

Deciphering user fashion sensibility

A proposal for Bestseller



目录

1. 背景
2. 解决方案
3. 算法流程
4. 划分结果
5. 结果分析



1. 背景：BS的产品粒度

➤ Bestseller的产品粒度：

1. 到**大品类**：BS服装产品分为饰品，西装，针织衫、Polo等10余个大品类；
2. 到**单款产品**：BS服装产品可细分到单款产品，以产品ID来标注

图1：BS各品类下产品数量

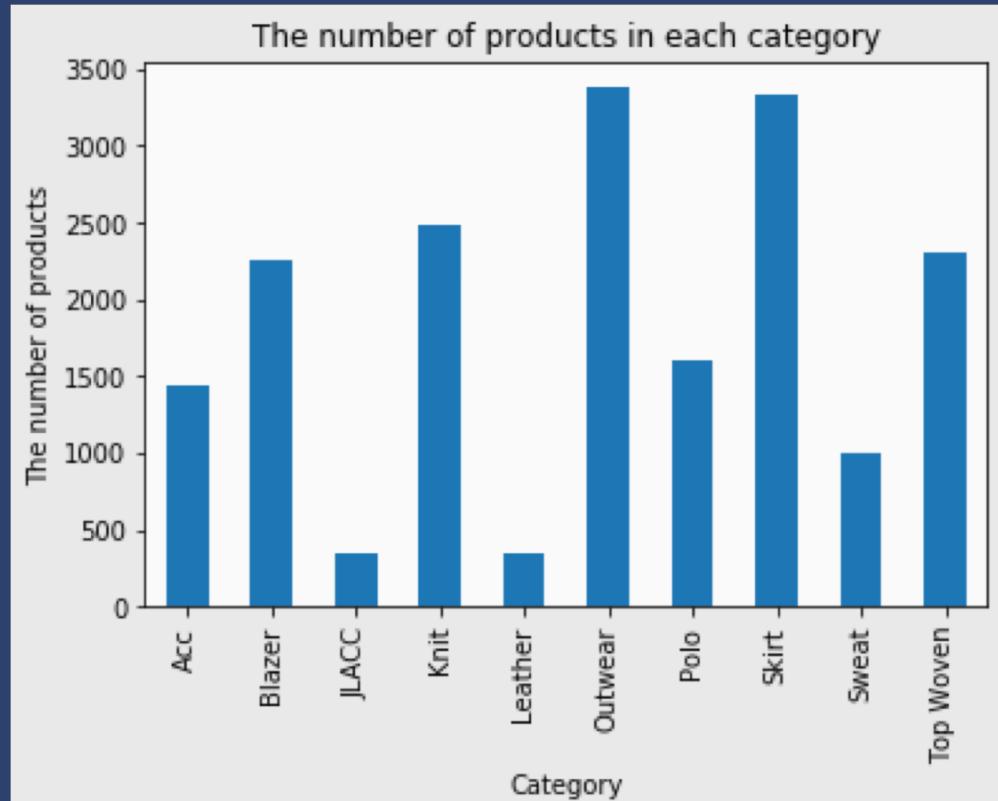


图2：BS单款产品销量

15位产品号	Sales volume
11711H512S01000	[3241]
11711H529A14000	[1778]
11711H526A14000	[843]
11711H506E03000	[813]
41641Q505106000	[575]
11711H510S01000	[573]
11711H507S01000	[568]
417106507C39380	[441]
115385012010000	[436]
11711H502D08000	[351]
11538501207B000	[344]
11528500708F000	[332]
316187005010320	[313]
0520102000181947	[313]
000380100285211	[335]
000BT0510286211	[444]
00080PS06HHTTTR	[321]
0000010510286211	[304]

1. 背景：以品类为基本单位的缺点

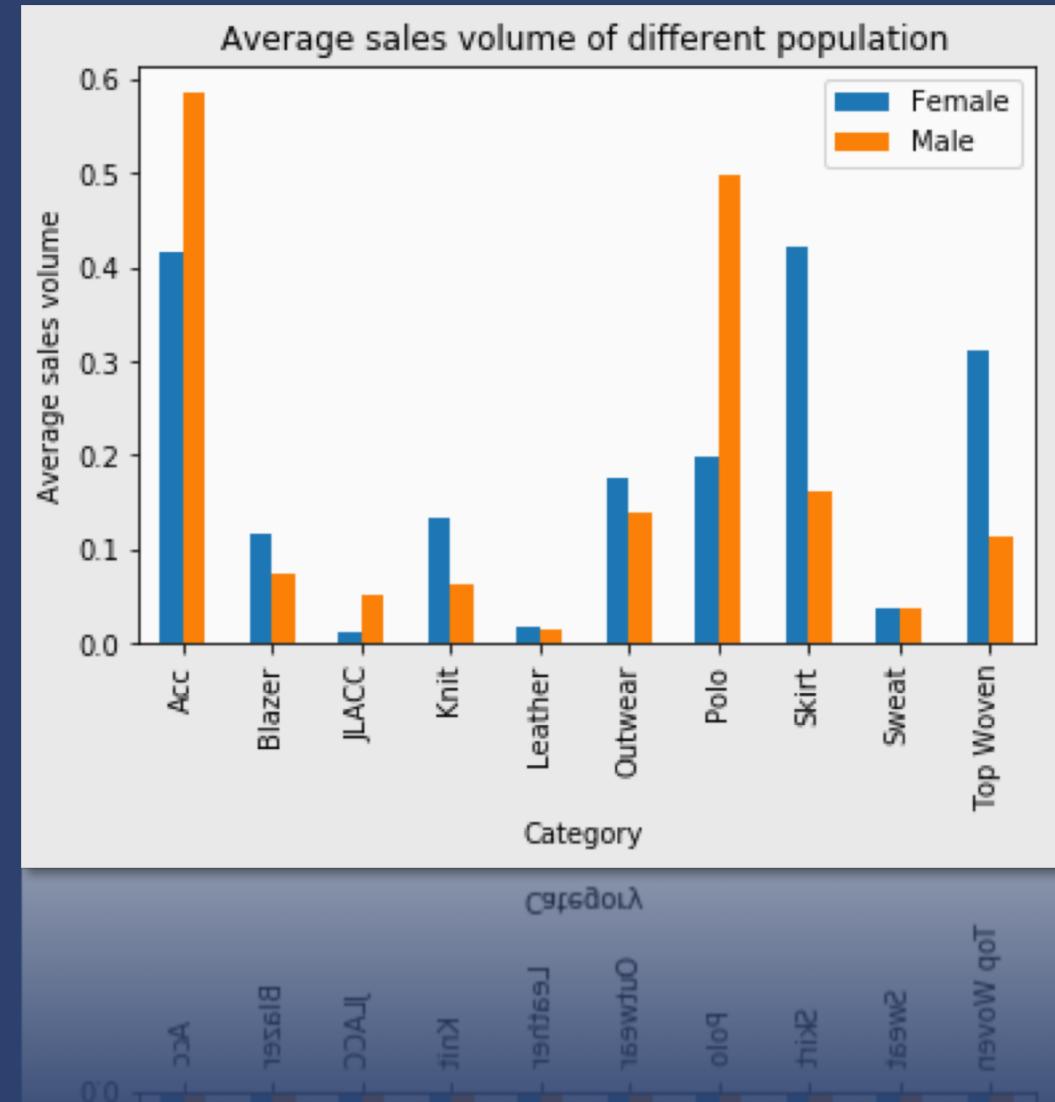
➤ 以品类作为基本单位，可获得：

1. 各品类的销售状况
 2. 各品类的价格分布
 3. 不同客群在各品类上的整体购买情况
-

➤ 缺点：粒度过大

1. 无法深入区别品类下的产品
 - 无法刻画一个品类下服装间的差异性
 - 无法捕捉一个品类下衣服间的相似性
 2. 无法进行更深入的销售分析和用户购买倾向分析
 3. 无法为服装搭配提出建议
-

图1：BS男女客群各品类平均销量



1. 背景：以单款产品为基本单位的缺点

➤ 以产品ID作为基本单位，可获得

1. 单款的销量，可识别爆款
2. 不同客群的单品购买情况

.....

➤ 缺点：粒度过小

1. 丢失了单款产品间的相似性和差异性
2. 基于用户的协同过滤 -> 仅限于推荐系统
3. 虽能识别爆款，但无法找到和爆款相似的产品

.....

图1：BS单款产品销量排名

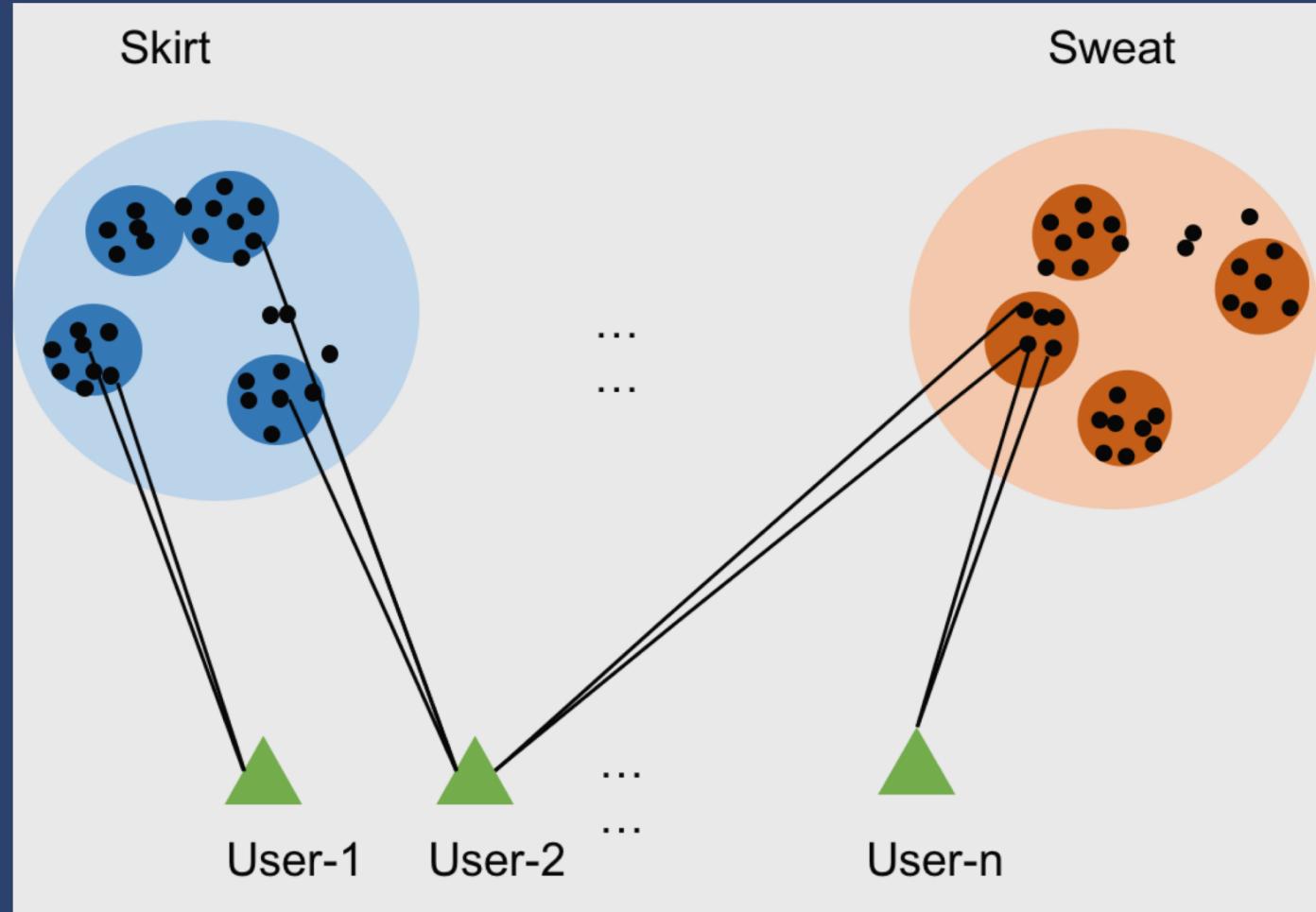
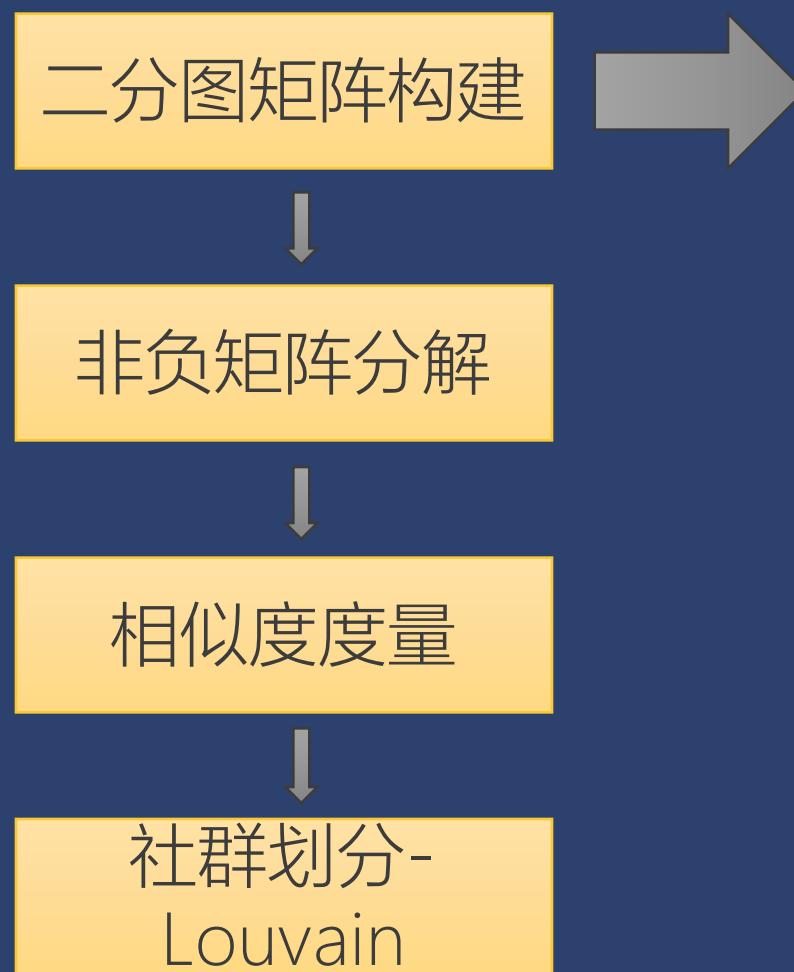
Sales volume	
15位产品号	
.....	
11711H512S01000	[3241]
11711H529A14000	[1778]
11711H526A14000	[843]
11711H506E03000	[813]
41641Q505106000	[575]
11711H510S01000	[573]
11711H507S01000	[568]
417106507C39380	[441]
115385012010000	[436]
11711H502D08000	[351]
11538501207B000	[344]
11238201501B000	[344]

2. 解决方案：对每一个品类的服装进行粒度细化

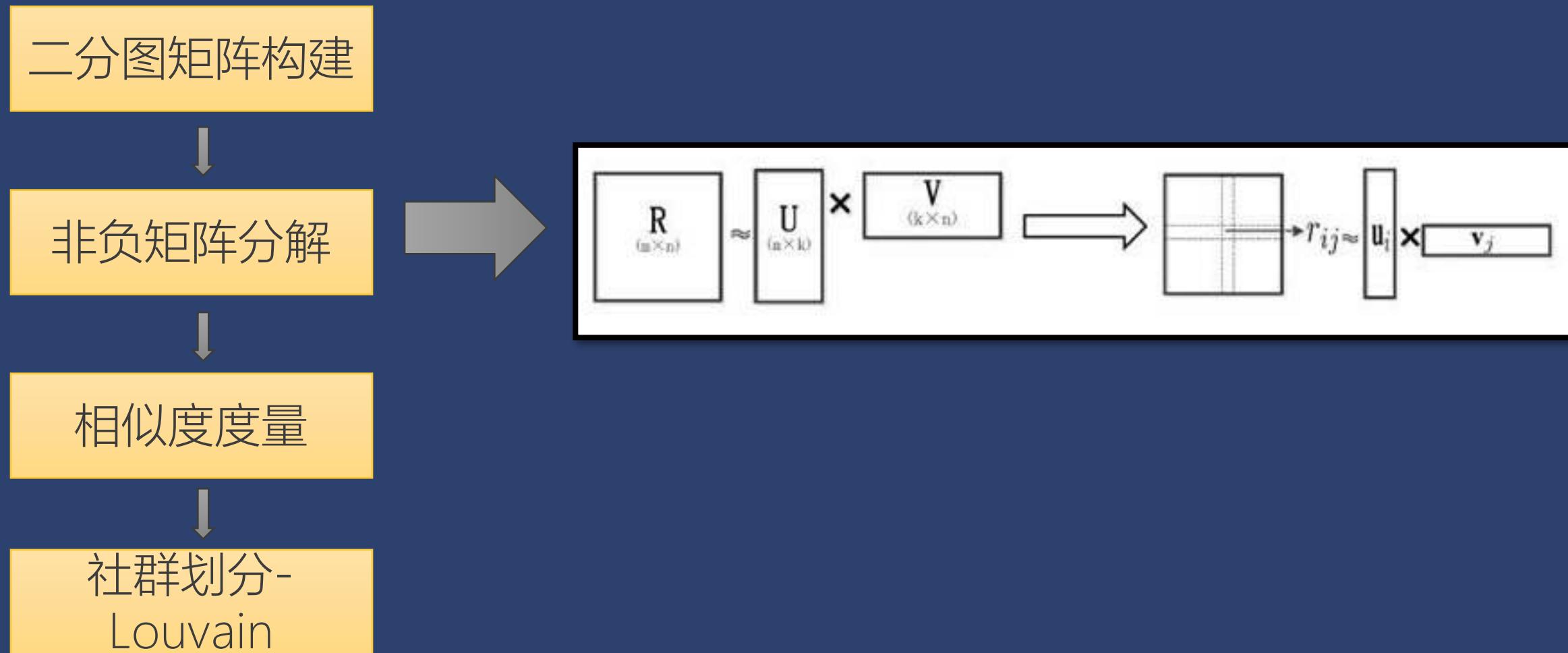
➤ 粒度细化方式：

1. Computer Vision : 模型难以搭建，计算资源大
2. 协同过滤 + 社群划分算法 : 业界成熟，稳定性高

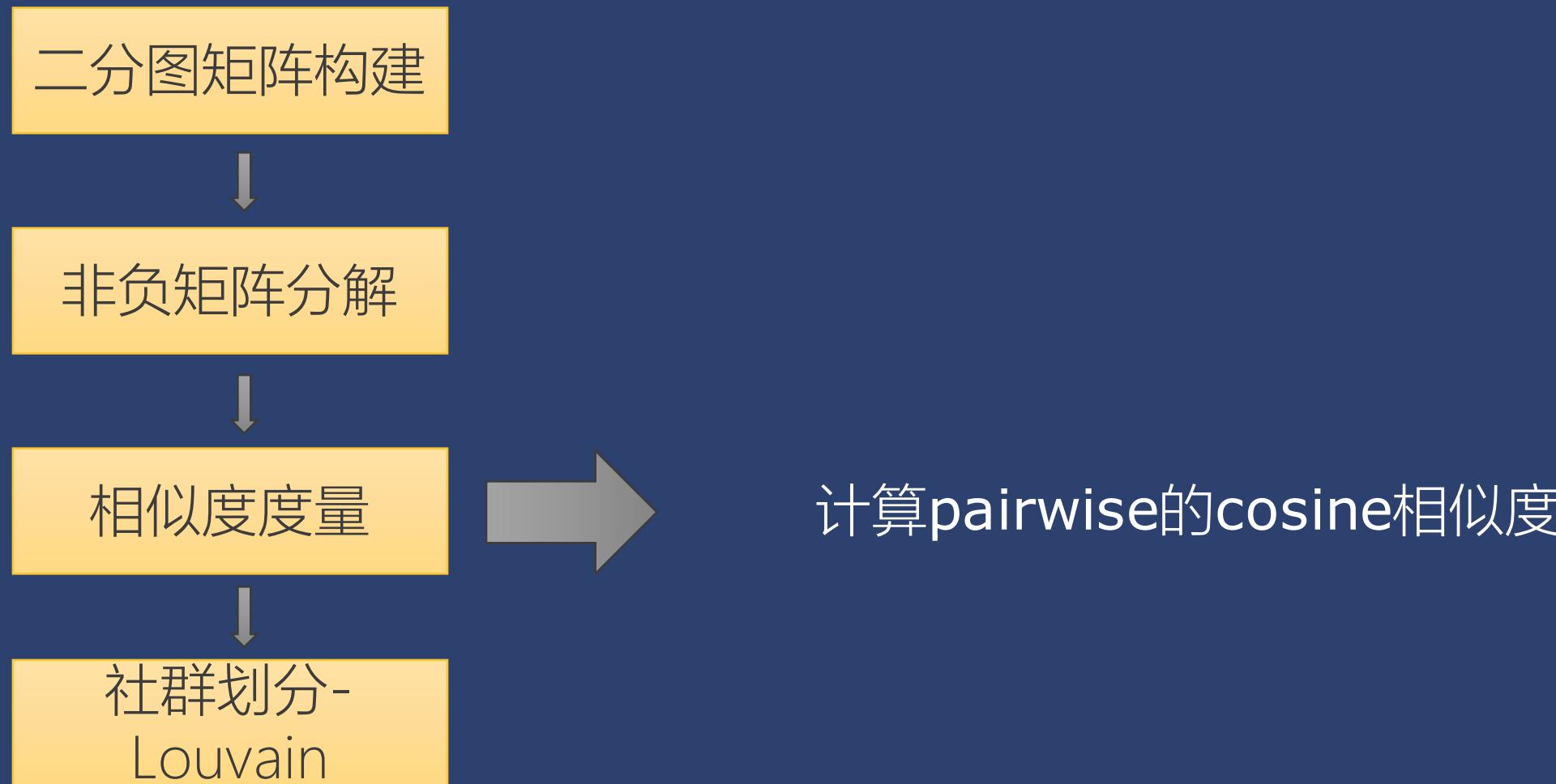
3. 算法流程



3. 算法流程



3. 算法流程



3. 算法流程

二分图矩阵构建



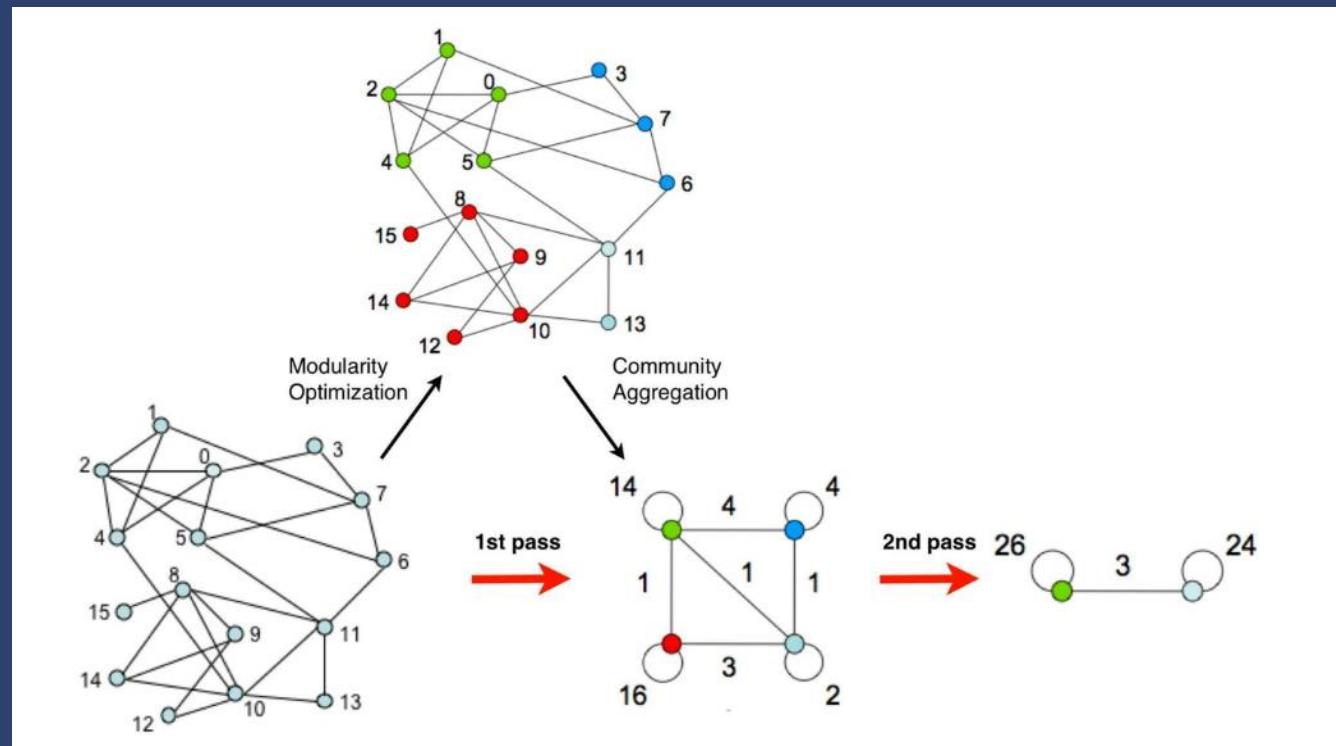
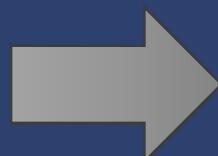
非负矩阵分解



相似度度量



社群划分-
Louvain



4. 划分结果

- 平均每个品类被分为3个左右子类
- 每个子类间价格相近，区分度主要由**用户品味**造成
- 去除产品数量小于100的子类

图1：子类产品数量

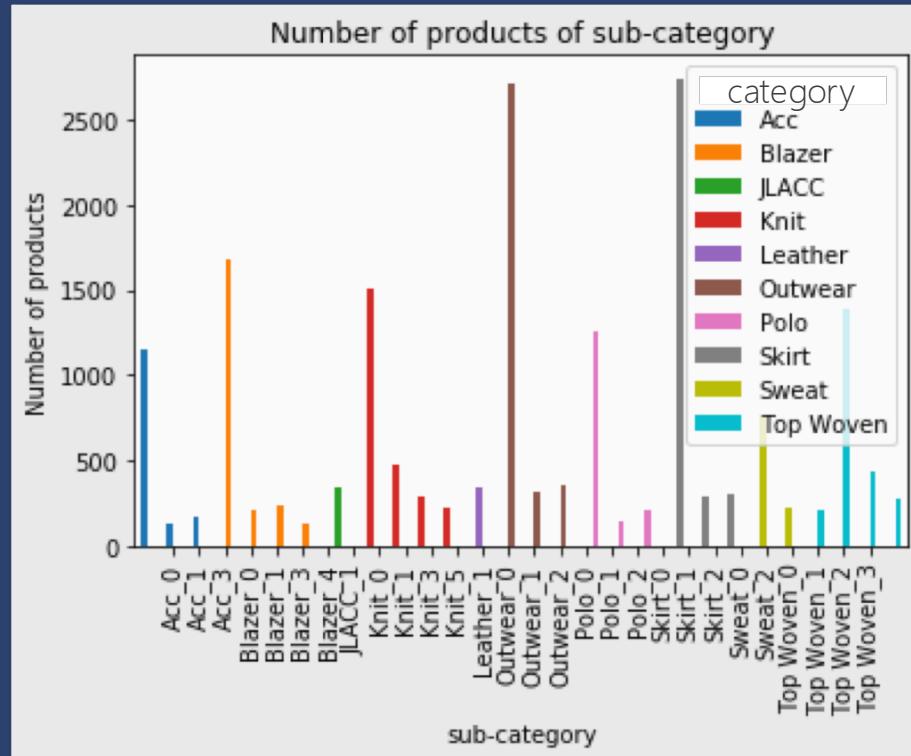
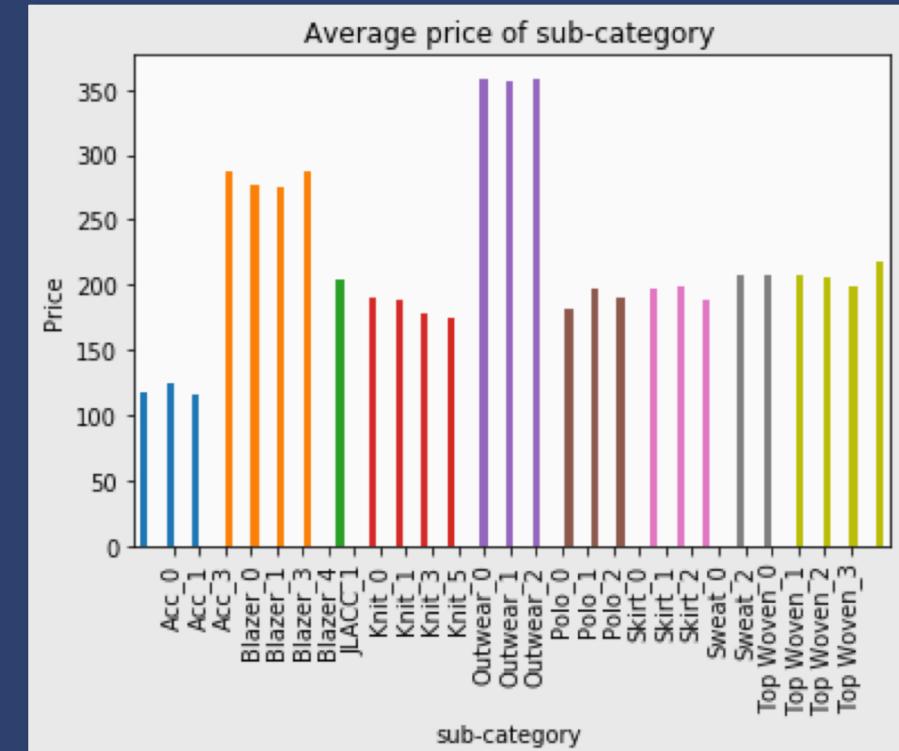


图2：子类产品平均价格

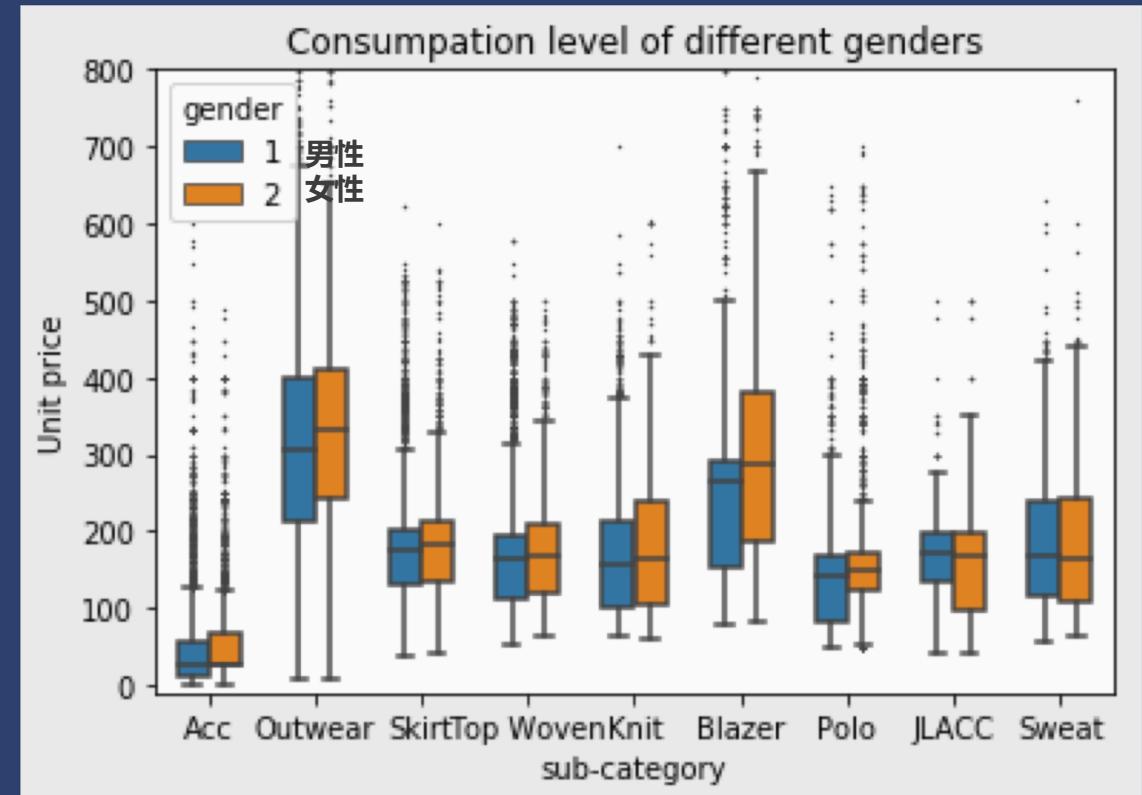


5. 结果分析 – 消费等级

分类前：

- 女性在任何一种品类上的消费等级都略高于男性
- 女性在Polo上的价格覆盖范围明显低于男性
- 无法甄别品类内的具体消费差别

图1：不同性别客群消费等级



5. 结果分析 – 消费等级

分类后：

- 女性在Polo上的价格覆盖范围低主要是由Polo_0子类上引起
- 女性在Acc_0子类上消费价格明显高于男性，而在Acc_1，Acc_3子类上缺低于男性

图1：不同性别客群Polo子类消费等级

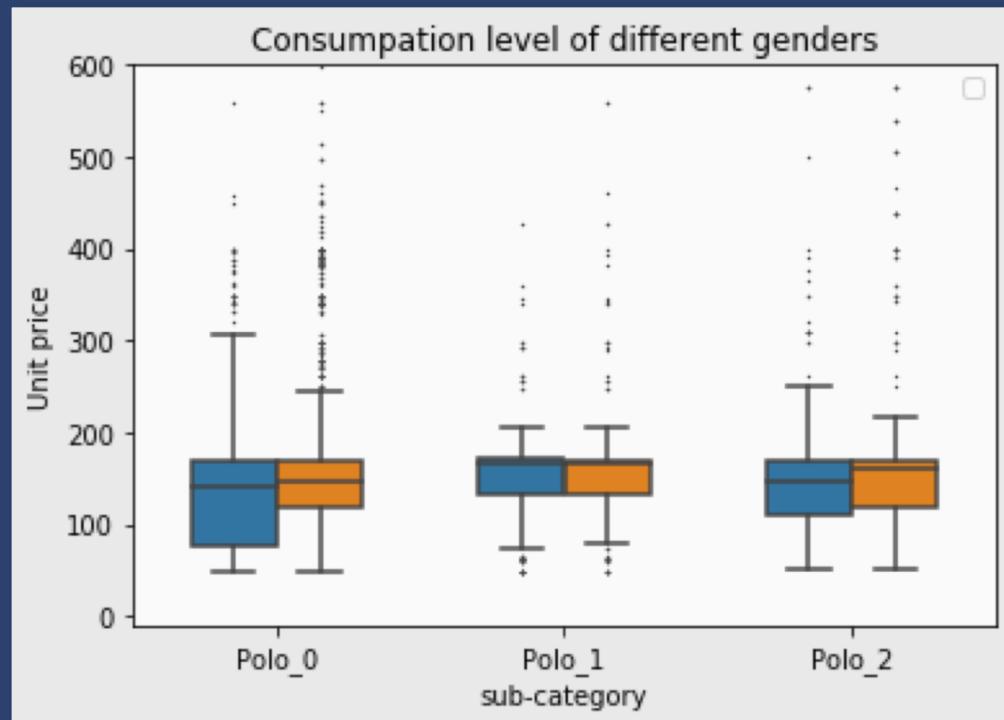
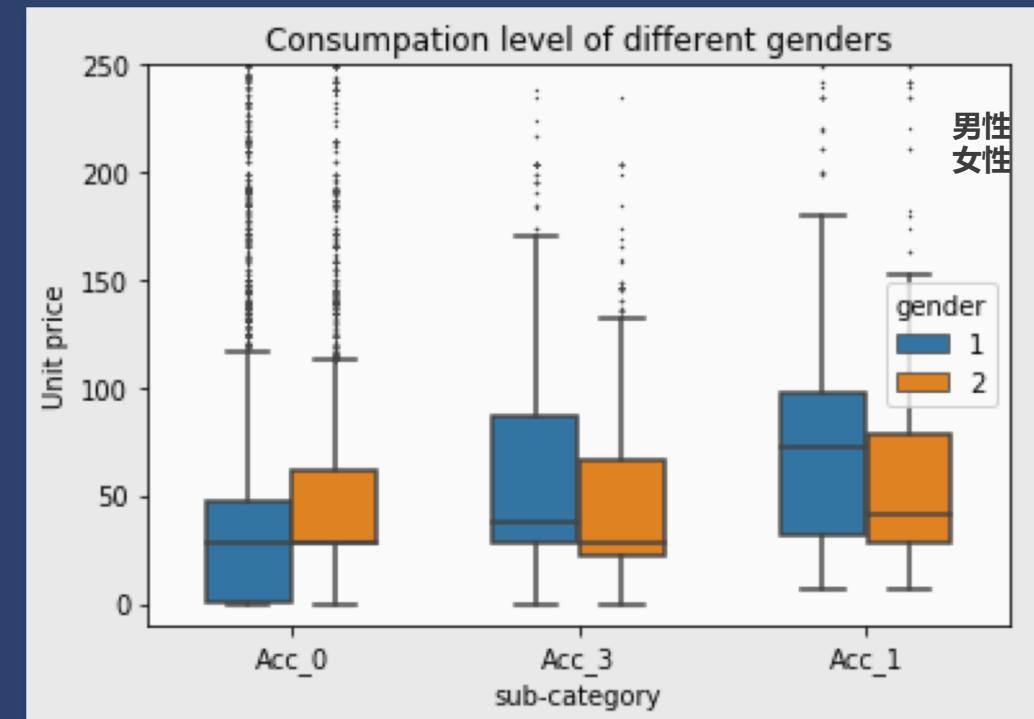


图2：不同性别客群Acc子类消费等级



5. 结果分析 – 消费等级

分类后：

- 女性在所有Blazer产品上的消费等级都高于男性
- 女性在Blazer_1子类上的消费价格跨度较广，价格高于600的部分有大量女性购买但鲜有男性购买

图1：不同性别客群Blazer子类消费等级

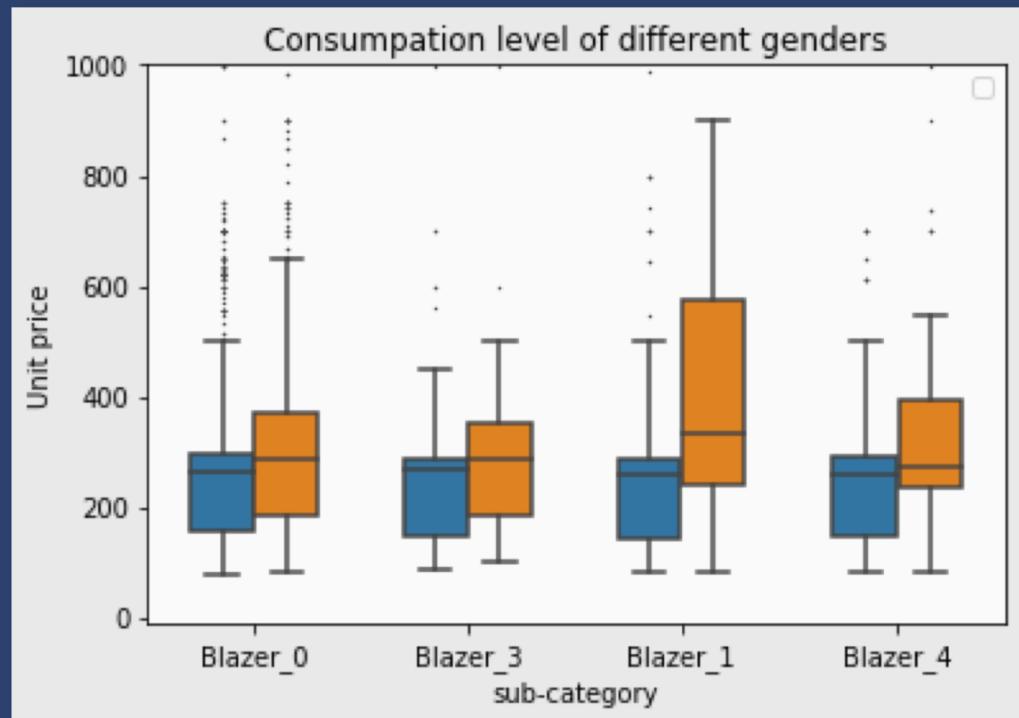
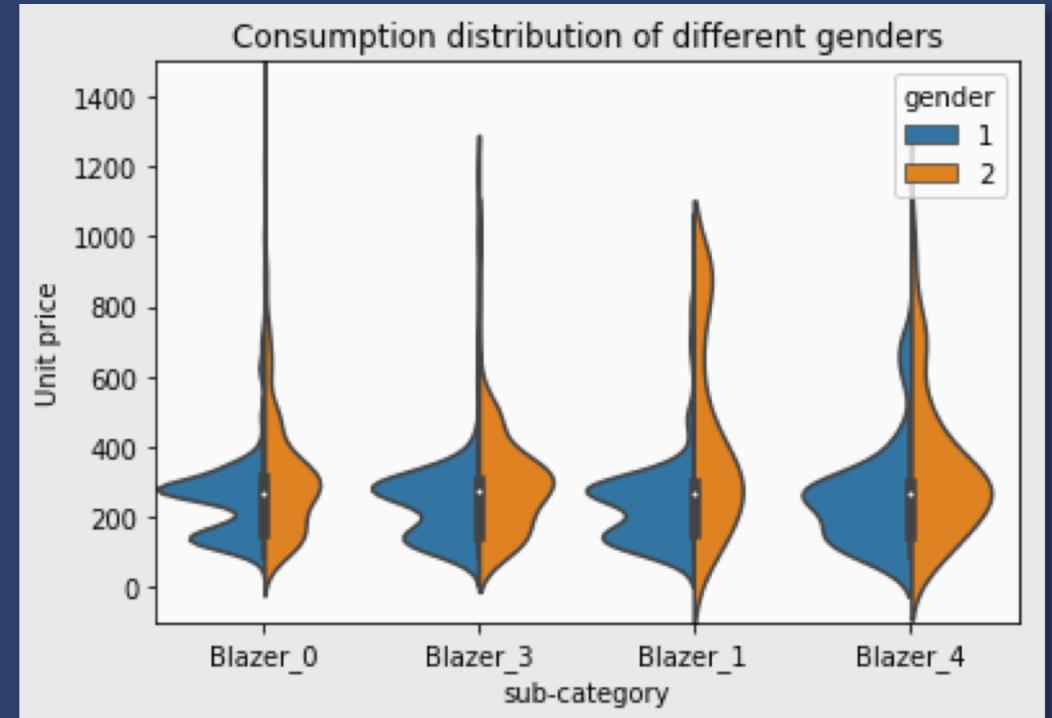


图2：不同性别客群Blazer子类消费贡献



5. 结果分析 – 购买倾向

分类前：

- 从所有购买记录（左图）来看，被购买数量较多的是Acc, JLACC, Polo。
- 对比准会员与金卡会员（右图）来看，金卡会员与左图购买倾向类似，但是准会员的购买倾向更分散，并且程度都有所提高
- 无法观察人群在更小粒度的购买倾向

图1：品类Attraction level

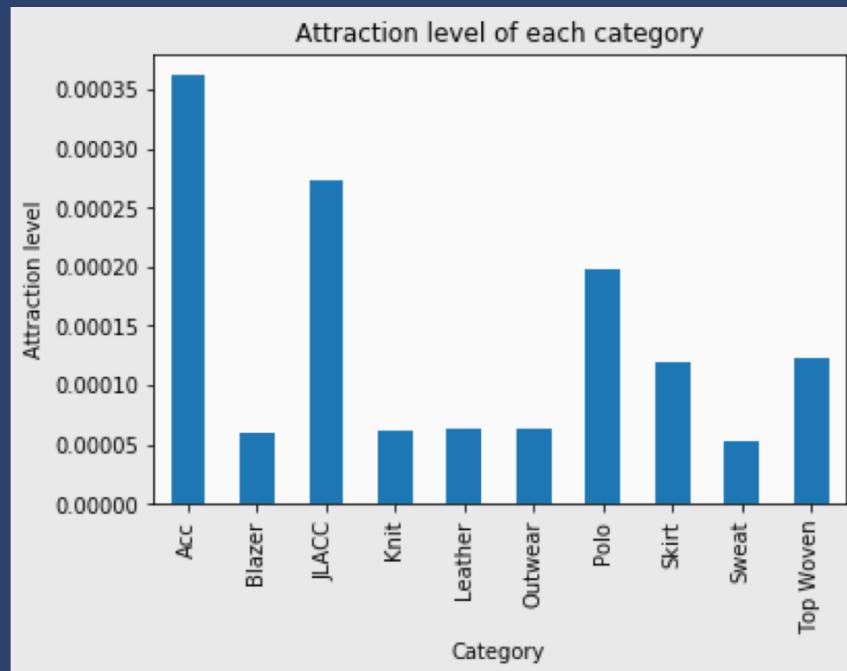
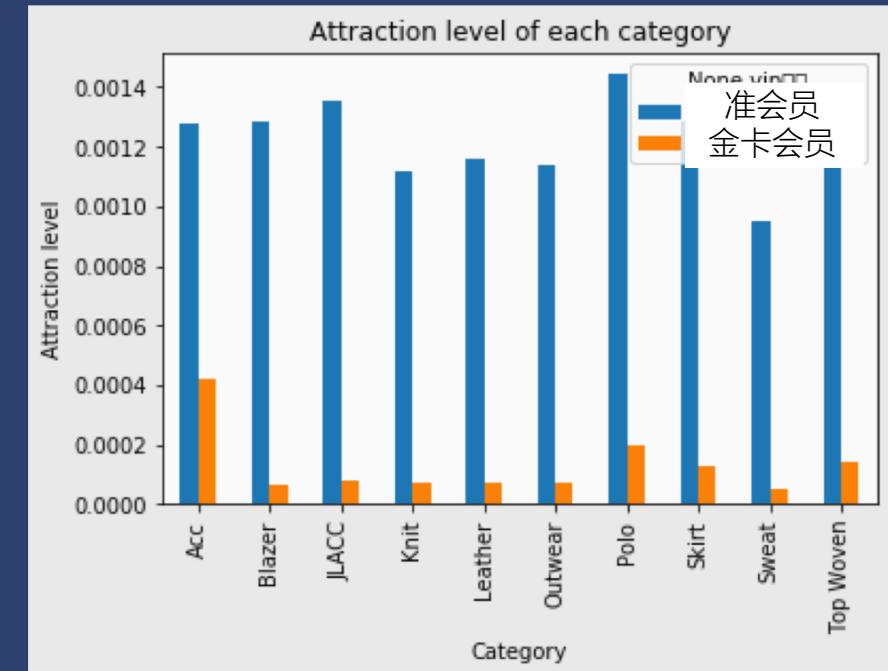


图2：不同性别客群-品类Attraction level



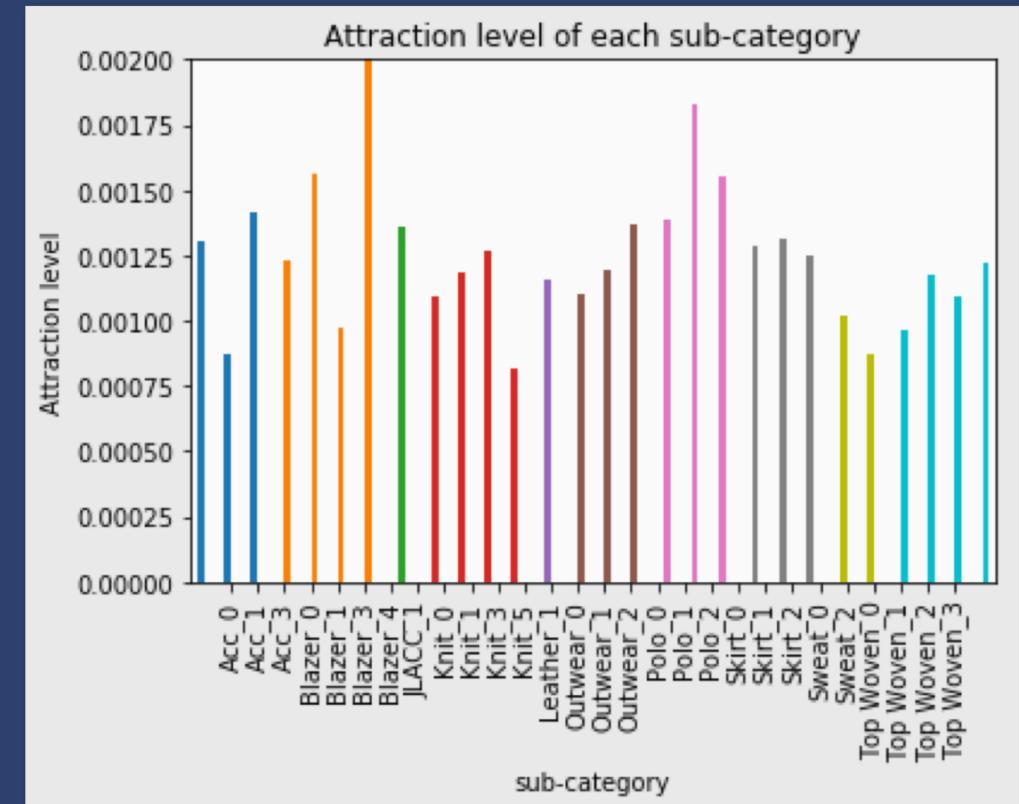
$$\text{Attraction level} = \frac{\text{该类别内产品总销量}}{\text{该类别内产品数量} * \text{人数}}$$

5. 结果分析 – 购买倾向

分类后：

- 可以观察更小粒度的购买倾向
- 准会员在Acc_1, Blazer_3, Knit_5等子类上购买倾向较低

图1：准会员-子类Attraction level

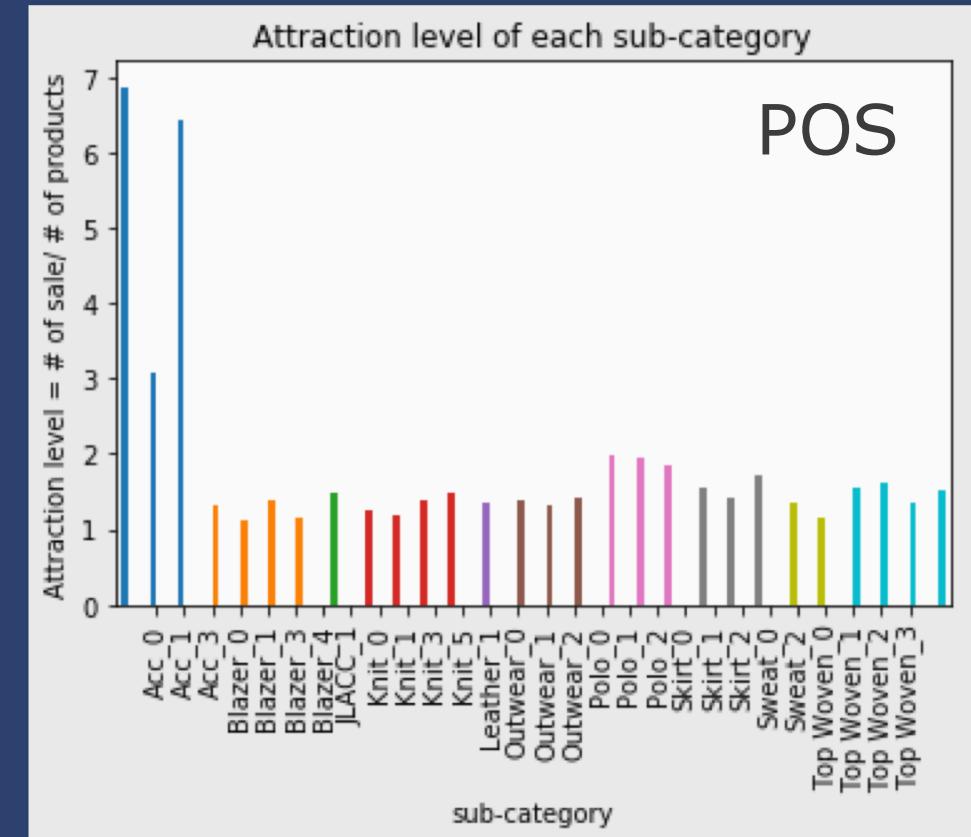


5. 结果分析 – 购买倾向

分类后：

- POS(point of sale)注册用户对Acc_0 与Acc_3子类上有较强的购买意愿

图1：POS注册用户-子类Attraction level

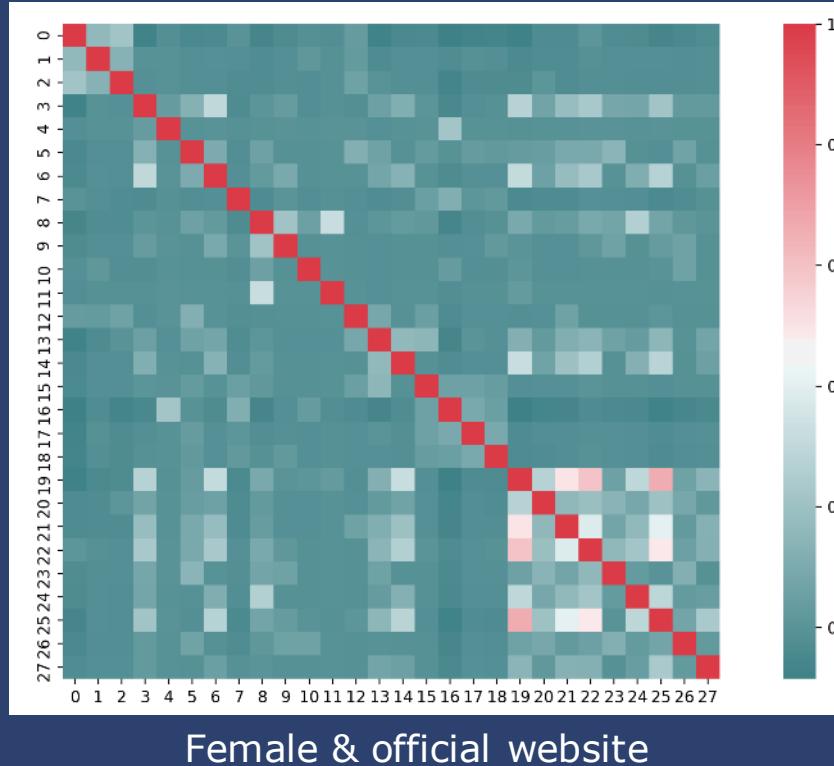


5. 结果分析 – 产品相关性

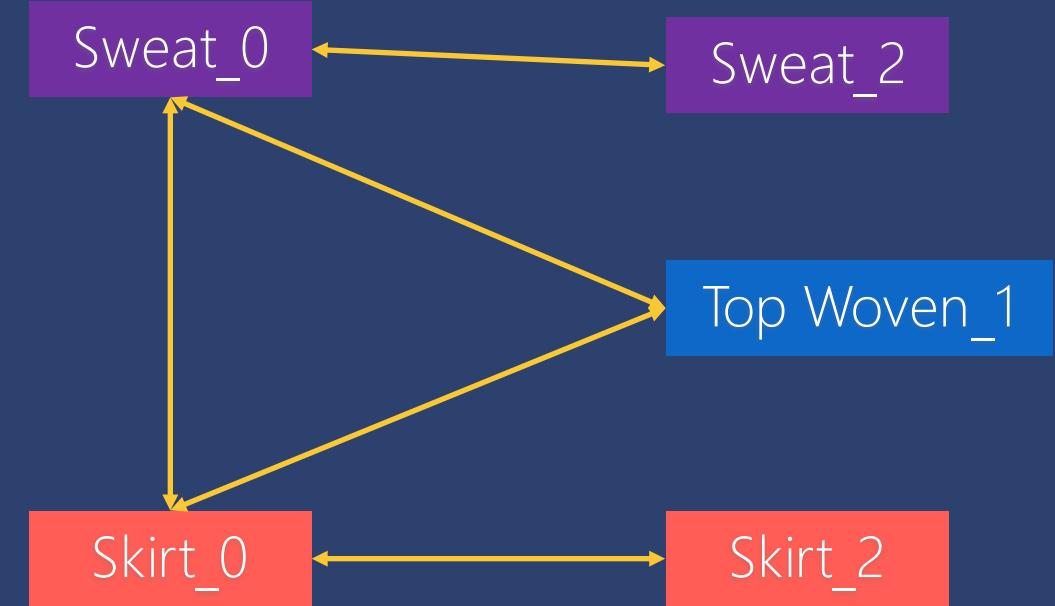
找出**跨类别相关性**高的产品

- 了解客群品味
- 进行服装搭配推荐

Correlation matrix between sub-category



- 相关性高的产品

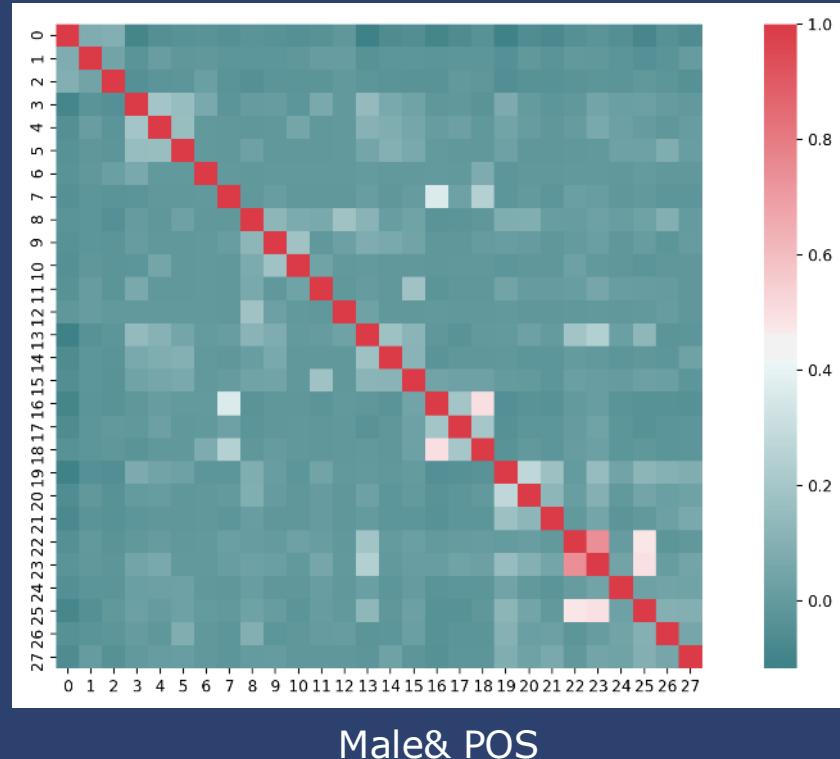


5. 结果分析 – 产品相关性

找出**跨类别相关性**高的产品

- 了解客群品味
- 进行服装搭配推荐

Correlation matrix between sub-category



- 相关性高的产品



6. 下一步计划

