

15 exploratory data analysis (EDA) questions that you can use to analyze an Amazon bestseller books dataset: [Amazon Top 50 Bestselling Books 2009 - 2019 | Kaggle](#)

1. How many records are in the dataset?
2. What is the distribution of book genres?
3. Who are the top 5 authors with the most books in the dataset?
4. What is the average rating of the books in the dataset?
5. What is the average price of the books?
6. What is the range of years in which the books were published?
7. What are the top 5 most reviewed books?
8. What is the distribution of book ratings?
9. How many books have a rating greater than 4.5?
10. What is the most common price range for books?
11. What is the average number of reviews for books in each genre?
12. How does the average price vary across different genres?
13. What is the relationship between the number of reviews and the rating of books?
14. What are the top 5 books with the highest ratings and the highest number of reviews?
15. How has the number of bestseller books published per year changed over time?