

UNIVERZA V LJUBLJANI
EKONOMSKA FAKULTETA

MAGISTRSKO DELO

**IZDELAVA POSLOVNO INTELIGENČNE NADZORNE PLOŠČE ZA
VIZUALIZACIJO PODATKOV O REDNIH ZOBOZDRAVSTVENIH
PREGLEDIH ZAVAROVANCEV V IZBRANI ZAVAROVALNICI**

Ljubljana, september 2024

LINA FENEVA

IZJAVA O AVTORSTVU

Podpisana Lina Feneva, študentka Ekonomske fakultete Univerze v Ljubljani, avtorica predloženega dela z naslovom Izdelava poslovno inteligenčne nadzorne plošče za vizualizacijo podatkov o rednih zobozdravstvenih pregledih zavarovancev v izbrani zavarovalnici, pripravljenega v sodelovanju s svetovalcem doc. dr. Luko Tomatom

IZJAVLJAM

1. da sem predloženo delo pripravila samostojno;
2. da je tiskana oblika predloženega dela istovetna njegovi elektronski obliki;
3. da je besedilo predloženega dela jezikovno korektno in tehnično pripravljeno v skladu z Navodili za izdelavo zaključnih nalog Ekonomske fakultete Univerze v Ljubljani, kar pomeni, da sem poskrbela, da so dela in mnenja drugih avtorjev oziroma avtoric, ki jih uporabljam oziroma navajam v besedilu, citirana oziroma povzeta v skladu z Navodili za izdelavo zaključnih nalog Ekonomske fakultete Univerze v Ljubljani;
4. da se zavedam, da je plagiatorstvo – predstavljanje tujih del (v pisni ali grafični obliki) kot mojih lastnih – kaznivo po Kazenskem zakoniku Republike Slovenije;
5. da se zavedam posledic, ki bi jih na osnovi predloženega dela dokazano plagiatorstvo lahko predstavljalo za moj status na Ekonomski fakulteti Univerze v Ljubljani v skladu z relevantnim pravilnikom;
6. da sem pridobila vsa potrebna dovoljenja za uporabo podatkov in avtorskih del v predloženem delu in jih v njem jasno označila;
7. da sem pri pripravi predloženega dela ravnala v skladu z etičnimi načeli in, kjer je to potrebno, za raziskavo pridobila soglasje etične komisije;
8. da soglašam, da se elektronska oblika predloženega dela uporabi za preverjanje podobnosti vsebine z drugimi deli s programsko opremo za preverjanje podobnosti vsebine, ki je povezana s študijskim informacijskim sistemom članice;
9. da na Univerzo v Ljubljani neodplačno, neizključno, prostorsko in časovno neomejeno prenašam pravico shranitve predloženega dela v elektronski obliki, pravico reproduciranja ter pravico dajanja predloženega dela na voljo javnosti na svetovnem spletu preko Repozitorija Univerze v Ljubljani;
10. da hkrati z objavo predloženega dela dovoljujem objavo svojih osebnih podatkov, ki so navedeni v njem in v tej izjavi;
11. da sem preverila verodostojnost informacij, ki izhajajo iz zapisov na podlagi uporabe orodij umetne inteligence.

V Ljubljani, dne _____

Podpis študentke: _____

KAZALO

1	UVOD	1
2	POSLOVNA INTELIGENCA	4
2.1	Opis in zgodovina poslovno inteligenčnih sistemov	5
2.2	Kakovost podatkov	7
2.3	Masovni podatki.....	9
2.3.1	Metode obdelave masovnih podatkov	11
2.3.2	Tehnike analize masovnih podatkov	12
2.3.3	Izzivi masovnih podatkov.....	13
2.4	Podatkovno skladišče.....	15
3	VIZUALIZACIJA PODATKOV	16
3.1	Pomen vizualizacije podatkov za podjetja.....	17
3.2	Pomen zbiranja in priprave podatkov za podjetja	17
3.3	Vizualizacija podatkov s poslovno inteligenčno nadzorno ploščo	19
3.4	Komponente poslovno inteligenčne nadzorne plošče	20
4	PREDSTAVITEV IZBRANEGA PODJETJA.....	21
4.1	Predstavitev osnovne dejavnosti.....	21
4.2	Predstavitev zobozdravstvenih storitev	22
5	IZDELAVA POSLOVNO INTELIGENČNE NADZORNE PLOŠČE ZA VIZUALIZACIJO PODATKOV V IZBRANI ZAVAROVALNICI.....	23
5.1	Metodologija.....	25
5.1.1	Pridobivanje podatkov	25
5.1.2	Izbira orodja.....	26
5.1.2.1	Jezik DAX.....	28
5.1.2.2	Jezik Power Query M.....	29
5.1.3	Urejanje podatkov	29
5.1.3.1	Čiščenje podatkov	30
5.1.3.2	Integracija podatkov	30
5.1.3.3	Transformacija podatkov	30
5.1.4	Nalaganje podatkov	30
5.2	Poslovno inteligenčna arhitektura v orodju Power BI	31

5.3 Koraki izdelave poslovno inteligenčne nadzorne plošče za vizualizacijo podatkov v izbrani zavarovalnici	32
5.3.1 Proces pridobivanja podatkov	32
5.3.2 Urejanje podatkov	33
5.3.2.1 Opis podatkov	33
5.3.2.2 Čiščenje podatkov	34
5.3.2.3 Dodajanje povezovalne tabele	36
5.3.3 Nalaganje podatkov	37
5.4 Podatkovni model.....	38
5.4.1 Razmerje med tabelami	41
5.4.2 Kardinalnost	41
5.5 Izgradnja poslovno inteligenčne nadzorne plošče za vizualizacijo podatkov o rednih zobozdravstvenih pregledih zavarovancev v izbrani zavarovalnici.....	42
6 DISKUSIJA	54
6.1 Ključne ugotovitve za izbrano zavarovalnico.....	54
6.2 Pomen ugotovitve za izbrano zavarovalnico	55
6.3 Omejitve raziskave.....	56
7 SKLEP.....	57
LITERATURA IN VIRI	58

KAZALO TABEL

Tabela 1: Primerjava različnih tehnik analize masovnih podatkov	13
Tabela 2: Opis specifikacije podatkov.....	33
Tabela 3: Analiza zavarovancev po starosti in spolu.....	50
Tabela 4: Analiza zavarovancev Skupine 1 po opravljenih storitvah pri petih najbolj pogosto obiskanih izvajalcih.....	52
Tabela 5: Analiza zavarovancev Skupine 1 po opravljenih storitvah pri petih najmanj pogosto obiskanih izvajalcih	52
Tabela 6: Analiza zavarovancev po datumu sklenitve zavarovanja	53

KAZALO SLIK

Slika 1: Primeri vizualizacijskih grafikonov	21
Slika 2: Power BI poslovno inteligenčna arhitektura	31

Slika 3: Primer prečiščenih podatkov za zavarovance Skupine 2 v tabeli Zobje Zavarovalniška pogodba	35
Slika 4: Primer podatkov v tabeli Zobje škode IZV.....	35
Slika 5: Primer podatkov v tabeli Dopolnilno škode	35
Slika 6: Datum veljavnosti zavarovanja.....	36
Slika 7: Tabela Zavarovalna vrsta	36
Slika 8: Tabela Stomatološki pregled	36
Slika 9: Tabela Produkt	37
Slika 10: Tabela Izvajalci storitev	37
Slika 11: Primer nalaganja podatkov.....	38
Slika 12: Podatkovni model	39
Slika 13: Primer tabele dimenzij	41
Slika 14: Nadzorna plošča 1: Analiza osnovne sklenjene pogodbe zavarovancev Skupine 1 in zavarovancev Skupine 2.....	43
Slika 15: Nadzorna plošča 2: Analiza stomatoloških pregledov zavarovancev Skupine 1	44
Slika 16: Nadzorna plošča 3: Analiza stomatoloških pregledov Skupine 2 in primerjava med skupinami zavarovancev	44
Slika 17: Nadzorna plošča 4: Analiza zavarovancev Skupine 1 in Skupine 2 po produktih....	45
Slika 18: Nadzorna plošča 5: Analiza zavarovancev Skupine 1 po izvajalcih storitev.....	46
Slika 19: Nadzorna plošča 6: Analiza zavarovancev Skupine 2 po izvajalcih storitev.....	47
Slika 20: Nadzorna plošča 7: Analiza zavarovancev Skupine 1 in zavarovancev Skupine 2 po izvajalcih storitev	48
Slika 21: Starost zavarovancev Skupine 1 po spolu.....	49
Slika 22: Število zavarovancev Skupine 1 po opravljenih storitvah pri izvajalcih	50
Slika 23: Število zavarovancev Skupine 2 po opravljenih storitvah pri izvajalcih	51
Slika 24: Skupno število zavarovancev Skupina 1 in Skupina 2	53
Slika 25: Zavarovanci Skupine 1 po datumu sklenitve pogodbe in spolu	54

SEZNAM KRATIC

angl. – angleško

AI – (angl. Artificial Intelligence); umetna inteligenca

BI – (angl. Business Intelligence); poslovna inteligenca

COVID-19 – (angl. coronavirus disease 2019); koronavirusna bolezen 2019

CRM – (angl. Consumer Relationship Management); upravljanje odnosov s strankami

DAX – (angl. Data Analysis Expressions); jezik z izrazi za analizo podatkov

DM – (angl. Data Mart); področno podatkovno skladišče

DW – (angl. Data Warehouse); podatkovno skladišče

ERP – (angl. Enterprise Resource Planning); celovite programske rešitve

ETL – (angl. Extract, Transform, Load); postopek pridobivanja, preoblikovanja in nalaganja podatkov

IS – (angl. Information System); informacijski sistem

IT – (angl. Information Technology); informacijska tehnologija

KPI – (angl. Key Performance Indicator); ključni kazalniki uspešnosti

OLAP – (angl. Online Analytical Processing); strežnik za spletno analitično obdelavo podatkov

RDBMS – (angl. Relational Database Management System); sistem za upravljanje relacijske zbirke podatkov

SCM – (angl. Supply Chain Management); upravljanje dobavne verige

VBA – (angl. Visual Basic for applications); jezik Visual Basic za aplikacije

1 UVOD

Sistemi za poslovno inteligenco (angl. business intelligence, v nadaljevanju BI) združujejo operativne podatke z uporabo analitičnih orodij za predstavitev kompleksnih in obsežnih informacij izbranim načrtovalcem in odločevalcem. Cilj uporabe sistema je izboljšati kakovost in pravočasnost vhodnih podatkov za proces odločanja. BI se uporablja za razumevanje zmogljivosti, ki so na voljo v podjetju, tehnik, trendov in prihodnjih usmeritev na trgu ter v tehnologijah in regulativnem okolju, v katerem podjetje tekmuje. Sistem se pojavi kot posledica različnih vrst prejšnjih sistemov, ki so bili zasnovani za podporo sprejemanju odločitev. Pojav podatkovnega skladišča (angl. data warehouse, v nadaljevanju DW) kot shrambe vodi do napredka na področju čiščenja podatkov, nato pa do večje zmogljivosti strojne in programske opreme, ki bo zagotovila prevladujoč uporabniški vmesnik, skupaj pa bodo posledično ustvarili enotno BI okolje (Negash, 2004).

BI pokriva skladiščenje podatkov v realnem času, podatkovno rudarjenje, samodejno odkrivanje anomalij in izjem, proaktivno opozarjanje s samodejnim prejemnikom, brezhibno spremljanje delovnega poteka, samodejno učenje, geografske informacijske sisteme (angl. information system, v nadaljevanju IS) in vizualizacijo podatkov (Negash, 2004).

Pri BI so pomembni tudi metapodatki. Izraz metapodatki je bil v preteklosti pogosto uporabljen in zlorabljen. Upravljalci datotek pri svojem delu uporabljajo metapodatke, kot so imena polj in imena datotek. Uporaba metapodatkov skupaj z dejanskimi podatki je močno vpeta v tehnologijo upravljanja področnih podatkovnih skladišč (angl. data mart, v nadaljevanju DM) (Sen, 2004).

Ko govorimo o podatkih, jih je v organizacijah večina zajeta v nestrukturirani obliki, ta pa se ne ujema s tipično obliko že shranjenih podatkov. Pogosto se zgodi, da so podatki shranjeni, vendar niso analizirani. Informacije, ki so skrite ali shranjene v nestrukturiranih oblikah, imajo ključno vlogo pri sprejemanju odločitev, razumevanju in izpolnjevanju predpisov ter izvajanju poslovnih aktivnosti. Povezovanje podatkov, shranjenih v strukturiranih in nestrukturiranih oblikah, lahko ima v organizaciji pomembno dodano vrednost (Gupta in Rathore, 2013).

Veliko podjetij in družb uporablja veliko količino informacij v obliki besedilno bogatih dokumentov v različnih tipih datotek: pdf, word, elektronska sporočila, klepetane datoteke, blogi, forumi organizacij in številne druge. Danes je svetovni splet največji vir informacij, organizacije pa lahko najdejo številne informacije o svojem poslovnem okolju na internetu, kar je sicer koristno, vendar pa je prav zaradi tega približno osemdeset odstotkov razpoložljivih podatkov v nestrukturiranih oblikah, kar pomeni, da jih je težko analizirati in shraniti v DW (Gupta in Rathore, 2013).

Sistemi BI omogočajo prepoznavanje napak in njihovo posledično odpravo. Sisteme BI je mogoče uporabiti tudi pri zbiranju potrebnih informacij iz besedil. Celotno DW pa se lahko uporabi kot okvir za učinkovite in natančne poslovne odločitve (Gupta in Rathore, 2013).

Večina organizacij uporablja različne informacijske rešitve, kot so sistem celovite programske rešitve (angl. enterprise resource planning, v nadaljevanju ERP), sistem upravljanja odnosov s strankami (angl. consumer relationship management, v nadaljevanju CRM) in sistem upravljanja dobavne verige (angl. supply chain management, v nadaljevanju SCM), v svojih procesih odločanja pa BI še vedno ni pogosto uporabljen. Zato lahko model in tehnike za ocenjevanje in vrednotenje ravni BI sistemov izboljšajo podporo pri odločanju. Prav ti sistemi pretvarjajo in shranjujejo podatke v svojih DM, zato jih je mogoče uporabiti kot zbirko podatkov za podporo odločitvam in raziskovanje uporabnega znanja. Zaradi možnosti pridobivanja konkurenčne prednosti pri sprejemanju pomembnih odločitev je ključno, da se podpora pri odločanju vključi v okolje podjetja in delovne sisteme. BI je mogoče vgraditi v te sisteme podjetja, da bi pridobili konkurenčno prednost (Ghazanfari in drugi, 2011).

Namen magistrskega dela je preko prikaza razvoja BI nadzorne plošče podjetjem v zavarovalniški dejavnosti izboljšati razumevanje koristi, ki jih prinaša vizualizacij podatkov. V okviru magistrskega dela sem za izbrano zavarovalnico izdelala BI nadzorno ploščo, ki pripomore k hitremu in učinkovitemu analiziranju podatkov s pomočjo vizualnega prikaza. Oblikovana BI nadzorna plošča za vizualizacijo podatkov izbrani zavarovalnici omogoča hitrejši prikaz podatkov zavarovancev ter učinkovitejši pregled in nadzor opravljanja storitev kot tudi možnost razvoja novih storitev.

Cilji za doseganje namena magistrskega dela so naslednji:

- prepoznati koristi in slabosti vizualizacije v BI sistemih za podjetja;
- predstaviti potek razvoja nadzorne plošče BI;
- predstaviti pomen podatkov, ki so potrebni za razvoj nadzorne plošče BI;
- predstaviti podrobne korake pri povezovanju podatkov na ravni izdelave in implementacije BI nadzorne plošče na konkretnem primeru;
- opredeliti koristi uporabe BI orodij za vizualizacijo podatkov v izbrani zavarovalnici.

Magistrsko delo vsebuje teoretični in praktični sklop. Na začetku je teoretični sklop, v katerem sem na podlagi kritičnega pregleda tuje in domače znanstvene in strokovne literature opisala BI sistem ter opredelila značilnosti, komponente in orodja z namenom izdelave čim bolj primerne nadzorne plošče za vizualizacijo podatkov za izbrano zavarovalnico. Podrobno sem predstavila tudi vpliv in pomen zobozdravstvenih pregledov na področju zavarovalništva.

V praktičnem sklopu magistrskega dela sem na podlagi metodologije iz literature izdelala BI nadzorno ploščo za vizualizacijo podatkov o rednih zobozdravstvenih pregledih

zavarovancev v izbrani zavarovalnici. Uporabila sem podatke izbrane zavarovalnice. Na koncu magistrskega dela sem ustrezno združila ugotovitve na podlagi vsebine teoretičnega in praktičnega sklopa ter predstavila ključne ugotovitve. Magistrsko delo je sestavljeno iz petih poglavij. Kratak opis poglavij in poudarkov sledi v nadaljevanju.

V drugem poglavju sem proučevala pomen BI, ki je predstavljala osrednjo temo magistrskega dela. Poseben poudarek je bil na zagotavljanju visokokakovostnih informacij upravljanja pri odločanju, kar predstavlja značilni pristop, ki vključuje identifikacijo, zbiranje, shranjevanje in vzdrževanje podatkov. V drugem sklopu sem opisala način pridobivanja masovnih podatkov ter uporabo tehnik analize masovnih podatkov. V tretjem sklopu pa je večji poudarek na pomenu zbiranja podatkov v DW, ki pomaga pri procesu odločanja.

Tretje poglavje predstavi vizualizacijo podatkov. Poseben poudarek je na koristi vizualnega prikaza, ki zagotavlja znatno zmanjšanje organizacijskih stroškov, poenostavitev procesov oskrbe in učinkovito analizo ter obdelavo podatkov. V tem poglavju sem proučevala tudi izbor komponent BI nadzornih plošč kot pomemben del vizualizacije, ki uporabnikom dejansko olajša prepoznavanje vzorcev in medsebojno primerjanje vrednosti.

V četrtem poglavju sem proučevala osnovne značilnosti izbrane zavarovalnice. Poudarek je bil na tem, da zavarovalnica nudi enostavne in zanesljive storitve, ki prinašajo številne prednosti ter izboljšajo kakovost življenja. V drugem delu je bil poudarek na kratki predstavitvi zobozdravstvene panoge.

V petem poglavju sem podrobneje predstavila izdelavo nadzorne plošče po korakih. Začela sem z metodologijo in opisom korakov izdelave nadzorne plošče, nato sem predstavila izbor orodij Power BI. Na kratko sem predstavila tudi arhitekturo BI sistema s sedmimi glavnimi komponentami orodja Power BI. V četrtem sklopu je prikazana posebnost priprave podatkov za izdelavo nadzornih plošč: profiliranje podatkov, nalaganje podatkov, čiščenje podatkov, integracija podatkov in transformacija podatkov. V petem sklopu sem opisala korake izdelave nadzorne plošče za vizualizacijo podatkov v izbrani zavarovalnici. Nato je sledil šesti sklop, v katerem sem prikazala ustvarjen podatkovni model z vsemi tabelami z vključenimi atributi in povezave med tabelami. V sedmem sklopu poglavja sem ustvarila sedem nadzornih plošč, ki imajo v povprečju šest vizualizacij na posamezno nadzorno ploščo, smiselno za določene analize. V osmem sklopu poglavja pa je bil poudarek na interpretaciji rezultatov posameznih grafov in analizi možnih scenarijev. Pri analizi sem upoštevala naslednje: starost in spol zavarovancev, obdobje sklenitve zavarovanja ter dostopnost izvajalcev zdravstvenih storitev.

V šestem poglavju sem predstavila ključne ugotovitve izvedene analize, nato sem prikazala ugotovitve za izbrano zavarovalnico, ki bodo pozitivno vplivale na njeno delovanje, predvsem oddelka za razvoj produktov. V naslednjih sklopih pa sem prikazala slabosti in tveganja raziskave ter možne izboljšave za nadaljnje delo.

2 POSLOVNA INTELIGENCA

BI predstavlja postopek pridobivanja, preoblikovanja, upravljanja in analiziranja velikih količin podatkov z namenom pridobivanja informacij in znanja, ki pomaga pri sprejemanju odločitev v kompleksnih situacijah. Elementi BI so DW, rudarjenje podatkov in sistem za podporo odločanju (Fitriana in drugi, 2011).

V literaturi se v splošnem omenja štiri osnovne pristope BI, ki se jih lahko interpretira na več različnih načinov. Prvi pristop je koncept upravljanja, katerega cilj je zagotoviti, da vodje dobijo pravočasne informacije ustrezne kakovosti. Po drugem pristopu je BI informacijska tehnološka (angl. information technology, v nadaljevanju IT) rešitev, kar pomeni, da gre za namenske aplikacije, ki omogočajo napredno analizo podatkov (poizvedbe, poročila in analize). Tretji pristop ga obravnava kot sistemsko arhitekturo, ki obsega predvsem DW, orodja za poizvedovanje in poročanje, spletne analitične obdelave, statistično analizo in rudarjenje podatkov. Četrty in zadnji pristop pa ga vidi kot rešitev sistema, ki izhaja iz tesnega sodelovanja IT in poslovanja, kjer BI ni le orodje za analizo podatkov, ampak tudi za pripravo in izvajanje te analize.

BI se razlikuje od sistemov za podporo odločanju in sistemov za izvršilne informacije po širšem tematskem obsegu, večkratni analizi, polstrukturiranih podatkih iz različnih virov in večdimenzionalni predstavitvi podatkov. Namen BI je omogočiti odločevalcem na vseh ravneh upravljanja, da sprejemajo boljše in pravočasne odločitve. Predpostavlja se, da lahko BI podpira odločanje na vseh ravneh upravljanja, ne glede na stopnjo njihove strukturiranosti (Olszak in Ziemba, 2007).

Glavni cilj sistema BI je običajno omogočiti odločevalcem dostop do pravih podatkov ob pravem času, kar omogoča proaktivno sprejemanje odločitev. Uporabniki sistemov BI pričakujejo dostop in predstavitev več informacij preko enostavno razumljivega vmesnika, ki naj bo intuitiven, interaktiven ter naj zagotavlja kontekst in vgrajeno znanje (Keeling in Morris, 2010).

Posebna pozornost v sistemih BI je namenjena uporabniškemu vmesniku, ki bi moral uporabniku med drugim omogočiti: filtriranje, urejanje in analiziranje podatkov, oblikovanje ad hoc ali vnaprej določenih poročil, ustvarjanje alternativnih scenarijev, izdelovanje raziskovanih grafikonov in grafov. Uporaba grafičnih vmesnikov je primarna značilnost nove generacije aplikacij BI.

Številne analitične aplikacije BI uporabljajo orodja za vizualizacijo podatkov, da organizacijam nudijo ključne poslovne informacije. Za končne uporabnike je plastičnost predstavljenega sloja najbolj kritičen vidik sistema BI, saj široko oblikuje njihovo osnovno razumevanje prikazanih podatkov (Wise, 2008). Raziskava še posebej kaže na to, da so IS, ki jih je upravljanje vrednotilo kot slabe, predstavljali informacije le v obliki besedila in

tabele. Vloga vizualizacije podatkov v okviru BI zato postaja vedno pomembnejša (Dudyys, 2010).

2.1 Opis in zgodovina poslovno inteligenčnih sistemov

Proces BI se začne z zbiranjem informacij, predstavljenih kot prva povezava med podjetjem in okoljem, pri čemer lahko razume dogajanje in ostane pozoren na spremembe. Tradicionalno se je faza zbiranja podatkov izvajala iz odprtih in človeških virov, vendar se je s pojavom interneta soočila z izzivom informacijske preobremenjenosti. Zaradi oblike podatkov je namreč prišlo do več slabih odločitev vodij, zato je bilo treba podatke prilagoditi in preoblikovati. Analitiki BI so uporabili računalniške sisteme za podporo pri odločanju, da bi pripravili zahtevane obveščevalne podatke za vodstvene delavce. Pravzaprav so takšne odločitvene pomoči sčasoma spodbudile oblikovanje IS za izvršilne organe z namenom pridobivanja informacij, povezanih z notranjim poslovanjem in poslovnim okoljem.

S postopkom pridobivanja, preoblikovanja in nalaganja podatkov (angl. extract, transform, load, v nadaljevanju ETL) se izvede podroben pregled faze preobrazbe, tako da se tako strukturirani kot nestrukturirani podatki pridobijo iz operativnih in zunanjih virov, nato pa se naložijo v DW za kasnejše združevanje v podatkovne zbirke (Talaoui in Kohtamaki, 2021).

Veliko podatkov v različnih podjetjih je običajno nestrukturiranih. To so podatki, ki se ne prilegajo tipični obliki obdelave podatkov. Največja slabost je v tem, da ostanejo shranjeni, vendar ne analizirani, informacije, ki so skrite ali shranjene v nestrukturirani obliki, pa bi lahko imele ključno oz. pomembno vlogo pri sprejemanju ključnih odločitev. Nestrukturirani podatki in posledično informacije, ki niso analizirane, vplivajo na nepravilno izvajanje poslovnih nalog. Velika dodana vrednost podjetju je povezovanje podatkov, shranjenih v strukturiranih in nestrukturiranih oblikah. Na boljši vpogled v podatke vpliva analiza in obdelava podatkov, integrirani podatki pa opredelijo infrastrukturo celotnega DW, s tem pa lahko organizacija pridobi tako imenovano enotno različico resnice (Gupta in Rathore, 2013).

BI je že več let ključen element na področju IT v številnih panogah. Na to vezana področja, kot so programi in izdelki, se hitro in nemoteno razvijajo. V zadnjem času se pojavljajo novi trendi, ki so tesno povezani z BI, kot je npr. upravljanje velikih količin podatkov, ki so prispevali k trajnostnemu razvoju programske opreme za BI. Sodobni sistemi BI se razlikujejo od prejšnjih sistemov za podporo pri odločanju zaradi več razlogov. Vključujejo namreč sistematično integracijo, združevanje in upravljanje strukturiranih in nestrukturiranih podatkov v DW. Drugi razlog je to, da se rešitve BI danes ukvarjajo z zelo velikimi podatki, ki se še povečujejo in se zanašajo na eksponentno naraščanje zmogljivosti obdelave (vključno s podatki v pomnilniku), kar je ustvarilo nove možnosti za odkrivanje znanja, kot je podatkovno rudarjenje. Tretji razlog pa so rešitve BI, ki ustrezno uporabljajo nove načine poizvedovanja po podatkih in zagotavljanja informacij, kot je samodejna

distribucija ali samopostrežna uporaba podatkov, ki je vseprisotna v različnih računalniških napravah (Gupta in Rathore, 2013).

Proces BI predstavlja analitičen, tehnološko podprt proces. Zbira in preoblikuje razdrobljene podatke organizacij in trgov v informacije ali znanje o ciljih, priložnostih, slabostih in napredku procesa v organizaciji. Programska oprema BI predstavlja programske izdelke, ki so najprej namenjeni podpori analitičnemu procesu, kot so programska oprema za podatkovno rudarjenje, programska oprema za DW, podatkovna programska oprema za digitalne nadzorne plošče itn. Orodja BI so aplikacije, ki predstavljajo programske izdelke BI, konfigurirane in uporabne v organizacijah, ter rešitev BI kot skupek orodij BI in povezanih tehnologij ter procesov, ki se uporabljajo za podporo ciljem BI (Gupta in Rathore, 2013).

Pri procesih BI ne gre le za programsko opremo in sisteme, temveč za celoten proces upravljanja podatkov, ki večinoma podpira vodstveno odločanje. Na področju pomembnih dejavnikov uspeha projektov DW je ustrezno upravljanje BI. Celoten koncept razvoja BI vključuje tudi upravljanje kot pomembno dimenzijo. S teoretičnega vidika je mogoče upravljanje BI razložiti kot proces učenja in odraz virov, ki so potrebni za združevanje programske opreme BI in organizacijske strategije v rešitve BI, ter nenehno doseganje ciljev, povezanih s procesom BI. Na drugi strani so na trgu na voljo programski izdelki BI kot sredstva, podobno pa so lahko na voljo storitve uvajanja programske opreme BI in tekoče vzdrževanje rešitev BI, ki jih je mogoče oddati tudi v zunanje izvajanje. Ključno je, da uspešno upravljanje BI zahteva tesno uskladitev IT in poslovanja v celotnem poslovnem ciklu (Gupta in Rathore, 2013).

Kakovost upravljanja BI je pozitivno povezana s kakovostjo vodstvenega odločanja. Vpliv na slednjo imajo zagotavljanje in vzdrževanje rešitev BI kot podpora učinkovitega prepoznavanja problemov in priložnosti, kritično odločanje in oblikovanje strategije. Izvajanja in vrednotenja ni mogoče v celoti oddati v zunanje upravljanje, temveč so za to potrebni notranji viri, ki ne pokrivajo le oddelka IT. Le če so viri IT in poslovne zahteve usklajeni z ustreznim upravljanjem BI, lahko organizacije hitreje uresničijo potencialne koristi sistemov BI (Wieder in Ossimitz, 2015).

Ključni sestavni del sistemov BI je DW. To je vsebinsko usmerjena, integrirana in časovno spremenljiva ogromno DM, ki pomaga pri spremljanju odločitev. Namenjena je poizvedovanju in analiziranju strukturiranih podatkov, ki jih lahko razdelimo na dejstva in dimenzije. Strukturirani podatki imajo vnaprej določeno obliko in so usmerjeni v zapise, medtem ko so nestrukturirani podatki obsežni, prosto oblikovani in obstajajo v različnih oblikah. Organizacije uspešno uporabljajo tehnologije DW in strežnik za spletno analitično obdelavo podatkov (angl. online analytical processing, v nadaljevanju OLAP) za izgradnjo sistemov za podporo odločanju organizacije in analizo velikih količin strukturiranih podatkov, ki jih podjetja hranijo v svojih DM.

DW shranjujejo zgodovinske in trenutne podatke o podjetniških okoljih in tako hranijo vse podatke na enem mestu. Pri shranjevanju je strukturirane podatke mogoče enostavno in preprosto shraniti v DW, medtem ko je shranjevanje nestrukturiranih podatkov bolj zahtevno. Do potreb po upravljanju nestrukturiranih podatkov pride, ker je več kot tri četrtine informacij na internetu nestrukturiranih. Prednost spletnih strani je, da z upravljanjem nestrukturiranih podatkov pridobimo poslovno vrednost, boljše, pravočasne in relevantne informacije (Wieder in Ossimitz, 2015).

Po eni strani DW običajno uporablja relacijske podatke (angl. relational database management system, v nadaljevanju RDBMS) za shranjevanje podatkov in hitro izvajanje poizvedb v širokem obsegu podatkov. Po drugi strani pa je DW potrjeno z OLAP, ki je zadolžen za pripravo, filtriranje in temeljito analizo podatkov, ta pa se nato posreduje v uporabo vmesniku v obliki nadzorne plošče, preglednice itn. Glavni izziv predstavlja pridobivanje pravih informacij iz podatkov, preoblikovanje podatkov in njihovo poznavanje, analiza za iskanje vzorcev in trendov, shranjevanje informacij za učinkovit dostop, upravljanje delovnih nalog in končna priprava uporabnih poročil BI (Talaoui in Kohtamaki, 2021).

Za uspešno raziskavo je nujno potrebna analiza podatkov iz več formatov in iz različnih virov v različnih časovnih okvirih. Integrirana informacijska arhitektura omogoča boljši vpogled v večdimenzionalne informacije za ciljne skupine subjektov. Omogoča tudi boljši vpogled, zmogljivejše statistične in semantične zmogljivosti poročanja. Hitrejša razumljivost branja in pisanja zagotavlja zbrane podatke v realnem času analitične zmogljivosti. Zahteva je, da je DW sposobno obdelave velikih podatkovnih nizov, ki so zahtevni za shranjevanje, analizo, iskanje, upravljanje in vizualizacijo podatkov (Wieder in Ossimitz, 2015).

2.2 Kakovost podatkov

Hitro širjenje računalniških IS povečuje pomen kakovosti podatkov tako za izdelovalce sistemov kot tudi za uporabnike. Vendar ne obstaja niti uveljavljen okvir niti skupni besednjak za preučevanje kakovosti podatkov. Pri masovnih podatkih je ključna težava v njihovi kakovosti. Obstajajo številne značilnosti masovnih podatkov, ki neposredno vplivajo na njihovo kakovost, ključne pa so natančnost, popolnost, doslednost in aktualnost podatkov (Taleb in drugi, 2016).

Podatki so najdragocenejše premoženje, s katerim podjetja razpolagajo. Ko so njihova kakovost poslabša, so posledice nepredvidljive in lahko privedejo do popolnoma napačnih ugotovitev. V kontekstu velikih podatkov je ocenjevanje kakovosti podatkov zahtevno in ga je treba opraviti pred vsako analizo velikih podatkov, hkrati pa je potrebno imeti nekaj zaupanja v kakovost podatkov. Glede na velikost podatkov in njihovo hitro generacijo so potrebni mehanizmi in strategije za hitro in učinkovito ocenjevanje kakovosti podatkov. Vendar pa je preverjanje kakovosti podatkov zelo drag postopek, če se uporablja na celotnih

podatkih. Vzorčenje bo zmanjšalo velikost podatkov na reprezentativne vzorčne populacije za hitro ocenjevanje kakovosti. Ocena je usmerjena v nekatere dimenzije kakovosti podatkov, kot sta popolnost in doslednost. Eksperimenti so bili izvedeni na naboru podatkov o motnjah spanja z uporabo tehnik vzorčenja velikih podatkov s ponavljanjem. Rezultati so pokazali, da je povprečna kakovost ocene vzorcev reprezentativna za izvirne podatke, in poudarili pomembnost vzorčenja za zmanjšanje računalniških stroškov pri ocenjevanju kakovosti velikih podatkov (Taleb in drugi, 2016).

Literatura ponuja širok nabor tehnik za ocenjevanje in izboljšanje kakovost podatkov. Zaradi raznolikosti in kompleksnosti teh tehnik se je raziskovanje osredotočilo na opredelitev metodologij, ki pomagajo pri izbiri, prilagajanju in uporabi tehnik ocenjevanja in izboljševanja kakovosti podatkov. Metodologije primerjajo po več dimenzij, vključno z metodološkimi fazami in koraki, strategijami in tehnikami, dimenzijami kakovost podatkov, vrstami podatkov in nazadnje vrstami IS (Batini in drugi, 2009).

Kakovost podatkov lahko opišemo z več vidikov. Prvi vidik so vrste podatkov, občutljive za težave s kakovostjo, pogosti razlogi za slabo kakovost podatkov in potencialni stroški ter vplivi nizke kakovosti podatkov. Drugi vidik obravnava dimenzijo kakovosti podatkov in merljive metrike, ki jih uporabljajo strokovnjaki za upravljanje informacij, da ocenijo in izboljšajo kakovost podatkov v primerjavi z dobro opredeljenim standardom. Tretji in četrti vidik se poglobljata v upravljanje in vladavino kakovosti podatkov, raziskujeta strategije, ki so ključne za upravljanje in izboljšanje kakovosti podatkov. Peti vidik poudarja vlogo strokovnjakov za kakovost pri kakovosti podatkov, vključno s postavljanjem standardov za podatke, spremljanjem kakovosti podatkov in usposabljanjem osebja za najboljše prakse upravljanja podatkov. Zadnji vidik vsebuje vpogled v prihodnost in obravnava prihajajoče trende v kakovosti podatkov, vključno z vplivom novih tehnologij, spremembami zakonodaje in premiki v poslovnih praksah (Yang, 2024).

Kakovost podatkov in informacij ima na področju IS dolgo in obsežno zgodovino. Zelo poznana je npr. raziskava DeLone in McLean (2003), ki opisuje model uspeha IS, ki je v zadnjih dveh desetletjih prejel največ pozornosti in je pritegnil številne sledilce. Večina raziskovalcev IS uporablja izraza »podatki in informacije« kot sinonima, medtem ko informacijska teorija, znanost o upravljanju in znanost o odločanju jasno ločujejo med podatki in informacijami.

Podatki so dejstva, ki so zbrana in shranjena ter posredovana na način, ki prispeva k znanju prejemnika. V nasprotju s podatki se lahko informacije nanašajo na prihodnost in so zato pomembne za odločanje. Informacije zmanjšujejo negotovost odločevalca, saj mu pomagajo pri opredelitvi razpoložljivih alternativ ali napovedanih posledic izbire alternative. V skladu s tem je kakovost informacij pozitivno povezana s kakovostjo vodstvenih odločitev (Wieder in Ossimitz, 2015).

Kakovost podatkov se nanaša na kakovost predstavitve relevantnih dejstev. Razlika med podatki in informacijami, ki se nanaša na njihovo kakovost, je še posebej očitna v kontekstu BI. Glavni cilji BI je zagotoviti visokokakovostne informacije za upravljanje odločanja. To dosežemo z uporabo predvsem dvostopenjskega pristopa. Prvi del pristopa vključuje identifikacijo, zbiranje, shranjevanje in vzdrževanje podatkov. Drugi del pristopa pa vključuje pridobivanje, obdelavo in posredovanje podatkov na način, ki je uporaben za prejemnika/odločevalca, na primer z uporabo tehnologije OLAP, poročil, nadzornih plošč ali orodij za rudarjenje podatkov (Wieder in Ossimitz, 2015).

Hierarhični odnos med podatki in informacijami pomeni, da je kakovost podatkov predpogoj, vendar ne zagotovi za kakovost informacij. Pričakujemo, da visokokakovostnimi podatki, ki so učinkovito shranjeni, privedejo do boljše informacij, ali drugače povedano, kakovost podatkov je pozitivno povezana s kakovostjo informacij (Wieder in Ossimitz, 2015).

2.3 Masovni podatki

Masovni podatki so podatki, ki so tako obsežni in kompleksni ter izhajajo iz tako širokega razpona virov pri tako visoki hitrosti, da presegajo zmogljivost tradicionalnih orodij in infrastrukture za zajemanje, shranjevanje in analizo. Opredelitev značilnosti masovnih podatkov zajema tri elemente: kakovost, hitrost in raznolikost (Levin in drugi, 2015). Kombinacija teh treh lastnosti naredi delo z masovnimi podatki zahtevno, čeprav težava najpogosteje izhaja bolj iz obsega in hitrosti podatkov kot iz raznolikosti. Vendar pa je zanesljivost potrebna subjektivna lastnost vseh podatkov, ne glede na velikost ali kompleksnost, in ni edinstvena za masovne podatke.

Naslednja predlagana opredelitev masovnih podatkov se nanaša na opis velikosti v smislu števila virov, ki otežujejo uporabo poizvedovanja, in kompleksnih medsebojnih odnosov, ki otežujejo čiščenje podatkov. Masovni podatki so lahko z vidika naborov podatkov tudi majhni. Ta nekoliko protislovna izjava zajema koncept, da tako kot velikost tudi pomanjkanje strukture in kompleksnost odnosov med podatkovnimi elementi otežujeta upravljanje z masovnimi podatki (Levin in drugi, 2015).

Eno od vprašanj, ki je bilo v razpravah o masovnih podatkih splošno slabo naslovljeno, je kakovost podatkov. Dejansko obstaja stališče, da je dovolj dobro dovolj dobro, kar pomeni, da je določena stopnja izgube podatkov in netočnosti sprejemljiv kompromis za vpogled, pridobljen iz masovnih nizov podatkov. Čeprav to lahko velja za neklinične aplikacije, je sprejetje takšnega pristopa v zdravstvenem okolju lahko problematično. Izguba podatkov in netočnost nista sprejemljivi za drobne analize na ravni posameznega bolnika. Kljub temu lahko uporaba tehnik statistične analize, zasnovanih za obvladovanje večjih podatkov, še vedno omogoči uporabo masovnih podatkov z nižjo kakovostjo podatkov. Uporaba takšnih metod zahteva zgodnje in tesno sodelovanje s formalno usposobljenimi statistiki in

bioinformatiki, saj so te tehnike matematično kompleksne in presegajo raven statističnega izobraževanja večine zdravnikov (Levin in drugi, 2015).

Še ena pomembna komponenta, ki jo je potrebno upoštevati, so metapodatki, ki so opredeljeni kot podatki o podatkih. Metapodatki so pomemben del podatkov v zdravstveni oskrbi, saj zagotavljajo kontekst za podatke, generirane v predoperativni oskrbi, ki jih je mogoče izkoristiti na netradicionalne načine za razumevanje delovnega procesa. Imajo obsežno količino in hitrost, vendar nimajo pomembne raznolikosti (Levin in drugi, 2015).

Metapodatki so vseprisotni v IS in se pojavljajo v različnih oblikah. Osnovne značilnosti večine programske opreme, ki jo uporabljano vsak dan, so metapodatki. Vsa vsebina je opremljena z njimi, kot so informacije o ustvarjanju elementa, imenu, temi, značilnostih in podobno. So ključni za delovanje sistemov, ki hranijo vsebino, omogočajo uporabnikom iskanje zanimivih elementov, beleženje bistvenih informacij o njih in deljenje teh informacij z drugimi (Riley, 2017).

Podjetja rutinsko shranjujejo metapodatke o iskanjih in transakcijah, kar jim omogoča analizo prodajnih trendov, napovedovanje prihodnjega povpraševanja, plačevanje davkov na prodajo in še več. Metapodatki omogočajo podjetjem bolj personalizirane nakupovalne izkušnje s funkcijami, kot so zgodovina nakupov, priporočila izdelkov itn. Tudi v turistični dejavnosti se zanašajo na metapodatke o potnikih, strankah, rezervacijah itn. Mediji uporabljajo metapodatke za sledenje dogodkom, pokritost in objavljeno vsebino. Vsa podjetja uporabljajo metapodatke za funkcije človeških virov, kot so zaposlovanje, obračun plač itn. (Riley, 2017).

Obstaja več vrst metapodatkov. Največja razlika je med opisnimi metapodatki in ostalimi vrstami. Upravni metapodatki so okvirni izraz, ki se nanaša na informacije, potrebne za upravljanje vira, ali se nanaša na njihovo ustvarjanje. V okviru področja upravnih metapodatkov so tehnični metapodatki predstavljeni kot informacije o digitalnih datotekah, potrebnih za dekodiranje in prikazovanje različnih atributov, kot je vrsta datoteke. Opisni in upravni metapodatki so drugačni od strukturiranih metapodatkov, ki opisujejo razmerja med deli virov. Zadnja kategorija metapodatkov so označevalni jeziki. Ti jeziki združujejo metapodatke in vsebino, praksa, ki se uporablja le včasih pri drugih oblikah metapodatkov. Pri besedilnih virih to pomeni označevanje strukturiranih elementov, kot so odstavki, označevanje besed s semantičnimi informacijami. Te različne kategorije metapodatkov podpirajo različne uporabne primere v IS. Odkrivanje je morda najpogostejše, pri čemer strukturirani metapodatki omogočajo uporabnikom iskanje ali brskanje po virih ali informacijah, ki jih zanimajo (Riley, 2017).

Metapodatke lahko najdemo v različnih oblikah in kodiranjih. V tradicionalnem oblikovanju IS se lahko shranijo kot polja v tabelah relacijskih podatkovnih baz. Večino vrst metapodatkov lahko shranimo na ta način. V tem scenariju bi se metapodatki običajno nalagali v serijah preko prilagojenih procesov ali pa bi jih ročno vnašali preko uporabniških

vmesnikov, ki so zasnovani za določen namen, vsak nadzorovan s prilagojenim programiranjem (Riley, 2017).

Težave se pojavijo pri masovnih podatkih, ki vključujejo informacijsko onesnaženost, navidezne korelacije in merilne napake. Akumulacija informacijske onesnaženosti se nanaša na povečanje količine manjkajočih in navideznih podatkov, ki postanejo prisotni, ko postaneta velikost in dimenzionalnost (število spremenljivk) nabora podatkov zelo visoki. To lahko zmanjša razmerja signala ter oteži prepoznavanje pravih in pozitivnih rezultatov. Visoka dimenzionalnost nabora masovnih podatkov lahko prav tako privede do navideznih korelacij, kjer se nepovezane naključne spremenljivke zdijo visoko in vzročno povezane, vendar v resnici niso (Levin in drugi, 2015).

Metode strojnega učenja in rudarjenja podatkov se pogosto uporabljajo za velike nize podatkov, da bi pomagale premagati težave, ki se pojavijo pri masovnih podatkih. Oba izraza se pogosto mešata in se na mnoge načine prekrivata. Strojno učenje je osredotočeno na napovedovanje novih podatkov na podlagi znanih lastnosti iz obstoječih podatkov, medtem ko se rudarjenje podatkov ukvarja z odkrivanjem prej neznanih lastnosti in povezav med podatki. Vsekakor ni dvoma, da lahko rudarjenje podatkov zagotovi dragocene vpoglede v niz masovnih podatkov in lahko koristi pri generiranju različnih hipotez. Rudarjenje podatkov lahko zajema povzemanje, odkrivanje odstopanj, modeliranje odvisnosti, klasifikacijo, gručenje in prilagajanje regresije. Te tehnike so koristne pri reševanju težave z nizkim razmerjem informacijske onesnaženosti, omenjene zgoraj (Levin in drugi, 2015).

Prava omejitev tradicionalnih relacijskih podatkovnih baz in analitičnih orodij postane očitna, ko postane obseg podatkov, ki jih je treba analizirati, zelo velik in visoko dinamičen. To se pogosto zgodi pri podatkih, kjer je dnevni obseg podatkov milijarde ali celo trilijone podatkovnih točk. Temeljni pristop k obvladovanju takšnih naborov podatkov je, da jih najprej razdelimo na manjše dele, analiziramo vsak del in nato sestavimo rezultate. Novost v zadnjem desetletju je uporaba tega pristopa s pomočjo tisočih ali celo milijonov računalnikov (Levin in drugi, 2015).

2.3.1 Metode obdelave masovnih podatkov

Posamezniki in podjetja se zelo pogosto osredotočajo na to, kako hitro pridobiti dragocene informacije iz velikih količin podatkov, zato uporabljajo učinkovite metode obdelave masovnih podatkov. Zgoščevanje (angl. hashing) je metoda obdelave masovnih podatkov, ki se uporablja pri veliki strukturi podatkovne baze, ko pridobivanje bloka prek iskanja po indeksu ni izvedljivo, saj indeksno iskanje izvaja celotno iskanje na disku, da bi našlo želene podatke. Zgoščevalna funkcija je preslikovalna funkcija, ki sprejeme vrednost kot vhod in pretvori to vrednost v ključ (k). Vrednost k označuje, kje so podatki nameščeni. Ena od prednosti zgoščevanja je hitro branje podatkov. Vendar pa zgoščevanje ni primerno, ko so podatki organizirani v določenem vrstnem redu (Yaqoob in drugi, 2016).

Indeksiranje je metoda obdelave masovnih podatkov, namenjena za hitro lociranje podatkov iz obsežnih količin kompleksnih podatkovnih nizov. Ročno raziskovanje takšnih zapisov ni izvedljivo in le pristopi k indeksiranju z visokim pretokom lahko izpolnijo zahteve po zmogljivosti shranjevanja masovnih podatkov (Yaqoob in drugi, 2016).

Paralelno računanje je metoda obdelave masovnih podatkov, ki pomaga izkoristiti več virov hkrati za dokončanje naloge. Za masovne podatke zagotavlja infrastrukturo za paralelno računanje na distributiven način. Izboljšuje procesno moč s skupno rabo iste datoteke med več strežniki. Vsebuje več procesnih elementov znotraj ene same naprave. Prednost te metode obdelave je hitro procesiranje, razdelitev kompleksnih nalog in manjša poraba energije pa sta njeni pomanjkljivosti (Yaqoob in drugi, 2016).

2.3.2 Tehnike analize masovnih podatkov

Za učinkovito analizo masovnih podatkov v omejenem časovnem obdobju so potrebne izjemne tehnike analize masovnih podatkov. Uporabiti je mogoče različne tehnike, ki pripomorejo k pomembnim odločitvam. V nadaljevanju jih bom predstavila. Poleg tega je v tabeli 1 predstavljena primerjava tehnik analize masovnih podatkov (Yaqoob in drugi, 2016).

Tehnike rudarjenja podatkov se uporabljajo za povzemanje podatkov v smiselne informacije. Med tehnikami so analiza skupin, učenje asociiranih pravil, klasifikacija in regresija. Rudarjenje podatkov uporablja metode strojnega učenja in statistike za pridobitev informacij. Nove tehnike rudarjenja masovnih podatkov so potrebne, ker se stopnja pridobivanja podatkov hitro povečuje (Yaqoob in drugi, 2016).

Podatkovno rudarjenje je tehnika, ki se uporablja za odkrivanje vzorcev v podatkih. Rudarjenje spleta razkriva neznano znanje o spletni strani in uporabnikih za izvajanje analize podatkov. To je tehnika, ki pomaga oceniti učinkovitost določene spletne strani (Mughal, 2018).

Strojno učenje omogoča računalnikom, da razvijejo vedenje na podlagi empiričnih podatkov. Obstoječe tehnike strojnega učenja, tako nadzorovane kot nenadzorovane, so potrebne za prilagajanje velikim količinam podatkov. Poleg tega se umetna nevronska mreža uporablja pri prepoznavanju vzorcev, prilagodljivem nadzoru, analizi in na drugih področjih. Temelji na statističnih ocenah in teoriji nadzora (Mahdavinejad in drugi, 2018).

Vizualizacijske metode se uporabljajo za ustvarjanje tabel in diagramov, s čimer olajšajo razumevanje podatkov. Vizualizacija masovnih podatkov je bolj zahtevna kot tradicionalna vizualizacija majhnih podatkov zaradi kompleksnosti komponent. Pri vizualizaciji masovnih podatkov raziskovalci najpogosteje uporabljajo programske rešitve v seriji, da dosežejo najvišjo ločljivost podatkov na vzporedni način (Yaqoob in drugi, 2016).

Tabela 1: Primerjava različnih tehnik analize masovnih podatkov

Tehnike analize masovnih podatkov	Opis	Uporaba na določenih področjih	Razpoložljiva orodja
Podatkovno rudarjenje	Iskanje vzorcev in/ali sistematičnih povezav med spremenljivkami	Biomedicina Zdravstvo	– Excel – Rapid-I – Rapidminer-R – KNMINE
Spletno rudarjenje	Odkrivanje vzorcev uporabe iz velikih spletnih skladišč	E-učenje Digitalne knjižnice	– KXEN – LION Solver – Dataiku
Strojno učenje	Razvijanje vedenja na podlagi empiričnih podatkov	Zdravstvo Storitve za stranke	– Weka – Scikit-Learn – PyMc – Shogun
Pristopi vizualizacije	Predstavitev podatkov z uporabo vizualnih elementov	Bančništvo Proizvodnja in javne službe	– Data wrapper – Highcharts JS – MAPBox
Metode optimizacije	Reševanje kvantitativnih problemov	Znanost o družbenih omrežjih Računalniška biologija	– Matlab
Analiza družbenih omrežij	Pogled na socialne odnose z vidika teorije omrežij	Antropologija Družbena omrežja	– Cytoscape – Gephi – Cuttlefish

Vir: prirejeno po Yaqoob in drugi (2016).

Metode optimizacije se uporabljajo za reševanje merljivih problemov in so razširjene v multidisciplinarnih področjih. Za reševanje globalnih optimizacijskih problemov se uporabljajo različne strategije, kot so simulirano ohlajanje, kvantno ohlajanje, optimizacija roja in genetski algoritmi. Te strategije je treba prilagoditi različnim okoljem ali zahtevam, kar omogoča njihovo učinkovito uporabo v različnih razsežnostih in težavnostih problemov. Skaliranje strategij je lahko pomembno pri soočanju z večjimi problemi ali pri izboljšanju njihove učinkovitosti v različnih situacijah.

Tehnika analize družbenih omrežij se uporablja za opazovanje socialnih odnosov v teoriji družbenih omrežij. Velik pomen je pridobila na področju računalništva. Obravnava visoko dimenzionalne podatke, ki so težavni za učinkovito analizo (Yaqoob in drugi, 2016).

2.3.3 Izzivi masovnih podatkov

Masovni podatki ponujajo številne privlačne priložnosti. Kljub temu se raziskovalci in strokovnjaki soočajo s številnimi izzivi pri raziskovanju nizov masovnih podatkov in pri pridobivanju vrednosti in znanja iz njih. Te težave se pojavljajo na različnih ravneh, vključno z zajemanjem podatkov, shranjevanjem, iskanjem, deljenjem, analizo, upravljanjem in vizualizacijo. Dejansko se velikost masovnih podatkov eksponentno povečuje, vendar je trenutna tehnološka zmogljivost za obdelavo in raziskovanje nizov masovnih podatkov še

vedno na relativno nižjih ravneh petabajtov, eksabajtov in zetabajtov podatkov (Oussous in drugi, 2018).

Raziskovalci se pri delu z masovnimi podatki srečujejo z mnogimi izzivi. Eden od izzivov je, kako izbrati, integrirati in shraniti izjemno velike nize podatkov, ki nastajajo iz različnih virov, ob manjših zahtevah po strojni in programski opremi. Drugi izziv je upravljanje masovnih podatkov, ki je ključno za olajšanje pridobivanja zanesljivih vpogledov in optimizacijo stroškov. Dejansko dobro upravljanje podatkov predstavlja temelj za analizo masovnih podatkov. To pomeni, da je treba zagotoviti učinkovito shranjevanje masovnih podatkov in dostop na podlagi vlog do več distribuiranih končnih točk. Cilj upravljanja masovnih podatkov je zagotoviti zanesljive podatke, ki so enostavno dostopni, obvladljivi, pravilno shranjeni in varovani (Oussous in drugi, 2018).

Tretji izziv pri masovnih podatkih je, kako obvladati kompleksnost njihove narave (hitrost, obseg in raznolikost) in jo obdelati v distribuiranem okolju z mešanico aplikacij. Za zanesljive rezultate analize je bistveno preveriti zanesljivost virov in kakovost podatkov, preden se angažirajo viri. Zadnji izziv je, kako očistiti takšne ogromne nize podatkov in kako odločiti, kateri podatki so zanesljivi in uporabni (Oussous in drugi, 2018).

Večinoma ni dovolj analizirati podatke, ki se generirajo znotraj organizacij. Da bi pridobili dragocene vpoglede in znanje, je pomembno iti korak dlje in združiti notranje podatke z zunanjimi viri podatkov. Zunanji podatki lahko vključujejo npr. vire tretjih oseb, informacije o nihanju na trgu, vremensko napoved in prometne pogoje. Zelo pomembna je uravnoteženost sistema, saj lahko neravnovesje zmogljivosti sistema upočasni dostop do podatkov in vpliva na učinkovitost ter razširjenost aplikacij masovnih podatkov. Z drugega zornega kota lahko opazimo različne zmogljivosti omrežnih naprav, kot so senzorji, diski, pomnilniki itn., kar lahko upočasni zmogljivost sistema (Oussous in drugi, 2018).

Še en izziv je klasifikacija neuravnoteženega nabora podatkov. Dejansko lahko aplikacije v resničnem svetu ustvarijo različne porazdelitve razredov. Klasične tehnike učenja niso prilagojene neuravnoteženim nizom podatkov. To je zato, ker je konstrukcija modela temeljila na globalnih iskalnih merah, ne da bi upoštevala število primerov. Globalna pravila so pogosto privilegirana namesto specifičnih pravil, zato se manjšinski razredi zanemarijo med konstrukcijo modela (Oussous in drugi, 2018).

Masovni podatki prinašajo velike priložnosti in so potencial za različne sektorje. Za razumevanje odnosov med značilnostmi in raziskovanjem podatkov je potrebna napredna analiza podatkov. Na primer, analiza podatkov omogoča organizaciji dragocene vpoglede in spremljanje vzorcev, ki lahko na poslovanje podjetja vplivajo pozitivno ali negativno. Kljub temu je analiza masovnih podatkov še vedno izziv, in sicer iz več razlogov: njihova kompleksna narava, potreba po razširjenosti in zmogljivosti za analizo tako ogromnih heterogenih nizov podatkov z odzivnostjo v realnem času (Oussous in drugi, 2018).

2.4 Podatkovno skladišče

DW je predmetno usmerjena, integrirana, časovno spremenljivo in trajno DM, ki pomaga pri procesu odločanja. Načeloma je DW zasnovano za poizvedovanje in analiziranje strukturiranih podatkov, ki so razdeljeni na dejstva in dimenzije. DW je ogromna zbirka podatkov, ki shranjuje zgodovinske in trenutne podatke poslovnega sveta ter tako ohranja vse podatke na enem mestu. Strukturirani podatki se enostavno shranijo v DW, medtem ko nestrukturirani podatki predstavljajo težavo pri takem shranjevanju. Potreba po upravljanju nestrukturiranih podatkov izhaja iz dejstva, da je velika količina informacij na internetu nestrukturiranih (Gupta in Rathore, 2013).

Klasični viri nestrukturiranih podatkov so zelo obsežni. Pojavlja se veliko težav v povezavi s pridobivanjem pravih informacij iz njih, pretvarjanjem, analiziranjem za odkrivanje vzorcev in trendov, shranjevanjem informacij za hiter in učinkovit dostop, upravljanjem delovnega procesa ter nazadnje z izdelavo uporabnih poročil BI. Za preiskovanje dejstev je potrebna analiza podatkov več formatov iz različnih virov v različnih časovnih okvirih (Eberendu, 2016).

Integrirana arhitektura informacij omogoča boljše vpoglede v večdimenzionalne informacije za ciljne entitete. Prav tako zagotavlja boljše vpoglede, močnejše statistične, semantične in korelacijske sposobnosti. Hitre sposobnosti branja in pisanja omogočajo zbiranje podatkov s sposobnostmi analize v skoraj realnem času. Pri tem je zahteva, da je DW sposobno za obdelavo velikih nizov podatkov, ki so izziv za shranjevanje, analiziranje, iskanje, vizualizacijo, deljenje in upravljanje (Pasha, 2016).

Postopek zajemanja BI informacij iz nestrukturiranih podatkov poteka v dveh fazah. V prvi fazi se dodaja struktura nestrukturiranim podatkom preko ekstrakcije imenovanih entitet. Nato se v drugi fazi rezultati integrirajo s strukturiranimi podatki. Celotno DW postane osnovni okvir za aplikacije BI. Prednosti BI pri uporabi metodologije analize besedila pospešijo identifikacijo napak in njihovo posledično popravilo ali zamenjavo z integracijo nestrukturiranih podatkov, ki načeloma vnašajo strukturirane podatke, shranjene v relacijski bazi podatkov.

Označevanje besedila in uporaba zaznamkov imata učinkovito vlogo pri povezovanju strukturiranih in nestrukturiranih podatkov. To se lahko uporabi kot okvir za učinkovite in natančne poslovne odločitve. Tako nastalo skupno DW postane osnovni okvir za aplikacije BI (Gupta in Rathore, 2013).

Upravljanje BI je odgovorno za načrtovanje, izvajanje in delovanje tako imenovane faze podatkov kot tudi faze informacij v BI procesu. Načrtovanje se načeloma začne z analizo zahtev, ki temeljijo na postavljanju pravih vprašanj. Zahteve se nato pretvorijo v potrebe po informacijah, ki zahtevajo podatke. Implementacijski postopek sledi nasprotni smeri, vendar oba procesa zahtevata večfunkcionalno upravljanje, da se zagotovi tesno sodelovanje med uporabniki informacij, tako imenovanimi odločevalci in IT, ter upoštevanje standardov za

podatke in informacije. V skladu s tem je kakovost upravljanja BI pozitivno povezana s kakovostjo podatkov in informacij.

Za pravilno upravljanje, širjenje in koristnost BI je pomembna kakovost podatkov in/ali kakovost informacij. Poleg tega je upravljanje podatkov za pravilnost, doslednost, celovitost, preglednost in posledično zaupanje v podatke pomemben predpogoj za doseganje visokih ravni kakovosti informacij, vendar so za odličnost na tem področju potrebna ustrezna orodja za njihovo relevantnost in aktualnost. Uvedba obsežnih rešitev za BI lahko prinese koristi, vendar v ospredju ni količina, temveč sama kakovost podatkov in še posebej informacij (Wieder in Ossimitz, 2015).

Prednosti upravljanja nestrukturiranih podatkov vključujejo poslovno vrednost, pravočasne in relevantne informacije, ki lahko dajejo obsežne vpogled. Zelo koristno jih je imeti, a zahtevajo ustrezen način zajemanja, shranjevanja in upravljanja (Foropon in Ranjan, 2021).

3 VIZUALIZACIJA PODATKOV

Vizualizacija je proces predstavljanja podatkov kot vizualne slike. Odločevalcem omogoča, da uporabljajo svoje naravne prostorske/vizualne sposobnosti za določanje, kje naj se opravi nadaljnje raziskovanje. To pomeni, da lahko vizualizacija, kadar je uporabljena ustrezno, odločevalcu omogoči, da v podatkih hitreje najde informacije (Dudycs, 2010).

Vizualizacija podatkov omogoča, da človeški vizualni sistem pridobi informacije iz podatkov, in zagotavlja pregled kompleksnega nabora podatkov. Pomaga pri prepoznavanju izjem v podatkih, spodbuja uporabnike k dejanjem in prepoznavanju strukture, vzorcev, trendov, anomalij in odnosov v podatkih (Tominski in Schumann, 2020).

Tehnologije vizualizacije so bile uporabljene na mnogih poslovnih področjih, saj grafična predstavitev podatkov olajša analizo in pomaga uporabniku pri sprejemanju informiranih in hitrih odločitev. Vizualizacija omogoča enostavno prepoznavanje nenavadnosti v podatkih, izolacijo le-teh in preiskovanje njihovega izvora veliko hitreje kot tradicionalne analitične metode (Dudycs, 2010).

Z vizualizacijo lahko predstavimo različne rešitve, med katere spadajo: nadzorne plošče za spremljanje uspešnosti, poročila, grafikoni, kazalniki za prikaz operativnih podatkov in podatkov uspešnosti, opozorila, sprožilci in indikatorji, ki se lahko dostavljajo preko številnih kanalov, kot so namizja, e-pošte, brezžične povezave, diagrami vpliva, zemljevidi strategije, drevesni zemljevidi, hiperbolični 3D in tematski zemljevidi. Osnovne lastnosti, potrebne za dobre aplikacije vizualizacije informacij, so: interaktivnost (uporaba neposredne manipulacije uporabniškega vmesnika za izvajanje operacij kot je filtriranje), učinkovitost (uporaba vizualizacije za prikaz v realnem času) in fokalizacija (uporaba vizualizacije osredotočenja in konteksta kot sta popačenje in dinamično povečevanje) (Dudycs, 2010).

3.1 Pomen vizualizacije podatkov za podjetja

Vizualizacija podatkov pomaga izboljšati razumevanje podatkov in zmanjšuje njihovo kompleksnost, s čimer izboljšuje način, kako ljudje dojemajo podatke, ne glede na to ali gre za javnost, ki spremlja napovedi na borzi, ali za vodjo marketinga, ki opazuje trend prodaje izdelkov podjetja v zadnjem desetletju in odkrije podrejeni trend. Na ta način vizualizacija podatkov pomaga ljudem najti neko vrsto skritega trenda, ki ga je mogoče uporabiti za napovedovanje in statistično, prilagojeno potrebam ljudi.

Vizualizacija podatkov je zelo pomembna pri odločanju. Organizacije generirajo ogromne količine podatkov in odločevalcem postaja vedno težje raziskovati in razumeti te podatke, vizualizacija pa prinaša rešitev za to težavo (Shao in drugi, 2022).

Z ustvarjanjem računalniško generiranih grafičnih prikazov podatkov lahko odločevalci analizirajo podatke v vizualni obliki, kar jim omogoča hitro prepoznavanje vzorcev in razumevanje kompleksnih informacij. To je še posebej pomembno v znanstvenem raziskovanju, kjer je sposobnost vizualizacije podatkov ključna. Z uporabo vizualizacije podatkov lahko odločevalci lažje razumejo in oblikujejo informirana mnenja, kar vodi v boljše odločanje (Sadiku in drugi, 2016).

Vizualizacija podatkov ni le preoblikovanje števil v vizualno obliko, ampak tudi zmožnost napovedovanja nekaterih osnovnih trendov, ki jih lahko vidimo le grafično, saj izboljšuje kognicijo in povezuje podatke s časom, kar je zelo težko videti v številkah. To je še posebej vidno v podjetjih pri sprejemanju določenih odločitev ob upoštevanju različnih kritičnih dejavnikov, ko lahko majhni in zapleteni detajli vplivajo na to, kako podjetje oblikuje svoje poti delovanja in politike v prihodnosti. Zato imajo podatki, ki so hrbtenica trenutnih podjetij, sposobnost spodbujanja pozitivne rasti in preprečevanja krize (Park in drugi, 2021).

3.2 Pomen zbiranja in priprave podatkov za podjetja

V zadnjih letih je bilo ustvarjenih veliko podatkov na različnih področjih, tudi v zavarovalnicah. Ti podatki ne izhajajo samo iz profesionalnih zdravstvenih sistemov, ampak tudi iz nosljivih naprav. To veliko količino podatkov bi lahko združili in uporabljali za optimizacijo zdravljenja vsakega posameznega pacienta. Da bi bilo to mogoče, je nujno, da bolnišnice, akademski sektor in dejavnost sodelujejo. Vendar bolnišnice in akademski sektor pogosto ne želijo deliti svojih podatkov z drugimi stranmi, čeprav je pacient dejanski lastnik svojih zdravstvenih podatkov. Bolnišnice vlagajo veliko časa v postavljanje kliničnih preizkusov in zbiranje podatkov ter želijo biti prve, ki objavijo članke na podlagi teh podatkov (Hulsen, 2020).

Obstajajo nekateri javno dostopni nabori podatkov, vendar so običajno na voljo šele po zaključku raziskave, kar pomeni ogromno zamudo v mesecih ali celo letih, preden lahko

podatke analizirajo tudi drugi. Zato eno izmed rešitev predstavlja spodbuda bolnišnicam, da delijo svoje podatke z različnimi akademskimi sektorji (Hulsen, 2020).

Pri razpravi o deljenju podatkov je pomembno vedeti, da ni soglasja o tem, kdo je dejanski lastnik teh podatkov. Prav ta del na kratko obravnava vprašanje lastništva podatkov v luči zakonov o zasebnosti. Institucije pogosto verjamejo, da so lastnice podatkov pacienta, saj so jih zbirale. Vendar so te institucije dejansko le varuhi podatkov, ker so podatki last pacienta in dostop ter uporaba teh podatkov izven klinične ustanove običajno zahtevata soglasje pacienta. To omejuje izkoriščanje velike količine podatkov v kliničnih zapisih, saj bi morali biti podatki po koncu študije uničeni (ali zadostno anonimizirani) (Hulsen, 2020).

Tehnike masovnih podatkov, kot sta npr. strojno in globoko učenje, uporabljajo od tisoče do milijone podatkovnih točk, kar lahko zahteva obsežno obdelavo. Škoda bi bilo izgubiti tako dragocene podatke ob koncu projekta, zato je priporočljivo, da bolnišnice in klinične ustanove prosijo pacienta za soglasje za shranjevanje in uporabo njegovih podatkov za prihodnje znanstvene raziskave. Čeprav na ta način ni mogoče uporabiti podatkov iz velikega števila retrospektivnih nizov, bo to zagotovilo, da se lahko vsaj prospektivno zbrani podatki uporabljajo v prihodnjih raziskavah. Ideja, da bi morali biti podatki čim bolj deljeni, da bi omogočili znanstveni napredek, pridobiva na veljavi predvsem zaradi moči analiz masovnih podatkov, strojnega učenja, globokega učenja ipd. (Hulsen, 2020).

Sistem zvezne baze podatkov (angl. data federation) predstavlja razvoj v medicinski znanosti in možno rešitev za dilemo med deljenjem podatkov in zasebnostjo pacienta. Pri uporabi metod strojnega učenja na zdravstvenih podatkih so potrebne velike vzročne velikosti, ki jih pogosto dosežejo le s kombiniranjem podatkov iz več raziskav. Toda to vrsto združevanja informacij otežujejo zahteve po zasebnosti pacienta in varstvu podatkov. Splošna ideja za porazdeljeno učenje je, da ustanove delijo statistični model in njegove parametre namesto občutljivih podatkov (Hulsen, 2020).

Deljenje podatkov lahko bi spremenilo celo zdravstveno dejavnost. Razpravlja o partnerstvu, ki ga je Apple objavil s trinajstimi večjimi zdravstvenimi sistemi, vključno z Univerzo Johnsa Hopkinsa in Univerzo v Pensilvaniji, mu bo omogočilo prenos elektronskih zdravstvenih podatkov pacientov na njegove naprave (po pridobljenem soglasju pacienta) (Hulsen, 2020).

Ta vrsta deljenja podatkov bi lahko preoblikovala zdravstveno dejavnost, saj bi omogočila nove načine in izboljšala oskrbo. Lahko bi celo zmanjšala organizacijske stroške s poenostavitvijo procesov oskrbe, saj bi zaposleni v bolnišnicah potrebovali manj časa za zagotavljanje dostopnosti podatkov pacientom. Umetna inteligenca (angl. artificial intelligence, v nadaljevanju AI) bi lahko uporabljala podatke pacientov za odgovarjanje na njihova vprašanja in jih usmerjala k zdravstvenim storitvam, ki jih potrebujejo (Hulsen, 2020).

Institucije, kot so zdravstvene zavarovalnice, ki uporabljajo velike količine zdravstvenih podatkov pacientov, se zanašajo na možne tehnične rešitve, kot je uporaba sistema zvezne baze podatkov. Prav ta s pomočjo algoritma omogoča dostop do podatkov brez potrebe po njihovem združevanju. Koristi predstavljajo znatno zmanjšanje organizacijskih stroškov, poenostavitev procesov oskrbe in učinkovito analizo ter obdelavo podatkov.

3.3 Vizualizacija podatkov s poslovno inteligenčno nadzorno ploščo

Organizacije se pogosto srečujejo s težavami pri zbiranju podatkov, pridobivanju informacij in spremljanju odločitev na podlagi izvlečenih informacij. Proces odločanja vključuje oblikovanje presoj glede na več naložb in virov na podlagi kvantitativnih in kvalitativnih podatkov. Obstaja več različnih sistemov, kot so DW, ERP in drugi, ki jih organizacije trenutno uporabljajo za procese odločanja.

Odločitve v katerikoli organizaciji spremljajo ljudje in ne sistemi, zato ima predstavitev podatkov pomembno vlogo v procesu odločanja, pri katerem različni oddelki v organizaciji pogosto uporabljajo nadzorne plošče za analizo podatkov (Janssen in drugi, 2020).

Nadzorna plošča podjetju omogoča koncentrirano vizualizacijo uspešnosti. Če je ustrezno zasnovana, je enostavna za uporabo, saj grafično predstavlja katerikoli preprost tip za razumevanje kvalitativnih ali kvantitativnih podatkov. Na globalnem trgu temelji uspeh vsakega oddelka na njegovi sposobnosti hitrega sprejemanja pametnih odločitev v zvezi s kompleksnimi problemi. Vsak zaposlen, odgovoren za sprejemanje odločitev, se mora spopasti z veliko količino podatkov, zato nadzorne plošče olajšajo razumevanje velikih količin podatkov. V primeru poslovne dejavnosti, ki nima nadzorne plošče, mora vodja, če želi primerjati podatke in sprejeti odločitev na njihovi podlagi, skozi dolg postopek pridobivanja relevantnih podatkov za primerjavo. Pogosto se podatki predstavljajo v različnih formatih, kar ustvarja težave pri njihovi združitvi (Hansonti, 2010).

Prednost nadzorne plošče je, da pomaga izkoristiti prihodke in optimizirati poslovne procese. V časih gospodarske krize je optimizacija poslovnih procesov za zmanjšanje stroškov in časa izjemno pomembna. Nadzorna plošča je pravzaprav uporabniški vmesnik, običajno sestavljen iz enega zaslona, ki ima statično strukturo in omogoča dostop do informacij ob pravem času s pomočjo kazalnikov.

Ker se nadzorna plošča lahko uporablja za različne funkcije s strani različnih ljudi, mora biti visoko prilagodljiva za različne uporabnike in različne scenarije. Prikazovala naj bi le ustrezne informacije, ki jih odločevalci potrebujejo pri svojem delu. Pomembno vlogo pri odločanju igra tudi mehanizem prikaza. Način, kako so informacije predstavljene, je zelo pomemben. Izbor barv ima ključno vlogo pri oblikovanju nadzorne plošče. Nadzorne plošče, uporabljene za vizualno rudarjenje, uporabljajo barve, kot so zelena, rumena in rdeča, da prikažejo dobro, zadovoljivo in slabo uspešnost (Hansonti, 2010).

Glavne prednosti uporabe nadzorne plošče so: predstavljanje informacij o uspešnosti v grafični obliki, kar omogoča uporabnikom hitro prepoznavanje težav pri uspešnosti; analiza informacij o uspešnosti na strukturiran način v okviru ciljev in nalog; pomoč pri boljšem odločanju in zmanjšanju stroškov, ki nastanejo pri ročnem zbiranju informacij, ter pomoč pri zmanjševanju stroškov, povezanih z administrativnim ali ročnim delom (Hansonti, 2010).

Kakovost odločitev je močno odvisna od kakovosti podatkov, sem pa sodi tudi dostopnost podatkov. Podatki na nadzorni plošči običajno prihajajo iz več virov, vključno z DW, preglednicami, internimi aplikacijami, zunanjimi storitvami in samostojnimi aplikacijami. ERP sistemi so prav tako eno od orodij, ki naj bi izboljšalo sposobnost odločanja na izvršilni ravni. Organizacije so se za analizo in poročanje zanašale na finančne sisteme, kot je ERP sistem. Čeprav so bili močni pri obdelavi transakcij, niso bili učinkoviti za izvajanje poizvedb in generiranje poročil. Ti sistemi niso bili zasnovani za oblikovanje poročil, specifičnih za uporabnika, vendar pa so takšna poročila potrebna za proces odločanja v katerikoli organizaciji (Hansonti, 2010). V primerjavi z drugimi sistemi BI so nadzorne plošče poenostavile postopek analize podatkov za analitike, saj zagotavljajo podatke preko enotnega uporabniškega vmesnika. (Temitope, 2021).

Različni oddelki v organizacijah uporabljajo nadzorne plošče na različne načine, prilagojene specifičnim potrebam. Način, kako oddelki uporabljajo nadzorno ploščo, lahko naredi razliko med neuspehom in uspehom. V prodaji in trženju običajno uporabljajo nadzorno ploščo za napovedovanje prodaje, zbiranje informacij o potencialnih poslih, naročilnih enotah in cenah (Malik, 2005).

V oskrbovalni verigi po navadi uporabljajo nadzorne plošče za upravljanje distribucije, zalog, logistike in spremljanje stopnje vračil (Malik, 2005). Oddelki za IT po navadi uporabljajo nadzorno ploščo za upravljanje virov, načrtovanje in izračun stroškov in časa, porabljenega v različnih projektih (Hansonti, 2010).

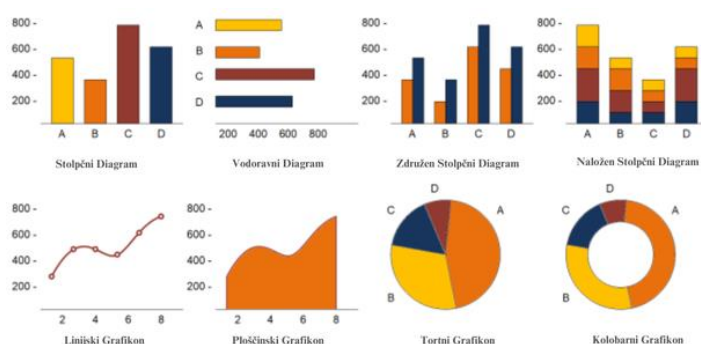
3.4 Komponente poslovno inteligenčne nadzorne plošče

V sistemih BI običajno podatkovne baze delujejo kot vir podatkov, medtem ko DW in DM služita kot ciljni podatkovni shrambi. Tradicionalno so vgrajeni z uporabo relacijskih ali kombiniranih sistemov za upravljanje baz podatkov. Dejansko se DW zanašajo na podatkovne zbirke, saj večina IS poslovne podatke pridobiva iz splošnega DW podjetja. Prav tako je značilno, da DW in DM optimizirata potrebo DW z uporabo podatkovnih struktur, kot sta shema zvezde in shema snežinke (Orlovskyi in Kopp, 2020).

Pri oblikovanju vmesnika je treba upoštevati nekatere funkcije, ki jih mora nadzorna plošča omogočati. Glavno je, da nadzorna plošča omogoča filtriranje, razvrščanje in analiziranje podatkov. Imeti mora funkcionalnost »povleci in spusti« (angl. drag-and-drop). Omogočati mora pregledne grafe, ustvarjanje alternativnih scenarijev in je lahko večjezična (Hansonti, 2010).

Pri izboru komponent oziroma grafov nadzorne plošče je pomemben namen grafikonom, ki uporabnikom dejansko olajša prepoznavanje vzorcev in primerjanje vrednosti med seboj. Ker ljudje vrednotijo vrednosti natančno na dvodimenzionalni ravnini, so najboljše vizualizacije stolpčni grafiki, linijski grafiki in krožni grafiki. Zmožnost analize podatkov upada, ko grafiki uporabljajo druge vizualne značilnosti, kot so velikost, barva, radialna razdalja itn. Krožni grafiki se štejejo za manj učinkovite kot stolpčni in linijski grafiki zaradi potrebe po primerjavi radialnih razdalj in površin. Nekaj primerov najučinkovitejših vizualizacijskih grafikonomov v različnih oblikah (npr. vodoravni in navpični stolpčni grafiki, kolobarni in tortni grafiki itn.) je prikazanih na sliki 1 (Orlovskyi in Kopp, 2020).

Slika 1: Primeri vizualizacijskih grafikonomov



Vir. prirejeno po Orlovskyi in Kopp (2020).

Zgoraj prikazane primere grafikonomov podjetja uporabljajo najpogosteje. Stolpčni grafikon uporabljajo za prikaz primerjave ob različnih časih, mestih in pogojih. Linijski grafikon za prikaz trendov v času glede na količine, velikost in stopnje. Krožne grafikone pa uporabljajo za prikaz delov, ki sestavljajo celoto, da lahko primerjajo relativne velikosti teh delov.

4 PREDSTAVITEV IZBRANEGA PODJETJA

V nadaljevanju so predstavljene osnovne dejavnosti izbrane zavarovalnice s poudarkom na zobozdravstvenih storitvah. Vključena je podrobna analiza zdravstvenih storitev in učinkovitost preventivnih pregledov.

4.1 Predstavitev osnovne dejavnosti

Zdravje prebivalstva v Sloveniji se je v zadnjih nekaj desetletjih izboljšalo. Pričakovana življenjska doba, tako za ženske kot za moške, je podobna povprečju EU. Zadovoljstvo z zdravstveno oskrbo je visoko, vendar so se podaljšali časi čakanja na nekatere specializirane ambulantne storitve, zato je več poudarka na preventivnih ukrepih in boljši koordinaciji oskrbe, zlasti za tiste s kroničnimi stanji. Relativno visoke stopnje doplačil za mnoge storitve, zajete s sistemom univerzalnega obveznega zdravstvenega zavarovanja, izravnava

prostovoljno zdravstveno zavarovanje. Slovenija je nekoliko edinstvena med državami s socialnim zdravstvenim zavarovanjem, saj se skoraj izključno zanaša na prispevke delodajalcev za financiranje obveznega sistema zdravstvenega zavarovanja. To naredi dohodke v zdravstvenem sektorju zelo občutljive na gospodarska in tržna nihanja (Albreht in drugi, 2016).

Izbrana zdravstvena zavarovalnica sodeluje z mrežo priznanih zdravstvenih izvajalcev. Njihov cilj je celostna zdravstvena obravnava v različnih obdobjih, neodvisno od starosti oseb in vrst bolečin. Nudijo enostavne, odzivne in zanesljive storitve, ki prinašajo številne prednosti ter izboljšajo kakovost življenja. Pri oblikovanju ponudb storitev po meri sodobnega časa in potrebah in željah zavarovancev skrbno sodeluje skupina vrhunskih strokovnjakov. Ponujajo širok spekter zdravstvenih storitev, zdravstvenih nasvetov in specialistov na določenem področju (Izbrano podjetje, 2024).

4.2 Predstavitev zobozdravstvenih storitev

Ker v magistrskem delu predstavim razvoj BI nadzorne plošče za vizualizacijo podatkov o rednih zobozdravstvenih pregledih zavarovancev v izbrani zavarovalnici, sledi kratka predstavitev zobozdravstvene panoge in zobozdravstvenih storitev.

Tradicionalno so zobozdravniki spodbujali tako paciente z nizkim tveganjem, kot tudi paciente z visokim tveganjem za razvoj zobnih bolezni, da obiskujejo svoje zobozdravstvene ambulate za redne šestmesečne preventivne preglede. Kljub temu pa je pacientom in zobozdravnikom na voljo malo dokazov o najboljšem časovnem intervalu med zobnimi pregledi za ohranjanje ustnega zdravja (Clarkson in drugi, 2020).

Za namen boljšega ocenjevanja stopnje napredovanja zobnih bolezni pri posameznikih obstajajo različni kazalniki, vendar pa se je pokazalo, da je uporaba klinične presoje zdravnika za prepoznavanje ljudi, ki so ogroženi zaradi zobne gnilobe, enako dobra ali celo boljša, kot katerikoli drugi način izbire ali preizkušanja. Tako se zdi, da lahko presoja ustreznih intervalov za zdravljenje zobne gnilobe temelji le na individualni osnovi. Splošna priporočila bi pripeljala do tega, da bi nekateri bolni zobje ostali neopaženi, mnogi zdravi zobje pa bi bili nepotrebno ali pogosto preiskovani, pri čemer bi se porabili viri brez koristi (Kay, 1999).

Za paciente je pomembno, da redno obiskujejo svoje zobozdravnike, da bi zmanjšali verjetnost nujnih zobozdravstvenih storitev. Odločitve o optimalnem intervalu za ponovni pregled pa lahko, tako kot pri večini kliničnih odločitev, sprejmemo le, ko imamo na voljo ustrezne informacije. Edina oseba, ki bo imela te informacije, je zobozdravnik. Če želimo služiti najboljšim interesom pacientov, je bistveno, da zdravstveni in zobozdravstveni delavci uskladijo in prilagodijo nasvete, ki jih dajejo pacientom. Novi način primarne zdravstvene oskrbe ponuja mehanizem, s katerim lahko zdravnik in zobozdravnik podpirata

to holistično stališče na način, ki bo koristil obema poklicema in seveda pacientom (Kay, 1999).

Pravočasna in ustrezna zobozdravstvena oskrba je ključna za preprečevanje in natančno zdravljenje ustnih bolezni. Ključno je, da so vodstveni kadri, zdravstveni strokovnjaki in ponudniki zdravstvenih storitev v celoti ozaveščeni o dejavnikih, ki spodbujajo uporabo zobozdravstvenih storitev. Uporaba zdravstvenih storitev izhaja iz interakcije med posameznikovimi in kontekstualnimi dejavniki, vključno z dostopom do zdravstvenih storitev in organizacijo zdravstvenega sistema. Vsekakor imajo vsi ljudje pravico do prejemanja zdravstvene oskrbe ne glede na svojo sposobnost plačila, zato je pravična distribucija zdravstvenih storitev resna odgovornost zdravstvenih politik in administratorjev (Zardak in drugi, 2023).

Izbrana zdravstvena zavarovalnica v okviru svoje ponudbe vključuje tudi zobozdravstvene storitve. Poudarek je na preventivnih pregledih. Pogosti razlogi za obisk zobozdravnika so karies, vnetje in krvavenje dlesni, zdravljenje zob in živcev ter ustvarjanje implantatov, protez in mostičkov. Zobozdravniki svetujejo pacientom, da jih redno obiskujejo, da lahko ti spremljajo razvoj čeljusti, izraščanje zob in jih naučijo ustne higiene. Tudi če ni prisotne nobene bolečine, priporočajo kontrolni pregled vsaj enkrat letno za preverbo zdravja ustne votline in oskrbo manjših težav, še preden se razvijejo v kaj bolj resnega (Interno gradivo zavarovalnice, 2024).

5 IZDELAVA POSLOVNO INTELIGENČNE NADZORNE PLOŠČE ZA VIZUALIZACIJO PODATKOV V IZBRANI ZAVAROVALNICI

Nadzorne plošče so v zdravstvu vse bolj popularne, saj so vse bolj napredne in zagotavljajo ključne upravljalne informacije o učinkovitosti poslovanja. Vendar pa so upravljalne nadzorne plošče več kot le jasen in neposreden prikaz uspešnosti v kontekstu zdravstva in opravljanja zdravstvenih storitev – gre za interaktiven in iterativen proces, ki vključuje vse zainteresirane stranke (van Elten in drugi, 2022).

Nadzorne plošče zagotavljajo vizualni pregled informacij, potrebnih za upravljanje uspešnosti, s čimer olajšujejo odločanje in upravljanje projektov. Tako je v zdravstvenem sektorju uporaba nadzornih plošč raznolika. Pomembne so predvsem za pregled ključnih kazalnikov uspešnosti (angl. key performance indicator, v nadaljevanju KPI) in omogočanje upravljanja uspešnosti v zdravstvenih ustanovah, zato se zelo pogosto uporabljajo kot upravljalni instrumenti za pridobivanje vpogledov v napredek in uspešnost (van Elten in drugi, 2022).

Izdelava nadzorne plošče ni enostaven, ampak dinamičen proces, ki vključuje različne faze. Vsak sektor zdravja potrebuje interaktivna orodja za upravljanje uspešnosti, kot so upravljalne nadzorne plošče, da bi lahko uspešneje merili in spremljali poslovanje ter tako izboljšali učinkovitost. Zdravstvene organizacije se vse bolj srečujejo z izzivom

zagotavljanja visokokakovostnih storitev po dostopnih stroških. Učinkovito upravljanje in izboljšanje uspešnosti takšnih sistemov zahteva identifikacijo in optimizacijo več spremenljivk. Statični sistemi poročanja o uspešnosti niso sposobni popolnoma zadovoljiti potreb zdravstvenih managerjev po podpori odločanju, zato so nadzorne plošče orodje, ki omogoča razvoj interaktivnih elementov za prenos, analizo in prikaz podatkov o uspešnosti v realnem času in primerjavo s prejšnjim obdobjem (Ghazisaeidi in drugi, 2015).

Upravljalne nadzorne plošče predstavljajo sistem dostave informacij na različnih ravneh, tako da na enem zaslonu prikažejo najpomembnejše informacije o doseganju strateških ciljev, kar omogoča managerjem, da učinkovito merijo in spremljajo poslovanje ter izboljšajo uspešnost. Če so podprte z ustrezno IT in dobro organizirane na osnovi načel merjenja uspešnosti, jim nadzorne plošče omogočajo, da se osredotočajo na bolj pomembne dejavnosti, identificirajo problematična področja, ki potrebujejo korektivne ukrepe, analizirajo osnovne vzroke slabe uspešnosti, napovedujejo trende in vzpostavljajo referenčne vrednosti (Ghazisaeidi in drugi, 2015).

Glavne značilnosti nadzornih plošč vključujejo obvestila v realnem času, opozorila, možnost poglobljenega pregleda ter analizo scenarijev. Poleg tega imajo dodatne funkcije poročanja, kot so razširjanje in krčenje skupin, interaktivno razvrščanje podatkov, označevanje strani in parametrizacijo, ter elemente interaktivnosti, ki omogočajo uporabniku manipulacijo s prikazom poročila med izvajanjem (Ghazisaeidi in drugi, 2015).

Funkcionalne značilnosti nadzornih plošč se morajo ujemati z njihovim namenom, sicer lahko pride do napačnih odločitev. Vse nadzorne plošče imajo skupni cilj – hitro predstaviti pravilne informacije za pravilno odločanje. Nadzorne plošče vključujejo širok nabor namenov, kot so dosledno merjenje, spremljanje, načrtovanje in večdimenzionalna analiza. Poleg tega se glede na naravo problema, ki ga je treba rešiti, ti nameni razlikujejo od statičnega poročanja do odkrivanja znanja, zato je natančna razmejitev med nameni ključni določevalec funkcionalnih značilnosti nadzornih plošč (Ghazisaeidi in drugi, 2015).

V celoti razvoj visokokakovostnih nadzornih plošč v kompleksnem kontekstu zdravstvenega sektorja z njegovim podatkovno intenzivnim in tehnološko usmerjenim okoljem zahteva reševanje več vprašanj glede razvoja KPI, virov podatkov, kakovosti podatkov, integracije nadzorne plošče s sistemom vira podatkov ter predstavitve podatkov uporabnikom. Nadzorne plošče omogočajo ustvarjanje dinamičnih poročil, ki bodo pomagala dosledno meriti uspešnost, nenehno spremljati KPI za zaznavanje odstopanj, globoko analizirati vzroke nesprejemljive uspešnosti in učinkovito načrtovati za prihodnost (Ghazisaeidi in drugi, 2015).

Osredotočajo se na pobude, ki jih uvajajo najvplivnejše zavarovalnice. Na splošno velja, da imajo zavarovalnice prostor in sposobnost za inovacije, v mnogih primerih pa uporabljajo tehnološke aplikacije in platforme za kritje novih in obstoječih tveganj.

Digitalizacija je postala pomemben sprožilec inovacij. Vpliv digitalizacije na vrednostno verigo zavarovalniškega sektorja je viden predvsem pri interakciji s strankami, prilagajanju njihovemu vedenju, avtomatizaciji poslovnih procesov in odločitvah, izboljšavah obstoječih izdelkov ter ponudbi novih izdelkov (Lanfranchi in Grassi, 2022).

Nov koncept, imenovan InsurTech, je fenomen, ki temelji na tehnoloških inovacijah v zavarovalniški dejavnosti, ter pridobiva zanimanje znotraj zavarovalniškega sektorja, ki ga poganjata povečano zadovoljstvo strank in učinkovitost. Učinki koncepta InsurTecha se čutijo v različnih vrstah zavarovanj. Na primer, zdravstvene zavarovalnice se morajo spoprijeti z vzponom novih medicinskih tehnologij in nosljivih naprav, ki jih je mogoče uporabiti za zbiranje koristnih, vendar občutljivih podatkov o pacientih, ter pretvoriti tveganje za fizično zdravje, ki ga prej ni bilo mogoče zavarovati, v tveganje, ki ga je mogoče zavarovati (Lanfranchi in Grassi, 2022).

InsurTech lahko pomaga izboljšati obstoječe izdelke, storitve in procese ter omogočiti nove poslovne modele. Na primer, napredna tehnologija podpira modele zavarovanja, ki segajo od cen na podlagi vedenja, široko preučevanih v avtomobilski dejavnosti zavarovanja, do personalizacije, povezane s podatki, pridobljenimi iz nosljivih naprav. Modeli zavarovanja med vrstniki so še en primer, kjer si ljudje lahko delno delijo tveganja. Ti lahko prispevajo k obnovi zaupanja v zavarovalniško dejavnost z zmanjšanjem konfliktov interesov, saj ti običajno ne vključujejo entitet, ki bi imele koristi od zavrnjenih zahtevkov (Lanfranchi in Grassi, 2022).

5.1 Metodologija

Obstajajo različne tehnike, ki jih uporablja BI, med katerimi so prevladujoče: opisna analiza, diagnostična analiza, podatkovno rudarjenje itd. Za potrebe magistrskega dela sem uporabila postopek ETL, posamezne faze pa so predstavljene v nadaljevanju.

5.1.1 Pridobivanje podatkov

Treba je pregledati že dostopne podatke, oceniti njihovo ustreznost za zastavljeno nalogo in preučiti izvedljivost pridobivanja novih podatkov, zbranih posebej za želeno nalogo. Prav tako je pomembno oceniti, ali je na voljo dovolj podatkov za realno doseganje zelenih rezultatov (Abdallah in drugi, 2016).

Natančnost, kakovost in veljavnost so kritične lastnosti podatkov ter postanejo še posebej pomembne, ko se v okviru BI več podatkovnih virov integrira in združuje. Da bi bili podatki natančni, morajo biti vrednosti podatkov pravilne in v pravilni obliki, sicer bi lahko bile analize napačne. Omejitve in postopki preverjanja podatkov pomagajo zagotoviti, da so čisti in pravilni podatki dodani v nize podatkov. Nekatera orodja za pripravo podatkov omogočajo premikanje podatkov skozi integriran delovni tok, ki izboljšuje kakovost v nizu korakov.

Nekatera orodja ponujajo analitikom možnost, da pregledajo izvirne podatke, medtem ko prihajajo skozi delovni tok, jih preverijo in določijo, kako jih želijo preoblikovati glede na svoje zahteve (Stodder, 2016).

Natančnost je agregirana vrednost nad kriteriji kakovosti: integriteto, doslednostjo in gostoto. Vsak kriterij natančnosti je opisan z različnih vidikov. Integriteta predstavlja celotno zbirko podatkov, ki vsebuje predstavitev vseh entitet, zahteva tako popolnost kot veljavnost podatkov. Popolnost podatkov pomeni zagotovitev celovite predstavitve podatkov, ki ne vsebujejo manjkajočih vrednosti. Popolnost v celoti dosežemo pri čiščenju podatkov z odpravljanjem anomalij. Prav tako je mogoče generirati dodatne podatke, ki predstavljajo obstoječe entitete, a trenutno niso predstavljene v podatkih. Veljavnost pomeni, da so podatki veljavni in ni kršena nobena omejitev. Obstaja veliko mehanizmov za povečanje veljavnosti, vključno z obveznimi polji, izvajanjem edinstvenih vrednosti ter podatkovno shemo. Doslednost se nanaša na sintaktične anomalije ali protislovja. Glavni izziv je izbira, kateremu viru podatkov zaupamo za zanesljiv dogovor med podatki iz različnih virov. Usklajenost sheme velja za relacijske sisteme baz podatkov, kjer se skladnost oblik domene zanaša na uporabnika. Enotnost je neposredno povezana z nepravilnostmi. Gostota se nanaša na delež manjkajočih vrednosti v podatkih (Abdallah in drugi, 2016). Pravočasnost je še en kriterij, ki je pomemben pri kakovosti podatkov. Nanaša se na aktualnost podatkov, ki zahteva ažurnost (Abdallah in drugi, 2016).

Za potrebe izdelave nadzorne plošče sem pridobila podatke s strani izbrane zavarovalnice, natančneje iz oddelka za razvoj produktov. Za bolj natančno analizo in bolj jasen prikaz podatkov sem pridobila podatke za več zaporednih let. Pridobivanje podatkov izbrane zavarovalnice je podrobneje predstavljeno v poglavju 5.3.1.

5.1.2 Izbira orodja

Hitro napredovanje tehnologij v zadnjih letih je naredilo podatke za bistveno in strateško sredstvo, ki določa uspeh ali neuspeh organizacije. Dostop do pravih informacij v realnem času in z visoko selektivnostjo lahko predstavlja konkurenčno prednost v poslovnem okolju. Sistemi BI pomagajo izvršnim direktorjem, poslovnim vodjem in drugim operativnim delavcem pri spremljanju boljših in bolj informiranih poslovnih odločitev (Gonçalves in drugi, 2023).

Nadzorne plošče omogočajo sintetiziran, agregiran in interaktiven način vizualizacije informacij, ki so ključne za odločevalce v poslovnem svetu, da jih proaktivno opozorijo na obstoječe težave. Na strateški ravni je s pomočjo sistemov BI mogoče določiti glavne cilje in strateške KPI organizacije in jih tudi spremljati. Spremljanje ključnih dejavnikov je za podjetja pomembno z vidika izboljšanja poslovnih rezultatov. Takšna naloga se začne z opredelitvijo ključnih dejavnosti, ki jih predstavljajo KPI. Načeloma ključni KPI pomagajo podjetjem meriti, kako dosegajo svoje glavne cilje in strategije. Odkrivanje, kaj ne gre dobro

v kritičnih poslovnih operacijah, je koristno tudi za izboljšanje aktivnosti, ki povzročajo izgube (Gonçalves in drugi, 2023).

Poslovni analitiki in odločevalci potrebujejo dostop do informacij v razumljivi in dinamični obliki. Pogosto to vključuje nadzorno ploščo s predstavitvijo vseh ključnih KPI podjetja. Nadzorna plošča je večplastna aplikacija, zgrajena na infrastrukturi BI in integraciji podatkov, ki omogoča organizacijam, da učinkoviteje merijo in spremljajo poslovanje ter tako izboljšajo poslovno uspešnost podjetja (Gonçalves in drugi, 2023). Uporaba platform za BI je trenutno ogromna prednost za mnoge vodje kot tudi podjetja. Uporaba BI sistemov je postala strateška pobuda in jo zdaj priznavajo tudi znana podjetja kot temelj za učinkovitost in inovacije podjetja. Omogočajo analizo poslovnih podatkov za podporo in izboljšanje odločanja v različnih poslovnih dejavnostih (Gonçalves in drugi, 2023).

Večina aplikacij za BI je trenutno usmerjena v storitve, zasnovane kot rešitev za obvladovanje sprememb v poslovanju. Implementacija aplikacije za BI, ki je integrirana z upravljavskimi aplikacijami, kot je ERP, skupaj z implementacijo v arhitekturo usmerjenih storitev, predstavlja eno od rešitev za premagovanje omejitev upravljavskih zmogljivosti v podjetju (Widjaja in Mauritsius, 2019). BI uporablja vizualizacijo za bolj razumljivo predstavitev podatkov končnim uporabnikom. Odločevalci brskajo po vmesnikih za podatke, analizirajo podatke in preverjajo organizacijsko uspešnost v realnem času (Widjaja in Mauritsius, 2019).

Ker je izbrana zavarovalnica pri izbiri orodja za izdelavo BI nadzorne plošče kolebala med programom Power BI in Microsoft Excel, v nadaljevanju oba na kratko predstavim.

Microsoftov program Power BI je namenjen samooskrbni BI in je platforma, ki olajša ustvarjanje vizualno privlačnih poročil na podlagi različnih podatkov. Prav tako se lahko uporablja za povezovanje s podatki iz več različnih virov, preoblikovanje in modeliranje podatkov, ustvarjanje grafikonov itn. Ena od prednosti Power BI je ustvarjanje estetsko prijetnih poročil in nadzornih plošč. Ta programska oprema je zlasti uporabna za hitre analize in izračune.

Jezik z izrazi za analizo podatkov (angl. data analysis expressions, v nadaljevanju DAX) znotraj programa Power BI omogoča programiranje poizvedb, dinamičnih agregacij in izračunov po meri. Povezana so tudi različna orodja in predloge, ki omogočajo ustvarjanje poročil in nadzornih plošč iz podatkov. Na drugi strani jezik Visual Basic za aplikacije (angl. visual basic for applications - VBA) znotraj programa Microsoft Excel omogoča avtomatizacijo nalog kot tudi oblikovanje poizvedb. Na voljo je tudi več brezplačnih predlog za ustvarjanje nadzornih plošč.

Microsoft Excel in Power BI uporabljata Power Query, ki je komponenta, namenjena obdelavi podatkov. Na splošno je Excel bolj prijazen za raziskovanje podatkov in ustvarjanje modelov, medtem ko je Power BI bolj primeren za predstavitev podatkov.

Power BI omogoča integracijo v oblak, ki ponuja zmožnosti DW, kot so priprava podatkov, odkrivanje podatkov in interaktivna nadzorna plošča. Orodje Power BI ima tudi sposobnost kombiniranja različnih baz podatkov, datotek in spletnih storitev, kar omogoča hitro spremljanje ali popravljanje podatkov in samodejno reševanje težav. To je bil eden izmed ključnih razlogov za uporabo prav tega orodja. Njegovo prednost vidim v tem, da zagotavlja tudi varnost pri objavi poročil znotraj podjetja in avtomatično posodablja vir podatkov z aktualnimi podatki.

Pri delu s podatki v okolju Microsoft, sta na voljo dva različna jezika, jezik Power Query M in DAX, ki se ju lahko uporablja za filtriranje, upravljanje in vizualizacijo podatkov. DAX je analitični jezik za izračun podatkov, ki se lahko uporablja za poglobljeno analizo podatkov v fazi prikaza podatkov. Na drugi strani je Power Query M namenjen formulam za poizvedbe, ki ga lahko uporabljamo v urejevalniku poizvedb Power BI za pripravo podatkov, preden se lahko naložijo v modelu. En od drugega nista odvisna, imata popolnoma različno strukturo in logiko ter različno sintakso osnovne kode. Ne moreta se uporabljati hkrati, saj se Power Query M uporablja v urejevalniku poizvedb, medtem ko se DAX večinoma uporablja v modelu za prikaz podatkov. Na podlagi primerjave obeh orodij sem se odločila, da za izgradnjo nadzorne plošče za BI uporabim orodje Power BI.

5.1.2.1 Jezik DAX

Power BI kot del postopka modeliranja podatkov omogoča dostop do DAX ali tako imenovanega jezika za izražanje analize podatkov. Obdelava podatkov je zapleten postopek, ki vsebuje različne metodologije, orodja in logična zaporedja za pridobivanje, preoblikovanje, nalaganje in ustrezno vizualizacijo podatkov. Danes obstajajo različna orodja, povezana z BI, ki delujejo tako na ravni namizne programske opreme in tudi kot storitev v oblaku. Ena od najbolj prepoznavnih programskih rešitev za popolno preoblikovanje podatkov v kombinaciji z velikim številom možnosti za njihovo vizualizacijo je orodje Power BI. Rezultat uporabe jezika DAX je optimizirano interaktivno poročilo, ki temelji na novih podatkovnih stolpcih, izračunih in vizualnih komponentah (Naneva in Stefanova, 2022).

DAX omogoča razširitev analitične moči podatkovnega modela z ustvarjanjem izračunanih stolpcev, mer in tabel. Vsak od teh elementov je ustvarjen z uporabo DAX, ki je uporabniku prijazen jezik, podoben Excelovemu jeziku formul. Izraža razlike med izračunanimi stolpci, merami in tabelami (Pearson in drugi, 2020).

Pisanje in odpravljanje napak pri nekaterih izračunih DAX je lahko za podatkovnega analitika izziv. Pogosto se zgodi, da zahtevani kompleksni izračuni vključujejo pisanje sestavljenih ali zapletenih izrazov. Sestavljeni izrazi lahko vključujejo uporabo številnih vgnezenih funkcij in po možnosti ponovno uporabo logike izraza. Uporaba spremenljivk v formulah lahko pomaga pri pisanju kompleksnejših in učinkovitejših izračunov.

Spremenljivke lahko izboljšujejo zmogljivost, zanesljivost, berljivost in zmanjšajo zapletenost (Myers in Buck, 2022).

Skozi izdelavo magistrskega dela sem uporabila tudi DAX, in sicer v prvi fazi čiščenja podatkov, da bi prišla do čim bolj čistih in jasnih podatkov za analizo. Nato sem ga uporabila še v drugi fazi pri opredelitvi stolpca DATUM, ki prikazuje datum oziroma določeno obdobje trajanja aktivnosti. Glavna prednost uporabe DAX je bolj izrazita jasnost in preciznost podatkov, s katero bi prišla do čim bolj jasne analize, prikazane z vizualizacijo podatkov.

5.1.2.2 Jezik Power Query M

Microsoft Power Query zagotavlja zmogljivo izkušnjo pridobivanja podatkov, ki vključuje številne funkcije. Osnovna zmožnost poizvedovanja v Power Query je filtriranje in združevanje, ki pomeni združevanje podatkov iz enega ali več podatkovnih virov. Vsako tako združevanje podatkov je izraženo z uporabo formularnega jezika M (Klopfenstein in Howell, 2023).

Model vrednotenja Power Query M je oblikovan po modelu vrednotenja, ki ga pogosto vidimo v preglednicah. Vključuje skupen nabor definicij, ki so na voljo za uporabo izraza, imenovanega standardna knjižnica. Te definicije so sestavljene iz nizov poimenovanih vrednosti. Imena vrednosti, ki jih zagotavlja knjižnica, so na voljo za uporabo v izrazu, ne da bi jih izraz izrecno opredelil.

Metapodatki so informacije, povezane z vrednostjo. Predstavljeni so kot vrednost zapisa, ki se imenuje metapodatkovni zapis. Polja metapodatkovnega zapisa se lahko uporabljajo za shranjevanje metapodatkov za vrednost. Vsaka vrednost ima metapodatkovni zapis. Če vrednost metapodatkovnega zapisa ni določena, je metapodatkovni zapis prazen oziroma nima polj. Metapodatkovni zapisi zagotavljajo način za poizvedovanje dodatnih informacij s katerokoli vrsto vrednosti na nevsiljiv način. Povezava metapodatkovnega zapisa z vrednostjo ne spremeni vrednosti ali njenega obnašanja (Klopfenstein in Howell, 2023).

Pri izdelavi nadzorne plošče v okviru magistrskega dela sem uporabila jezik Power Query M. V prvi fazi sem ga uporabila za filtriranje, agregiranje in dodajanje stolpcev v tabeli. Tako sem na primeru tabele »Stomatološki pregled« ustrezno naredila poizvedbo za filtriranje podatkov po šifri in ustrezen opis te šifre oziroma dodaten opis storitve za lažje in enostavno razumevanje nadzorne plošče.

5.1.3 Urejanje podatkov

V nadaljevanju so predstavljeni koraki urejanja podatkov, potrebni za izdelavo BI nadzorne plošče za vizualizacijo podatkov, urejanje konkretnih podatkov izbrane zavarovalnice pa je predstavljeno v poglavju 5.3.2.

5.1.3.1 Čiščenje podatkov

Čiščenje podatkov je opredeljeno kot postopek zaznavanja in popravljanja pokvarjenih ali netočnih zapisov iz nabora zapisov, tabele ali baze podatkov. Čiščenje podatkov se lahko imenuje tudi usklajevanje podatkov. Bolj natančno lahko čiščenje podatkov pojasnim kot štiristopenjski proces.

Prva stopnja je določitev in identifikacija napak v podatkih, kot so nepopolnost, napačnost, netočnost ali irelevantnost. Druga stopnja je čiščenje in popravljanje teh napak z zamenjavo, spreminjanjem ali brisanjem. Tretja stopnja je dokumentacija primerov napak in vrst napak. Zadnja ali četrta stopnja je merjenje in preverjanje, ali čiščenje izpolnjuje določene tolerance uporabnika glede čistoče (Abdallah in drugi, 2016).

5.1.3.2 Integracija podatkov

Obstoječe podatke je mogoče obogatiti, kar. pogosto vključuje pridobivanje dodatnih informacij o podatkovnih točkah, glede katerih že obstajajo podatki. Uvoženi podatki morajo biti integrirani z drugimi podatki za celovit prikaz vseh enot podatkov. Integracija podatkov je ključna naloga pri pripravi podatkov. Kombiniranje podatkov iz različnih virov ni enostavno, še posebej pri obdelavi velikih količin podatkov in heterogenih virov. Podatki so pogosto predstavljeni v različnih oblikah (strukturirani, polstrukturirani ali nestrukturirani), prav tako iz različnih virov, ki se lahko shranjujejo lokalno ali razpršeno. Poleg tega lahko strukturirani podatki iz enega vira vsebujejo različne sheme (Abdallah in drugi, 2016).

5.1.3.3 Transformacija podatkov

Pogosto je treba podatke preoblikovati iz enega prikaza v drugega. Obstaja veliko razlogov za spreminjanje prikazov. Med najbolj pogostimi je generiranje simetričnih porazdelitev namesto prvotnih asimetričnih. Preoblikovanje izboljša vizualizacijo podatkov, ki so lahko testno združeni v primerjavi z nekaj izjemami. Podatki se preoblikujejo za doseganje boljše razumljivosti in za izboljšanje združljivosti podatkov z domnevami, ki temeljijo na modeliranju procesa (Abdallah in drugi, 2016).

5.1.4 Nalaganje podatkov

Nalaganje podatkov je postopek kodiranja in prenosa podatkov iz podatkovnega vira v DW ali katerikoli drug ciljni shranjevalni sistem. Nalaganje podatkov lahko vključuje tudi izvajanje preverjanja omejitev določenih shem podatkovne baze, kot so enoličnost, referenčna celovitost in obvezna polja. To ni le operacija, prisotna v tradicionalnih bazah podatkov, ampak postaja tudi osnovni korak v mnogih sodobnih sistemih za analizo podatkov. Pri nalaganju surovih podatkov ti sistemi pogosto pretvorijo podatke v določene podatkovne formate. Čeprav nekateri sistemi omogočajo obdelavo podatkov brez nalaganja,

njihova učinkovitost ni zadovoljiva. To je zato, ker lahko ti sistemi brez uporabnih metapodatkov ali optimizacije postavitve podatkov obdelujejo surove podatke samo na način surove sile (angl. brute force), kot je na primer ponovljivo preiskovanje celotnih datotek za vsako poizvedbo (Tian, 2017).

Podatkovni analitiki želijo analizirati nabor podatkov čim prej, počasno nalaganje podatkov pa tako postane ozko grlo v analizi podatkov. Zlasti pri analizi začasnih naborov podatkov, ki bodo opuščeni po izvedbi nekaj poizvedb, nalaganje podatkov ni primerno. Jasno je, da potrebujemo pristop, ki se izogiba počasnemu nalaganju podatkov in je sposoben doseči učinkovito izvajanje poizvedb (Tian, 2017).

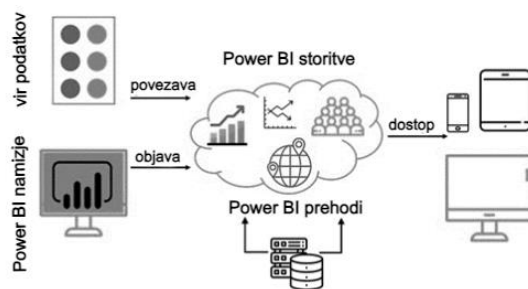
Nalaganje podatkov izbrane zavarovalnice je podrobneje predstavljeno v poglavju 5.3.3.

5.2 Poslovno inteligenčna arhitektura v orodju Power BI

Orodje Power BI ima nabor storitev, kot so aplikacije, povezave in storitve, ki pretvarjajo zbrane surove podatke v razumljive dosledne in vizualno interaktivne informacije. Glavni cilj Power BI je zagotoviti potrebe posameznih uporabnikov ali organizacij. Lahko obvladuje povezovanje obsežnih virov podatkov v različnih formatih, ki se lahko nahajajo lokalno ali v oblaku. Microsoftova orodja imajo ogromno različnih virov podatkov, ki so na voljo za povezavo s Power BI. Ti viri podatkov se razlikujejo od Excel, CSV/Text, JSON, XML datotek do relacijskih in nerelacijskih podatkovnih baz. Obvladuje tudi podatke iz Oracle, MySQL, PostgreSQL, Sybase, Teradata, SAP HANA, IBM DB2 in Azure SQL. Poleg tega se povezuje tudi s strežnikom SAP Business Warehouse, Facebook in Github (Sousa in drugi, 2021).

Ena od glavnih prednosti Power BI je njegova sposobnost obvladovanja prilagojenega vmesnika, v katerem lahko uporabniki zagotovijo več parametrov glede na svojo vlogo in dostop v organizaciji, ali celo sposobnost deljenja podatkov med zaposlenimi. Še ena velika prednost je sposobnost ustvarjanja nadzorne plošče in poročil vseh povezanih virov podatkov, s čimer se vsi pomembni podatki združujejo na enem mestu za vse uporabnike. Osnovna BI arhitektura Power BI je vidna na sliki 2 (Sousa in drugi, 2021).

Slika 2: Power BI poslovno inteligenčna arhitektura



Vir: prirejeno po Sousa in drugi (2021).

Orodje Microsoft Power BI vključuje veliko komponent, ki imajo pomembno vlogo na platformi. Te komponente je mogoče uporabiti v kombinaciji ali posamezno, zato je treba za boljše razumevanje same platforme poznati vse komponente Power BI in njihovo delovanje. Glavne komponente orodja Power BI so (Sousa in drugi, 2021):

- Namizje (angl. Power Desktop): komponenta, odgovorna za enostavno analizo s ključnim ciljem ustvarjanja poročil. Z njeno funkcijo povleci in spusti (angl. drag and drop) je enostavno postaviti vsebino natančno na želeno mesto na prilagodljivi in tekoči nadzorni plošči;
- Orodje za pretvarjanje podatkov (angl. Power Query): komponenta, odgovorna za ETL proces;
- Tehnologija za upodobitev podatkov (angl. Power View): komponenta, ki uporabnikom zagotavlja visoko interaktivno orodje za vizualizacijo z vmesnikom povleci in spusti (angl. drag and drop) za ustvarjanje in poenostavitev vizualizacij podatkov v Excelovih delovnih zvezkih;
- Orodje za tridimenzionalno (3D) ponazoritev podatkov (angl. Power Map): komponenta, ki se uporablja v Excelu, je tridimenzionalni model za vizualizacijo podatkov, ki uporabnikom pomaga razumeti informacije na nove in različne načine;
- Vprašanja in odgovori (angl. Power Q&A): komponenta, ki si prizadeva identificirati vnesene besede in iskati, kje najti odgovor;
- Power BI spletna stran: komponenta, ki omogoča uporabnikom, da svoja poročila iz programa Power BI Desktop objavijo na spletni strani Power BI. Poleg tega uporabnikom omogoča ustvarjanje različnih nadzornih plošč za njihova poročila in deljenje z drugimi uporabniki;
- mobilna aplikacija Power BI: komponenta, ki predstavlja mobilno rešitev, tako za Android kot za iOS, ima glavni cilj, da uporabnikom omogoči interaktivni pregled njihovih poročil na mobilnih napravah.

5.3 Koraki izdelave poslovno inteligenčne nadzorne plošče za vizualizacijo podatkov v izbrani zavarovalnici

V nadaljevanju so predstavljeni koraki, potrebni za izdelavo BI nadzorne plošče za vizualizacijo podatkov v izbrani zavarovalnici.

5.3.1 Proces pridobivanja podatkov

Empirični podatki so bili pridobljeni s strani izbrane zavarovalnice. Za bolj natančno analizo in bolj jasen prikaz podatkov sem pridobila podatke za več zaporednih let. Zbiranje podatkov je potekalo znotraj oddelka za razvoj produktov v izbrani zavarovalnici. Podatki vsebujejo informacije stomatološkega pregleda kot dela ponudbe izbrane zavarovalnice. Prvotno so bili shranjeni v Excel datoteki. Gre za podatke o skupnem številu zavarovancev Skupine 1 in Skupine 2, ki so ali niso opravili rednega stomatološkega pregleda. Skupina 1 so

zavarovanci, ki imajo sklenjeno zavarovanje Zobje. Skupina 2 so zavarovanci, ki imajo sklenjeno dopolnilno zavarovanje. Uporabljeni nabor podatkov je razporejen v sedem ločenih datotek in je na voljo v formatu Excel. Vsaka datoteka vsebuje en delovni zavihek.

5.3.2 Urejanje podatkov

Od izbrane zavarovalnice sem pridobila veliko podatkov zavarovancev, ki so ali niso opravili rednega zobozdravstvenega pregleda. Pridobila sem podatke o skupnem številu zavarovancev Skupine 1 in Skupine 2, ki so ali niso opravili rednega stomatološkega pregleda med letoma 2019 in 2023. Za čim bolj natančno analizo sem primerjala več zaporednih let. Zavarovance sem ločila po posameznih skupinah zaradi lažje razporeditve njihovih produktov in police zavarovanja glede na nastale škode v določenem obdobju.

5.3.2.1 Opis podatkov

V tabeli 2 je prikazan kratek opis specifikacije podatkov, ki so bili pridobljeni od izbrane zavarovalnice in uporabljeni v analizi. Predstavljen je pomen različnih atributov za lažje razumevanje pri nadaljnji analizi podatkov. V tabeli so predstavljeni vsi atributi, ki jih je potrebno opredeliti za razumevanje nadaljnje analize.

Za proces pridobivanja podatkov šifer stomatoloških storitev in šifer njihovih podskupin sem dne 8. 3. 2024 uporabila podatke z URL naslova <https://partner.zzzs.si/sifranti/>, ki sem jih prenesla v Microsoft Excel obliko in tako pridobila vse potrebne šifre, ki so vključene v podatkovno bazo. Zaradi boljšega pregleda podatkov sem povezala skupino z ustrezno podskupino. Nato sem dodala opis podskupin storitve.

Tabela 2: Opis specifikacije podatkov

Naziv atributa	Opis atributa	Tip podatka	Vrednost
OSEBA_ID	edinstvena številka zavarovanca	celo število (int)	desetmestno število
DATUM_STORITVE_OD	datum, ko je zavarovanec opravil storitev	datum (date)	od novembra leta 2019 do junija leta 2023
SIFRA_STORITVE_OZ_SKODE	edinstveno število stomatološke storitve	celo število (int)	petmestno število
OPIS_STORITVE	kratak opis stomatološke storitve	besedilo (text)	besedilo
OZZ_SIFRA	število storitve zobozdravstvene dejavnosti, stomatološki pregled za odrasle – zdravljenje in protetika	celo število (int)	trimestno število

se nadaljuje

Tabela 2: Opis specifikacije podatkov (nad.)

Naziv atributa	Opis atributa	Tip podatka	Vrednost
OZZ_DOD_SIFRA	število podskupine storitve zobozdravstvene dejavnosti, stomatološki pregled za odrasle – zdravljenje in protetika	celo število (int)	trimestno število
MinOfDATUM_ZACETKA_OSNOVNE_POGODBE	datum, ko je zavarovanec sklenil osnovno pogodbo zavarovanja	datum (date)	od novembra leta 2016 do decembra leta 2023
ZAVAROVALNA VRSTA_ID	edinstveno število za ločitev podatkov Skupine 1 ali Skupine 2	celo število (int)	šestmestno število (020111/020449)
SPOL	spol zavarovanca	besedilo (text)	ena črka (M/Z)
DATUM_ROJSTVA	datum rojstva zavarovanca	datum (date)	datum
POSTA	poštna številka zavarovanca v Sloveniji	število (number)	štirimestno število
DATUM_VELJ_KRIT_OD	datum, ko je začela veljati določena polica zavarovanja	datum (date)	datum
DATUM_VELJ_KRIT_DO	datum, do katerega je veljala določena polica zavarovanja	datum (date)	datum
POLICA	številka police zavarovanja, sklenjene s strani zavarovanca za določen produkt	število (number)	desetmestno število
PRODUKT	paket storitev, ki je del Skupine 1 ali Skupine 2	besedilo (text)	vrsta produktov po skupinah
IZV_NAZIV	naziv izvajalcev zobozdravstvenih storitev	besedilo (text)	besedilo
ST_NAROCILNICE	število naročilnic zavarovanca pri posameznih izvajalcih	število (number)	osemestno število

Vir: lastno delo.

5.3.2.2 Čiščenje podatkov

Iz nabora podatkov, potrebnih za analizo, sem identificirala določene stolpce znotraj tabele za zavarovance iz Skupine 2, ki imajo veljavnost police zavarovanja od začetka leta 2004 do leta 2023, kar se ne ujema s Skupino 1, kjer sem zajela podatke za veljavnost police zavarovanja od novembra leta 2016 do konca leta 2023, zato sem prečistila podatke za zavarovance Skupine 2 in jih ponastavila, da so veljavne police enake za obe skupini

zavarovancev, to je od novembra leta 2016 do decembra leta 2023. Prečiščeni podatki v tabeli so vidni na sliki 3.

Slika 3: Primer prečiščenih podatkov za zavarovance Skupine 2 v tabeli Zobje Zavarovalniška pogodba

OSEBA_ID	MinOfDATUM_ZACETKA_OSNOVNE_POGODBE	ZAVAROVALNA_VRSTA_ID	SPOL	DATUM_ROJSTVA	POSTA
1000004049	01/04/2019	20111	M	01/04/1959	2212
1000004055	01/07/2018	20111	Ž	11/09/1983	2252
1000004618	01/11/2019	20111	M	03/10/1980	8222
1000004701	01/12/2019	20111	Ž	20/02/1963	6250

Vir: lastno delo.

Nato sem pogledala stolpec, imenovan Posta, kjer so shranjene lokacije prebivanja zavarovancev. Po pregledu podatkov sem zaključila, da je verjetno prišlo do napačnega zapisovanja številke in da jih je nekaj zapisanih s šifro, ki ni ustrezna, niti ni enako zapisana v primerjalnih dokumentih. Za bolj učinkovito analizo sem prečistila podatke, tako da sem iz Skupine 1 izbrisala dve šifri, ki nista bili ustrezno zapisani. Nato sem iz Skupine 2 izbrisala devet šifer, ki niso bile ustrezno zapisane.

Pri tem sem pogledala ostale številčno zapisane vrednosti v stolpcu Posta in naletela na nekaj vrednosti »0«, ki sem jih nato odstranila iz vseh dokumentov, ker ne bi bile pomembne za mojo analizo v nadaljevanju. Tako sem tudi nastavila, da so številke predstavljene s štirimi ciframi, saj so poštna številke v Sloveniji predstavljene s štirimi ciframi. Podatki so prikazani na sliki 4.

Slika 4: Primer podatkov v tabeli Zobje škode IZV

OSEBA_ID	DATUM_STORITVE_OD	SIFRA_STORITVE_OZ_SKODE	OPIS_STORITVE	IPS_NAZIV	ST_NAROCILNICE	SPOL	DATUM_ROJSTVA	ZAV_VRSTA_ID
1000697613	28/06/2023	P14	Kompozitna zalitka na žilostokih	DENTALNI CENTER TATALOVIČ ZOBODRAVSTVO D.O.O.	23040138	Ž	22/01/1988	20449
1000166976	28/06/2023	N0002	Kompozitna zalitka na eni plošči	MALINGER MARTIN LUKA STOMATOLOŠKA ORDIN.	23040136	Ž	29/12/1967	20449
1000029813	03/05/2023	N0033	fuži zalitka	URBANIČ POLONICA ZOBODRAVSTVO	23035567	Ž	10/07/1973	20449
1000521109	12/05/2023	2013	Bela kompozitna zalitka na več pl. ZDRAVSTVENI ZAVOD DENTOR		23040084	Ž	28/07/1980	20449

Vir: lastno delo.

V stolpcu OZZ_SIFRA sem opazila, da obstaja veliko zapisov, ki nimajo zapisane šifre storitve zobozdravstvene dejavnosti. Tako sem vse prazne vrednosti izbrisala iz podatkovne baze. Podatki so prikazani na sliki 5.

Slika 5: Primer podatkov v tabeli Dopolnilno škode

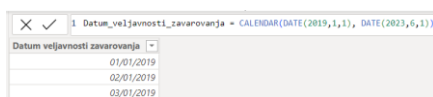
OSEBA_ID	DATUM_VELJ_KRIT_OD	SIFRA_STORITVE_OZ_SKODE	OZZ_SIFRA	OZZ_DOD_SIFRA	PRODUKT	ZAV_VRSTA_ID
1000697644	14/10/2019	1007	404	101	Paket Rumeni	20111
1777789015	23/10/2019	1007	404	101	Paket Rumeni	20111
1000229058	24/10/2019	1007	404	101	Paket Rumeni	20111
1000414025	14/10/2019	1007	404	101	Paket Rumeni	20111

Vir: lastno delo.

5.3.2.3 Dodajanje povezovalne tabele

Ker so se podatki nahajali v različnih tabelah, jih je bilo potrebno ustrezno povezati. V treh povezovalnih tabelah sem tako prikazala datume, ki se nanašajo na različno trajanje določenih aktivnosti. V prvi tabeli sem prikazala datume pogodbe, trajanje pogodbe pa sem upoštevala od 1. novembra leta 2016 do 1. junija leta 2023. Druga in tretja tabela prikazujeta različne aktivnosti, čeprav imajo pogodbe isto obdobje trajanja. Drugo tabelo sem poimenovala Datum veljavnosti zavarovanja, upoštevano pa je obdobje trajanja zavarovanja, ki je veljavno od 1. januarja leta 2019 do 1. junija leta 2023, kot je vidno na sliki 6. Tretjo tabelo, ki ima isto obdobje trajanja aktivnosti kot datum veljavnosti zavarovanja, sem poimenovala Datum storitve in se nanaša na obdobje opravljene storitve.

Slika 6: Datum veljavnosti zavarovanja

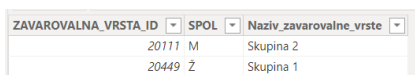


Datum veljavnosti zavarovanja
01/01/2019
02/01/2019
03/01/2019

Vir: lastno delo.

Nato sem ustvarila tabelo, imenovano Zavarovalna vrsta prikazano na sliki 7, v kateri sem upoštevala dva atributa, spol in zavarovalno vrsto ID. Za boljše razumevanje nadzorne plošče sem dodala atribut opis zavarovalne vrste, pri katerem bo vidno, za katero zavarovalno vrsto je ustvarjena nadzorna plošča. Če bi jo uporabil uporabnik, ki ne pozna šifre storitve, bi lahko brez težav razumel vizualizacijo in podatke na nadzorni plošči.

Slika 7: Tabela Zavarovalna vrsta

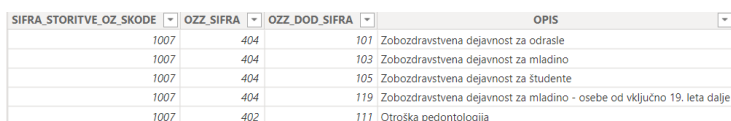


ZAVAROVALNA_VRSTA_ID	SPOL	Naziv_zavarovalne_vrste
20111	M	Skupina 2
20449	Ž	Skupina 1

Vir: lastno delo.

V tabeli Stomatološki pregled prikazano na sliki 8 sem prikazala šifro storitve oziroma škode, dodatne šifre storitve in opis storitev. Tabelo sem ustvarila zaradi lažjega povezovanja med zavarovanci Skupine 1 in zavarovanci Skupine 2. Kljub enaki storitvi sem želela prikazati razlike med dodatnimi storitvami, ki so vključene v različne dodatne stomatološke skupine. Analizirala sem štiri skupine stomatoloških storitev in sedem dodatnih podskupin storitev, ki so vključene znotraj glavnih skupin.

Slika 8: Tabela Stomatološki pregled



SIFRA_STORITVE_OZ_SKODE	OZZ_SIFRA	OZZ_DOD_SIFRA	OPIS
1007	404	101	Zobozdravstvena dejavnost za odrasle
1007	404	102	Zobozdravstvena dejavnost za mladino
1007	404	105	Zobozdravstvena dejavnost za študente
1007	404	119	Zobozdravstvena dejavnost za mladino - osebe od vključno 19. leta dalje
1007	402	111	Otroška pedontologija

Vir: lastno delo.

V tabeli Produkt razvidno na sliki 9 sem prikazala štiri različne pakete storitev kot produkt vsake zavarovalne skupine. S pomočjo tabele sem želela prikazati zavarovance Skupine 1, ki imajo v svojem naboru zavarovalne storitve štiri pakete storitve, Paket 1, Paket 2, Paket 3 in Paket 4. Zavarovanci Skupine 2 imajo v svojem naboru zavarovalnih storitev le en paket storitev, paket Rumeni.

Slika 9: Tabela Produkt

PRODUKT	ZAVAROVALNA_VRSTA_ID
Paket 1	20449
Paket 2	20449
Paket 3	20449
Paket 4	20449
Paket Rumeni	20111

Vir: lastno delo.

Zadnja dodana povezovalna tabela je prikaz izvajalcev stomatoloških storitev za obe skupini skupaj prikazano na sliki 10. S pomočjo te tabele sem želela prikazati 960 izvajalcev stomatoloških storitev, s katerimi ima zavarovalnica sklenjeno pogodbo o sodelovanju. Tabelo sem poimenovala Izvajalci storitev.

Slika 10: Tabela Izvajalci storitev

IZS_NAZIV
ZASEBNA ZOBOZDRAVSTVENA AMBULANTA - BARBARA PIRHER, dr. dent. med
VRANKIČ ZOBOZDRAVSTVO IN HIPNOTERAPIJA D.O.O.
STRLIČ IRIS - ZOBOZDRAVSTVENA ORDINACIJA
DENT MEDICO TERAN d.o.o.
JANADENTAL D.O.O.
ZDRAVSTVENI DOM ILIRSKA BISTRICA
KATAVIČ MIRNA - DOKTOR STOMATOLOGIJE

Vir: lastno delo.

5.3.3 Nalaganje podatkov

Nalaganje podatkov je postopek pakiranja podatkov in premikanja v določeno DW. Na začetku te prehodne faze je treba dobro poznati podatke, ki so sestavni del vsake datoteke. Nalaganje zahteva določanje atributov, stolpcev in vrstic ter opredelitev agregacijske funkcije.

Za uvoz podatkov v Power BI sem podatke iz vsake Excelove datoteke formirala v tabele, nato sem pri uvozu podatkov zaradi lažjega pregleda vsako tabelo ustrezno poimenovala in uvozila podatke v Power BI.

V t. i. navigatorju sem pregledala stolpce, ne pa vseh vrednosti, ki jih vsebujejo. Na ta način sem za vsako tabelo in vsak stolpec posebej pregledala vrsto poimenovanja stolpcev,

vrednosti podatkov in manjkajoče vrednosti. Po tem koraku sem naložila podatke, primer pa je viden na sliki 11.

Slika 11: Primer nalaganja podatkov

Navigator

Display Options

- Dopolnilno_T_skode.xl...
- Skupina 2_T_skode
- Dopolnilno_T_sko...

Dopolnilno_T_skode

OSEBA_ID	DATUM_VELJ_KRIT_OD	SIFRA_STORITVE_OZ_SKODE	OZZ_SIFRA	OZZ_DOD_SIFRA	PRODUKT
1000697644	14/10/2019	1007	404	101	Paket Rur
1777789015	23/10/2019	1007	404	101	Paket Rur
1000229058	24/10/2019	1007	404	101	Paket Rur
1000414025	14/10/2019	1007	404	101	Paket Rur
1000018834	21/10/2019	1007	404	101	Paket Rur
1000472455	10/10/2019	1007	404	101	Paket Rur
1000327494	11/10/2019	1007	404	101	Paket Rur
1000678085	17/10/2019	1007	404	101	Paket Rur
1000236550	24/10/2019	1007	404	101	Paket Rur
1000730055	24/10/2019	1007	404	101	Paket Rur
1000580647	23/10/2019	1007	404	101	Paket Rur
1000657291	10/10/2019	1007	404	101	Paket Rur
1000451054	15/10/2019	1007	404	101	Paket Rur
1000432059	09/10/2019	1007	404	101	Paket Rur
1000446476	17/10/2019	1007	404	101	Paket Rur
1000232474	03/10/2019	1007	404	101	Paket Rur
1000522060	01/10/2019	1007	404	101	Paket Rur
1000111165	15/10/2019	1007	404	101	Paket Rur
1000705621	01/10/2019	1007	404	101	Paket Rur
1000066840	01/10/2019	1007	404	101	Paket Rur
1000078800	08/10/2019	1007	404	101	Paket Rur
1000731505	07/10/2019	1007	404	101	Paket Rur
1000186729	16/10/2019	1007	404	101	Paket Rur

Load Transform Data Cancel

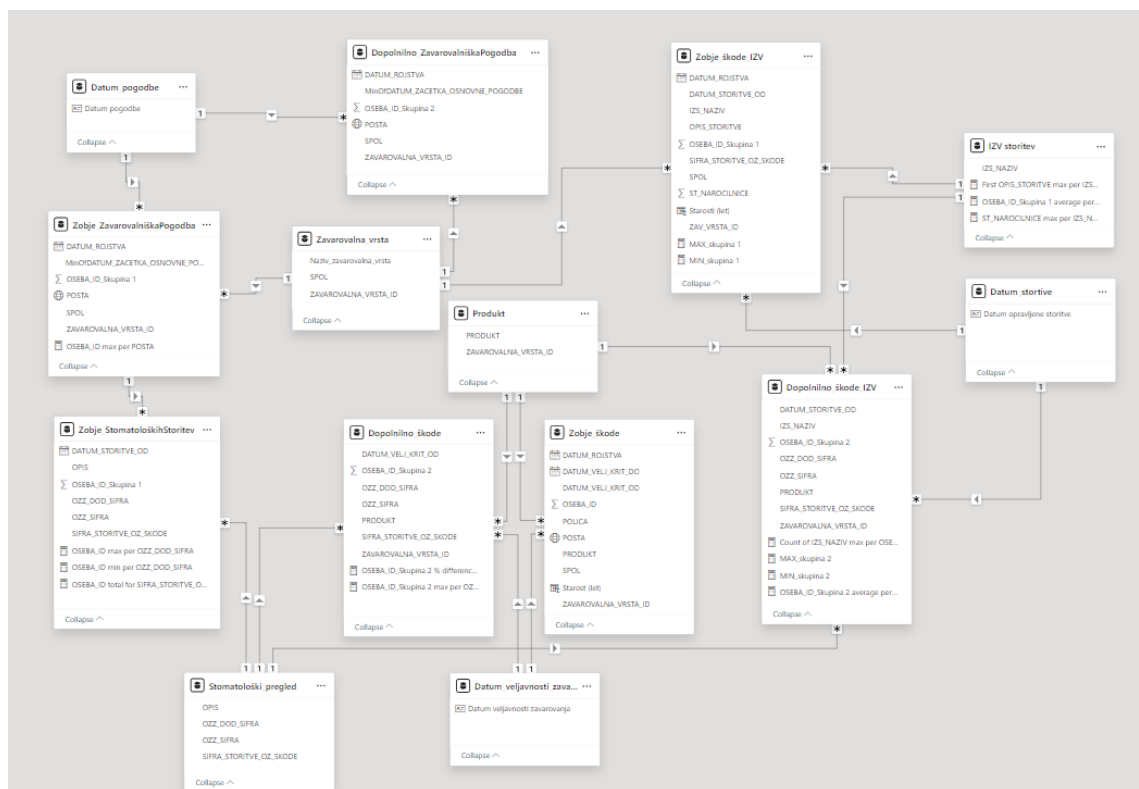
Vir: lastno delo.

5.4 Podatkovni model

V današnjem podatkovno usmerjenem poslovnem svetu je vizualizacija podatkov ključna, saj se veliko uporablja za pomoč pri odločanju, kar lahko vpliva na prihodke številnih podjetij. Kljub temu pa zaradi visokega povpraševanja po obdelavi podatkov glede na obseg, hitrost in zanesljivost podatkov obstaja naraščajoča potreba po načrtovalcih podatkovne baze, ki bi pomagali pri učinkoviti vizualizaciji podatkov. Zato je pomembno izboljšanje podatkovnega modela za uporabniško izkušnjo pri gledanju in interakcijah z vizualizacijo podatkov v realnem času, neodvisno od velikosti zapisov (Qin in drugi, 2020).

Za potrebe poslovne analitike je običajno smotno imeti t.i. zvezdno shemo, zaradi narave podatkov in povezav med entitetami pa v okviru izdelave svojega magistrskega dela transformacija podatkovnega modela v obliko zvezdne sheme ni bila izvedljiva. Na sliki 12 sem prikazala razvit podatkovni model z vsemi tabelami vključujočih atributov in povezav med njimi. Ker nisem želela, da pride do težave pri vizualizaciji na nadzorni plošči, sem ustvarila tri datumske tabele, da datumi določene aktivnosti, na katerih ni bila izvedena nobena aktivnost za določene storitve, ne bi bili prikazani. Tabele datumov sem razdelila, da bi v končni fazi vizualizacije prikazala čim bolj jasne in razumljive podatke za uporabnika.

Slika 12: Podatkovni model



Vir: lastno delo.

Podatkovni model je sestavljen iz tabel dejstev in iz dimenzijskih tabel, razloženih v nadaljevanju. Tabela dejstev (angl. fact table) je glavna tabela v dimenzijskem modelu, kjer se shranjujejo numerične vrednosti. Pogosto gre za shranjene vrednosti o merjenju, ki izhajajo iz poslovnega procesa. Vrstica v tabeli dejstev ustreza merjenju. Vsa merjenja v tabeli morajo biti na isti granulaciji. Dejstvo je lahko določena numerična vrednost. Teoretično je mogoče, da je merjeno dejstvo besedilno, vendar se to zelo redko pojavi. V večini primerov je besedilno merjenje opis nečesa in izhaja iz diskretne liste vrednosti. Analitik bi moral v tem primeru besedilna merila vključiti v dimenzije, saj jih je mogoče boljše povezati z drugimi besedilnimi dimenzijskimi atributi in zavzamejo veliko manj prostora. V tabelah dejstev se ne shranjujejo redundantne besedilne informacije. Če je besedilo edinstveno za vsako vrstico v tabeli dejstev, sodi v dimenzijsko tabelo. Pravo besedilno dejstvo je v DW redko, saj nepredvidljiva vsebina besedilnega dejstva skorajda onemogoča analizo (Kimball in Ross, 2013).

V tabeli dejstev so shranjeni kvantitativni podatki, ki jih je mogoče združevati in analizirati na različne načine. Pogosto vsebujejo informacije o določenih dogodkih, storitvah, transakcijah ali meritvah in predstavljajo osrednji del podatkovnega modela.

V razvitem podatkovnem modelu se tabele dejstev nanašajo na obe skupini zavarovancev. Skupina 1 so zavarovanci, katerih podatki so vključene v podatkovni model v tabelah z

oznako Zobje. To je manjša skupina zavarovancev, ki imajo v naboru storitve vključene štiri vrste paketov v svoji zavarovalni polici, vezani na redne zobozdravstvene preglede. Skupina 2 so zavarovanci, katerih podatki so vključeni v podatkovni model v tabelah z oznako Dopolnilno. To je večja skupina zavarovancev, ki imajo v svoji zavarovalni polici vključen le en paket storitve, vezan na redne zobozdravstvene preglede.

V tabeli dimenzij (angl. dimension table) so prisotne dodatne podrobnosti za numerične podatke, shranjene v tabeli dejstev. Dimenzijske tabele pogosto vsebujejo razmeroma majhno število vrstic, vendar so široke z večjim številom velikih stolpcev. Tabele dejstev pa lahko vsebujejo zelo veliko število vrstic in se sčasoma še povečujejo. Dimenzijska tabela je povezana s tabelo dejstev preko razmerja. S pomočjo obeh tabel lahko uporabniki analizirajo podatke v tabeli dejstev s filtriranjem, grupiranjem ali agregiranjem podatkov na podlagi atributov v tabeli dimenzij.

Dimenzijske tabele lahko vsebujejo tudi besedilne opise. Dimenzijski atributi služijo kot primarni vir poizvedbenih omejitev, združevanj in oznak poročil. V poizvedbi ali zahtevi za poročilo se atributi posledično prepoznajo. Atributi dimenzijske tabele igrajo ključno vlogo v DW. Ker so vir omejitev, so ključni za uporabnost in razumljivost DW. Moč DW je neposredno sorazmerna s kakovostjo in globino dimenzijskih atributov. DW je boljše le v primeru, da je za zagotavljanje atributov s podrobnimi poslovnimi izrazi porabljenega več časa. Več časa kot je porabljenega za polnjenje in zagotavljanje kakovosti vrednosti v stolpcu atributov, bolj učinkovito je DW (Kimball in Ross, 2013).

Dimenzijske tabele so vhodne točke za tabelo dejstev. Robustni dimenzijski atributi zagotavljajo robustne analitične sposobnosti rezanja in urejanja. Dimenzije implementirajo uporabniški vmesnik do DW (Kimball in Ross, 2013).

V razvitem podatkovnem modelu je prisotnih sedem dimenzijskih tabel. Tabela Datum pogodbe, s katero se povežeta dve tabeli dejstev: na eni strani Zobje_ZavarovalniškaPogodba oz. Skupina 1 in na drugi strani Dopolnilno_ZavarovalniškaPogodba oz. Skupina 2. Njihova skupna dimenzijska tabela je tudi tabela dejstev, imenovana Zavarovalna_vrsta, v kateri so prisotni atributi, vezani na obe tabeli dejstev. Tabela Stomatološki_pregled je dimenzijska tabela, ki omogoča prikaz podatkov: tabela Zobje_StomatološkihStoritev podatke, vezane na zavarovance Skupine 1, in tabela Dopolnilno_škode, vezana na zavarovance Skupine 2. Tabela Produkt je dimenzijska tabela, ki omogoča povezavo podatkov o produktih in se nanaša na tabeli dejstev: na tabelo Dopolnilno_škode in na tabelo Zobje_škode. Prisotna je še njihova dodatna dimenzijska tabela Datum_veljavnosti_zavarovanja. Tabela Datum_storitve je dimenzijska tabela, povezana s tabelama: Dopolnilno_škode_IZV in Zobje_škode_IZV. Hkrati je prisotna še njihova dodatna dimenzijska tabela Izvajalci_storitev za dodatno pojasnitev in naziv izvajalcev storitev. Na sliki 13 je vidno, da tabela Dopolnilno_ZavarovalniškaPogodba predstavlja tabelo dimenzij, v kateri so prisotni atributi, vezani le na to skupino zavarovancev. Pri tem vsaka vrstica predstavlja en dogodek.

Slika 13: Primer tabele dimenzij



Vir: lastno delo.

5.4.1 Razmerje med tabelami

Po naloženih tabelah in urejeni podatkovni bazi sem ustvarila relacije med tabelami. Vsaka tabela ima dodeljen glavni ali tuji ključ. V tem procesu je pomembno opredeliti skupni stolpec, ki povezuje več tabel. Pravzaprav so glavni in tuji ključi stolpci v tabeli, ki ustvarjajo razmerje med tabelami.

Na primeru podatkovnega modela glavni ključ predstavlja vrstico v tabeli Datum_pogodbe, ki enolično identificira vrstico v tabeli Zobje_ZavarovalniškaPogodba in v tabeli Dopolnilno_ZavarovalniškaPogodba. Tuji ključ, znan tudi kot sekundarni ključ, je stolpec v tabeli, katerega vrednosti ustrezajo vrednostim glavnega ključa v drugi tabeli. V podatkovnem modelu je vidno, da so tuji ključi prisotni v dimenzijskih tabelah. Na primeru razvitega podatkovnega modela je tabela Zobje_ZavarovalniškaPogodba povezana s tabelo dejstev Datum_pogodbe na podlagi atributa MinOfDATUM_ZACETKA_OSNOVNE_POGODBE, zato bi ta stolpec v tabeli Datum_pogodbe deloval kot tuji ključ.

5.4.2 Kardinalnost

Tabele dimenzij in tabele dejstev so povezane z glavnim in tujim ključem, vendar je treba podrobneje opredeliti tudi njihovo razmerje. Vsako razmerje v modelu je opredeljeno s tipom kardinalnosti. Ta se nanaša na število edinstvenih vrednosti v eni tabeli, povezanih s številom edinstvenih vrednosti v drugi tabeli. V podatkovnem modeliranju v orodju Power BI obstajajo tri vrste razmerij kardinalnosti.

Ena proti ena (1:1) je relacija med tabelami, ko je en zapis v prvi tabeli povezan z enim samim atributom v drugi tabeli. Ta vrsta razmerja je pri modeliranju podatkov v orodju Power BI razmeroma redka.

Ena proti mnogim (1:N) je razmerje med tabelami, ko je en zapis v prvi tabeli lahko povezan z več zapisi v drugi tabeli. Je najpogostejša vrsta razmerja pri modeliranju podatkov v orodju Power BI. To vrsto kardinalnosti sem uporabila za povezovanje dejanske tabele z eno ali več

dimenzijskimi tabelami, pri čemer so dimenzijske tabele pogosto enostranska razmerja, medtem ko je dejanska tabela pogosto v razmerju mnogo.

Glavna prednost orodja Power BI, ki se pojavlja pri opredelitvi kardinalnosti, je, da orodje ponuja možnost avtomatičnega prikazovanja kardinalnosti med tabelami. Tudi v primeru, če je ta vrsta relacij nemogoča, ponuja obrazložitev zaradi nezmožnosti povezave med tabelami.

Zadnja vrsta kardinalnosti je mnogo proti mnogo (M:N). To je relacija, ki se zgodi, ko je veliko zapisov v prvi tabeli povezanih z veliko zapisi v drugi tabeli. Ta vrsta razmerja ni priporočena pri modeliranju podatkov v orodju Power BI in se zelo redko uporablja.

Pri pripravi podatkovnega modela sem kardinalnost čim bolj smiselno upoštevala. Tako je kardinalnost v podatkovnem modelu usmerjena zgolj v razmerje ena proti mnogim (1:N), ki predstavlja razmerje med tabelami, ko je zapis iz ene tabele povezan z več zapisi v drugi tabeli, kot je vidno v podatkovnem modelu v poglavju 4.6.

5.5 Izgradnja poslovno inteligenčne nadzorne plošče za vizualizacijo podatkov o rednih zobozdravstvenih pregledih zavarovancev v izbrani zavarovalnici

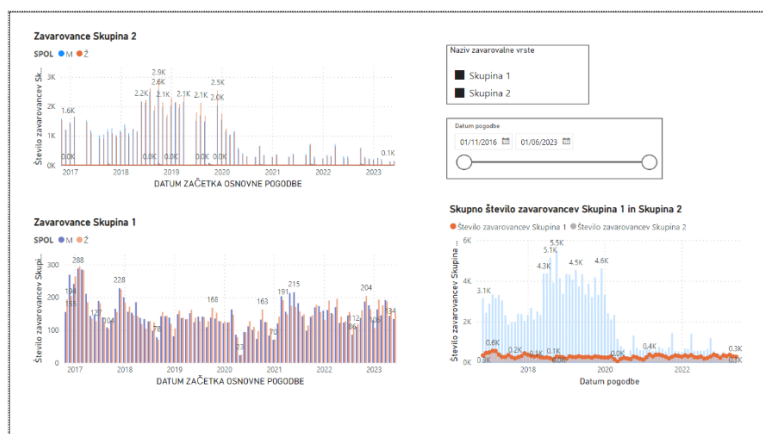
Pri izgradnji nadzorne plošče v orodju Power BI sem upoštevala potrebe izbrane zavarovalnice in sledila dobrim smernicam iz literature ter izkušnjam, ki sem jih nabrala z opravljanjem dela v analitičnem oddelku izbrane zavarovalnice. Najprej sem podatke analizirala pred nalaganjem v orodje Power BI, tako da sem dobro pregledala Excelove datoteke in njihove stolpce ter posamezne vrstice zaradi boljšega ustvarjanja vizualne predstavitve podatkov.

Svojo analizo sem začela pripravljati leta 2023, podatki pa obravnavajo dve skupini zavarovancev, ki sem jih v vizualnem prikazu ustrezno anonimizirala. Omenila sem že, da gre za zavarovance Skupine 1 s podoznako Dopolnilno zavarovanje in zavarovance Skupine 2 s podoznako Zobje zavarovanje.

Magistrsko delo obravnava eno izmed njih, in sicer Dopolnilno zavarovanje, ki ga od leta 2024 ni več. Te podatke sem uporabila namenoma, saj je lahko, kljub temu da dopolnilnega zavarovanja ni več, analiza dogajanja za podjetje še vedno smiselna, saj morajo mestoma analizirati tudi podatke za nazaj, kar jim pride prav za nadaljnje analize in pripravo njihovih storitev ter ustvarjanje novih produktov.

Ustvarila sem sedem nadzornih plošč, ki imajo v povprečju šest vizualizacij na posamezno nadzorno ploščo, kar je smiselno za določene analize. Na sliki 14 je prikazana nadzorna plošča 1, imenovana Analiza osnovne sklenjene pogodbe zavarovancev Skupine 1 in zavarovancev Skupine 2.

Slika 14: Nadzorna plošča 1: Analiza osnovne sklenjene pogodbe zavarovancev Skupine 1 in zavarovancev Skupine 2

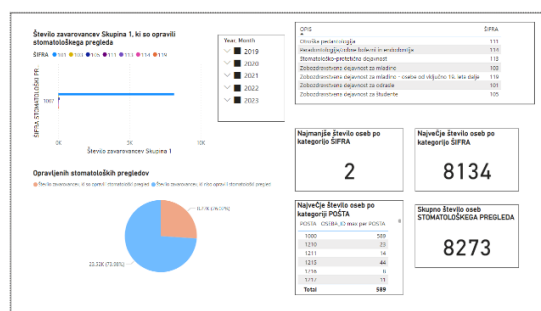


Vir: lastno delo.

Na nadzorni plošči 1 je predstavljena analiza osnovne sklenjene pogodbe za izbrano zavarovalnico. Uporabniku omogoča interaktivno uporabo filtrov glede na zavarovance Skupine 1 in zavarovance Skupine 2 glede na določen datum sklenjene osnovne pogodbe zdravstvenega zavarovanja. Nadzorna plošča zajema vizualni prikaz, ki prikazuje primerjavo med skupnim številom zavarovancev Skupine 1 in zavarovancev Skupine 2 glede na datum sklenjene pogodbe. Bolj podrobno sem prikazala ločene vizualne prikaze zavarovancev določene skupine, glede na to, da sem znotraj prikaza vključila tudi kategorijo SPOL, kjer sem zavarovance moškega spola označila z modro barvo, ženskega pa z oranžno. Pri filtru Naziv zavarovalne vrste sem v prvi fazi prikazala skupine z njihovimi identifikacijskimi številkami (Zavarovalna vrsta_ID), nato sem zaradi boljšega razumevanja uporabnikov in hitrejšega premikanja v nadzorni plošči dodala stolpec Naziv_zavarovalna_vrsta, s katerim sem jih bolj precizno označila kot Skupina 1 in Skupina 2. Z razdelkom Datum pogodbe sem prikazala obdobje od 1. 11. 2016 do 1. 6. 2023. Uporabniku omogoča premikanje po obdobjih, ustrezni prikaz podatkov za obdobje petih let ter primerjavo med določenimi datumi sklenjenih pogodb zavarovancev.

S pomočjo nadzorne plošče 1 lahko ugotovim, da je število zavarovancev Skupine 2 večje. V Skupino 1 štejemo zavarovance s podoznako Zobje in zavarovance Skupine 2 s podoznako Dopolnilno. Kot je jasno vidno, je največ zavarovancev Skupine 2 sklenilo osnovno pogodbo v letu 2018, in sicer so bili to zavarovanci ženskega spola. Največ zavarovancev Skupine 1 je sklenilo osnovno pogodbo v letu 2017 – tudi ti zavarovanci so bili ženskega spola. Vizualni prikaz z nazivom Skupno število zavarovancev Skupine 1 in Skupine 2 prikazuje primerjavo med številom zavarovancev po skupinah, iz česar je mogoče sklepati, da je večje število zavarovancev Skupine 2 sklenilo osnovno pogodbo zdravstvenega zavarovanja. Naslednja nadzorna plošča, analiza stomatoloških pregledov zavarovancev Skupine 1, je predstavljena na sliki 15.

Slika 15: Nadzorna plošča 2: Analiza stomatoloških pregledov zavarovancev Skupine 1

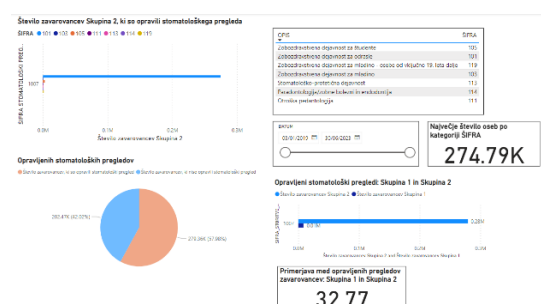


Vir: lastno delo.

Na nadzorni plošči 2 je predstavljena analiza stomatoloških pregledov zavarovancev Skupine 1. Uporabniku omogoča interaktivno uporabo filtrov po letih in možnosti izbiranja po mesecih ter selektivno izbiro določene SIFRE (s pripadajočim opisom), kar se nanaša na šifre škode stomatološkega pregleda. Z namenom bolj podrobne analize sem na drugi nadzorni plošči predstavila le zavarovance Skupine 1, ker sem hotela bolj podrobno prikazati njihove značilnosti in primerjave. Vidno je, da ima uporabnik celotni vpogled v vse storitve, ki pripadajo skupini stomatološkega pregleda, zato sem dodala OPIS storitve. Dodatno sem v vizualni prikaz dodala še šifre škode zaradi lažjega prikazovanja dodatnih značilnosti in primerjave med skupinami storitve.

S pomočjo nadzorne plošče 2 lahko ugotovim, da je največje število zavarovancev Skupine 1 uporabilo zobozdravstveno dejavnost za odrasle v obdobju od leta 2019 do leta 2023, in to kar v veliki razliki glede na ostale vrste škode. Stomatološki pregled je opravilo 23 % zavarovancev glede na vse iz Skupine 1, ki so imeli sklenjeno osnovno pogodbo o zdravstvenem zavarovanju. Nato je vidno, da največje število zavarovancev Skupine 1 v podkategoriji POŠTA spada pod poštno številko 1000, kar pomeni, da je največje število zavarovancev opravilo stomatološki pregled v Ljubljani. Naslednja nadzorna plošča, analiza stomatoloških pregledov Skupine 2 in primerjava med skupinama zavarovancev, je predstavljena na sliki 16.

Slika 16: Nadzorna plošča 3: Analiza stomatoloških pregledov Skupine 2 in primerjava med skupinami zavarovancev

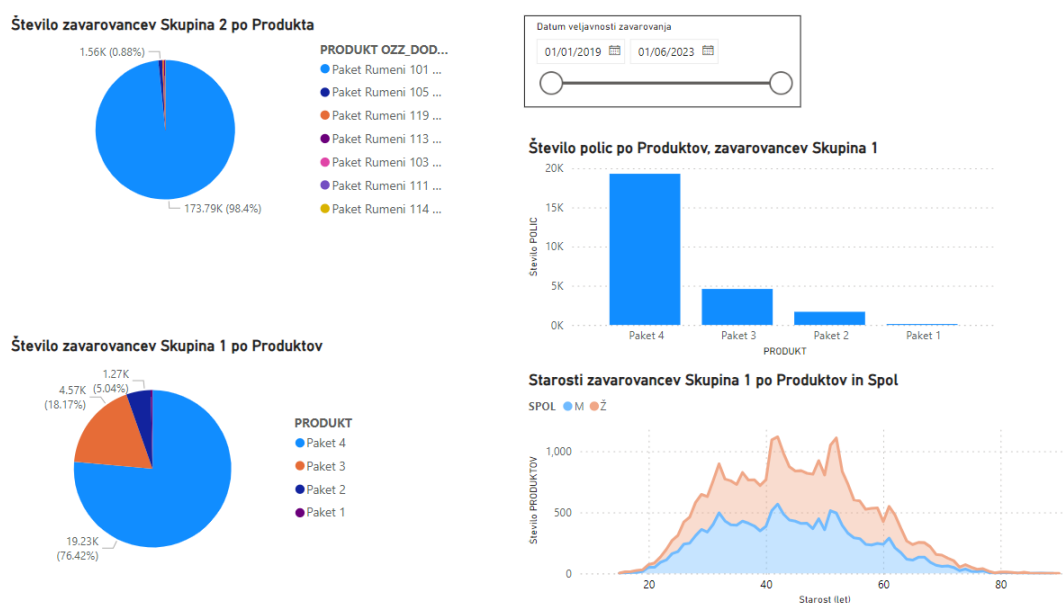


Vir: lastno delo.

Na nadzorni plošči 3 sta predstavljeni analiza stomatoloških pregledov zavarovancev Skupine 2 in primerjava med skupinami zavarovancev. Uporabniku omogoča interaktivno uporabo filtrov po letih in možnosti izbiranja po mesecih ter selektivno izbiro določene SIFRE, kar se nanaša na šifre škode stomatološkega pregleda, z možnostjo branja opisov storitve. V spodnjem razdelku je vidno, da ima uporabnik vpogled v storitve, ki pripadajo skupini stomatološkega pregleda. Tako ima uporabnik vpogled v število zavarovancev Skupine 2, ki so opravili stomatološki pregled, označeno z oranžno barvo, ter število zavarovancev Skupine 2, ki niso opravili stomatološkega pregleda, označeno z modro barvo.

S pomočjo nadzorne plošče 3 lahko ugotovim, da je glede na vrsto škode v obdobju od leta 2019 do leta 2023 največje število zavarovancev Skupine 2 uporabilo zobozdravstveno dejavnost za odrasle. Stomatološki pregled je opravilo 42,02 % vseh, ki so imeli sklenjeno osnovno pogodbo o zdravstvenem zavarovanju Skupine 2. Vizualno sem predstavila primerjavo opravljenih stomatoloških pregledov zavarovancev Skupine 1 in Skupine 2. Iz tega sklepam, da so zavarovanci Skupine 2 v obdobju od leta 2019 do leta 2023 opravili 32,77-krat več stomatoloških pregledov kot zavarovanci Skupine 1. Dodatno sem predstavila največje število zavarovancev po kategoriji SIFER. Iz tega je mogoče sklepati, da je pri šifri s številom 101 največ zavarovancev v skupnem največjem številu 274.790 iz Skupine 2. Naslednja nadzorna plošča, analiza zavarovancev Skupine 1 in Skupine 2 po produktih, je predstavljena na sliki 17.

Slika 17: Nadzorna plošča 4: Analiza zavarovancev Skupine 1 in Skupine 2 po produktih



Vir: lastno delo.

Na nadzorni plošči 4 je predstavljena analiza zavarovancev Skupine 1 in Skupine 2 po produktih. Uporabniku omogoča interaktivno uporabo filtrov po letih in možnosti izbiranja po mesecih. V prvem delu uporabniku omogoča vpogled v število zavarovancev Skupine 2

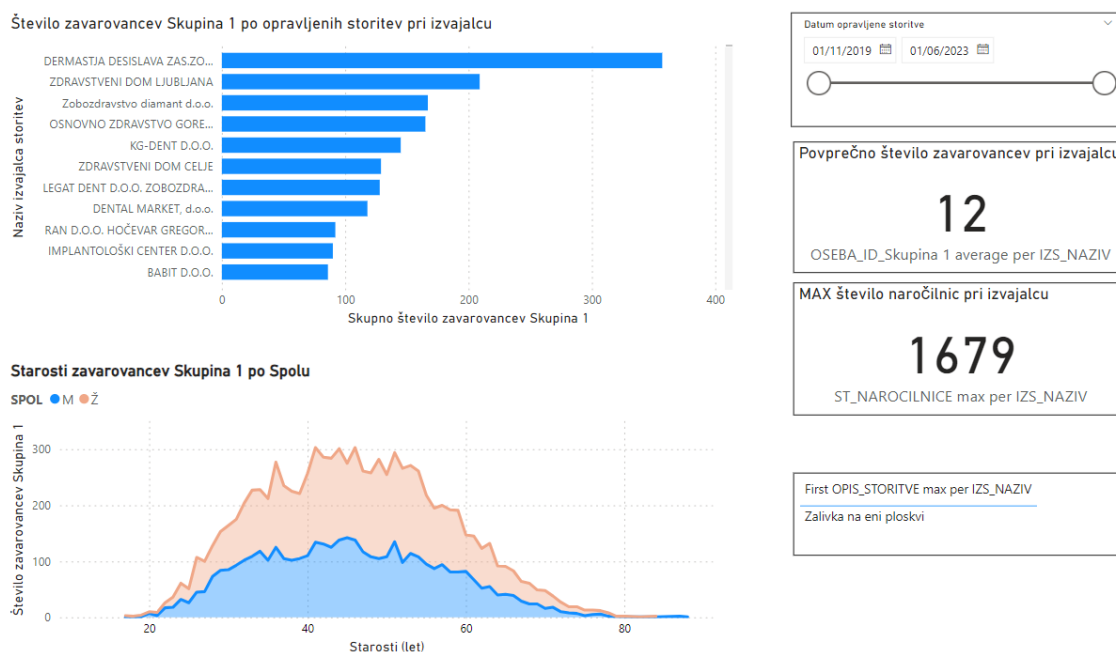
po enem samem produktu, imenovanem Paket Rumeni. Zaradi tega sem ga povezala z atributom SIFRA, ki omogoča ločitev med dodatnimi vsebujočimi storitvami.

Nadzorna plošča uporabniku omogoča vpogled po produktih za zavarovance Skupine 1, skupno po štirih produktih. Dodatno sem dodala podrobno analizo produktov za zavarovance Skupine 1 po številu polic ter starosti in spolu zavarovancev.

S pomočjo nadzorne plošče 4 lahko ugotovim, da je glede na vrsto šifre v obdobju od leta 2019 do leta 2023 največje število zavarovancev Skupine 2, glede na šifro 101 in produkt Rumeni, v največji meri uporabilo zobozdravstveno dejavnost za odrasle. Število zavarovancev Skupine 1 je v največji meri uporabilo produkt, imenovan Paket 4, in sicer za 76,42 % več kot se je uporabilo ostalih treh produktov storitev.

Dodatno sem predstavila število zavarovancev Skupine 1 po številu izdanih polic ter starosti zavarovancev. Iz tega je mogoče sklepati, da je največ zavarovancev v skupnem številu 19.302 polic pri produktu Paket 4. Naslednja nadzorna plošča, analiza zavarovancev Skupine 1 po izvajalcih storitev, je predstavljena na sliki 18.

Slika 18: Nadzorna plošča 5: Analiza zavarovancev Skupine 1 po izvajalcih storitev



Vir: lastno delo.

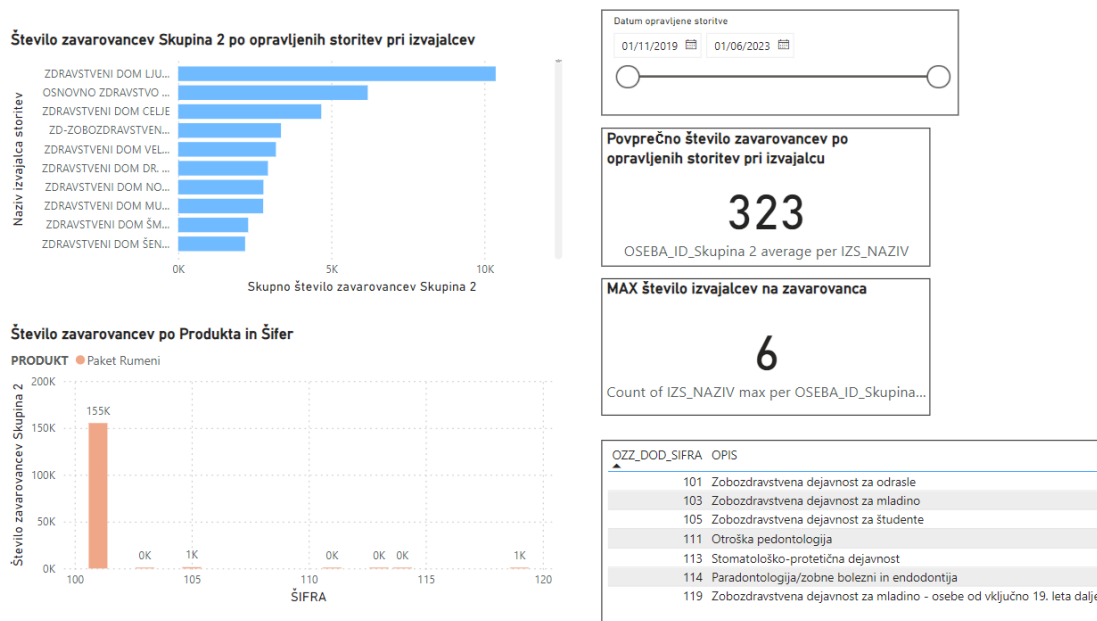
Na nadzorni plošči 5 je predstavljena analiza zavarovancev Skupine 1 po kategoriji izvajalcev stomatoloških storitev. Uporabniku omogoča interaktivno uporabo filtrov po letih in možnosti izbiranja po mesecih, posledično lažji prikaz povprečnega števila zavarovancev pri posameznem izvajalcu storitev ter največje število naročilnic pri posameznem izvajalcu.

Vidno je, da ima uporabnik celotni vpogled v prve najpogostejše obiskane izvajalce storitve. Dodatno ima vpogled v opis storitve glede na izvajalca in naziv storitve. Za podrobnejšo analizo pa še vpogled v starost in spol zavarovancev Skupine 1, ki so opravili storitev pri izvajalcu.

S pomočjo nadzorne plošče 5 lahko ugotovim, da je v obdobju od leta 2019 do leta 2023 največ zavarovancev Skupine 1 opravilo storitev pri izvajalcu Dermastija Desislava, Zasebna zobozdravstvena ordinacija. Drugi najbolj obiskan izvajalec storitve je Zdravstveni dom Ljubljana. Podatki kažejo, da so zavarovanci v največji meri ženske.

Dodatno sem predstavila še opis najbolj pogoste storitve glede na najbolj obiskanega izvajalca, kar se je velikokrat ponavljalo. Najpogostejši razlog za obisk je bila zalivka na eni ploskvi. Naslednja nadzorna plošča, analiza zavarovancev Skupine 2 po izvajalcih storitev, je predstavljena na sliki 19.

Slika 19: Nadzorna plošča 6: Analiza zavarovancev Skupine 2 po izvajalcih storitev



Vir: lastno delo.

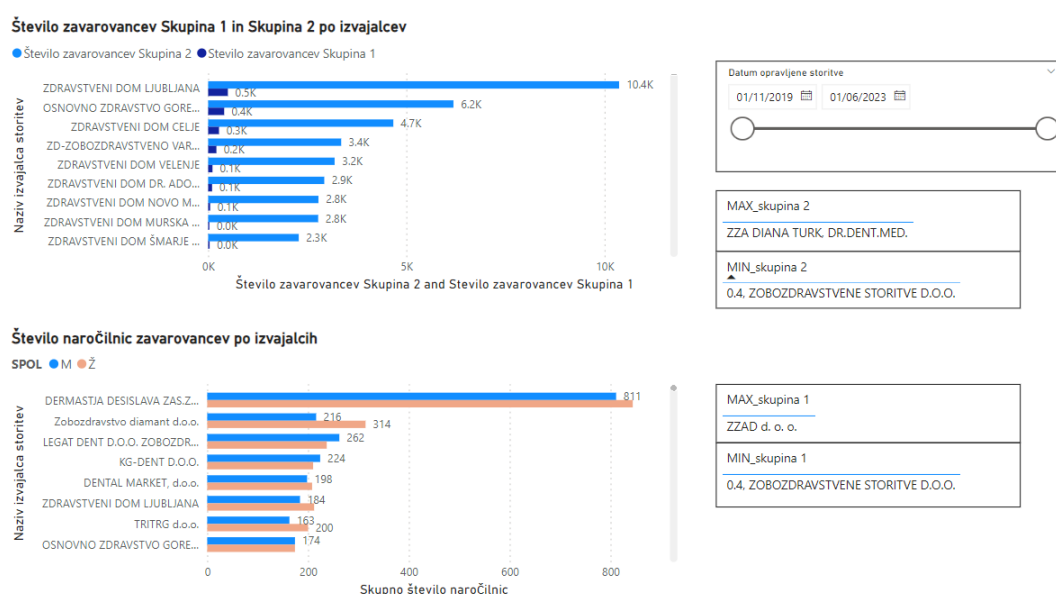
Na nadzorni plošči 6 je predstavljena analiza zavarovancev Skupine 2 po kategoriji Naziv izvajalcev stomatoloških storitev. Uporabniku omogoča interaktivno uporabo filtrov po letih in možnosti izbiranja po mesecih, posledično lažji prikaz povprečnega števila zavarovancev pri posameznemu izvajalcu storitev ter največje število izvajalcev na določenega zavarovanja.

Za bolj jasen pregled podatkov ima uporabnik vpogled v število zavarovancev po produktu in šifri storitev. Dodatno so zapisani še opisi posameznih šifer, ki bolj natančno opisujejo določene aktivnosti.

S pomočjo nadzorne plošče 6 lahko ugotovim, da je v obdobju od leta 2019 do leta 2023 največ zavarovancev Skupine 2 opravilo storitev pri izvajalcu Zdravstveni dom Ljubljana. Drugi najbolj obiskan izvajalec storitve je Osnovno zdravstvo Gorenjske.

Podatki kažejo, da so zavarovanci v največji meri obiskali izvajalca v okviru zobozdravstvene dejavnosti za odrasle, in to kar 155.243 zavarovancev Skupine 2. Dodatno sem predstavila še izračun največjega števila izvajalcev na posameznega zavarovanca (šest izvajalcev na zavarovanca). Naslednja nadzorna plošča, analiza zavarovancev Skupine 1 in Skupine 2 po izvajalcih storitev, je predstavljena na sliki 20.

Slika 20: Nadzorna plošča 7: Analiza zavarovancev Skupine 1 in zavarovancev Skupine 2 po izvajalcih storitev



Vir: lastno delo.

Na nadzorni plošči 7 je predstavljena analiza zavarovancev Skupine 1 in zavarovancev Skupine 2 po kategoriji Naziv izvajalcev stomatoloških storitev. Uporabniku omogoča interaktivno uporabo filtrov po letih in možnosti izbiranja po mesecih. Tako je predstavljena primerjava med skupnim številom zavarovancev Skupine 1 in skupnim številom zavarovancev Skupine 2 po nazivu izvajalca stomatoloških storitev. Uporabnik ima možnost vpogleda v število naročilnic zavarovancev po izvajalcih glede na spol zavarovanca. Dodatno še vpogled v najbolj in najmanj obiskane izvajalce storitev po posameznih skupinah zavarovancev.

S pomočjo nadzorne plošče 7 lahko ugotovim, da je v obdobju od leta 2019 do leta 2023 največ zavarovancev Skupine 2 opravilo storitev pri izvajalcu Zdravstveni dom Ljubljana. Drugi najbolj obiskan izvajalec storitve pa je Osnovno zdravstvo Gorenjske, kjer je vidna razlika števila zavarovancev Skupine 2, ki so obiskali tega izvajalca v večji meri glede na

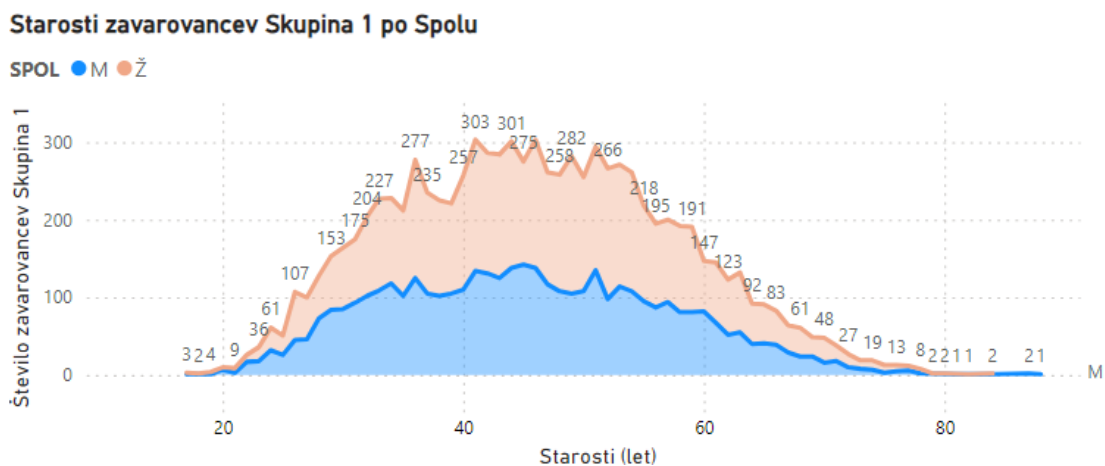
zavarovance Skupine 1, s tem da je ta razlika manj vidna kot pri Zdravstvenem domu Ljubljana.

V drugem delu je predstavljen izračun najbolj in najmanj obiskanega izvajalca s strani zavarovancev obeh skupin. Iz tega je mogoče sklepati, da sta Skupina 1 in Skupina 2 imeli največjo razliko pri najbolj pogosto obiskanem izvajalcu storitve. Pri zavarovancih Skupine 2 je bil to ZZA Diana Turk, dr. dent. med, pri zavarovancih Skupine 1 pa je to bil ZZAD d.o.o. Najmanj opravljenih storitev pa so zavarovanci obeh skupin opravili pri izvajalcu 0,4 Zobozdravstvene storitve, d.o.o.

Zgoraj sem predstavila vizualizacijo sedmih integriranih nadzornih plošč z uporabo orodja Power BI, kjer je med drugim opisana analiza posameznih podatkov, vezanih na število zavarovancev in opravljene redne zobozdravstvene storitve. V tem poglavju sledi podrobna interpretacija rezultatov posameznih grafov in analiza možnih scenarijev.

Vizualni prikaz starosti in spola zavarovancev Skupine 1 sem pripravila tako, da sem s pomočjo DAX pripravila poizvedbo, s katero sem s pomočjo atributa datum rojstva izračunala starosti posameznih zavarovancev (v letih). Dodatno sem dodala še spol zavarovancev. Na sliki 21 je predstavljen naloženi ploščinski grafikon starosti zavarovancev Skupine 1 po spolu.

Slika 21: Starost zavarovancev Skupine 1 po spolu



Vir: lastno delo.

V obdobju od leta 2019 do leta 2023 se je starost zavarovancev zelo razlikovala. Največje število oseb, ki so opravile stomatološki pregled, je bilo ženskega spola pri starosti 41 let. Prav tako so ženske najbolj pogosto obiskale zobozdravstvenega izvajalca storitev.

Tabela 3 prikazuje število opravljenih rednih zobozdravstvenih pregledov v obdobju od leta 2019 do leta 2023.

Analiza podatkov kaže na to, da so v povprečju največ obiskov opravili zavarovanci obeh spolov pri starosti 46 let. Najmlajša oseba, ki je v analiziranem obdobju opravila pregled, je bila stara 17 let in je bila ženskega spola. Najstarejša oseba je bila stara 88 let in je bila moškega spola. Največja obiskanost in opravljanje rednih stomatoloških pregledov temeljita na tem, da so osebe v obdobju od 40 do 50 let starosti najbolj izpostavljene poškodbam zob, ki bi lahko nastale zaradi starosti. Menim, da so zavarovanke ženskega spola bolj odgovorne in skrbne za svoje zdravje, zato bolj pogosto opravljajo redne stomatološke preglede kot moški.

Tabela 3: Analiza zavarovancev po starosti in spolu

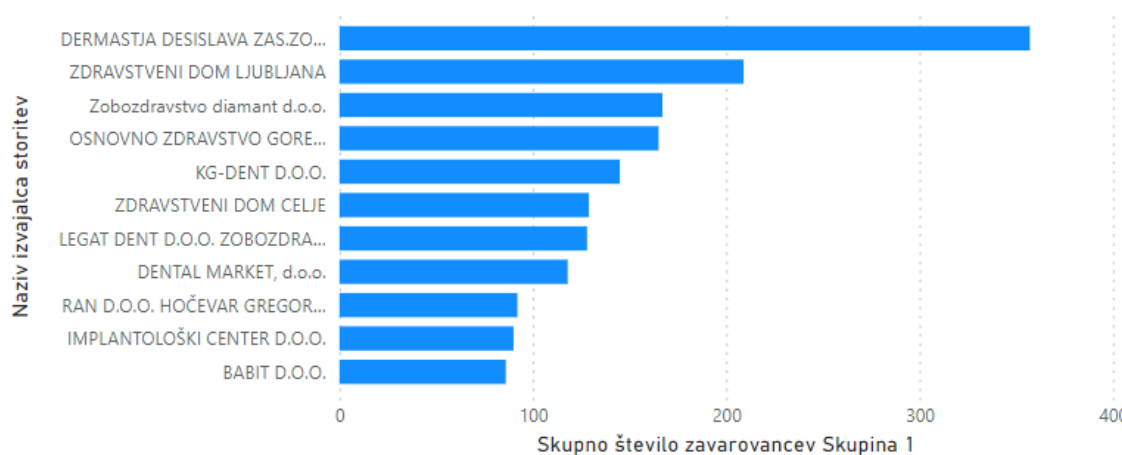
Obdobje (leta)	2019–2020	2020–2021	2021–2022	2022–2023
Ženski	69 oseb (53 let)	83 oseb (44 let)	110 oseb (46 let)	81 oseb (41 let)
Moški	58 oseb (53 let)	81 oseb (44 let)	86 oseb (46 let)	70 oseb (41 let)
Največje število oseb v obdobju	127 oseb	164 oseb	196 oseb	151 oseb

Vir: lastno delo.

Vizualni prikaz opravljenih rednih zobozdravstvenih pregledov zavarovancev Skupine 1 in zavarovancev Skupine 2 sem pripravila s pomočjo jezika M oziroma uporabe orodja Power Query. Namen je bilo ustvariti novo tabelo, poimenovano IZV storitev (izvajalci storitev), in prikazati vse izvajalce storitev ter jih tako ustrezno povezati v povezovalno tabelo za obe skupini zavarovancev, nato pa jih ustrezno primerjati med seboj. Dodatno sem v sklopu tega prikaza zaradi bolj jasnega prikaza podatkov dodala še kategorijo spol. Na sliki 22 je predstavljen gručasti palični grafikon, imenovan število zavarovancev Skupine 1 po opravljenih storitvah po nazivu izvajalcev storitve.

Slika 22: Število zavarovancev Skupine 1 po opravljenih storitvah pri izvajalcih

Število zavarovancev Skupina 1 po opravljenih storitvah pri izvajalcu



Vir: lastno delo.

V obdobju od leta 2019 do leta 2023 so zavarovanci Skupine 1 obiskali različne izvajalce storitev. Največje število zavarovancev je opravilo storitev pri izvajalcu Dermastija Desislava, Zasebna zobozdravstvena ordinacija, 148 zavarovancev več glede na drugega najbolj pogosto obiskanega izvajalca storitev. Drugi najbolj obiskani izvajalec storitve pa je Zdravstveni dom Ljubljana. Podatki kažejo, da so zavarovanci v največji meri ženskega spola. Povprečna starost zavarovancev, ki so najbolj pogosto obiskali katerega izmed izvajalcev, je bila 41 let. Na sliki 23 je predstavljen gručasti palični grafikon, ki prikazuje skupno število zavarovancev Skupine 2 po nazivu izvajalcev storitev.

Slika 23: Število zavarovancev Skupine 2 po opravljenih storitvah pri izvajalcih



Vir: lastno delo.

V obdobju od leta 2019 do leta 2023 so zavarovanci Skupine 2 obiskali različne izvajalce storitev. Največje število zavarovancev je opravilo zobozdravstvene preglede pri izvajalcu Zdravstveni dom Ljubljana, 4179 zavarovancev več glede na drugega najbolj pogosto obiskanega izvajalca storitev. Drugi najbolj obiskani izvajalec storitve pa je Osnovno zdravstvo Gorenjske. Podatki kažejo, da so bile storitve v največji meri opravljene iz naslova zobozdravstvene dejavnosti za odrasle, v najmanjšem številu pa iz naslova parodontologije/zobnih bolezni in endodontije.

Tabela 4 prikazuje, da je bilo v obdobju od leta 2019 do leta 2023 največ obiskanih izvajalcev na področju zobozdravstvenih storitev za Skupino 1 prav v Ljubljani. Bolj podrobno analiza pravi, da je pri zavarovancih Skupine 1 sorazmerno padajoče število zavarovancev med izbranimi petimi najpogostejše obiskanimi zavarovanci. Med njimi so zavarovanci najbolj pogosto obiskali izvajalce v Ljubljani, najmanj pa v Celju.

Tabela 4: Analiza zavarovancev Skupine 1 po opravljenih storitvah pri petih najbolj pogosto obiskanih izvajalcih

Naziv	Dermastija Desislava, Zasebna zobozdravstvena ordinacija	Zdravstveni dom Ljubljana	Zobozdravstvo Diamant d.o.o.	Osnovno zdravstvo Gorenjske (Zdravstveni dom Radovljica)	Zdravstveni dom Celje
Skupina 1	357 oseb	209 oseb	167 oseb	165 oseb	129 oseb
Pošta in kraj opravljene storitve	1000 Ljubljana	1000 Ljubljana	1000 Ljubljana	4000 Kranj	3000 Celje

Vir: lastno delo.

Tabela 5 prikazuje, da je bilo v obdobju od leta 2019 do leta 2023 najmanj obiskanih izvajalcev na področju zobozdravstvenih storitev Skupine 1 na območju občine Ruše. Med vsemi zavarovanci Skupine 1 so najmanj pogosto obiskali izvajalce v medkrajevnih območnih enotah zaradi težje dostopnosti do večjih mest.

Tabela 5: Analiza zavarovancev Skupine 1 po opravljenih storitvah pri petih najmanj pogosto obiskanih izvajalcih

Naziv	Zobozdravstvo VIHER d.o.o.	Dental Studio Idea Nova, d.o.o.	B.D.D. - Dental d.o.o	ALFA Dental d.o.o.	Ambrožič Selina – zasebna ordinacija
Skupina 1	1 oseba	2 osebi	3 osebe	4 osebe	5 oseb
Pošta in kraj opravljene storitve	2342 Ruše	5000 Nova Gorica	2000 Maribor	2000 Maribor	5270 Ajdovščina

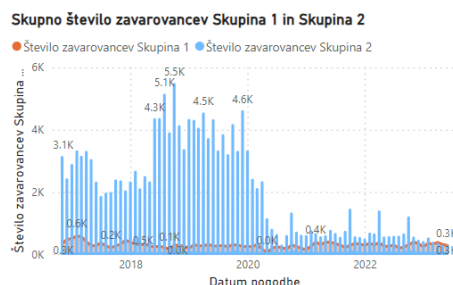
Vir: lastno delo.

Iz analize zavarovancev Skupine 1 lahko ugotovim, da imajo zavarovanci, ki živijo v socialno prikrajšanih skupnostih, kot je medkrajevno območje Ruše, manjši dostop do večjega nabora izvajalcev različnih vrst zobozdravstvenih storitev. To je en izmed razlogov za slabše zdravstveno stanje, kot ga imajo tisti iz bolj premožnih skupnosti ali tisti iz večjih mest, ki imajo več izbire zobozdravstvenih izvajalcev storitev. Kot je vidno v tabeli, je v tej isti regiji prisotno večje mesto Maribor, ki ima zelo velik nabor izvajalcev zdravstvenih storitev. Na drugi strani je območje Nova Gorica, ki je trinajsto največje mesto v Sloveniji in ima značilno manjši nabor izvajalcev glede na Ajdovščino, obe pa spadata v Severno Primorsko regijo.

Vizualni prikaz obdobja sklenjenih zavarovanj zavarovancev Skupine 1 in zavarovancev Skupine 2 sem pripravila s pomočjo dvojnega štetja zavarovancev obeh skupin. Dodatno

sem v sklopu tega prikaza dodala še kategorijo spol zaradi boljšega prikaza podatkov. Na sliki 24 je predstavljen območni grafikon skupnega števila zavarovancev Skupine 1 in zavarovancev Skupine 2 po spolu in sklenitvi zavarovalne pogodbe.

Slika 24: Skupno število zavarovancev Skupina 1 in Skupina 2



Vir: lastno delo.

V obdobju od leta 2016 do leta 2023 se je datum sklenitve zavarovanja med skupinama zavarovancev zelo razlikoval. Največje število zavarovancev, ki so sklenili zavarovanje, je bilo iz Skupine 2 v letu 2018. Vidno je, da se je od leta 2020 število zavarovancev Skupine 2, ki so sklenili zobozdravstveno zavarovanje v izbrani zavarovalnici, zmanjševalo. Razlog za padec števila novih zavarovancev temelji na pojavu bolezni pandemije koronavirusne bolezni 2019 (angl. coronavirus disease 2019, v nadaljevanju COVID-19) in različnih ukrepov, ki so bili posledično v veljavi. Tabela 6 prikazuje, da je bilo v izbrani zavarovalnici največ sklenjenih pogodb zavarovanja na področju zobozdravstvenih storitev v obdobju od leta 2017 do leta 2018. Bolj podrobna analiza pravi, da je število zavarovancev Skupine 2 v primerjavi z zavarovanci Skupine 1 večje. Analizo sem opravila za vsako skupino posebej, nato pa sledi primerjava med njima.

Tabela 6: Analiza zavarovancev po datumu sklenitve zavarovanja

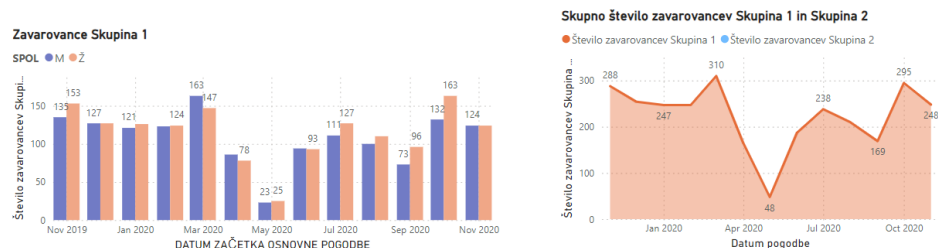
Obdobje (leta)	2016–2017	2017–2018	2018–2019	2019–2020	2020–2021	2021–2022	2022–2023
Skupina 1	582 oseb	452 oseb	307 oseb	310 oseb	394 oseb	391 oseb	380 oseb
Skupina 2	3298 oseb	5461 oseb	4514 oseb	4587 oseb	1422 oseb	1367 oseb	523 oseb
Največje število zavarovancev obeh skupin	3880 oseb	5913 oseb	4821 oseb	4897 oseb	1816 oseb	1758 oseb	903 oseb

Vir: lastno delo.

Zavarovanci Skupine 1 so v največji meri sklenili zavarovalno pogodbo v obdobju med letoma 2016 in 2017. Vidno je, da so več pogodb sklenili med letoma 2016 in 2018, kar pa pomeni, da so imeli od leta 2018 do leta 2023 približno sorazmerno naraščajoče število novih sklenjenih pogodb. Glede na kategorijo spola se med analizo sedmih let pokaže, da so

zavarovanke ženskega spola v večji meri sklenile pogodbo glede na zavarovance moškega spola. V največjem številu je bilo to 294 zavarovank ženskega spola v začetku leta 2017. Največja posebnost je nastala med letoma 2019 in 2020. Podrobnosti zavarovancev Skupine 1 sem prikazala na sliki 25.

Slika 25: Zavarovanci Skupine 1 po datumu sklenitve pogodbe in spolu



Vir: lastno delo.

Na sliki 27 je vidno, da je nastala največja posebnost pri sklenitvah pogodb v maju leta 2020, ko se je število sklenitev zavarovanj zmanjšalo od skupnega števila zavarovancev Skupine 1 – aprila je bilo skupno število zavarovancev 164, maja pa 48. Junija se je število zavarovancev ponovno povečalo na skupno število 187. Kot je vidno, se je število novih sklenjenih pogodb potem takoj spet zvišalo.

Zavarovanci Skupine 2 so v največji meri sklenili zavarovalno pogodbo v obdobju med letoma 2017 in 2018. Vidno je, da so več pogodb sklenili med letoma 2016 in 2020, kar pa pomeni, da so imeli od leta 2020 do leta 2023 približno sorazmerno padajoče število novih sklenjenih pogodb.

6 DISKUSIJA

V nadaljevanju so predstavljene ključne ugotovitve analize podatkov. Poudarek je na pomenu ugotovitve za izbrano zavarovalnico oz. predstavitvi koristi, ki jih uporaba BI nadzorne plošče lahko prinese zavarovalnicam. Podane so tudi omejitve raziskave ter možnosti za nadaljnje delo.

6.1 Ključne ugotovitve za izbrano zavarovalnico

Vizualizacijo podatkov sem pripravila, da bi prišla do ključnih ugotovitev, vezanih na opravljene redne zobozdravstvene preglede zavarovancev Skupine 1 in zavarovancev Skupine 2. Glede na interpretacijo rezultatov in možne scenarije ugotavljam, da so glavni razlogi za redno opravljene zobozdravstvene preglede vezani na starost, spol, obdobje sklenitve zdravstvene pogodbe ter dostopnost do izvajalcev storitve.

Proces ustvarjanja nadzorne plošče je potekal tako, da sem zbrala in prečistila podatke, nato pa skladno z izbrano metodologijo izdelala nadzorno ploščo z jasno in enostavno predstavljenimi pomembnimi podatki, ki lahko pomagajo odločevalcem pri sprejemanju boljših odločitev v prihodnosti.

V okviru magistrskega dela sem prepoznala tudi ključne koristi za zavarovalnice, ki pri svojem poslovanju uporabljajo nadzorne plošče, predvsem pa je pomembno, da se lahko z njihovo uporabo izboljša kakovost odločitev, kar lahko pozitivno vpliva na poslovanje v prihodnosti. Nadzorne plošče BI lahko uporabnikom znotraj podjetja pomagajo, da se usmerijo na področja oziroma storitve, ki so bile zapostavljene, in tako izboljšajo ponudbo storitev za svoje zavarovance. Na ta način bodo zavarovalnice imele večji nadzor nad vsemi storitvami različnih skupin strank in bodo lahko analizirale njihovo obnašanje skozi čas. Prihodnje odločitve bodo na ta način ustrezno izboljšane in bodo lahko privedle do zadovoljnejših strank in večje uspešnosti podjetja.

Med razvojem nadzorne plošče sem prepoznala tudi nekaj izzivov, ki izhajajo predvsem iz podatkov. Podjetje se sooča s pomanjkljivimi in neprečiščenimi podatki, pa tudi konkretnih podatkov o preteklih dogodkih oz. zavarovancih je razmeroma malo, kar zavarovalnici ne omogoča tako kakovostne analize podatkov, kot bi lahko bila.

Poleg konkretnih ugotovitev na izbranem primeru pa dodano vrednost magistrskega dela predstavlja tudi prikaz postopka obdelave podatkov ter ustvarjanje nadzorne plošče BI, kar lahko koristi organizacijam v zavarovalniški panogi, ki bodo pri svojem delu želele uporabljati nadzorne plošče oz. izkoristiti potencial, ki ga v sodobnem svetu prinaša poslovanje na podlagi BI.

6.2 Pomen ugotovitve za izbrano zavarovalnico

Izbrana zavarovalnica se v zadnjih časih spopada z večjimi spremembami na trgu. Pripravljena analiza in vizualizacija podatkov bosta tako pozitivno vplivali na delovanje zavarovalnice, predvsem v oddelku za razvoj produktov.

Glavne ugotovitve, ki vplivajo na zavarovalnico, dajejo večji pomen zavarovancem oz. njihovemu boljšemu poznavanju. Predvsem je to pokazala analiza moških zavarovancev, kjer se je izkazalo, da se ne zavedajo, da so redni pregledi pomembni za zdravstveno stanje njihovih zob. Pomembno je tudi razumevanje pomena rednih pregledov v različnih letih življenja. To pomeni, da posameznike za zdravje svojih zob ne more skrbeti zgolj v trenutku, ko se pojavi zobobol, ampak da so pripravljeni resno skrbeti zanj skozi vsa leta svojega življenja.

Zavarovalnica mora čim bolj skrbeti za ohranjanje števila zavarovancev, predvsem zaradi morebitnih nepričakovanih izgub velikega števila zavarovancev, kot se je to npr. zgodilo v

času pojava bolezni COVID-19. Potrebe zavarovancev je potrebno natančno analizirati in pripraviti čim bolj ustrezne zavarovalniške produkte, ki bodo dosegli čim večje število ljudi.

Dostopnost izvajalcev storitev je pomembna za zavarovalnico, glede na to, da je njihov cilj doseči čim večje število zavarovancev. Na izbiro izvajalca vpliva prebivališče zavarovancev, zato je treba skrbeti in obnavljati mrežo izvajalcev zobozdravstvenih storitev, da bodo lahko zavarovanci čim bolj redno opravljali zobozdravstvene storitve.

Izboljšava analize in vizualizacije podatkov na področju panoge, kot je zavarovalništvo, je možna v večji meri in na različnih ravneh. Prihodnje raziskave bi se lahko osredinile na še boljše prepoznavanje potreb zavarovancev in posledično ustvarjanju novih produktov oziroma storitev, namenjenih zavarovancem. Zavarovalnica bi lahko npr. tudi ustvarila aplikacijo, s katero bi prepoznala potrebe zavarovancev, podatke pa bi lahko pridobivala preko anket ali vprašalnikov znotraj aplikacije po vsaki opravljeni storitvi. Tako bi zavarovalnica lahko svojim zavarovancem ponudila še bolj personalizirane produkte, prilagojene njihovim potrebam.

Izboljšave vidim tudi na področju dostopnosti zavarovancev do izvajalcev storitve. Zavarovalnica se je bolj osredotočila na večja mesta in pridobivanje več zavarovancev na tem območju, je pa očitno dana premajhna pozornost na zavarovance iz medkrajevnih mest. Ti nimajo velikega dostopa do izbire izvajalcev storitve, za izbiro večih izvajalcev pa porabijo več denarja. Menim, da je treba redno vzdrževati mrežo izvajalcev zdravstvene storitve in skrbeti za vse zavarovance, predvsem za te, ki so v premoženjski stiski ali težje dostopajo do večjih mest. Predlagam, da zavarovalnica ponuja dodatni popust sklenitve zdravstvenega zavarovanja za zavarovance, ki so dve leti zaporedoma redni zavarovanci in prebivajo v medkrajevnih mestih.

Možne izboljšave vidim tudi na področju zavarovancev, ki opravljajo storitve v različnih skupinah starosti. Kot je bilo vidno v analizi, so to v največji meri zavarovanci med 40 in 50 let. Zavarovalnica bi se lahko bolj osredotočila na osebe, mlajše od 40 let, kot tudi na osebe, starejše od 50 let. Za posameznike, mlajše od 40 let, predlagam aplikacijo, ki bo pripomogla k ustvarjanju rednih zaporednih pregledov in dodatnih storitev, če se zavarovanec redno udeležuje rednih pregledov. V aplikacijo je možno dodati npr. opomnik in ustvariti nadzorno ploščo, kjer bi vsak posameznik pogledal svojo uspešnost opravljanja rednih pregledov, in dodatne storitve kot bonus za uspešno opravljene preglede zavarovanja. Na drugi strani pa bi bilo dobro za uporabo aplikacije motivirati tudi starejše od 50 let.

6.3 Omejitve raziskave

Raziskava je bila narejena leta 2023. V analizi podatkov sem obravnavala eno izmed skupin zavarovancev, ki se nanaša na dopolnilno zdravstveno zavarovanje, ki ga od leta 2024 ni več v obravnavi zaradi določbe zavarovalnice. Kljub temu da dopolnilno zdravstveno zavarovanje ne obstaja več, je analiza dogajanja za podjetje še vedno lahko smiselna, saj

analiza obravnava tudi podatke, vezane na vrste storitve, ki jih zavarovalnica še vedno ponuja. Poleg tega izbrana zavarovalnica še vedno sodeluje z mrežo izvajalcev, s katerimi je sodelovala v letu 2023, zato sem analizo dogajanja v letu 2023 za podjetje še vedno dojemala kot smiselno, saj so nekateri uporabljeni podatki še vedno del redne ponudbe oz. njihovih produktov.

Omejitev raziskave je tudi, da so zavarovanci, ki so bili del te skupine, trenutno preskočili v Skupino 1 ali so prekinili svoje pogodbe za zavarovanja za naslednja leta. Predstavlja pa dobro podlago za nadaljnje analize z možnimi spremembami na podlagi podatkov, ki so še vedno veljavni.

Naslednja omejitev raziskave se nanaša na manjši obseg podatkov, zaradi česar je analiza bolj omejena, kot bi lahko bila. Če bi imela več podatkov o zavarovancih in njihovih podrobnostih, vezanih na opravljene redne preglede, bi lahko prišla do še bolj univerzalnih ugotovitev, ki bi koristile nadaljnjemu razvoju poslovanja oz. produktov.

7 SKLEP

Uporaba BI in vizualizacije podatkov igra veliko vlogo pri sprejemanju pomembnih odločitev, saj omogoča vpoglede v podatke oz. razkriva povezave med njimi, kar lahko vodi v kakovostnejše odločanje. Učinkovita vizualizacija podatkov omogoča pretvorbo podatkov iz nestrukturirane v bolj jasno in sprejemljivo obliko in tako omogoča uporabnikom bolj kakovosten pogled nad podatki.

V okviru magistrskega dela sem na primeru izbrane zavarovalnice prikazala, kako poteka razvoj BI nadzorne plošče in prikazala koristi, ki izhajajo iz uporabe različnih vizualizacij podatkov, kar podjetjem omogoča boljši pregled nad svojim poslovanjem in posledično sprejemanju boljših poslovnih odločitev.

Vizualizacija podatkov prinaša mnogo koristi, kar je razvidno tudi iz magistrskega dela. Pri tem velja izpostaviti številne vplive, od poenostavljanja postopkov analize podatkov do izboljšanja praks komunikacije in vpliv na proces odločanja. Vidno je, da se izbrana zavarovalnica sooča z velikimi težavami pri odločitvah glede uvajanja novih storitev oziroma produktov na področju zobozdravstvenih zavarovanj. Zato sta nujno potrebni analiza podatkov in jasna vizualizacija že vnaprej pripravljenih analiz. Vizualizacija podatkov, prikazana v tem magistrskem delu, predstavlja eno izmed možnosti, za nov način razmišljanja o potencialnih produktih in poenostavitve načinov odločanja. Omogoča tudi hitro prepoznavanje slabih odločitev na podlagi povezav med podatki, ki so jih lahko odločevalci prej spregledali.

V okviru magistrskega dela sem za primer izbrane zavarovalnice ugotovila, da sta se obe skupini zavarovancev, uporabljeni v analizi, zelo razlikovali med seboj. Zavarovanci Skupine 1 so se v četrtinskem deležu odločili za redno opravljanje zobozdravstvenega

pregleda, zavarovanci Skupine 2 pa so v veliki meri opravili redni zobozdravstveni pregled. Na njihovo odločitev za opravljanje rednih pregledov je v veliki meri vplivalo več dejavnikov, predvsem pa starost, spol, obdobje sklenitve zavarovalne pogodbe ter dostopnost do izvajalcev storitve. Če bi izbrana zavarovalnica želela višjo verjetnost, da se bodo zavarovanci odpravili na redne preglede, bi bilo koristno razmisliti o uvedbi novih tehnolog, kot je npr. mobilna aplikacija za uporabnike, ki bi preko različnih funkcionalnosti pripomogla k višjemu številu opravljenih rednih pregledov zavarovancev obeh skupin.

LITERATURA IN VIRI

1. Abdallah, Z. S., Du, L. in Webb, G. I. (2016). Data preparation. V C. Sammut in G. I. Webb (ur.), *Encyclopedia of machine learning and data mining* (str. 318–327). Humana Press.
2. Albreht, T., Pribaković Brinovec, R., Jošar, D., Poldrugovac, M., Kostnapfel, T., Zaletel, M., Panteli, D. in Maresso, A. (2016). Slovenia health system review. *Health Systems in Transition*, 18(3), 1–207.
3. Batini, C., Cappiello, C., Francalanci, C. in Maurino, A. (2009). Methodologies for data quality assessment and improvement. *ACM Computing Surveys*, 41(3), 1–52.
4. Bhargava, M. G., Phani Surya Kiran, K. T. in Rajeswara Rao, D. (2018). Analysis and design of visualization of educational institution database using power BI tool. *Global Journal of Computer Science and Technology*, 18(4), 1–8.
5. Clarkson, J. E., Pitts, N. B., Goulao, B., Boyers, D., Ramsay, C. R., Floate, R., Braid, H. J., Fee, P. A., Ord, F. S., Worthington, H. V., van der Pol, M., Young, L., Freeman, R., Gouick, J., Humphris, G. M., Mitchell, F. E., McDonald, A. M., Dt Norrie, J., Sim, K. ... Ricketts, D. (2020). Risk-based, 6-monthly and 24-monthly dental check-ups for adults: the INTERVAL three-arm RCT. *Health Technology Assessment*, 24(60), 1–138.
6. Dudycs, H. (2010). Visualisation methods in business intelligence systems – an overview. *Prace Naukowe Uniwersytetu Ekonomicznego we Wrocławiu. Informatyka Ekonomiczna*, 16(104), 9–24.
7. Eberendu, A. C. (2016). Unstructured data: an overview of the data od big data. *International journal of computer trends and technology (IJCTT)*, 38(1).
8. Fitriana, R., Eriyatno, T. D. in Djatna, T. (2011). Progress in business intelligence system research: a literature review. *International Journal of Basic & Applied Sciences IJBAS-IJENS*, 11(3), 96–105.
9. Foropon, C. in Ranjan, J. (2021). Big data analyitics in building the competitive intelligence of organisations. *Internal Journal of Information Management*, 56, 10223.
10. Ghazanfari, M., Jafari, M. in Rouhani, S. (2011). A tool to evaluate the business intelligence of enterprise systems. *Scientia Iranica*, 18(6), 1579–1590.
11. Ghazisaeidi, M., Safdari, R., Torabi, M., Mirzaee, M., Farzi, J. in Goodini, A. (2015). Development of performance dashboards in healthcare sector: key practical issues. *Acta Informatica Medica*, 23(5), 317–321.

12. Gonçalves, C. T., Gonçalves, M. J. A. in Campante, M. I. (2023). Developing integrated performance dashboards visualisations using power BI as a platform. *Information*, 14(614), 1–16.
13. Gupta, V. in Rathore, N. (2013). Deriving business intelligence from unstructured data. *International Journal of Information and Computation Technology*, 3(9), 971–976.
14. Hansonti, B. (2010). *Business intelligence dashboard in decision making*. College of Technology Directed Projects.
15. Hulslen, T. (2020). Sharing is caring – data sharing initiatives in healthcare. *International Journal of Environmental Research and Public Health*, 17(9), 1–12.
16. Izbrano podjetje. (2024). *Interno gradivo podjetja za leto 2024*. Ljubljana.
17. Janssen, M., Matheus, R. in Maheshwari, D. (2020). Data science empowering the public: Data driven dashboards for transparent and accountable decision-making in smart cities. *Government information quarterly*, 37(3), 101284.
18. Kay, E. J. (1999). How often should we go to the dentist? *BMJ*, 319(7204), 204–205.
19. Keeling, T. L. in Morris, R. (2010). Clinical research: using business intelligence framework. *Issues in Information Systems*, 11(1), 372-376.
20. Kimball, R. in Ross, M. (2013). *The data warehouse toolkit: the complete guide to dimensional modeling* (3. izd.). John Wiley & Sons.
21. Klopfenstein, D. in Howell, J. W. (2023, 21. julij). *Introduction*. <https://learn.microsoft.com/en-us/powerquery-m/m-spec-introduction>
22. Lanfranchi, D. in Grassi, L. (2022). Examining insurance companies' use of technology for innovation. *The Geneva Papers on Risk and Insurance Issues and Practice*, 47(3), 520–537.
23. Levin, M. A., Wanderer, J. P. in Ehrenfeld, J. M. (2015). Data, big data, and metadata in anesthesiology. *Anesthesia & Analgesia*, 121(6), 1661–1667.
24. Mahdavinejad M. S., Adibi, P., Barekatain, M., Barnaghi, P., Rezvan, M. in Sheth, A.P. (2018). Digital Communication and networks. *Machine learning for internet of things data analysis: a survey*, 4, 161-175.
25. Malik, S. (2005). Enterprise Dashboards. *Design and Best Practices for IT*, 1, 95-113.
26. Mughal Hamid, J. M. (2018). Data mining: web data mining techniques, tools and algorithms. *Internal journal of advanced computer science and applications*, 9(6), 208-215.
27. Myers, P. & Buck, A. (2022, 31. oktober). Learn. *Use variables to improve your DAX formulas*. <https://learn.microsoft.com/en-us/dax/best-practices/dax-variables>
28. Naneva, V. in Stefanova, K. (2022). Programming with DAX. *AIP Conference Proceedings*, 2505(1), 1–7.
29. Negash, S. (2004). Business intelligence. *Communications of the Association for Information Systems*, 13, 177–195.
30. Olszak, C. M. & Ziemba, E. (2007). *Approach to building and implementing business intelligence systems*, 2, 135-148.

31. Orlovskiy, D. in Kopp, A. (2020). A business intelligence dashboard design approach to improve data analytics and decision making. V *International conference information technology and interactions* (str. 48–59). ICITI.
32. Oussous, A., Benjelloun, F. Z., Ait Lahcen, A. in Belfkih, S. (2018). Big data technologies: a survey. *Journal of King Saud University - Computer and Information Sciences*, 30(4), 431–448.
33. Park, S., Bekemeier, B., Flaxman, A. in Schultz, M. (2021). Impact of data visualization on decision-making and its implications for public health practice: a systematic literature review. *Informatics for Health and Social Care*, 47(4), 1–19.
34. Pasha, M. R. (2016). Advanced data warehouse. *Data warehouse and the unstructured data*, 3-A.
35. Pearson, M., Knight, B., Knight, D. in Quintana, M. (2020). Extending Your data model with DAX calculations. V M. Pearson, B. Knight, D. Knight in M. Quintana (ur.), *Pro Microsoft power platform: solution building for the citizen developer* (str. 245–257). Apress.
36. Qin, X., Luo, Y., Tang, N. in Li, G. (2020). Making data visualization more efficient and effective: a survey. *The VLDB Journal*, 29, 93–117.
37. Riley, J. (2017). *Understanding metadata – what is metadata, and what is it for?* National Information Standards Organization.
38. Sadiku, M., Shadare, A. E., Shadare, A. E., Musa, S. M., Akujuobi, C. M. in Perry, R. (2016). Data visualisation. *International Journal of Engineering Research and Advanced Technology (ijerat)*, 2(12), 11-16.
39. Sen, A. (2004). Metadata management: past, present and future. *Decision Support Systems*, 37(1), 151–173.
40. Shao, C., Yang, Y., Juneja, S. in Gseetharam, T. (2022). Information processing & management. *IoT data visualization for business intelligence in corporate finance*, 59 (1), 102736.
41. Sousa, R., Miranda, R., Moreira, A., Alves, C., Lori, N. in Machado, J. (2021). Software tools for conducting real-time information processing and visualization in industry: an up-to-date review. *Applied Sciences*, 11(4800), 1–19.
42. Stodder, D. (2016). *Improving data preparation for business analytics: Applying technologies and methods for establishing trusted data assets for more productive users*. http://www.redpoint.net/wpcontent/uploads/2016/10/TDWI_BPReport_Q316_RedPoint_F_rev2_code_Final.pdf
43. Talaoui, Y. in Kohtamaki, M. (2021). 35 years of research on business intelligence process: a synthesis of a fragmented literature. *Management Research Review*, 44(5), 677–717.
44. Taleb, I., Kassabi, H. T. E., Serhani, M. A., Dssouli, R. in Bouhaddioui, C. (2016). Big data quality: a quality dimensions evaluation. V *2016 Intl IEEE conferences on ubiquitous intelligence & computing, advanced and trusted computing, scalable computing and communications, cloud and big data computing, internet of people, and smart world congress* (str. 759–765). IEEE.

45. Temitope Olubunmi, A. (2021). Interactive dashboard design for manager, data analyst and data scientist perspective, *11*(19), 175-185.
46. Tian, Y. (2017). *Accelerating data preparation for big data analytics*. <https://pastel.hal.science/tel-01791498/>
47. Tominski, C. in Schumann, H. (2020). Interactive visual data analysis. *Visualization methods and techniques*, *1*(3), 49-126.
48. van Elten, H. J., Sülz, S., van Raaij, E. M. in Wehrens, R. (2022). Big data health care innovations: performance dashboarding as a process of collective sensemaking. *Journal of Medical Internet Research*, *24*(2), e30201.
49. Widjaja, S. in Mauritsius, T. (2019). The development of performance dashboard visualisation with power BI as platform. *International Journal of Mechanical Engineering and Technology (IJMET)*, *10*(5), 235–249.
50. Wieder, B. in Ossimitz, M. L. (2015). The impact of business intelligence on the quality of decision making – a mediation model. *Procedia Computer Science*, *64*, 1163–1171.
51. Yang, K. (2024). *Quality in the era of industry 4.0: integrating tradition and innovation in the age of data and AI*. John Wiley & Sons.
52. Yaqoob, I., Abaker Targio, H. I., Gani, A., Mokhtar, S., Ahmed, E., Badrul Anuar, N. in Vasilakos, A. V. (2016). Big data: from beginning to future. *International Journal of Information Management*, *36*(6), 1231–1247.
53. Zardak, A. N., Amini-Rarani, M., Abdollahpour, I., Eslamipour, F. in Tahani, B. (2023). Utilization of dental care among adult populations: a scoping review of applied models. *BMC Oral Health*, *23*(1), 596.