Harry Potter Book Data Analysis

Aly Sewelam

April 2024

1 Pre-processing Steps

In early steps first we chose the columns relevant to our analysis that included at first titles, authors, ratings and lastly total ratings

In early stages of filtering the data, an outlier popped up which was a fan made book that further explained origins of some stories but was not part of the franchise, hence I have implemented further filtering by both name and author

2 Analysis Results

From the analysis on the dataset the top rated book in the series was "Harry Potter and the Sorcerer's Stone" with the most 4 and 5 star ratings, however it has the least `average` rating as "Harry Potter and the Deathly Hallows" has the highest average rating of 4.87

Overall the book franchise has an average rating of 4.82 across all books, though this includes box sets and collections