1. 将数据中的n，y替换成0，1。？用每一列的均值替换。区分不读不大，为了提高精确性，不降维。
2. 数据第一列为党派，与后续数据有所不同，所以不处理
3. 添加第17列为分类列，先初始化成-1，然后后续分类了就是类的序号
4. 使用K-means方法对除党派信息以外的1-16列16个问题进行聚类。常识告诉我们对这16个问题的回答可能跟所属党派有关，而党派总共有两个，所以聚类也聚两类。
5. 结果分析，看看聚类和党派的关系，可以发现有超过八成与党派相符！
6. 结果中剩余不到百分之20预测不符未必是算法不精确的问题，将在第三题进行误差分析