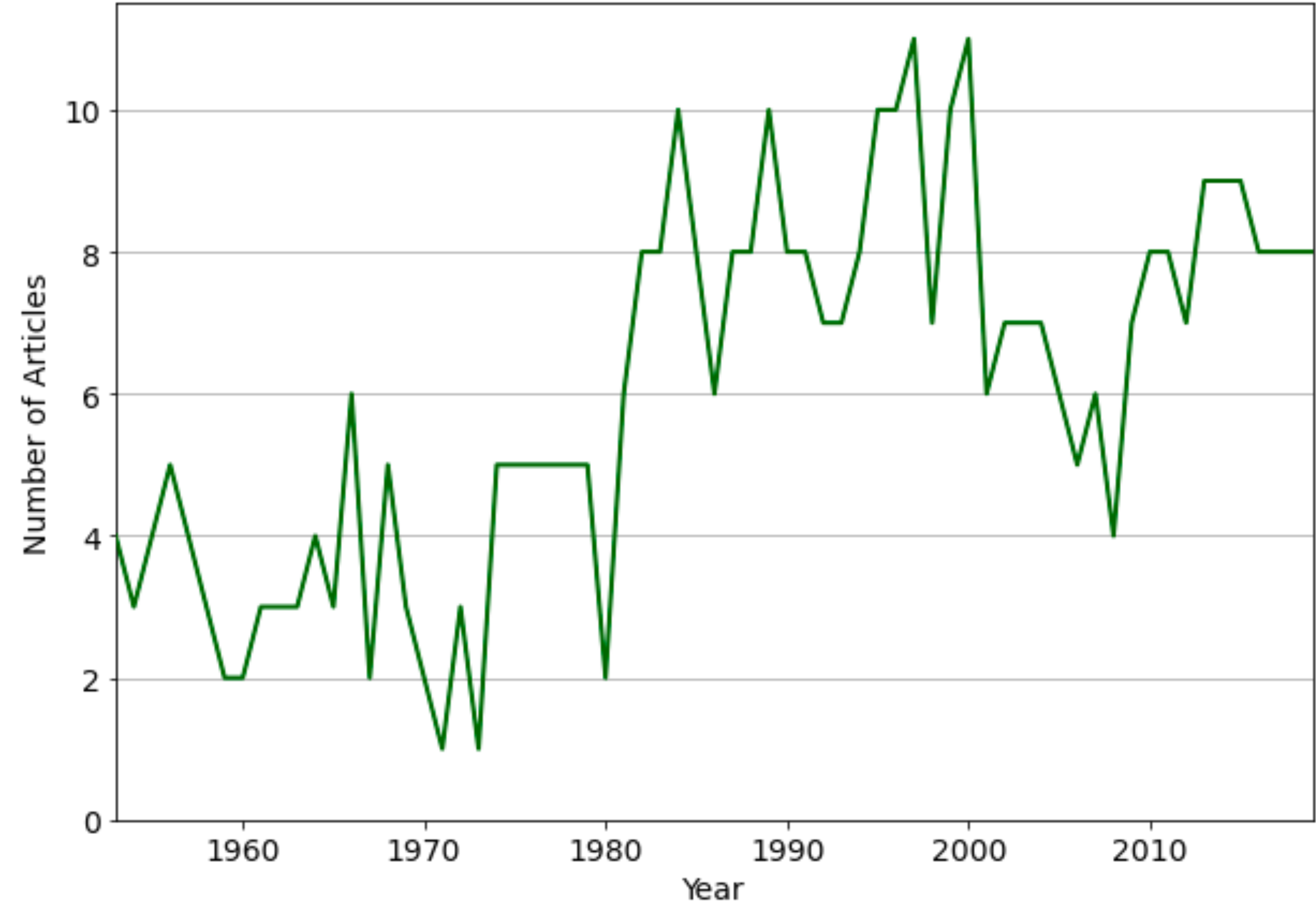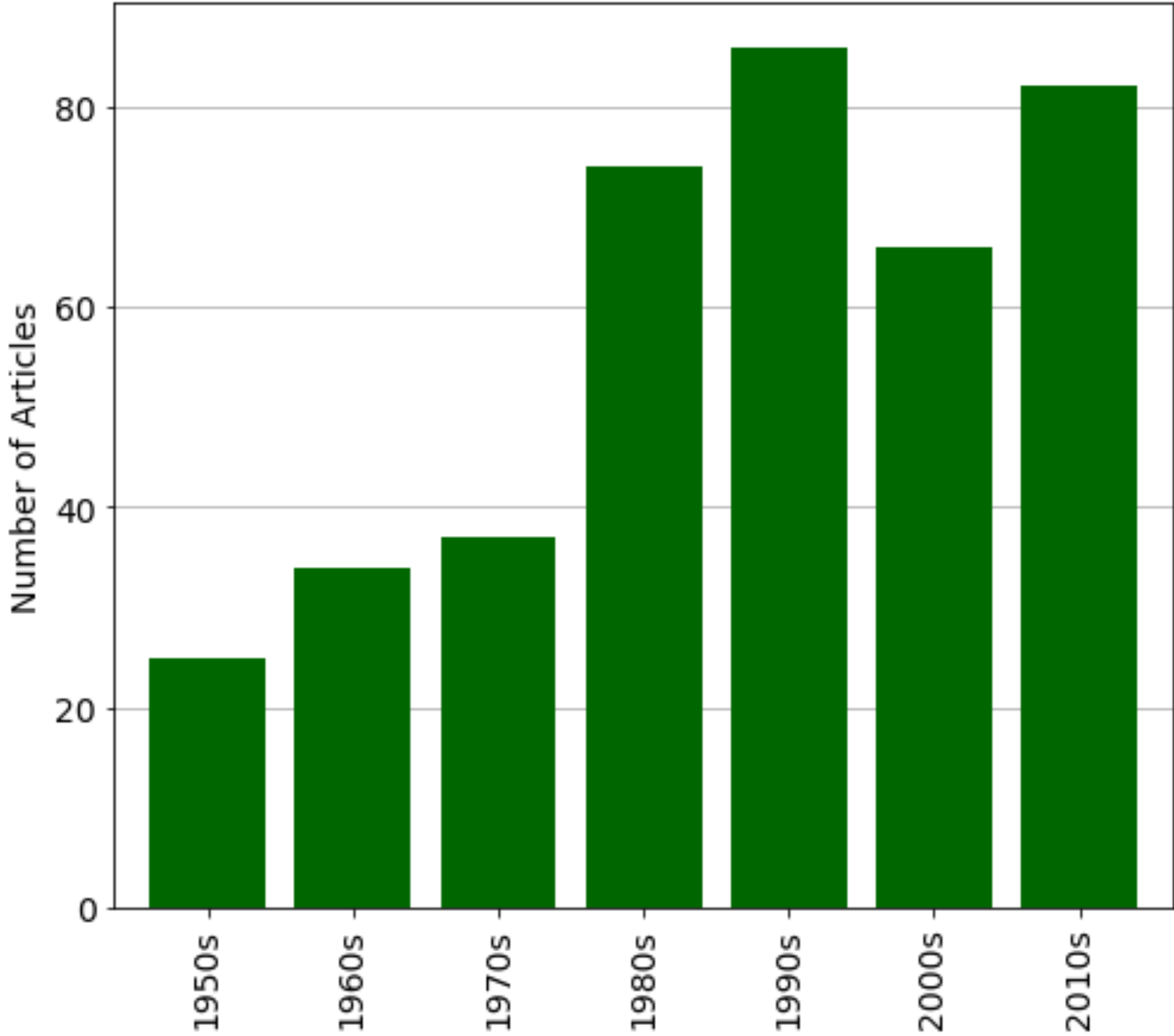# Corpus Analysis: Articles over Time

Number of Articles Per Year

Number of Articles Per Decade

Corpus Analysis: Word Frequencies

Top 50 words by raw frequency count for entire corpus (after stopword filtering)

Rank	Word	Count		Rank	Word	Count
1	write	18782		26	part	2177
2	book	8792		27	remember	2030
3	work	7701		28	friend	1985
4	read	6822		29	language	1922
5	writer	5996		30	young	1875
6	story	5571		31	experience	1853
7	life	4987		32	person	1848
8	poem	4396		33	american	1840
9	live	3874		34	publish	1807
10	man	3749		35	leave	1765
11	world	3211		36	line	1745
12	start	3037		37	fact	1744
13	long	3008		38	mind	1650
14	word	2860		39	form	1645
15	begin	2824		40	reader	1637
16	love	2711		41	father	1599
17	play	2708		42	child	1586
18	interest	2687		43	early	1582
19	character	2665		44	school	1573
20	poetry	2547		45	fiction	1563
21	idea	2497		46	understand	1547
22	end	2482		47	learn	1526
23	poet	2273		48	mother	1508
24	sense	2232		49	english	1486
25	woman	2207		50	turn	1476

Corpus Analysis: TF-IDF Scores

Top 50 words by TF-IDF scores for entire corpus (after stopword filtering)

Rank	Word	Score
1	write	101.05
2	book	47.86
3	work	41.55
4	read	36.32
5	writer	34.04
6	story	31.79
7	poem	31.55
8	life	26.98
9	man	20.94
10	live	20.81
11	world	17.47
12	poetry	17.32
13	character	16.99
14	start	16.88
15	long	16.34
16	play	15.96
17	word	15.78
18	poet	15.55
19	begin	15.52
20	love	15.39
21	interest	15.14
22	idea	14.27
23	woman	14.11
24	end	13.84
25	sense	12.29

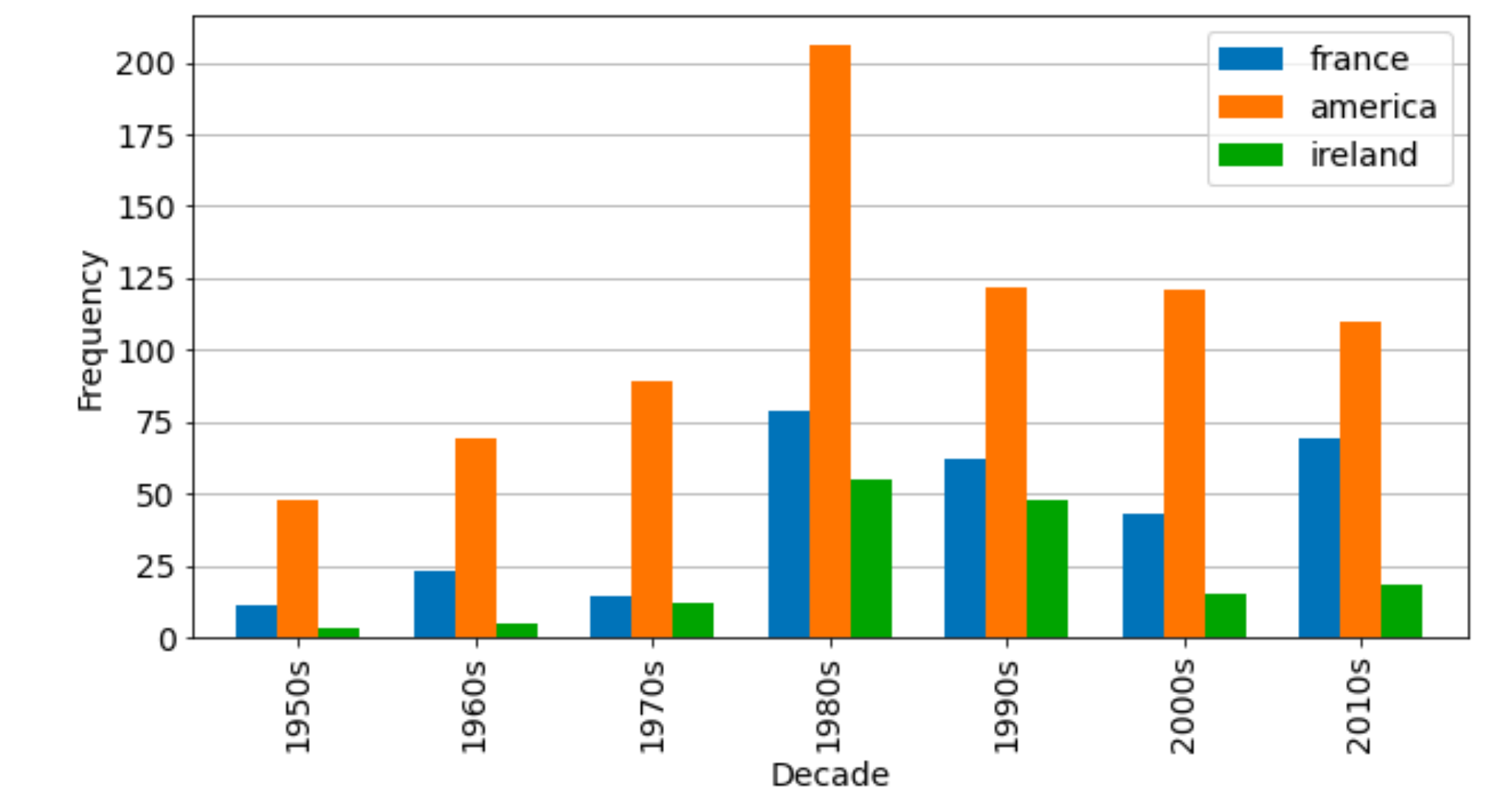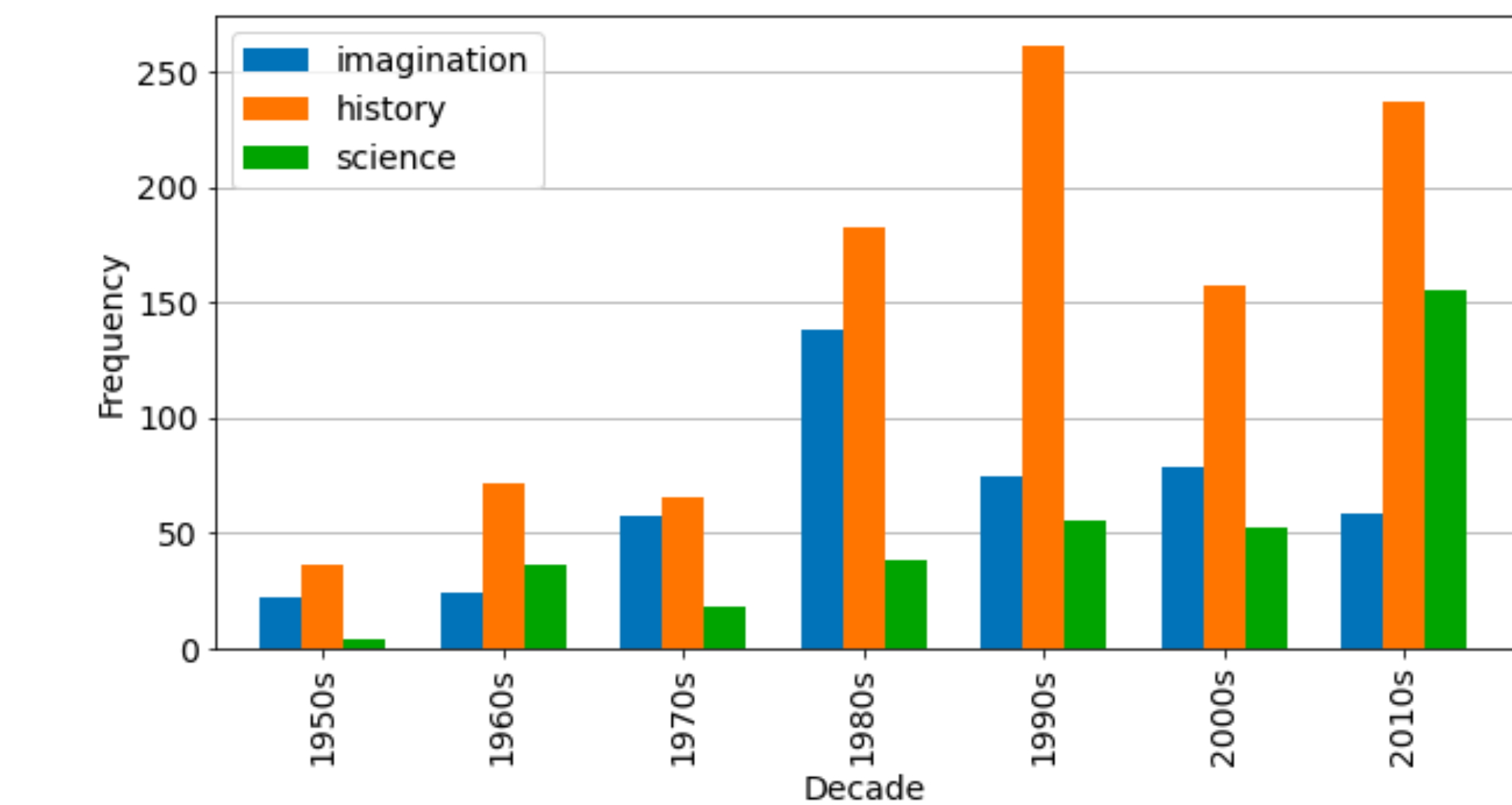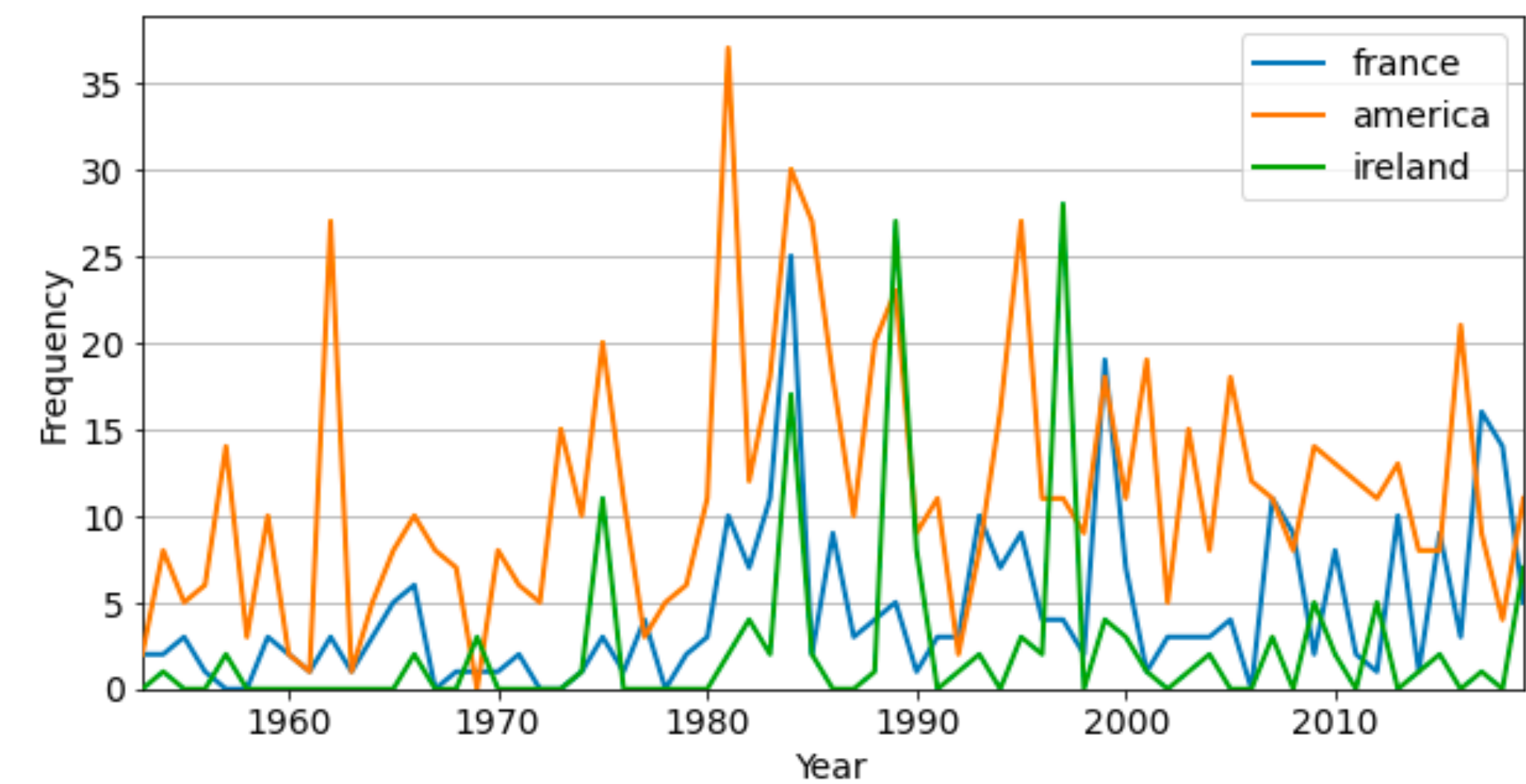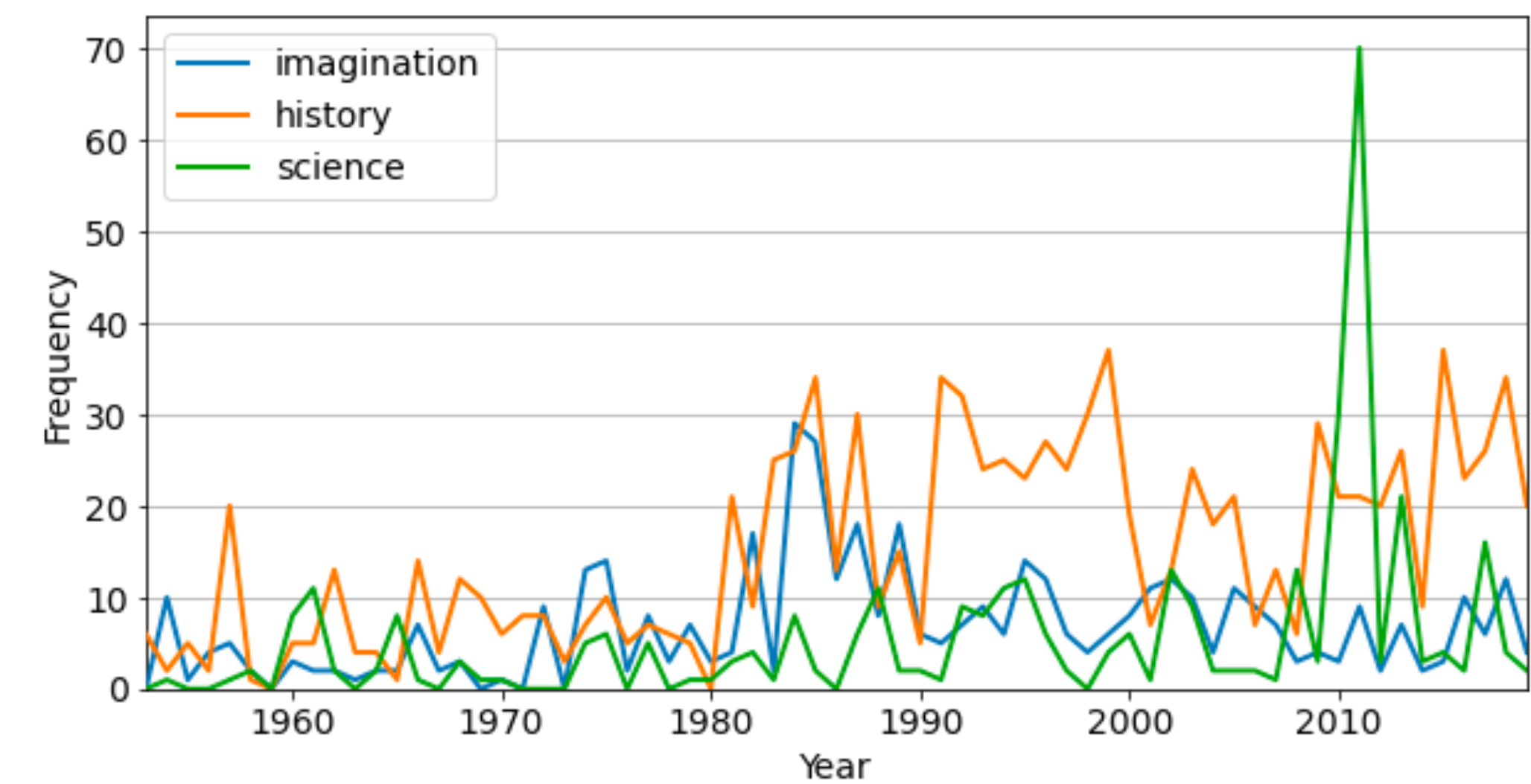| Rank	Word	Score
26	language	12.20
27	part	11.88
28	remember	11.63
29	experience	11.31
30	fiction	11.27
31	american	11.25
32	person	11.23
33	friend	11.15
34	young	10.86
35	publish	10.84
36	father	10.57
37	reader	10.50
38	mother	10.35
39	fact	10.28
40	war	10.18
41	child	10.15
42	line	10.13
43	form	10.04
44	leave	9.86
45	mind	9.61
46	school	9.61
47	english	9.59
48	art	9.52
49	learn	9.27
50	early	9.24

Corpus Analysis: Word Frequencies Over Time

Top 20 words by raw frequency count per decade (after stopword filtering)

Rank	1950s	1960s	1970s	1980s	1990s	2000s	2010s
1	write	write	write	write	write	write	write
2	writer	work	book	book	book	book	book
3	book	read	work	work	work	work	work
4	work	writer	writer	writer	read	read	read
5	story	man	read	read	writer	story	story
6	man	book	poem	story	life	poem	writer
7	read	play	life	life	poem	life	life
8	character	life	man	poem	story	writer	live
9	life	word	story	man	play	live	start
10	live	poem	live	live	live	start	world
11	long	story	poetry	long	man	world	love
12	play	interest	poet	begin	world	man	interest
13	world	begin	interest	world	poetry	love	long
14	word	live	long	poetry	word	end	idea
15	idea	long	love	poet	start	long	word
16	form	idea	begin	woman	long	interest	man
17	begin	world	world	love	begin	word	end
18	start	sense	character	word	character	idea	character
19	young	character	start	play	woman	sense	person
20	interest	end	word	interest	poet	begin	begin

Corpus Analysis: Word Frequencies Over Time

Examples of raw word frequency counts over time (note: not normalised)

Corpus Analysis: TF-IDF Scores Over Time

Top 20 words by TF-IDF scores per decade (after stopword filtering)

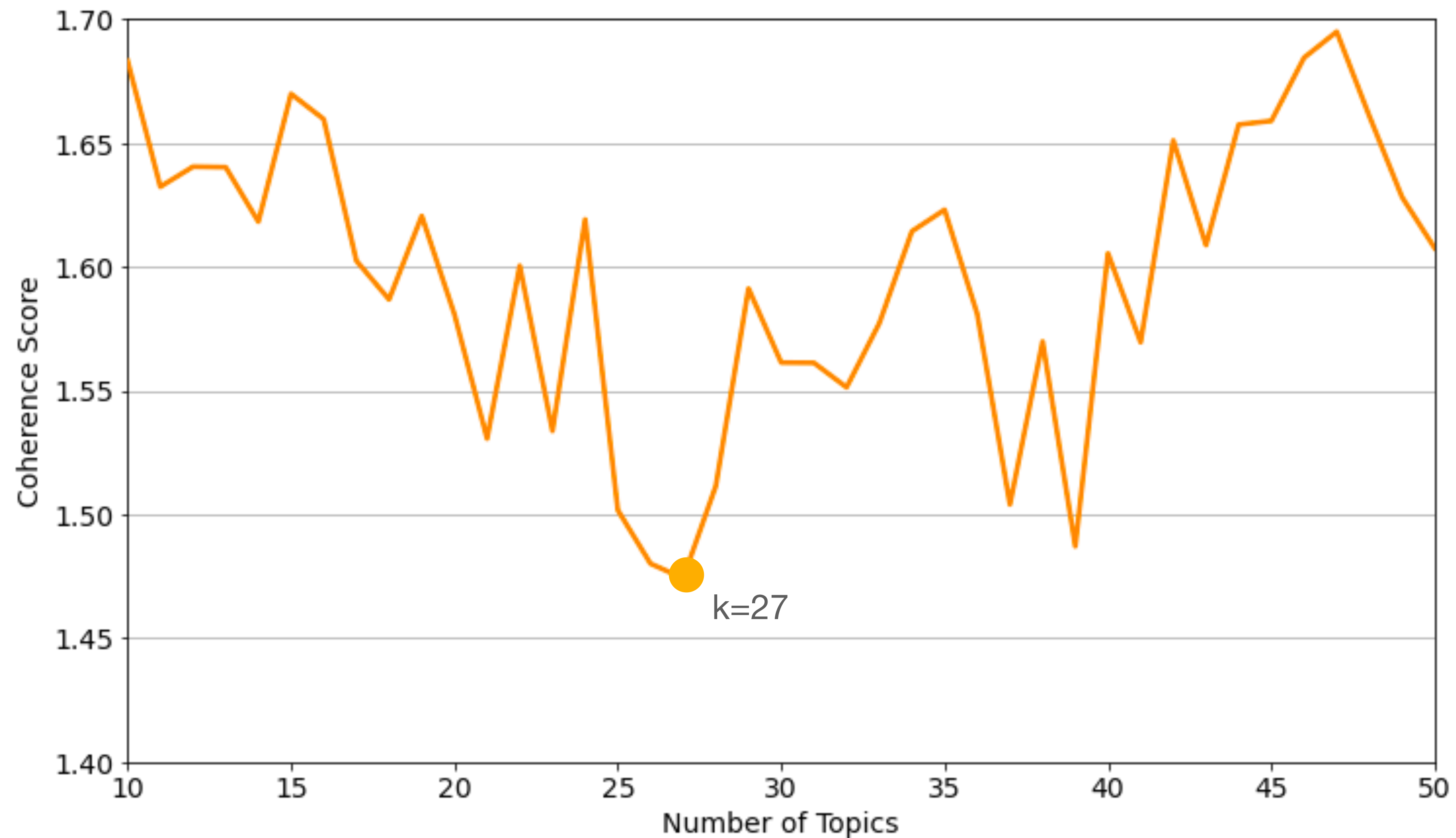| Rank	1950s	1960s	1970s	1980s	1990s	2000s	2010s
1	write	write	write	write	write	write	write
2	book	work	book	book	book	book	book
3	writer	writer	work	work	work	work	work
4	work	read	poem	poem	writer	poem	read
5	story	book	writer	writer	read	read	story
6	character	man	read	read	poem	story	writer
7	man	poem	life	story	story	life	life
8	read	life	story	life	life	writer	live
9	life	play	man	man	play	live	start
10	live	word	poetry	poetry	poetry	start	poem
11	play	story	poet	live	live	world	love
12	long	interest	live	poet	man	love	world
13	idea	live	love	long	character	poetry	interest
14	negro	begin	interest	world	world	man	man
15	world	idea	character	begin	woman	interest	mother
16	start	poet	long	woman	word	end	long
17	begin	character	begin	play	start	mother	idea
18	form	world	start	love	long	father	character
19	word	long	world	word	begin	long	word
20	artist	poetry	remember	american	poet	word	woman

Corpus Analysis: JSD Scores Over Time

Top 20 words by Jensen–Shannon Divergence (JSD) scores per decade (after stopword filtering)

Rank	1950s	1960s	1970s	1980s	1990s	2000s	2010s
1	long	poem	book	work	story	write	story
2	live	story	begin	begin	read	writer	work
3	play	play	interest	fiction	life	poem	reader
4	art	writer	poetry	remember	friend	publish	long
5	word	begin	read	friend	sound	remember	poetry
6	begin	american	remember	word	publish	woman	part
7	love	english	american	long	man	sense	man
8	remember	start	friend	end	war	love	school
9	human	poet	fact	poetry	end	war	father
10	experience	love	public	interest	person	poetry	remember
11	influence	character	paris	home	open	poet	job
12	hear	long	open	hand	learn	bring	head
13	head	world	bear	bad	begin	age	true
14	close	remember	subject	problem	difficult	hear	hard
15	hand	poetry	wonderful	novelist	country	short	george
16	french	film	mine	instance	paper	party	human
17	sentence	critic	hard	history	force	novelist	move
18	full	study	learn	image	black	instance	science
19	paper	modern	author	important	couple	intellectual	poet
20	woman	simple	discover	hope	review	bear	class

Topics Analysis: Selection of Number of Topics

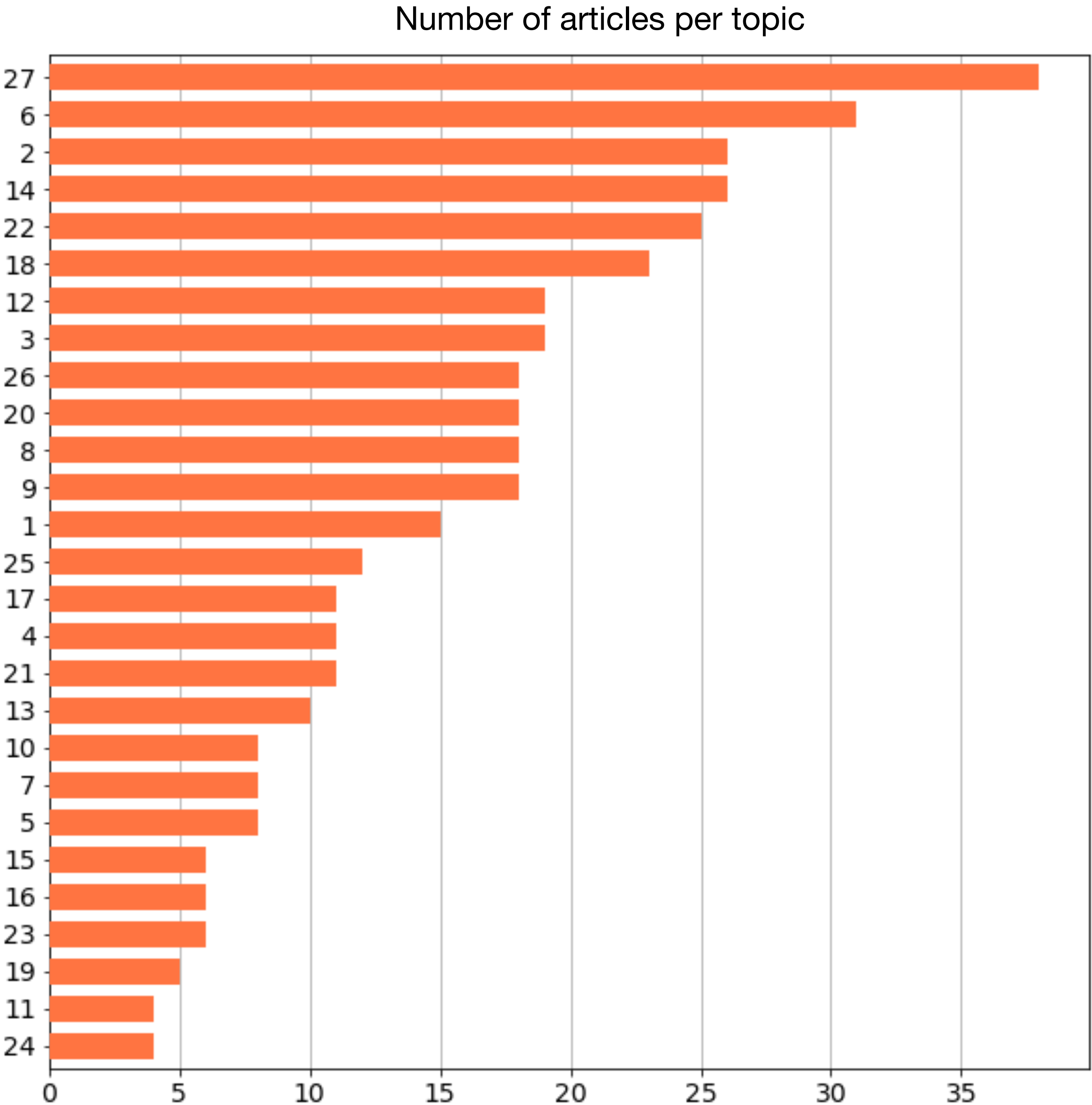Selection of number of topics for topic modelling using *min-max* coherence score

Topics Analysis: Top Word Descriptors

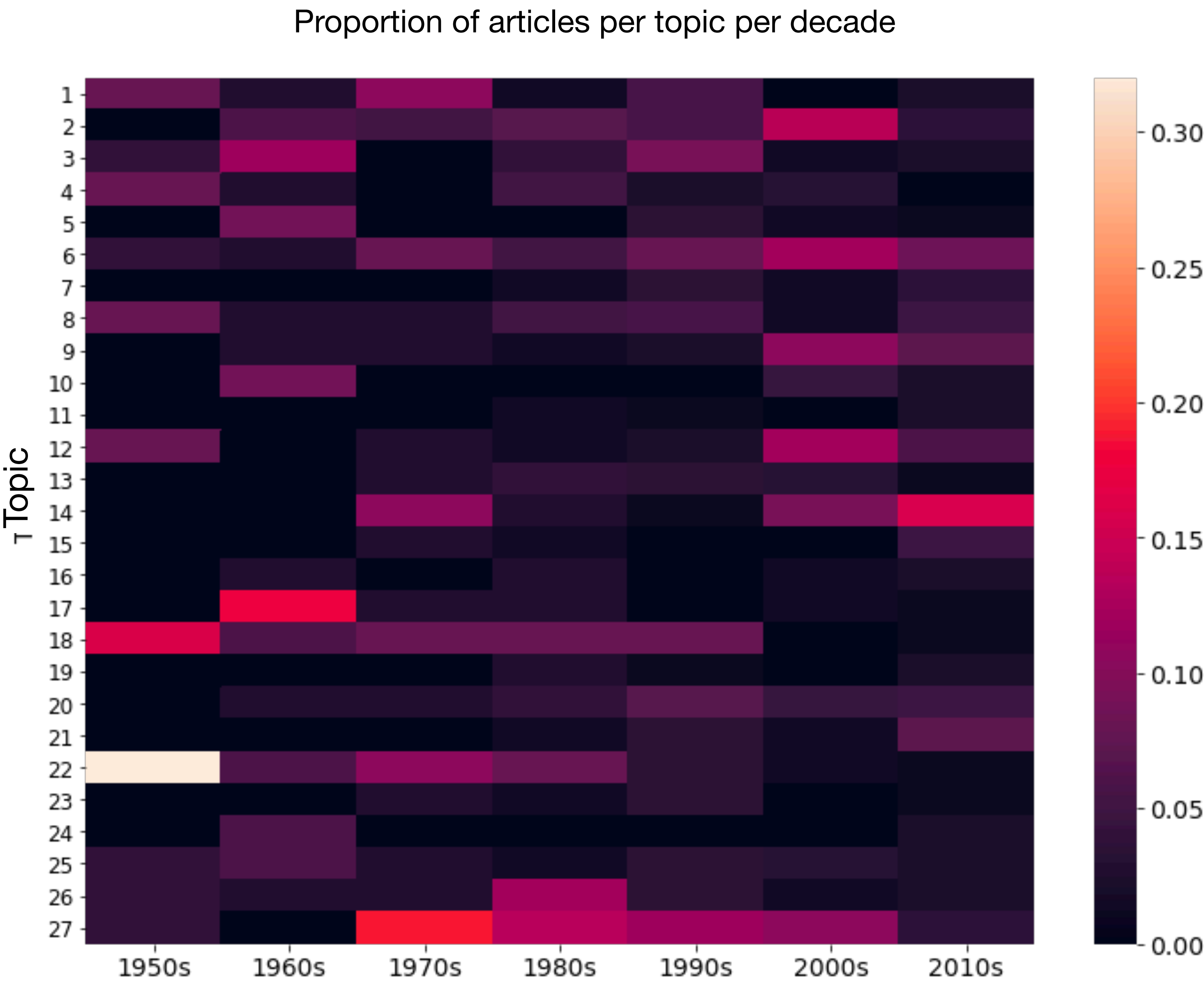| Topic	Top 10 Words	Next 10 Words
1	write, writer, read, life, interest, work, start, idea, word, england	story, experience, live, book, remember, begin, fact, english, child, play
2	poem, write, poet, line, poetry, stanza, lowell, rhyme, frost, long	sense, meter, read, sonnet, sound, love, end, language, begin, teacher
3	play, theater, write, playwright, actor, audience, stage, act, rehearsal, director	production, character, drama, start, broadway, critic, work, playwriting, scene, comedy
4	work, life, world, sense, live, man, experience, human, simply, word	god, form, moment, art, long, question, artist, begin, create, term
5	ginsberg, allen, kerouac, burroughs, francisco, rexroth, miller, publish, paris, laughlin	grove, cop, buddha, york, lsd, jack, publisher, light, bookstore, city
6	story, write, work, short, character, read, writer, fiction, life, long	begin, live, start, love, publish, end, family, friend, reader, chekhov
7	hebrew, jewish, israel, israeli, jew, arab, arabic, jerusalem, palestine, write	language, translation, yiddish, father, egyptian, holocaust, child, al, literature, political
8	yorker, editor, ross, writer, magazine, work, piece, york, write, humor	word, edit, office, manuscript, journalism, hara, publish, esquire, funny, sentence
9	mother, read, father, love, work, live, school, life, sister, die	remember, write, kid, start, job, house, home, money, walk, class
10	translate, translation, english, russian, read, translator, italian, word, language, french	german, russia, book, dialect, work, dictionary, tolstoy, music, scholar, rome
11	film, work, viewer, sequence, movie, edit, shoot, cinema, script, documentary	interest, comedy, read, funny, actor, choice, direct, school, tv, reality
12	book, write, publish, read, writer, character, life, world, reader, editor	publisher, author, review, start, love, city, house, long, war, plot
13	mexico, spanish, write, mexican, latin, writer, cuban, country, revolution, live	borges, spain, city, book, literature, political, reality, life, baroque, freedom
14	write, work, read, book, writer, character, interest, sentence, reader, start	person, experience, fiction, language, voice, part, gay, sense, end, school
15	biography, biographer, james, write, vera, shelley, letter, shaw, archive, coleridge	subject, life, strachey, wharton, volume, biographical, edith, woolf, book, johnson
16	science, fiction, write, world, technology, art, idea, future, read, city	culture, experience, states, book, present, question, television, hero, notion, american
17	pound, ezra, williams, bill, read, ski, eliot, work, harvard, review	write, canto, harry, england, interest, poetry, poet, meet, man, verse
18	writer, write, man, faulkner, read, american, young, critic, literature, hemingway	yiddish, book, literary, life, novelist, joyce, talent, fiction, drink, long
19	irish, ireland, dublin, catholic, nun, child, joyce, read, protestant, begin	northern, sense, english, mary, dickinson, yeats, trinity, church, writer, england
20	german, camp, war, communist, poland, regime, auschwitz, polish, language, nazi	write, nazis, man, prisoner, germany, literature, czech, spanish, village, book
21	black, write, white, book, read, african, baldwin, work, american, jazz	woman, editor, language, slave, person, bill, slavery, music, ellison, important
22	character, novelist, man, work, sartre, life, book, proust, art, technique	balzac, today, literature, influence, moral, modern, french, interest, question, world
23	homer, greek, iliad, odyssey, english, virgil, iris, pope, verse, latin	work, poem, read, translation, johnson, war, hand, word, teach, line
24	comic, cartoonist, draw, strip, art, cartoon, artist, cleveland, lsd, early	start, color, visual, memory, charles, picture, work, paper, funny, medium
25	movie, studio, director, picture, hollywood, screenplay, script, write, work, scene	actor, producer, man, screenwriter, writer, shoot, line, joan, girl, star
26	woman, man, write, love, life, child, book, mother, live, husband	male, father, woolf, girl, read, marry, war, french, feminist, long
27	poetry, poet, poem, read, auden, write, verse, eliot, language, poetic	work, whitman, yeats, publish, american, stevens, lowell, live, form, shakespeare

Topics Analysis: Article Counts

Topic	Top 10 Words
27	poetry, poet, poem, read, auden, write, verse, eliot, language, poetic
6	story, write, work, short, character, read, writer, fiction, life, long
2	poem, write, poet, line, poetry, stanza, lowell, rhyme, frost, long
14	write, work, read, book, writer, character, interest, sentence, reader, start
22	character, novelist, man, work, sartre, life, book, proust, art, technique
18	writer, write, man, faulkner, read, american, young, critic, literature, hemingway
3	play, theater, write, playwright, actor, audience, stage, act, rehearsal, director
12	book, write, publish, read, writer, character, life, world, reader, editor
8	yorker, editor, ross, writer, magazine, work, piece, york, write, humor
9	mother, read, father, love, work, live, school, life, sister, die
20	german, camp, war, communist, poland, regime, auschwitz, polish, language, nazi
26	woman, man, write, love, life, child, book, mother, live, husband
1	write, writer, read, life, interest, work, start, idea, word, england
25	movie, studio, director, picture, hollywood, screenplay, script, write, work, scene
4	work, life, world, sense, live, man, experience, human, simply, word
17	pound, ezra, williams, bill, read, ski, eliot, work, harvard, review
21	black, write, white, book, read, african, baldwin, work, american, jazz
13	mexico, spanish, write, mexican, latin, writer, cuban, country, revolution, live
5	ginsberg, allen, kerouac, burroughs, francisco, rexroth, miller, publish, paris, laughlin
7	hebrew, jewish, israel, israeli, jew, arab, arabic, jerusalem, palestine, write
10	translate, translation, english, russian, read, translator, italian, word, language, french
15	biography, biographer, james, write, vera, shelley, letter, shaw, archive, coleridge
16	science, fiction, write, world, technology, art, idea, future, read, city
23	homer, greek, iliad, odyssey, english, virgil, iris, pope, verse, latin
19	irish, ireland, dublin, catholic, nun, child, joyce, read, protestant, begin
11	film, work, viewer, sequence, movie, edit, shoot, cinema, script, documentary
24	comic, cartoonist, draw, strip, art, cartoon, artist, cleveland, lsd, early

Number of articles per topic

Topics Analysis: Date Ranges

Topic	Year Start	Year End	Span
1	1958	2018	60
2	1961	2019	58
3	1956	2012	56
4	1955	2007	52
5	1966	2019	53
6	1957	2019	62
7	1984	2017	33
8	1955	2019	64
9	1964	2018	54
10	1964	2019	55
11	1989	2018	29
12	1955	2016	61
13	1971	2018	47
14	1974	2018	44
15	1977	2017	40
16	1961	2013	52
17	1960	2017	57
18	1953	2017	64
19	1984	2019	35
20	1969	2018	49
21	1984	2018	34
22	1953	2010	57
23	1970	2016	46
24	1962	2014	52
25	1956	2014	58
26	1956	2018	62
27	1959	2019	60

Proportion of articles per topic per decade

Topics Analysis: Exemplar Articles
Articles with highest similarity to topic centroids

Topic	Topic Top 10 Words	Author	Year	Article
1	write, writer, read, life, interest, work, start, idea, word, england	kadare ismail	1998	1990/1998_kadare_ismail.txt
2	poem, write, poet, line, poetry, stanza, lowell, rhyme, frost, long	collins billy	2001	2000/2001_collins_billy.txt
3	play, theater, write, playwright, actor, audience, stage, act, rehearsal, director	guare john	1992	1990/1992_guare_john.txt
4	work, life, world, sense, live, man, experience, human, simply, word	winterson jeanette	1997	1990/1997_winterson_jeanette.txt
5	ginsberg, allen, kerouac, burroughs, francisco, rexroth, miller, publish, paris, laughlin	kerouac jack	1968	1960/1968_kerouac_jack.txt
6	story, write, work, short, character, read, writer, fiction, life, long	munro alice	1994	1990/1994_munro_alice.txt
7	hebrew, jewish, israel, israeli, jew, arab, arabic, jerusalem, palestine, write	amichai yehuda	1992	1990/1992_amichai_yehuda.txt
8	yorker, editor, ross, writer, magazine, work, piece, york, write, humor	maxwell william	1982	1980/1982_maxwell_william.txt
9	mother, read, father, love, work, live, school, life, sister, die	erdrich louise	2010	2010/2010_erdrich_louise.txt
10	translate, translation, english, russian, read, translator, italian, word, language, french	pevear richard	2015	2010/2015_pevear_richard.txt
11	film, work, viewer, sequence, movie, edit, shoot, cinema, script, documentary	haneke michael	2014	2010/2014_haneke_michael.txt
12	book, write, publish, read, writer, character, life, world, reader, editor	rushdie salman	2005	2000/2005_rushdie_salman.txt
13	mexico, spanish, write, mexican, latin, writer, cuban, country, revolution, live	márquez gabriel	1981	1980/1981_márquez_gabriel.txt
14	write, work, read, book, writer, character, interest, sentence, reader, start	barrett andrea	2003	2000/2003_barrett_andrea.txt
15	biography, biographer, james, write, vera, shelley, letter, shaw, archive, coleridge	lee hermione	2013	2010/2013_lee_hermione.txt
16	science, fiction, write, world, technology, art, idea, future, read, city	guin ursula	2013	2010/2013_guin_ursula.txt
17	pound, ezra, williams, bill, read, ski, eliot, work, harvard, review	laughlin james part 2	1983	1980/1983_laughlin_james_part_2.txt
18	writer, write, man, faulkner, read, american, young, critic, literature, hemingway	vidal gore	1974	1970/1974_vidal_gore.txt
19	irish, ireland, dublin, catholic, nun, child, joyce, read, protestant, begin	trevor william	1989	1980/1989_trevor_william.txt
20	german, camp, war, communist, poland, regime, auschwitz, polish, language, nazi	steiner george	1995	1990/1995_steiner_george.txt
21	black, write, white, book, read, african, baldwin, work, american, jazz	morrison toni	1993	1990/1993_morrison_toni.txt
22	character, novelist, man, work, sartre, life, book, proust, art, technique	gardner john	1979	1970/1979_gardner_john.txt
23	homer, greek, iliad, odyssey, english, virgil, iris, pope, verse, latin	logue christopher	1993	1990/1993_logue_christopher.txt
24	comic, cartoonist, draw, strip, art, cartoon, artist, cleveland, lsd, early	ware chris	2014	2010/2014_ware_chris.txt
25	movie, studio, director, picture, hollywood, screenplay, script, write, work, scene	price richard	1996	1990/1996_price_richard.txt
26	woman, man, write, love, life, child, book, mother, live, husband	gordimer nadine	1983	1980/1983_gordimer_nadine.txt
27	poetry, poet, poem, read, auden, write, verse, eliot, language, poetic	ashbery john	1983	1980/1983_ashbery_john.txt

Topics Analysis: Coherence
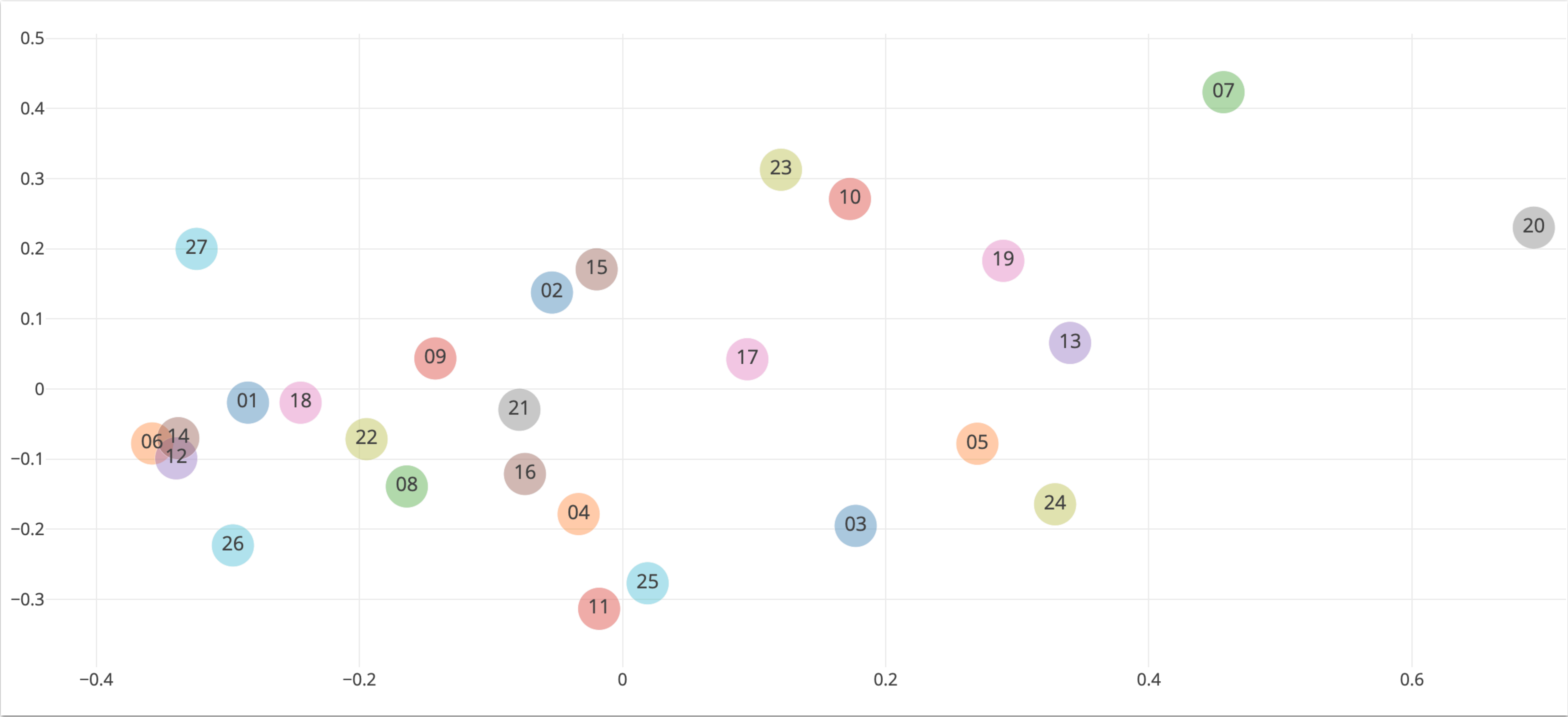
Ranking of topics using *min-max* score

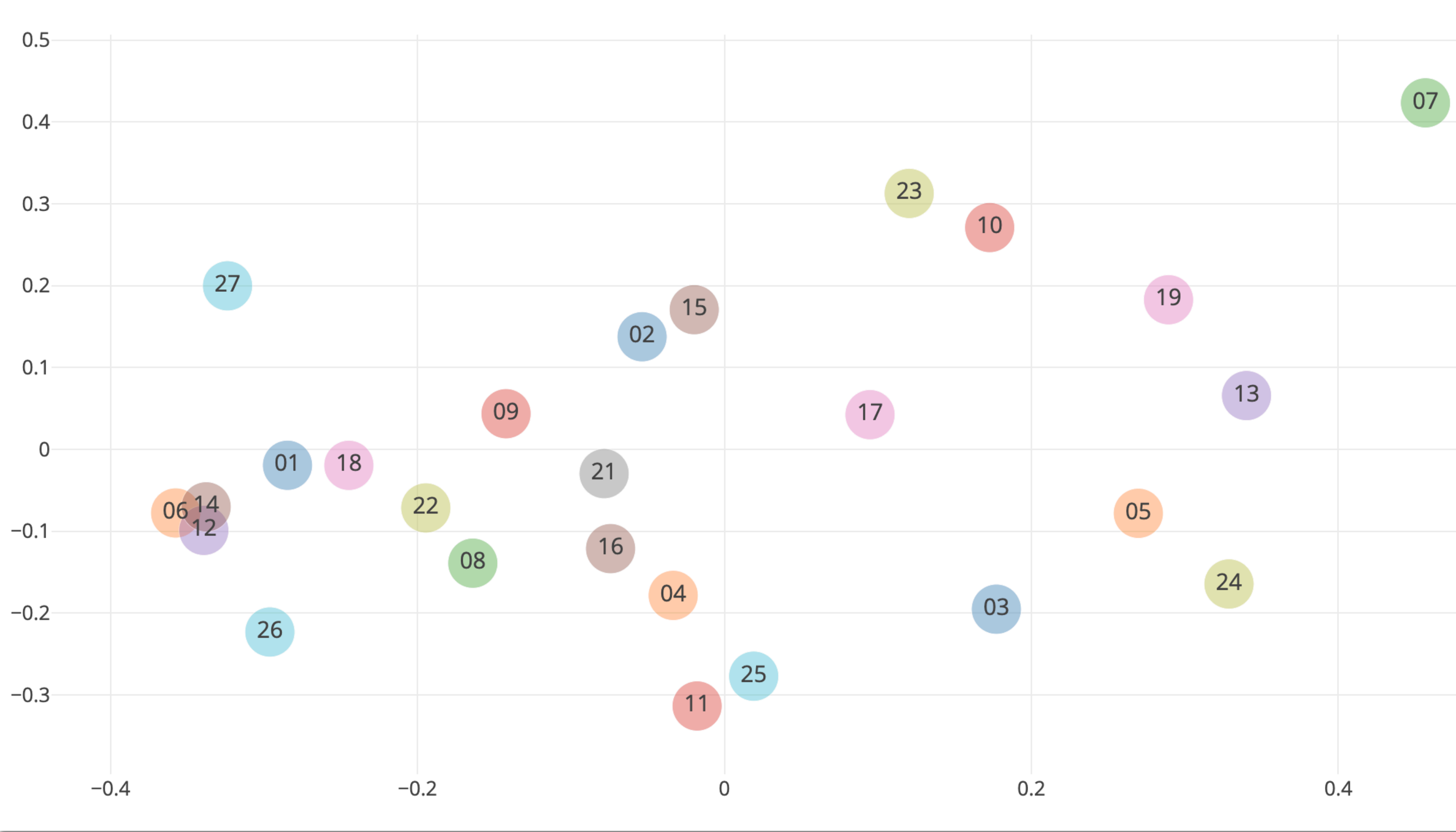| Topic	Topic Top 10 Words	Score
20	german, camp, war, communist, poland, regime, auschwitz, polish, language, nazi	0.785
7	hebrew, jewish, israel, israeli, jew, arab, arabic, jerusalem, palestine, write	0.796
10	translate, translation, english, russian, read, translator, italian, word, language, french	1.294
19	irish, ireland, dublin, catholic, nun, child, joyce, read, protestant, begin	1.302
27	poetry, poet, poem, read, auden, write, verse, eliot, language, poetic	1.325
5	ginsberg, allen, kerouac, burroughs, francisco, rexroth, miller, publish, paris, laughlin	1.337
11	film, work, viewer, sequence, movie, edit, shoot, cinema, script, documentary	1.352
23	homer, greek, iliad, odyssey, english, virgil, iris, pope, verse, latin	1.357
13	mexico, spanish, write, mexican, latin, writer, cuban, country, revolution, live	1.396
3	play, theater, write, playwright, actor, audience, stage, act, rehearsal, director	1.404
26	woman, man, write, love, life, child, book, mother, live, husband	1.429
24	comic, cartoonist, draw, strip, art, cartoon, artist, cleveland, lsd, early	1.449
4	work, life, world, sense, live, man, experience, human, simply, word	1.473
25	movie, studio, director, picture, hollywood, screenplay, script, write, work, scene	1.495
9	mother, read, father, love, work, live, school, life, sister, die	1.567
16	science, fiction, write, world, technology, art, idea, future, read, city	1.589
15	biography, biographer, james, write, vera, shelley, letter, shaw, archive, coleridge	1.590
2	poem, write, poet, line, poetry, stanza, lowell, rhyme, frost, long	1.614
17	pound, ezra, williams, bill, read, ski, eliot, work, harvard, review	1.629
8	yorker, editor, ross, writer, magazine, work, piece, york, write, humor	1.644
12	book, write, publish, read, writer, character, life, world, reader, editor	1.649
22	character, novelist, man, work, sartre, life, book, proust, art, technique	1.677
6	story, write, work, short, character, read, writer, fiction, life, long	1.691
21	black, write, white, book, read, african, baldwin, work, american, jazz	1.733
18	writer, write, man, faulkner, read, american, young, critic, literature, hemingway	1.734
1	write, writer, read, life, interest, work, start, idea, word, england	1.744
14	write, work, read, book, writer, character, interest, sentence, reader, start	1.750

Topics Analysis: Coherence

Ranking of topics using *embedding coherence* score

Topic	Topic Top 10 Words	Score
27	poetry, poet, poem, read, auden, write, verse, eliot, language, poetic	0.473
7	hebrew, jewish, israel, israeli, jew, arab, arabic, jerusalem, palestine, write	0.459
26	woman, man, write, love, life, child, book, mother, live, husband	0.441
11	film, work, viewer, sequence, movie, edit, shoot, cinema, script, documentary	0.388
12	book, write, publish, read, writer, character, life, world, reader, editor	0.381
6	story, write, work, short, character, read, writer, fiction, life, long	0.374
14	write, work, read, book, writer, character, interest, sentence, reader, start	0.354
4	work, life, world, sense, live, man, experience, human, simply, word	0.350
20	german, camp, war, communist, poland, regime, auschwitz, polish, language, nazi	0.339
25	movie, studio, director, picture, hollywood, screenplay, script, write, work, scene	0.338
8	yorker, editor, ross, writer, magazine, work, piece, york, write, humor	0.337
1	write, writer, read, life, interest, work, start, idea, word, england	0.337
18	writer, write, man, faulkner, read, american, young, critic, literature, hemingway	0.334
9	mother, read, father, love, work, live, school, life, sister, die	0.331
16	science, fiction, write, world, technology, art, idea, future, read, city	0.330
22	character, novelist, man, work, sartre, life, book, proust, art, technique	0.328
10	translate, translation, english, russian, read, translator, italian, word, language, french	0.326
3	play, theater, write, playwright, actor, audience, stage, act, rehearsal, director	0.306
5	ginsberg, allen, kerouac, burroughs, francisco, rexroth, miller, publish, paris, laughlin	0.305
21	black, write, white, book, read, african, baldwin, work, american, jazz	0.300
2	poem, write, poet, line, poetry, stanza, lowell, rhyme, frost, long	0.289
23	homer, greek, iliad, odyssey, english, virgil, iris, pope, verse, latin	0.282
15	biography, biographer, james, write, vera, shelley, letter, shaw, archive, coleridge	0.281
19	irish, ireland, dublin, catholic, nun, child, joyce, read, protestant, begin	0.277
13	mexico, spanish, write, mexican, latin, writer, cuban, country, revolution, live	0.272
17	pound, ezra, williams, bill, read, ski, eliot, work, harvard, review	0.263
24	comic, cartoonist, draw, strip, art, cartoon, artist, cleveland, lsd, early	0.230

Topics Analysis: Topic-Level MDS Plot

Topics Analysis: Topic-Level MDS Plot

This scatter plot shows inter-topics distances, visualised in 2 dimensions using Multidimensional scaling (MDS), for all 27 topics in the current model. The distance between a pair of topics is calculated as the average distance between the terms appearing in their descriptors, where distances are calculated relative to a *word2vec* word embedding constructed on a dump of Wikipedia abstracts.

Topic	Top 10 Words
1	write, writer, read, life, interest, work, start, idea, word, england
2	poem, write, poet, line, poetry, stanza, lowell, rhyme, frost, long
3	play, theater, write, playwright, actor, audience, stage, act, rehearsal, director
4	work, life, world, sense, live, man, experience, human, simply, word
5	ginsberg, allen, kerouac, burroughs, francisco, rexroth, miller, publish, paris, laughlin
6	story, write, work, short, character, read, writer, fiction, life, long
7	hebrew, jewish, israel, israeli, jew, arab, arabic, jerusalem, palestine, write
8	yorker, editor, ross, writer, magazine, work, piece, york, write, humor
9	mother, read, father, love, work, live, school, life, sister, die
10	translate, translation, english, russian, read, translator, italian, word, language, french
11	film, work, viewer, sequence, movie, edit, shoot, cinema, script, documentary
12	book, write, publish, read, writer, character, life, world, reader, editor
13	mexico, spanish, write, mexican, latin, writer, cuban, country, revolution, live
14	write, work, read, book, writer, character, interest, sentence, reader, start
15	biography, biographer, james, write, vera, shelley, letter, shaw, archive, coleridge
16	science, fiction, write, world, technology, art, idea, future, read, city
17	pound, ezra, williams, bill, read, ski, eliot, work, harvard, review
18	writer, write, man, faulkner, read, american, young, critic, literature, hemingway
19	irish, ireland, dublin, catholic, nun, child, joyce, read, protestant, begin
20	german, camp, war, communist, poland, regime, auschwitz, polish, language, nazi
21	black, write, white, book, read, african, baldwin, work, american, jazz
22	character, novelist, man, work, sartre, life, book, proust, art, technique
23	homer, greek, iliad, odyssey, english, virgil, iris, pope, verse, latin
24	comic, cartoonist, draw, strip, art, cartoon, artist, cleveland, lsd, early
25	movie, studio, director, picture, hollywood, screenplay, script, write, work, scene
26	woman, man, write, love, life, child, book, mother, live, husband
27	poetry, poet, poem, read, auden, write, verse, eliot, language, poetic

Topics Analysis: Term-Level MDS Plot

This scatter plot shows the distances between all pairs of terms which appear in the 27 topic descriptors in the current model, visualised in 2 dimensions using *Multidimensional scaling* (MDS).