练习三（11月22日）

第一题：在文件Demo\_4.txt 中存储了一组序列，推荐使用numpy完成

1. 读取文件，判断该序列共有多少个不重复的数字；
2. 计算该序列的最大值、最小值、方差、中位数和众数；
3. 计算该序列的偏度和峰度，判断其分布是否服从于正态分布；
4. 假设该数据存在某种周期性的变化规则，请你找到其变化周期。

第二题：在文件Demo\_3.csv 中存储了沈阳市某连锁书店2024年的销售量，现请您分析该连锁书店的销售情况，推荐使用pandas完成：

1. 使用pandas库读取Demo\_3.csv；
2. 已知文件中数据存在部分缺失（存在缺失值），请您找到缺失数据，并对其进行适当的处理，输出存在缺失的数据行数；
3. 已知文件中部分数据不准确（存在异常值），请您找到异常数据，并对其进行修正，输出异常的数据行数；
4. 该连锁书店共有多少种商品，销量最好的商品是什么；
5. 已知，每种商品的原价相同，但出售时的折扣总是变化的，请你计算每个品类在每个月的平均折扣；
6. 请您提取每个商品的价格表（包括原价、平均折扣、平均卖价）并对原价和平均卖价数据进行标准化，标准化公式：



1. 计算每个商品原价和平均卖家的Pearson相关系数，公式：



1. 请你预测2025年1月各门店的销售额。