

# Glasovanje za Pesem Evrovizije

Matea Lenček (63140392)

21. oktober 2017

## 1 Podatki

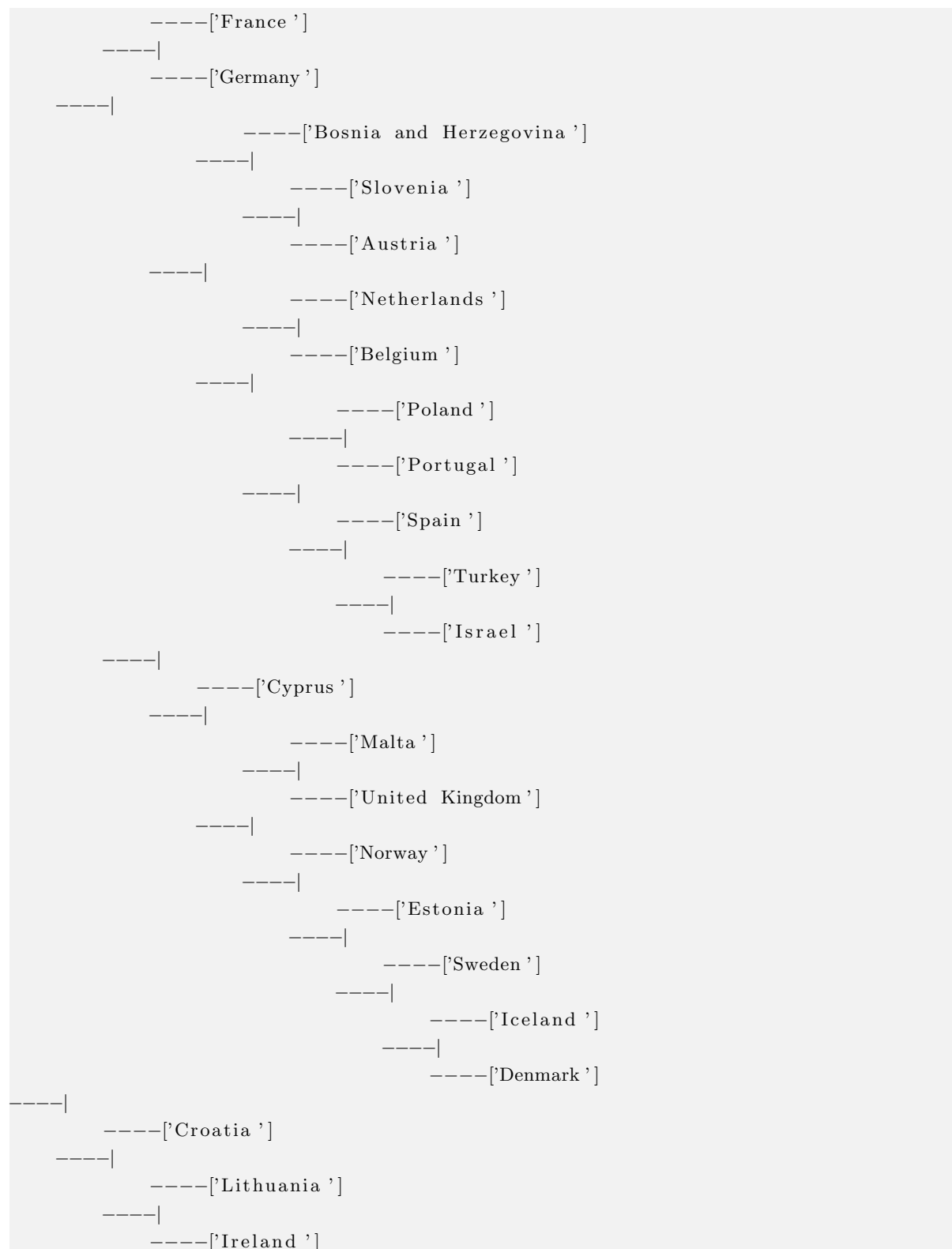
Podatke za analizo sem pridobila na portalu Kaggle. Analizirala sem le podatke finalnih tekmovalcev vseh let, ki sem jih imela na voljo (1998 - 2009). Uporabila sem podatek za katero državo gre in vektor vrednosti kako je ta država glasovala. To sem izluščila iz celotnih podatkov tako, da sem uporabila le določene stolpce (od 16. stolpca dalje, brez zadnjih praznih stolpcev), ki sem jih spremenila v celoštevilski tip (int). Manjkajoče vrednosti (") sem v tem koraku spremenila v -1.

## 2 Računanje razdalj

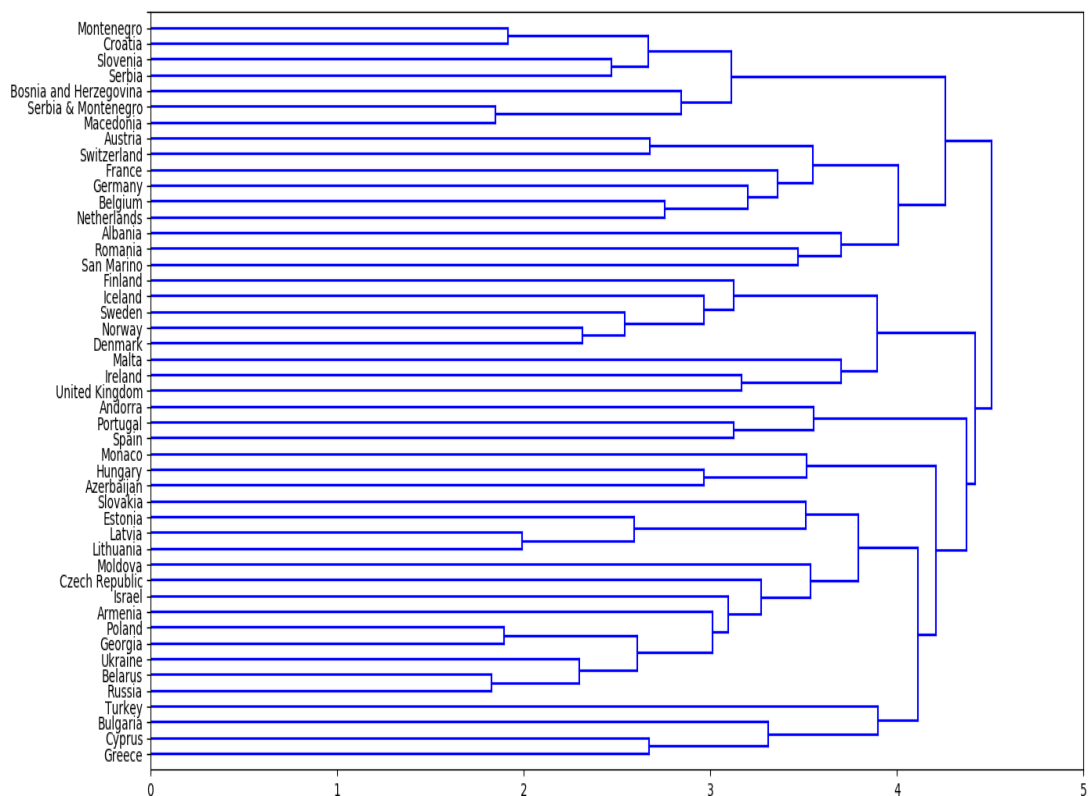
Za računanje razdalj med posameznimi profili držav sem uporabila evklidsko razdaljo. Pare z eno ali dvema manjkajočimi vrednostmi (na tem mestu -1) sem pri računanju ignorirala. Zato sem vsoto pred korenjenjem normalizirala s številom parov, katere sem dejansko uporabila za izračun. Za razdaljo med posameznimi skupinami pa sem uporabila kar povprečno razdaljo med vsemi kombinacijami parov obeh skupin in združila skupini z najmanjšo povprečno razdaljo. Pri tem sem ignorirala pare držav, kjer njune razdalje nisem mogla izračunati zaradi manjkajočih vrednosti.

### 3 Dendrogram

Za boljšo preglednost prilagam tekstovni izris dendrograma le za leto 1999, ko je na tekmovanju glasovalo najmanj držav.

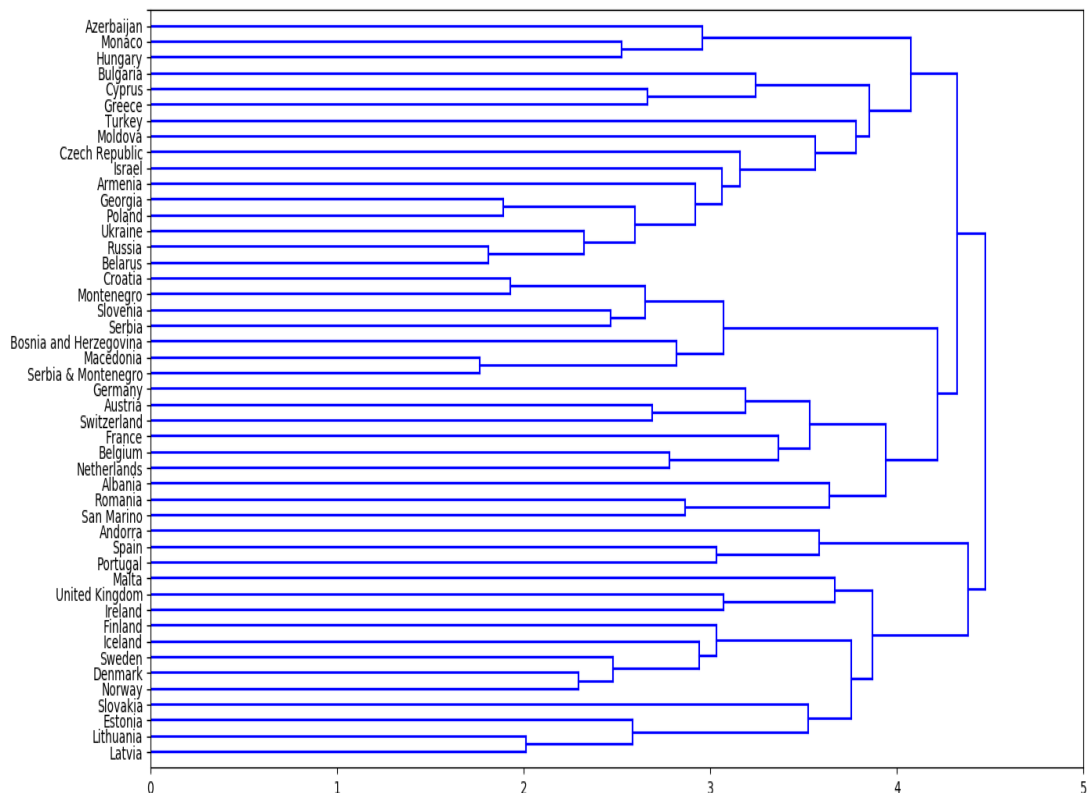


Slika 1 prikazuje grafični izris dendrograma za podatke vseh let finalnih tekmovanj, ki sem ga implementirala sama.



Slika 1: Grafični izris dendrograma.

Poskusila sem tudi z idejo, da pri hierarhičnem razvrščanju v skupine za vsako leto ne upoštevam zmagovalne pesmi, saj ponavadi le-ta dobi največ visokih točk od velike večine držav. Rezultat tega je bil skoraj identičen prvemu, saj podobnost pri točkah dodeljenih zmagovalcu, ne vpliva prav dosti na razdaljo. To podobnost s prvim dendrogramom lahko vidimo na sliki 2



Slika 2: Grafični izris dendrograma, brez upoštevanja zmagovalnih pesmi pri hierarhičnem razvrščanju v skupine.

## 4 Skupine in njihove preferenčne izbire

Skupine sem iz rezultatov razvrščanja določila tako, da sem dendrogram odrezala na razdalji, katero sem določila tam, kjer se je meni zdelo smiselno. Veliko skupin se namreč združi zelo pozno (na velikih razdaljah), kar po mojem mnenju ne prikazuje resnih podobnosti med njimi. Določila sem razdaljo 3.79, kjer sem dobila skupine predstavljene v tabeli 1, kjer so tudi skupinam pripadajoče preferenčne države in države, ki so jim skupine namenile najmanj točk. Preferenčne države sem določila tako, da sem preštela koliko točk so posamične države v skupini namenile določeni državi, pri tem pa sem štela le države, katerim so namenile 10 ali 12 točk (rezultati, ko sem upoštevala zraven še 8 točk, so se od prikazanih razlikovali minimalno). Izbrala sem le top 3 države. Za države, ki jim je skupina namenila najmanj točk, pa sem izbrala 3 države, ki so dobile od skupine največkrat 0 točk.

Tabela 1: Skupine, njihove preferenčne države (10, 12 točk) in najslabše ocenjene države

skupina	preferenčne države (10, 12 točk)	države, ki jim namenijo najmanj točk
Turkey	Bosnia and Herzegovina, Armenia, Greece	France, Spain, United Kingdom
Croatia, Montenegro, Slovenia, Serbia, Bosnia and Herzegovina, Serbia & Montenegro, Macedonia	Bosnia and Herzegovina, Croatia, Serbia	Spain, Germany, France
Finland, Iceland, Sweden, Norway, Denmark	Sweden, Denmark, Norway	Spain, United Kingdom, Croatia
Bulgaria, Greece, Cyprus	Greece, Cyprus, Armenia	Germany, France, Sweden
Slovakia, Estonia, Latvia, Lithuania	Russia, Latvia, Estonia	Turkey, Spain, Bosnia and Herzegovina
Monaco, Hungary, Azerbaijan	Greece, Turkey, Azerbaijan	Spain, United Kingdom, Germany
Moldova, Czech Republic, Israel, Armenia, Georgia, Poland, Ukraine, Belarus, Russia	Russia, Ukraine, Armenia	Germany, Spain, Bosnia and Herzegovina
Switzerland, Austria, France, Germany, Belgium, Netherlands	Turkey, Germany, Israel	United Kingdom, France, Romania
Andorra, Portugal, Spain	Spain, Romania, Germany	Bosnia and Herzegovina, Turkey, Croatia
Albania, San Marino, Romania	Greece, Turkey, Moldova	United Kingdom, Croatia, Spain
Malta, Ireland, United Kingdom	Sweden, Malta, Latvia	Spain, France, Croatia

Glede na to, da so razdalje na katerih se večina skupin združi precej velike, ne vem če je smiselno govoriti v nekih velikih podobnostih. Morda bi bilo rez bolje narediti že nekoliko prej, s čimer bi sicer dobili veliko osamelcev in le nekaj skupin držav, ki pa bi jim zaradi manjših razdalj ob združenju lahko pripisali večjo podobnost, predvsem geografsko (s tem tudi zgodovinsko). Nekaj teh skupin bi tako bilo: države bivše Jugoslavije, Rusija in njene sosednje države, pribaltske države, skandinavske države z Islandijo in Dansko.

## 5 Odvisnost rezultatov od leta glasovanja

Z analizo podatkov finalnih tekmovanj po posamičnih letih sem opazila, da se dendrogrami do leta 2004 iz leta v leto precej razlikujejo. Razlog za to bi bil lahko manjša udeležba držav na tekmovanju za Pesem Evrovizije, kar opazimo že po samem številu glasujočih držav. Od leta 2004 dalje, pa so rezultati hierarhičnega združevanja v skupine iz leta v leto zelo podobni (in hkrati podobni tudi analizi podatkov vseh let), saj se je množica glasujočih držav močno povečala in s tem so bolj opazne podobnosti med zopet predvsem geografsko in zgodovinsko povezanimi državami.