

1. 加载数据集

`seeds.csv`

2. 查看数据集，分别提出品种列和特征数据，并将特征数据转化为 `numpy` 数组
3. 对数据多次运行非负矩阵分解 (`rank=3`, `repeat_times=20`)
4. 计算每次聚类结果的轮廓系数，并绘制折线图，标注最佳分数
5. 将最佳分数对应的聚类结果转化为 `Dataframe`，做品种和最佳聚类结果的交叉表
`pd.crosstab(df1, df2)`