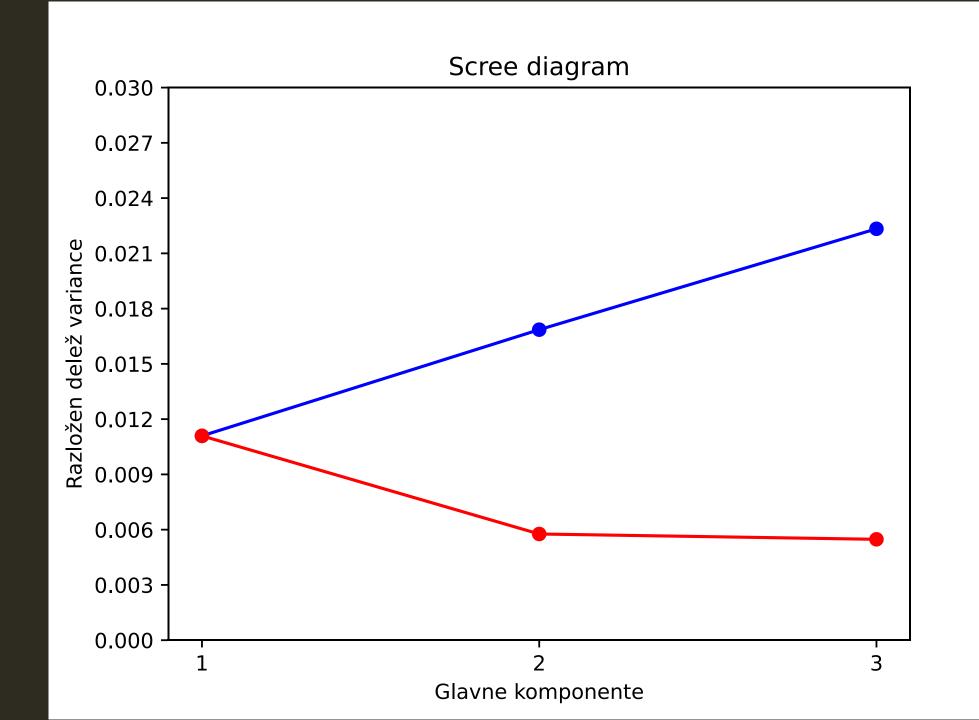
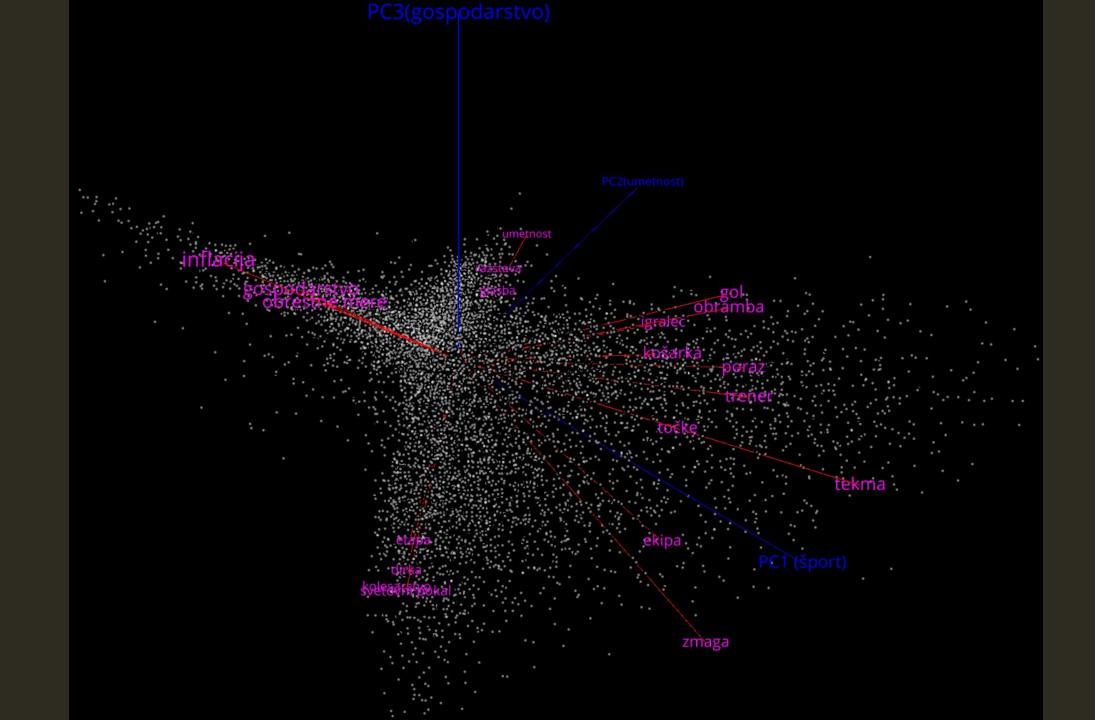
OBDELAVA PODATKOV

'košarka'	'nba'	'video'	'denver'
0.157	0.204	0.285	0.272
0.000	0.000	0.000	0.000
0.000	0.000	0.000	0.000
0.000	0.000	0.000	0.000
0.000	0.000	0.000	0.000

- lower(), unique()
- tf-idf pred rezanjem besed
- članki brez ključnih besed
- redka matrika, a dovolj majhna
- pickle





INTERPRETACIJA

Izbranih 20 z največjo normo.

Pojavijo se tudi v glavnih komponentah samih.

Najbolj zaznamovane besede v glavnih komponentah:

- 1. 'tekma', 'zmaga', 'trener', 'ekipa'
- 2. 'umetnost', 'razstava', 'glasba', 'festival'
- 3. 'inflacija', 'gospodarstvo', 'obrestne mere', 'rast'

Glavne komponente očitno sledijo temam športa, umetnosti in gospodarstva.

Šport se deli na kolesarstvo in ostalo.

Kolesarstvo negativni koeficient v glavni komponenti gospodarstva.