

完成以下两个任务，提交源代码和实验报告。请在实验报告中记录你的实验过程、实验结果和思考。

任务1 聚类（对应文件house-votes-84.csv）

任务描述：数据中给定议员对于16项不同议题的投票情况（属性表示为支持Y、反对N、弃权？），数据的第一列为议员所属的党派类别，请使用聚类方法对上述数据进行聚类分析：

1. 尝试定量的评价的聚类结果；
2. 尝试比较不同聚类方法或者不同初始情况对聚类结果的影响；

数据来源于：<https://archive.ics.uci.edu/ml/datasets/Congressional+Voting+Records>

任务2 关联分析（对应文件 car.data.csv）

任务描述：给定二手车评估情况的数据，第一行为各类属性名称，最后一列为二手车的评估情况分类，以比例分为如下4类：

等级类别	样例数	占比（%）
unacc	1210	70.023
Acc	384	22.222
Good	69	3.993
V-good	65	3.762

各属性的可能取值如下（并不全是数值）：

buying: v-high, high, med, low

maint: v-high, high, med, low

doors: 2, 3, 4, 5more

persons: 2, 4, more

lug_boot: small, med, big

safety: low, med, high

1. 请对数据进行关联分析，计算哪些属性的哪些值对于每个分类结果有较强的关联（请自定义置信度和支持的阈值）。
2. 尝试用前述聚类方法结果对该数据进行聚类，尝试分析聚类结果与关联分析的联系。

数据来源于：<https://archive.ics.uci.edu/ml/datasets/Car+Evaluation>

任务3 异常检测

任务描述：针对上述两个任务中的某一个任务的数据，尝试使用某种异常检测方法分析数据的情况，判断数据集中是否存在异常数据点。