

Natural Language Processing

Bible Vs Koran

by Alex Dance

Background

Have completed the ability to compare multiple aspects of text analysis on large documents / text in tables or from web scraping.

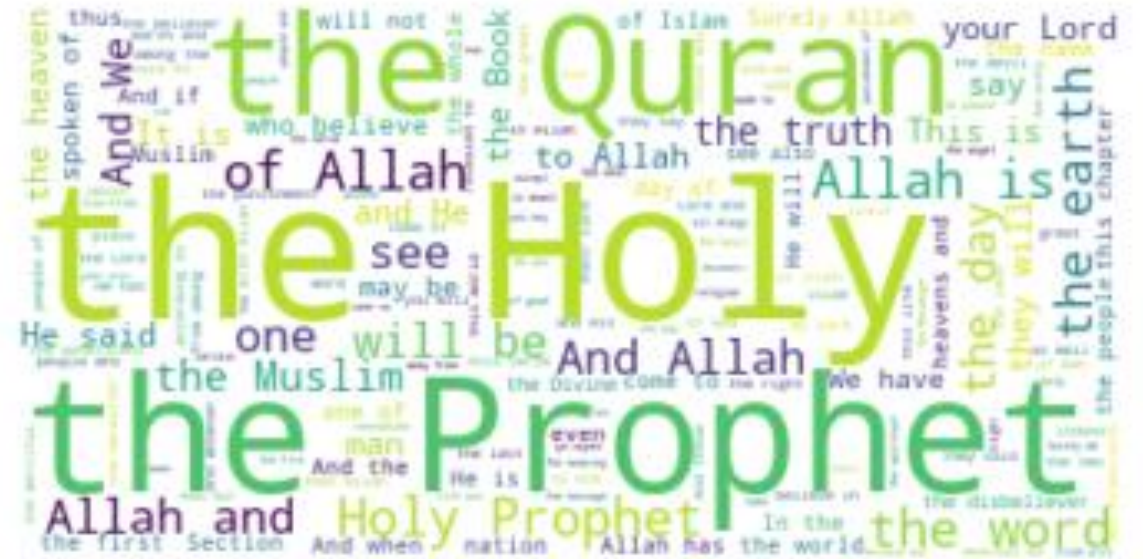
Why Bible Vs Koran: Am not particularly interested in religion or the benefits of a religion vs the other. This is purely statistical and use the tools. Seeing results on their own is partially interesting – but a comparison is more interesting. An alternative comparison could be something like Emily Bronte Vs JK Rowling.

WordCloud

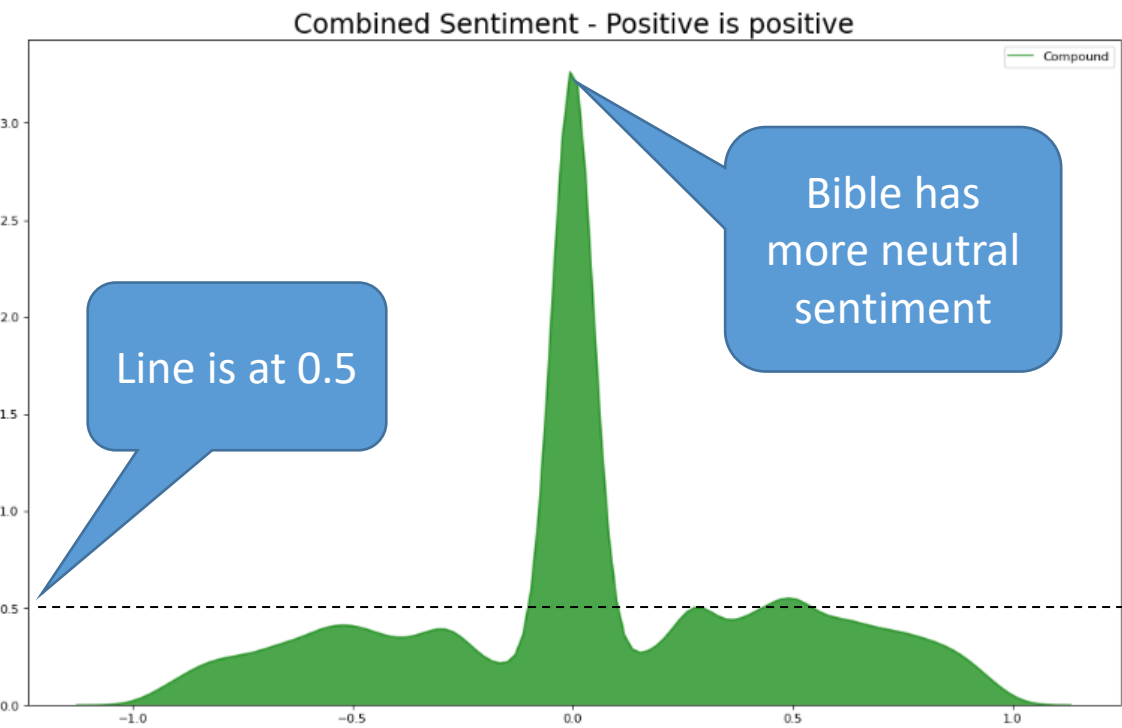
Bible



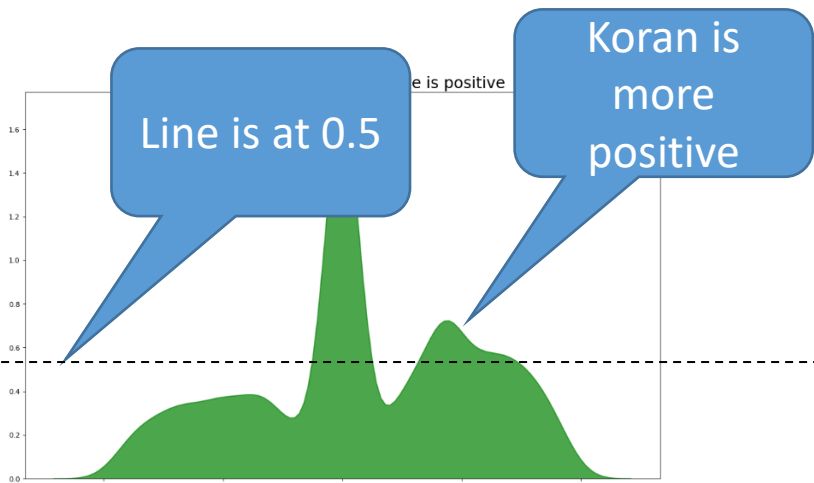
Koran



Sentiment Bible

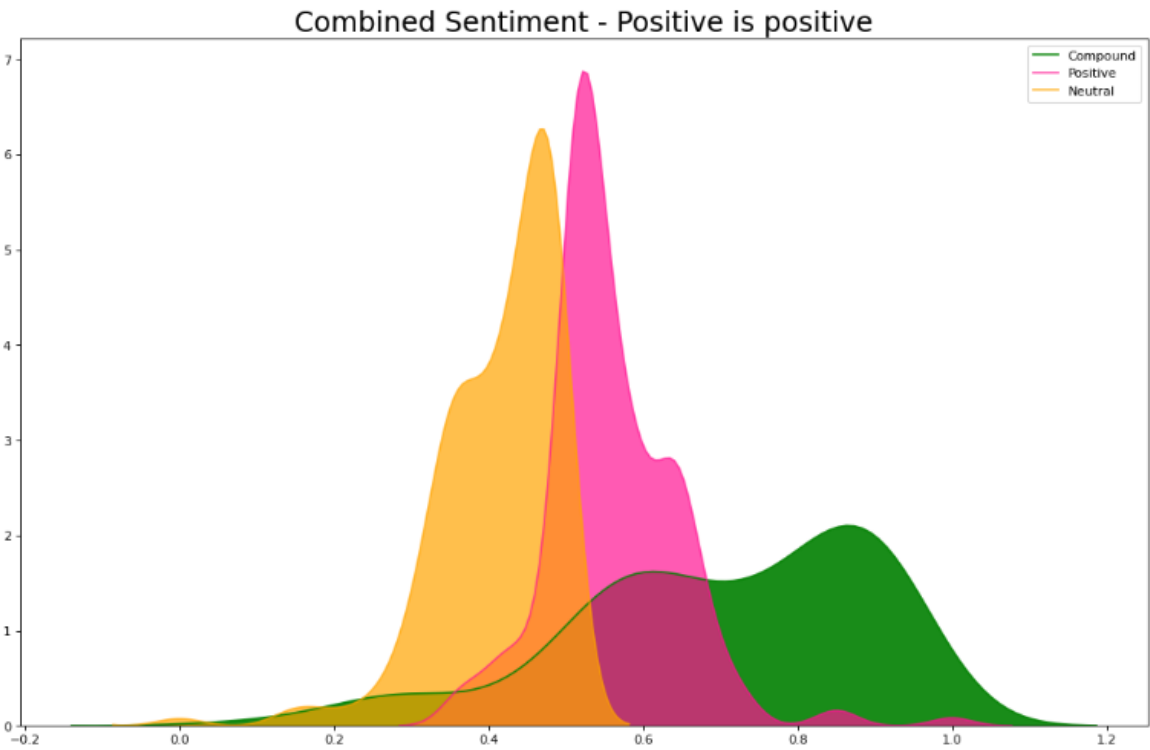


Koran

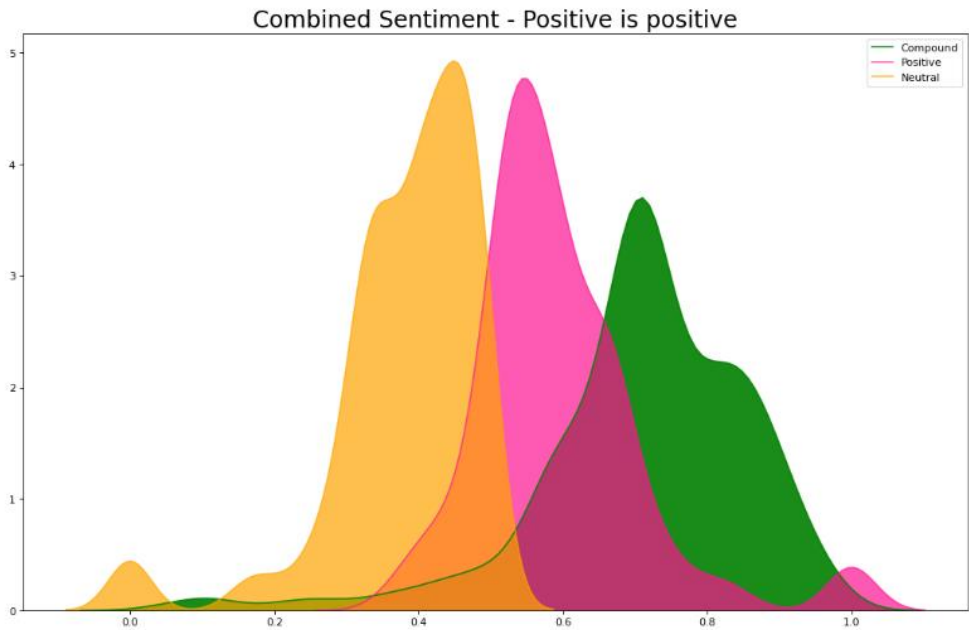


Sentiment

Bible



Koran



Grouped Words and phrases

Bible

		phrase	count	word_count
the son of	1286			
the children of	1274			
the house of	816			
out of the	730			
the land of	611			

	phrase	count	word_count
	And it came to pass,	170	5
	And it came to pass	157	5
	it came to pass, when	135	5
	out of the land of	100	5
	And it shall come to	91	5
	Yahweh, the God of Israel,	88	5

	phrase	count	word_count
	And it came to pass, when	115	6
	the word of Yahweh came to	79	6
And Yahweh spoke to Moses, saying,		72	6
	of Yahweh came to me, saying,	50	6
	word of Yahweh came to me,	50	6

	phrase	count	word_count
	did that which was evil in the sight of	40	9
	in the book of the chronicles of the kings	34	9
	written in the book of the chronicles of the	34	9
	the book of the chronicles of the kings of	34	9

Koran

the Holy Prophet	283
of the Holy	230
the heavens and	195
the day of	176
And Allah is	174

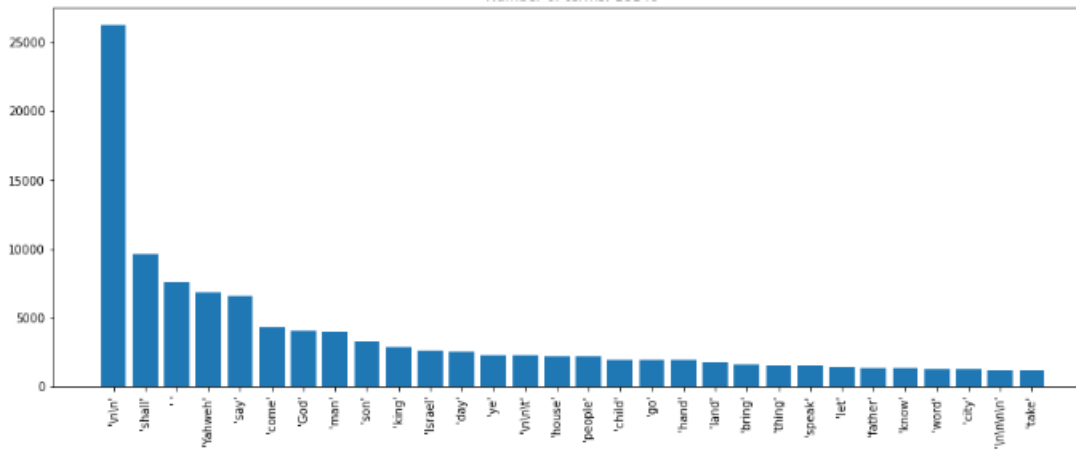
	phrase	count	word_count
	of Allah, the Beneficent, the	117	5
	the name of Allah, the	116	5
	name of Allah, the Beneficent,	116	5
	Allah, the Beneficent, the Merciful.	114	5
	In the name of Allah,	114	5
	the Beneficent, the Merciful. 1	108	5

	phrase	count	word_count
	name of Allah, the Beneficent, the	116	6
	the name of Allah, the Beneficent,	116	6
	In the name of Allah, the	114	6
	of Allah, the Beneficent, the Merciful.	114	6
	Allah, the Beneficent, the Merciful. 1	108	6

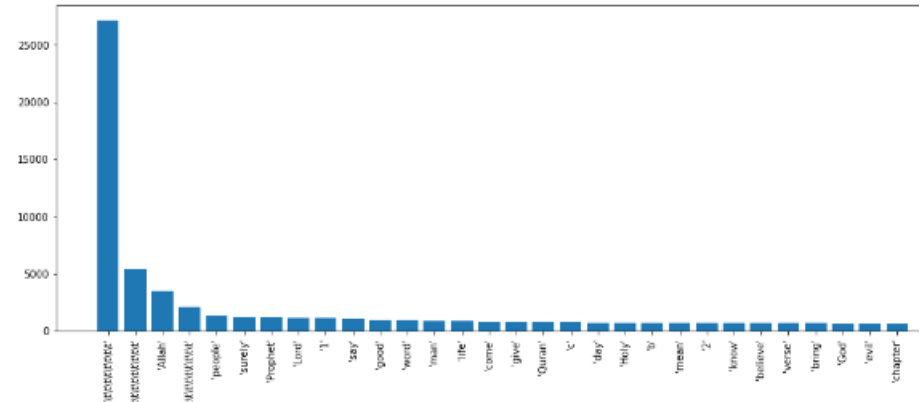
	phrase	count	word_count
	In the name of Allah, the Beneficent, the Merc...	113	9
	the name of Allah, the Beneficent, the Mercifu...	108	9
	of the bounties of your Lord will you deny?	29	9
	then of the bounties of your Lord will you	29	9
	Which then of the bounties of your Lord will	29	9

Main Plot Words (excluding stop words eg the)

Bible

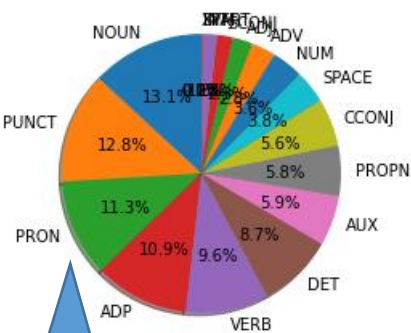


Koran



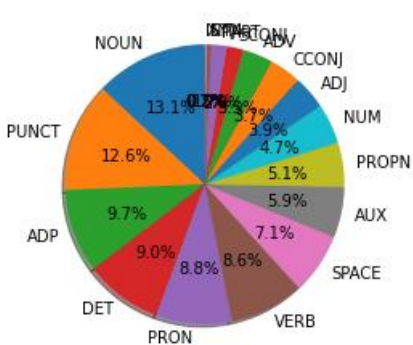
Note: includes some spacing in document eg"/n" etc that can be ignored

Parts of Speech Bible



NOUN	130059
PUNCT	127074
PRON	112158
ADP	108429
VERB	95214
DET	86378
AUX	57964
PROP	57775
CCONJ	55084
SPACE	38014
NUM	35795
ADV	27972
ADJ	22935
SCONJ	18139
PART	15660
INTJ	1920
X	161
SYM	24

Koran



NOUN	64092
PUNCT	61513
ADP	47530
DET	44066
PRON	42881
VERB	42263
SPACE	34616
AUX	28721
PROP	25124
NUM	22926
ADJ	18917
CCONJ	17928
ADV	16927
SCONJ	9762
PART	8566
SYM	2578
X	800
INTJ	344

Bible has higher % of Pronouns

Sentences and Length

Bible

Bible is about
twice as long

```
sentence_length  
word_count  
word_density  
punctuation_count  
title_word_count  
uppercase_word_count
```

```
4291038  
821188  
154854  
158760  
99579  
10423
```

	sentence_length	word_count	word_density	punctuation_count	title_word_count	uppercase_word_count
count	31383.000000	31383.000000	31383.000000	31383.00000	31383.000000	31383.000000
mean	136.731288	26.166651	4.934312	5.05879	3.173024	0.332122
std	103.256286	19.040236	0.547576	4.81070	3.279108	0.749297
min	4.000000	1.000000	2.000000	1.00000	0.000000	0.000000
25%	76.000000	14.000000	4.650000	3.00000	1.000000	0.000000
50%	113.000000	22.000000	4.945455	4.00000	2.000000	0.000000
75%	170.000000	33.000000	5.236842	6.00000	4.000000	0.000000
max	3438.000000	455.000000	47.095890	505.00000	307.000000	12.000000

Koran

Koran has
longer
sentences

```
sentence_length  
word_count  
word_density  
punctuation_count  
title_word_count  
uppercase_word_count
```

```
2698745  
387049  
108537  
78441  
45476  
3497
```

	sentence_length	word_count	word_density	punctuation_count	title_word_count	uppercase_word_count
count	17269.000000	17269.000000	17269.000000	17269.000000	17269.000000	17269.000000
mean	156.276854	22.412936	6.285101	4.542301	2.633389	0.202502
std	212.242314	21.727390	1.619516	16.860872	2.346220	0.751373
min	1.000000	1.000000	0.500000	1.000000	0.000000	0.000000
25%	76.000000	12.000000	5.647059	2.000000	1.000000	0.000000
50%	125.000000	19.000000	6.250000	3.000000	2.000000	0.000000
75%	194.000000	28.000000	6.840000	5.000000	3.000000	0.000000
max	11812.000000	1123.000000	34.715385	972.000000	97.000000	11.000000

Main People Bible

	Detail
Yahweh	2593
David	1110
God	571
Jesus	519
Aaron	276
Saul	268
Jacob	267
O Yahweh	250
Moses	223
Solomon	218
Judah	201
Paul	180
Thus	172
Joseph	167
Babylon	161
John	149
Joshua	145
Peter	135

Yahweh is the most mentioned person in the bible

Koran

	Detail
Allah	2394
Islam	244
Jesus	161
Muslims	161
Him	154
God	144
Abraham	113
Pharaoh	85
Joseph	75
Divine	74
Who	74
Solomon	62
Lord	57
Mary	55
Moses	53

Jesus is 3rd most mentioned person in Koran

Places

Bible

Detail	
Israel	2523
Yahweh	2323
Jerusalem	661
Egypt	608
Judah	519
God	459
Moses	415
Behold	211
Jordan	194
Christ	153
Saul	142
Babylon	123
Moab	121
Pharaoh	119
Assyria	113
Samaria	110
Abraham	103
Ephraim	93
Jacob	91
Benjamin	91

Koran

Detail	
Allah	852
Quran	329
Divine	295
Islam	252
Truth	199
Moses	137
God	128
Him	120
Makkah	114
Us	97
Wise	90
Arabia	84
Madinah	79
Abraham	79
Israel	75
Quraish	68
Mighty	67

Categorisation is
not perfect

Location

Bible

Detail	
Syria	72
Philistines	65
Israel	27
Damascus	21
Assyria	8
Haran	5
Philadelphia	3
Samaria	3
Sidon	1
Gehenna	1

Koran

Detail	
Syria	17
Us	4
North Africa	2
Praiseworthy	2
Constantine	1
Damascus	1
Israel	1

Facilities

Bible

Detail	
Hittite	30
Amorite	23
Canaanite	22
Perizzite	15
Netophathite	8
Girgashite	7
Gittite	6
Shunammite	5
Ahohite	5
Carmelite	5
Agagite	5
Cushite	5
Hushathite	5
Jezreelite	5
Madianite	4

Koran

Detail	
Madinah	26
Garden	12
Israelite	5
Gardens	4
Best Judge	1
Israelites	1
White	1

Orgs

Bible

Detail	
Lord	547
Yahweh	320
Son	159
Lord Yahweh	145
Spirit	113
PhilistinesS	93
Holy Spirit	86
Levites	86
Come	79
God	72
Gentiles	65
Lord Jesus Christ	62
Pharisees	62
Father	50
Philistines	48
Lord Jesus	47
Jerusalem	46
Sabbath	43

Koran

Detail	
Holy Prophet	445
Prophet	418
Muslims	235
Lord	218
Book	217
Beneficent	179
Muslim	173
Merciful	159
Quran	134
THE	112
Makkah	108
Fire	90
Holy Quran	88
Holy	86
Hour	56

Types

Bible

GPE	18545
PERSON	16793
ORGANIZATION	4349
FACILITY	243
GSP	206
LOCATION	43

Koran

PERSON	6285
ORGANIZATION	5179
GPE	4827
FACILITY	50
LOCATION	33
GSP	28

Some Misspelled words (by current standards)

Bible

Koran

'oznites', 'kengkreia', '"', 'admah', '52this', 'perverseness', 'asahel', 'jezrahiah', 'maaz', '67then', '35see', 'hodaviah', '7for', '54he', 'jeshaiah', 'theudas'ginath', 'lampstands', 'vaniah', 'haft', 'intimidations', 'quaestiones', 'birsha', 'jether', '62the', 'baladan', 'harhur', 'zophar', 'iturea',

regene', 'hady', 'saq', 'yukhlifu', 'rejecters', 'ishmaelites', 'hawazin', 'inviter', 'dighth', 'showjou', 'gression', 'charac', 'ihram', 'fited', 'shment', 'nize', '22w', 'thamma', 'suraqa', 'lishing', 'ference', '59m', 'fection', 'disburses', 'cance',

Thanks

Alex Dance



Copy of This Presentation and code

<https://github.com/alexdance2468/>

Plus other data science projects completed

Contact Details

www.linkedin.com/in/alex-dance/

alexdance2468@gmail.com

0414 293 628