

|  |  |
| --- | --- |
| **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



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

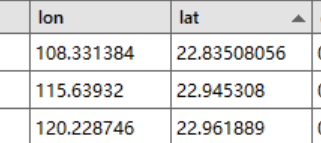


图2.1.1定位异常

（2）种类繁杂，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。

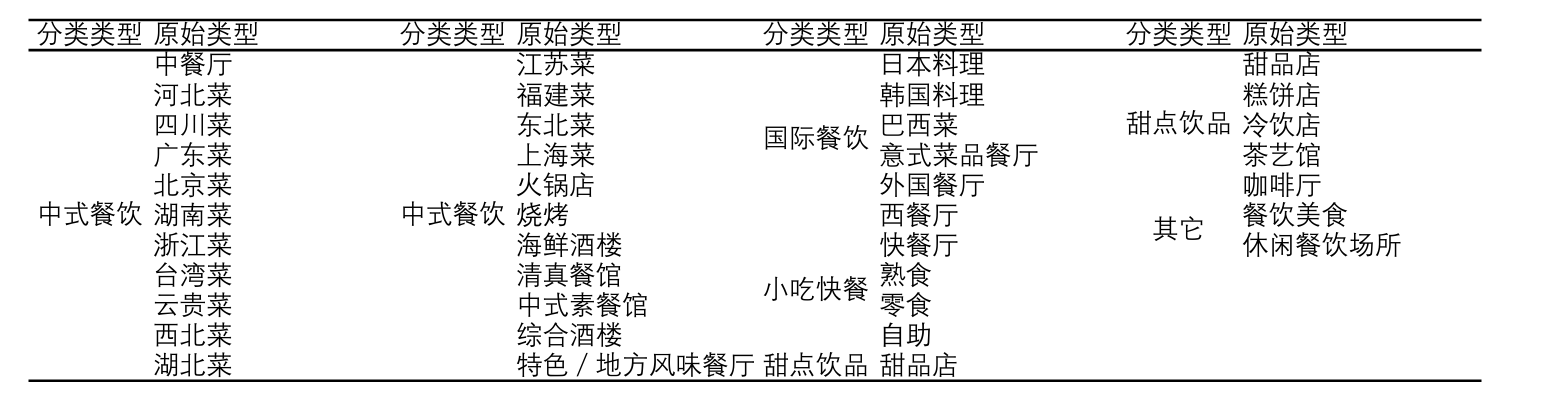


图2.2.2 餐饮分类对照表

（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。

（7）重新投影，重采样栅格数据。

3.3处理结果

（1）得到石家庄住宅区数据点

（2）得到分类后的石家庄餐饮数据点

（3）得到重分类和统一坐标系、像元大小后的道路密度、GDP、人口、夜间灯光数据

4 空间分布分析

4.1 分布情况

明显可知餐饮点集中分布在中心城区（新华区、长安区、桥西区、裕华区），并且中式餐饮占大多数，其中中式餐饮5742个、其它1516个、甜点饮品1328个、小吃快餐1302个、国际餐饮1954个。

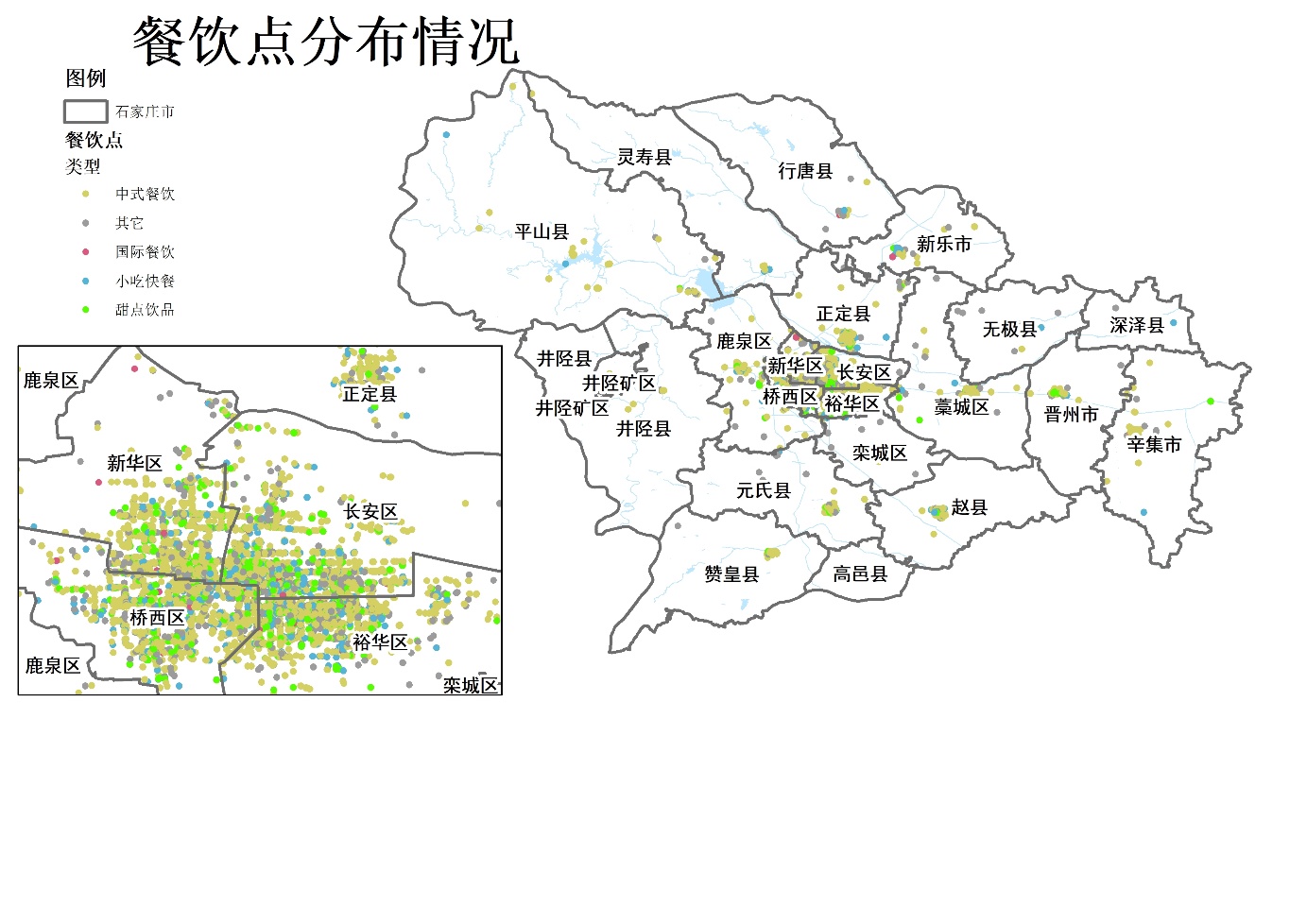


图3.1 餐饮点分布情况

4.2 K-means

通过k-means算法对lat、lon字段进行空间聚类，根据肘部图，拐点差不多在4左右，故选k值为4，得到结果如下图，可以明显看到餐饮点分成中部、北部、南部、东部四类，其中中部餐饮点最为密集，其余呈零星分布

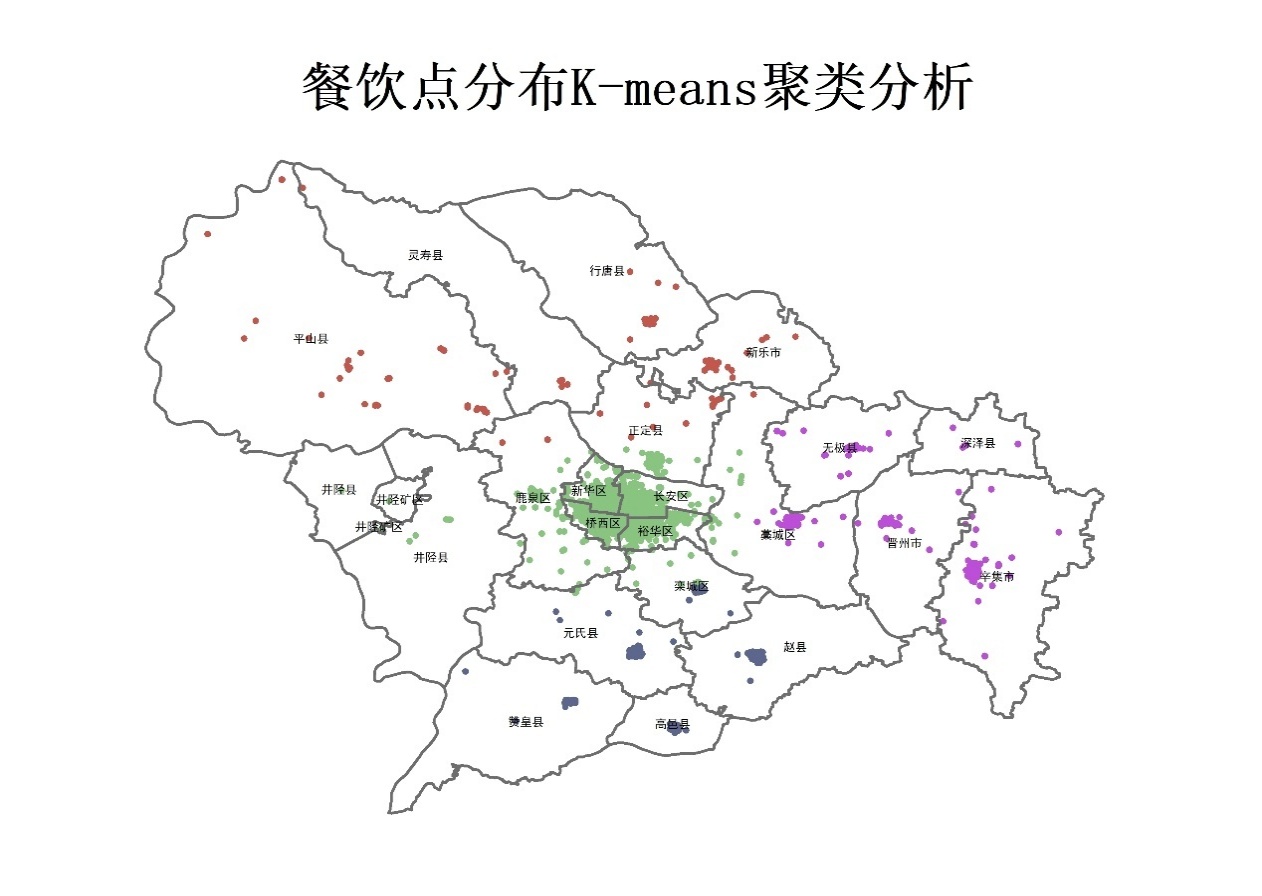
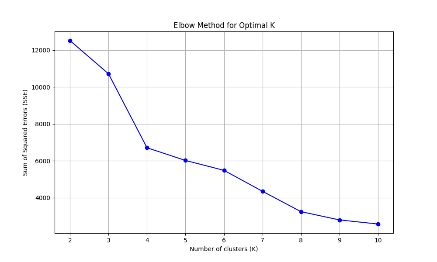


图3.2 餐饮点分布聚类分析

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。

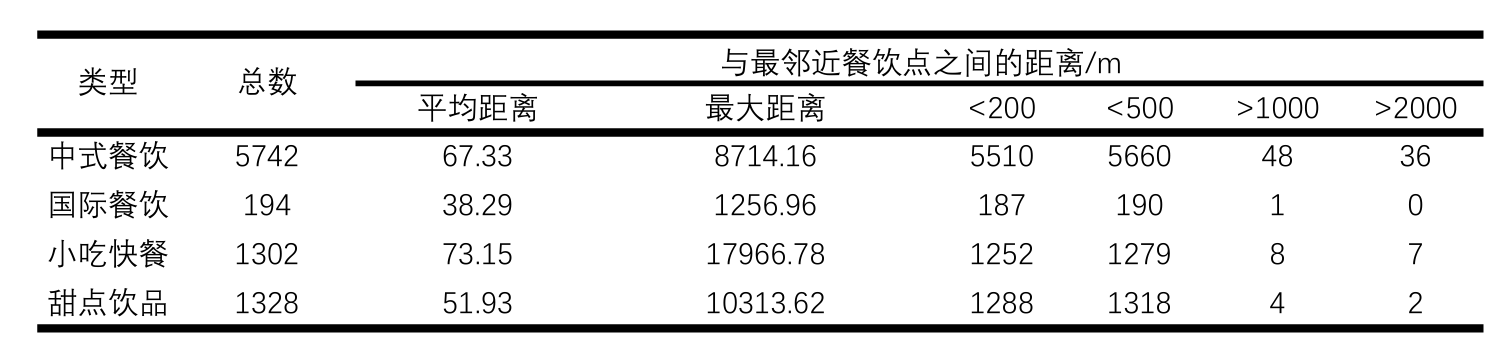
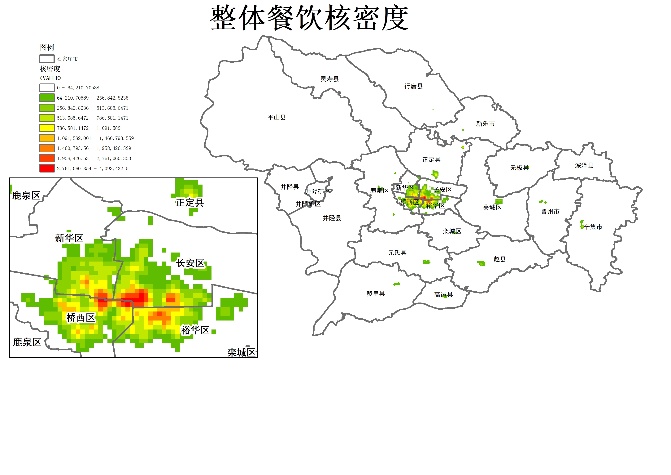


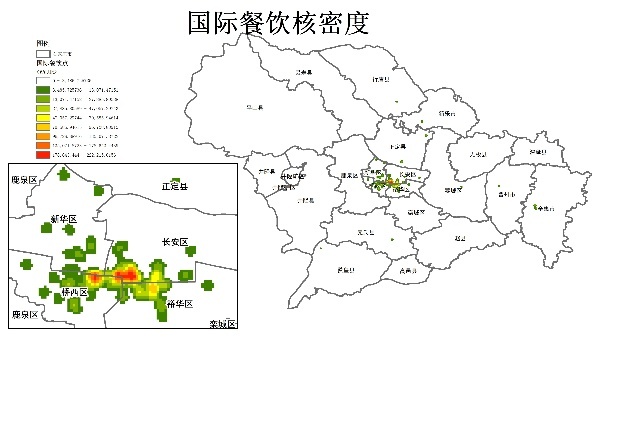
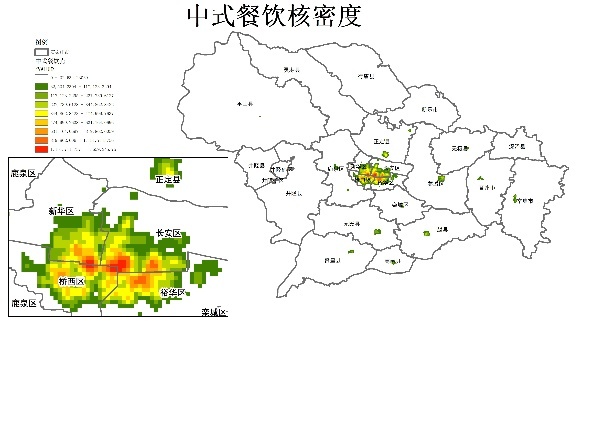
图3.3 邻近表

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





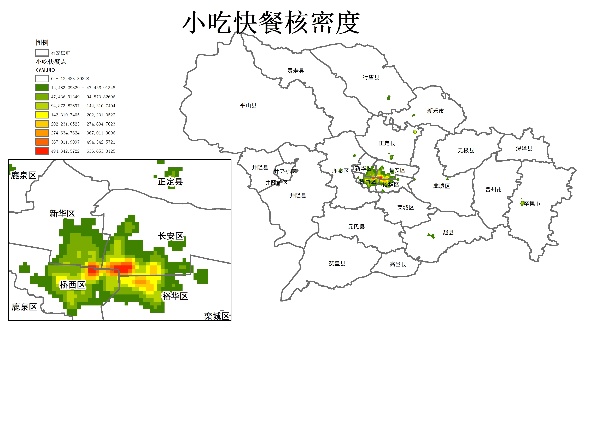
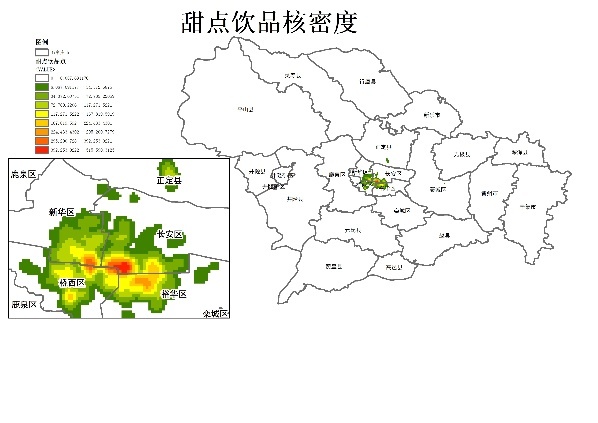


图3.4 核密度分析

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。

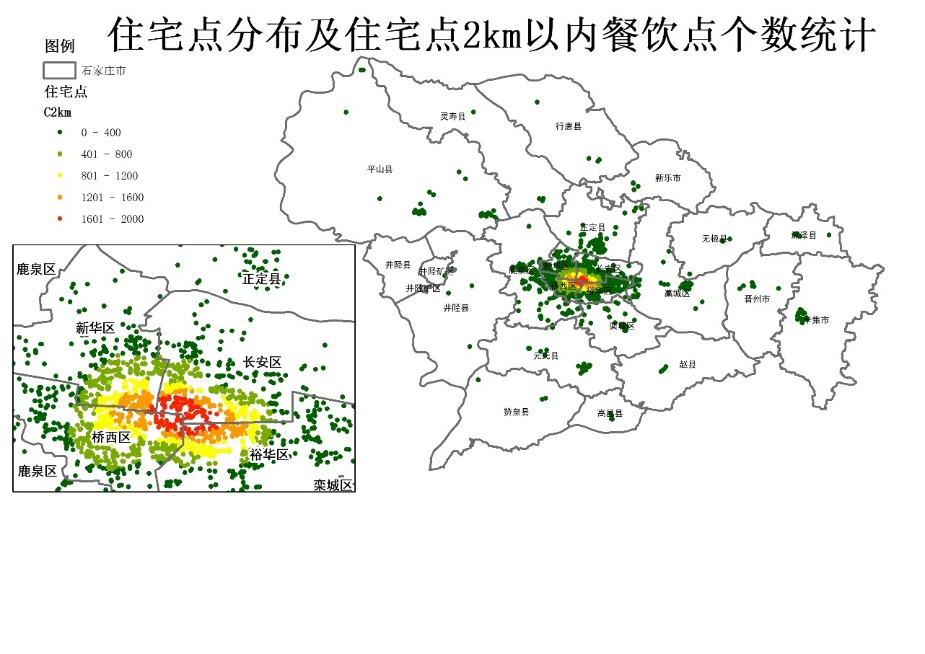


图3.1 各个住宅点2km以内餐饮点个数统计

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。

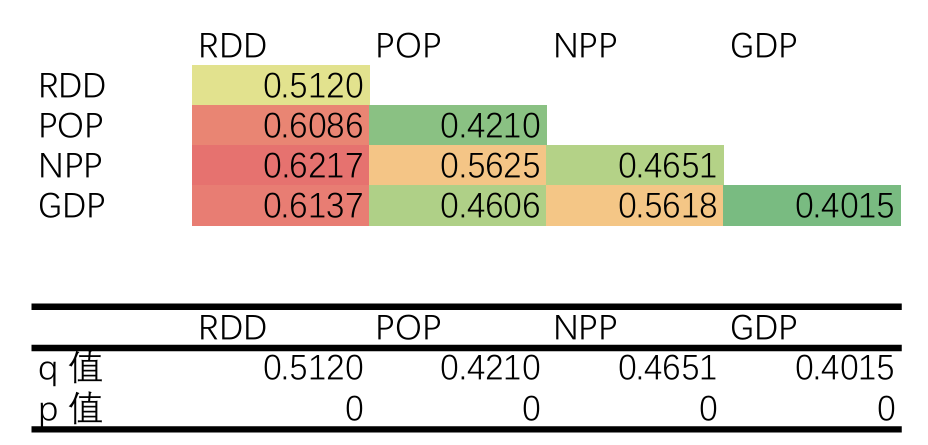


图 3.2 地理探测器

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。