

Text Stats from KSU Web Crawler

doc_len: 681.769

emails:

('registrar@kennesaw.edu', 45)
('jgraise1@kennesaw.edu', 40)
('ksuinfo@kennesaw.edu', 33)
('educ_grad@kennesaw.edu', 32)
('tuitionclassification@kennesaw.edu', 23)
('lhaskett@kennesaw.edu', 23)
('cburto32@kennesaw.edu', 23)
('bcruz3@kennesaw.edu', 20)
('globalsafety@kennesaw.edu', 18)
('iet@kennesaw.edu', 18)

perc: 0.294

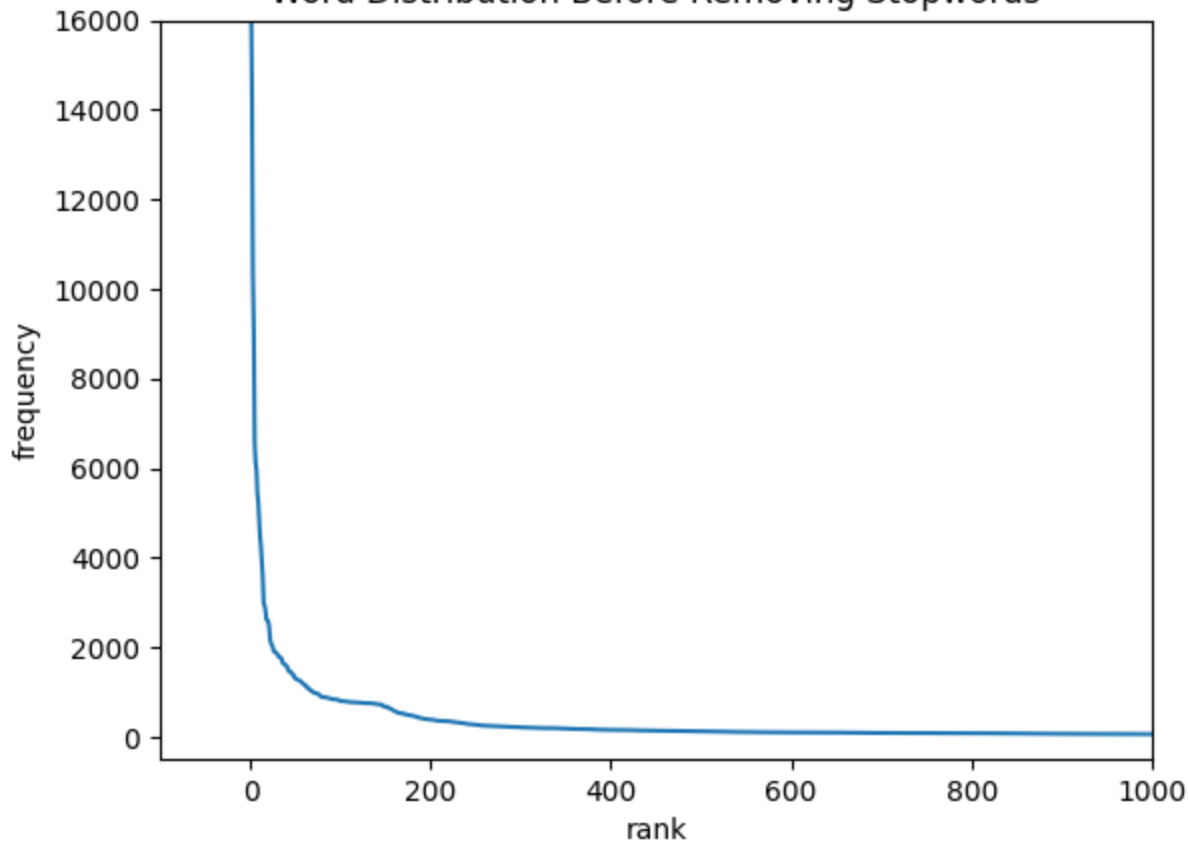
With Stopwords:

rank	term	freq.	perc.
1	.	20398	0.038
2	and	16590	0.031
3	the	14381	0.027
4	of	10395	0.019
5	to	9082	0.017
6	&	6554	0.012
7	for	6127	0.011
8	/	5960	0.011
9	in	5446	0.010
10	a	5237	0.010
11	kennesaw	4803	0.009
12	campus	4459	0.008
13	ksu	4239	0.008
14	university	3932	0.007
15	students	3550	0.007
16	state	2975	0.006
17	on	2935	0.006
18	student	2820	0.005
19	you	2617	0.005
20	search	2608	0.005
21	is	2573	0.005
22	or	2474	0.005
23	with	2201	0.004
24	information	2093	0.004
25	-	2057	0.004
26	community	2015	0.004
27	this	1919	0.004
28	faculty	1902	0.004
29	contact	1897	0.004
30	marietta	1875	0.004

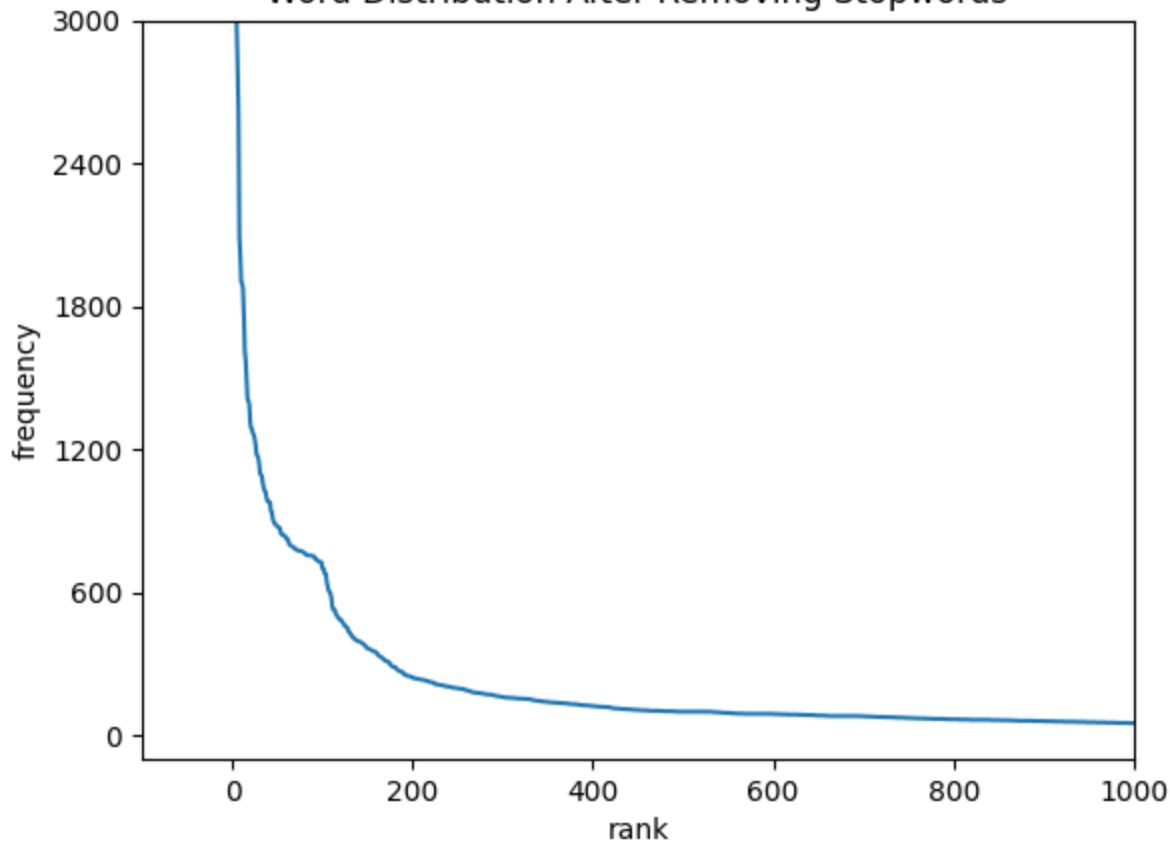
Without Stopwords:

rank	term	freq.	perc.
-----	-----	-----	-----
1	kennesaw	4803	0.009
2	campus	4459	0.008
3	ksu	4239	0.008
4	university	3932	0.007
5	students	3550	0.007
6	state	2975	0.006
7	student	2820	0.005
8	search	2608	0.005
9	information	2093	0.004
10	community	2015	0.004
11	faculty	1902	0.004
12	contact	1897	0.004
13	marietta	1875	0.004
14	global	1763	0.003
15	ga	1613	0.003
16	financial	1575	0.003
17	research	1496	0.003
18	program	1413	0.003
19	diversity	1403	0.003
20	programs	1389	0.003
21	academic	1315	0.002
22	resources	1289	0.002
23	online	1284	0.002
24	aid	1267	0.002
25	registration	1262	0.002
26	staff	1244	0.002
27	visit	1220	0.002
28	human	1180	0.002
29	learning	1174	0.002
30	college	1163	0.002

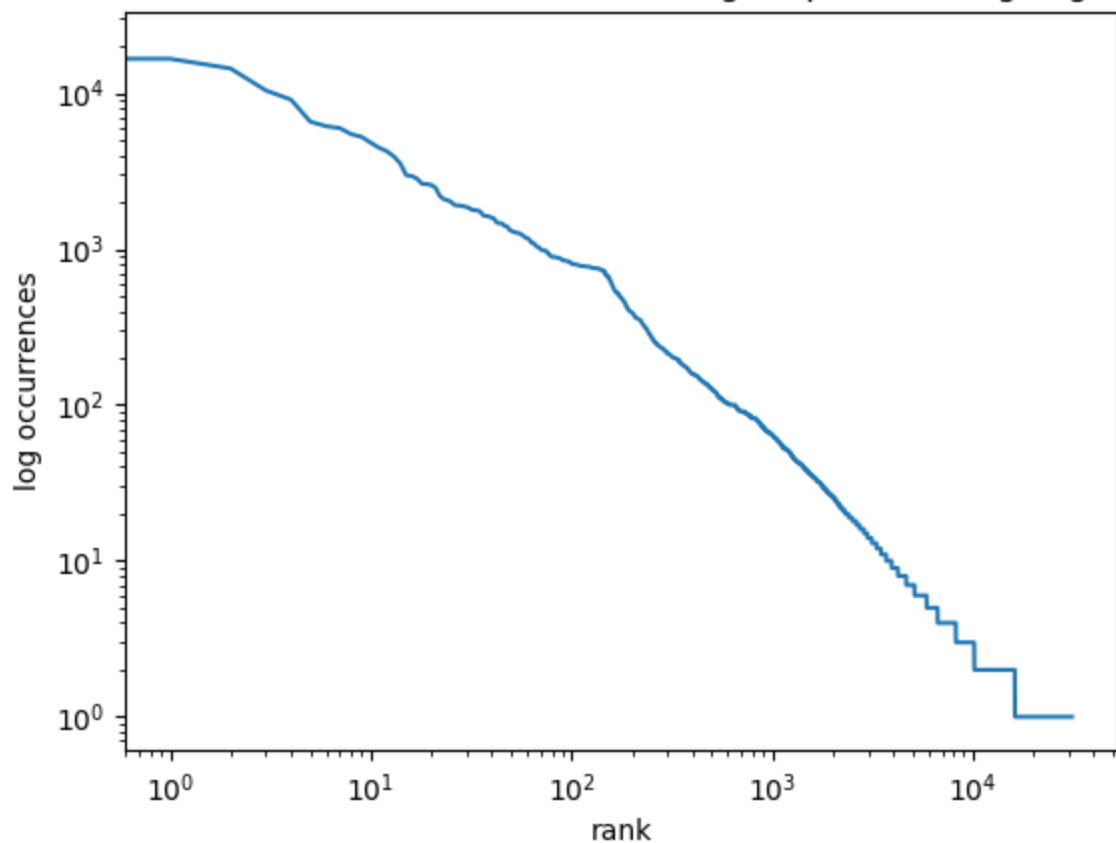
Word Distribution Before Removing Stopwords



Word Distribution After Removing Stopwords



Word Distribution Before Removing Stopwords - Log-Log



Word Distribution After Removing Stopwords - Log-Log

