

P2 results

Prof Bill - Feb 2020

Here are my results for P2. This is a snapshot as of **Mon Feb 24**. I may still tweak my P2, causing these results to change a bit.

Important: Your mileage may vary. Don't worry about matching my results exactly. Look for two things: 1) you should be close, and 2) you should have the same Freq results over all your data structures.

thanks...yow, bill

Word counts

Here are my word count stats. I get the same results for each data structure. (pew)

	ex1 sentence	ex2 lincoln	ex3 wizard of oz	ex4 war and peace
Total words	5	283	39,603	562,601
Diff words	4	144	3,140	21,429
"war" count	0	2	0	289
"peace" count	0	0	2	109

Top 20

I print the Top 20 words by count for each file.

File: **ex1_sentence.txt**

```
the=2, in=1, paris=1, spring=1
```

File: **ex2_lincoln.txt**

```
that=13, the=11, we=10, here=8, to=8, ---=7, a=7, and=6, can=5, for=5,  
have=5, it=5, nation=5, not=5, of=5, dedicated=4, in=4, this=4, are=3,  
dead=3, great=3
```

File: ex3_wizard_oz.txt

the=3001, and=1665, to=1116, of=847, a=804, i=593, was=504, you=487,
in=481, he=451, it=417, her=407, they=401, she=398, that=388,
dorothy=342, as=329, said=328, for=327, so=307, but=301

File: ex4_war_and_peace.txt

the=34381, and=22074, to=16624, of=14873, a=10435, he=9807, in=8870,
his=7966, that=7788, was=7326, with=5646, had=5354, it=5143, her=4685,
not=4636, him=4536, at=4519, i=4080, but=4005, on=3978, as=3973

Performance

How fast does each data structure run on our largest examples?

File: ex3_wizard_oz.txt (39K words)

- ArrayList = 0.27 seconds
- Binary search tree = 0.07 seconds
- HashMap = 0.07 seconds

File: ex4_war_and_peace.txt (562K words)

- ArrayList = 5.84 seconds
- Binary search tree = 0.45 seconds
- HashMap = 0.44 seconds