



Chetan Mali &lt;2016ccchetan4885@poornima.edu.in&gt;

---

**Mini Project - Vocabulary**

1 message

**Forsk Labs** <forsklabs@gmail.com>

Wed, Dec 12, 2018 at 11:30 AM

To: Avi Agrawal <aviagrawal512@gmail.com>, 2016ccchetan4885@poornima.edu.in, Krishna Parsarampuria <krishna.agarwal961@gmail.com>, Rohit Bhardwaj <rohit.bhardwaj9792@gmail.com>, Sampat Aheer <sampat489@gmail.com>, surendrakumaryadav3085@gmail.com  
Cc: Devendra singh shekhawat <shekhawat.dev313@gmail.com>, Kunal Vaid <vaidkunal09@gmail.com>

Novel = james\_joyce\_ulysses.txt

1. Find the total number of words in the novel
2. Quality of a novel is the number of different words.  
Find the number of different words used
3. look at the other novels and find the total words and unique words for comparison

novels = ['sons\_and\_lovers\_lawrence.txt',  
'metamorphosis\_kafka.txt',  
'the\_way\_of\_all\_flash\_butler.txt',  
'robinson\_crusoe\_defoe.txt',  
'to\_the\_lighthouse\_woolf.txt',  
'james\_joyce\_ulysses.txt',  
'moby\_dick\_melville.txt']








4. Special Words in Ulysses novel by comparing with others, words which are only used in Ulysses, store it in a file

5. Common Words - Find the words which occur in every book.

Hint: Use Sets, Regex, File Handling

---

**7 attachments**

-  **metamorphosis\_kafka.txt**  
139K
-  **to\_the\_lighthouse\_woolf.txt**  
978K
-  **the\_way\_of\_all\_flash\_butler.txt**  
895K
-  **moby\_dick\_melville.txt**  
1241K
-  **sons\_and\_lovers\_lawrence.txt**  
923K
-  **james\_joyce\_ulysses.txt**  
1544K
-  **robinson\_crusoe\_defoe.txt**  
641K