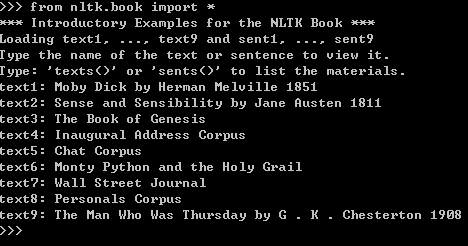
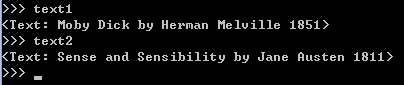
装好nltk，下载好book以后

打开Python cmd，输入 from nltk.book import \*



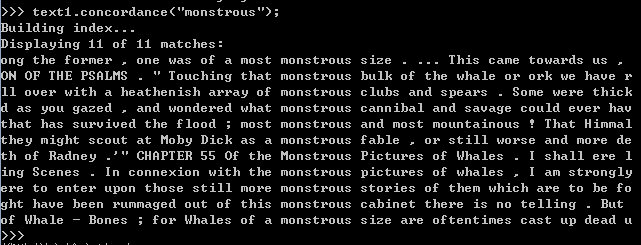
text1-9 对应不同的语料



搜索文本：

1. concordance函数用来查看制定的文字出现的位置和上下文

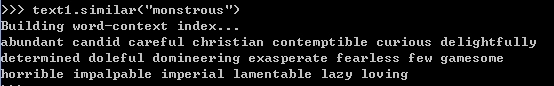
text1.concordance(“monstrous”)



由此我们可以看出monstrous是如何和其他单词搭配的

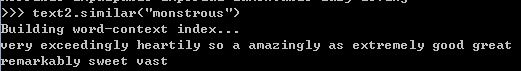
2. similar函数可以用来查看那些单词和monstrous出现在相似的语句中

text1.similar(“monstrous”)



这是text1中monstrous的相似单词

在text2中，monstrous的用法有很大不同



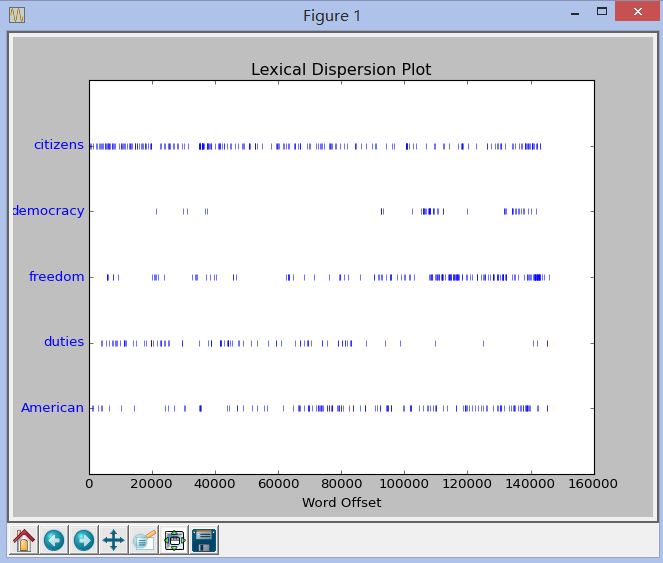
text2中，monstrous是用作一个表示非常，极其之意的副词

3. common\_contexts函数可以用来查看两个及以上单词都会出现的上下文

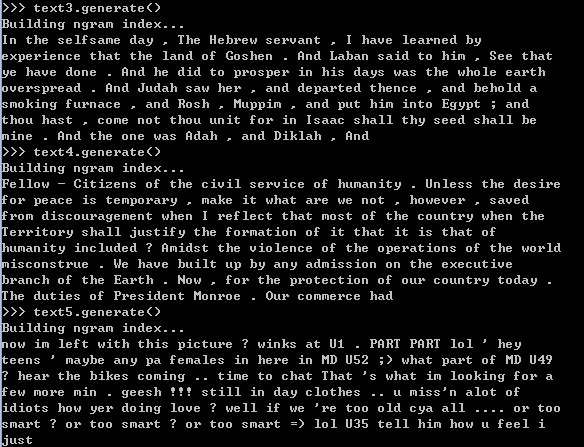


可见，monstrous lucky和very lucky在text2中是通用的

4. dispersion\_plot()函数可以用来看一个单词在文中出现的位置（需要安装matplotlib，有可能需要numpy）  

5. generate()函数会产生一些类似于所选语料的随机文本。 虽然生成的文本完全没有逻辑，但是能看出语料的风格



6. 计数词汇

len（）函数获取文本长度



set() 看文本中有多少不同的标志符（一组字符的序列）出现多次只算一次

set（text1）不好意思。。。刷屏了。。。



所以看出有很多单词是重复的。

sorted（）函数可以用来排序a-z

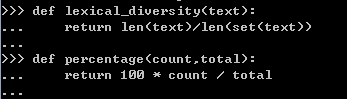
7. 看看每个单词平均用了几次

注意： from后面是两个下划线。。。

8. 统计单词的出现次数



9. 定义一个函数



用def 定义，注意第二行要打一个tab，不能直接输





好了，这一节结束，去吃饭了