Contemporary Data Processing Technology (CCOD)

Lab 10 (October 28, 2016)

Kolesnikov Dmitry, AS-36

In this work I took the 10 Russian letters: **й,ц,у,к,е,н,г,ш,щ,з.**

$\alpha = 0.5$

Table of similarity percentage of letters

	й	ц	у	к	е	н	г	ш	щ	з	
й	1	0.6	0.5	0.3	0.1	0.8	0.2	0.3	0.3	0.1	
ц	0.6	1	0.2	0.1	0.3	0.6	0.6	0.8	0.9	0.1	
у	0.5	0.2	1	0.4	0.1	0.1	0.1	0.2	0.2	0.1	
к	0.3	0.1	0.4	1	0.1	0.7	0.4	0.2	0.2	0.1	
е	0.1	0.3	0.1	0.1	1	0.2	0.4	0.1	0.1	0.5	R0
н	0.8	0.6	0.1	0.7	0.2	1	0.4	0.3	0.3	0.1	
г	0.2	0.6	0.1	0.4	0.4	0.4	1	0.6	0.6	0.2	
ш	0.3	0.8	0.2	0.2	0.1	0.3	0.6	1	0.9	0.3	
щ	0.3	0.9	0.2	0.2	0.1	0.3	0.6	0.9	1	0.6	
з	0.1	0.1	0.1	0.1	0.5	0.1	0.2	0.3	0.6	1	

Four steps of modifying this table:

	й	ц	у	к	е	н	г	ш	щ	з	
й	1	0.6	0.5	0.7	0.3	0.8	0.6	0.6	0.6	0.3	
ц	0.6	1	0.5	0.6	0.4	0.6	0.6	0.9	0.9	0.6	
у	0.5	0.5	1	0.4	0.2	0.5	0.4	0.3	0.3	0.2	
к	0.7	0.6	0.4	1	0.4	0.7	0.4	0.4	0.4	0.2	
е	0.3	0.4	0.2	0.4	1	0.4	0.4	0.4	0.5	0.5	R1
н	0.8	0.6	0.5	0.7	0.4	1	0.6	0.6	0.6	0.3	
г	0.6	0.6	0.4	0.4	0.4	0.6	1	0.6	0.6	0.6	
ш	0.6	0.9	0.3	0.4	0.4	0.6	0.6	1	0.9	0.6	
щ	0.6	0.9	0.3	0.4	0.5	0.6	0.6	0.9	1	0.6	
з	0.3	0.6	0.2	0.2	0.5	0.3	0.6	0.6	0.6	1	

	й	ц	у	к	е	н	г	ш	щ	з	
й	1	0.6	0.5	0.7	0.5	0.8	0.6	0.6	0.6	0.6	
ц	0.6	1	0.5	0.6	0.5	0.6	0.6	0.9	0.9	0.6	
у	0.5	0.5	1	0.5	0.4	0.5	0.5	0.5	0.5	0.5	
к	0.7	0.6	0.5	1	0.4	0.7	0.6	0.6	0.6	0.6	
е	0.5	0.5	0.4	0.4	1	0.5	0.5	0.5	0.5	0.5	R2
н	0.8	0.6	0.5	0.7	0.5	1	0.6	0.6	0.6	0.6	
г	0.6	0.6	0.5	0.6	0.5	0.6	1	0.6	0.6	0.6	
ш	0.6	0.9	0.5	0.6	0.5	0.6	0.6	1	0.9	0.6	
щ	0.6	0.9	0.5	0.6	0.5	0.6	0.6	0.9	1	0.6	
з	0.6	0.6	0.5	0.6	0.5	0.6	0.6	0.6	0.6	1	

	й	ц	у	к	е	н	г	ш	щ	з	
й	1	0.6	0.5	0.7	0.5	0.8	0.6	0.6	0.6	0.6	
ц	0.6	1	0.5	0.6	0.5	0.6	0.6	0.9	0.9	0.6	
у	0.5	0.5	1	0.5	0.5	0.5	0.5	0.5	0.5	0.5	
к	0.7	0.6	0.5	1	0.5	0.7	0.6	0.6	0.6	0.6	
е	0.5	0.5	0.5	0.5	1	0.5	0.5	0.5	0.5	0.5	R3
н	0.8	0.6	0.5	0.7	0.5	1	0.6	0.6	0.6	0.6	
г	0.6	0.6	0.5	0.6	0.5	0.6	1	0.6	0.6	0.6	
ш	0.6	0.9	0.5	0.6	0.5	0.6	0.6	1	0.9	0.6	
щ	0.6	0.9	0.5	0.6	0.5	0.6	0.6	0.9	1	0.6	
з	0.6	0.6	0.5	0.6	0.5	0.6	0.6	0.6	0.6	1	

	й	ц	у	к	е	н	г	ш	щ	з	
й	1	0.6	0.5	0.7	0.5	0.8	0.6	0.6	0.6	0.6	
ц	0.6	1	0.5	0.6	0.5	0.6	0.6	0.9	0.9	0.6	
у	0.5	0.5	1	0.5	0.5	0.5	0.5	0.5	0.5	0.5	
к	0.7	0.6	0.5	1	0.5	0.7	0.6	0.6	0.6	0.6	
е	0.5	0.5	0.5	0.5	1	0.5	0.5	0.5	0.5	0.5	R4
н	0.8	0.6	0.5	0.7	0.5	1	0.6	0.6	0.6	0.6	
г	0.6	0.6	0.5	0.6	0.5	0.6	1	0.6	0.6	0.6	
ш	0.6	0.9	0.5	0.6	0.5	0.6	0.6	1	0.9	0.6	
щ	0.6	0.9	0.5	0.6	0.5	0.6	0.6	0.9	1	0.6	
з	0.6	0.6	0.5	0.6	0.5	0.6	0.6	0.6	0.6	1	

In the final table, where all values less than α and main diagonal will become zeros.

	й	ц	у	к	е	н	г	ш	щ	з	
й	0	0.6	0	0.7	0	0.8	0.6	0.6	0.6	0.6	
ц	0.6	0	0	0.6	0	0.6	0.6	0.9	0.9	0.6	
у	0	0	0	0	0	0	0	0	0	0	
к	0.7	0.6	0	0	0	0.7	0.6	0.6	0.6	0.6	
е	0	0	0	0	0	0	0	0	0	0	Final
н	0.8	0.6	0	0.7	0	0	0.6	0.6	0.6	0.6	
г	0.6	0.6	0	0.6	0	0.6	0	0.6	0.6	0.6	
ш	0.6	0.9	0	0.6	0	0.6	0.6	0	0.9	0.6	
щ	0.6	0.9	0	0.6	0	0.6	0.6	0.9	0	0.6	
з	0.6	0.6	0	0.6	0	0.6	0.6	0.6	0.6	0	

1) $I = \{1, 2, 3, 4, 5, 6, 7, 8, 9, 10\}$ $C_1 = \{\}$

$a_{28} = a_{82} = 0.9$ $C_1 = \{2, 8\}$

$Max = sum_9 \rightarrow \quad C_1 = \{2, 8, 9\}$

$Max = sum_1 \rightarrow \quad C_1 = \{2, 8, 9, 1\}$

$Max = sum_6 \rightarrow \quad C_1 = \{2, 8, 9, 1, 6\}$

$Max = sum_4 \rightarrow \quad C_1 = \{2, 8, 9, 1, 6, 4\}$

$Max = sum_7 \rightarrow \quad C_1 = \{2, 8, 9, 1, 6, 4, 7\}$

$$Max = sum_{10} \rightarrow \quad C_1 = \{2, 8, 9,1,6,4,7,10\}$$

	й	ц	у	к	е	н	г	ш	щ	з	
й	0	0.6	0	0.7	0	0.8	0.6	0.6	0.6	0.6	
ц	0.6	0	0	0.6	0	0.6	0.6	0.9	0.9	0.6	
у	0	0	0	0	0	0	0	0	0	0	Final
к	0.7	0.6	0	0	0	0.7	0.6	0.6	0.6	0.6	
е	0	0	0	0	0	0	0	0	0	0	
н	0.8	0.6	0	0.7	0	0	0.6	0.6	0.6	0.6	
г	0.6	0.6	0	0.6	0	0.6	0	0.6	0.6	0.6	
ш	0.6	0.9	0	0.6	0	0.6	0.6	0	0.9	0.6	
щ	0.6	0.9	0	0.6	0	0.6	0.6	0.9	0	0.6	
з	0.6	0.6	0	0.6	0	0.6	0.6	0.6	0.6	0	

2) $I = \{3,5\}$ $C_2 = \{3\}$

3) $I = \{5\}$ $C_3 = \{5\}$

$C = \{\{2, 8, 9,1,6,4,7,10\},\{3\},\{5\}\}$

OR

$C = \{\{ц,ш,щ,й,н,к,г,з\},\{у\},\{е\}\}$