

PODATKI

- Samo podatki iz finalov
- Načini izluščevanja profilov:
 - Po glasujočih državah zlaganje glasov vsakega leta v en dolg vektor
 - Za vsa leta
 - Od 2000 do 2019
 - Od 2000 do 2008
 - Po glasujočih državah povprečje glasov vsaki državi za vsa leta

```
prep3years = range(2000,2020)
prep3orig = prep3orig[prep3orig["Year"] >= 2000]
prep3orig = prep3orig[prep3orig["Year"] <= 2019]
```

PARAMETRI RAZVRŠČANJA

Metrika razdalje: Euklidska razdalja

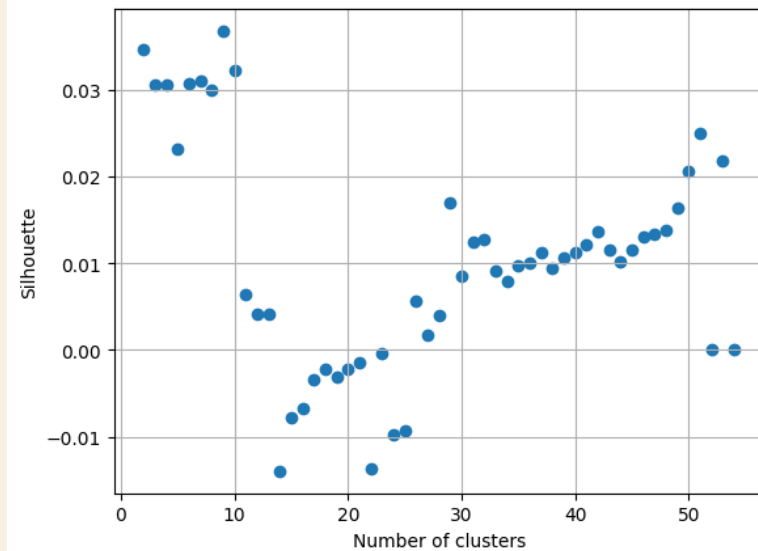
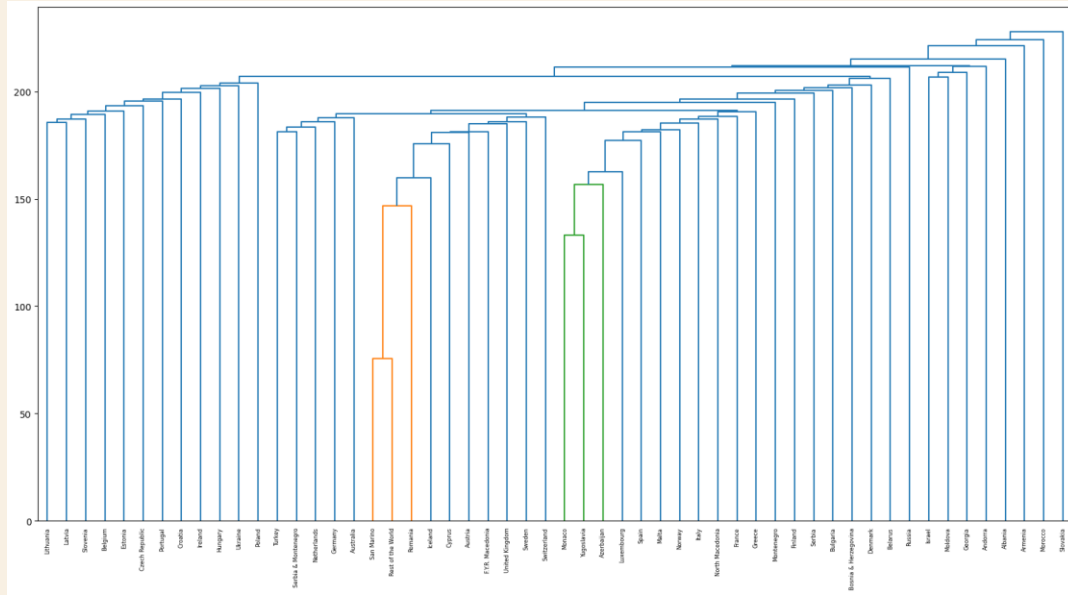
- Implementirana tudi za manjkajoče vrednosti
- Bolj ustrezna kot Manhattanska

Linkage: single

- Dober tudi za podolgovate gruče
- Implementirano tudi za manjkajoče vrednosti

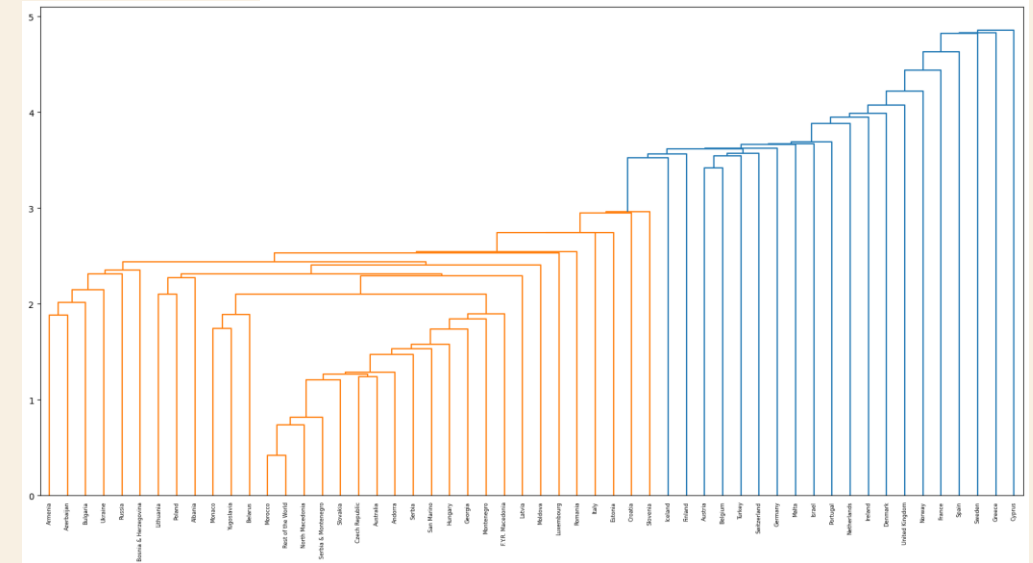
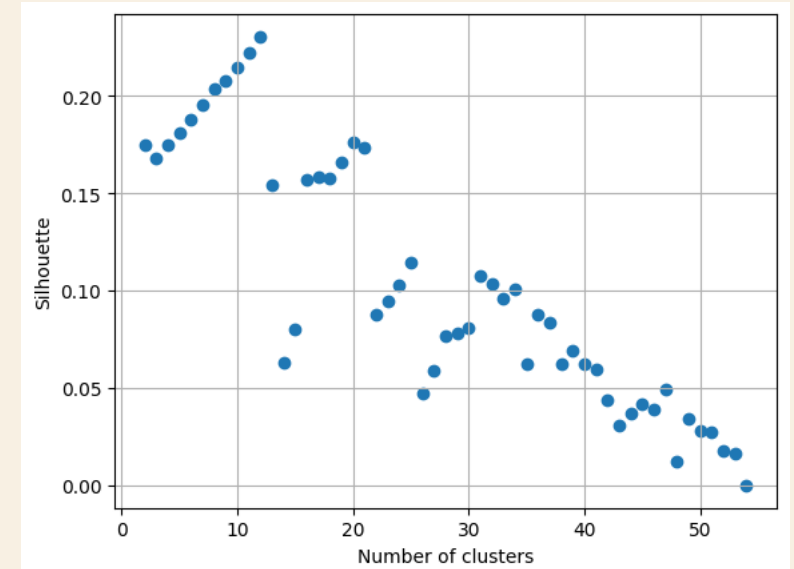
```
#params
distance1:=hw0.euclidean_dist
linkage1:=hw0.single_linkage
distFun1:=lambda c1, c2: linkage1(c1, c2, distance1)
silhFun1:=silhouette_average
```

DENDROGRAM (IN SILHUETA)



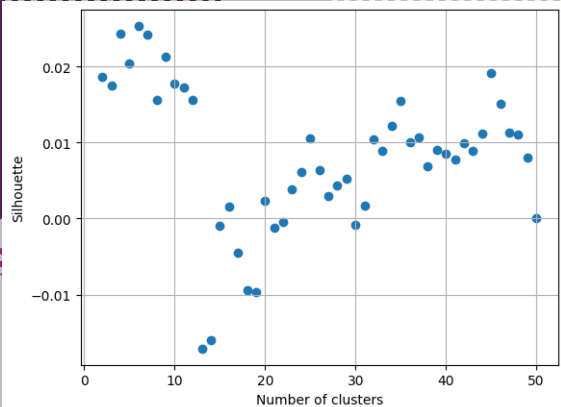
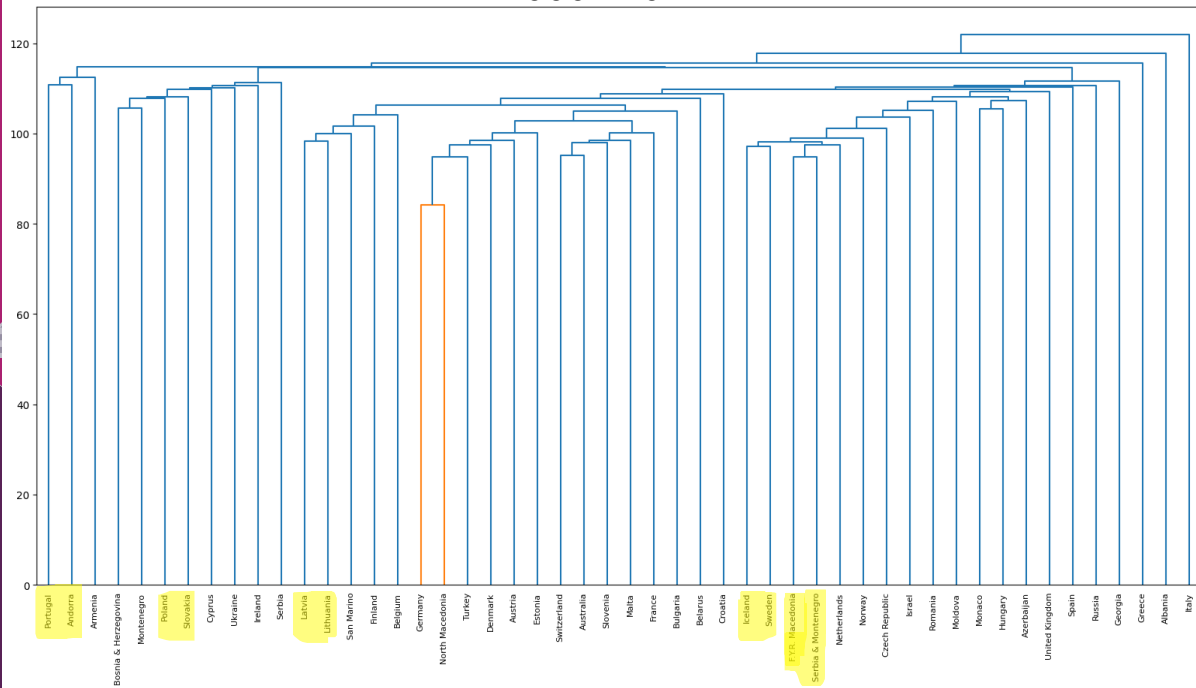
Leta
zložena v
vektor

Povprečje
let



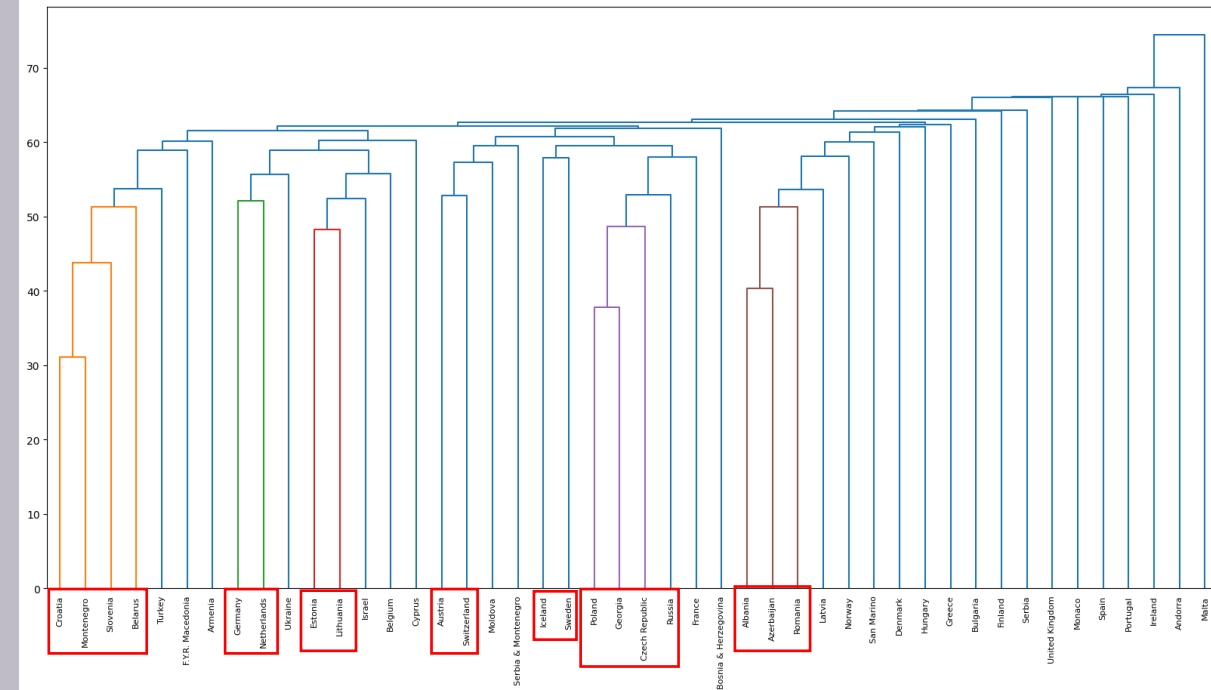
ARGUMENTACIJA SKUPIN

2000 - 2019

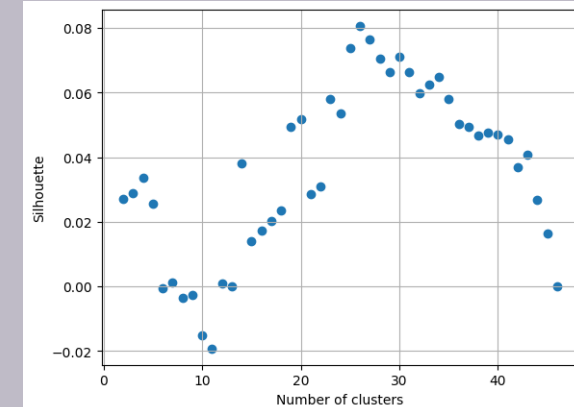


- Izboljšanje v primerjavi s podatki za vsa leta
 - Silhueta
 - Bolj smiselne skupine

2000 - 2008



- Bolj smiselne skupine (geografsko, kulturno,...)
- Silhueta kaže boljše prileganje



ZANIMIVE SKUPINE

