

DATA-ORIENTED PROGRAMMING PARADIGMS

EXERCISE 2 / GROUP 01

An Inspection of Amazon Reviews

Hardt Julian, Jabary Yahya, Oster Johannes, Weber Martin

AGENDA.

1	Research Questions & Dataset
---	------------------------------

2	Deepdive Into Each Question
---	-----------------------------

3	Summary
---	---------

4	Q&A
---	-----

01

Research Questions & Dataset

Exploring 571M Reviews Across 33 Product Categories

01

CATEGORY BIAS

Are reviews for some categories of product on Amazon overall more positive than for other Categories?

02

SUBJECTIVITY

Are reviews more subjective for some classes of products than for others?

03

IMPORTANT PRODUCT CHARACTERISTICS

Which aspects of different classes of products are the most important in the reviews?

04

RATING PREDICTION FROM TEXT

Can one predict the star rating from the review text?



Amazon Reviews'23¹ with **571.54M reviews** from May 1996 - Sept 2023 across **33 + “unknown” categories**

Provides 2 components:

- User Reviews Data
- Item Metadata

SIMPLE DATA PREPROCESSING:

- Select 1000 random samples per category
- Removed HTML tags from text and title fields
- Stripped whitespaces
- Eliminated reviews with empty text or titles
- Dropped unused columns

FIELD	TYPE	DESCRIPTION
rating	float	Rating of the product (1.0 - 5.0)
title	str	Title of the user review
text	str	Text body of the user review
main_category	str	Main category of the product

Final data fields used

1. <https://huggingface.co/datasets/McAuley-Lab/Amazon-Reviews-2023>

02

Deepdive Into Each Question

Category Bias, Subjectivity, Product Characteristics and Rating Predictions

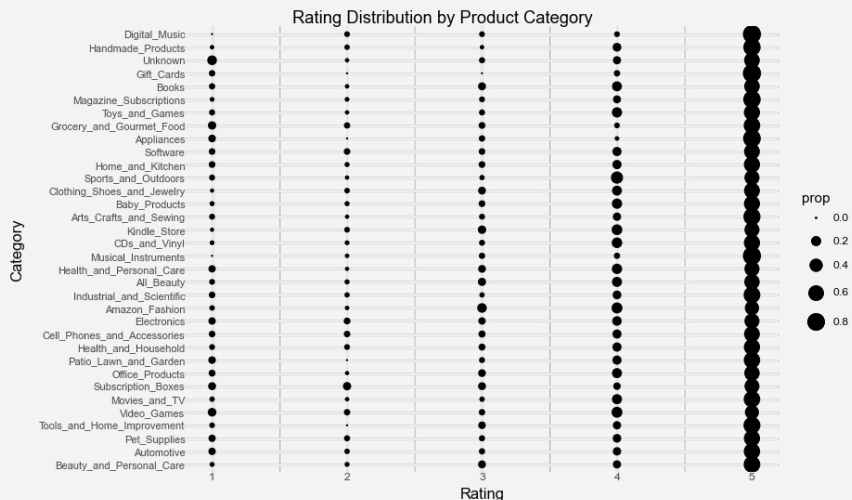


Are reviews for some categories of product on Amazon overall more positive than for other Categories?

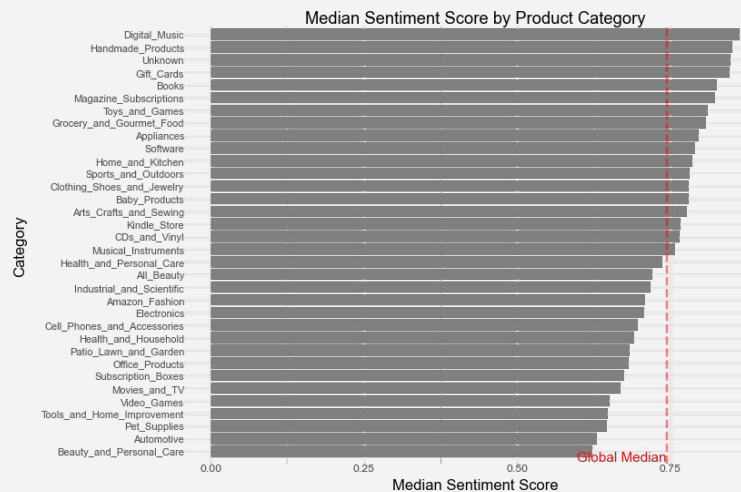


Yes: *Digital Music, Handmade Products, Gift Cards and Books lead in positive sentiment, while Beauty and Automotive rank lowest, though differences between categories are modest with most clustering near the global median.*

Continuous sentiment score by BERT-based¹ model for text and title of reviews



Star rating distribution with dot size indicating relative frequency within each category



1. <https://huggingface.co/lxyuan/distilbert-base-multilingual-cased-sentiments-student>

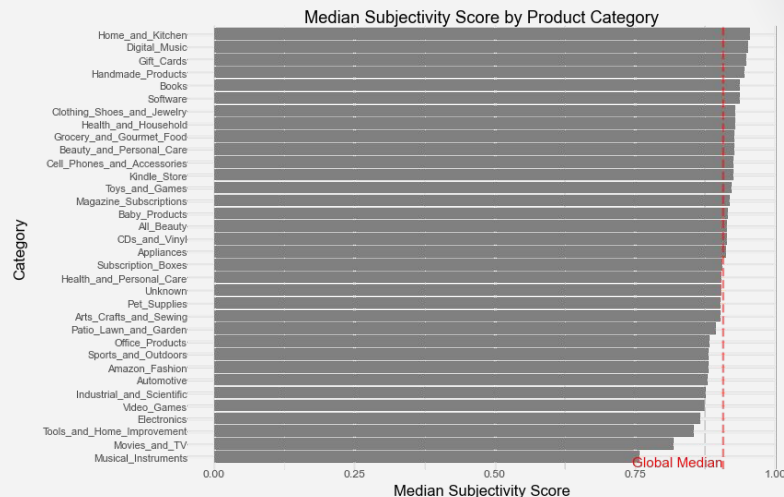
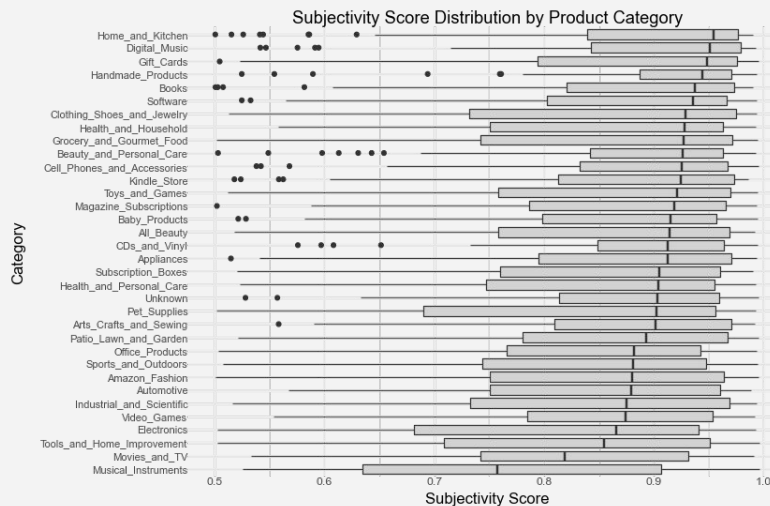


Are reviews more subjective for some classes of products than for others?



Yes: *Home/kitchen, digital music, and gift cards* demonstrate the most subjective reviews, while electronics, tools, and musical instruments exhibit less subjective review content. All categories show consistently high subjectivity scores.

Continuous *subjectivity score* by BERT-based¹ model for text and title of reviews sorted by median subjectivity score.



1. <https://huggingface.co/cffl/bert-base-styleclassification-subjective-neutral>

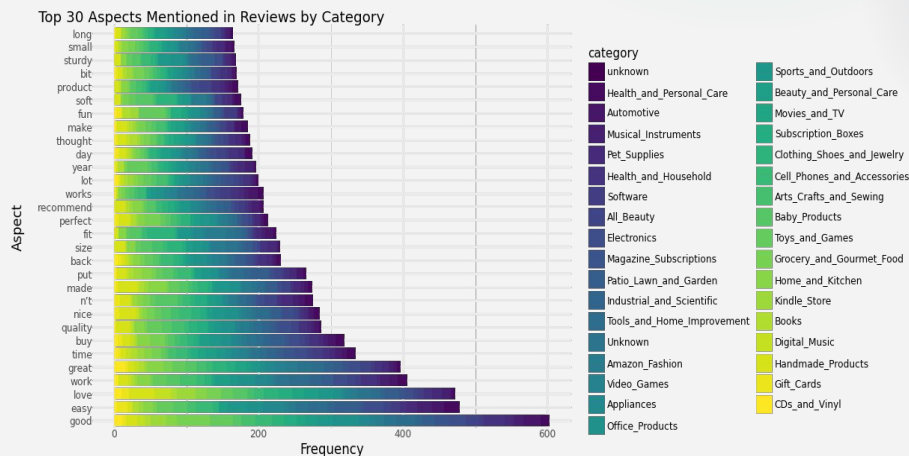


Which aspects of different classes of products are the most important in the reviews?



While aspect extraction identified frequently mentioned characteristics like "size" and "quality," the analysis failed to reveal meaningful product-specific insights, instead surfacing generic descriptors across all categories.

- Applied **unsupervised keyword extraction** tool: YAKE¹
- Category-specific top-k analysis failed to reveal unique product features, with top terms **remaining general** across all categories
- Most meaningful aspects were physical attributes ("size", "small") and customer experience terms ("easy", "works", "recommend")

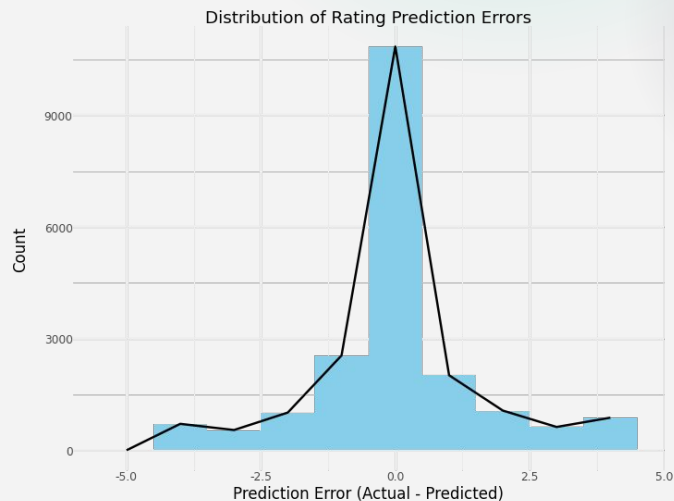
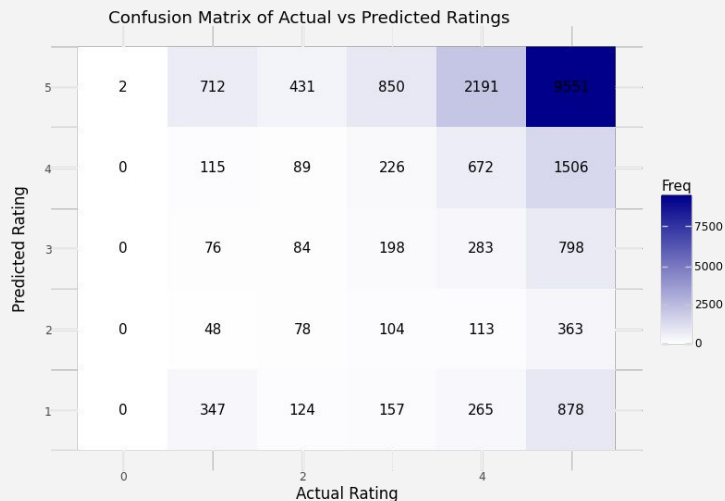




Can one predict the star rating from the review text?



BERT classifier¹ produces accuracy of around 68% with highest error rates in 1-star and 4-star ratings. However, as not thorough training-dataset² analysis as been carried out this has to be taken with a grain of salt.



1. <https://huggingface.co/LiYuan/amazon-review-sentiment-analysis>

2. <https://www.kaggle.com/datasets/cynthiarempel/amazon-us-customer-reviews-dataset>

03

Summary

01

CATEGORY BIAS

Digital Music, Handmade, Gift Cards and Books show highest sentiment scores, with Beauty and Automotive lowest, though variations are small.

02

SUBJECTIVITY

Home/kitchen, digital music, and gift cards have most subjective reviews; electronics, tools, and instruments have least. Overall subjectivity is high.

03

IMPORTANT PRODUCT CHARACTERISTICS

Aspect extraction found common terms ("size", "quality") but lacked category-specific insights.

04

RATING PREDICTION FROM TEXT

BERT classifier achieved ~70% accuracy, struggling with 1- and 4-star ratings. Results need validation due to limited training data analysis.

Q&A!

APPENDIX.

1	Review Ratings
---	----------------

2	Sentiment / Subjectivity Distributions
---	--

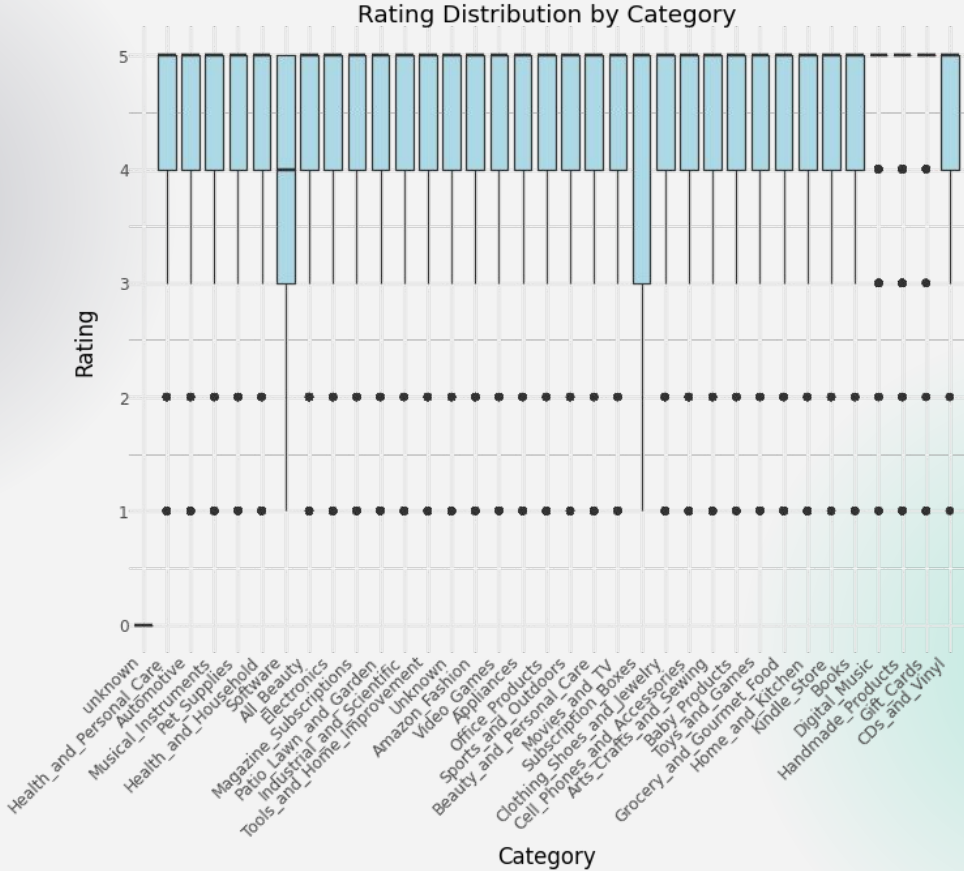
3	Aspect Extraction per Keyword
---	-------------------------------

A1

Data Insights

Appendix.





A2

Sentiment / Subjectivity Distributions

Appendix.

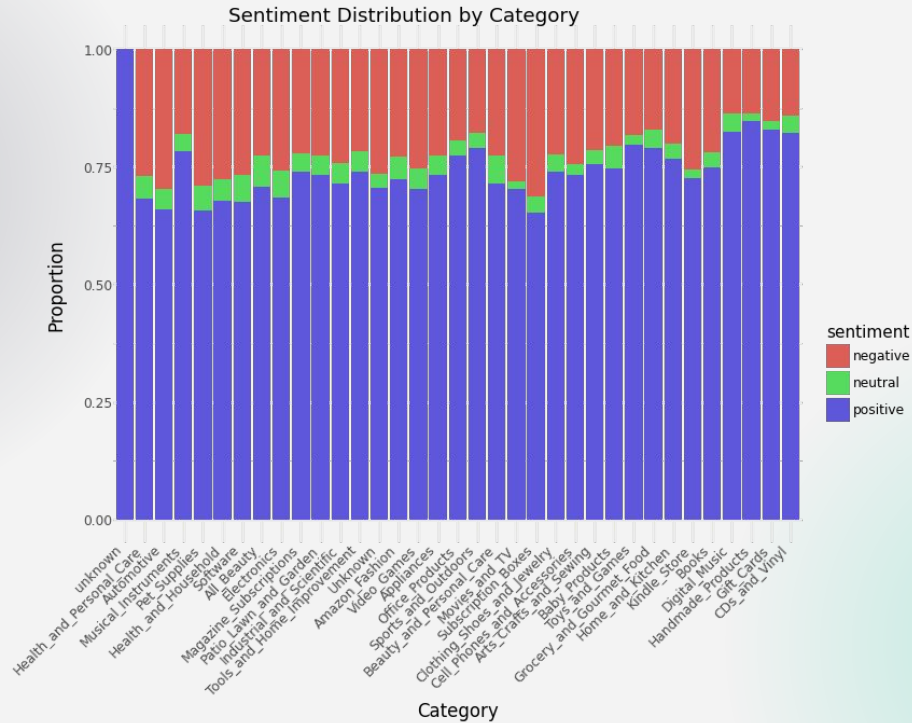
SENTIMENT / SUBJECTIVITY DISTRIBUTIONS



Besides a continuous sentiment score the BERT-based¹ model also categorized each sentiment into one of 3 classes.

1. <https://huggingface.co/lxyuan/distilbert-base-multilingual-cased-sentiments-student>

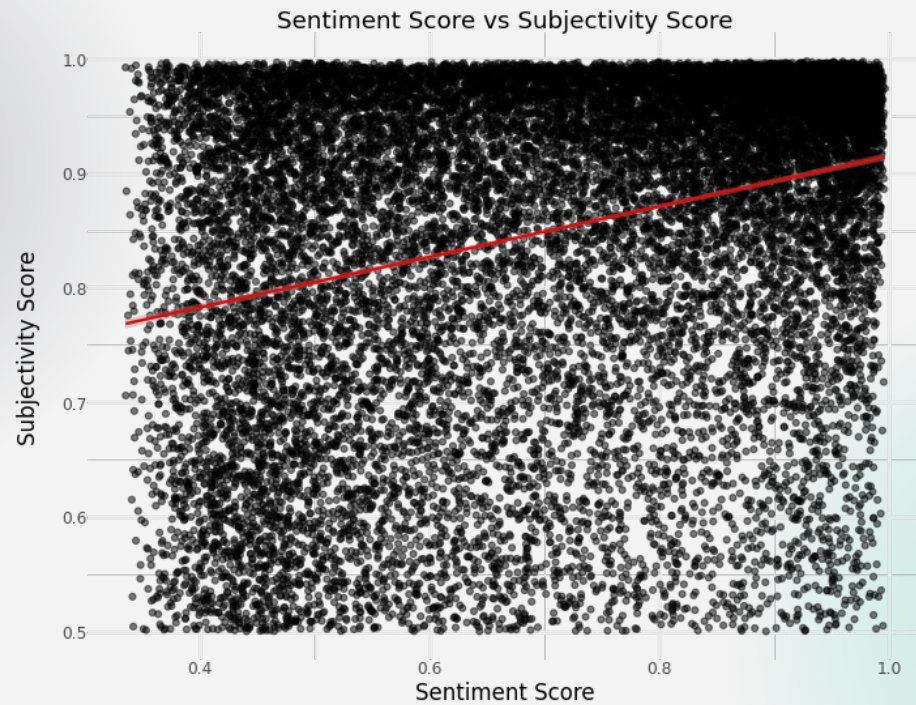
SENTIMENT / SUBJECTIVITY DISTRIBUTIONS



Besides a continuous sentiment score the BERT-based¹ model also categorized each sentiment into one of 3 classes.

1. <https://huggingface.co/lxyuan/distilbert-base-multilingual-cased-sentiments-student>

SENTIMENT / SUBJECTIVITY DISTRIBUTIONS



A3

Aspect Extraction Per Keyword

Appendix.

ASPECT EXTRACTION PER KEYWORD

