The correlation between Amazon product review and Amazon products sales

Katie Tran
Department of Computational and Data Sciences, George Mason University, Fairfax, VA

Background or Abstract

Reviews are still important for marketing products on Amazon, its reviews specifically rate whether the product met the customer's expectations. Reviews can be an important part of product listing optimization strategy, not only in converting traffic that comes to product listing, but also in increasing or decreasing traffic that comes from the search engine result page (SERP).

Figure 1: Amazon Product Price & Reviews

Research Questions

1. What category of product on Amazon is the best seller? 2. Amazon product reviews affects Amazon product sales in positive or negative ways? 3. Compare Amazon marketplace versus Walmart marketplace

Related Research

Amazon recommends at least **25 reviews** for a listing before sellers start scaling advertising. If a product is very niche, it may only need **five to 25 reviews** to be retail-ready.

Data

Amazon reviews data is collected from **2013 to 2017**. The reviews are picked randomly and the corpus has nearly 1600 reviews of different customers. The dataset is 27 columns including 99% of Amazon product brand and 1% of Moshi.
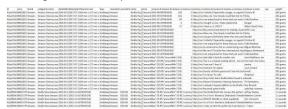

Table1: Amazon product reviews dataset

Preprocessing

The processing of Web Scraping: query data directly from API to get the dataset in csv.file:
  1. Inspect the XHR network section of the URL to get URL
  2. Input data into Python to create a data format I want

Figure 2: Inspect the XHR Network Section of the URL

Figure 3: Input Data into Python to Get a csv.file

Methods / Analysis

Using Natural Language Processing (NLP) to analyze texts, allowing machines to understand how humans speak.

Figure 4 : Understand of Natural Language Processing

Figure 5: Components of NLP

Figure 6: Steps to Practice NLP

DATA ANALYSIS METHOD IN PYTHON:
  • Load dataset in csv file into Python
  • Use **nltk** function to stop unnecessary words
  • Performing lemmatization,
  • Extract sentences to words,
  • Dictionary to look up words and their frequency,
  • Building a topic model,
  • Compute model and coherence score
  • Judge how good the given topic model is,
  • Coherence score: 0.41,
  • Data visualization.

Results
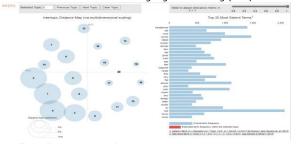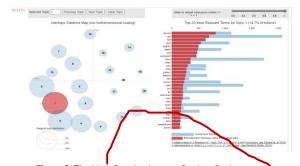
The results of this Natural Language Processing (NLP)

Figure 7: Result Map of Amazon Product Reviews

Figure 8: The Most Popular Amazon Product Reviews

Figure 9: Amazon vs Walmart — Report Highlights Battle over Grocery and Tech in 2022

Discussion

  • Most popular reviews: blue devices, long-time battery tablets, headphones, sound devices.
  • Positive reviews increase sales by 20%
  • Walmart has a bigger net sales than Amazon for everyday basics.
  • Post Covid-19, people love to go to physical stores.
  • Increasing the volume of visits in Walmart.
  • Amazon's sale growth is slowing down.

Conclusion

A positive reviews are strongly converting traffic to product listing and increasing traffic that comes from the search engine result page (SERP) on Amazon. Online sales around the world hitting $2 trillion per year, a product that has just one review is 65% more likely to be purchased than a product that has none, according to Power Reviews. The impact of reviews is immense in sale revennues.

Future Work

I. Analyze Amazon product reviews in physical stores (Wholes Food)
II. Analyze Walmart product reviews and online shipping policy
III. Compare the net sales of both pre and post covid-19 era

References

1. Wu, S. et aL 2022. Web scraping basics
2. Johnson, R. et aL 2020 Amazon Product Review
3. Vega, N n.d Here's Why
4. H, M. Y. et aL 2022 Amazon Product Reviews Dataset.
5.
6. Vibha16 et aL 2022 Topic modelling
7. Z, M. et aL 2022 Amazon vs Walmart