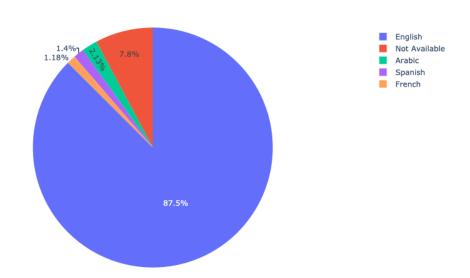
## GOODREADS PROJECT VISUALIZATIONS

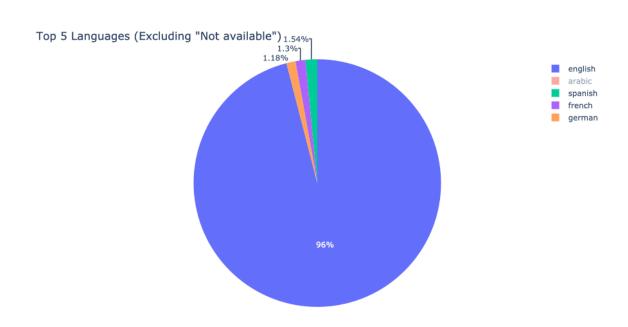
Link for the Dataset - <a href="https://www.kaggle.com/datasets/thedevastator/comprehensive-overview-of-52478-goodreads-best-b">https://www.kaggle.com/datasets/thedevastator/comprehensive-overview-of-52478-goodreads-best-b</a>

#### 1. Top 5 Languages

Top 5 Languages

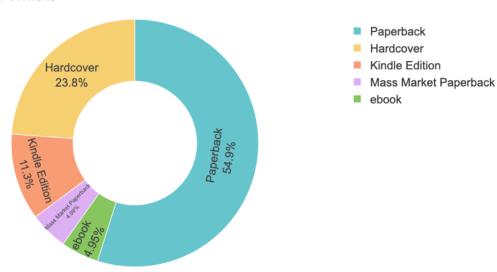


2. Top 5 Languages without Missing values - As "Not Available (Missing Values)" turned out to be second most common one, that was removed from consideration to determine the top 5 languages.

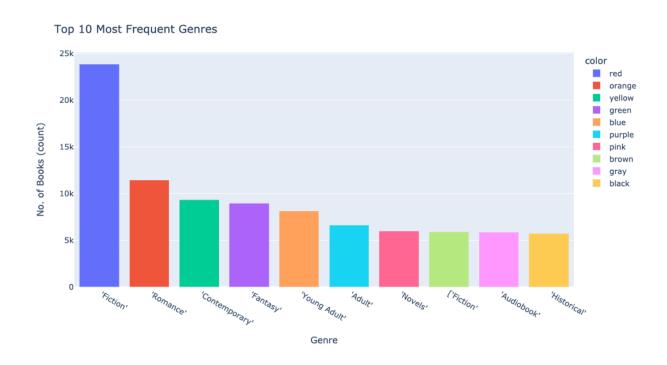


### 3. Top 5 Book Formats.

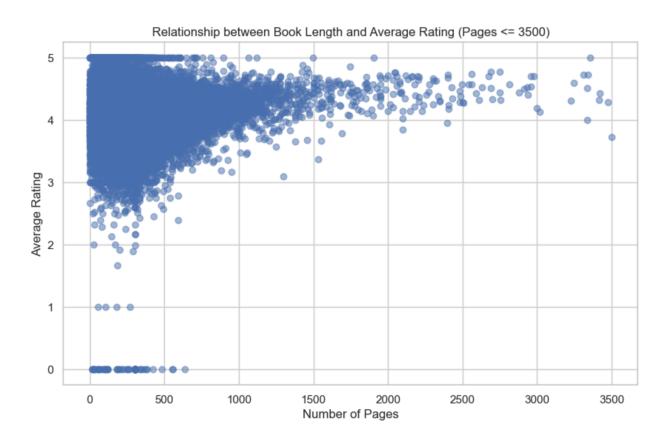
Top 5 Book Formats



### 4. Top 10 Most Frequent Genres.



## 5. Scatterplot between Average Rating and Number of Pages(<=3500).

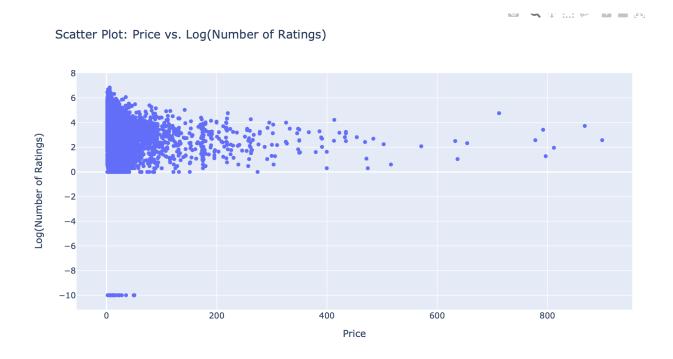


### 6. Correlation Heatmap between numeric variables.

Correlation Heatmap (Including numRatings)

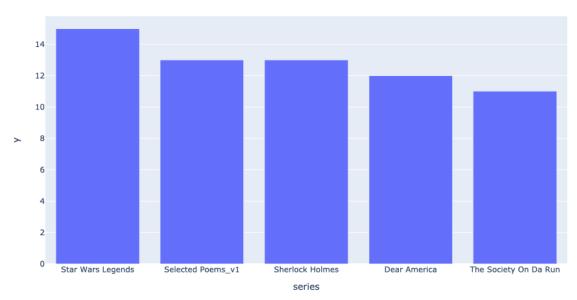
	rating	pages	bbeScore	bbeVotes	price	numRatings	
numRatings	0.01	0.04	0.71	0.72	-0.02	1.0	1
price	0.08	0.09	-0.01	-0.01	1.0	-0.02	0.8
bbeVotes	0.02	0.03	1.0	1.0	-0.01	0.72	0.6
bbeScore	0.02	0.02	1.0	1.0	-0.01	0.71	0.4
pages	0.1	1.0	0.02	0.03	0.09	0.04	0.2
rating	1.0	0.1	0.02	0.02	0.08	0.01	0

## 7. ScatterPlot between Price and Log of Number of Ratings (As Number of Ratings is in Millions too)



## 8. Top 5 Series with most Books.





# 9. Top 10 Publishers with High Average Ratings (95th percentile of numRatings)

Top 10 Publishers with High Average Ratings (95th percentile of numRatings)

