Some Interesting Analysis

Pride and Prejudice – Volume I

Sanaz Khosravi

CodePath

Jane Austen

PRIDE & PREJUDICE

DOVER THRIFT EDITIONS

Table of Contents

Github Info

My project code can be found here!

Total words in novel

The first volume of *Pride and Prejudice* has **40997**words!

Total unique words in novel

Out of the 40997 words, **4307**are unique words!

Most frequent words

The 20 most frequent words in this volume are as follow:

Word	Frequency
the	**1494**
to	**1402**
of	**1211**
and	**1182**
a	**719**
i	**703**
her	**681**
in	**614**
was	**547**
that	**517**
not	**501**
he	**483**
she	**480**
it	**478**
You	**473**
His	**449**
With	**403**
Be	**392**
mr	**392**
as	**373**

Most interesting frequent words

Probably you are not interested in knowing how many times "the" is appeared in the text, that's why you can find a list of interesting frequent words as follow:

Word	Frequency
i	**703**

not	**501**
Mr	**392**
very	**201**
elizabeth	**189**
bingley	**156**
bennet	**150**
darcy	**146**
miss	**143**
me	**143**
such	**133**
much	**132**
am	**125**
mrs	**124**
any	**121**
must	**94**
should	**93**
jane	**91**
collins	**85**
think	**77**

20 Least Frequent words

Word	Frequency
prepare	1
acutely	1
mince	1
withdrawn	1
discourage	1
hall	1
gloomy	1
unfeeling	1
daring	1
agreeableallowing	1
meanwhile	1
reserves	1
require	1
compassionate	1
applies	1
reserved	1
poorly	1

sonnet	1
pleasantness	1
bennetshe	1

Interesting word usage pattern

Do you want to know which chapters some of the most and least frequent words appeared along with the number of times they appeared? Here is some of the results:

Volume I has 23 chapters:

* * The first row contains the number of chapter.

Elizabeth

Daring

Bennet

1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23
0	1	4	3	1	11	13	13	8	15	12	4	2	1	5	18	8	23	10	9	8	10	10
0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0
6	8	12	3	3	2	10	5	11	5	4	4	8	10	8	3	2	10	3	14	2	5	12

Favorite quote

I can search for different quotes in my text and it will return the chapter number in case in can find it in my text. Otherwise, it will return -1.

"I am more obliged to you than I can express."
Chapter 22

Generated sentence

Here is a few of the randomly generated sentences:

The supposition did before." She would have more intolerable. The dishes at all--and I was determined to the whole evening

The three miles. I have you were grieved, how he could have no more deceitful," said Miss Elizabeth, "that this

The hint of trade. Mr. Collins's letter of her daughter's favour, without seeming to them all, is impossible. My dear.