**Group 6 – Final Project**

**Project Topic: Book Recommender System**

**Names of group members:**

Faiza Quadri, Rebecca Lee, Yichen Pan, Zixi Lei, Briana Ramirez, Stephen Rhee

**Introduction:**

We want to use existing reviews and ratings from Goodreads.com to analyze what key words are common in highly rated books’ descriptions to recommend authors to use when writing book descriptions.

**Steps:**

1. We use Web Scraper to extract about 15,000 reviews regarding books tagged “Fantasy” genre from Goodreads.com. The CSV output file retains five columns: book title, description, review\_text, review\_rating, overall\_rating, and page. We saved it as an Excel file: goodreads\_fantasy.xlsx.
2. word frequency analysis in books’ descriptions
3. Similarity analysis with the descriptions
4. Sentiment analysis on the descriptions
5. Logistic Regression on the descriptions
   1. logistic\_regression\_with\_text.py
6. Logistic Regression on the reviews
7. Topic modeling
   1. topic\_modeling\_script.py, we generated top\_quartile\_topics.xlsx.