大数据分析第一次作业：

SA19225404 吴语港

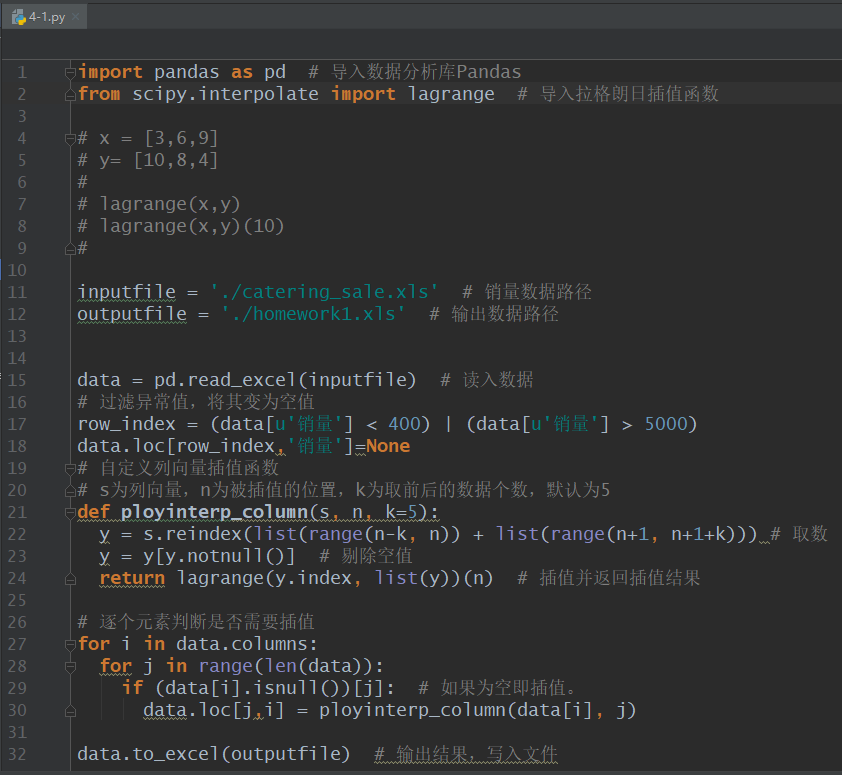
任务：

优化拉格朗日插值函数，避免出现不合理的值

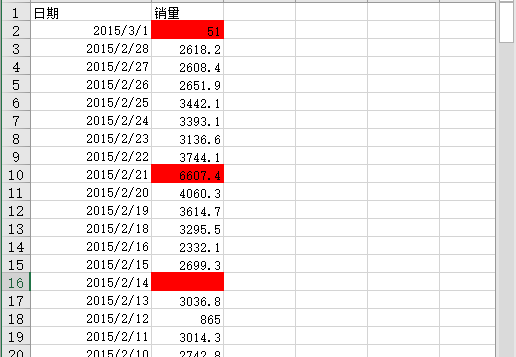
要求提交一份以“学号+姓名”命名的word文档，里面包含解题思路、代码、运行结果（只需要插补的6个值）

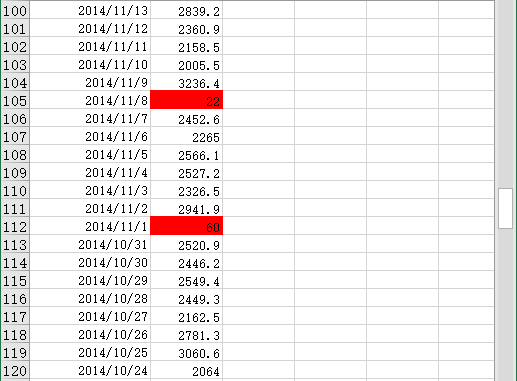
1. 研究问题：

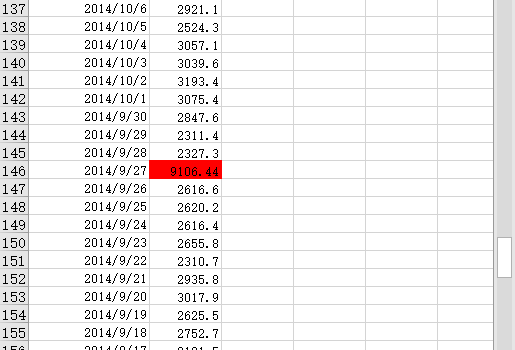
原版程序



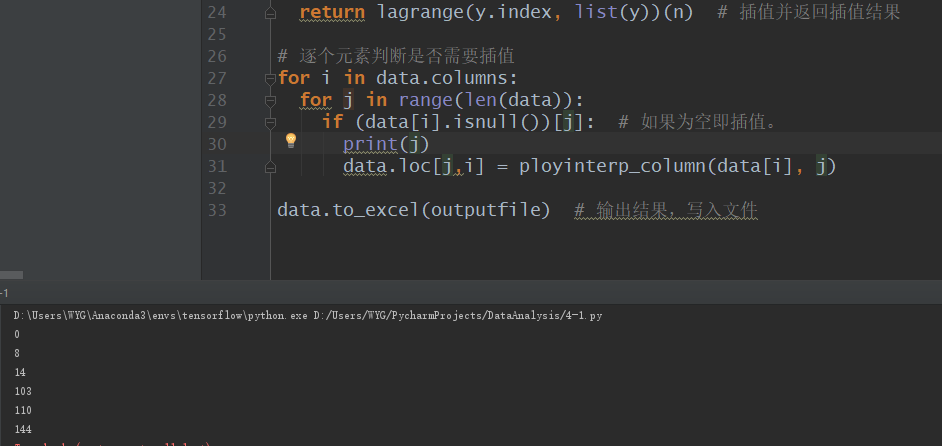
插值前出现问题的值，包括缺失和异常的值



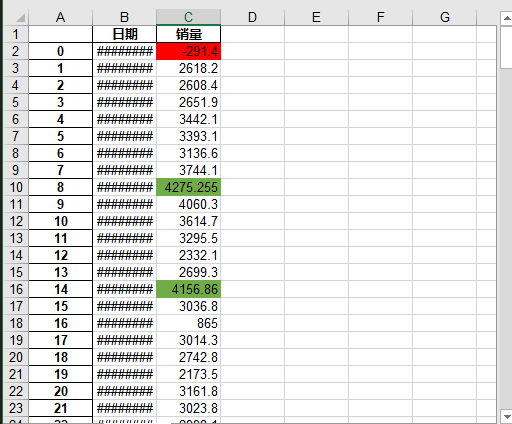


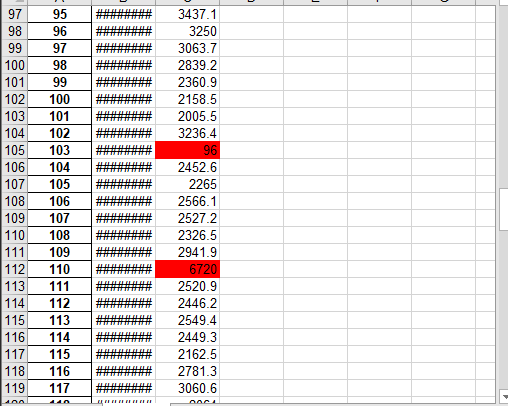


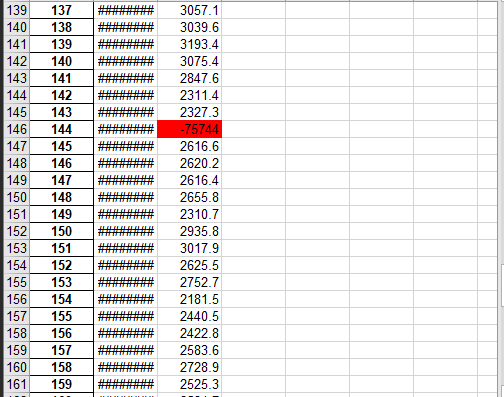
出现问题的值的索引为 0、8、14、103、110、144.



我们来看插补后的值





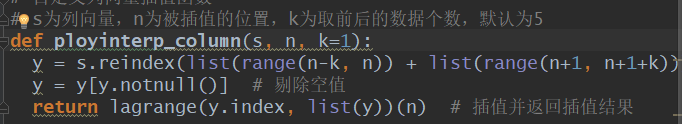


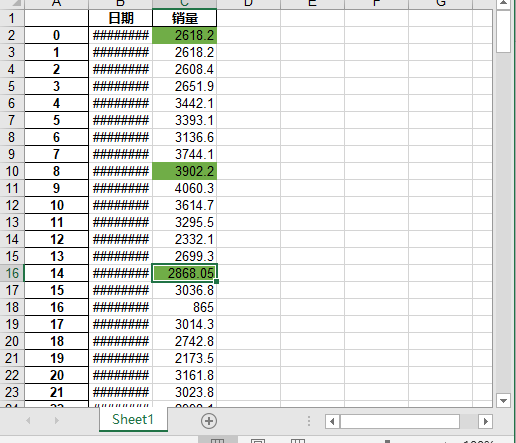
绿色的为插值后正常的值，红色的为插值后异常的值。有4个值出现了异常，两个正常了，那么为什么会出现插值后还是异常呢？

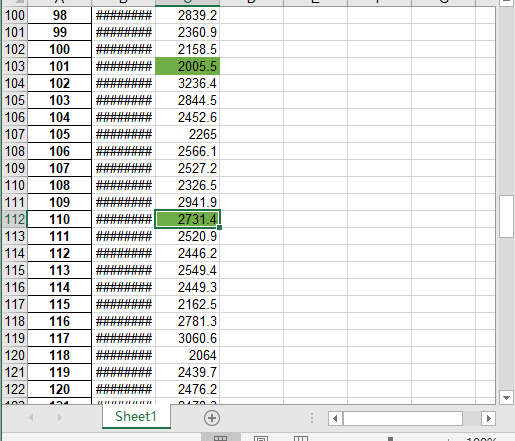
1. 解决问题：

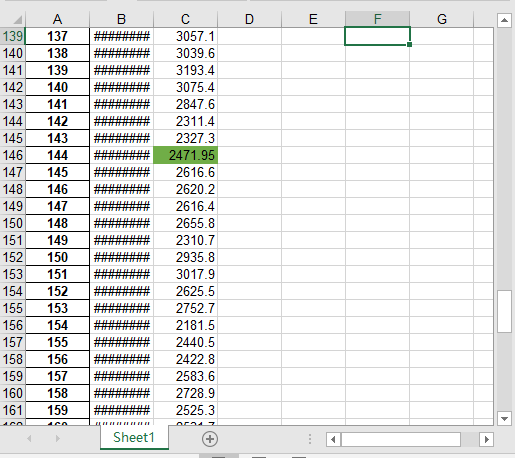
方案一：

先将插值函数的k改为1，然后运行程序





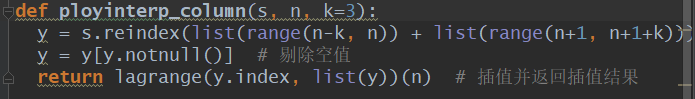


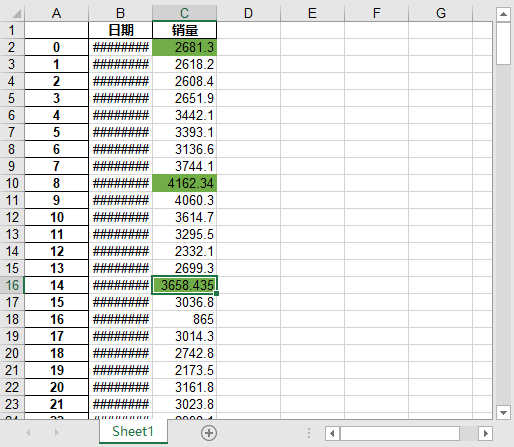


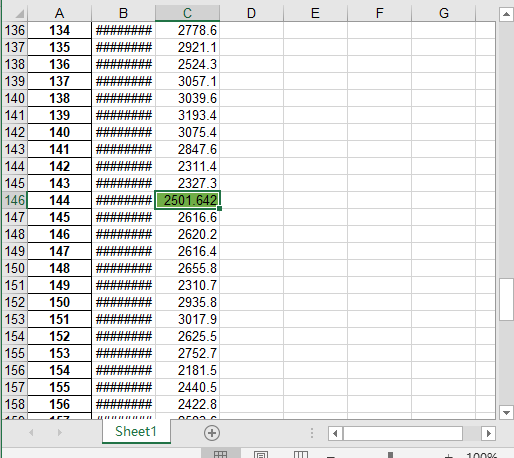
这个时候插值函数的输出结果全部通过了范围，可能是k太大了会造成拉格朗日函数次数过高，从而有急剧振荡的现象，不过老师给的默认k = 5， 可能是更高的次数对非线性的函数拟合效果好一些，那么我们再看一看当k = 3时，函数的输出结果。

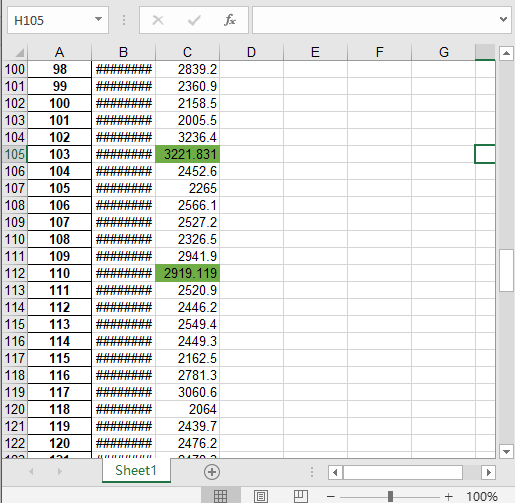
方案二：

K = 3 时的输出结果



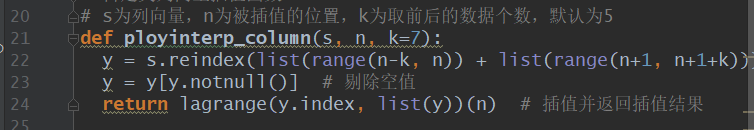


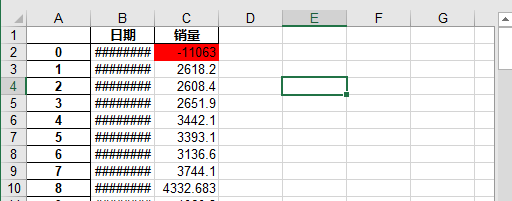




同样可以进行正常的插值，没有结果超出范围

再看一看k调整为7会怎么样：





还是出现了同样的问题

看来k不能太大，不然拉格朗日函数会非常曲折，会出现异常值