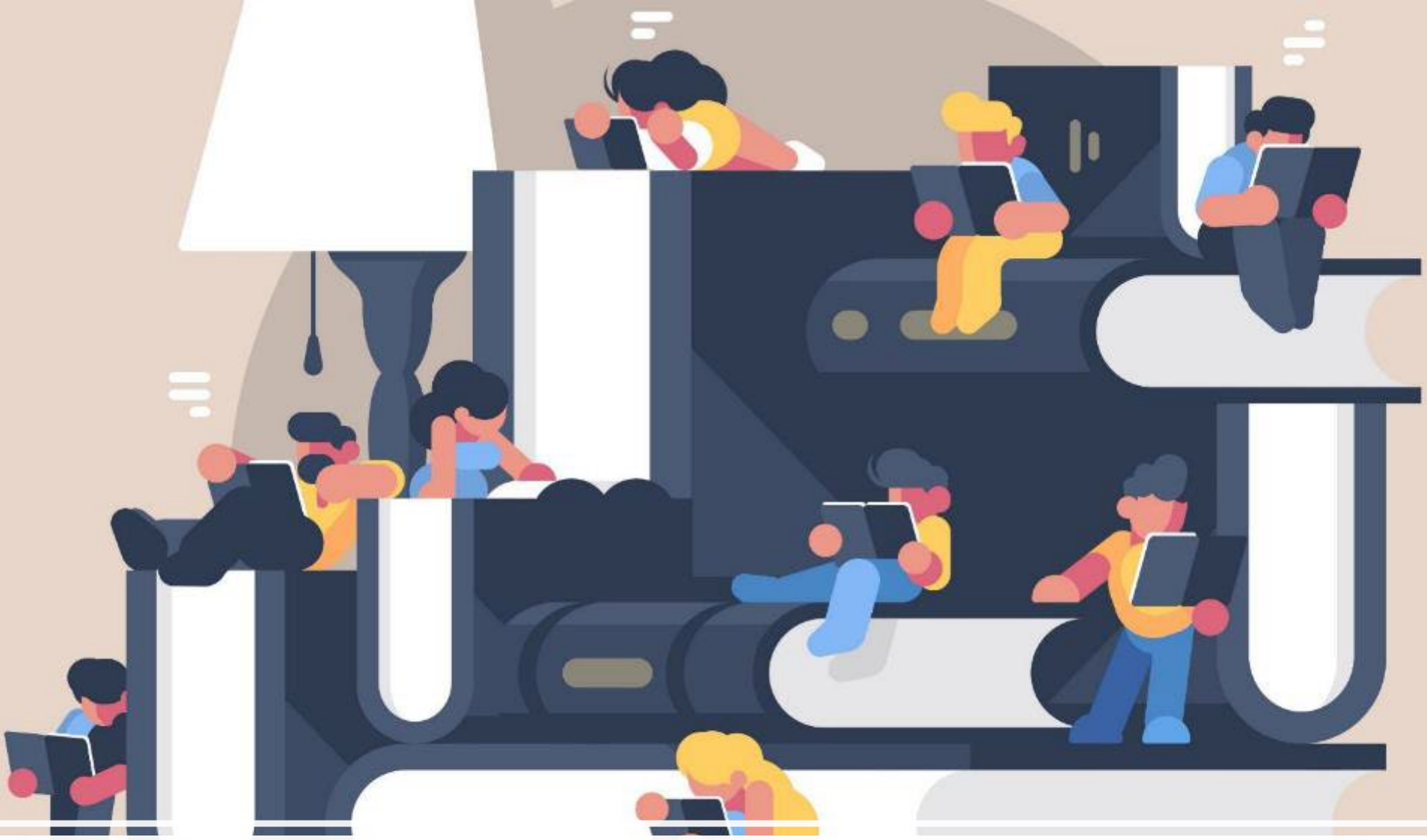
Combining ER and DR(Distant Reading) Using Literary Texts with Computer Coding

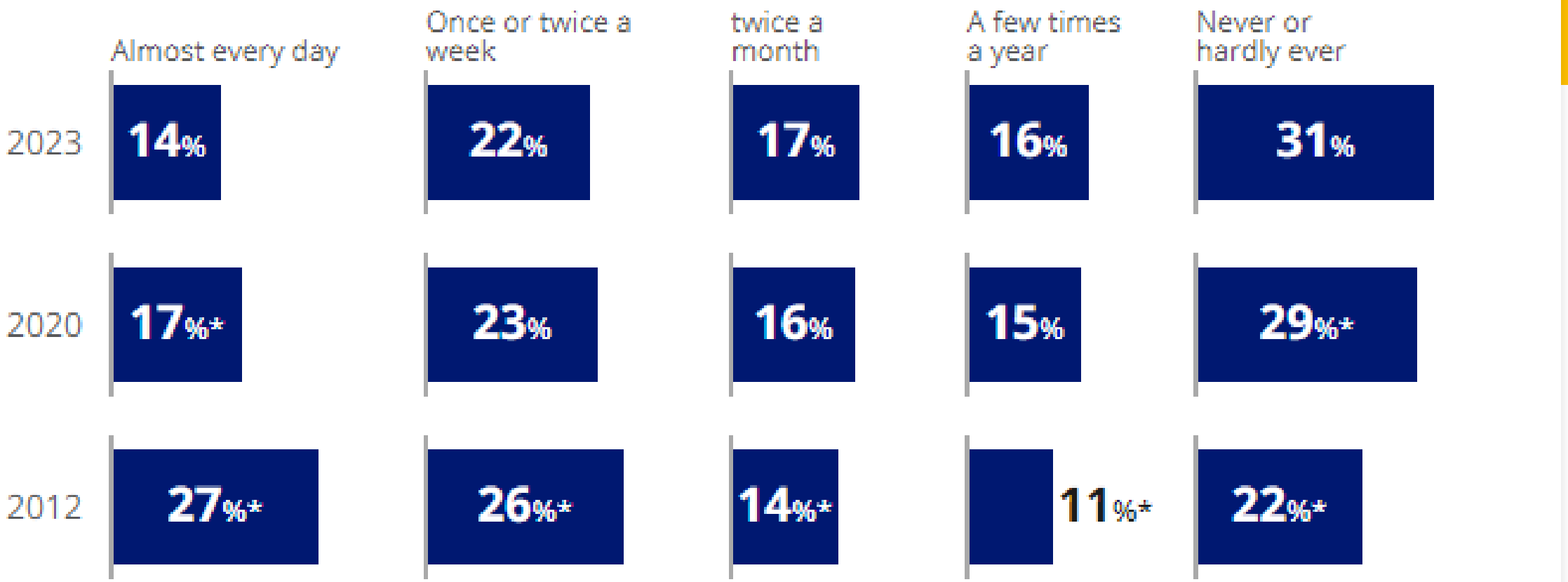
- Youngwoo KIM
- Hyunkee AHN
- Yongkook WON
- Kwanghyun PARK
- Junghye Fran CHOI
- Joohae KIM



Reading is important, but …

Americans read fewer books.

https://news.gallup.com/poll/388541/americans-reading-fewer-books-past.aspx



Among many U.S. children, reading for fun has become less common.

https://www.nationsreportcard.gov/highlights/ltt/2023/

Children are reading less than ever before, research reveals

Just a quarter of under-18s read each day, study shows in run-up to World Book Day

Just 53% of children said they enjoyed reading 'very much' or 'quite a lot'. Photograph: Danny Lawson/PA

Children are reading less than ever before in U.K.

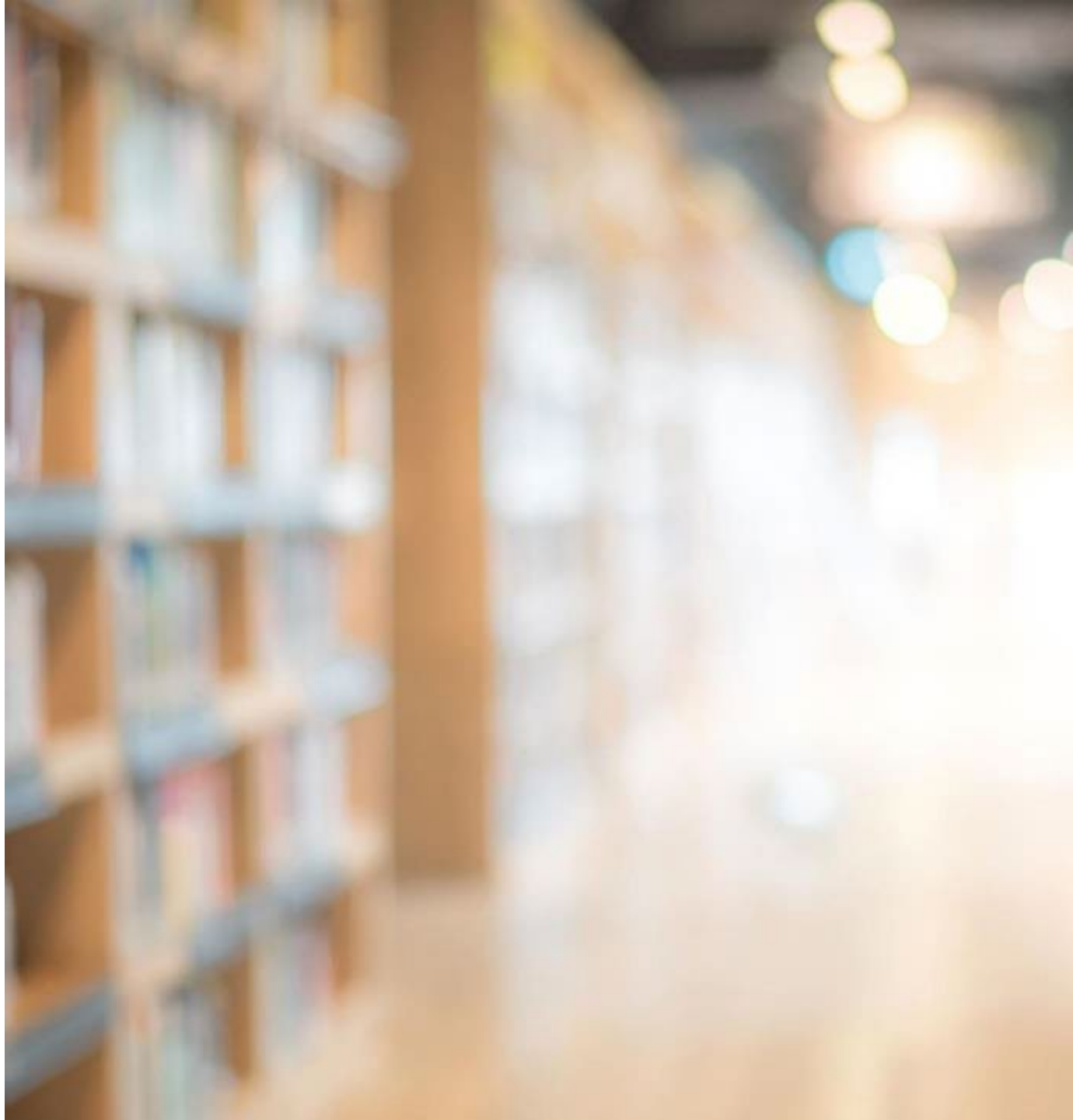
https://www.theguardian.com/education/2020/feb/29/children-reading-less-says-new-research

What about people in your country?

Reading is good for

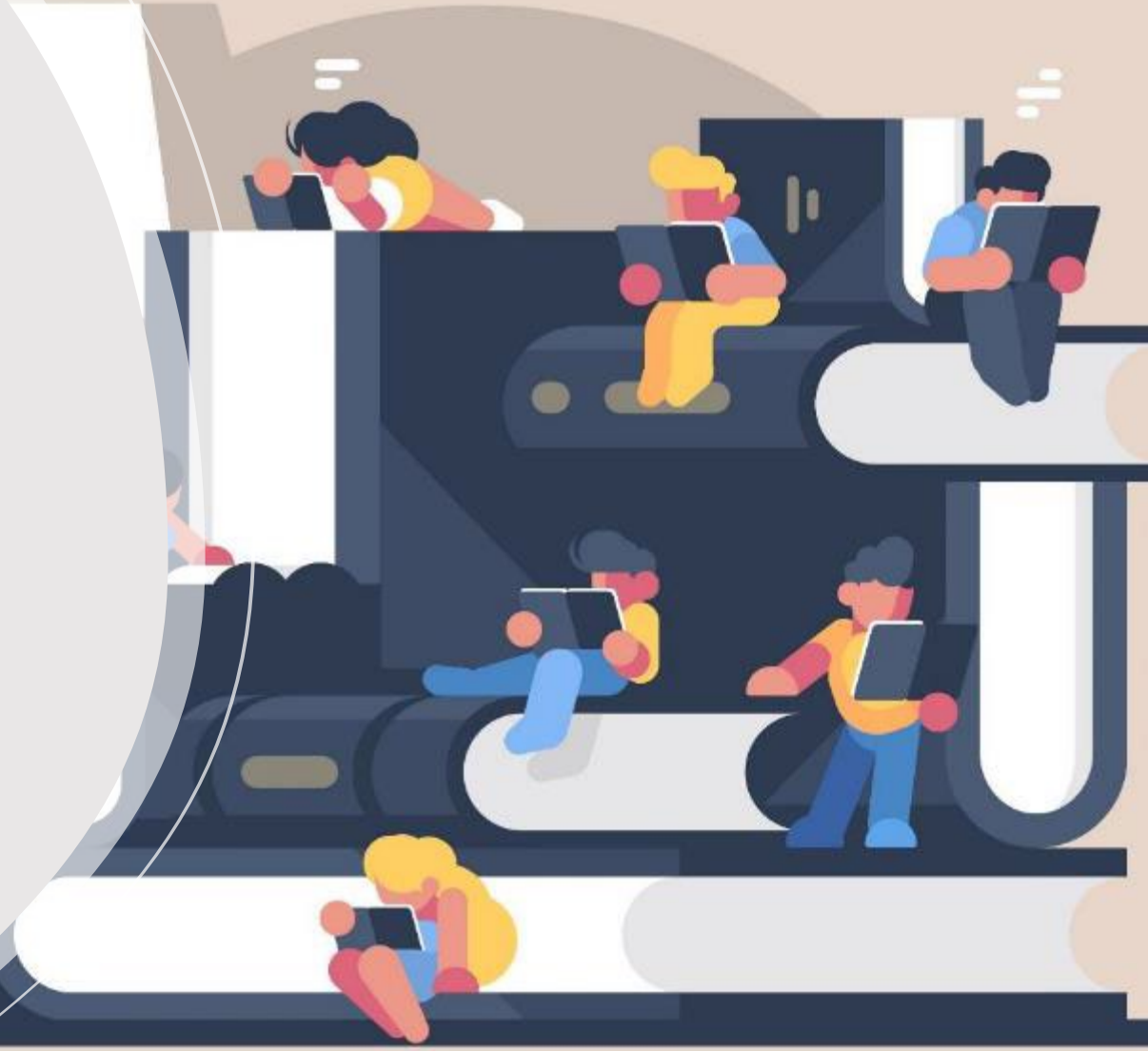
- fun,
- experience, and
- knowledge.

However, other activities such as watching videos, playing games, and so on, can also provide enjoyment, experiences, and knowledge.



Then, why we still need to read?

There are many reasons. Among them,

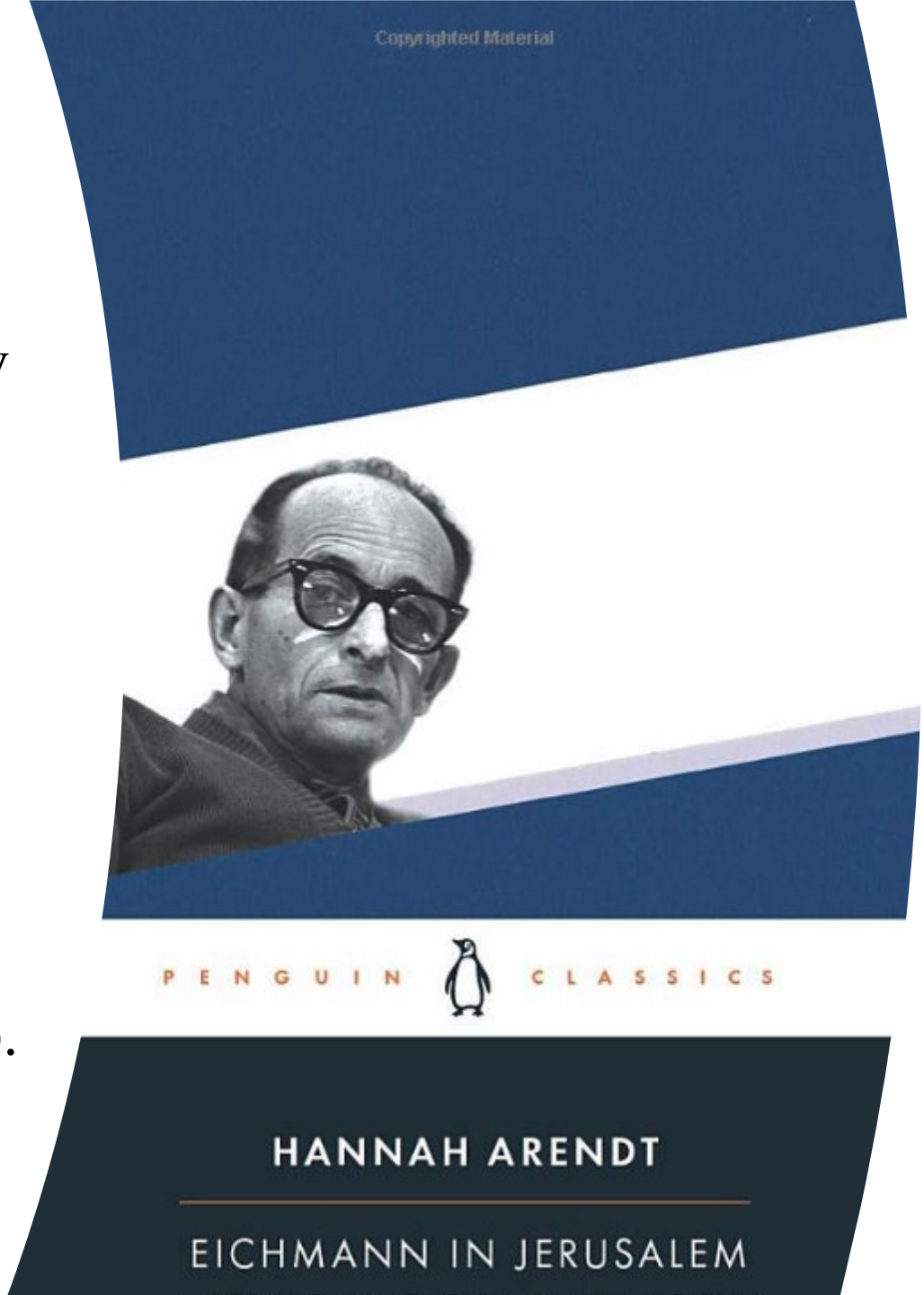


Reading ensures time for thinking.

And Extensive Reading is good for thinking.

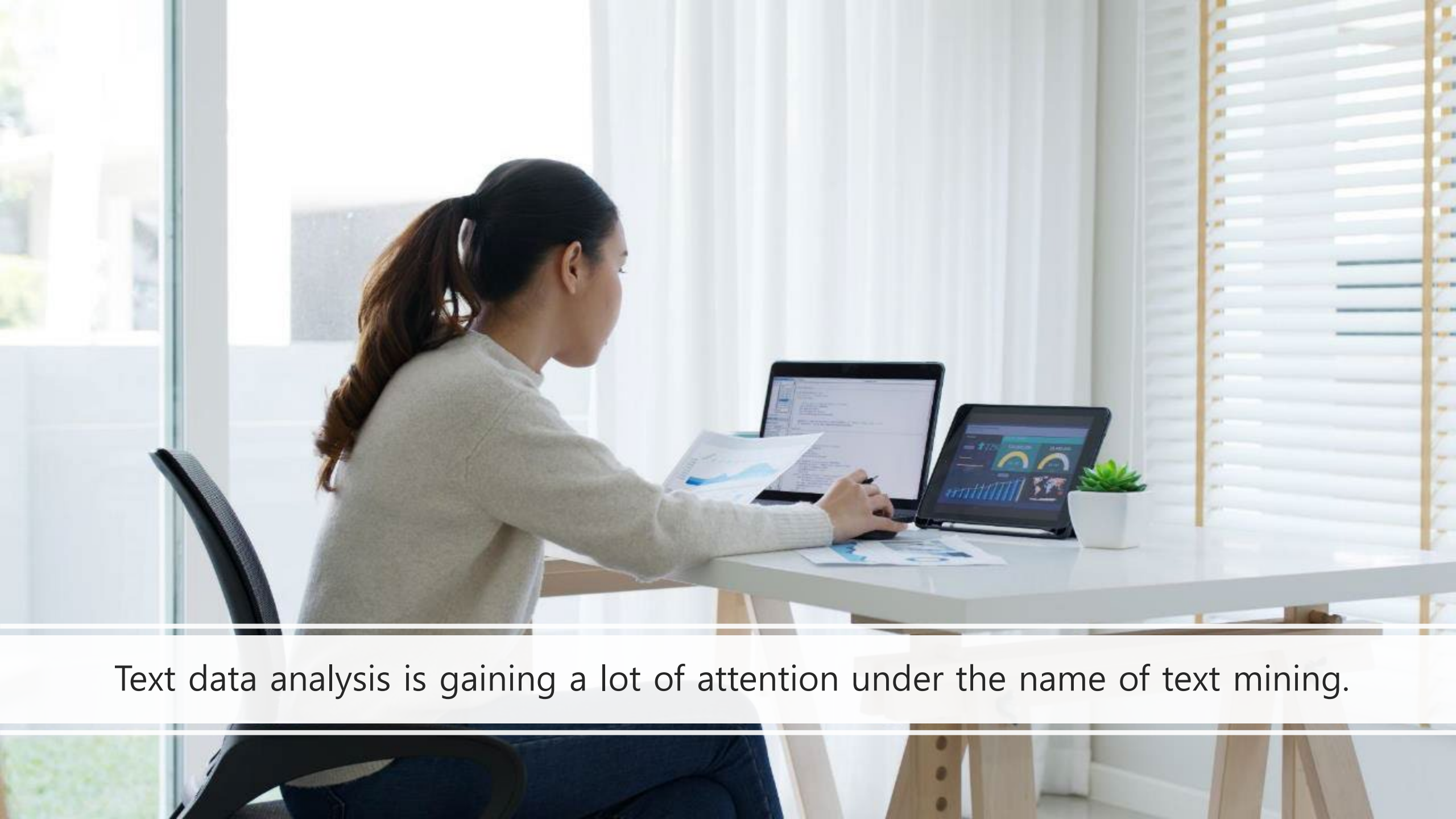
Because ER provides a variety of inputs.

She [Arendt] concluded that Eichmann's inability to speak coherently in court was connected with his incapacity to think, or to think from another person's point of view. His shallowness was by no means identical with stupidity (p.14).



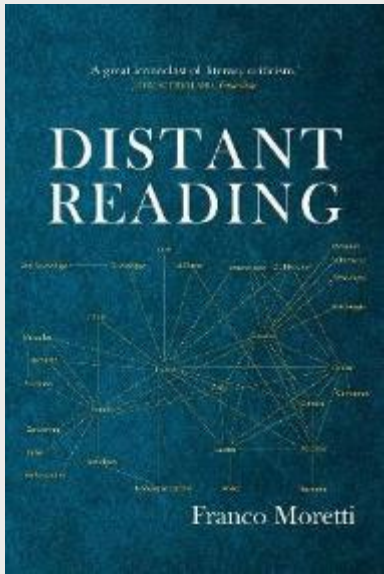
Evil comes from a failure to think. It defies thought for as soon as thought tries to engage itself with evil and examine the premises and principles from which it originates, it is frustrated because it finds nothing there. That is the banality of evil (p.15).

However, ER requires a significant time commitment that is precious in this era.

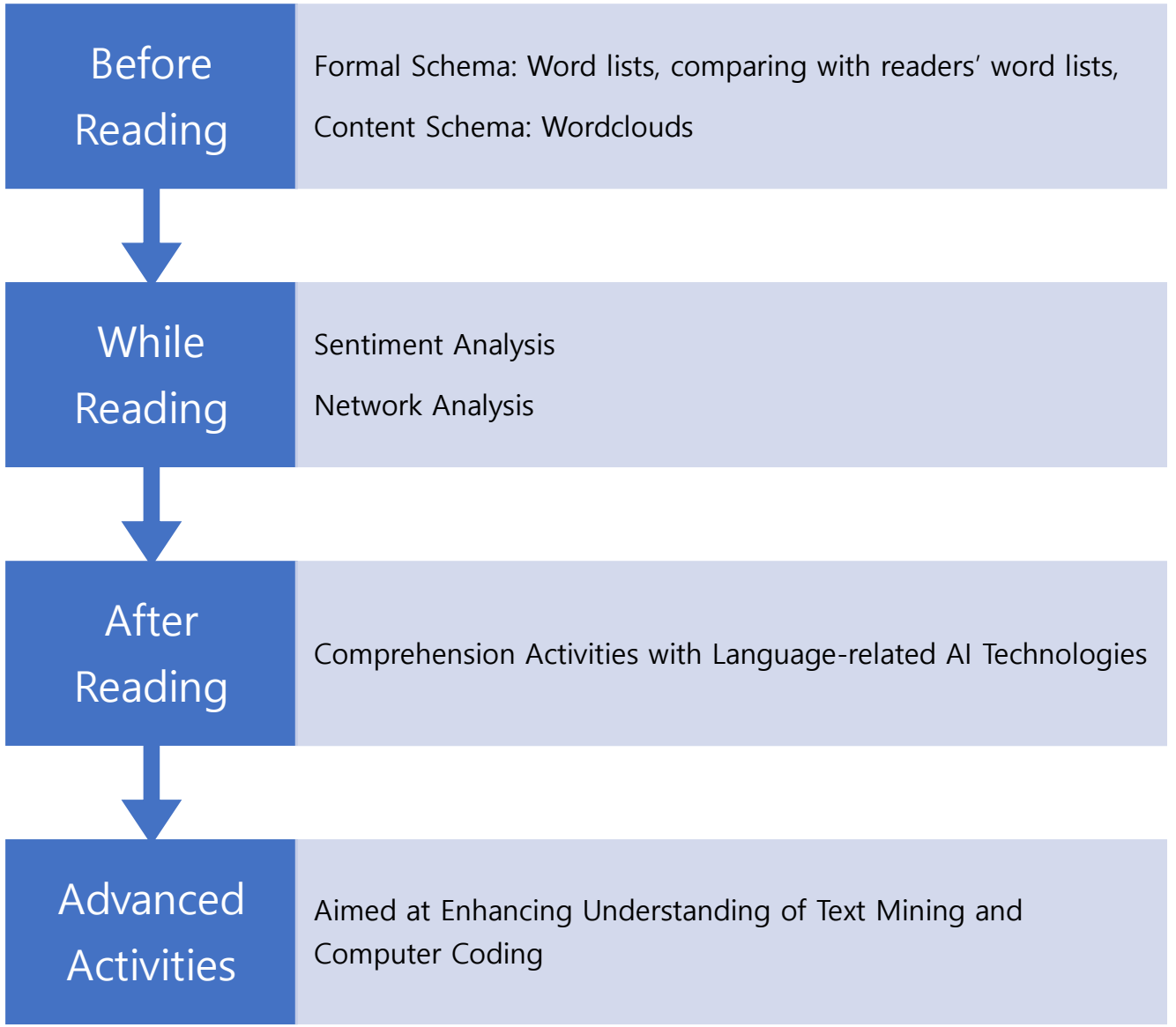


Text data analysis is gaining a lot of attention under the name of text mining.

Distant reading with text mining can be combined with ER.



Here are some text mining methods for ER.



Frequency analysis enables the creation of a list of words

that appear frequently in literary texts.

Such list of words can contribute to the formation of a formal schema, which assists readers in understanding literary texts.

	word	n
1	gatsby	197
2	tom	177
3	daisy	150
4	house	95
5	eyes	89
6	looked	85
7	time	83
8	car	73
9	night	73
10	moment	70
11	door	68
12	gatsby's	67
13	jordan	64
14	hand	63
15	people	62
16	wilson	62
17	voice	58
18	girl	47
19	half	47
20	sport	47

Words from literary texts can be compared with

words from readers' word lists.

Readers can improve their vocabulary skills cumulatively by comparing the words of literary texts with the words from their own word lists.

	A	B	C
	word	n	sourc
5	sport	47	gls2000
0	afternoon	42	gls2000
0	suddenly	40	gls2000
8	baker	38	gls2000
1	garage	28	gls2000
8	inquired	26	gls2000
4	feet	23	gls2000
6	hair	22	gls2000
3	yellow	22	gls2000
4	chair	19	gls2000
3	inside	19	gls2000
7	telephone	19	gls2000
6	dozen	18	gls2000
9	grey	18	gls2000
2	island	18	gls2000
9	slowly	18	gls2000
9	rain	17	gls2000

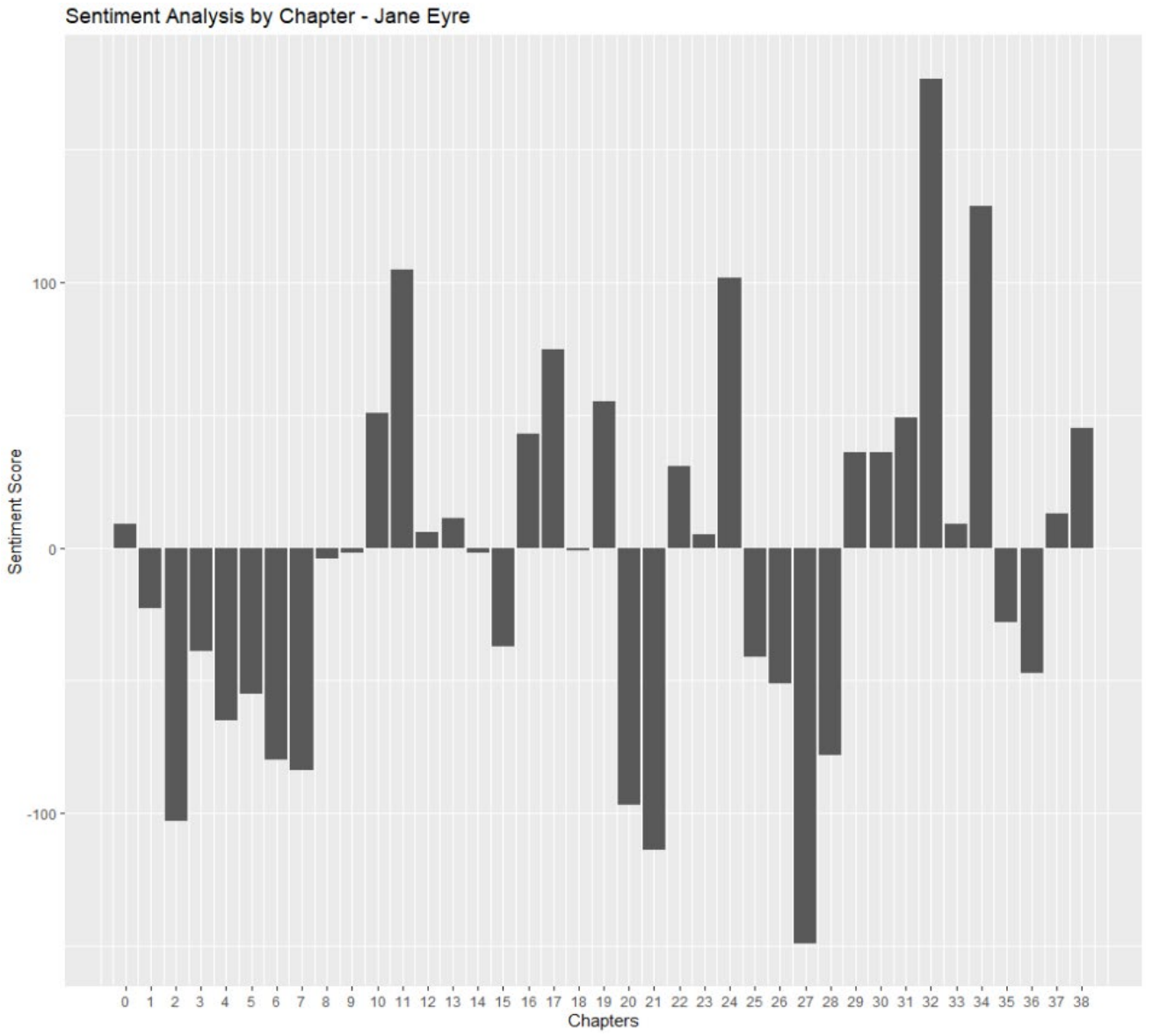
Wordclouds can be created

using words from literary texts.

Wordclouds can assist in forming content schemas, which aid readers in understanding literary texts.

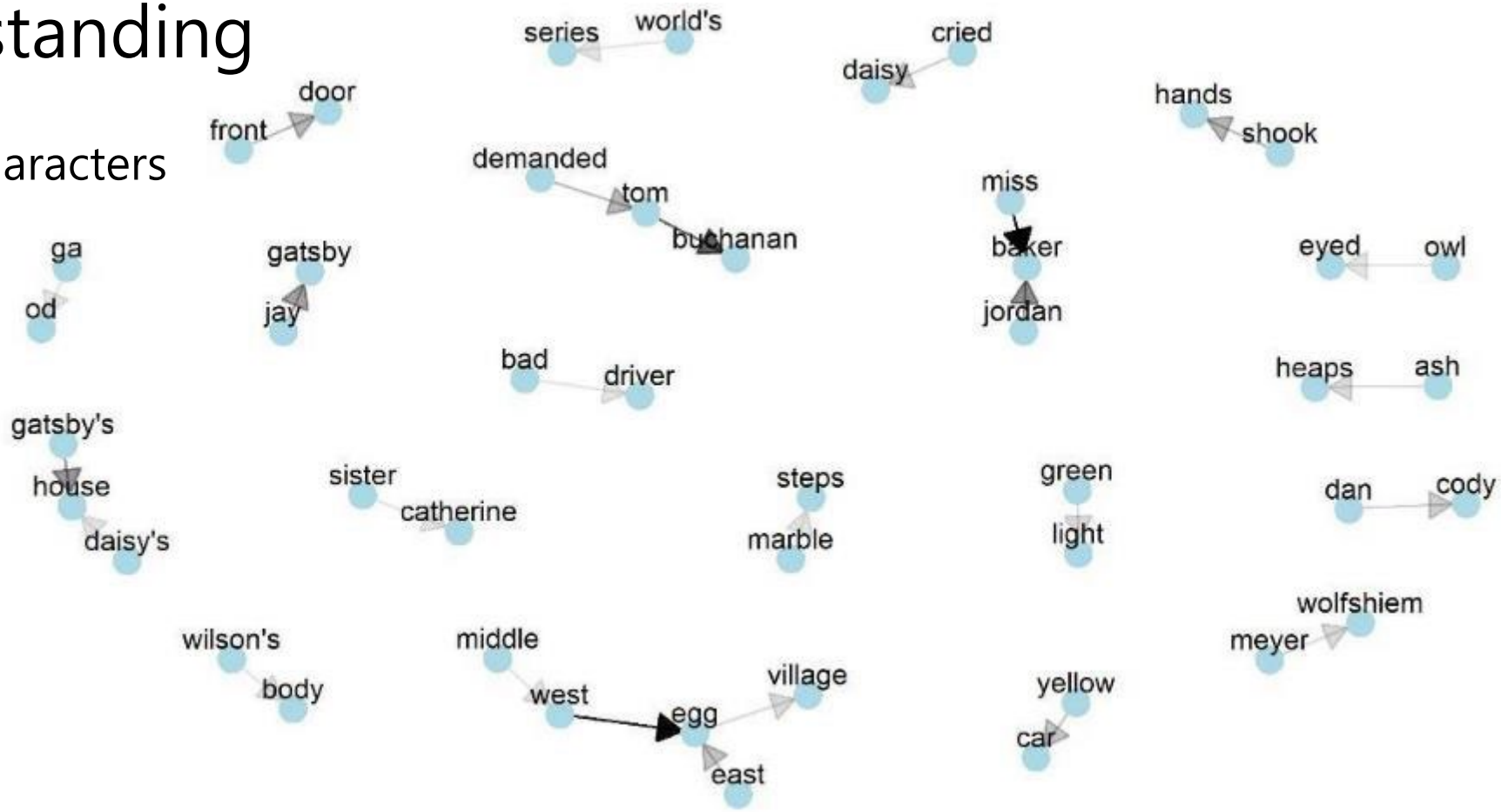
By conducting sentiment analysis according to the chapter,

it is possible to gain an understanding of the development process of a literary work.



Network analysis can assist in understanding

the characteristics of characters and the relationships within literary texts.

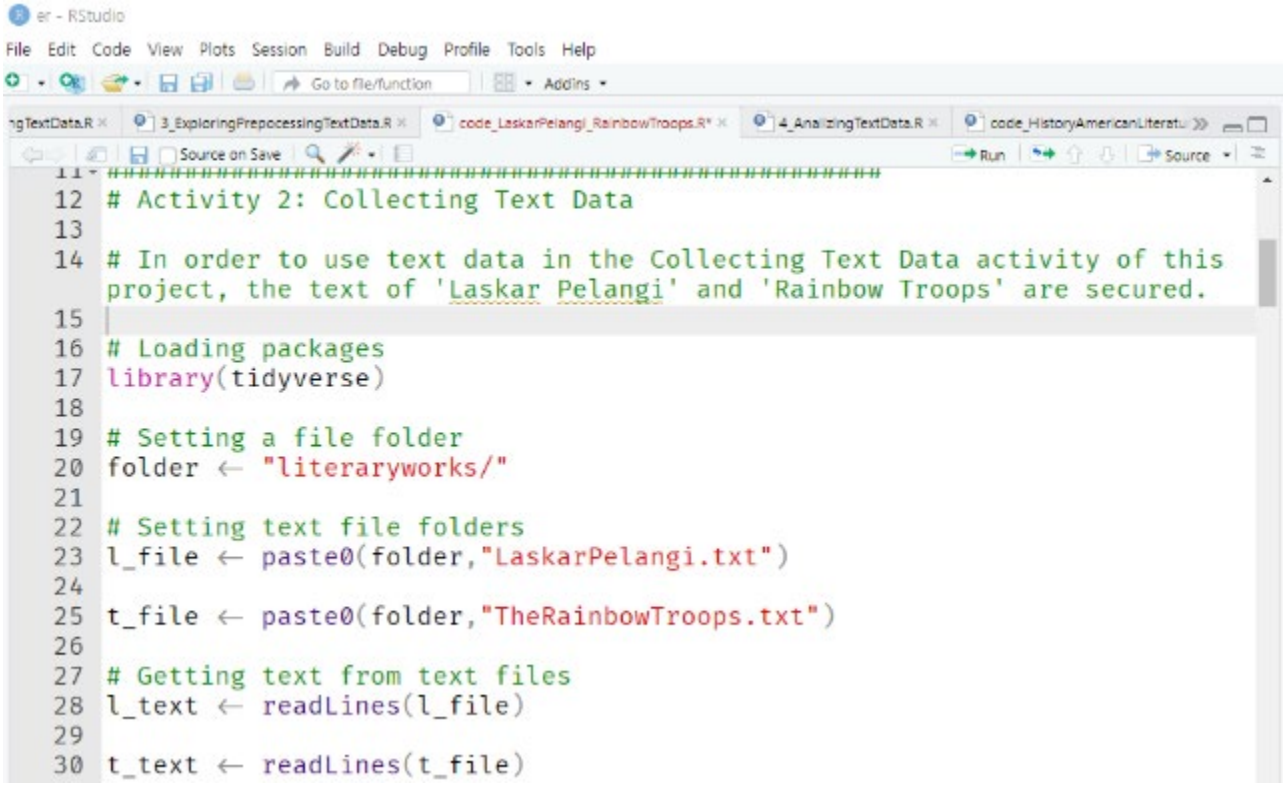


Improving the comprehension of literary texts can be achieved

through various language-related AI technologies.

Through advanced activities, readers can acquire the skills

to analyze the language data of literary texts using computer coding techniques.



Here are sample literary works

- Jane Eyre
- The Great Gatsby
- The Red Badge of Courage
- Through the Arc of the Rain Forest
- Rainbow Troops / Laskar Pelangi

Demonstration

DEMO

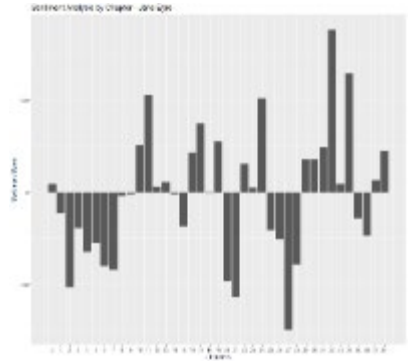
Extensive Reading and Language Data Activities

1. CHOOSE A BOOK AND READ IT

2. UTILIZE THE BOOK'S TEXT DATA

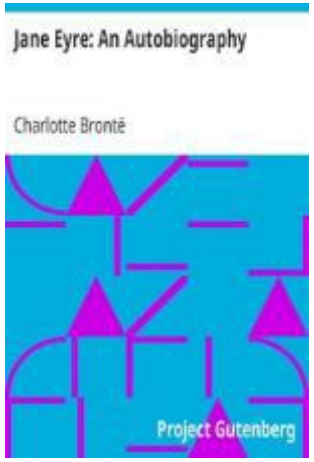
	A	B	C	D
1	no	chapter	text	word_count
2	1	0	JANE EYRE	2
3	2	0	AN AUTOBIOGRAPHY	2
4	3	0	by Charlotte Brontë	3
5	4	0	ILLUSTRATED BY F. H. TOWNSEND	5
6	6	0	London	1
7	7	0	SERVICE & PATON	3
8	8	0	5 HENRIETTA STREET	3
9	9	0	1897	1
10	11	0	The illustrations	2
11	12	0	in this Volume are the copyright of	7
12	13	0	SERVICE & PATON, London	4
13	15	0	TO	1
14	16	0	W. M. THACKERAY, ESQ.,	4
15	18	0	This Work	2
16	19	0	IS RESPECTFULLY INSCRIBED	3
17	21	0	BY	1
18	22	0	THE AUTHOR	2
19	24	0	PREFACE	1
20	25	0	A preface to the first edition of "Jane Eyre" being unnecessary, I gave none: this second edition demands a few words both of acknowledgment and miscellaneous remark.	27

3. MINING THE TEXT



DEMO

ER and Language Data Activities



1. Choose a book and read it

   1) Choose a book to read. An out-of-copyright book is preferable.

      Sample: Jane Eyre: An Autobiography by Charlotte Brontë

   2) Obtain the book to read. For ebooks, get the file.

      Sample: Jane Eyre at Project Gutenberg (https://www.gutenberg.org/ebooks/1260)

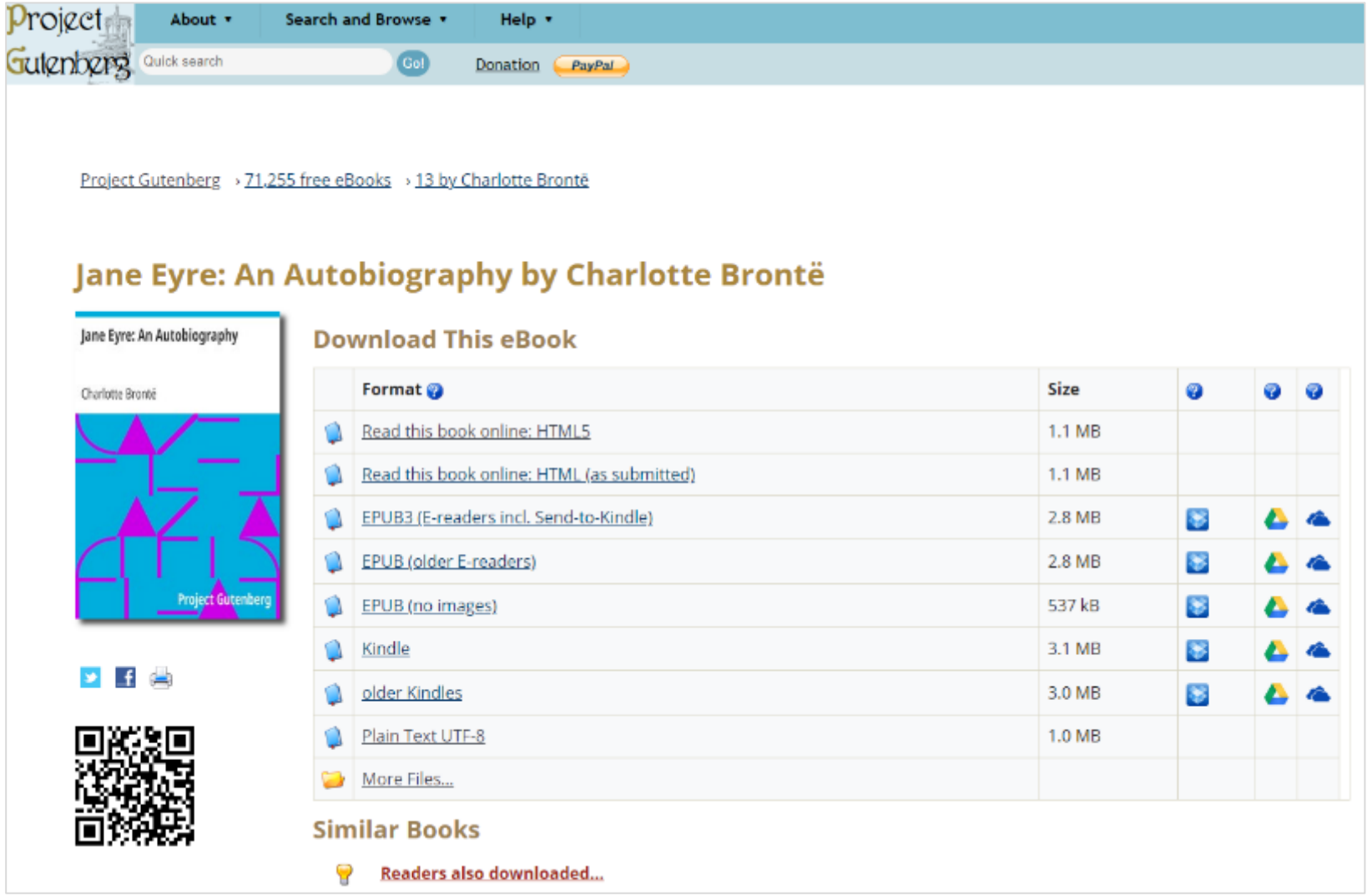
   3) Read the book.

      For Kindle: https://www.amazon.com/gp/help/customer/display.html?nodeId=G5WYD9SAF7PGXRNA

      Consider using a reading tracking app to record your book reading activity.

      Ex: Goodreads, StoryGraph, Basmo, Italic Type, Bookshelf, Bookly

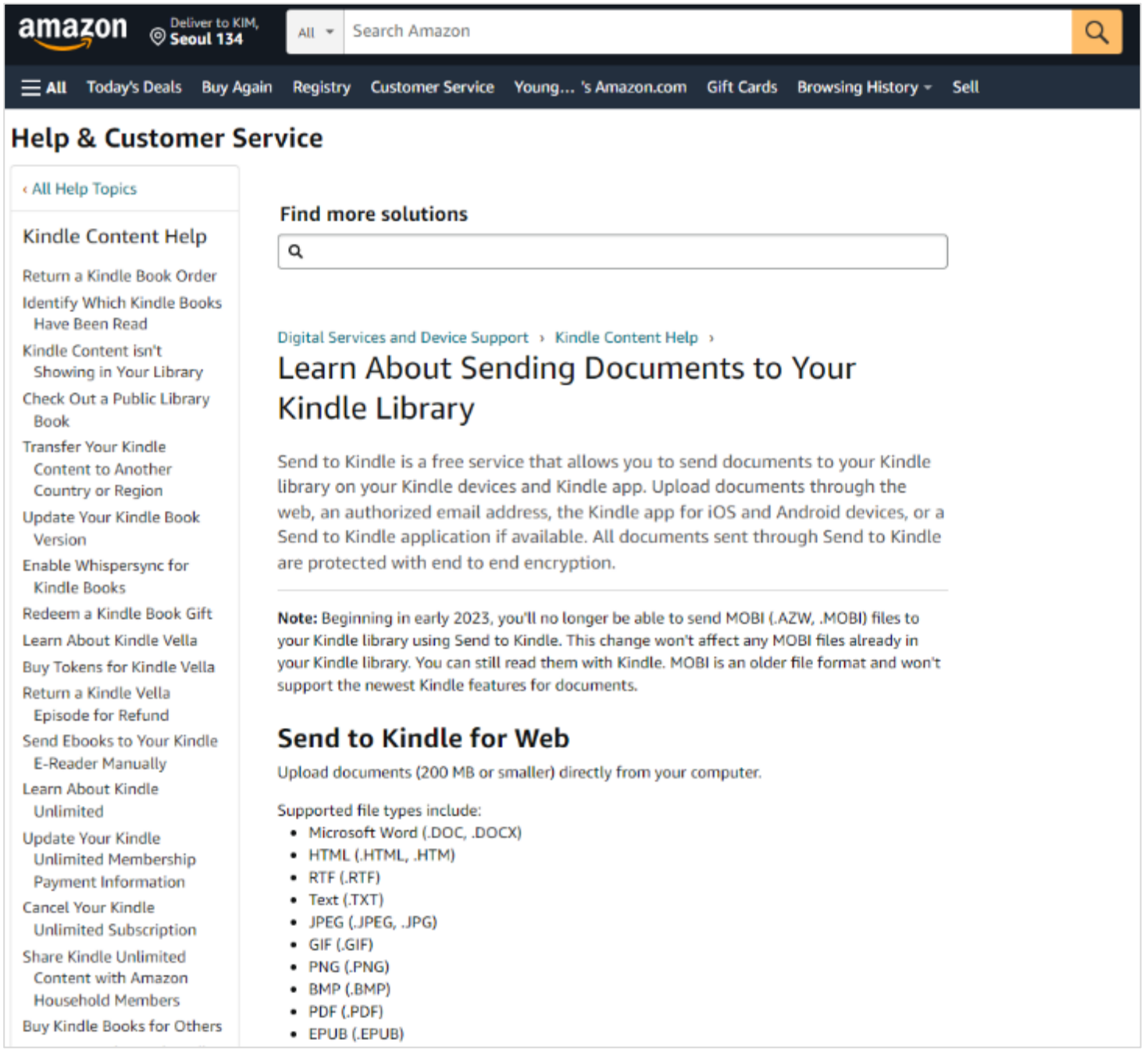
Jane Eyre at Project Gutenberg



https://www.gutenberg.org/ebooks/1260

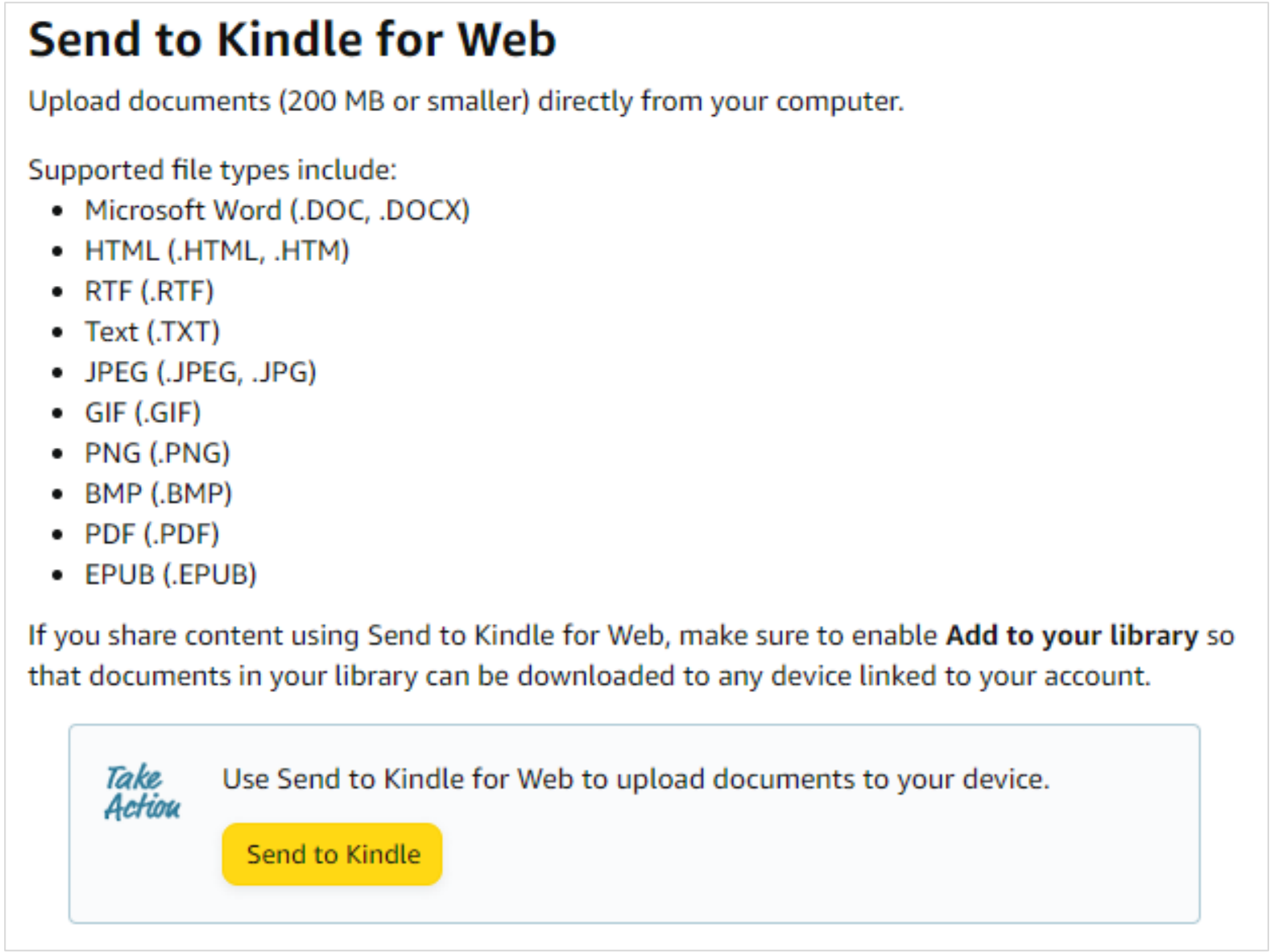
Amazon Kindle

https://www.amazon.com/gp/help/customer/display.html?nodeId=G5WYD9SAF7PGXRNA

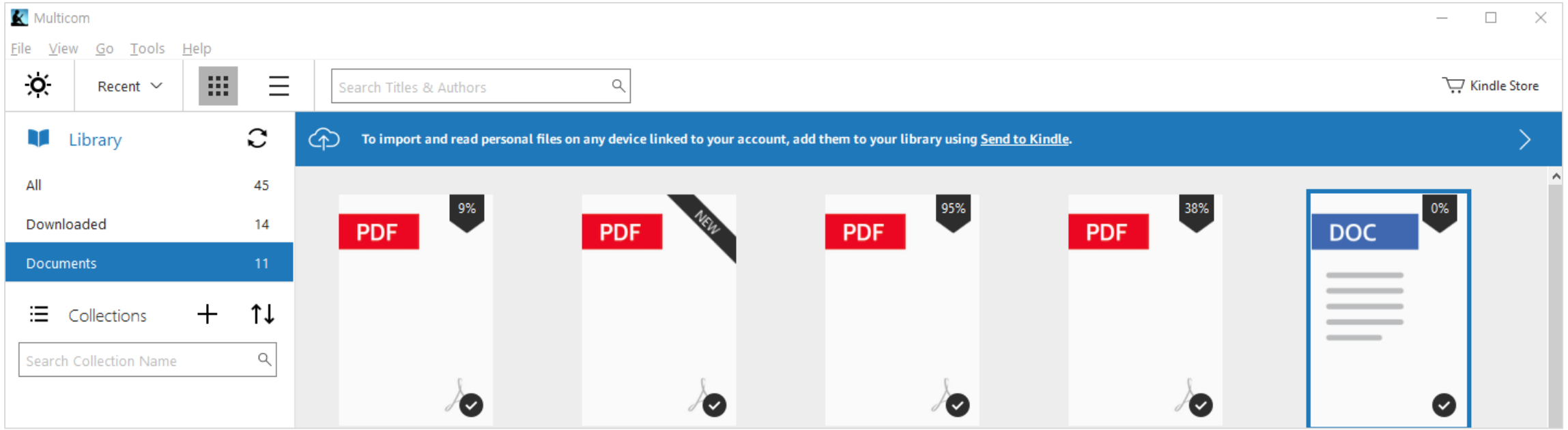


Amazon Kindle

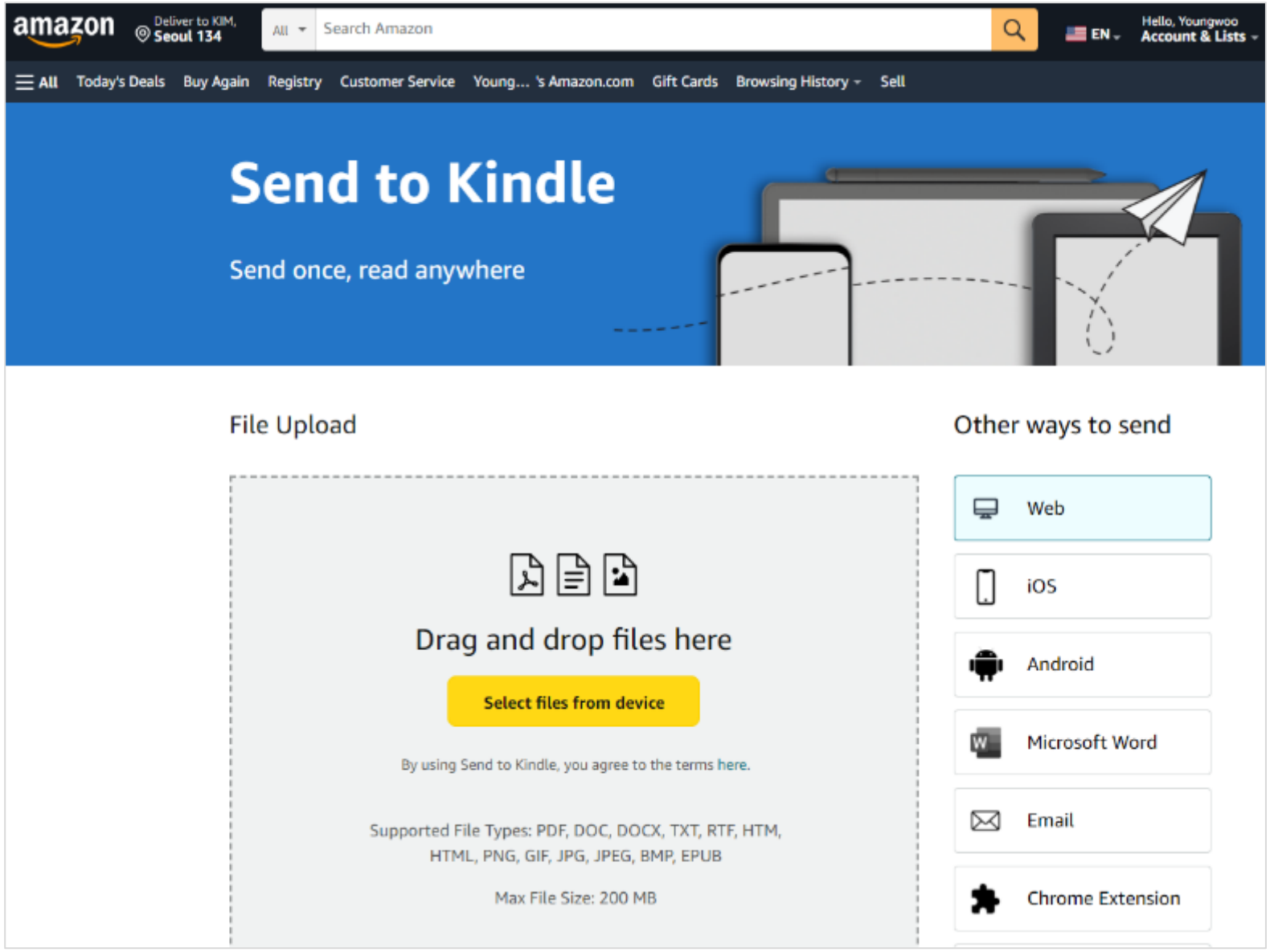
https://www.amazon.com/gp/help/customer/display.html?nodeId=G5WYD9SAF7PGXRNA



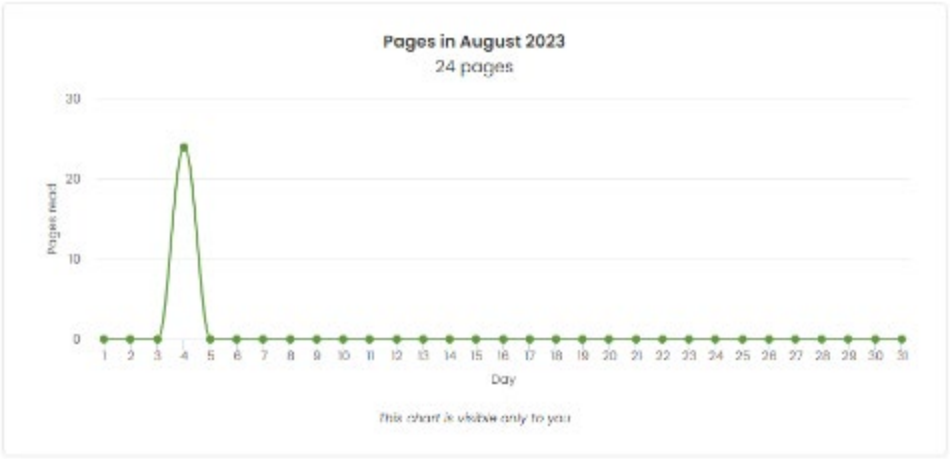
Amazon Kindle

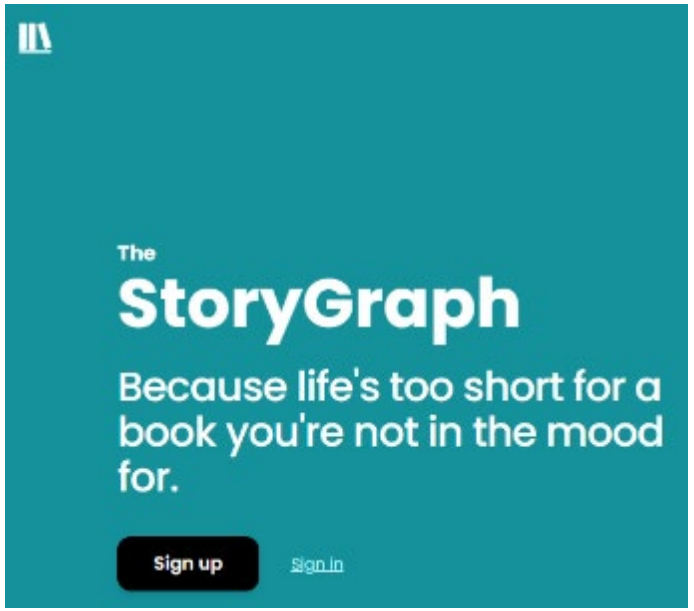


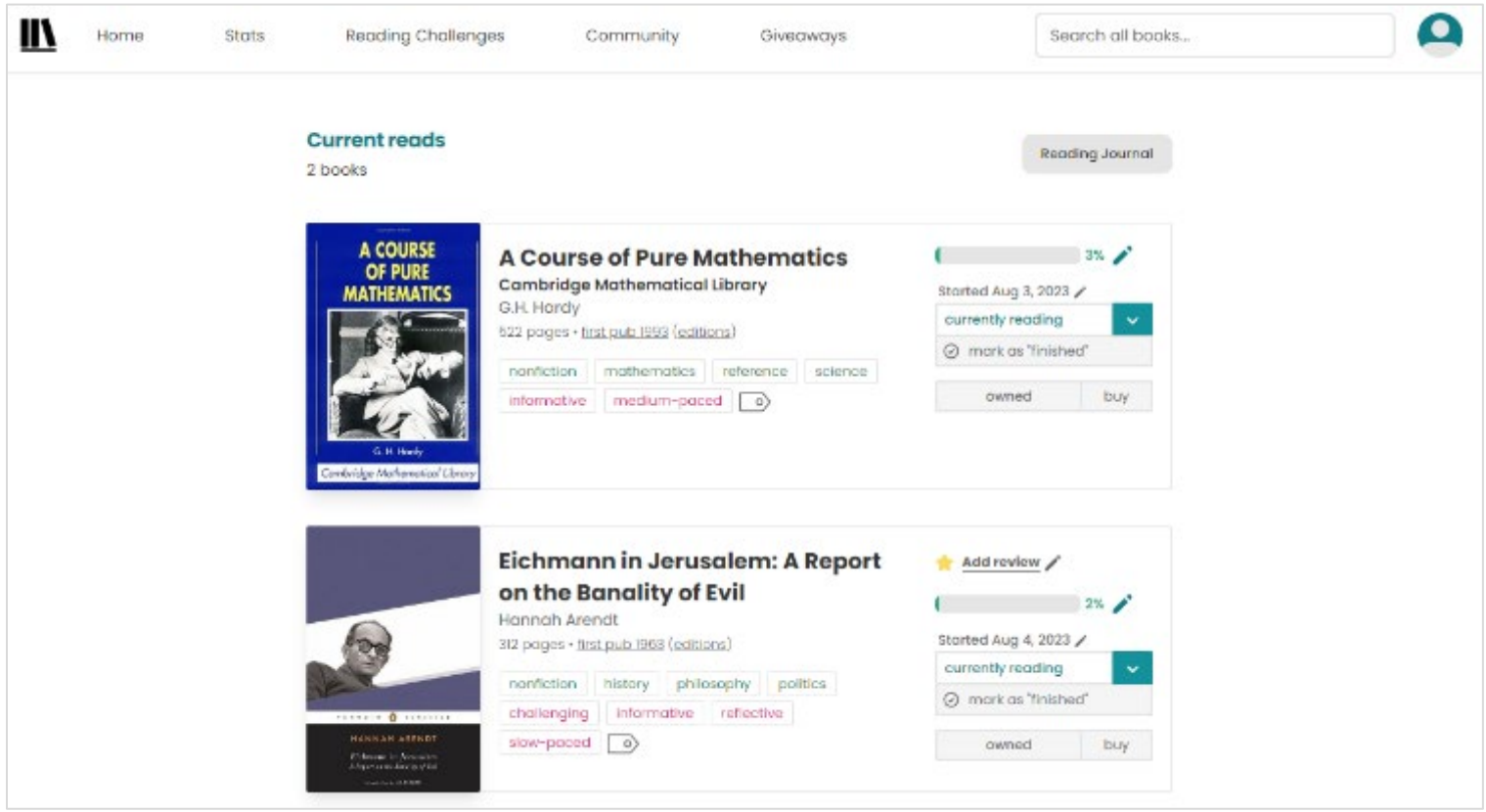
Amazon Kindle



StoryGraph


Pages in August 2023
24 pages
Pages read
Day
This chart is visible only to you


The
StoryGraph
Because life's too short for a book you're not in the mood for.
Sign up
Sign in


Home
Stats
Reading Challenges
Community
Giveaways
Search all books...
Current reads
2 books
Reading Journal
A Course of Pure Mathematics
Cambridge Mathematical Library
G.H. Hardy
522 pages • first pub 1959 (editions)
nonfiction
mathematics
reference
science
informative
medium-paced
3%
Started Aug 3, 2023
currently reading
mark as "finished"
owned
buy
Eichmann in Jerusalem: A Report on the Banality of Evil
Hannah Arendt
312 pages • first pub 1963 (editions)
nonfiction
history
philosophy
politics
challenging
informative
reflective
slow-paced
Add review
2%
Started Aug 4, 2023
currently reading
mark as "finished"
owned
buy

DEMO

ER and Language Data Activities

2. Utilize the book's text data

   1) Obtain an Excel file with the book's text data.

      Sample: JaneEyre_Results_Text.xlsx

   2) Handle the Excel file in a computer.

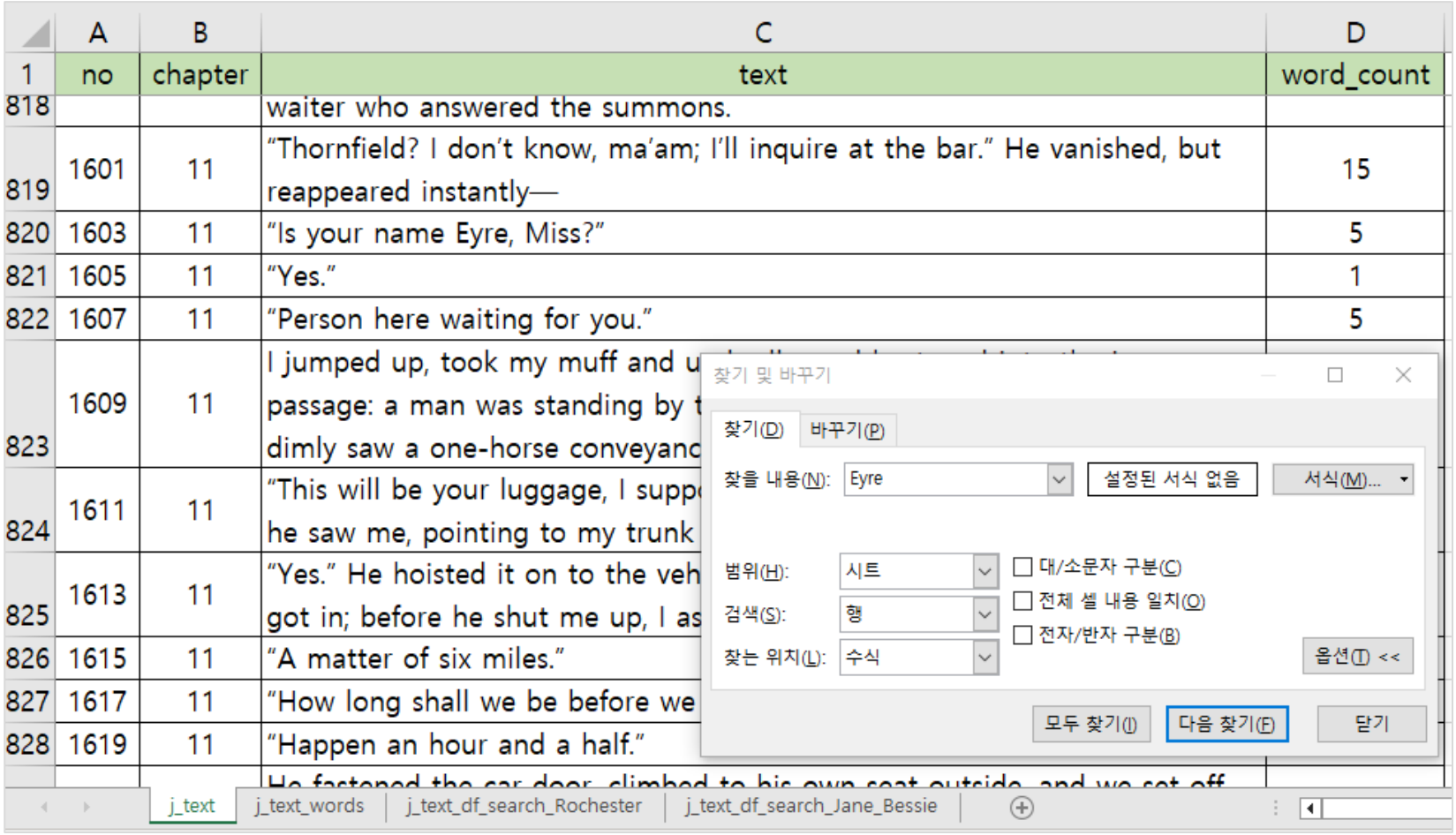
      Or use Google Sheets.

   3) Utilize Excel's features:

      Searching, sorting, tabulation, etc.

	A	B	C	D
1	no	chapter	text	word_count
2	1	0	JANE EYRE	2
3	2	0	AN AUTOBIOGRAPHY	2
4	3	0	by Charlotte Brontë	3
5	4	0	ILLUSTRATED BY F. H. TOWNSEND	5
6	6	0	London	1
7	7	0	SERVICE & PATON	3
8	8	0	5 HENRIETTA STREET	3
9	9	0	1897	1
10	11	0	The Illustrations	2
11	12	0	in this Volume are the copyright of	7
12	13	0	SERVICE & PATON, London	4
13	15	0	TO	1
14	16	0	W. M. THACKERAY, ESQ.,	4
15	18	0	This Work	2
16	19	0	IS RESPECTFULLY INSCRIBED	3
17	21	0	BY	1
18	22	0	THE AUTHOR	2
19	24	0	PREFACE	1
20	25	0	A preface to the first edition of “Jane Eyre” being unnecessary, I gave none: this second edition demands a few words both of acknowledgment and miscellaneous remark.	27

Excel Search



	A	B	C	D
1	no	chapter	text	word_count
818			waiter who answered the summons.	
819	1601	11	“Thornfield? I don’t know, ma’am; I’ll inquire at the bar.” He vanished, but reappeared instantly—	15
820	1603	11	“Is your name Eyre, Miss?”	5
821	1605	11	“Yes.”	1
822	1607	11	“Person here waiting for you.”	5
823	1609	11	I jumped up, took my muff and u… passage: a man was standing by t… dimly saw a one-horse conveyanc…	
824	1611	11	“This will be your luggage, I suppo… he saw me, pointing to my trunk	
825	1613	11	“Yes.” He hoisted it on to the veh… got in; before he shut me up, I as…	
826	1615	11	“A matter of six miles.”	
827	1617	11	“How long shall we be before we	
828	1619	11	“Happen an hour and a half.”	
			He fastened the car door, climbed to his own seat outside, and we set off	

찾기 및 바꾸기

찾기(D) 바꾸기(P)

찾을 내용(N): Eyre — 설정된 서식 없음 — 서식(M)...

범위(H): 시트 — 대/소문자 구분(C)

검색(S): 행 — 전체 셀 내용 일치(O)

찾는 위치(L): 수식 — 전자/반자 구분(B)

옵션(T) <<

모두 찾기(I) 다음 찾기(F) 닫기

j_text | j_text_words | j_text_df_search_Rochester | j_text_df_search_Jane_Bessie

Excel

Sort

	A	B	C
1	word	n	
2	the	7856	
3	i	7169	
4	and	6632	
5	to	5238	
6	a	4470	
7	of	4368	
8	you	2970	
9	in	2769	
10	was	2526	
11	it	2406	
12	my	2236	
13	me	2056	
14	he	1901	
15	her	1716	
16	that	1659	
17	as	1593	

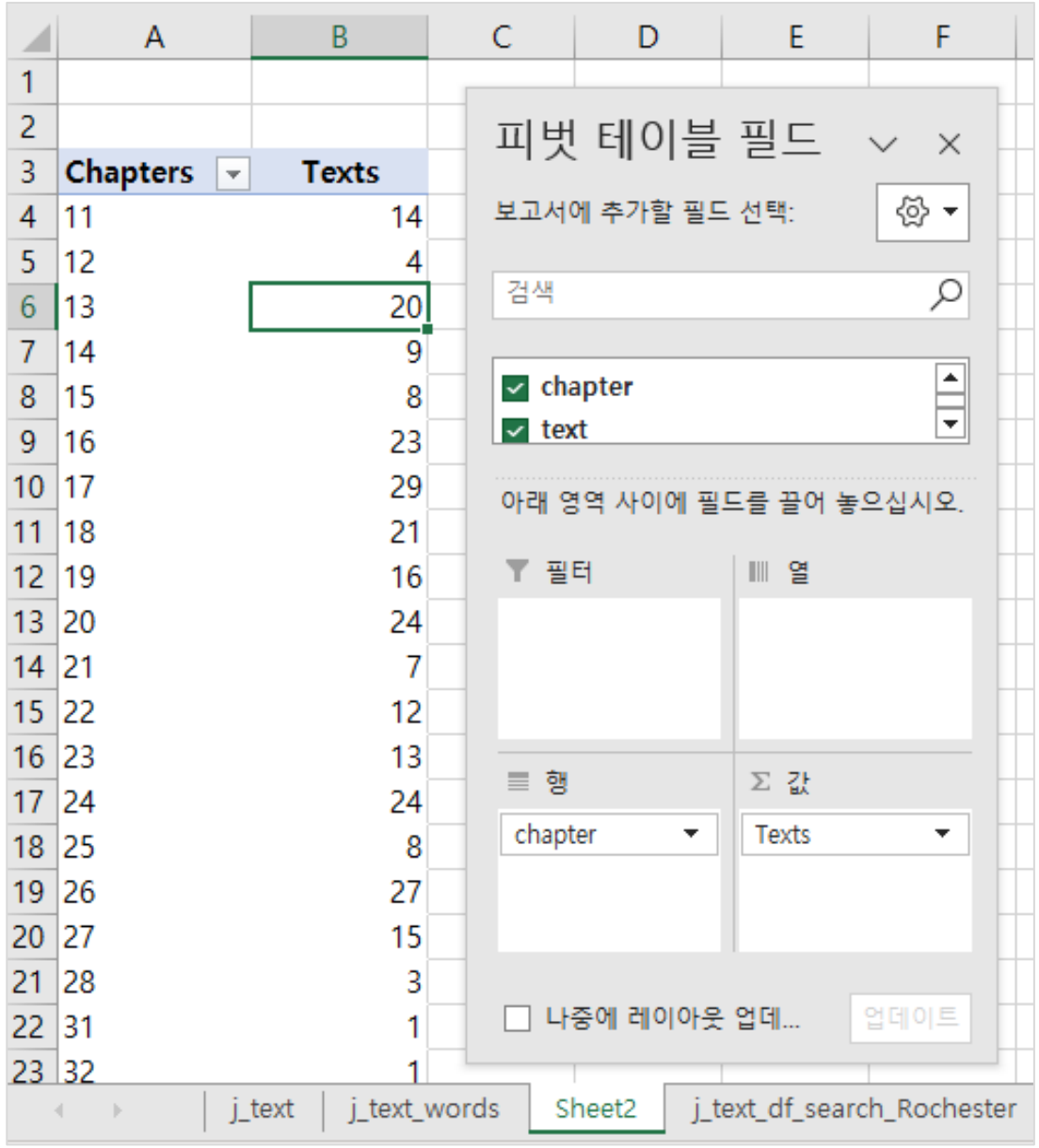
j_text | j_text_words | j_text_df_search_Roch

	A	B	C
1	word	n	length
2	characteristically	1	18
3	conventionalities	2	=LEN(A3)
4	conscientiousness	1	17
5	disadvantageously	1	17
6	inconsiderateness	1	17
7	misunderstandings	1	17
8	incomprehensible	3	16
9	disproportionate	2	16
10	responsibilities	2	16
11	indiscriminately	1	16
12	misunderstanding	1	16
13	personifications	1	16
14	superciliousness	1	16
15	superintendent's	1	16
16	unimpressionable	1	16
17	unsubstantiality	1	16

j_text | j_text_words | j_text_df_search_Roch

Excel

Making Tables



Chapters	Texts
11	14
12	4
13	20
14	9
15	8
16	23
17	29
18	21
19	16
20	24
21	7
22	12
23	13
24	24
25	8
26	27
27	15
28	3
31	1
32	1

Excel

Copy Text

	A	B	C
1	no	chapter	text
6			the impression made on me by her nursery tales. She was pretty too, if my recollections of her face and person are correct. I remember her as a slim young woman, with black hair, dark eyes, very nice features, and good, clear complexion; but she had a capricious and hasty temper, and indifferent ideas of principle or justice: still, such as she was, I preferred her to any one else at Gateshead Hall.
	3982	21	"Missis had been out of health herself for some time: she had got very stout, but was not strong with it; and the loss of money and fear of poverty were quite breaking her down. The information about Mr. John's death and the manner of it came too suddenly: it brought on a stroke. She was three days without speaking; but last Tuesday she seemed rather better: she appeared as if she wanted to say something, and kept making signs to my wife and mumbling. It was only yesterday morning, however, that Bessie understood she was pronouncing your name; and at last she made out the words, 'Bring Jane—fetch Jane Eyre: I want to speak to her.' Bessie is not sure whether she is in her right mind, or means anything by the words; but she told Miss Reed and Miss Georgiana, and advised them to send for you. The young ladies put it off at first; but their mother grew so restless, and said, 'Jane, Jane,' so many times, that at last they consented. I left

j_text | j_text_words | Sheet2 | j_text_df_search_Rochester | j_text_df_search_Jane_Bessie

ER and Language Data Activities

3. Mining the text (to be applied in additional/advanced stages)

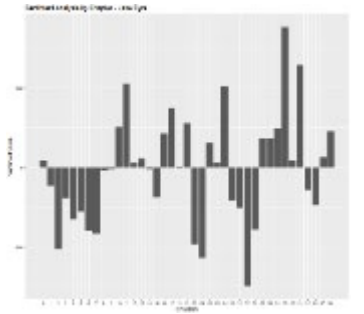
   1) Obtain the text data of the book as a text file. An excel file will do. Sample: JaneEyre.txt

   2) Prepare a coding environment.
      For R, use Posit Cloud (https://posit.cloud/),
      For Python, use Google Colab (https://colab.google/).

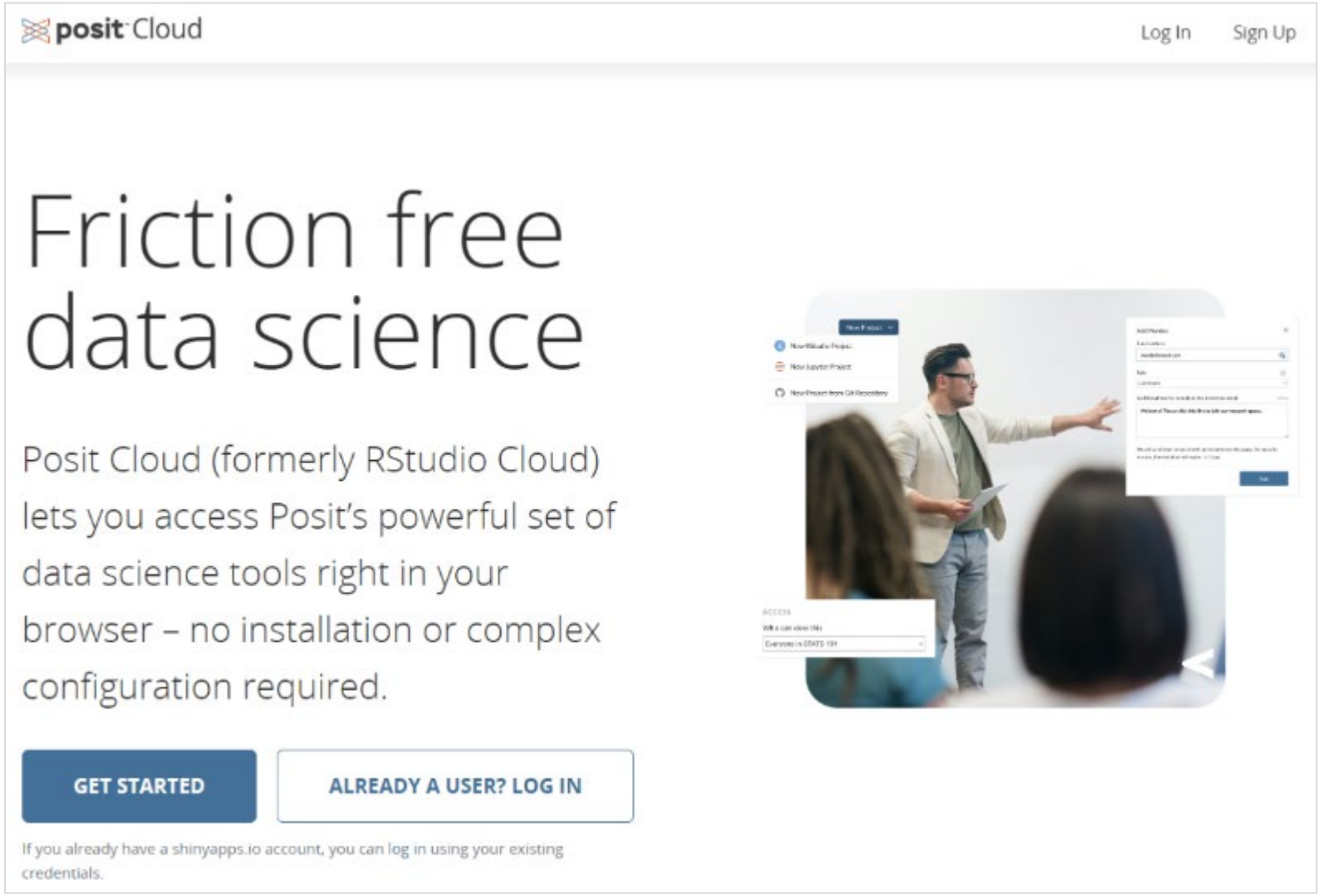
   3) Obtain text mining code written in R or Python.
      Sample: https://posit.cloud/content/6256020

   4) Mine text of the book with codes provided.

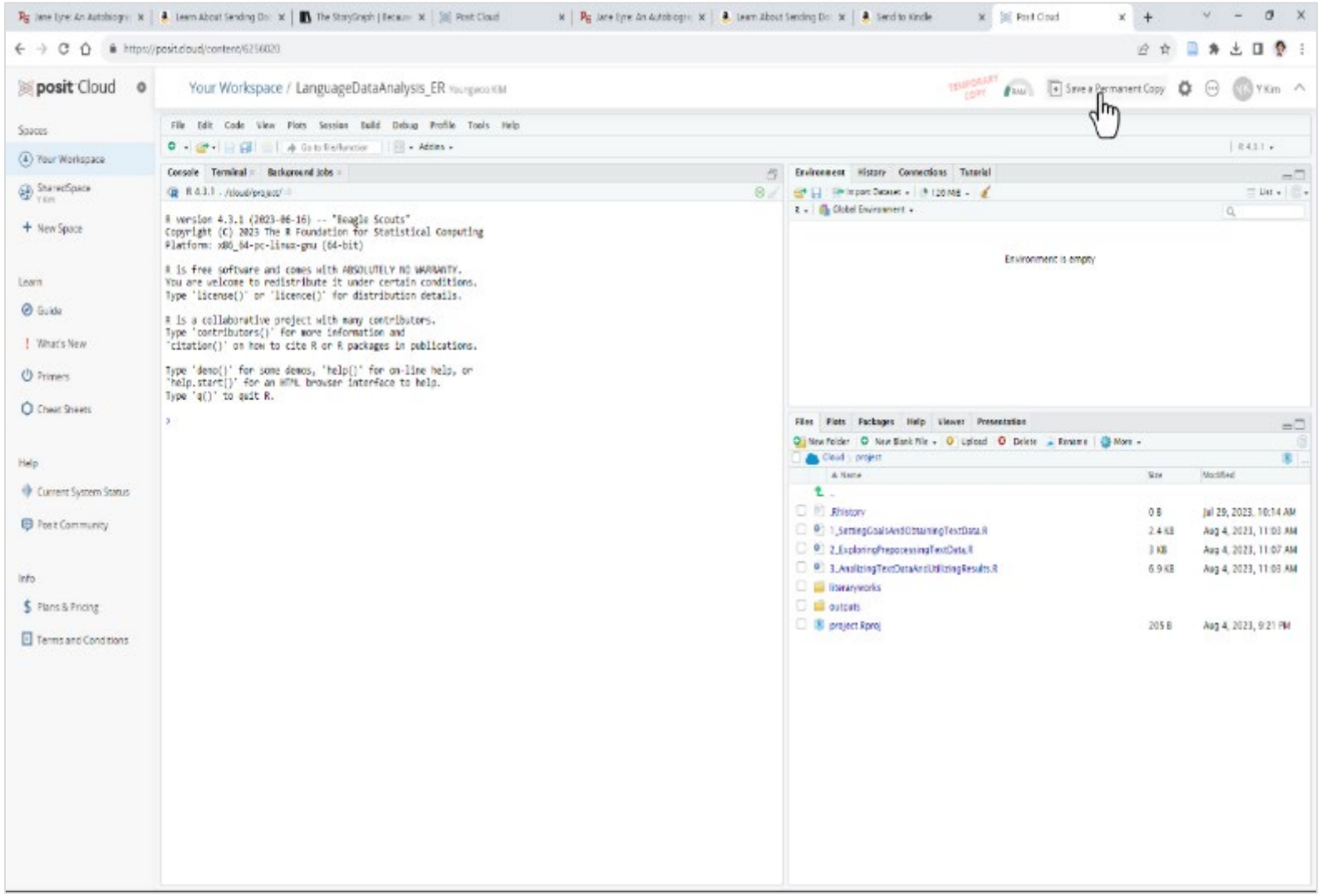
   5) Utilize the text mining results (excel file, visuals, etc.) for your reading activity.
      JaneEyre_Results_Text.xlsx, JaneEyre_Results_Visual.zip



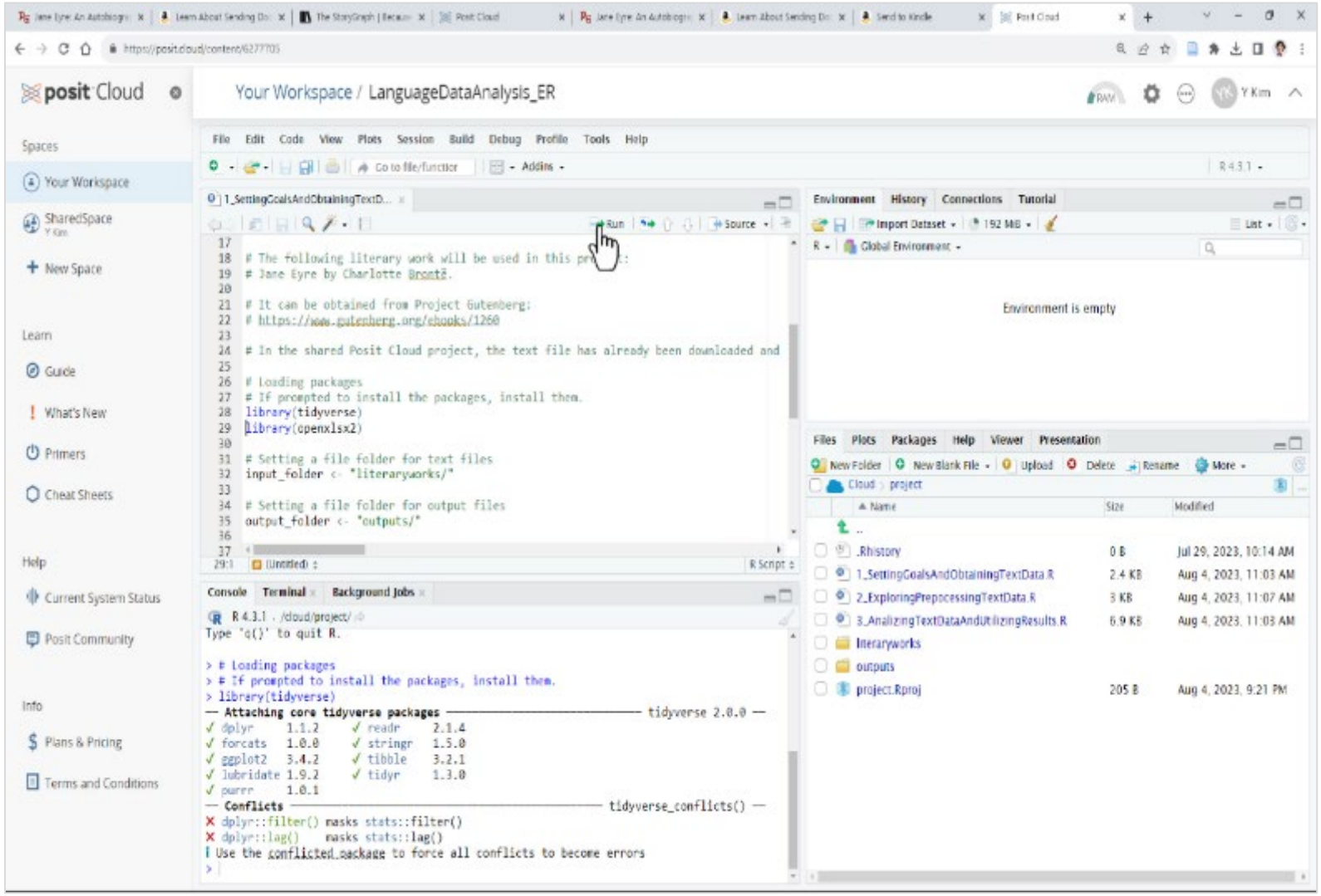
Posit Cloud


posit Cloud
Log In
Sign Up
Friction free data science
Posit Cloud (formerly RStudio Cloud) lets you access Posit's powerful set of data science tools right in your browser – no installation or complex configuration required.
GET STARTED
ALREADY A USER? LOG IN
If you already have a shinyapps.io account, you can log in using your existing credentials.

Posit Cloud Project

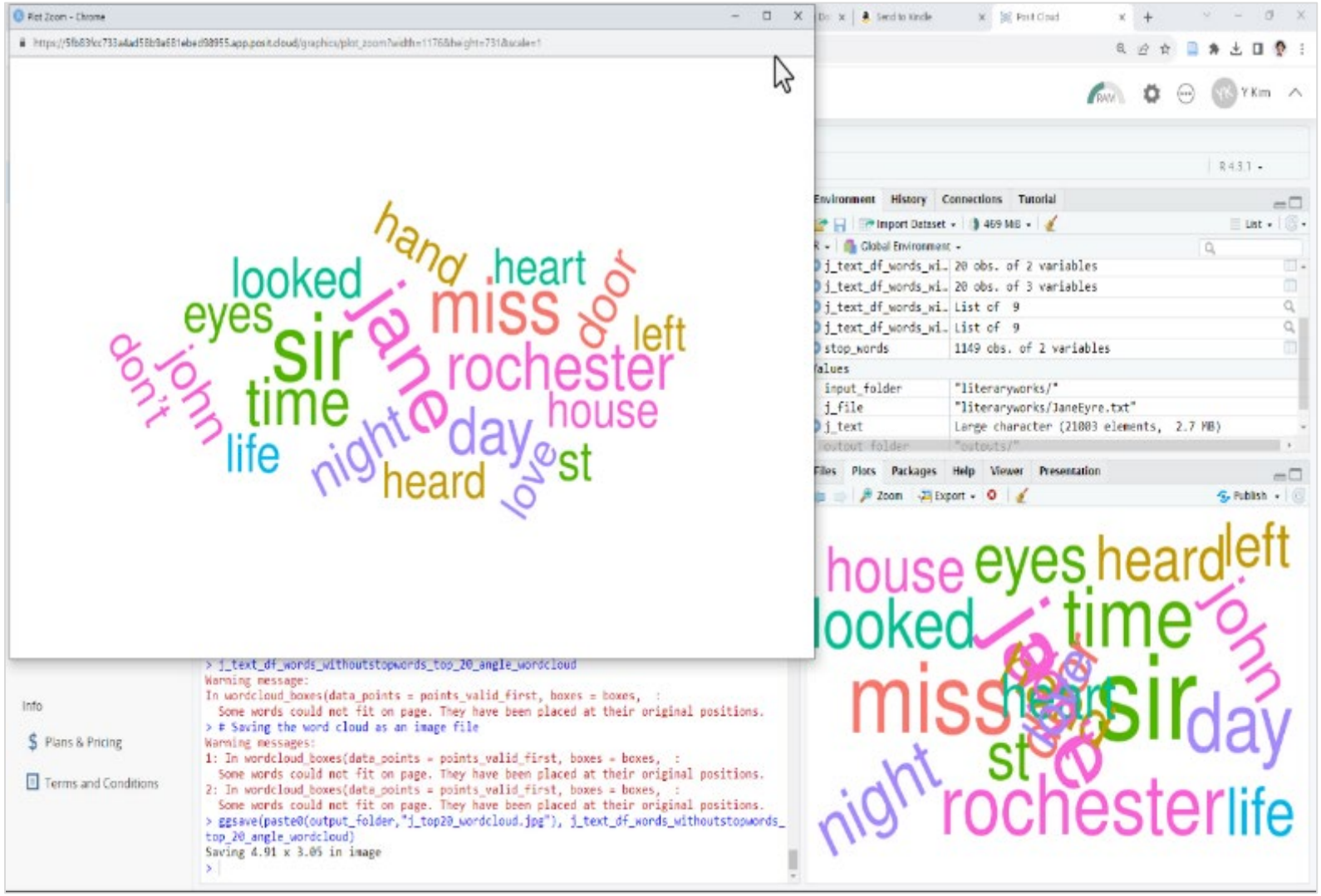


Posit Cloud Running R Script



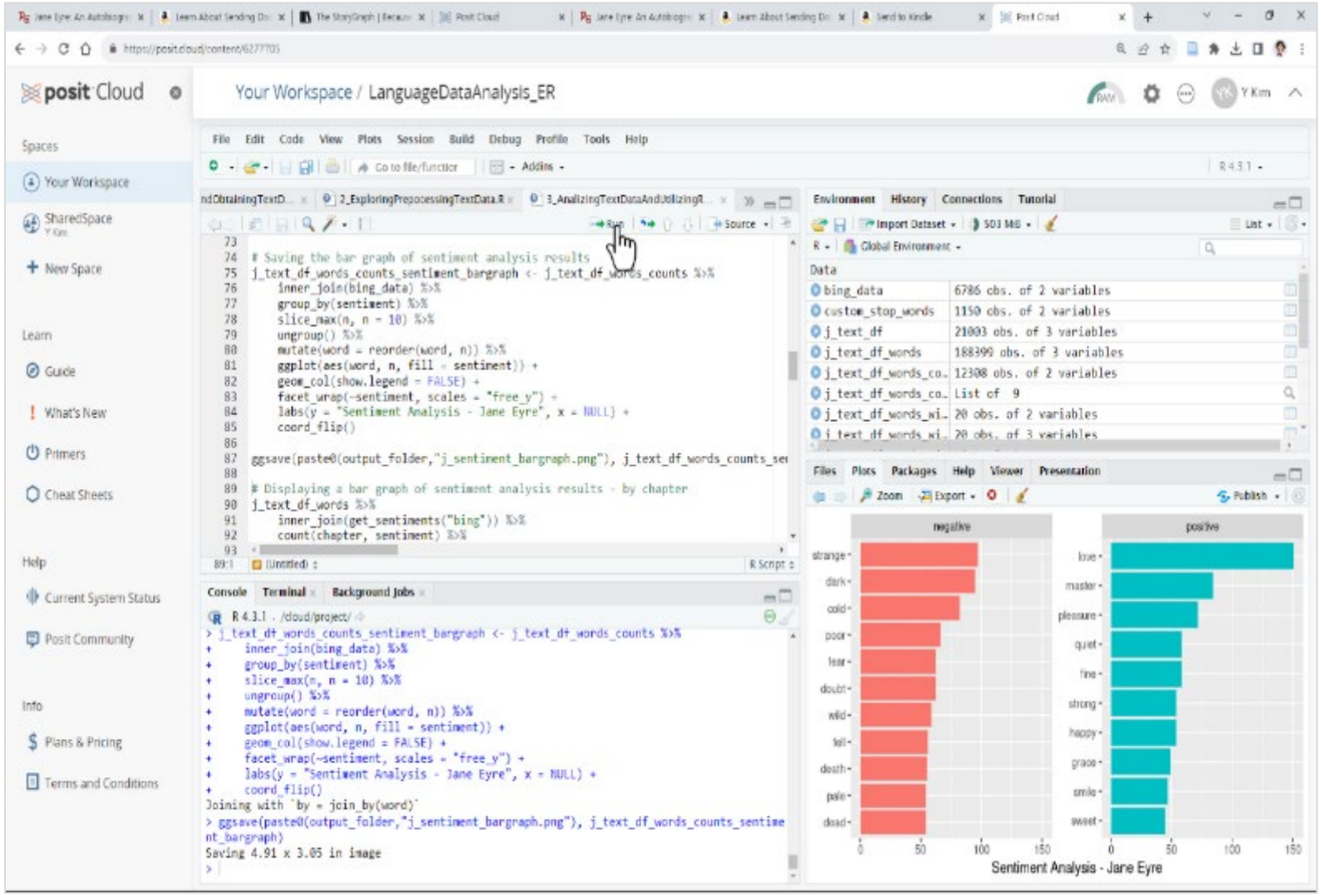
Posit Cloud

Word Cloud

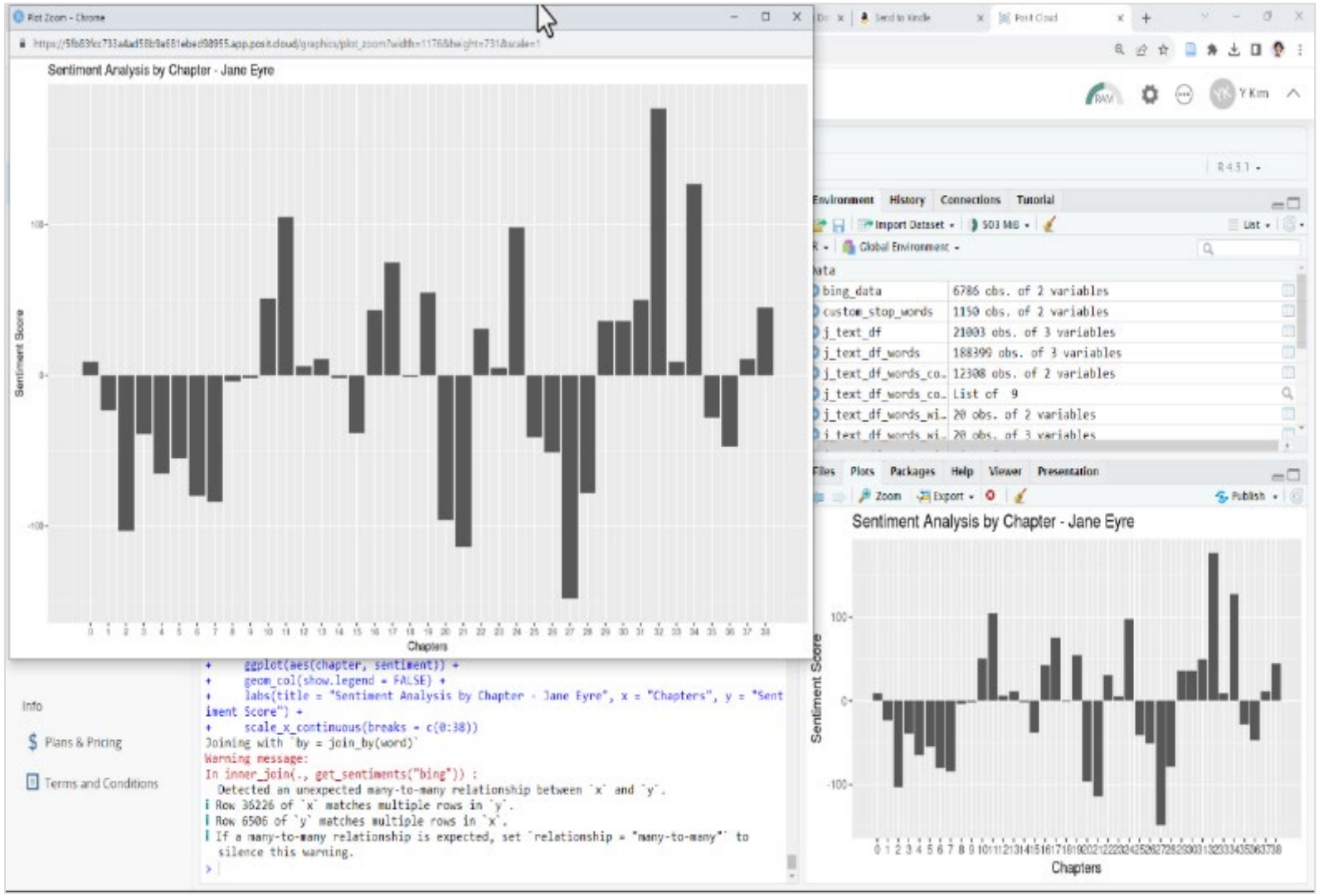


Posit Cloud

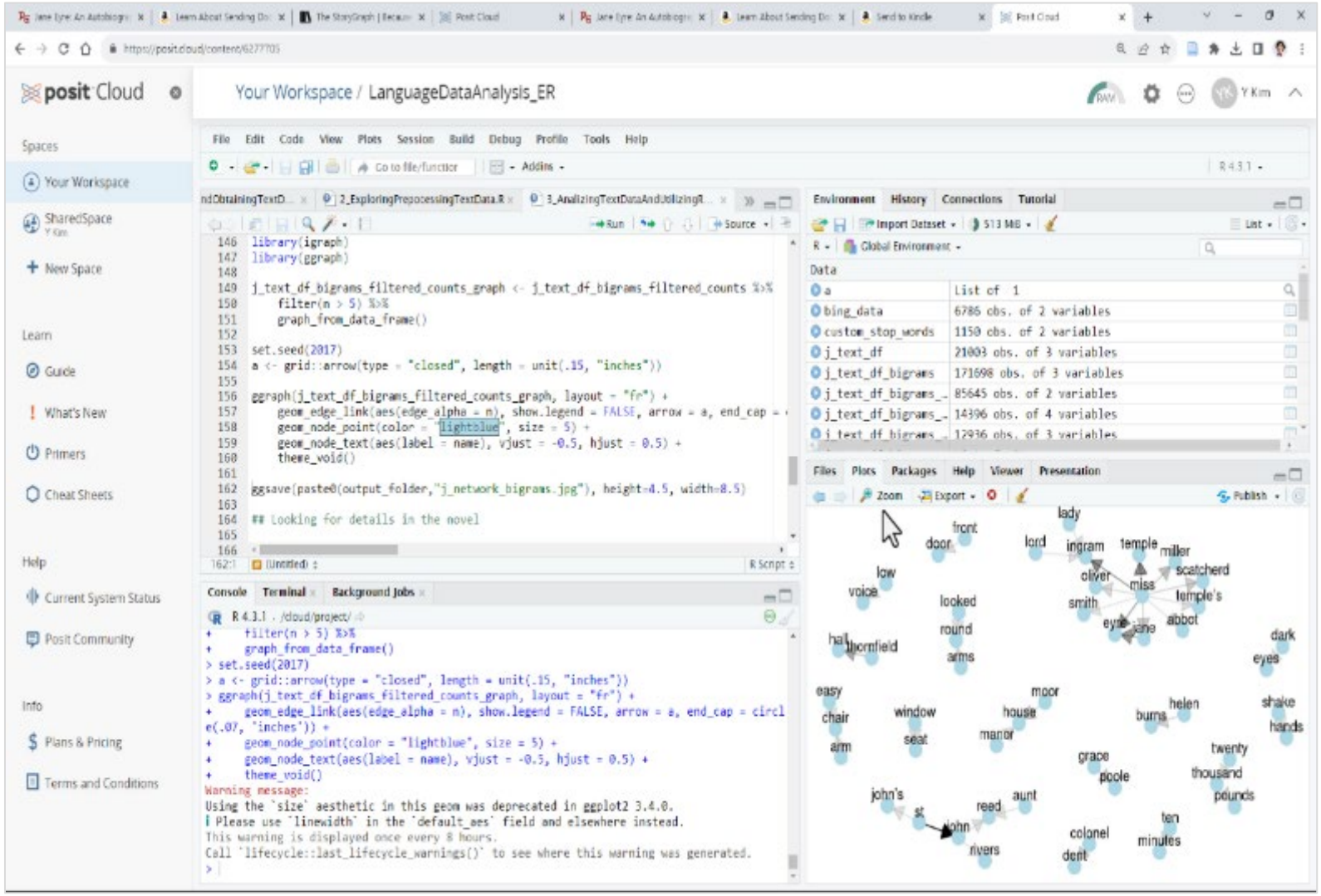
Bar Chart



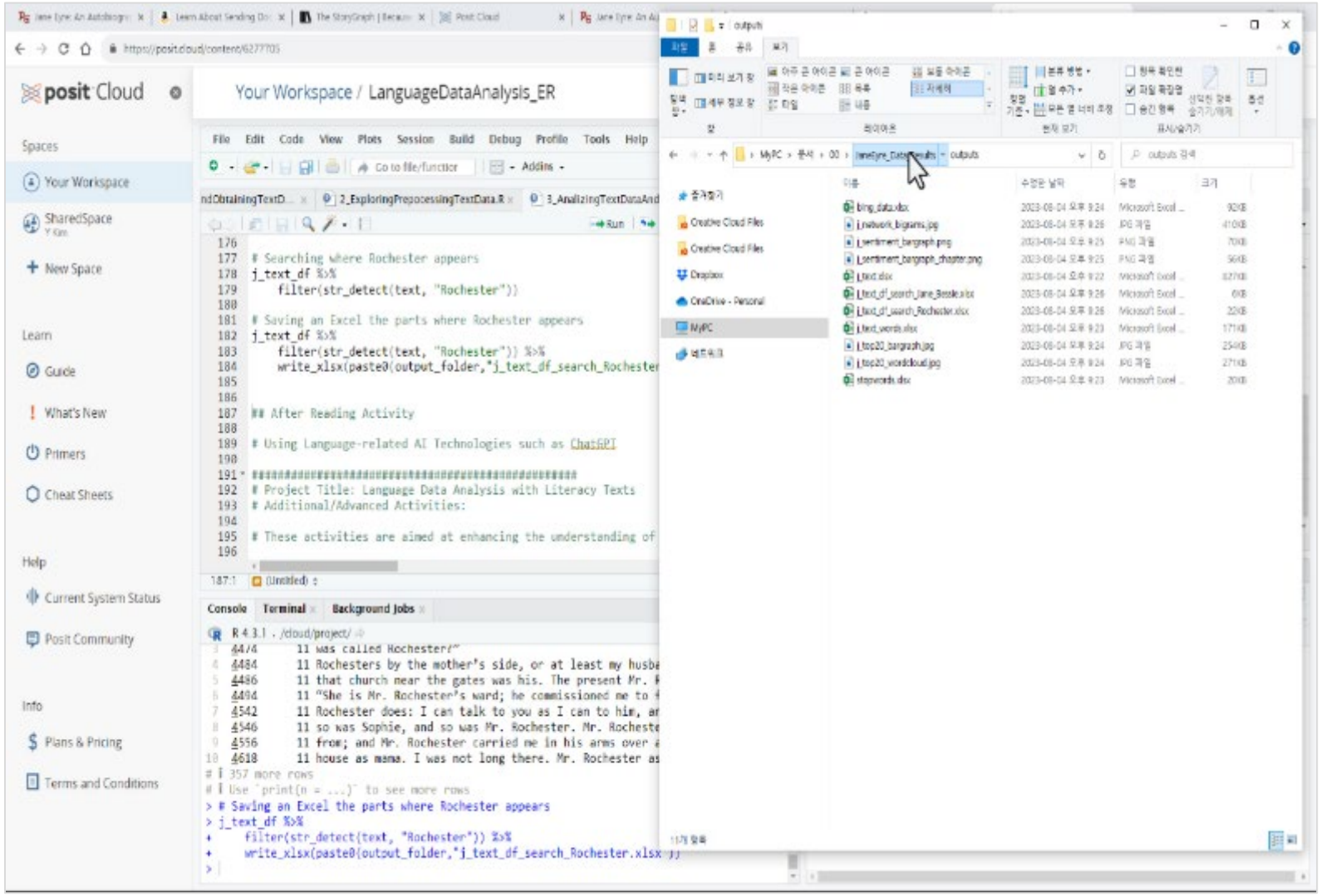
Posit Cloud Sentiment Analysis



Posit Cloud Word Network



Posit Cloud Outcomes



Combining ER and DR(Distant Reading) Using Literary Texts with Computer Coding

Youngwoo KIM
Yongkook WON
Junghye Fran CHOI
Hyunkee AHN
Kwanghyun PARK
Joohae KIM

Contact: youngwoolearning@gmail.com

Terima Kasih

Matur Suksma