EDA for Amazon Books reviews

- Exploratory Data Analysis (EDA) for Amazon Books reviews involves examining and summarizing the dataset to understand its structure, patterns, and insights.

- It is essential for transforming raw data to get actionable insights, driving informed decision-making, and ultimately it improves business outcomes related to book sales and customer satisfaction.

Use Cases

- Most number of books in each genres

- Most frequent Words in over 3 rated Reviews

- Most frequent Words in less than 2 rated Reviews

- Most Reviewed Books

- Highest Rated Books with over 3000 ratings each Book

- Number of reviews and Number of ratings on top 10 books

- In which Genres readers give Positive and Negative Ratings

- Top 10 Authors with 5 star Ratings

Business Use Cases

- Understand Genre Preferences

- Customer Sentiment Analysis

- Identify Areas for Improvement

- Popular Product Identification

- Showcase Quality Products

- Performance Assessment

- Genre-specific Insights
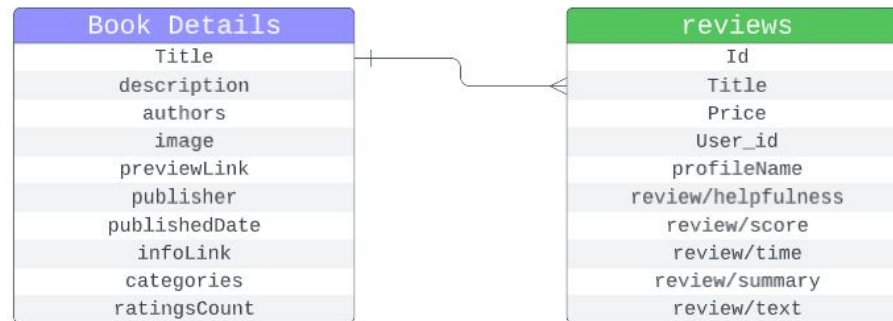
- Author Performance Evaluation

DataSet

- Website link

https://www.kaggle.com/datasets/mohamedbakhet/amazon-books-reviews/data

Data Model

Solution Architecture

- CSV Files -> Python(Jupiter Notebook, PyCharm)

Data Cleaning and Preprocessing

- Handle missing or duplicate data.

- Convert data types if necessary.

- Remove irrelevant columns that are not needed for analysis.