

Analiza financijske ovisnosti starijeg stanovništva o radno aktivnom stanovništvu korištenjem otkrivanja znanja iz baza podataka

Molnar, Nika

Master's thesis / Diplomski rad

2023

Degree Grantor / Ustanova koja je dodijelila akademski / stručni stupanj: **University of Zagreb, Faculty of Economics and Business / Sveučilište u Zagrebu, Ekonomski fakultet**

Permanent link / Trajna poveznica: <https://um.nsk.hr/um:nbn:hr:148:188739>

Rights / Prava: [Attribution-NonCommercial-ShareAlike 3.0 Unported / Imenovanje-Nekomercijalno-Dijeli pod istim uvjetima 3.0](#)

Download date / Datum preuzimanja: **2024-07-11**



Repository / Repozitorij:

[REPEFZG - Digital Repository - Faculty of Economics & Business Zagreb](#)



Sveučilište u Zagrebu

Ekonomski fakultet

Integrirani preddiplomski i diplomski sveučilišni studij

Poslovna ekonomija – smjer Menadžerska informatika

**ANALIZA FINANCIJSKE OVISNOSTI STARIJEG
STANOVNIŠTVA O RADNO AKTIVNOM STANOVNIŠTVU
KORIŠTENJEM OTKRIVANJA ZNANJA IZ BAZA
PODATAKA**

Diplomski rad

Nika Molnar

Zagreb, rujan, 2023.

Sveučilište u Zagrebu

Ekonomski fakultet

Integrirani preddiplomski i diplomski sveučilišni studij

Poslovna ekonomija – smjer Menadžerska informatika

**ANALIZA FINANCIJSKE OVISNOSTI STARIJEG
STANOVNIŠTVA O RADNO AKTIVNOM STANOVNIŠTVU
KORIŠTENJEM OTKRIVANJA ZNANJA IZ BAZA
PODATAKA**

**ANALYSIS OF FINANCIAL DEPENDANCE OF ELDERLY
POPULATION ON WORKING POPULATION USING
KNOWLEDGE DISCOVERY FROM DATABASES**

Diplomski rad

Student: Nika Molnar

JMBAG studenta: 0066255540

Mentor: prof. dr. sc. Mirjana Pejić Bach

Zagreb, rujan, 2023.

IZJAVA O AKADEMSKOJ ČESTITOSTI

Izjavljujem i svojim potpisom potvrđujem da je diplomski rad isključivo rezultat mog vlastitog rada koji se temelji na mojim istraživanjima i oslanja se na objavljenu literaturu, a što pokazuju korištene bilješke i bibliografija.

Izjavljujem da nijedan dio rada nije napisan na nedozvoljen način, odnosno da je prepisan iz necitiranog rada, te da nijedan dio rada ne krši bilo čija autorska prava. Izjavljujem, također, da nijedan dio rada nije iskorišten za bilo koji drugi rad u bilo kojoj drugoj visokoškolskoj, znanstvenoj ili obrazovnoj ustanovi.

U Zagrebu, rujan, 2023.



Nika Molnar

STATEMENT ON THE ACADEMIC INTEGRITY

I hereby declare and confirm by my signature that the final thesis is the sole result of my own work based on my research and relies on the published literature, as shown in the listed notes and bibliography.

I declare that no part of the thesis has been written in an unauthorized manner, i.e., it is not transcribed from the non-cited work, and that no part of the thesis infringes any of the copyrights. I also declare that no part of the thesis has been used for any other work in any other higher education, scientific or educational institution.

In Zagreb, september, 2023.



Nika Molnar

SAŽETAK I KLJUČNE RIJEČI

Radom se proučavaju trend starenja stanovništva, načini izračuna omjera ovisnosti starijeg stanovništva i posljedice koje nosi demografsko starenje. Starenje stanovništva značajno utječe na financijske izdatke države, stoga njegovo je praćenje i razumijevanje ključno za ostvarenje dobrobiti stanovništva. Pomoću softverskog alata *Weka* provodi se otkrivanje znanja u bazama podataka Ujedinjenih naroda i Organizacije za ekonomsku suradnju i razvoj. Grupiranjem država u klustere na temelju sličnosti, Hrvatska se pozicionira na svjetskoj razini. Prognoziranje vremenskih serija prikazuje predviđene vrijednosti omjera ovisnosti starijeg stanovništva u Hrvatskoj u budućnosti. Cilj je predvidjeti kretanje omjera kako bi se razumjeli raspodjela resursa, planiranje infrastrukture i razvoja te formuliranje budućih ciljeva i politika, osobito onih vezanih uz mirovinski sustav.

Ključne riječi: starenje stanovništva, omjer ovisnosti starog stanovništva, financijska ovisnost starog stanovništva, otkrivanje znanja iz baza podataka, rudarenje podataka, klaster analiza, analiza vremenskih serija

SUMMARY AND KEYWORDS

This study analyzes the trend of population ageing, ways of calculating old age dependance ratio and future impacts that demographic ageing holds. Population ageing significantly affects financial expenditures of the state, therefore, monitoring and understanding it are essential for achieving well-being of the population. The *Weka* software tool is used for knowledge discovery in the databases of the United Nations and The Organization for Economic Co-operation and Development. By grouping countries into clusters based on similarities, Croatia is positioned on a global scale. Time series forecasting presents predicted values of the old age dependance ratio in Croatia in the future. The goal is to predict movement of the ratio in order to understand resource allocation, infrastructure and development planning, as well as formulation of future goals and policies, especially those related to pension system.

Keywords: population ageing, old age dependency ratio, elderly population financial dependance, data discovery from databases, data mining, cluster analysis, time series analysis

Sadržaj

1. UVOD	1
1.1. Predmet i cilj rada	1
1.2. Izvori podataka i metode prikupljanja	1
1.3. Sadržaj i struktura rada	2
2. OVISNOST STAROG STANOVNIŠTVA	3
2.1. Demografski resursi i njihova obilježja	3
2.2. Starenje stanovništva	5
2.3. Omjer ovisnosti starog stanovništva.....	6
2.4. Posljedice starenja stanovništva.....	10
3. OTKRIVANJE ZNANJA U BAZAMA PODATAKA	12
3.1. Proces otkrivanja znanja iz baza podataka	12
3.2. Primjena otkrivanja znanja iz baza podataka.....	15
3.3. Metode otkrivanja znanja iz baza podataka.....	16
3.3.1. Klaster analiza.....	18
3.3.2. Analiza vremenskih serija.....	19
3.4. Alati za otkrivanje znanja u bazama podataka.....	20
4. ANALIZA OVISNOSTI STAROG STANOVNIŠTVA KORIŠTENJEM OTKRIVANJA ZNANJA IZ BAZA PODATAKA	21
4.1. Metodologija istraživanja	21
4.2. Klaster analiza.....	21
4.2.1. Korišteni podaci	21
4.2.2. Rezultati istraživanja	23
4.3. Analiza vremenskih serija.....	33
4.3.1. Korišteni podaci	33
4.3.2. Rezultat istraživanja	34
5. ZAKLJUČAK.....	41
LITERATURA.....	43
POPIS SLIKA.....	46
POPIS TABLICA.....	46
POPIS GRAFIKONA.....	46
PRILOZI	47

1. UVOD

1.1. Predmet i cilj rada

Značajan broj znanstvenika i političara pridaje veliku važnost procesu starenja stanovništva. Ono je jedan od čimbenika koji imaju značajan utjecaj na socioekonomsku održivost. Poznata je činjenica da se posljednjih desetljeća povećao broj starog stanovništva zbog produljenja životnog vijeka i smanjenje stope fertiliteta. Starenje utječe na ekonomiju, odnosno stvara velike troškove. Troškovi koji se uglavnom povezuju uz starenje troškovi su za mirovine, zdravstvo, dugotrajnu skrb te obrazovni troškovi i naknade za nezaposlenost (Puljiz, 2015.). Prema tome, zaključuje se kako je starije stanovništvo ovisno o radno sposobnom stanovništvu.

Predmet rada omjer je ovisnosti starog stanovništva (u dobi od 65 godina i više) o radno aktivnom stanovništvu (dobi između 15 i 64 godine). Dovodeći u vezu skupinu stanovništva koja će najvjerojatnije biti ekonomski aktivna sa skupinom koja će biti ekonomski ovisna, omjer ovisnosti starog stanovništva pokazuje potencijalne učinke dobne strukture stanovništva na društveni i ekonomski razvoj.

Cilj ovog rada je provesti analizu ovisnosti starog stanovništva korištenjem otkrivanja znanja iz baza podataka. Analiza će se provesti pomoću alata za rudarenje podataka, *Weke* uz pomoć klaster analize i analize vremenskih serija. Rezultati analize ovisnosti starijeg stanovništva bit će korisni za sastavljanje plana korištenja resursa i formuliranje budućih ciljeva politika, kao i za planiranje infrastrukture i razvoja.

1.2. Izvori podataka i metode prikupljanja

Kao izvor podataka koristit će se baza podataka Razvojnog programa Ujedinjenih naroda (*engl. United Nations Development Programme*), točnije Izvješće o ljudskom razvoju iz 2020. godine (*engl. Human Development Report 2020*). Razvojni program radi na iskorjenjivanju siromaštva te pomaže zemljama da razviju politike, partnerstva, vještine i institucije kako bi ostvarile napredak. Baze podataka ažuriraju se svake godine te obuhvaćaju podatke o kvaliteti ljudskog razvoja, rodnim razlikama, osnaživanju žena, ekološkoj održivosti i socioekonomskoj održivosti, pod što spada i omjer ovisnosti starog stanovništva.

Osim UN-ove baze, koristit će se i baza Organizacije za ekonomsku suradnju i razvoj (*engl. Organization for Economic Co-operation and Development*)(OECD) koja sadrži podatke o

ukupnom stanovništvu, mladom i starom stanovništvu, stanovništvu u radnoj dobi kao i podatke o ovisnosti starijeg stanovništva za sve države svijeta.

1.3.Sadržaj i struktura rada

Rad je podijeljen u pet poglavlja. U prvom, uvodnom, navode se predmet i cilj rada kao i izvori podataka na temelju kojih će se provoditi analiza. U drugom su poglavlju opisani demografski resursi, pojam starenja stanovništva, objašnjeni su načini izračuna omjera ovisnosti starijeg stanovništva kao i posljedice koje nosi demografsko starenje. Trećim je poglavljem opisan proces otkrivanja znanja u bazama i njegova raznovrsna primjena. Također, opisane su metode korištene u analizi u narednom poglavlju, a to su klaster analiza i analiza vremenskih serija. Četvrto se poglavlje bavi analizom omjera ovisnosti starijeg stanovništva. Ono obuhvaća metodologiju i podatke korištene u analizi te analizu i interpretaciju rezultata grupiranjem država po klasterima i predviđanjem vrijednosti kretanja omjera ovisnosti starijeg stanovništva u budućnosti. U zaključnom se poglavlju daje osvrt na rad i istraživanje provedeno u prethodnom poglavlju.

2. OVISNOST STAROG STANOVNIŠTVA

2.1. Demografski resursi i njihova obilježja

A. Wertheimer-Baletić (1999.) opisuje demografiju kao društvenu znanost koja proučava stanovništvo s kvantitativnog i kvalitativnog aspekta. Dakle, predmet demografije je stanovništvo, njegovo kretanje i druge pojave, kao i njegove strukture koje se zasnivaju na individualnim karakteristikama pojedinaca (Breznik, 1980.).

Demografski resursi važna su sastavnica ljudskog razvoja i obuhvaćaju ukupna kvalitativna i kvantitativna, stvarna i potencijalna društvena i biološka obilježja stanovništva u određenom vremenu i prostoru (Oliveira-Roca, 1991.). Važni su za razne dionike, uključujući vladine institucije, istraživače, planere razvoja, organizacije civilnog društva i ostale koji se bave pitanjima populacije. Razumijevanje obilježja demografskih resursa omogućuje donositeljima odluka formiranje programa koji su usklađeni s potrebama i izazovima stanovništva. Na primjer, temeljem informacija o dobnoj strukturi stanovništva, mogu se planirati obrazovne politike prilagođene potrebama djece i mladih ili politike za starije stanovništvo koje potiču socijalnu skrb. Promjene u obilježjima demografskih resursa utječu na društvenu strukturu i društvene odnose (Nejašmić i Mišketić, 2010.).

Demografske je resurse moguće definirati uz pomoć navedenih indikatora (Poston, Bouvier, 2010.):

1. broja stanovnika – ukupnog broja ljudi koji žive na određenom području,
2. dobne strukture – raspodjele stanovništva prema dobnim skupinama kao što su djeca, mladi, starije stanovništvo, radno sposobno stanovništvo i sl.,
3. spolne strukture – raspodjele prema spolu (muškarci i žene),
4. nataliteta – broja živorođene djece u određenom razdoblju (obično se izražava kao broj živorođene djece na 1000 stanovnika u godini dana i pokazuje stopu rodnosti),
5. mortaliteta – broja smrtnih slučajeva u određenom razdoblju (obično se izražava kao broj umrlih na 1000 stanovnika u godini dana i pokazuje stopu smrtnosti),
6. migracija – unutarnjih migracija (kretanje stanovništva unutar zemlje) i vanjskih migracija (kretanje stanovništva izvan granica države),
7. distribucije populacije – raspodjele stanovništva u određenom geografskom prostoru,
8. obrazovanju i strukturi rada – stupnju obrazovanja, strukturi radne snage, zanimanju i nezaposlenosti te

9. zdravlju i socijalnoj zaštiti – životnom vijeku, bolesti, pristupu zdravstvenim i socijalnim uslugama.

Demografska inercija pojam je koji objašnjava da se demografske promjene poput nataliteta, mortaliteta ili migracija događaju sporije od očekivanog (Wertheimer-Baletić, 1999.). Njome se, dakle, ubrzavaju negativni demografski trendovi kao što su niska stopa fertiliteta, visoka stopa smrtnosti, migracije odlaska, nejednakost u raspodjeli stanovništva i negativne promjene u demografskoj strukturi. Primjerice, promjene u demografskoj strukturi stanovništva, kao što je starenje populacije, mogu dovesti do povećanja troškova zdravstvene skrbi i opterećenja mirovinskog sustava. Zbog toga je neophodno upravljati resursima kako bi se postigao održiv i inkluzivan razvoj društva.

Ujedinjeni narodi (2022.) u svom Programu održivog razvoja (*engl. United Nations Development Programme – UNDP*) navode kako promjene u strukturi stanovništva pokazuju socioekonomski idikatori. Oni se dijele na:

- pokazatelje ekonomske održivosti i
- pokazatelje socijalne održivosti.

Pokazatelji ekonomske održivosti su:

- prilagođena neto štednja,
- ukupno servisiranje duga,
- bruto investicije,
- kvalificirana radna snaga,
- indeks koncentracije izvoza i
- rashodi za istraživanje i razvoj.

Pokazatelji socijalne održivosti su:

- omjer ovisnosti starog stanovništva,
- vojni izdaci kao udio u BDP-u,
- omjer rashoda za obrazovanje i zdravstvo u odnosu na vojne rashode,
- prosječna godišnja promjena u ukupnom gubitku HDI vrijednosti zbog nejednakosti,
- prosječna godišnja stopa promjene indeksa rodne nejednakosti i
- prosječna godišnja stopa promjene udjela prihoda 40% najsiromašnijeg stanovništva.

2.2. Starenje stanovništva

Starenje stanovništva ili demografsko starenje trend je koji se odvija dugi niz godina i čest je predmet istraživanja i analiza. Puljiz (2016.) navodi dvije determinante starenja stanovništva: starenje odozgo i starenje odozdo. Starenje odozgo obuhvaća produženje životnog vijeka, to jest, povećanje prosječne životne dobi populacije. To se događa zbog napretka u medicini, poboljšanja zdravstvene skrbi, bolje higijene i prehrambenih uvjeta te napretka u znanosti i tehnologiji općenito. Produženje životnog vijeka obično prati smanjenje stope fertiliteta, što je druga determinanta zvana starenje odozdo. Smanjenje stope fertiliteta odnosi se na pad prosječnog broja djece koje žene rađaju tijekom svog reproduktivnog razdoblja. Događa se zbog obrazovanja i emancipacije žena, povećanog pristupa kontracepciji te urbanizacije i promjena u načinu života i društvenih vrijednosti. Taj trend također ima utjecaj na demografski jaz između starije i mlađe populacije.

Pojam staro stanovništvo obuhvaća osobe od 65 i više godina, a pomoću indikatora starenja stanovništva mjeri se demografski proces starenja pojedine populacije. U pojedinoj se literaturi pod starije stanovništvo ubrajaju osobe iznad navršenih 60 godina života, ali u ovom se radu u starim stanovništvom smatraju osobe od 65 i više godina. Živić (2003.) navodi tri glavna indikatora starenja stanovništva. Prvi je koeficijent starosti koji izražava udio starijih osoba u ukupnom broju stanovnika. Povećanje koeficijenta starosti ukazuje na povećan udio starijih osoba u populaciji. Slijedi odnos starijeg prema mladom stanovništvu koji izražava omjer osoba od 65 godina i više u odnosu na broj mladih osoba (obično u dobi 0-14 godina). Kao treći indikator navodi se omjer ovisnosti starog stanovništva prema radno sposobnom stanovništvu koji uspoređuje broj starijeg stanovništva s radnim kontingentom (radno sposobnim osobama u dobi 16-64 godine). Osim navedenih, bitno je spomenuti stopu očekivane životne dobi kojom se mjeri očekivano trajanje života u populaciji. Može se računati za muškarce, žene ili za ukupnu populaciju. Prema World Health Organization (WHO), globalno očekivano trajanje života se povećalo za više od 6 godina između 2000. i 2020. godine i trenutno iznosi nešto preko 73 godine.

Prije nego što su se počele primjenjivati moderne medicinske tehnologije i javne zdravstvene intervencije, očekivana životna dob bila je relativno niska, a stope smrtnosti, posebno u djetinjstvu, bile su vrlo visoke. U toj fazi starenje stanovništva nije bio demografski fenomen na koji treba obratiti pažnju. Tijekom 19. i 20. stoljeća dolazi do značajnog medicinskog napretka, poboljšanih sanitarnih uvjeta i javnog zdravstva koji su utjecali na smanjenje stope

smrtnosti. S demografskom tranzicijom, koja se dogodila zbog urbanizacije i socioekonomskog razvoja u 20. stoljeću, smanjuje se stopa mortaliteta, ali i fertiliteta. Došlo je do promjena u društvenim vrijednostima, obrazovanja žena i povećanju kontracepcijskih metoda zbog kojih se smanjio udio djece. Očekivana životna dob do danas je nastavila rasti, a stopa fertiliteta se održava na niskoj razini. Zbog toga se danas puno pažnje posvećuje izučavanju procesa starenja stanovništva jer ono stavlja velik pritisak na zdravstveni i mirovinski sustav, radnu snagu i ostale aspekte društvenog funkcioniranja.

Važno je napomenuti da su različite regije svijeta u različitim fazama starenja stanovništva. Dok su neke zemlje već suočene s visokim postotkom starijeg stanovništva, druge se još uvijek nalaze u ranijim fazama demografske tranzicije. Povijest starenja stanovništva nastavlja se razvijati kako se društvo mijenja i demografske se karakteristike prilagođavaju novim trendovima i promjenama.

2.3. Omjer ovisnosti starog stanovništva

Raniji koncepti izračuna starosti stanovništva uzimali su u obzir samo odnos između starijih i mlađih dobnih skupina. Međutim, kako bi se dobili što precizniji podaci, potrebno je uzeti u obzir životni vijek stanovništva te kombinirati ekonomske i demografske podatke. Na taj se način uspoređuju ekonomski doprinos i ovisnost sa dobnom strukturom (UN, 2019.).

Ujedinjeni Narodi (*engl. United Nations – UN*) navode tri izračuna starosti stanovništva koji ispituju promjene u dobnjoj strukturi i međugeneracijsku podršku: koeficijent ovisnosti starog stanovništva, prospektivni omjer ovisnosti starog stanovništva i ekonomsku ovisnost starog stanovništva.

Koeficijent ovisnosti starog stanovništva (*engl. Old-age dependency ratio – OADR*) jedan je od najčešćih alata korištenih za praćenje dobne strukture populacije. Pokazuje broj starijih uzdržavanih osoba (u dobi od 65 godina i više) na 100 radno sposobnih osoba (u dobi 16-64 godine). Koristi se za procjenu potencijalnog opterećenja radno sposobnog stanovništva u smislu financijskog uzdržavanja pružanjem potpora za mirovine, zdravstvenu skrb i socijalne usluge za starije osobe. OADR se izračunava na slijedeći način:

$$OADR = \frac{\text{broj osoba u dobi } \geq 65 \text{ godina}}{\text{broj osoba u dobi 16-64 godine}} \times 100 \quad (1)$$

Viši OADR podrazumijeva veći udio starih osoba u odnosu na radno sposobno stanovništvo, što sugerira veće opterećenje za podršku i njegu starije populacije. UN (2019.) pretpostavlja da će ovisnosti u starijoj dobi rasti u svim dijelovima svijeta, osobito u Istočnoj i Jugoistočnoj Aziji, Latinskoj Americi i na Karibima. Treba napomenuti kako se ovaj koeficijent značajno razlikuje u pojedinim zemljama zbog različitih demografskih trendova, socijalnih politika i ekonomskih čimbenika.

Prema UN-ovim podacima za 2019. godinu, najveći OADR imao je Japan s 51 osobom od 65 i više godina na 100 radno sposobnih osoba. Slijede Finska i Italija u kojima je koeficijent ovisnosti starog stanovništva 39, te Portugal sa koeficijentom 38.

Glavni nedostatak koeficijenta ovisnosti starog stanovništva je baziranost na fiksnoj dobnoj granici (obično 60 ili 65 godina) koja ne obuhvaća složenost ovisnosti unutar populacije. Potrebno je prepoznati raznolikost ekonomske aktivnosti starijeg stanovništva, kao što su sudjelovanje u radnoj snazi i funkcionalnu sposobnost. Osim toga, nisu svi pojedinci u dobi između 15 i 64 godine uključeni u radnu snagu te su i sami ekonomski ovisni. Zbog toga, iako je prikladan za prikazivanje promjena u strukturi stanovništva, glavni mu je nedostatak to što ne pruža sveobuhvatno razumijevanje dinamike ovisnosti unutar populacije.

Alternativni način koji računa omjer ovisnosti starog stanovništva uzima u obzir mlađu populaciju, točnije osobe do 20 godina starosti. Pritom se ovaj put uspoređuje omjer mladog stanovništva (u dobi do 20 godina) i radno aktivnog stanovništva (u dobi 20-64 godine). Dugoročno gledano, pad omjera mladog u radno aktivnom stanovništvu još je jedan od značajnih pokazatelja starenja. Izračun koeficijenta mladog stanovništva prikazan je u nastavku:

$$\text{Koeficijent mladog stanovništva} = \frac{\text{broj osoba u dobi } \leq 19 \text{ godina}}{\text{broj osoba u dobi 20-64 godine}} \times 100 \quad (2)$$

S povećanjem očekivane životne dobi i povećanjem dobne granice za mirovinu te sve većom vitalnosti starije populacije, izračun omjera ovisnosti starog stanovništva ne zadovoljava sve potrebe istraživanja. Zbog toga su Sanderson i Scherbov (2007.) osmislili izračun omjera starosti stanovništva, odnosno „**Prospektivni omjer ovisnosti starog stanovništva**“ (*engl. Prospective old-age dependency ratio – POADR*). Prospektivni omjer za izračun koristi prospektivnu dob, godinu indeksa (interesne godine) i standardne godine. Ovisno o promjeni

prospektivnih godina i standardnih godina moguće je dobiti izračun koji je prikladan na odgovaranje za različita pitanja. POADR se izračunava na sljedeći način:

$$POADR = \frac{\text{broj osoba prospektivne dobi}}{\text{broj osoba između 20 godina i prospektivne dobi}} \times 100 \quad (3)$$

Za usporedbu, konvencionalnim načinom izračuna dobio bi se omjer broja ljudi u dobi od 65 i više godina u usporedbi s onima između 20 i 64 godine. Prospektivnim omjerom računa se broj ljudi u prospektivnoj dobi u usporedbi s brojem ljudi između 20 i prospektivne (očekivane) dobi. OADR pretpostavlja da starost počinje sa 65 godina dok POADR pretpostavlja da počinje u dobi kada je preostali životni vijek 15 godina. S povećanjem očekivane životne dobi, prospektivni omjer starog stanovništva niži je od konvencionalnog. Sanderson i Scherbov (2007.) navode kako konvencionalni omjer ovisnosti starog stanovništva preuveličava buduće poteškoće i donositeljima politika može pružiti pogrešne informacije. Navode kako je bolji način razumijevanja izazova starenja populacije korištenje njihovog prospektivnog omjera starosti stanovništva dopunjenog informacijama o promjenama očekivane životne dobi. Osim toga, izračun može biti koristan u sociologiji i ekonomiji, gdje se koristi za proučavanje vremenskog usklađivanja događaja tijekom životnog ciklusa, poput umirovljenja. Promjene u prospektivnoj medijalnoj dobi također se mogu koristiti u predviđanju promjena u ponašanju štednje i ulaganja, pohađanja škole u starijoj dobi te troškova zdravstvene skrbi.

Treći način koji UN (2019.) navodi je **ekonomska ovisnost starog stanovništva** (*engl. Economic old-age dependency ratio*). Ekonomski pokazatelji ovisnosti starog stanovništva računa se uz pomoć Nacionalnih računa prijenosa (UN, 2013). Prema UN-u (2019.) Nacionalni računi prijenosa (*enl. National transfer accounts – NTA*) sveobuhvatan su sustav starosno usmjerenog računovodstva ekonomskih tokova koji se koristi za proučavanje raspodjele ekonomskih resursa među pojedincima različitih dobnih skupina kroz obitelj, vladu i tržište.

Pristup koristi analitičke metode demografije i ekonomije kako bi se razumjelo ekonomsko ponašanje povezano sa starenjem populacije. Pruža informacije o sudjelovanju radne snage, potrošnji i proizvodnji u određenoj ekonomiji te pomaže u razumijevanju ekonomskih aspekata starenja populacije. Izračun NTA prikazan je u nastavku:

$$NTA = C + S + T \quad (4)$$

Vrlo pojednostavljena formula za NTA glasila bi: C+S+T, gdje C (*engl. consumption*) predstavlja potrošnju ili konzumaciju u određenoj dobnoj skupini, S (*engl. savings*) predstavlja štednju ili akumulaciju resursa u određenoj dobnoj skupini, a T (*engl. transfer*) označava prijenos resursa između dobnih skupina kao što su mirovine, socijalna pomoć i slično. Složenije formule primjenjuju se na različite dobne skupine unutar populacije kako bi se analizirao način na koji se ekonomski resursi preusmjeravaju između generacija i različitih sektora ekonomije.

Uz pomoć NTA, ekonomski omjer ovisnosti starog stanovništva (ekonomski OADR) koristi informacije o populaciji, potrošnji i proizvodnji u određenoj ekonomiji, razgranate prema dobnim skupinama. Računa se dijeljenjem efektivnog broja potrošača u dobi od 65 godina ili više s efektivnim brojem radno sposobnog stanovništva svih dobnih skupina. Ekonomski OADR izračunava se na slijedeći način:

$$\text{Ekonomski OADR} = \frac{\sum_{x=65}^w c(x)N(x)}{\sum_{x=0}^w y(x)N(x)} \quad (5)$$

U Jednadžbi 5 varijabla $c(x)$ označava potrošnju *per capita* u x dobi, $y(x)$ označava dohodak od rada u dobi x dok je $N(x)$ populacija u dobi x . Omjer prikazuje potrebe za resursima, odnosno potrošnju starije populacije u odnosu na resurse koje čini radni dohodak radno sposobnog stanovništva bez obzira na njihovu dob (UN, 2019.). Porast ekonomskog OADR-a ukazuje na povećanje efektivnih starijih potrošača po efektivnom radniku što može uzrokovati promjenu potražnje i financiranja dobara i usluga kao što su mirovine i zdravstvena skrb.

Prema UN-ovim podacima za 2019. godinu zemlje s najvećom ekonomskom ovisnosti starog stanovništva najčešće su u Europi i Sjevernoj Americi. Međutim, na prvom mjestu nalazi se Japan kao država s najstarijim stanovništvom, gdje je 78 osoba starijih od 65 godina ekonomski ovisno o 100 zaposlenih osoba. Slijede Finska, Francuska, Njemačka i Danska koje se nalaze u rasponu 53-57 ekonomski ovisnih starijih osoba na 100 zaposlenih. Rješenja koja UN navodi za održavanje određene razine potrošnje u starosti uključuju zapošljavanje starijih osoba, korištenje prihoda od vlastite imovine, dobivanje ekonomske podrške od obitelji ili javnih programa te smanjenje vlastite potrošnje.

2.4. Posljedice starenja stanovništva

Starenje stanovništva je proces u kojem se prosječna dob osobe u nekoj populaciji povećava, a udio starijih osoba u ukupnom broju stanovnika raste. Ono ima brojne implikacije na društvo, ekonomiju i politiku. Fultz (2010.) navodi da starenje stanovništva znači povećanje broja korisnika mirovinskih prava i veće rashode za zdravstvenu zaštitu, a istodobno i manji broj osoba u radnom kontingentu. Koliko god je produljenje ljudskog vijeka „pobjeda čovječanstva“ (Nejašmić i Toskić, 2013.) toliko je i proces koji za sobom povlači brojne negativne posljedice i izazove.

U literaturi se uz posljedice starenja stanovništva najčešće vežu pitanja smanjenja radne snage, opterećenosti mirovinskih sustava, financijske potpore, potrebe za prilagodbom politika te izazovi vezani uz zdravstvenu skrb, socijalnu sigurnost i društvene promjene.

I. Smanjenje radne snage

Smanjenjem stope fertiliteta smanjuje se broj radno sposobnih osoba, a povećanjem očekivane životne dobi raste broj osoba ovisnih o radno aktivnom stanovništvu. Milinković (1994.) navodi kako dugoročno starenje povećava opterećenje radno aktivnog stanovništva u obliku povećanja poreza i doprinosa pomoću kojih se osigurava određena razina socijalne sigurnosti stare populacije.

II. Opterećenost mirovinskih sustava

Povećanje stare populacije ima direktan utjecaj na mirovinske sustave jer smanjen broj radno aktivnog stanovništva treba financirati sve veću količinu umirovljenika. Samodol (2021.) navodi opće pravilo prilikom povezivanja demografskih trendova i mirovinskih sustava: neovisno o strukturi mirovinskog sustava, bilo kakav nedostatak radne snage uzrokovan starenjem stanovništva vrlo će vjerojatno smanjiti razinu mirovina. To dovodi do potrebe za prilagođavanjem politika kako bi se osigurala financijska održivost.

III. Prilagodba politika

Kao posljedica starenja javlja se potreba za reformom mirovinskog sustava. Države moraju prilagoditi svoje politike zapošljavanja kako bi se nosile s demografskim promjenama. Jedan od načina uključuje poticanje starijih radnika na produženje radne dobi i fleksibilnijih oblika zaposlenja. Međutim, stariji radnici mogu iskusiti smanjenje produktivnosti zbog fizičkih i kognitivnih ograničenja povezanih sa starenjem.

Prilagodba politika također uključuje i mjere za poticanje nataliteta, ulaganje u zdravstvenu skrb i politike za sprječavanje odljeva mozgova.

IV. Opterećenost zdravstvenog sustava

S rastom životne dobi povećava se potreba za zdravstvenom skrbi. Starije osobe imaju složenije zdravstvene potrebe koje zahtijevaju multidisciplinarnu zdravstvenu skrb, imaju veći rizik od invaliditeta te su podložnije depresiji, demenciji i drugim mentalnim bolestima. Osiguranje adekvatnog planiranja i financiranja skrbi za starije osobe ključno je za osiguranje kvalitete iste.

V. Društveni izazovi

Povećanje životne dobi sa sobom nosi potrebe za različitim oblicima skrbi osim financijske – obiteljske i socijalne. Obiteljska dinamika mijenja se zbog povećanog „tereta“ za članove obitelji koji brinu o starijima. Starije stanovništvo zahtjeva podršku, a s obzirom na smanjenje broja djece u obitelji raste rizik od socijalne izolacije i usamljenosti.

Sve posljedice starenja stanovništva očituju se u financijskim izdacima i naglašavaju važnost planiranja, prilagodbe politika i resursa te promicanje inkluzivnih i održivih rješenja kako bi se osigurala dobrobit starijih osoba i cjelokupnog društva.

3. OTKRIVANJE ZNANJA U BAZAMA PODATAKA

3.1. Proces otkrivanja znanja iz baza podataka

Otkrivanje znanja u bazama podataka naziva se još i rudarenjem podataka (engl. *data mining*), a uključuje otkrivanje korisnih informacija, obrazaca i zakonitosti iz velike količine podataka.

Kantardžić (2011.) opisuje rudarenje podataka kao otkrivanje različitih modela, sažetaka i izvedenih vrijednosti iz dane zbirke podataka. Tehnika otkrivanja znanja u bazama podataka koristi statističke i matematičke metode, strojno učenje, algoritme i tehnike izvlačenja znanja kako bi pronašla skrivene uzorke, trendove, povezanosti ili pravilnosti koje se mogu primijeniti u donošenju poslovnih odluka, istraživanju, marketingu i drugim područjima.

Rudarenje podataka uključuje korištenje sofisticiranih alata za analizu podataka u otkrivanju prethodno nepoznatih, valjanih obrazaca i odnosa u velikim skupovima podataka (Seifert, 2006). Kantardžić (2011.) također navodi kako je rudarenje podataka zajednički napor ljudi i računala, a najbolji se rezultati postižu upravo spajanjem znanja stručnjaka u opisivanju problema i ciljeva s mogućnostima pretraživanja računala. Cilj rudarenja podataka pretvoriti je sirove podatke u koristan skup koji može pridonijeti donošenju informiranih odluka i stvaranju vrijednosti.

Pejić Bach i Kerep (2011.) navode kako se proces otkrivanja znanja iz baza podataka sastoji od četiri koraka:

1. Definiranje poslovnog problema

Za vjerodostojno istraživanje potrebno je imati jasno definiran problem i definirane ciljeve za koje se očekuje da će proces rudarenja podataka postići (Gupta, 2014.). Kantardžić (2011.) ističe važnost znanja i iskustva vezanih uz područje istraživanja kako bi se formulirala smisljena definicija poslovnog problema. Dobro definirani ciljevi dovode do jednostavnije evaluacije projekta.

2. Priprema podataka

Parr Rud (2003.) navodi kako su koraci vazani uz pripremu podataka kritičniji za cjelokupni projekt nego tehnika korištena za izradu modela. U praksi, proces pronalaska i pripreme podataka zna biti teži i dugotrajniji od procesa modeliranja. Potrebno mu je posvetiti posebnu pažnju jer loša kvaliteta podataka može dovesti do netočnih rezultata i interpretacije.

Tijekom pripreme, potrebno je provjeriti da nema nepravilnosti kao što su prazne ćelije te transformirati podatke kako bi se stvorili novi atributi ili skupovi podataka relevantni za analizu. Nadalje, ako su podaci u različitim mjerilima ili rasponima, skaliraju se kako bi se osiguralo da svi atributi imaju jednak utjecaj na analizu. Podaci s anomalijama ili ekstremima (tzv. *outliers*) otklanjanju se kako bi rezultati bili što vjerodostojniji.

3. Modeliranje

Proces modeliranja odnosi se na razvoj i testiranje modela koji će se koristiti za otkrivanje obrazaca, predviđanje budućih događaja ili donošenje informiranih odluka na temelju analiziranih podataka. Prvi korak u modeliranju čini odabir odgovarajuće metode. Prema Pejić Bach i Kerep (2011.) metode modeliranja su:

- klasifikacija (stablo odlučivanja, neuronske mreže i logistička regresija),
- predviđanje vrijednosti (linearna regresija, neuronske mreže i analiza vremenskih serija) i
- grupiranje (klaster analiza).

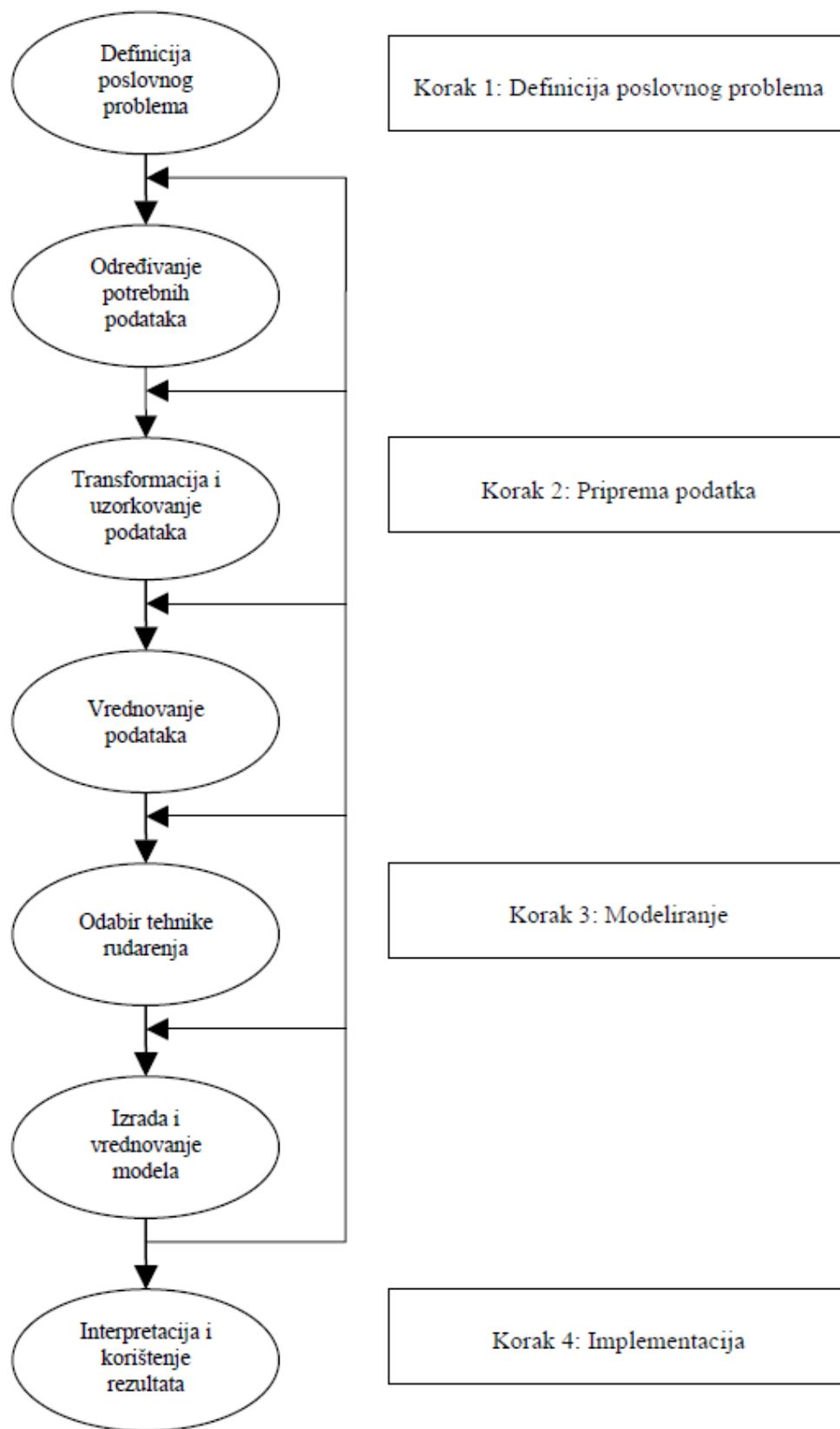
Svaka od navedenih metoda ima svoje prednosti i slabosti te će bolje funkcionirati u različitim kontekstima. Važno je zbog toga testirati i usporediti različite metode kako bi se odabrala najprikladnija za konkretan problem.

4. Implementacija

Posljednji je korak u procesu otkrivanja znanja u bazama podataka implementacija koja obuhvaća interpretaciju dobivenih rezultata i njihovu primjenu.

Cilj rudarenja podataka pomoći je u donošenju odluka, stoga se modeli moraju moći interpretirati kako bi bili korisni (Kantardžić, 2011.). Podatke je potrebno sažeti i vizualizirati jer dobro prezentirani podaci olakšavaju donošenje odluka te potiču provedbu sličnih procesa u budućnosti. Nakon interpretacije, donositelji odluka mogu implementirati model za korištenje u informacijski sustav, a jedan od načina može biti i razvoj softvera.

Slika 1 Proces otkrivanja znanja iz baza podataka



Izvor: Pejić Bach, M. (2005.) Rudarenje podataka u bankarstvu. Zbornik Ekonomskog fakulteta u Zagrebu, 3 (1), 181-193. Preuzeto s <https://hrcak.srce.hr/26220> (2005., str. 184.)

3.2. Primjena otkrivanja znanja iz baza podataka

Oko nas se svakodnevno generira velika količina podataka. S obzirom na to da je takva količina nepregledna, od nje nema prevelike koristi ako ju se pravilno ne obradi i analizira. Gupta (2014.) ističe kako je tehnike rudarenja moguće koristiti za manje količine podataka, ali što su podaci veći, veća je mogućnost pronalaska nečeg novog i zanimljivog.

Kantardžić (2011.) navodi kako je potreba za razumijevanjem velikih složenih skupova podataka bogatih informacijama zajednička gotovo u svim područjima poslovanja, znanosti i inženjerstva. Objašnjava da su korporativni podaci i podaci o klijentima prepoznati kao strateška imovina.

Rudarenje podataka rasprostranjeno je i u privatnom i u javnom sektoru. Omogućuje otkrivanje skrivenih obrazaca, donošenje informiranih odluka i dobivanje dubljeg uvida iz velike količine podataka. U energetskom se sektoru uz pomoć rudarenja podataka optimizira potrošnja energije te predviđa potrošnja i upravljanje resursima. Magazzino i sur. (2021.) su tako uz pomoć strojnog učenja analizirali odnos između obnovljivih izvora energije i ekonomskog rasta u vrijeme pandemije koronavirusa u Brazilu. Brazil je jedan od najvećih proizvođača i potrošača obnovljive energije u svijetu i zbog toga bilježi značajan gospodarski rast. Međutim, dolaskom koronavirusa bile su neizvjesne posljedice na gospodarski rast. Uz pomoć umjetnih neuronskih mreža zaključilo se kako je Brazil tijekom pandemije mogao očekivati negativne učinke koji će se odraziti na ekonomski sustav. Ovim su se istraživanjem dale smjernice donositeljima odluka vezane uz pravilno formiranje politika i mjera za ublažavanje negativnih posljedica.

U znanstvenim istraživanjima otkrivanje znanja iz baza podataka analizira eksperimentalne rezultate i otkriva nove veze u kompleksnim sustavima, a u zdravstvenom se sektoru rudarenje podataka koristi pri dijagnosticiranju bolesti, praćenju epidemija i personaliziranju medicinske skrbi. Zbog opterećenosti jedinica intenzivne njege u bolnicama, strojno učenje pruža podršku i olakšava zdravstvenim djelatnicima donošenje odluka. Shillan i sur. (2019.) strojnim učenjem predviđaju smrtnost, poboljšavaju prognostičke modele i klasificiraju subpopulaciju. Cilj im je predvidjeti komplikacije koje se svakodnevno događaju, duljinu liječenja i poboljšanje zdravlja. Međutim, zaključuje se je veličina uzorka istraživanja premala da bi se iskoristio puni potencijal metoda.

Rudarenje podataka i strojno učenje mogu biti snažan alat za analizu demografskih podataka. Iz demografskih je podataka moguće dobiti informacije korisne za sektore poput gospodarstva, zdravstva ili obrazovanja. Međutim, istražitelji često ne znaju koju metodu odabrati pa testiraju

više njih. Rose (2012.) koristi *super learner* pristup strojnog učenja koji kombinira više algoritama u jednom (takozvana *stacking* metoda). Rezultat je najbolji alat za predviđanje s najboljom srednjom kvadratnom pogreškom. Poboljšanja koja je *super learner* napravio u odnosu na R2 algoritam i neuronske mreže bila su otprilike dvostruka.

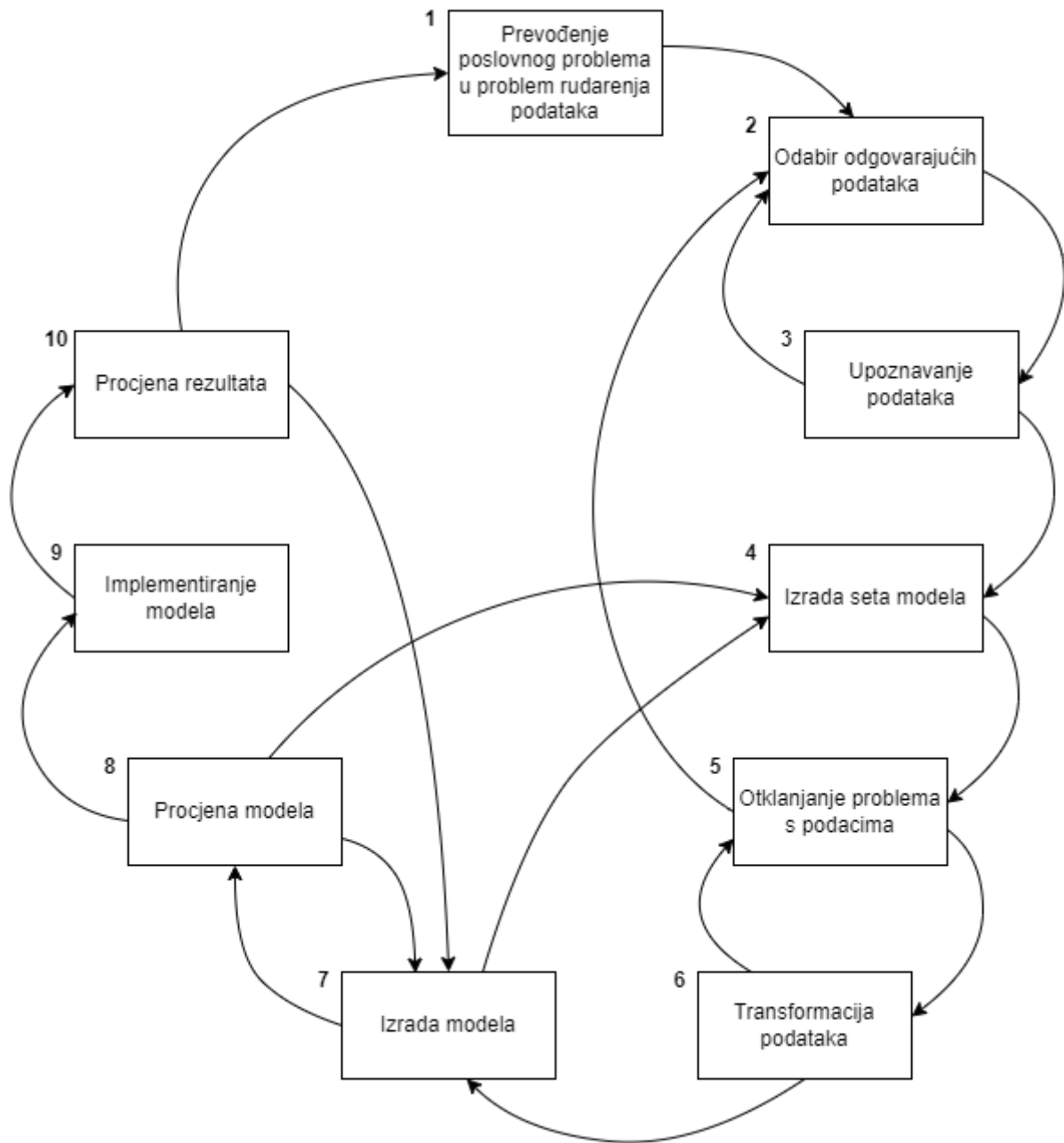
Rudarenje podataka ima mnoge primjene i u ekonomiji. Omogućava dublje razumijevanje ekonomske dinamike, identificira trendove, pomaže pri donošenju odluka i optimizira poslovne procese. Industrije poput bankarstva, osiguranja i maloprodaje redovito se služe rudarenjem podataka kako bi smanjile troškove, unaprijedile istraživanje i povećale prodaju (Seifert, 2006.). Ono može analizirati vremenske serije podataka kao što su podaci o inflaciji, kamatnim stopama, nezaposlenosti i BDP-u te predviđati ekonomsku stabilnost ili recesiju. Beutel i sur. (2019.) u svom radu uspoređuju učinkovitost različitih modela ranog upozoravanja na bankarske krize koristeći podatke razvijenih gospodarstava u posljednjih 45 godina. Navode kako su konvencionalni logistički modeli toliko napredovali u korištenju informacija da bi mogli predvidjeti financijsku krizu iz 2007./2008. godine. Modeli identificiraju skokove cijena imovine i vanjske oscilacije kao ključne prediktore bankarskih kriza. Rad sugerira potrebu za daljnjim poboljšanjima modela strojnog učenja za rano upozoravanje kako bi se dostigla razina mogućnosti predviđanja buduće bankarske krize.

Poduzećima pomaže prilikom analize potrošačkih preferencija, trendova kupovine i konkurentske dinamike, što je važno kod strategije prodaje i marketinga. Osim toga, služi za upravljanje zalihama, formiranje cijena, identifikaciju rizika te analizu performansi investicija. Navedeni primjeri ilustriraju samo neke od mogućnosti korištenja. Stvarna primjena u ekonomiji vrlo je široka i donosi poboljšanje poslovnih rezultata.

3.3. Metode otkrivanja znanja iz baza podataka

Kantardžić (2011.) objašnjava rudarenje podataka kao ponavljajući proces. Podaci se proučavaju i ispituju pomoću određene analitičke tehnike, a ako ih se ispita na drugačiji način ili modificira, dolazi do drugačijih podataka. Kada bi se vratili na početak i na istim podacima primijenili drugi alat za analizu podataka, dobili bi se bolji ili drugačiji rezultati. Takav se proces može primijeniti nebrojeno mnogo puta, a svaka se tehnika koristi za ispitivanje malo drugačijih aspekta podataka. Ono što se ovdje opisuje putovanje je otkrića koje moderno rudarenje podataka čini uzbudljivim.

Slika 2 Primjer procesa otkrivanja znanja iz baza podataka



Izvor: autorski rad prema Berry, M.J.A., Linoff, G.S. (2011). Mastering Data Mining. Wiley, Chichester (2011., str. 160.)

Hann i sur. (2012.) navode kako je rudarenje podataka spoj više disciplina. Pri tome navode kako ono obuhvaća strojno učenje, statistiku, prepoznavanje uzoraka, obradu prirodnog jezika, baze podataka, vizualizaciju i interakciju čovjeka s računalom, algoritme, računalstvo, društvene znanosti i sl. Ističu kako upravo interdisciplinarna priroda istraživanja i razvoja rudarenja podataka pridonosi uspjehu i širokoj primjeni.

Pejić Bach i Kerep (2011.) navode tri metode za otkrivanje znanja u bazama podataka: klasifikaciju, predviđanje vrijednosti i grupiranje.

Klasifikacija je proces razvrstavanja instanci u predefinirane kategorije ili klase na temelju njihovih karakteristika ili značajki. Postoji nekoliko metoda za klasifikaciju, a najčešće su stablo odlučivanja (*engl. Decision tree*), neuronske mreže, linearni klasifikator, Bayesov klasifikator i metoda k-najbližih susjeda.

Metode za predviđanje vrijednosti još su poznate kao regresijske metode, a koriste se kako bi predvidjele numeričke ili kontinuirane vrijednosti na temelju značajki ili karakteristika. Najčešće korištena je linearna regresija, ali koriste se još i metoda neuronskih mreža kao i analiza vremenskih serija.

Metode za otkrivanje podataka u literaturi se spominju i kao metode za grupiranje ili klasteriranje, a koriste se za identificiranje prirodnih skupova sličnih instanci ili obrazaca unutar skupa podataka. Primjeri takvih metoda su klaster analiza, asocijativna pravila i metoda tržišne košarice.

S obzirom na to da će se u ovom radu koristiti klaster analiza i analiza vremenskih serija, one će u nastavku biti detaljnije opisane.

3.3.1. Klaster analiza

Klaster analiza (*engl. Cluster Analysis*) tehnika je koja se koristi za grupiranje sličnih instanci ili podataka zajedno u klasterima ili skupinama. Na primjer, ako bismo skupili niz kamenčića s obale potoka, zabilježili njihove attribute veličine, oblika i boje te razvrstali slične kamenčiće u iste hrpe, fizički bismo izvodili analizu klastera, a svaka hrpa sitnih kamenčića bio bi jedan klaster (Romesburg, 2004). Aristotel je tako klasificirao životinje podijelivši ih na one koje imaju crvenu krv (kralježnjaci) i one koji nemaju (beskralježnjaci). Kasnije ih je dijelio prema načinu razmnožavanja i drugim karakteristikama. Hann i sur. (2012.) kažu da se objekti klasteriraju ili grupiraju na temelju načela maksimiziranja unutarklasne sličnosti i minimiziranja međurazredne sličnosti. Klasteri objekata formirani su tako da objekti unutar klastera imaju veliku sličnost u usporedbi jedni s drugima, ali su prilično različiti od objekata u drugim klasterima. Klasifikacija predstavlja metodu prikladnu za organiziranje velikog skupa podataka radi lakšeg razumijevanja i učinkovitijeg dohvata informacija (Everitt i sur., 2001.).

Romesburg (2004.) objašnjava kako je klaster analiza primjenjiva u gotovo svim profesijama i najosnovnija je metoda za procjenu sličnosti. U ekonomiji je osobito korisna za identifikaciju homogenih skupina potrošača, ali i, kao u ovom radu, za usporedbu država po čimbenicima poput omjera ovisnosti starog stanovništva.

3.3.2. Analiza vremenskih serija

Prognoziranje vremenskih serija (*engl. Time Series Forecasting*) tehnika je analize podataka koja se koristi za predviđanje budućih vrijednosti u nizu podataka koji su uređeni prema vremenskom slijedu. Vremenski niz je skup podataka koji se prikupljaju ili mjere tijekom određenih vremenskih intervala kao što su dani, mjeseci ili godine. Bahovac i Crkvenac (1995.) kao glavni cilj prognoziranja vremenskih serija navode pronalaženje modela koji odražava obrasce ponašanja promatranog dinamičkog sustava te predviđanje njegova budućeg stanja temeljem poznatih stanja u sadašnjosti i prošlosti. Drugim riječima, ova se metoda koristi kako bismo gledanjem sadašnjosti i prošlosti pokušali shvatiti što će se dogoditi u budućnosti. Analizom vremenskih nizova pokušavaju se otkriti ponavljajući obrasci ili trendovi te se uz pomoć matematike stvaraju modeli koji objašnjavaju te obrasce.

Shmueli i Lichtendahl Jr. (2016.) objašnjavaju kako se predviđanje vremenskih serija provodi u gotovo svakoj organizaciji: maloprodajne trgovine predviđaju prodaju, energetske tvrtke predviđaju rezerve, proizvodnju, potražnju i cijene, obrazovne ustanove predviđaju upis, međunarodne financijske organizacije poput Svjetske banke i Međunarodnog monetarnog fonda predviđaju inflaciju i gospodarsku aktivnost. Tvrtke za prijevoz putnika koriste vremenske serije za predviđanje budućih putovanja dok tvrtke rizičnog kapitala predviđaju tržišni potencijal za procjenu poslovnih planova.

Ovisno o vrsti vremenskog niza, odabir pravilnog modela i parametara može predstavljati velik izazov. Ako se vremenski niz ponaša nepredvidivo ili ima složene uzroke, može biti potrebno koristiti napredne tehnike poput neuronskih mreža.

3.4. Alati za otkrivanje znanja u bazama podataka

Rudarenje podataka u osnovi označava pronalazak korisnih informacija u velikim bazama podataka. Ono se obavlja raznim tehnikama uz pomoć programa za rudarenje podataka. Cilj je pronaći skrivene obrasce i informacije koje nam mogu pomoći razumjeti što se događa i predvidjeti što bi se moglo dogoditi u budućnosti.

Alati za otkrivanje znanja u bazama podataka uz pomoć strojnog učenja, matematike, statistike i vizualizacije ljudima prikazuju informacije na razumljiv način. Oni filtriraju skrivene obrasce, trendove i informacije unutar velikih skupina podataka te omogućuju korisnicima da donose proaktivne odluke temeljene na znanju.

Na tržištu postoji velik broj komercijalnih, kao i programa otvorenog koda koji se koriste za rudarenje. Od komercijalnih alata najčešće su spominjani *Statistica Data Miner*, *SAS Enterprise Miner*, *Synapse* i *Advanced Miner*, dok su najpoznatiji besplatni alati *Rapid Miner*, *Alpha Miner*, *Weka*, *Orange*, *R* i *Kinme*. Za potrebe istraživanja u ovom radu, koristit će se program *Weka*.

Weka (*Waikato Environment for Knowledge Learning*) je program razvijen na Sveučilištu Waikato na Novom Zelandu. To je aplikacija otvorenog koda potpuno implementirana u programskom jeziku *Java* i stoga se izvodi na gotovo svakoj suvremenoj računalnoj platformi. Sadrži zbirku alata za pripremu podataka, klasteriranje, klasifikaciju, regresiju, vizualizaciju i prediktivno modeliranje. Njezina prednost leži u dostupnosti alata kao i jednostavnosti uporabe u otkrivanju znanja u bazama podataka.

4. ANALIZA OVISNOSTI STAROG STANOVNIŠTVA KORIŠTENJEM OTKRIVANJA ZNANJA IZ BAZA PODATAKA

4.1. Metodologija istraživanja

Alat korišten za istraživanje u ovom radu bit će program otvorenog koda *Weka* pomoću kojeg će se provoditi klaster analiza i analiza vremenskih serija.

U klaster analizi koristit će se skup podataka Razvojnog programa Ujedinjenih naroda u kojem su sadržani podaci o omjeru ovisnosti starijeg stanovništva za sve države svijeta u 2019. godini. Države će biti podijeljene u tri klastera: one s visokim, srednjim i niskim omjerom ovisnosti. Cilj klaster analize je grupirati države prema omjeru ovisnosti te pozicionirati Hrvatsku na svjetskoj razini. Podaci istraživanja mogu se koristiti za usporedbu država, analiziranje onih s niskim omjerom ovisnosti starijeg stanovništva te obrazovanjem o metodama smanjenja omjera u budućnosti.

Prognoziranjem vremenskih serija prikazat će se predviđene vrijednosti omjera ovisnosti starijeg stanovništva u budućnosti. U ovom će se istraživanju koristiti podaci Organizacije za ekonomsku suradnju i razvoj (*OECD*), a analiza će se provoditi za Hrvatsku do 2050. godine. Cilj je predvidjeti kretanje omjera ovisnosti starog stanovništva kao demografskog čimbenika koji ima značajan utjecaj na život. Pomoću takvih istraživanja ističe se važnost pojava poput starenja stanovništva i promjena koje ono donosi.

4.2. Klaster analiza

4.2.1. Korišteni podaci

Izvor podataka korištenih za klaster analizu baza je podataka Razvojnog programa Ujedinjenih naroda (*engl. United Nations Development Programme*), točnije Izvješće o ljudskom razvoju iz 2020. godine (*engl. Human Development Report 2020*). Iz navedene je baze izrađena nova tablica s podacima važnim za ovo istraživanje. Tablica 1 pokazuje odabrane attribute, njihove opise, formate, modalitete te minimalne i maksimalne vrijednosti i prosjeke.

Izvješće o ljudskom razvoju analizira ukupno 195 zemalja, no u ovom je istraživanju sudjelovalo njih 183. Preostale su države izdvojene zbog nedostatka podataka: Lihtenštajn,

Andora, Saint Kitts and Nevis, Dominica, Marshall Islands, Monako, Nauru, San Marino, Tuvalu, Koreja i Somalia.

Tablica 1 Atributi, formati i modaliteti za klaster analizu

Naziv atributa	Opis atributa	Format atributa (numerički, binomni, nominalni)	Minimum, maksimum i prosjeak atributa
Country	država	nominalni	
HDI	indeks ljudskog razvoja	numerički	min: 0.3944 max: 0.957 prosjeak: 0.72
Life expectancy at birth	očekivani životni vijek pri rođenju	numerički	min: 53.3 max: 84.9 prosjeak: 72.567
Total population	ukupna populacija (u milijunima)	numerički	min: 0.1 max: 1433.78 prosjeak: 41.748
Average annual population growth 2005/2021	prosječni rast populacije između 2005. i 2010. (izražen u postotku)	numerički	min: -1.4 max: 15.3 prosjeak: 1.624
Average annual population growth 2015/2020	prosječni rast populacije između 2015. i 2020. (izražen u postotku)	numerički	min: -1.5 max: 4.3 prosjeak: 1.293

Population under age 5	populacija mlađa od 5 godina (u milijunima)	numerički	min: 0 max: 116.8 prosjek: 3.666
Population ages 15-64	populacija u dobi 15-64 (u milijunima)	numerički	min: 0.1 max: 1014 prosjek: 27.251
Population ages 65 and older	populacija u dobi od 65 godina i starija (u milijunima)	numerički	min: 0 max: 164.5 prosjek: 3.798
Median population age	medijan dobi populacije	numerički	min: 15.2 max: 48.4 prosjek: 30.094
Young age (0-14) dependency ratio	omjer ovisnosti mladog stanovništva	numerički	min: 16 max: 104.8 prosjek: 45.399
Old age (65 and older) dependency ratio	omjer ovisnosti starijeg stanovništva	numerički	min: 1.4 max: 47.1 prosjek: 13.727

Izvor: autorska izrada

4.2.2. Rezultati istraživanja

Podaci za istraživanje preuzeti su u .xlsx obliku, a potom spremljeni u .csv oblik. Kako bi bili prikladni za obradu u Weki, bilo ih je potrebno urediti pomoću Bloka za pisanje. Podaci su uređeni na način da su države s nepotpunim podacima uklonjene iz istraživanja, uklonjene su praznine i razmaci, decimalni zarezi (.) zamijenjeni su točkama (.), a točke sa zarezom (;) zamijenjene su zarezima (,).

Za bolju preglednost, podaci odvojeni zarezom kasnije su stavljeni u odvojene ćelije u Excelu odabirom stupca A i odlaskom na karticu Podaci, zatim Tekst u stupce te odabirom zareza (,) kao graničnika.

Za početak, podatke je potrebno u *Weki* pretvoriti u .arff format. Ulaskom u program *Weka* otvara se početni izbornik s opcijama *Explorer*, *Experimenter*, *KnowledgeFlow*, *Workbeench* i *Simple CLI*. U ovom će radu bit korištena opcija *Explorer* sa panelima: *Preprocess*, *Cluster*, i *Forecast*.

Analiza započinje učitavanjem .csv datoteke u panelu *Preprocess*. Klikom na *Open file* i odabirom željene datoteke podaci se učitavaju u program te se potom u gornjem desnom kutu spremaju u .arff oblik (*engl. Attribute-Relation File Format*) kao što je prikazano na Slici 3. Nakon toga se ponovno učitavaju u *Weku* na isti način, samo ovog puta u .arff umjesto u .csv formatu.

Slika 3 Weka – panel *Preprocess* za klaster analizu

The screenshot displays the Weka Explorer interface in the *Preprocess* tab. The main window title is "Weka Explorer". The menu bar includes "Preprocess", "Classify", "Cluster", "Associate", "Select attributes", "Visualize", and "Forecast". Below the menu bar are buttons for "Open fil...", "Open UR...", "Open D...", "Generate...", "Undo", "Edit...", and "Save...".

The "Filter" section shows a "Choose" button and a dropdown menu set to "None", with "Apply" and "Stop" buttons.

The "Current relation" section shows: "Relation: Klaster analiza - u...", "Attributes: 12", "Instances: 183", and "Sum of weights: 183".

The "Attributes" section has buttons for "All", "None", "Invert", and "Pattern". A list of 12 attributes is shown, with the 12th attribute, "Old_age_(65_and_older)_dependency_ratio", selected and highlighted in blue. A "Remove" button is located below the list.

The "Selected attribute" section shows: "Name: Old_age_(65_and_older)_dep...", "Type: ...", "Missing: 0 (0%)", "Distinct: 123", and "Unique: 8...". Below this is a table of statistics:

Statistic	Value
Minimum	1.4
Maximum	47.1
Mean	13.727
StdDev	9.764

Below the statistics table is a dropdown menu set to "Class: Old_age_(65_and_older)_dep..." and a "Visualize All" button. A histogram is displayed below, showing the distribution of values for the selected attribute. The x-axis ranges from 1.4 to 47.1. The y-axis represents frequency, with bars labeled with their counts: 67, 46, 19, 15, 18, 15, 2, and 1.

The status bar at the bottom shows "Status OK" and a "Log" button.

Izvor: autorski rad

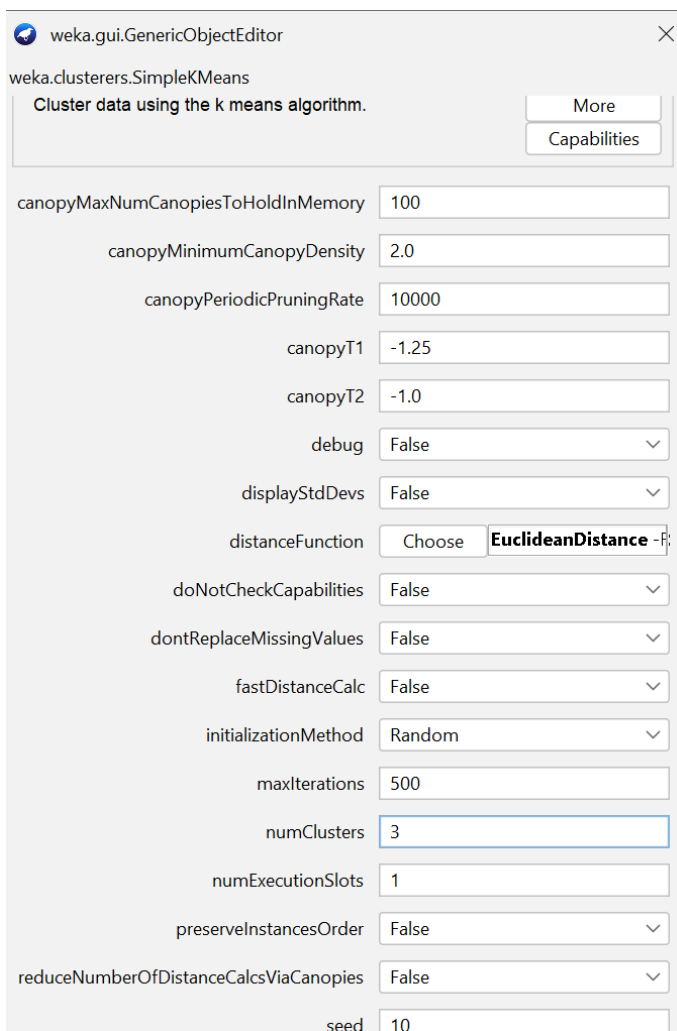
Na Slici 3 primjećuje se kako panel *Preprocess* s lijeve strane prikazuje atribute, a s desne njihove formate, modalitete te minimalne i maksimalne vrijednosti ovisno o tome radi li se o nominalnim ili numeričkim vrijednostima, kao i grafički prikaz u donjem desnom dijelu sučelja.

Po završetku *Preprocess* procesa odabire se željena metoda analize, u ovom slučaju klaster analiza. Odlaskom na panel *Cluster* otvaraju se mogućnosti klaster analize. Za potrebe istraživanja u odjeljku *Clusterer* odabire se opcija *SimpleKMeans* s odabranim vrijednostima

prikazanim na Slici 4. Algoritam KMeans jednostavan je za korištenje i najbolje radi kada podaci imaju jasne klasterne. Rezultat koji daje slične su točke podudaranja u grupama klastera. Kako bi rezultati bili vjerodostojni, potrebno je odabrati pravilan broj klastera. Ovisno o istraživanju, on može biti odabran po želji ili uz pomoć jednadžbe prikazane ispod, u kojoj se broj klastera dobije dijeljenjem broja atributa (n) s dva. Procjenom potreba ovog istraživanja, države će biti podijeljene u tri klastera.

$$k \approx \sqrt{\frac{n}{2}} \quad (6)$$

Slika 4 Postavke algoritma SimpleKMeans



Izvor: autorski rad

U odjeljku *Cluster mode* odabire se opcija *Classes to clusters evaluation*, a ispod se nalazi opcija za odabir atributa koji će se ignorirati u istraživanju. Atribut korišten u istraživanju je *Old age (65 and older) dependency ratio*, odnosno omjer ovisnosti starijeg stanovništva, a ignorirani su ostali atributi:

- država,
- indeks ljudskog razvoja,
- očekivani životni vijek pri rođenju,
- ukupna populacija,
- prosječni rast populacije između 2005. i 2010.,
- prosječni rast populacije između 2015. i 2020.,
- populacija mlađa od 5 godina,
- populacija u dobi 15-64 godine,
- populacija u dobi od 65 godina i starija,
- medijan dobi populacije,
- omjer ovisnosti mladog stanovništva,
- omjer ovisnosti starijeg stanovništva.

Odabirom algoritma i opcija te pritiskom na gumb *Start* vrši se proces klasteriranja.

U odjeljku *Output* prikazani su rezultati istraživanja. Države su podijeljene u klaster te kraj sebe imaju oznake 0 i 1. Oznaka 0 označava nepripadnost, dok 1 označava pripadnost države pojedinom klasteru. Slika 5 daje vizualni prikaz pripadnosti država klasterima. X i y osi prikazuju odabrane atribute koje je moguće mijenjati radi lakšeg razumijevanja odnosa između atributa. Na x osi prikazan je omjer ovisnosti starijeg stanovništva koji se kreće u rasponu od 1.4% do 47.1%, a na osi y prikazane su države. Svaki je klaster radi preglednosti označen određenom bojom pa je tako klaster 1 označen crvenom, klaster 2 zelenom te klaster 0 plavom bojom. Klikom na određenu točku na grafu prikazuju se informacije o njoj kao što je prikazano na Slici 6.

Slika 5 Vizualni prikaz pripadnosti klasteru po omjeru ovisnosti starog stanovništva



Izvor: autorski rad

Slika 6 Informacije o točki na grafu

```
Weka: Instance info
Plot : 19:11:09 - SimpleKMeans (Uređeni podaci)
Instance: 42
Instance_number : 41.0
Country : Croatia
HDI : 0.851
Life_expectancy_at_birth : 78.5
Total_population : 4.13
Average_annual_population_growth_2005/2010 : -0.2
Average_annual_population_growth_2015/2020 : -0.6
Population_under_age_5 : 0.2
Population_ages_15-64 : 2.7
Population_ages_65_and_older : 0.9
Median_population_age : 44.3
Young_age_(0-14)_dependency_ratio : 22.6
Old_age_(65_and_older)_dependency_ratio : 32.3
Cluster : cluster0
```

Izvor: autorski rad

U Tablici 2 prikazane su države raspoređene u klaster. Klaster 0 predstavlja države s visokim omjerom ovisnosti starog stanovništva te u njega pripada 45 država, odnosno 24,59%. Klaster započinje s vrijednosti 23,1% i obuhvaća sve države koje imaju omjer ovisnosti od navedenog broja do maksimalne vrijednosti, a vrijednost centroida klastera iznosi 28,7867.

Klaster 1 označava države s niskim omjerom ovisnosti starog stanovništva čije se vrijednosti kreću od minimalnog omjera koji iznosi 1,4% do omjera 9,9%. Ovaj klaster obuhvaća 94 od 183 države, odnosno 51,37%. Vrijednost centroida klastera iznosi 6,2266.

Klasteru 2 pripadaju države sa srednjim omjerom ovisnosti starijeg stanovništva koji se kreće u rasponu između 10% i 23%. Centroid klastera iznosi 14,3477, a u njemu su obuhvaćene 44 države, odnosno njih 24,04%.

Tablica 2 Pripadnost država klasteru

Visoka ovisnost starijeg stanovništva	Niska ovisnost starijeg stanovništva	Srednja ovisnost starijeg stanovništva
Australia	Afghanistan	Albania
Austria	Angola	Algeria
Barbados	Azerbaijan	Antigua and Barbuda
Belarus	Bahrain	Argentina
Belgium	Bangladesh	Armenia
Bosnia and Herzegovina	Belize	Bahamas
Bulgaria	Benin	Bolivia
Canada	Bhutan	Brazil
Croatia	Botswana	Chile
Cuba	Brunei Darussalam	China
Czechia	Burkina Faso	Colombia
Denmark	Burundi	Costa Rica
Estonia	Cabo Verde	Cyprus
Finland	Cambodia	Dominican Republic
France	Cameroon	Ecuador
Georgia	Central African Republic	El Salvador
Germany	Chad	Grenada
Greece	Comoros	Guyana
Hong Kong	Congo	Israel
Hungary	Cote d'Ivoire	Jamaica
Iceland	Democratic Republic of Congo	Kazakhstan
Ireland	Djibouti	Korea
Italy	Egypt	Lebanon
Japan	Equatorial Guinea	Luxembourg
Latvia	Eritrea	Mauritius
Lithuania	Eswatini	Mexico
Malta	Ethiopia	Moldova
Montenegro	Fiji	Morocco
Netherlands	Gabon	North Macedonia
New Zealand	Gambia	Panama
Norway	Ghana	Paraguay
Poland	Guatemala	Peru
Portugal	Guinea	Saint Lucia
Romania	Guinea Bissau	Saint Vincent and the Grenadines
Russian Federation	Haiti	Seychelles
Serbia	Honduras	Singapore
Slovakia	India	Sri Lanka
Slovenia	Indonesia	Suriname
Spain	Iran	Thailand
Sweden	Iraq	Trinidad and Tobago
Switzerland	Jordan	Tunisia
UK	Kenya	Turkey

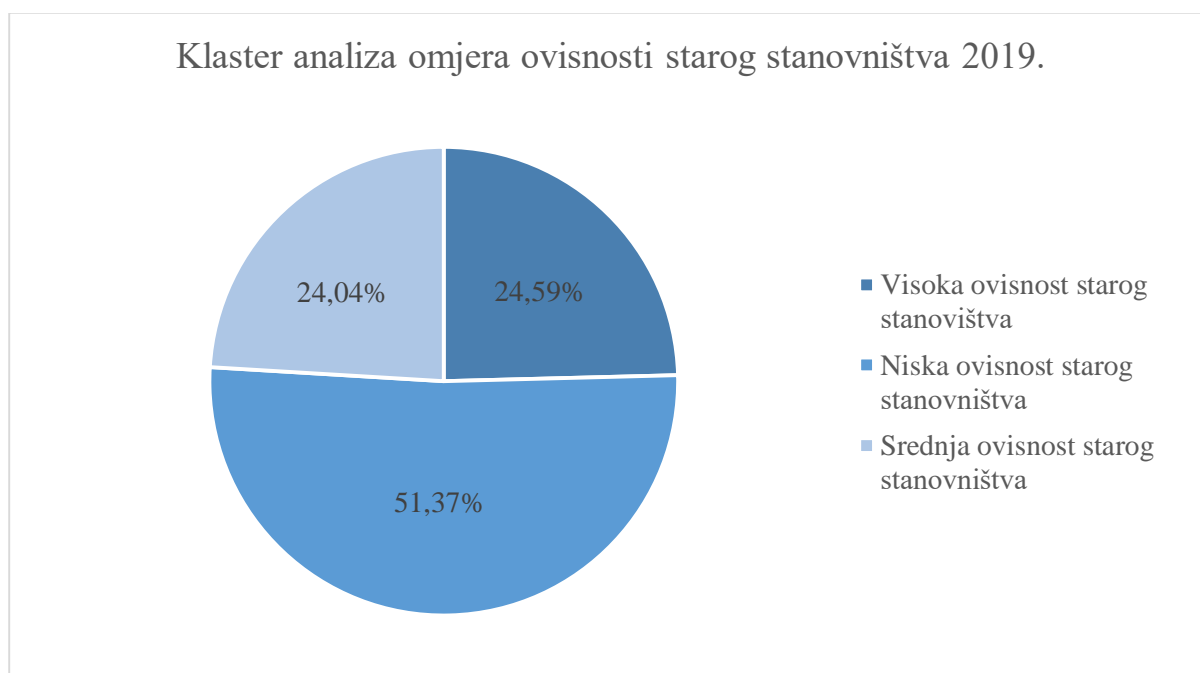
Ukraine	Kiribati	Venezuela
United States	Kuwait	Vietnam
Uruguay	Kyrgyzstan	
	Lao People's Democratic Republic	
	Lesotho	
	Liberia	
	Libya	
	Madagascar	
	Malawi	
	Malaysia	
	Maldives	
	Mali	
	Mauritania	
	Micronesia	
	Mongolia	
	Mozambique	
	Myanmar	
	Namibia	
	Nepal	
	Nicaragua	
	Niger	
	Nigeria	
	Oman	
	Pakistan	
	Palestine	
	Papua New Guinea	
	Philippines	
	Qatar	
	Rwanda	
	Samoa	
	Sao Tome and Principe	
	Saudi Arabia	
	Senegal	
	Sierra Leone	
	Solomon Islands	
	South Africa	
	South Sudan	
	Sudan	
	Syrian Arab Republic	
	Tajikistan	
	Tanzania	
	Timor Leste	
	Togo	
	Tonga	
	Turkmenistan	
	UAE	

	Uganda	
	Uzbekistan	
	Vanuatu	
	Yemen	
	Zambia	
	Zimbabwe	

Izvor: autorski rad

Postotak država po klasterima prikazan je na Grafikonu 1. Kao što je navedeno, više od polovine država ima nizak omjer ovisnosti starijeg stanovništva, no na popisu država u klasteru 1 se primjećuje kako je uglavnom riječ o slabije razvijenim zemljama. Područja u kojima još uvijek traje proces demografske tranzicije zbog slabije razvijenosti imaju niske stope fertiliteta i visoke stope mortaliteta stoga je omjer ovisnosti starijeg stanovništva nizak. Državama s visokim omjerom starijeg stanovništva pripada veći dio Europe, uključujući Hrvatsku, a najveću ovisnost, od 47,1%, ima Japan. Razlozi su tako visokog omjera niska stopa fertiliteta, produženi životni vijek i dobro razvijeni sustavi socijalne zaštite. Visok omjer ovisnosti starijeg stanovništva predstavlja izazove za održivost mirovinskih sustava i pružanje adekvatne skrbi za starije osobe.

Grafikon 1 Veličina klastera



Izvor: autorski rad

4.3. Analiza vremenskih serija

4.3.1. Korišteni podaci

Podaci korišteni za analizu vremenskih serija preuzeti su sa stranice Organizacije za ekonomsku suradnju i razvoj (*engl. Organization for Economic Co-operation and Development*). OECD baza podataka sadrži podatke o ukupnom stanovništvu, mladom i starom stanovništvu, stanovništvu u radnoj dobi kao i podatke o ovisnosti starijeg stanovništva za sve države svijeta. Za potrebe ovog istraživanja preuzeta je baza podataka europskih država sa podacima od 2000. do 2023. godine. Podaci su preuzeti u .csv formatu i uređeni na isti način kao za klaster analizu. Osim toga, uklonjene su sve države osim Hrvatske. Cilj analize je otkriti trend promjene, odnosno predvidjeti kretanje omjera uz pomoć do sad prikupljenih podataka.

U Tablici 3 prikazani su atributi, njihovi formati i modaliteti. U analizi sudjeluju samo dva atributa zato što se prognoziraju vrijednosti omjera ovisnosti starijeg stanovništva u Hrvatskoj do 2050. godine.

Tablica 3 Atributi, formati i modaliteti za analizu vremenskih serija

Naziv atributa	Format atributa (numerički, binomni, nominalni)	Modaliteti atributa	Minimum, maksimum i prosjek atributa
Godine	numerički	Minimum, Maksimum, Mean, StdDev	min: 2000 max: 2023 prosjeck: 2011.5
Old_age_dep_ratio_(CRO)	numerički	Minimum, Maksimum, Mean, StdDev	min: 25.8 max: 39.8 prosjeck: 31.117

Izvor: autorski rad

4.3.2. Rezultat istraživanja

Za početak je potrebno instalirati proširenje u Weki koje omogućuje analize prognoziranje vremenskih serija. Pokretanjem Weke i odabirom opcije *Tools*, zatim *Package Manager* otvara se prozor s tražilicom u koju se upisuje riječ *forecasting*. Od ponuđenih paketa potrebno je instalirati opciju *timeseriesForecasting* i ponovno pokrenuti Weku kako bi panel bio vidljiv i spreman za korištenje.

Analiza vremenskih serija započinje na isti način kao i klaster analiza, učitavanjem skupa podataka u .csv formatu te spremanjem istih u .arff format. Bitno je napomenuti kako je za ovakvu vrstu istraživanja potrebno urediti podatke na način da države budu u stupcima s vrijednostima po godinama ispod njih. Nakon učitavanja datoteke u .arff obliku, odabire se novoinstalirani panel *Forecast*.

S obzirom na to da su Ujedinjeni narodi i Organizacija za ekonomsku suradnju i razvoj već proveli predviđanje omjera ovisnosti starijeg stanovništva za pojedine svjetske zemlje i regije, cilj ovog istraživanja bio je dobiti predviđanja omjera za Hrvatsku. Analiza vremenskih serija provedena je s omjerima ovisnosti starijeg stanovništva u Hrvatskoj u rasponu od 2000. do 2023. godine, a predviđanja su se vršila 27 godina u budućnost.

Odlaskom na panel *Forecast* predviđaju se vremenske serije. Pod odjeljkom *Basic configuration* odabiru se atributi iz modela na kojima će se provoditi analiza. Budući da su u ovom modelu izdvojeni podaci samo za Hrvatsku, odabran je atribut omjer ovisnosti starog stanovništva (*Old_age_dep_ratio_(CRO)*). Ako postoji više atributa, analizu je moguće istovremeno provoditi za više njih. U desnom dijelu, na kartici *Parameters*, podešavaju se postavke modela. Broj jedinica za predviđanje namješten je na 27 tako da se dobiju predviđanja do 2050. godine. Važno je napomenuti da što je veća vrijednost jedinica za predviđanje, to će predviđanje biti manje točno. Za vremensku oznaku predviđanja odabrane su godine kao i za periodičnost koja može poprimiti vrijednosti po satima, danima, tjednima, mjesecima, kvartalima i godinama. Osim toga, izvršenje procjene označeno je kvačicom kako bi se procijenio model koristeći podatke trening modela.

Osim osnovnih postavki (*engl. Basic configuration*), postoje i napredne postavke (*engl. Advanced configuration*).

Na kartici *Base learner* odabire se algoritam za analizu vremenskih serija, u ovom slučaju *Multilayer Perceptron*. To je neuronska mreža s više slojeva neurona, ulaznim, jednim ili više skrivenih i izlaznim neuronom. Algoritam analizira dane podatke te uz pomoć strojnog učenja generira outpute, odnosno predviđa kretanje danih podataka.

Na kartici *Output* postavljene su opcije izlaznih podataka tako da su za svaki korak prikazane stvarna i predviđena vrijednost kao i stupanj pogreške (*Output predictions at step*) te opcija *Output future predictions beyond end of series* kojom se prikazuju podaci za budućnost, odnosno podaci izvan dostupnih podataka. Za grafičke opcije odabrani su *Graph predictions at step* koji prikazuje podatke po vremenskim koracima i *Graph future predictions beyond end of series* na kojem će se vidjeti kretanje predviđanja omjera ovisnosti starog stanovništva u budućnosti.

U metrici na kartici *Evaluation* odabrane su opcije *Mean absolute error* (MAE) i *Root mean squared error* (RMSE). MAE je kratica za srednju apsolutnu pogrešku i procjenjuje koliko su prognoze ili modeli udaljeni od stvarnih vrijednosti. Što je srednja apsolutna pogreška manja znači da je manja pogreška u predviđanjima i to je model bolji. RMSE predstavlja prosječnu udaljenost između stvarnih i predviđenih vrijednosti u kvadratnom obliku i kao kod MAE, niža vrijednost ukazuje na bolju točnost modela.

Odabirom željenih postavki i pritiskom gumba Start provodi se predviđanje. U kartici *Output/Visualisation* vidljivi su rezultati provedene analize.

Kao što je već navedeno, algoritam *Multilayer Perceptron* ima ulazni, izlazni i jedan ili više skrivenih slojeva. Čvor 0 označava izlazni čvor, a na Slici 7 prikazan je *Sigmoid Node 7*, jedan od skrivenih čvorova.

Slika 7 Isječak procesa učenja u analizi vremenskih serija

Sigmoid Node 7	
Inputs	Weights
Threshold	-0.520559328406496
Attrib Godine	-0.1682287630408007
Attrib Lag_Old_age_dep_ratio_(CRO)-1	0.07868464795754773
Attrib Lag_Old_age_dep_ratio_(CRO)-2	0.22601407375472013
Attrib Lag_Old_age_dep_ratio_(CRO)-3	0.08589442361780243
Attrib Lag_Old_age_dep_ratio_(CRO)-4	0.15227973120886243
Attrib Lag_Old_age_dep_ratio_(CRO)-5	0.04319473551592846
Attrib Godine^2	-0.1384308629441236
Attrib Godine^3	-0.2091390710681013
Attrib Godine*Lag_Old_age_dep_ratio_(CRO)-1	0.15110956610460874
Attrib Godine*Lag_Old_age_dep_ratio_(CRO)-2	0.2102564598830242
Attrib Godine*Lag_Old_age_dep_ratio_(CRO)-3	0.14396098082058897
Attrib Godine*Lag_Old_age_dep_ratio_(CRO)-4	0.11528961102153544
Attrib Godine*Lag_Old_age_dep_ratio_(CRO)-5	0.03807759369515644

Izvor: autorski rad

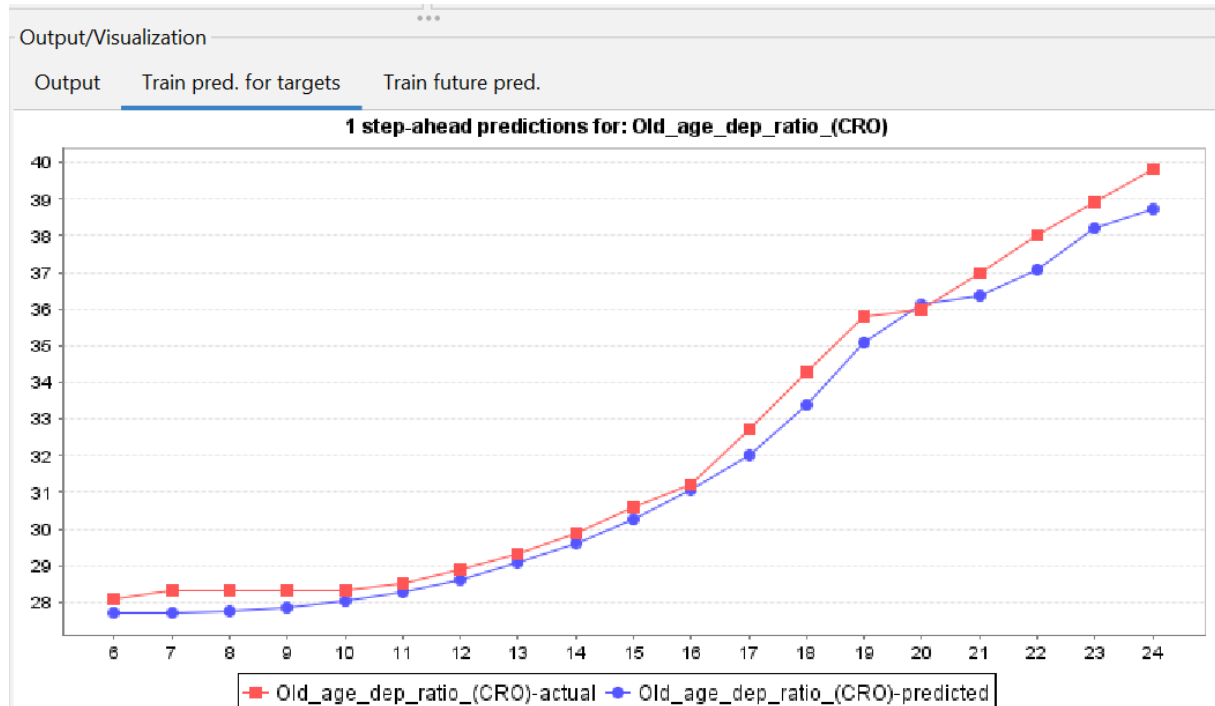
Tablica 4 Stvarne i predviđene vrijednosti i pogreške

Instanca	Stvarna vrijednost	Predviđena vrijednost	Pogreška
6	28,1	27,7098	-0,3902
7	28,3	27,7038	-0,5962
8	28,3	27,7493	-0,5507
9	28,3	27,8514	-0,4486
10	28,3	28,0215	-0,2785
11	28,5	28,2688	-0,2312
12	28,9	28,6213	-0,2787
13	29,3	29,0863	-0,2137
14	29,9	29,6186	-0,2814
15	30,6	30,2441	-0,3559
16	31,2	31,0868	-0,1132
17	32,7	32,0081	-0,6919
18	34,3	33,3728	-0,9272
19	35,8	35,0799	-0,7201
20	36	36,108	0,108
21	37	36,3623	-0,6377
22	38	37,0772	-0,9228
23	38,9	38,2112	-0,6888
24	39,8	38,7236	-1,0764

Izvor: autorski rad

U Tablici 4 prikazane su stvarne i predviđene vrijednosti omjera ovisnosti starijeg stanovništva kao i pogreške dobivene u analizi. Može se primijetiti kako instance započinju od broja šest, odnosno nedostaje prvih pet vrijednosti, a to je zbog toga što ih algoritam koristi za trening.

Slika 8 Grafički prikaz stvarnog i predviđenog kretanja omjera ovisnosti starog stanovništva



Izvor: autorski rad

Na Slici 8 nalazi se grafički prikaz stvarnog kretanja omjera ovisnosti starijeg stanovništva i vrijednosti predviđene u analizi. Stvarne vrijednosti prikazane su crvenom bojom, a plava boja prikazuje one koje je predvidjela *Weka*.

Nakon podataka koji dolaze iz učitane baze, algoritam navodi vrijednosti u budućnosti za ranije definirane godine. U Tablici 5 prikazane su stvarne vrijednosti omjera ovisnosti starijeg stanovništva, odnosno vrijednosti za razdoblje od 2000. do 2023. godine, a u Tablici 6 navode se predviđanja do 2050. godine. Vrijednosti označene zvjezdicom (*) predstavljaju predviđanja zadanog atributa u budućnosti koja su postavljena na 27 godina.

Tablica 5 Omjer ovisnosti starijeg stanovništva od 2000. do 2023.

Godina	Omjer ovisnosti starog stanovništva
2000	25,80
2001	26,40
2002	27,00
2003	27,50
2004	27,90
2005	28,10
2006	28,30
2007	28,30
2008	28,30
2009	28,30
2010	28,50
2011	28,90
2012	29,30
2013	29,90
2014	30,60
2015	31,20
2016	32,70
2017	34,30
2018	35,80
2019	36,00
2020	37,00
2021	38,00
2022	38,90
2023	39,80

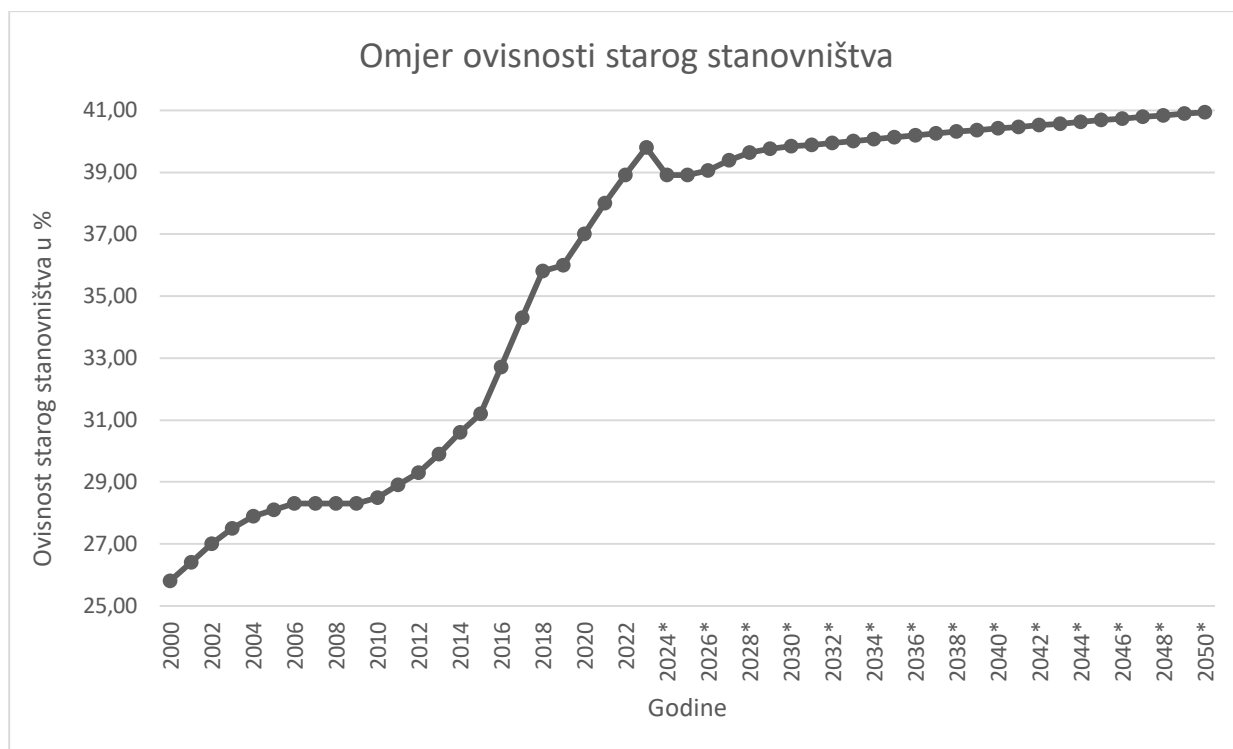
Izvor: autorski rad

Tablica 6 Predviđene vrijednosti omjera ovisnosti starijeg stanovništva

Godina	Omjer ovisnosti starog stanovništva
2024*	38,91
2025*	38,91
2026*	39,06
2027*	39,39
2028*	39,63
2029*	39,76
2030*	39,84
2031*	39,89
2032*	39,94
2033*	40,00
2034*	40,07
2035*	40,13
2036*	40,19
2037*	40,25
2038*	40,30
2039*	40,36
2040*	40,41
2041*	40,47
2042*	40,52
2043*	40,57
2044*	40,63
2045*	40,68
2046*	40,73
2047*	40,78
2048*	40,84
2049*	40,89
2050*	40,94

Izvor: autorski rad

Grafikon 2 Kretanje omjera ovisnosti od 2000. do 2050. godine



Izvor: autorski rad

Grafikon 2 vizualizira kretanje omjera ovisnosti starijeg stanovništva do 2050. godine. Iz podataka se primjećuje kako je omjer od 2000. godine do danas u kontinuiranom porastu, a predviđanja pokazuju da će se taj trend nastaviti. Od 2000. do 2014. godine indeks se godišnje u prosjeku povećavao za 0,36%, a nakon 2014. se počeo povećavati prosječno 1,08% godišnje. Algoritam *Multilayer Perceptron* predviđa kako će se omjer ovisnosti starijeg stanovništva kontinuirano povećavati do 2050. godine kada bi trebao doseći vrijednost 40,94%. Takav postotak govori da će 2050. godine 41 starija osoba biti financijski ovisna o 100 radno sposobnih osoba.

Važno je napomenuti kako su se u klaster analizi i analizi vremenskih serija pronašli podaci o omjeru ovisnosti starijeg stanovništva u Hrvatskoj u 2019. godini, međutim vrijednosti podataka nisu jednake. Vrijednost omjera ovisnosti u klaster analizi dolazi iz baze podataka Razvojnog programa Ujedinjenih naroda dok su se u analizi vremenskih serija koristili podaci OECD-a. UN-ov omjer ovisnosti za 2019. godinu iznosi 32,3% dok OECD-ov iznosi 36%. Razlika leži u tome što UN računa omjer ovisnosti starijeg stanovništva kao broj osoba u dobi

65 i više godina na 100 osoba u dobi od 15 do 64 godine, a OECD na 100 osoba raspona godina od 20 do 64.

Ne može se zaključiti kako je jedan ili drugi izračun pravilan ili nepravilan. Bitno je pravilno iskoristiti podatke ovisno o vrsti istraživanja. Rezultati analize kao u ovom diplomskom radu pružaju okvir za razumijevanje demografskih promjena u društvu.

5. ZAKLJUČAK

Starenje stanovništva demografski je trend koji u svijetu traje desetljećima. Zbog svojih značajnih utjecaja, predmet je mnogih analiza i istraživanja. Do demografskog starenja dolazi zbog povećanja očekivanog trajanja života i smanjenja stope nataliteta. Ono predstavlja izazov za društvo jer stvara pritisak na mirovinske sustave, povećava troškove zdravstvene skrbi i zahtijeva prilagodbe javnih politika kako bi se osigurala dobrobit građana i održiva budućnost. Starim stanovništvom smatraju se osobe od 65 i više godina, a pomoću indikatora starenja stanovništva mjeri se proces demografskog starenja populacije. Jedan od ključnih indikatora je omjer ovisnosti starijeg stanovništva kojim se prikazuje broj osoba u dobi od 65 i više godina na 100 radno sposobnih osoba u dobi od 15 do 64 godine.

Organizacije poput Ujedinjenih naroda i Organizacije za ekonomsku suradnju i razvoj neprestano prikupljaju podatke o ukupnom stanovništvu, ljudskom razvoju, rodnom razlikama, ekološkoj održivosti te socioekonomskoj održivosti koja uključuje omjer ovisnosti starijeg stanovništva. Podaci pomažu državama u razvijanju politika, partnerstva, vještina i institucija kako bi ostvarile napredak. S obzirom na veličinu baza podataka potrebni su posebni alati i analize kako bi se došlo do željenih zaključaka. Proces otkrivanja znanja iz baza podataka pomaže pretvoriti sirove podatke u koristan skup koji može pridonijeti donošenju informiranih odluka i stvaranju vrijednosti. Rudarenje podataka nije ništa drugo nego otkrivanje korisnih informacija, obrazaca i zakonitosti iz velike količine podataka. U ovom diplomskom radu prikazan je proces rudarenja koristeći klaster analizu za grupiranje država prema omjeru ovisnosti starijeg stanovništva i analizu vremenskih serija za predviđanje budućih kretanja omjera u Hrvatskoj.

Cilj diplomskog rada bio je provedba analize financijske ovisnosti starijeg stanovništva o radno aktivnom stanovništvu. Klaster analizom države su grupirane u tri skupine prema razini ovisnosti starijeg stanovništva – države s visokim, srednjim i niskim omjerom ovisnosti. Cilj je bio pozicionirati Hrvatsku na svjetskoj razini, vidjeti koje se države nalaze u istom klasteru, a koje su u preostalima. Hrvatska se nalazi u klasteru 0, odnosno klasteru s visokim omjerom ovisnosti starijeg stanovništva zajedno s većinom europskih zemalja. Primjećuje se kako se države iz istog klastera često nalaze u istim geografskim regijama. Razlog tome je taj što različite regije svijeta imaju drugačiji stupanj razvoja te se neke još uvijek nalaze u demografskoj tranziciji dok je u drugima tranzicija završila. Europske države iz klastera 0 karakterizira niska stopa fertiliteta, produljen životni vijek i uređen sustav socijalne zaštite dok

nerazvijene afričke države iz klastera 1, koje su još uvijek u demografskoj tranziciji, imaju visoke stope nataliteta, niske stope mortaliteta i nerazvijen sustav socijalne zaštite. Hrvatska se, kao i ostale države iz klastera 0, suočava s izazovom održivosti mirovinskih sustava i pružanja adekvatne skrbi starijim osobama.

Analiza vremenskih serija provedena je za Hrvatsku s podacima od 2000. do 2023. godine s ciljem predviđanja budućeg kretanja omjera ovisnosti starijeg stanovništva. Predviđanjem su se dobili podaci o kretanju omjera do 2050. godine korištenjem strojnog učenja. Analiza pokazuje kontinuiran rast omjera ovisnosti starijeg stanovništva do 2050. godine kada je predviđeno da on iznosi 40,94%. Takav trend rasta u budućnosti za Hrvatsku nije održiv s obzirom na to da su mirovinski, zdravstveni i drugi sustavi već sada opterećeni.

Može se zaključiti kako donositelji odluka, ali i društvo imaju težak zadatak prilikom oblikovanja sustava koji se suočava s konstantnim povećanjem broja financijski ovisnih starijih osoba. Međutim, stvaranjem politika i programa moguće je ublažiti posljedice demografskog starenja te osigurati socijalnu sigurnost i održiv gospodarski rast.

LITERATURA

1. Bahovac, V., Crkvenac, M. (1995). Teorijski pristup analizi vremenskih serija i mogućnosti primjene u prognoziranju gospodarske kretnje u Hrvatskoj. *Privredna kretanja i ekonomska politika*, 5(37), 38-49. Preuzeto s <https://hrcak.srce.hr/33750> [15. rujna 2023].
2. Bejaković, P. (2011). Elaine Fultz: Dynamic Social Security for Europe: Choice and Responsibility. *Revija za socijalnu politiku*, 18(2), str. 232-236. <https://doi.org/10.3935/rsp.v18i2.1022>
3. Berry, M.J.A., Linoff, G.S. (2011.). *Mastering Data Mining*. Wiley, Chichester.
4. Beutel, J., List, S., Von Schweinitz, G. (2019.) Does machine learning help us predict banking crises?. *Journal of Financial Stability*, 45. Preuzeto s: <https://doi.org/10.1016/j.jfs.2019.100693> [15. rujna 2023].
5. Breznik, D. (1980.). *Demografija: analiza, metodi i modeli*. 2. izd. Institut društvenih nauka, Centar za demografska istraživanja
6. Everitt, B.S., Landau, S., Leese, M. (2001.). *Cluster analysis*. London: Arnold; Cham: Palgrave Macmillan.
7. Gupta, G.K. (2015.). *Introduction to data mining with case studies*. 3 izd. Delhi: PHI Learning Private Limited.
8. Han, J., Kamber, M., Pei, J. (2012.) *Data mining: Concepts and techniques*, 3 izd. Waltham: Morgan Kaufmann Publishers.
9. Kantardžić, M. (2011.). *Data mining: concepts, models, methods, and algorithms*. 2 izd. Hoboken: Wiley-Interscience.
10. Magazzino, C., Mele, M., Morelli, G. (2021.). The Relationship between Renewable Energy and Economic Growth in a Time of Covid-19: A Machine Learning Experiment on the Brazilian Economy. *Sustainability*, 13(3), 1285. MDPI AG. Preuzeto s: <http://dx.doi.org/10.3390/su13031285> [15. rujna 2023].
11. Milinković, D. (1994.). Demografska struktura i mirovinski sustav. *Revija za socijalnu politiku*, 1(2), str. 141-149. <https://doi.org/10.3935/rsp.v1i2.598>
12. Nejašmić, I., Mišetić, R. (2010.). Sintetični pokazatelj demografskih resursa: doprinos tipologiji hrvatskog prostora. *Hrvatski geografski glasnik*, 72. (1.), 49-60. <https://doi.org/10.21861/hgg.2010.72.03>

13. Nejašmić, I., Toskić, A. (2013.). Starenje stanovništva u Hrvatskoj–sadašnje stanje i perspektive. *Hrvatski geografski glasnik*, 75.(1.), str. 89-110.
<https://doi.org/10.21861/HGG.2013.75.01.05>
14. OECD (2022). Demography – old-age dependency ratio – OECD data. Preuzeto s: <https://data.oecd.org/pop/old-age-dependency-ratio.htm> [25. listopada 2022].
15. Oliveira-Roca, M. (1991.). Demografski resursi regija Hrvatske: prijedlog konceptualno-metodološkog okvira istraživanja. U: Seferagić, D.(ur.): Društvene promjene u prostoru. Zagreb, Institut za društvena istraživanja, 43-69.
16. Parr Rud, O. (2003.) *Data Mining Cookbook*. Wiley, Chichester.
17. Pejić Bach, M. & Kerep, I. (2011). *Weka – alat za otkrivanje znanja iz baza podataka*. Zagreb: Mikrorad.
18. Pejić Bach, M. (2005.) Rudrenje podataka u bankarstvu. *Zbornik Ekonomskog fakulteta u Zagrebu*, 3 (1), 181-193. Preuzeto s: <https://hrcak.srce.hr/26220>
19. Poston, D. L. Jr., Bouvier, L. F. (2010.). *Population and society: an introduction to demography*. New York: Cambridge University Press.
20. Puljiz, V. (2016.) Starenje stanovništva – izazov socijalne politike. *Revija za socijalnu politiku*, 23(1), str. 81-98. <https://doi.org/10.3935/rsp.v23i1.1281>
21. Puljiz, V. i sur. (2005.). *Socijalna politika: povijest, sustavi, pojmovnik*. Pravni fakultet Sveučilišta u Zagrebu.
22. Romesburg, H.C. (2004.). *Cluster analysis for researchers*. Morrisville, North Carolina: Lulu Press.
23. Rose, S. (2013.). Mortality risk score prediction in an elderly population using machine learning. *American journal of epidemiology*, 177(5), 443-452
24. Samodol, A. (2021.). Demografske promjene i mirovinski sustavi u Europskoj uniji: primjer Hrvatske. *Međunarodne studije*, XXI(1), str. 93-125.
<https://doi.org/10.46672/ms.21.1.5>
25. Sanderson, W.C., Scherbov, S. (2007.) A new perspective on population aging, *Demographic Research*. Preuzeto s: <https://www.demographic-research.org/volumes/vol16/2/default.htm> [28. svibnja 2023].
26. Seifert, J.W. (2006.) *Data Mining: An Overview*. U: *National Security Issues*, ur. Daniel D. Pergakov. Nova Science Publishers, Inc., str. 201-217.
27. Shillan, D., Sterne, J.A.C., Champneys, A., Gubbson, B. (2019.) Use of machine learning to analyse routinely collected intensive care unit data: a systematic

- review. *Critical Care*, 23(284) Preuzeto s: <https://doi.org/10.1186/s13054-019-2564-9> [15. rujna 2023].
28. Shmueli, G., Lichtendahl Jr, K.C. (2016.). *Practical Time Series Forecasting with R: A Hands-On Guide*, 2. izd. Axelrod Schnall Publishers
29. United Nations (2013.). *National Transfer Accounts Manual: Measuring and Analysing the Generational Economy*. Sales No.: E.13.XIII.6. Preuzeto s: <https://doi.org/10.18356/9df3caed-en>. [28. svibnja 2023].
30. United Nations Development Programme (2022.). *Data center, Human Development Reports*. Preuzeto s: <https://hdr.undp.org/data-center> [29. listopada 2022].
31. United Nations, Department of Economic and Social Affairs, Population Division (2019.). *World population ageing 2019 highlights*. Preuzeto s: <https://doi.org/10.18356/9df3caed-en>. [25. listopada 2022].
32. Wertheimer-Baletić, A. (1999.). *Stanovništvo i razvoj*. Zagreb: Mate
33. Živić, D. (2003.). *Demografske odrednice i posljedice starenja stanovništva Hrvatske*. *Revija za socijalnu politiku*, 10 (3), 307-319. <https://doi.org/10.3935/rsp.v10i3.107>

POPIS SLIKA

Slika 1 Proces otkrivanja znanja iz baza podataka.....	14
Slika 2 Primjer procesa otkrivanja znanja iz baza podataka.....	17
Slika 3 Weka – panel Preprocess za klaster analizu.....	25
Slika 4 Postavke algoritma SimpleKMeans.....	26
Slika 5 Vizualni prikaz pripadnosti klasteru po omjeru ovisnosti starog stanovništva	28
Slika 6 Informacije o točki na grafu	29
Slika 7 Isječak procesa učenja u analizi vremenskih serija	36
Slika 8 Grafički prikaz stvarnog i predviđenog kretanja omjera ovisnosti starog stanovništva	37

POPIS TABLICA

Tablica 1 Atributi, formati i modaliteti za klaster analizu	22
Tablica 2 Pripadnost država klasteru.....	30
Tablica 3 Atributi, formati i modaliteti za analizu vremenskih serija.....	33
Tablica 4 Stvarne i predviđene vrijednosti i pogreške	36
Tablica 5 Omjer ovisnosti starijeg stanovništva od 2000. do 2023.	38
Tablica 6 Predviđene vrijednosti omjera ovisnosti starijeg stanovništva.....	38

POPIS GRAFIKONA

Grafikon 1 Veličina klastera	32
Grafikon 2 Kretanje omjera ovisnosti od 2000. do 2050. godine	39

PRILOZI

Country,HDI,Life_expectancy_at_birth>Total_population
,Average_annual_population_growth_2005/2010
,Average_annual_population_growth_2015/2020,Population_under_age_5
,Population_ages_15-64 ,Population_ages_65_and_older,Median_population_age,Young
age_(0-14)_dependency_ratio ,Old_age_(65_and_older)_dependency_ratio

Norway,0.957,82.4,5.38,1.1,0.8,0.3,3.5,0.9,39.8,26.7,26.4

Ireland,0.955,82.3,4.88,1.9,1.2,0.3,3.2,0.7,38.2,32.7,22.0

Switzerland,0.955,83.8,8.59,1.1,0.8,0.4,5.7,1.6,43.1,22.6,28.4

Hong_Kong,0.949,84.9,7.44,0.6,0.8,0.3,5.2,1.3,44.8,17.6,24.9

Iceland,0.949,83.0,0.34,1.6,0.7,0.0,0.2,0.1,37.5,30.2,23.3

Germany,0.947,81.3,83.52,-0.2,0.5,4.0,54.0,18.0,45.7,21.3,33.4

Sweden,0.945,82.8,10.04,0.8,0.7,0.6,6.2,2.0,41.1,28.3,32.5

Australia,0.944,83.4,25.20,1.9,1.3,1.7,16.3,4.0,37.9,29.7,24.6

Netherlands,0.944,82.3,17.10,0.4,0.2,0.9,11.0,3.4,43.3,24.6,30.4

Denmark,0.940,80.9,5.77,0.5,0.4,0.3,3.7,1.2,42.3,25.8,31.4

Finland,0.938,81.9,5.53,0.4,0.2,0.3,3.4,1.2,43.1,25.9,35.8

Singapore,0.938,83.6,5.80,3.7,0.9,0.2,4.4,0.7,42.2,16.4,16.5

UK,0.932,81.3,67.53,1.0,0.6,4.0,43.1,12.5,40.5,27.7,29.0

Belgium,0.931,81.6,11.54,0.7,0.5,0.6,7.4,2.2,41.9,26.7,29.7

New_Zealand,0.931,82.3,4.78,1.1,0.9,0.3,3.1,0.8,38.0,30.4,24.8

Canada,0.929,82.4,37.41,1.2,0.9,2.0,24.9,6.6,41.1,23.8,26.5

United_States,0.926,78.9,329.06,0.9,0.6,19.6,214.7,53.3,38.3,28.4,24.8

Austria,0.922,81.5,8.96,0.4,0.7,0.4,6.0,1.7,43.5,21.6,28.7

Israel,0.919,83.0,8.52,2.4,1.6,0.8,5.1,1.0,30.5,46.5,20.4

Japan,0.919,84.6,126.86,0.0,-0.2,4.9,75.4,35.5,48.4,21.2,47.1

Slovenia,0.917,81.3,2.08,0.5,0.1,0.1,1.3,0.4,44.5,23.3,31.2

Korea,0.916,83.0,51.23,0.3,0.2,2.0,37.0,7.7,43.7,17.7,20.9

Luxembourg,0.916,82.3,0.62,2.1,2.0,0.0,0.4,0.1,39.7,22.4,20.4

Spain,0.904,83.6,46.74,1.3,0.0,2.0,30.7,9.2,44.9,22.2,29.9

France,0.901,82.7,65.13,0.6,0.3,3.7,40.3,13.3,42.3,28.8,33.0

Czechia,0.900,79.4,10.69,0.5,0.2,0.6,6.9,2.1,43.2,24.4,30.7

Malta,0.895,82.5,0.44,0.5,0.4,0.0,0.3,0.1,42.6,22.1,32.1
 Estonia,0.892,78.8,1.33,-0.4,0.2,0.1,0.8,0.3,42.4,25.9,31.5
 Italy,0.892,83.5,60.55,0.4,0.0,2.4,38.6,13.9,47.3,20.6,36.1
 UAE,0.890,78.0,9.77,12.4,1.3,0.5,8.2,0.1,32.6,17.5,1.4
 Greece,0.888,82.2,10.47,-0.6,-0.4,0.4,6.7,2.3,45.6,21.7,34.2
 Cyprus,0.887,81.0,1.20,1.6,0.8,0.1,0.8,0.2,37.3,24.1,20.3
 Lithuania,0.882,75.9,2.76,-1.4,-1.5,0.1,1.8,0.6,45.1,23.4,31.1
 Poland,0.880,78.7,37.89,0.0,-0.1,1.9,25.3,6.9,41.7,22.8,27.2
 Latvia,0.866,75.3,1.91,-1.2,-1.1,0.1,1.2,0.4,43.9,25.7,32.1
 Portugal,0.864,82.1,10.23,0.2,-0.3,0.4,6.6,2.3,46.2,20.6,34.7
 Slovakia,0.860,77.5,5.46,0.0,0.1,0.3,3.7,0.9,41.2,22.7,23.7
 Hungary,0.854,76.9,9.68,-0.3,-0.2,0.5,6.4,1.9,43.3,21.9,29.9
 Saudi_Arabia,0.854,75.1,34.27,2.8,1.9,3.0,24.6,1.2,31.8,34.7,4.8
 Bahrain,0.852,77.3,1.64,6.7,4.3,0.1,1.3,0.0,32.5,23.7,3.2
 Chile,0.851,80.2,18.95,1.1,1.2,1.2,13.0,2.3,35.3,28.4,17.3
 Croatia,0.851,78.5,4.13,-0.2,-0.6,0.2,2.7,0.9,44.3,22.6,32.3
 Qatar,0.848,80.2,2.83,15.3,2.3,0.1,2.4,0.0,32.3,16.0,1.8
 Argentina,0.845,76.7,44.78,1.0,1.0,3.7,28.7,5.0,31.5,38.3,17.5
 Brunei_Darussalam,0.838,75.9,0.43,1.2,1.1,0.0,0.3,0.0,32.3,31.3,7.2
 Montenegro,0.829,76.9,0.63,0.3,0.0,0.0,0.4,0.1,38.8,27.3,23.2
 Romania,0.828,76.1,19.36,-0.9,-0.7,0.9,12.7,3.6,43.2,23.7,28.6
 Kazakhstan,0.825,73.6,18.55,1.1,1.3,1.9,11.8,1.4,30.7,45.5,12.1
 Russian_Federation,0.824,72.6,145.87,0.0,0.1,9.4,97.4,22.0,39.6,27.2,22.6
 Belarus,0.823,74.8,9.45,-0.3,0.0,0.6,6.4,1.4,40.3,25.1,22.4
 Turkey,0.820,77.7,83.43,1.3,1.4,6.7,55.9,7.3,31.5,36.3,13.0
 Uruguay,0.817,77.9,3.46,0.2,0.4,0.2,2.2,0.5,35.8,31.7,23.1
 Bulgaria,0.816,75.1,7.00,-0.7,-0.7,0.3,4.5,1.5,44.6,22.9,33.2
 Panama,0.815,78.5,4.25,1.8,1.7,0.4,2.8,0.4,29.7,41.3,12.8
 Bahamas,0.814,73.9,0.39,1.8,1.0,0.0,0.3,0.0,32.3,31.3,10.6
 Barbados,0.814,79.2,0.29,0.4,0.1,0.0,0.2,0.0,40.5,25.6,24.3

Oman,0.813,77.9,4.97,3.8,3.6,0.5,3.7,0.1,30.6,29.8,3.3
 Georgia,0.812,73.8,4.00,-0.5,-0.2,0.3,2.6,0.6,38.3,30.9,23.2
 Costa_Rica,0.810,80.3,5.05,1.3,1.0,0.4,3.5,0.5,33.5,30.5,14.3
 Malaysia,0.810,76.2,31.95,1.9,1.3,2.6,22.2,2.2,30.3,34.2,10.0
 Kuwait,0.806,75.5,4.21,5.5,2.1,0.3,3.2,0.1,36.8,28.5,3.7
 Serbia,0.806,76.0,8.77,-0.4,-0.3,0.4,5.8,1.6,41.6,23.6,28.5
 Mauritius,0.804,75.0,1.27,0.4,0.2,0.1,0.9,0.2,37.5,24.4,17.0
 Seychelles,0.796,73.4,0.10,0.6,0.7,0.0,0.1,0.0,34.2,34.7,11.4
 Trinidad_and_Tobago,0.796,73.5,1.39,0.5,0.4,0.1,1.0,0.2,36.2,29.5,16.2
 Albania,0.795,78.6,2.88,-0.9,-0.1,0.2,2.0,0.4,36.4,25.4,20.8
 Cuba,0.783,78.8,11.33,-0.1,0.0,0.6,7.8,1.8,42.2,23.5,22.8
 Iran,0.783,76.7,82.91,1.1,1.4,7.6,57.2,5.3,32.0,35.7,9.2
 Sri_Lanka,0.782,77.0,21.32,0.7,0.5,1.7,13.9,2.3,34.0,36.7,16.6
 Bosnia_and_Herzegovina,0.780,77.4,3.30,-0.3,-0.9,0.1,2.2,0.6,43.1,21.6,25.3
 Grenada,0.779,72.4,0.11,0.3,0.5,0.0,0.1,0.0,32.0,35.6,14.5
 Mexico,0.779,75.1,127.58,1.5,1.1,11.0,84.7,9.5,29.2,39.4,11.2
 Ukraine,0.779,72.1,43.99,-0.5,-0.5,2.2,29.6,7.3,41.2,23.7,24.8
 Antigua_and_Barbuda,0.778,77.0,0.10,1.6,0.9,0.0,0.1,0.0,34.0,31.8,13.1
 Peru,0.777,76.7,32.51,0.8,1.6,2.8,21.6,2.7,31.0,38.1,12.7
 Thailand,0.777,77.2,69.63,0.5,0.3,3.6,49.3,8.6,40.1,23.8,17.5
 Armenia,0.776,75.1,2.96,-0.7,0.3,0.2,2.0,0.3,35.4,30.7,17.0
 North_Macedonia,0.774,75.8,2.08,0.1,0.0,0.1,1.4,0.3,39.1,23.6,20.3
 Colombia,0.767,77.3,50.34,1.2,1.4,3.7,34.5,4.4,31.3,33.0,12.8
 Brazil,0.765,75.9,211.05,1.0,0.8,14.6,147.2,19.5,33.5,30.1,13.3
 China,0.761,76.9,1433.78,0.6,0.5,85.0,1014.0,164.5,38.4,25.2,16.2
 Ecuador,0.759,77.0,17.37,1.6,1.7,1.7,11.3,1.3,27.9,42.7,11.4
 Saint_Lucia,0.759,76.2,0.18,1.3,0.5,0.0,0.1,0.0,34.5,25.4,14.0
 Azerbaijan,0.756,73.0,10.05,1.1,1.0,0.8,7.0,0.6,32.3,33.4,9.2
 Dominican_Republic,0.756,74.1,10.74,1.3,1.1,1.0,7.0,0.8,28.0,42.6,11.2
 Moldova,0.750,71.9,4.04,-0.4,-0.2,0.2,2.9,0.5,37.6,22.0,16.7

Algeria,0.748,76.9,43.05,1.6,2.0,5.0,27.1,2.8,28.5,48.6,10.4
 Lebanon,0.744,78.9,6.86,1.1,0.9,0.6,4.6,0.5,29.6,38.1,10.8
 Fiji,0.743,67.4,0.89,0.9,0.6,0.1,0.6,0.1,27.9,45.0,8.6
 Maldives,0.740,78.9,0.53,2.7,3.4,0.0,0.4,0.0,29.9,26.0,4.7
 Tunisia,0.740,76.7,11.69,1.0,1.1,1.0,7.9,1.0,32.8,36.1,12.8
 Saint_Vincent_and_the_Grenadines,0.738,72.5,0.11,-0.1,0.3,0.0,0.1,0.0,32.9,32.6,14.3
 Suriname,0.738,71.7,0.58,1.2,1.0,0.1,0.4,0.0,29.0,40.7,10.6
 Mongolia,0.737,69.9,3.23,1.5,1.8,0.4,2.1,0.1,28.2,47.4,6.4
 Botswana,0.735,69.6,2.30,2.0,2.1,0.3,1.4,0.1,24.0,54.6,7.1
 Jamaica,0.734,74.5,2.95,0.5,0.5,0.2,2.0,0.3,30.7,34.8,13.2
 Jordan,0.729,74.5,10.10,4.6,1.9,1.1,6.3,0.4,23.8,53.7,6.2
 Paraguay,0.728,74.3,7.04,1.4,1.3,0.7,4.5,0.5,26.3,45.5,10.3
 Tonga,0.725,70.9,0.10,0.6,1.0,0.0,0.1,0.0,22.4,59.5,10.1
 Libya,0.724,72.9,6.78,1.3,1.4,0.6,4.6,0.3,28.8,41.6,6.6
 Uzbekistan,0.720,71.7,32.98,1.5,1.6,3.4,22.0,1.5,27.8,43.2,6.9
 Bolivia,0.718,71.5,11.51,1.7,1.4,1.2,7.1,0.8,25.6,49.4,11.8
 Indonesia,0.718,71.7,270.63,1.3,1.1,23.9,183.3,16.4,29.7,38.7,8.9
 Philippines,0.718,71.2,108.12,1.7,1.4,10.8,69.4,5.7,25.7,47.5,8.3
 Belize,0.716,74.6,0.39,2.6,1.9,0.0,0.3,0.0,25.5,45.4,7.4
 Samoa,0.715,73.3,0.20,0.7,0.5,0.0,0.1,0.0,21.8,66.2,8.6
 Turkmenistan,0.715,68.2,5.94,1.4,1.6,0.7,3.8,0.3,26.9,47.6,7.1
 Venezuela,0.711,72.1,28.52,1.5,-1.1,2.4,18.5,2.2,29.6,42.2,11.7
 South_Africa,0.709,64.1,58.56,1.3,1.4,5.8,38.4,3.2,27.6,44.1,8.3
 Palestine,0.708,74.1,4.98,2.5,2.4,0.7,2.9,0.2,20.8,66.4,5.5
 Egypt,0.707,72.0,100.39,1.8,2.0,12.8,61.1,5.3,24.6,55.6,8.7
 Vietnam,0.704,75.4,96.46,1.0,1.0,7.9,66.8,7.3,32.5,33.5,10.9
 Gabon,0.703,66.5,2.17,3.1,2.7,0.3,1.3,0.1,22.5,62.7,6.0
 Kyrgyzstan,0.697,71.5,6.42,1.3,1.8,0.8,4.0,0.3,26.0,51.7,7.3
 Morocco,0.686,76.7,36.47,1.2,1.3,3.4,24.0,2.7,29.5,41.0,11.1
 Guyana,0.682,69.9,0.78,0.1,0.5,0.1,0.5,0.1,26.7,42.7,10.3

Iraq,0.674,70.6,39.31,2.0,2.5,5.4,23.0,1.3,21.0,64.9,5.8
 El_Salvador,0.673,73.3,6.45,0.4,0.5,0.6,4.2,0.5,27.6,41.5,13.1
 Tajikistan,0.668,71.1,9.32,2.1,2.4,1.3,5.6,0.3,22.4,61.9,5.2
 Cabo_Verde,0.665,73.0,0.55,1.2,1.2,0.1,0.4,0.0,27.6,42.5,7.0
 Guatemala,0.663,74.3,17.58,2.2,1.9,2.1,10.8,0.9,22.9,55.3,8.1
 Nicaragua,0.660,74.5,6.55,1.4,1.3,0.7,4.2,0.4,26.5,46.2,8.4
 Bhutan,0.654,71.8,0.76,1.1,1.2,0.1,0.5,0.0,28.1,37.0,8.9
 Namibia,0.646,63.7,2.49,1.8,1.9,0.3,1.5,0.1,21.8,62.0,6.1
 India,0.645,69.7,1366.42,1.5,1.0,116.8,915.6,87.1,28.4,39.7,9.5
 Honduras,0.634,75.3,9.75,2.2,1.7,1.0,6.2,0.5,24.3,48.7,7.5
 Bangladesh,0.632,72.6,163.05,1.2,1.1,14.4,110.2,8.4,27.6,40.3,7.7
 Kiribati,0.630,68.4,0.12,2.2,1.5,0.0,0.1,0.0,23.0,59.7,6.8
 Sao_Tome_and_Principe,0.625,70.4,0.22,2.7,1.9,0.0,0.1,0.0,18.6,76.8,5.4
 Micronesia,0.620,67.9,0.11,-0.6,1.1,0.0,0.1,0.0,24.4,48.9,6.5
 Lao_Peoples_Democratic_Republic,0.613,67.9,7.17,1.7,1.5,0.8,4.6,0.3,24.4,50.8,6.6
 Eswatini,0.611,60.2,1.15,0.7,1.0,0.1,0.7,0.0,20.7,65.1,6.9
 Ghana,0.611,64.1,30.42,2.5,2.2,4.1,18.1,0.9,21.5,62.8,5.2
 Vanuatu,0.609,70.5,0.30,2.4,2.5,0.0,0.2,0.0,21.1,67.1,6.3
 Timor_Leste,0.606,69.5,1.29,1.9,1.9,0.2,0.8,0.1,20.8,63.8,7.3
 Nepal,0.602,70.8,28.61,1.0,1.5,2.7,18.5,1.7,24.6,45.7,8.9
 Kenya,0.601,66.7,52.57,2.8,2.3,7.0,30.7,1.3,20.1,67.1,4.2
 Cambodia,0.594,69.8,16.49,1.5,1.5,1.8,10.6,0.8,25.6,48.5,7.4
 Equatorial_Guinea,0.592,58.7,1.36,4.6,3.7,0.2,0.8,0.0,22.3,61.0,4.0
 Zambia,0.584,63.9,17.86,2.8,2.9,2.9,9.5,0.4,17.6,83.2,4.0
 Myanmar,0.583,67.1,54.05,0.7,0.6,4.5,36.8,3.2,29.0,38.1,8.8
 Angola,0.581,61.2,31.83,3.7,3.3,5.7,16.3,0.7,16.7,91.1,4.3
 Congo,0.574,64.6,5.38,3.3,2.6,0.8,3.0,0.1,19.2,74.5,4.9
 Zimbabwe,0.571,61.5,14.65,1.0,1.5,2.1,8.0,0.4,18.7,76.8,5.4
 Solomon_Islands,0.567,73.0,0.67,2.3,2.6,0.1,0.4,0.0,19.9,71.4,6.5
 Syrian_Arab_Republic,0.567,72.7,17.07,3.0,-0.6,1.8,11.0,0.8,25.6,48.3,7.3

Cameroon,0.563,59.3,25.88,2.7,2.6,4.1,14.2,0.7,18.7,77.1,5.0
 Pakistan,0.557,67.3,216.57,2.3,2.0,27.7,131.3,9.4,22.8,57.8,7.1
 Papua_New_Guinea,0.555,64.5,8.78,2.4,2.0,1.1,5.4,0.3,22.4,58.2,5.8
 Comoros,0.554,64.3,0.85,2.4,2.2,0.1,0.5,0.0,20.4,68.1,5.3
 Mauritania,0.546,64.9,4.53,2.9,2.8,0.7,2.6,0.1,20.1,70.0,5.5
 Benin,0.545,61.8,11.80,2.8,2.7,1.9,6.4,0.4,18.8,77.4,6.0
 Uganda,0.544,63.4,44.27,3.2,3.6,7.7,22.8,0.9,16.7,90.2,3.8
 Rwanda,0.543,69.0,12.63,2.5,2.6,1.8,7.2,0.4,20.0,69.5,5.3
 Nigeria,0.539,54.7,200.96,2.6,2.6,33.4,107.7,5.5,18.1,81.6,5.1
 Cote_dIvoire,0.538,57.8,25.72,2.2,2.5,4.1,14.3,0.7,18.9,75.3,5.2
 Tanzania,0.529,65.5,58.01,2.9,3.0,9.5,31.1,1.5,18.0,81.9,4.9
 Madagascar,0.528,67.0,26.97,2.9,2.7,4.0,15.3,0.8,19.6,71.3,5.4
 Lesotho,0.527,54.3,2.13,0.0,0.8,0.3,1.3,0.1,24.0,51.9,7.9
 Djibouti,0.524,67.1,0.97,1.4,1.6,0.1,0.6,0.0,26.6,44.1,7.0
 Togo,0.515,61.0,8.08,2.7,2.5,1.2,4.5,0.2,19.4,73.1,5.1
 Senegal,0.512,67.9,16.30,2.7,2.8,2.6,8.8,0.5,18.5,79.2,5.7
 Afghanistan,0.511,64.8,38.04,2.6,2.5,5.6,20.9,1.0,18.4,77.3,4.8
 Haiti,0.510,64.0,11.26,1.6,1.3,1.3,7.0,0.6,24.0,52.9,8.2
 Sudan,0.510,65.3,42.81,2.2,2.4,6.2,24.1,1.6,19.7,71.4,6.5
 Gambia,0.496,62.1,2.35,3.0,2.9,0.4,1.3,0.1,17.8,82.8,4.8
 Ethiopia,0.485,66.6,112.08,2.8,2.6,16.6,62.9,3.9,19.5,71.8,6.3
 Malawi,0.483,64.3,18.63,2.8,2.7,2.9,10.0,0.5,18.1,80.7,4.9
 Democratic_Republic_of_Congo,0.480,60.7,86.79,3.3,3.2,15.5,44.2,2.6,17.0,90.2,5.9
 Guinea_Bissau,0.480,58.3,1.92,2.5,2.5,0.3,1.1,0.1,18.8,76.7,5.2
 Liberia,0.480,64.1,4.94,3.8,2.5,0.7,2.8,0.2,19.4,72.8,5.9
 Guinea,0.477,61.6,12.77,2.2,2.8,2.1,6.8,0.4,18.0,81.1,5.5
 Yemen,0.470,66.1,29.16,2.8,2.4,4.1,16.9,0.8,20.2,67.8,5.0
 Eritrea,0.459,66.3,3.50,2.3,1.2,0.5,1.9,0.2,19.2,77.1,8.4
 Mozambique,0.456,60.9,30.37,2.8,2.9,5.0,16.0,0.9,17.6,84.1,5.5
 Burkina_Faso,0.452,61.6,20.32,3.0,2.9,3.4,10.8,0.5,17.6,84.4,4.5

Sierra_Leone,0.452,54.7,7.81,2.6,2.1,1.1,4.4,0.2,19.4,72.3,5.2

Mali,0.434,59.3,19.66,3.3,3.0,3.5,9.9,0.5,16.3,94.2,5.0

Burundi,0.433,61.6,11.53,3.3,3.1,2.0,6.0,0.3,17.3,86.8,4.4

South_Sudan,0.433,57.9,11.06,4.7,0.9,1.7,6.1,0.4,19.0,75.5,6.1

Chad,0.398,54.2,15.95,3.4,3.0,2.9,8.1,0.4,16.6,92.4,4.9

Central_African_Republic,0.397,53.3,4.75,1.7,1.4,0.7,2.5,0.1,17.6,82.5,5.3

Niger,0.394,62.4,23.31,3.8,3.8,4.6,11.1,0.6,15.2,104.8,5.