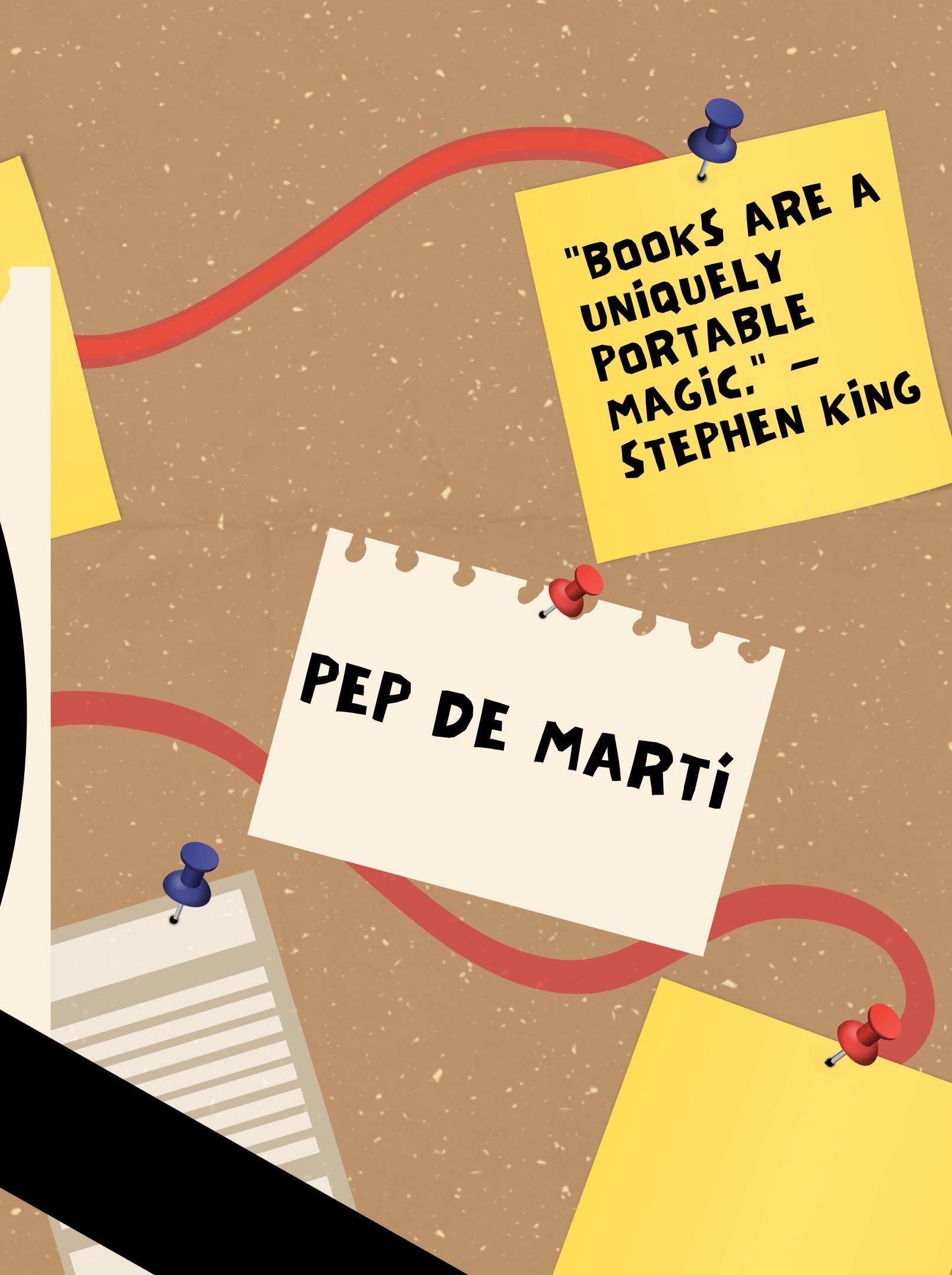




FINAL PROJECT

CREATING A

LIBRARY



PEP DE MARTÍ

**"BOOKS ARE A
UNIQUELY
PORTABLE
MAGIC." —
STEPHEN KING**

Welcome to the ironHack Library

Today, we're uncovering insights from Goodreads data to help build the perfect bookstore! By analyzing book ratings, genres, and reviews, we'll discover which books are the best to stock for success. Let's dive in and crack the code to a thriving bookstore! 🕵️‍♂️🌟

"A room without books is like a body without a soul." - Cicero



Today's Agenda:

DataBase creation

We'll start by looking at how we built a relational database from Goodreads data, organizing key information into separate tables for effective analysis.

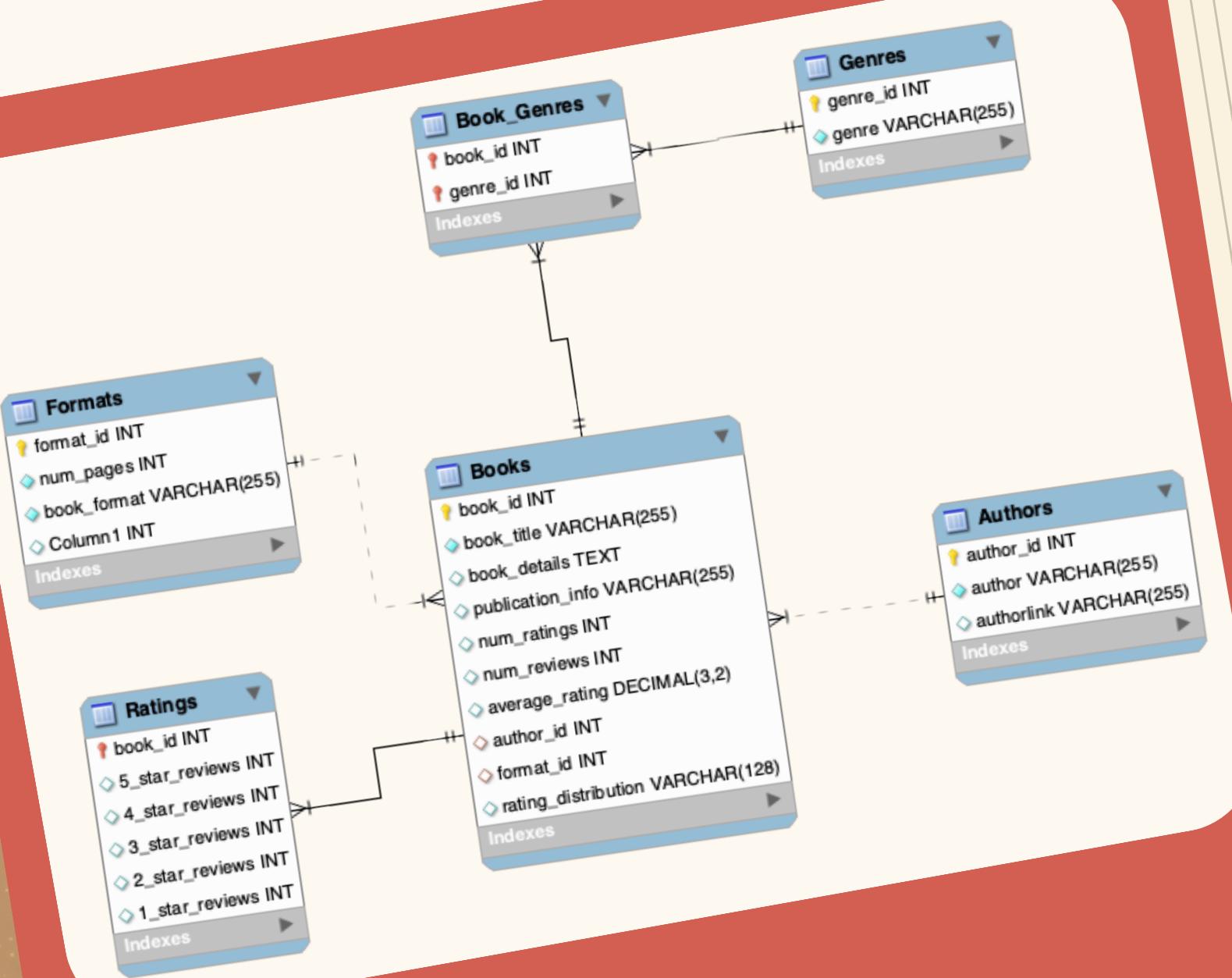
Asking the right questions

Next, we'll discuss the essential business questions we need to answer to help shape our bookstore's strategy and ensure it's stocked with the best titles.

Data Analysis

We'll dive into the analysis process, revealing insights from the data that answer our key questions and guide our business decisions.

Database Diagram



Creating the database

I began by sourcing a comprehensive dataset from Kaggle, containing detailed information about books from Goodreads. This raw data was then systematically organized into multiple related tables, including Authors, Books, Formats, Ratings, and Genres. By creating these interconnected tables, I built a robust relational database that allows for efficient data retrieval and analysis, laying the foundation for our subsequent business insights.

Database creation challenges

Finding the Right Dataframe: It took me three days to find the perfect dataframe on Kaggle, as I couldn't find any useful API's

Creating Foreign Keys: Creating foreign keys posed a significant challenge, as I initially created some that did not correctly align with the primary keys, resulting in an incoherent and inoperable database.

Importing Subtables to MySQL: It took me more than two days to import all data into MySQL due to the inconsistency between Foreign Keys and Primary Keys.

1. Which books have the highest number of 5-star reviews?

Identifying books with the most 5-star reviews helps us understand which titles are most loved by readers, providing key insights into potential bestsellers that should be prominently featured in the bookstore.

5. Which formats are the most in-demand (e.g., hardcover, paperback)?

Understanding the preferred book formats of customers (like paperback vs. hardcover) ensures the bookstore can stock the most popular formats, optimizing inventory and meeting customer expectations.

2. Which books have high average ratings but a low number of reviews (less than 100)?

Books with high ratings but few reviews may be hidden gems. These titles could be excellent candidates for highlighting in marketing campaigns to increase their visibility and attract more readers.

3. Which authors (with more than 10 books) have the highest average ratings?

Knowing which authors consistently produce highly rated books allows us to focus on stocking their works, as they are likely to attract a loyal customer base and drive repeat sales.

4. Which books have the highest average rating?

Highlighting books with the highest average ratings ensures that the bookstore is known for offering top-quality reads, helping to build a reputation for excellence among book lovers.

6. Which genres are the most popular among the highest-rated books?

Popular genres among top-rated books indicate what customers are most interested in. Focusing on these genres can help curate a collection that aligns with current market trends, maximizing customer satisfaction and sales.

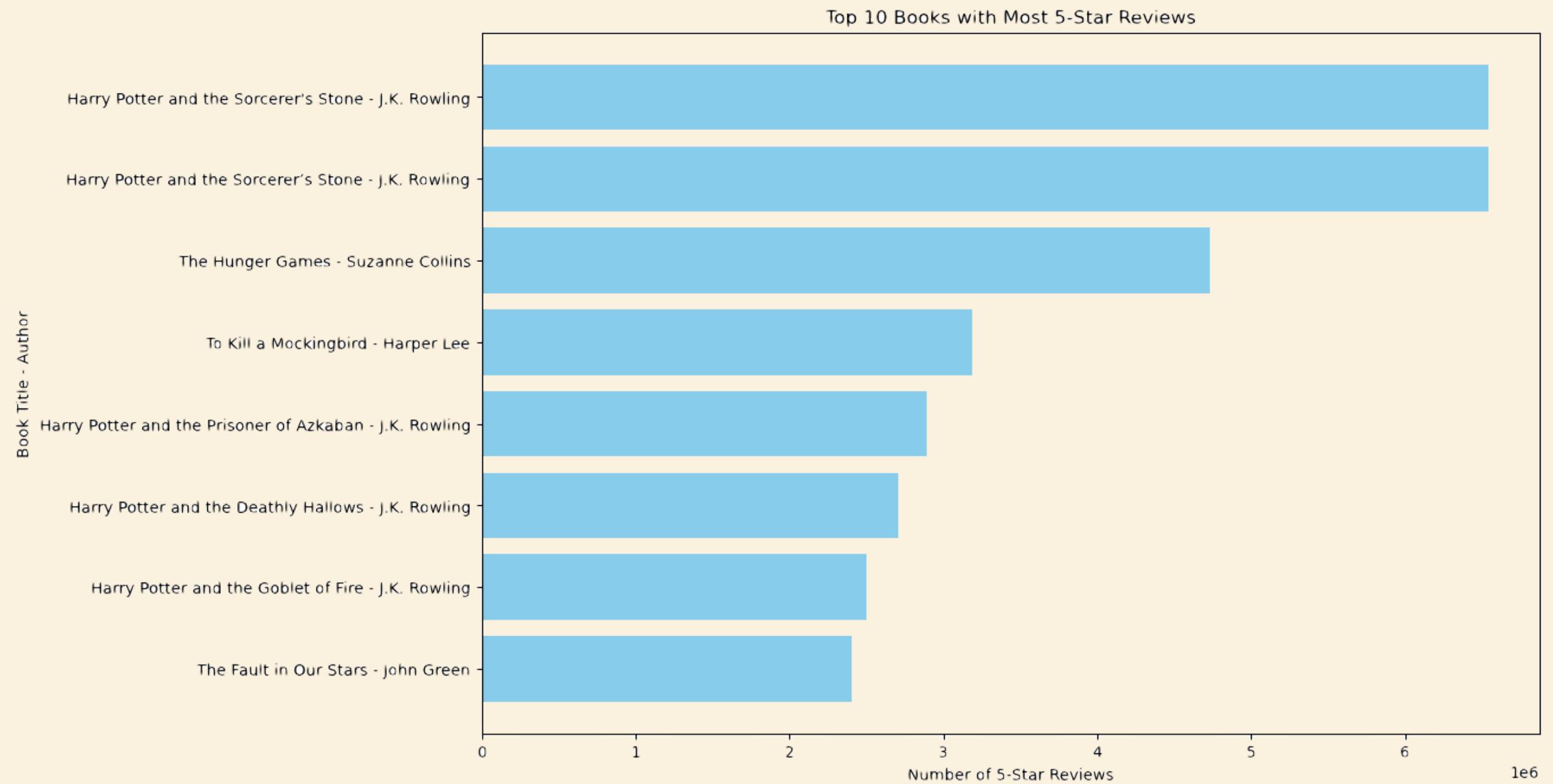
7. Which genres have the highest number of reviews?

Genres with the highest number of reviews typically represent the most engaged and active reader communities. Catering to these genres can enhance customer engagement and foster a vibrant, loyal customer base.

8. Do the number of pages correlate with the Average Rating of the book?

Understanding whether book length influences ratings can help in selecting titles that match reader preferences, whether they lean toward shorter reads or more in-depth, longer works.

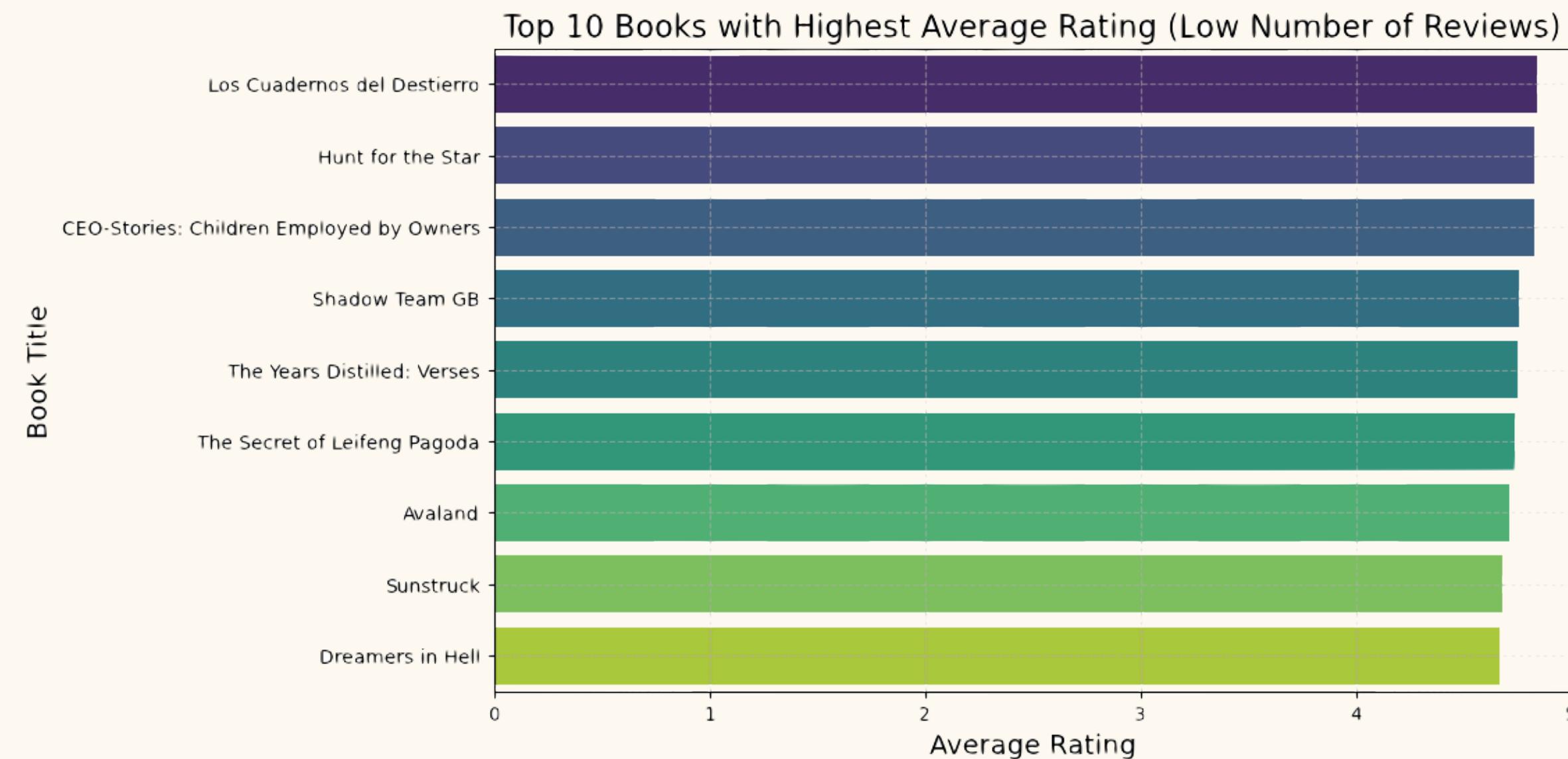
1. Books with the Highest Number of 5-Star Reviews:



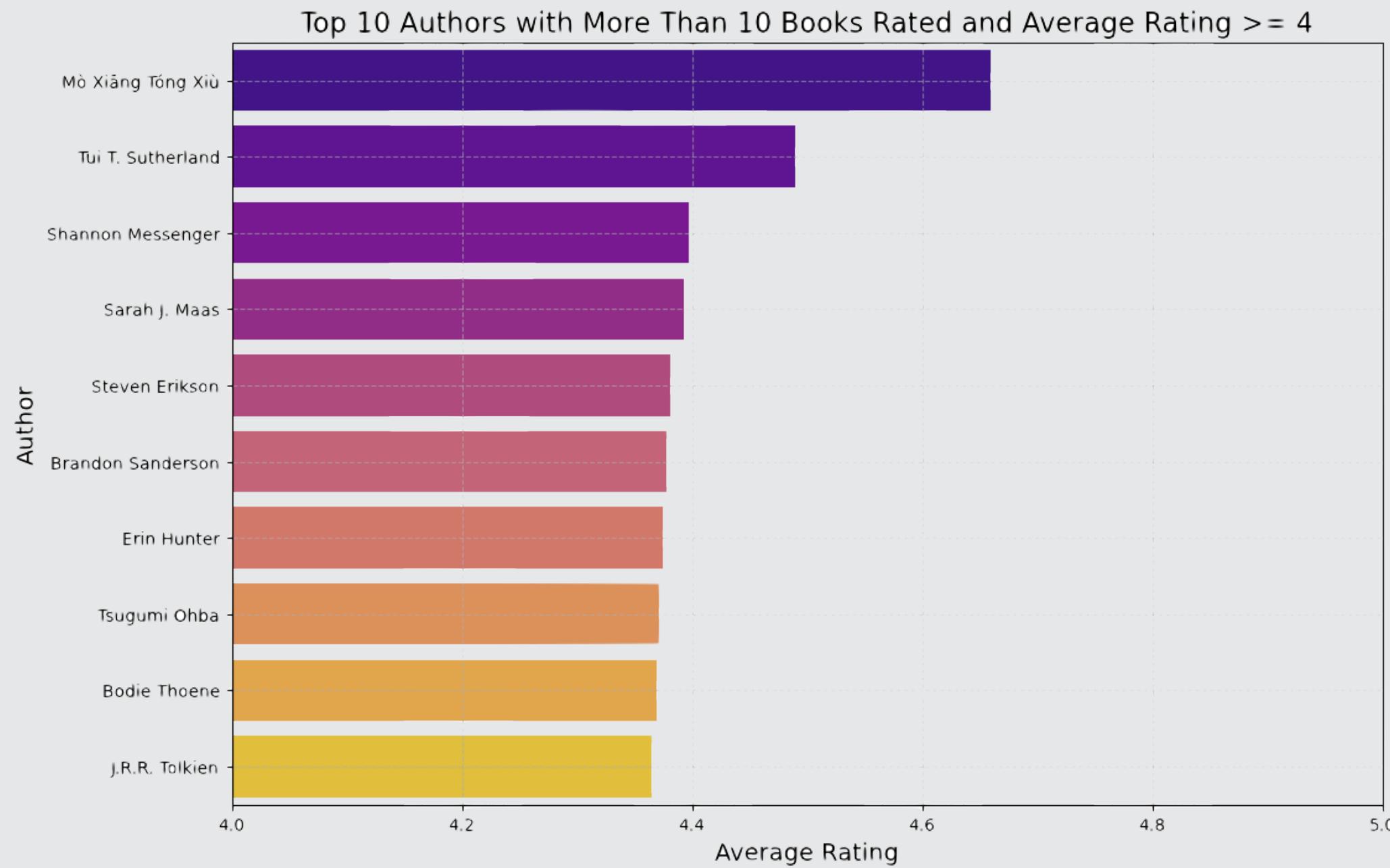
The data reveals that Harry Potter and the Sorcerer's Stone by J.K. Rowling and The Hunger Games by Suzanne Collins are among the books with the highest number of 5-star reviews. These books have a substantial fan base, evident from the millions of 5-star reviews.

2. Books with High Average Ratings but a Low Number of Reviews:

The analysis identifies several books with high average ratings but a relatively low number of reviews. These include *Los Cuadernos del Destierro* and *Hunt for the Star*, which have stellar ratings despite limited visibility.



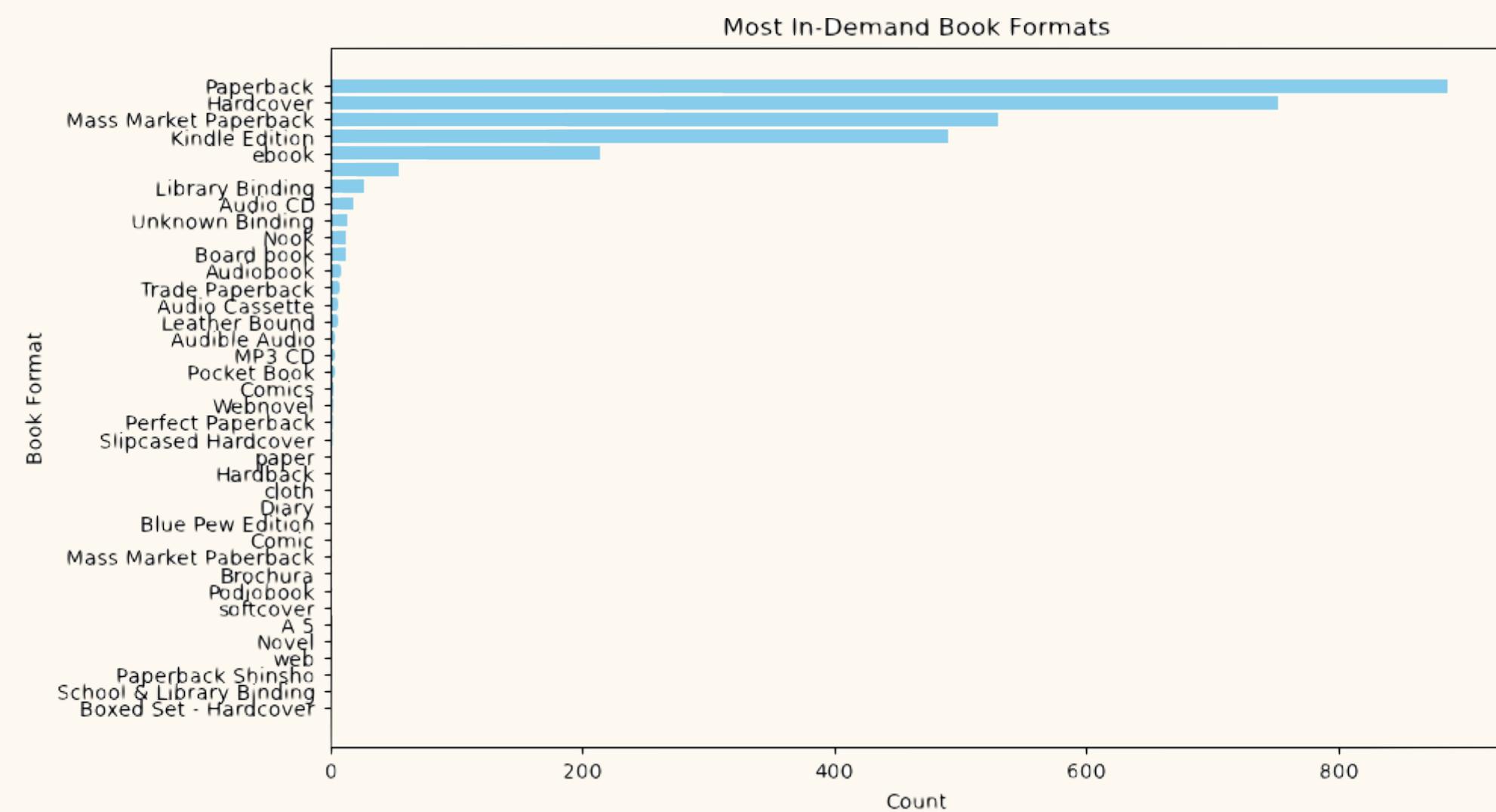
3. Authors with More Than 10 Books Rated and High Average Ratings:



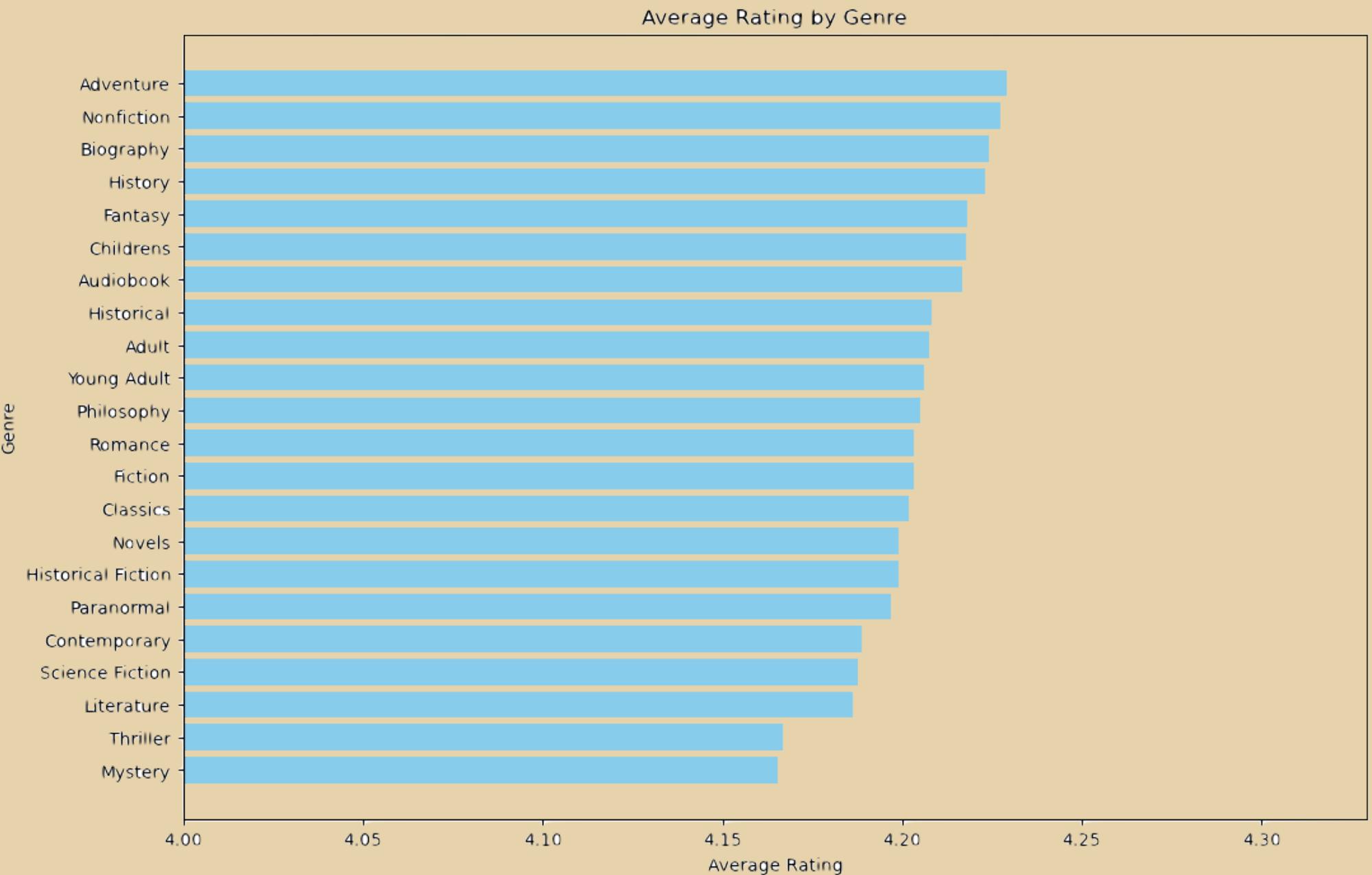
Authors like Mò Xiāng Tóng Xiù and Tui T. Sutherland stand out with high average ratings across a significant number of books. These authors have consistently delivered quality content that resonates with readers.

4. Most in-Demand Book Formats:

The most in-demand formats are Paperback, Hardcover, and Mass Market Paperback, followed by Kindle Edition and eBooks. These formats dominate the market and are preferred by a wide range of readers.



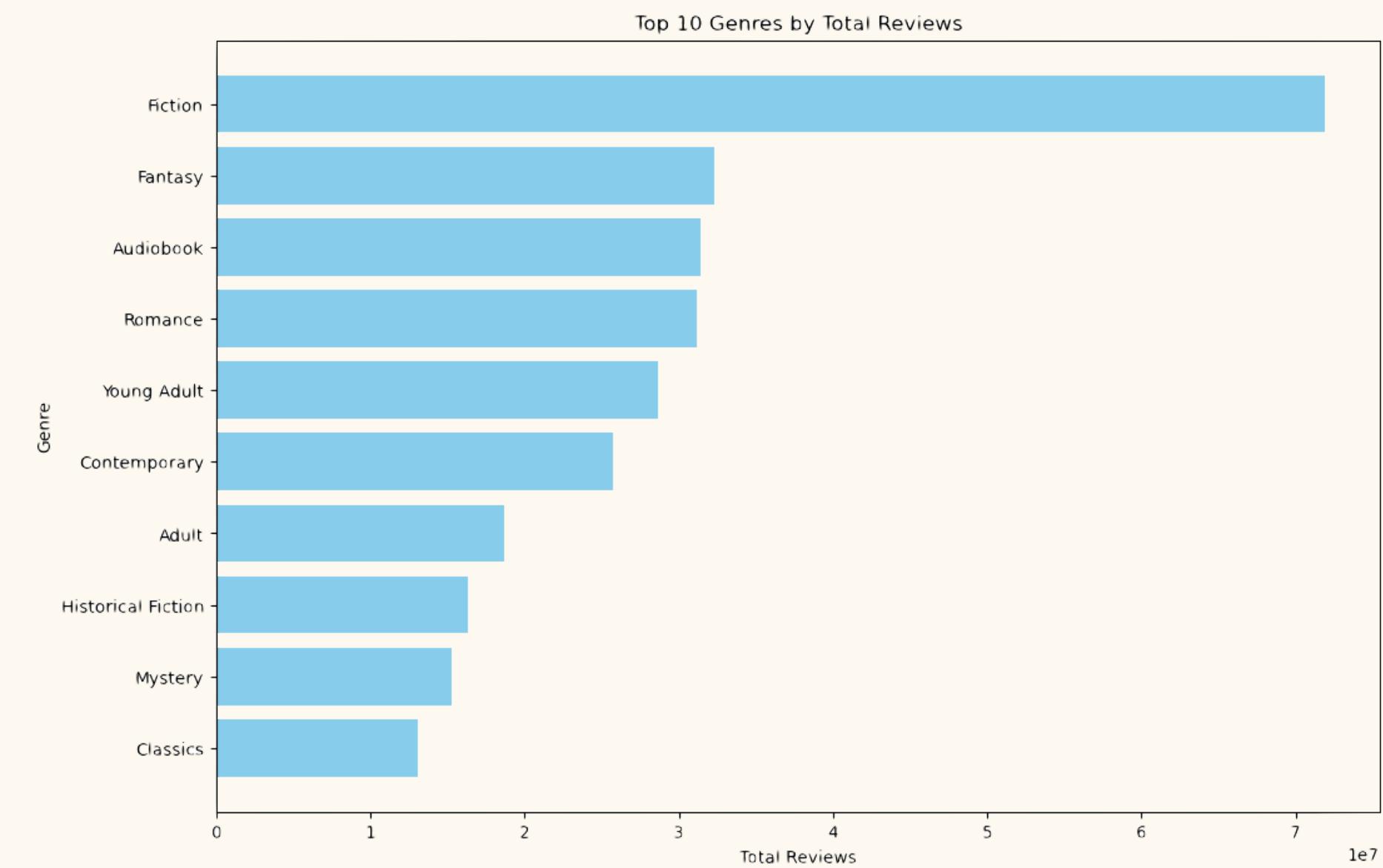
5. Genres Popular Among the Highest-Rated Books:



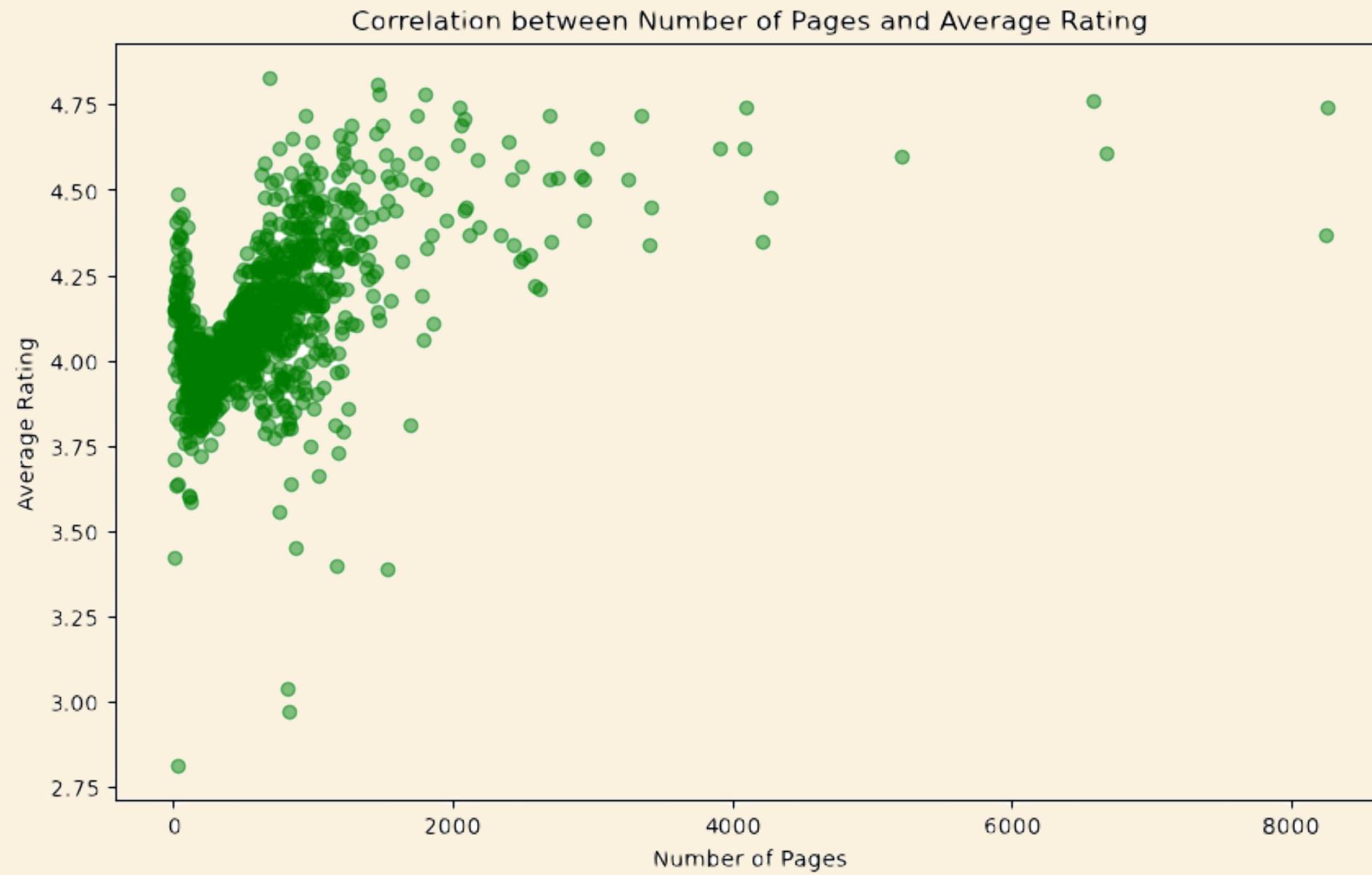
Genres such as Adventure, Nonfiction, and Biography top the list of highest-rated categories. These genres consistently receive high average ratings from readers, indicating their strong appeal.

6. Genres with the Highest Number of Reviews:

Fiction, Fantasy, and Audiobook (though not a genre, it is classified as one by Goodreads) receive the highest number of reviews. These categories not only attract readers but also engage them enough to leave reviews.



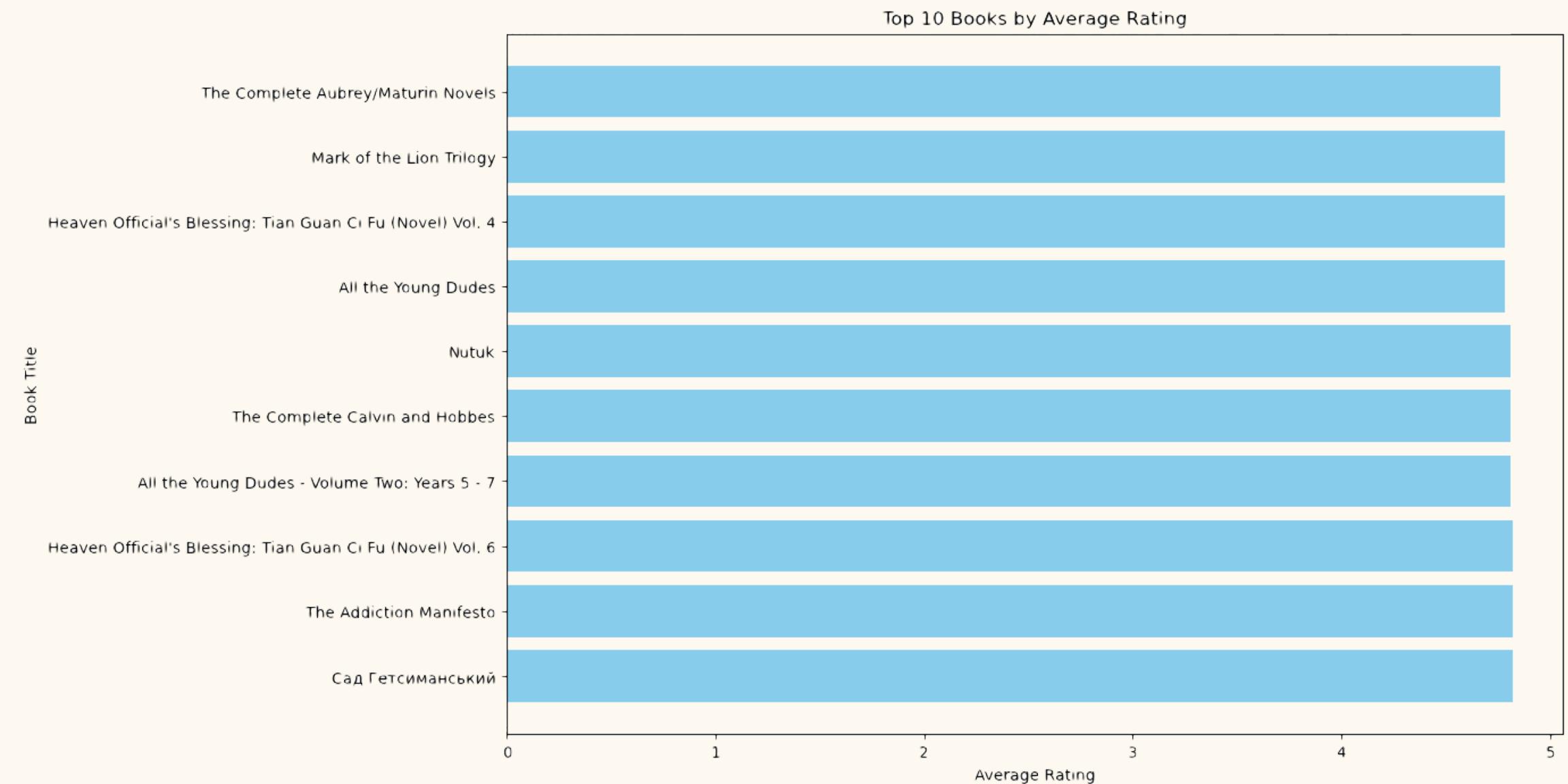
7. Correlation Between Number of Pages and Average Rating:



The correlation analysis shows a poor positive relationship (0.5313) between the number of pages and average rating. Books with a higher page count tend to receive better ratings, although this is not a strict rule.

8. Books with the Highest Average Rating:

The top 10 books with the highest average ratings include titles like The Complete Aubrey/Maturin Novels and Mark of the Lion Trilogy. These books are highly regarded by those who have read them, with ratings close to 5.



Applying the Data to the business:

-Based on the data analysis, our bookstore should prioritize stocking and promoting popular and beloved titles like **Harry Potter** and **The Hunger Games**, which have proven to be bestsellers. We should also spotlight **hidden gems** with high ratings but low visibility, such as **Los Cuadernos del Destierro**, in special promotions or "**Staff Picks**" sections. Additionally, featuring authors with consistently high ratings, like Mò Xiāng Tóng Xiù, through **author-centric displays** or events can attract their loyal fan base. It's essential to ensure that in-demand formats like **Paperback** and **Hardcover** are well-stocked, and to curate a robust collection in top-rated genres such as **Adventure** and **Nonfiction**. Promoting popular genres with the highest review counts, like **Fiction** and **Fantasy**, through bestsellers lists or recommendations will further enhance customer engagement. Finally, highlighting longer, **highly rated books** can appeal to readers who enjoy in-depth stories, and featuring top-rated books in a "**Highly Rated**" section will attract quality-seeking customers.

Any
Questions?

Link_GitHub