Analysis of Amazon Top 100 Best-Selling Books (2009-2021)

🆔 Fariha Sultana * α

α Department of English, University of Barishal

ABSTRACT

: This project analyzes the Amazon Top 100 Best-Selling Books from 2009 to 2021 to uncover insights into book pricing, ratings, cover types, and genre preferences. Using Microsoft Excel, pivot tables were utilized to answer various analytical questions such as price trends, rating correlations, and the impact of book attributes on rankings. The findings provide valuable insights for readers, publishers, and authors by identifying patterns in consumer preferences and market trends.

5

1 INTRODUCTION

The book industry has experienced significant changes over the years, with e-commerce platforms like Amazon shaping consumer purchasing behavior. Understanding the trends behind best-selling books can offer valuable insights for authors, publishers, and retailers. This project aims to analyze the key factors influencing book sales and rankings, such as price, ratings, cover type, and genre popularity. The study leverages data from Amazon's Top 100 Best-Selling Books from 2009 to 2021 and utilizes Microsoft Excel to perform a comprehensive analysis.

10

2 DATASET

15

The dataset used in this project consists of Amazon's Top 100 Best-Selling Books from 2009 to 2021. It includes the following attributes:

- Book Title
- Author
- Genre
- Cover Type (Hardback, Paperback, Kindle, etc.)
- Price
- Rating
- Number of Reviews
- Year of Release
- Rank in Overall List

20

25

3 METHODOLOGY

3.0.1 *Data Collection*

o The dataset was compiled from various online sources that track Amazon's best-seller lists.

*✉ fariha765@gmail.com

3.1   Pivot Table Analysis

o Several pivot tables were created to answer key analytical questions, including:

Average price by genre

Correlation between price and rating

Rating variation by cover type

Impact of reviews on book ratings

Popularity of genres based on reviews

4   CHALLENGES

- Data inconsistencies and missing values required manual cleaning.

- Some books had multiple editions with different prices and ratings.

- Correlating subjective factors such as ratings with objective measures like price proved to be challenging.

- Limited data for certain years affected trend analysis.

5   RESULTS

The analysis yielded several key insights, including:

- The average price of books varies significantly across genres, with textbooks being the most expensive.

- There is a weak correlation between price and rating, indicating that higher prices do not necessarily translate to better ratings.

- Paperback books tend to have higher ratings compared to hardbacks and e-books.

- The number of reviews is positively correlated with book ratings.

- Fiction genres, particularly mystery and thriller, tend to receive the highest number of reviews.

- Certain authors consistently achieve high average ratings across their books.

6   CONCLUSION

This study provides valuable insights into the factors influencing book sales and popularity on Amazon. The findings suggest that genre, cover type, and price influence a book's success, while ratings and reviews play a crucial role in consumer perception. Future work could expand the analysis by incorporating additional data sources or utilizing advanced analytical tools such as machine learning to predict best-seller trends.

COPYRIGHT NOTICE