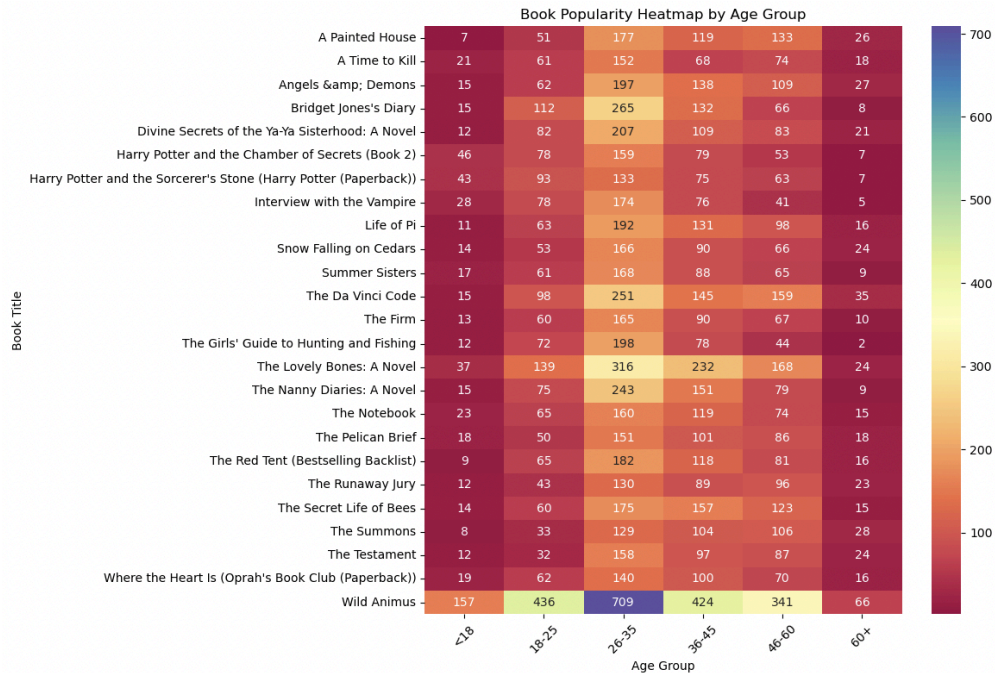


Visualizing Book Popularity Across Age Groups and Authors

By: Anjali Patel

Heatmap highlighting book popularity by age group

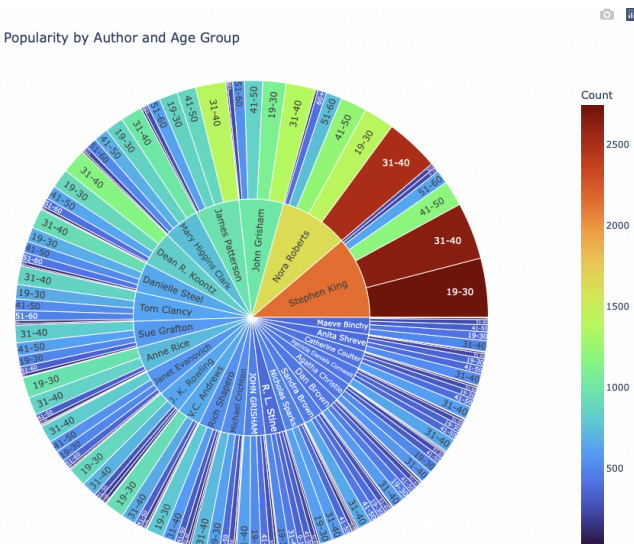


1. Legend

- X-axis: Age Groups
- Y-axis: Book Titles
- Color Intensity: Number of ratings (the darker the shade, the more popular the book within that age group)

I made two sunburst charts. The first one shows the top 25 popular books by author and age group. I thought this looked messy, so I organized it better shown in the second sunburst chart.

Top 25 Book Popularity by Author and Age Group



- d. Despite being crowded, the chart highlighted clusters of authors popular within the same age bracket, supporting genre-based targeting for similar readers.

Data and Method

5. Dataset Source: Kaggle Books Dataset
<https://www.kaggle.com/datasets/saurabhbhagchi/books-dataset>
6. Preprocessing:
 - a. Removed duplicates and missing values
 - b. Categorized users into age groups
 - c. Grouped ratings data by age and book title (for the heatmap) and by age and author (for the sunburst chart)
7. Tools and methods used:
 - a. Python, Pandas for data manipulation
 - b. Seaborn for the heatmap visualization
 - c. Plotly for interactive and clean sunburst chart rendering
 - d. Since the dataset was so big I had to narrow down to top 25
8. Color Tuning:
 - a. Used `plotly.colors.qualitative.Vivid` to make categories visually distinct and intuitive in the sunburst chart

Significance Statement

This project helps us see how different age groups engage with books, which is something that can be really useful for people in publishing, education, or even marketing. Instead of just guessing who likes what, we can actually see patterns—like how the 26–35 age group seems to be the most active when it comes to reading, or how certain authors (like Stephen King or Nora Roberts) attract a wide age range while others are more niche. The heatmap gives a straightforward look at which specific books are popular across age groups, while the sunburst chart goes a level deeper by breaking it down by author and age. That makes it easier to notice things like which authors have broad appeal and which ones are more targeted toward a certain generation. This kind of visualization doesn't just make the data easier to understand—it also shows how we can use data to learn more about people's interests and habits. We can learn from the sunburst chart and heat map to see what genres of books the most popular authors are publishing and be able to recommend books in those genres to people in the corresponding age groups. Additionally, we can see which age groups aren't reading as much, which gives us a chance to dig deeper into why that is and think about how to suggest or create books they'd actually enjoy. Books connect with people on a personal level, but looking at trends like these helps us understand how culture, age, and storytelling come together.