```
import nltk
nltk.download('brown')
from nltk.corpus import brown

print ('File ids of brown corpus\n',brown.fileids())
'''Let's pick out the first of these texts — Emma by Jane Austen — and give it a short
name, emma, then find out how many words it contains:'''
ca01 = brown.words('ca01')
# display first few words
print('\nca01 has following words:\n',ca01)
# total number of words in ca01
print('\nca01 has',len(ca01),'words')
#categories or files
print ('\n\nCategories or file in brown corpus:\n')
print (brown.categories())
'''display other information about each text, by looping over all the values of fileid
corresponding to the brown file identifiers listed earlier and then computing statistics for
each text.'''
print ('\n\nStatistics for each text:\n')
print('AvgWordLen\tAvgSentenceLen\tno.ofTimesEachWordAppearsOnAvg\t\tFileName')
for fileid in brown.fileids():
    # Indent the lines within the for loop
    num_chars = len(brown.raw(fileid))
    num_words = len(brown.words(fileid))
    num_sents = len(brown.sents(fileid))
    num_vocab = len(set([w.lower() for w in brown.words(fileid)]))
    print (int(num_chars/num_words),'\t\t\t', int(num_words/num_sents),'\t\t\t',
           int(num_words/num_vocab),'\t\t\t', fileid) # Fix indentation and remove HTML entities
```