本题通过对产品数据集的处理，包括去除无用数据、去除无效评价、数据集转化、评论合并等。分析了不同产品的星级与评价之间的关系，并利用python的NLTK包对商品的评论进行了分词、去除标点、去除停用词、词干提取、词性还原并最后进行了情感分析，对每条评论给出了1~5分的评价，1分表示评论为完全负面，5分表示评论为完全正面。分析了每一类产品的每一星级的评论占比，通过对星级和评论情感得分进行建模得到了商品的综合得分，统计出了每天产品获得的评价的数量情况并将其可视化，同时也细化出每天不同星级的评价情况并可视化。我们还分析了评论词数和评论点赞率之间的关系，发现往往评论字数越多的评论点赞率也越高，同时我们发现拥有vine认证的用户往往也拥有较高的点赞率。

我们分别对积极评论和消极评论进行了关键词提取，分别得到了各自的关键词，并制作成了词云，我们通过对关键词的分析发现好评的产品往往在价格、质量方面有优势，而差评的产品往往在外观、使用方式等方面存在不足，厂家也可以根据各自不同产品的词云来进行相关的改善。