

Data Appendix

- reviews_cleaned.csv

Each observation (row) of *reviews_cleaned* contains the data associated with a unique Amazon review regarding a specific book

- user_id: unique identifier of reviewer
- profileName: profile name of reviewer
- book_title: unique book identifier, name of book in review
- book_price: price of book in review
- helpfulness: ratio of rated helpful/total rates
- rated_helpful: number of times rated helpful
- rated_total: number of times rated today
- rated_unhelpful: number of times rated unhelpful
- text_word_count: word count of body text of review
- summary_word_count: word count of summary text of review
- vader_sentiment: Vader polarity decimal score (-1 to 1)

	book_price	score	time	rated_helpful	rated_total	rated_unhelpful	text_word_count	summary_word_count	vader_sentiment
count	481171.000000	481171.000000	4.811710e+05	481171.000000	481171.000000	481171.000000	481171.000000	481171.000000	481171.000000
mean	21.762656	4.223733	1.148213e+09	5.931978	8.154669	2.222690	140.128547	4.816207	0.607443
std	26.206541	1.211909	1.359182e+08	17.386096	21.169098	7.888277	164.434314	2.833631	0.547979
min	1.000000	1.000000	8.061984e+08	0.000000	0.000000	-1.000000	1.000000	1.000000	-1.000000
25%	10.780000	4.000000	1.041638e+09	0.000000	0.000000	0.000000	47.000000	3.000000	0.526700
50%	14.930000	5.000000	1.144368e+09	2.000000	3.000000	0.000000	89.000000	4.000000	0.866600
75%	23.950000	5.000000	1.264896e+09	6.000000	8.000000	2.000000	169.000000	6.000000	0.957600
max	995.000000	5.000000	1.362355e+09	2927.000000	3616.000000	689.000000	6810.000000	34.000000	1.000000

- books_cleaned.csv

Each observation (row) of *books_cleaned* contains the data associated with a unique book offered on Amazon

- title: unique identifier of book
- description: descriptive summary of book
- authors: author(s) of book
- previewLink: URL for book preview
- publisher: publishing company of book (contains NULL values)
- publishedDate: year the book was published
- categories: categorical classification of book
- ratingsCount: number of recorded reviews of book

	description	authors	publisher	publishedDate	categories	title
count	45127	49035	43396	49597	47797	49751
unique	40583	31419	5305	7422	2598	49751
top	The final book of the Bible, Revelation prophe...	["Louis L'Amour"]	Simon and Schuster	2004	["Fiction"]	The Church of Christ: A Biblical Ecclesiology ...
freq	22	104	2119	1021	11703	1