Project Gitenberg

-Greeshma Avadhootha

In this project, I have done some interesting analysis on some of the most famous literature works while honing my skills working with strings, arrays, hash tables, trees, and more!

I have analyzed the novel **Jane Austen** for the project. I have implemented the following methods on the text of the novel.

```
1. getTotalNumberOfWords()
2. getTotalUniqueWords()
3. get20MostFrequentWords()
4. get20MostInterestingFrequentWords()
5. get20LeastFrequentWords()
6. getFrequencyOfWord()
7. getChapterQuoteAppears()
8. generateSentence()
```

The following is my output :

Total num of words : 121541

Total num of unique words : 12883

Top 20 freq words :

1. to 4067

2. the 4047

3. of 3581

4. and 3341

5. her 2106

6. a 1888

7. was 1832

8. in 1772

9. I 1750

10. that 1474

11. not 1386

12. she 1376

13. be 1229

14. it 1218

15. his 1185

16. had 1148

17. as 1113

18. he 1092

19. you 1083

20. with 1016

Top 20 Least Frequent words :

1. nicely 1

2. costing 1

3. abrupt 1

4. pollution 1

5. superlatively 1

6. spreading 1

7. coincide 1

8. here's 1

9. avoidance 1

10. Promise 1

11. newspaper 1

12. B 1

13. hang 1

14. fearless 1

15. P 1

16. it—unless 1

17. barbarously 1

18. resume 1

19. no—it 1

20. poverty—comparative 1

Top 20 Interesting words :

1. mr 766

2. elizabeth 587

3. darcy 349

4. mrs 338

5. "i 297

6. bennet 278

7. jane 253

8. bingley 242

9. herself 221

10. being 175

11. without 168

12. wickham 155

13. collins 149

14. into 141

15. however 133

16. lydia 121

17. himself 117

18. away 112

19. cannot 111

20. catherine 109

Interesting word usage pattern:

The word Bennet is interesting since it was used frequently at the beginning and less in the middle. However, the usage increases in the last few chapters.

Number of usages in each chapter:

6 8 11 3 4 2 10 5 11 5 4 4 9 10 8 3 2 10 3 15 2 5 12 5 5 1 0 0 2 1 3
1 0 0 1 0 2 0 3 2 5 2 0 1 2 0 7 11 4 15 4 2 12 6 11 18 1 2 10 1 0

Quote :

"Happiness in marriage is entirely a matter of chance."

Chapter number : 6

Generated sentence :

the evening of that there is such people do me if not but in your character we know when at