国际餐饮竞争最为激烈，进一步印证其选址对区位经济条件的敏感性。

影响因素分析：

地理探测器结果表明，交通路网密度是影响餐饮分布的首要因素（q 值最大），其次为夜间灯光强度与 GDP 水平，人口密度相关性相对较弱。这说明餐饮选址更倾向于交通便利、经济发达的区域，与商业活力及消费能力密切相关。

住宅点与餐饮点的空间相关性显示，中心城区交界处的住宅区就餐便利性最高，便利性呈放射性递减趋势，体现了“居住-商业”空间耦合特征。

方法与技术验证：

通过数据清洗、分类映射、栅格重分类、栅格重采样等技术，解决了原始数据的定位异常、种类冗余、范围偏差等问题，验证了多源异构数据整合的可行性。

K-means 聚类、核密度分析与地理探测器的综合应用，有效揭示了空间分布模式及驱动机制，为城市商业规划提供了定量化支持。

局限与改进：

实验数据时间跨度较大（2014 年 POI 数据与 2024 年路网数据），可能影响分析结果的时效性。

地理探测器仅分析了线性相关性，未考虑空间异质性影响。后续可引入地理加权回归（GWR）等模型深化机制解析。

未来可以使用 folium 库进行多时序大数据可视化，动态展示多时间段的动态变化。