

UNIVERZA V LJUBLJANI  
EKONOMSKA FAKULTETA

MAGISTRSKO DELO

**IZGRADNJA NAPREDNEGA SISTEMA POROČANJA KOT POMOČ  
PRI POSLOVNEM ODLOČANJU**

Ljubljana, marec 2023

EVA NOVAK

## IZJAVA O AVTORSTVU

Podpisana Novak Eva, študentka Ekonomske fakultete Univerze v Ljubljani, avtorica predloženega dela z naslovom Izgradnja naprednega sistema poročanja kot pomoč pri poslovnem odločanju, pripravljenega v sodelovanju s svetovalcem doc. dr. Urošem Godnovom

### 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 plagiatstvo – 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 plagiatstvo 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.

V Ljubljani, dne \_\_\_\_\_

Podpis študentke: \_\_\_\_\_

# KAZALO

UVOD .....	1
1 RAZISKAVA.....	2
1.1 Namen in cilj raziskave .....	2
1.2 Metodologija.....	2
1.3 Raziskovalna vprašanja .....	3
2 ANALIZA PODATKOV – DEL KLJUČNIH AKTIVNOSTI V PODJETJIH.....	4
2.1 Zakaj je analiza podatkov sploh pomembna.....	4
2.2 Orodja za izvajanje analize podatkov in njihove pomanjkljivosti.....	5
2.3 Prvi koraki k uspešni implementaciji analize podatkov v podjetju –kultura	10
2.4 Posledice pomanjkanja analize podatkov in poročanja .....	13
3 PODROČJA NAPREDNE ANALIZE PODATKOV .....	14
3.1 Načini zbiranja podatkov .....	15
3.2 Shranjevanje in dostop do podatkov.....	16
3.3 Obdelovanje podatkov.....	18
3.4 Uporaba DAX programskega jezika pri izgradnji ključnih kazalnikov uspešnosti.....	20
3.5 Vizualizacija podatkov .....	23
3.6 Avtomatizacija in prilagoditve .....	25
3.7 Napredni sistemi poročanja kot pomoč pri poslovnih odločitvah .....	27
3.8 Trendi uporabe analize podatkov v prihodnosti.....	29
4 PRIMER IZ PRAKSE .....	31
4.1 Pridobivanje podatkov preko vmesnika uporabniškega programa.....	31
4.2 Čiščenje in priprava podatkov v programu Power BI .....	32
4.3 Priprava kazalnikov v DAX-u .....	35
4.4 Vizualizacija podatkov .....	41
4.5 Dodeljevanje pravic in avtomatizacija.....	47
4.6 Ključne ugotovitve in priporočila.....	48
4.7 Možnosti za nadaljnji razvoj naprednega sistema poročanja .....	49
SKLEP .....	50
LITERATURA IN VIRI .....	51
PRILOGE.....	55

## KAZALO TABEL

Tabela 1: Povzetek orodij za izdelavo analiz podatkov .....	8
Tabela 2: Prikaz in razlaga različnih zapisov vrednosti .....	20

## KAZALO SLIK

Slika 1: Vpletenost strokovnjakov iz treh različnih področij v poslovanje podjetja.....	12
Slika 2: Formula za izračun marže v DAX-u .....	21
Slika 3: Formula za izračun seštevka vrednosti prodanih produktov.....	22
Slika 4: Formula za izračun prejšnjega obdobja .....	22
Slika 5: Osveževanje podatkov .....	26
Slika 6: Prednosti, ki jih je podjetjem prinesla uporaba analitike .....	28
Slika 7: Povezovanje do strežnika preko uporabniškega umesnika .....	32
Slika 8: Koraki pri čiščenju in pripravi podatkov v tabeli stranke.....	33
Slika 9: Čiščenje in priprava podatkov pri kreiranju nove tabele .....	34
Slika 10: Relacijski model v programu Power BI.....	34
Slika 11: DAX - Življenjska vrednost stranke .....	35
Slika 12: DAX - Nove stranke .....	36
Slika 13: DAX - Povprečna vrednost nakupa strank, ki se vračajo .....	36
Slika 14: DAX - Zaporedna številka naročila .....	37
Slika 15: DAX - Število dni med nakupi .....	37
Slika 16: DAX - Povprečno število nakupov na stranko .....	38
Slika 17: DAX - Povezave med kupljenimi produkti .....	38
Slika 18: DAX - Stopnja zaupanja .....	39
Slika 19: DAX - Verjetnost pojava kombinacij (lift).....	39
Slika 20: Model za predikcijo spola glede na ime stranke.....	40
Slika 21: Kreiranje dinamičnih vrednosti.....	40
Slika 22: Prva stran nadzorne plošče, s filtri in kazalom .....	41
Slika 23: Izbira vizualizacije - Razmerje med končnimi kupci in trgovci .....	42
Slika 24: Življenjska vrednost strank .....	42
Slika 25: Prikaz števila novih strank po državah .....	43
Slika 26: Dodajanje vrednosti v vizualizacijo - Povprečne vrednosti nakupov .....	43
Slika 27: Preimenovanje osi na grafih - Povprečen čas med nakupi .....	44
Slika 28: Povprečno število nakupov po državah in prikaz največjih kupcev .....	44
Slika 29: Produkti, ki jih kupci kupujejo skupaj .....	45
Slika 30: Urejanje naslova na vizualizaciji – Produkti, po katere se vračajo kupci.....	46
Slika 31: Dodajanje legende - Količina kupljenih produktov v določenem nakupu.....	46
Slika 32: Dinamične vizualizacije - Status naročila in vrsta plačila .....	47
Slika 33: Osveževanje in deljenje nadzorne plošče .....	48

## KAZALO PRILOG

Priloga 1: Vprašalnik za intervju z direktorjem izbranega podjetja ..... 1

## SEZNAM KRATIC

angl. – angleško

**AI** – (angl. Artificial Intelligence); Umetna inteligenca

**API** – (angl. Application Programming Interface); Umesnik uporabniškega programa

**BI** – (angl. Business Intelligence); Poslovna inteligenca

**CAO** – (angl. Chief Analytics Officer); Vodja analitike

**CDO** – (angl. Chief Data Officer); Vodja podatkov

**CSV** – (angl. Comma Separated Values); Vrednosti ločene z vejicami

**DAX** – (angl. Data Analysis Expressions)

**FTP** – (angl. File Transfer Protocol); Protokol za prenos datotek

**GE** – (angl. General Electric); General Electric

**HiPPO** – (angl. Highest paid person's opinion); Mnenje najboljše plačane osebe

**IBM** – (angl. Institute for Business Value); Inštitut za poslovno vrednost

**IDC** – (angl. International Data Corporation); International Data Corporation

**IoT** – (angl. Internet Of Things); Internet stvari

**JSON** – (angl. Java Script Object Notation); Objektna notacija za JavaScript

**KPI** – (angl. Key Performance Indicators); Ključni kazalniki uspeha

**ML** – (angl. Machine Learning); Strojno učenje

**NA** – (angl. not available); Ni na voljo

**NLP** – (angl. Natural Language Processing); Obdelava naravnega jezika

**PwC** – (angl. PricewaterhouseCoopers); PricewaterhouseCoopers

**SQL** – (angl. Structured Query Language); Strukturiran povpraševalni jezik za delo s podatkovnimi bazami

**t. i.** – tako imenovan

**TSV** – (angl. Tab Separated Values); Vrednost ločena z zavihki

**URL** – (angl. Uniform Resource Locator); Enolični krajevnik vira

**XML** – (angl. Extensive Markup Language); Razširljiv označevalni jezik



## UVOD

V začetku 20. stoletja je za najdragocenejšo dobrino veljala nafta, ki je izbranim državam omogočila naraščajoče dobičke in spremenila razmerje moči med svetovnimi državami (History.com Editors, 2018). Danes pa si marsikdo upa trditi, da so naftni industriji ob bok stopile korporacije, ki že desetletja zbirajo in analizirajo podatke svojih uporabnikov. Google, Amazon, Apple, Facebook in Microsoft so ena izmed najvrednejših podjetij na svetu in najbolj poznani zbiratelji podatkov (The Economist Newspaper Limited, 2017).

Ko govorimo o podatkovno vodenih podjetjih, ne govorimo samo o tem, da ima podjetje velike količine podatkov, vendar o celotnem procesu vzpostavljanja učinkovite in dobro zakoreninjene kulture podjetja. Tovrstna podjetja vzpostavljajo napovedne modele, ki na primer pomagajo optimizirati hitrost oglasov in oskrbovalne verige ali pa zmanjšajo odlive strank. Z njihovo pomočjo zaposleni generirajo priporočila in napovedi ter ustvarijo zgodbo iz končnih analiz, ki odločevalcem pomaga pri sprejemanju odločitev (Anderson, 2015). Postati podatkovno vodeno podjetje pa vsekakor ni lahko, saj proces že v začetnih fazah zahteva dobro načrtovano zbiranje podatkov, ki omogočijo kvalitetno izvajanje analize in posledično njeno pravilno razumevanje. V magistrskem delu bosta podrobneje razloženi problematiki zadnjih dveh korakov, saj je potrebno poudariti, da podjetje ni podatkovno vodeno, če se rezultati analiz ne upoštevajo pri odločitvah in analize niso usmerjene dolgoročno. Brynjolfsson, Hitt in Kim (2011) so v svojem raziskovalnem delu ugotovili, da so podatkovno vodena podjetja za pet do šest odstotkov bolj produktivna kot tista, ki to niso. Prav tako bolje izkoriščajo svoje vire, imajo višji donos na kapital in višjo tržno vrednost. Prepogosto pa prihaja tudi do tega, da podjetja sicer izvajajo analitiko, vendar večini ne uspe postati podatkovno vodeno podjetje v celotnem pomenu besede. Večina ima težave pri vzpostavljanju kulture, ki temelji na podatkih. Tu nazadujejo predvsem večja podjetja, ki so prepočasna pri prilagajanju in omenjene kulture niso vzpostavila že od samega začetka poslovanja, kot to na primer opazimo pri zagonskih (angl. start-up) podjetjih (Davenport & Bean, 2018). Ključni za spodbujanje kulture, ki temelji na podatkih, so generalni direktorji. Slednji morajo spodbujati redne informativne pogovore, omogočiti delavnice in dodatna izobraževanja za najvišje odločevalce in vodje podatkovnih baz v podjetju (Diaz, Rowshankish & Seleh, 2018). Poleg tega vse večjo težavo predstavlja zaposlovanje podatkovnih znanstvenikov (angl. data scientists). V zadnjih letih je povpraševanje močno preseгло ponudbo, zato so marsikatera podjetja pripravljena ponuditi visoke plače za vrhunske talente (Maffeo, 2020).

Ob prebiranju literature in obstoječih raziskav lahko opazimo, da imajo vse težave skupni imenovalec. Podjetja preprosto nimajo dovolj znanja o naprednih sistemih poročanja, da bi v njih lahko videla dodano vrednost. Pomanjkanje razumevanja podatkov in percepcija, da so podatki zgolj skrb informacijskega oddelka, ne pa dragoceno sredstvo, ki daleč presega tehnične analize, so le nekateri izmed argumentov, zakaj je o tem potrebno govoriti. Že samo preprosta postavitev naprednega sistema poročanja s pomočjo programa R in Power BI lahko vzbudi navdušenje v zaposlenih in pomaga pri prvih korakih k podatkovno vodenemu

podjetju. Do takrat pa lahko sledimo sporočilu vodje znanosti o korporativnem odločanju v NBCUniversal, Cameron Daviesu: »Ne gre za podatke same. Ne gre samo za analitiko – kot jemanje vitaminov, samo zato, da lahko trdite, da ste vsako jutro uspešno vzeli tableto. Ko gre za analitiko, se moramo zavedati, da je končni cilj pomagati sprejemati boljše odločitve, bolj pogosto.« (Diaz, Kayvaun, & Seleh, 2018).

# **1 RAZISKAVA**

## **1.1 Namen in cilj raziskave**

V prvem delu magistrskega dela bo predstavljeno teoretično ozadje naprednih sistemov poročanja. Namen je bralca seznaniti z osnovnimi pojmi kot so množični podatki, napreden sistem poročanja in vizualizacija podatkov. S pomočjo znanstvene literature bo predstavljen pomen zbiranja podatkov, njihovega obdelovanja in pomembnost vključevanja naprednega sistema poročanja v vsakdanje procese organizacij. Poudarek bo predvsem na vključevanju tovrstnih sistemov v proces odločanja. Prav tako bo s pomočjo računalniške dokumentacije in druge literature predstavljeno orodje Power BI. V drugem, empiričnem delu, pa bo teoretično znanje aplicirano na primer podatkov izbranega podjetja. Sistematičen prikaz korakov pri postavitvi naprednega sistema za poročanje bo lahko v pomoč tako študentom kot tudi kateremkoli podjetju, ki se želi seznaniti s tem, kakšni so prvi koraki pri obdelavi podatkov. Delo bo prispevalo k boljšemu razumevanju poslovanja v izbranem podjetju in pomagalo pri poslovnih odločitvah na različnih področjih, predvsem v marketingu in razvoju.

Cilj magistrskega dela je s pomočjo literature in aplikacije pridobljenega znanja na primeru izbranega podjetja preučiti razsežnosti naprednega sistema poročanja (v tem primeru programa Power BI) in predstaviti pomembnost vključevanja tovrstnih sistemov v procese odločanja.

## **1.2 Metodologija**

Teoretičen del magistrskega dela bo temeljil na pregledu izbrane literature, ki bo zajemala tako znanstvene dokumentacije, kot tudi dokumentacije izbranih računalniških programov. Obravnaval bo tematiko načina postavitve naprednega sistema poročanja, ki je v pomoč podjetjem pri sprejemanju poslovnih odločitev. Opisna oziroma deskriptivna metoda bo pomagala pri analizi obstoječih raziskav in druge literature, komparativna metoda pa bo uporabljena za primerjavo ugotovitev različnih avtorjev ter odkrivanje podobnosti oziroma različnosti ugotovitev. V prvem poglavju bodo predstavljena štiri področja, ki pomembno vplivajo na uspešnost uporabe podatkov v podjetju. Predstavljena bo pomembnost analiz, orodja za njihovo izvajanje, podatkovna kultura in posledice pomankanja poročanja. V nadaljevanju pa bodo predstavljeni koraki za uspešno implementacijo naprednega sistema



poročanja v vsakdanje aktivnosti podjetja. Podrobneje bodo opisani načini zbiranja podatkov, njihovo shranjevanje in obdelovanje, uporaba jezika DAX (angl. Data analysis expressions), možnosti vizualiziranja podatkov ter avtomatizacija in prilagoditve. Ob koncu poglavja bomo predstavili kako tovrstni napredni sistemi poročanj pomagajo pri sprejemanju poslovnih odločitev in kakšni so trendi za prihodnost.

V empiričnem delu bo predstavljen napreden sistem poročanja, na primeru izbranega podjetja z množičnimi podatki. Podjetje je proizvajalec in prodajalec kozmetičnih izdelkov. Kljub svoji majhnosti svoje izdelke uspešno prodaja tako v Evropi, kot tudi v Ameriki. Že od samega začetka podjetje zbira podatke svojih kupcev in je v želji po izboljšanju svojih produktov izpostavilo 11 ključnih kazalnikov uspešnosti. Interni podatki podjetja bodo tako analizirani v programu Power BI, kjer bodo izračunani tudi omenjeni kazalniki. Kazalniki se nanašajo na nakupne navade obstoječih strank, vrednosti nakupov, nove stranke, asociacije med produkti, vračila in plačilna sredstva. Do podatkov bomo dostopali preko vmesnika uporabniškega pograma (angl. application programming interface, v nadaljevanju API) in jih nato v programu Power BI pripravili za nadaljnje analize. Omenjene kazalnike bomo izračunali s pomočjo DAX-a, ki se uporablja za ustvarjanje formul in izrazov. Z njegovo pomočjo podatke lahko filtriramo in združujemo na različne načine. Poleg tega pa omogoča tudi izvajanje kompleksnejših izračunov, ki jih bomo potrebovali za nekatere kazalnike.

Ključne ugotovitve in priporočila iz empiričnega dela bodo predstavljene skupaj z ugotovitvami iz literature, kjer bomo predstavili tudi ugotovitve direktorja izbranega podjetja. S slednjim bo narejen intervju o uporabnosti naprednega sistema poročanja v njihovem vsakdanjem poslovanju, podatkovni kulturi podjetja, izzivi s katerimi se spopadajo in nadaljnjem razvoju sistema. V sklepnem delu bodo povzete ključne ugotovitve in omejitve raziskave.

### **1.3 Raziskovalna vprašanja**

Z odgovori na sledeča raziskovalna vprašanja želimo pomagati bralcem razumeti, zakaj so napredni sistemi poročanja pomembni, kako lahko pomagajo pri poslovanju podjetji in kaj lahko pričakujemo v prihodnosti.

**Raziskovalno vprašanje 1:** Ali napreden sistem poročanja pripomore k sprejemanju boljših poslovnih odločitev?

**Raziskovalno vprašanje 2:** Kako lahko s pomočjo naprednega sistema poročanja predvidimo prihodnost poslovanja?

**Raziskovalno vprašanje 3:** Kako napreden sistem poročanja rešuje težave v zvezi s preveliko količino podatkov, ne-aktualnimi analizami in nerazumevanje podatkov?

## **2 ANALIZA PODATKOV – DEL KLJUČNIH AKTIVNOSTI V PODJETJIH**

Količina podatkov, ki jih proizvedejo tako podjetja kot tudi posamezniki, vztrajno raste. Če sta še nekaj desetletij nazaj veljali intuicija in izkušnje za najbolj zanesljivi veščini zaposlenih na najvišjih položajih, to mesto danes pripada analitiki. Vse več podjetij se močno zanaša na množične podatke in analitiko. Potreba po podatkovno vodenih podjetjih in posledično sovpadajoči kulturi postaja vse prej kot le modna muha (Bean & Davenport, 2019). Število orodij in programov, s katerimi je moč analizirati podatke, postaja vse večje, s tem pa se porajajo vprašanja o primernosti orodij za različne analize.

### **2.1 Zakaj je analiza podatkov sploh pomembna**

Sprejemanje odločitev je ena izmed glavnih nalog generalnih direktorjev, ki pa ni vedno lahka. Sodeč po dosedanjih raziskavah se odločevalci zanašajo tako na svojo intuicijo kot tudi na podatkovne analize (Mikuškova, 2017). Podjetja, ki v svoje poslovanje že vključujejo analize, beležijo izboljšane storitve za stranke, nižje stroške poslovanja (Davenport & Bean, 2018), boljše razumevanje trga, konkurenčno prednost, bolj natančno targetiranje strank, lažje optimiziranje svojega poslovanja (Bughin, 2016), bolj natančno razporeditev sredstev, rast prodaje, povečan tržni delež, povečano finančno uspešnost, večjo donosnost sredstev in krajši obrat zalog (Mihai & Borza, 2019). V študiji, ki je bila izvedena v 176 velikih ameriških javnih podjetjih, je bilo ugotovljeno, da imajo podjetja, ki za svoje odločitve uporabljajo podatkovne analize, od pet do šest odstotno povečanje produktivnosti v obdobju petih let. Prav tako je bilo dokazano, da obstajajo povezave med podjetji, ki uporabljajo podatkovne analize v procesu odločanja, in tržno vrednostjo ter nekaterimi kazalniki dobičkonosnosti, kot sta donosnost kapitala in izkoriščenost sredstev (Brynjolfsson, Hitt & Kim, 2011). Z analizo podatkov tako ne dobimo samo vpogleda v preteklo poslovanje, ampak tudi lažje odkrijemo vzroke določenih težav, mesto njihovega nastanka in morebitne rešitve. Ko pa stopimo stopničko višje, nam lahko zbrani podatki pomagajo pri predvidevanju prihodnjih trendov in optimizaciji poslovanja. Da pa se lahko povzpne na to raven, moramo najprej podatke spremeniti v informacije, da dobijo pomen, in jih nato umestiti v celoten kontekst. Zadnji korak nam tako omogoči, da dobljeno znanje pravilno apliciramo v vsakdanje odločitve.

Študija, ki jo je med 3000 generalnimi direktorji naredila organizacija IBM (Anderson, 2015), je razkrila, da bodo podjetja, ki so se označila za bistveno boljša od ostalih v panogi, petkrat verjetneje uporabljala analitiko na različnih področjih svojega poslovanja. Poleg tega pa bodo tudi:

- trikrat verjetneje postala napreden uporabnik analitike,
- dvakrat verjetneje uporabljala analitiko kot pomoč pri vsakodnevem poslovanju,
- dvakrat verjetneje uporabljala analitiko kot usmeritev za razvoj strategije.

Avtorji študije so glede na izsledke tako določili tri nivoje analitične zmogljivosti, in sicer ambiciozna, izkušena in preoblikovana podjetja. Zadnja se nahajajo na najvišjem nivoju, kjer analitiko uporabljajo na vseh področjih poslovanja (upravljanje s tveganji, skrb za stranke, planiranje, vodenje podjetja, upravljanje blagovne znamke in trženje). Anderson (2015) je ugotovil, da imajo v primerjavi z ambicioznimi podjetji preoblikovana podjetja:

- štirikrat večjo verjetnost, da bodo zelo dobro zajemala informacije,
- devetkrat večjo verjetnost, da bodo zelo dobro združila informacije,
- osemkrat večjo verjetnost, da bodo zelo dobro analizirala informacije,
- desetkrat večjo verjetnost za razširitev informacij in spoznanj,
- 63 % večjo verjetnost, da bodo uporabljala centraliziran sistem analitike kot primarni vir analitike v svoji organizaciji.

Naslednji vidik, ki ga lahko omenimo, je strateške narave. Napredni sistemi poročanja nam na podlagi preteklih analiz in podatkov, pridobljenih iz zunanjih virov, pomagajo pri boljšem razumevanju kupcev. Brazilsko podjetje, ki ga omenja McKinsey v poročilu o pomembnosti napredne analitike pri upravljanju prihodkov, je s pomočjo naprednega sistema poročanja definiralo optimalne cene za določene produkte na specifičnih trgih. Le nekaj manjših popravkov v cenah je rezultiralo v 2,5-odstotnem povišanju donosnosti od prodaje v samo nekaj tednih. S tem povezana dejavnost marsikaterega podjetja je tudi oglaševanje. Z uspešno implementacijo analitike lahko analize povežemo z zunanjimi podatki, kot so vreme, oglaševanje konkurentov in variiranje cen substitutov. Na tovrsten način lahko natančneje definiramo oglaševalski proračun, ciljne skupine kupcev, oglaševalske kanale in oglaševane produkte oziroma storitve. Zadnja komponenta strateške narave, ki jo bomo omenili, se osredotoča na skladiščenje. Ko definiramo, kateri trgovci naročujejo bolj ali manj pogosto, katere izdelke in koliko ter kateri produkti vplivajo na povečano prodajo drugih produktov, tudi skladiščenje postane del naprednih sistemov poročanja. Predvsem pa poveča razpoložljivost polic, optimizira postavitev produktov in poveča natančnost planiranja. Če povzamemo primer iz lepote industrije, kjer se je podjetje odločilo za ukinitve določenega produkta, vidimo, da je to povzročilo upad prodaje še dveh drugih izdelkov. Zaradi njihove komplementarnosti so ti produkti zasedali isto skladiščno mesto, kar pa se je z ukinitvijo enega izdelka seveda spremenilo (Breuer, Elliott & Rickert, 2013). Bill Ruh, podpredsednik za programsko opremo pri General Electric (GE), je o svoji izkušnji z analitiko povedal: »Ugotavljamo, da več podatkov, kot jih imamo, bolj smo inovativni v analitiki in začnemo delati stvari, za katere nismo mislili, da jih lahko.« (Economist Intelligence Unit, 2012).

## **2.2 Orodja za izvajanje analize podatkov in njihove pomanjkljivosti**

Orodja za izvajanje analize podatkov so ena glavnih komponent v procesu izvajanja analiz množičnih podatkov. Z njimi se srečujemo od samega začetka, ko podatke zbiramo, in vse do zadnjega koraka, ko pripravimo vizualizacije (Kumar, Aruna & Suchethana, 2022). V

nadaljevanju bomo predstavili najbolj poznana orodja, ki se uporabljajo v procesu izgradnje naprednih sistemov poročanja. Proces navadno vključuje štiri glavne korake:

- priprava podatkov,
- raziskovanje in analiziranje podatkov,
- vizualizacija podatkov,
- izgradnja nadzorne plošče poslovnega obveščanja.

Za pripravo podatkov lahko uporabimo orodja, ki so v večji meri namenjena pripravi in čiščenju podatkov. Poznamo pa tudi orodja, ki imajo poleg omenjene funkcije tudi funkcije analiziranja in vizualizacije podatkov. Verjetno najbolj poznan in najpogosteje uporabljen **Excel** je odlična izbira, ko imamo manjšo količino podatkov, ki so že dobro prečiščeni. Na ta način lahko izvedemo lažje analize s pomočjo že vgrajenih funkcij. Verjetno ena bolj poznanih je tudi funkcija VLOOKUP, ki omogoča združevanje dveh tabel. Pri uporabi te funkcije pa moramo biti še posebej pozorni, saj je dovzetna za napake in nepravilnosti (Mannerings, 2021). Na drugi strani je Excel zelo priročen, kadar ga je potrebno deliti in uporabljati z drugimi uporabniki, ki drugih specializiranih programskih oprem ne poznajo. Prav tako omogoča enostavno shranjevanje v druge oblike zapisa, kot so na primer .xls, .txt, .csv ali .pdf. Excel kljub omenjenim prednostim ni vedno najboljša izbira. Ima namreč omejeno število vrstic (1.048.576) (Microsoft Corporation, 2022b), zato pri delu z veliko količino podatkov ni primeren. Prav tako nima naprednih orodij, ki bi omogočala čiščenje in pripravo podatkov, vsebuje le osnovne vizualizacije in ne omogoča naprednih statističnih analiz. Alternativno orodje, ki ima tako rekoč enake funkcije, je Google Sheets, kjer pa je pomembno omeniti ključno razliko – omejitev ni podana v številu vrstic, ampak v številu celic. Slednja ne more presegati deset milijonov celic oziroma 18.278 stolpcev (Google, 2022).

Sledeča orodja, ki jih bomo omenili, temeljijo na različnih programskih jezikih, zato je njihova uporaba omejena na uporabnike s specifičnim znanjem. Strukturiran povpraševalni jezik za delo s podatkovnimi bazami (angl. Structured Query Language, v nadaljevanju **SQL**) je programski jezik, ki je zasnovan za delo z (velikimi) bazami podatkov. Pri analizah se uporablja za izvajanje kompleksnih poizvedb in združevanje velikih naborov podatkov. Ker ne zahteva veliko znanja programiranja, omogoča izvajanje osnovnih funkcij tudi tistim, ki s programiranjem nimajo izkušenj. Filtriranje in obdelava podatkov na različne načine sta zagotovo dve izmed glavnih prednosti omenjenega programa. Prav tako omogoča združevanje več tabel iz različnih virov, njihovo urejanje in združevanje podatkov za namen statističnih povzetkov ali pa vrtilnih tabel. Pomembno je tudi poudariti, da z njim lahko pridobimo le podatke, ki jih potrebujemo za določeno analizo, kar predstavlja večjo učinkovitost pri nadaljnjih operacijah. SQL nam torej omogoča na razmeroma preprost način pridobiti in obdelovati velike količine podatkov, poleg tega pa je sposoben tudi nekoliko kompleksnejših analiz (Yildirim, 2021). Ker SQL ni imel toliko fleksibilnosti kot programska jezika Python in R, ki ju bomo omenili v nadaljevanju, je Microsoft v novejših verzijah omogočil izvajanje njunih skript znotraj SQL Serverja. Storitve je tako poenostavila

izvajanje določenih vrst analiz in prilagajanje specifičnim potrebam. Nekatere preproste funkcije, kot je na primer povprečje, niso vključene, zato jih je potrebno izračunati, kar pa nam lahko povzroči veliko težav (Peck, 2020). V kolikor imamo podatke že dobro pripravljene, bo za čiščenje in obdelovanje SQL popolnoma zadoščal. V nasprotnem primeru pa bodo za lažje delo primernejši drugi programski jeziki ali pa specializirana orodja. Tudi SQL, tako kot Excel, generira osnovne vizualizacije, medtem ko so za bolj kompleksne primernejši drugi programi. Programski jezik, ki se prav tako veliko uporablja za analizo podatkov, je jezik **R**. Predvsem je specializiran za obdelavo velikih količin podatkov in statistične analize. Zaradi svoje prepoznavnosti in široke uporabe je dobro poznan med statistiki in analitiki, ki svoje znanje delijo na temu namenjenih skupnostih. Uporabniki lahko prispevajo tudi k razvoju in izboljšavam, saj je R odprtokoden program. Njegova velika prednost je zagotovo v tem, da je primeren za delo z velikimi količinami podatkov in kompleksnimi podatki, saj ima ogromen nabor funkcij in knjižnice za obdelavo. Že osnovne knjižnice vključujejo funkcije za uvoz, čiščenje, preoblikovanje in povzemanje podatkov. Je odličen program za statistične analize, saj vključuje napredne funkcije, kot so statistični testi, modelirne funkcije in algoritmi za strojno učenje. Podpira datoteke, ki so shranjene v različnih formatih, omogoča iterativno raziskovanje in analizo podatkov ter sprotno spreminjanje analize. Tako kot že prej omenjena programa ima tudi R možnost vizualizacij. Slednje lahko nekoliko bolj prilagajamo svojim potrebam, vendar še vedno nima zmogljivosti za popolno prilagoditev, kot na primer Tableau, ki ga bomo omenili v nadaljevanju. Ker je sintaksa precej kompleksna, lahko vzame nekoliko več časa pri učenju in zahteva dobro razumevanje konceptov programiranja.

Podjetja v namen poslovnega obveščanja (angl. Business intelligence) uporabljajo različne portale in nadzorne plošče, ki se lahko prilagodijo potrebam uporabnika. Pogosto pa so nadzorne plošče razvite brez predhodnih nastavitvev, kar omogoči popolno prilagoditev potrebam. Uporabnik lahko tako vanje vključi različne filtre in ključne kazalnike uspešnosti, ki so specifični za njegov posel (RSM US, 2020). V nadaljevanju bomo na kratko predstavili Tableau in Power BI, ki se v podjetjih vse pogosteje pojavljata kot orodji poslovnega obveščanja.

Programska oprema, specializirana za vizualizacijo in poslovno inteligenco, **Tableau**, omogoča ustvarjanje interaktivnih nadzornih plošč z grafi in analizami podatkov. Zasnovan je tako, da je uporabnikom prijazen in ne zahteva posebnega znanja programiranja (Tableau Software, 2022). Tableau lahko pridobi podatke iz različnih virov in nudi širok izbor orodij in lastnosti za pripravo podatkov, analize in vizualizacije. Omogoča enostavno deljenje nadzorne plošče preko Tableau serverja, mobilnega telefona, spletne strani ali PDF dokumenta. Njegova največja prednost, v primerjavi s prej omenjenimi programi, pa so zagotovo grafične zmožnosti. Uporabniku dovoljujejo ustvarjanje visoko kakovostnih in popolnoma prilagodljivih vizualizacij. Tableau pa ni najprimernejši za delo z zelo velikimi količinami podatkov in kompleksnimi podatki. Prav tako ne nudi toliko fleksibilnosti pri prilagajanju in avtomatizaciji ter ne nudi toliko možnosti pri čiščenju in obdelovanju

podatkov kot prej omenjena programska jezika. **Power BI** je orodje, namenjeno poslovni inteligenci, ki omogoča delo z množičnimi podatki, tako na ravni obdelave podatkov kot tudi vizualizacij. Zasnovan je tako, da uporabnikom pomaga pridobiti vpogled v njihovo poslovanje s pomočjo popolnoma prilagodljivih nadzornih plošč in jim posledično pomaga sprejemati boljše ter hitreje odločitve. Power BI omogoča povezovanje s številnimi viri podatkov in programskimi opremami, kot sta tudi že prej omenjena SQL in R. Vključuje tudi funkcije, ki omogočajo pretakanje podatkov v realnem času, procesiranje naravnega jezika in deljenje ter souporabo poročil preko spletne verzije ali mobilnega telefona. Slednje omogoča upravljanje dovoljenj za različne uporabnike in deljenje tako znotraj kot tudi zunaj organizacije (Microsoft Corporation, 2022c). Za namen kreiranja poročil ne zahteva posebnega tehničnega znanja, kljub temu pa omogoča veliko možnosti za prilagoditev tako vizualizacij kot tudi poročila v celoti. Za uporabo vseh funkcij je potrebno kupiti plačljivo verzijo, ki pa lahko za mala podjetja predstavlja velik strošek. Njegova uporaba lahko postane zahtevnejša ob analizah velikih količin podatkov, kjer je potrebno dodatno znanje jezika DAX., razumevanje vrste podatkov in njihovih razmerij, napredno filtriranje, združevanje podatkov in odpravljanje napak (angl. debugging). K Power BI-ju se bomo vrnili v tretjem poglavju, kjer si bomo podrobneje ogledali korake za postavitve naprednega sistema poročanja na primeru izbranega podjetja. Mnoga podjetja se v naslednjem koraku odločijo tudi za integracijo nadzorne plošče poslovnega obveščanja v svoj informacijski sistem, t. i. Intranet. Omenimo lahko SharePoint, ki poleg klasičnih lastnosti Intraneta nudi tudi možnosti poročanja, analitike in dostave informacij na povsem varen način (RSM US, 2020). Spodnja tabela 1 prikazuje povzetek omenjenih orodij. Prednosti, slabosti, zahtevnost uporabe in obdelovanje določene količine podatkov so ocenjeni za namen analize množičnih podatkov in glede na povprečnega uporabnika, ki nima posebnih znanj programiranja.

*Tabela 1: Povzetek orodij za izdelavo analiz podatkov*

	<b>Prednosti</b>	<b>Slabosti</b>	<b>Zahtevnost uporabe</b>
Excel	<ul style="list-style-type: none"> <li>• Integrirane funkcije</li> <li>• Enostavno deljenje in souporaba</li> <li>• Enostavno shranjevanje v drugih oblikah</li> </ul>	<ul style="list-style-type: none"> <li>• Omejena količina podatkov</li> <li>• Pomanjkanje naprednih orodij za čiščenje in pripravo podatkov</li> <li>• Omejena prilagoditev vizualizacij</li> <li>• Ne omogoča naprednih statističnih analiz</li> </ul>	Nizka

se nadaljuje

Tabela 1: Povzetek orodij za izdelavo analiz podatkov (nad.)

	<b>Prednosti</b>	<b>Slabosti</b>	<b>Zahtevnost uporabe</b>
SQL	<ul style="list-style-type: none"> <li>• Zasnovan za množične podatke</li> <li>• Osnovne funkcije enostavne tudi za začetnike</li> <li>• Veliko možnosti filtriranja in obdelovanja podatkov</li> <li>• Združevanje več tabel iz različnih virov</li> <li>• Omogoča kompleksnejše analize</li> </ul>	<ul style="list-style-type: none"> <li>• Manj fleksibilen za specifične potrebe</li> <li>• Omejen pri čiščenju in pripravi podatkov</li> <li>• Omejena prilagoditev vizualizacij</li> </ul>	Zahtevna
R	<ul style="list-style-type: none"> <li>• Zasnovan za množične podatke</li> <li>• Omogoča statistične analize</li> <li>• Ima veliko skupnost</li> <li>• Odprtokoden program</li> <li>• Ogromen nabor funkcij in knjižnic</li> <li>• Vključuje napredne funkcije in algoritme</li> <li>• Sprotno spreminjanje analiz</li> </ul>	<ul style="list-style-type: none"> <li>• Omejena prilagoditev vizualizacij</li> <li>• Zahteva dobro razumevanje konceptov programiranja</li> <li>• Zahteva več časa pri prvih korakih učenja</li> </ul>	Zahtevna
Tableau	<ul style="list-style-type: none"> <li>• Ustvarjanje interaktivnih nadzornih plošč</li> <li>• Uporabniku prijazen</li> <li>• Pridobivanje podatkov iz različnih virov</li> <li>• Širok izbor orodij</li> <li>• Enostavno deljenje</li> <li>• Popolna prilagodljivost vizualizacij</li> </ul>	<ul style="list-style-type: none"> <li>• Ni primeren za delo s kompleksnejšimi podatki</li> <li>• Manj fleksibilen pri prilagajanju in avtomatizaciji</li> <li>• Malo možnosti za čiščenje in obdelavo podatkov</li> </ul>	Nizka
Power BI	<ul style="list-style-type: none"> <li>• Zasnovan za množične podatke</li> <li>• Popolnoma prilagodljiva nadzorna plošča</li> <li>• Primeren za sprejemanje vsakodnevnih odločitev</li> <li>• Povezovanje z različnimi viri podatkov</li> <li>• Pretakanje podatkov v realnem času</li> <li>• Enostavno deljenje in souporaba</li> <li>• Ne zahteva veliko tehničnega znanja</li> </ul>	<ul style="list-style-type: none"> <li>• Plačljiva spletna verzija za souporabo in deljenje</li> <li>• Dodatno znanje jezika DAX za analize in obdelavo podatkov</li> <li>• Razumevanje vrst podatkov in razmerij</li> </ul>	Srednja

Vir: Povzeto po Microsoft Corporation (2022b), Google (2022), Yildirim (2021), Peck (2020), Tableau Software (2022), Microsoft Corporation (2022c).

### 2.3 Prvi koraki k uspešni implementaciji analize podatkov v podjetju – kultura

Mnoga podjetja se zavedajo pomembnosti zbiranja in analiziranja podatkov ter prednosti, ki jih slednje prinaša. Kljub temu pa ne uspejo uspešno implementirati naprednih sistemov obveščanja v svoje vsakdanje poslovanje. Mnoga podjetja investirajo ogromno sredstev v podatkovno infrastrukturo, programsko opremo, strokovna znanja in podatkovne modele. Velikokrat pa pozabijo, da je potrebno uskladiti način dela in uvesti podatkovno kulturo, ki v veliki meri podpira uspešno implementacijo naprednih sistemov poročanja na ravni celotnega podjetja in ne le posameznih projektov (Fountaine, McCarthy & Saleh, 2019).

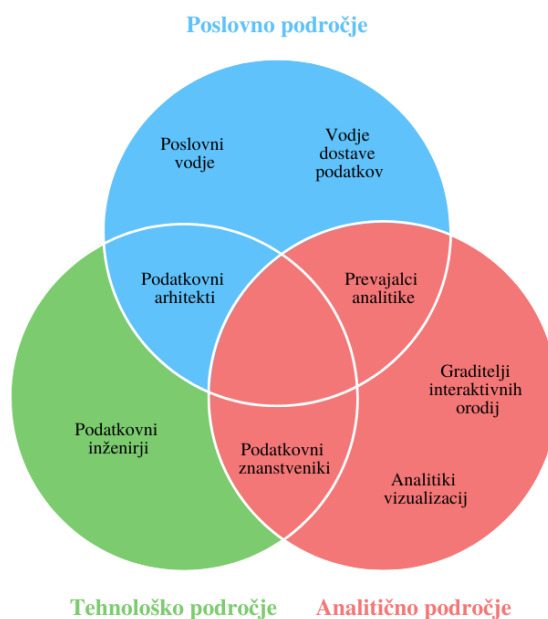
V raziskavi, ki je bila narejena med 64 vodji tehnologije in direktorji v velikih podjetjih, kot so American Express, Ford Motors, General Electric in drugi, so našli zaskrbljujoče rezultate. Kar 72 % sodelujočih je odgovorilo, da še niso vzpostavili podatkovne kulture, 69 %, da se še ne tretirajo kot podatkovno vodeno podjetje in 53 %, da podatkov še ne obravnavajo kot poslovno sredstvo. Ko so se vprašanja dotaknila kulturnega vidika, je kar 93 % sodelujočih odgovorilo, da so zaposleni ena izmed glavnih težav pri vzpostavitvi podatkovno vodenega podjetja (Bean & Davenport, 2019). Podjetja se velikokrat soočajo s tako imenovanim pojavom HiPPO (angl. Highest paid person's opinion), mnenje najbolje plačane osebe v podjetju. Navadno so to strokovnjaki z dolgoletnimi izkušnjami, ki jim ni mar, kaj pravijo podatki, saj imajo jasno začrtane načrte za prihodnost. Poleg tega pa so navadno tudi direktorji ali zaposleni na visokih položajih. Financial Times jih imenuje uničevalci posla, saj svoje odločitve sprejemajo na bazi nerazumljivih meritev ali ugibanj (Anderson, 2015). V nadaljevanju si bomo pogledali, kateri so ključni elementi pri izgradnji kulture podatkovno vodenega podjetja.

V prvi fazi postavljanja kulture je pomemben pristop vodilnih v podjetju. Spodbuda mora priti iz višjih položajev, ki aktivno spremljajo in podpirajo podatkovno vodene projekte. Najprej morajo prepoznati potrebo za strukturiranje podatkov in analitiko ter zagotoviti potrebna sredstva. Prepoznavanje analitičnih priložnosti, pridobitev ustreznih notranjih in zunanjih podatkov ter postavitev in delovanje analitike so prvi koraki, ki vodijo k nadaljnjemu razvoju podatkovne kulture podjetja (Grossman & Siegel, 2014). Generalni direktorji morajo spremeniti svoj način odločanja in se osredotočiti na sprejemanje odločitev na podlagi podatkov in dejstev. Intuicija in izkušnje naj tako postanejo le en izmed faktorjev odločanja (Mihai & Borza, 2019). V idealnem primeru bo podjetje zaposlilo vodjo podatkov (angl. chief data officer – CDO) ali pa vodjo analitike (angl. Chief analytics officer – CAO) (Anderson, 2015). V raziskavi, ki jo je opravilo podjetje MicroStrategy, je 58 % vprašanih podjetij odgovorilo, da že imajo vodjo tehnologij, 23 % pa, da razmišljajo o ustanovitvi omenjenega delovnega mesta (Microstrategy, 2020). Glavna vloga vodje podatkov je, da oblikuje kulturo in s svojim vplivom spreminja odnos do uporabe podatkov v celotni organizaciji. Vodja analitike pa uporablja podatke v strateške namene in večinoma komunicira z analitiki, podatkovnimi znanstveniki in drugimi podatkovno-tehničnimi vodji. Podatkovno kulturo spodbuja s pomočjo ustvarjanja resnične poslovne vrednosti na podlagi analitike podatkov (Anderson, 2015). Ob vzpostavljanju oddelka, ki se ukvarja s podatki,



ima podjetje več možnosti. Prvi model je centraliziran, kar pomeni, da podatkovne znanstvenike in analitike postavimo v svoj, popolnoma ločen oddelk. Ta način je najenostavnejši za pridobitev ustreznih podatkov, postavitev podatkovne infrastrukture, pridobitev ustreznega znanja in kakovostnega kadra ter razvoj ustreznih statističnih ter napovednih modelov. Izziv, ki ga prinaša centraliziran sistem, je oddaljenost oddelka od drugih enot, kar oteži razumevanje celotnega poslovanja podjetja in potreb drugih oddelkov. Naslednji model postavlja podatkovne znanstvenike in analitike v druge oddelke, tako delujejo skupaj in imajo več možnosti sodelovanja. Tu pa se lahko pojavi vprašanje o zadostni količini znanja za specifična področja in številu strokovnjakov v določenem oddelku. Zadnji model združuje oba prej omenjena modela. Hibriden način omogoča zadostno število strokovnjakov v oddelku za upravljanje s podatki, obenem pa vključuje strokovnjake tudi v druge oddelke. Slednje omogoča dobro integracijo analitike v vsakdanje poslovno odločanje. Hibriden model morda ni najboljša izbira za manjša podjetja, saj zahteva večje število strokovnjakov in povzroči več stroškov (Grossman & Siegel, 2014). Po poročilu, ki ga je izdal McKinsey po pogovorih z vodilnimi na položajih podatkovne analitike, se v začetnih fazah postavljanja analitike podjetja večinoma poslužujejo centraliziranega modela zaradi lažje organizacije, kasneje pa je hibridni model najprimernejši, saj vključuje strokovnjake tako iz poslovnega kot tudi analitičnega področja (Diaz, Kayvaun, & Seleh, 2018). V idealnem primeru podjetje zaposluje strokovnjake iz treh različnih področij - poslovnega, tehnološkega in analitičnega. Strokovnjaki iz poslovnega področja imajo znanje o industriji, v kateri podjetje deluje, produktih, storitvah in celotnem delovanju podjetja. Strokovnjaki iz tehnološkega področja dobro poznajo orodja in infrastrukturo, potrebno za izdelovanje poročil na podlagi analiziranih podatkov. Analitiki pa so specializirani za obravnavo podatkov in analitiko (Grossman & Siegel, 2014). Slika 1 prikazuje vpletenost različnih strokovnjakov v omenjena področja. **Poslovni vodje** (angl. Business leaders) so odgovorni za vzpostavitev analitike v celotnem podjetju. **Vodje dostave podatkov** (angl. Delivery managers) dostavljajo podatke in analitične vpoglede končnim uporabnikom. **Graditelji interaktivnih orodij** (angl. Workflow integrators) gradijo interaktivna orodja, ki pomagajo pri iskanju rešitev in odločanju. **Analitiki vizualizacij** (angl. Visualization analysts) vizualizirajo podatke in gradijo poročila ter nadzorne plošče. **Podatkovni inženirji** (angl. Data engineers) zbirajo, strukturirajo in analizirajo podatke. **Podatkovni arhitekti** (angl. Data architects) so zadolženi za kakovost in doslednost sedanjih in bodočih podatkovnih tokov. **Prevajalci analitike** (angl. Analytics translators) zagotavljajo, da analitika rešuje ključne poslovne težave. **Podatkovni znanstveniki** (angl. Data scientist) pa razvijajo statistične modele in algoritme (Diaz, Kayvaun & Seleh, 2018).

Slika 1: Vpletenost strokovnjakov iz treh različnih področij v poslovanje podjetja



Prirejeno po Diaz, Rowshankish & Seleh (2018) in lastno delo.

Kako zahtevno je v resnici uvajanje podatkovne kulture v podjetje, pove podatek, da skoraj 90 % podjetij, ki so bila uspešna pri njeni uveljavitvi, porabi več kot polovico proračuna, namenjenega analitiki, za dejavnosti, ki spodbujajo sprejemanje analitike v vsakodnevnem poslovanju. Podjetja se največkrat poslužujejo internih izobraževanj, ki vključujejo delavnice, usposabljanja na delovnem mestu in obiske podobnih podjetij z izkušnjami na določenem področju. Najpogosteje se uporabljajo štirje tipi izobraževanj. Izobraževanje, namenjeno generalnim direktorjem in zaposlenim na višjih položajih, se osredotoča na razumevanje naprednih orodij in analitike ter prepoznavanje priložnosti. Poleg tega predstavi tudi vplive na zaposlene, ovire pri sprejemanju podatkovne kulture, razvoju talentov in predstavi smernice za pomoč pri kulturni spremembi. Naslednje izobraževanje, namenjeno analitikom, se osredotoča na izpopolnjevanje trdih in mehkih veščin podatkovnih znanstvenikov, inženirjev, arhitektov in drugih zaposlenih, ki se ukvarjajo s podatkovno analitiko, upravljanjem podatkov in izgradnjo rešitev. Na področju prevajalstva analitike se pogosto dogaja, da zaposleni prihajajo iz poslovnega področja in potrebujejo več tehničnega usposabljanja. Izobraževanje tako vključuje uporabo analitičnih pristopov k poslovnim težavam in razvoju ter uporabi orodij. Pogosto se učijo s pomočjo spletnih tečajev in primerov iz prakse na delovnem mestu pod nadzorom bolj izkušenega sodelavca. Še zadnje izobraževanje pa se osredotoča na končne uporabnike analitike. Zaposleni na položajih, ki niso strateški, navadno potrebujejo le splošno predstavitev novih orodij in usposabljanje na delovnem mestu, medtem ko strateški nosilci, kot so tržniki in finančniki, potrebujejo usposabljanje na višji ravni, ki vključuje uporabljanje orodij na primerih iz prakse, za potrebe izboljšanja poslovnih odločitev (Fountaine, McCarthy & Saleh, 2019).

Popolno preoblikovanje podjetja pri večini podjetij traja med 18 in 36 mesecev, nekatera pa zaradi različnih težav potrebujejo tudi do pet let. Vodje lahko pripomorejo k hitrejšemu in uspešnejšemu preoblikovanju na različne načine. V nadaljevanju bomo omenili štiri.

- **Biti vzgled** – vodje lahko svojo zavzetost za razvoj podatkovno vodenega podjetja pokažejo tako, da se udeležujejo delavnic in seminarjev ter se aktivno vključujejo v aktivnosti na omenjenem področju. Pomembno je tudi, da se ne bojijo neuspehov in se iz njih znajo učiti. Najboljši vodje se na dnevni bazi pogovarjajo z zaposlenimi in jim postavljajo vprašanja, kot sta, kako pogosto imamo prav in kateri podatki podpirajo današnje odločitve.
- **Poslovati odgovorno** – jasno razločevanje med analitiko in njeno implementacijo v poslovanje je pomemben dejavnik uspeha. Analitika je namreč le sredstvo, ki pomaga pri reševanju izzivov. Zato mora odgovornost za uspešno implementacijo analitike v vsakdanje odločitve prevzeti poslovni oddelek in ne analitični oddelek, ki je odgovoren za proces postavitve analitike. Vodja preoblikovanja naj spremlja celoten proces in zagotovi tudi ustrezno merjenje uspešnosti.
- **Spremljati in poenostaviti preoblikovanje** – zaposleni so pogosto skeptični pri sprejemanju novih načinov poslovanja, zato je primerjanje rezultatov odločanja s pomočjo analitike in brez nje eden izmed pobudnikov za njeno sprejemanje.
- **Zagotoviti spodbude za spremembe** – nagrade in priznanja za zaposlene, ki na kakršenkoli način pripomorejo k razvoju analitike in njene uporabe, so dobra pobuda za vzpostavljanje kulture. S tem podjetja dobijo tudi vpogled v ravnanje zaposlenih in njihovo pripadnost podatkovni kulturi.

S pomočjo omenjenih načinov zaposleni začenjajo razmišljati širše in reševati bolj kompleksne probleme na ravni celotnega podjetja. Hitrost inovacij na področju analitike se poveča in s tem zaposleni pričnejo zaupati v »preizkusi in se nauči« princip. Novi pristopi bodo v podjetjih povzročili spremembe pri delu, zaposlenih in kulturi. Vendar bodo uspešna podjetja, v katerih bodo ljudje skupaj s stroji oblikovali prihodnost, imela veliko prednost, v nasprotju s podjetji, v katerih bodo ljudje ali stroji delali vsak le zase (Fountaine, McCarthy & Saleh, 2019).

#### 2.4 Posledice pomanjkanja analize podatkov in poročanja

Implementacija podatkovne kulture prinese s seboj veliko novih spoznanj. Med drugim je bilo 9 od 10 vodilnih zaposlenih mnenja, da bi bile odločitve, ki so jih sprejeli v zadnjih treh letih, boljše, če bi imeli na razpolago ustrezne podatke. V isti raziskavi je 40 % vprašanih odgovorilo, da bi bile njihove odločitve bistveno boljše, če bi imeli na razpolago vse strukturirane in nestrukturirane podatke za podporo pri odločitvah (Economist Intelligence Unit, 2012).

S pomočjo analitike in njene pravilne uporabe lahko podjetja najdejo nove priložnosti, učinkoviteje poslujejo, povečajo dobičke in imajo več zadovoljnih strank. Medtem ko se konkurenti na podlagi analitike in interaktivnih nadzornih plošč odločajo v realnem času, ostala ne-podatkovno vodena podjetja zaostajajo in lahko sprejemajo napačne odločitve. Kako doseči najučinkovitejše oglaševanje za določene skupine strank, se odločiti o najboljši ceni za določen produkt in spremljati konkurenčna podjetja pri njihovih vsakdanjih spremembah, so le nekatera izmed vprašanj, ki so jih podjetja dolga leta reševala na podlagi intuicije. Učinkovitost analitičnih orodij pa danes daleč prekaša intuicijo vodilnih in osebna mnenja strank. Podjetja z analitičnim pristopom so v povprečju za 5 % bolj produktivna in za 6 % bolj profitabilna kot njihovi konkurenti (Alsghaier, Akour, Shehabat & Aldiabat, 2017).

V kolikor podjetja postavitve analitike ne jemljejo resno, se lahko med procesom pojavi kar nekaj težav. Te vodijo v nezadovoljstvo in razočaranje nad naprednimi analitičnimi sistemi. Podjetja so v raziskavi, ki so jo opravili Fountaine, McCarthy in Saleh (2019), izpostavila sledeče težave:

- pomanjkanje razumevanja napredne analitike in zaposlovanje ključnega kadra,
- slaba ocena izvedljivosti in časovnega okvirja,
- neupoštevanje priložnosti in groženj napredne analitike,
- slabo določevanje ključnih nalog zaradi pomanjkanja razumevanja analitičnih spretnosti,
- pomanjkanje podatkovnih prevajalcev, ki bi znali identificirati dodano vrednost analitike in jo na učinkovit način predstaviti vodilnim v podjetju,
- izolacija analitičnega oddelka od celotnega poslovanja,
- neuskkljena priprava podatkov,
- nepremišljeno postavljanje analitičnih sistemov in nadzornih plošč,
- pomanjkanje meritev za oceno uspešnosti projektov,
- pozabljanje na etična, socialna in regulatorna področja.

Poleg omenjenih težav pa se je potrebno zavedati, da podjetje lahko zamudi veliko priložnosti, saj jih brez analitičnega pristopa ne zna identificirati. S tem lahko izgubi tržni delež, saj so konkurenčna, podatkovno vodena podjetja bolj natančna in hitrejša pri sprejemanju odločitev. Bill Ruh, podpredsednik programske opreme pri podjetju General Motors, je po uvedbi analitike povedal: »Kar smo odkrili, je, da več kot imamo podatkov, bolj postajamo inovativni v analitiki in pričenjamo delati stvari, za katere prej nismo mislili, da jih zmoremo.« (Economist Intelligence Unit, 2017).

### **3      PODROČJA NAPREDNE ANALIZE PODATKOV**

Za vzpostavitev naprednih sistemov poročanja potrebujemo veliko več kot le podatkovni model. V tem poglavju bomo predstavili vse korake, ki so potrebni za učinkovito nadzorno ploščo oziroma poročilo. Predstavili bomo različne načine zbiranja podatkov in uporabo

različnih orodij, pri čemer se bomo osredotočili na orodja, povezana s programom Power BI. V drugi polovici poglavja se bomo dotaknili tudi dodatnih funkcij, ki jih ponujajo programi za postavitev naprednih sistemov poročanja. Slednji postajajo pomemben element odločanja in se uporabljajo v vedno več oddelkih v podjetjih. Predstavili bomo tudi nekaj primerov, kako podjetja uporabljajo omenjene sisteme in kakšni so trendi za prihodnost.

### 3.1 Načini zbiranja podatkov

Anderson (2015) v svoji knjigi navaja Thomsona Nguyena, ustanovitelja in direktorja finančno-tehnološkega podjetja Nearside, ki je dejal: »Dobri podatki bodo vedno premagali dobre modele.« (str. 19). Podatki so temelj vsakršnih analiz in njihove uporabe, zato morajo biti ustrezni, verodostojni, zbrani na pravi način, v pravi obliki, na pravem mestu in ob pravem času (Anderson, 2015). Kvaliteta podatkov je z dobo množičnih podatkov postala vse kompleksnejša. V preteklosti so bili uporabniki podatkov obenem tudi proizvajalci podatkov, kar je zagotovilo lažje preverjanje kakovosti. Z dobo množičnih podatkov pa se je pojavilo tudi mnogo različnih virov, ki so težje preverljivi (Cai & Zhu, 2015).

Podjetja naj se v splošnem držijo načela zbiranja vseh podatkov, ki so v določenem trenutku na voljo. Nikoli ne morejo namreč vedeti, kateri in kdaj jim bodo prišli prav. Več podatkov pomeni tudi bolj natančne modele in boljše razumevanje strank. Če z zbiranjem podatkov odlašamo, se lahko zgodi, da določeni podatki niso več dostopni, ko jih potrebujemo. Zbiranje vseh dostopnih podatkov pa ima lahko tudi pomanjkljivosti, saj jih je potrebno standardizirati in očistiti. Slednje vzame veliko časa, zahteva veliko spretnosti zaposlenih na tem področju in povzroči stroške v povezavi s shranjevanjem in vzdrževanjem velikih količin podatkov. Zavedati se je potrebno tudi dejstva, da je proces zbiranja podatkov dolgotrajen in počasen, zato morajo zaposleni dobro poznati vrednost podatkov, ki jih želijo pridobiti, in se zavedati njihovega vpliva na celotno podjetje. Za vsak oddelek je priporočljivo, da ima glaven set podatkov. Oddelek za stranke navadno zbira osebne podatke strank, podatke o elektronskih naslovih, telefonskih številkah in naročilih strank. Poleg glavnih podatkov imajo podjetja običajno tudi podporne sete podatkov. Slednji so zagotovo dragoceni, niso pa tako pomembni kot glavni seti podatkov. Podporni podatki pomagajo na primer razumeti obnašanje strank, z njimi lahko marketinški oddelek izvede A/B teste in ugotovi, kje imajo stranke težave pri razumevanju njihovih kampanj.

Za mnoge podatkovne vire podjetja uporabljajo sistematičen pristop, kjer pridobijo vse razpoložljive podatke. Podatke lahko pridobijo preko vmesnika uporabniškega programa protokola za prenos datotek (angl. file transfer protocol – FTP) ali pa z izvlečenjem podatkov iz spletnih strani (angl. scraping). Pred pričetkom dela je potrebno dobro razumeti namen analize in določiti parametre, saj je od tega odvisen tudi način zbiranja podatkov. Ali za končno nadzorno ploščo potrebujemo surove ali agregirane podatke, je eno izmed pomembnih vprašanj. Za določene analize je smiselno vzorčenje podatkov, za druge potrebujemo čim več surovih podatkov, na primer, ko delamo poglobljene analize iz

kvalitetnih podatkov. Naslednji način je, da podjetja kupijo podatke, ki se jih pogosto uporablja kot dodatek k ostalim, saj dodajo analizi širši kontekst. Pri nakupu je potrebno dobro oceniti ceno, kvaliteto, ekskluzivnost podatkov, možnost posodabljanja, zanesljivost, varnost, pogoje uporabe, format, dokumentacijo, obseg in granularnost podatkov (Anderson, 2015). Pogosto se podatke pridobiva tudi iz socialnih medijev, nekateri vsebujejo integrirana orodja za analize, s katerimi lahko spremljamo vedenje sledilcev. Lahko pa se povežemo tudi preko zunanjih orodij, ki nam pogosto omogočajo več svobode pri analizi in dovoljujejo podrobnejše analize. Točno določene podatke lahko pridobimo tudi z anketami, ki sicer vzamejo več časa, vendar omogočajo pridobitev točno določenih odgovorov. Eksperiment in opazovanje sta še dva načina zbiranja podatkov, ki pa zagotovo zahtevata več časa in sredstev. Prav tako zahtevata več strokovnega znanja s področja vodenja tovrstnih raziskav. Kot eden izmed temeljnih trendov načina zbiranja podatkov se je pojavil internet stvari (angl. internet of things, v nadaljevanju IoT). Z njegovo pomočjo bodo predvsem tovarniška podjetja lahko v veliki meri optimizirala proizvodno dejavnost in učinkoviteje upravljala s sredstvi in zaposlenimi (Chui, Collins & Patel, 2021).

### 3.2 Shranjevanje in dostop do podatkov

Preden začnemo z izbiro načina shranjevanja podatkov, moramo poznati vrsto podatkov, ki jih imamo. Glede na vrsto in količino podatkov so tudi načini shranjevanja različni. Glede na to, kako in kam shranimo podatke ter kakšni so ti podatki, do njih tudi dostopamo na različne načine.

V **ploskih zbirkah podatkov** se shranjujejo tekstovni in binarni dokumenti, ki so lahko strukturirani ali pa nestrukturirani. Navadno so uporabljeni za shranjevanje manjših količin podatkov, ki ne potrebujejo dodatnega kompleksnejšega obdelovanja. Uporabljajo se tudi za izmenjevanje podatkov med različnimi sistemi in aplikacijami. Ploske zbirke podatkov so enostavne za kreiranje, branje in pisanje z osnovnimi programskimi jeziki in orodji. Do njih navadno dostopamo preko besedilnega urejevalnika ali pa vgrajenega programskega jezika. Najprepoznavnejši predstavniki ploskih zbirk podatkov so CSV (angl. comma separated values), TSV (angl. tab separated values), JSON (angl. java script object notation) in XML (angl. extensible markup language). Ko se podatkovna baza poveča, jo je težko urejati, zato je v tem primeru smiselno uporabiti **relacijske zbirke podatkov**. Podatki so v tem primeru shranjeni v tabelah, kot vrstice in stolpci v strukturirani obliki. Vsaka tabela predstavlja drugačno vrsto podatkov, vsaka vrstica predstavlja en zapis teh podatkov in vsak stolpec predstavlja posamezna polja podatkov. Relacijske baze temeljijo na relacijskih modelih, ki omogočajo strukturiranje podatkov v tabele, ki imajo medsebojne relacije. Najbolj uporaben jezik za interakcije z relacijskimi bazami je SQL. Uporablja se ga za vstavljanje, posodabljanje, pridobivanje in brisanje podatkovnih baz. Relacijske baze so primerne tako za uporabo v manjših podjetjih kot tudi večjih poslovnih sistemih. Omogočajo celovitost in konsistenco čez celotno bazo podatkov, podpirajo kompleksnejše poizvedbe in združevanja, razširitve in prilagoditve ter transakcije med različnimi podatkovnimi bazami.

Najprepoznavnejše relacijske baze so MySQL, SQL Server, SQLite in Oracle. Slednje niso najboljša izbira, če imamo zelo velike količine podatkov, zelo nestrukturirane podatke in podatke, ki se hitro spreminjajo. V tem primeru uporabimo NoSQL, ki je bolj fleksibilen in ne uporablja tabel in vrstic. Narejen je predvsem za shranjevanje in obdelovanje velike količine nestrukturiranih ter delno strukturiranih podatkov. Navadno se NoSQL baze podatkov uporabljajo za spletne aplikacije v realnem času, analitiko v realnem času, mobilne aplikacije in v sistemih za upravljanje vsebin. V primeru, ko imamo popolnoma strukturirane podatke, potrebujemo kompleksne poizvedbe in združevanja ter popolno konsistenco, je bolj smiselno uporabiti relacijske baze podatkov (Geeks for geeks, 2022). V dosedanjih primerih smo omenili shranjevanje podatkov v obliki dokumentov, podatke pa lahko shranimo tudi v obliki **predmetov**, ki vključujejo podatke same in metapodatke, ki slednje opisujejo. Dostopni so preko enoličnih identifikatorjev, kot sta ključ in URL (angl. uniform resource locator). Navadno se uporabljajo za shranjevanje nestrukturiranih podatkov, kot so medijske vsebine in varnostne kopije, in pri storitvah za shranjevanje podatkov v oblaku. Zasnovani so tako, da se prilagajajo velikim količinam podatkov in velikemu številu sočasnih uporabnikov ter so odporni na napake. Njihove prednosti so zagotovo možnost razširitve, stroškovna učinkovitost, kompatibilnost s storitvami v oblaku in enostavna prilagodljivost metapodatkov. Najpopularnejši sistemi za shranjevanje predmetov so Amazon S3, Google Cloud Storage in Microsoft Azure Blob Storage (International Business Machines Corporation, 2023b). Zadnji način shranjevanja podatkov, ki ga bomo omenili, je **shranjevanje podatkov v oblaku**. V tem primeru so podatki shranjeni na oddaljenem strežniku, do katerega dostopamo preko interneta. Uporablja se tako za shranjevanje osebnih datotek kot tudi datotek velikih poslovnih sistemov. V vsakem primeru nam nudi enostavno prilagajanje količine prostora, ki ga potrebujemo, enostaven dostop do podatkov, varnostno kopiranje in obnovitev podatkov ter stroškovno učinkovitost, saj plačamo le za prostor, ki ga potrebujemo. Shranjujemo lahko različne vrste datotek, vključno z vsemi tistimi, ki smo jih že omenili v tem poglavju. Pri shrambi podatkov v oblaku pa je potrebna določena previdnost. Ko izbiramo ponudnika, je potrebno razmišljati o pravni varnosti, ki jo storitev ponuja, skladnosti s predpisi in drugimi standardi, ki se lahko pojavijo v specifični industriji. Največji ponudniki tovrstne storitve so zagotovo Amazon, Google in Microsoft. V letu 2019 je kar 47 % vprašanih podjetij uporabljalo analitične nadzorne plošče v oblaku, kar je za 8 % več kot leto prej (Microstrategy, 2020).

Ne glede na način hranjenja podatkov in dostopnost mora podjetje dobro določiti, kdo v podjetju ima dostop do katerih podatkov. Zlasti ko gre za osebne in druge občutljive podatke, je v procesu zbiranja načina shranjevanja podatkov potrebno razmisliti o potencialnih težavah. Dostopnost, deljenje, točnost, varnost in zasebnost podatkov zahtevajo kompleksne varnostne ukrepe, ki pogosto vključujejo šifriranje, postavitve požarnih zidov, nadzornih sistemov za dostope, jasno politiko o hrambi podatkov in načrtovanje ukrepov v primeru vdorov.

### 3.3 Obdelovanje podatkov

Ena največjih težav, s katero se soočajo podjetja, je pomanjkanje kadra, natančneje ljudi, ki bi analizirali podatke. Sklepamo lahko, da imajo podjetja zaradi pomanjkanja omenjenega kadra veliko težav pri interpretaciji nestrukturiranih podatkov. Slednje potrjuje podatek, da kar 42 % anketiranih v raziskavi meni, da je nestrukturirane podatke pretežno interpretirati (Economist Intelligence Unit, 2012).

Predpostavka za uspešno obdelovanje podatkov so zaposleni. Specialisti na določenih področjih imajo pomembno vlogo pri naprednih sistemih poročanja. Podatkovno vodeno podjetje ima najverjetneje več različnih strokovnjakov za posamezna področja analitike. Podatkovni analitiki obdelujejo podatke na različnih ravneh, od zbiranja in priprave podatkov do grajenja zelo specifičnih analiz. Njihovo delo ponavadi zajema tako analiziranje kot tudi poročanje. Glede na podjetje, v katerem delajo, in delo, ki ga opravljajo, pa se tudi njihovo tehnično znanje lahko močno razlikuje. Podatkovni inženirji obdelujejo podatke v smislu pridobivanja, čiščenja in preoblikovanja v obliko, ki je primerna za nadaljnje analiziranje. Te podatke lahko prevzamejo podatkovni znanstveniki, ki pripravljene podatke obdelujejo v povezavi s statističnimi analizami. V zadnjem koraku obdelovanja podatkov slednje prevzamejo strokovnjaki za vizualizacijo. Analizirane podatke pretvorijo v nadzorne plošče, ki so razumljive tudi zaposlenim iz drugih oddelkov. Kot lahko opazimo iz opisov nalog, so pri obdelovanju podatkov potrebne veščine iz poslovanja, matematike, strojnega učenja oziroma množičnih podatkov, programiranja in statistike.

Obdelovanje podatkov se prične z njihovo pripravo, ki je eno najpomembnejših in hkrati tudi najbolj zamudnih opravil. Priprava podatkov terja med 50 in 70 % celotnega časa, ki je potreben za določen projekt. Navadno vključuje združevanje nizov podatkov in/ali zapisov, izbiro vzorčnih podatkov, izpeljevanje novih atributov, razvrščanje podatkov, odstranjevanje ali zamenjavo praznih in/ali manjkajočih vrednosti ter razdelitev podatkov v učne podatkovne množice in testne podatke (International Business Machines Corporation, 2021). Že kar nekajkrat smo omenili, kako pomembna je kvaliteta podatkov. Kvaliteto lahko zagotovimo v procesu čiščenja podatkov, ki vključuje pregled, raziskovanje in urejanje podatkov. Omenjeni koraki nam omogočijo najti preproste probleme, kot so nedoslednosti pri številih in črkah ter manjkajoče vrednosti. V tem koraku tudi izločimo nepotrebne podatke, torej tiste, ki ne sodijo v kontekst problema, ki ga želimo z analizo rešiti (Elgabry, 2019). Za odkrivanje bolj kompleksnih težav pa je potrebno narediti raziskovalno analizo podatkov, ki pomaga prepoznati vzorce v naboru podatkov. V prvi fazi, fazi pregledovanja, navadno naletimo na štiri tipe napak:

- **Težave z vrsto podatkov:** v podatkovni znanosti poznamo dve glavni vrsti podatkov, kategorično in numerično. Kategorična predstavlja kvalitetne lastnosti, kot so ime, spol, naziv, narodnost in podobno, medtem ko numerična predstavlja kvantitativne značilnosti, kot so na primer število prebivalcev, število transakcij in višina oseb.



Napačno določena vrsta podatkov lahko privede do napačnih rezultatov, zrušitev sistema, neučinkovitosti analiz in drugih težav (Huang, 2019).

- **Manjkajoče in podvojene vrednosti:** tovrstne napake so enostavne za odkrivanje, vendar lahko povzročijo veliko težav pri odpravljanju. Podvojene vrednosti navadno samo izločimo iz podatkov, medtem ko moramo pri manjkajočih poiskati originale ali pa metapodatke. Če podatkov ne najdemo na ta način, naredimo raziskovalno analizo, ki lahko pomaga zapolniti manjkajoče vrednosti. V primeru, da podatkov z omenjenimi metodami vseeno ne najdemo, lahko numerične manjkajoče podatke zamenjamo z »nič« (slednjih ne smemo upoštevati pri izračunih), kategorične podatke pa lahko zamenjamo z »manjkajoče« (angl. not available – NA) (Elgabry, 2019).
- **Izjeme:** med raziskovanjem podatkov lahko zasledimo vrednosti, ki so tako drugačne od vseh ostalih, da je skoraj neverjetno, da bi bile pravilne. Kljub temu jih je pred odstranitvijo potrebno raziskati in se prepričati, da ne spadajo v bazo podatkov. Nekateri modeli, kot je na primer linearna regresija, so izjemno občutljivi na izjeme in jih lahko prikažejo v napačni luči (Elgabry, 2019).
- **Nekonsistenca in konfliktni podatki:** slednja težava se največkrat pojavi pri združevanju dveh baz podatkov, ko so isti podatki zapisani v različnih formatih. V nekaterih primerih je tovrstno nekonsistentnost težko rešiti v prvi fazi obdelovanja podatkov. V tem primeru problematične podatke označimo in se k njim ponovno vrnemo, ko imamo boljši pregled nad celotnimi podatki (Elgabry, 2019).

V fazi raziskovanja uporabimo raziskovalno analizo, ki jo lahko naredimo na več načinov, vendar je eden izmed najpogosteje uporabljenih načinov vizualizacija, ki pomaga pri boljšem razumevanju baze podatkov. Analiza lahko v tem primeru poteka enodimenzionalno, kjer podatke predstavimo s pomočjo grafičnega prikaza kvantilov (angl. boxplot) in/ali histograma. Pri dvodimenzionalni analizi podatke prikažemo z grafom raztrosa (angl. scatter plot) in/ali stolpčnim grafom. Omenjeni analizi sta precej enostavni, medtem ko so pri večdimenzionalni analizi podatki običajno zmanjšani na več dimenzij. Pri slednji uporabimo metodo razčlenbe načelnih sestavin (angl. principal component analysis), ki analizira korelacije med večvariantnimi numeričnimi vrednostmi. Poleg omenjenih metod obstaja še veliko drugih, tako vizualnih kot tudi nevizualnih, vendar so omenjene najbolj uporabljene in obenem tudi najenostavnejše. Omogočajo nam prepoznavanje izstopajočih vrednosti, povezav med podatki in prepoznavanje nenavadnih vzorcev ter trendov. Ko identificiramo vse težave, nam preostane le še njihovo odpravljanje. Podatke lahko urejamo z različnimi orodji, kot sta na primer programska jezika R in Python. Oba omogočata dobro analizo in odpravljanje težav, tako enostavnih kot tudi bolj kompleksnih (Huang, 2019). Spomnimo se lahko tudi na druga orodja, ki smo jih omenili v poglavju 1.2.

Pri pripravi podatkov se je potrebno zavedati nepravilnosti, kjer sta dve človeku enaki vrednosti zapisani na dva različna načina. Slednje računalniški programi razumejo kot dve različni vrednosti, ki ju je potrebno poenotiti. V spodnji tabeli 2 je zapisanih nekaj tovrstnih primerov.

Tabela 2: Prikaz in razlaga različnih zapisov vrednosti

Zapis vrednosti – način 1	Zapis vrednosti – način 2	Razlaga
» hello world «	»hello world«	Dodatni prazni prostori pred ali po vrednosti naj bodo odstranjeni.
727	000727	Nekatere številske kode so zapisane z ničlami zaradi zagotavljanja enakega števila števk. Ne glede na to je zapisa potrebno poenotiti.
Male	m	Nize je mogoče vnesti na različne načine, tako lahko v primeru spolov dobimo namesto enega niza (moški spol) dva in več nizov.
25.2.2023	2/25/2023	Zapis datumov je možen na različne načine. Potrebna je standardizacija.

Vir: lastno delo.

V kolikor bomo tekom analize uporabili statistične metode, je potrebno podatke tudi normalizirati. Torej, spremeniti moramo lestvico vrednosti v obseg med 0 in 1, s čimer dobimo normalno porazdelitev podatkov. Tekom celotne priprave pa je podatke potrebno tudi dobro razumeti. Pride lahko namreč do napak v zapisih in med samim naborom podatkov. Plača zaposlenega, na primer, ne more biti nižja od obračunanih davkov (Elgabry, 2019). Nerazumevanje konteksta lahko privede do popolnoma napačnih rezultatov in posledično napačnih odločitev v kasnejših fazah, ko se analize uporabljajo za namen odločanja.

### 3.4 Uporaba DAX programskega jezika pri izgradnji ključnih kazalnikov uspešnosti

»Preštejte, kar je števno, izmerite, kar je merljivo, in kar ni merljivo, naredite merljivo.« (Anderson, 2015, str. 111). Galileo Galilei je s temi besedami nazorno pokazal, kako pomembno je merjenje rezultatov. Podatkovno vodena podjetja se morajo tega načela držati na vsakem koraku, saj lahko le tako vidijo, ali gre podjetje v pravo smer. Poleg »držanja prave smeri« pa so pokazatelj napredka tudi ključni kazalniki uspešnosti, s katerimi navadno merimo uspeh podjetja. Kazalniki morajo kazati prave informacije in biti natančni, zato je zelo pomembno, kako so postavljeni. Ključna načela, ki se jih podjetja držijo pri razvoju kazalnikov, so enostavnost, standardiziranost, točnost, robustnost in direktnost. Kazalniki so **enostavni**, kadar jih zaposleni razumejo in v njih ne vzbujajo zmede, so enostavni za implementacijo in uporabni tudi za druge oddelke. **Standardizacija** kazalnikov pomeni, da uporabimo kazalnike, ki so že uveljavljeni, namesto da razvijamo svoje. Tako jih bodo vsi zaposleni v podjetju, ne glede na oddelek, lažje razumeli. Izven podjetja pa bomo rezultate lažje primerjali s konkurenti. Kazalniki izražajo **točnost**, kadar so njihovi rezultati natančni in obenem izmerijo enak rezultat ob morebitni ponovitvi. V kolikor so kazalniki občutljivi na ekstreme, lahko dobimo zavajajoče rezultate, zato moramo podatke dobro raziskati že med raziskovalno analizo, da lahko kasneje razvijemo **robustne** kazalnike, ki so neobčutljivi

na ekstremne vrednosti. Med merjenjem si želimo čim manj motečih elementov, zato poskušamo izbrati ali razviti kazalnike, ki bodo čim bolj **direktno** merili zelen proces. Poleg omenjenih načel je dobro razmišljati tudi o uporabi relativnih oziroma absolutnih kazalnikov. V kolikor imamo število strank, ki se vračajo, izraženo samo v deležu, je ta vrednost lahko zavajajoča. Ko v določenem obdobju podjetje pridobi veliko novih strank, se delež tistih, ki se vračajo, zmanjša, kar ne pomeni nujno, da se je število omenjenih zmanjšalo (Anderson, 2015).

Ko uspešno določimo kazalnike, ki jih potrebujemo, jih lahko izračunamo na več načinov in s pomočjo različnih orodij. Ker bomo v tretjem delu magistrske naloge predstavili primer iz prakse, v katerem je bil uporabljen programski jezik DAX, bomo le-tega tudi predstavili. DAX je zbirka funkcij, operatorjev in konstant, ki jih uporabljamo v formulah ali izrazih, s pomočjo katerih izračunavamo vrednosti. Pomaga nam ustvariti nove informacije iz podatkov, ki jih imamo v našem modelu (Microsoft Corporation, 2022a). Funkcije so razdeljene v več kategorij, ki vključujejo združevanja, datume in čase, filtre, finančne, informacijske, logične, matematične, statistične, relacijske, tekstovne in druge funkcije (SQLBI Corporation, 2022). Najdemo ga v programu Power BI, Power Pivotu v Excelu in Analysis Servicu, kjer je namenjen predvsem za uporabo v poslovni inteligenci ter za modeliranje podatkov. Za boljše razumevanje njegovega delovanja si bomo pogledali tri ključne in zelo pomembne koncepte: sintakso, funkcije in kontekst. **Sintaksa** je sestavljena iz več elementov, ki skupaj predstavljajo formulo. Za lažje razumevanje si pogledjmo primer za izračun marže na sliki 2.

*Slika 2: Formula za izračun marže v DAX-u*

```
Margin = 'Table'[TotalSales] - 'Table'[TotalCosts]
```

*Vir: lastno delo.*

- Ime formule je lahko poljubno, vendar naj bo enostavno in naj jasno označuje izračunano vrednost.
- Prvi operator je enačaj (=), ki označuje začetek formule in vrne rezultate oziroma vrednost. Uporabimo ga pri vseh formulah, ki izračunavajo vrednosti.
- [TotalSales] predstavlja stolpec, iz katerega želimo vzeti vrednosti za izračun. Sklic na stolpec v formuli je vedno obdan z oglatimi oklepaji. Potrebno se je zavedati, da se formule DAX vedno sklicujejo na stolpce in ne na celice (kar velja za formule v Excelu).
- Minus (-) je matematični operator.
- Stolpec [TotalCosts] vsebuje vrednosti, ki jih želimo odšteti od vrednosti v stolpcu [TotalSales].

Poglejmo si še en primer, prikazan na sliki 3, ki predstavlja drugačno vrsto formule, zapisane v DAX-u.

Slika 3: Formula za izračun seštevka vrednosti prodanih produktov

```
Sum of Sales Amount = SUM(FactSales[SalesAmount])
```

Vir: lastno delo.

- Tako kot prejšnja, se tudi ta formula začne z imenom in operatorjem enačajem (=).
- Funkcija SUM sešteje vsa števila v stolpcu [SalesAmount].
- Okrogli oklepaj (()) obdaja enega ali več argumentov. Vse funkcije zahtevajo vsaj en argument, saj z njim funkciji posredujemo informacije (vrednosti), ki jih uporabi za izračune.
- FactSales je navedena tabela.
- [SalesAmount] je naveden stolpec, ki ga najdemo v tabeli FactSales. SUM funkcija torej sešteje vrednosti iz stolpca [SalesAmount].

Prva in druga formula se kljub kar nekaj podobnostim med seboj razlikujeta. V drugi smo predstavili funkcijo, ki je že vnaprej napisana in nam s tem olajša kompleksnejše izračune in ravnanje z različnimi tipi podatkov. V prvi formuli smo navedli samo ime stolpca, iz katerega želimo izračunati vrednosti, medtem ko smo v drugi formuli navedli tudi ime tabele, kjer se stolpec nahaja. Stolpci, na katere se sklicuje ista tabela, ne zahtevajo vključitve imena tabele v formulo. Kljub temu je navedba imena tabele zaželena, saj se s tem izognemo morebitni navedbi napačnih stolpcev. Pri pisanju formule je zelo pomemben pravilen zapis sintakse. V nasprotnem primeru bo vrnjena napaka ali pa bodo vrednosti napačne. Tretji primer, ki ga bomo predstavili, vključuje uporabo funkcije CALCULATE, ki je ena najmočnejših funkcij v DAX-u. Uporablja se jo predvsem pri pisanju kompleksnejših formul.

Slika 4: Formula za izračun prejšnjega obdobja

```
Previous Quarter Sales =  
CALCULATE(  
    SUM(FactSales[Sales]),  
    PREVIOUSQUARTER(DimDate[DataKey])  
)
```

Vir: lastno delo.

Že ob pogledu na zgornjo sliko vidimo, da je funkcija daljša, poleg tega pa vsebuje dve funkciji. PREVIOUSQUARTER je ugnezdena kot argument v funkcijo CALCULATE. Obenem smo uporabili tudi filter, ki zoži nabor izračunanih vrednosti. V tem primeru je izbran argument filter, ki je pravzaprav druga funkcija (Microsoft Corporation, 2023).

Predstavili smo tri različne primere izračunov, ki jih lahko vsakdo uporabi pri kreiranju kazalnikov. Ker DAX vključuje veliko vnaprej napisanih formul, je enostaven za uporabo

in ne zahteva veliko predznanja. Standardne funkcije za združevanje, kot so SUM, AVERAGE, COUNT, DISTINCT COUNT, MAX, MIN in druge, omogočajo dobro razumevanje podatkov in za kreiranje ne vzamejo veliko časa. Na ta način tudi zmanjšamo možnosti za pojav napak. Ker je podoben formulam v Excelu, lahko njegove osnove hitro osvojimo. Potrebno se je zavedati, da je kreiranje kompleksnejših formul bolj zahtevno in vzame več časa pri učenju. Druge prednosti, ki jih lahko izpostavimo, vključujejo:

- **Fleksibilnost:** z DAX-om lahko kreiramo tabele in stolpce ter izračunavamo kazalnike s pomočjo širokega nabora vnaprej napisanih funkcij.
- **Časovne funkcije:** omogočajo izračunavanje kazalnikov skozi različna časovna obdobja.
- **Dinamično agregacijo:** omogočajo izračune na nizu podatkov, ki se spreminjajo glede na trenutni filter. Primeri funkcij so SUMX, AVERAGEX, COUNTX in MINX.
- **Poizvedbe in filtriranje podatkov:** s funkcijami, kot so FILTER, ALL in CALCULATETABLE, podatke v podatkovnem modelu enostavno filtriramo.
- **Učinkovitost:** DAX je učinkovit pri izračunih in agregacijah na podatkovnih modelih. Kazalnike lahko računamo na več različnih načinov, kar optimizira hitrost izračunov.

Poleg omenjenih prednosti pa ima tudi nekaj slabosti, ki jih je potrebno upoštevati pri uporabi. Pri zelo velikih količinah podatkov ima DAX lahko težave, saj potrebuje več časa pri izračunavanju. V kolikor imamo zelo velike količine podatkov, je smiselno razmišljati o drugih tehnologijah, kot je zbirka odprtokodnih programskih pripomočkov Apache Hadoop (International Business Machines Corporation, 2023a). DAX ni tako univerzalen kot nekateri drugi jeziki za obdelovanje podatkov, saj ima omejeno število vrst izračunov, ki jih lahko izvede. V kolikor potrebujemo napredne matematične in statistične funkcije, strojno učenje, obdelavo naravnega jezika in podobno, bomo potrebovali druge programske jezike, kot sta na primer R in Python. Kreiranje kompleksnejših izračunov zahteva več časa in znanja ter zelo dobro razumevanje formul in celotne sintakse. Pri slednjih se lahko pojavijo tudi napake, ki jih je velikokrat težko odkriti in popraviti. Ker je DAX specifičen za Microsoft okolje, ga izven tega ni mogoče uporabljati, kar lahko predstavlja omejenost uporabe.

### **3.5 Vizualizacija podatkov**

V prejšnjih podpoglavjih smo se vselej ukvarjali s surovimi podatki, kako jih zbiramo, shranjujemo in obdelujemo. Z omenjenimi koraki smo podatke pripravili tako, da jih lahko razumemo na enostaven način. V tem podpoglavju si bomo pogledali razlike med dvema najpogosteje uporabljenima orodjema, kako pripravljene podatke prikažemo na enostaven in prijazen način, kako izberemo pravo vizualizacijo ter kaj je pri celotnem procesu vizualizacije pomembno.

Tako kot pri vseh prejšnjih korakih priprave podatkov, tudi pri vizualizaciji poznamo več različnih orodij. Omenili smo že Power BI in Tableau, ki sta verjetno najbolj prepoznavna

in najbolj enostavna za uporabo ter zmoreta delo z večjimi količinami podatkov. Med njima pa je tudi nekaj razlik, ki jih je potrebno poznati, preden se odločimo, katero orodje bomo uporabili. Power BI je tesno povezan z drugimi Microsoft produkti, kar pripomore k lažji uporabi in integraciji z drugimi Microsoft orodji. Obe orodji imata možnost povezave tudi z drugimi viri in strežniki, vendar ima Tableau nekoliko širši nabor. Za namen vizualizacij je Power BI uporabniku nekoliko prijaznejši, saj je uporabniški vmesnik narejen po zgledu Microsoft Office 365, ki je večini uporabnikom dobro poznan. Tableau pa za nekatere vizualizacije zahteva nekoliko več tehničnega znanja. Ko govorimo o združevanju podatkov v vizualizacijah, je Tableau nekoliko bolj fleksibilen, saj ne zahteva združevanja podatkov iz različnih virov v fazi priprave. Vgrajene funkcije za sodelovanje, skupno rabo nadzornih plošč in možnost dela na poročilih v realnem času so prednosti, ki jih prinaša Power BI. Slednje ponuja tudi Tableau, vendar niso tako dobro integrirane v celoten proces. Tableau ima široko paleto interaktivnih vizualizacij in naprednih možnosti, medtem ko je nabor v Power BI-ju nekoliko bolj omejen. Še zadnja razlika, ki jo bomo omenili, je strošek omenjenih orodij. Power BI je za manjša podjetja stroškovno bolj ugoden, nudi tako plačevanje na uporabnika kot tudi letno naročnino. Tableau je pri obeh omenjenih možnostih dražji. Že 13. leto zapored je Gartner prepoznal Microsoft kot vodilnega na področju platform za analitiko in poslovno inteligenco (Biswal, 2023). V nadaljevanju si bomo tako podrobneje ogledali možnosti vizualizacij s pomočjo orodja Power BI in pridobljeno znanje v zadnjem delu naloge aplicirali na primer izbranega podjetja.

V Power BI-ju lahko kreiramo popolnoma svoja poročila ali pa uporabimo že obstoječe predloge, ki jih prilagodimo glede na svoje potrebe. Ne glede na to, katero opcijo izberemo, imamo na voljo kar nekaj standardnih vizualizacij, ki jih lahko vidimo na sliki 5. Še preden pa se posvetimo vizualizacijam in njihovemu kreiranju, se moramo vprašati nekaj ključnih vprašanj o uporabniku:

- Kako uporablja poročilo?
- Katere vizualizacije olajšajo njegov proces sprejemanja odločitev?
- Ali obstajajo kakršnekoli predpostavke, ki vplivajo na izbiro vizualizacij?
- Katere informacije potrebuje za uspešno opravljanje svojega dela?
- Kje bo poročilo oziroma nadzorna plošča prikazana?

Ko si odgovorimo na zgornja vprašanja, bodimo pozorni tudi na to, kam postavimo določene informacije. V zahodni kulturi beremo od zgoraj navzdol in iz leve proti desni, najpomembnejše informacije torej postavimo v zgornji levi kot, podrobnosti pa postavimo nižje (Microsoft Corporation, 2022d). Ob kreiranju poročila se je dobro držati načela »manj je več«. Če imamo veliko komponent, je smiselno razmisliti, če zares vse potrebujemo, saj lahko postavitev veliko različnih vizualizacij uporabnika zmede. Poleg tega se lahko zgodi, da nam zaradi specifičnosti podatkov standardne vizualizacije ne ustrezajo. V tem primeru jih prilagodimo ali pa kreiramo sami. Držimo se tudi načela enovitosti; barve, oblike in pisave naj bodo skozi celotno poročilo usklajene. Kadar želimo poleg glavnih kategorij prikazati tudi podkategorije, uporabimo funkcijo vrtnje v globino (angl. drill down).

Slednja bo prilagodila vizualizacijo izbrani podkategoriji (Ferrari & Russo, 2016). Ko postavimo vizualizacije na mesto, kjer jih želimo, poskrbimo, da so pravilno poimenovane. Imena vrednosti, naslov, imena osi in podobno naj bodo označene tako, da bo vizualizacija uporabniku popolnoma razumljiva. V kolikor imamo na eni strani veliko vizualizacij, premislimo, kako bi jih lahko postavili v smiselno celoto. Lahko jih razdelimo v tri stolpce, ki vsebujejo vizualizacije s podobnimi kategorijami. V prvi stolpec lahko tako postavimo vizualizacije, ki se nanašajo na prodajo, v drugega vizualizacije s stroški in v zadnjega vizualizacije z zaposlenimi. V tem primeru tudi zmanjšamo količino teksta, odstranimo nepotrebne legende in uporabimo ikone, ki razložijo koncepte.

Smernice, ki smo jih predstavili v tem podpoglavju, so le izhodišče za ustvarjanje uporabnih poročil. Skozi izkušnje in prebiranje nasvetov ostalih uporabnikov lahko svoje znanje nadgradimo in poročila vse bolj prilagajamo potrebam uporabnikov. Vsi podatki, ki smo jih predstavili v poročilu ali nadzorni plošči, govorijo zgodbo, ki jo odločevalci lahko uporabijo za sprejemanje odločitev na višjem nivoju. Kvaliteta delovanja podjetja se tako zviša, saj imamo boljši nadzor nad poslovanjem. Odgovori na vprašanja, ki smo jih prej podajali na podlagi občutkov, dobijo oprijemljive argumente, ki zaključijo celotno zgodbo podatkov (Anderson, 2015).

### **3.6 Avtomatizacija in prilagoditve**

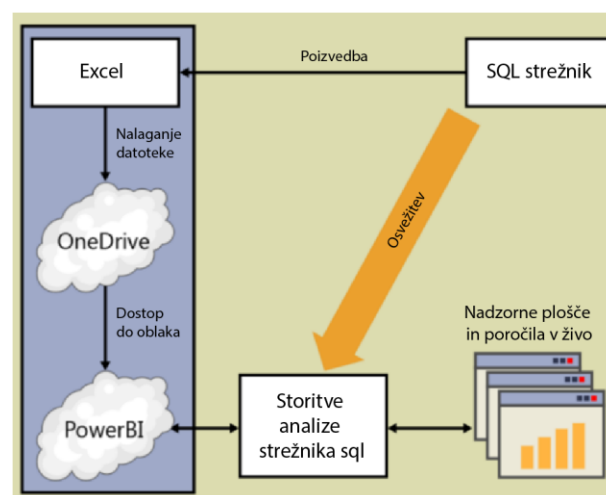
Podatki na splošno in poročila oziroma nadzorne plošče, ki jih kreiramo, se uporabljajo za pomoč pri odločanju ali pa za avtomatizirano odločanje. V povprečju podjetja uporabljajo podatke kot pomoč za odločanje v 58 %, za avtomatizacijo pa le v 29 %. Ko se odločamo za avtomatizacijo, se je potrebno zavedati nivoja tveganja, saj je od tega odvisno, do kolikšne mere bomo glede na podatke avtomatizirali določen proces. Pri bolj kompleksnih procesih in pomembnejših odločitvah, kot je na primer odobritev visokih kreditov, je avtomatizacija lahko vzpostavljena le v začetnih korakih, kasneje pa proces prevzame bančnik.

Orodja za avtomatizacijo odločanja so se razvila iz preprostih programskih stavkov, kot je »če se zgodi X, naredi Y«, in postajajo vse bolj sofisticirana. Učijo se iz vsakega primera posebej in sestavljajo kompleksne algoritme, ki se v mnogih primerih znajo odločiti bolje kot zaposleni (Economist Intelligence Unit, 2012). PricewaterhouseCoopers (PwC) raziskava je pokazala, da podjetja porabijo med 30 in 40 % manj ur za obdelavo rutinske dokumentacije, že pri uporabi najosnovnejših tehnik avtomatske ekstrakcije podatkov. Kljub temu podatku pa je le 28 % vodstvenih delavcev prioritiziralo uporabo umetne inteligence in strojnega učenja. Razlogov za to je lahko več, nekateri preprosto nimajo dovolj sredstev za vzpostavitev omenjenih orodij, drugi so skeptični o njeni resnični uporabnosti, tretji pa nimajo dovolj znanja, da bi lahko o tovrstnih rešitvah razmišljali (Bernard & Rao, 2021). Poznamo več načinov, kako lahko avtomatiziramo procese analitike. Omenili bomo nekaj najenostavnejših, ki se jih lahko poslužujejo podjetja, ki so v začetnih fazah postavitve naprednih sistemov poročanja.

- **Avtomatizirano zbiranje in integracija podatkov:** z uporabo različnih orodij in tehnologij lahko avtomatsko zbiramo podatke. To storimo s spletnim strganjem (angl. scraping) ali pa preko aplikacijskih programskih vmesnikov. Slednje znatno skrajša čas, ki je potreben za zbiranje in pripravo podatkov.
- **Avtomatizacija delovnih tokov:** orodja, kot so Alteryx, KNIME in RapidMiner, omogočajo avtomatizacijo ponavljajočih se nalog in delovnih tokov, kot so čiščenje in priprava podatkov ter izdelava modelov. S tem zagotovimo doslednost in natančnost ter zmanjšamo možnost za pojav nepravilnosti.
- **Uporaba strojnega učenja in umetne inteligence:** uporabiti jih je mogoče za avtomatizacijo različnih analitičnih nalog, kot so čiščenje podatkov, izbira funkcij, izdelava modelov in napovedi v realnem času.
- **Vnaprej izdelana programska oprema:** slednje je mogoče prilagoditi potrebam podjetja. Navadno se uporabljajo za avtomatizacijo različnih analitičnih nalog, kot so čiščenje podatkov, statistične analize in vizualizacije. S programi, kot sta Power BI in Tableau, lahko poročila popolnoma prilagodimo našim potrebam.
- **Izgradnja analitičnih orodij po meri:** s pomočjo programskih jezikov, kot sta R in Python, lahko podjetja razvijejo svoja orodja in jih integrirajo v obstoječe sisteme.
- **Avtomatizirano ustvarjanje poročil in nadzornih plošč:** slednje je mogoče načrtovati za samodejno posodabljanje in deljenje v določenem časovnem zamiku. S tem zmanjšamo potrebo po ročnem ustvarjanju poročil.

Bolj podrobno si bomo pogledali, kako delujeta avtomatizacija in posodabljanje v programu Power BI. Prišli smo do točke, ko imamo nadzorno ploščo postavljeno in na njej vidimo uporabne kazalnike. Da pa bodo informacije uporabne tudi čez pol leta ali eno leto, moramo zagotoviti posodabljanje nadzorne plošče oziroma podatkov. To še toliko bolj velja za podjetja, ki poslujejo sezonsko. Najprej se bomo posvetili osveževanju podatkov, ki je prikazano na sliki 6.

*Slika 5: Osveževanje podatkov*



*Prيرهjeno po Ferrari & Russo (2016).*



Večina uporabnikov pridobi nove podatke, na primer prodaje, iz računalniških sistemov in jih nato ročno kopira v glavno datoteko, ki jo posodobi v shrambi v oblaku. Power BI se nato samodejno poveže do shrambe in posodobi poročilo. Takšen način avtomatizira samo zadnji korak, še vedno pa je dovzeten za napake in časovno zamuden. Da zmanjšamo omenjena tveganja, si želimo avtomatizirati tudi prva dva koraka. Na sliki 6 je prikazano, kako se želimo izogniti ročnemu urejanju podatkov in s tem preskočiti korake 2, 3 in 4. SQL Server Analysis Services (v nadaljevanju SSAS) je modul z bazo podatkov v Power BI-ju, ki preko uporabniškega vmesnika izračuna vrednosti v nadzorni plošči oziroma poročilu. Od vira podatkov gremo tako direktno k SSAS-u in preskočimo korake, ki so obarvani z modro. Za tovrstno avtomatizacijo bomo morali prenesti in nastaviti še en del Power BI programske opreme – Personal Gateway. Slednji bo dostopal do naših dokumentov na lokalnem računalniku in jih samodejno posodabljal. Power BI ima tako vse informacije, ki jih potrebuje za osveževanje podatkov, tako na zahtevo kot tudi glede na urnik (Ferrari & Russo, 2016).

Mnoga podjetja si želijo optimizirati poslovanje tako, da sprejemajo odločitve v trenutku, ko opazijo spremembe pri obnašanju svojih kupcev. Za tovrstno odločanje je najprej potrebna dobra analitika, kasneje pa tudi dostopnost analitike na vsakem koraku. V programu Power BI lahko prilagodimo poročila in nadzorne plošče tudi mobilnim napravam. Slednje omogoča sledenje poslovanju tudi, ko nismo v službi ali za računalnikom. Ob postavljanju vizualizacij moramo biti previdni in se zavedati majhnosti ekrana mobilnih naprav. Razen omenjene omejitve, poročila in nadzorne plošče delujejo enako kot na namizni oziroma spletni verziji.

### **3.7 Napredni sistemi poročanja kot pomoč pri poslovnih odločitvah**

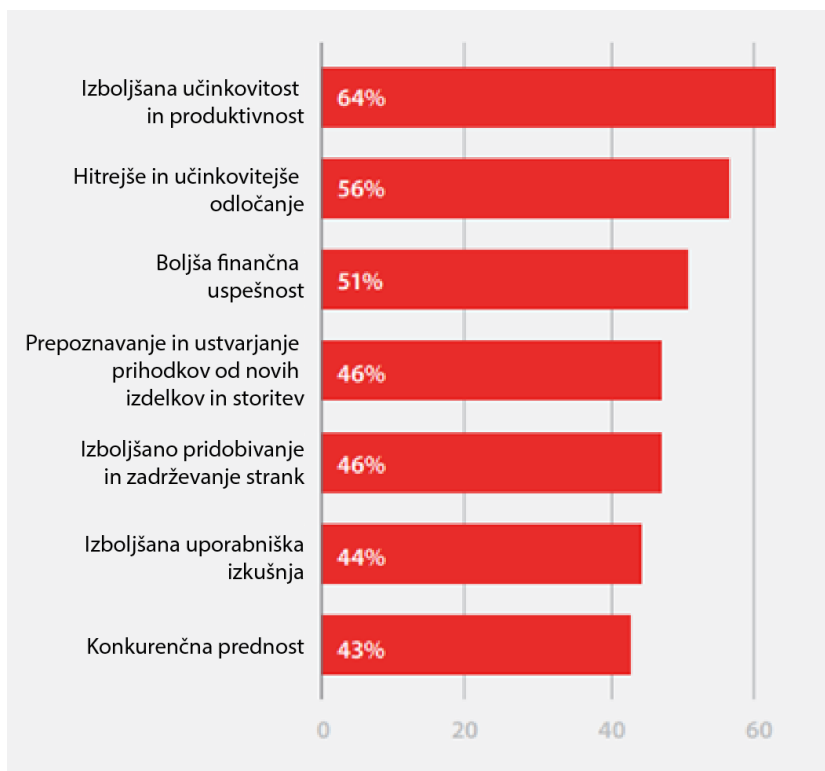
Kot smo že nekajkrat omenili, analitika igra pomembno vlogo pri poslovanju podjetij, še posebej pa pri sprejemanju odločitev. Raziskava, ki je analizirala številne študije, ki so proučevale pozitivne učinke analitike množičnih podatkov, je potrdila, da slednja znatno pripomore k uspehu podjetja. Analitika množičnih podatkov je bila v analiziranih podjetjih vzpostavljena v različnih oddelkih in na različnih nivojih, kar je pripomoglo k dodatnemu razvoju analitike in njene uporabe v podjetju. Zelo pomemben učinek je povečana učinkovitost poslovnih odločitev. Procesi v podjetju se poenostavijo in prilagodijo, strankam lahko omogočimo boljšo storitev, povečamo tržni delež, izboljšamo produktivnost dela, donosnost sredstev in pospešimo obrat zalog. To so le nekateri primeri, ki so povezani z vzpostavitvijo dobre analitike oziroma uporabo nadzornih plošč. Iz strateškega vidika so tovrstna podjetja bolj agilna, lažje predvidijo in se prilagodijo spremembam ter sčasoma pridobijo konkurenčno prednost. Kot lahko opazimo v zadnjih letih, analitika in avtomatizacija na različnih področjih obetata novo industrijsko revolucijo. Da pa nam podatkovni modeli zares pomagajo, potrebujemo tudi druga sredstva. V prejšnjih poglavjih smo že govorili o specialistih na posameznih področjih. Tukaj pa bomo omenili zaposlene, ki jim moramo zagotoviti dovolj znanja, da bodo znali na pravi način uporabljati nadzorne

plošče in jih bodo tudi razumeli (Mihai & Borza, 2019). Različni oddelki potrebujejo različne informacije o poslovanju, ki so navadno zbrani v enem poročilu ali nadzorni plošči. Cote (2021) omenja štiri oddelke in njihove potrebe v analitiki.

- **Marketing:** zaposleni v marketingu uporabljajo analitiko predvsem za pridobivanje informacij o trendih v industriji in za analitiko preteklih kampanj, s pomočjo katerih gradijo marketinške strategije.
- **Produkcija:** potrebuje informacije o trgih, industriji in uporabnikih, na podlagi katerih lahko izboljšajo prodajne produkte ali storitve.
- **Finance:** s pomočjo preteklih podatkov uspešnosti in podatkov iz industrije lahko lažje ter bolj natančno ocenijo bodoče finančno stanje podjetja.
- **Kadrovske službe:** s pomočjo mnenj zaposlenih, njihovih vpogledov, motivatorjev in vedenj ter trendov v industriji lahko prispevajo k pomembnim spremembam v podjetju.

Na sliki 7 lahko vidimo, da so podjetja kot najpomembnejšo prednost uporabe analitike navedla izboljšano učinkovitost in produktivnost, sledijo ji hitrejša in učinkovitejša sprejemanje odločitev, boljša finančna uspešnost, prepoznavanje in ustvarjanje novih produktov ter storitev, izboljšano pridobivanje in zadrževanje strank, izboljšana izkušnja strank in konkurenčna prednost (Microstrategy, 2020).

*Slika 6: Prednosti, ki jih je podjetjem prinesla uporaba analitike*



*Prirjeno po MicroStrategy (2020).*

Poleg že omenjenih prednosti naprednih sistemov poročanja je poročilo o globalnem stanju poslovne analitike identificiralo še nekaj novih prednosti. Skoraj polovica podjetij, ki so sodelovala v globalni raziskavi, je s pomočjo naprednih sistemov prepoznala in ustvarila nove produkte in posledično z njimi povečala prihodke. S pomočjo uporabe podatkov in analitike je 45 % podjetij razvilo tudi nove poslovne modele. Sklepamo lahko, da so s pomočjo slednjih podjetja spremenila svoj način dela in bila zmožna nadaljevati z uporabo napredne in napovedne analitike. Uporaba teh se je namreč glede na leto prej povečala za 7 %. Velika večina podjetij (94 %) navaja, da je uporaba analitike močno povezana z digitalno transformacijo (Microstrategy, 2020).

Za boljše razumevanje podatkovnih modelov za pomoč pri poslovnih odločitvah bomo predstavili tri znane primere. **Google** je s pomočjo analitike identificiral običajna vedenja zelo uspešnih vodij in ustvaril program usposabljanja za razvoj posebnih kompetenc, ki ga je imenoval »analitika ljudi«. V sklopu programa je pridobil podatke iz več kot 10.000 ocen in jih primerjal s stopnjo zadržanja zaposlenih. **Starbucks** sodeluje s podjetji na področju analitike lokacij, ki jim pomaga določiti najboljšo lokacijo za njihove trgovine na podlagi demografskih in prometnih vzorcev strank. Prav tako pa na podlagi regionalnih podatkov določi verjetnost za uspeh nove lokacije, preden vanjo investira. **Amazon** uporablja strojno učenje za priporočanje izdelkov na podlagi predhodnih nakupov uporabnikov in njihovega načina iskanja izdelkov. McKinsey je ocenil, da lahko 35 % nakupov v letu 2017 povežemo z Amazonovim sistemom priporočanja (Stobierski, 2019).

### 3.8 Trendi uporabe analize podatkov v prihodnosti

Paul Nannetti, nekdanji direktor prodaje in portfelja v družbi za storitve informacijske tehnologije in svetovanje – Capgemini, meni, da so danes podatki četrti proizvodni dejavnik, enako pomemben kot zemlja, delo in kapital (Economist Intelligence Unit, 2017). Podjetja, ki so sposobna analizirati množične podatke in iz njih z uporabo naprednih analitičnih tehnik generirati napovedi v realnem času, bodo imela v prihodnosti zagotovo veliko konkurenčno prednost (Economist Intelligence Unit, 2012). International data corporation (IDC) predvideva, da bo do leta 2025 60 % od 175 zeta bajtov obstoječih podatkov ustvarjenih s strani podjetij, medtem ko jih je bilo leta 2015 le 30 %. Slednje lahko izhaja tudi iz napovedi, da 65 % podjetij, ki so sodelovala v raziskavi o globalnem stanju poslovne analitike, načrtuje povečane investicije v analitiko v naslednjih letih. Hitrost, s katero napreduje tehnologija, bo v naslednjih letih vsaj tako velika kot danes, še verjetneje pa se zdi, da bo vedno večja. Računalništvo v oblaku, IoT, umetna inteligenca (angl. Artificial Intelligence – AI), strojno učenje (angl. Machine Learning – ML) in množični podatki izgledajo kot največji trendi, s katerimi se bodo morala podjetja soočiti v naslednjih petih letih. Skupaj z razvojem tehnologije in uporabo naprednih sistemov poročanja se bo spremenil tudi način dela. Poročilo o globalnem stanju podjetniške analitike, ki ga je objavilo podjetje Microstrategy (2020), izpostavlja, da bodo generacije, ki prihajajo v podatkovno vodena podjetja, imele tri ključne naloge:

- **Razmišljati odprto in sprejemati okolje, obkroženo z mnogimi orodji.** Podjetja se že danes ne morejo več zadovoljiti z uporabo enega orodja, ki bi pokrilo vse njihove potrebe. Zaposlenim je zato potrebno dovoliti raziskovanje podatkov in orodij, ki jim pomagajo pri ustvarjanju večje vrednosti v podjetju, tako danes kot tudi v prihodnosti.
- **Omogočiti, da so vsi procesi in aplikacije maksimalno izkoriščeni.** Ko se odločitve sprejemajo trenutek za trenutkom, morajo biti vpogledi dostopni vsako sekundo. Orodja, aplikacije in storitve morajo zato biti oblikovani glede na potrebe podjetja oziroma specifičnega oddelka. Dosegljiva morajo biti preko različnih medijev, na enostaven in jasn način.
- **Začeti s podatki in vpogledi.** Nadzorne plošče, ki so dostopne v oblaku, preproste za uporabo in varne, naj predstavljajo vir vpogledov v realnem času, tako v celotnem podjetju kot tudi v posameznih oddelkih.

Priča bomo tudi razvoju orodij in njihovi uporabi. Napovedi pravijo, da bo v ospredju pet faktorjev, s katerimi se bodo podjetja soočala in jih morala razviti za ohranjanje konkurenčne prednosti. Pri malih in srednjih podjetjih bomo opazili premik **BI orodij v oblak**, saj jim možnost plačila glede na uporabo bolj odgovarja. Poleg tega tovrstni sistemi omogočajo tudi povečanje prostora za upravljanje podatkov in analiz, v kolikor to podjetje potrebuje. Rast trga analitike v oblaku spodbujajo nizki stroški, enostavna implementacija, enostaven dostop do orodij in možnosti shranjevanja podatkov ter analiz. **Upravljanje podatkov** je gonilna sila poslovne analitike, ki se nanaša na formalne procese, ki zagotavljajo enotno kakovost podatkov in njihovo zaščito. V prihodnosti pričakujemo naraščanje regulativnih zahtev, kar bo zahtevalo več sodelovanja med ekipami. Prav tako se bo povečala količina množičnih, predvsem nestrukturiranih podatkov, za kar bo potreben pravilen pristop k njihovem upravljanju. Podjetja, ki bodo uspešno uporabljala programske rešitve za sledenje, pridobivanje, shranjevanje, vzdrževanje in odstranjevanje poslovnih podatkov, bodo poslovala uspešneje. K slednjemu bodo prispevala tudi **samopostrežna BI orodja**, ki ustvarjajo jasne vpogleds iz različnih vrst podatkov. S tem vodjem pomagajo razumeti poslovanje in zaradi enostavnosti uporabe graditi podatkovno kulturo v podjetju. Na drugi strani pa s pomočjo glasovnih pomočnikov lažje razumemo kupce in s tem prilagodimo prodajno strategijo. **Orodja za analizo govora** nam pomagajo tudi pri napovedih o namerah strank in uspešnosti naših storitev. Pričakuje se, da se bo v primerjavi z letom 2020 vrednost svetovnega trga govorne analitike do leta 2025 več kot podvojila. K temu bo pripomogla vse večja uporaba glasovnih pomočnikov za brskanje po spletu, nakupovanje in opravljanje storitev za stranke. Veliko pomoč pri razumevanju strank pa bo prinesla tudi **umetna inteligenca**, ki lahko že danes do določene mere samodejno analizira podatke iz več virov in prepozna trende. Prav tako pomaga pri čiščenju podatkov, ustvarjanju poročil, olajša razumevanje podatkov, zmanjša čas, ki je potreben za obdelavo velikih količin nestrukturiranih podatkov, in ustvari natančnejše vpogleds. Po mnenju Gartnerja bo do leta 2025 kar 75 % podjetij popolnoma implementiralo umetno inteligenco in s tem petkrat povečalo podatkovno ter analitično infrastrukturo. Pričakujemo lahko tudi povečano povpraševanje po internetu stvari, strojnem učenju, obdelavi naravnega jezika (angl. Natural

language processing – NLP) in drugih tehnologijah, ki bodo postale potrebne za podporo vsakdanjih procesov in analiz (Gitanjali, 2021).

Ne glede na vse trende, ki smo jih omenili, in vse ostale, ki jih bomo zasledili v prihodnosti, se moramo vedno zavedati, da niso samo množični podatki dragoceni. Bodimo pozorni tudi na ostale, ki nam lahko pomagajo pri razumevanju analiz. Tehnologije se bodo vedno spreminjale in razvijale, naša naloga pa je, da smo na te spremembe nenehno pozorni. Kajti, kar je veliko danes, je lahko jutri popolnoma nepomembno (Anderson, 2015).

## **4 PRIMER IZ PRAKSE**

Do sedaj smo govorili o tem, kakšne so možnosti za postavitev naprednih sistemov poročanja in kaj vse je mogoče narediti s pomočjo določenih orodij. V zadnjem delu pa bomo predstavili, kako je potekala postavitev nadzorne plošče za izbrano podjetje. Opisali bomo vse korake od načina pridobivanja podatkov, njihove priprave, izračuna kazalnikov s pomočjo DAX-a do postavitev vizualizacij, dodeljevanja pravic in avtomatizacije. Izkušnje in opažanja vodje podjetja bomo izpostavili na koncu poglavja, kjer predstavimo tudi ključne ugotovitve, priporočila in možnosti nadaljnjega razvoja.

Slovensko podjetje, ustanovljeno leta 2018, se ukvarja s proizvodnjo in prodajo kozmetičnih izdelkov, ki jih prodaja preko lastne spletne trgovine tako v Evropi kot tudi v Ameriki. Podjetje ima dva lastnika in zaposluje predvsem mlade sodelavce. Nenehno sledi trendom v svoji industriji in razvija nove izdelke. S svojim inovativnim pristopom, tako na področju ustvarjanja novih izdelkov, kot tudi na področju naprednih sistemov poročanja svoje poslovanje neprestano izboljšuje. V ta namen sta lastnika, skupaj z zaposlenimi in zunanji sodelavci, določila 11 ključnih kazalnikov uspešnosti, ki jih bomo vključili v njihov sistem poročanja.

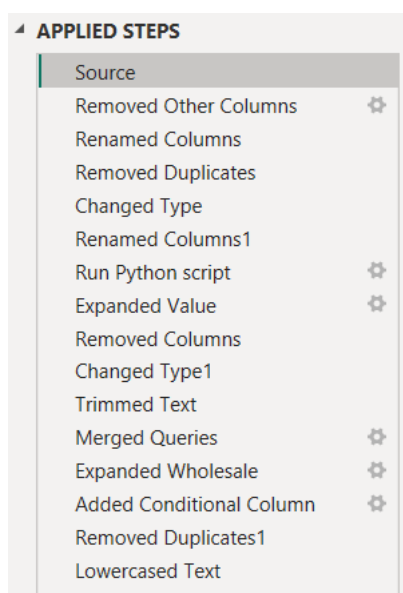
### **4.1 Pridobivanje podatkov preko vmesnika uporabniškega programa**

Izbrano podjetje uporablja program, ki omogoča vodenje podatkov in zagotavlja varno shranjevanje. Množične podatke ima shranjene na strežniku, s katerim se bomo povezali preko uporabniškega vmesnika (angl. application programming interface – API). Uporabili bomo knjižnico »requests«, ki pomaga pri pošiljanju zahteve na povezavo, od katere želimo pridobiti vsebino. Knjižnica »pandas« je vsestranska knjižnica v Pythonu, ki podatke predstavi v strukturirani obliki (as pd) in nudi mnogo možnosti za obdelovanje podatkov in pripravo analiz. Prav tako omogoča vhodni in izhodni vmesnik uporabniškega programa, ki ga v našem primeru potrebujemo. Za preprosto upravljanje vhodnih in izhodnih operacij bomo uporabili knjižnico »io«, ki omogoča branje in pisanje iz in v različne vire vnosov ter vključuje orodja za delo z različnimi vrstami podatkov (Python Software Foundation, 2023). V naslednjem koraku bomo s pomočjo tajnega ključa in ID podjetja dostopali do podatkov, ki jih potrebujemo za analizo. Funkcija »requests.post« bo poslala zahtevo za obdelavo teh



pred ali po besedilu. V naslednjem koraku s funkcijo »merge queries« združimo poizvedbe s pomočjo levega spoja (angl. left outer joint), ki združi vse vrstice iz leve (prve) tabele z desno (drugo) tabelo. Ker želimo pri strankah razločevati med končnimi strankami in trgovci, s funkcijo »expand table columns« izvlečemo ugnezdene podatke in stolpec »Wholesale.Email« umestimo v tabelo. Nato dodamo pogojni stolpec, ki izpiše »B2C« v primeru, da podatek v celici »Wholesale.Email« obstaja, v nasprotnem primeru pa izpiše »B2B«. V zadnjem koraku spremenimo vse velike črke v male s pomočjo funkcije »Text.Lower«.

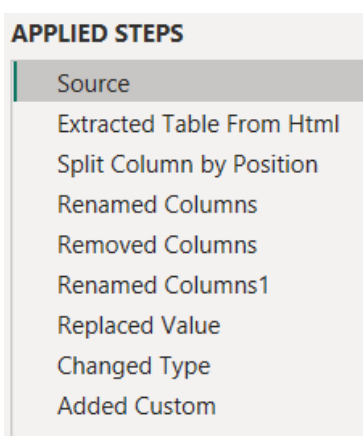
Slika 8: Koraki pri čiščenju in pripravi podatkov v tabeli stranke



*Vir: lastno delo.*

Izbrano podjetje posluje v različnih državah, tako evropskih kot tudi neevropskih. Zaradi poslovanja v različnih valutah je potrebno analize prilagoditi tako, da prikazujejo realno stanje. Na sliki 10 lahko vidimo vse korake, potrebne za pripravo nove tabele, ki jo pridobimo iz spletnega mesta. S pomočjo funkcije »Web.BrowserContents« dostopamo do spletnega vira, v našem primeru je to spletna stran SKB banke, kjer so objavljeni tečaji. Iz nje izvlečemo podatke iz tabele (funkcija »Html.Table«), pri čemer določimo katere parametre želimo. V naslednjem koraku razdelimo en stolpec v tabeli v več stolpcev. S funkcijo »Table.SplitColumn« določimo, kateri stolpec želimo razdeliti, na katerem mestu in kako novo nastale stolpce poimenujemo. Nato odstranimo stolpce, ki jih ne potrebujemo več in zamenjamo določene vrednosti, kot so ločila in naslovi vrednosti. Tako kot v prejšnji tabeli, tudi tukaj preverimo vrsto podatkov in jih po potrebi s funkcijo »Table.TransformColumnTypes« spremenimo v ustrezno vrsto. V zadnjem koraku dodamo stolpec, kjer izračunamo srednje vrednosti nakupnega in prodajnega tečaja.

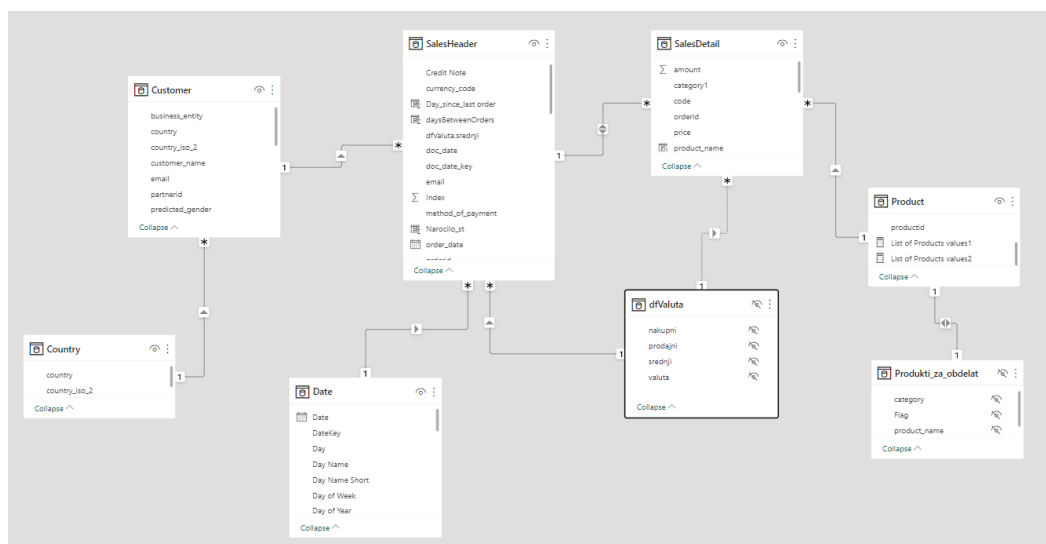
Slika 9: Čiščenje in priprava podatkov pri kreiranju nove tabele



Vir: lastno delo.

Na zgoraj opisan način pripravimo vse tabele glede na to, katere podatke bomo potrebovali in v kakšni obliki. V naslednjem koraku jih je potrebno povezati, da lahko kombiniramo podatke iz različnih tabel. Relacijski model postavimo tako, da v vsaki tabeli poiščemo ujemajoče-se podatke. Najpogostejša je enosmerna povezava s kardinalnostjo mnogo proti ena (angl. many to one). Slednjo lahko vidimo na sliki 11 v večini primerov, na primer med tabelama »SalesHeader« in »Customers«, ki sta povezani s podatkom »partnerid«. V primerih, ko imamo v eni tabeli zapis, ki se ujema samo z enim edinstvenim zapisom v drugi tabeli, uporabimo kardinalnost ena proti ena (angl. one to one). To povezavo prikazujeta tabeli »Product« in »Produkti\_za\_obdelat«. V nadaljevanju bomo omenili še eno tabelo z imenom »dfProducts«, ki jo v tem relacijskem modelu ne vidimo. Slednja vsebuje vse podatke, ki jih sicer vidimo na računu (ime in kodo produkta, številko računa, valuto, znesek in količine). Tabela bo pomembna pri štetju edinstvenih transakcij.

Slika 10: Relacijski model v programu Power BI



Vir: lastno delo.



### 4.3 Priprava kazalnikov v DAX-u

Priprava kazalnikov oziroma KPI-jev je zelo pomembna naloga zaposlenih. V izbranem podjetju so se že pred vzpostavitvijo naprednega sistema poročanja zavedali potrebe po tovrstnih kazalnikih, vendar jih nikoli prej niso opredelili, omenja direktor podjetja. Zato so skupaj z zaposlenimi in zunanji partnerji začeli z zbiranjem idej (angl. brainstorming). Ker so zaposleni predvsem mladi, so dobro razumeli pomembnost kreiranja dobrih kazalnikov in prispevali veliko spodbudnih idej. Glavni namen je bil, da se pripravi kazalnike, ki bodo pomagali zaposlenim na najvišjih položajih pri strateških odločitvah, marketingu pri čim boljšem poznavanju strank in lastnikom pri boljšem razumevanju celotnega poslovanja. V ta namen so določili 11 kazalnikov, ki zajemajo stranke, vrednosti naročil, produkte in povezave med njimi, plačila in nezaključena naročila. Vse kazalnike, ki si jih bomo podrobneje ogledali v nadaljevanju, bomo v končni nadzorni plošči prikazali po vseh trgih skupaj, po posameznih trgih, v različnih obdobjih in naredili predikcijo spola glede na ime strank. Vsaka stran nadzorne plošče je interaktivna, kar pomeni, da se s klikom na določene vrednosti ostale vizualizacije prilagodijo izbranemu podatku. Vsi izrezki zaslona, ki bodo predstavljeni v nadaljevanju, so vzeti iz realnega poročila izbranega podjetja. Prikazane vrednosti so zaradi zagotavljanja anonimnosti podjetja izmišljene.

**Življenjska vrednost strank** (angl. Customer life time value) je prvi in po besedah direktorja izbranega podjetja najpomembnejši kazalnik, ki predstavlja seštevek vseh nakupov posamezne stranke. S pomočjo kazalnika podjetje lažje segmentira stranke glede na frekvenco njihovih nakupov in posledično optimizira marketinške kampanje. Življenjsko vrednost stranke smo izračunali v DAX-u s pomočjo funkcije »countrows«, ki prešteje naročila vseh strank. Funkcija »summarize« v virtualni tabeli pa združi število naročil in stranke (McKay, 2018). Celoten izračun lahko vidimo na sliki 12.

*Slika 11: DAX - Življenjska vrednost stranke*

```
1 LTV = COUNTROWS(SUMMARIZE(SalesHeader, SalesHeader[email], "Total Purchases", COUNTROWS(SalesHeader)))
```

*Vir: lastno delo.*

**Število novih strank in strank, ki se vračajo**, sta dva kazalnika, ki pomagata pri razumevanju strukture strank in načrtovanju marketinških kampanj. Pomembno je tudi ravnovesje med novimi in ponavljajočimi strankami, saj so stroški pridobivanja novih strank navadno veliko višji, v primerjavi s sredstvi, ki jih namenimo za ponavljajoče. Število novih strank izračunamo v DAX-u, tako kot vidimo na sliki 13. S pomočjo te kode preštejemo število vrstic (konkretno elektronskih naslovov), pri čemer upoštevamo prvi datum, ko je vsaka stranka opravila nakup. Nekoliko enostavneje je izračunati ponavljajoče stranke, saj preprosto preštejemo vse stranke in odštejemo tiste, ki so nove.

Slika 12: DAX - Nove stranke

```
1 NewCustomers =
2     COUNTROWS (
3         FILTER (
4             CALCULATETABLE (
5                 ADDCOLUMNS (
6                     VALUES ( Customer[email] ),
7                     "DateOfFirstBuy", CALCULATE ( MIN ( SalesHeader[order_date] ) )
8                 ),
9                 ALL ( 'Date' )
10            ),
11            CONTAINS (
12                VALUES ( 'Date'[Date] ),
13                'Date'[Date],
14                [DateOfFirstBuy]
15            )
16        )
17    )
```

Vir: lastno delo.

**Povprečna vrednost nakupa za nove in ponavljajoče stranke** se izračuna na zelo podoben način, kot smo videli na sliki 13, saj za nove stranke prav tako potrebujemo datum prvega nakupa in njihov elektronski naslov. Namesto da preštejemo naslove, izračunamo povprečno vrednost nakupov s pomočjo funkcij »calculate« in »average«, ki ju postavimo na začetek DAX kode. Ko pa računamo povprečno vrednost nakupa za ponavljajoče stranke, ustvarimo nekaj novih spremenljivk, ki jih vidimo na sliki 14. Določimo datum prvega nakupa, dodelimo največjo in najmanjšo vrednost stolpca »order\_date«, izberemo elektronske naslove in z zadnjo spremenljivko določimo zadnji datum nakupa. S pomočjo omenjenih spremenljivk v predzadnjem koraku filtriramo stranke, ki so v izbranem obdobju opravile več kot en nakup. Na koncu s funkcijo »averagex« izračunamo povprečno vrednost nakupa.

Slika 13: DAX - Povprečna vrednost nakupa strank, ki se vračajo

```
1 AOV_returning_customer =
2 var _firstorderdate=SalesHeader[First order date]
3 var _minselecteddate=min(SalesHeader[order_date])
4 var _maxselecteddate=max(SalesHeader[order_date])
5 VAR currentCustomers = VALUES(SalesHeader[email])
6 VAR currentDate = MAX(SalesHeader[order_date])
7
8 VAR returningCustomers = filter(CALCULATETABLE(VALUES(SalesHeader[email])),
9     _firstorderdate<=currentDate &&
10     COALESCE(SalesHeader[St_narocil],0)>1)
11
12 RETURN AVERAGEX(returningCustomers, CALCULATE(AVERAGE(SalesHeader[sum_all_EUR])))
```

Vir: lastno delo.

Ko govorimo o strankah, ki se vračajo, se nam vedno porajajo še dodatna vprašanja. Naslednji kazalnik tako obravnava **čas med prvim in drugim nakupom in ostalimi nakupi**. Ta informacija je zanimiva predvsem za marketinški oddelek, saj lahko z oglaševanjem vpliva na pogostejše nakupe strank. Da bomo kasneje lahko pravilno prikazali omenjen kazalnik, moramo v tem koraku izračunati dve stvari. Najprej moramo določiti prvi, drugi, tretji in vsak naslednji nakup, nato pa moramo izračunati, koliko je bilo v povprečju dni med posameznimi nakupi. S pomočjo novih spremenljivk »cur\_customer« in »tmp« filtriramo tabelo »SalesHeader« tako, da prešteje samo elektronske naslove, ki se ujemajo s prvo spremenljivko – »cur\_customer«. V zadnjem koraku kode s funkcijo »RANKX« vsaki vrstici v tabeli dodelimo številko naročila, glede na datum naročila, v naraščajočem vrstnem redu. Funkcija »IF« nam prikaže rezultat, razen če trenutna stranka (cur\_customer) ne obstaja. Opisane korake lahko vidimo na sliki 15.

*Slika 14: DAX - Zaporedna številka naročila*

```

1 Narocilo_st =
2 VAR cur_customer = SalesHeader[email]
3 VAR tmp =
4 |   FILTER ( ALLSELECTED ( SalesHeader ), SalesHeader[email] = cur_customer )
5 RETURN
6 |   if (cur_customer="",1,RANKX ( tmp,SalesHeader[order_date],, ASC, DENSE))

```

*Vir: lastno delo.*

Razliko med posameznimi nakupi izračunamo s pomočjo spremenljivke »razlika«, ki izračuna razliko med naročili v dnevih. S funkcijo »IF« preverimo, ali je trenutno naročilo prvo naročilo določene stranke. V kolikor je, dobimo rezultat »0«, v nasprotnem primeru pa izračunano razliko. Funkcija »return« prikaže rezultat z dvema decimalnima mestoma, kar smo določili z drugo spremenljivko. Na sliki 16 lahko vidimo celoten postopek.

*Slika 15: DAX - Število dni med nakupi*

```

1 daysBetweenOrders =
2 var Razlika=if (SalesHeader[Narocilo_st]=1,0,SalesHeader[order_date]-SalesHeader[PreviousOrderDate])
3 var Ralika=FORMAT(Razlika,"Fixed")
4
5 return value(Ralika)

```

*Vir: lastno delo.*

Naslednji kazalnik, ki ga vidimo na sliki 17, prikazuje **povprečno število naročil na stranko**. Najprej ustvarimo novo mero, pri čemer izračunamo število naročil tako, da uporabimo funkcijo »COUNTROWS«. To mero nato uporabimo v našem končnem izračunu, kjer delimo število naročil z elektronskimi naslovi strank. Funkcija »DISTINCTCOUNT« omogoči, da se ponavljajoči naslovi ne upoštevajo, saj ne želimo podvajanj. V vizualizacijah bomo ta kazalnik pomnožili s 1000 in tako dobili podatke o povprečnem številu nakupov na 1000 strank.

Slika 16: DAX - Povprečno število nakupov na stranko

```
1 Average number of orders = divide(SalesHeader[St_narocil],DISTINCTCOUNT(SalesHeader[email]))
```

*Vir: lastno delo.*

Z naslednjim kazalnikom želimo ugotoviti, kakšne so **povezave med različnimi produkti**. Za iskanje povezav med dvema produktoma najprej potrebujemo vrednosti iz kataloga produktov. V našem primeru jih označimo z »item« in »item2«. Nato med njima naredimo povezave tako, da poiščemo vse možne kombinacije, slednje shranimo pod imenom »Basket«. Nato izračunamo, kako pogosto se določene kombinacije pojavljajo. Izračun imenujemo »Support basket« in ga izračunamo tako, kot vidimo na sliki 18. Najprej z novima spremenljivkama določimo, kateri produkti nas zanimajo. Nato s funkcijo »SELECTCOLUMNS« pripravimo tabelo, ki vključuje vse transakcije, kjer se je pojavil »item1«. V tabeli »dfProducts« s funkcijo »FILTER« poiščemo samo tiste vrstice, kjer se stolpec »category1« in »item1« ujemata. Spomnimo se, da tabela »dfProducts« vsebuje vse transakcije, ki so bile kadarkoli narejene, zato v njej lahko poiščemo ujemajoče se produkte. Enako naredimo za »item2«. Nato s funkcijo »INTERSECT« ustvarimo tabelo, ki vključuje samo transakcije, ki se nahajajo v obeh prej ustvarjenih tabelah. V zadnjem koraku delimo število transakcij, ki so vključevale »item1« in »item2«, s skupnim številom transakcij v tabeli »dfProducts«.

Slika 17: DAX - Povezave med kupljenimi produkti

```
1 Support basket =  
2  
3 var item1=[Item]  
4 var item2=[Item2]  
5  
6 var transactionsWithItem1 = SELECTCOLUMNS(FILTER('dfProducts', 'dfProducts'[category1]=item1), "transactionID", [mk_id])  
7  
8 var transactionsWithItem2 = SELECTCOLUMNS(FILTER('dfProducts', 'dfProducts'[category1]=item2), "transactionID", [mk_id])  
9  
10 var transactionsWithBothItems = INTERSECT(transactionsWithItem1, transactionsWithItem2)  
11  
12 RETURN COUNTROWS(transactionsWithBothItems) / DISTINCTCOUNT(dfProducts[mk_id])
```

*Vir: lastno delo.*

Nato izračunamo interval zaupanja (angl. confidence), ki nam pove, v koliko odstotkih je bil drugi produkt kupljen skupaj s prvim oziroma obratno. Iz slike 19 je razvidno, da nas v tem primeru zanima »Item«, ki je naša prva spremenljivka, druga spremenljivka pa je »numberOfTransactions«, ki prešteje edinstvene transakcije v tabeli »dfProducts«. V tretjem koraku izračunamo »supportItem1« tako, da delimo število transakcij, v katerih je bil kupljen »item1«, s skupnim številom transakcij. V zadnjem koraku pa delimo »Support basket«, ki smo ga predhodno izračunali, s »supportItem1«. Končni izračun predstavlja zaupanje, da je »item1« del košarice, ko ta že vsebuje »item2«. Enak postopek ponovimo za »Confidence Item 2«.

Slika 18: DAX - Stopnja zaupanja

```
1 Confidence Item 1 -> Basket =
2
3 var item1 = [Item]
4 var numberOfTransactions = DISTINCTCOUNT(dfProducts[mk_id])
5
6 var supportItem1 = COUNTROWS(FILTER('dfProducts', 'dfProducts'[category1]=item1))/ numberOfTransactions
7
8 RETURN [Support basket]/supportItem1
```

*Vir: lastno delo.*

Še zadnji izračun pri kazalniku, ki predstavlja povezave med produkti, je »Lift«, ki ponazarja verjetnost pojava določene kombinacije. Izračuna se na enak način, kot je prikazano na sliki 19, z razliko v zadnjem koraku, kjer v imenovalcu upoštevamo tako »supportItem1« kot tudi »supportItem2«. Celoten postopek lahko vidimo na sliki 20.

Slika 19: DAX - Verjetnost pojava kombinacij (lift)

```
1 Lift =
2
3 var item1 = [Item]
4 var item2 = [Item2]
5
6 var numberOfTransactions = DISTINCTCOUNT(dfProducts[mk_id])
7
8 var supportItem1 = COUNTROWS(FILTER('dfProducts', 'dfProducts'[category1]=item1))/numberOfTransactions
9
10 var supportItem2 = COUNTROWS(FILTER('dfProducts', 'dfProducts'[category1]=item2))/numberOfTransactions
11
12 RETURN [Support basket]/(supportItem1*supportItem2)
```

*Vir: lastno delo.*

Osmi kazalnik prikazuje **povezavo med strankami, ki se vračajo, in produkti**, ki jih kupujejo. Ker smo na začetku že izračunali število novih strank, v tem koraku ne potrebujemo dodatnih izračunov, razen ko želimo ponazoriti število strank v odstotkih. V tem primeru s funkcijo »DIVIDE« delimo število strank, ki se vračajo, s številom vseh strank. Tudi za naslednji kazalnik imamo vse vrednosti že izračunane. Zanima nas, **koliko katerega izdelka je kupljenega v prvem, drugem in vsakem naslednjem nakupu**. Za prikaz potrebujemo številko nakupa (izračun na sliki 15) in količine kupljenih izdelkov, ki jih imamo že podane v osnovni tabeli. Za zadnja dva kazalnika ne potrebujemo dodatnih izračunov, saj imamo vse potrebne podatke že pripravljene v osnovnih tabelah. Kazalnika prikazujeta **odstotek strank glede na vrsto plačila in odstotek vračil oziroma neprevzetih paketov glede na tip plačila**.

Za zaključek želi podjetje videti tudi razlike med nakupnimi navadami moških in ženskih strank. V ta namen naredimo **predikcijo spola glede na ime stranke**. Za predikcijo spola bomo uporabili Python, kamor najprej uvozimo knjižnico »pandas« in »gender\_guesser.detector«. Knjižnica razlikuje med šestimi spoli, glede na to, kolikšna je verjetnost, da ime predstavlja določen spol. Ker tako natančnih podatkov ne potrebujemo, določimo, da je spol »mostly\_female« enak spolu »female«, »mostly\_male« enak »male« in

»andy« enak »unknown«. Spol »andy« sicer pomeni, da se ime lahko uporablja tako za moškega kot tudi za žensko, zato spola ni mogoče določiti, medtem ko vrednost »unknown« pomeni, da imena ni bilo mogoče najti v bazi imen (Python Software Foundation, 2016). Celoten postopek za predikcijo spola glede na ime stranke lahko vidimo na sliki 21. Ko je model vzpostavljen, stolpec z imenom »predicted\_gender« umestimo v tabelo »Customer«.

Slika 20: Model za predikcijo spola glede na ime stranke

```
import pandas as pd
import gender_guesser.detector as gender

def uniteGenders(gender):
    if gender == "mostly_female":
        result = "female"
    elif gender == "mostly_male":
        result = "male"
    elif gender == "andy":
        result = "unknown"
    else:
        result = gender
    return result

d = gender.Detector()

dataset = pd.DataFrame(dataset) # convert dataset to a pandas DataFrame

dataset['ime'] = dataset['customer_name'].str.split(' ').str[0]
dataset['predicted_gender'] = dataset['ime'].apply(lambda x: d.get_gender(x))
dataset = dataset.drop(columns=['ime'])
dataset['predicted_gender'] = dataset['predicted_gender'].apply(lambda x: uniteGenders(x))

dataset = dataset.values.tolist() # convert dataset back to a list
```

Vir: lastno delo.

S pomočjo DAX-a kreiramo tudi dinamično kazalo, ki prikazuje trenutno uporabljene filtre. To kazalo se bo prikazovalo na dnu vsake strani nadzorne plošče. Ker uporabljamo sedem filtrov (gender, month, year, sales type, status code, products in country), naredimo sedem kazalnikov, za vsako vrednost enega. Kot primer vzemimo vrednost »Year« na sliki 22. Funkcija »ISFILTERED« preveri, ali je vrednost filtrirana, če v kazalo doda »Year« in filtrirane vrednosti, ki jih loči z vejico. V nasprotnem primeru vrne »ALL«, kar ponazarja, da so izbrane vse vrednosti. Ko na ta način izračunamo vrednosti za vse filtre, kreiramo besedilo na dnu strani, kjer namesto besedila izberemo »vstavi dinamično vrednost«. Dinamična vrednost se bo avtomatsko posodabljala, ko bomo izbirali vrednosti v filterih.

Slika 21: Kreiranje dinamičnih vrednosti

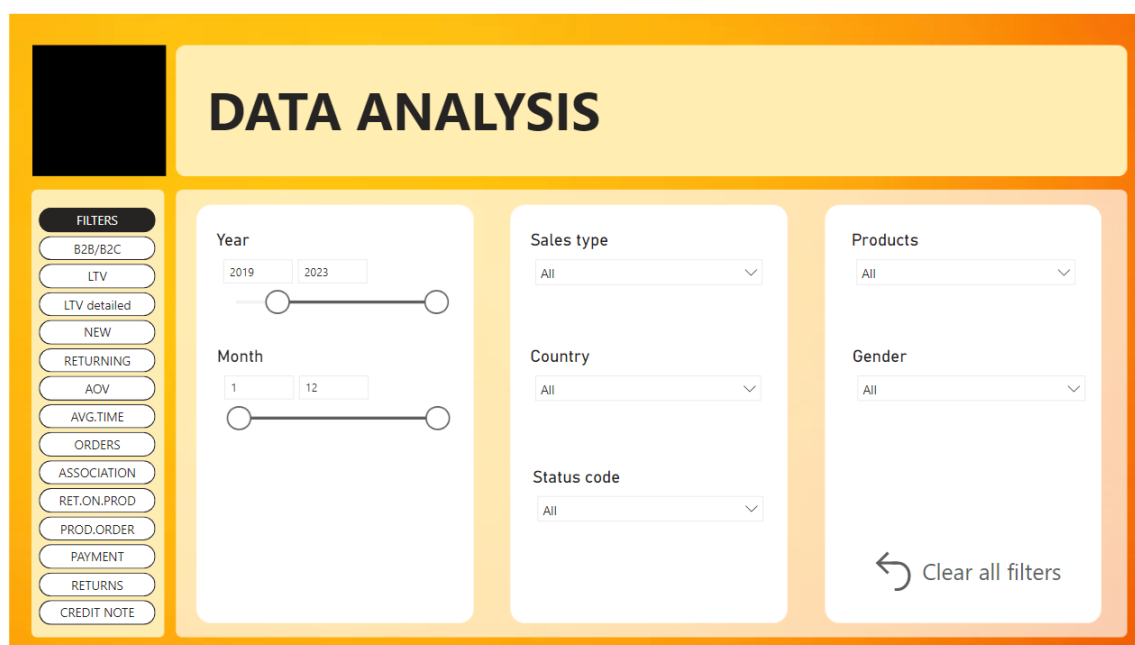
```
1 List of Year values =
2 IF (
3     ISFILTERED ( 'Date'[Year] ),
4     "Year: " &
5     CONCATENATEX (
6         VALUES ( 'Date'[Year] ),
7         'Date'[Year],
8         ", "
9     ),
10    "ALL"
11 )
```

Vir: lastno delo.

#### 4.4 Vizualizacija podatkov

Prispeli smo do stopnje, kjer imamo vse podatke pripravljene za vizualizacijo. V tem koraku je pomembno, da izberemo prave vizualizacije za določen tip podatkov. Poleg tega moramo omogočiti tudi čim bolj enostavno uporabo nadzorne plošče. V ta namen na prvi strani kreiramo filtre, za katere želimo, da vplivajo na celotno nadzorno ploščo, in kazalo vseh strani, slednje lahko vidimo na sliki 23.

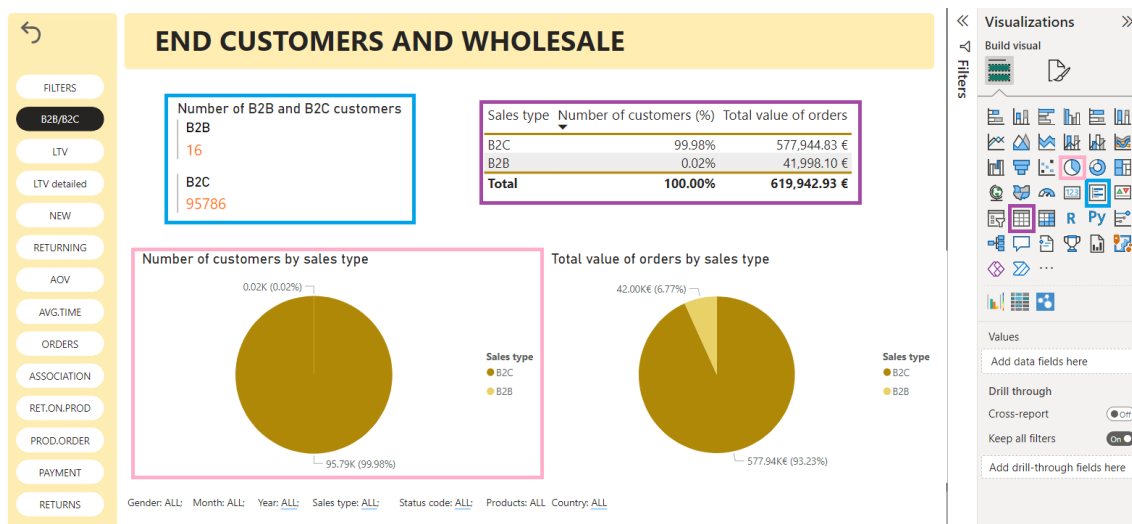
Slika 22: Prva stran nadzorne plošče, s filtri in kazalom



Vir: lastno delo.

Druga stran nadzorne plošče prikazuje delež končnih kupcev (angl. Business-to-customer – B2C) in veleprodajo (angl. Business-to-business – B2B). Kot vidimo na sliki 24, imamo predstavljeno tako število posameznih kupcev kot tudi vrednost nakupov. Iz tortnih diagramov vidimo, da trgovci po številu predstavljajo zelo majhen delež, kljub temu pa vrednost njihovih nakupov predstavlja skoraj 7 % vseh nakupov. Za boljšo predstavo smo dodali tudi tabelo z vrednostmi nakupov v evrih. Vsako vizualizacijo smo naredili tako, da smo jo izbrali na desni strani v zavihku »Visualizations« in nato med podatki izbrali vrednosti, ki jih želimo prikazati. Vsako posebej tudi dodatno uredimo, dodamo naslov, spremenimo barve, velikosti podatkov in po potrebi dodamo legende ter uredimo prikazane vrednosti. Ikone in končne vizualizacije so označene z barvnimi pravokotniki na sliki 24. Kot smo že omenili na koncu prejšnjega podpoglavja, smo na dno strani dodali dinamično kazalo, ki se samodejno posodablja glede na izbrane vrednosti v filtrih.

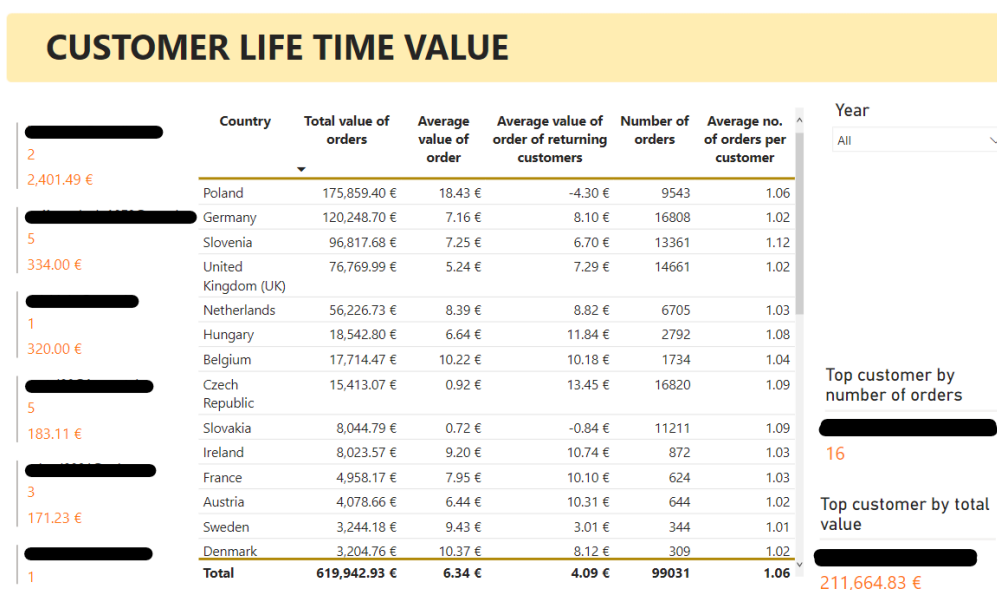
Slika 23: Izbira vizualizacije - Razmerje med končnimi kupci in trgovci



Vir: lastno delo.

Nadaljujemo s sliko 25, ki prikazuje kazalnik življenjske vrednosti kupcev. Na levi strani je prikazanih šest največjih strank po vrednosti nakupov, v tabeli na sredini vidimo vrednost vseh nakupov, povprečno vrednost nakupa, povprečno vrednost nakupa strank, ki se vračajo, število nakupov in povprečno število nakupov na stranko po določenih državah. Tabela ima možnost podrobnejšega ogleda s funkcijo »drill through«, ki prikaže število naročil, vrednost naročil in povprečno vrednost naročil po posameznih strankah za določeno državo. Na tej strani imamo možnost dodatnega filtriranja po letih.

Slika 24: Življenjska vrednost strank

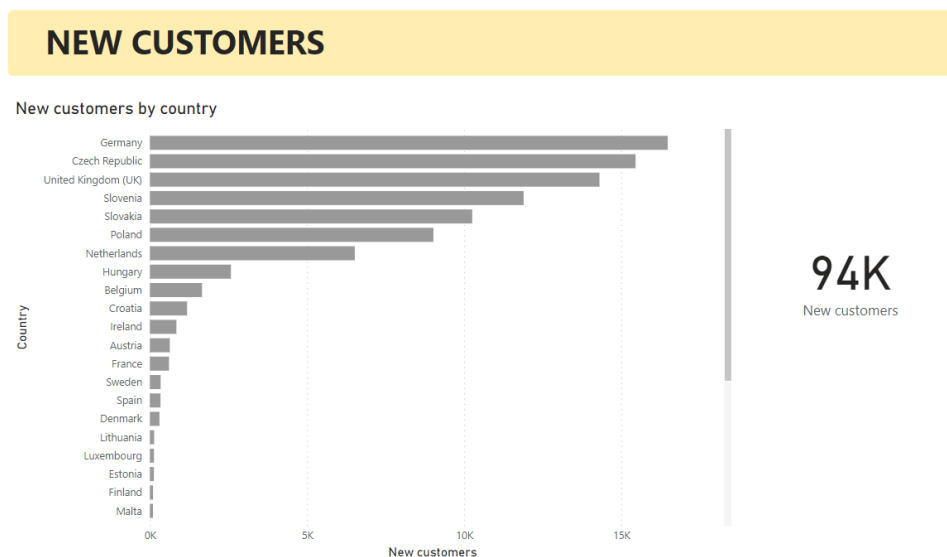


Vir: lastno delo.



Za prikaz števila novih strank uporabimo prikaz po državah in kartico na desni strani, ki prikazuje skupno število vseh novih strank. V kolikor želimo videti točno število za posamezno državo, kliknemo nanjo in s tem posodobimo tudi kartico na desni strani. Prikaz tega kazalnika je viden na sliki 26. Na enak način prikažemo tudi število strank, ki se vračajo.

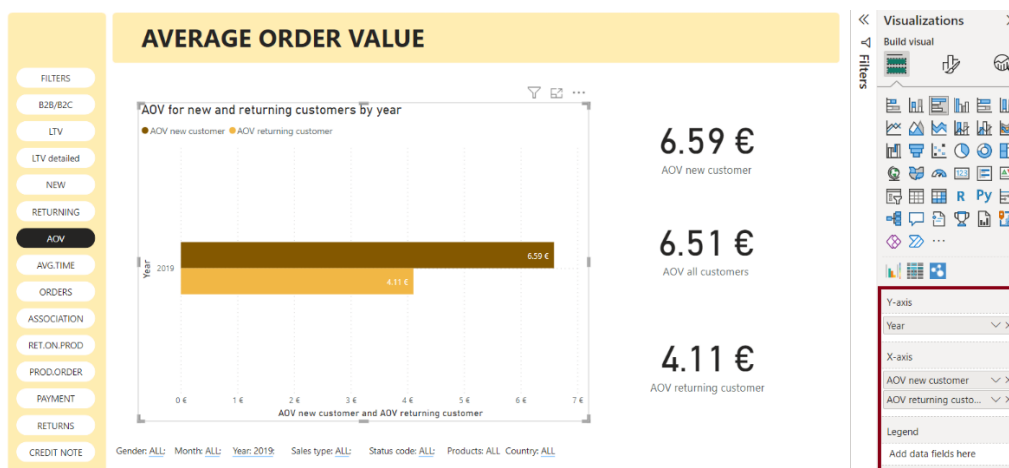
Slika 25: Prikaz števila novih strank po državah



Vir: lastno delo.

V prejšnjem podpoglavju smo izračunali povprečne vrednosti nakupov tako za nove stranke kot tudi za tiste, ki se vračajo. Obe vrednosti prikažemo na isti vizualizaciji, da hitreje in lažje opazimo razliko med njimi. Na sliki 27 vidimo palični grafikon, kjer so prikazane vrednosti za leto 2019. V rdečem pravokotniku pa so desno spodaj označena polja, kamor pripeljemo vrednosti, ki jih želimo v vizualizaciji. To polje in njegove vrednosti se spreminjajo glede na to, katere podatke določena vizualizacija potrebuje.

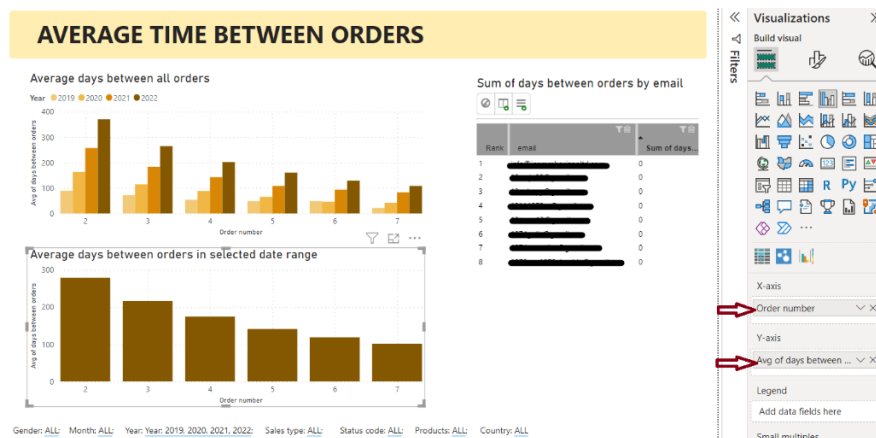
Slika 26: Dodajanje vrednosti v vizualizacijo - Povprečne vrednosti nakupov



Vir: lastno delo.

Na naslednji strani nadzorne plošče z dvema grafoma prikažemo povprečen čas med nakupi. Prvi graf, kljub nastavljenim filtrom za obdobje med 2019 in 2022 preveri, ali je kupec kupoval že pred tem obdobjem, medtem ko drugi graf striktno upošteva obdobje, ki smo ga določili v filtrih. Na obeh grafih preimenujemo tudi osi. To najlažje naredimo kar v poljih, kamor smo pripeljali podatke. Na sliki 28 je slednje označeno z rdečima puščicama. Na drugem delu strani dodatno ustvarimo še tabelo, ki rangira kupce glede na njihovo pogostost nakupovanja. Nič dni med nakupi pomeni, da je kupec oddal dva naročila v istem dnevu oziroma v najkrajšem možnem intervalu, ki ga podjetje upošteva, ko zabeleži naročilo.

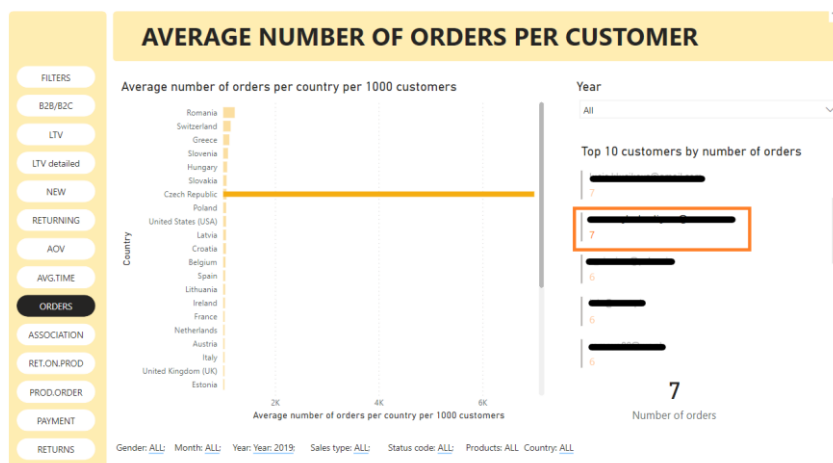
Slika 27: Preimenuvanje osi na grafih - Povprečen čas med nakupi



Vir: lastno delo.

Povprečno število nakupov na stranko prikažemo po državah, zraven pa dodamo še 10 največjih kupcev. Ko kliknemo na enega izmed njih, na sliki 29, kjer je označen drugi kupec, se tudi graf na levi strani posodobi. Tako vidimo, iz katere države prihaja kupec in primerjavo s povprečnim številom nakupov za njegovo državo. Ob kliku se posodobi tudi kartica, desno spodaj, ki prikazuje število nakupov glede na izbrane vrednosti.

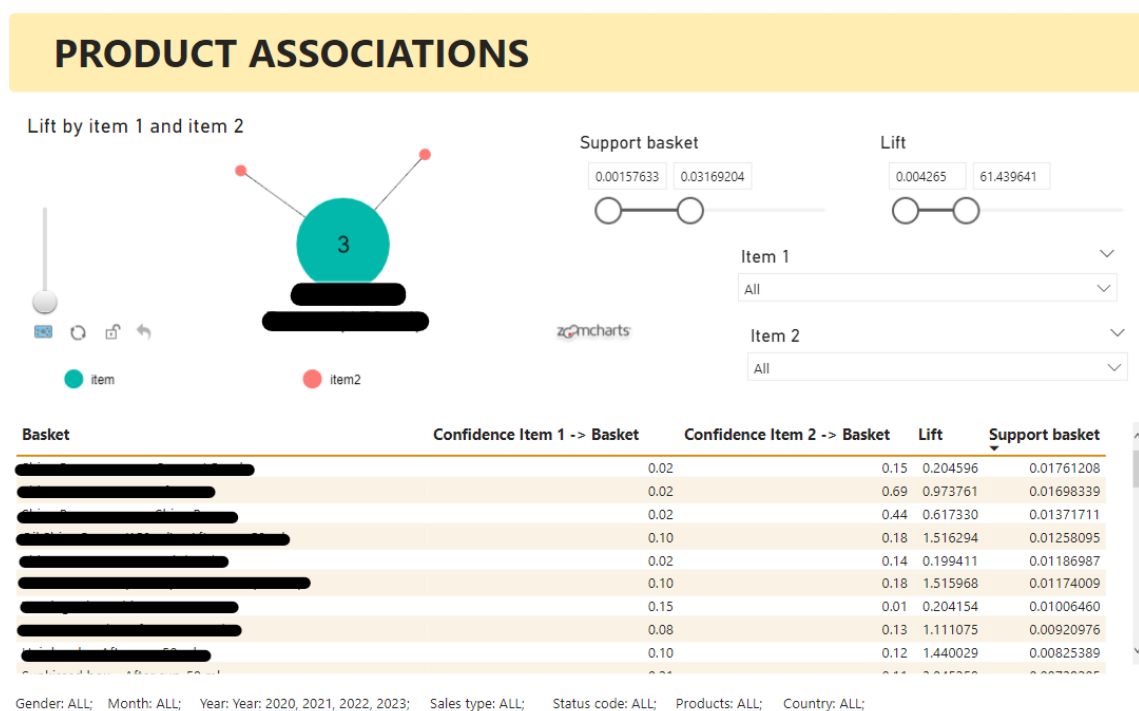
Slika 28: Povprečno število nakupov po državah in prikaz največjih kupcev



Vir: lastno delo.

Na naslednji strani postavimo model za prikaz povezanih produktov in tabelo vrednosti. Da bodo zaposleni lažje iskali povezave, postavimo tudi filtre za support, lift in iskanje produktov. Naredili smo jih tako, da smo izbrali vizualizacijo »slicer« in nato v nastavitvah vizualizacije določili vrsto filtra. Zaradi zagotavljanja anonimnosti podjetja smo imena produktov skrili. Slednje bi sicer videli v prvem stolpcu (Basket). Najprej je naveden prvi produkt, ki ga ima strank v košarici, nato pa še drugi, ki ga je poleg prvega izbrala. Kombinacija bi bila zapisana na sledeč način: Tonik – Maskara. Confidence Item 1 nam bi v tem primeru povedal, da je stranka najprej izbrala tonik in za tem v košarico dodala še maskaro. Tovrstna kombinacija (v omenjenem zaporedju) se je zgodila v 0.02 %. Confidence Item 2 pa nam pove, da se je enaka kombinacija, vendar v obratnem zaporedju (Maskara – Tonik), zgodila v 0.15 %. Lift prikazuje verjetnost za pojav prvotne kombinacije, support pa kako pogosto se slednja pojavi. Nastavitve vizualizacije lahko vidimo na sliki 30.

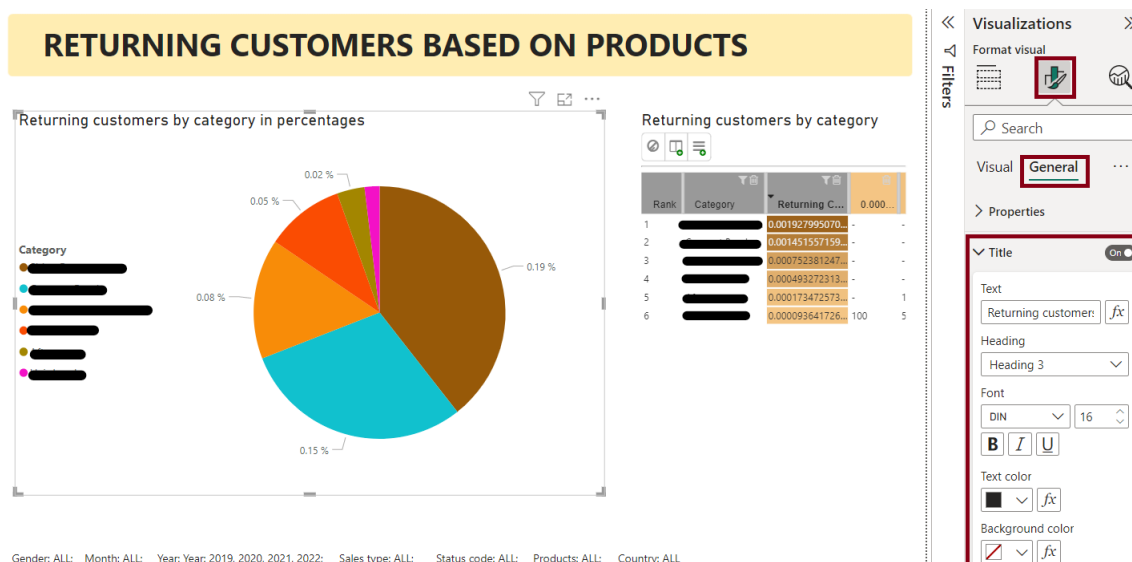
Slika 29: Produkti, ki jih kupci kupujejo skupaj



Vir: lastno delo.

Na spodnjem tortnem diagramu so prikazani produkti, po katere se stranke vračajo. Na sliki 31 lahko vidimo, da smo diagramu dodali tudi naslov. V nastavitvah na desni strani imamo pod osnovnimi nastavitvami tudi naslov. Power BI nam ga glede na kategorije, ki so v vizualizaciji, predlaga sam, vendar ga lahko kadarkoli popravimo. V našem primeru smo ga preimenovali in povečali velikost črk.

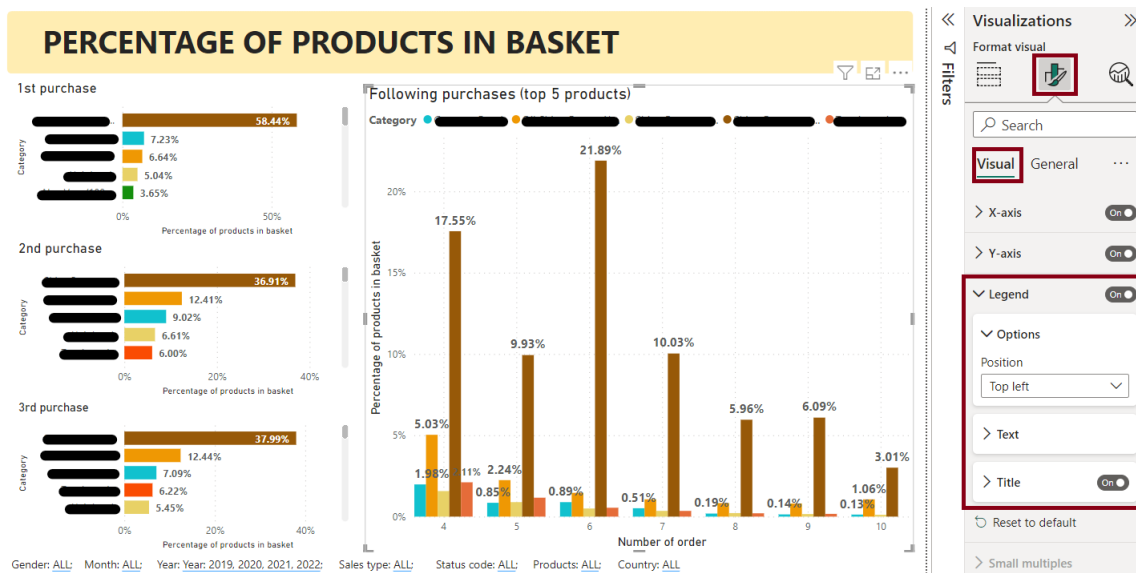
Slika 30: Urejanje naslova na vizualizaciji – Produkti, po katere se vračajo kupci



Vir: lastno delo.

Slika 32 predstavlja količino kupljenih produktov v določenem nakupu. Na desnem grafu smo poleg naslova in imena osi dodali še legendo. Slednjo lahko postavimo na različna mesta ob grafu, uredimo napis in njen naslov.

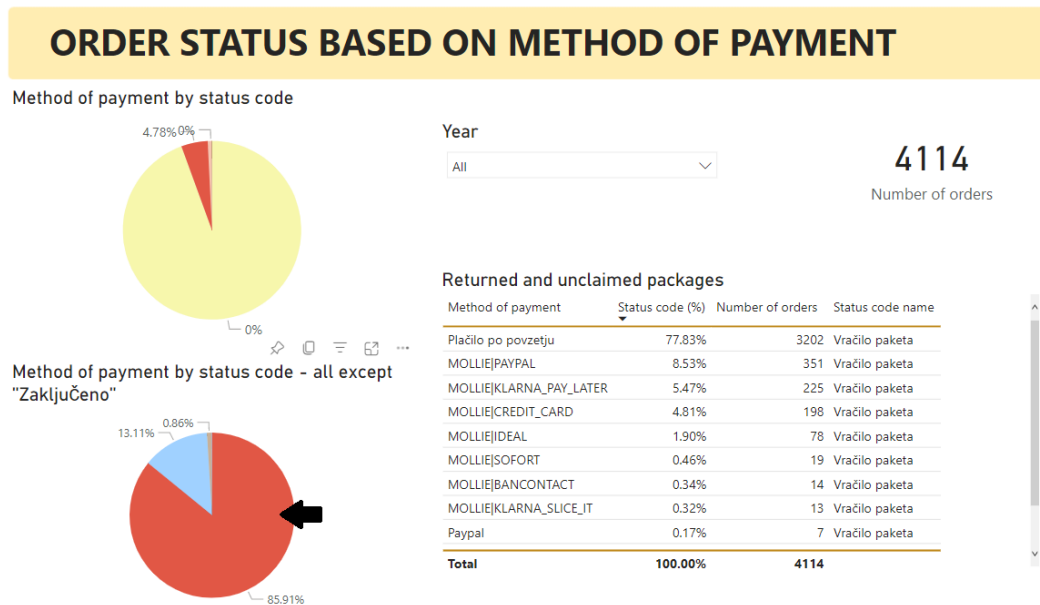
Slika 31: Dodajanje legende - Količina kupljenih produktov v določenem nakupu



Vir: lastno delo.

Na zadnji strani nadzorne plošče prikažemo še razmerje med statusom naročil ter vrsto plačila. Na sliki 33 lahko vidimo, da smo s klikom na levi spodnji diagram (označeno s črno puščico) posodobili tudi desno tabelo. Slednja je izpisala vse kombinacije, ki vključujejo vračila paketov.

Slika 32: Dinamične vizualizacije - Status naročila in vrsta plačila



Vir: lastno delo.

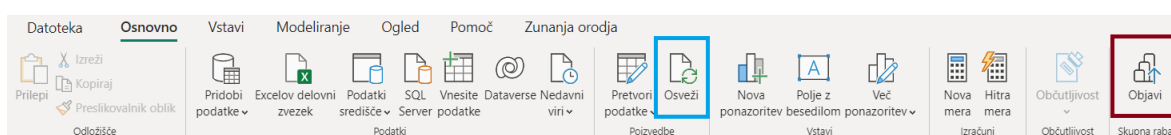
#### 4.5 Dodeljevanje pravic in avtomatizacija

V izbranem podjetju bodo napreden sistem poročanja, po besedah direktorja, uporabljali predvsem marketinški oddelek in lastnika. Ker je bila sama analitika zasnovana v ta namen, se podjetje zaenkrat še ni odločilo za omejevanje pravic uporabnikov. Do spletne verzije nadzorne plošče tako lahko dostopajo vsi zaposleni, ki poročilo potrebujejo za vsakdanje poslovanje in sprejemanje odločitev. Vsekakor pa se lahko zgodi, da bo v nekem trenutku potreba po dodeljevanju različnih ravni dostopov do poročila narasla. V primeru, ko podjetje zraste in zaposli strokovnjake znotraj različnih področij, je smiselno omejiti uporabo na tiste strani poročila, ki so namenjene točno določeni funkciji zaposlenega.

Eden izmed ciljev tovrstnih nadzornih plošč je zagotovo povečati avtomatizacijo. Trenutna nadzorna plošča je v večji meri avtomatizirana, manjše ročne posodobitve pa so potrebne pri naboru podatkov. Eden izmed korakov pri pripravi podatkov namreč vključuje tudi Python, ki ga je možno avtomatizirati samo preko personal gateway prehoda. Podatki se zbirajo v One Drive-u podjetja, iz tam pa jih Power BI avtomatsko prenese v nadzorno ploščo in izračunava kazalnike. Korak, ki ga je potrebno narediti ročno, pa je prenos podatkov iz strežnika, kjer so ti shranjeni. Za to uporabimo program, kjer izberemo datum, od katerega želimo podatke pridobiti in shraniti. Slednji se shranijo v temu namenjeno mapo, mi pa jih moramo samo še prenesti na One Drive. Da pa bodo podatki resnično v naši nadzorni plošči, moramo odpreti namizno verzijo poročila in podatke osvežiti. Na sliki 34 je pod zavihkom »Osnovno« (angl. Home) gumb »Osveži« (angl. Refresh) (označen z modro), ki osveži podatke. Zaradi velike količine lahko proces traja tudi do ene ure. Večina uporabnikov v podjetju uporablja nadzorno ploščo le za spremljanje analitike in ne tudi za urejanje, zato

uporabljajo spletno verzijo. Ko torej osvežimo podatke, moramo zadnjo verzijo še objaviti. To storimo s klikom na gumb »Objavi« (angl. Publish), ki je na sliki 34 označen z rdečo barvo.

Slika 33: Osveževanje in deljenje nadzorne plošče



Vir: Power BI desktop (2023).

#### 4.6 Ključne ugotovitve in priporočila

Ob vzpostavitvi novega, drugačnega načina poslovanja se podjetja navadno srečujejo z mnogimi izzivi. Že ob vzpostavljanju naprednega sistema poročanja so se zaposleni in lastniki naučili veliko novega o svojih podatkih. Kje in kako so shranjeni, v kakšnih oblikah in na kakšne načine jih lahko obdelujemo. Med pogovorom o uporabi naprednega sistema poročanja je direktor, ki je hkrati tudi lastnik podjetja, izpostavil tri ključne dodane vrednosti poročila. Sedaj se zaposleni zavedajo, katere podatke imajo in kje jih najdejo, nadzorna plošča pa jim pomaga, da o poslovanju ne govorijo več na pamet, ampak imajo na razpolago natančne podatke, predvsem podatke o posameznih kupcih oziroma segmentih kupcev. Druga velika dodana vrednost pa je uporaba tovrstnih podatkov v marketinške aktivnosti. Z njihovo pomočjo veliko lažje načrtujejo kampanje in optimizirajo proračun. K temu pripomore predvsem sekcija poročila, ki obravnava kupljene produkte in njihove povezave. Zadnja dodana vrednost pa je strateškega vidika. Z njim se ukvarjata predvsem lastnika, ki sedaj bolje vidita celotno sliko podjetja in hitreje sprejemata odločitve. Celoten proces izgradnje naprednega sistema poročanja je za podjetje pomenil korak naprej, tako v tehnološkem kot tudi kulturnem smislu. Direktor je med pogovorom izpostavil: »Projekt nam je dal dodatno samozavest, da se podjetje razvija v pravo smer, pokazal nam je novo dimenzijo, ki jo sedaj šele spoznavamo, v prihodnosti pa jo bomo zagotovo občutili v polnem pomenu.« Ob zaključku projekta smo skupaj ugotovili, da ob tovrstnih sistemih ne smemo zaspati. Podjetje mora za njihov razvoj skrbeti, se z njihovim vzdrževanjem stalno ukvarjati ter zaposlenim nuditi dovoljšno podporo ter znanje za njihovo učinkovito uporabo.

Po pogovoru z direktorjem podjetja smo ugotovili, da zaposleni potrebujejo podporo pri vzdrževanju nadzorne plošče, njeni uporabi in razumevanju nekaterih kompleksnejših kazalnikov. V ta namen smo pripravili pisna navodila z razlago posameznih kazalnikov, njihovih elementov in vizualizacij. Poleg tega smo posneli tudi video navodila za uporabo spletne verzije Power BI-ja in korake za osvežitev podatkov. Kljub vsem navodilom in predstavitvi celotnega sistema poročanja predlagamo njegovo redno vzdrževanje s strani usposobljenega strokovnjaka. Ker si podjetje želi še naprej razvijati napredne sisteme poročanja, predlagamo, da se izdelano poročilo vsakodnevno uporablja, saj bodo tako

zaposleni raziskali vse njegove možnosti, obenem pa dobili zaupanje v tovrstno analitiko. Podjetju predlagamo tudi mesečne sestanke z vzdrževalcem sistema. Na ta način lahko zagotovimo najvišjo kvaliteto poročila, hitro odkrivanje in odpravljanje napak ter hitrejši razvoj tega in vseh nadaljnjih sistemov poročanja. Kot smo ugotovili skupaj z direktorjem podjetja, morajo investirati tudi v znanje zaposlenih na področju analitike in naprednih sistemov poročanja.

Prej ali slej bo v vsakem podjetju analitika postala del ključnih funkcij poslovanja, napredni sistemi poročanja pa se bodo uporabljali pri vsakodnevnih odločitvah. Z direktorjem podjetja smo se pogovarjali tudi o tem, kako naj podjetja naredijo prve korake k podatkovno vodenemu poslovanju. Prva in verjetno najpomembnejša odločitev je miselni preskok. Zavedanje, da analitika postane enakovreden del podjetja, prav tako, kot so to na primer finance, marketing, prodaja in tako dalje. Pomembno se je zavedati, da je postavitve naprednih sistemov poročanja proces, ki se ga je potrebno lotiti, tudi če ga danes ne potrebujemo. Že jutri nam namreč lahko prinese veliko konkurenčno prednost. Drug, nič manj pomemben korak so zaposleni. Nov način dela jim je potrebno predstaviti na pravilen način, ki jih ne bo prestrašil, marveč spodbudil k raziskovanju in uporabi tovrstnih sistemov. Direktor je izbral princip tedenskih sestankov, na katerih se z zaposlenimi pogovarjajo o analitiki in uporabi poročila. S tem jih spodbuja k njihovi uporabi in razmišljanju, kot pravi sam, jim pomaga, da jim analitika pride v kri. Tako kot pri razvoju sistemov in njihovem posodabljanju, tudi pri zaposlenih ne smemo zaspati. Potrebno jim je dati orodja in znanje, s katerimi lahko sprejemajo odločitve, ki bodo najoptimalnejše za podjetje. S tem poenostavimo njihove delovne naloge in zagotovimo kakovostnejše poslovanje.

#### **4.7 Možnosti za nadaljnji razvoj naprednega sistema poročanja**

Carl Anderson je v svoji knjigi, *Creating a data-driven organization*, zapisal: »*Data doesn't have to be big to be valuable, and what is big today will be small tomorrow.*« S tem stavkom je želel povedati, da v podjetju nikoli ne smemo zaspati, pa čeprav govorimo o analitiki in podatkih, ki se samodejno zbirajo na temu namenjenem strežniku. Direktor izbranega podjetja je analitiko primerjal s podjetništvom, kjer je produkte potrebno venomer dopolnjevati in razvijati ter biti na preži za konkurenti. Zaradi podjetniške žilice se dobro zaveda pomembnosti razvoja naprednega sistema poročanja, ki smo ga skupaj razvili v preteklih mesecih. Ob pogovorih o prihodnosti smo se dotaknili kar nekaj različnih smeri, kamor bi sistem poročanja lahko razširili. Najenostavnejša je zagotovo dopolnitev obstoječega sistema s kazalniki, za katere bi se izkazalo, da so potrebni pri uporabi poročila v marketinške namene. V dosedanjem poročilu je zajeto omejeno število podatkov, torej le tisti, ki so bili potrebni za izračun točno določenih kazalnikov. V prihodnosti bi v to, ali ločeno poročilo, lahko dodali razširjene podatke, ki bi prinesli večjo natančnost in učinkovitost poročila. Slednji bi lahko samo dopolnjevali poročilo in omogočali, da so vsi podatki zbrani na enem mestu. V kolikor bi se poslužili tovrstnega načina, bi uredili tudi omejitve dostopa do naprednega sistema poročanja. Tako bi vsak zaposlen imel dostop le do

tistega dela poročila, ki je potreben za izpolnjevanje njegovih delovnih nalog. V naslednjem koraku razvoja še vedno ostajamo pri marketingu. Potencial, ki ga lastniki vidijo v tovrstni analitiki, se dotika oglaševalskih kanalov. Kljub temu, da imajo danes vse platforme svoje analitične plošče, bi si želeli njihove podatke pripeljati v skupno poročilo in jih tam primerjati. V ta namen bi kreirali novo nadzorno ploščo, ki bi bila povezana z vsemi oglaševalskimi kanali. Analizirala bi učinkovitost oglasov in optimizirala vse elemente, ki so prisotni v oglasu. V nadzorni plošči bi tako potekala analiza teksta, slik in gumbov. V nekem trenutku bi tovrstni algoritmi lahko vsaj delno zamenjali zaposlene, saj bi že sami dovolj dobro sestavili oglas. Takrat se bo, po besedah direktorja, struktura zaposlenih v marketingu bistveno spremenila. Namesto t. i. uploaderjev (zaposleni, ki pripravljajo oglase in jih razporejajo po družabnih omrežjih), bodo v oddelku prevladovali analitiki in strokovnjaki za poslovno inteligenco. Drugo področje, kjer bi bil podoben sistem uporaben, so finance. Na enak način bi kreirali napreden sistem poročanja, določili potrebne kazalnike in jih umestili v poročilo. Ker slednje vsebuje veliko bolj občutljive podatke kot sedanje poročilo, si želi podjetje pred njegovo izgradnjo pridobiti še nekoliko več izkušenj in znanja.

## **SKLEP**

Ne glede na to, ali govorimo o prvem delu magistrskega dela – pregledu literature, ali pa o primeru iz prakse, naletimo na zelo podobne ugotovitve. Podjetja potrebujejo zelo veliko časa, da se zavejo pomembnosti analitike in naprednih sistemov poročanja. Srečujejo se z mnogimi izzivi, kot so pomanjkanje časa za tovrstne projekte, pomanjkanje znanja s področja analitike in neodobravanje izgradnje naprednih sistemov poročanja s strani zaposlenih. V prvem delu magistrskega dela smo izpostavili kar nekaj raziskav, ki potrjujejo, da so podatkovno vodena podjetja na dolgi rok veliko bolj uspešna od tistih, ki to niso. Tako literatura kot tudi primer iz prakse pa pravita, da so začetki vedno težki. Večina podjetij pozna in uporablja le Excel, ki je za množične podatke neprimerno orodje. Torej se že v samem začetku soočajo z vprašanji, kje in kako začeti, kje pridobiti dovolj znanja in katera orodja so sploh primerna za izvajanje določenih analiz. Velikokrat je med zaposlenimi prisoten strah pred novimi orodji, saj so mnenja, da njihovo delo ne bo več potekalo nemoteno. Poleg vsakodnevnih nalog se bodo morali priučiti novih veščin, za katere menijo, da jim bodo vzele več časa, kot prinesle koristi. Vzpostavitev podatkovne kulture v podjetju je torej velik izziv, predvsem to velja za večja in starejša podjetja. V nekaterih primerih lahko sprejemanje podatkovne kulture med zaposlenimi traja tudi več let, zato je izredno pomembno, da se vodje tovrstnih projektov posvetijo tudi izobraževanju zaposlenih. Slednji morajo imeti dovolj znanja za razumevanje naprednih sistemov poročanja. Poleg tega morajo zgraditi zaupanje v podatke ter ponotranjiti, da so sistemi na svojem mestu zato, da jim pomagajo, in ne zato, da bi jim odvzeli delo. Podjetja, ki se ne bodo odločila za podatkovno vodenje, se bodo med drugim v prihodnosti soočala s slabšim razumevanjem svojega poslovanja, zamudila priložnosti za razvoj in napredek ter izgubila konkurenčno prednost.



Ob pričetku izgradnje naprednih sistemov poročanja se je potrebno dobro zavedati, kje imamo podatke, na kakšen način so shranjeni, in predvsem, kaj želimo doseči z analizami. Glede na postavljene cilje se odločimo za orodja, ki jih bomo uporabili, in način dela. Izbrano podjetje se je skozi proces postavitve naprednega sistema poročanja naučilo uporabe orodja Power BI in veliko bolje spoznalo svoje podatke. Zaradi vzpostavitve sistema so bolj samozavestni pri poslovanju, pogovarjajo se o točnih podatkih, medtem ko so prej govorili številke na pamet, veliko bolje načrtujejo marketinške aktivnosti in veliko bolje poznajo svoje kupce. Največja dodana vrednost je, po besedah direktorja, prav strateške narave. Z analizami so dosegli, da navade njihovih kupcev poznajo bolje, kot poznajo slednji sebe. Podatke lahko uporabijo za segmentacijo kupcev in targetiranje na različnih oglaševalskih kanalih. S projektom je podjetje naredilo miselni preskok v smer podatkovno vodenega podjetja in vstopilo na prvi nivo analitične zmogljivosti – postalo je ambiciozno podjetje. Z možnostmi nadaljnega razvoja sistemov tudi na druga področja poslovanja in razširitev obstoječih sistemov lahko v prihodnjih letih postane eno izmed vodilnih preoblikovanih podjetij v svoji panogi v Sloveniji.

Magistrsko delo predstavlja vse pomembne prve korake pri vzpostavitvi naprednih sistemov obveščanja. Slednje je lahko v pomoč vsem podjetjem, ki so še na začetku izgradnje tovrstne analitike oziroma o njej še razmišljajo. Ključni koraki so opisani na enostaven način, ki je primeren tudi za popolne začetnike, tako tiste, ki prihajajo iz prakse, kot tudi tiste na področju znanosti. V magistrskem delu smo predstavili le najbolj poznana in najenostavnejša orodja za izgradnjo naprednih sistemov poročanja. Zavedati pa se je potrebno, da obstaja še mnogo drugih, alternativnih orodij, ki so lahko primernejša za določena podjetja. Prav tako smo opisali le ključne naloge orodij, medtem ko so njihove zmožnosti veliko večje. V prihodnosti pričakujemo še hitrejši razvoj orodij in pojav novih poklicev, ki bodo s specifičnim znanjem nudili pomoč podjetjem na različnih nivojih podatkovne pismenosti. Nadaljnje raziskovanje lahko tako temelji na omenjenih dveh področjih in nadaljnjem razvoju naprednih sistemov poročanja.

## LITERATURA IN VIRI

1. Alsghaier, H., Akour, M., Shehabat, I. & Aldiabat, S. (2017). The Importance of Big Data Analytics in Business: A Case Study. *American Journal of Software Engineering and Applications*, 6(4).
2. Anderson, C. (2015). *Creating a Data-Driven Organization: Practical Advice from the Trenches*. O'Reilly Media, Inc.
3. Bean, R. & Davenport, T. (2019). Companies Are Failing in Their Efforts to Become Data-Driven. *Harvard Business Review*. 4. Pridobljeno 9. septembra 2022 iz <https://hbr.org/2019/02/companies-are-failing-in-their-efforts-to-become-data-driven>
4. Bernard, R., & Rao, A. (2021). PwC. *It's time to get excited*. Pridobljeno 26. januarja 2023 iz <https://www.pwc.com/gx/en/issues/data-and-analytics/artificial-intelligence/publications/ai-automation-data-extraction.html>

5. Biswal, A. (2023). Simplilearn. *Power BI Vs Tableau: Difference and Comparison*. Pridobljeno 25. januarja 2023 iz <https://www.simplilearn.com/tutorials/power-bi-tutorial/power-bi-vs-tableau>
6. Breuer, P., Elliott, B. & Rickert, S. (2013). McKinsey & Company. *The power of advanced analytics in revenue management*. Pridobljeno 7. septembra 2022 iz [https://www.mckinsey.com/~/media/mckinsey/dotcom/client\\_service/retail/articles/perspectives%20-%20winter%202013/power\\_of\\_advanced\\_analytics\\_in\\_revenue\\_management\\_periscope%20vf.pdf](https://www.mckinsey.com/~/media/mckinsey/dotcom/client_service/retail/articles/perspectives%20-%20winter%202013/power_of_advanced_analytics_in_revenue_management_periscope%20vf.pdf)
7. Brynjolfsson, E., Hitt, L. & Kim, H. H. (2011). *Strength in Numbers: How Does Data-Driven Decisionmaking Affect Firm Performance?* Pridobljeno 8. septembra 2022 iz [https://papers.ssrn.com/sol3/papers.cfm?abstract\\_id=1819486](https://papers.ssrn.com/sol3/papers.cfm?abstract_id=1819486)
8. Bughin, J. (2016). Big data: Getting a better read on performance. *McKinsey Quarterly*. Pridobljeno 7. septembra 2022 iz <https://www.mckinsey.com/industries/technology-media-and-telecommunications/our-insights/big-data-getting-a-better-read-on-performance>
9. Cai, L. & Zhu, Y. (2015). The Challenges of Data Quality and Data Quality Assessment in the Big Data Era. *Data Science Journal*, 14, 2.
10. Chui, M., Collins, M. & Patel, M. (2021). McKinsey & Company. *IoT value set to accelerate through 2030: Where and how to capture it*. Pridobljeno 11. januarja 2023 iz <https://www.mckinsey.com/capabilities/mckinsey-digital/our-insights/iot-value-set-to-accelerate-through-2030-where-and-how-to-capture-it>
11. Cote, C. (2021). *4 Types of data analytics to improve decision-making* [objava na blogu]. Pridobljeno 31. januarja 2023 iz <https://online.hbs.edu/blog/post/types-of-data-analysis>
12. Davenport, T. & Bean, R. (2018). Big Companies Are Embracing Analytics, But Most Still Don't Have a Data-Driven Culture. *Harvard business review*. Pridobljeno 9. septembra 2022 iz <https://hbr.org/2018/02/big-companies-are-embracing-analytics-but-most-still-dont-have-a-data-driven-culture>
13. Diaz, A., Kayvaun, R. & Seleh, T. (2018). Why data culture matters. *McKinsey Quarterly*. Pridobljeno 9. septembra 2022 iz <https://www.mckinsey.com/~/media/McKinsey/Business%20Functions/McKinsey%20Analytics/Our%20Insights/Why%20data%20culture%20matters/Why-data-culture-matters.ashx>
14. Economist Intelligence Unit. (2012). *The Deciding Factor: Big data & decision making*. Pridobljeno 6. septembra 2022 iz [https://www.capgemini.com/wp-content/uploads/2017/07/The\\_Deciding\\_Factor\\_Big\\_Data\\_\\_Decision\\_Making.pdf](https://www.capgemini.com/wp-content/uploads/2017/07/The_Deciding_Factor_Big_Data__Decision_Making.pdf)
15. Elgabry, O. (2019). The ultimate guide to data cleaning. *Towards Data Science*. Pridobljeno 19. januarja 2023 iz <https://towardsdatascience.com/the-ultimate-guide-to-data-cleaning-3969843991d4>
16. Ferrari, A. & Russo, M. (2016). Introducing Microsoft Power BI. *Microsoft Press*. Pridobljeno 7. septembra 2022 iz <https://www.microsoftpressstore.com/store/introducing-microsoft-power-bi-9781509302284>

17. Fountaine, T., McCarthy, B. & Saleh, T. (2019). Building the AI-Powered Organization. *Harvard Business Review*. Pridobljeno 9. septembra 2022 iz <https://hbr.org/2019/07/building-the-ai-powered-organization>
18. Geeks for geeks. (2022). *Introduction to NoSQL*. Pridobljeno 13. januarja 2023 iz <https://www.geeksforgeeks.org/introduction-to-nosql/>
19. Gitanjali, M. (2021). *Future of Business Analytics: 5 Predictions for 2025 and Beyond*. Pridobljeno 5. februarja 2023 iz <https://www.getapp.com/resources/predictions-for-the-future-of-business-analytics/>
20. Google. (2022). *Google Files you can store in Google Drive*. Pridobljeno 18. decembra 2022 iz Google Help: <https://support.google.com/drive/answer/37603>
21. Grossman, R. & Siegel, K. (2014). Organizational models for big data and analytics. *Journal of Organization Design*, 3(1), 20-25.
22. History.com Editors. (2018). *Oil industry*. Pridobljeno 7. septembra 2022 iz <https://www.history.com/topics/industrial-revolution/oil-industry>
23. Huang, F. (2019). *Data cleansing*. Pridobljeno 19. januarja 2023 iz [https://doi.org/10.1007/978-3-319-32001-4\\_300-1](https://doi.org/10.1007/978-3-319-32001-4_300-1)
24. International Business Machines Corporation. (2021). *Data preparation overview*. Pridobljeno 18. januarja 2023 iz <https://www.ibm.com/docs/en/spss-modeler/saas?topic=preparation-data-overview>
25. International Business Machines Corporation. (2023a). *Apache Hadoop*. Pridobljeno 24. januarja 2023 iz <https://www.ibm.com/analytics/hadoop>
26. International Business Machines Corporation. (2023b). *What is object storage*. Pridobljeno 13. januarja 2023 iz <https://www.ibm.com/topics/object-storage>
27. Kumar, P., Aruna, K. & Suchethana, H. (2022). Literature Survey on Big-Data Analytics and Tools. *JNNCE Journal of Engineering & Management*, 6.
28. Maffeo, L. (2020). Most businesses don't use data the right way [Research]. *Towards Data Science* Pridobljeno 9. septembra 2022 iz <https://towardsdatascience.com/most-businesses-dont-use-data-the-right-way-research-e1dcc81a5627>
29. Mannerings, A. (2021). *Four Options to Store Data Other Than Excel* [LinkedIn]. Pridobljeno 18. decembra 2022 iz <https://www.linkedin.com/pulse/four-options-store-data-other-than-excel-alexandra-mannerings>
30. McKay, S. (2018). *Discover Multiple Product Purchases Using DAX in Power BI* [objava na blogu]. Pridobljeno 19. oktobra 2022 iz <https://blog.enterprisedna.co/discover-the-amount-of-customers-who-purchase-multiple-products-w-power-bi/>
31. Microsoft Corporation. (2022a). *Advanced analytics with Power BI*. Pridobljeno 7. septembra 2022 iz <http://info.microsoft.com/rs/157-GQE-382/images/EN-CNTNT-PowerBI-Advanced-Analytics-with-PowerBI-white-paper.pdf>
32. Microsoft Corporation. (2022b). *Excel specifications and limits*. Pridobljeno 18. decembra 2022 iz Microsoft: <https://support.microsoft.com/en-us/office/excel-specifications-and-limits-1672b34d-7043-467e-8e27-269d656771c3>

33. Microsoft Corporation. (2022c). *Share Power BI reports and dashboards with coworkers and others*. Pridobljeno 27. decembra 2022 iz <https://learn.microsoft.com/en-us/power-bi/collaborate-share/service-share-dashboards>
34. Microsoft Corporation. (2022d). *Tips for designing a great Power BI dashboard*. Pridobljeno 25. januarja 2023 iz <https://learn.microsoft.com/en-us/power-bi/create-reports/service-dashboards-design-tips>
35. Microsoft Corporation. (2023). *QuickStart: Learn DAX Basics in 30 Minutes*. Pridobljeno 24. januarja 2023 iz <https://support.microsoft.com/en-us/office/quickstart-learn-dax-basics-in-30-minutes-51744643-c2a5-436a-bdf6-c895762bec1a>
36. Microstrategy. (2020). *2020 Global state of enterprise analytics*. Tysons. Pridobljeno 2. februarja 2023 iz <https://www3.microstrategy.com/getmedia/db67a6c7-0bc5-41fa-82a9-bb14ec6868d6/2020-Global-State-of-Enterprise-Analytics.pdf>
37. Mihai, B. & Borza, A. (2019). *Big Data Analytics and Organizational Performance*. Pridobljeno 7. septembra 2022 iz <https://doi.org/10.24818/mer/2019.12-06>
38. Mikuškova, E. (2017). Intuition in managers' decision-making: qualitative study. *Journal of East European Management Studies*, 22, 313-333.
39. Peck, J. (2020). *SQL vs R. Which to use for data analysis?* [objava na blogu] Pridobljeno 21. decembra 2022 iz <https://dataform.co/blog/sql-vs-r>
40. Python Software Foundation. (2016). *Gender-guesser 0.4.0*. Pridobljeno 13. februarja 2023 iz <https://pypi.org/project/gender-guesser/>
41. Python Software Foundation. (2023). *Core tools for working with streams*. Pridobljeno 10. februar 2023 iz <https://docs.python.org/3/library/io.html>
42. RSM US. (2020). The importance of the right reporting, analytics and information delivery. *Data & Digital transformation*, 5.
43. SQLBI Corporation. (2022). *The DAX language*. Pridobljeno 9. septembra 2022 iz <https://dax.guide/>
44. Stobierski, T. (2019). *The advantages of data-driven decision-making* [objava na blogu]. Pridobljeno 31. januarja 2023 iz <https://online.hbs.edu/blog/post/data-driven-decision-making>
45. Tableau Software. (2022). *What is Tableau?* Pridobljeno 23. decembra 2022 iz <https://www.tableau.com/why-tableau/what-is-tableau>
46. The Economist Newspaper Limited. (2017). *The world's most valuable resource is no longer oil, but data*. Pridobljeno 8. septembra 2022 iz <https://www.economist.com/leaders/2017/05/06/the-worlds-most-valuable-resource-is-no-longer-oil-but-data>
47. Yildirim, S. (2021). SQL as a data analysis tool. *Towards Data Science*. Pridobljeno 21. decembra 2022 iz <https://towardsdatascience.com/sql-as-a-data-analysis-tool-a18bf698a9cd>

## **PRILOGE**



## **Priloga 1: Vprašalnik za intervju z direktorjem izbranega podjetja**

1. Če začneva čisto na začetku, kako ste sploh v podjetju prišli do tega, da potrebujete analitiko?
2. Kako ste izbrali orodje, v katerem smo postavili analitiko, v našem primeru Power BI?
3. Na kakšen način ste kreirali mere, ki smo jih vključili v poročilo?
4. Katere zaposlene ste vključili v proces postavljanja naprednega sistema poročanja?
5. Kako so zaposleni sprejeli idejo in sodelovanje v tem projektu, ste se soočali s kakšnim odporom do uporabe analitike v vsakdanjem poslovanju?
6. Kateri oddelki poročilo največ uporabljajo?
7. Sedaj, ko ste poročilo začeli uporabljati, kje vidite njegovo največjo dodano vrednost?
8. Kako vam analitika pomaga pri procesu odločanja? Mi lahko zaupate kakšen konkreten primer?
9. Imate občutek, da ste zaradi naprednega sistema poročanja boljši od konkurence?
10. Kaj ste se tekom procesa postavitve analitike in njene uporabe naučili?
11. Na kakšen način menite, da bo tovrstna analitika spremenila vaše poslovanje?
12. Kje vidite nadaljnje možnosti razvoja tovrstnega sistema poročanja?
13. Imate kakšen nasvet za podjetja, ki še niso začela s postavitvijo tovrstnih analiz? Kje in kako začeti, na kaj se osredotočiti, kako pripraviti zaposlene na drugačen način dela in podobno?