

ESRI 中国（北京）培训中心

ESRI China(Beijing) Learning Center

发掘地图背后的秘密

——探析ArcGIS空间统计工具

讲座人：王珏

wangjue@lreis.ac.cn

内容提要

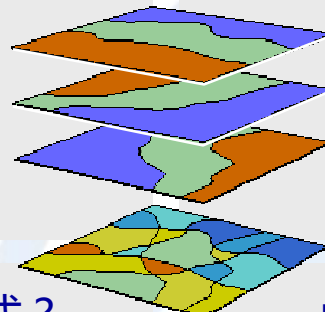
- 什么是空间统计学
- 为什么要用空间统计
- 空间统计学基础
- 空间统计工具简介

什么是空间统计学

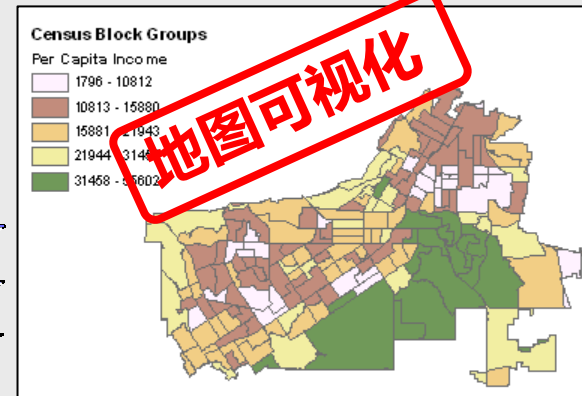
年收入大于20万的家庭占总数的百分比？

TOTPOP00	HHPOP00	POPFAMHH	PCI00
1059	998	747	9027
1126	710	587	7559
918	918	778	21340
997	997	794	9806
511		87	8660
1312	1312	1147	5595
1167		988	15059
1468	1468	1251	17253
2185	2185	1715	15707

传统的统计学



年收入大于20万的家庭分布在哪些地方？



年收入大于20万的家庭分布有何模式？

分析

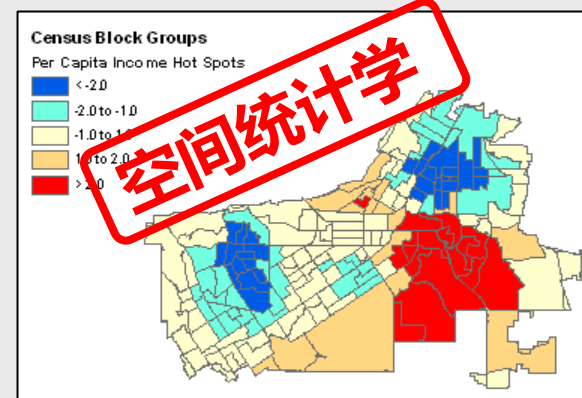
哪里是高收入人群聚居的热点区域？



Dispersed

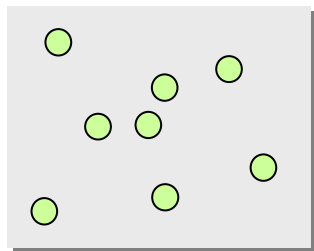
Clustered

空间统计学

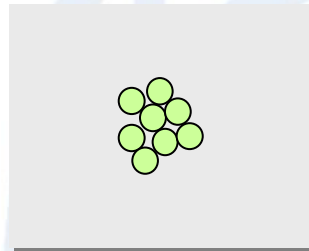


什么是空间统计学

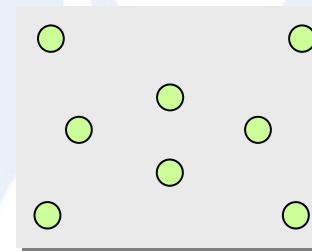
- 制图和传统数理统计的强化和补充
- 用于：
 - 发现数据的空间分布模式、趋势、过程和空间关联
 - 对人文、社科、经济类数据做研究分析



随机分布

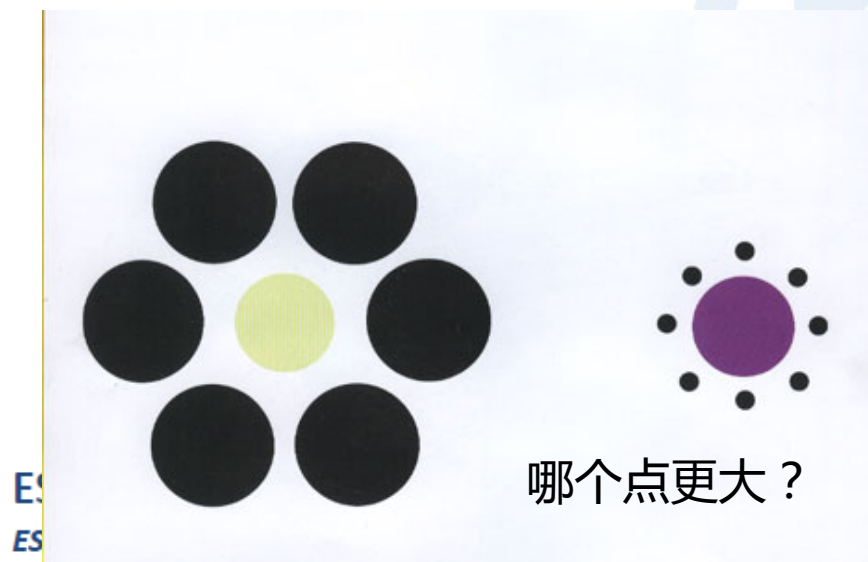
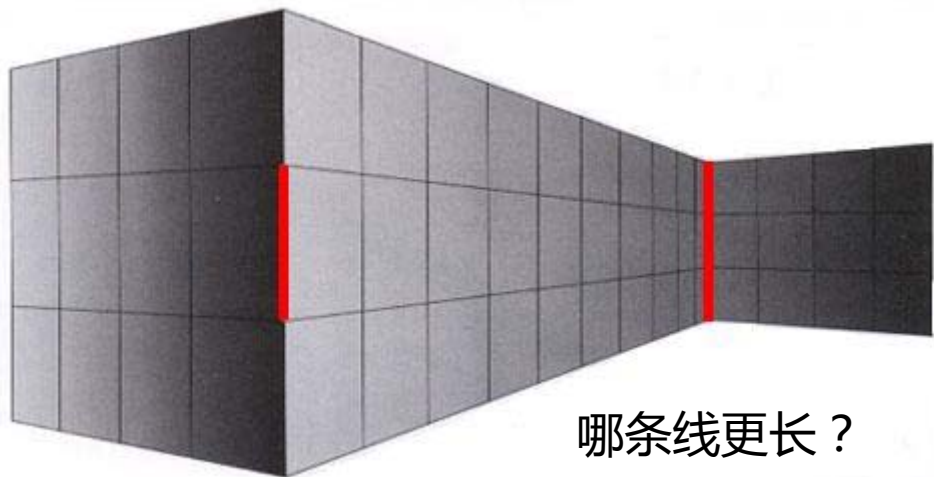


聚类分布



离散分布

为什么要用空间统计 (1) 眼见一定为实吗？

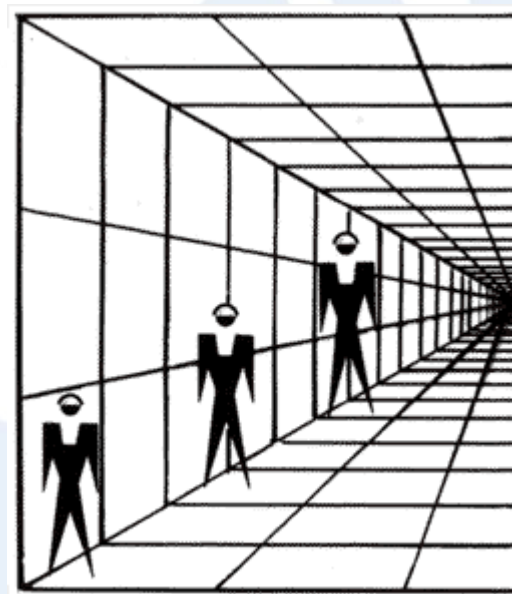


图片会产生**错觉**

眼睛会**欺骗**大脑

用**数字**说话才可靠

哪个人更高？



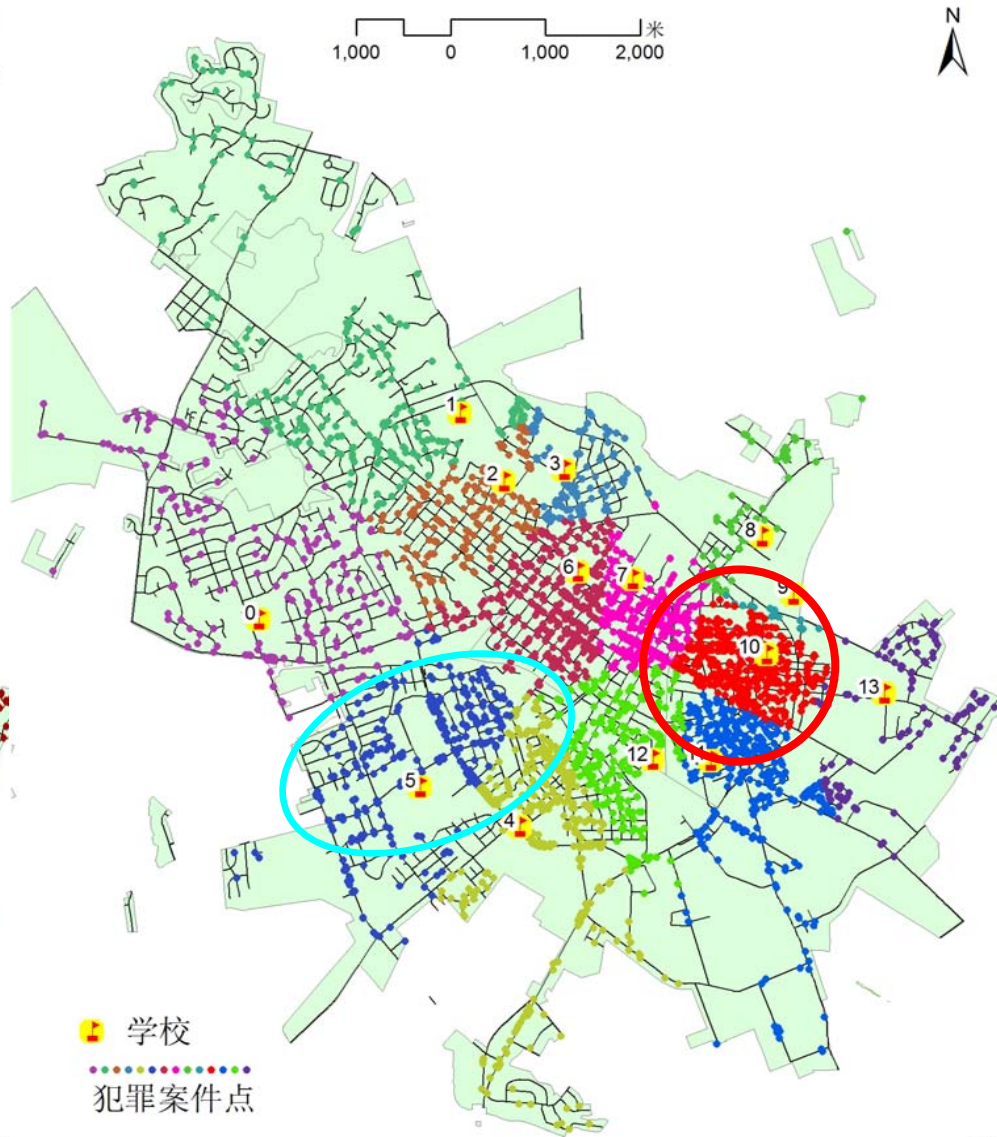
为什么要用空间统计 (2)

地图不能说明问题吗？

校园周边犯罪案件分布图 (一)



校园周边犯罪案件分布图 (二)

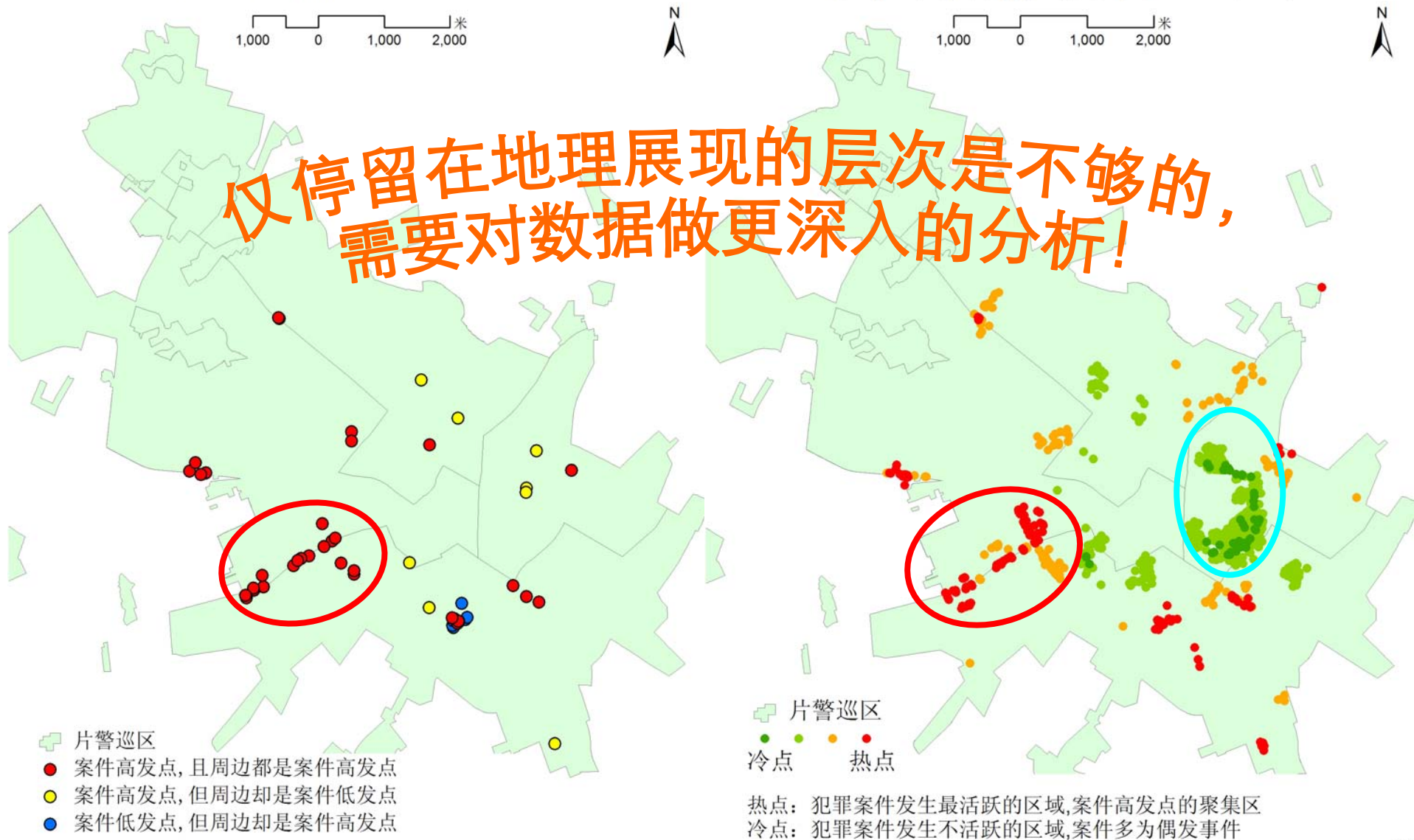


为什么要用空间统计 (3)

深入发掘隐藏在数据背后的秘密

犯罪案件热度分析图 (一)

犯罪案件热度分析图 (二)

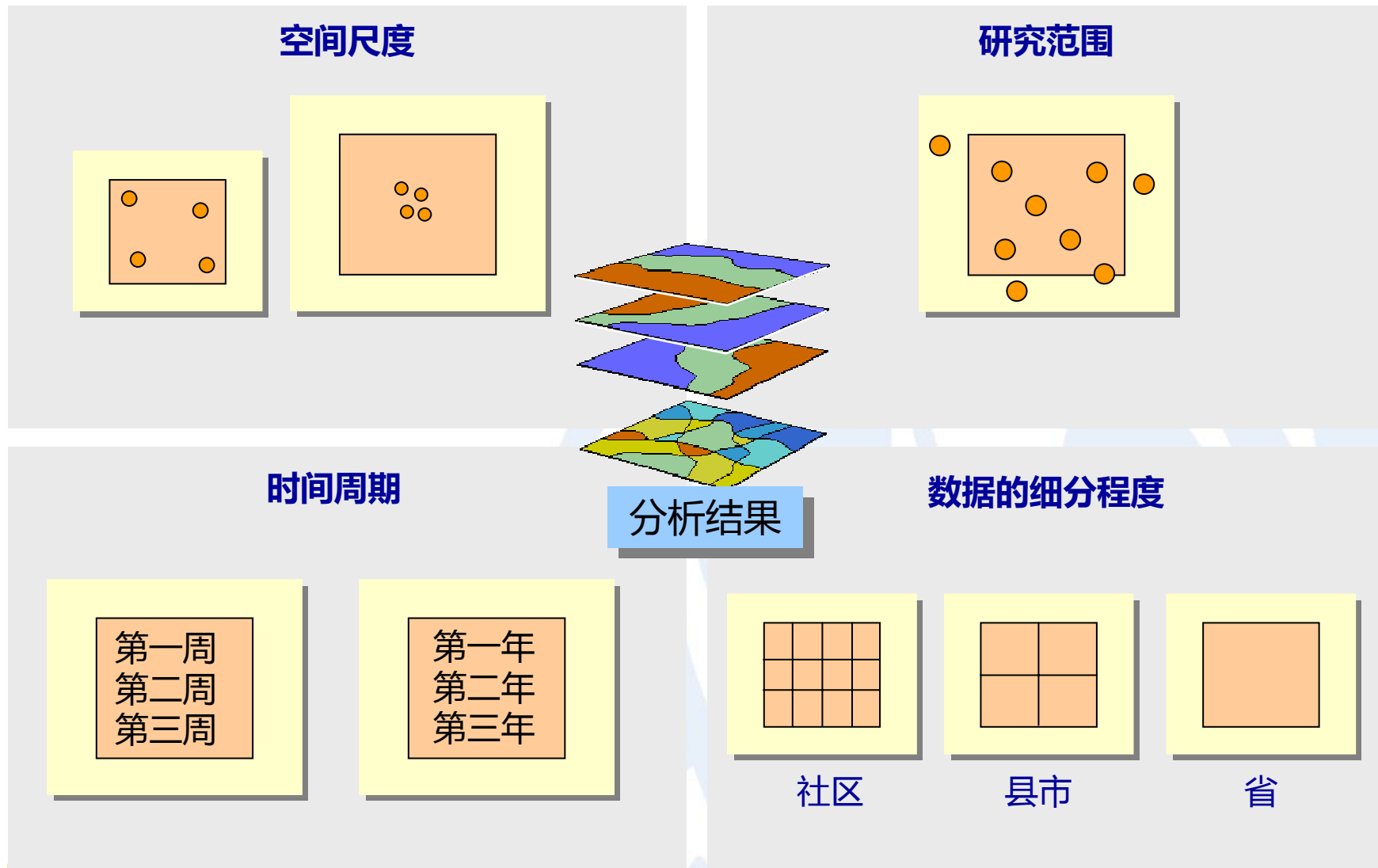


空间统计学基础

- 影响分析结果的几个因素
- Z值与p值
- 模拟空间关系
- 距离与面积

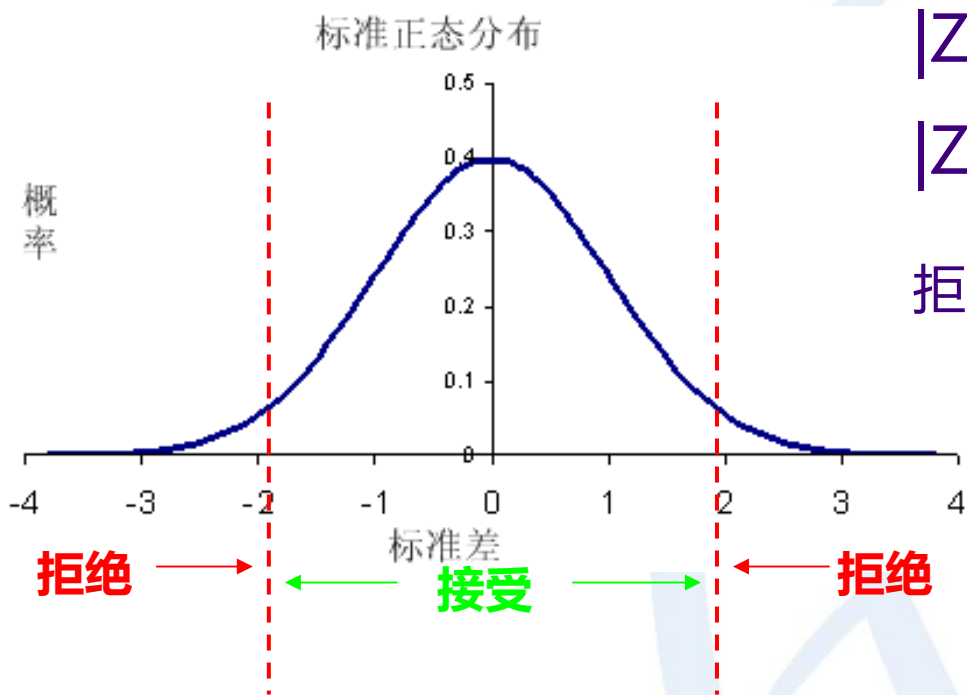


影响分析结果的几个因素



Z值与p值

- 零假设
 - 空间统计中的零假设是支持随机分布
- Z值与p值：帮助判断是否能够推翻零假设



|Z|越小，p越大：**接受**零假设

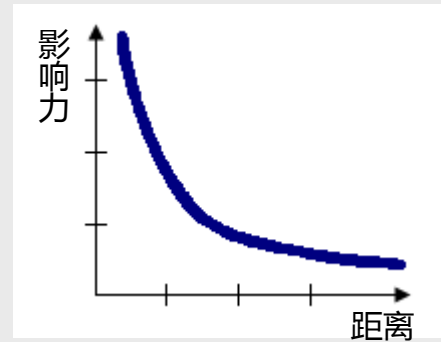
|Z|越大，p越小：**拒绝**零假设

拒绝域的临界值，由**置信水平**决定

模拟空间关系

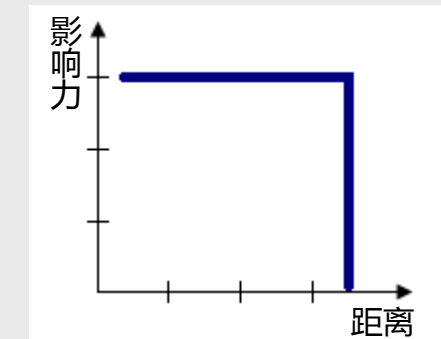
◆ 反距离/反距离平方

- ◆ 所有要素都对目标要素产生影响
- ◆ 距离越远，影响越小



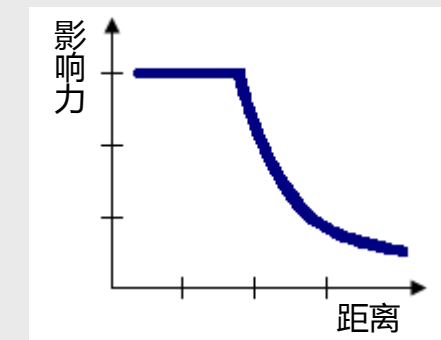
◆ 固定距离段

- ◆ 只有指定距离之内的要素对目标要素产生影响
- ◆ 影响范围内的所有要素，影响力相同



◆ 无差异带

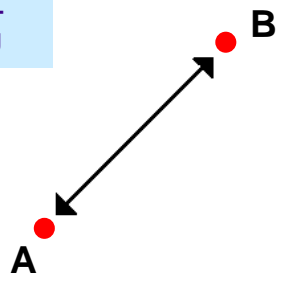
- ◆ 反距离与固定距离段的综合
- ◆ 指定范围之内的要素具有相同的影响力，超出范围的要素其影响力随距离递减



距离与面积

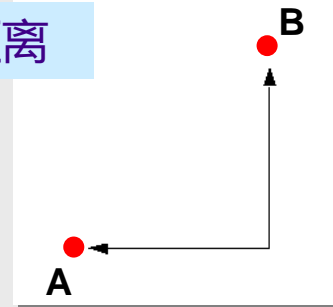
- Projected Coordinate System
 - 使用以米为单位的坐标系统，不要使用经纬度为单位的坐标系统

直线距离



$$D = \sqrt{(X_1 - X_2)^2 + (Y_1 - Y_2)^2}$$

曼哈顿距离

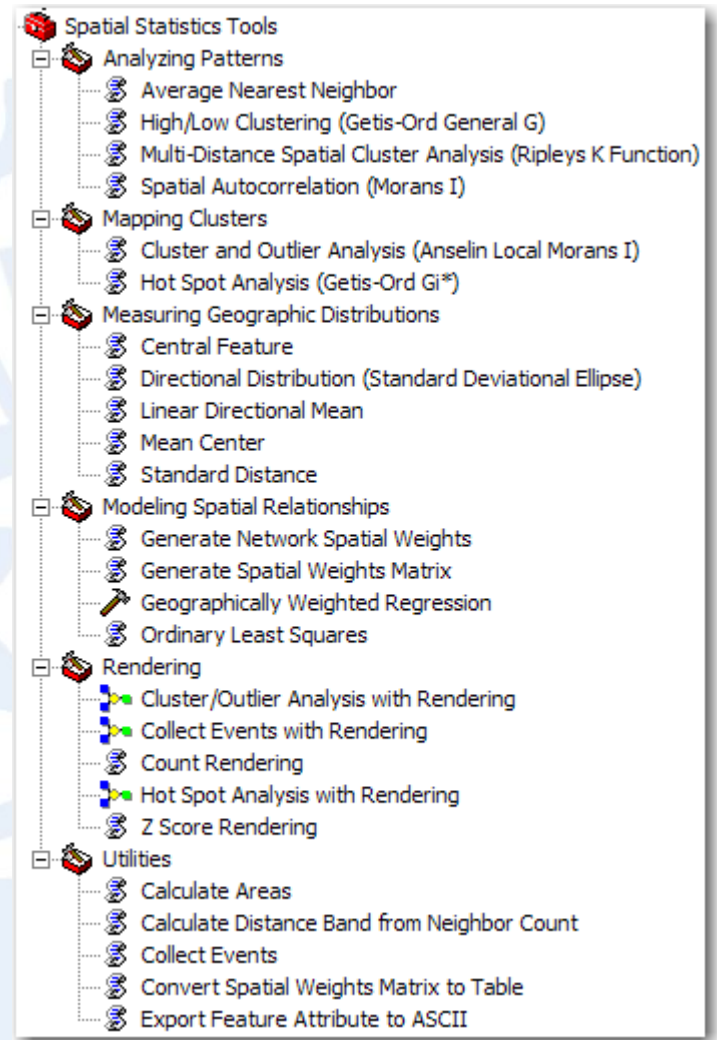


$$D = |X_1 - X_2| + |Y_1 - Y_2|$$

Spatial Statistics

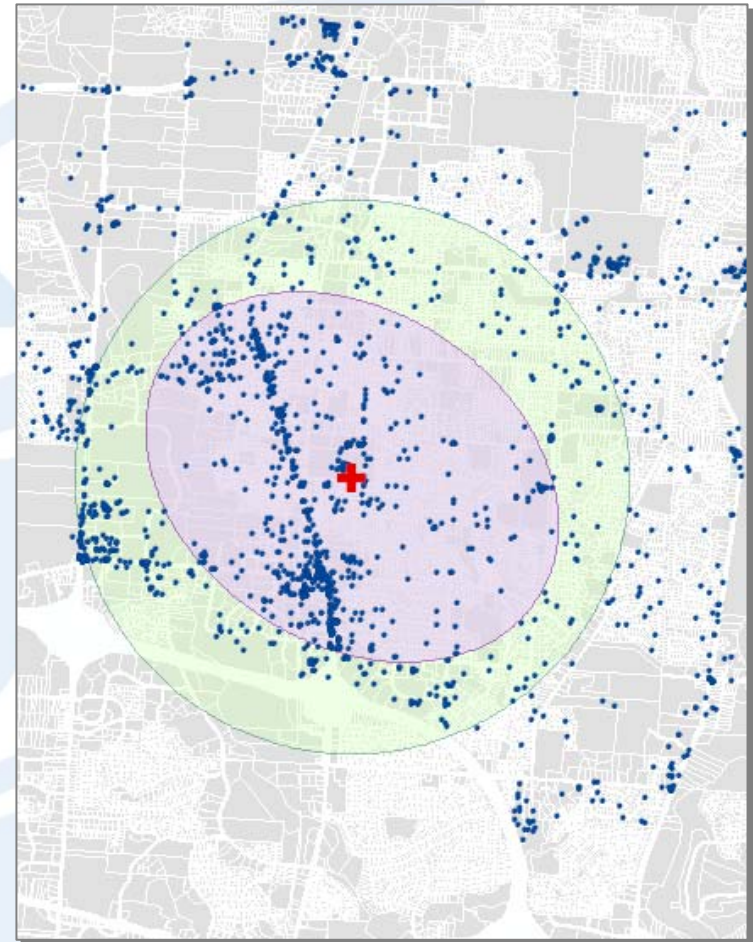
空间统计工具简介

- 所有license等级均可使用
- 使用python编写
 - 源代码公开
 - 鼓励用户利用源代码学习、修改、扩展、分享这些分析工具



Measuring Geographic Distributions 测量地理分布工具集

- 识别分布特征
 - 哪里是中心
 - 向中心聚集还是发散
 - 分布趋势如何
- 用于：
 - 验证可视化分析
 - 对比不同类别的要素
 - 检查随时间的变迁



Mean Center 平均中心工具

- 计算研究区域的平均xy坐标
- 用于:
 - 跟踪分布中心的变化
 - 对比不同类别要素的分布中心

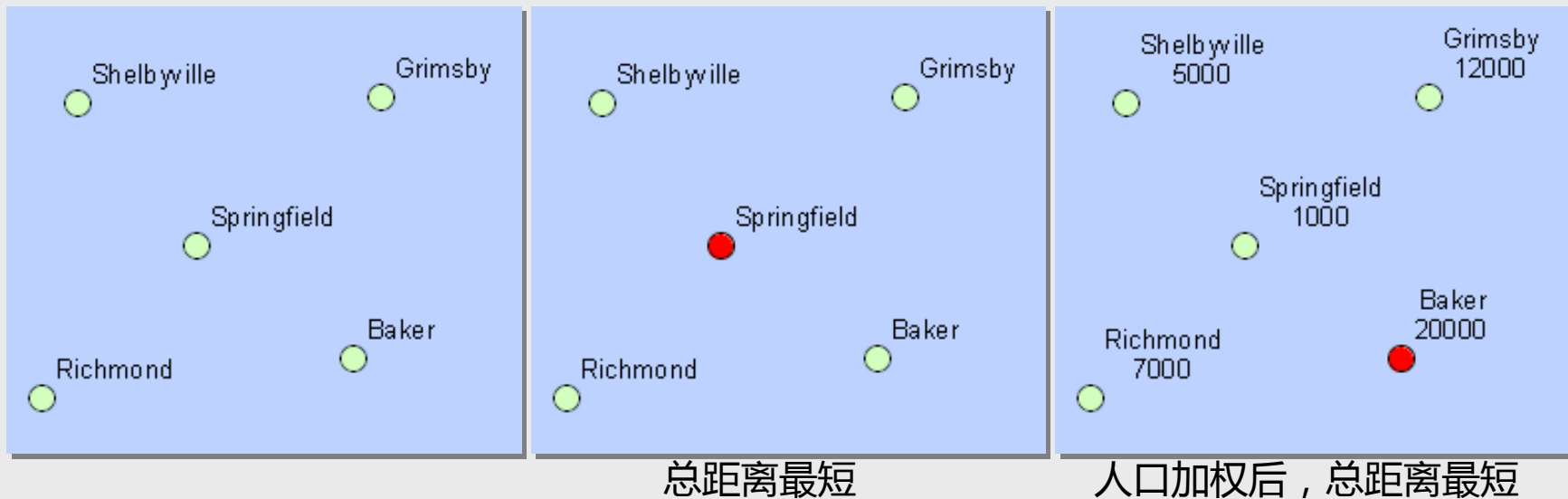


Central Feature

中心要素工具

- 找出位于最中心的要素
 - 到其他所有要素的距离和最短的那个要素
- 用于：
 - 在已有要素中寻找通达条件最佳的

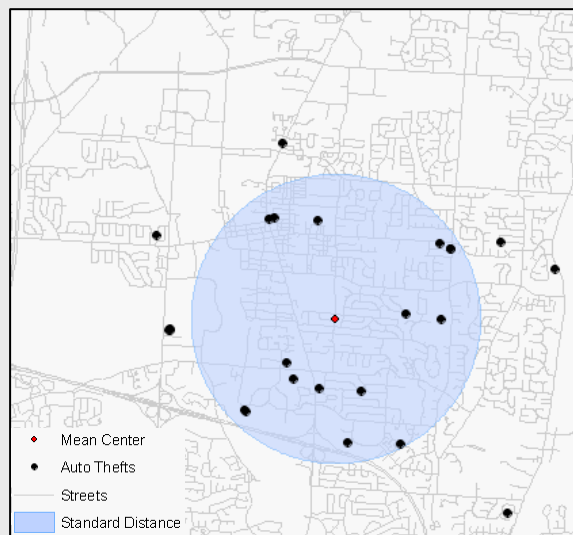
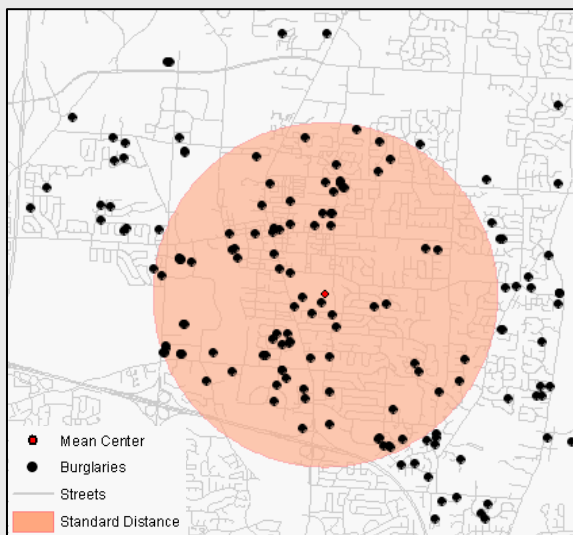
在哪个小区建一个社区图书馆最佳？



Standard Distance 标准差距离工具

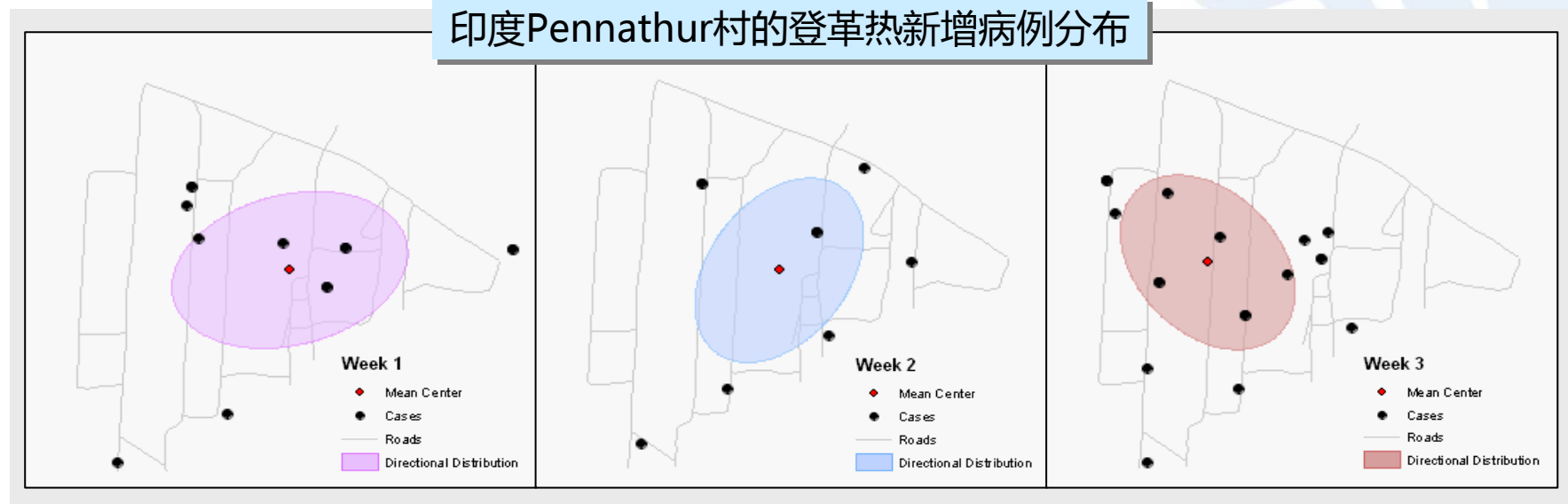
- 衡量要素距离平均中心的离散程度
 - 结果以距离的形式简单直观地展现出来
- 对异常点敏感

入室盗窃案与汽车盗窃案的分布对比



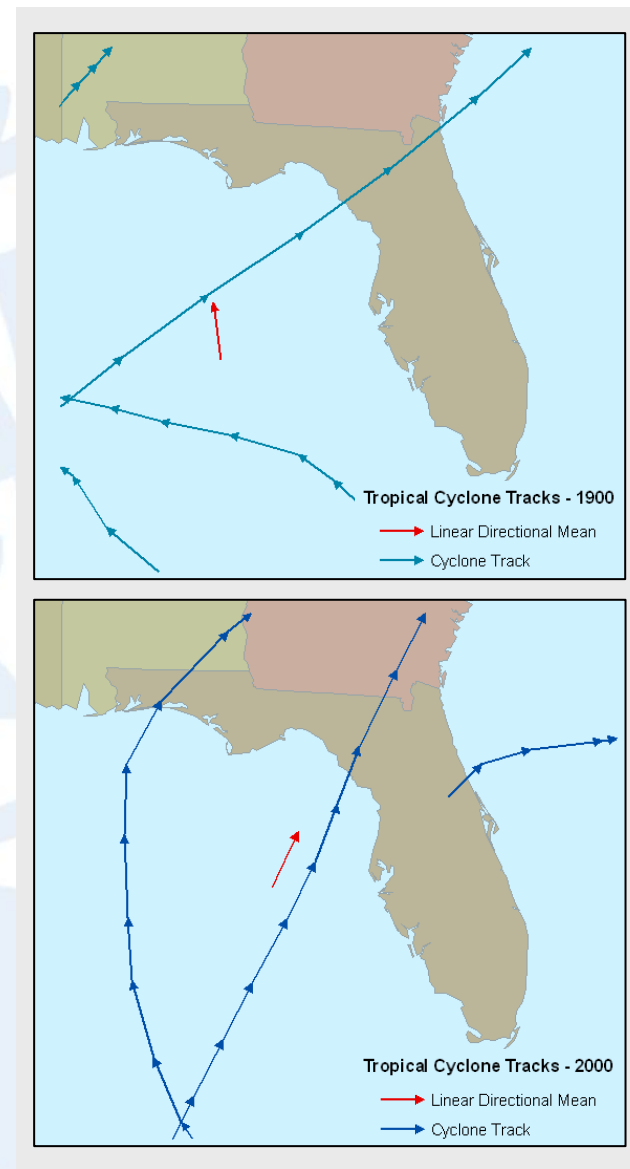
Directional Distribution (Standard Deviational Ellipse) 方向性分布（标准差椭圆）工具

- 提炼要素离散的空间趋势
- 用于：
 - 对比不同类别要素的分布
 - 对比不同时相的同类要素的分布变化
 - 显示出要素离散的方向性趋势



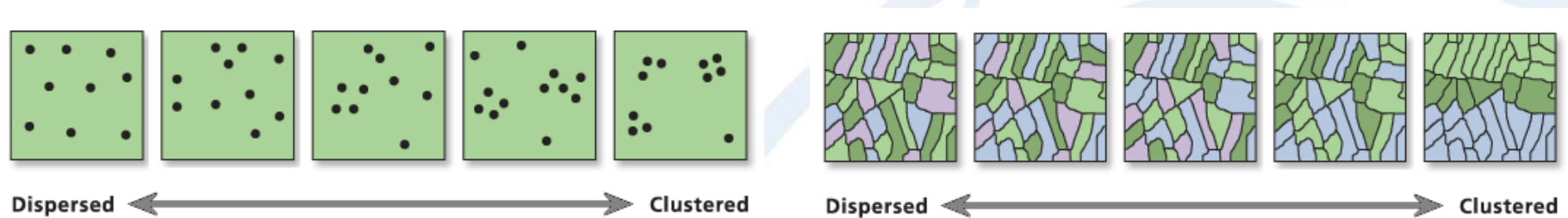
Linear Directional Mean 线性方向均值工具

- 提炼出所有被分析的线要素的平均角度/方向
- 计算从起点到终点的角度
- 用于：
 - 对比多组线要素
 - 对比不同时相的线要素
 - 计算方向性趋势，更直观地显示其空间关系



Analyzing Patterns 分析模式工具集

- 确定研究区域内的要素聚类、离散或随机分布的程度

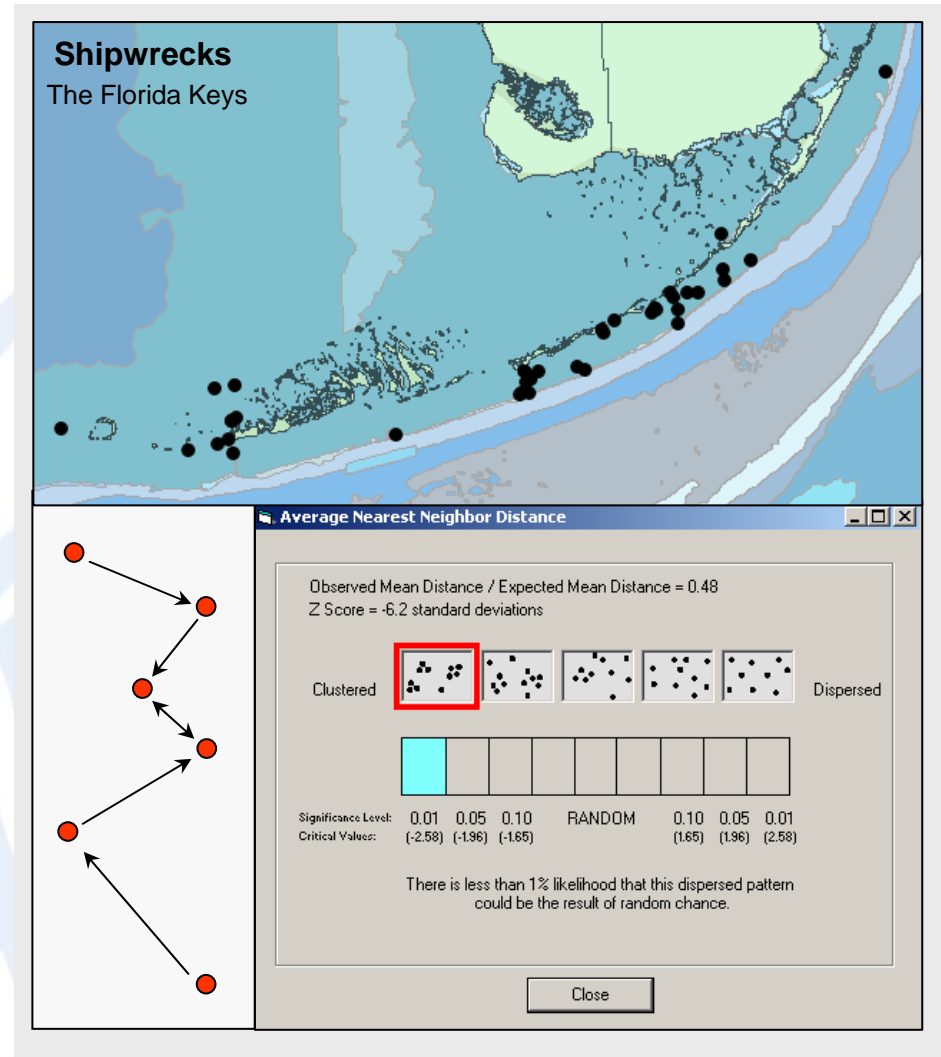


- 全局计算定量分析要素的地理模式
- 用于：
 - 更好地了解要素的分布
 - 研究区域内要素的聚类或离散程度
 - 对比不同要素的分布
 - 对比不同时相的要素分布的变化

Average Nearest Neighbor 平均最近邻工具

- 计算每一个要素到其最近邻要素的平均距离
- 基于面积
- 分析结果：

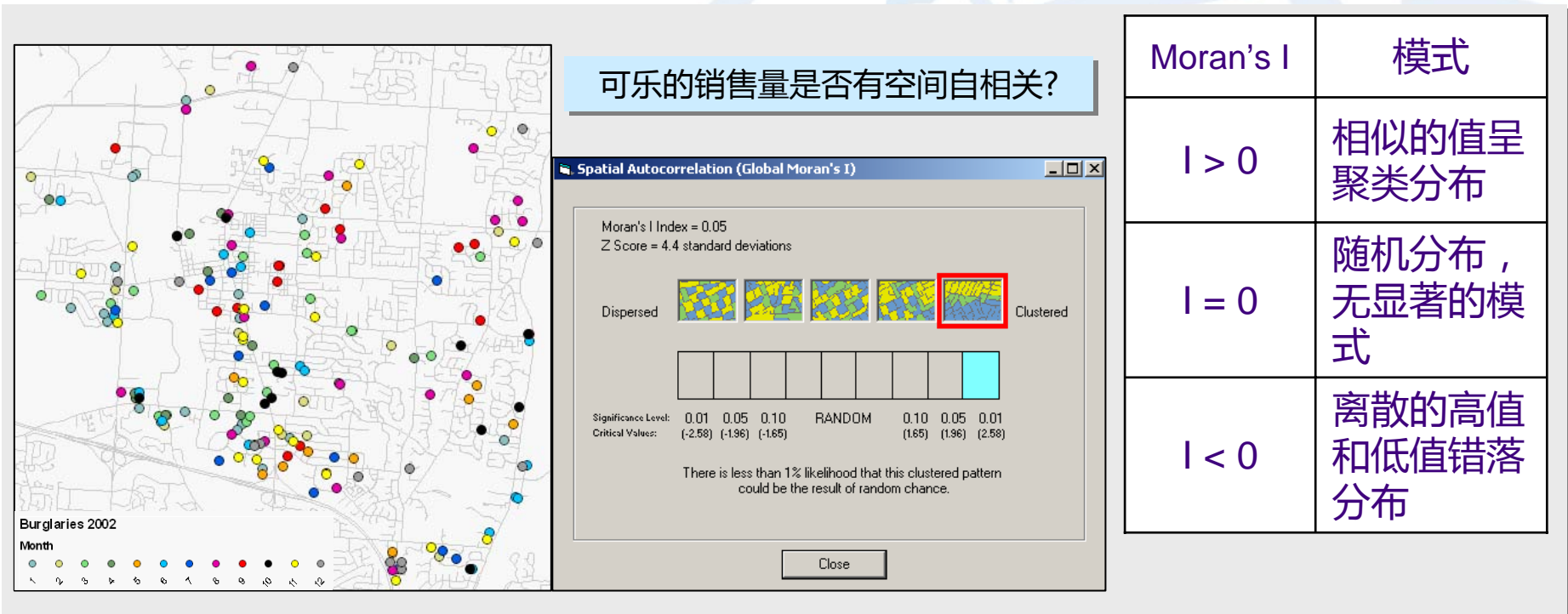
最近邻比值 (观测值/期望值)	分布模式
Ratio = 1	随机
Ratio < 1	聚类
Ratio > 1	离散



Spatial Autocorrelation (Moran's I)

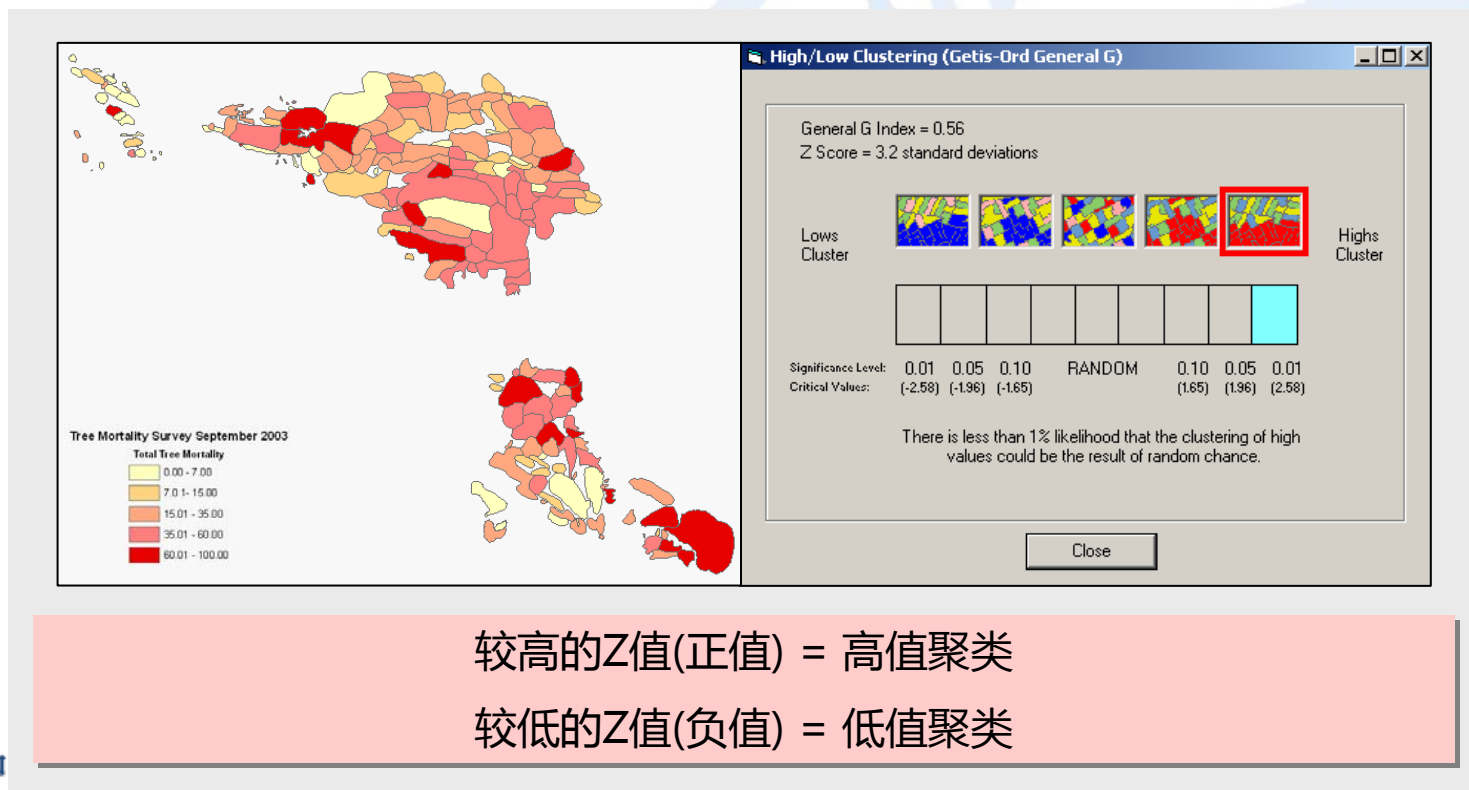
空间自相关工具

- 衡量要素的相似性
 - 距离越近的事物越容易具有相似的特征
- 既考虑空间位置，又考虑属性值



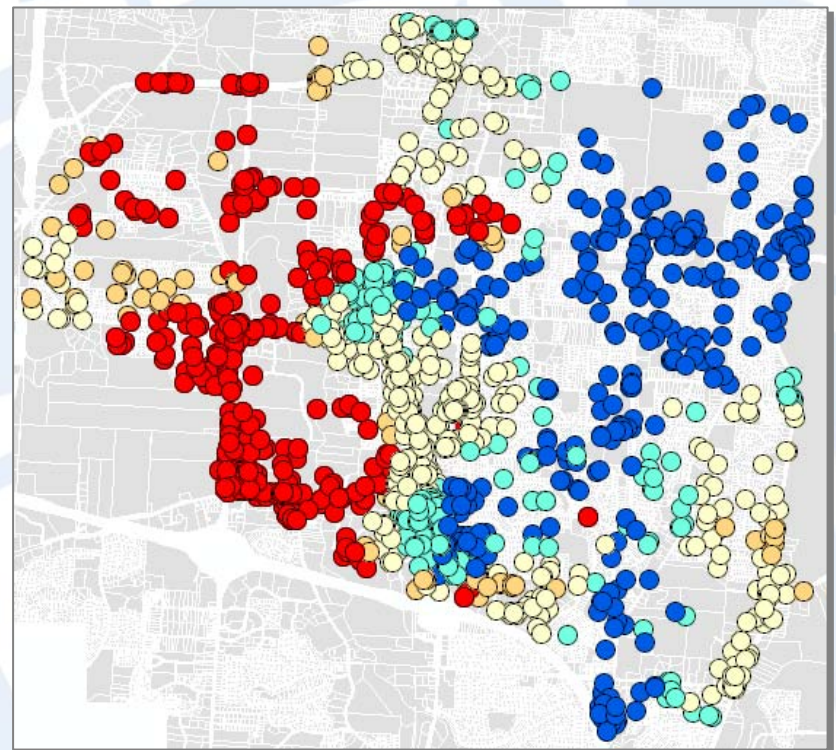
High/Low Clustering (Getis-Ord General G) 高低聚类工具

- 衡量研究区域内高值和低值的聚集程度
- 基于用户指定的邻域距离来计算



Mapping Clusters 绘制聚类工具集

- 对比研究聚类的要素和区域，有助于更好地理解为什么会出现聚类
- 展示研究区域内的聚类或离散区域
- 局部计算
 - 计算每一个要素
 - 展示出研究区域内的差异
- 用于：
 - 定位目标市场
 - 预防犯罪

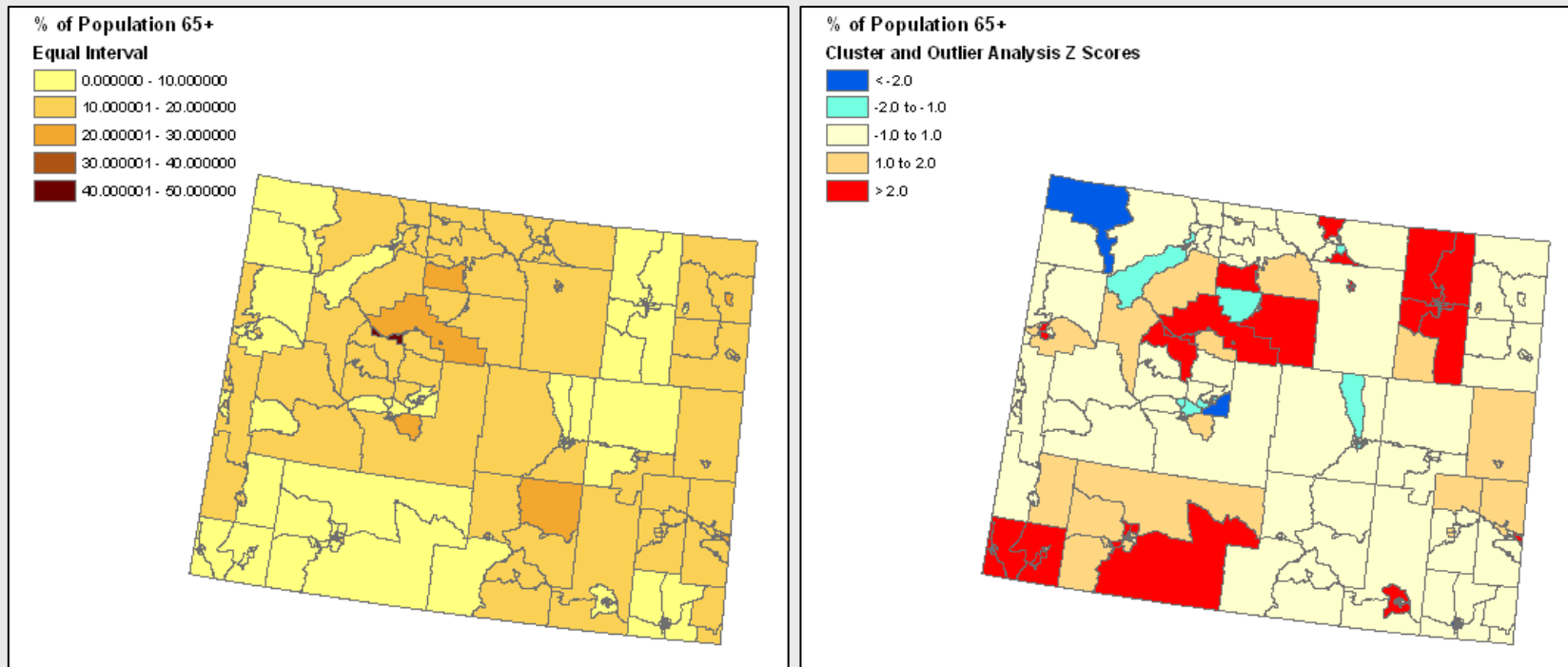


Cluster and Outlier Analysis (Anselin Local Moran's I)

聚类与离群值分析工具

- 强调研究区域内要素值的特殊性
- 帮助发现数据中的离群值
- 计算结果为Moran's I索引值和Z值

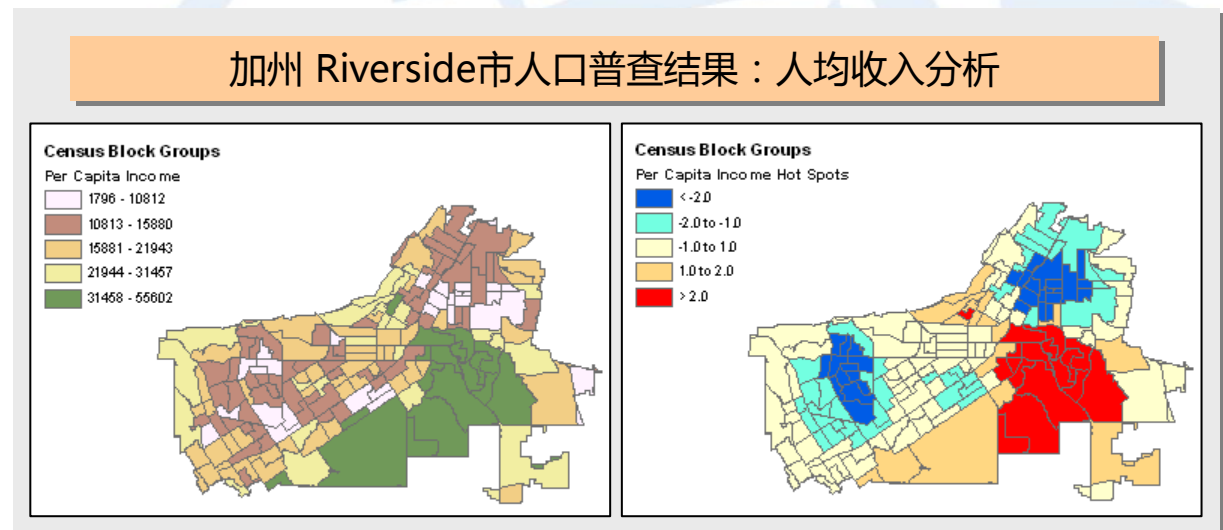
怀俄明州2000年人口普查结果：人口老龄化程度分析



Hot Spot Analysis (Getis-Ord G_i^*)

热点分析

- 指示每个要素被相似的高值或低值包围的程度
 - 确定研究区域中哪儿存在高值或低值的聚类
- 对每一个要素计算Z值
- 用于：
 - 预防犯罪
 - 定位目标市场
 - 查找流行病的传染源



总结

- 空间统计：
 - 制图和传统数理统计的强化和补充
 - 发现数据的空间分布模式、趋势、过程和空间关联
 - 对人文、社科、经济类数据做研究分析
- ArcGIS提供了很多有效的空间统计工具

THANK YOU!

王珏 wangjue@lreis.ac.cn

欢迎关注相关ArcGIS培训课程 <http://training.esrichina-bj.cn/>

ESRI 中国（北京）培训中心
ESRI China(Beijing) Learning Center