问题重述：

第一问：

我们需要做一下数据的预处理，对没有实际意义的数据做合适的处理，然后对数据进行可视化观察。不仅要观察星级评论，文本评论，帮助等级之间的关系，也要观察它们内部的情况，比如作出评论的人是否购买了此产品，是否为vine

第二问

1. 我们需要自定义一个度量标准，从而根据客户本次消费的评级以及评论等来对此次评论进行一个综合评价，可以采用线性加权的方法来得到综合评价的结果
2. 考虑综合评价随时间的变化趋势，从而说明声誉的趋势
3. 基于产品声誉和时间的关系，找出潜在成功或失败的评级评论以及其他可能影响因素的度量组合，
4. 观察时间节点前的星级评价对时间节点后的评论之间的关系
5. 统计文本评论中出现的高频词汇并选取其中的部分词汇进行研究，观察时间节点前选定的这些高频词汇出现的频率与时间节点后的星级评论之间的关系

第三问

总结本次的分析结果，说明数据及分析过程的合理性，最终阐述结论，并根据结论提出合理的建议

The retelling of the question:

First question:

We need to pre-process the data, do the appropriate processing of the data that doesn't make sense, and then visualize the data. Not only do you have to look at =relationships between star ratings, text reviews,and helpfulness ratings, but also look inside them, such as whether the person making the comment perchase the product and whether he is one of the Amazon Vine members

Second question

A. We need to customize a metric to make a comprehensive review of the review based on the customers' ratings and reviews, etc., and we can use a linearly weighted approach to get the results of the comprehensive review

B. Consider trends in comprehensive evaluation over time to illustrate trends in reputation

C. Identify a combination of text-based measure(s) and ratings-based measures and other possible factors that can indicate future sales trends for products, then we can determine whether a product is successful or failed.

D. Observe the relationship between star ratings before time nodes and comments after time nodes

E. Statistical high-frequency words appearing in text reviews and selecting some of them for research, observing the relationship between the frequency of these high-frequency words selected before the time node and the star rating review after the time node

Third question

Summarize the analysis results, explain the reasonableness of the data and analysis process, finally expound the conclusion, and make reasonable suggestions according to the conclusions.