Dear Marketing Director,

Our team has analyzed your requirements and completed the task.

To meet the first three requirements, we established the SRR model which is short for Sentiment-Reliability-Renown Model. It is based on an assumption that the most three factors that determine whether a product is successful or not is reviewers’ sentiment, the reliability of the reviews and the renown of the products. The reviewers’ sentiment relies on the star rating, and the sentiment score based on the content of the review headline and the review body. The reliability is based on whether the reviewer is a member of Amazon Vine Voices, whether the reviewer has actually bought this product, the helpfulness rating, the total votes, and the length of the review. And the renown relies on the total number of the reviews and the density of reviews which denotes average increasing number of the reviews per day. You just need to input these data crawled on the Internet, and then you can get an assessment of the product’s market and potential trend of development. The rising of the integrated score denotes that that a product’s reputation is increasing and vice versa. The inflection point of the figure denotes that the product is a potentially successful or failing product. You can gain insights into the markets in which they participate, the timing of that participation, and the potential success of product design feature choices by adopting this model.

To meet the fourth requirements, we analyze the statistical significance between the star rating and the density of the reviews. We find that the star rating and the density of reviews have the statistical significance. Then considering that the inciting process is usually kept for a long period of time. We plan to analyze the correlation between a star rating and next months’ total number of reviews. Then we find that one star can incite more reviews, while 5 stars can’t.

To meet the last requirements, we pretreat the raw text and then choose the most frequent words to calculate their tf-idf. Then analyze the correlation between tf-idf of these words and star rating. Finally, we find that there is association between some quality descriptors and star rating.

The fourth and the last work we do prove the relationship between rating and review, which in turn prove that our model is rational.