

# A study on Amazon's distinctive features

**Raphael Barman**

raphael.barman@epfl.ch

**Hakim Invernizzi**

hakim.invernizzi@epfl.ch

**Albane Descombes**

albane.descombes@epfl.ch

## Abstract

This document describes and discusses the results obtained in the context of the authors' project in ADA2017. The project aims at deepening the understanding of Amazon's framework. The analysis is centred around three different aspects .

First, the review feature, with the goal to describe its usage from the userbase and answer the question of whether reviews are generally biased or not.

Second, the adverting feature which is composed of the 'also bought', 'also viewed' and 'bought together' features. The intra-correlation among the sub-features as well as their impact on the whole system are the subject of study.

Lastly, the categorisation feature and its impact on the user perception of a product.

The three features represent three different starting points for the analysis which should ideally converge in a single conclusion.

## 1 Introduction

Explain the relevant information about metadata here.

## 2 Review feature

The first step in trying to describe Amazon's reviews was trying to detect potential correlation between its features. This was done by defining two metric: the "wordcount", which is the length of a review in number of words, and the "helpfulness", which is a measure of the helpfulness of a review in the range  $[-1, 1]$ . By observing the distribution of these two metrics it was possible to remark that the majority of reviews are less than 100 words long and are considered unhelpful by

users [add figure?]. The correlation between these two metric and other descriptive features of the review such as the product ranking was studied, however no relevant result was found. The most interesting result was noticing the presence of a weak Spearman correlation (0.23) between helpfulness and word count. This hint at the presence of a monotonic component in the relationship, which makes sense: if a review hasn't a certain minimum length, then it is unlikely it will be considered helpful.

The second step was trying to quantify the bias from the reviewers . A first analysis showed that the average number of review written per user is low [add figure?], therefore the sample will be composed of users who have written at least 5 reviews, a sample whose size is of approximately 110000 reviewers. The average rating for each reviewer was computed [add figure], then further averaged to obtain the mean average rating. This value is of 4.23 out of 5, with a mean standard deviation of 0.9. It can be argued that the mean average rating is high while the mean standard deviation is low. Furthermore by varying the sample based on the standard deviation, it was noted that the weakest the standard deviation, the higher the average for the reviewer [add figure?]. This is a peculiar result that could be argued as pointing towards a general bias in the review system, since the only situation that justify these results is that all product are excellent, an hypothesis which doesn't hold. A possible explanation to this behaviour could be that Amazon users tend to rate the product they're satisfied with and not those they're unsatisfied with.

The third step was studying the users behaviour with the regard to the brand. The approach was to select reviewers with at least 5 reviews concerning a same brand and analyze the mean average

rating and mean standard deviation. This value are 4.5 and 0.63 respectively on a sample of [size]. While these values are not directly comparable to the ones obtained in the second part because they refer to sample of different size, the higher mean average rating and the lower mean standard deviation let think of an even higher bias in the evaluation.

Furthermore, the correlation between the ranking of a product and its reviews' quantity, average helpfulness and average rating were studied without meaningful results.

### 3 Advertising features

The three advertising features subject of the study are the 'also bought', 'bought together' and 'also viewed' features. As a first approach, correlation between these three features and the product ranking have been computed. While there were no relevant results concerning 'also bought' and 'also viewed', a Spearman correlation of 0.41 has been detected in the relationship between the ranking of two product that were 'bought together'. This result implies the presence of a monotonic component in the ranking of bought together items, which means the ranking of one product influences the ranking of the other.

Another goal of the research was to examine the intra-correlation between the three features, however due to the array-like nature of these attributes and the consequent exponential growth in the size of the dataframe this was not possible.

### 4 Categorisation feature

The third part of this analysis aims at assessing the impact of the product categorisation on the whole Amazon's system. To do this, products are divided by category. Then, the number of reviews per category as well as the average helpfulness and average product rating per categories are computed.

As expected, category which are conceptually broader ('Automotive') have many more reviews compared to niche category ('Pizza Kits'). The mean average helpfulness is -0.32, highlighting the fact that helpfulness of the reviews is judged in a slightly negatively way. Concerning the mean average rating, it is 4.16, a

result which is line with the result obtained in the analysis of reviews.

Furthermore, there are interesting results from the point of view of Spearman correlations.

The correlation between the average helpfulness of reviews and the average rating of reviews per category is -0.24, which means that there's a weak monotonic component in this relationship, similarly to what was observed in the first part of this analysis. This results means that if a review gives an high rating to the product, then it is slightly more likely to have a [lower/higher] helpfulness.

The correlation between average helpfulness and number of reviews per category is 0.23. The interpretation of this result is that if a category has many reviews, then it is slightly more likely that the reviews are considered helpful in average.

## 5 Conclusions

try to sum up things here.

## References

Laerd Statistics.

<https://statistics.laerd.com/statistical-guides/spearmans-rank-order-correlation-statistical-g.php>.

Last visited on 17.12.2017.

Minitab Express Support.

<http://support.minitab.com/en-us/minitab-express/1/help-and-how-to/modeling-statistics/regression/how-to/correlation/interpret-the-results/>.

Last visited on 17.12.2017.