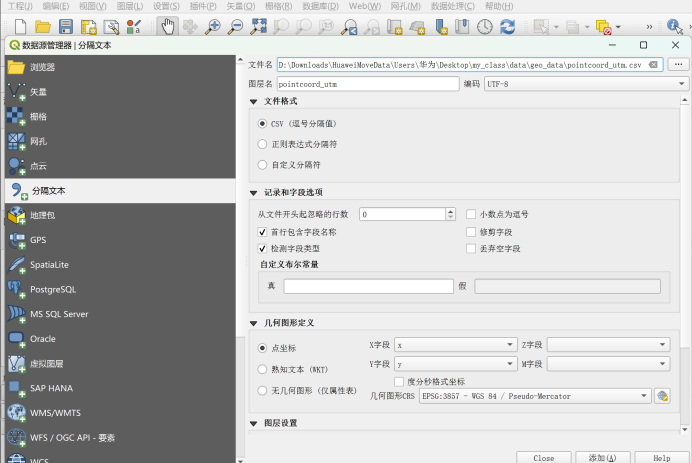
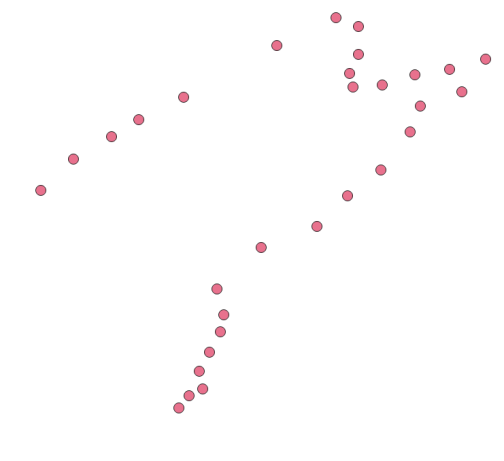
**Homework-08**

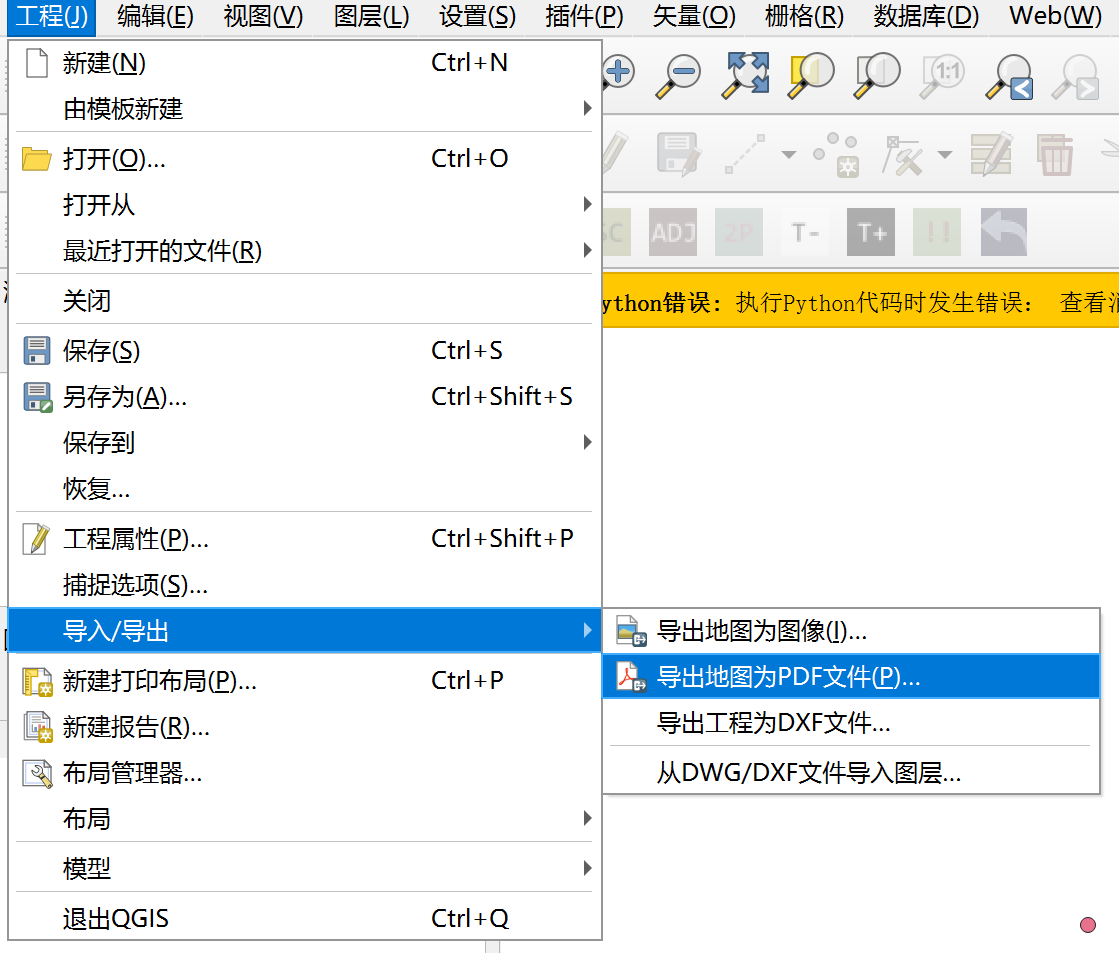
问题：

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。

（1）GGIS中图层——添加图层——添加分隔文本图层实现pointcoord\_utm.csv采样点数据的导入，导入后如右图所示

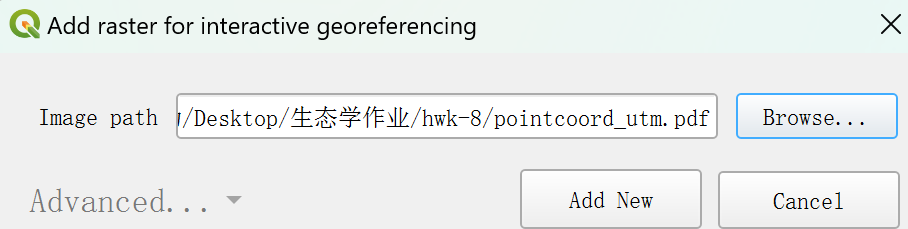


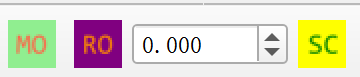
1. 选择工程—导出—导出为pointcoord\_utm.pdf用于后续配准

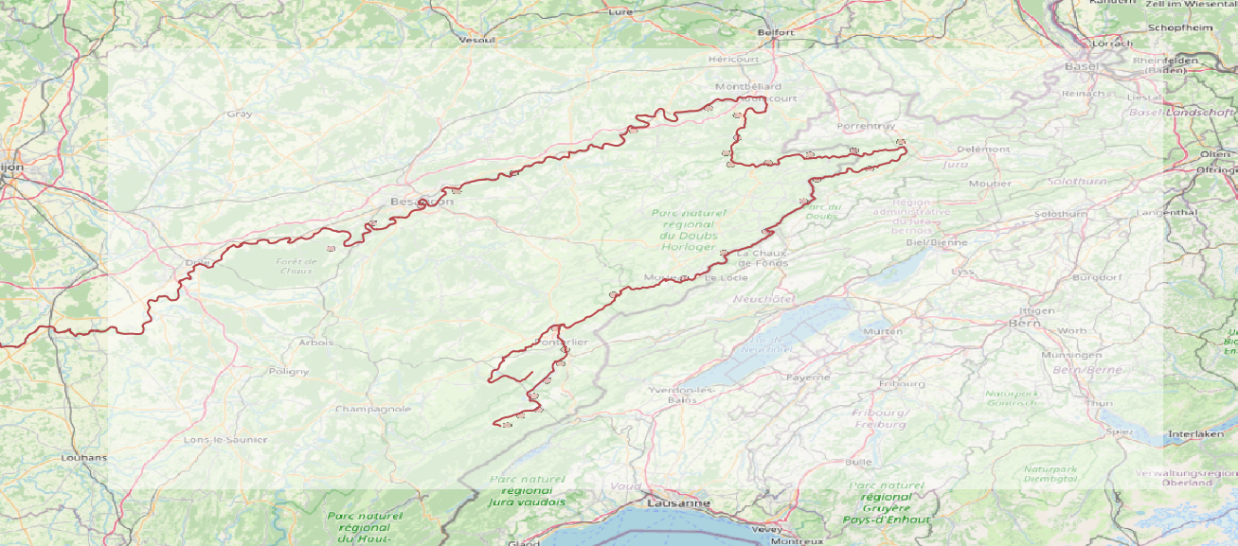


1. 将hwk-7的doubs\_river.geojson导入：图层 → 添加图层 → 添加矢量图层，并添加一个地图底图：Web → QuickMapServices → OSM → OSM Standard。

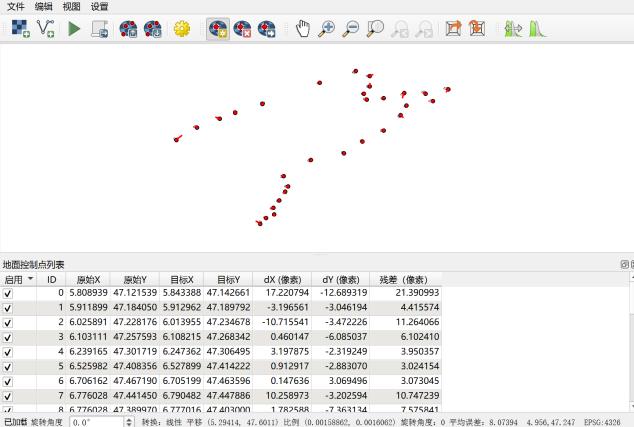
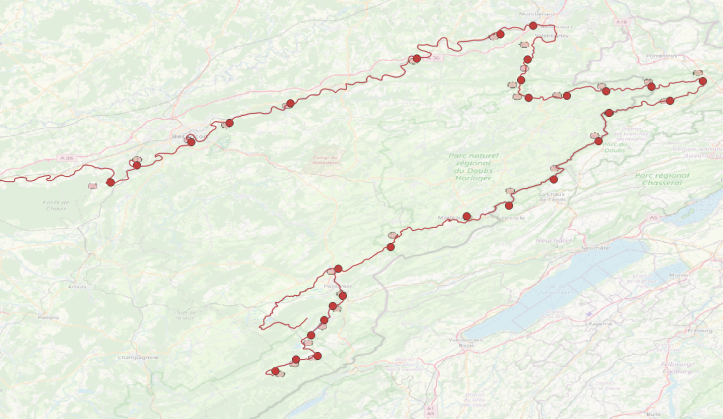


1. 安装插件 georeferencer，将pointcoord\_utm.pdf文件导入，导入后利用Freehand Raster Georeferencer的移动（MO）和缩放（SC）功能，将搜索到的河流数据与导入的数据点进行重叠。

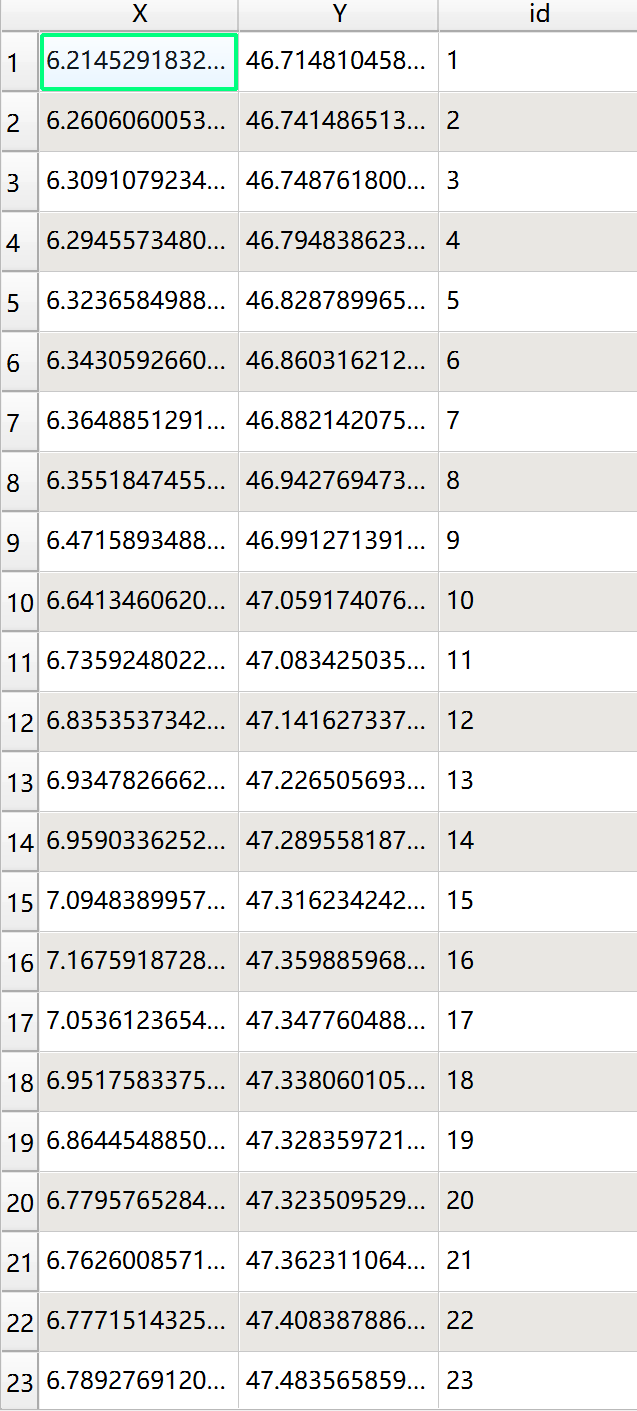
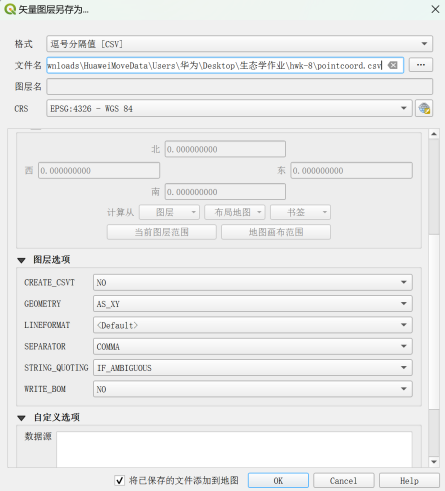
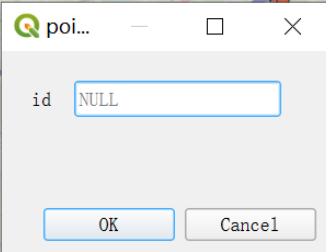




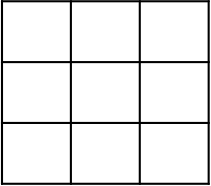
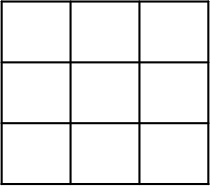
1. 使用配准工具，根据doubs\_river河流数据，对前面导出的pointcoord\_utm.pdf文件进行配准：点击图层 → 配准工具 → 打开栅格添加配准文件 → 对PDF中的每一个点进行配准，尽量使得点落在河流线段上。点击开始校准，获得pointcoord\_utm已修改.tif文件，在图层中可见校准好的数据点。



（6）每个采样点添加id，保存地理坐标为csv，注意geometry选在AS XY，输出结果如图所示



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。



左图Rook邻近：仅考虑共边单元的相邻关系，即上下左右邻接

右图Queen邻近：考虑共边和共顶点的相邻关系，上下左右对角线邻接

空间权重矩阵：

左图（rook邻近）

中间蓝色点的空间权重向量为 0 1 0 1 0 1 0 1 0

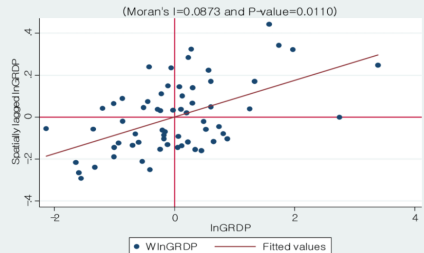
右图（queen邻近）

中间蓝色点的空间权重向量为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？

Moran’s scatter plot 是通过**将每个观察值与其空间邻域的平均值进行比较来判断空间自相关的**。图中的四个象限分别表示不同的空间自相关模式：

第一象限（高-高）：高值邻近高值区域，表示正空间自相关。

第二象限（低-高）：低值邻近高值区域，表示负空间自相关。

第三象限（低-低）：低值邻近低值区域，表示正空间自相关。

第四象限（高-低）：高值邻近低值区域，表示负空间自相关。

因此如果大多数点集中在第一和第三象限，说明存在正空间自相关

1. 两图中都有一直线，此直线分别表达的意思是什么？

在Moran’s scatter plot中，是一条**线性拟合回归线**，其斜率等于 Moran’s I 指数。

斜率 > 0：表示正空间自相关（高-高、低-低聚集）。

斜率 < 0：表示负空间自相关（高-低或低-高分布）。

斜率 = 0：表示无空间自相关。

这条直线显示了**整体变量与其空间滞后值之间的线性关系强度**。

在Lagged mean plot中，是观察值与其空间滞后值之间的**回归趋势线，**是一条45°的固定对角线，即x=y线。这条对角线作为参考线，代表某一单元周边所有相邻单元的平均值，测量某一单元值与其相邻单元平均值的偏离程度，即用于初步判断该单元值与周围单元值的相似程度。**当点刚好落在参考线上或集中于参考线附近时**（即存在一定的线性趋势），偏离程度小，**局部空间自相关程度高**，若偏离参考线，则局部自相关程度低。

与 Moran’s scatter plot 不同，该图强调了**变量与其空间平均值之间的实际趋势关系。**

|  |  |  |
| --- | --- | --- |
|  | 轴 | 含义 |
| Moran’s scatter plot | X轴 | 原始值 |
| Y轴 | 邻域的加权平均值(即空间滞后值) |
| Lagged Mean Plot | X轴 | 原始观测值 |
| Y轴 | 当前观测值前面的若干滞后值的平均值 |

1. 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。

一．直接引入坐标（X, Y）作为模型特征

将每个采样点的投影坐标（x, y）作为普通的数值型特征纳入模型

二．用已知点和预测点间的距离即buffer distance作为特征，提高模型精度：将鱼类丰度点数据分层后，计算各层上点到各格网的缓冲区距离（即已知点-预测点距离，buffer\_distance），然后将各层buffer\_distance数据作为新的features引入机器学习建模，评估并筛选较为重要的缓冲区距离变量。

要求：

一周内上传至各自GitHub托管的homework中。