第二次作业及讨论要求  
1、根据所给数据集data，完成用户聚类分析，发现具有相同偏好的**用户簇**；  
2、data数据集包括三个字段：用户编号，商品编号，用户对商品的评分（评分范围：1-5），5是最高分；  
3、设计的聚类算法可以是单一聚类方法的改进，也可以是多种聚类方法的融合（特别注意：用户分组数量未知，也需要在算法中分析确定）；  
4、为了方便大家验证所设计聚类模型的效果，提供了测试数据集test, 但该数据的结果不允许在聚类算法设计与实现阶段使用，即聚类分析时，用户的偏好分组的数量和具体情况是未知的。最后我们会比较各组同学最后的聚类精度。精度度量指标选用精确率P，召回率R和F1-Measure。  
5、实验输出包括文本和可视化两种方式，都要实现，具体形式可自行决定。  
6、本次作业在4月30日课堂讨论报告，报告后提交作业。