

数据财富

亚马逊在其创建的线上商城中为客户提供了对交易进行评分和评价的机会。个人评级也被称为“星级”，使得购买者可以使用 1（低评级，低满意度）到 5（高评级，高满意度）的等级来表示他们对产品的满意度。此外，客户可以提交基于文本的消息（称为“评论”），以表达有关产品的更多意见和信息。其他客户可以在这些评论中提交有帮助或无帮助的评分（称为“帮助评分”），以辅助他们自己做出产品购买决策。公司使用这些数据来深入了解其参与经营的市场，了解的内容包括参与的时间以及产品设计功能选择。

阳光公司计划在在线市场上推出和销售三种新产品：微波炉，婴儿奶嘴和吹风机。他们已聘请您的团队作为顾问，以在与其他竞争产品相关的客户提供的过去评级和评论中识别关键模式，关系，度量和参数，以：1）告知其在线销售策略；2）识别潜在的重要设计功能，以增强产品的满意度。阳光公司过去曾使用数据为销售策略提供信息，但他们以前从未使用过这种特殊的组合和数据类型。阳光公司特别感兴趣的是这些数据中的基于时间的模式，以及它们是否以有助于该公司制造成功产品。

为了协助您，阳光公司的数据中心为您提供了该项目的三个数据文件：hair_dryer.tsv，microwave.tsv 和 pacifier.tsv。这些数据代表在数据指示的时间段内在亚马逊市场上出售的微波炉，婴儿奶嘴和吹风机的客户提供的评分和评论。还提供了数据标签定义的词汇表。提供的数据文件是您解决此问题的唯一数据来源。

要求：

1. 分析提供的三个产品数据集，以数学证据来鉴定，描述和支持有意义的定量和/或定性模式，关系，测度和参数。这些将在星级，评论和帮助等级之内和之间进行。并且这将有助于阳光公司在三个新的在线市场产品中取得成功。

2. 使用您的分析解决阳光公司市场总监的以下问题和要求：

A. 一旦三种产品在在线市场上出售后，就可以根据评级和评论确定阳光公司的最佳测量度。

B. 在每个数据集中识别并讨论基于时间的度量和模式，这些度量和模式可能表明产品在在线市场中的声誉在上升或下降。

C. 确定最能表明潜在成功或失败产品的基于文本的度量和基于评级的度量的组合。

D. 特定星级会引起更多评论吗？例如，在看到一系列低星级评级之后，客户是否更有可能撰写某种类型的评论？e. 诸如“热情”，“失望”之类的描述是否与其他的得分紧密

欢迎关注微信公众号：科研交流，回复“思路预定”可预定题完整解题思路

相连？

您提交的内容应包括：

- 一页摘要表
- 目录
- 一页至两页的文章
- 您的解决方案不超过 20 页，摘要，目录和文章总计不得超过 24 页。

注意：参考文献和任何附录不计入页数限制，应在完成解决方案后出现。您不应使用未经版权法限制使用的未经授权的图像和材料。确保您引用了想法的来源和报告中使用的材料。

Helpfulness Rating: 表明在决定是否购买该产品时特定产品评论的价值。

Pacifier: 一种橡胶或塑料的舒缓装置，通常是乳头状的，供婴儿吮吸或咬咬。

Review: 产品的书面评估。

Star Rating: 在系统中给出的分数，该分数使人们可以对具有多个星级的产品进行评分
数据集见附件

更多点评和模型参考资料请进内部群，群号：622264434，进不去可联系 QQ（1634063196）

