

---

---

# NLP

# CORPUS DESCRIPTION WITH SKETCH ENGINE

Daniel, Chidinma, Oppong  
30/10/2024

---

# Corpus

## Novel: Things Fall Apart

Author: Chinua Achebe (1958)

The book talks about the impact of colonization on the African mind. Okonkwo - the main author's world was turned upside down, and he goes from being an influential leader in his community to a paragraph in a colonialist book.

The book portrays the Igbo culture of Nigeria and sheds light on the complexities of identity and power dynamics. The book talks about the colonization of Nigeria by the British, specifically with the area called the Igboland, which is the southeastern part of Nigeria where the Igbo people predominantly live.

---

# Description

## General Info

### COUNTS ⓘ

<b>Tokens</b>	59,452
<b>Words</b>	51,776
<b>Sentences</b>	3,818
<b>Documents</b>	1

### LEXICON SIZES ⓘ

<b>word?</b>	5,084
<b>tag</b>	59
<b>lempos?</b>	4,081
<b>pos</b>	9
<b>lemma</b>	3,692
<b>lempos_lc ⓘ</b>	4,017
<b>lemma_lc ⓘ</b>	3,614
<b>lc ⓘ</b>	4,732

---

---

# Using SketchEngine KEYWORDS

Lemma	Lemma	Lemma	Lemma	Lemma
1 okonkwo ...	11 chielo ...	21 kiaga ...	31 iyi-uwa ...	41 clan ...
2 umuofia ...	12 unoka ...	22 machete ...	32 ezeudu ...	42 odukwe ...
3 ekwefi ...	13 palm-wine ...	23 abame ...	33 ojiugo ...	43 umuru ...
4 obierika ...	14 uchendu ...	24 ogbuefi ...	34 mbaino ...	44 maduka ...
5 ezinma ...	15 mbanta ...	25 obiageli ...	35 kwenu ...	45 bride-price ...
6 nwoye ...	16 obi ...	26 priestess ...	36 hut ...	46 crier ...
7 ikemefuna ...	17 uzowulu ...	27 nwakibie ...	37 ogbanje ...	47 enoch ...
8 yam ...	18 kola ...	28 cowry ...	38 palm-oil ...	48 okoye ...
9 agbala ...	19 goatskin ...	29 kinsman ...	39 chukwu ...	49 ofoedu ...
10 egwugwu ...	20 okagbue ...	30 foo-foo ...	40 ibo ...	50 ozo ...

---

---

# Using SketchEngine WordList

	Noun	Frequency ? ↓
31	head	59 ...
32	thing	58 ...
33	daughter	54 ...
34	story	53 ...
35	tree	52 ...
36	ikemefuna	49 ...
37	life	47 ...
38	crowd	46 ...
39	spirit	46 ...
40	home	45 ...

	Noun	Frequency ? ↓
121	leaf	20 ...
122	song	20 ...
123	kind	20 ...
124	week	20 ...
125	peace	20 ...
126	moment	20 ...
127	harvest	20 ...
128	sun	19 ...
129	okagbue	19 ...
130	smith	19 ...

---

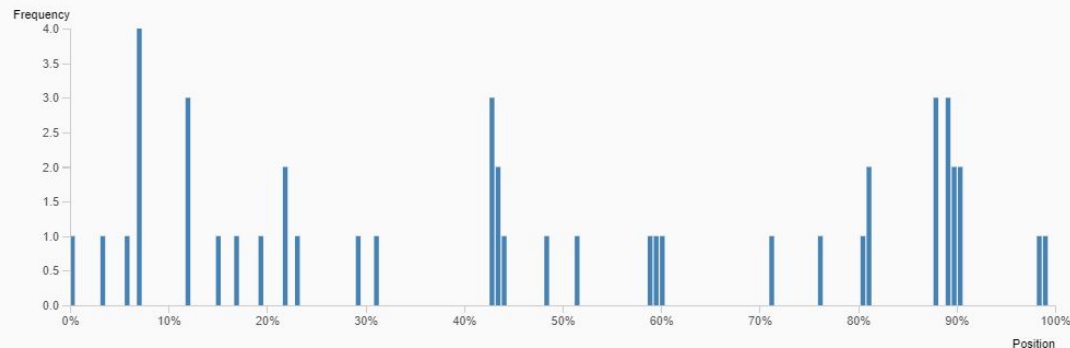
# Theme 1: Spirit

## RESULT DETAILS

simple **spirit**

Number of hits	46
Number of hits per million tokens	773.73
Percent of whole corpus	0.07737%
Corpus size (tokens)	59,452

Distribution of hits in the corpus



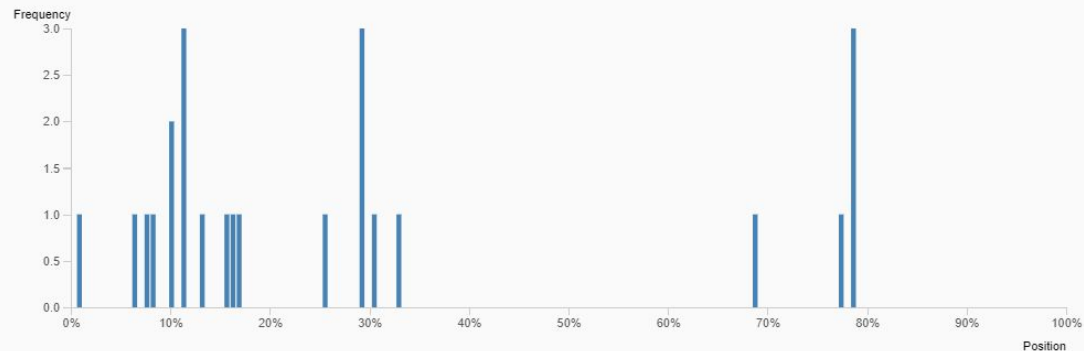
# Theme 2: Harvest

## RESULT DETAILS

simple **harvest**

Number of hits	24
Number of hits per million tokens	403.69
Percent of whole corpus	0.04037%
Corpus size (tokens)	59,452

Distribution of hits in the corpus



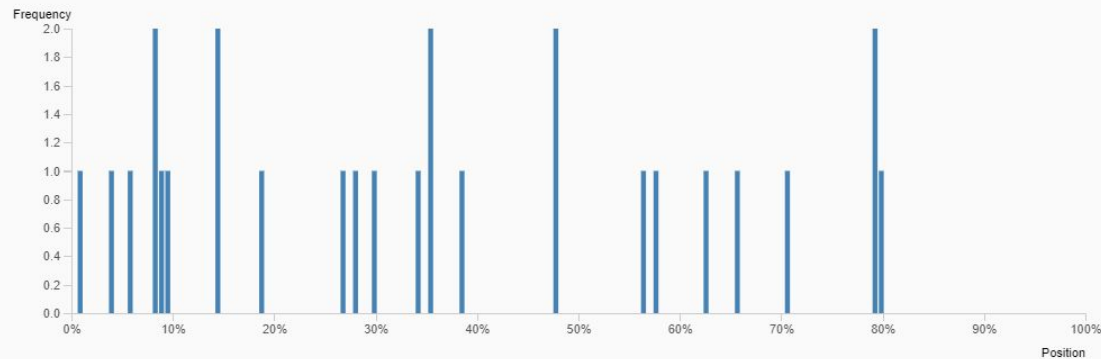
# Theme 3: Palm Wine

## RESULT DETAILS

simple **palm-wine**

Number of hits	27
Number of hits per million tokens	454.15
Percent of whole corpus	0.04541%
Corpus size (tokens)	59,452

Distribution of hits in the corpus





# Using SketchEngine

## WORD SKETCH

Corpus - CA - TFA



clan as noun 116x



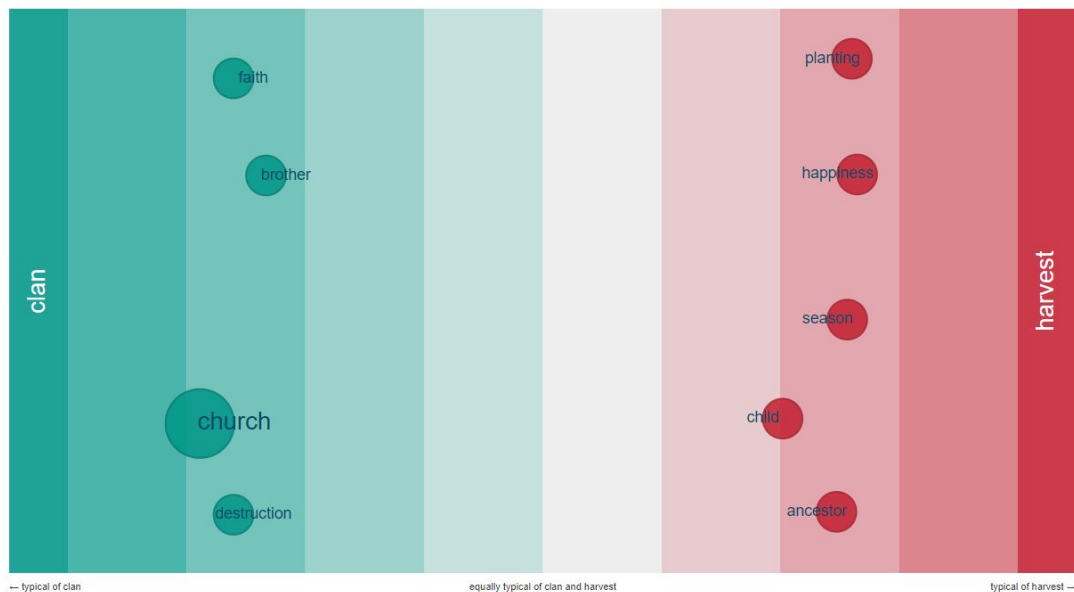
modifiers of "clan"	nouns modified by "clan"	verbs with "clan" as object	verbs with "clan" as subject
<b>whole</b> ... the whole clan	<b>meeting</b> ... clan meeting	<b>neighbour</b> ... the neighbouring clans	<b>know</b> ... clans who naturally knew
<b>distant</b> ... distant clan		<b>pacify</b> ... clan was pacified	<b>undergo</b> ... clan had undergone
<b>womanly</b> ... womanly clan		<b>annoy</b> ... annoy the clan	<b>assume</b> ... clan had assumed
<b>enemy</b> ... enemy clan		<b>ruin</b> ... ruin the whole clan	<b>disobey</b> ... clan had disobeyed
<b>warlike</b> ... warlike clan		<b>worry</b> ... clan was worried	<b>respect</b> ... respected even by the clan
<b>single</b> ... single clan		<b>name</b> ... naming the clan	<b>bury</b> ... clan buried
<b>different</b> ... different clan		<b>break</b> ... break their clan	<b>gather</b> ... clan gathers
<b>other</b> ... other clans		<b>surround</b> ... surrounding clans	<b>reply</b> ... clan replied
<b>many</b> ... many other clans		<b>lead</b> ... lead his warlike clan	<b>decide</b> ... clan had decided
		<b>be</b> ... is not his clan	<b>hold</b> ... clan holds
			<b>do</b> ... the clan did not
			<b>see</b> ... clan saw

# Using Word Sketch Difference

clan 116x harvest 20x

Show relation

"clan /harvest" and/or ...



# Conclusion

1. Sketch Engine provides an analysis of selected text in the corpus.
2. Insights on the usage of the selected terms in different chapters of the corpus.

---