数据预处理实验 (一)

实验一：缺失数据填充

餐饮系统中的销售数据可能会出现缺失，catering\_sale\_missing.xls为某餐厅一段时间的销量，其中，有部分数据缺失。利用拉格朗日插值对缺失数据进行插补。要求：

* 前后取值个数k=3
* 读出并显示原始数据
* 插值后的数据另存为missing\_data\_processing.xls
* 读出并显示插值后数据

实验二：相关性分析

数据集catering\_sale\_all.xls是某餐饮企业的菜品日销售统计，利用Python中统计函数，分析这些菜品销售之间的相关性。要求：

* 用拉格朗日插值法，填充cantering\_sale\_all.xls中的缺失数据（保留一位小数）
* 计算任意两款菜式之间的相关系数
* 计算“百合酱蒸凤爪”与其他菜式的相关系数
* 计算“百合酱蒸凤爪”与“翡翠蒸香茜饺”的相关系数

实验三：数据规范化

使用python 相关库完成给定数据集的据规范化和转换。要求：

* 读取normalization\_data.xls
* 最小-最大规范化数据
* Z-score规范化(零-均值规范化)
* 小数定标规范化
* 输出所有规范化结果

实验四：数据离散化

* 读取discretization\_data.xls，使用padans.cut()对数据进行0-3的离散处理；
* 处理后的数据保存在data\_descret.xls中

实验五：数据转化

* 读取sales\_data.xls，将字符数据转换成数值型数据，并将转换结果保存为data\_tras.xls（data\_transfer）