

UNIVERZA V LJUBLJANI
FAKULTETA ZA RAČUNALNIŠTVO IN INFORMATIKO

Vid Rotar

**Vizualizacija podatkov o porabi
zdravil na recept**

DIPLOMSKO DELO

UNIVERZITETNI ŠTUDIJSKI PROGRAM
PRVE STOPNJE
RAČUNALNIŠTVO IN INFORMATIKA

MENTOR: viš. pred. dr. Alenka Kavčič

SOMENTOR: as. Matevž Pesek

Ljubljana, 2018

COPYRIGHT. Rezultati diplomske naloge so intelektualna lastnina avtorja in Fakultete za računalništvo in informatiko Univerze v Ljubljani. Za objavo in koriščenje rezultatov diplomske naloge je potrebno pisno privoljenje avtorja, Fakultete za računalništvo in informatiko ter mentorja.

Besedilo je oblikovano z urejevalnikom besedil L^AT_EX.

Fakulteta za računalništvo in informatiko izdaja naslednjo nalogo:

Vizualizacija podatkov o porabi zdravil na recept

Tematika naloge:

Nacionalni inštitut za javno zdravje že več let redno spremlja porabo zdravil v Sloveniji. Njihova obsežna zbirka podatkov o porabi zdravil, ki so predpisana na recept, vsebuje podatke o porabi po letih za posamezne slovenske regije za različne skupine zdravil. Vendar so velike količine številčnih podatkov zelo nepregledne, težko razumljive in ne omogočajo dobre interpretacije ali širšega pregleda nad njimi. Za lažje razumevanje in interpretacijo podatkov je zato nujno potreben primeren prikaz podatkov, najbolje v grafični obliki. V okviru diplomske naloge zasnujte in izdelajte spletno aplikacijo za vizualizacijo podatkov o porabi zdravil v Sloveniji v preteklih letih. Aplikacija naj uporabniku omogoča prikaz podatkov za izbrano regijo v izbranem letu ter medsebojno primerjavo podatkov dveh regij. Za izbrana zdravila oziroma bolezni pripravite tudi prikaz porabe po posameznih regijah Slovenije. Vse vizualne predstavitve naj bodo izbrane tako, da bodo najbolj ustrezale vrsti prikazanih podatkov ter da bodo uporabne in zanimive za čim širši krog uporabnikov. Pri tem upoštevajte smernice teorije vizualnega kodiranja ter oblikovne principe vizualizacij. Pri zasnovi in realizaciji aplikacije uporabite sodobne spletne tehnologije in orodja.

Rad bi se zahvalil svoji mentorici dr. Alenki Kavčič in somentorju as. Matježu Pesku za pomoč in svetovanje pri izdelavi diplomske naloge. Zahvaljujem se tudi Alešu Korošču iz Nacionalnega inštituta za javno zdravje za pridobljene podatke in nasvete ter Eleni Plahuti za pomoč pri oblikovanju spletne aplikacije.

Kazalo

Povzetek

Abstract

1	Uvod	1
1.1	Namen interaktivne vizualizacije	1
1.2	Klasifikacijski sistem ATC	3
1.3	Struktura klasifikacijskega sistema ATC	3
1.4	Željena funkcionalnost aplikacije	5
2	Pregled sorodnih rešitev	9
2.1	Vizualizacija z uporabo krogov	9
2.2	Barvni zemljevid	10
3	Uporabljene tehnologije	13
3.1	HTML in CSS	13
3.2	Javascript	14
3.3	D3.js	14
3.4	Bootstrap	14
3.5	ElasticSearch	15
4	Vizualizacija porabe zdravil na recept	17
4.1	Vnos v bazo	17
4.2	Struktura aplikacije	18

5 Sklepne ugotovitve	27
Literatura	30

Seznam uporabljenih kratic

kratica	angleško	slovensko
ATC	Anatomical Therapeutic Chemical	Anatomsko-terapevtsko-kemična
SZO	World Health Organization	Svetovna zdravstvena organizacija
DDD	Defined daily dose	Definirani dnevni odmerek
DID	Defined daily dose per 1000 inhabitants per day	Definirani dnevni odmerek na 1000 prebivalcev na dan
NIJZ	National Institute of Public Health	Nacionalni inštitut za javno zdravje

Povzetek

Naslov: Vizualizacija podatkov o porabi zdravil na recept

Avtor: Vid Rotar

Cilj diplomske naloge je bil izdelati interaktivno vizualizacijo podatkov o izdanih zdravilih na recept. Podatke je priskrbel Nacionalni inštitut za javno zdravje. Podatkov o zdravilih na recept je ogromno, poleg tega pa so razporejeni v več kategorij in ravni. Pred implementacijo smo raziskali že narejene interaktivne vizualizacije. Ugotovili smo, da na našo temo ni implementiranih veliko vizualizacij. Za predstavitev podatkov smo se odločili implementirati tri različne vizualizacije. Vsaka predstavlja podatke iz drugega zornega kota. Z interaktivnostjo vizualizacij je uporabniku omogočeno lažje raziskovanje podatkov o zdravilih, izdanih na recept. Vse tri vizualizacije smo vključili v spletno aplikacijo, ki je dostopna širši javnosti.

Ključne besede: interaktivna vizualizacija, zdravila na recept, HTML, CSS, JavaScript, ElasticSearch, klasifikacija ATC.

Abstract

Title: Visualization of data on prescription drugs consumption

Author: Vid Rotar

The aim of this thesis was to create an interactive data visualisation of prescription medicine. We used data that was provided by the National Institute of Public Health. The structure of provided data was divided in several groups and levels. One of the first steps of creating our application was to research already made similar visualisations. There were almost no similar visualisations based on prescription medicine. We implemented 3 different visualisations. All visualisations were interactive, which enabled easier data research. Each visualisation shows data from a different angle. All visualisations were implemented into a web application.

Keywords: interactive visualisation, prescription medicine, HTML, CSS, JavaScript, ElasticSearch, ATC classification.

Poglavje 1

Uvod

Podatki so v današnjem svetu izjemno pomembni in koristni. Z njihovo obdelavo lahko dobimo ogromno informacij. Problem nastane, ko je podatkov, ki jih proučujemo ogromno. Podatkov o porabi zdravil na recept je veliko, razporejeni so v več skupin in ravni. Čeprav so podatki o zdravilih dostopni vsakomur (na voljo so na podatkovnem portalu Nacionalnega inštituta za javno zdravje [19]), jih zaradi njihove strukture in števila težko interpretiramo. Zato smo v sodelovanju z Nacionalnim inštitutom za javno zdravje ustvarili interaktivno vizualizacijo podatkov o izdanih zdravilih na recept. Vizualizacijo smo implementirali v obliki spletne aplikacije. Ta vizualizacija uporabniku pomaga razumeti in si boljše predstavljati velike količine podatkov o zdravilih, izdanih na recept.

1.1 Namen interaktivne vizualizacije

Predstavitev in razumevanje ogromnega števila podatkov je za človeka težka naloga. Vizualizacija predstavlja učinkovit način predstave informacij in zgodbe, ki jo podatki pripovedujejo. Preučevanje vizualizacij uporabniku pomaga prepoznati različne trende ali odstopanja. Poleg tega ga vzpodbujajo k raziskovanju po vizualiziranih podatkih [20, 17].

Vizualizacija podatkov je definirana kot predstavitev podatkov z določenimi

lastnostmi.

- Je narisana algoritmično (večinoma je izrisana s pomočjo računalniških metod).
- Lahko se jo uporabi za izris različnih podatkov.
- Estetsko je manj bogata.
- Je podatkovno bogata.

Vizualizacije podatkov splošno delimo na dve kategoriji: vizualizacije, ki nam podatke predstavljajo ter tiste, po katerih lahko raziskujemo. Vizualizacije za raziskovanje uporabimo, ko je na voljo veliko število podatkov, a še ne vemo, kaj vsebujejo. Večinoma je vizualizacija za raziskovanje podatkov del faze analize, s katero najdemo zgodbo, ki nam jo podatki prikazujejo. Vizualizacije za predstavljanje podatkov pridejo v poštev, ko že imamo idejo o tem, kakšno zgodbo nam podatki skušajo prikazati [17].

Vsaki vizualizaciji lahko določimo število dimenzij podatkov, ki jih vsebuje. Vsaka dimenzija predstavlja svoj tip informacije, ki je prisotna v vizualizaciji. Število dimenzij predstavlja nivo kompleksnosti vizualizacije. Višje ko je število dimenzij, težje je vizualizacijo oblikovati ter iz nje dobiti informacije. Zaradi tega imajo na splošno vizualizacije največ tri do štiri dimenzije [17].

Posamezno dimenzijo predstavimo z vizualnimi spremenljivkami. Vizualne spremenljivke je opisal francoski kartograf Jacques Bertin, ki je razvil teorijo vizualnega kodiranja, zato najpogosteje uporabljamo kar Bertinove vizualne spremenljivke. Predstavil je kanale, vizualne spremenljivke, ki dajejo izgled oznakam glede na vrednosti njihovih atributov. Oznake predstavljajo entitete in povezave. Pod oznake je uvrstil točke, črte, površine in volumen, pod kanale pa položaj, velikost, obliko, svetlost, orientacijo, barvo, teksturo in premik [17, 13].

1.2 Klasifikacijski sistem ATC

Opis klasifikacijskega sistema ATC (anatomsko-terapevtsko-kemična) je povzet po Javni agenciji Republike Slovenije za zdravila in medicinske pripomočke (JAZMP) [11]. ATC klasifikacija zdravil je mednarodni, univerzalni klasifikacijski sistem za zdravila. Služi kot orodje v raziskavah o porabi zdravil, pri čemer je osnovni namen tovrstnih raziskav izboljšanje kakovosti uporabe zdravil. Sestavni del tako imenovanih utilizacijskih študij je tudi prikaz in primerjava statističnih podatkov o porabi zdravil na mednarodnih ravneh.

Klasifikacijski sistem ATC razvija Center Svetovne zdravstvene organizacije (SZO) za statistično obdelavo zdravil v Oslu. Ravna se po strokovnih navodilih in smernicah Mednarodne delovne skupine SZO za statistično obdelavo zdravil.

Poglavitni cilj Centra SZO za statistično obdelavo zdravil in Mednarodne delovne skupine SZO za statistično obdelavo zdravil je ohraniti stabilno ATC oznako skozi daljši čas. Zaradi tega lahko raziskovalci proučijo trende porabe zdravil, ne da bi pri tem prišlo do prevelikih zapletov zaradi pogostih sprememb sistema.

Definirani dnevni odmerek (DDD) je domnevni povprečni vzdrževalni odmerek zdravila na dan za njegovo glavno indikacijo pri odraslih. DDD določa SZO za statistično obdelavo zdravil v Oslu za zdravila, ki jim je bila oznaka ATC že dodeljena [10].

Definirani dnevni odmerek na 1000 prebivalcev na dan (DID) je po SZO priporočen prikaz porabe ambulantno predpisanih zdravil. Izračuna se na podlagi celoletne porabe zdravil v določeni skupini ATC (v DDD) in števila prebivalcev v določenem letu (na dan 1. 7.) [18].

1.3 Struktura klasifikacijskega sistema ATC

Razumevanje strukture klasifikacijskega sistema ATC je bil pomemben del pri implementaciji končne rešitve diplomske naloge. Opis klasifikacijskega sistema ATC je povzet po viru JAZMP [12].

Klasifikacijski sistem ATC razvršča zdravila v različne skupine glede na mesto delovanja ter njihove terapevtske in kemične lastnosti. Zdravila najprej razvrsti v 14 glavnih, anatomskih skupin. Znotraj vsake skupine obstaja pet ravni, med katerimi je anatomska skupina prva raven. Vsaka anatomska skupina je omejena na določen organski sistem, razen anatomskih skupin H, J in P. Anatomsko skupino označuje velika tiskana črka. Anatomske skupine klasifikacije ATC so vidne na sliki 1.1.

A	Zdravila za bolezni prebavil in presnove
B	Zdravila za bolezni krvi in krvotvornih organov
C	Zdravila za bolezni srca in ožilja
D	Zdravila za bolezni kože in podkožnega tkiva
G	Zdravila za bolezni sečil in spolovil ter spolni hormoni
H	<i>Hormonska zdravila za sistemsko zdravljenje – razen spolnih hormonov in insulinov</i>
J	<i>Zdravila za sistemsko zdravljenje infekcij</i>
L	Zdravila z delovanjem na novotvorbe in imunomodulatorji
M	Zdravila za bolezni mišično-skeletnega sistema
N	Zdravila z delovanjem na živčevje
P	<i>Antiparazitiki, insekticidi in repeleti</i>
R	Zdravila za bolezni dihal
S	Zdravila za bolezni čutil
V	Razna zdravila

Slika 1.1: Anatomske skupine klasifikacije ATC.

Druga raven klasifikacije ATC je lahko glavna terapevtska skupina, razvidna na sliki 1.2 ali pa glavna farmakološka skupina, razvidna na sliki 1.3. Drugo raven označuje dvomestno število. Tretja raven je terapevtska oziroma terapevtsko-farmakološka podskupina, označena z veliko tiskano črko. Klasifikacija ATC vključuje še dve dodatni ravni. Četrta raven je terapevtsko-farmakološko-kemična podskupina, ki jo označuje velika tiskana črka. Na peti ravni je navedena učinkovina, ki jo označuje dvomestno število. Oznako ATC vsakega zdravila torej sestavlja 7 številsko-črkovnih znakov.

C	ZDRAVILA ZA BOLEZNI SRCA IN OZILJA (<i>prva raven, anatomska skupina</i>)
C02	ANTIHIPERTENZIVI (<i>druga raven, terapevtska skupina</i>)
C02A	ANTIADRENERGIKI S CENTRALNIM DELOVANJEM (<i>trejta raven, farmakološka podskupina</i>)
C02A C	AGONISTI IMIDAZOLINSKIH RECEPTORJEV (<i>četrti raven, farmakološka podskupina</i>)
C02A C01	KLONIDIN (<i>peta raven, učinkovina – INN</i>)

Slika 1.2: Struktura klasifikacijskega sistema ATC; druga raven je glavna terapevtska skupina.

C	ZDRAVILA ZA BOLEZNI SRCA IN OZILJA (<i>prva raven, anatomska skupina</i>)
C08	ZAVIRALCI KALCIJEVIH KANALCKOV (<i>druga raven, farmakološka skupina</i>)
C08D	SELEKTIVNI ZAVIRALCI KALCIJEVIH KANALCKOV Z DIREKTNIM DELOVANJEM NA PREVODNI SISTEM SRCA (<i>trejta raven, terapevtsko-farmakološka podskupina</i>)
C08D A	FENILALKILAMINSKI DERIVATI (<i>četrti raven, kemična podskupina</i>)
C08D A01	VERAPAMIL (<i>peta raven, učinkovina – INN</i>)

Slika 1.3: Struktura klasifikacijskega sistema ATC; druga raven je glavna farmakološka skupina.

1.4 Željena funkcionalnost aplikacije

Pred implementacijo spletne aplikacije smo se z Nacionalnim inštitutom za javno zdravje (NIJZ) domenili, kaj naj bi aplikacija vsebovala.

1. Izdelek bo v obliki spletne aplikacije.
2. Podatki o zdravilih bodo prikazani do ravni klasifikacije ATC 3.
 - Podatki bodo vezani na regije Slovenije.
 - Zajeti bodo podatki od leta 2009 do 2017.

3. Vizualizacije v aplikaciji naj bodo interaktivne.
 - Vizualizacija, ki bo predstavljala vse tri ravni klasifikacije ATC.
 - Vizualizacija, ki bo predstavljala porabo zdravil po regijah.
 - Vizualizacija, ki bo predstavljala trend porabe zdravil po letih.
4. Uporabniku bo omogočena primerjava med dvema regijama Slovenije.
5. Aplikacija bo uporabniku omogočila izpostavitve njemu zanimivih zdravil.

Podatke o zdravilih smo predstavili do ravni klasifikacije ATC 3. Tako smo se odločili zaradi ogromnega števila podatkov. Kot je bilo opisano v opisu klasifikacije ATC, so zdravila najprej razvrščena v 14 anatomskih skupin. V vsaki skupini je 5 ravni. Vsaka naslednja raven ima več različnih kategorij. Zaradi tega smo se odločili, da bi bila vizualizacija, ki bi predstavljala vseh 5 ravni vseh skupin, nepregledna. NIJZ nam je pripravil podatke za vse skupine do 3. ravni klasifikacije ATC. Domenili smo se, da bodo podatki vezani na določeno regijo Slovenije in določeno leto. Za vsako regijo so bili podatki za vsako zdravilo do 3. ravni klasifikacije ATC zbrani od leta 2009 do 2017. Poleg tega so nam za vsako zdravilo pripravili tudi podatek o definirani dnevni dozi, številu izdanih škatel tega zdravila, vrednosti vseh škatel v evrih in številu izdanih receptov zanj.

Interaktivnost naših vizualizacij je uporabniku omogočila različne poglede na podatke. Za raziskovanje po skupinah zdravil in ravneh smo uporabili vizualizacijo s krogi. Vsaka skupina zdravil je predstavljena z enim barvnim krogom. Večja velikost kroga pomeni večjo porabo zdravila. Na takšen način smo uporabniku predstavili vse skupine in ravni zdravil na recept.

Med podatki o vsakem zdravilu sta bila tudi določeno leto in regija, na katero se je poraba zdravila nanašala. Zaradi tega smo se odločili implementirati vizualizacijo, ki bi prikazovala porabo zdravil glede na ta dva podatka. Za prikaz porabe zdravil po regijah smo uporabili DID, ker je vrednost normirana na število prebivalcev. Vsako regijo Slovenije smo obarvali z eno izmed

petih odtenkov zelene barve. Bolj ko je odtenek zelene intenziven, večja je poraba zdravila. Z uporabo različnih intenzitet barve lahko uporabnik hitro vidi, v katerih regijah je poraba zdravila, na katerega se podatek nanaša, visoka ali nizka.

Ena izmed željenih predstavitev podatkov je bila poraba zdravil skozi čas. V vizualizaciji je bilo potrebno prikazati število izdanih škatel zdravil in ustrezno denarno vrednost le-teh po posameznih letih. Za prikaz teh podatkov smo se odločili implementirati raztreseni grafikon. Na x-os smo nanegli letnice, na y-os pa število škatel. Za prikaz podatkov v posameznem letu smo uporabili piko določene površine. Večja površina pike je predstavljala višjo denarno vrednost za število škatel, izdanih v izbranem letu. S tem smo uporabniku na enostaven način predstavili trend porabe zdravil skozi čas.

Da bi uporabnika pritegnili pri raziskovanju aplikacije, smo v sodelovanju z NIJZ k vizualizacijam dodali posebne možnosti prikaza določenih zdravil. Izbrali smo 5 uporabniku zanimivih zdravil ali bolezni, za katere se zdravila uporabljajo. Primer takega zdravila so antibiotiki. Uporabnik ob izbiri zdravila antibiotiki lahko vidi, v katerih regijah se antibiotiki največ uporabljajo. Poleg tega na vizualizaciji porabe zdravil skozi čas lahko vidi trend porabe antibiotikov.

Poglavje 2

Pregled sorodnih rešitev

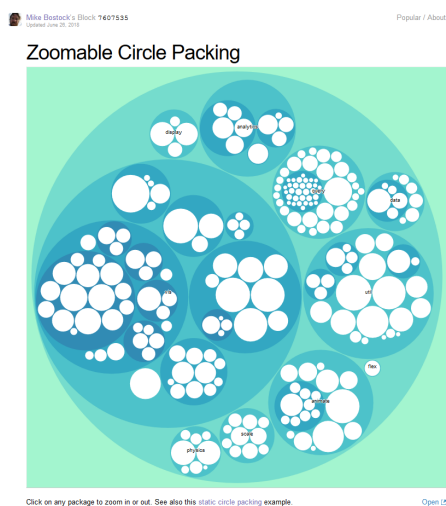
Pred izdelavo spletne vizualizacije smo najprej raziskali že oblikovane sorodne rešitve. Na naše presenečenje smo našli le eno interaktivno vizualizacijo na temo zdravil na recept (vizualizacija je opisana v poglavju 2.2). Zaradi tega smo poiskali splošne vizualizacije, ki na učinkovit način prikazujejo podobne podatke. Glavni vir že implementiranih vizualizacij je bila spletna galerija Mika Bostocka [14]. Na tej strani je že implementiranih ogromno vizualizacij. Eden izmed problemov pri iskanju ustreznih vizualizacij za našo diplomsko nalogo pa je bila interaktivnost. Mnoge od implementiranih vizualizacij niso bile interaktivne.

2.1 Vizualizacija z uporabo krogov

Zdravil na recept je ogromno število. Kot je bilo omenjeno v poglavju o strukturi klasifikacijskega sistema ATC, so zdravila razdeljena na pet ravni. V pogovoru z NIJZ smo se domenili, da bodo v vizualizacijah predstavljena zdravila do tretje ravni klasifikacije ATC. Dogovor je bil sklenjen zaradi ogromnih količin podatkov za predstavitev. Če bi uporabili vse podatke do ravni 5, bi vizualizacija postala nepregledna.

Ena izmed željenih funkcionalnosti aplikacije je bila interaktivna vizualizacija, ki bi predstavljala vse tri ravni klasifikacije ATC. Zaradi tega smo

poiskali že oblikovano vizualizacijo, ki prikazuje veliko število kategorij podatkov. Prikazane kategorije so morale biti tudi razporejene po določeni hierarhiji. Našli smo primer ustrezne interaktivne vizualizacije za omenjene zahteve [15]. Že implementirana vizualizacija je vidna na sliki 2.1.



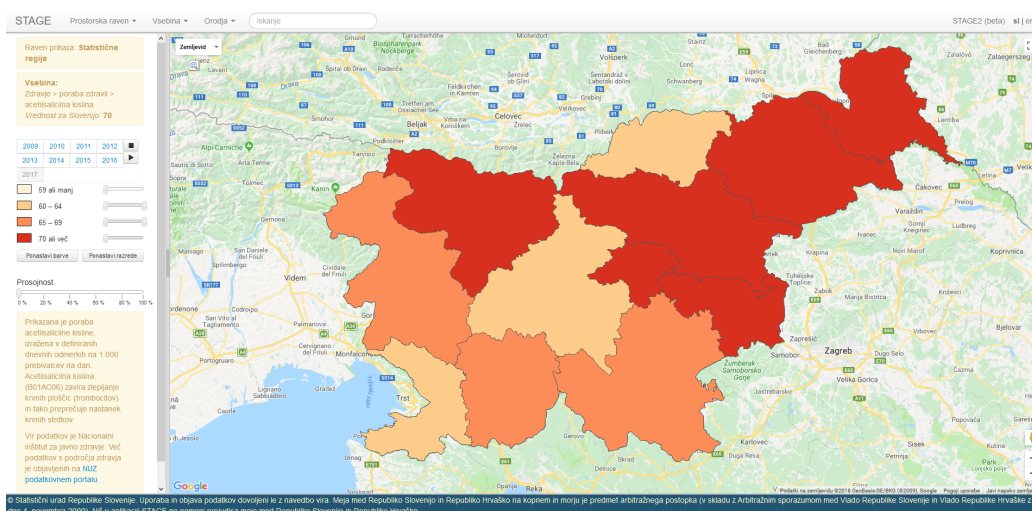
Slika 2.1: Implementirana vizualizacija krogov.

Vizualizacija predstavlja veliko število krogov. Krogi so različnih velikosti in barv. Poleg tega so nekateri krogi sestavni del drugih, večjih krogov. Taka predstavitev podatkov je po našem mnenju dober način prikaza hierarhije podatkov. Ob pritisku miške na določen krog, ga vizualizacija uporabniku približa. Poleg tega se v sredini kroga prikaže poljuben napis. Za naš primer napisi v krogih niso prišli v poštev zaradi predolгих imen zdravil. Poleg tega bi zaradi ogromnega števila besed v krogih vizualizacija postala nepregledna.

2.2 Barvni zemljevid

Edina interaktivna vizualizacija, ki prikazuje podobne podatke, je bila vizualizacija barvnega zemljevida v spletni aplikaciji STAGE [21]. Vizualizacija je implementirana kot zemljevid sveta, na katerem območja Slovenije spre-

minjajo barvo glede na izbrane podatke. Uporabnik lahko izbira izmed več različnimi temami podatkov, na primer število podjetij v Sloveniji, povprečna plača prebivalstva, zdravila itd. Poleg tega lahko izbira prikaz podatkov glede na različno velika področja Slovenije. Možnost ima izbirati med statističnimi regijami, občinami, naselji in manjšimi območji Slovenije. Primer vizualizacije barvnega zemljevida je viden na sliki 2.2.



Slika 2.2: Primer rešitve barvnega zemljevida.

Ta vizualizacija je dober vzor implementacije prikaza podatkov po regijah Slovenije. Uporabnik lahko izbira leta od 2009 do 2017. Sestavni del zemljevida je legenda, ki pojasnjuje odtenke prikazanih barv. Kot dodatno funkcijo lahko izbere prosojnost zemljevida. S tem lahko na primer podrobno vidi, katera mesta spadajo pod obarvano regijo. Ob prekritju neke regije z miško se pojavi okno s točno številko definiranega dnevnega odmerka na 1000 prebivalcev na dan (DID) za to regijo. Kot bo razloženo v poglavju o strukturi aplikacije, smo podobno funkcijo implementirali tudi v naši vizualizaciji barvnega zemljevida.

Poglavje 3

Uporabljene tehnologije

Pri izdelavi diplomske naloge smo za spletno stran uporabili označevalna jezika HTML in CSS, ter programski jezik Javascript. Ti jeziki so najpogosteje uporabljena orodja za izdelavo spletnih strani. Za uvažanje in iskanje po podatkih smo uporabili podatkovno bazo Elasticsearch, ki je razmeroma enostavna za uporabo. Poleg tega omogoča hitro iskanje in možnost kombiniranja iskalnih nizov po vnešenih podatkih. Glavno orodje za naše vizualizacije je bila knjižnica D3.js, ki je najpopularnejša za ustvarjanje interaktivnih vizualizacij za spletne strani. Poleg tega je odprtokodna in učinkovita pri prilagajanju vizualizacij za potrebe uporabnika. Za pomoč pri ustvarjanju spletne aplikacije in vizualizacij smo uporabili tudi knjižnico jQuery. Z njo smo si olajšali programiranje v programskem jeziku Javascript.

3.1 HTML in CSS

Hypertext Markup Language (HTML) je standardni jezik za izdelavo spletnih strani in spletnih aplikacij. V naši aplikaciji smo uporabili najnovejšo verzijo, HTML 5. Ta verzija ima dodatno podporo za multimedijske vsebine, kot na primer video in zvočne vsebine. Poleg tega podpira uporabo SVG elementov (scalable vector graphics), s pomočjo katerih smo ustvarjali vizualizacije. Poleg tega smo za izgled spletne aplikacije uporabili jezik CSS3, ki skrbi

za prezentacijo spletnih strani. Z njim definiramo stil HTML elementov v smislu pravil, kako naj se ti prikažejo na strani. Jezik lahko določa barve, velikosti, odmike, poravnave in vrsto drugih atributov, prav tako pa lahko nadzira aktivnosti, ki jih uporabnik nad elementi strani izvaja (na primer prekritje povezave z miško) [5, 8, 3].

3.2 Javascript

Javascript je objektni skriptni programski jezik za ustvarjanje interaktivnih spletnih strani. Poleg HTML in CSS je Javascript ena izmed glavnih tehnologij svetovnega spleta. Uporabniku omogoča dinamično spreminjanje spletnih in multimedijskih vsebin [7]. Za ta jezik je implementiranih veliko knjižnic, kot na primer D3.js, ki smo jo uporabili mi.

3.3 D3.js

D3.js (Data Driven Documents) je knjižnica za programski jezik Javascript. Ustvaril jo je Mike Bostock. Namenjena je za izdelavo dinamičnih in interaktivnih vizualizacij z uporabo standardov HTML, SVG in CSS. Knjižnica je izjemno hitra in omogoča delovanje z velikim številom podatkov. Podpira jo večino sodobnih brskalnikov. [4].

3.4 Bootstrap

Pri izdelavi aplikacije smo uporabili tudi ogrodje Bootstrap. To ogrodje je odprtokodno. Uporablja se ga za izdelavo in oblikovanje spletnih strani in spletnih aplikacij. Vsebuje že narejene predloge, implementirane v jezikih HTML in CSS. Primeri teh predlogov so gumbi, sezname, obrazci in druge komponente aplikacij. Podpira večino modernih brskalnikov, kot so Firefox, Chrome, Opera in Microsoft Edge [2].

3.5 ElasticSearch

Iskalnik ElasticSearch je odprtokodni iskalnik. Zgrajen je na osnovi hitrega in kompleksnega iskalnika Apache Lucene. ElasticSearch je napisan v programskem jeziku Java. Zaradi vgrajenega preprostega RESTful API-ja je uporaba prvotno kompleksnega iskalnika poenostavljena.

ElasticSearch ni le iskalnik, lahko se ga opiše kot:

- porazdeljena realno časovna shramba dokumentov, kjer je vsako polje indeksirano in omogoča iskanje po njem,
- porazdeljeni iskalnik, ki omogoča analizo v realnem času,
- sposoben uporabljati na stotine strežnikov in petabajte strukturiranih in nestrukturiranih podatkov.

Vse našteje funkcionalnosti združi v samostojni strežnik, ki komunicira preko preprostega RESTful API-ja [16].

Poglavje 4

Vizualizacija porabe zdravil na recept

4.1 Vnos v bazo

Kot je bilo že omenjeno, smo dobili podatke o zdravilih od Nacionalnega inštituta za javno zdravje. Podatke smo dobili v obliki excel datoteke. Za uvoz v bazo Elasticsearch smo morali podatke pretvoriti v format JSON [6]. Za preoblikovanje podatkov smo napisali ustrezno skripto v jeziku Python. Primer podatkov za zdravilo je prikazan na sliki 4.1.

```
{ "index": { "_id": "0" } }  
{ "id": "0", "year": 2009, "region_id": 9, "region_name": "Gorenjska", "atc_code": "A", "atc_name":  
"Zdravila za bolezni prebavil in presnove", "DDD": 10945047.76, "DID": 148.61, "boxes_sum":  
412201, "money_worth": 5364952.82, "perscriptions_sum": 156551, "level": 1 }
```

Slika 4.1: Primer podatkov za zdravilo.

V excelovi tabeli, ki smo jo dobili od NIJZ, je za vsako zdravilo manjkal podatek o DID zdravila. Manjkajoč podatek smo izračunali sami. Za pravilen izračun DID smo potrebovali DDD zdravila ter število prebivalcev v določeni regiji. Podatke o številu prebivalcev posamezne regije Slovenije smo dobili na spletni strani statističnega urada [22]. Za pravilen izračun DID

smo uporabili enačbo (4.1). Enačbo smo pridobili iz ustreznega dokumenta NIJZ [18]. Št. DDD v enačbi predstavlja celoletno porabo določenega zdravila. Izračunani DID smo v vizualizacijah uporabljali za primerjavo porabe zdravil med regijami.

$$\text{DID} = \frac{(\text{št. DDD}) * 1000}{(\text{št. prebivalcev}) * 365} \quad (4.1)$$

4.2 Struktura aplikacije

Celotna aplikacija je sestavljena iz treh različnih interaktivnih vizualizacij. Vsaka uporabniku predstavi podatke iz različnega zornega kota. Pri oblikovanju spletne aplikacije smo sodelovali s študentko Naravoslovnotehniške fakultete. Spletno aplikacijo smo implementirali uporabniku prijazno. Ob vsaki vizualizaciji so podana ustrezna navodila, ki uporabniku omogočajo boljše razumevanje in lažjo uporabo vizualizacij. Položaji gumbov in besedil so pri večini vizualizacij postavljeni na podobnih mestih. Tako je aplikacija preglednejša in urejena.

4.2.1 Uvodna stran

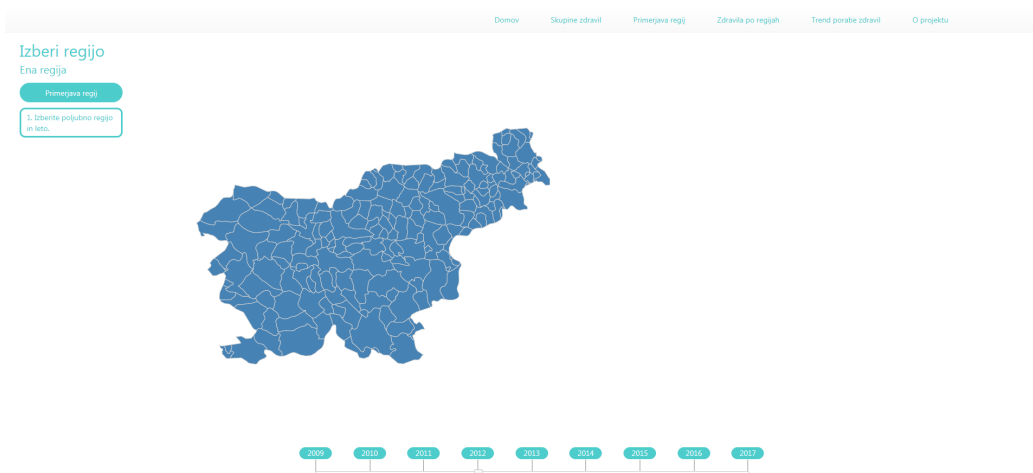
Primer uvodne strani je viden na sliki 4.2. Na uvodni strani je simbolična slika zdravil ter uvodno besedilo. Besedilo vsebuje kratek povzetek namena spletne aplikacije. Pod besedilom so trije gumbi: skupine zdravil, zdravila po regijah, trend porabe zdravil. Ob pritisku na določen gumb, je uporabnik preusmerjen na določeno vizualizacijo. Na zgornji strani aplikacije je podan meni gumbov. Meni je izdelan s pomočjo ogrodja Bootstrap [1].



Slika 4.2: Uvodna stran spletne aplikacije.

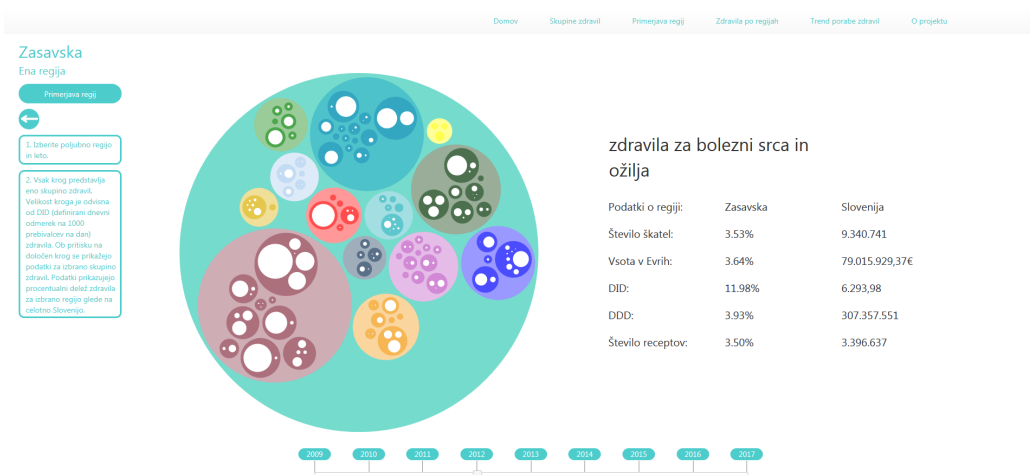
4.2.2 Vizualizacija porabe zdravil v izbrani regiji po skupinah

Ob pritisku na gumb "skupine zdravil" aplikacija uporabnika preusmeri na novo stran, vidno na sliki 4.3. Vizualizacija omogoča izbiro določene regije in leta. Zemljevid Slovenije, razdeljen po regijah, smo dobili v spletni učilnici Fakultete za računalništvo in informatiko [9].



Slika 4.3: Izbira regije in leta.

Na levi strani aplikacije je gumb "Primerjava regij", ki uporabnika preusmeri na primerjavo dveh regij. Pod tem gumbom je besedilo, ki uporabniku predstavi, kaj lahko na tej strani naredi. Na spodnji strani zemljevida Slovenije je časovnica, na kateri so razporejena leta od 2009 do 2017. S pritiskom na modro polje z izbranim letom, uporabnik lahko izbere leto. Nato izbere regijo.

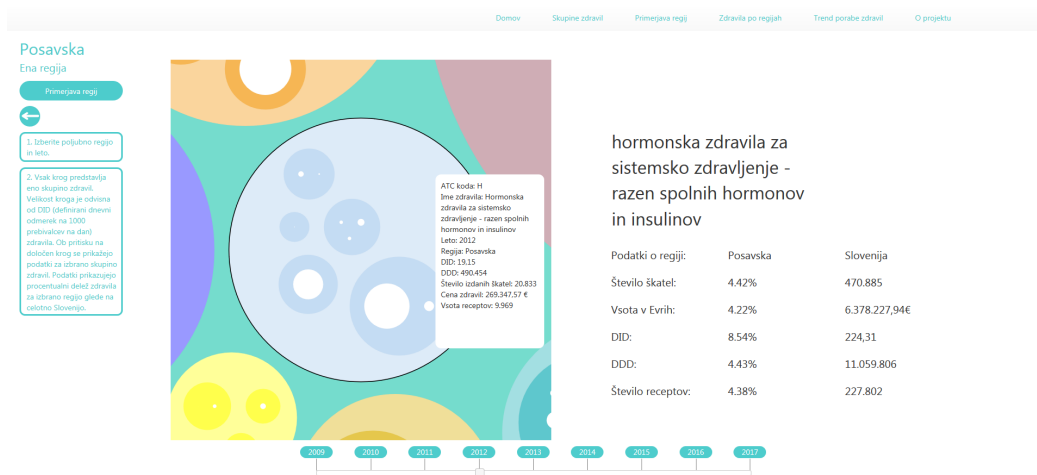


Slika 4.4: Vizualizacija skupin zdravil za zasavsko regijo za leto 2012. Izbrani so podatki o skupini zdravil "zdravila za bolezni srca in ožilja".

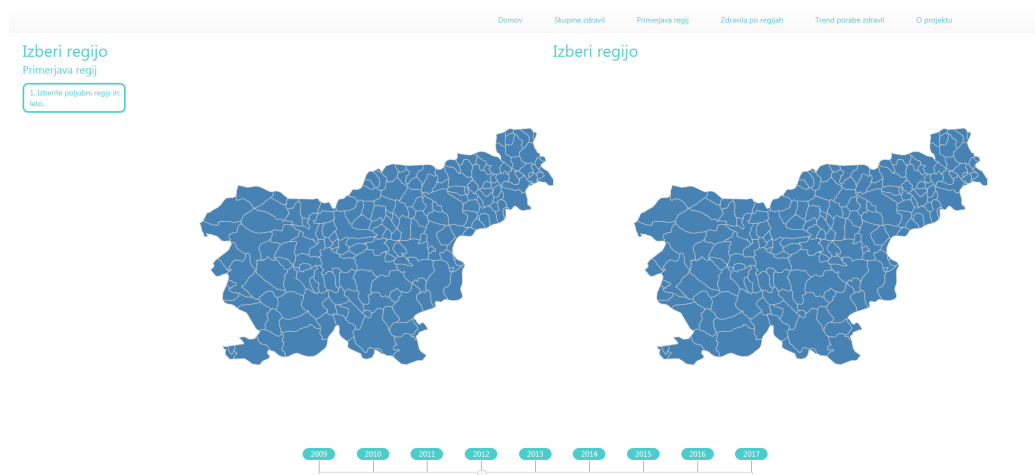
Ob izbiri regije se prikaže vizualizacija, vidna na sliki 4.4. Vizualizacija predstavlja porabo zdravil po skupinah za izbrano regijo in izbrano leto. Na zgornji levi strani se izpiše ime regije. Pod njo je okrogel gumb s puščico, ki uporabnika preusmeri nazaj na izbiro regije. Poleg tega se pojavi dodatno besedilo z opisom, kaj vizualizacija predstavlja. V primeru na sliki 4.4 je prikazana poraba zdravil za zasavsko regijo za leto 2012.

Vsaka izmed 14 skupin zdravil klasifikacije ATC je predstavljena kot skuppek krogov z določenimi barvami. Ravni skupin so predstavljene kot manjši krogi, narisani znotraj večjih. Vsaka raven je obarvana drugače, da uporabnik prepozna strukturo posameznih skupin. Prva raven skupine je predstavljena kot največji krog določene barve. Drugo raven zdravil predstavljajo krogi znotraj kroga prve ravni. Krogi te ravni so obarvani s temnejšim odtenkom barve prve ravni. Krogi tretje ravni so narisani znotraj krogov druge ravni. Barva krogov tretje ravni je bela. Velikost krogov je odvisna od DID zdravila. Ob naslonitvi miške na določen krog se okoli kroga prikaže črna obroba. Poleg tega se desno od miške prikaže bel prostor, v katerem so izpisani podrobni podatki o zdravilu za določeno leto in regijo.

Na desni strani vizualizacije so navedeni nekateri podatki za zdravilo: število izdanih škatel zdravila, denarna vrednost izdanih škatel v evrih, DDD, DID in število izdanih receptov. Podatki so podani za izbrano regijo in celotno Slovenijo. Uporabnik lahko na izbrani krog v vizualizaciji pritisne z miško. Ob pritisku na poljuben krog se spremenijo predstavljeni podatki o zdravilu. Za izbrano zdravilo se izračunajo podatki o porabi zdravila za celotno Slovenijo v določenem letu. Poleg tega se izračunajo ustrezni deleži porabe zdravila za izbrano regijo. Primer prikaza ob pritisku na poljuben krog je viden na sliki 4.5. Uporabnik lahko za trenutno izbrano regijo poljubno spreminja leto na spodnji časovnici. Ob zamenjavi leta se vizualizacija ponovno izriše za izbrano regijo in leto. Uporabniku tako vizualizacija omogoča raziskovanje po podatkih o porabi zdravil na recept skozi vse tri ravni klasifikacije ATC.

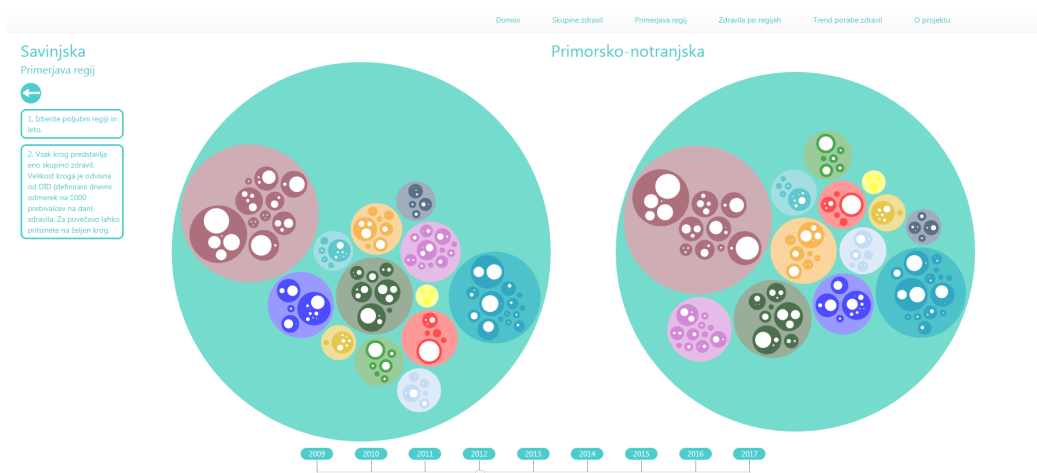


Slika 4.5: Primer izbrane skupine zdravil "hormonska zdravila za sistemsko zdravljenje - razen spolnih hormonov in insulinov".



Slika 4.6: Stran primerjave dveh regij.

Kot je bilo že omenjeno, lahko uporabnik izbere tudi primerjavo dveh regij. Na strani za primerjavo dveh regij uporabnik najprej izbere dve poljubni regiji iz dveh zemljevidov Slovenije. Stran primerjave dveh regij je prikazana na sliki 4.6. Ob izbiri obeh regij, se prikažeta dve vizualizaciji o porabi zdravil po skupinah za vsako regijo posebej. Ob pritisku na vsako vizualizacijo posebej se uporabniku približujejo in izpisujejo podatki za vsako regijo. Uporabnik lahko tako primerja porabo zdravil med dvema regijama. Uporabljene barve krogov za skupine zdravil so za obe regiji enake. Uporabniku je tako lažje primerjati porabo med enakimi skupinami zdravil. Primer primerjave med regijama Savinjska in Primorsko-notranjska za leto 2012, je viden na sliki 4.7.

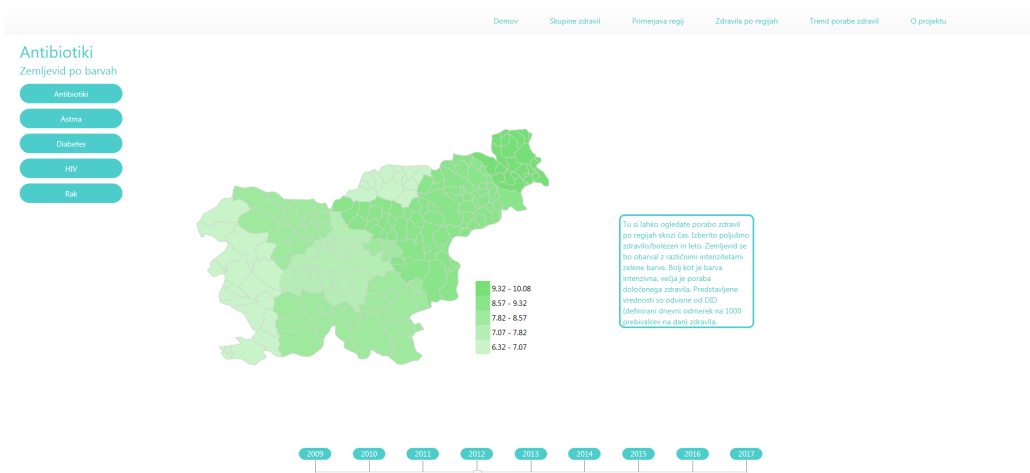


Slika 4.7: Primerjava porabe zdravil po skupinah dveh regij.

4.2.3 Vizualizacija porabe zdravil po regijah

Na uvodni strani aplikacije ima uporabnik možnost pritisniti na gumb "zdravila po regijah", ki uporabnika preusmeri na stran, prikazano na sliki 4.8. Vizualizacija je predstavljena kot zemljevid Slovenije, ki je obarvan v različne odtenke zelene barve. Bolj ko je odtenek zelene intenziven, večja je poraba skupine zdravil za določeno regijo.

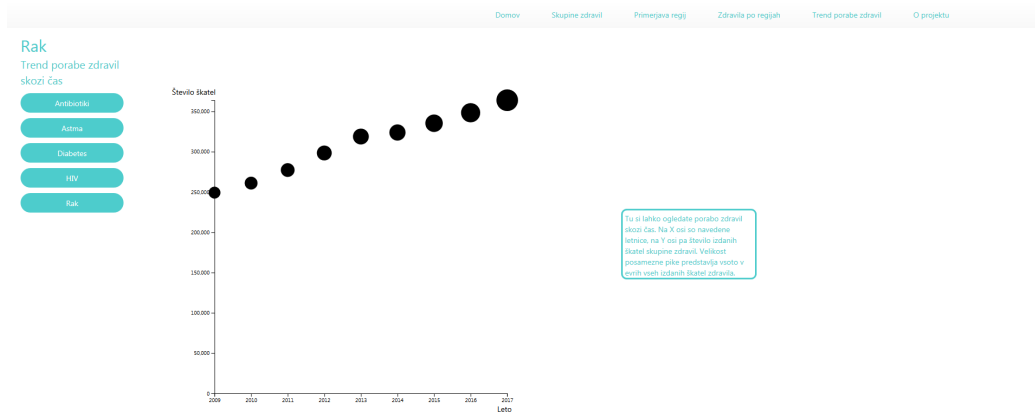
Na levi strani je 5 gumbov. V notranjosti vsakega gumba je napisana določena bolezen ali skupina zdravil. Ob pritisku na izbrani gumb se zemljevid Slovenije obarva glede na porabo izbranega zdravila ali zdravila za izbrano bolezen. Na tak način je uporabniku predstavljenih 5 uporabniku zanimivih zdravil ali bolezni. Glede izbire zdravil in bolezni smo se uskladili z NIJZ. S tem smo hoteli uporabnika vzpodbuditi k raziskovanju sprememb porabe zdravil za določene bolezni.



Slika 4.8: Vizualizacija porabe zdravil "antibiotiki" po regijah.

Primer zdravil, ki so zanimiva za uporabnike, so antibiotiki, katerih porabo po regijah v letu 2012 prikazuje vizualizacija na sliki 4.8. Uporabnik lahko ob izbranem gumbu zdravila ali bolezni izbira tudi določeno leto na časovnici spodaj, od leta 2009 do 2017. Vsak izbrani gumb zdravila ali bolezni uporabniku predstavi, kolikšna je poraba izbranega zdravila za izbrano bolezen v posameznih regijah. Višja intenziteta zelene barve pomeni večjo porabo zdravil. Poraba zdravil za vsako obravnavano bolezen je prikazana glede na DID zdravila za posamezno regijo in posamezno leto. Na vsaki takšni strani je tudi legenda, ki pojasnjuje pomen posameznih odtenkov zelene barve v obliki številčne vrednosti DID. Desno od zemljevida Slovenije sta opis vizualizacije in navodila za njeno uporabo. S tem uporabniku izboljšamo uporabniško izkušnjo naše aplikacije.

4.2.4 Vizualizacija trenda porabe skozi čas



Slika 4.9: Vizualizacija trenda porabe zdravil za bolezni raka.

Do tretje vizualizacije uporabnik lahko dostopa s pritiskom na gumb "trend porabe zdravil" ali iz uvodne strani aplikacije ali menija na zgornji strani aplikacije. Vizualizacija predstavlja trend porabe zdravil skozi čas v obliki raztresenega grafikona. Na levi strani so podani enaki gumbi kot pri vizualizaciji o porabi zdravil po regijah. Na x-osi grafa so letnice od 2009 do 2017, na y-osi pa število izdanih škatel za izbrano skupino zdravil. Točka na grafikonu predstavlja skupno število izdanih škatel v posameznem letu. Velikost točke predstavlja vrednost izdanih škatel v evrih. Zaradi velikih razlik med denarnimi vrednostmi v posameznih letih smo velikosti točk normalizirali glede na skupno denarno vrednost izdanih škatel. S pomočjo knjižnice D3.js smo določili najmanjšo in največjo dovoljeno površino glede na minimalno in maksimalno denarno vsoto. Na desni strani grafikona je besedilo, v katerem je opisan namen in navodila za uporabo vizualizacije. Primer strani z raztresenim grafikonom je na sliki 4.9.

Poglavje 5

Sklepne ugotovitve

Cilj diplomske naloge je bil na enostaven in uporabniku prijazen način predstaviti ogromno število podatkov o porabi zdravil na recept. V obliki spletne aplikacije smo omogočili širšo dostopnost implementiranih vizualizacij.

Med implementacijo spletne aplikacije smo naleteli tudi na več težav. Ena izmed njih je bil prikaz aplikacije na različnih napravah. Spletna aplikacija je bila primarno narejena namiznim računalnikom. Poleg tega smo glede na implementirane vizualizacije in obliko aplikacije ugotovili, da je najmanjša primerna resolucija na zaslonu 1810 x 770 slikovnih pik. Če je resolucija manjša od te meje, se položaj nekaterih elementov aplikacije spremeni. Poleg tega bi ob zmanjšanju velikosti vizualizacij le-te zaradi velike količine prikazanih podatkov postale nepregledne. Iz naštetih razlogov spletna aplikacija ni primerna za uporabo na mobilnih telefonih.

Poleg tega smo se soočili s problemom dostopnosti spletne aplikacije. Omrežje Nacionalnega inštituta za javno zdravje je namreč imel odprta le vrata 80. Naša aplikacija pa je za dostop do podatkovne baze uporabljala tudi vrata 9200. Problem smo rešili z zamenjavo uporabe vrat 9200 na vrata 80.

Menimo, da je implementiran izdelek uspešen, poleg tega pa ima veliko prostora za nadgradnjo. Za prikaz podatkov se lahko implementira še nove vizualizacije. Tudi dodajanje novih podatkov za prihodnja leta je z

manjšimi spremembami aplikacije mogoče. Možno bi bilo implementirati skripto, preko katere bi enostavno dodajali podatke o porabi zdravil za nadaljnja leta. Za zahtevnejše uporabnike bi se lahko implementiralo iskalnik po zdravilih, preko katerega bi lahko našli in vizualizirali porabo izbranega zdravila. Ob zanimanju za te nadgradnje se spletno aplikacijo lahko še naprej razvija.

Literatura

- [1] Bootstrap. Dosegljivo: <https://getbootstrap.com/>, 2018. [Dostopano 31. 8. 2018].
- [2] Bootstrap (front-end framework). Dosegljivo: [https://en.wikipedia.org/wiki/Bootstrap_\(front-end_framework\)](https://en.wikipedia.org/wiki/Bootstrap_(front-end_framework)), 2018. [Dostopano 31. 8. 2018].
- [3] CSS. Dosegljivo: <https://sl.wikipedia.org/wiki/CSS>, 2018. [Dostopano 8. 8. 2018].
- [4] D3 Data-Driven Documents. Dosegljivo: <https://d3js.org>, 2018. [Dostopano 31. 8. 2018].
- [5] HTML5. Dosegljivo: <https://en.wikipedia.org/wiki/HTML5>, 2018. [Dostopano 8. 8. 2018].
- [6] Introducing JSON. Dosegljivo: <https://www.json.org/>, 2018. [Dostopano 14. 8. 2018].
- [7] JavaScript. Dosegljivo: <https://en.wikipedia.org/wiki/JavaScript>, 2018. [Dostopano 9. 8. 2018].
- [8] Scalable Vector Graphics. Dosegljivo: https://sl.wikipedia.org/wiki/Scalable_Vector_Graphics, 2018. [Dostopano 8. 8. 2018].
- [9] Učilnica FRI 17/18. Dosegljivo: <https://ucilnica.fri.uni-lj.si/mod/folder/view.php?id=21068>, 2018. [Dostopano 22. 8. 2018].

- [10] Javna agencija Republike Slovenije za zdravila in medicinske pripomočke. Definirani dnevni odmerek. Dosegljivo: <http://www.jazmp.si/humana-zdravila/podatki-o-zdravilih/atchum-klasifikacija/definirani-dnevni-odmerek-defined-daily-dose-ddd/>, 2018. [Dostopano 21. 8. 2018].
- [11] Javna agencija Republike Slovenije za zdravila in medicinske pripomočke. Namen klasifikacijskega sistema ATC. Dosegljivo: <http://www.jazmp.si/humana-zdravila/podatki-o-zdravilih/atchum-klasifikacija/>, 2018. [Dostopano 19. 7. 2018].
- [12] Javna agencija Republike Slovenije za zdravila in medicinske pripomočke. Struktura klasifikacijskega sistema ATC. Dosegljivo: <http://www.jazmp.si/humana-zdravila/podatki-o-zdravilih/atchum-klasifikacija/struktura-klasifikacijskega-sistema-atc/>, 2018. [Dostopano 20. 7. 2018].
- [13] Jacques Bertin. *Semiology of Graphics: Diagrams, Networks, Maps*. 1983.
- [14] Mike Bostock. Gallery. Dosegljivo: <https://github.com/d3/d3/wiki/gallery>, 2018. [Dostopano 18. 7. 2018].
- [15] Mike Bostock. Zoomable Circle Packing. Dosegljivo: <https://bl.ocks.org/mbostock/7607535>, 2018. [Dostopano 3. 8. 2018].
- [16] Clinton Gormley and Zachary Tong. *Elastisearch: The Definitive Guide*. O'Reilly Media, 2015.
- [17] Noah Iliinsky and Julie Steele. *Designing data visualisations*. O'Reilly Media, 2011.
- [18] Nacionalni inštitut za javno zdravje. Izdana zdravila na recept. Dosegljivo: https://podatki.nijz.si/docs/6_61_RECEPTI_2016.pdf, 2018. [Dostopano 30. 7. 2018].

- [19] Nacionalni inštitut za javno zdravje. Podatkovni portal. Dosegljivo: https://podatki.nijz.si/pxweb/sl/NIJZ%20podatkovni%20portal/?px_language=sl&px_db=NIJZ+podatkovni+portal&rxid=b0abbc97-77ad-42dd-b9b9-6e94c16c665d, 2018. [Dostopano 21. 8. 2018].
- [20] Scott Murray. *Interactive Data Visualisation for the Web*. O'Reilly Media, 2011.
- [21] Statistični urad republike Slovenije. Stage. Dosegljivo: <http://gis.stat.si/>, 2018. [Dostopano 2. 8. 2018].
- [22] Statistični urad republike Slovenije. Število prebivalcev, statistične regije. Dosegljivo: http://pxweb.stat.si/pxweb/Database/Dem_soc/05_prebivalstvo/10_stevilo_preb/10_05C20_prebivalstvo_stat_regije/10_05C20_prebivalstvo_stat_regije.asp, 2018. [Dostopano 30. 7. 2018].