ISOM5160 Group Report

Dataset: amazon_food_reviews.csv

Group7

CAO, Xi	21271664	LI, Heyi	21265689
LIAO, Jingyu	21262106	LIN, Chuwei	21237955
YE, Chenwei	21199517	ZHANG, Ziyang	21266920

October 9, 2025





- 1 Introduction
- 2 Literature Review
- 3 Correlation Between Ratings and Product Descriptions
- 4 Results

- Introduction
- 2 Literature Review
- 3 Correlation Between Ratings and Product Descriptions
- 4 Results

Background and Motivation

Why to give a presentation:

- show the main arguments and results of your work
- produce interest to read the full paper/report
- goal: be educational and also entertaining

Advantages of using LaTeX with the beamer package:

- very easy if the report is already written in LaTeX
- · different themes which are usable in practice
- possibility to create handouts using beamerarticle

Research question

this is a Research question.



- Introduction
- 2 Literature Review
- 3 Correlation Between Ratings and Product Descriptions
- 4 Results

Research gap

this is a Literature gap.



Research question

this is a Literature Review.



8 / 21

- Introduction
- 2 Literature Review
- **3** Correlation Between Ratings and Product Descriptions
- 4 Results



Overview

1. Visual Information Correlation

Correlation with the number of sample images

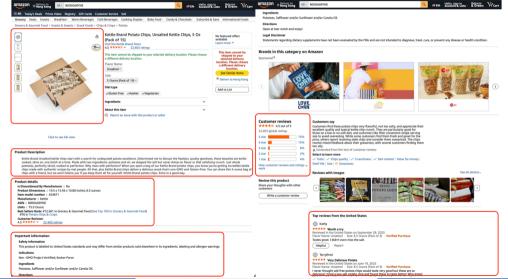
2. Text Information Correlation

- Correlation with description length
- Correlation with reading ease
- Correlation with the marketing tone of the description
- Correlation with product description items beamerarticle

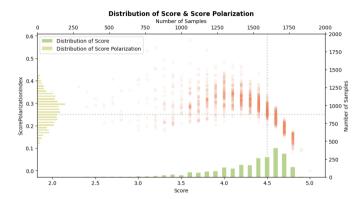
Correlation computation: We use spearmant to compute Correlation



Supplementary data (Scraped from Amazon's website)



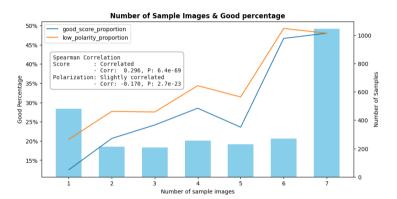
Rating features



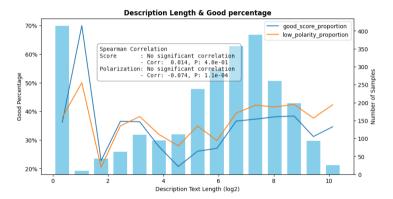
- **10 Mean Score:** The score displayed on the product detail page.
- **②** Score Polarization: Indicates whether the ratings for this product are polarized.
- **Standards of Good Score:** Score >= 4.5; Polarization <= 0.25



Correlation with the number of sample images



Conclusion: The is a **positive correlation** in product ratings and descriptions. The more images included in the description, the greater the likelihood of the product receiving positive reviews.

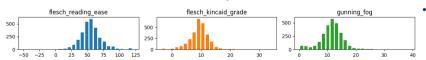


Conclusion: There is **no significant correlation** between product ratings and the length of product descriptions.

Correlation with reading ease

ease_index	compare_with	corr	p_value	conclusion
flesch_reading_ease flesch_reading_ease flesch_kincaid_grade flesch_kincaid_grade gunning_fog	Score Polarization Score Polarization Score	0.0163 -0.0187 0.0123 -0.0130 0.0066	0.3863 0.3214 0.5135 0.4900 0.7270	No significant correlation No significant correlation No significant correlation No significant correlation No significant correlation
gunning_fog	Polarization	-0.0076	0.6870	No significant correlation





- flesch reading ease: The higher the score, the easier it is to read.
- flesch kincaid grade:
 The required grade level to read; The higher the number, the more difficult the reading level.
- gunning fog: Long word ratio; The higher the ratio, the harder it is to understand.

Conclusion: Product ratings show **no significant correlation** with reading difficulty.



HKUST ISOM5160 Group Report October 9, 2025.

Correlation with the marketing tone of the description

sentiment_type	compare_with	corr	p_value	conclusion	_
marketing_tone_score	Score	0.0588	0.0007	No significant correlation	-
marketing_tone_score	Polarization	-0.0469	0.0066	No significant correlation	
sentiment_score	Score	0.0788	0.0000	No significant correlation	
sentiment_score	Polarization	-0.0568	0.0010	No significant correlation	



Types of marketing tone (Measuring Criteria):

- **1 marketing tone score:** The proportion of exaggerated marketing language in product descriptions. The higher the score, the more exaggerated the marketing claims become.
- **2 sentiment score by language model**: model is distilbert-base-uncased-finetuned-sst-2-english, The higher the score, the more positive the emotion.

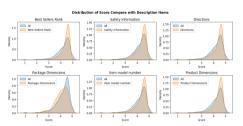
Conclusion:

- · Product ratings show no significant correlation with marketing tone/sentiment.
- *However, since the metrics used to measure the emotional orientation of product descriptions do not follow a normal distribution, conclusions drawn from this basis may be unreliable.



Correlation with product description items

item	compare_with	corr	p_value	conclusion
Best Sellers Rank	Score	0.3809	2.9e-116	Highly correlated
Directions	Score	0.2211	2.12e-38	Correlated
Item model number	Score	0.2721	4.95e-58	Correlated
Package Dimensions	Score	-0.2473	6.66e-48	Correlated
Package Dimensions	Polarization	0.1717	1.32e-23	Slightly correlated
Product Dimensions	Score	0.2715	9.67e-58	Correlated
Safety Information	Score	0.1618	4.03e-21	Slightly correlated



Products with the following description items might have a better rating:

- Best Sellers Rank: Only high-quality goods carry this label.
- Item model number & Product Dimensions: Products with model numbers may be more formal and reliable.
- Directions: Products accompanied by directions may reduce negative reviews caused by users' inability to use or misuse the product.
- · Safety Information: Products with safety information may reduce negative reviews caused by user allergies.

Products with the following description items might have a lower rating:

Package Dimensions: May receive negative reviews due to logistics-related issues.



October 9, 2025

Chapter Conclusion

For Sellers

- Increase the number of product images;
- Make product descriptions as professional as possible;
- Provide more detailed safety information and instructions;

For Buyers

- Choose products with more pictures;
- Choose products that appear more professional;
- Carefully read the product descriptions to avoid potential issues;

- Introduction
- 2 Literature Review
- 3 Correlation Between Ratings and Product Descriptions
- 4 Results

- different themes
- different themes
- different themes
- different themes

Thank you for listening!

Group 7

@connect.hkust-gz.edu.cn