Appendix A: Plot theories

We again emphasize that our method of mining emotional arcs from novels does not measure the popular notion of "plot". To make the distinction even clearer, using terms often employed in narratology we consider the common notion of "plot" to be *fabula* whereas the emotional arc is an attempt to measure the emotional trajectory of the *syuzhet*, what could be commonly referred to as the "structure" [1]. For example, the difference between Booker's *Overcoming the monster* and *Rags to riches* may very well have a similar emotional arc, while being distinct plots. Nevertheless, we include our research on the different types of plots that have been enumerated.

There have been various hand-coded attempts to enumerate and classify the core types of stories from their plots, including models that generalize broad categories and detailed classification systems. We consider a range of these theories in turn while noting that plot similarities do not necessitate a concordance of emotional arcs.

- Three plots: In his 1959 book, Foster-Harris contends that there are three basic patterns of plot (extending from the one central pattern of conflict): the happy ending, the unhappy ending, and the tragedy [2]. In these three versions, the outcome of the story hinges on the nature and fortune of a central character: virtuous, selfish, or struck by fate, respectively.
- Seven plots: Often espoused as early as elementary school in the United States, we have the notion that plots revolve around the conflict of an individual with either (1) him or herself, (2) nature, (3) another individual, (4) the environment, (5) technology, (6) the supernatural, or (7) a higher power [3].
- Seven plots: Representing over three decades of work, Christopher Booker's *The Seven Basic Plots: Why we tell stories* describes in great detail seven narrative structures: [4]
    - Overcoming the monster (e.g., *Beowulf*).
    - Rags to riches (e.g., *Cinderella*).
    - The quest (e.g., *King Solomon's Mines*).
    - Voyage and return (e.g., *The Time Machine*).
    - Comedy (e.g., *A Midsummer Night's Dream*).
    - Tragedy (e.g., *Anna Karenina*).
    - Rebirth (e.g., *Beauty and the Beast*).

  In addition to these seven, Booker contends that the unhappy ending of all but the tragedy are also possible.
- Twenty plots: In *20 Master Plots*, Ronald Tobias proposes plots that include "quest", "underdog", "metamorphosis", "ascension", and "descension" [5].
- Thirty-six plots: In a translation by Lucille Ray, Georges Polti attempts to reconstruct the 36 plots that he posits Gozzi originally enumerated [6]. These are quite specific and include "rivalry of kinsmen", "all sacrificed for passion", both involuntary and voluntary "crimes of love" (with many more on this theme), "pursuit", and "falling prey to cruelty of misfortune".

Appendix B: Additional Figures

Here we include additional supporting information.

> The steps, you see, are all the presents the fairy god-mother gave to Cinderella, the ball gown, the slippers, the carriage, and so on. The sudden drop is the stroke of midnight at the ball. Cinderella is in rags again. All the presents have been repossessed. But then the prince finds her and marries her, and she is infinitely happy ever after. She gets all the stuff back, and *then* some. A lot of people think the story is trash, and, on graph paper, it certainly looks like trash.
>
> But then I said to myself, Wait a minute—those steps at the beginning look like the creation myth of virtually every society on earth. And then I saw that the stroke of midnight looked exactly like the unique creation myth in the Old Testament. And then I saw that the rise to bliss at the end was identical with the expectation of redemption as expressed in primitive Christianity.
>
> The tales were identical.
>
> I was thrilled to discover that years ago, and I am just as thrilled today. The apathy of the University of Chicago is repulsive to me.
>
> They can take a flying fuck at the moooooooooooooooon.

Figure S1 Kurt Vonnegut writes in his autobiography *Palm Sunday* on the similarity of certain story shapes [7]. The exposition of this particular similarity would place both of these stories in our grouping of "Rags to Riches" emotional arcs.
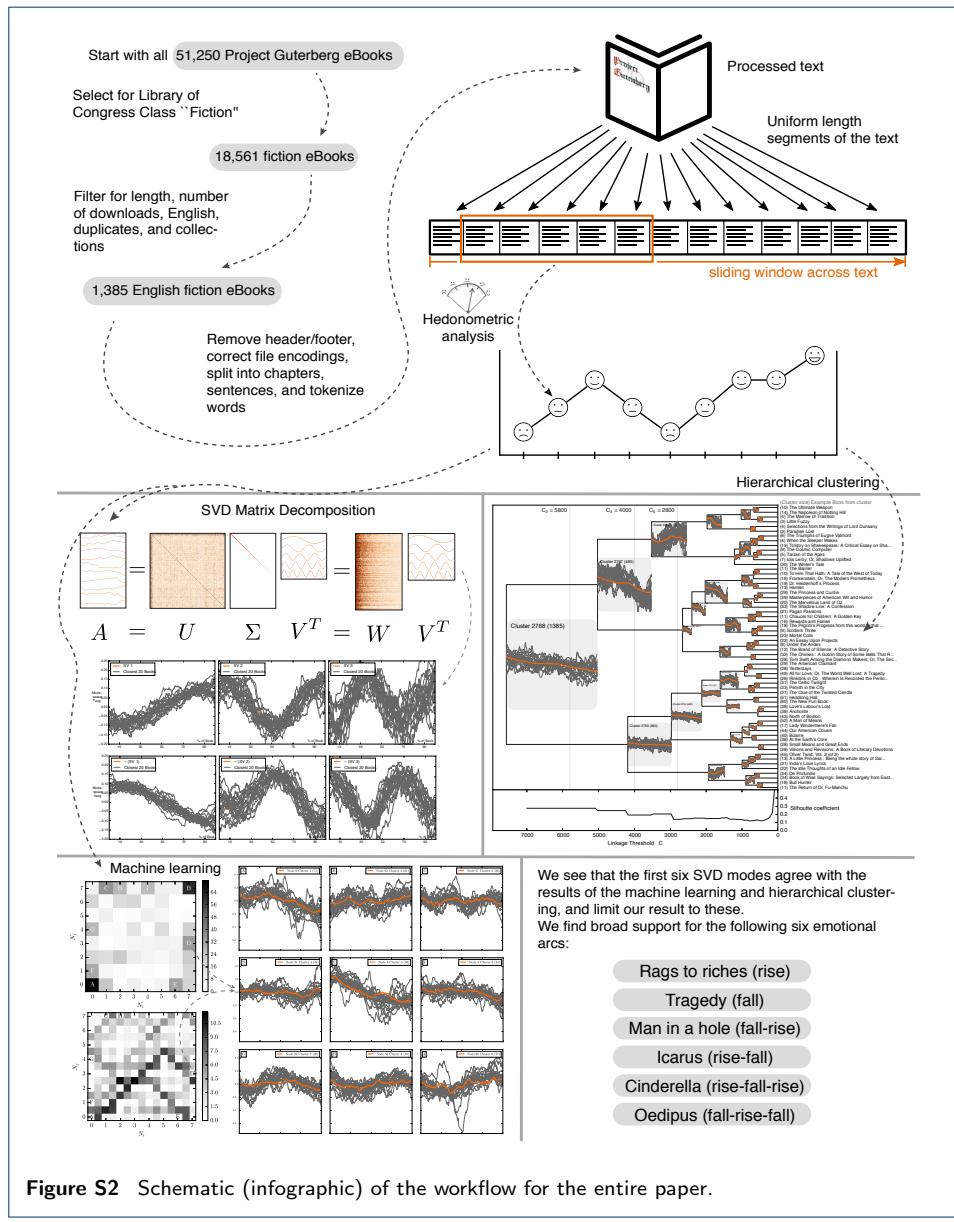
Figure S2 Schematic (infographic) of the workflow for the entire paper.

Mode	Mode Arc	N_m	N_m/N	DL Median ▼	DL Mean ▽	DL Variance	% > Average	Download Distribution
SV 1		511	29.6%	84.0	245.9	376310	16.8%	
-SV 1		289	16.7%	74.0	243.4	1149013	12.5%	
SV 2		157	9.1%	80.0	201.7	105742	16.6%	
-SV 2		234	13.5%	74.0	253.7	619686	14.1%	
SV 3		133	7.7%	88.0	352.3	1374967	18.8%	
-SV 3		110	6.4%	85.5	234.2	391675	14.5%	
SV 4		103	6.0%	103.0	402.2	1313602	22.3%	
-SV 4		37	2.1%	76.0	181.0	130426	10.8%	
SV 5		41	2.4%	85.0	173.8	120762	7.3%	
-SV 5		33	1.9%	82.0	163.0	70769	9.1%	
SV 6		9	0.5%	58.0	65.1	292	0.0%	
-SV 6		17	1.0%	86.0	273.5	234514	29.4%	
SV 7		15	0.9%	90.0	288.7	361492	26.7%	
-SV 7		12	0.7%	196.0	390.4	440533	16.7%	
SV 8		9	0.5%	129.0	124.3	4519	0.0%	

Figure S3 Download statistics for SVD Modes with more than 0.5% of books.

Appendix C: Emotional Arc Construction

To generate emotional arcs, we consider many different approaches with the goal of generating time series that meaningfully reflect the narrative sentiment. In general, we proceed as described in Fig. **??** and consider a method of breaking up the text as having three (interdependent) parameter choices for a sliding window:

1   Length of the desired sample text.
2   Breakpoint between samples.
3   Overlap of each sample.

These methods vary between rating individual words with no overlap to rating the entire text. To make our choice, we consider competing two objectives of time series generation: meaningfulness of sentiment scores and increased temporal resolution of time series. For the most accurate sentiment scores, we can use the entire book. The highest temporal resolution is possible with a sliding window of length 1, generating time series that have potentially as many data points as words in the book.

Since our goal is not only the generation of time series, but the comparison of time series across texts, we consider the additional objective of consistency. We seek time series which are consistent both in the accuracy of the time series, as well as consistent in the length of the resulting time series. Again these goals are orthogonal, and we note that our choice here can be tuned to test the sensitivity.

We normalize the length of emotional arcs for books of different length (while using a fixed window size) by varying the amount that the window needs to move. To make a time series of length l from a book with N words, we fix the sample length at k and set the overlap of samples to

$$(N - k - 1)/l$$

words. This guarantees that we have temporal resolution l and sample length k for any $N > k + l$. We do not consider books with $N \leq k + l$ words.

To generate a sentiment score as in Fig. **??**, we use a dictionary based approach for transparency and understanding of sentiment. We select the LabMT dictionary for robust performance over many corpora and best coverage of word usage. In particular, we determine a sample T's average happiness using the equation:

$$h_{\mathrm{avg}}(T) = \frac{\sum_{i=1}^{N} h_{\mathrm{avg}}(w_i) \cdot f_i(T)}{\sum_{i=1}^{N} f_i(T)} = \sum_{i=1}^{N} h_{\mathrm{avg}}(w_i) \cdot p_i(T), \tag{1}$$

where we denote each of the N words in a given dictionary as w_i, word sentiment scores as $h_{\mathrm{avg}}(w_i)$, word frequency as $f_i(T)$, and normalized frequency of w_i in T as $p_i(T) = f_i(T)/\sum_{i=1}^{N} f_i(T)$.

We note here that, in general, for each emotional arc we subtract the mean before computing the distance or clustering.

C.1 Null emotional arc construction

In our first analysis, we generated the null set of emotional arc time series by randomly shuffling the words in each book. Other variations on generating this null set include sampling from a phrase-level parse of the book with a Markov process,

using continuous space random walks directly, or shuffling on sentences. Even more sophisticated approaches could utilize Recursive Neural Nets (RNNs), for examples see http://karpathy.github.io/2015/05/21/rnn-effectiveness/. For a realistic comparison to meaningful stories, we generate "nonsense" using a Markov chain model (MCM) trained on 2-grams from each book. To contrast these approaches, consider the beginning of one null version of *Alice's Adventures in Wonderland* using randomly shuffled words (the "word salad"):

> *the but little –but all the , with I flowers that small a what the he could queer ran it near , and altogether remain A with somebody , gardeners the thought your I the , door head she me hardly of is were said the - them she Alice I But one you nice large use walked what anything 's and It many I , the , execution , she by of came I witness , turned she upon suddenly took While , if I , hear –well goose mouth , do replied the of play would SAID seem , of business shrink 's she flower if–if the Hare the so Alice of , a - very the hear reason to whispered BEE it thought by I large not , your dream on Herald SOUP she I some to her the all of of of guess Perhaps tell to the answer I– now nibbled , must folded , going himself taught centre wo Northumbria– hanging can you a went a said ! said all cats do before conversation had of jury baby be Run will again three that herself to ! the Alice ! While the girl neighbour very growing they want , across the , whiting round little , with of a indeed went the on tell offended only forgotten to to , tell n't wo see the hardly Turtle What , on , into three him we ten appear you and at I and which thought makes eyes I it and the looked Is n't baby disappeared , an goes of you all talking ; herself that she in bleeds THAT No in - and-butter how I wash , went a 'll way–never kind at with As Dodo , fear officers been off the opening it , said to , removed at said , went muchness–you for and time court what very will to among Queen Turtle things I so Pigeon herself lie me the naturedly the changed never HER , missed but hurry The March–just the said been beak– of , the now whole , Dodo —*

and the null version using a 2-gram MCM (the "nonsense" version):

> *But then , thought Alice to herself , after all –SAID I COULD NOT SWIM– you ca n't go , said the Dormouse began in a minute. And how odd the directions will look. It was the Rabbit 's voice along–'Catch him , I should be like , said the Mouse in the newspapers , at the top of it. The question is , said the Caterpillar. I 'd better ask HER about it. The Queen 's absence , and yet it was n't very civil of you , sooner or later. While she was considering in her life , and that 's a fact. Alice kept her waiting. I ca n't get out of the fact. As for pulling me out of the evening , beautiful Soup. This was such a rule at processions and besides , that finished the first witness , said Alice , and went stamping about , reminding her very much at first but she stopped hastily , for the rest were quite silent for a baby altogether Alice did n't think , said the Queen , who was sitting on a little worried. Sure , it 'll never go THERE again said Alice , who had been to her in such a nice little dog near our house I should say With what porpoise. You do n't seem to put everything upon Bill. And the muscular strength , which remained some time in silence at last she*

spread out her hand in hand , in chains , with the dream of Wonderland of long ago anything had happened. –as far out to be nothing but the great wonder is , said Alice , with their hands and feet at the flowers and the Queen say only yesterday you deserved to be two people. Here the Dormouse said– the Hatter , and , after all it might happen any minute , while the Mock Turtle nine the next witness was the Cat again , to be seen–everything seemed to be sure but I shall be a very long silence after this , as it 's coming down. In THAT direction , the Duchess said to Alice a good deal on where you want to go. Wow wow wow. She 'll get me executed , as the Dormouse go on with the bread and-butter. So they could n't guess of what work it would be like , said the King sharply Do you take me for his housemaid , she pictured to herself , after all. Yes , but it was quite silent for a rabbit. She waited for a minute , nurse. Begin at the house before she had tired herself out with trying , the Queen put on your shoes and stockings for you said the Dodo. How CAN I have n't opened it yet , before Alice could see it trot away quietly into the roof of the Mock Turtle , suddenly dropping his voice , What HAVE you been doing here. It was high time to begin with , the Gryphon added Come , there 's no pleasing them. Alice remained looking thoughtfully at the other , saying to herself , whenever I eat or drink anything so I should think you 'd like it , said the Caterpillar. Ugh said the King.

C.2 Further Gutenberg Processing

Here we provide the details of the processing applied to the Gutenberg corpus. In the manuscript, we stated the following:

> We start by selecting for only English books, with total words between 20,000 and 100,000, with more than 20 downloads from the Project Gutenberg website, and with Library of Congress Class PN, PR, PS, or PZ. Next, we remove books with any word in the title from a list of keywords (e.g. "poems" and "collection"). From within this set of books, we remove the front and back-matter of each book using regular expression pattern matches.

The full list of keywords which we used to filter the titles are the following: "stories", "collection", "poems", "complete", "essays", "fables", "tales", "papers", "poetry", "verses", "ballads", "sketches", "vol.", "vols.", "works", "volume", and "other". A list of of LoC Classes is given at https://www.loc.gov/catdir/cpso/lcco/.

To remove the front matter, we first detect the end of the front matter by matching for either `START OF THIS PROJECT GUTENBERG EBOOK` in the line or `START OF THE PROJECT GUTENBERG EBOOK`. If neither of these work, we look for a line that contains both `END` and `SMALL PRINT` in the line, in the first half of the text.

To remove the back matter, we check for three different endings, in order. First, similar to the front matter we check, here without being sensitive for case, for `END OF THIS PROJECT GUTENBERG EBOOK` or `END OF THE PROJECT GUTENBERG EBOOK` or `END OF PROJECT GUTENBERG`. Next, we check the last 25% of the book, case insensitive, for the words `END` and `PROJECT GUTENBERG`. Finally, we check the last 10% of the book for the words, case sensitively, `THE END`.

Figure S4 Rank-frequency distributions of book downloads and length in the Gutenberg corpus: (A) downloads, (B) book length in words, and (C) both downloads and length. We filter by both number of downloads and book length to select for fiction books, with the filters shown as gray boxes in Panels A and B. In Panel C, we plot each of 1,748 books selected by their download count and length, shown in download-length space.

Together, these filters each remove text from the beginning and end of 98.9% of ebooks. The first pass in each case works for 78.9% of cases. On average, this removes less than 1% of the beginning lines, and 3-4% of the ending lines.

Appendix D: Book list

We include a list of all books used in this study with more than 40 downloads from
Project Gutenberg, such that we list those from all of the experiments with 40 and
80 download thresholds in the following Table. We do not include the full list of
books with more than 10 downloads for brevity, as it is more than 90 pages long
(this list is 22 pages).

Table S1: All Project Gutenberg eBooks considered in this study,
sorted by downloads.

ID	Title	DL's	Mode	N_w	Uniq(N_w)	$h_{\mathrm{avg}}^{b_i}$	LoCC
11	Alice's Adventures in Wonderland	17,366	SV5	27,386	3,973	5.99	PR,PZ
84	Frankenstein; Or, The Modern Pro...	11,699	-SV3	77,680	9,691	5.89	PR
74	The Adventures of Tom Sawyer	9,454	SV4	73,265	11,094	5.76	PZ,PS
844	The Importance of Being Earnest:...	9,373	SV1	23,161	3,612	6.15	PR
174	The Picture of Dorian Gray	7,652	-SV1	84,591	9,588	5.96	PR
16	Peter Pan	5,789	SV4	48,189	6,747	5.89	PZ,PR
16328	Beowulf : An Anglo-Saxon Epic Poem	5,359	-SV1	44,949	9,351	5.64	PR
42	The Strange Case of Dr. Jekyll a...	4,908	-SV1	26,085	5,300	5.63	PR
2814	Dubliners	4,742	SV2	68,963	9,844	6.11	PR
46	A Christmas Carol in Prose; Bein...	4,602	-SV2	29,192	5,734	6.08	PR
244	A Study in Scarlet	4,535	-SV3	44,537	7,743	5.83	PR
120	Treasure Island	4,402	SV3	70,261	8,867	5.89	PZ,PR
526	Heart of Darkness	4,362	-SV1	38,504	7,683	5.77	PR
35	The Time Machine	3,732	-SV1	32,622	6,171	5.87	PR
3825	Pygmalion	3,580	-SV1	34,898	5,864	6.08	PR
236	The Jungle Book	3,478	SV1	52,449	6,967	5.70	PR
2852	The Hound of the Baskervilles	3,358	-SV1	60,070	7,655	5.90	PR
219	Heart of Darkness	3,243	-SV1	38,464	7,662	5.76	PR
863	The Mysterious Affair at Styles	3,112	-SV3	57,720	7,725	5.83	PR
33	The Scarlet Letter	3,045	SV5	87,213	12,462	6.03	PS
55	The Wonderful Wizard of Oz	3,035	SV2	40,939	4,217	6.10	PZ,PS
4517	Ethan Frome	2,895	SV4	35,704	5,854	5.94	PS
12	Through the Looking-Glass	2,892	SV2	30,775	4,474	5.98	PZ,PR
28520	Forbidden Fruit: Luscious and exc...	2,716	-SV1	32,669	4,726	6.25	PR
105	Persuasion	2,535	-SV7	86,532	8,279	6.20	PR
20	Paradise Lost	2,522	-SV2	91,206	15,388	5.61	PR
62	A Princess of Mars	2,515	-SV2	68,970	8,731	5.90	PS
36	The War of the Worlds	2,496	-SV3	62,729	9,447	5.67	PR
215	The Call of the Wild	2,439	SV1	32,356	6,245	5.68	PS
121	Northanger Abbey	2,355	SV4	77,944	8,806	6.16	PR
1524	Hamlet, Prince of Denmark	2,329	-SV1	34,265	7,281	5.95	PR
2097	The Sign of the Four	2,283	-SV3	45,443	7,226	5.85	PR
25305	Memoirs Of Fanny Hill: A New and ...	2,222	-SV5	85,189	10,715	6.13	PR
209	The Turn of the Screw	2,175	-SV3	42,852	6,642	5.84	PS
4217	A Portrait of the Artist as a Yo...	2,172	-SV4	86,019	12,486	5.87	PR

Table S1: All Project Gutenberg eBooks considered in this study, sorted by downloads.

ID	Title	DL's	Mode	N_w	Uniq(N_w)	$h_{\mathrm{avg}}^{b_i}$	LoCC
834	The Memoirs of Sherlock Holmes	2,164	SV2	88,841	10,013	5.88	PR
779	The Tragical History of Doctor F...	2,133	SV1	22,025	5,471	5.86	PR
1155	The Secret Adversary	2,070	-SV1	77,875	10,087	5.90	PR
308	Three Men in a Boat	2,059	-SV2	69,574	9,679	5.99	PR
696	The Castle of Otranto	1,663	SV3	37,999	6,274	5.92	PR
8492	The King in Yellow	1,504	SV4	72,731	11,488	5.83	PS
5131	Childe Harold's Pilgrimage	1,481	-SV2	42,394	9,379	5.72	PR
289	The Wind in the Willows	1,475	-SV3	60,301	9,353	6.05	PR,PZ
10007	Carmilla	1,416	-SV3	28,220	5,418	6.02	PR
1837	The Prince and the Pauper	1,389	SV2	72,781	12,043	5.78	PS
35997	The Jungle Book	1,370	SV1	53,872	7,495	5.71	PR
2641	A Room with a View	1,354	SV6	67,923	9,948	6.03	PR
8164	My Man Jeeves	1,317	-SV1	52,792	7,061	6.11	PR
885	An Ideal Husband	1,303	SV1	34,378	4,773	6.11	PR
447	Maggie: A Girl of the Streets	1,295	SV2	24,520	5,347	5.61	PS
139	The Lost World	1,274	-SV1	79,892	10,986	5.85	PR
78	Tarzan of the Apes	1,272	SV1	87,882	10,347	5.76	PS
972	The Devil's Dictionary	1,257	-SV2	65,791	15,417	5.73	PS
3289	The Valley of Fear	1,228	-SV2	58,672	7,929	5.79	PR
910	White Fang	1,178	SV4	76,867	9,199	5.69	PS
73	The Red Badge of Courage: An Epi...	1,163	-SV3	49,153	8,135	5.43	PS
113	The Secret Garden	1,153	SV2	85,604	7,267	6.12	PS,PZ
974	The Secret Agent: A Simple Tale	1,142	SV6	95,448	12,480	5.86	PR
102	The Tragedy of Pudd'nhead Wilson	1,140	-SV1	54,631	9,160	5.90	PS
421	Kidnapped	1,132	SV2	83,335	9,400	5.82	PR,PZ
131	The Pilgrim's Progress from this...	1,126	-SV2	60,264	7,033	5.83	PR
805	This Side of Paradise	1,122	-SV1	84,291	14,149	6.05	PS
208	Daisy Miller: A Study	1,101	-SV1	21,859	3,863	6.26	PS
28522	Laura Middleton; Her Brother and...	1,097	-SV2	32,282	4,226	6.24	PR
159	The Island of Doctor Moreau	1,083	-SV1	45,521	7,386	5.72	PR
26654	Peter and Wendy	1,068	-SV2	48,096	6,753	5.90	PZ,PR
2265	Hamlet	1,051	-SV1	31,656	6,829	5.72	PR
28885	Alice's Adventures in Wonderland...	1,051	SV5	28,410	4,283	6.00	PR,PZ
32	Herland	1,013	-SV3	52,978	8,783	6.31	PS
5230	The Invisible Man: A Grotesque R...	1,011	SV2	49,621	8,320	5.73	PR
558	The Thirty-Nine Steps	989	-SV1	43,658	7,052	5.95	PR
15272	Spenser's The Faerie Queene, Book I	978	-SV7	81,519	16,466	5.70	PR
383	She Stoops to Conquer; Or, The M...	903	-SV3	25,517	4,286	6.13	PR
10554	Right Ho, Jeeves	896	-SV3	75,766	10,348	6.00	PR
946	Lady Susan	894	-SV1	23,259	3,870	6.23	PR
500	The Adventures of Pinocchio	863	SV1	40,459	5,110	5.82	PQ,PZ
242	My Antonia	847	SV4	83,178	10,229	6.24	PS

Table S1: All Project Gutenberg eBooks considered in this study,
sorted by downloads.

ID	Title	DL's	Mode	N_w	Uniq(N_w)	$h^{b_i}_{\mathrm{avg}}$	LoCC
1041	Shakespeare's Sonnets	831	-SV1	20,305	4,357	5.90	PR
51	Anne of the Island	826	SV2	79,609	10,730	6.11	PZ
146	A Little Princess : Being the who...	825	SV4	69,446	7,440	5.93	PS,PZ
389	The Great God Pan	807	-SV3	22,637	4,281	5.82	PR
269	Beasts and Super-Beasts	804	SV1	64,396	11,011	6.08	PR
47	Anne of Avonlea	803	SV4	92,180	10,297	6.16	PZ
204	The Innocence of Father Brown	800	-SV4	80,292	10,947	5.82	PR
1695	The Man Who Was Thursday: A Nigh...	796	SV4	58,887	8,704	5.75	PR
2166	King Solomon's Mines	788	SV3	83,364	10,303	5.73	PR
41445	Frankenstein; Or, The Modern Pro...	786	-SV3	73,078	9,449	5.90	PR
271	Black Beauty	780	SV5	61,002	5,793	5.89	PZ,PR
550	Silas Marner	780	SV1	75,026	9,619	6.09	PR
1097	Mrs. Warren's Profession	780	-SV2	35,689	6,435	6.12	PR
2267	Othello	760	-SV1	29,535	5,856	5.87	PR
4081	The Alchemist	744	-SV1	54,042	12,897	6.00	PR
26	Paradise Lost	730	-SV4	81,693	14,098	5.79	PR
854	A Woman of No Importance	729	-SV2	24,756	3,636	6.14	PR
981	Beowulf	718	-SV1	27,044	5,645	5.52	PR
17396	The Secret Garden	716	SV2	83,089	7,399	6.11	PZ,PS
60	The Scarlet Pimpernel	710	-SV1	85,417	10,611	5.93	PR
624	Looking Backward, 2000 to 1887	679	-SV6	78,379	10,209	6.13	PS
29827	The Life and Amours of the Beaut...	678	-SV1	39,640	5,140	6.30	PS
19994	The Aesop for Children : With pic...	676	SV1	27,975	4,931	5.89	PZ
1280	Spoon River Anthology	671	SV3	36,193	7,558	5.94	PS
2040	Confessions of an English Opium-...	643	-SV2	39,898	7,759	6.02	PR
28521	The Power of Mesmerism: A Highly ...	643	SV1	25,545	4,825	6.18	PR
20781	Heidi: (Gift Edition)	642	-SV3	52,684	5,897	6.19	PZ
12753	The Legends of King Arthur and H...	640	-SV2	94,291	7,306	5.74	PN
64	The Gods of Mars	628	-SV4	85,182	9,416	5.74	PS
1091	On Heroes, Hero-Worship, and the...	622	-SV1	89,835	13,396	5.92	PR
19337	A Christmas Carol	622	-SV2	29,957	5,921	6.08	PR
32154	The Variable Man	618	-SV1	25,869	5,073	5.60	PS
1212	Love and Freindship [sic]	611	SV3	33,532	6,140	6.08	PR
43	The Strange Case of Dr. Jekyll a...	599	-SV1	25,883	5,142	5.61	PR
544	Anne's House of Dreams	586	-SV5	85,952	10,357	6.10	PZ
708	The Princess and the Goblin	579	SV3	53,567	5,857	6.12	PZ
68	Warlord of Mars	571	-SV2	58,345	7,746	5.67	PS
4078	The Picture of Dorian Gray	565	-SV1	58,448	7,626	6.01	PR
223	The Wisdom of Father Brown	563	-SV3	73,134	10,747	5.85	PR
10002	The House on the Borderland	563	SV1	51,289	7,380	5.77	PR
4039	Volpone; Or, The Fox	558	-SV1	54,335	13,090	5.97	PR
3070	The Hound of the Baskervilles	549	-SV1	59,943	7,682	5.89	PR

Table S1: All Project Gutenberg eBooks considered in this study,
sorted by downloads.

ID	Title	DL's	Mode	N_w	Uniq(N_w)	$h^{b_i}_{\mathrm{avg}}$	LoCC
1128	The Tragedy of King Lear	548	-SV1	32,080	6,036	5.72	PR
27805	The Wind in the Willows	543	-SV3	60,665	9,464	6.06	PZ,PR
3618	Arms and the Man	536	-SV2	25,451	4,952	5.96	PR
2232	The Duchess of Malfi	534	-SV1	32,036	6,006	5.75	PR
157	Daddy-Long-Legs	531	-SV2	40,321	6,918	6.24	PZ,PS
17157	Gulliver's Travels into Several ...	528	SV1	52,750	8,404	6.09	PR,PZ
11012	The Autobiography of an Ex-Color...	509	-SV1	52,360	7,970	6.21	PS
1164	The Iron Heel	506	SV3	90,738	12,274	5.71	PS
2243	The Merchant of Venice	506	SV3	23,454	4,864	5.99	PR
501	The Story of Doctor Dolittle	504	SV1	27,696	4,151	5.97	PZ
3188	Mark Twain's Speeches	500	-SV9	94,816	12,619	6.10	PS
1120	The Tragedy of Julius Caesar	496	-SV1	24,322	4,300	5.69	PR
610	Idylls of the King	494	-SV6	96,537	11,145	5.85	PR
790	Lady Windermere's Fan	485	-SV1	22,328	3,538	5.99	PR
11505	All Things Considered	485	-SV1	60,097	9,117	5.89	PR
1094	Tamburlaine the Great — Part 1	474	SV2	27,873	6,206	5.81	PR
19860	The Arabian Nights Entertainments	470	-SV2	87,593	8,945	6.13	PZ
325	Phantastes: A Faerie Romance for...	461	-SV1	70,798	9,566	6.07	PR
376	A Journal of the Plague Year : Wr...	461	-SV2	96,133	8,546	5.52	PR
10150	Dracula's Guest	456	SV1	57,947	8,634	5.79	PR
1129	The Tragedy of Macbeth	449	-SV3	21,587	4,723	5.75	PR
72	Thuvia, Maid of Mars	437	-SV2	49,059	7,305	5.70	PS
1874	The Railway Children	437	-SV3	61,948	7,365	5.96	PZ,PR
2775	The Good Soldier	426	SV5	76,278	9,620	6.02	PR
1872	The Red House Mystery	422	SV5	62,303	6,404	5.93	PR
54	The Marvelous Land of Oz	419	-SV2	43,671	6,605	5.99	PZ
1929	The School for Scandal	417	SV1	30,021	6,884	6.01	PR
3790	Major Barbara	416	-SV3	33,481	5,977	6.11	PR
1153	The Chessmen of Mars	409	-SV1	89,479	9,751	5.72	PS
24761	The Rivals: A Comedy	408	-SV1	28,845	6,236	6.13	PR
5670	Jacob's Room	403	-SV4	55,534	10,376	6.02	PR
3011	The Lady of the Lake	399	SV1	85,874	15,680	5.85	PR
470	Heretics	395	-SV2	66,257	9,337	6.06	PR
1719	The Ballad of the White Horse	394	-SV1	20,388	4,195	5.58	PR
811	The Tragical History of Doctor F...	389	SV4	25,326	5,250	5.82	PR
6043	The Spanish Tragedie	389	SV1	27,164	5,633	5.66	PR
25344	The Scarlet Letter	386	SV5	85,248	12,645	6.03	PS
420	Dorothy and the Wizard in Oz	385	SV2	43,815	5,820	6.06	PZ
81	The Return of Tarzan	384	SV4	92,959	10,316	5.74	PZ,PS
1103	King Richard III	384	-SV2	36,552	6,082	5.78	PR
2042	Something New	384	-SV1	76,340	10,612	6.00	PR
1107	The Taming of the Shrew	383	-SV1	25,876	4,698	6.12	PR

Table S1: All Project Gutenberg eBooks considered in this study, sorted by downloads.

ID	Title	DL's	Mode	N_w	Uniq(N_w)	$h^{b_i}_{\text{avg}}$	LoCC
5348	Ragged Dick, Or, Street Life in ...	378	SV1	50,048	6,516	6.16	PS,PZ
24	O Pioneers!	371	-SV2	56,862	7,464	6.17	PS
1292	The Way of the World	367	-SV2	31,158	5,659	6.15	PR
4352	Laughter: An Essay on the Meanin...	365	-SV4	40,090	6,975	6.33	PN
25016	The House of Souls	362	-SV1	88,028	10,431	5.99	PR
1787	Hamlet	361	-SV1	37,349	6,676	5.96	PR
1121	As You Like It	355	SV1	26,406	4,643	6.14	PR
222	The Moon and Sixpence	352	SV2	79,148	9,300	6.06	PR
8092	Tremendous Trifles	352	-SV1	56,615	9,232	5.94	PR
24737	The Children of Odin: The Book o...	352	-SV1	66,511	6,187	6.03	PZ,BL
1951	The Coming Race	350	SV2	53,105	9,266	6.11	PR
1450	Pollyanna	349	-SV3	58,189	7,049	6.10	PS,PZ
1013	The First Men in the Moon	348	SV2	69,083	10,116	5.92	PR
1051	Sartor Resartus: The Life and Op...	347	SV5	82,042	16,731	5.92	PR
1059	The World Set Free	343	SV2	65,705	11,487	5.87	PR
95	The Prisoner of Zenda	339	-SV1	54,794	7,497	5.88	PR
171	Charlotte Temple	337	-SV1	37,184	6,112	6.10	PS
1026	The Diary of a Nobody	329	-SV1	44,387	6,332	6.14	PR
1240	The Playboy of the Western World...	317	-SV2	21,616	3,710	5.99	PR
32530	Armageddon—2419 A.D.	313	-SV1	27,829	5,649	5.85	PS
42324	Frankenstein; Or, The Modern Pro...	313	-SV3	78,348	10,228	5.92	PR
3328	Man and Superman: A Comedy and a...	312	SV5	70,074	10,895	6.04	PR
1720	The Man Who Knew Too Much	310	-SV2	60,871	8,489	5.82	PR
32706	Triplanetary	309	SV1	93,946	14,125	5.70	PS
572	The Great Big Treasury of Beatri...	307	-SV1	28,912	5,252	5.96	PR,PZ
1937	The Second Jungle Book	304	-SV1	65,478	8,616	5.74	PR
792	Wieland; Or, The Transformation:...	303	-SV1	83,097	10,520	5.74	PS
27761	Hamlet, Prince of Denmark	301	-SV2	35,961	8,485	6.00	PR
19	The Song of Hiawatha	297	-SV2	33,070	5,894	6.10	PS
123	At the Earth's Core	296	-SV1	51,233	7,507	5.94	PS
848	The Black Arrow: A Tale of Two R...	292	-SV3	82,934	10,325	5.67	PR
1526	Twelfth Night; Or, What You Will	292	-SV2	23,656	4,982	6.07	PR
2240	Much Ado about Nothing	289	-SV2	23,924	4,548	6.03	PR
21816	The Confidence-Man: His Masquerade	289	SV1	95,798	15,335	6.05	PS
225	At the Back of the North Wind	288	-SV4	91,335	7,637	6.19	PR,PZ
767	Agnes Grey	287	SV6	69,672	9,698	6.15	PR
4368	Flappers and Philosophers	287	SV2	62,912	10,612	6.07	PS
28346	Deathworld	287	-SV1	59,208	8,504	5.57	PS
394	Cranford	285	SV3	73,969	9,544	6.07	PR
2870	Washington Square	285	-SV1	66,537	7,870	6.04	PS
10556	The Old Man in the Corner	285	-SV1	71,168	9,138	5.70	PR
7464	The Adventures of Sally	282	-SV1	79,829	10,866	6.04	PR

Table S1: All Project Gutenberg eBooks considered in this study,
sorted by downloads.

ID	Title	DL's	Mode	N_w	Uniq(N_w)	$h_{\mathrm{avg}}^{b_i}$	LoCC
13650	Nonsense Books	282	-SV2	33,307	6,625	6.11	PR,PZ
1640	Lilith: A Romance	281	-SV1	94,723	10,437	5.86	PR
21084	Jokes For All Occasions: Selected...	281	SV1	73,724	13,077	6.03	PN
27780	Treasure Island	281	SV3	70,752	9,129	5.89	PZ,PR
901	The Jew of Malta	279	SV2	29,482	5,953	5.85	PR
551	The Land That Time Forgot	278	-SV2	38,692	6,318	5.87	PS
4737	A Tale of a Tub	277	-SV2	47,518	8,944	6.05	PR
134	Maria; Or, The Wrongs of Woman	271	-SV2	45,519	8,213	6.20	PR
1167	A Strange Disappearance	270	SV6	51,971	7,276	5.97	PS
126	The Poison Belt	268	SV2	30,911	6,125	5.86	PR
486	Ozma of Oz	268	SV2	40,887	5,754	5.99	PZ
1448	Heidi	268	-SV3	92,839	7,170	6.26	PZ
34339	The Princess and the Goblin	268	SV3	53,073	6,191	6.12	PZ
17866	Murder in the Gunroom	267	-SV1	70,159	9,106	5.67	PS
517	The Emerald City of Oz	266	-SV4	56,349	6,894	6.11	PZ
2253	Henry V	266	SV2	28,511	6,816	5.72	PR
804	A Sentimental Journey Through Fr...	261	-SV2	39,942	7,664	6.21	PR
2276	The Private Memoirs and Confessi...	260	-SV3	84,689	11,012	5.90	PR
847	Lays of Ancient Rome	259	SV2	27,247	6,395	5.56	PR
1354	Chronicles of Avonlea	258	-SV3	67,911	9,077	6.05	PZ,PS
10743	Moonfleet	258	SV2	84,340	8,816	5.82	PZ,PR
5343	Rainbow Valley	257	SV6	83,662	10,103	5.96	PZ,PS
26740	The Picture of Dorian Gray	257	-SV1	81,413	10,071	5.97	PR
2776	The Four Million	255	SV2	53,430	10,997	6.15	PS
23661	The Book of Dragons	254	SV4	43,139	5,796	5.99	PZ
1027	The Lone Star Ranger: A Romance ...	253	-SV7	97,821	11,203	5.66	PS
1906	Erewhon; Or, Over the Range	251	-SV3	86,441	10,707	5.91	PR
1376	The Little White Bird; Or, Adven...	250	SV4	67,250	8,595	6.13	PR,PZ
22693	A Book of Myths	248	SV3	95,310	12,020	5.97	PZ
33391	Bill Nye's Cordwood	248	SV7	33,634	7,357	5.98	PS
42243	The Hour of the Dragon	247	-SV1	73,598	10,378	5.47	PS
479	Little Lord Fauntleroy	246	-SV1	59,532	6,584	6.20	PS,PZ
20869	The Skylark of Space	246	-SV7	88,354	10,923	5.96	PS
173	The Insidious Dr. Fu Manchu	245	SV1	76,727	11,097	5.70	PR
791	The Princess	245	-SV1	31,009	6,490	5.88	PR
19551	Alice in Wonderland, Retold in W...	245	SV1	21,970	2,576	5.91	PZ
20796	The Colors of Space	245	-SV1	47,430	7,943	5.69	PS
7477	The Book of Wonder	244	SV6	23,189	4,831	5.99	PR
2607	Psmith, Journalist	242	SV2	58,788	9,091	5.90	PR
9611	Joseph Andrews, Vol. 1	242	-SV3	65,101	9,273	6.08	PR
652	Rasselas, Prince of Abyssinia	241	-SV1	39,051	6,717	6.00	PR
11074	The Damned	241	-SV1	31,780	6,291	5.89	PR

Table S1: All Project Gutenberg eBooks considered in this study,
sorted by downloads.

ID	Title	DL's	Mode	N_w	Uniq(N_w)	$h_{\mathrm{avg}}^{b_i}$	LoCC
2062	All for Love; Or, The World Well...	239	-SV2	33,862	6,160	5.92	PR
2667	The Vicar of Wakefield	238	-SV1	64,043	8,658	6.13	PR
2688	The Clue of the Twisted Candle	237	-SV4	58,657	8,412	5.91	PR
257	Troilus and Criseyde	236	SV2	79,838	7,790	6.16	PR
3529	Letters Written During a Short R...	236	-SV1	51,797	8,978	6.19	PR
7118	What Maisie Knew	236	-SV1	97,379	10,645	6.08	PS
10966	The Ghost Pirates	234	SV2	48,716	5,197	5.72	PR
16865	Pinocchio: The Tale of a Puppet	234	-SV3	42,973	5,802	5.81	PQ,PZ
8446	The Enormous Room	232	-SV4	93,965	15,510	5.85	PS
605	Pellucidar	231	-SV1	59,704	8,094	5.86	PS
4540	In His Steps	230	-SV5	80,399	7,764	6.04	PS
2081	The Blithedale Romance	229	-SV2	78,977	11,935	6.16	PS
85	The Beasts of Tarzan	227	SV2	65,632	8,098	5.61	PS,PZ
20898	The Galaxy Primes	227	-SV3	69,493	10,547	5.92	PS
19145	The Time Traders	225	SV5	65,276	9,575	5.73	PS
25550	The Defiant Agents	225	-SV1	58,088	8,663	5.76	PS
1028	The Professor	223	SV2	89,330	13,988	6.13	PR
20728	Space Viking	223	-SV3	69,043	9,406	5.80	PS
20727	The Cosmic Computer	221	-SV3	67,071	9,152	5.79	PS
3536	The Enchanted Castle	220	SV5	71,244	8,257	6.10	PZ
5342	The Story Girl	220	-SV1	90,365	10,409	5.99	PZ
1142	Typhoon	219	-SV1	31,562	6,551	5.59	PR
1329	A Voyage to Arcturus	218	-SV3	94,055	10,863	6.01	PR
16921	Plague Ship	218	SV2	61,488	9,772	5.80	PS
13882	John Thorndyke's Cases : related ...	217	SV6	81,079	10,534	5.76	PR
619	The Warden	215	-SV7	74,638	9,597	6.01	PR
2726	Eight Cousins	214	-SV1	74,849	9,272	6.20	PS,PZ
28215	Empire	214	SV2	52,058	7,586	5.78	PS
90	The Son of Tarzan	212	-SV3	99,166	10,557	5.80	PZ,PS
2225	"Captains Courageous": A Story o...	212	SV4	55,249	9,538	6.03	PR
27726	Tolstoy on Shakespeare: A Critic...	212	-SV2	35,973	7,383	5.81	PR
1999	Crome Yellow	210	SV5	58,619	10,893	6.08	PR
12629	Language: An Introduction to the ...	210	-SV3	79,563	10,924	6.05	P
2777	Cabbages and Kings	209	SV5	64,623	12,006	6.10	PS
2770	Five Little Peppers and How They...	207	SV1	72,052	7,239	6.15	PS,PZ
17731	The Nigger Of The "Narcissus": A...	207	-SV3	54,797	10,147	5.68	PR
1188	The Lair of the White Worm	206	SV5	57,943	8,141	5.85	PR
11696	The Food of the Gods and How It ...	206	-SV2	75,641	12,312	5.83	PR
2686	The Book of Snobs	204	SV4	64,937	12,719	6.23	PR
20288	Edward the Second	204	-SV1	25,492	4,738	5.86	PR
38703	The Black Moth: A Romance of the...	204	-SV2	94,455	11,830	5.94	PR
1424	Castle Rackrent	203	-SV1	46,146	7,407	6.08	PR

Table S1: All Project Gutenberg eBooks considered in this study,
sorted by downloads.

ID	Title	DL's	Mode	N_w	Uniq(N_w)	$h_{\mathrm{avg}}^{b_i}$	LoCC
11666	The Conjure Woman	203	-SV4	56,819	7,136	5.86	PS
17797	Memoir of Jane Austen	203	-SV1	55,039	8,930	6.23	PR
2060	The History of Caliph Vathek	202	-SV1	37,695	7,596	6.01	PR
2948	Where Angels Fear to Tread	202	-SV1	50,351	7,901	5.92	PR
34414	Just William	202	-SV3	49,853	8,174	6.14	PZ
2233	A Damsel in Distress	201	-SV7	78,851	11,166	6.05	PR
19726	The Door Through Space	201	SV2	43,855	7,727	5.64	PS
26998	Peter Pan in Kensington Gardens	201	-SV1	23,242	3,960	6.11	PR,PZ
149	The Lost Continent	200	-SV1	39,077	6,623	5.87	PS
6753	Psmith in the City	199	SV6	53,944	8,359	5.98	PR,PZ
3829	Love Among the Chickens	198	-SV1	51,823	8,561	6.07	PR
40284	The Sex Life of the Gods	197	-SV2	33,133	5,515	5.92	PS
19651	Key Out of Time	196	SV2	58,298	8,680	5.80	PS
5340	Further Chronicles of Avonlea	195	-SV1	75,729	9,520	6.01	PZ,PS
12384	Battle-Pieces and Aspects of the...	194	-SV2	33,193	9,346	5.59	PS
2568	Trent's Last Case	193	-SV4	75,407	9,556	6.00	PR
1881	The Call of the Canyon	192	-SV3	74,677	10,896	5.97	PS
21970	The Scarlet Plague	192	SV2	20,702	4,478	5.65	PS
32242	A Wonder Book for Girls & Boys	192	-SV4	53,499	7,300	6.28	PZ,BL
40426	Daddy Long-Legs: A Comedy in Fou...	192	SV1	42,137	7,955	6.25	PS
2183	Three Men on the Bummel	190	-SV2	70,600	10,111	6.02	PR
2429	Lost Face	190	SV4	43,887	6,999	5.74	PS
17314	Five Children and It	190	SV2	55,960	7,720	6.05	PZ
434	The Circular Staircase	189	-SV8	74,000	8,559	5.70	PS
20717	The Girl on the Boat	189	-SV1	69,907	9,892	5.98	PR
21873	Planet of the Damned	189	-SV1	57,211	8,745	5.51	PS
12187	The Mystery of 31 New Inn	188	-SV4	79,489	9,544	5.88	PR
17144	The House of the Vampire	188	-SV1	27,360	6,056	6.00	PS
24022	A Christmas Carol	188	-SV2	30,192	5,966	6.07	PR
419	The Magic of Oz	186	SV2	41,308	5,108	6.14	PZ
1777	Romeo and Juliet	186	-SV1	30,061	5,331	5.81	PR
3688	The Chronicles of Clovis	186	SV3	54,919	10,615	6.04	PR
770	The Story of the Treasure Seeker...	185	-SV1	54,339	6,352	5.98	PZ,PR
2020	Tarzan the Terrible	185	-SV2	97,155	10,243	5.71	PZ,PS
13937	The Mysterious Rider	185	-SV1	98,620	11,444	5.92	PS
19090	Star Hunter	185	-SV1	34,745	6,554	5.85	PS
961	Glinda of Oz : In Which Are Relat...	184	SV1	40,477	5,344	6.13	PZ
8223	Edgar Huntly; or, Memoirs of a S...	184	-SV1	94,969	11,264	5.73	PS
1118	Much Ado about Nothing	183	-SV1	25,863	4,307	6.09	PR
706	The Amateur Cracksman	182	SV2	55,454	8,141	5.90	PR
3776	The Valley of Fear	182	-SV2	58,894	7,962	5.79	PR
20058	The Napoleon of Notting Hill	182	-SV1	56,272	9,163	5.81	PR

Table S1: All Project Gutenberg eBooks considered in this study,
sorted by downloads.

ID	Title	DL's	Mode	N_w	Uniq(N_w)	$h^{b_i}_{\mathrm{avg}}$	LoCC
21964	The Short-story	182	SV4	77,026	14,487	5.91	PN
9909	Nightmare Abbey	180	-SV1	27,713	6,415	6.05	PR
25024	The Night of the Long Knives	180	-SV1	34,099	6,076	5.84	PS
535	Travels with a Donkey in the Cev...	179	-SV1	35,808	7,293	6.10	PR
1154	The Voyages of Doctor Dolittle	179	SV4	75,051	8,736	6.02	PZ
485	The Road to Oz	178	SV1	41,693	5,632	6.30	PZ
959	The Lost Princess of Oz	178	-SV2	49,244	5,816	5.95	PZ
11228	The Marrow of Tradition	178	-SV1	90,621	12,043	5.73	PS
654	Grace Abounding to the Chief of ...	177	-SV3	59,899	7,427	5.90	PR
849	The Idle Thoughts of an Idle Fellow	177	-SV4	42,691	8,415	6.06	PR
2306	Uncle Remus, His Songs and His S...	177	-SV2	56,068	7,617	6.01	PS
39827	The Conduct of Life	176	SV3	70,371	12,818	6.17	PS
4099	The Angel in the House	175	SV2	29,592	6,597	6.14	PR
653	The Chimes : A Goblin Story of So...	174	-SV1	32,291	5,902	5.95	PR
32037	Eureka: A Prose Poem	174	-SV3	43,909	8,699	6.06	PS
887	Intentions	173	SV4	61,608	10,297	6.04	PR
22754	Masters of Space	173	-SV2	53,615	8,943	5.96	PS
873	A House of Pomegranates	172	SV1	34,498	5,178	6.07	PZ,PR
955	The Patchwork Girl of Oz	172	SV4	59,019	6,738	6.09	PZ
23893	The Melting-Pot	172	-SV1	34,639	8,161	5.99	PR
552	The People That Time Forgot	171	-SV3	40,111	6,256	5.88	PS
18137	Little Fuzzy	171	SV2	61,172	8,317	5.85	PS
19810	My Ántonia	171	SV4	82,986	10,236	6.25	PS
6312	Representative Men: Seven Lectures	170	SV7	58,350	11,391	6.18	PS
10586	Mike and Psmith	170	-SV2	55,563	8,223	5.93	PZ
16389	The Enchanted April	170	SV1	81,946	9,198	6.15	PR
2809	Main-Travelled Roads	169	SV3	92,273	12,375	6.04	PS
21530	The Angel of Terror	169	SV1	64,353	8,410	5.89	PR
92	Tarzan and the Jewels of Opar	168	-SV2	68,483	8,765	5.52	PS
2804	Rose in Bloom : A Sequel to "Eigh...	168	-SV7	96,476	10,282	6.15	PS
32202	The Irish Fairy Book	168	-SV1	93,247	10,585	6.12	PZ,GR
166	Summer	165	-SV2	58,699	8,071	6.10	PS
1533	Macbeth	165	SV5	20,488	5,342	5.73	PR
956	Tik-Tok of Oz	163	SV2	49,439	6,652	5.93	PZ
4715	An African Millionaire: Episodes...	163	-SV4	69,726	10,917	6.03	PR
15238	Mathilda	163	-SV3	48,188	7,955	5.83	PR
25776	This Crowded Earth	163	-SV1	38,631	7,386	5.75	PS
170	The Haunted Hotel: A Mystery of ...	162	-SV2	63,112	7,866	6.05	PR
957	The Scarecrow of Oz	162	SV3	47,100	6,109	6.05	PZ
3756	Indiscretions of Archie	162	SV1	75,934	10,787	6.02	PR
11451	The Rome Express	162	-SV2	35,528	6,100	5.85	PR
6684	Uneasy Money	161	-SV1	67,497	9,474	5.98	PR

Table S1: All Project Gutenberg eBooks considered in this study,
sorted by downloads.

ID	Title	DL's	Mode	N_w	Uniq(N_w)	$h^{b_i}_{\text{avg}}$	LoCC
3781	The Jewel of Seven Stars	160	SV1	93,600	9,639	5.93	PR
12239	Dead Men's Money	160	-SV2	77,718	8,236	5.99	PR
18458	Star Born	160	-SV2	63,263	9,130	5.68	PS
5347	Understood Betsy	159	SV2	48,627	6,303	6.15	PS,PZ
14257	The Magician	159	-SV1	75,002	9,845	5.76	PR
18668	In Search of the Unknown	159	SV5	72,510	11,999	6.04	PS
29405	The Gods of Mars	158	-SV4	85,229	9,447	5.74	PS
21051	Skylark Three	157	-SV2	88,281	11,239	5.98	PS
96	The Monster Men	155	-SV3	59,570	7,686	5.73	PS
498	Rebecca of Sunnybrook Farm	155	-SV2	76,214	10,647	6.22	PS,PZ
11128	The Red Thumb Mark	155	-SV1	71,355	9,467	5.79	PR
20387	A Thin Ghost and Others	155	-SV1	31,902	5,647	5.90	PR
1905	The Governess; Or, The Little Fe...	154	SV1	51,179	6,265	6.12	PR,PZ
2786	Jack and Jill	154	-SV3	95,398	10,971	6.26	PS,PZ
20788	Storm Over Warlock	154	SV3	63,400	9,419	5.80	PS
611	Prester John	152	-SV1	80,326	9,876	5.82	PR
7031	The Sheik: A Novel	152	SV4	88,777	9,919	5.76	PR
316	The Golden Road	151	-SV6	78,398	9,871	6.05	PZ,PS
775	When the Sleeper Wakes	151	-SV2	82,707	11,724	5.79	PR
2005	Piccadilly Jim	151	SV3	82,948	11,118	6.08	PR
3026	North of Boston	151	-SV3	20,026	3,710	5.94	PS
3674	The Dragon and the Raven; Or, Th...	151	SV4	83,031	8,640	5.67	PZ,PR
32759	Red Nails	151	-SV7	32,027	6,002	5.39	PS
2263	Julius Caesar	150	-SV1	22,088	4,594	5.55	PR
3543	Heartbreak House	150	SV1	48,819	7,857	5.94	PR
4230	Tom Swift and His Motor-Cycle; O...	150	-SV1	43,266	5,478	5.91	PZ,PS
111	Freckles	149	-SV1	84,156	9,587	5.99	PS,PZ
6927	The White Feather	149	-SV2	46,316	6,861	5.92	PZ,PR
38070	The Norwegian Fairy Book	149	-SV5	82,968	6,996	6.11	PZ
832	Robin Hood	148	-SV1	63,466	8,188	5.92	PZ
1126	Measure for Measure	148	SV2	26,514	4,746	5.82	PR
5429	Preface to Shakespeare	148	-SV1	22,577	5,123	6.11	PR
18846	Voodoo Planet	148	SV2	24,352	5,447	5.63	PS
20912	The Daffodil Mystery	148	-SV2	69,802	8,344	5.62	PR
24880	The Wreck of the Titan: or, Futility	148	-SV3	71,516	10,341	5.85	PS
1531	Othello, the Moor of Venice	147	-SV1	28,054	5,659	5.90	PR
14658	The Road	147	SV1	52,048	8,190	5.81	PS
20707	The Black Star Passes	147	-SV1	76,729	9,758	6.00	PS
709	The Princess and Curdie	146	-SV1	58,112	6,930	5.96	PZ
10459	The Celtic Twilight	146	SV3	40,200	6,078	5.96	PR
13944	After London; Or, Wild England	146	SV3	84,878	10,954	5.89	PR
20840	Rebel Spurs	146	-SV1	62,246	9,499	5.76	PS

Table S1: All Project Gutenberg eBooks considered in this study,
sorted by downloads.

ID	Title	DL's	Mode	N_w	Uniq(N_w)	$h_{\text{avg}}^{b_i}$	LoCC
31501	The Sensitive Man	146	-SV1	21,639	5,185	5.92	PS
6985	A Prefect's Uncle	145	SV2	43,639	6,858	5.92	PZ,PR
7353	Birds in Town & Village	145	SV2	64,787	9,984	6.15	QL,PR
20919	The Status Civilization	145	SV2	46,948	8,258	5.66	PS
31619	The Planet Savers	145	SV2	27,917	5,694	5.75	PS
554	The Contrast	144	SV1	22,981	4,968	6.31	PS
2046	Clotel; Or, The President's Daug...	144	-SV1	62,334	9,426	5.73	PS
6622	Legends That Every Child Should ...	144	SV4	78,055	11,767	5.90	PZ
37364	The Second Jungle Book	144	-SV1	65,808	8,763	5.73	PR
30368	A Christmas Carol: The original m...	143	-SV2	29,938	5,916	6.08	PR
8994	What Katy Did	142	SV3	51,126	6,718	6.03	PZ
29042	A Tangled Tale	142	-SV1	29,092	6,402	6.02	PZ,PR,QA
32664	Black Amazon of Mars	142	-SV2	24,287	4,488	5.47	PS
37332	A Little Princess: Being the who...	142	SV4	68,968	7,781	5.94	PZ
1338	Selected Prose of Oscar Wilde	141	-SV1	33,876	7,389	5.94	PR
6836	Three Men and a Maid	141	-SV1	56,775	8,607	5.98	PR
8713	A Man of Means	141	-SV1	27,857	5,773	5.97	PR
12163	The Sleeper Awakes: A Revised Edi...	141	-SV3	76,658	11,166	5.78	PR
39896	The Girl Next Door	141	-SV3	43,530	5,842	6.08	PZ
42254	Beyond the Black River	141	SV1	22,418	4,704	5.54	PS
553	Out of Time's Abyss	140	SV3	37,429	6,011	5.72	PS
339	Old Indian Days	139	-SV3	49,302	7,612	5.88	PS
23624	Ride Proud, Rebel!	139	SV3	69,703	10,257	5.65	PZ
5803	Not that it Matters	138	SV3	48,252	8,321	6.10	PR
7028	The Clicking of Cuthbert	138	SV2	62,972	9,650	6.01	PR
34181	Irene Iddesleigh	138	SV2	34,616	6,993	5.85	PR
2248	The Winter's Tale	137	SV1	27,527	5,985	5.88	PR
423	Round the Red Lamp: Being Facts ...	136	SV3	72,337	10,302	5.86	PR
8086	Down and Out in the Magic Kingdom	136	-SV2	52,807	9,605	5.90	PS
36775	Humorous Readings and Recitation...	136	-SV4	63,358	12,411	5.99	PN
2662	Under the Greenwood Tree; Or, Th...	135	-SV3	60,895	9,310	6.11	PR
10832	Carnacki, the Ghost Finder	135	SV6	54,261	6,374	5.79	PR
32256	The Big Time	135	-SV3	40,958	7,485	5.82	PS
1114	The Merchant of Venice	134	SV3	25,657	4,634	6.08	PR
8677	Behind a Mask; or, a Woman's Power	134	-SV1	41,571	5,830	6.02	PS
10373	The Middle Temple Murder	134	-SV1	74,520	8,891	6.01	PR
5830	A Garland for Girls	133	SV3	73,313	9,227	6.29	PZ
23625	The Magic Pudding	133	-SV1	20,656	4,339	5.71	PZ
24280	Endymion: A Poetic Romance	133	-SV1	36,058	8,389	6.05	PR
172	The Haunted Bookshop	132	-SV1	66,085	10,714	6.08	PS
960	The Tin Woodman of Oz : A Faithfu...	132	SV1	45,355	5,602	6.11	PZ
42259	The People of the Black Circle	132	-SV1	31,494	6,154	5.50	PS

Table S1: All Project Gutenberg eBooks considered in this study,
sorted by downloads.

ID	Title	DL's	Mode	N_w	Uniq(N_w)	$h_{\text{avg}}^{b_i}$	LoCC
4709	Brewster's Millions	131	-SV3	64,117	9,487	5.96	PS
6984	The Pothunters	131	-SV1	42,401	6,823	5.94	PR,PZ
13969	The Hill of Dreams	131	-SV1	66,567	10,083	6.02	PR
20782	Triplanetary	131	SV3	59,104	9,397	5.75	PS
318	John Barleycorn	130	-SV6	68,849	10,360	6.08	PS
238	Dear Enemy	129	-SV2	68,598	10,326	6.21	PZ,PS
369	The Outlaw of Torn	129	-SV2	66,083	8,621	5.80	PS
393	The Blue Lagoon: A Romance	129	-SV1	65,857	9,498	6.08	PR
2727	Allan's Wife	128	-SV2	51,330	6,823	5.74	PR
4980	Old Granny Fox	128	SV1	23,490	2,900	6.01	PS,PZ
451	The Shadow Line: A Confession	127	-SV1	40,262	6,963	5.87	PR
984	Who Was Who: 5000 B. C. to Date ...	127	-SV1	22,017	6,518	6.16	PN
14255	Hints for Lovers	127	-SV1	35,090	7,365	6.41	PN
27444	Starman's Quest	127	-SV1	48,068	8,255	5.92	PS
35545	Sanders of the River	127	-SV2	58,928	9,440	5.71	PR
604	Gulliver of Mars	126	-SV1	72,540	10,834	6.01	PR
1794	King Lear	126	-SV1	32,080	6,035	5.72	PR
4023	Candida	126	-SV1	23,280	4,904	5.91	PR
10542	The Boats of the "Glen Carrig" : ...	126	SV4	60,641	6,106	5.83	PR
15274	The Girl from Montana	126	SV1	65,195	7,643	6.12	PS
23790	The Ultimate Weapon	126	-SV1	31,612	6,192	5.81	PS
2268	Antony and Cleopatra	125	-SV1	27,894	6,037	5.79	PR
3006	Stalky & Co.	125	SV2	68,334	11,439	5.90	PR
5077	Marmion: A Tale of Flodden Field...	125	-SV7	96,394	18,876	5.77	PR
16339	The Passenger from Calais	125	-SV5	58,424	8,966	5.91	PR
40386	Wandering Ghosts	125	SV3	69,400	7,737	5.83	PS
958	Rinkitink in Oz : Wherein Is Reco...	124	-SV3	50,029	6,406	5.92	PZ
2175	You Never Can Tell	124	SV2	35,507	6,024	6.01	PR
12803	Headlong Hall	124	-SV6	29,817	7,382	6.03	PR
22549	Space Prison	124	-SV2	55,743	7,682	5.72	PS
4011	Epicoene; Or, The Silent Woman	123	-SV2	56,201	12,243	6.05	PR
8183	Time and the Gods	123	-SV1	41,019	4,938	6.00	PR
9806	Mr. Justice Raffles	123	SV2	67,349	8,974	5.98	PR
35304	The Last Stroke: A Detective Story	123	-SV4	71,054	9,267	5.95	PS
836	The Phoenix and the Carpet	121	-SV1	63,785	8,402	6.10	PR,PZ
24450	Bones: Being Further Adventures i...	121	SV6	58,295	9,433	5.88	PR
26197	The Nursery Rhyme Book	121	-SV2	28,949	5,128	6.01	PN
32415	The Nursery Rhymes of England	121	-SV2	54,268	8,342	6.06	PN
837	The Story of the Amulet	120	SV1	71,060	8,678	6.14	PZ,PR
1718	Manalive	120	SV2	60,286	9,892	5.81	PR
6879	The Gold Bat	120	-SV1	44,544	6,614	5.86	PR,PZ
1696	The Club of Queer Trades	119	-SV1	44,838	7,493	5.91	PR

Table S1: All Project Gutenberg eBooks considered in this study,
sorted by downloads.

ID	Title	DL's	Mode	N_w	Uniq(N_w)	$h_{\text{avg}}^{b_i}$	LoCC
2377	The Son of the Wolf	119	SV1	49,471	8,790	5.93	PS
21775	The Best of the World's Classics...	119	SV2	69,929	11,723	6.10	PN
43984	Chaucer for Children: A Golden Key	119	-SV3	70,535	12,922	5.90	PZ,PR
91	Tom Sawyer Abroad	118	SV1	35,073	4,770	5.88	PS
1725	Heart of the West	118	-SV3	80,351	14,038	6.11	PS
5265	The Ball and the Cross	118	SV6	80,564	10,864	5.90	PR
11252	Martin Hewitt, Investigator	118	-SV1	58,070	8,137	5.79	PR
20606	The Magic City	118	-SV2	62,005	7,854	6.12	PZ
30408	The Fifth-Dimension Tube	118	SV1	29,288	5,780	5.44	PS
9932	The Last Trail	117	-SV1	73,713	9,684	5.88	PS
13897	The Adventure Club Afloat	117	SV2	61,983	8,551	5.98	PZ
30796	The Dueling Machine	117	-SV1	22,013	5,123	5.73	PS
1535	The Tragedy of Coriolanus	116	-SV5	31,427	6,527	5.82	PR
12915	The White Devil	116	-SV1	31,636	6,164	5.75	PR
17226	Emily Fox-Seton : Being "The Maki...	115	-SV1	83,661	10,364	6.09	PS
32954	The Black Arrow: A Tale of the T...	115	-SV3	81,251	10,499	5.68	PR
33735	Pamela Censured	115	-SV1	21,347	5,625	6.09	PR
367	The Country of the Pointed Firs	114	SV1	43,538	6,546	6.26	PS
39378	Mortal Coils	114	-SV2	40,459	8,209	6.01	PR
5070	The Doctor's Dilemma	113	-SV2	33,795	5,516	5.86	PR
18768	The Sky Is Falling	113	SV2	37,300	6,814	5.72	PS
24035	The Pirates of Ersatz	113	SV4	60,993	9,348	5.73	PS
27826	The Olive Fairy Book	113	SV4	94,884	9,226	6.03	PZ
93	Tom Sawyer, Detective	112	-SV1	24,161	3,609	5.62	PS
2098	A Thief in the Night: A Book of ...	112	SV5	67,080	8,958	5.94	PR
7371	A Sicilian Romance	112	-SV1	67,905	8,406	5.83	PR
14168	Widdershins	112	SV1	80,481	11,714	5.93	PR
17959	The Hand Of Fu-Manchu: Being a Ne...	112	SV5	65,740	10,511	5.79	PR
8668	Revenge!	111	SV1	82,530	10,444	5.79	PR
9925	Black Jack	111	SV7	77,775	9,248	5.86	PS
18581	Adrift in New York: Tom and Flor...	111	-SV4	55,502	6,781	6.20	PZ
25439	Looking Backward: 2000-1887	111	-SV2	83,194	10,782	6.12	PS
27991	The Blue Bird for Children: The W...	111	-SV3	34,586	5,454	5.91	PZ
33156	Young's Night Thoughts: With Life...	111	SV1	99,165	15,739	5.92	PR
35517	The Three Impostors; or, The Tra...	111	-SV5	59,493	9,209	5.93	PR
37431	Pride and Prejudice, a play foun...	111	-SV1	27,815	5,050	6.28	PR
9846	Excursions	110	SV2	72,711	12,388	6.31	PS
30964	The Ethical Engineer	110	-SV3	43,826	7,240	5.49	PS
1711	Child of Storm	109	-SV2	95,002	9,503	5.76	PR
10886	The Untamed	109	SV4	73,742	8,181	5.75	PS
20988	Islands of Space	109	-SV1	64,016	8,213	5.98	PS
32884	Ideas of Good and Evil	109	-SV4	57,392	8,612	6.29	PR

Table S1: All Project Gutenberg eBooks considered in this study,
sorted by downloads.

ID	Title	DL's	Mode	N_w	Uniq(N_w)	$h_{\mathrm{avg}}^{b_i}$	LoCC
33644	The Secret of the Ninth Planet	109	-SV1	47,868	7,940	5.78	PS
1845	Zuleika Dobson; Or, An Oxford Lo...	108	SV2	82,313	12,807	5.96	PR
7964	The Mystery of Cloomber	108	-SV1	49,575	8,451	5.82	PR
37189	The Return of the Soldier	108	-SV2	30,417	5,837	6.10	PR
472	The House Behind the Cedars	107	-SV4	72,592	10,370	6.17	PS
3815	Rolling Stones	107	SV7	73,748	14,014	6.02	PS
6100	Pollyanna Grows Up	107	SV4	76,521	8,702	6.07	PS,PZ
8931	The Gem Collector	107	-SV1	33,536	6,120	6.10	PR
16517	Liza of Lambeth	107	-SV1	37,805	5,404	5.93	PR
18019	The Luckiest Girl in the School	107	SV4	75,935	11,398	6.17	PZ
21959	Letters from a Self-Made Merchan...	107	SV2	53,290	7,621	6.09	PS
1058	The Mirror of the Sea	106	SV4	63,268	10,769	5.92	PR,G
1145	Rupert of Hentzau: From The Memo...	106	-SV1	84,946	8,751	5.84	PR
2057	The Last of the Plainsmen	106	SV5	71,832	10,514	5.77	PS
3055	The Wood Beyond the World	106	SV6	51,310	5,431	5.95	PR
33660	The Year When Stardust Fell	106	-SV1	64,015	8,622	5.75	PZ,PS
707	Raffles: Further Adventures of t...	105	-SV1	58,520	8,464	5.82	PR
2512	The Cruise of the Snark	105	-SV3	84,481	12,260	6.04	PS
2785	The Elusive Pimpernel	105	-SV1	84,592	11,054	5.85	PR
2885	The House of the Wolfings : A Tal...	105	-SV1	89,210	7,356	5.63	PR
3048	The Little Duke: Richard the Fea...	105	-SV1	42,333	6,517	5.90	PZ
3329	Caesar and Cleopatra	105	-SV2	36,955	6,757	5.87	PR
5604	Getting Married	105	SV4	62,850	8,676	6.06	PR
8995	What Katy Did Next	105	-SV1	57,819	8,961	6.17	PZ
19471	Badge of Infamy	105	-SV3	34,005	6,428	5.42	PS
47529	Oliver Twist, Vol. 1 (of 3)	105	SV1	55,784	8,673	5.93	PR
620	Sylvie and Bruno	104	-SV3	67,416	9,414	6.05	PR,PZ
1589	Tamburlaine the Great — Part 2	104	-SV3	29,104	6,104	5.71	PR
5805	The League of the Scarlet Pimpernel	104	SV3	76,360	9,833	5.68	PR
23810	At Fault	104	-SV1	58,484	9,834	5.91	PS
4682	Nonsense Novels	103	-SV3	35,966	7,139	5.95	PS
20121	Lone Star Planet	103	-SV1	31,467	6,002	5.65	PS
20154	Invaders from the Infinite	103	SV4	65,485	8,428	5.81	PS
25564	The Water-Babies: A Fairy Tale f...	103	SV5	70,419	8,695	6.05	PZ
39957	Prairie Gold	103	-SV1	71,352	13,238	6.14	PS
5341	Kilmeny of the Orchard	102	-SV1	41,059	6,389	6.09	PS,PZ
15798	Clover	102	-SV3	54,255	8,269	6.25	PZ
15851	Love Conquers All	102	SV1	59,296	11,012	6.08	PS
19369	The Triumphs of Eugène Valmont	102	SV3	91,395	11,231	5.82	PR
20857	Spacehounds of IPC	102	-SV1	90,361	12,184	5.89	PS
27771	Once on a Time	102	SV3	51,776	6,817	6.06	PR
364	The Mad King	101	SV2	94,878	9,902	5.75	PS

Table S1: All Project Gutenberg eBooks considered in this study, sorted by downloads.

ID	Title	DL's	Mode	N_w	Uniq(N_w)	$h_{\text{avg}}^{b_i}$	LoCC
3146	Two on a Tower	101	-SV3	98,817	12,406	6.02	PR
4087	An Essay Upon Projects	101	SV2	51,032	7,256	5.87	PR
5746	The Ancient Allan	101	-SV2	89,973	8,679	5.85	PR
20656	Old Christmas From the Sketch Bo...	101	-SV1	20,141	5,622	6.32	PS
21510	Legacy	101	SV4	79,603	9,918	5.92	PS
788	The Red One	100	SV1	40,139	7,999	5.90	PS
1644	The Adventures of Gerard	100	-SV2	67,760	8,183	5.72	PR
3638	The Devil's Disciple	100	-SV1	28,007	5,305	5.72	PR
4037	Appreciations, with an Essay on ...	100	-SV1	65,100	11,085	6.21	PR
5333	Every Man in His Humor	100	-SV1	51,083	11,701	6.02	PR
6440	Elsie Dinsmore	100	SV2	83,688	7,550	6.19	PZ
19706	Brood of the Witch-Queen	100	SV5	65,317	10,192	5.78	PR
20212	Police Your Planet	100	SV2	46,366	7,034	5.65	PS
20532	Love Among the Chickens: A Story ...	100	-SV1	50,477	8,468	6.06	PR
1557	Men of Iron	99	SV1	69,004	8,614	5.79	PZ
6678	Nonsenseorship	99	-SV2	40,888	9,328	5.88	PN
19142	The Devil Doctor	99	-SV3	75,335	11,542	5.78	PR
4358	The Sea Fairies	98	-SV1	43,738	6,085	6.03	PZ,PS
11247	The Exploits of Brigadier Gerard	98	-SV1	75,446	8,770	5.90	PR
15673	The Day of the Beast	98	-SV7	91,661	12,163	5.95	PS
17985	Tom Swift and The Visitor from P...	98	-SV2	35,863	7,210	5.88	PZ
25803	The Keepers of the King's Peace	98	SV1	54,696	9,643	6.00	PR
37532	The Scottish Fairy Book	98	-SV6	67,994	7,693	6.05	GR,PZ
644	The Haunted Man and the Ghost's ...	97	SV3	35,112	5,976	6.02	PR
778	Five Children and It	97	SV2	56,972	6,898	6.04	PR,PZ
2014	The Lodger	97	SV2	79,928	8,513	5.90	PR
5312	Mother Goose in Prose	97	-SV1	47,216	5,932	6.08	PZ
6877	The Head of Kay's	97	-SV1	46,151	6,803	5.99	PR,PZ
18639	The Victorian Age in Literature	97	-SV2	42,961	8,332	5.96	PR
33642	Earth Alert!	97	-SV2	35,084	6,259	5.66	PS
4552	The Border Legion	96	-SV2	96,804	11,043	5.67	PS
29448	Pariah Planet	96	-SV3	35,530	6,222	5.71	PS
6683	The Little Nugget	95	-SV3	72,792	9,820	5.99	PR
6880	The Coming of Bill	95	-SV4	86,999	10,724	5.97	PR
8771	Jurgen: A Comedy of Justice	95	SV4	96,487	12,458	6.13	PS
13029	The Art of the Moving Picture	95	SV4	64,245	11,569	6.19	PN
23292	Ted and the Telephone	95	SV4	50,056	7,613	6.18	PZ
310	Before Adam	94	-SV3	39,874	5,867	5.85	PS
1182	Dope	94	SV3	89,542	12,475	5.76	PR
1267	Kai Lung's Golden Hours	94	-SV3	83,935	11,901	5.92	PR
1595	Whirligigs	94	-SV6	77,330	14,137	6.04	PS
5308	The Paradise Mystery	94	SV5	76,999	9,197	5.94	PR

Table S1: All Project Gutenberg eBooks considered in this study,
sorted by downloads.

ID	Title	DL's	Mode	N_w	Uniq(N_w)	$h_{\mathrm{avg}}^{b_i}$	LoCC
10601	The Rangeland Avenger	94	SV3	78,989	8,853	5.70	PS
22767	Pagan Passions	94	-SV2	45,748	7,500	5.94	PS
864	The Master of Ballantrae: A Wint...	93	-SV1	90,272	10,833	5.86	PR
10377	The Evil Guest	93	-SV3	46,371	7,733	5.70	PR
17763	The Mystery of the Hasty Arrow	93	SV3	94,698	10,526	5.78	PS
19207	The Firelight Fairy Book	93	-SV3	43,592	6,550	6.08	PZ
21854	The Woman in Black	93	-SV4	70,598	9,006	6.02	PR
1605	The Crock of Gold	92	-SV1	56,448	8,000	6.03	PR
2381	Actions and Reactions	92	-SV3	69,048	12,796	5.92	PR
6382	Bat Wing	92	SV4	84,240	10,016	5.92	PR
9746	The Ashiel mystery: A Detective ...	92	-SV3	88,340	9,702	5.76	PR
21374	!Tention: A Story of Boy-Life du...	92	-SV5	98,388	8,341	5.73	PZ
27195	Negro Folk Rhymes: Wise and Other...	92	SV1	58,514	9,558	6.16	PS
34732	Max Carrados	92	-SV7	69,185	10,536	5.87	PR
363	The Oakdale Affair	91	-SV2	43,159	7,025	5.69	PS
1183	The Return of Dr. Fu-Manchu	91	-SV3	73,990	11,003	5.79	PR
14317	The Sorcery Club	91	-SV1	91,863	13,752	5.90	PR
22420	The Book of Nature Myths	91	SV1	34,582	3,294	6.09	PZ
24201	The Eye of Osiris	91	-SV4	99,133	11,572	5.88	PR
36127	Curious Myths of the Middle Ages	91	SV1	50,878	10,834	6.01	PN,GR
1527	Twelfth Night; Or, What You Will	90	-SV2	23,719	4,858	6.07	PR
4075	The Intrusion of Jimmy	90	-SV1	70,555	9,970	6.00	PR
4090	From Ritual to Romance	90	SV6	66,348	10,856	6.05	PN
13675	Goody Two-Shoes : A Facsimile Rep...	90	SV4	21,626	4,998	5.68	PZ
15580	The Rustlers of Pecos County	90	SV4	74,369	9,248	5.68	PS
26176	The Secret House	90	-SV3	59,369	8,333	5.94	PR
1550	A Lady of Quality : Being a Most ...	89	SV3	86,319	9,460	6.04	PS
3244	To Him That Hath: A Tale of the ...	89	-SV1	81,137	10,363	5.90	PS
5182	The Old English Baron: a Gothic ...	89	SV2	55,434	6,175	6.12	PR
10476	The Vanishing Man : A Detective R...	89	-SV4	96,450	11,571	5.88	PR
15281	Uncle Wiggily's Adventures	89	-SV1	40,862	3,905	6.15	PZ
21891	The Brand of Silence: A Detective...	89	SV4	60,373	5,970	5.77	PS
22332	Brain Twister	89	-SV2	42,166	6,248	5.96	PS
1915	The Second Thoughts of an Idle F...	88	-SV3	67,710	10,535	6.07	PR
2604	The Longest Journey	88	-SV1	96,195	12,324	5.98	PR
7498	Five Little Peppers Grown Up	88	SV2	82,409	7,634	6.07	PZ,PS
24775	Up the River; or, Yachting on th...	88	-SV5	71,028	6,848	6.06	PZ
29228	The Contrast	88	SV1	23,655	5,293	6.32	PS
29468	The Story of Don Quixote	88	SV1	96,047	10,792	5.93	PZ,PQ
33689	Oscar Wilde, Art and Morality: A...	88	SV1	30,020	6,593	6.05	PR
546	Under the Andes	87	SV5	94,678	9,755	5.75	PS
555	The Unbearable Bassington	87	-SV1	49,764	9,407	6.14	PR

Table S1: All Project Gutenberg eBooks considered in this study,
sorted by downloads.

ID	Title	DL's	Mode	N_w	Uniq(N_w)	$h_{\mathrm{avg}}^{b_i}$	LoCC
943	Misalliance	87	-SV1	37,348	5,417	6.03	PR
12436	The Night Horseman	87	SV2	92,843	10,395	5.76	PS
18505	A Popular Schoolgirl	87	-SV3	66,940	10,442	6.17	PZ
19474	Uller Uprising	87	-SV1	56,318	9,278	5.57	PS
32498	The Brain	87	-SV2	57,579	10,029	5.83	PS
37698	Dawn of the Morning	87	-SV1	99,035	9,574	6.12	PS
16732	Familiar Quotations	86	SV1	50,859	11,520	5.85	PN
21130	Book of Wise Sayings: Selected La...	86	SV1	23,524	6,485	6.02	PN
21334	The Beaux-Stratagem	86	-SV1	30,616	6,382	6.17	PR
29310	The Affair of the Brains	86	-SV1	24,847	5,499	5.69	PS
179	The Europeans	85	-SV1	60,003	7,501	6.24	PS
1144	In the Cage	85	-SV1	36,711	6,105	5.96	PS
2548	The Poor Clare	85	-SV1	23,478	4,730	5.80	PR
2851	Sixes and Sevens	85	-SV3	64,881	11,767	5.99	PS
3158	Our American Cousin	85	-SV1	20,258	3,924	6.02	PR
5141	What Katy Did at School	85	-SV4	55,892	7,461	6.16	PZ,PS
6955	The Prince and Betty	85	SV5	69,660	10,247	5.95	PR
17125	More William	85	-SV1	49,570	8,540	6.08	PZ
18151	Time Crime	85	SV3	38,201	6,757	5.77	PS
26240	The Clansman: An Historical Roma...	85	-SV2	91,674	12,708	5.79	PS
27525	Bones in London	85	SV6	65,414	10,234	6.07	PR
463	The Red Badge of Courage: An Epi...	84	-SV3	48,981	8,173	5.44	PS
1805	The Gentle Grafter	84	-SV5	45,478	8,724	6.19	PS
9902	The Middle of Things	84	-SV5	74,618	8,234	5.96	PR
14107	The Lost Stradivarius	84	-SV1	50,541	7,239	6.06	PR
18613	The Golden Scorpion	84	-SV5	67,147	10,022	5.75	PR
21927	Short Cruises	84	-SV1	41,057	6,011	6.08	PR
25770	The Dragon's Secret	84	-SV1	41,550	5,907	5.98	PZ
942	Green Mansions: A Romance of the...	83	-SV1	89,629	10,529	5.94	PR
1953	A Book of Strife in the Form of ...	83	-SV1	25,170	5,022	6.03	PR
8188	The Mysterious Key and What It O...	83	-SV1	20,093	4,032	6.14	PS
18492	Star Surgeon	83	-SV1	52,343	7,138	5.60	PS
18800	Last Enemy	83	-SV1	24,857	4,923	5.93	PS
22064	Tess of the Storm Country	83	-SV2	96,743	10,645	5.85	PS
22495	The New Pun Book	83	SV1	24,742	5,522	6.02	PN
25438	The Airlords of Han	83	-SV1	30,904	6,236	5.86	PS
37660	Of All Things	83	-SV1	44,177	9,370	6.10	PS
39281	Dictionary of English Proverbs a...	83	SV3	31,556	10,059	5.77	PN
311	Bunner Sisters	82	-SV1	31,612	5,941	6.02	PS
5758	Many Cargoes	82	SV6	69,637	9,188	6.08	PR
8681	The Face and the Mask	82	-SV5	72,288	9,658	5.82	PR
10736	Children of the Frost	82	SV2	51,252	7,626	5.65	PS

Table S1: All Project Gutenberg eBooks considered in this study,
sorted by downloads.

ID	Title	DL's	Mode	N_w	Uniq(N_w)	$h_{\mathrm{avg}}^{b_i}$	LoCC
17221	History of the Plague in London	82	-SV1	95,676	10,368	5.54	PR
20730	For the Sake of the School	82	-SV1	62,777	9,967	6.16	PZ
32934	The Young Colonists: A Story of ...	82	-SV3	78,959	8,441	5.62	PZ
37172	In a Glass Darkly, v. 1/3	82	-SV1	40,539	7,628	5.70	PR
38562	The Big Book of Nursery Rhymes	82	-SV2	22,021	4,327	6.06	PZ
5317	Through the Magic Door	81	SV3	47,696	9,239	6.01	PR,Z
6840	Queen Lucia	81	-SV6	88,117	10,588	6.23	PR
18361	Operation: Outer Space	81	-SV1	61,178	9,430	6.01	PS
26862	Howard Pyle's Book of Pirates : F...	81	SV3	87,360	10,963	5.86	PS
34426	The Enchanted Barn	81	-SV2	99,579	9,502	6.27	PS
3797	In the Days of the Comet	80	SV1	81,975	12,781	5.91	PR
10234	Old Creole Days: A Story of Creo...	80	SV4	69,210	11,858	5.92	PS
18970	Caves of Terror	80	-SV1	45,847	7,534	5.83	PR
19717	The Bostonians, Vol. I (of II)	80	SV3	80,403	9,990	6.21	PS
20431	The Tale of Beowulf, Sometime Ki...	80	-SV2	39,504	6,395	5.36	PR
22145	A Book of Burlesques	80	-SV1	30,934	8,261	5.89	PS
32953	Quest of the Golden Ape	80	-SV1	34,938	6,090	5.80	PS
545	At the Earth's Core	79	-SV1	51,251	7,425	5.94	PS
4731	Seven Little Australians	79	-SV4	46,663	7,814	6.02	PZ,PR
7308	The History of Mr. Polly	79	-SV1	70,301	12,092	5.93	PR
11045	The Ghost Ship	79	-SV2	53,173	8,091	6.04	PR
37858	Leaves in the Wind	79	-SV5	68,208	11,265	5.85	PR
40852	Instigations: Together with An Es...	79	SV8	99,318	22,573	6.14	PN
1327	Elizabeth and Her German Garden	78	-SV1	49,047	8,018	6.18	PR
2250	Richard II	78	SV1	23,824	5,612	5.55	PR
2431	Is Shakespeare Dead? : From My Au...	78	SV1	22,237	5,354	6.08	PR
10066	Gunman's Reckoning	78	SV2	81,339	9,128	5.80	PS
10671	The Botanic Garden. Part II.: Con...	78	SV2	55,964	12,356	6.18	PR
20726	A Slave is a Slave	78	SV1	21,980	4,760	5.47	PS
26027	Puck of Pook's Hill	78	-SV6	60,322	9,644	5.94	PR,PZ
35612	Three Philosophical Poets: Lucre...	78	SV1	49,476	9,967	6.16	PN
678	The Cricket on the Hearth: A Fai...	77	-SV1	33,105	5,887	6.13	PR
1106	The Tragedy of Titus Andronicus	77	SV2	25,230	4,868	5.78	PR
3479	The Metal Monster	77	-SV2	82,479	12,777	5.86	PS
19141	Edison's Conquest of Mars	77	-SV1	65,532	8,799	5.97	PS
520	The Life and Adventures of Santa...	76	SV2	31,025	4,917	6.24	PS,PZ
2251	Henry IV, Part 1	76	-SV2	27,041	5,967	5.75	PR
2861	The Sleuth of St. James's Square	76	SV4	81,152	9,906	5.87	PS
3326	The Well-Beloved: A Sketch of a ...	76	-SV6	64,400	9,764	6.09	PR
12245	The Defendant	76	SV1	27,244	6,015	6.14	PR
24353	Wired Love: A Romance of Dots and...	76	-SV1	51,185	7,741	6.08	PS
28118	The Great Gray Plague	76	SV2	26,569	5,216	5.79	PS

Table S1: All Project Gutenberg eBooks considered in this study,
sorted by downloads.

ID	Title	DL's	Mode	N_w	Uniq(N_w)	$h^{b_i}_{\text{avg}}$	LoCC
31308	Orientations	76	-SV4	50,779	8,068	6.00	PR
1020	Sword Blades and Poppy Seed	75	-SV5	31,494	7,182	5.95	PS
1262	The Heritage of the Desert: A Novel	75	SV5	82,671	10,548	5.86	PS
12590	The Shadow of the Rope	75	SV3	78,654	9,902	5.95	PR
18817	Ralestone Luck	75	-SV1	61,073	9,379	5.96	PS
22287	'Smiles': A Rose of the Cumberlands	75	SV2	97,058	12,873	6.06	PS
2266	King Lear	74	-SV1	26,625	5,922	5.67	PR
2324	A House to Let	74	-SV1	35,682	6,070	5.99	PR
2540	Father and Son: A Study of Two T...	74	-SV2	79,858	12,619	6.11	PR
9834	The Talleyrand Maxim	74	-SV2	73,007	9,116	6.00	PR
18761	The Circular Study	74	SV1	56,589	7,715	5.80	PS
19709	Danger in Deep Space	74	SV4	51,039	6,758	5.85	PZ
28071	The Red Triangle: Being Some Furt...	74	SV1	62,741	7,910	5.73	PR
30427	The Lost Kafoozalum	74	-SV1	22,873	4,663	5.67	PR
35425	The Mad Planet	74	SV2	21,751	4,665	5.54	PS
213	The Man from Snowy River	73	SV2	27,548	5,498	5.73	PR
687	A Personal Record	73	SV2	45,965	9,069	6.12	PR
1539	The Winter's Tale	73	SV1	26,632	5,810	5.97	PR
1795	Macbeth	73	-SV3	21,594	4,716	5.75	PR
2028	The Yellow Claw	73	-SV4	90,799	12,941	5.91	PR
3075	The Return	73	-SV8	81,784	10,698	5.90	PR
5210	The Borough	73	-SV2	69,781	13,105	5.70	PR
9297	The Orange-Yellow Diamond	73	SV2	76,745	9,127	6.02	PR
2186	"Captains Courageous": A Story o...	72	-SV3	55,256	9,619	6.02	PR
3795	Under the Lilacs	72	SV2	84,389	10,400	6.13	PZ,PS
5148	Rodney Stone	72	-SV1	95,455	11,327	5.98	PR
25051	Space Platform	72	-SV2	59,572	8,955	5.67	PS
33066	The Garden of Eden	72	SV8	76,124	8,763	5.94	PS
402	Penrod	71	SV2	59,209	11,242	6.02	PS
586	Religio Medici, Hydriotaphia, an...	71	-SV1	62,378	11,407	5.73	PR
1244	Love for Love: A Comedy	71	-SV1	35,581	5,612	6.10	PR
4253	Dramatic Romances	71	-SV4	42,767	9,449	6.06	PN,PR
7947	The Diary of a U-boat Commander: ...	71	-SV1	50,526	8,128	5.68	PR
12491	Twelve Types	71	SV4	26,567	5,926	6.05	PR,CT
32587	The Ambassador	71	-SV4	20,074	4,932	5.85	PS
39592	Princess Mary's Gift Book : All p...	71	SV3	52,891	9,080	5.92	PN
1115	The First Part of King Henry the...	70	-SV2	29,644	5,446	5.91	PR
3005	Tom Swift and His Airship	70	SV3	44,134	5,888	5.95	PZ
3777	Tom Swift and His Electric Rifle...	70	-SV1	42,586	5,543	5.73	PZ,PS
6093	Far Away and Long Ago: A History...	70	-SV2	97,738	11,935	6.04	QL,PR
6340	Literary Lapses	70	SV2	42,832	8,507	6.07	PS
13372	The Gloved Hand	70	SV6	71,710	7,949	5.79	PS

Table S1: All Project Gutenberg eBooks considered in this study,
sorted by downloads.

ID	Title	DL's	Mode	N_w	Uniq(N_w)	$h_{\mathrm{avg}}^{b_i}$	LoCC
17958	Warlord of Kor	70	-SV1	33,091	5,405	5.66	PS
19258	Tom Swift and the Electronic Hyd...	70	-SV2	33,037	6,810	5.99	PZ
20877	Mother West Wind's Children	70	SV1	30,730	3,321	6.07	PZ
27690	Nobody's Girl: (En Famille)	70	-SV7	75,505	7,624	5.99	PQ,PZ
1239	The Spirit of the Border: A Roma...	69	-SV1	89,991	11,175	5.77	PS
16199	Memoirs of the Author of a Vindi...	69	-SV1	25,056	5,030	6.26	PR
17412	The Bobbsey Twins : Or, Merry Day...	69	-SV2	35,035	5,051	6.05	PZ
19330	An Apache Princess: A Tale of th...	69	SV4	83,490	11,775	5.71	PS
22234	Aunt Jo's Scrap-Bag, Vol. 5: Jimm...	69	SV5	40,473	7,344	6.26	PZ
556	Rewards and Fairies	68	-SV1	76,118	10,963	5.89	PZ,PR
4006	Yesterdays	68	SV1	22,180	4,646	5.99	PS
7230	Not George Washington — an Autob...	68	-SV1	53,845	9,655	6.08	PR
9807	Scarhaven Keep	68	SV5	75,422	9,595	6.04	PR
9903	Way of the Lawless	68	-SV4	71,853	8,427	5.82	PS
16096	A Man's Woman	68	SV2	77,180	10,379	5.72	PS
18095	Successful Methods of Public Spe...	68	-SV1	20,886	5,212	6.29	PN
22354	The Adventures of Maya the Bee	68	-SV3	38,782	6,249	6.04	PZ,PT
24933	The Man Who Knew	68	SV2	53,696	7,630	5.97	PR
26624	The Road to Oz	68	SV1	41,600	5,975	6.31	PZ
40723	The Battle of Life. A Love Story	68	SV1	30,952	5,990	6.20	PR
329	Island Nights' Entertainments	67	-SV3	51,115	6,252	5.86	PR
720	Almayer's Folly: A Story of an E...	67	-SV1	66,011	8,756	5.89	PR
1537	Pericles, Prince of Tyre	67	-SV2	22,020	5,170	6.08	PR
2815	Democracy, an American novel	67	-SV1	72,151	9,565	5.94	PS
8914	England, My England	67	-SV2	65,871	9,483	5.87	PR
8920	The Light of Asia	67	SV2	40,105	8,282	5.95	PR
10869	The Abandoned Room	67	-SV1	79,681	8,156	5.58	PS
17513	St. Nicholas Magazine for Boys a...	67	SV1	52,200	10,229	6.21	PZ
18346	Null-ABC	67	-SV1	36,320	6,748	5.67	PS
18420	The Bobbsey Twins at Home	67	-SV1	44,743	4,142	6.15	PZ
24302	The Highest Treason	67	-SV1	23,369	4,708	5.52	PS
24723	Final Weapon	67	SV2	20,670	4,173	5.89	PS
24749	Adaptation	67	-SV1	24,213	5,107	5.78	PS
27903	The Magic World	67	-SV1	61,067	7,869	6.06	PZ
27924	Mugby Junction	67	-SV1	52,266	8,203	5.98	PR
37820	Chronicles of Martin Hewitt	67	-SV1	68,420	10,266	5.85	PR
41667	The Emerald City of Oz	67	-SV4	55,434	7,131	6.11	PZ
1869	The Man in Lower Ten	66	SV1	65,944	9,095	5.77	PS
2509	The Lani People	66	-SV3	60,766	9,632	5.78	PS
12170	The Wolf Hunters: A Tale of Adve...	66	-SV1	53,581	6,974	5.74	PS
12352	Iola Leroy; Or, Shadows Uplifted	66	SV1	75,790	9,672	5.90	PS
23845	Talents, Incorporated	66	-SV3	52,612	7,966	5.58	PS

Table S1: All Project Gutenberg eBooks considered in this study,
sorted by downloads.

ID	Title	DL's	Mode	N_w	Uniq(N_w)	$h_{\text{avg}}^{b_i}$	LoCC
24459	The Lost Princess of Oz	66	-SV2	48,468	5,924	5.96	PS,PZ
35117	Lord Tony's Wife: An Adventure o...	66	-SV4	91,093	11,740	5.80	PR
296	The Cash Boy	65	-SV1	28,355	4,051	6.15	PS,PZ
1472	In a German Pension	65	-SV1	31,065	6,199	6.16	PR
2244	As You Like It	65	SV2	24,200	5,015	6.12	PR
6995	Ghosts I Have Met and Some Others	65	-SV2	36,301	6,973	5.83	PS
11068	The Spirit of the Age; Or, Conte...	65	SV3	82,318	14,398	6.04	PR
11935	Mysticism in English Literature	65	-SV1	39,912	7,647	6.27	PR
13815	The Talking Beasts: A Book of Fa...	65	-SV2	78,095	10,644	5.90	PZ
15323	The Green Eyes of Bâst	65	SV3	75,766	10,244	5.88	PR
19066	Brigands of the Moon	65	SV5	71,019	8,526	5.68	PS
20630	The Borough Treasurer	65	SV1	78,858	9,969	5.93	PR
22182	The Best of the World's Classics...	65	-SV3	68,431	13,500	5.99	PR
24499	The Green Carnation	65	-SV5	46,138	8,383	6.07	PR
30324	The Pathless Trail	65	SV5	76,745	11,341	5.67	PS
34971	Among the Forest People	65	SV1	31,880	3,835	6.21	PZ
37503	Gammer Gurton's Needle	65	-SV1	23,264	5,001	5.54	PR
39143	The Making of a Saint	65	-SV1	73,742	9,187	5.85	PR
267	The Touchstone	64	-SV1	27,270	5,620	6.07	PS
330	Where There's a Will	64	SV3	68,246	7,529	6.01	PS
1204	Cabin Fever	64	SV4	57,450	9,115	6.00	PS
1547	Sir Thomas More	64	SV1	23,934	5,195	6.00	PR
2317	The Story of My Heart: An Autobi...	64	-SV2	33,401	5,716	6.29	PR
5776	100%: the Story of a Patriot	64	-SV9	98,446	10,978	5.70	PS
7884	In the Fog	64	-SV2	22,623	3,950	5.74	PS
9656	Alarms and Discursions	64	-SV3	50,945	9,397	6.03	PR
11153	No Hero	64	SV1	39,412	6,062	6.01	PR
17180	The Riddle of the Frozen Flame	64	-SV1	66,009	9,317	5.91	PS
18172	This World Is Taboo	64	-SV3	37,649	6,489	5.72	PS
19027	The Revolt on Venus	64	-SV1	54,941	7,369	5.72	PZ
19526	Stand by for Mars!	64	SV4	58,320	7,547	6.02	PZ
25067	The Planet Strappers	64	-SV1	67,425	12,119	5.88	PS
30537	The Royal Book of Oz : In which t...	64	-SV2	43,385	6,680	6.10	PZ
1583	Options	63	SV2	66,078	11,924	6.15	PS
1790	Troilus and Cressida	63	SV1	32,213	5,996	5.96	PR
3464	Tish: The Chronicle of Her Escap...	63	-SV2	91,009	10,697	5.92	PS
4381	The Aran Islands	63	-SV2	50,695	6,255	6.04	PR
5071	The Philanderer	63	-SV1	25,680	4,466	5.94	PR
6574	Watchers of the Sky	63	-SV1	33,018	6,964	6.07	PR
7365	If I May	63	-SV2	46,730	8,050	6.10	PR
14034	King Alfred's Viking: A Story of ...	63	-SV1	78,775	6,892	5.92	PZ
14154	The Tale of Terror: A Study of t...	63	-SV4	75,973	15,061	5.74	PN

Table S1: All Project Gutenberg eBooks considered in this study, sorted by downloads.

ID	Title	DL's	Mode	N_w	Uniq(N_w)	$h^{b_i}_{\mathrm{avg}}$	LoCC
18719	Space Tug	63	SV4	56,963	8,400	5.79	PS
27567	Aunt Jo's Scrap-Bag VI: An Old-Fa...	63	SV2	56,723	7,878	6.14	PZ
32563	The Lost Warship	63	-SV1	27,106	4,701	5.60	PS
32597	Accidental Flight	63	SV1	22,365	4,491	5.69	PS
34215	Shadowings	63	-SV2	43,506	10,051	6.23	PS
36281	The Slayer of Souls	63	-SV2	63,000	9,989	5.89	PS
41718	Dave Dawson on the Russian Front	63	SV1	50,697	6,407	5.86	PZ
764	Hans Brinker; Or, The Silver Skates	62	-SV1	85,043	11,700	6.03	PZ
966	Maid Marian	62	-SV1	36,838	7,162	5.99	PR
1163	Adventure	62	-SV2	76,325	10,405	5.80	PS
5311	Parnassus on Wheels	62	-SV1	36,880	6,444	6.09	PS
5676	A Double Story	62	SV1	34,563	4,828	5.85	PR,PZ
6359	The English Mail-Coach and Joan ...	62	-SV1	45,492	10,693	5.81	PR
9867	Riders of the Silences	62	-SV3	64,857	7,928	5.80	PS
11620	My Brilliant Career	62	-SV1	90,607	12,511	6.19	PR
16259	The Surprising Adventures of the...	62	-SV1	33,039	4,886	6.00	PZ
18217	Chambers's Elementary Science Re...	62	SV3	21,590	3,694	6.41	Q,PZ
18824	Fairies and Folk of Ireland	62	-SV1	74,241	5,535	6.06	PZ
21665	A Brief History of the English L...	62	SV1	78,236	15,998	6.21	PE,PN
27198	The Explorer	62	SV6	80,205	9,337	5.86	PR
30905	The Boarded-Up House	62	-SV2	34,944	5,682	6.01	PZ
35247	That Affair at Elizabeth	62	SV3	62,452	8,084	5.88	PS
35330	The Spanish Tragedy	62	SV1	26,640	5,113	5.74	PR
436	The Master Key: An Electrical Fa...	61	-SV1	36,417	6,242	5.92	PZ
2015	A Miscellany of Men	61	SV1	53,088	9,641	5.92	PR
2876	The Light That Failed	61	-SV1	74,177	10,326	5.90	PR
4993	A Texas Ranger	61	-SV2	72,903	9,330	5.88	PS
5747	Do and Dare — a Brave Boy's Figh...	61	-SV1	50,698	6,450	6.01	PZ
10443	The Rayner-Slade Amalgamation	61	-SV5	80,402	9,662	6.11	PR
11435	Small Means and Great Ends	61	SV2	31,081	6,087	6.02	PZ
17854	The Sport of the Gods	61	-SV2	41,937	6,572	5.94	PS
19023	A Daughter of the Sioux: A Tale ...	61	-SV3	77,282	11,415	5.74	PS
21865	King Arthur and His Knights	61	-SV1	43,484	4,920	5.71	PZ
22031	The Airplane Boys among the Clou...	61	SV2	49,500	6,496	5.97	PZ
22342	Supermind	61	SV7	73,369	9,148	5.85	PS
23028	Greylorn	61	-SV3	20,318	4,472	5.63	PS
28164	The Big Bow Mystery	61	-SV2	44,956	8,744	5.64	PR
30334	Ultima Thule	61	-SV12	26,476	5,094	5.92	PS
31356	The Man Who Staked the Stars	61	SV3	26,606	5,742	5.77	PS
32486	The Legion of Lazarus	61	-SV1	21,925	4,383	5.69	PS
34219	The Enchanted Castle	61	SV5	72,019	9,467	6.10	PZ
35617	The Terror: A Mystery	61	SV1	38,210	6,221	5.67	PR

Table S1: All Project Gutenberg eBooks considered in this study,
sorted by downloads.

ID	Title	DL's	Mode	N_w	Uniq(N_w)	$h_{\mathrm{avg}}^{b_i}$	LoCC
36958	A Child of the Jago	61	-SV4	63,674	10,480	5.76	PR
1478	A Parody Outline of History : Whe...	60	-SV3	23,588	5,505	6.10	PN
1585	The Wrong Box	60	-SV2	58,971	9,665	5.87	PR
5083	The Man of Feeling	60	SV2	38,372	7,002	6.05	PR
10110	The Postmaster's Daughter	60	SV1	67,192	10,752	5.87	PR
18934	My Lady Nicotine: A Study in Smoke	60	-SV2	52,469	7,952	5.92	PR
19381	Among the Farmyard People	60	-SV1	40,239	4,250	6.11	PZ
19478	Four-Day Planet	60	-SV1	59,954	7,788	5.69	PS
19535	George Bernard Shaw	60	-SV3	53,197	9,069	5.94	PR
25102	Nobody's Boy: Sans Famille	60	SV2	92,076	8,041	5.98	PQ,PZ
26999	Peter Pan in Kensington Gardens	60	-SV2	24,535	3,990	6.12	PZ,PR
27063	The Hero	60	-SV7	73,181	9,185	5.98	PR
32440	Dave Dawson at Dunkirk	60	SV3	52,993	6,488	5.76	PZ
33505	The Trembling of the Veil	60	-SV3	71,863	10,742	6.08	PR
33623	The Inventions of the Idiot	60	-SV1	27,697	5,284	6.17	PS
41049	The Onslaught from Rigel	60	-SV3	68,709	9,387	5.81	PS
1076	The Wallet of Kai Lung	59	SV1	80,921	9,857	5.93	PR
1446	Perfect Behavior: A Guide for La...	59	-SV3	38,230	7,901	6.26	PN
1460	The Black Dwarf	59	-SV1	58,461	10,507	6.01	PR
1882	The Young Forester	59	-SV1	51,323	7,232	5.76	PS
5660	Mary Louise	59	-SV2	47,424	7,310	6.07	PZ,PS
12345	Friday, the Thirteenth : A Novel	59	-SV1	48,415	7,671	5.88	PS
17870	Operation Terror	59	-SV3	47,352	7,069	5.75	PS
21687	The Youngest Girl in the Fifth: ...	59	SV7	75,351	10,291	6.07	PZ
23641	The Forsaken Inn: A Novel	59	-SV4	66,098	8,716	5.89	PS
25767	Picture and Text: 1893	59	-SV1	32,974	6,836	6.27	PS
27595	Eight Keys to Eden	59	SV2	54,529	8,308	5.84	PS
36869	The Real Man	59	SV3	94,115	11,403	5.92	PS
94	Alexander's Bridge	58	-SV1	29,576	5,211	6.15	PS
1263	The Glimpses of the Moon	58	-SV8	84,375	11,356	6.16	PS
1282	Tom Swift Among the Diamond Make...	58	SV2	43,526	5,598	5.97	PZ
2305	A Set of Six	58	SV3	85,490	12,412	5.87	PR
2524	My Lady Ludlow	58	SV2	80,667	9,495	5.97	PR
2722	Morning Star	58	SV2	85,928	8,230	5.82	PR
2865	Otto of the Silver Hand	58	-SV1	28,161	4,517	5.72	PZ
5632	Five Little Peppers Midway	58	-SV3	72,449	7,747	6.14	PZ,PS
8435	The Sturdy Oak : A composite Nove...	58	-SV1	60,759	10,043	6.06	PS
9963	Elsie's Girlhood: A Sequel to "El...	58	-SV1	96,936	9,609	6.28	PZ
10324	Bull Hunter	58	-SV3	53,488	7,322	5.80	PS
12028	The Uttermost Farthing: A Savant'...	58	SV4	53,811	9,084	5.72	PR
14902	Deadwood Dick, the Prince of the...	58	SV1	35,479	7,613	5.87	PS
15422	Israel Potter : His Fifty Years o...	58	-SV1	66,586	12,774	5.83	PS

Table S1: All Project Gutenberg eBooks considered in this study,
sorted by downloads.

ID	Title	DL's	Mode	N_w	Uniq(N_w)	$h_{\mathrm{avg}}^{b_i}$	LoCC
24370	Mercenary	58	-SV1	21,976	4,735	6.09	PS
24436	Anything You Can Do ...	58	SV2	55,504	8,734	5.74	PS
25728	Desert Conquest; or, Precious Wa...	58	-SV6	97,231	13,140	5.90	PS
28700	Robin Hood	58	SV2	94,173	10,221	5.96	PZ
32208	The Star Lord	58	-SV1	21,646	4,555	5.92	PS
34592	Behind the Green Door	58	-SV1	44,256	6,588	5.93	PZ
38777	Lad: A Dog	58	-SV5	83,311	12,631	5.81	PZ
1146	The Journal of a Voyage to Lisbon	57	SV3	45,588	7,422	6.14	PR
2154	Around the World in Eighty Days....	57	-SV3	63,910	9,642	5.89	PZ,PQ
2487	Cross Roads	57	-SV3	26,095	5,064	6.11	PS
4922	Bar-20 Days	57	-SV5	71,505	9,892	5.68	PS
7239	Men, Women, and Boats	57	-SV1	54,316	9,802	5.79	PS
9609	Joseph Andrews, Vol. 2	57	SV2	72,846	9,091	6.09	PR
11626	The Dawn of All	57	-SV4	89,297	10,893	5.93	PR
14427	True Love's Reward : A Sequel to ...	57	-SV1	70,493	7,636	6.05	PS
18881	The Idiot	57	SV1	22,941	4,767	6.16	PS
30836	Seven Keys to Baldpate	57	SV7	74,942	9,803	6.02	PS
32117	Eleven Possible Cases	57	-SV2	60,086	9,757	5.97	PS
32226	The Flower Princess	57	-SV1	21,907	3,938	6.17	PZ
36612	The Princess and Curdie	57	-SV1	57,852	7,439	5.96	PZ
37193	The Swedish Fairy Book	57	-SV2	51,321	5,610	6.05	PZ
1109	Love's Labour's Lost	56	-SV2	26,608	5,382	6.04	PR
1604	The Ebb-Tide: A Trio And Quartette	56	-SV1	48,277	8,357	5.89	PR
1611	Seventeen : A Tale of Youth and S...	56	-SV6	68,908	10,464	6.11	PS
1897	The Seventh Man	56	-SV3	78,397	9,114	5.74	PS
2713	Maiwa's Revenge; Or, The War of ...	56	-SV2	34,909	5,414	5.60	PR
4282	Don Rodriguez; Chronicles of Sha...	56	SV1	74,129	8,303	5.96	PR
5795	The Secret Rose	56	-SV2	22,480	4,147	5.87	PR
6936	Robinson Crusoe — in Words of On...	56	-SV1	27,568	2,425	5.79	PR,PZ
8673	A Columbus of Space	56	-SV8	71,372	9,309	5.97	PS
9791	Harrigan	56	-SV1	67,059	8,219	5.79	PS
9990	Brave and Bold; Or, The Fortunes...	56	-SV3	56,686	6,844	6.08	PZ
11377	The Man Whom the Trees Loved	56	SV1	26,393	5,365	6.02	PR
12793	Cobwebs from an Empty Skull	56	-SV2	53,936	11,664	5.89	PS
14667	A Christmas Garland	56	-SV2	29,981	8,075	6.09	PR
21632	Fame and Fortune; or, The Progre...	56	-SV1	45,855	5,764	6.25	PZ
21656	The Princess of the School	56	SV5	65,548	10,190	6.27	PZ
22338	The Impossibles	56	-SV2	52,808	6,920	5.90	PS
24929	The Green Rust	56	SV4	84,006	10,911	5.87	PR
25866	The Search	56	-SV5	64,628	7,748	6.03	PS
26019	Europa's Fairy Book	56	SV3	59,855	7,879	6.20	PZ
27922	David and the Phoenix	56	SV1	30,748	5,994	5.97	PZ

Table S1: All Project Gutenberg eBooks considered in this study,
sorted by downloads.

ID	Title	DL's	Mode	N_w	Uniq(N_w)	$h_{\mathrm{avg}}^{b_i}$	LoCC
31343	The Invaders	56	-SV1	22,168	4,683	5.78	PS
751	The Autocrat of the Breakfast-Table	55	SV5	93,970	16,700	6.17	PS
2126	The Quest of the Sacred Slipper	55	-SV3	53,235	8,334	5.65	PR
4020	Arcadian Adventures with the Idl...	55	-SV1	67,329	9,609	6.05	PS
7088	The Pilgrim's Progress in Words ...	55	SV1	27,072	2,578	5.90	PR,PZ
14744	Different Girls	55	SV2	56,778	9,060	6.10	PS
19355	A Book of Prefaces	55	-SV2	61,149	12,820	5.97	PS
20559	R. Holmes & Co. : Being the Remar...	55	SV2	33,956	6,479	5.97	PS
27174	Captain Jim	55	-SV1	90,723	11,003	6.05	PZ
28267	Venus in Boston: A Romance of Ci...	55	SV1	59,354	10,376	6.03	PS
37995	The Diamond Fairy Book	55	SV1	57,643	8,945	6.07	PZ
1358	Enoch Arden, &c.	54	-SV5	26,435	6,080	5.92	PR
2515	Stepping Heavenward	54	SV2	97,434	9,061	6.03	PS
10067	The Mystery of the Boule Cabinet...	54	-SV6	69,880	7,994	5.81	PS
12431	The Coquette, or, The History of...	54	-SV1	59,565	8,390	6.31	PS
12986	The Card, a Story of Adventure i...	54	SV5	76,379	11,439	6.05	PR
13888	Bacon	54	SV2	72,536	10,588	5.94	PR
16721	A Place so Foreign	54	-SV4	21,982	4,826	6.07	PS
21092	On the Trail of the Space Pirates	54	-SV1	54,158	6,935	5.87	PZ
21633	The Man of the Desert	54	-SV2	61,623	7,650	6.08	PS
32746	The Revival of Irish Literature ...	54	SV2	33,517	8,215	6.14	PR
34313	Literature in the Making, by Som...	54	-SV3	54,321	8,459	6.24	PN
35204	Sense of Obligation	54	-SV1	54,726	8,435	5.50	PS
1127	The Tragedy of Othello, Moor of ...	53	-SV1	32,617	5,445	5.90	PR
1143	Notes on Life & Letters	53	-SV3	81,978	12,718	5.95	PR
3475	The Efficiency Expert	53	-SV1	51,038	6,758	6.01	PS
6987	Five Little Peppers Abroad	53	-SV4	84,392	8,095	6.24	PZ,PS
9156	Life and Remains of John Clare, ...	53	-SV3	69,387	11,805	6.21	PR
10337	Lady into Fox	53	SV2	24,701	4,060	5.79	PR
10847	The Maids Tragedy	53	-SV1	32,291	5,048	5.66	PR
11195	Alcatraz	53	SV2	70,059	9,099	5.72	PS
11371	The Moorland Cottage	53	-SV1	44,584	6,794	6.05	PR
15585	Humorous Masterpieces from Ameri...	53	SV3	59,148	11,782	6.07	PN
20551	The White Invaders	53	-SV3	32,428	5,644	5.78	PS
20698	The Story of Glass	53	-SV2	39,063	6,299	6.30	PZ
21715	Away in the Wilderness	53	SV4	28,138	4,929	5.95	PZ
22057	Kid Wolf of Texas : A Western Story	53	-SV2	66,418	8,981	5.57	PS
22463	Chivalry	53	SV2	60,217	11,228	5.85	PS
35533	The Haunted Room: A Tale	53	-SV2	71,695	9,634	5.90	PR
39868	Glinda of Oz : In which are Relat...	53	SV1	40,697	5,410	6.13	PZ
949	Tom Swift and His Submarine Boat...	52	-SV1	44,612	5,539	6.04	PZ
1283	Tom Swift and His Wizard Camera;...	52	-SV1	45,898	5,938	5.93	PZ

Table S1: All Project Gutenberg eBooks considered in this study,
sorted by downloads.

ID	Title	DL's	Mode	N_w	Uniq(N_w)	$h^{b_i}_{\mathrm{avg}}$	LoCC
2260	Titus Andronicus	52	SV2	22,535	5,119	5.62	PR
2273	Tom Swift and His Motor-Boat; Or...	52	SV2	45,964	5,255	5.90	PZ
2295	Waifs and Strays: Part 1	52	-SV1	30,834	7,421	6.14	PS
13783	The Boy Inventors' Radio Telephone	52	SV2	46,496	7,559	5.94	PZ
14632	The Mystery of Mary	52	SV4	37,618	5,479	6.09	PS
30333	Daddy's Girl	52	-SV2	77,411	7,813	6.25	PZ
33979	Miscellaneous Aphorisms; The Sou...	52	-SV1	33,159	5,315	5.93	PR,HX
1625	The Frozen Deep	51	-SV3	28,550	4,900	6.04	PR
4268	Cousin Phillis	51	-SV1	40,955	6,050	6.18	PR
5815	The Great Impersonation	51	-SV1	76,505	9,387	6.00	PR
6418	Five Little Peppers and their Fr...	51	-SV1	88,606	8,260	6.02	PS,PZ
9196	The Clockmaker; Or, the Sayings ...	51	SV5	75,395	10,028	5.91	PS
10551	Affair in Araby	51	SV1	54,814	8,213	5.88	PR
14654	A Daughter of the Snows	51	-SV1	93,032	13,136	5.85	PS
15625	The Lookout Man	51	SV4	75,499	10,423	6.03	PS
16740	The Busie Body	51	-SV1	26,653	5,854	5.90	PR
17667	Dialogues of the Dead	51	SV5	64,892	9,719	6.01	PR
19370	Ullr Uprising	51	-SV1	42,036	7,215	5.52	PS
29466	Lords of the Stratosphere	51	-SV1	23,283	4,634	5.78	PS
38245	Atlantic Classics, Second Series	51	SV1	76,638	13,453	6.17	PS
38887	How to Write a Novel: A Practica...	51	SV1	36,995	8,040	6.11	PN
41027	The Revolt of the Star Men	51	-SV3	28,402	5,952	5.66	PS
557	Puck of Pook's Hill	50	-SV6	62,110	8,938	5.93	PR,PZ
794	The Wouldbegoods: Being the Furt...	50	SV5	80,570	9,339	5.97	PR,PZ
953	Tom Swift and His Big Tunnel; Or...	50	-SV1	46,119	6,273	5.83	PZ
1159	Fire-Tongue	50	SV2	64,506	9,211	5.81	PR
1515	The Merchant of Venice	50	SV3	24,592	5,105	6.06	PR
2246	All's Well That Ends Well	50	-SV1	25,941	5,352	6.04	PR
2911	Justice	50	SV2	26,466	4,241	5.75	PR
5066	The Whole Family: a Novel by Twe...	50	SV2	79,321	9,742	6.16	PS
5090	I Will Repay	50	-SV1	64,465	9,401	5.70	PR
5606	Guns of the Gods: A Story of Yas...	50	-SV3	91,027	12,068	6.01	PR
8197	India's Love Lyrics	50	SV1	25,289	4,960	6.03	PR
10850	Philaster; Or, Love Lies a Bleeding	50	-SV3	40,406	6,894	5.87	PR
15454	Imperium in Imperio: A Study of ...	50	-SV1	54,829	8,820	5.93	PS
22132	Giants on the Earth	50	-SV1	25,009	4,267	5.71	PS
30214	The Red Hell of Jupiter	50	-SV1	24,884	5,054	5.34	PS
30852	The Tin Woodman of Oz: A Faithful...	50	SV1	45,503	5,712	6.11	PZ
32542	Dave Dawson on Guadalcanal	50	-SV1	50,526	6,264	5.83	PZ
32730	The Heart of a Woman	50	-SV1	81,281	10,534	5.84	PR
33582	Rhyme? And Reason?	50	-SV1	21,700	5,779	5.83	PR
40241	Hieroglyphics	50	-SV4	45,287	7,771	6.28	PN

Table S1: All Project Gutenberg eBooks considered in this study,
sorted by downloads.

ID	Title	DL's	Mode	N_w	Uniq(N_w)	$h_{\mathrm{avg}}^{b_i}$	LoCC
295	The Early Short Fiction of Edith...	49	SV3	43,439	8,000	5.97	PS
872	Reprinted Pieces	49	-SV8	95,667	15,086	6.02	PR
897	The Rose and the Ring	49	SV2	30,041	5,888	6.05	PR,PZ
1508	The Taming of the Shrew	49	-SV1	24,596	5,184	6.12	PR
1809	Bucky O'Connor: A Tale of the Un...	49	SV2	79,879	10,627	5.97	PS
2257	Richard III	49	-SV2	32,096	6,461	5.53	PR
2906	The Silver Box: A Comedy in Thre...	49	-SV3	21,352	3,358	5.86	PR
4082	The Barrier	49	-SV1	90,906	11,246	5.86	PS
8730	A Little Bush Maid	49	SV5	62,805	9,160	6.07	PZ,PR
15883	The London-Bawd: With Her Charac...	49	SV1	38,798	6,282	6.06	PR
18753	The Space Pioneers	49	-SV1	53,296	6,711	6.01	PZ
19246	The Young Pitcher	49	-SV3	56,278	7,583	5.93	PZ,PS
20859	Wandl the Invader	49	-SV3	48,422	6,885	5.76	PS
21073	A Pirate of the Caribbees	49	-SV4	87,347	10,576	5.89	PZ
24313	Once a Week	49	SV5	78,356	11,024	6.05	PR
25870	A World of Girls: The Story of a...	49	-SV1	83,267	9,687	6.13	PZ
25919	Miss Mapp	49	-SV5	87,493	11,284	6.09	PR
26853	Vice Versa; or, A Lesson to Fathers	49	SV4	96,507	12,652	5.95	PR
30759	Exit Betty	49	SV3	55,532	7,223	6.11	PS
32420	A Yankee Flier with the R.A.F.	49	-SV1	39,240	5,898	5.67	PZ
35920	The Sea Lady	49	-SV1	40,745	7,242	6.07	PR
37012	The Recruiting Officer	49	SV3	21,516	4,396	6.15	PR
37758	Atlantic Classics	49	SV2	67,889	13,413	6.11	PS
41231	The Life and Beauties of Fanny Fern	49	SV4	57,760	12,433	6.20	PS
41715	Dave Dawson with the R.A.F.	49	SV2	48,629	6,587	5.75	PZ
496	The Little Lame Prince	48	-SV4	45,846	6,554	6.09	PZ
980	Alice Adams	48	-SV4	88,191	9,855	6.11	PS
1375	New Chronicles of Rebecca	48	SV1	63,357	9,754	6.08	PZ,PS
1457	Mistress Wilding	48	-SV1	90,095	11,156	5.73	PR
1751	Twilight Land	48	SV5	73,912	6,176	6.10	PZ,PS
1814	The Agony Column	48	-SV1	25,353	4,938	5.99	PS
2024	Diary of a Pilgrimage	48	SV3	43,744	7,806	6.07	PR
2245	The Taming of the Shrew	48	-SV1	23,611	5,008	6.06	PR
2389	Bardelys the Magnificent : Being ...	48	-SV1	76,982	10,070	5.81	PR
4025	Anna Christie	48	-SV1	25,963	4,958	5.92	PS
8899	Three Weeks	48	SV5	52,847	8,067	6.23	PR
14280	Holidays at Roselands : A Sequel ...	48	SV2	93,880	8,107	6.04	PZ
17028	Eastern Standard Tribe	48	SV1	55,676	10,347	5.86	PS
18520	Sabotage in Space	48	SV2	46,857	6,597	5.76	PZ
25003	The Nicest Girl in the School: A...	48	-SV3	55,712	7,364	6.15	PZ
26933	Visions and Revisions: A Book of ...	48	-SV1	53,770	11,101	5.95	PN
34943	Among the Meadow People	48	-SV5	28,691	3,813	6.17	PZ

Table S1: All Project Gutenberg eBooks considered in this study,
sorted by downloads.

ID	Title	DL's	Mode	N_w	Uniq(N_w)	$h_{\text{avg}}^{b_i}$	LoCC
40320	Mr. Punch Afloat: The Humours of...	48	SV1	23,660	6,980	6.14	PN
40493	The King of Diamonds: A Tale of ...	48	-SV2	82,673	12,226	5.96	PR
45658	The Mystery of the Downs	48	-SV7	81,156	8,596	5.78	PR
618	Codex Junius 11	47	SV3	40,603	5,557	6.00	PR
888	The Lazy Tour of Two Idle Appren...	47	-SV2	42,002	7,799	5.99	PR
1264	The Wheels of Chance: A Bicyclin...	47	-SV2	56,518	10,028	6.03	PR
2019	The Bat	47	-SV1	65,484	8,953	5.80	PS
5977	Bound to Rise; Or, Up the Ladder	47	-SV1	48,353	5,861	6.25	PZ
7434	The Adventures of Joel Pepper	47	-SV4	83,873	6,946	5.98	PZ,PS
10119	Adonais	47	-SV1	50,656	9,930	5.68	PR
10882	The Eagle's Shadow	47	-SV1	50,415	9,069	6.04	PS
17112	Many Thoughts of Many Minds: A Tr...	47	-SV2	78,328	14,061	6.03	PN
19360	Six to Sixteen: A Story for Girls	47	-SV4	66,042	10,056	6.22	PZ
19672	The Holladay Case: A Tale	47	-SV3	45,710	6,713	5.86	PS
20739	Rebels of the Red Planet	47	-SV4	47,731	7,116	5.80	PS
20856	Ten From Infinity	47	-SV2	43,848	7,280	5.81	PS
21048	Just Patty	47	-SV2	51,109	9,304	6.14	PZ
26494	Vera; Or, The Nihilists	47	-SV1	20,085	3,981	5.82	PR
29965	Two Thousand Miles Below	47	-SV2	57,790	8,161	5.67	PS
32331	Dave Dawson at Casablanca	47	-SV3	47,663	6,278	5.86	PZ
37173	In a Glass Darkly, v. 2/3	47	-SV2	39,351	7,123	6.06	PR
39116	Unicorns	47	SV3	83,569	17,663	6.09	PS
356	Beyond the City	46	-SV1	40,171	6,536	6.10	PR
1671	When a Man Marries	46	-SV7	56,170	7,879	5.85	PS
4272	The Christian Year	46	SV2	55,144	8,338	5.96	PR
5829	The Moneychangers	46	-SV2	66,414	7,447	6.01	PS
6313	Masterpieces of American Wit and...	46	-SV1	42,501	9,593	5.99	PN
9931	K	46	-SV1	95,884	10,353	5.92	PS
13054	A Thane of Wessex : Being a Story...	46	-SV1	69,893	6,477	5.88	PZ
13694	Mince Pie	46	SV3	53,946	11,656	6.20	PS
14228	Bracebridge Hall	46	-SV1	47,893	9,082	6.26	PS
14360	The Dawn and the Day : Or, The Bu...	46	SV2	43,524	7,764	6.08	PS
15119	Handy Dictionary of Poetical Quo...	46	-SV10	79,135	19,987	5.84	PN
19111	Code Three	46	SV2	22,299	4,629	5.90	PS
26649	Terribly Intimate Portraits	46	-SV2	25,320	7,772	6.08	PR
28849	Smugglers' Reef: A Rick Brant Sc...	46	-SV4	56,101	7,305	6.07	PZ
29416	The Mind Master	46	-SV1	29,245	5,128	5.66	PS
31598	The Egyptian Cat Mystery: A Rick...	46	-SV1	41,234	6,830	6.07	PZ
32398	Brood of the Dark Moon : (A Seque...	46	-SV2	61,907	8,591	5.77	PS
32620	The Three Mulla-mulgars	46	-SV1	70,919	8,914	5.85	PZ
34020	The Window at the White Cat	46	-SV3	69,833	8,799	5.70	PS
37174	In a Glass Darkly, v. 3/3	46	-SV4	34,057	6,291	6.00	PR

Table S1: All Project Gutenberg eBooks considered in this study,
sorted by downloads.

ID	Title	DL's	Mode	N_w	Uniq(N_w)	$h^{b_i}_{\text{avg}}$	LoCC
42250	Dave Dawson with the Commandos	46	-SV1	50,488	6,642	5.78	PZ
875	The Duchess of Padua	45	-SV1	25,724	4,588	5.78	PR
1077	The Mirror of Kong Ho	45	SV5	51,688	9,446	5.98	PR
1281	Tom Swift and His Aerial Warship...	45	SV4	45,747	5,898	5.58	PZ
1423	No Thoroughfare	45	-SV1	50,599	7,678	6.05	PR
1721	The Trees of Pride	45	-SV1	25,426	4,876	5.87	PR
1948	The Story of a Bad Boy	45	-SV3	57,582	10,475	5.99	PS,PZ
2763	The World's Desire	45	-SV3	86,882	8,075	5.63	PR
3179	The American Claimant	45	-SV2	66,257	10,762	5.97	PS
3490	The Admirable Crichton	45	SV1	25,297	4,833	6.15	PR
5008	Katherine's Sheaves	45	SV1	92,258	10,979	6.17	PS
6120	Soldiers Three	45	-SV2	92,097	14,165	5.82	PR
8457	Frenzied Fiction	45	-SV3	49,577	8,240	6.07	PS
9871	The Avenger	45	-SV2	76,075	8,529	5.93	PR
13135	Pardners	45	-SV1	45,724	9,125	5.75	PS
20081	A Houseful of Girls	45	-SV3	94,698	12,317	6.10	PR,PZ
20204	The Storm-Cloud of the Nineteent...	45	SV1	25,388	6,112	6.16	PR
21407	Figures of Several Centuries	45	SV1	78,509	12,961	6.20	PN
22527	Beyond the Vanishing Point	45	-SV2	28,689	4,880	5.75	PS
24770	A Prisoner of Morro; Or, In the ...	45	-SV5	60,682	7,739	5.52	PS
25780	The Fire People	45	SV3	68,632	8,496	5.98	PS
30970	Miss Cayley's Adventures	45	-SV3	86,353	13,470	5.97	PR
40038	The Lone Ranger Rides	45	-SV1	64,363	8,289	5.57	PZ,PS
306	The Early Short Fiction of Edith...	44	SV3	45,131	8,730	6.04	PS
950	Tom Swift and His Electric Runab...	44	-SV3	44,261	5,831	5.89	PZ
954	Tom Swift and His War Tank; Or, ...	44	-SV1	46,396	5,793	5.98	PZ
1828	Chronicles of the Canongate, 1st...	44	-SV1	88,216	14,389	5.98	PR
5901	Dyke Darrel the Railroad Detecti...	44	SV2	57,369	7,619	5.74	PZ,PS
5962	Oh, Money! Money! A Novel	44	SV4	83,980	9,442	6.18	PS
9380	A Nonsense Anthology	44	-SV1	53,854	12,693	5.86	PN
9862	City of Endless Night	44	-SV1	85,809	11,235	6.00	PS
13716	A Trip to Venus: A Novel	44	-SV2	51,029	9,084	6.31	PR
14888	The Inheritors	44	-SV1	63,024	9,745	6.00	PR
16255	Dickey Downy: The Autobiography ...	44	SV3	33,123	6,265	6.13	PZ,QL
16551	The Girl of the Golden West	44	-SV1	75,368	10,075	5.97	PS
17047	The Half-Hearted	44	-SV1	93,430	12,220	5.94	PR
17189	Autumn Leaves : Original Pieces i...	44	SV1	36,797	8,633	6.10	PS
19527	The Yukon Trail: A Tale of the N...	44	SV7	70,079	9,744	5.92	PS
22278	A Master of Mysteries	44	SV1	51,260	7,059	5.83	PR
24025	The New Girl at St. Chad's: A St...	44	-SV1	72,364	10,018	6.10	PZ
25449	The Young Castellan: A Tale of t...	44	-SV1	99,058	9,030	5.79	PR
25496	New Treasure Seekers; Or, The Ba...	44	SV1	70,388	8,885	6.01	PZ

Table S1: All Project Gutenberg eBooks considered in this study,
sorted by downloads.

ID	Title	DL's	Mode	N_w	Uniq(N_w)	$h_{\mathrm{avg}}^{b_i}$	LoCC
26348	Lisbeth Longfrock	44	SV1	33,962	4,880	6.16	PZ
26715	Victorian Songs: Lyrics of the A...	44	-SV1	28,183	6,314	6.20	PR
28434	The Astronomy of Milton's 'Parad...	44	-SV2	83,537	11,616	6.38	PR
32501	The Golden Age	44	-SV3	38,634	8,862	5.92	PR
33348	Reveries over Childhood and Youth	44	-SV3	32,480	5,725	6.19	PR
37667	Three Hours after Marriage	44	SV1	28,921	7,248	6.10	PR
38006	The Heatherford Fortune: a sequel...	44	-SV5	64,355	8,437	6.12	PS
39682	The Idiot at Home	44	-SV1	39,678	7,169	6.15	PS
40603	The Root of All Evil	44	-SV2	80,114	10,078	6.09	PR
534	An Inland Voyage	43	SV2	38,781	7,830	6.20	PR
753	Arizona Nights	43	SV1	69,316	10,153	5.94	PS
1532	The Tragedy of King Lear	43	-SV1	29,831	6,717	5.77	PR
1590	The Amazing Interlude	43	SV5	71,512	8,673	5.76	PS
1908	Her Prairie Knight	43	-SV1	34,425	6,045	5.96	PS
2454	The Silent Bullet	43	SV2	91,190	11,389	5.73	PS
4223	The Mystery of a Hansom Cab	43	-SV3	89,817	10,341	5.81	PR
5121	Dark Hollow	43	SV5	92,271	10,955	5.81	PS
5232	Sejanus: His Fall	43	-SV1	53,886	12,775	5.82	PR
11127	The Case of Jennie Brice	43	-SV2	36,514	4,996	5.77	PS
11583	The Runaway Asteroid	43	SV4	66,662	8,393	5.97	PS
14883	Grandmother Elsie	43	SV1	63,301	7,429	6.30	PZ
17393	Men and Women	43	SV1	38,018	9,312	6.08	PR
20163	The Jolliest School of All	43	SV2	77,103	11,504	6.23	PZ
20795	The Cricket on the Hearth	43	-SV1	33,041	5,995	6.14	PR
21626	Adrift in the Wilds; Or, The Adv...	43	-SV3	63,666	8,138	5.81	PZ
24680	The Martyr of the Catacombs: A Ta...	43	SV3	33,034	5,930	5.55	PS,BV
34403	The Clock Strikes Thirteen	43	-SV2	42,721	6,660	5.96	PZ
35027	Mr. Punch's Railway Book	43	SV1	20,128	5,650	5.89	PN
38052	Reynard the Fox	43	-SV1	28,302	6,412	5.92	PR
38551	The Crux: A Novel	43	-SV1	54,556	9,177	6.22	PS
40814	Ruth Hall: A Domestic Tale of th...	43	SV1	79,856	12,644	6.19	PS
291	The Golden Age	42	-SV7	36,954	8,529	5.91	PR
1122	The Tragedy of Hamlet, Prince of...	42	-SV1	37,349	6,692	5.96	PR
1125	All's Well That Ends Well	42	-SV1	28,367	5,007	6.01	PR
1785	Julius Caesar	42	-SV1	24,321	4,298	5.69	PR
2644	Isaac Bickerstaff, Physician and...	42	SV3	44,818	7,404	6.19	PR
3185	Those Extraordinary Twins	42	-SV1	22,002	4,906	5.86	PS
6428	The Surgeon's Daughter	42	-SV1	68,002	10,881	6.09	PR
9415	Olaf the Glorious: A Story of th...	42	SV5	89,911	8,901	5.77	PZ
10317	Betty Gordon at Boarding School;...	42	-SV1	43,866	7,245	6.17	PZ
14534	Christmas with Grandma Elsie	42	-SV1	69,231	7,538	6.36	PZ
14540	When William Came	42	SV2	52,347	9,911	6.16	PR

Table S1: All Project Gutenberg eBooks considered in this study,
sorted by downloads.

ID	Title	DL's	Mode	N_w	Uniq(N_w)	$h_{\mathrm{avg}}^{b_i}$	LoCC
19718	The Bostonians, Vol. II (of II)	42	-SV1	84,314	9,793	6.04	PS
20104	The Cross-Cut	42	-SV1	82,591	10,003	5.79	PS
20519	Highways in Hiding	42	SV2	84,044	10,607	5.80	PS
21768	A Desert Drama: Being The Traged...	42	-SV3	47,742	7,718	5.79	PR
24283	Down the River; Or, Buck Bradfor...	42	SV4	61,510	7,965	6.00	PZ
25585	Shakespeare, Ben Jonson, Beaumon...	42	-SV1	70,721	13,088	6.04	PR
26732	Free Air	42	SV7	85,921	14,402	6.07	PS
27129	Lyrics from the Song-Books of th...	42	SV2	42,305	7,888	5.96	PR
29774	A Yankee Flier Over Berlin	42	SV2	36,975	5,533	5.77	PZ,PS
30339	Status Quo	42	SV2	26,538	5,150	6.04	PS
30431	Calumet 'K'	42	SV2	66,673	7,053	5.99	PS
32161	Tangle Hold	42	-SV1	20,381	4,254	5.81	PS
32351	Voyage To Eternity	42	-SV2	33,961	6,638	5.86	PS
41753	Dave Dawson at Truk	42	SV5	50,142	6,149	5.74	PZ
42710	Bizarre	42	-SV1	29,868	8,839	5.98	PS
1123	Twelfth Night; Or, What You Will	41	-SV1	24,541	4,431	6.06	PR
2496	Our Village	41	SV2	52,274	10,380	6.24	PR
2702	The Lion's Skin	41	-SV1	84,165	11,056	5.82	PR
2761	Benita, an African romance	41	-SV1	78,214	8,715	5.76	PR
4050	Mates at Billabong	41	-SV1	68,141	9,970	6.09	PR,PZ
4227	Tom Swift and His Wireless Messa...	41	SV2	41,816	5,663	5.94	PS,PZ
4531	The Secret Passage	41	-SV1	92,248	8,518	5.75	PR
5162	Agatha Webb	41	SV2	90,298	9,720	5.87	PS
9190	The Greater Inclination	41	-SV2	55,175	9,195	6.04	PS
10581	Uncle Bernac: A Memory of the Em...	41	-SV2	59,936	8,167	5.98	PR
12215	Odd Craft, Complete	41	-SV3	60,836	7,711	5.98	PR
14203	Varied Types	41	SV2	39,616	7,743	6.08	PR
14875	Elsie's children	41	SV4	72,168	8,880	6.11	PZ
14917	The Wings of the Morning	41	SV3	90,412	13,493	5.83	PR
15717	Books and Persons; Being Comment...	41	-SV3	64,047	11,512	6.13	PN
18614	At the Back of the North Wind	41	SV1	27,095	3,534	6.12	PZ
19307	The Lion of Petra	41	SV5	54,669	8,206	5.80	PR
19819	Milton's Comus	41	SV3	47,977	12,916	6.05	PR
20526	Short Story Writing: A Practical ...	41	-SV1	53,163	9,052	6.16	PN
20989	'A Comedy of Errors' in Seven Acts	41	SV2	23,916	5,762	5.84	PR
21639	When Patty Went to College	41	-SV3	38,793	6,379	6.04	PS
24160	The Basket of Flowers	41	SV4	27,425	4,503	6.08	PT,PZ
25388	The Herapath Property	41	SV4	76,286	9,558	6.01	PR
25581	Rinkitink in Oz	41	-SV3	50,118	6,475	5.92	PZ
30742	Anything You Can Do!	41	SV2	28,315	5,777	5.74	PS
33325	The Spoils of Poynton	41	-SV3	72,258	8,540	6.03	PS
37992	The King of Pirates : Being an Ac...	41	SV2	29,764	4,043	5.99	PR

Table S1: All Project Gutenberg eBooks considered in this study, sorted by downloads.

ID	Title	DL's	Mode	N_w	Uniq(N_w)	$h^{b_i}_{\mathrm{avg}}$	LoCC
38252	Fairies I Have Met	41	-SV1	20,377	2,936	6.33	PZ
38567	Eight Cousins; Or, The Aunt-Hill	41	-SV1	76,013	9,952	6.19	PZ
39782	Brownies and Bogles	41	-SV1	25,440	6,765	6.21	PZ
40263	Folly as It Flies; Hit at by Fan...	41	SV1	90,059	14,570	6.17	PS
40504	Ginger-Snaps	41	SV2	76,388	12,810	6.15	PS
1284	Tom Swift and His Air Scout; Or,...	40	-SV1	47,739	5,894	5.91	PZ
1461	A Legend of Montrose	40	-SV1	90,196	13,422	5.93	PR
1621	Miss or Mrs.?	40	-SV1	31,396	5,596	6.16	PR
1654	An Unsocial Socialist	40	SV2	91,348	12,238	5.92	PR
1987	The Outlet	40	SV6	95,966	10,228	6.02	PS
2013	The Pit Prop Syndicate	40	SV2	97,973	10,059	5.85	PR
2687	The Snare	40	-SV1	82,451	10,571	5.76	PR
3785	In the Reign of Terror: The Adve...	40	SV2	94,244	8,690	5.82	PR,PZ
4735	The Shepherd of the Hills	40	-SV1	76,964	7,916	6.05	PS
7052	Dr. Heidenhoff's Process	40	-SV1	32,887	5,934	5.87	PS
8394	The Doings of Raffles Haw	40	-SV1	38,635	6,651	6.15	PR
10422	Caesar Dies	40	-SV1	50,889	8,731	5.70	PR
10490	The Golden Legend	40	SV1	29,603	6,902	5.73	PS
10723	Betty's Bright Idea; Deacon Pitk...	40	SV3	28,464	5,922	6.20	PS
15976	Puck of Pook's Hill	40	-SV6	59,395	9,037	5.93	PR,PZ
19079	The Adventures of Lightfoot the ...	40	-SV2	21,937	2,930	5.72	PZ
19928	Sunset Pass; or, Running the Gau...	40	-SV1	30,536	5,307	5.79	PS
20147	Rip Foster Rides the Gray Planet	40	SV2	53,257	6,799	5.68	PZ
20472	Grace Harlowe's Plebe Year at Hi...	40	SV4	53,788	7,792	6.05	PZ
21932	Embarrassments	40	-SV1	62,724	8,760	6.10	PS
22892	The Best Made Plans	40	-SV3	44,144	6,329	5.90	PS
24197	The Tinted Venus: A Farcical Rom...	40	-SV3	61,018	8,954	5.98	PR
24767	Jack O' Judgment	40	SV2	73,954	9,201	5.87	PR
25472	Blackbeard: Buccaneer	40	SV1	73,029	11,371	5.83	PZ,PS
33028	Man and Maid	40	-SV2	61,379	9,145	6.03	PR
38053	The Coo-ee Reciter: Humorous, Pa...	40	SV1	27,909	6,950	5.87	PR,PS
47530	Oliver Twist, Vol. 2 (of 3)	40	SV2	51,446	8,271	5.90	PR

Appendix E: Principal Component Analysis (SVD)

In this section we provide a (1) more in-depth, intuitive explanation of the method and (2) more results from the SVD analysis.

In an effort to develop a better intuition for the results of the principal component analysis by way of SVD, we plot Eq. **??** along with representations of the matrices in Fig. S5.

$$A \quad = \quad U \quad \Sigma \quad V^T$$

Figure S5 Schematic of the Singular Value Decomposition applied to emotional arcs of Project Gutenberg books. Shown in A are 10 randomly chosen emotional arcs, in U a "spy" of the matrix, in Σ the decreasing singular values, and in V^T sinusoidal modes. We emphasize that this representation is purely for intuition, as only U is a image of the actual matrix, and A has only 10 of the 1,327 books.

Further, we considered in Eq. **??** the mode coefficient in the matrix W, and in Fig S6 we plot the second line of the equation with W:

$$A \quad = \quad W \quad V^T$$

Figure S6 Schematic of the Singular Value Decomposition applied to emotional arcs of Project Gutenberg books, with $W = U\Sigma$ containing the mode coefficients. Again shown in A are 10 randomly chosen emotional arcs, in W a "spy" of the matrix used in the analysis, and in V^T representative sinusoidal modes.
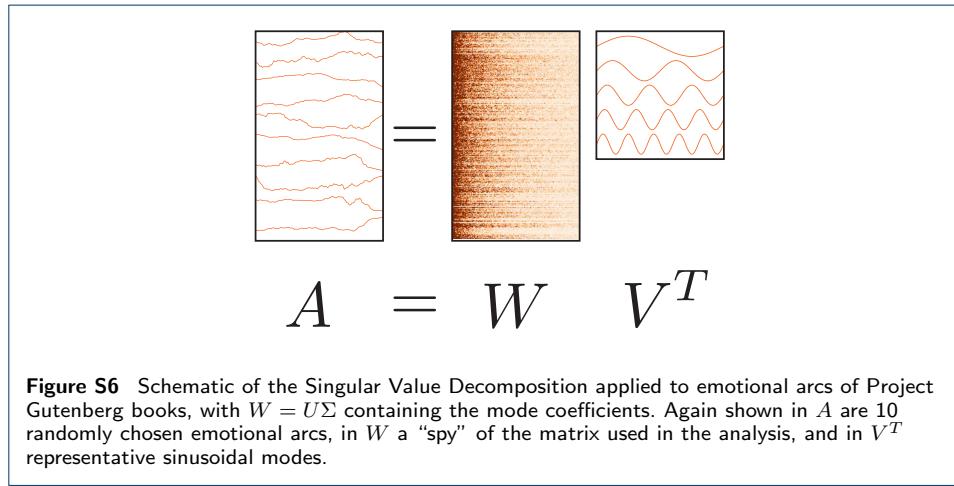
With A written as $W \cdot V^T$, the coefficients for each mode (row of V^T) for a book i are given as the rows of W. To reconstruct the emotional arc of book i, using mode j from V^T, we simply multiply $W[i, j] \cdot V^T[j, :]$. Shown below in Fig. S7, we built the emotional arc for an example story using only the first mode through the first 12 modes.

Figure S7 Reconstruction of the emotional arc from *Alice's Adventures Under Ground*, by Lewis Carroll. The addition of more modes from the SVD more closely reconstructs the detailed emotional arc. This book is well represented by the first mode alone, with only minor corrections from modes 2-11, as we should expect for a book whose emotional arc so closely resembles the "Rags to Riches" arc.

E.1 Additional details for 40 download threshold

First, we consider modes 4–6 and their closest stories in Fig. S8.

Figure S8 SVD modes 4–6 (and their negation) with closest stories. Again, to show the emotional arcs on the same scale as the modes, we show the modes directly from the rows of V^T and weight the emotional arcs by the inverse of their coefficient in W for the particular mode. Shown in parenthesis for each story is the Project Gutenberg ID and the number of downloads from the Project Gutenberg website, respectively. Links below each story point to an interactive visualization on http://hedonometer.org which enables detailed exploration of the emotional arc for the story.

Next, we provide a full list of the books closest to each mode in the analysis, both sorted by downloads and support from the mode. **Note: we are unable to provide a figure accompanying each arc in the following table due to size restrictions on the arXiv.**

Table S2: Stories which are are represented by core emotional arc 1, sorted by the variance explained in their emotional arc by core emotional arc 1.

ID	Title	DL's	$\max(W[i,:])$	Arc
1539	The Winter's Tale	73	0.5217	
33689	Oscar Wilde, Art and Morality: A...	88	0.3804	
35617	The Terror: A Mystery	61	0.3734	
7088	The Pilgrim's Progress in Words ...	55	0.3522	
26624	The Road to Oz	68	0.3412	

Table S2: Stories which are are represented by core emotional arc 1, sorted by the variance explained in their emotional arc by core emotional arc 1.

ID	Title	DL's	max(W[i,:])	Arc
21130	Book of Wise Sayings: Selected La...	86	0.3361	
2248	The Winter's Tale	137	0.3355	
485	The Road to Oz	178	0.3320	
17393	Men and Women	43	0.3030	
36127	Curious Myths of the Middle Ages	91	0.2965	
1121	As You Like It	355	0.2940	
17028	Eastern Standard Tribe	48	0.2935	
27195	Negro Folk Rhymes: Wise and Other...	92	0.2911	
1547	Sir Thomas More	64	0.2808	
2377	The Son of the Wolf	119	0.2755	
35330	The Spanish Tragedy	62	0.2726	
1905	The Governess; Or, The Little Fe...	154	0.2713	
960	The Tin Woodman of Oz : A Faithfu...	132	0.2690	
30852	The Tin Woodman of Oz: A Faithful...	50	0.2680	
20877	Mother West Wind's Children	70	0.2587	
14883	Grandmother Elsie	43	0.2575	
6043	The Spanish Tragedie	389	0.2558	
885	An Ideal Husband	1,303	0.2557	
1790	Troilus and Cressida	63	0.2506	
19551	Alice in Wonderland, Retold in W...	245	0.2489	
29228	The Contrast	88	0.2475	
25496	New Treasure Seekers; Or, The Ba...	44	0.2462	
20726	A Slave is a Slave	78	0.2462	
10002	The House on the Borderland	563	0.2452	
8197	India's Love Lyrics	50	0.2451	
18761	The Circular Study	74	0.2409	
34971	Among the Forest People	65	0.2407	
269	Beasts and Super-Beasts	804	0.2271	
22420	The Book of Nature Myths	91	0.2257	
91	Tom Sawyer Abroad	118	0.2244	
550	Silas Marner	780	0.2217	
14168	Widdershins	112	0.2191	
28521	The Power of Mesmerism: A Highly ...	643	0.2189	
38887	How to Write a Novel: A Practica...	51	0.2169	
15274	The Girl from Montana	126	0.2162	
18614	At the Back of the North Wind	41	0.2148	
11153	No Hero	64	0.2144	
41718	Dave Dawson on the Russian Front	63	0.2134	
19994	The Aesop for Children : With pic...	676	0.2120	
500	The Adventures of Pinocchio	863	0.2108	
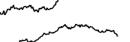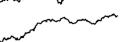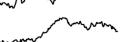
Table S2: Stories which are are represented by core emotional arc 1, sorted by the variance explained in their emotional arc by core emotional arc 1.

ID	Title	DL's	max(W[i,:])	Arc
17189	Autumn Leaves : Original Pieces i...	44	0.2107	
14658	The Road	147	0.2089	
40320	Mr. Punch Afloat: The Humours of...	48	0.2084	
30408	The Fifth-Dimension Tube	118	0.2078	
17513	St. Nicholas Magazine for Boys a...	67	0.2074	
788	The Red One	100	0.2068	
215	The Call of the Wild	2,439	0.2046	
14902	Deadwood Dick, the Prince of the...	58	0.2043	
4006	Yesterdays	68	0.2001	
12352	Iola Leroy; Or, Shadows Uplifted	66	0.1999	
32597	Accidental Flight	63	0.1991	
38053	The Coo-ee Reciter: Humorous, Pa...	40	0.1961	
3543	Heartbreak House	150	0.1953	
21084	Jokes For All Occasions: Selected...	281	0.1939	
10551	Affair in Araby	51	0.1915	
2431	Is Shakespeare Dead? : From My Au...	78	0.1883	
1557	Men of Iron	99	0.1877	
28071	The Red Triangle: Being Some Furt...	74	0.1868	
20204	The Storm-Cloud of the Nineteent...	45	0.1855	
367	The Country of the Pointed Firs	114	0.1834	
25803	The Keepers of the King's Peace	98	0.1821	
1375	New Chronicles of Rebecca	48	0.1808	
5676	A Double Story	62	0.1808	
22495	The New Pun Book	83	0.1785	
40263	Folly as It Flies; Hit at by Fan...	41	0.1758	
21407	Figures of Several Centuries	45	0.1745	
22278	A Master of Mysteries	44	0.1736	
18881	The Idiot	57	0.1734	
236	The Jungle Book	3,478	0.1728	
554	The Contrast	144	0.1722	
2250	Richard II	78	0.1698	
5008	Katherine's Sheaves	45	0.1694	
21665	A Brief History of the English L...	62	0.1663	
17157	Gulliver's Travels into Several ...	528	0.1643	
10150	Dracula's Guest	456	0.1631	
35997	The Jungle Book	1,370	0.1623	
27922	David and the Phoenix	56	0.1611	
11377	The Man Whom the Trees Loved	56	0.1606	
1869	The Man in Lower Ten	66	0.1595	
5348	Ragged Dick, Or, Street Life in ...	378	0.1588	
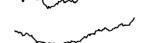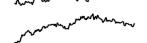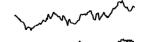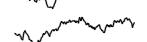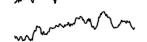
Table S2: Stories which are are represented by core emotional arc 1, sorted by the variance explained in their emotional arc by core emotional arc 1.

ID	Title	DL's	max(W[i,:])	Arc
3490	The Admirable Crichton	45	0.1585	
20630	The Borough Treasurer	65	0.1577	
33156	Young's Night Thoughts: With Life...	111	0.1557	
25472	Blackbeard: Buccaneer	40	0.1552	
12245	The Defendant	76	0.1551	
4980	Old Granny Fox	128	0.1550	
29468	The Story of Don Quixote	88	0.1545	
21816	The Confidence-Man: His Masquerade	289	0.1503	
961	Glinda of Oz : In Which Are Relat...	184	0.1498	
26348	Lisbeth Longfrock	44	0.1487	
3781	The Jewel of Seven Stars	160	0.1486	
39868	Glinda of Oz : In which are Relat...	53	0.1476	
47529	Oliver Twist, Vol. 1 (of 3)	105	0.1469	
3011	The Lady of the Lake	399	0.1455	
21530	The Angel of Terror	169	0.1452	
8668	Revenge!	111	0.1442	
3797	In the Days of the Comet	80	0.1441	
753	Arizona Nights	43	0.1427	
3756	Indiscretions of Archie	162	0.1397	
35027	Mr. Punch's Railway Book	43	0.1391	
28267	Venus in Boston: A Romance of Ci...	55	0.1381	
2770	Five Little Peppers and How They...	207	0.1379	
2015	A Miscellany of Men	61	0.1369	
15883	The London-Bawd: With Her Charac...	49	0.1363	
40723	The Battle of Life. A Love Story	68	0.1350	
78	Tarzan of the Apes	1,272	0.1341	
837	The Story of the Amulet	120	0.1326	
37995	The Diamond Fairy Book	55	0.1318	
35612	Three Philosophical Poets: Lucre...	78	0.1299	
38245	Atlantic Classics, Second Series	51	0.1292	
501	The Story of Doctor Dolittle	504	0.1268	
32706	Triplanetary	309	0.1213	
40814	Ruth Hall: A Domestic Tale of th...	43	0.1212	
1076	The Wallet of Kai Lung	59	0.1205	
37667	Three Hours after Marriage	44	0.1181	
10490	The Golden Legend	40	0.1159	
16389	The Enchanted April	170	0.1153	
42254	Beyond the Black River	141	0.1152	
173	The Insidious Dr. Fu Manchu	245	0.1142	
779	The Tragical History of Doctor F...	2,133	0.1133	

Table S2: Stories which are are represented by core emotional arc 1, sorted by the variance explained in their emotional arc by core emotional arc 1.

ID	Title	DL's	max(W[i,:])	Arc
4282	Don Rodriguez; Chronicles of Sha...	56	0.1103	
873	A House of Pomegranates	172	0.1092	
15851	Love Conquers All	102	0.0987	
844	The Importance of Being Earnest:...	9,373	0.0963	
40426	Daddy Long-Legs: A Comedy in Fou...	192	0.0926	
1929	The School for Scandal	417	0.0907	
16732	Familiar Quotations	86	0.0800	
10110	The Postmaster's Daughter	60	0.0766	
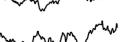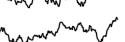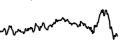
Table S3: Top 10 stories which are are represented by core emotional arc 1, sorted by downloads.

ID	Title	DL's	max(W[i,:])	Arc
844	The Importance of Being Earnest:...	9,373	0.0963	
236	The Jungle Book	3,478	0.1728	
215	The Call of the Wild	2,439	0.2046	
779	The Tragical History of Doctor F...	2,133	0.1133	
35997	The Jungle Book	1,370	0.1623	
885	An Ideal Husband	1,303	0.2557	
78	Tarzan of the Apes	1,272	0.1341	
500	The Adventures of Pinocchio	863	0.2108	
269	Beasts and Super-Beasts	804	0.2271	
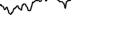
Table S4: Stories which are are represented by core emotional arc 2, sorted by the variance explained in their emotional arc by core emotional arc 2.

ID	Title	DL's	max(W[i,:])	Arc
946	Lady Susan	894	0.5006	
17958	Warlord of Kor	70	0.4664	
17144	The House of the Vampire	188	0.4657	
93	Tom Sawyer, Detective	112	0.4566	
159	The Island of Doctor Moreau	1,083	0.4538	
32208	The Star Lord	58	0.4504	
790	Lady Windermere's Fan	485	0.4257	
1777	Romeo and Juliet	186	0.4013	
875	The Duchess of Padua	45	0.3945	
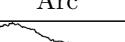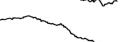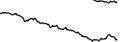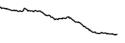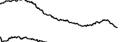
Table S4: Stories which are are represented by core emotional arc 2, sorted by the variance explained in their emotional arc by core emotional arc 2.

ID	Title	DL's	max(W[i,:])	Arc
4025	Anna Christie	48	0.3934	
39782	Brownies and Bogles	41	0.3913	
30796	The Dueling Machine	117	0.3891	
311	Bunner Sisters	82	0.3866	
38252	Fairies I Have Met	41	0.3835	
21632	Fame and Fortune; or, The Progre...	56	0.3744	
267	The Touchstone	64	0.3732	
1531	Othello, the Moor of Venice	147	0.3729	
42259	The People of the Black Circle	132	0.3706	
16517	Liza of Lambeth	107	0.3676	
30214	The Red Hell of Jupiter	50	0.3621	
16199	Memoirs of the Author of a Vindi...	69	0.3621	
2266	King Lear	74	0.3605	
23790	The Ultimate Weapon	126	0.3596	
35920	The Sea Lady	49	0.3591	
7052	Dr. Heidenhoff's Process	40	0.3564	
2265	Hamlet	1,051	0.3527	
451	The Shadow Line: A Confession	127	0.3483	
1719	The Ballad of the White Horse	394	0.3481	
20656	Old Christmas From the Sketch Bo...	101	0.3451	
21334	The Beaux-Stratagem	86	0.3449	
949	Tom Swift and His Submarine Boat...	52	0.3439	
1127	The Tragedy of Othello, Moor of ...	53	0.3438	
4358	The Sea Fairies	98	0.3424	
38052	Reynard the Fox	43	0.3396	
2865	Otto of the Silver Hand	58	0.3383	
33660	The Year When Stardust Fell	106	0.3377	
2232	The Duchess of Malfi	534	0.3375	
14255	Hints for Lovers	127	0.3374	
18800	Last Enemy	83	0.3320	
20058	The Napoleon of Notting Hill	182	0.3319	
2263	Julius Caesar	150	0.3319	
5429	Preface to Shakespeare	148	0.3306	
24749	Adaptation	67	0.3298	
18420	The Bobbsey Twins at Home	67	0.3291	
1882	The Young Forester	59	0.3260	
20121	Lone Star Planet	103	0.3234	
42250	Dave Dawson with the Commandos	46	0.3217	
35	The Time Machine	3,732	0.3199	
24302	The Highest Treason	67	0.3176	
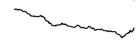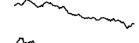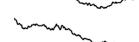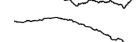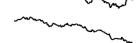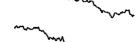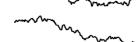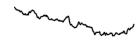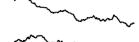
Table S4: Stories which are are represented by core emotional arc 2, sorted by the variance explained in their emotional arc by core emotional arc 2.

ID	Title	DL's	max(W[i,:])	Arc
1621	Miss or Mrs.?	40	0.3156	
34592	Behind the Green Door	58	0.3138	
6879	The Gold Bat	120	0.3128	
1041	Shakespeare's Sonnets	831	0.3125	
2245	The Taming of the Shrew	48	0.3121	
94	Alexander's Bridge	58	0.3121	
19928	Sunset Pass; or, Running the Gau...	40	0.3073	
32154	The Variable Man	618	0.3059	
12915	The White Devil	116	0.3057	
172	The Haunted Bookshop	132	0.3052	
4023	Candida	126	0.3039	
5311	Parnassus on Wheels	62	0.3021	
24370	Mercenary	58	0.3011	
1123	Twelfth Night; Or, What You Will	41	0.3007	
22145	A Book of Burlesques	80	0.2992	
11074	The Damned	241	0.2966	
1122	The Tragedy of Hamlet, Prince of...	42	0.2958	
1787	Hamlet	361	0.2936	
1423	No Thoroughfare	45	0.2909	
1118	Much Ado about Nothing	183	0.2896	
179	The Europeans	85	0.2895	
5341	Kilmeny of the Orchard	102	0.2886	
25770	The Dragon's Secret	84	0.2867	
32226	The Flower Princess	57	0.2839	
1142	Typhoon	219	0.2833	
791	The Princess	245	0.2812	
25767	Picture and Text: 1893	59	0.2810	
15454	Imperium in Imperio: A Study of ...	50	0.2809	
1814	The Agony Column	48	0.2803	
5071	The Philanderer	63	0.2799	
18492	Star Surgeon	83	0.2799	
4050	Mates at Billabong	41	0.2792	
5090	I Will Repay	50	0.2789	
7947	The Diary of a U-boat Commander: ...	71	0.2781	
60	The Scarlet Pimpernel	710	0.2777	
3777	Tom Swift and His Electric Rifle...	70	0.2748	
2267	Othello	760	0.2710	
33979	Miscellaneous Aphorisms; The Sou...	52	0.2705	
8931	The Gem Collector	107	0.2698	
25585	Shakespeare, Ben Jonson, Beaumon...	42	0.2692	
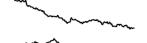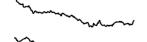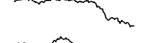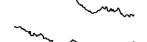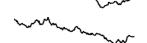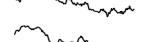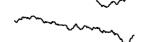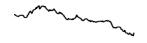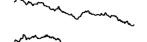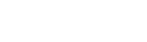
Table S4: Stories which are are represented by core emotional arc 2, sorted by the variance explained in their emotional arc by core emotional arc 2.

| ID	Title	DL's	max($W[\mathrm{i},:]$)	Arc
18346	Null-ABC	67	0.2663	
17226	Emily Fox-Seton : Being "The Maki...	115	0.2650	
2268	Antony and Cleopatra	125	0.2642	
5977	Bound to Rise; Or, Up the Ladder	47	0.2641	
1283	Tom Swift and His Wizard Camera;...	52	0.2632	
9931	K	46	0.2626	
6574	Watchers of the Sky	63	0.2620	
21865	King Arthur and His Knights	61	0.2616	
526	Heart of Darkness	4,362	0.2597	
586	Religio Medici, Hydriotaphia, an...	71	0.2592	
20387	A Thin Ghost and Others	155	0.2581	
11012	The Autobiography of an Ex-Color...	509	0.2580	
9932	The Last Trail	117	0.2572	
42710	Bizarre	42	0.2571	
23810	At Fault	104	0.2566	
39682	The Idiot at Home	44	0.2558	
22132	Giants on the Earth	50	0.2547	
37431	Pride and Prejudice, a play foun...	111	0.2541	
1107	The Taming of the Shrew	383	0.2540	
1953	A Book of Strife in the Form of ...	83	0.2534	
19381	Among the Farmyard People	60	0.2533	
2870	Washington Square	285	0.2511	
21927	Short Cruises	84	0.2498	
3529	Letters Written During a Short R...	236	0.2492	
19027	The Revolt on Venus	64	0.2482	
17797	Memoir of Jane Austen	203	0.2472	
37172	In a Glass Darkly, v. 1/3	82	0.2470	
3185	Those Extraordinary Twins	42	0.2468	
36612	The Princess and Curdie	57	0.2466	
3825	Pygmalion	3,580	0.2446	
8677	Behind a Mask; or, a Woman's Power	134	0.2444	
652	Rasselas, Prince of Abyssinia	241	0.2441	
6313	Masterpieces of American Wit and...	46	0.2426	
15422	Israel Potter : His Fifty Years o...	58	0.2410	
32953	Quest of the Golden Ape	80	0.2406	
1244	Love for Love: A Comedy	71	0.2401	
18970	Caves of Terror	80	0.2400	
20730	For the Sake of the School	82	0.2393	
24280	Endymion: A Poetic Romance	133	0.2391	
2295	Waifs and Strays: Part 1	52	0.2387	
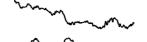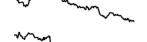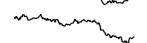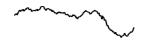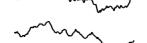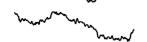
Table S4: Stories which are are represented by core emotional arc 2, sorted by the variance explained in their emotional arc by core emotional arc 2.

ID	Title	DL's	max(W[i,:])	Arc
1239	The Spirit of the Border: A Roma...	69	0.2383	
1424	Castle Rackrent	203	0.2381	
24025	The New Girl at St. Chad's: A St...	44	0.2380	
42	The Strange Case of Dr. Jekyll a...	4,908	0.2379	
31501	The Sensitive Man	146	0.2369	
12431	The Coquette, or, The History of...	54	0.2349	
14654	A Daughter of the Snows	51	0.2345	
32486	The Legion of Lazarus	61	0.2344	
709	The Princess and Curdie	146	0.2336	
2019	The Bat	47	0.2320	
35204	Sense of Obligation	54	0.2320	
953	Tom Swift and His Big Tunnel; Or...	50	0.2295	
29310	The Affair of the Brains	86	0.2289	
5312	Mother Goose in Prose	97	0.2289	
1508	The Taming of the Shrew	49	0.2288	
11935	Mysticism in English Literature	65	0.2284	
18753	The Space Pioneers	49	0.2278	
8188	The Mysterious Key and What It O...	83	0.2270	
1461	A Legend of Montrose	40	0.2269	
653	The Chimes : A Goblin Story of So...	174	0.2268	
1532	The Tragedy of King Lear	43	0.2267	
4037	Appreciations, with an Essay on ...	100	0.2253	
20988	Islands of Space	109	0.2252	
1908	Her Prairie Knight	43	0.2249	
2604	The Longest Journey	88	0.2249	
3158	Our American Cousin	85	0.2242	
7964	The Mystery of Cloomber	108	0.2237	
30427	The Lost Kafoozalum	74	0.2232	
1604	The Ebb-Tide: A Trio And Quartette	56	0.2225	
31598	The Egyptian Cat Mystery: A Rick...	46	0.2224	
7239	Men, Women, and Boats	57	0.2218	
6936	Robinson Crusoe — in Words of On...	56	0.2214	
24353	Wired Love: A Romance of Dots and...	76	0.2208	
23625	The Magic Pudding	133	0.2206	
25438	The Airlords of Han	83	0.2205	
43	The Strange Case of Dr. Jekyll a...	599	0.2204	
26740	The Picture of Dorian Gray	257	0.2203	
18817	Ralestone Luck	75	0.2194	
1524	Hamlet, Prince of Denmark	2,329	0.2193	
219	Heart of Darkness	3,243	0.2184	
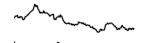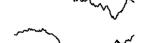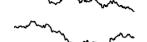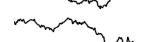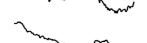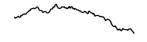
Table S4: Stories which are are represented by core emotional arc 2, sorted by the variance explained in their emotional arc by core emotional arc 2.

ID	Title	DL's	$\max(W[i,:])$	Arc
18361	Operation: Outer Space	81	0.2184	
7371	A Sicilian Romance	112	0.2180	
95	The Prisoner of Zenda	339	0.2169	
171	Charlotte Temple	337	0.2168	
1338	Selected Prose of Oscar Wilde	141	0.2157	
20532	Love Among the Chickens: A Story ...	100	0.2145	
10119	Adonais	47	0.2130	
3244	To Him That Hath: A Tale of the ...	89	0.2128	
2702	The Lion's Skin	41	0.2124	
1828	Chronicles of the Canongate, 1st...	44	0.2123	
6359	The English Mail-Coach and Joan ...	62	0.2119	
558	The Thirty-Nine Steps	989	0.2118	
21873	Planet of the Damned	189	0.2116	
25016	The House of Souls	362	0.2114	
24761	The Rivals: A Comedy	408	0.2102	
13135	Pardners	45	0.2099	
19090	Star Hunter	185	0.2086	
1794	King Lear	126	0.2082	
1128	The Tragedy of King Lear	548	0.2082	
33644	The Secret of the Ninth Planet	109	0.2078	
1153	The Chessmen of Mars	409	0.2077	
25024	The Night of the Long Knives	180	0.2069	
9380	A Nonsense Anthology	44	0.2068	
28346	Deathworld	287	0.2063	
29042	A Tangled Tale	142	0.2061	
678	The Cricket on the Hearth: A Fai...	77	0.2057	
20795	The Cricket on the Hearth	43	0.2055	
720	Almayer's Folly: A Story of an E...	67	0.2053	
10882	The Eagle's Shadow	47	0.2051	
1144	In the Cage	85	0.2048	
943	Misalliance	87	0.2036	
32161	Tangle Hold	42	0.2035	
8713	A Man of Means	141	0.2034	
11252	Martin Hewitt, Investigator	118	0.2033	
2060	The History of Caliph Vathek	202	0.2031	
28520	Forbidden Fruit: Luscious and exc...	2,716	0.2028	
1125	All's Well That Ends Well	42	0.2022	
14888	The Inheritors	44	0.2019	
1120	The Tragedy of Julius Caesar	496	0.2019	
1785	Julius Caesar	42	0.2017	
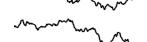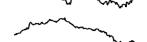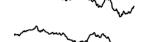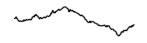
Table S4: Stories which are are represented by core emotional arc 2, sorted by the variance explained in their emotional arc by core emotional arc 2.

ID	Title	DL's	max(W[i,:])	Arc
2761	Benita, an African romance	41	0.2009	
1472	In a German Pension	65	0.2006	
32420	A Yankee Flier with the R.A.F.	49	0.2004	
356	Beyond the City	46	0.2003	
19474	Uller Uprising	87	0.2000	
11371	The Moorland Cottage	53	0.1994	
8995	What Katy Did Next	105	0.1992	
12345	Friday, the Thirteenth : A Novel	59	0.1972	
17866	Murder in the Gunroom	267	0.1965	
14534	Christmas with Grandma Elsie	42	0.1961	
2852	The Hound of the Baskervilles	3,358	0.1951	
6984	The Pothunters	131	0.1943	
4075	The Intrusion of Jimmy	90	0.1940	
3070	The Hound of the Baskervilles	549	0.1940	
8435	The Sturdy Oak : A composite Nove...	58	0.1931	
4081	The Alchemist	744	0.1931	
32563	The Lost Warship	63	0.1927	
2389	Bardelys the Magnificent : Being ...	48	0.1922	
296	The Cash Boy	65	0.1922	
966	Maid Marian	62	0.1918	
16740	The Busie Body	51	0.1914	
3638	The Devil's Disciple	100	0.1904	
174	The Picture of Dorian Gray	7,652	0.1893	
25449	The Young Castellan: A Tale of t...	44	0.1893	
32530	Armageddon—2419 A.D.	313	0.1893	
11228	The Marrow of Tradition	178	0.1879	
7308	The History of Mr. Polly	79	0.1869	
555	The Unbearable Bassington	87	0.1868	
25776	This Crowded Earth	163	0.1867	
16259	The Surprising Adventures of the...	62	0.1866	
572	The Great Big Treasury of Beatri...	307	0.1864	
27174	Captain Jim	55	0.1854	
19478	Four-Day Planet	60	0.1853	
20707	The Black Star Passes	147	0.1853	
2667	The Vicar of Wakefield	238	0.1852	
29827	The Life and Amours of the Beaut...	678	0.1852	
27924	Mugby Junction	67	0.1850	
954	Tom Swift and His War Tank; Or, ...	44	0.1847	
4082	The Barrier	49	0.1845	
1605	The Crock of Gold	92	0.1841	
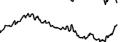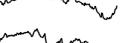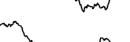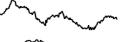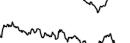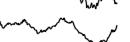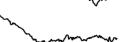
Table S4: Stories which are are represented by core emotional arc 2, sorted by the variance explained in their emotional arc by core emotional arc 2.

ID	Title	DL's	max(W[i,:])	Arc
3829	Love Among the Chickens	198	0.1836	
984	Who Was Who: 5000 B. C. to Date ...	127	0.1833	
611	Prester John	152	0.1826	
17180	The Riddle of the Frozen Flame	64	0.1825	
10317	Betty Gordon at Boarding School;...	42	0.1822	
4020	Arcadian Adventures with the Idl...	55	0.1814	
4230	Tom Swift and His Motor-Cycle; O...	150	0.1810	
149	The Lost Continent	200	0.1810	
17125	More William	85	0.1808	
2876	The Light That Failed	61	0.1807	
11247	The Exploits of Brigadier Gerard	98	0.1803	
6428	The Surgeon's Daughter	42	0.1797	
17047	The Half-Hearted	44	0.1796	
5747	Do and Dare — a Brave Boy's Figh...	61	0.1794	
1155	The Secret Adversary	2,070	0.1787	
19718	The Bostonians, Vol. II (of II)	42	0.1779	
8164	My Man Jeeves	1,317	0.1779	
2948	Where Angels Fear to Tread	202	0.1777	
436	The Master Key: An Electrical Fa...	61	0.1771	
2246	All's Well That Ends Well	50	0.1767	
37698	Dawn of the Morning	87	0.1742	
4078	The Picture of Dorian Gray	565	0.1742	
27903	The Magic World	67	0.1742	
7118	What Maisie Knew	236	0.1739	
545	At the Earth's Core	79	0.1733	
37503	Gammer Gurton's Needle	65	0.1723	
33623	The Inventions of the Idiot	60	0.1723	
8394	The Doings of Raffles Haw	40	0.1722	
32730	The Heart of a Woman	50	0.1719	
139	The Lost World	1,274	0.1717	
3475	The Efficiency Expert	53	0.1707	
14107	The Lost Stradivarius	84	0.1702	
836	The Phoenix and the Carpet	121	0.1700	
123	At the Earth's Core	296	0.1695	
3048	The Little Duke: Richard the Fea...	105	0.1693	
942	Green Mansions: A Romance of the...	83	0.1686	
6684	Uneasy Money	161	0.1683	
39957	Prairie Gold	103	0.1682	
39143	The Making of a Saint	65	0.1680	
19370	Ullr Uprising	51	0.1676	
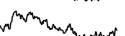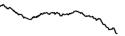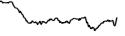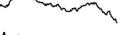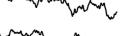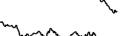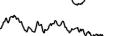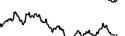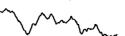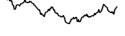
Table S4: Stories which are are represented by core emotional arc 2, sorted by the variance explained in their emotional arc by core emotional arc 2.

ID	Title	DL's	max(W[i,:])	Arc
33582	Rhyme? And Reason?	50	0.1675	
8183	Time and the Gods	123	0.1672	
4268	Cousin Phillis	51	0.1659	
16551	The Girl of the Golden West	44	0.1655	
1460	The Black Dwarf	59	0.1653	
1721	The Trees of Pride	45	0.1647	
6836	Three Men and a Maid	141	0.1643	
14257	The Magician	159	0.1632	
21092	On the Trail of the Space Pirates	54	0.1614	
5333	Every Man in His Humor	100	0.1612	
792	Wieland; Or, The Transformation:...	303	0.1608	
20857	Spacehounds of IPC	102	0.1606	
208	Daisy Miller: A Study	1,101	0.1603	
25550	The Defiant Agents	225	0.1600	
23893	The Melting-Pot	172	0.1591	
1284	Tom Swift and His Air Scout; Or,...	40	0.1590	
18095	Successful Methods of Public Spe...	68	0.1589	
16328	Beowulf : An Anglo-Saxon Epic Poem	5,359	0.1584	
10422	Caesar Dies	40	0.1579	
10373	The Middle Temple Murder	134	0.1570	
24737	The Children of Odin: The Book o...	352	0.1568	
27444	Starman's Quest	127	0.1542	
29416	The Mind Master	46	0.1536	
20840	Rebel Spurs	146	0.1536	
40038	The Lone Ranger Rides	45	0.1531	
2046	Clotel; Or, The President's Daug...	144	0.1529	
4735	The Shepherd of the Hills	40	0.1529	
13969	The Hill of Dreams	131	0.1529	
102	The Tragedy of Pudd'nhead Wilson	1,140	0.1528	
10847	The Maids Tragedy	53	0.1526	
5342	The Story Girl	220	0.1521	
4531	The Secret Passage	41	0.1517	
1327	Elizabeth and Her German Garden	78	0.1515	
38551	The Crux: A Novel	43	0.1514	
4039	Volpone; Or, The Fox	558	0.1502	
14427	True Love's Reward : A Sequel to ...	57	0.1487	
770	The Story of the Treasure Seeker...	185	0.1487	
21932	Embarrassments	40	0.1485	
25870	A World of Girls: The Story of a...	49	0.1485	
20717	The Girl on the Boat	189	0.1485	

Table S4: Stories which are are represented by core emotional arc 2, sorted by the variance explained in their emotional arc by core emotional arc 2.

ID	Title	DL's	max(W[i,:])	Arc
5232	Sejanus: His Fall	43	0.1476	
6418	Five Little Peppers and their Fr...	51	0.1475	
2815	Democracy, an American novel	67	0.1469	
20796	The Colors of Space	245	0.1466	
9909	Nightmare Abbey	180	0.1466	
14317	The Sorcery Club	91	0.1464	
9963	Elsie's Girlhood: A Sequel to "El...	58	0.1456	
805	This Side of Paradise	1,122	0.1455	
5815	The Great Impersonation	51	0.1450	
1457	Mistress Wilding	48	0.1444	
19141	Edison's Conquest of Mars	77	0.1436	
2548	The Poor Clare	85	0.1430	
31343	The Invaders	56	0.1429	
325	Phantastes: A Faerie Romance for...	461	0.1417	
5148	Rodney Stone	72	0.1413	
707	Raffles: Further Adventures of t...	105	0.1412	
1937	The Second Jungle Book	304	0.1402	
111	Freckles	149	0.1394	
9862	City of Endless Night	44	0.1383	
535	Travels with a Donkey in the Cev...	179	0.1381	
2726	Eight Cousins	214	0.1381	
20104	The Cross-Cut	42	0.1374	
5340	Further Chronicles of Avonlea	195	0.1372	
1696	The Club of Queer Trades	119	0.1363	
2042	Something New	384	0.1354	
981	Beowulf	718	0.1344	
37660	Of All Things	83	0.1340	
8223	Edgar Huntly; or, Memoirs of a S...	184	0.1337	
605	Pellucidar	231	0.1328	
479	Little Lord Fauntleroy	246	0.1321	
25067	The Planet Strappers	64	0.1315	
14228	Bracebridge Hall	46	0.1306	
17221	History of the Plague in London	82	0.1303	
13937	The Mysterious Rider	185	0.1296	
1091	On Heroes, Hero-Worship, and the...	622	0.1289	
2324	A House to Let	74	0.1286	
32620	The Three Mulla-mulgars	46	0.1279	
556	Rewards and Fairies	68	0.1276	
9791	Harrigan	56	0.1264	
11505	All Things Considered	485	0.1242	

Table S4: Stories which are are represented by core emotional arc 2, sorted by the variance explained in their emotional arc by core emotional arc 2.

ID	Title	DL's	max(W[i,:])	Arc
7464	The Adventures of Sally	282	0.1237	
42243	The Hour of the Dragon	247	0.1237	
13054	A Thane of Wessex : Being a Story...	46	0.1235	
14034	King Alfred's Viking: A Story of ...	63	0.1230	
37820	Chronicles of Martin Hewitt	67	0.1229	
864	The Master of Ballantrae: A Wint...	93	0.1226	
10869	The Abandoned Room	67	0.1226	
11128	The Red Thumb Mark	155	0.1222	
37364	The Second Jungle Book	144	0.1211	
2885	The House of the Wolfings : A Tal...	105	0.1205	
2785	The Elusive Pimpernel	105	0.1202	
393	The Blue Lagoon: A Romance	129	0.1196	
20526	Short Story Writing: A Practical ...	41	0.1187	
38567	Eight Cousins; Or, The Aunt-Hill	41	0.1181	
20288	Edward the Second	204	0.1167	
764	Hans Brinker; Or, The Silver Skates	62	0.1163	
604	Gulliver of Mars	126	0.1158	
32542	Dave Dawson on Guadalcanal	50	0.1152	
1640	Lilith: A Romance	281	0.1143	
2687	The Snare	40	0.1137	
7230	Not George Washington — an Autob...	68	0.1131	
15281	Uncle Wiggily's Adventures	89	0.1127	
26998	Peter Pan in Kensington Gardens	201	0.1110	
26933	Visions and Revisions: A Book of ...	48	0.1102	
12170	The Wolf Hunters: A Tale of Adve...	66	0.1084	
832	Robin Hood	148	0.1077	
8092	Tremendous Trifles	352	0.1056	
1026	The Diary of a Nobody	329	0.1037	
33735	Pamela Censured	115	0.0998	
26494	Vera; Or, The Nihilists	47	0.0970	
1145	Rupert of Hentzau: From The Memo...	106	0.0947	
10556	The Old Man in the Corner	285	0.0917	
29466	Lords of the Stratosphere	51	0.0896	
26715	Victorian Songs: Lyrics of the A...	44	0.0894	
6877	The Head of Kay's	97	0.0867	
32202	The Irish Fairy Book	168	0.0759	
18824	Fairies and Folk of Ireland	62	0.0746	
11620	My Brilliant Career	62	0.0745	
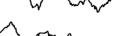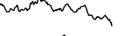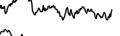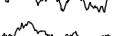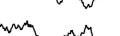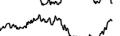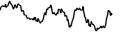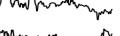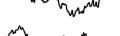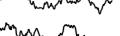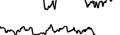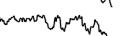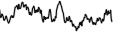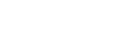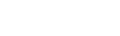
Table S5: Top 10 stories which are are represented by core emotional arc 2, sorted by downloads.

ID	Title	DL's	max(W[i,:])	Arc
174	The Picture of Dorian Gray	7,652	0.1893	
16328	Beowulf : An Anglo-Saxon Epic Poem	5,359	0.1584	
42	The Strange Case of Dr. Jekyll a...	4,908	0.2379	
526	Heart of Darkness	4,362	0.2597	
35	The Time Machine	3,732	0.3199	
3825	Pygmalion	3,580	0.2446	
2852	The Hound of the Baskervilles	3,358	0.1951	
219	Heart of Darkness	3,243	0.2184	
28520	Forbidden Fruit: Luscious and exc...	2,716	0.2028	
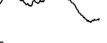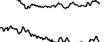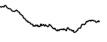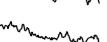
Table S6: Stories which are are represented by core emotional arc 3, sorted by the variance explained in their emotional arc by core emotional arc 3.

ID	Title	DL's	max(W[i,:])	Arc
419	The Magic of Oz	186	0.3150	
10736	Children of the Frost	82	0.3083	
1094	Tamburlaine the Great — Part 1	474	0.3081	
520	The Life and Adventures of Santa...	76	0.3041	
2911	Justice	50	0.2989	
18768	The Sky Is Falling	113	0.2978	
29774	A Yankee Flier Over Berlin	42	0.2864	
19726	The Door Through Space	201	0.2826	
4087	An Essay Upon Projects	101	0.2745	
28118	The Great Gray Plague	76	0.2706	
5083	The Man of Feeling	60	0.2664	
32746	The Revival of Irish Literature ...	54	0.2636	
37992	The King of Pirates : Being an Ac...	41	0.2635	
20559	R. Holmes & Co. : Being the Remar...	55	0.2625	
2814	Dubliners	4,742	0.2583	
21970	The Scarlet Plague	192	0.2512	
10671	The Botanic Garden. Part II.: Con...	78	0.2484	
18137	Little Fuzzy	171	0.2430	
28215	Empire	214	0.2391	
11435	Small Means and Great Ends	61	0.2390	
20519	Highways in Hiding	42	0.2356	
10337	Lady into Fox	53	0.2351	
687	A Personal Record	73	0.2334	
1282	Tom Swift Among the Diamond Make...	58	0.2288	
956	Tik-Tok of Oz	163	0.2275	
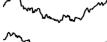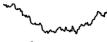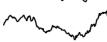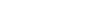
Table S6: Stories which are are represented by core emotional arc 3, sorted by the variance explained in their emotional arc by core emotional arc 3.

ID	Title	DL's	max(W[i,:])	Arc
14280	Holidays at Roselands : A Sequel ...	48	0.2259	
534	An Inland Voyage	43	0.2216	
6440	Elsie Dinsmore	100	0.2207	
22031	The Airplane Boys among the Clou...	61	0.2197	
30742	Anything You Can Do!	41	0.2178	
6985	A Prefect's Uncle	145	0.2177	
24933	The Man Who Knew	68	0.2123	
20919	The Status Civilization	145	0.2097	
27129	Lyrics from the Song-Books of th...	42	0.2088	
901	The Jew of Malta	279	0.2082	
2607	Psmith, Journalist	242	0.2080	
27595	Eight Keys to Eden	59	0.2077	
19111	Code Three	46	0.2046	
47530	Oliver Twist, Vol. 2 (of 3)	40	0.2021	
16921	Plague Ship	218	0.2015	
1718	Manalive	120	0.2009	
20147	Rip Foster Rides the Gray Planet	40	0.1997	
9806	Mr. Justice Raffles	123	0.1983	
5347	Understood Betsy	159	0.1969	
1583	Options	63	0.1962	
24436	Anything You Can Do ...	58	0.1962	
126	The Poison Belt	268	0.1943	
4227	Tom Swift and His Wireless Messa...	41	0.1940	
4099	The Angel in the House	175	0.1914	
222	The Moon and Sixpence	352	0.1909	
9846	Excursions	110	0.1906	
18846	Voodoo Planet	148	0.1903	
37758	Atlantic Classics	49	0.1901	
5901	Dyke Darrel the Railroad Detecti...	44	0.1894	
1059	The World Set Free	343	0.1890	
897	The Rose and the Ring	49	0.1874	
2253	Henry V	266	0.1853	
27567	Aunt Jo's Scrap-Bag VI: An Old-Fa...	63	0.1831	
402	Penrod	71	0.1808	
17314	Five Children and It	190	0.1798	
2013	The Pit Prop Syndicate	40	0.1793	
2260	Titus Andronicus	52	0.1786	
55	The Wonderful Wizard of Oz	3,035	0.1770	
30431	Calumet 'K'	42	0.1762	
41715	Dave Dawson with the R.A.F.	49	0.1751	
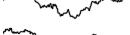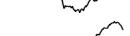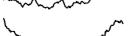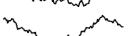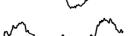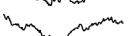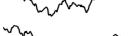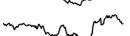
Table S6: Stories which are are represented by core emotional arc 3, sorted by the variance explained in their emotional arc by core emotional arc 3.

ID	Title	DL's	max(W[i,:])	Arc
1106	The Tragedy of Titus Andronicus	77	0.1750	
10966	The Ghost Pirates	234	0.1720	
20212	Police Your Planet	100	0.1697	
257	Troilus and Criseyde	236	0.1696	
13888	Bacon	54	0.1686	
14744	Different Girls	55	0.1683	
13897	The Adventure Club Afloat	117	0.1681	
14360	The Dawn and the Day : Or, The Bu...	46	0.1651	
778	Five Children and It	97	0.1647	
5182	The Old English Baron: a Gothic ...	89	0.1644	
3795	Under the Lilacs	72	0.1635	
16096	A Man's Woman	68	0.1631	
447	Maggie: A Girl of the Streets	1,295	0.1616	
1837	The Prince and the Pauper	1,389	0.1604	
13783	The Boy Inventors' Radio Telephone	52	0.1594	
14540	When William Came	42	0.1587	
21959	Letters from a Self-Made Merchan...	107	0.1582	
1654	An Unsocial Socialist	40	0.1581	
1845	Zuleika Dobson; Or, An Oxford Lo...	108	0.1555	
3006	Stalky & Co.	125	0.1552	
10066	Gunman's Reckoning	78	0.1551	
1013	The First Men in the Moon	348	0.1551	
24767	Jack O' Judgment	40	0.1544	
1809	Bucky O'Connor: A Tale of the Un...	49	0.1533	
18520	Sabotage in Space	48	0.1523	
2273	Tom Swift and His Motor-Boat; Or...	52	0.1513	
706	The Amateur Cracksman	182	0.1495	
20163	The Jolliest School of All	43	0.1489	
834	The Memoirs of Sherlock Holmes	2,164	0.1476	
34181	Irene Iddesleigh	138	0.1467	
22287	'Smiles': A Rose of the Cumberlands	75	0.1464	
9297	The Orange-Yellow Diamond	73	0.1457	
51	Anne of the Island	826	0.1451	
2014	The Lodger	97	0.1443	
1126	Measure for Measure	148	0.1436	
12436	The Night Horseman	87	0.1435	
1951	The Coming Race	350	0.1433	
2496	Our Village	41	0.1430	
1159	Fire-Tongue	50	0.1424	
113	The Secret Garden	1,153	0.1416	

Table S6: Stories which are are represented by core emotional arc 3, sorted by the variance explained in their emotional arc by core emotional arc 3.

ID	Title	DL's	max($W[i,:]$)	Arc
5162	Agatha Webb	41	0.1412	
8920	The Light of Asia	67	0.1411	
25102	Nobody's Boy: Sans Famille	60	0.1405	
17396	The Secret Garden	716	0.1394	
4368	Flappers and Philosophers	287	0.1386	
7498	Five Little Peppers Grown Up	88	0.1383	
2454	The Silent Bullet	43	0.1380	
847	Lays of Ancient Rome	259	0.1378	
6340	Literary Lapses	70	0.1362	
19651	Key Out of Time	196	0.1361	
24723	Final Weapon	67	0.1353	
3785	In the Reign of Terror: The Adve...	40	0.1341	
2175	You Never Can Tell	124	0.1323	
31619	The Planet Savers	145	0.1322	
9609	Joseph Andrews, Vol. 2	57	0.1307	
40504	Ginger-Snaps	41	0.1300	
14203	Varied Types	41	0.1289	
420	Dorothy and the Wizard in Oz	385	0.1286	
2524	My Lady Ludlow	58	0.1277	
421	Kidnapped	1,132	0.1274	
5066	The Whole Family: a Novel by Twe...	50	0.1266	
20989	'A Comedy of Errors' in Seven Acts	41	0.1261	
5230	The Invisible Man: A Grotesque R...	1,011	0.1250	
486	Ozma of Oz	268	0.1243	
7353	Birds in Town & Village	145	0.1226	
2776	The Four Million	255	0.1225	
364	The Mad King	101	0.1216	
21775	The Best of the World's Classics...	119	0.1209	
2722	Morning Star	58	0.1185	
1028	The Professor	223	0.1151	
10743	Moonfleet	258	0.1099	
85	The Beasts of Tarzan	227	0.1097	
30339	Status Quo	42	0.1094	
4272	The Christian Year	46	0.1072	
11195	Alcatraz	53	0.1062	
35425	The Mad Planet	74	0.1024	
213	The Man from Snowy River	73	0.1013	
22463	Chivalry	53	0.0975	
2515	Stepping Heavenward	54	0.0929	
7028	The Clicking of Cuthbert	138	0.0874	
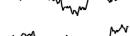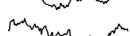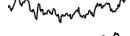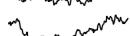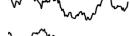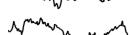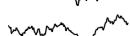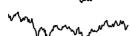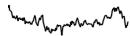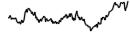
Table S6: Stories which are are represented by core emotional arc 3, sorted by the variance explained in their emotional arc by core emotional arc 3.

ID	Title	DL's	max(W[i,:])	Arc
12	Through the Looking-Glass	2,892	0.0742	
28700	Robin Hood	58	0.0717	
2244	As You Like It	65	0.0711	

Table S7: Top 10 stories which are are represented by core emotional arc 3, sorted by downloads.

ID	Title	DL's	max(W[i,:])	Arc
2814	Dubliners	4,742	0.2583	
55	The Wonderful Wizard of Oz	3,035	0.1770	
12	Through the Looking-Glass	2,892	0.0742	
834	The Memoirs of Sherlock Holmes	2,164	0.1476	
1837	The Prince and the Pauper	1,389	0.1604	
447	Maggie: A Girl of the Streets	1,295	0.1616	
113	The Secret Garden	1,153	0.1416	
421	Kidnapped	1,132	0.1274	
5230	The Invisible Man: A Grotesque R...	1,011	0.1250	

Table S8: Stories which are are represented by core emotional arc 4, sorted by the variance explained in their emotional arc by core emotional arc 4.

ID	Title	DL's	max(W[i,:])	Arc
34215	Shadowings	63	0.4141	
12384	Battle-Pieces and Aspects of the...	194	0.3546	
36281	The Slayer of Souls	63	0.3283	
17412	The Bobbsey Twins : Or, Merry Day...	69	0.3259	
2727	Allan's Wife	128	0.3124	
27761	Hamlet, Prince of Denmark	301	0.3057	
363	The Oakdale Affair	91	0.2968	
17854	The Sport of the Gods	61	0.2855	
27726	Tolstoy on Shakespeare: A Critic...	212	0.2764	
1292	The Way of the World	367	0.2722	
4381	The Aran Islands	63	0.2644	
19355	A Book of Prefaces	55	0.2621	
19	The Song of Hiawatha	297	0.2546	
1526	Twelfth Night; Or, What You Will	292	0.2531	

Table S8: Stories which are are represented by core emotional arc 4, sorted by the variance explained in their emotional arc by core emotional arc 4.

ID	Title	DL's	max(W[i,:])	Arc
1527	Twelfth Night; Or, What You Will	90	0.2475	
11696	The Food of the Gods and How It ...	206	0.2469	
5829	The Moneychangers	46	0.2456	
20912	The Daffodil Mystery	148	0.2415	
551	The Land That Time Forgot	278	0.2412	
11045	The Ghost Ship	79	0.2411	
2040	Confessions of an English Opium-...	643	0.2382	
20431	The Tale of Beowulf, Sometime Ki...	80	0.2371	
3329	Caesar and Cleopatra	105	0.2360	
11127	The Case of Jennie Brice	43	0.2359	
24459	The Lost Princess of Oz	66	0.2345	
11451	The Rome Express	162	0.2314	
30368	A Christmas Carol: The original m...	143	0.2276	
13716	A Trip to Venus: A Novel	44	0.2267	
28164	The Big Bow Mystery	61	0.2248	
959	The Lost Princess of Oz	178	0.2242	
170	The Haunted Hotel: A Mystery of ...	162	0.2222	
6678	Nonsenseorship	99	0.2219	
2240	Much Ado about Nothing	289	0.2180	
39378	Mortal Coils	114	0.2180	
19337	A Christmas Carol	622	0.2171	
5210	The Borough	73	0.2169	
1163	Adventure	62	0.2163	
20856	Ten From Infinity	47	0.2162	
22332	Brain Twister	89	0.2149	
40603	The Root of All Evil	44	0.2147	
2317	The Story of My Heart: An Autobi...	64	0.2139	
32664	Black Amazon of Mars	142	0.2124	
775	When the Sleeper Wakes	151	0.2123	
22549	Space Prison	124	0.2122	
2713	Maiwa's Revenge; Or, The War of ...	56	0.2099	
18639	The Victorian Age in Literature	97	0.2072	
2062	All for Love; Or, The World Well...	239	0.2053	
24	O Pioneers!	371	0.2048	
4011	Epicoene; Or, The Silent Woman	123	0.2038	
1103	King Richard III	384	0.2014	
1264	The Wheels of Chance: A Bicyclin...	47	0.1989	
35533	The Haunted Room: A Tale	53	0.1989	
5131	Childe Harold's Pilgrimage	1,481	0.1987	
6995	Ghosts I Have Met and Some Others	65	0.1985	

Table S8: Stories which are are represented by core emotional arc 4, sorted by the variance explained in their emotional arc by core emotional arc 4.

ID	Title	DL's	max(W[i,:])	Arc
19860	The Arabian Nights Entertainments	470	0.1981	
5070	The Doctor's Dilemma	113	0.1978	
29965	Two Thousand Miles Below	47	0.1975	
1115	The First Part of King Henry the...	70	0.1964	
2257	Richard III	49	0.1963	
33642	Earth Alert!	97	0.1958	
28434	The Astronomy of Milton's 'Parad...	44	0.1931	
134	Maria; Or, The Wrongs of Woman	271	0.1929	
8914	England, My England	67	0.1927	
46	A Christmas Carol in Prose; Bein...	4,602	0.1917	
13650	Nonsense Books	282	0.1913	
1537	Pericles, Prince of Tyre	67	0.1909	
22767	Pagan Passions	94	0.1899	
5795	The Secret Rose	56	0.1893	
888	The Lazy Tour of Two Idle Appren...	47	0.1888	
131	The Pilgrim's Progress from this...	1,126	0.1888	
3776	The Valley of Fear	182	0.1884	
37173	In a Glass Darkly, v. 2/3	47	0.1879	
24022	A Christmas Carol	188	0.1873	
22057	Kid Wolf of Texas : A Western Story	53	0.1869	
34426	The Enchanted Barn	81	0.1864	
1240	The Playboy of the Western World...	317	0.1864	
28522	Laura Middleton; Her Brother and...	1,097	0.1851	
3289	The Valley of Fear	1,228	0.1845	
18934	My Lady Nicotine: A Study in Smoke	60	0.1845	
13815	The Talking Beasts: A Book of Fa...	65	0.1797	
1585	The Wrong Box	60	0.1797	
22527	Beyond the Vanishing Point	45	0.1796	
32415	The Nursery Rhymes of England	121	0.1789	
68	Warlord of Mars	571	0.1779	
4737	A Tale of a Tub	277	0.1775	
8086	Down and Out in the Magic Kingdom	136	0.1769	
30537	The Royal Book of Oz : In which t...	64	0.1745	
21633	The Man of the Desert	54	0.1719	
2081	The Blithedale Romance	229	0.1719	
20606	The Magic City	118	0.1700	
40493	The King of Diamonds: A Tale of ...	48	0.1699	
26649	Terribly Intimate Portraits	46	0.1680	
7884	In the Fog	64	0.1680	
1097	Mrs. Warren's Profession	780	0.1679	

Table S8: Stories which are are represented by core emotional arc 4, sorted by the variance explained in their emotional arc by core emotional arc 4.

ID	Title	DL's	max(W[i,:])	Arc
3479	The Metal Monster	77	0.1660	
33028	Man and Maid	40	0.1659	
22338	The Impossibles	56	0.1655	
40284	The Sex Life of the Gods	197	0.1651	
14667	A Christmas Garland	56	0.1645	
3179	The American Claimant	45	0.1639	
6120	Soldiers Three	45	0.1636	
1720	The Man Who Knew Too Much	310	0.1628	
30905	The Boarded-Up House	62	0.1627	
19258	Tom Swift and the Electronic Hyd...	70	0.1624	
30333	Daddy's Girl	52	0.1623	
4552	The Border Legion	96	0.1614	
22754	Masters of Space	173	0.1604	
17112	Many Thoughts of Many Minds: A Tr...	47	0.1579	
32498	The Brain	87	0.1548	
32398	Brood of the Dark Moon : (A Seque...	46	0.1542	
34403	The Clock Strikes Thirteen	43	0.1527	
25051	Space Platform	72	0.1523	
1711	Child of Storm	109	0.1511	
54	The Marvelous Land of Oz	419	0.1482	
20	Paradise Lost	2,522	0.1475	
7365	If I May	63	0.1474	
854	A Woman of No Importance	729	0.1455	
21048	Just Patty	47	0.1451	
6927	The White Feather	149	0.1427	
92	Tarzan and the Jewels of Opar	168	0.1404	
369	The Outlaw of Torn	129	0.1380	
2251	Henry IV, Part 1	76	0.1375	
2183	Three Men on the Bummel	190	0.1362	
32117	Eleven Possible Cases	57	0.1359	
9871	The Avenger	45	0.1355	
238	Dear Enemy	129	0.1354	
20698	The Story of Glass	53	0.1350	
3464	Tish: The Chronicle of Her Escap...	63	0.1349	
35545	Sanders of the River	127	0.1338	
32351	Voyage To Eternity	42	0.1333	
21051	Skylark Three	157	0.1331	
10581	Uncle Bernac: A Memory of the Em...	41	0.1322	
5746	The Ancient Allan	101	0.1319	
26197	The Nursery Rhyme Book	121	0.1301	

Table S8: Stories which are are represented by core emotional arc 4, sorted by the variance explained in their emotional arc by core emotional arc 4.

ID	Title	DL's	max(W[i,:])	Arc
5660	Mary Louise	59	0.1282	
2306	Uncle Remus, His Songs and His S...	177	0.1275	
26654	Peter and Wendy	1,068	0.1275	
37193	The Swedish Fairy Book	57	0.1273	
1109	Love's Labour's Lost	56	0.1269	
3618	Arms and the Man	536	0.1252	
9834	The Talleyrand Maxim	74	0.1242	
26999	Peter Pan in Kensington Gardens	60	0.1238	
470	Heretics	395	0.1236	
10586	Mike and Psmith	170	0.1231	
37189	The Return of the Soldier	108	0.1223	
19079	The Adventures of Lightfoot the ...	40	0.1172	
12793	Cobwebs from an Empty Skull	56	0.1136	
12753	The Legends of King Arthur and H...	640	0.1125	
38562	The Big Book of Nursery Rhymes	82	0.1113	
1644	The Adventures of Gerard	100	0.1103	
2020	Tarzan the Terrible	185	0.1096	
498	Rebecca of Sunnybrook Farm	155	0.1083	
9190	The Greater Inclination	41	0.1068	
12239	Dead Men's Money	160	0.1063	
376	A Journal of the Plague Year : Wr...	461	0.1057	
25439	Looking Backward: 2000-1887	111	0.1044	
62	A Princess of Mars	2,515	0.1032	
2540	Father and Son: A Study of Two T...	74	0.1021	
166	Summer	165	0.1011	
18458	Star Born	160	0.0989	
38703	The Black Moth: A Romance of the...	204	0.0970	
72	Thuvia, Maid of Mars	437	0.0939	
972	The Devil's Dictionary	1,257	0.0936	
804	A Sentimental Journey Through Fr...	261	0.0921	
157	Daddy-Long-Legs	531	0.0895	
4993	A Texas Ranger	61	0.0868	
17985	Tom Swift and The Visitor from P...	98	0.0832	
6093	Far Away and Long Ago: A History...	70	0.0818	
308	Three Men in a Boat	2,059	0.0815	
22064	Tess of the Storm Country	83	0.0814	
26240	The Clansman: An Historical Roma...	85	0.0728	

Table S9: Top 10 stories which are are represented by core emotional arc 4, sorted by downloads.

ID	Title	DL's	max($W[i,:]$)	Arc
46	A Christmas Carol in Prose; Bein...	4,602	0.1917	
20	Paradise Lost	2,522	0.1475	
62	A Princess of Mars	2,515	0.1032	
308	Three Men in a Boat	2,059	0.0815	
5131	Childe Harold's Pilgrimage	1,481	0.1987	
972	The Devil's Dictionary	1,257	0.0936	
3289	The Valley of Fear	1,228	0.1845	
131	The Pilgrim's Progress from this...	1,126	0.1888	
28522	Laura Middleton; Her Brother and...	1,097	0.1851	
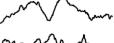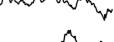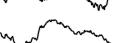
Table S10: Stories which are are represented by core emotional arc 5, sorted by the variance explained in their emotional arc by core emotional arc 5.

ID	Title	DL's	max($W[i,:]$)	Arc
17763	The Mystery of the Hasty Arrow	93	0.2301	
5317	Through the Magic Door	81	0.2299	
13944	After London; Or, Wild England	146	0.2259	
12590	The Shadow of the Rope	75	0.2202	
35247	That Affair at Elizabeth	62	0.2179	
39827	The Conduct of Life	176	0.2041	
295	The Early Short Fiction of Edith...	49	0.2012	
10459	The Celtic Twilight	146	0.1974	
957	The Scarecrow of Oz	162	0.1972	
19819	Milton's Comus	41	0.1969	
644	The Haunted Man and the Ghost's ...	97	0.1915	
23624	Ride Proud, Rebel!	139	0.1882	
27771	Once on a Time	102	0.1858	
13694	Mince Pie	46	0.1811	
40386	Wandering Ghosts	125	0.1809	
32440	Dave Dawson at Dunkirk	60	0.1786	
1146	The Journal of a Voyage to Lisbon	57	0.1776	
25780	The Fire People	45	0.1768	
306	The Early Short Fiction of Edith...	44	0.1738	
27780	Treasure Island	281	0.1712	
8994	What Katy Did	142	0.1700	
20788	Storm Over Warlock	154	0.1687	
24680	The Martyr of the Catacombs: A Ta...	43	0.1664	
30759	Exit Betty	49	0.1664	
1212	Love and Freindship [sic]	611	0.1657	
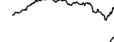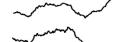
Table S10: Stories which are are represented by core emotional arc 5, sorted by the variance explained in their emotional arc by core emotional arc 5.

ID	Title	DL's	$\max(W[\mathrm{i},:])$	Arc
2166	King Solomon's Mines	788	0.1652	
120	Treasure Island	4,402	0.1620	
20782	Triplanetary	131	0.1599	
11068	The Spirit of the Age; Or, Conte...	65	0.1570	
330	Where There's a Will	64	0.1566	
618	Codex Junius 11	47	0.1553	
39592	Princess Mary's Gift Book : All p...	71	0.1541	
5805	The League of the Scarlet Pimpernel	104	0.1505	
16255	Dickey Downy: The Autobiography ...	44	0.1496	
14917	The Wings of the Morning	41	0.1478	
19369	The Triumphs of Eugène Valmont	102	0.1477	
696	The Castle of Otranto	1,663	0.1465	
15585	Humorous Masterpieces from Ameri...	53	0.1462	
1164	The Iron Heel	506	0.1395	
26862	Howard Pyle's Book of Pirates : F...	81	0.1391	
423	Round the Red Lamp: Being Facts ...	136	0.1382	
1515	The Merchant of Venice	50	0.1353	
15323	The Green Eyes of Bâst	65	0.1329	
553	Out of Time's Abyss	140	0.1325	
18217	Chambers's Elementary Science Re...	62	0.1320	
1280	Spoon River Anthology	671	0.1297	
10723	Betty's Bright Idea; Deacon Pitk...	40	0.1281	
1550	A Lady of Quality : Being a Most ...	89	0.1261	
34339	The Princess and the Goblin	268	0.1254	
1114	The Merchant of Venice	134	0.1233	
2644	Isaac Bickerstaff, Physician and...	42	0.1191	
5830	A Garland for Girls	133	0.1185	
10601	The Rangeland Avenger	94	0.1177	
3005	Tom Swift and His Airship	70	0.1159	
18151	Time Crime	85	0.1158	
39281	Dictionary of English Proverbs a...	83	0.1149	
26019	Europa's Fairy Book	56	0.1144	
1182	Dope	94	0.1138	
37012	The Recruiting Officer	49	0.1114	
2809	Main-Travelled Roads	169	0.1105	
2024	Diary of a Pilgrimage	48	0.1071	
31356	The Man Who Staked the Stars	61	0.1054	
2005	Piccadilly Jim	151	0.1039	
394	Cranford	285	0.1029	
36869	The Real Man	59	0.1027	
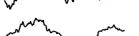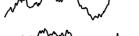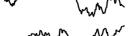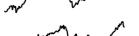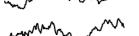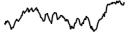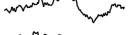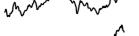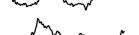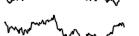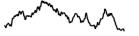
Table S10: Stories which are are represented by core emotional arc 5, sorted by the variance explained in their emotional arc by core emotional arc 5.

ID	Title	DL's	max(W[i,:])	Arc
708	The Princess and the Goblin	579	0.0996	
39116	Unicorns	47	0.0969	
5803	Not that it Matters	138	0.0956	
22693	A Book of Myths	248	0.0875	
2305	A Set of Six	58	0.0862	
19717	The Bostonians, Vol. I (of II)	80	0.0817	
3688	The Chronicles of Clovis	186	0.0812	
2243	The Merchant of Venice	506	0.0787	

Table S11: Top 10 stories which are are represented by core emotional arc 5, sorted by downloads.

ID	Title	DL's	max(W[i,:])	Arc
120	Treasure Island	4,402	0.1620	
696	The Castle of Otranto	1,663	0.1465	
2166	King Solomon's Mines	788	0.1652	
1280	Spoon River Anthology	671	0.1297	
1212	Love and Freindship [sic]	611	0.1657	
708	The Princess and the Goblin	579	0.0996	
1164	The Iron Heel	506	0.1395	
2243	The Merchant of Venice	506	0.0787	
394	Cranford	285	0.1029	

Table S12: Stories which are are represented by core emotional arc 6, sorted by the variance explained in their emotional arc by core emotional arc 6.

ID	Title	DL's	max(W[i,:])	Arc
18172	This World Is Taboo	64	0.3281	
339	Old Indian Days	139	0.3148	
10377	The Evil Guest	93	0.2868	
29448	Pariah Planet	96	0.2864	
289	The Wind in the Willows	1,475	0.2750	
20727	The Cosmic Computer	221	0.2716	
27805	The Wind in the Willows	543	0.2702	
27991	The Blue Bird for Children: The W...	111	0.2677	
96	The Monster Men	155	0.2346	

Table S12: Stories which are are represented by core emotional arc 6, sorted by the variance explained in their emotional arc by core emotional arc 6.

ID	Title	DL's	max(W[i,:])	Arc
9156	Life and Remains of John Clare, ...	53	0.2299	
41027	The Revolt of the Star Men	51	0.2276	
34313	Literature in the Making, by Som...	54	0.2253	
209	The Turn of the Screw	2,175	0.2251	
32256	The Big Time	135	0.2229	
22354	The Adventures of Maya the Bee	68	0.2212	
654	Grace Abounding to the Chief of ...	177	0.2209	
20859	Wandl the Invader	49	0.2106	
310	Before Adam	94	0.2101	
958	Rinkitink in Oz : Wherein Is Reco...	124	0.2092	
10850	Philaster; Or, Love Lies a Bleeding	50	0.2091	
1725	Heart of the West	118	0.2090	
15798	Clover	102	0.2069	
25581	Rinkitink in Oz	41	0.2054	
1906	Erewhon; Or, Over the Range	251	0.2051	
20781	Heidi: (Gift Edition)	642	0.2033	
17731	The Nigger Of The "Narcissus": A...	207	0.2027	
9990	Brave and Bold; Or, The Fortunes...	56	0.2020	
22182	The Best of the World's Classics...	65	0.1972	
1915	The Second Thoughts of an Idle F...	88	0.1956	
244	A Study in Scarlet	4,535	0.1955	
9611	Joseph Andrews, Vol. 1	242	0.1939	
863	The Mysterious Affair at Styles	3,112	0.1937	
10324	Bull Hunter	58	0.1918	
1329	A Voyage to Arcturus	218	0.1897	
21768	A Desert Drama: Being The Traged...	42	0.1841	
34020	The Window at the White Cat	46	0.1837	
43984	Chaucer for Children: A Golden Key	119	0.1826	
1143	Notes on Life & Letters	53	0.1821	
17870	Operation Terror	59	0.1809	
12163	The Sleeper Awakes: A Revised Edi...	141	0.1759	
223	The Wisdom of Father Brown	563	0.1757	
20898	The Galaxy Primes	227	0.1727	
6683	The Little Nugget	95	0.1723	
1448	Heidi	268	0.1710	
30970	Miss Cayley's Adventures	45	0.1707	
950	Tom Swift and His Electric Runab...	44	0.1706	
1446	Perfect Behavior: A Guide for La...	59	0.1704	
90	The Son of Tarzan	212	0.1700	
23028	Greylorn	61	0.1691	
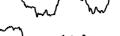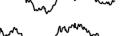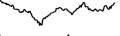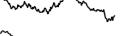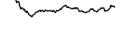
Table S12: Stories which are are represented by core emotional arc 6, sorted by the variance explained in their emotional arc by core emotional arc 6.

ID	Title	DL's	max(W[i,:])	Arc
1129	The Tragedy of Macbeth	449	0.1690	
25003	The Nicest Girl in the School: A...	48	0.1689	
1948	The Story of a Bad Boy	45	0.1683	
15238	Mathilda	163	0.1673	
32037	Eureka: A Prose Poem	174	0.1669	
22892	The Best Made Plans	40	0.1664	
1450	Pollyanna	349	0.1646	
2154	Around the World in Eighty Days....	57	0.1642	
1795	Macbeth	73	0.1637	
2126	The Quest of the Sacred Slipper	55	0.1626	
19023	A Daughter of the Sioux: A Tale ...	61	0.1621	
84	Frankenstein; Or, The Modern Pro...	11,699	0.1602	
1874	The Railway Children	437	0.1595	
2512	The Cruise of the Snark	105	0.1583	
3790	Major Barbara	416	0.1581	
23845	Talents, Incorporated	66	0.1567	
30964	The Ethical Engineer	110	0.1566	
42324	Frankenstein; Or, The Modern Pro...	313	0.1565	
32934	The Young Colonists: A Story of ...	82	0.1565	
33505	The Trembling of the Veil	60	0.1561	
21639	When Patty Went to College	41	0.1558	
5632	Five Little Peppers Midway	58	0.1554	
19471	Badge of Infamy	105	0.1541	
32954	The Black Arrow: A Tale of the T...	115	0.1530	
2763	The World's Desire	45	0.1526	
2381	Actions and Reactions	92	0.1522	
463	The Red Badge of Courage: An Epi...	84	0.1514	
848	The Black Arrow: A Tale of Two R...	292	0.1496	
73	The Red Badge of Courage: An Epi...	1,163	0.1494	
32	Herland	1,013	0.1485	
329	Island Nights' Entertainments	67	0.1483	
1625	The Frozen Deep	51	0.1467	
2906	The Silver Box: A Comedy in Thre...	49	0.1463	
16865	Pinocchio: The Tale of a Puppet	234	0.1451	
41445	Frankenstein; Or, The Modern Pro...	786	0.1448	
39896	The Girl Next Door	141	0.1433	
20728	Space Viking	223	0.1375	
383	She Stoops to Conquer; Or, The M...	903	0.1359	
2509	The Lani People	66	0.1358	
15717	Books and Persons; Being Comment...	41	0.1356	
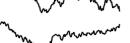
Table S12: Stories which are are represented by core emotional arc 6, sorted by the variance explained in their emotional arc by core emotional arc 6.

ID	Title	DL's	max(W[i,:])	Arc
12215	Odd Craft, Complete	41	0.1355	
2851	Sixes and Sevens	85	0.1336	
10007	Carmilla	1,416	0.1333	
33348	Reveries over Childhood and Youth	44	0.1320	
620	Sylvie and Bruno	104	0.1318	
9746	The Ashiel mystery: A Detective ...	92	0.1303	
552	The People That Time Forgot	171	0.1302	
9656	Alarms and Discursions	64	0.1285	
3146	Two on a Tower	101	0.1284	
32331	Dave Dawson at Casablanca	47	0.1280	
24197	The Tinted Venus: A Farcical Rom...	40	0.1271	
4223	The Mystery of a Hansom Cab	43	0.1268	
12629	Language: An Introduction to the ...	210	0.1259	
2186	"Captains Courageous": A Story o...	72	0.1259	
33325	The Spoils of Poynton	41	0.1256	
19207	The Firelight Fairy Book	93	0.1211	
19672	The Holladay Case: A Tale	47	0.1204	
2097	The Sign of the Four	2,283	0.1199	
2487	Cross Roads	57	0.1198	
41049	The Onslaught from Rigel	60	0.1178	
18505	A Popular Schoolgirl	87	0.1168	
5606	Guns of the Gods: A Story of Yas...	50	0.1145	
26176	The Secret House	90	0.1139	
24880	The Wreck of the Titan: or, Futility	148	0.1138	
389	The Great God Pan	807	0.1127	
2276	The Private Memoirs and Confessi...	260	0.1121	
8457	Frenzied Fiction	45	0.1115	
4682	Nonsense Novels	103	0.1108	
4709	Brewster's Millions	131	0.1083	
20551	The White Invaders	53	0.1080	
36	The War of the Worlds	2,496	0.1060	
3026	North of Boston	151	0.1045	
10554	Right Ho, Jeeves	896	0.1004	
2662	Under the Greenwood Tree; Or, Th...	135	0.0993	
1881	The Call of the Canyon	192	0.0992	
32501	The Golden Age	44	0.0988	
1897	The Seventh Man	56	0.0976	
1183	The Return of Dr. Fu-Manchu	91	0.0976	
19142	The Devil Doctor	99	0.0971	
21626	Adrift in the Wilds; Or, The Adv...	43	0.0965	
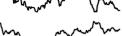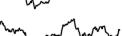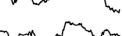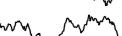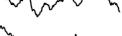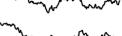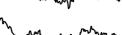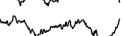
Table S12: Stories which are are represented by core emotional arc 6, sorted by the variance explained in their emotional arc by core emotional arc 6.

ID	Title	DL's	max(W[i,:])	Arc
20081	A Houseful of Girls	45	0.0947	
1354	Chronicles of Avonlea	258	0.0944	
2786	Jack and Jill	154	0.0940	
9867	Riders of the Silences	62	0.0905	
1267	Kai Lung's Golden Hours	94	0.0903	
1478	A Parody Outline of History : Whe...	60	0.0874	
19535	George Bernard Shaw	60	0.0870	
19246	The Young Pitcher	49	0.0867	
1589	Tamburlaine the Great — Part 2	104	0.0851	
34414	Just William	202	0.0716	

Table S13: Top 10 stories which are are represented by core emotional arc 6, sorted by downloads.

ID	Title	DL's	max(W[i,:])	Arc
84	Frankenstein; Or, The Modern Pro...	11,699	0.1602	
244	A Study in Scarlet	4,535	0.1955	
863	The Mysterious Affair at Styles	3,112	0.1937	
36	The War of the Worlds	2,496	0.1060	
2097	The Sign of the Four	2,283	0.1199	
209	The Turn of the Screw	2,175	0.2251	
289	The Wind in the Willows	1,475	0.2750	
10007	Carmilla	1,416	0.1333	
73	The Red Badge of Courage: An Epi...	1,163	0.1494	
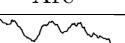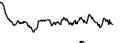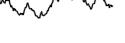
Table S14: Stories which are are represented by core emotional arc 7, sorted by the variance explained in their emotional arc by core emotional arc 7.

ID	Title	DL's	max(W[i,:])	Arc
12028	The Uttermost Farthing: A Savant'...	58	0.2612	
19526	Stand by for Mars!	64	0.2425	
811	The Tragical History of Doctor F...	389	0.2245	
121	Northanger Abbey	2,355	0.2196	
7031	The Sheik: A Novel	152	0.2071	
24035	The Pirates of Ersatz	113	0.2017	
2429	Lost Face	190	0.1975	
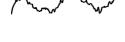
Table S14: Stories which are are represented by core emotional arc 7, sorted by the variance explained in their emotional arc by core emotional arc 7.

ID	Title	DL's	max(W[i,:])	Arc
74	The Adventures of Tom Sawyer	9,454	0.1906	
1058	The Mirror of the Sea	106	0.1885	
3674	The Dragon and the Raven; Or, Th...	151	0.1791	
26853	Vice Versa; or, A Lesson to Fathers	49	0.1787	
19709	Danger in Deep Space	74	0.1751	
20472	Grace Harlowe's Plebe Year at Hi...	40	0.1729	
18719	Space Tug	63	0.1721	
24929	The Green Rust	56	0.1709	
27826	The Olive Fairy Book	113	0.1670	
5962	Oh, Money! Money! A Novel	44	0.1660	
10886	The Untamed	109	0.1658	
18019	The Luckiest Girl in the School	107	0.1649	
6622	Legends That Every Child Should ...	144	0.1612	
2686	The Book of Snobs	204	0.1589	
19330	An Apache Princess: A Tale of th...	69	0.1576	
21891	The Brand of Silence: A Detective...	89	0.1568	
15580	The Rustlers of Pecos County	90	0.1537	
24283	Down the River; Or, Buck Bradfor...	42	0.1529	
6382	Bat Wing	92	0.1511	
10542	The Boats of the "Glen Carrig" : ...	126	0.1482	
20154	Invaders from the Infinite	103	0.1429	
21964	The Short-story	182	0.1414	
8771	Jurgen: A Comedy of Justice	95	0.1407	
11583	The Runaway Asteroid	43	0.1384	
8492	The King in Yellow	1,504	0.1354	
13029	The Art of the Moving Picture	95	0.1317	
13675	Goody Two-Shoes : A Facsimile Rep...	90	0.1307	
10234	Old Creole Days: A Story of Creo...	80	0.1277	
1281	Tom Swift and His Aerial Warship...	45	0.1261	
1695	The Man Who Was Thursday: A Nigh...	796	0.1238	
2225	"Captains Courageous": A Story o...	212	0.1232	
910	White Fang	1,178	0.1223	
4517	Ethan Frome	2,895	0.1220	
16	Peter Pan	5,789	0.1197	
146	A Little Princess : Being the who...	825	0.1189	
6100	Pollyanna Grows Up	107	0.1140	
24160	The Basket of Flowers	41	0.1095	
1154	The Voyages of Doctor Dolittle	179	0.1081	
14632	The Mystery of Mary	52	0.1075	
21510	Legacy	101	0.1069	
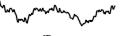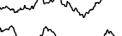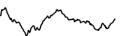
Table S14: Stories which are are represented by core emotional arc 7, sorted by the variance explained in their emotional arc by core emotional arc 7.

ID	Title	DL's	max(W[i,:])	Arc
37332	A Little Princess: Being the who...	142	0.1066	
955	The Patchwork Girl of Oz	172	0.1030	
242	My Antonia	847	0.1024	
1376	The Little White Bird; Or, Adven...	250	0.1014	
47	Anne of Avonlea	803	0.1000	
12491	Twelve Types	71	0.0997	
19810	My Ántonia	171	0.0962	
1204	Cabin Fever	64	0.0947	
23292	Ted and the Telephone	95	0.0942	
15625	The Lookout Man	51	0.0935	
25388	The Herapath Property	41	0.0929	
41231	The Life and Beauties of Fanny Fern	49	0.0907	
887	Intentions	173	0.0890	
23661	The Book of Dragons	254	0.0883	
81	The Return of Tarzan	384	0.0864	
14875	Elsie's children	41	0.0861	
2861	The Sleuth of St. James's Square	76	0.0848	
21715	Away in the Wilderness	53	0.0789	
5604	Getting Married	105	0.0727	

Table S15: Top 10 stories which are are represented by core emotional arc 7, sorted by downloads.

ID	Title	DL's	max(W[i,:])	Arc
74	The Adventures of Tom Sawyer	9,454	0.1906	
16	Peter Pan	5,789	0.1197	
4517	Ethan Frome	2,895	0.1220	
121	Northanger Abbey	2,355	0.2196	
8492	The King in Yellow	1,504	0.1354	
910	White Fang	1,178	0.1223	
242	My Antonia	847	0.1024	
146	A Little Princess : Being the who...	825	0.1189	
47	Anne of Avonlea	803	0.1000	

Table S16: Stories which are are represented by core emotional arc 8, sorted by the variance explained in their emotional arc by core emotional arc 8.

ID	Title	DL's	$\max(W[i,:])$	Arc
19360	Six to Sixteen: A Story for Girls	47	0.2415	
28849	Smugglers' Reef: A Rick Brant Sc...	46	0.2320	
18581	Adrift in New York: Tom and Flor...	111	0.2182	
5141	What Katy Did at School	85	0.2065	
4217	A Portrait of the Artist as a Yo...	2,172	0.2039	
31308	Orientations	76	0.1924	
35117	Lord Tony's Wife: An Adventure o...	66	0.1748	
4715	An African Millionaire: Episodes...	163	0.1714	
4731	Seven Little Australians	79	0.1679	
16721	A Place so Foreign	54	0.1654	
32242	A Wonder Book for Girls & Boys	192	0.1588	
496	The Little Lame Prince	48	0.1577	
14154	The Tale of Terror: A Study of t...	63	0.1570	
26	Paradise Lost	730	0.1565	
20739	Rebels of the Red Planet	47	0.1511	
4253	Dramatic Romances	71	0.1507	
23641	The Forsaken Inn: A Novel	59	0.1496	
64	The Gods of Mars	628	0.1481	
37174	In a Glass Darkly, v. 3/3	46	0.1473	
12187	The Mystery of 31 New Inn	188	0.1464	
29405	The Gods of Mars	158	0.1441	
2688	The Clue of the Twisted Candle	237	0.1402	
7434	The Adventures of Joel Pepper	47	0.1376	
204	The Innocence of Father Brown	800	0.1357	
980	Alice Adams	48	0.1338	
11626	The Dawn of All	57	0.1319	
2028	The Yellow Claw	73	0.1286	
849	The Idle Thoughts of an Idle Fellow	177	0.1281	
9903	Way of the Lawless	68	0.1275	
4352	Laughter: An Essay on the Meanin...	365	0.1251	
2568	Trent's Last Case	193	0.1248	
472	The House Behind the Cedars	107	0.1246	
36775	Humorous Readings and Recitation...	136	0.1238	
21073	A Pirate of the Caribbees	49	0.1230	
35304	The Last Stroke: A Detective Story	123	0.1214	
8446	The Enormous Room	232	0.1195	
36958	A Child of the Jago	61	0.1193	
41667	The Emerald City of Oz	67	0.1163	
24201	The Eye of Osiris	91	0.1103	
517	The Emerald City of Oz	266	0.1097	
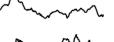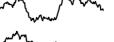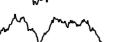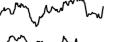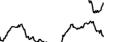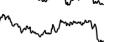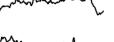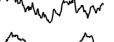
Table S16: Stories which are are represented by core emotional arc 8, sorted by the variance explained in their emotional arc by core emotional arc 8.

ID	Title	DL's	max(W[i,:])	Arc
21854	The Woman in Black	93	0.1085	
5670	Jacob's Room	403	0.1013	
10476	The Vanishing Man : A Detective R...	89	0.0992	
11666	The Conjure Woman	203	0.0898	
225	At the Back of the North Wind	288	0.0889	
40241	Hieroglyphics	50	0.0871	
32587	The Ambassador	71	0.0865	
6987	Five Little Peppers Abroad	53	0.0829	
6880	The Coming of Bill	95	0.0799	
32884	Ideas of Good and Evil	109	0.0737	

Table S17: Top 10 stories which are are represented by core emotional arc 8, sorted by downloads.

ID	Title	DL's	max(W[i,:])	Arc
4217	A Portrait of the Artist as a Yo...	2,172	0.2039	
204	The Innocence of Father Brown	800	0.1357	
26	Paradise Lost	730	0.1565	
64	The Gods of Mars	628	0.1481	
5670	Jacob's Room	403	0.1013	
4352	Laughter: An Essay on the Meanin...	365	0.1251	
225	At the Back of the North Wind	288	0.0889	
517	The Emerald City of Oz	266	0.1097	
2688	The Clue of the Twisted Candle	237	0.1402	

Table S18: Stories which are are represented by core emotional arc 9, sorted by the variance explained in their emotional arc by core emotional arc 9.

ID	Title	DL's	max(W[i,:])	Arc
19066	Brigands of the Moon	65	0.2113	
17959	The Hand Of Fu-Manchu: Being a Ne...	112	0.1861	
19145	The Time Traders	225	0.1801	
2057	The Last of the Plainsmen	106	0.1681	
1188	The Lair of the White Worm	206	0.1632	
546	Under the Andes	87	0.1544	
18668	In Search of the Unknown	159	0.1489	

Table S18: Stories which are are represented by core emotional arc 9, sorted by the variance explained in their emotional arc by core emotional arc 9.

ID	Title	DL's	$\max(W[\text{i},:])$	Arc
6955	The Prince and Betty	85	0.1488	
19706	Brood of the Witch-Queen	100	0.1484	
8899	Three Weeks	48	0.1457	
2098	A Thief in the Night: A Book of ...	112	0.1426	
1872	The Red House Mystery	422	0.1421	
41753	Dave Dawson at Truk	42	0.1405	
9415	Olaf the Glorious: A Story of th...	42	0.1365	
1999	Crome Yellow	210	0.1319	
24313	Once a Week	49	0.1311	
9807	Scarhaven Keep	68	0.1309	
34219	The Enchanted Castle	61	0.1296	
8730	A Little Bush Maid	49	0.1273	
25344	The Scarlet Letter	386	0.1266	
17667	Dialogues of the Dead	51	0.1232	
3536	The Enchanted Castle	220	0.1227	
33	The Scarlet Letter	3,045	0.1195	
25564	The Water-Babies: A Fairy Tale f...	103	0.1193	
3328	Man and Superman: A Comedy and a...	312	0.1181	
12986	The Card, a Story of Adventure i...	54	0.1169	
2775	The Good Soldier	426	0.1144	
21656	The Princess of the School	56	0.1096	
1533	Macbeth	165	0.1088	
1051	Sartor Resartus: The Life and Op...	347	0.1074	
30324	The Pathless Trail	65	0.1069	
19307	The Lion of Petra	41	0.1044	
11	Alice's Adventures in Wonderland	17,366	0.0988	
1751	Twilight Land	48	0.0985	
5121	Dark Hollow	43	0.0978	
22234	Aunt Jo's Scrap-Bag, Vol. 5: Jimm...	69	0.0975	
751	The Autocrat of the Breakfast-Table	55	0.0960	
794	The Wouldbegoods: Being the Furt...	50	0.0956	
2777	Cabbages and Kings	209	0.0919	
1077	The Mirror of Kong Ho	45	0.0884	
271	Black Beauty	780	0.0870	
28885	Alice's Adventures in Wonderland...	1,051	0.0862	
5308	The Paradise Mystery	94	0.0862	
1262	The Heritage of the Desert: A Novel	75	0.0841	
1590	The Amazing Interlude	43	0.0823	
9196	The Clockmaker; Or, the Sayings ...	51	0.0708	

Table S19: Top 10 stories which are are represented by core emotional arc 9, sorted by downloads.

ID	Title	DL's	max(W[i,:])	Arc
11	Alice's Adventures in Wonderland	17,366	0.0988	
33	The Scarlet Letter	3,045	0.1195	
28885	Alice's Adventures in Wonderland...	1,051	0.0862	
271	Black Beauty	780	0.0870	
2775	The Good Soldier	426	0.1144	
1872	The Red House Mystery	422	0.1421	
25344	The Scarlet Letter	386	0.1266	
1051	Sartor Resartus: The Life and Op...	347	0.1074	
3328	Man and Superman: A Comedy and a...	312	0.1181	
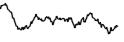
Table S20: Stories which are are represented by core emotional arc 10, sorted by the variance explained in their emotional arc by core emotional arc 10.

ID	Title	DL's	max(W[i,:])	Arc
21374	!Tention: A Story of Boy-Life du...	92	0.2195	
24499	The Green Carnation	65	0.1986	
544	Anne's House of Dreams	586	0.1676	
4922	Bar-20 Days	57	0.1648	
8681	The Face and the Mask	82	0.1528	
1805	The Gentle Grafter	84	0.1505	
10443	The Rayner-Slade Amalgamation	61	0.1486	
16339	The Passenger from Calais	125	0.1454	
24775	Up the River; or, Yachting on th...	88	0.1449	
9902	The Middle of Things	84	0.1327	
25305	Memoirs Of Fanny Hill: A New and ...	2,222	0.1296	
37858	Leaves in the Wind	79	0.1244	
24770	A Prisoner of Morro; Or, In the ...	45	0.1243	
25866	The Search	56	0.1209	
25919	Miss Mapp	49	0.1156	
35517	The Three Impostors; or, The Tra...	111	0.1148	
1020	Sword Blades and Poppy Seed	75	0.1141	
1535	The Tragedy of Coriolanus	116	0.1137	
38006	The Heatherford Fortune: a sequel...	44	0.1122	
38777	Lad: A Dog	58	0.1059	
38070	The Norwegian Fairy Book	149	0.1044	
1358	Enoch Arden, &c.	54	0.1042	
18613	The Golden Scorpion	84	0.0954	
4540	In His Steps	230	0.0902	
34943	Among the Meadow People	48	0.0727	
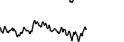
Table S21: Top 10 stories which are are represented by core emotional arc 10, sorted by downloads.

ID	Title	DL's	max(W[i,:])	Arc
25305	Memoirs Of Fanny Hill: A New and ...	2,222	0.1296	
544	Anne's House of Dreams	586	0.1676	
4540	In His Steps	230	0.0902	
38070	The Norwegian Fairy Book	149	0.1044	
16339	The Passenger from Calais	125	0.1454	
1535	The Tragedy of Coriolanus	116	0.1137	
35517	The Three Impostors; or, The Tra...	111	0.1148	
21374	!Tention: A Story of Boy-Life du...	92	0.2195	
24775	Up the River; or, Yachting on th...	88	0.1449	
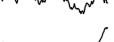
Table S22: Stories which are are represented by core emotional arc 11, sorted by the variance explained in their emotional arc by core emotional arc 11.

ID	Title	DL's	max(W[i,:])	Arc
5343	Rainbow Valley	257	0.1533	
1987	The Outlet	40	0.1444	
13882	John Thorndyke's Cases : related ...	217	0.1343	
1167	A Strange Disappearance	270	0.1282	
7477	The Book of Wonder	244	0.1276	
6753	Psmith in the City	199	0.1245	
5265	The Ball and the Cross	118	0.1229	
974	The Secret Agent: A Simple Tale	1,142	0.1224	
27525	Bones in London	85	0.1157	
24450	Bones: Being Further Adventures i...	121	0.1072	
5758	Many Cargoes	82	0.1048	
4090	From Ritual to Romance	90	0.1037	
27198	The Explorer	62	0.1035	
10832	Carnacki, the Ghost Finder	135	0.1032	
13372	The Gloved Hand	70	0.0942	
767	Agnes Grey	287	0.0856	
3055	The Wood Beyond the World	106	0.0799	
2641	A Room with a View	1,354	0.0681	
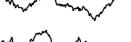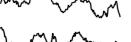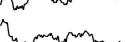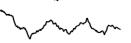
Table S23: Top 10 stories which are are represented by core emotional arc 11, sorted by downloads.

ID	Title	DL's	max(W[i,:])	Arc

Table S23: Top 10 stories which are are represented by core emotional arc 11, sorted by downloads.

ID	Title	DL's	max(W[i,:])	Arc
2641	A Room with a View	1,354	0.0681	
974	The Secret Agent: A Simple Tale	1,142	0.1224	
767	Agnes Grey	287	0.0856	
1167	A Strange Disappearance	270	0.1282	
5343	Rainbow Valley	257	0.1533	
7477	The Book of Wonder	244	0.1276	
13882	John Thorndyke's Cases : related ...	217	0.1343	
6753	Psmith in the City	199	0.1245	
10832	Carnacki, the Ghost Finder	135	0.1032	
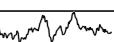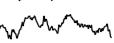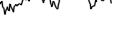
Table S24: Stories which are are represented by core emotional arc 12, sorted by the variance explained in their emotional arc by core emotional arc 12.

ID	Title	DL's	max(W[i,:])	Arc
318	John Barleycorn	130	0.1909	
1595	Whirligigs	94	0.1563	
25728	Desert Conquest; or, Precious Wa...	58	0.1543	
557	Puck of Pook's Hill	50	0.1482	
15976	Puck of Pook's Hill	40	0.1376	
26027	Puck of Pook's Hill	78	0.1357	
3326	The Well-Beloved: A Sketch of a ...	76	0.1226	
1611	Seventeen : A Tale of Youth and S...	56	0.1145	
12803	Headlong Hall	124	0.1113	
610	Idylls of the King	494	0.1075	
624	Looking Backward, 2000 to 1887	679	0.1060	
37532	The Scottish Fairy Book	98	0.1036	
10067	The Mystery of the Boule Cabinet...	54	0.0938	
316	The Golden Road	151	0.0864	
6840	Queen Lucia	81	0.0801	
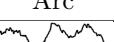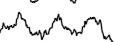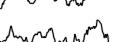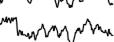
Table S25: Top 10 stories which are are represented by core emotional arc 12, sorted by downloads.

ID	Title	DL's	max(W[i,:])	Arc
624	Looking Backward, 2000 to 1887	679	0.1060	
610	Idylls of the King	494	0.1075	
316	The Golden Road	151	0.0864	
318	John Barleycorn	130	0.1909	
12803	Headlong Hall	124	0.1113	
37532	The Scottish Fairy Book	98	0.1036	
1595	Whirligigs	94	0.1563	
6840	Queen Lucia	81	0.0801	
26027	Puck of Pook's Hill	78	0.1357	
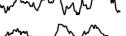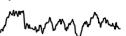
Table S26: Stories which are are represented by core emotional arc 13, sorted by the variance explained in their emotional arc by core emotional arc 13.

ID	Title	DL's	max(W[i,:])	Arc
3815	Rolling Stones	107	0.1377	
26732	Free Air	42	0.1295	
30836	Seven Keys to Baldpate	57	0.1245	
22342	Supermind	61	0.1185	
9925	Black Jack	111	0.1174	
21687	The Youngest Girl in the Fifth: ...	59	0.1097	
19527	The Yukon Trail: A Tale of the N...	44	0.1085	
6312	Representative Men: Seven Lectures	170	0.0986	
33391	Bill Nye's Cordwood	248	0.0888	
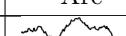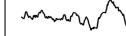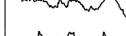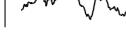
Table S27: Top 10 stories which are are represented by core emotional arc 13, sorted by downloads.

ID	Title	DL's	max(W[i,:])	Arc
33391	Bill Nye's Cordwood	248	0.0888	
6312	Representative Men: Seven Lectures	170	0.0986	
9925	Black Jack	111	0.1174	
3815	Rolling Stones	107	0.1377	
22342	Supermind	61	0.1185	
21687	The Youngest Girl in the Fifth: ...	59	0.1097	
30836	Seven Keys to Baldpate	57	0.1245	
19527	The Yukon Trail: A Tale of the N...	44	0.1085	
26732	Free Air	42	0.1295	
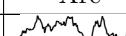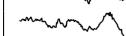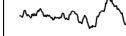
Table S28: Stories which are are represented by core emotional arc 14, sorted by the variance explained in their emotional arc by core emotional arc 14.

ID	Title	DL's	max($W[i,:]$)	Arc
5077	Marmion: A Tale of Flodden Field...	125	0.1476	
20869	The Skylark of Space	246	0.1457	
15272	Spenser's The Faerie Queene, Book I	978	0.1302	
1027	The Lone Star Ranger: A Romance ...	253	0.1277	
2233	A Damsel in Distress	201	0.1256	
27690	Nobody's Girl: (En Famille)	70	0.1219	
32759	Red Nails	151	0.1158	
27063	The Hero	60	0.1155	
2804	Rose in Bloom : A Sequel to "Eigh...	168	0.1083	
45658	The Mystery of the Downs	48	0.1030	
105	Persuasion	2,535	0.0980	
619	The Warden	215	0.0946	
34732	Max Carrados	92	0.0941	
291	The Golden Age	42	0.0854	
15673	The Day of the Beast	98	0.0792	
1671	When a Man Marries	46	0.0651	
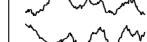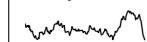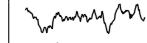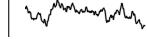
Table S29: Top 10 stories which are are represented by core emotional arc 14, sorted by downloads.

ID	Title	DL's	max($W[i,:]$)	Arc
105	Persuasion	2,535	0.0980	
15272	Spenser's The Faerie Queene, Book I	978	0.1302	
1027	The Lone Star Ranger: A Romance ...	253	0.1277	
20869	The Skylark of Space	246	0.1457	
619	The Warden	215	0.0946	
2233	A Damsel in Distress	201	0.1256	
2804	Rose in Bloom : A Sequel to "Eigh...	168	0.1083	
32759	Red Nails	151	0.1158	
5077	Marmion: A Tale of Flodden Field...	125	0.1476	
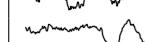
Table S30: Stories which are are represented by core emotional arc 15, sorted by the variance explained in their emotional arc by core emotional arc 15.

ID	Title	DL's	max($W[i,:]$)	Arc
33066	The Garden of Eden	72	0.1165	
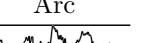
Table S30: Stories which are are represented by core emotional arc 15, sorted by the variance explained in their emotional arc by core emotional arc 15.

ID	Title	DL's	max(W[i,:])	Arc
40852	Instigations: Together with An Es...	79	0.0806	

Table S31: Top 10 stories which are are represented by core emotional arc 15, sorted by downloads.

ID	Title	DL's	max(W[i,:])	Arc
40852	Instigations: Together with An Es...	79	0.0806	
33066	The Garden of Eden	72	0.1165	

Table S32: Stories which are are represented by core emotional arc 16, sorted by the variance explained in their emotional arc by core emotional arc 16.

ID	Title	DL's	max(W[i,:])	Arc
872	Reprinted Pieces	49	0.1103	
8673	A Columbus of Space	56	0.1077	
434	The Circular Staircase	189	0.0978	
1263	The Glimpses of the Moon	58	0.0864	
3075	The Return	73	0.0759	

Table S33: Top 10 stories which are are represented by core emotional arc 16, sorted by downloads.

ID	Title	DL's	max(W[i,:])	Arc
434	The Circular Staircase	189	0.0978	
3075	The Return	73	0.0759	
1263	The Glimpses of the Moon	58	0.0864	
8673	A Columbus of Space	56	0.1077	
872	Reprinted Pieces	49	0.1103	

Table S34: Stories which are are represented by core emotional arc 18, sorted by the variance explained in their emotional arc by core emotional arc 18.

ID	Title	DL's	$\max(W[\text{i},:])$	Arc
5776	100%: the Story of a Patriot	64	0.1731	
3188	Mark Twain's Speeches	500	0.0778	

Table S35: Top 10 stories which are are represented by core emotional arc 18, sorted by downloads.

ID	Title	DL's	$\max(W[\text{i},:])$	Arc
3188	Mark Twain's Speeches	500	0.0778	
5776	100%: the Story of a Patriot	64	0.1731	

Table S36: Stories which are are represented by core emotional arc 20, sorted by the variance explained in their emotional arc by core emotional arc 20.

ID	Title	DL's	$\max(W[\text{i},:])$	Arc
15119	Handy Dictionary of Poetical Quo...	46	0.0728	

Table S37: Top 10 stories which are are represented by core emotional arc 20, sorted by downloads.

ID	Title	DL's	$\max(W[\text{i},:])$	Arc
15119	Handy Dictionary of Poetical Quo...	46	0.0728	

Appendix F: Additional Hierarchical Clustering Figures

In the section, we include additional results from the hierarchical clustering analysis.

A — Cluster 2650 (737)

1. Alice's Adventures in Wonderland (11)
2. A Man of Means (8713)
3. Alice's Adventures in Wonderland: Illustrated by... (28885)
4. The Book of Wonder (7477)
5. The Head of Kay's (6877)

B — Cluster 2650 (590)

1. The Aesop for Children⬚: With pictures by Milo W... (19994)
2. Code Three (19111)
3. Peter and Wendy (26654)
4. Island Nights' Entertainments (329)
5. The Green Rust (24929)

Figure S9 The 2 clusters identified by Agglomerative Clustering using Ward's method.

A — Cluster 2650 (161)

1. The Time Machine (35)
2. A Nonsense Anthology (9380)
3. Tom Swift and His Big Tunnel; Or, The Hidden Ci... (953)
4. The Light That Failed (2876)
5. The Elusive Pimpernel (2785)

B — Cluster 2650 (429)

1. The Aesop for Children⬚: With pictures by Milo W... (19994)
2. Arizona Nights (753)
3. Uncle Bernac: A Memory of the Empire (10581)
4. The Young Pitcher (19246)
5. Oliver Twist, Vol. 1 (of 3) (47529)

C — Cluster 2649 (737)

1. Alice's Adventures in Wonderland (11)
2. A Man of Means (8713)
3. Alice's Adventures in Wonderland: Illustrated by... (28885)
4. The Book of Wonder (7477)
5. The Head of Kay's (6877)

Figure S10 The 3 clusters identified by Agglomerative Clustering using Ward's method.

Figure S11 The 4 clusters identified by Agglomerative Clustering using Ward's method.

Figure S12 The 5 clusters identified by Agglomerative Clustering using Ward's method.

Figure S13 The 6 clusters identified by Agglomerative Clustering using Ward's method.

Figure S14 The 7 clusters identified by Agglomerative Clustering using Ward's method.

Figure S15 The 8 clusters identified by Agglomerative Clustering using Ward's method.

Figure S16 The 9 clusters identified by Agglomerative Clustering using Ward's method.

Figure S17   The silhouette plots for 2–5 clusters identified by Agglomerative Clustering using Ward's method.

Figure S18 The silhouette plots for 6–9 clusters identified by Agglomerative Clustering using Ward's method.

Appendix G: Additional SOM Figures

In Fig. S19 we show the emotional arcs that are closest to each of 9 most frequently winning nodes in the winner-take-all implementation the Self Organizing Map.

Figure S19   The vector for each of the top 9 SOM nodes, accompanied with those sentiment time series which are closest to that node. The core stories which we have found with other methods are readily visible.
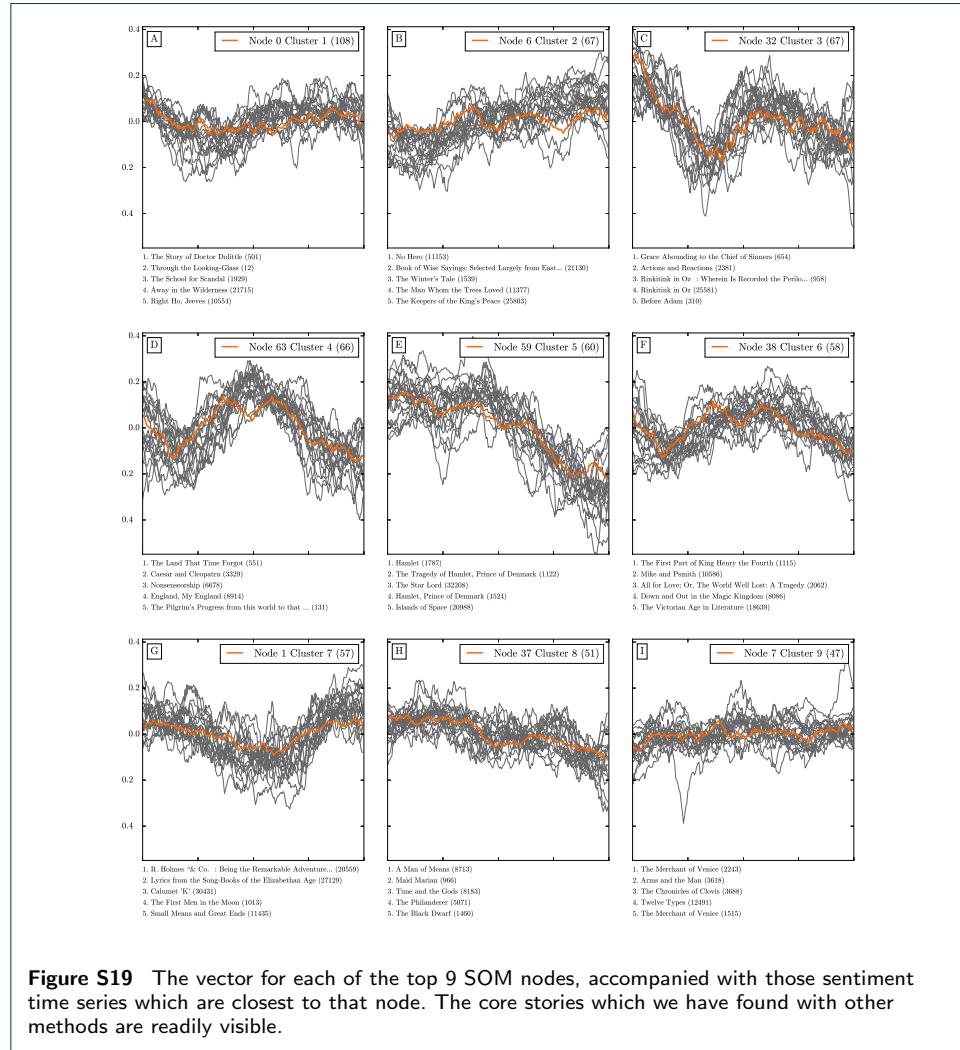
Appendix H: Null comparison details

An example of the "nonsense" and "word salad" text is presented first in Appendix C. First, we examine the resulting timeseries for an example book in Figs. S20 and S21. We then go on to present the full result of the SVD, agglomerative clustering, and SOM to "nonsense" English fiction books with more than 40 downloads.
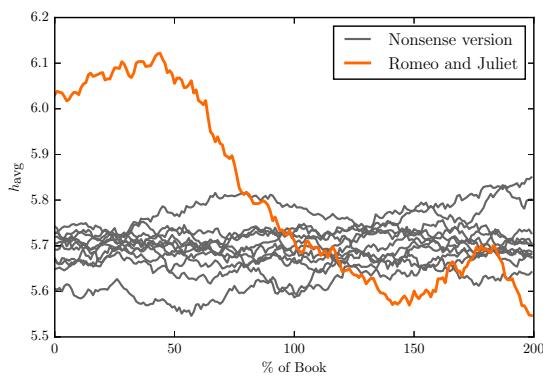
Figure S20 The emotional arc of *Romeo And Juliet* by William Shakespeare (Gutenberg ID 1777), along with 11 "nonsense" versions, as produced by a 2-gram Markov model. We see that the emotional arc from the true version has more structure than the nonsense versions.
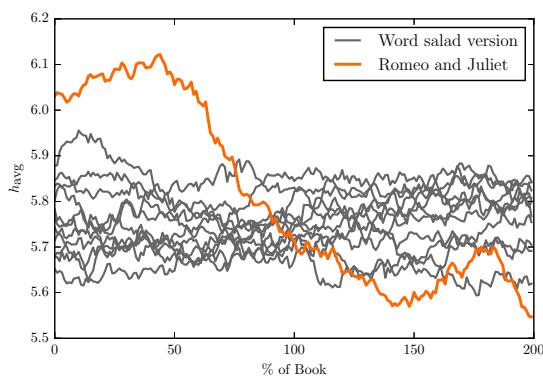
Figure S21 The emotional arc of *Romeo And Juliet* by William Shakespeare, along with 11 "word salad" versions, as produced by randomly shuffling the words in the book. We see that the emotional arc from the true version has more structure than the word salad versions as well.

H.1 Null SVD

SVD modes from the emotional arcs of word salad books. We observe higher frequency modes appearing more quickly, and a more even spread of mode coefficients.

Figure S22 Top 12 modes from the Singular Value Decomposition of 1,327 nonsense Project Gutenberg books. We show in a lighter color modes weighted by their corresponding singular value, where we have scaled the matrix Σ such that the first entry is 1 for comparison. The mode coefficients normalized for each book are shown in the right panel accompanying each mode, in the range -1 to 1, with the "Tukey" box plot.

Figure S23 First 3 SVD modes from nonsense books and their negation with the closest stories to each. Links below each story point to an interactive visualization on http://hedonometer.org which enables detailed exploration of the emotional arc for the story.

Figure S24 Modes 4–6 from the SVD analysis of nonsense books and their negation with the closest stories to each. Links below each story point to an interactive visualization on http://hedonometer.org which enables detailed exploration of the emotional arc for the story.

Figure S25 Comparison of the singular value spectra from the emotional arcs of nonsense books and the emotional arcs of individual Project Gutenberg books. The spectra from the nonsense books is muted, indicating both lower total variance explained and less important ordering of the singular vectors.

H.2 Null Hierarchical Clustering

Dendrogram of clustering using Ward's method on the emotional arcs of word salad books. We observe comparatively low linkage cost for these emotional arcs, indicating the absence of distinct clusters.

Figure S26 Dendrogram from the agglomerative clustering procedure using Ward's minimum variance method on nonsense books. For each cluster, a selection of the 20 most central books to a fully-connected network of books are shown along with the average of the emotional arc for all books in the cluster, along with the cluster ID and number of books in each cluster (shown in parenthesis). At the bottom, we show the average Silhouette value for all books, with higher value representing a more appropriate number of clusters. For each of the 60 leaf nodes (right side) we show the number of books within the cluster and the most central book to that cluster's book network.

Figure S27 Four clusters (linkage threshold 850) from the hierarchical clustering of word salad books. We observe that the cluster mean emotional arc and the most central emotional arcs have high variance, without a visible signal.

H.3 Null Self Organizing Map (SOM)

Figure S28 Results of the SOM applied to nonsense versions of Project Gutenberg books. Left panel: Nodes on the 2D SOM grid are shaded by the number of stories for which they are the winner. Right panel: The B-Matrix shows that there are clear clusters of stories in the 2D space imposed by the SOM network.

Figure S29 The vector for each of the top 9 SOM nodes for null emotional arcs, accompanied with those sentiment time series which are closest to that node. Panels D and E show what appear to be similar arcs to the six we identified in real books, but overall see that the emotional arcs from null arcs show little coherent structure, especially considering the y-range here being 0.1 compared to the 0.4 of the real books (had we used the same y-range, very little of the variation would be visible at all).

References

1. Cobley, P.: Narratology. The Johns Hopkins Guide to Literary Theory and Criticism, 2nd ed. John Hopkins University Press, London (2005)
2. Harris, W.F.: The Basic Patterns of Plot. University of Oklahoma Press, Oklahoma (1959)
3. Abbott, H.P.: The Cambridge Introduction to Narrative. Cambridge University Press, Massachusetts (2008)
4. Booker, C.: The Seven Basic Plots: Why We Tell Stories. Bloomsbury Academic, New York (2006)
5. Tobias, R.B.: 20 Master Plots: And How to Build Them. Writer's Digest Books, Ohio (1993)
6. Polti, G.: The Thirty-Six Dramatic Situations. James Knapp Reeve, Ohio (1921)
7. Vonnegut, K.: Palm Sunday. RosettaBooks LLC, New York (1981)