Matthew Branson

ENG 777-900

Dr. Rhonda Stanton

March 1, 2025

Term Paper Topic: Analyzing the Kaggle Goodreads Dataset

I downloaded the Kaggle Goodreads datasets containing information on tens of thousands of books and millions of book reviews. My premise is to examine sentiment in book reviews, to model trends in ratings, genres, and author popularity, and to perform temporal analyses to see how literary trends evolve over time or in coincidence with world events. Many of the analysis techniques employed will be those directly taught in my data science courses, and I have countless resources demonstrating more advanced approaches. To keep the subject within a reasonable length for the project, I may constrain the scope to focus on answering a single specific question. My working title in that case is: "Anatomy of a Critic: Using Data Science to Identify and Analyze Goodreads' Most Negative Reviewer", which we could humorously subtitle as "Who is this jerk and what is wrong with them?" The feasibility of this focused premise though will depend on the findings of the analysis, but I'm confident in any case that I can find an engaging narrative somewhere in the numbers.