Sample Queries from moodle for Project 2 of COMP 479 by Silas Kalinowski ID: 40256077

Query 'copper' on uncompressed index:

104 documents contain the word 'copper': ['22', '793', '800', '816', '1148', '1184', '1552', '1607', '2006', '2074', '2186', '2764', '2782', '2880', '3454', '3613', '3862', '4291', '4431', '4744', '4816', '5203', '5209', '5435', '5788', '5827', '5888', '6025', '6225', '6395', '6846', '6927', '7552', '7724', '7775', '8569', '8601', '11273', '11945', '12024', '12215', '12223', '12243', '12489', '12513', '12633', '12641', '12651', '12683', '12692', '12725', '12857', '12861', '12910', '12926', '12963', '12980', '12992', '13003', '13670', '13675', '13694', '13877', '13897', '14274', '14297', '14302', '14398', '14476', '14499', '14572', '14687', '14805', '14852', '15166', '15264', '15556', '15831', '15932', '15988', '16123', '16971', '17007', '17118', '17714', '18225', '18280', '18317', '18339', '18392', '18394', '18430', '18457', '18631', '18691', '18758', '18849', '18952', '19502', '19781', '19786', '19878', '21293', '21327']

Query 'copper' on compressed index:

116 documents contain the word 'copper': ['22', '793', '800', '816', '922', '1148', '1184', '1552', '1607', '2006', '2074', '2186', '2764', '2782', '2880', '3454', '3613', '3862', '4291', '4373', '4431', '4744', '4816', '5203', '5209', '5435', '5788', '5827', '5888', '6025', '6225', '6395', '6846', '6927', '7552', '7724', '7775', '8569', '8601', '11053', '11273', '11945', '12024', '12215', '12223', '12243', '12484', '12489', '12513', '12633', '12641', '12651', '12683', '12692', '12725', '12857', '12861', '12910', '12926', '12963', '12980', '12992', '13003', '13465', '13495', '13608', '13665', '13670', '13675', '13687', '13694', '13757', '13877', '13897', '14136', '14274', '14297', '14302', '14398', '14476', '14499', '14572', '14687', '14805', '14852', '15166', '15264', '15556', '15831', '15932', '15988', '16110', '16123', '16971', '17007', '17118', '17714', '18225', '18280', '18317', '18339', '18392', '18394', '18430', '18457', '18631', '18691', '18758', '18849', '18952', '19502', '19781', '19786', '19878', '21293', '21327']

Explanation: The reason why the compressed index returns more documents could be that the documents that contain the word 'Copper' are new also returned, because of case folding. Also, some words could be stemmed to copper.

Query 'Samjens' on uncompressed index:

4 documents contained the term 'Samjens': ['17837', '17863', '18071', '19419']

Query 'Samjens' on compressed index:

No document contains the term 'Samjens'.

Explanation: On the compressed index no document is returned. That can be explained by the fact that because of case folding every term in the dictionary is now lowercase. Therefore, every capitalized word does not return any document.

Query 'Carmark' on uncompressed index:

One document contains the term 'Carmark': ['19758']

Query 'Carmark' on compressed index:

No document contains the term 'Carmark'.

Explanation: For the same reason as mentioned for 'Samjens' no document is returned on the compressed index.

Query 'Bundesbank' on uncompressed index:

144 documents contain the word 'Bundesbank': ['278', '926', '942', '950', '1540', '1959', '1971', '2110', '2197', '2662', '2686', '2979', '3020', '3514', '4113', '4297', '4830', '4873', '5176', '5201', '5241', '5253', '5290', '5371', '6108', '6426', '6446', '6453', '7031', '7051', '7310', '7555', '7767', '8070', '8137', '8145', '8444', '8677', '8714', '9295', '9787', '9946', '9975', '10344', '10359', '10382', '10697', '10741', '10751', '10765', '11212', '11234', '11238', '11776', '11818', '11820', '11900', '12084', '12447', '12456', '12471', '12780', '12801', '12875', '13244', '13379', '13404', '13461', '13471', '13496', '13512', '13531', '13571', '13577', '13629', '13949', '14109', '14770', '14849', '14931', '14956', '14987', '15048', '15253', '15364', '15384', '15444', '15539', '15625', '15763', '16128', '16268', '16942', '16976', '16977', '17064', '17065', '17088', '17119', '17139', '17193', '17210', '17230', '17242', '17243', '17246', '17248', '17249', '17259', '17265', '17272', '17306', '17396', '17448', '17598', '17620', '17821', '17883', '17899', '17906', '17915', '18086', '18090', '18095', '18250', '18299', '18452', '19557', '20038', '20071', '20080', '20158', '20377', '20447', '20500', '20764', '20868', '20893', '20907', '20925', '21277', '21280', '21422', '21508']

Query 'Bundesbank' on compressed index:

No document contains the term 'Bundesbank'.

For the same reason as mentioned for 'Samjens' no document is returned on the compressed index.

I certify that this submission is my original work and meets the Faculty's Expectations of Originality.

Signature   ID: 40256077   Date 2022-10-08