Week2： Item-Based CF

数据介绍：

animelist.csv 涵盖了所有用户对某部 Anime 的打分和观看状况，其中：
1. user_id: 随机生成的用户 ID
2. anime_id: 动画 ID，在整个数据集中用作动画的唯一标识符
3. score: 用户打分，范围 1 到 10，0 则是未评价。
4. watching_status: 观看状况，具体含义请查阅 watching_status.csv
5. watched_episodes: 用户观看的集数。

rating_complete.csv 是 animelist.csv 的子集，仅包含完整看完一部动画的所有剧集并且打分的条目。共计 57M 条目，而全集包含 109M 条目。
anime.csv 是每一部 Anime 的基本数据，其中：
1. MAL_ID: 唯一 ID
2. Name: 动画全名
3. Score: 平均分
4. Genres: 使用逗号分隔的类别列表
5. English name: 英文全名
6. Japanese name: 日文全名
7. Type: TV, movie, OVA 等
8. Episodes': 剧集数量
9. Aired: 上映日期
10. Premiered: 首映季节
11. Producers: 使用逗号分隔的制作人列表
12. Licensors: 使用逗号分隔的版权方列表
13. Studios: 使用逗号分隔的工作室列表
14. Source: 来源，例如原创、漫画、轻小说改编等
15. Duration: 每集时长
16. Rating: 年龄分级
17. Ranked: 根据评分的排名
18. Popularity: 热度
19. Members: 讨论组成员数量
20. Favorites: “喜欢”的数量
21. Watching: 正在观看人数
22. Completed: 已经观看完毕的人数
23. On-Hold: 待定的人数
24. Dropped: 弃坑的人数
25. Plan to Watch': 计划观看的人数
26. Score-x': 评分为 x 的人数

任务要求：

使用以上数据构建 item-based CF 模型，要求最后的模型能够做到：输入一个与 animelist.csv 结构相同但是没有 user_id 列的数据列表（代表单个用户的偏好），返回一组有 MAL_ID、Name、Score、Type、Source、synopsis 列的动画列表，列表至少包含 5 个条目。