

# SPRINT 9 T01: ANÀLISIS DE SENTIMENT I TEXTOS

Descripció

Exercicis d'anàlisi de sentiment i textos.

## EXERCICI 1

Agafa un text en anglès que vulguis, i calcula'n la freqüència de les paraules.

```
In [1]: # Després d'instal·lar el paquet necessari, el cridem juntament amb altres llibreries (no

import pandas as pd
import numpy as np
import sklearn
import matplotlib.pyplot as plt
import seaborn as sns
import warnings
import nltk
```

```
In [2]: #Importo un llibre en anglès d'una pàgina que permet llegir-los en format txt

from urllib import request
text = "https://www.gutenberg.org/files/68894/68894-0.txt"
response = request.urlopen(text)
rawText = response.read().decode('utf8')
print(type(rawText))
print(len(rawText))
print(rawText[:75])
```

```
<class 'str'>
480429
The Project Gutenberg eBook of Twenty tales by twenty women, by
Anonymous
```

El tipus ens diu que l'arxiu és de tipus str, la longitud és de 480.429 caràcters i imprimim els primers 75. A partir d'ara treballarem amb l'arxiu rawText ja que és el més "nu" possible i que ens permetrà treure'n informació

```
In [3]: # Tranquem el text en frases utilitzant tokenization

from nltk.tokenize import sent_tokenize
tokenFrase = sent_tokenize(rawText)
print(type(tokenFrase))
print(len(tokenFrase))
print(tokenFrase[:5])
```

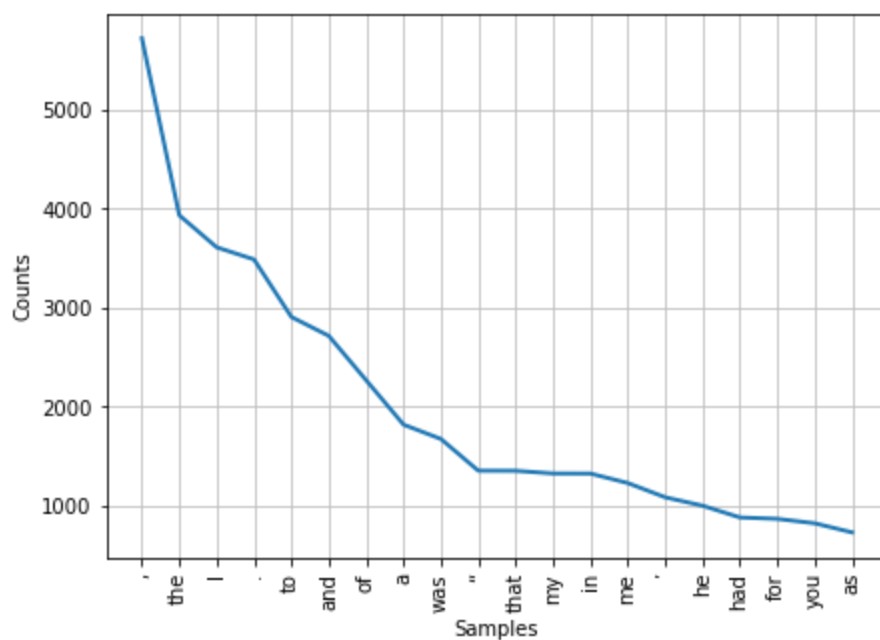
```
<class 'list'>
3693
['\ufe0ffThe Project Gutenberg eBook of Twenty tales by twenty women, by\r\nAnonymous\r\n\r\n\r\nThis eBook is for the use of anyone anywhere in the United States and\r\nmost other part
s of the world at no cost and with almost no restrictions\r\nwhatsoever.', 'You may copy i
t, give it away or re-use it under the terms\r\nof the Project Gutenberg License included
with this eBook or online at\r\nwww.gutenberg.org.', 'If you are not located in the United
States, you\r\nwill have to check the laws of the country where you are located before\r\n\r\nusing this eBook.', 'Title: Twenty tales by twenty women\r\n\r\nFrom real life in Chica
go\r\n\r\n\r\nAuthor: Anonymous\r\n\r\n\r\nRelease Date: September 1, 2022 [eBook #68894]\r\n\r\n\r\nL
anguage: English\r\n\r\n\r\nProduced by: MFR and the Online Distributed Proofreading Team at\r\n\r\n
https://www.pgdp.net (This file was produced from images\r\n\r\nge
```

[illegible]

```
In [4]: # Tranquem el text en paraules utilitzant tokenization
from nltk.tokenize import word_tokenize
tokenParaules = word_tokenize(rawText)
print(type(tokenParaules))
print(len(tokenParaules))
print(tokenParaules[:20])
```

```
In [6]: # Calculo les que es repeteixen més i quantes vegades
        frequencia.most common(10)
```

```
In [7]: # Faig el gràfic
plt.figure(figsize=(7,5))
frecuencia.plot(20,cumulative=False)
plt.show()
```



## EXERCICI 2

Treu les stopwords i realitza stemming al teu conjunt de dades.

```
In [8]: # Imprimeixo els stopwords, o paraules que no tenen molta importància dins un text i que
from nltk.corpus import stopwords
stop_words=set(stopwords.words("english"))
print(stop_words)
```

```
{'you'll', 'have', 'aren', 'she's', 'had', 'against', 'her', 'how', 'other', 'ain', 'shoul
dn't', 'my', 'between', 'own', 're', 'what', 'before', 'any', 'by', 'their', 'hadn't', 'i
n', 'has', 'i', 'only', 'does', 'out', 'aren't', 'didn', 'mightn't', 'your', 'for', 'has
n', 'didn't', 'him', 've', 'hasn't', 'who', 'can', 'so', 'wasn't', 'why', 'mightn', 'll',
'himself', 'doesn't', 'should've', 'where', 'haven', 'isn', 'once', 'doesn', 'themselves',
'should', 'theirs', 'itself', 'these', 'needn't', 'weren', 'a', 'they', 'am', 'but', 'bein
g', 'do', 'is', 'this', 'over', 'just', 'yourself', 'mustn't', 'ma', 'up', 'further', 'no
t', 'if', 'o', 'been', 'couldn', 'no', 'or', 'which', 'nor', 'are', 'of', 'during', 'was',
'while', 'off', 'we', 'each', 'mustn', 'it's', 'through', 'don't', 'doing', 'all', 'he',
'down', 's', 'from', 'hadn', 'them', 'below', 't', 'with', 'shouldn', 'its', 'both', 'her
s', 'needn', 'his', 'ourselves', 'that'll', 'about', 'here', 'you'd', 'ours', 'isn't', 'di
d', 'yours', 'again', 'now', 'you', 'yourselves', 'shan', 'into', 'our', 'd', 'few', 'woul
dn', 'wouldn't', 'then', 'and', 'there', 'to', 'y', 'because', 'couldn't', 'on', 'won', 'a
s', 'the', 'that', 'above', 'm', 'don', 'after', 'haven't', 'at', 'having', 'me', 'she',
'won't', 'some', 'when', 'it', 'you've', 'wasn', 'you're', 'weren't', 'than', 'until', 've
ry', 'too', 'under', 'herself', 'those', 'were', 'will', 'myself', 'an', 'shan't', 'more',
'same', 'such', 'most', 'be', 'whom'}
```

```
In [9]: # Elimino els stopwords del meu text
textFiltrat=[]
for w in tokenParaules:
    if w not in stop_words:
        textFiltrat.append(w)
print ("Text amb stopwords:", tokenParaules[:200])
print ("_____")
print ("Text sense stopwords:",textFiltrat[:200])
```

```
Text amb stopwords: ['\uffeffThe', 'Project', 'Gutenberg', 'eBook', 'of', 'Twenty', 'tale
s', 'by', 'twenty', 'women', ',', 'by', 'Anonymous', 'This', 'eBook', 'is', 'for', 'the',
'use', 'of', 'anyone', 'anywhere', 'in', 'the', 'United', 'States', 'and', 'most', 'othe
r', 'parts', 'of', 'the', 'world', 'at', 'no', 'cost', 'and', 'with', 'almost', 'no', 'res
trictions', 'whatsoever', '.', 'You', 'may', 'copy', 'it', ',', 'give', 'it', 'away', 'o
```

```
r', 're-use', 'it', 'the', 'terms', 'of', 'the', 'Project', 'Gutenberg', 'License', 'included', 'with', 'this', 'eBook', 'or', 'online', 'at', 'www.gutenberg.org', '.', 'If', 'you', 'are', 'not', 'located', 'in', 'the', 'United', 'States', ',', 'you', 'will', 'have', 'to', 'check', 'the', 'laws', 'of', 'the', 'country', 'where', 'you', 'are', 'located', 'before', 'using', 'this', 'eBook', '.', 'Title', ':', 'Twenty', 'tales', 'by', 'twenty', 'women', 'From', 'real', 'life', 'in', 'Chicago', 'Author', ':', 'Anonymous', 'Release', 'Date', ':', 'September', '1', ',', '2022', '[', 'eBook', '#', '68894', ']', 'Language', ':', 'English', 'Produced', 'by', ':', 'MFR', 'and', 'the', 'Online', 'Distributed', 'Proofreading', 'Team', 'at', 'https', ':', '://www.pgdp.net', '(', 'This', 'file', 'was', 'produced', 'from', 'images', 'generously', 'made', 'available', 'by', 'The', 'Internet', 'Archive', ')', '*', '*', '*', 'START', 'OF', 'THE', 'PROJECT', 'GUTENBERG', 'EBOOK', 'TWENTY', 'TALES', 'BY', 'TWENTY', 'WOMEN', '*', '*', '*', 'TWENTY', 'TALES', 'BY', 'Twenty', 'Women', 'From', 'Real', 'Life', 'in', 'Chicago', 'ANONYMOUS', 'ALL', 'RIGHTS', 'RESERVED', 'FOR', 'SALE', 'ONLY', 'BY', 'NOVELTY', 'PUBLISHING', 'CO.', 'CHICAGO', ',', 'ILL']
```

---

```
Text sense stopwords: ['\uffeffThe', 'Project', 'Gutenberg', 'eBook', 'Twenty', 'tales', 'twenty', 'women', ',', 'Anonymous', 'This', 'eBook', 'use', 'anyone', 'anywhere', 'United', 'States', 'parts', 'world', 'cost', 'almost', 'restrictions', 'whatsoever', '.', 'You', 'may', 'copy', ',', 'give', 'away', 're-use', 'terms', 'Project', 'Gutenberg', 'License', 'included', 'eBook', 'online', 'www.gutenberg.org', '.', 'If', 'located', 'United', 'States', ',', 'check', 'laws', 'country', 'located', 'using', 'eBook', '.', 'Title', ':', 'Twenty', 'tales', 'twenty', 'women', 'From', 'real', 'life', 'Chicago', 'Author', ':', 'Anonymous', 'Release', 'Date', ':', 'September', '1', ',', '2022', '[', 'eBook', '#', '68894', ']', 'Language', ':', 'English', 'Produced', ':', 'MFR', 'Online', 'Distributed', 'Proofreading', 'Team', 'https', ':', '://www.pgdp.net', '(', 'This', 'file', 'produced', 'images', 'generously', 'made', 'available', 'The', 'Internet', 'Archive', ')', '*', '*', '*', 'START', 'OF', 'THE', 'PROJECT', 'GUTENBERG', 'EBOOK', 'TWENTY', 'TALES', 'BY', 'TWENTY', 'WOMEN', '*', '*', '*', 'TWENTY', 'TALES', 'BY', 'Twenty', 'Women', 'From', 'Real', 'Life', 'Chicago', 'ANONYMOUS', 'ALL', 'RIGHTS', 'RESERVED', 'FOR', 'SALE', 'ONLY', 'BY', 'NOVELTY', 'PUBLISHING', 'CO.', 'CHICAGO', ',', 'ILL', '.', 'COPYRIGHT', ',', '1903', 'BY', 'NOVELTY', 'PUBLISHING', 'CO', '.', 'CONTENTS', '.', 'PAGE', 'PREFACE', '3', 'INTRODUCTION', '5', 'A', 'WOMAN', ',', 'S', 'ANGUISH', '13', 'TALE', 'ONE-The', 'Diary', 'Chicago', 'Girl', '17', 'TALE', 'TWO-The', 'Life', 'Story', 'Southern', 'Widow', '33', 'TALE', 'THREE-A', 'Story', 'Chicago', 'Ghetto', '53', 'TALE', 'FOUR-A', 'Woman', 'Thirty-eight', '71', 'TALE', 'FIVE-A', 'Forecast', '89', 'TALE', 'SIX-A', 'Daughter', 'Proud', 'Kentucky', '103', 'TALE', 'SEVEN-My']
```

```
In [10]: # Per veure més la diferència, imprimeixo el tamany dels dos textos per veure si ha canviat
print("Text sense filtrar =", len(tokenParaules), "paraules")
print("Text filtrat =", len(textFiltrat), "paraules")
```

```
Text sense filtrar = 103523 paraules
Text filtrat = 60542 paraules
```

```
In [11]: # Ara realitzaré stemming, que bàsicament vol dir reduir les paraules a la seva arrel principal
from nltk.stem import PorterStemmer
from nltk.tokenize import sent_tokenize, word_tokenize

ps = PorterStemmer()

textStemming=[]
for w in textFiltrat:
    textStemming.append(ps.stem(w))

print("Text sense stemming:",textFiltrat[:200])
print("_____")
print("Text amb stemming:",textStemming[:200])
```

```
Text sense stemming: ['\uffeffThe', 'Project', 'Gutenberg', 'eBook', 'Twenty', 'tales', 'twenty', 'women', ',', 'Anonymous', 'This', 'eBook', 'use', 'anyone', 'anywhere', 'United', 'States', 'parts', 'world', 'cost', 'almost', 'restrictions', 'whatsoever', '.', 'You', 'may', 'copy', ',', 'give', 'away', 're-use', 'terms', 'Project', 'Gutenberg', 'License', 'included', 'eBook', 'online', 'www.gutenberg.org', '.', 'If', 'located', 'United', 'States', ',', 'check', 'laws', 'country', 'located', 'using', 'eBook', '.', 'Title', ':', 'Twenty', 'tales', 'twenty', 'women', 'From', 'real', 'life', 'Chicago', 'Author', ':', 'Anonymous', 'Release', 'Date', ':', 'September', '1', ',', '2022', '[', 'eBook', '#', '68894', ']', 'Language', ':', 'English', 'Produced', ':', 'MFR', 'Online', 'Distributed', 'Proofreading', 'Team', 'https', ':', '://www.pgdp.net', '(', 'This', 'file', 'produced', 'images', 'generously', 'made', 'available', 'The', 'Internet', 'Archive', ')', '*', '*', '*', 'START', 'OF', 'THE', 'PROJECT', 'GUTENBERG', 'EBOOK', 'TWENTY', 'TALES', 'BY', 'TWENTY', 'WOMEN', '*', '*', '*', 'TWENTY', 'TALES', 'BY', 'Twenty', 'Women', 'From', 'Real', 'Life', 'Chicago', 'ANONYMOUS', 'ALL', 'RIGHTS', 'RESERVED', 'FOR', 'SALE', 'ONLY', 'BY', 'NOVELTY', 'PUBLISHING', 'CO.', 'CHICAGO', ',', 'ILL', '.', 'COPYRIGHT', ',', '1903', 'BY', 'NOVELTY', 'PUBLISHING', 'CO', '.', 'CONTENTS', '.', 'PAGE', 'PREFACE', '3', 'INTRODUCTION', '5', 'A', 'WOMAN', ',', 'S', 'ANGUISH', '13', 'TALE', 'ONE-The', 'Diary', 'Chicago', 'Girl', '17', 'TALE', 'TWO-The', 'Life', 'Story', 'Southern', 'Widow', '33', 'TALE', 'THREE-A', 'Story', 'Chicago', 'Ghetto', '53', 'TALE', 'FOUR-A', 'Woman', 'Thirty-eight', '71', 'TALE', 'FIVE-A', 'Forecast', '89', 'TALE', 'SIX-A', 'Daughter', 'Proud', 'Kentucky', '103', 'TALE', 'SEVEN-My']
```

```
ty', 'tales', 'twenty', 'women', 'From', 'real', 'life', 'Chicago', 'Author', ':', 'Anonym
ous', 'Release', 'Date', ':', 'September', '1', ',', '2022', '[', 'eBook', '#', '68894',
]', 'Language', ':', 'English', 'Produced', ':', 'MFR', 'Online', 'Distributed', 'Proofre
ading', 'Team', 'https', ':', 'http://www.pgdp.net', '(', 'This', 'file', 'produced', 'images',
'generously', 'made', 'available', 'The', 'Internet', 'Archive', ')', '*', '*', '*', 'STAR
T', 'OF', 'THE', 'PROJECT', 'GUTENBERG', 'EBOOK', 'TWENTY', 'TALES', 'BY', 'TWENTY', 'WOME
N', '*', '*', '*', 'TWENTY', 'TALES', 'BY', 'Twenty', 'Women', 'From', 'Real', 'Life', 'Ch
icago', 'ANONYMOUS', 'ALL', 'RIGHTS', 'RESERVED', 'FOR', 'SALE', 'ONLY', 'BY', 'NOVELTY',
'PUBLISHING', 'CO.', 'CHICAGO', ',', 'ILL', '.', 'COPYRIGHT', ',', '1903', 'BY', 'NOVELT
Y', 'PUBLISHING', 'CO', '.', 'CONTENTS', '.', 'PAGE', 'PREFACE', '3', 'INTRODUCTION', '5',
'A', 'WOMAN', "'", 'S', 'ANGUISH', '13', 'TALE', 'ONE-The', 'Diary', 'Chicago', 'Girl', '1
7', 'TALE', 'TWO-The', 'Life', 'Story', 'Southern', 'Widow', '33', 'TALE', 'THREE-A', 'Sto
ry', 'Chicago', 'Ghetto', '53', 'TALE', 'FOUR-A', 'Woman', 'Thirty-eight', '71', 'TALE',
'FIVE-A', 'Forecast', '89', 'TALE', 'SIX-A', 'Daughter', 'Proud', 'Kentucky', '103', 'TAL
E', 'SEVEN-My']
```

```
Text amb stemming: ['\ufe00the', 'project', 'gutenberg', 'ebook', 'twenti', 'tale', 'twent
i', 'women', ',', 'anonym', 'thi', 'ebook', 'use', 'anyon', 'anywher', 'unit', 'state', 'p
art', 'world', 'cost', 'almost', 'restrict', 'whatsoev', '.', 'you', 'may', 'copi', ',',
'give', 'away', 're-us', 'term', 'project', 'gutenberg', 'licens', 'includ', 'ebook', 'onl
in', 'www.gutenberg.org', '.', 'if', 'locat', 'unit', 'state', ',', 'check', 'law', 'count
ri', 'locat', 'use', 'ebook', '.', 'titl', ':', 'twenti', 'tale', 'twenti', 'women', 'fro
m', 'real', 'life', 'chicago', 'author', ':', 'anonym', 'releas', 'date', ':', 'septemb',
'1', ',', '2022', '[', 'ebook', '#', '68894', ']', 'languag', ':', 'english', 'produc',
':', 'mfr', 'onlin', 'distribut', 'proofread', 'team', 'http', ':', 'http://www.pgdp.net', '(',
'thi', 'file', 'produc', 'imag', 'gener', 'made', 'avail', 'the', 'internet', 'archiv',
')', '*', '*', '*', 'start', 'of', 'the', 'project', 'gutenberg', 'ebook', 'twenti', 'tal
e', 'by', 'twenti', 'women', '*', '*', '*', 'twenti', 'tale', 'by', 'twenti', 'women', 'fr
om', 'real', 'life', 'chicago', 'anonym', 'all', 'right', 'reserv', 'for', 'sale', 'onli',
'by', 'novelti', 'publish', 'co.', 'chicago', ',', 'ill', '.', 'copyright', ',', '1903',
'by', 'novelti', 'publish', 'co', '.', 'content', '.', 'page', 'prefac', '3', 'introduc',
'5', 'a', 'woman', "'", 's', 'anguish', '13', 'tale', 'one-th', 'diari', 'chicago', 'gir
l', '17', 'tale', 'two-th', 'life', 'stori', 'southern', 'widow', '33', 'tale', 'three-a',
'stori', 'chicago', 'ghetto', '53', 'tale', 'four-a', 'woman', 'thirty-eight', '71', 'tal
e', 'five-a', 'forecast', '89', 'tale', 'six-a', 'daughter', 'proud', 'kentucki', '103',
'tale', 'seven-mi']
```

```
In [12]: # Per veure més la diferència, imprimeixo el tamany dels dos textos per veure si ha canviat
print("Text sense stemming =", len(textFiltrat), "paraules")
print("Text amb stemming =", len(textStemming), "paraules")
```

```
Text sense stemming = 60542 paraules
Text amb stemming = 60542 paraules
```

Veiem que hem aconseguit un text amb paraules principals, totes en minúscula i amb el format més bàsic (en aquest cas no hi ha hagut canvis). Un text que és molt més fàcil per treballar-lo i comparar-lo. Per últim, tornaré a fer el gràfic de quines són les paraules amb més freqüència per veure si han canviat o no

```
In [13]: frequenciaStemming = FreqDist(textStemming)
print(frequenciaStemming)
```

```
<FreqDist with 5874 samples and 60542 outcomes>
```

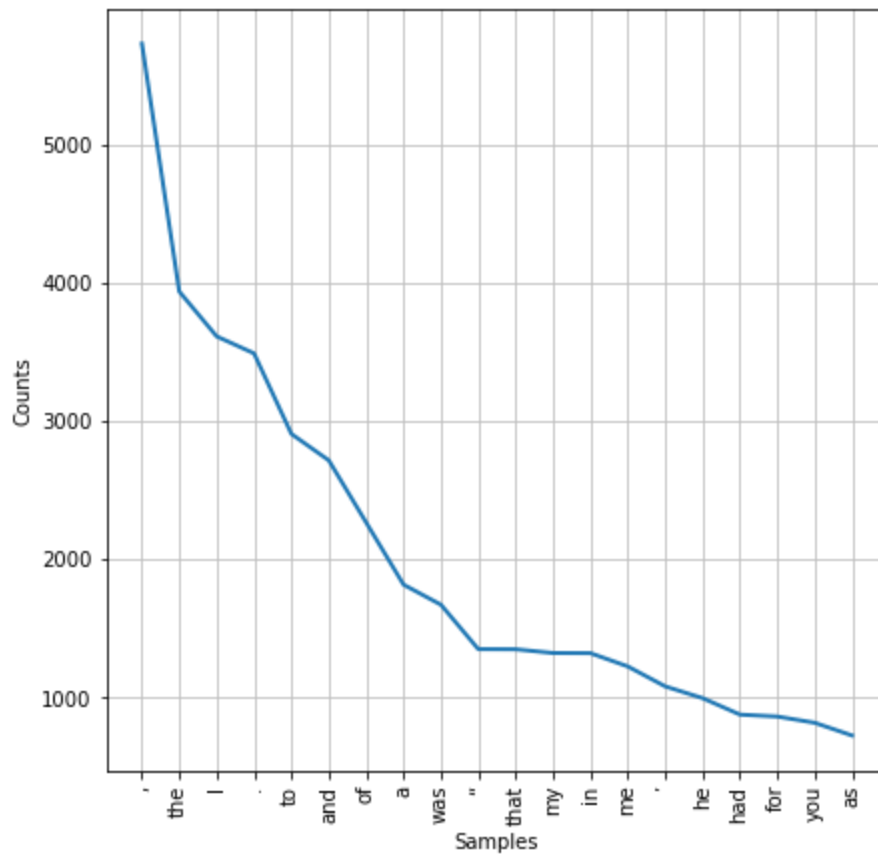
```
In [14]: frequenciaStemming.most_common(10)
```

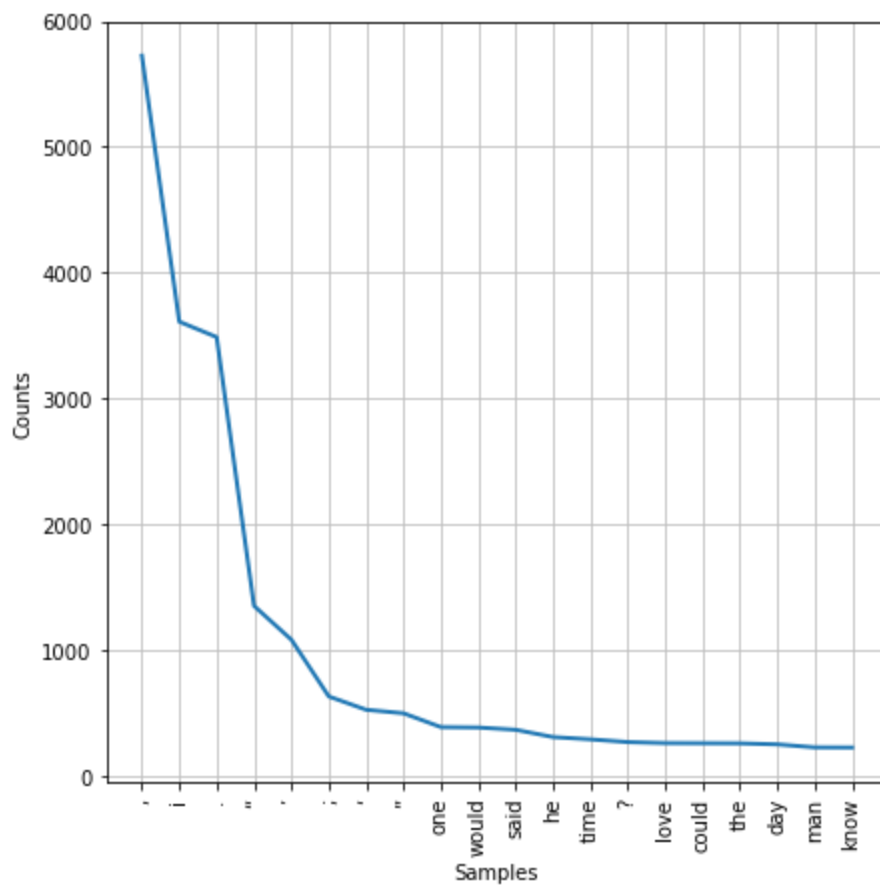
```
Out[14]: [(' ', 5726),
('i', 3611),
('.', 3488),
('\'', 1351),
('\'', 1082),
('; ', 631),
('\'', 525),
('\'', 497),
```

```
('one', 387),  
('would', 384)]
```

In [15]:

```
# Faig el gràfic  
plt.figure(figsize=(7,7))  
frequencia.plot(20,cumulative=False)  
plt.figure(figsize=(7,7))  
frequenciaStemming.plot(20, cumulative=False)  
plt.show()
```





## EXERCICI 3

Realitza sentiment analysis al teu conjunt de dades.

```
In [19]: # Instal·lo les llibreries necessàries
from nltk.sentiment.vader import SentimentIntensityAnalyzer
from nltk import sentiment
from nltk import word_tokenize
```

```
In [20]: #Creo l'analitzador de sentiments
analitzador = SentimentIntensityAnalyzer()
```

```
In [22]: #Aplico l'anàlisis a cada una de les frases que hem fet abans amb el tokenitzador per frase
for sentence in tokenFrase:
    print(sentence)
    scores = analitzador.polarity_scores(sentence)
    for key in scores:
        print(key, ': ', scores[key])
    print()
```

The Project Gutenberg eBook of Twenty tales by twenty women, by Anonymous

This eBook is for the use of anyone anywhere in the United States and most other parts of the world at no cost and with almost no restrictions whatsoever.

neg : 0.097

neu : 0.841

pos : 0.062

compound : -0.1531

You may copy it, give it away or re-use it under the terms  
of the Project Gutenberg License included with this eBook or online at  
[www.gutenberg.org](http://www.gutenberg.org).

neg : 0.0

neu : 1.0

pos : 0.0

compound : 0.0

If you are not located in the United States, you  
will have to check the laws of the country where you are located before  
using this eBook.

neg : 0.0

neu : 0.903

pos : 0.097

compound : 0.4215

Title: Twenty tales by twenty women  
From real life in Chicago

Author: Anonymous

Release Date: September 1, 2022 [eBook #68894]

Language: English

Produced by: MFR and the Online Distributed Proofreading Team at  
<https://www.pgdp.net> (This file was produced from images  
generously made available by The Internet Archive)

\*\*\* START OF THE PROJECT GUTENBERG EBOOK TWENTY TALES BY TWENTY  
WOMEN \*\*\*

TWENTY TALES  
BY  
Twenty Women

From Real Life in Chicago

ANONYMOUS  
ALL RIGHTS RESERVED

FOR SALE ONLY  
BY  
NOVELTY PUBLISHING CO.  
CHICAGO, ILL.

neg : 0.041

neu : 0.927

pos : 0.032



compound : -0.186

COPYRIGHT, 1903

BY  
NOVELTY PUBLISHING CO.

neg : 0.0

neu : 1.0

pos : 0.0

compound : 0.0

CONTENTS.

neg : 0.0

neu : 1.0

pos : 0.0

compound : 0.0

PAGE

PREFACE	3
INTRODUCTION	5
A WOMAN'S ANGUISH	13
TALE ONE—The Diary of a Chicago Girl	17
TALE TWO—The Life Story of a Southern Widow	33
TALE THREE—A Story of the Chicago Ghetto	53
TALE FOUR—A Woman of Thirty-eight	71
TALE FIVE—A Forecast	89
TALE SIX—A Daughter of Proud Kentucky	103
TALE SEVEN—My Lover's Bequest	129
TALE EIGHT—The Victim of a Drug	145
TALE NINE—What Happened to a Girl Who Flirted	163
TALE TEN—Sold at a Fixed Price	173
TALE ELEVEN—A Story of Suicide Bridge	181
TALE TWELVE—Two Babes and Two Mothers	193
TALE THIRTEEN—Not Guilty	205
TALE FOURTEEN—My Lover's Daughter	215
TALE FIFTEEN—As Told to a Clergyman	221
TALE SIXTEEN—A Story of Stage Life	231
TALE SEVENTEEN—A Trip Across the Lake	261
TALE EIGHTEEN—One Woman's Way	269

PREFACE.

neg : 0.106

neu : 0.873

pos : 0.021

compound : -0.9029

"It may be weeds I've gathered, too;  
But even weeds may be as fragrant,  
With some sweet memory.

neg : 0.0

neu : 0.81

pos : 0.19

compound : 0.6124

As the fairest flower."

Without apology this book goes forth.

neg : 0.113

neu : 0.887

pos : 0.0

compound : -0.0382

If it is productive of some good,  
it will have fulfilled its mission.

neg : 0.0

neu : 0.659

pos : 0.341

compound : 0.6908

In presenting this work it is with a feeling of restitution.

neg : 0.0

neu : 0.857

pos : 0.143

compound : 0.128

If I have  
digressed from, or stormed the barricaded citadel of formal literature,  
I have done so without hesitation, simply complying with an obeisance  
to civility toward my fellow men.

neg : 0.0

neu : 0.928

pos : 0.072

compound : 0.2715

I have pictured life as a man of the  
world is sometimes forced to see it, and not altogether as angels would  
transcribe it.

neg : 0.125

neu : 0.875

pos : 0.0

compound : -0.4588

If the manner in which the subjects are hereinafter treated and woven  
into stories, meets the approval of the public, the work will have served  
to indicate the power and simplicity of truth.—THE AUTHOR.

neg : 0.0

neu : 0.914

pos : 0.086

compound : 0.4767

"All truth is precious, if not divine,  
And what dilates the pow'rs must needs refine."

INTRODUCTION.

neg : 0.133

neu : 0.593

pos : 0.274

compound : 0.4724

"Without women, the beginning of our life would be helpless; the middle,  
devoid of pleasure, and the end of consolation."

"The very first  
Of human life must spring from woman's breast,  
Your first small words are taught you from her lips,  
Your first tears quench'd by her, and your last sighs  
Too often breathed out in a woman's hearing,  
When men have shrunk from the ignoble care  
Of watching the last hour of him who led them."

In London alone there are eighty thousand fallen women, and, while the  
number is infinitely smaller in Chicago, they all have a history, an  
excuse to offer, and a tale to tell.

neg : 0.08

neu : 0.838

pos : 0.081

compound : 0.0258

We have resided upon this terrestrial sphere just long enough to know  
that the reformation of a fallen woman rivals the labors of Hercules.

neg : 0.102

neu : 0.898

pos : 0.0

compound : -0.3612

All

men have a physical nature and must meet people who appeal to it.

neg : 0.0

neu : 1.0

pos : 0.0

compound : 0.0

The conditions are such that there has arisen in society, a figure that is certainly the most mournful, and, in some respects, the most awful, upon which the eye of the moralist can dwell.

neg : 0.153

neu : 0.693

pos : 0.154

compound : -0.2467

That unhappy being, whose

very name it is a shame to speak; who counterfeits, with a cold heart, the transports of affection and submits herself as the passive instrument of lust; who is scorned and insulted as the vilest of her sex, and doomed, for the most part, to disease and abject wretchedness of men, then death.

neg : 0.245

neu : 0.681

pos : 0.074

compound : -0.9201

Who will pity her?

neg : 0.423

neu : 0.577

pos : 0.0

compound : -0.296

A poor unknown, who shall be lowered into a grave of cold clay (and possibly in the potter's field), among slimy, creeping things that feed on foul air and putrid masses.

neg : 0.217

neu : 0.783

pos : 0.0

compound : -0.7351

Not even a slab to say,  
"Here lies."

With dreamy eyes and rum dulled brain, her companions take in the scene

without warning.

neg : 0.102

neu : 0.815

pos : 0.083

compound : -0.119

They shrink not from the horrors of the charnel house or the maggot filled grave; sin fascinates them as the cursed death giving flame does the foolish moth.

neg : 0.293

neu : 0.505

pos : 0.202

compound : -0.5862

They continue to cultivate avarice, defy all laws of nature and modesty, all rules of etiquette, and break down all barriers which ordinarily defend pure womanhood.

neg : 0.0

neu : 1.0

pos : 0.0

compound : 0.0

"She is a rag and a bone and a hank of hair."

Women of this class feel that they are social outcasts, that their sins are as scarlet; they believe that they are past reform.

neg : 0.088

neu : 0.912

pos : 0.0

compound : -0.4588

Herself, the supreme type of vice, she is usually the most efficient guardian of virtue.

neg : 0.0

neu : 0.558

pos : 0.442

compound : 0.8588

But for her, the unchallenged purity of countless homes would be polluted, and not a few, who, in the pride of their untested chastity, think of her with an indignant shudder, would have known the agony of remorse and despair.

neg : 0.32

neu : 0.621

pos : 0.058

compound : -0.9313

On that one degraded and ignoble form are concentrated the passions that

might have filled the world with shame.

neg : 0.235

neu : 0.637

pos : 0.127

compound : -0.4019

She remains while creeds and civilization rise and fall, the eternal priestess of humanity, blasted for the sins of the world.

neg : 0.13

neu : 0.87

pos : 0.0

compound : -0.4588

It is not our intention to perorate and dissertate on a theory calculated to turn the world into a miniature heaven, for we don't believe for the fractional part of a moment that a general reformation of the fallen is practicable or possible.

neg : 0.058

neu : 0.864

pos : 0.077

compound : 0.2023

It is not unusual that the men who deplore so loudly the existence of soiled doves are the very men who are responsible for their existence.

neg : 0.0

neu : 0.916

pos : 0.084

compound : 0.3182

The only practicable solution that we may be tempted to offer, would be for society to brand the men with the stigma of its contempt, the same as it does the women, when he sinks himself below her level in an attempt to pervert her purity.

neg : 0.135

neu : 0.821

pos : 0.044

compound : -0.7003

If the immoral men were ostracized from polite society with the same despatch that a weak woman is, society would be composed almost entirely of women.

neg : 0.204

neu : 0.796

pos : 0.0

compound : -0.7096

The world's fallen women are divided into two classes:

The woman whose nature is depraved, who is too coarse to realize or heed the depth of her own infamy, and the woman whose circumstances have forced her to a life of shame.

neg : 0.185

neu : 0.815

pos : 0.0

compound : -0.8225

Of the former, it is useless to take heed  
for she understands nothing outside of her own dep

**limit\_output extension: Maximum message size of 10000 exceeded with 25306 characters**

Per cada frase obtenim 3 puntuacions diferents:

- neg (negatiu): valor entre 0 i 1 que ens indica la negativitat de la frase
- neu (neutral): valor entre 0 i 1 que ens indica la neutralitat de la frase
- pos (positiu): valor entre 0 i 1 que ens indica la positivitat de la frase

compound: valor de -1 a 1 que ens indica la positivitat o negativitat de la frase, com més a prop de -1 més negatiu, a prop de 0 indica neutralitat, a prop de 1 indica positivisme.