**Homework-08**

问题：

1. 对于Doubs数据集，采样点坐标为千米，但没有参考系信息，只有通过QGIS确定各点地理坐标。请简述QGIS操作过程并给出对应点的地理坐标。

答：先将将Doubs数据集中的采样点坐标整理为CSV文件，打开QGIS，点击"Layer"，然后点击"Add Layer"，然后点击"Add Delimited Text Layer"，选择准备好的CSV文件，设置 X field 为 x ， Y field 为 y ，并指定坐标系为 EPSG:4326 ，导入数据后，QGIS会将采样点显示在地图画布上。右键点击图层名称，选择“打开属性表”，查看每个采样点的地理坐标。

校正坐标：点击"Layer"的“Georeferencer”工具，加载采样点数据。在地图画布上选择易于选择的点（如河流拐弯处等），作为地面控制点（GCP），并输入其准确坐标。根据输入的GCP点，选择合适的变换算法（如仿射变换），进行地理配准。

具体坐标点：因我的电脑QGIS无法正常运行，无法显示

2. 为确定各地点鱼群数量是否存在空间自相关，需要确定空间邻近，依据空间邻近及权重矩阵，得到空间滞后spatial lag。下图展示蓝色位置的Rook和queen相邻（黄色），请分别指出两种空间邻近的空间权重矩阵。



答：图一为Rook相邻，空间权重矩阵为：0 1 0 1 0 1 0 1 0

图二为Queen相邻，空间权重矩阵为：1 1 1 1 0 1 1 1 1

3. 在空间数据探索性分析中，常见Moran‘s scatter plot和Lagged mean plot两种图，依据图判断是否存在空间自相关。请回答如下问题：

1）根据Moran’s scatter plot图，如何分辨空间自相关？

* 两图中都有一直线，此直线分别表达的意思是什么？

答：Moran’s scatter plot是一种用于可视化空间自相关的工具，其横轴表示变量的观测值（如某个区域的疾病率、收入水平等），纵轴表示该观测值的空间滞后值（即其邻近区域的加权平均值）。正空间自相关：如果图中的点呈现出明显的正线性趋势，即观测值与其邻近区域的值呈正相关，说明存在正空间自相关。此时，高值区域倾向于与高值区域相邻，低值区域与低值区域相邻。负空间自相关：如果图中的点呈现出明显的负线性趋势，即观测值与其邻近区域的值呈负相关，说明存在负空间自相关。此时，高值区域倾向于与低值区域相邻。无空间自相关（随机分布）：如果图中的点分布较为随机，没有明显的线性趋势，说明空间自相关不显著。此外，Moran’s scatter plot还可以通过四个象限来进一步分析空间自相关：高-高（High-High）：位于右上象限，表示高值区域被高值区域包围。低-低（Low-Low）：位于左下象限，表示低值区域被低值区域包围。高-低（High-Low）：位于左上象限，表示高值区域被低值区域包围。低-高（Low-High）：位于右下象限，表示低值区域被高值区域包围。

Moran’s scatter plot中的直线：这条直线是回归线，其斜率对应于Moran’s I统计量。斜率的正负和大小反映了空间自相关的方向和强度。正斜率表示正空间自相关，负斜率表示负空间自相关[。斜率越接近1，表示空间自相关越强。Lagged mean plot中的直线：这条直线通常是 x = y 的参考线。如果数据点围绕这条直线呈现出明显的线性趋势，说明观测值与其邻近区域的平均值之间存在较强的空间自相关。如果数据点随机分布，且与 x = y 线没有明显关联，则表明空间自相关不显著

4. 探索性分析表明，doubs河流中鱼群多度（spe表格中采样点鱼个体数量）存在空间自相关，请简述如何将空间自相关纳入鱼群多度-环境要素关系模型中。

答：1. 使用空间回归模型：在模型中加入空间滞后项或空间误差项，以考虑空间自相关的影响。这些模型可以捕捉到邻近观测值之间的相关性，从而更准确地估计环境要素对物种多样性的影响[。2. 应用地理加权回归（GWR）：这是一种局部回归方法，它允许模型参数随地理位置变化，从而可以更细致地分析环境要素与物种多样性之间的关系，并考虑空间自相关的影响。3. 利用空间自相关和环境相似性：结合空间自相关和环境相似性，通过一个调节参数来确定两者的结合方式，从而更切实际、更准确地预测目标变量的空间分布。4. 采用空间插值方法：如克里金插值，可以用于估计未观测位置的环境要素值，这些方法可以考虑到空间自相关性，从而为模型提供更准确的环境要素数据。5. 使用空间自相关的统计检验：如Moran's I指数，来量化空间自相关的程度，并在模型中作为控制变量，以减少空间自相关对模型估计的偏差。

要求：

一周内上传至各自GitHub托管的homework中。