Usporedba alata za vizualizaciju podataka

Bižaca, Roko

Master's thesis / Diplomski rad

2017

Degree Grantor / Ustanova koja je dodijelila akademski / stručni stupanj: **University of Split, Faculty of economics Split / Sveučilište u Splitu, Ekonomski fakultet**

Permanent link / Trajna poveznica: https://urn.nsk.hr/urn:nbn:hr:124:046958

Rights / Prava: In copyright / Zaštićeno autorskim pravom.

Download date / Datum preuzimanja: **2024-03-28**

Repository / Repozitorij:

REFST - Repository of Economics faculty in Split

zir.nsk.hr

UNIVERSITY OF SPLIT

dabar
DIGITALNI AKADEMSKI ARHIVI I REPOZITORIJI

SVEUČILIŠTE U SPLITU
EKONOMSKI FAKULTET

DIPLOMSKI RAD

USPOREDBA ALATA ZA VIZUALIZACIJU PODATAKA

Mentor:
Doc. dr. sc. Maja Ćukušić

Student:
Roko Bižaca

Split, rujan 2017.

SADRŽAJ:

1. UVOD 4
1.1 Problem istraživanja 4
1.2 Predmet istraživanja 5
1.3 Istraživačka pitanja 7
1.4 Ciljevi istraživanja 8
1.5 Metode istraživanja 8
1.6 Doprinos istraživanja 10
1.7 Struktura diplomskog rada 10
2. POSLOVNA INTELIGENCIJA I VIZUALIZACIJA PODATAKA 12
2.1 Poslovna inteligencija 12
2.1.1 Podatak, informacija i znanje 12
2.1.2 Poslovna inteligencija i pojam *intelligence* 13
2.1.3 Važnost poslovne inteligencije u suvremenom poslovanju 14
2.1.4 Izvori i organizacija internih podataka 16
2.1.4.1 Relacijske baze podataka 17
2.1.4.2 Pretraživanje baze podataka 22
2.1.4.3 Skladišta podataka 25
2.1.5 Višedimenzijske strukture podataka 29
2.1.5.1 Višedimenzijski model podataka 29
2.1.5.2 OLAP 32
2.1.5.3 Rudarenje podataka 34
2.1.5.4 OLAM 36
2.2 Vizualizacija podataka 37
2.2.1 Pojam vizualizacije podataka 37
2.2.2 Važnost u suvremenom poslovanju 39
2.2.3 Najčešće vrste grafikona i njihovo korištenje u praksi 41
2.2.3.1 Točkasti grafikon 41
2.2.3.2 Stupčasti grafikon 44
2.2.3.3 Linijski grafikon 47
2.2.3.4 Ostali grafikoni 50
2.2.4 Nadzorne ploče 53
2.2.4.1 Uvod u nadzorne ploče 53
2.2.4.2 Prakse u izradi nadzornih ploča 54
3. ALATI ZA VIZUALIZACIJU PODATAKA 57
3.1 Uvod u tematiku alata za vizualizaciju podataka 57
3.2 Funkcionalnosti alata za vizualizaciju podataka 58
3.3 Tržište alata za vizualizaciju podataka 60
3.4 Najpopularnija rješenja na tržištu 62
3.4.1 Tableau 62
3.4.2 QlikView 62
3.4.3 Microsoft Power BI 64
3.4.4 Looker 64

3.4.5 Domo ... 65
3.4.6 Dodaci za programske jezike ... 66
3.4.6.1 D3.js ... 66
3.4.6.2 Chart.js ... 67
3.4.6.3 Fusion Charts ... 67
3.4.6.4 R ... 67
3.5 Usporedba alata za vizualizaciju ... 68
4. KOMPARACIJA ALATA TABLEAU I ALATA POWER BI ... 69
4.1 Pregled značajki alata Tableau ... 69
4.2 Pregled značajki alata Power BI ... 71
4.3 Komparacija funkcionalnosti alata ... 72
4.4 Testiranje i komparacija korisničkog iskustva ... 74
4.4.1 Postavljanje zadataka ... 74
4.4.2 Određivanje sudionika u istraživanju ... 76
4.4.3 Provođenje eksperimenta ... 76
4.4.3.1 Praćenje računalnog miša, tipkovnice i vremena ... 77
4.4.4 Postavljanje i testiranje hipoteza ... 79
4.4.5 Interpretacija empirijskih rezultata ... 82
4.4.6 Intervjuiranje sudionika ... 83
5. ZAKLJUČAK ... 84
SAŽETAK ... 86
SUMMARY ... 86
LITERATURA ... 87
POPIS SLIKA I TABLICA ... 91

1. UVOD

1.1 Problem istraživanja

U modernom poslovanju informacije predstavljaju ključni resurs kojim se generira konkurentska prednost. Međutim, definirati informaciju kao resurs neke organizacije, ili kao robu na tržištu, je vrlo teško. Resursi se obično percipiraju kao dobra koja nosa mjerljivu ili potencijalnu korist korisniku. Informacija se može smatrati vrstom nematerijalnog dobra, iako i ona sama ima materijalne karakteristike (oblik u kojemu je pohranjena može biti materijalan: papir, serverski prostor itd.). Informacija ima karakteristike i obnovljivog resursa: „Podaci i informacije kao gradivni element znanja mogu se u nedogled kombinirati i obnavljati. Jednom iskorištena informacija može biti vrijedan čimbenik kolektivne ili individualne memorije, a osim toga definitivno proširuje repertoar za kreiranje novih informacija, te je samim tim višekratno iskoristiva.“ (Pfeiffer, 2013.) Jednom iskorištena informacija drastično gubi na svojoj vrijednosti u svom osnovnom obliku, međutim ukoliko se ista informacija kombinira s ostalim informacijama može imati daleko veću vrijednost od inicijalne.

U suvremenom informacijskom svijetu nerijetko dolazi do informacijske krize - nemogućnost pojedinaca i poslovnih sustava da u svakom trenutku mogu pribaviti i koristiti potrebne informacije, te problemi uvođenja računala za podršku poslovnim i ostalim aktivnostima odraz su informacijske krize. (Strahonja, 1992.). Informacijska kriza javlja se iz nemogućnosti procjene relevantnih informacija u mnoštvu podataka. Prema tome, informacija (podatak) predstavlja anomaliju kroz perspektivu resursa jer kriza ne označava manjak, već višak. Poslovni subjekti ulažu napore u svrhu pronalaska načina organizacije i razumijevanja mnoštva podataka koji su generirani kroz njihovo poslovanje. Mogućnost i brzina saznavanja informacije utječe direktno na poslovne odluke, a samim time i na boljitak organizacije. Potreba za brzinom saznavanja informacije ali i njenom razumijevanju iznimno je poželjna mogućnost. Informacije pohranjene u bazama podataka i skladištima podataka tako se prikazuju u nekom jednostavnom vizualnom obliku jer on pruža daleko jasniji pogled od brojčanog tabličnog prikaza. (Bižaca, Dobrović, 2016.)

Brojni proizvođači softverskih poslovnih rješenja prepoznali su nužnost za vizualizacijom podataka pa su u svoje poslovne sustave uvrstili funkcionalnosti vizualizacije kroz nadzorne

ploče (*dashboard*-ove). Daljnjim razvojem poslovnih sustava vizualizacija je postala jedna od važnijih funkcionalnosti da bi se na kraju razvili posebni alati koji služe isključivo za vizualizaciju podataka. Upravo ti alati područje su ovog rada. Analitika podataka prebačena je sa viših razina menadžmenta na operativnu razinu što stavlja naglasak na jednostavnost i intuitivnost korištenja navedenih alata (kako bi se alat uspješno mogao koristiti neovisno o iskustvu korisnika). Ovaj rad uspoređuje dva popularna alata iz ove domene kroz prizmu:

- Funkcionalnosti
- Dostupnosti
- Subjektivne metrike korisnosti
- Objektivne metrike korisnosti.

Alati za vizualizaciju podataka moraju zadovoljiti eksperte u analizi podataka, ali i krajnje korisnike iz ostalih domena, stoga raspolažu visokim brojem funkcionalnosti. Aplikacije s mnoštvom funkcionalnosti obično budu komplicirane za korištenje. Krivulja učenja kod takvih aplikacija ima eksponencijalni oblik što upućuje da je u početnim vremenskim razdobljima (početak korištenja softvera) napredak poprilično spor, dok stjecanjem iskustva učenje postaje lakše. (Pindyck et.al., 2005.)

Ispitivanje korisničkog iskustva korištenjem alata za vizualizaciju podataka predstavlja osnovnu problematiku ovog diplomskog rada. „*User experience* (UX, ili korisničko iskustvo) odnosi se na emocije i sveobuhvatno iskustvo koje kod korisnika izaziva interakcija s određenim sustavom, uslugom ili proizvodom. Ono uključuje i proučavanje praktičnih, iskustvenih, afektivnih, smislenih i vrijednih aspekata u odnosu između čovjeka i računala. Na njega utječe i percepcija aspekata sustava poput korisnosti, učinkovitosti i jednostavnosti uporabe.“ (Sušac, 2015.) Načini mjerenja korisničkog iskustva biti će objašnjeni u nastavku.

1.2 Predmet istraživanja

Alati za vizualizaciju podataka postali su neizostavan dio velikih poslovnih sustava. Napredni alati za vizualizaciju podataka bilježe porast korisnika u razvijenim ekonomijama upravo zbog spoznaje vrijednosti podataka koji se na dnevnoj bazi generiraju unutar neke organizacije. Počevši od Microsoft Excela, odnosno njegove Pivot Tables opcije, do naprednijih specijaliziranih alata kao što su QlikView ili Tableau, softverski alati neizbježni su kod stvaranja poslovnih izvještaja. Jednostavni alati koriste se za izgradnju jednostavnih grafičkih

prikaza i info grafika uglavnom iz već očišćenih izvora podataka. S druge strane, složenija rješenja za vizualizaciju podataka omogućavaju direktne upite na jedan ili više složenih izvora podataka (baze podataka, skladišta podataka…) u realnom vremenu. (Rist, 2016.)

Alati koji će biti uspoređeni u ovom diplomskom radu spadaju u napredna rješenja za vizualizaciju. Oni moraju zadovoljavati opcije *on-demand* upita direktno u bazu (skladište) podataka bez prijašnjeg „čišćenja" *dataseta*. Složena rješenja, prema tome, moraju raditi na principu nekog upitnog jezika (pr. SQL-a) kako bi navedeni *on-demand* upiti bili mogući. Korisnik mora imati opciju odabira između mnoštva različitih grafičkih prikaza (za razliku od jednostavnih rješenja koja se uglavnom temelje na bar i linijskim grafikonima). Kontrolne ploče još jedan su uvjet koji softver mora posjedovati kako bi se smatrao ozbiljnim, složenim rješenjem za vizualizaciju. Njihov izgled inspiriran je takvim pločama koje se nalaze primjerice u automobilima ili avionima. Kontrolna ploča ili *Dashboard* služi kao vizualni prikaz najvažnijih informacija potrebnih za postizanje jednog ili više ciljeva, konsolidiranih i uređenih na jednom zaslonu tako da se sve informacije mogu pratiti na prvi pogled (Few, 2004).
Međutim, osnovna funkcija složenih alata za vizualizaciju podataka odnosi se na jednostavnost korištenja. Korisnik mora biti u stanju uspješno postavljati složene upite na bazu podataka bez prijašnjeg iskustva s bazama podataka. Naglasak **na korisničko iskustvo** i **na krivulju učenja** kojom korisnik relativno brzo savladava osnovne funkcionalnosti ključno je za softver koji pokriva jednu relativno mladu granu. Pretpostavka ovih alat je da ih neće koristiti samo eksperti iz domene analize podataka, već menadžeri na svim razinama kojima je potrebna pravovremena informacija. Upravo zbog toga predmet istraživanja ovog rada je korisničko iskustvo u korištenju alata.
Analizom tržišta naprednih salata za vizualizaciju podataka pronađena su najpopularnija rješenja: Tableau, SAP Lumira, QlikView, Power BI, Looker itd. Neće svi alati biti detaljno razmotreni u ovom radu pa su na posljetku izabrana dva: **Tableau**, kao definitivni tržišni predvodnik te **Power BI** kao moderno rješenje za poslovnu inteligenciju koje trenutno predstavlja najveću prijetnju zavidnom položaju Tableau-a, a dolazi iz redova Microsoft-a. Ova dva rješenja izabrana su prema iskustvima i preporukama online *community*-ja korisnika alata za vizualizaciju podataka.

1.2 Istraživačka pitanja

S obzirom na prethodno postavljen problem i predmet istraživanja, u ovom radu postavljeno je nekoliko istraživačkih pitanja, na koja će se odgovoriti po završenom istraživanju:

1. Postoji li razlika u funkcionalnostima odabranih rješenja za vizualizaciju podataka? (osnovne funkcije, izgled sučelja, dizajn grafikona…)
2. Postoji li razlika s obzirom na konektivnost rješenja sa ostalim *3rd party* rješenjima za vizualizaciju podataka
3. Postoji li razlika s obzirom na konektivnost rješenja sa različitim izvorima podataka i različitima oblicima datoteka u kojima su podaci spremljeni (SQL baze, Access baze, Google Analytics baze, .csv format, .xml format…)
4. Postoji li razlika u *cross-platform* dostupnost pojedinog rješenja (Desktop verzija, Cloud verzija, mobilna verzija, Windows verzija, Mac verzija…)
5. Postoji li razlika u cijeni pojedinog rješenja?
6. Postoji li razlika u Korisničkom iskustvu?
   - Za zadnje pitanje postavljena je i hipoteza koja glasi:

     H_0 .....Ne postoji statistički značajna razlika u korisničkom iskustvu korištenja alata

     H_1.....Postoji statistički značajna razlika u korisničkom iskustvu korištenja alata

   Hipoteze će biti ispitane temeljem brojčanih podataka prikupljenih praćenjem miša korisnika prilikom rada u jednom, odnosno drugom alatu.

1.3 Ciljevi istraživanja

S obzirom na postavljen problem, predmet te pitanja ovog istraživanja postavljaju se i njegovi ciljevi. Osnovni cilj ovog istraživanja dijeli se na dva djela: teorijski i empirijski.

Cilj teorijskog dijela je postavljanje izrada teorijske podloge koja će upotpuniti empirijski dio istraživanja. Ta teorijska podloga formira se prikupljanjem sekundarnih podataka iz raznih znanstvenih izvora kao što su knjige, članci, časopisi itd. Osim znanstvenih izvora, veliku ulogu u stvaranju teorijske slike imaju i tzv. ekspertni izvori podataka koji se prikupljaju uglavnom on-line čitanjem blogova, preporuka i diskusija. Teorijski dio služi kao podloga koja objašnjava koncept poslovne inteligencije, koncept vizualizacije podataka te na kraju koncept alata za vizualizaciju podataka.

Cilj empirijskog dijela istraživanja predstavlja dubinsku analizu oba softvera za vizualizaciju podataka. Analiza će biti odrađena prvenstveno eksperimentom u kojemu će sudjelovati autor, ali i eksperti iz raznih poslovnih grana. Korištenjem softvera na objektivan i na subjektivan način analizirat će se pojedine funkcionalnosti. Nakon detaljne analize slijedi komparacija oba rješenja koja je na posljetku osnovni cilj empirijskog dijela ovog rada.

1.4 Metode istraživanja

Metode istraživanja odnose se na sve postupke i procese koji se koriste u znanstvenom istraživačkom radu u svrhu pronalaska rezultata istraživanja unutar određenog znanstvenog područja. (Zelenika, 2000). Istraživanje unutar ovog rada temelji se na dvije vrste podataka: sekundarnim i primarnim podacima. Sekundarni su podaci oni koji su prikupljeni iz nekih vanjskih izvora (unutar knjiga, znanstvenih članaka, online izvora itd.) a koji se koriste u rješavanju istraživačkog problema. Prednost sekundarnih podataka je njihova dostupnost, dok je potencijalni problem njihova zastarjelost. Primarni podaci, s druge strane, predstavljaju podatke koje autor sam prikuplja tijekom provođenja istraživanja. Ovaj način prikupljana zahtijeva daleko veće napore istraživača, ali zato pruža ažurne i točne podatke.

Metode istraživanja koje se koriste u ovom radu, bez obzira radi li se o primarnim ili sekundarnim podacima, su sljedeće:

Induktivna metoda Induktivna metoda je sustavna primjena induktivnog načina zaključivanja kojim se na temelju analize činjenica dolazi do zaključka. Induktivnom metodom istraživanja zaključuje se općenita korisnost pojedinog softverskog alata njegovim korištenjem u kontekstu poslovne problematike analize podataka.

Deduktivna metoda temelji se na analizi općih činjenica, a bazira se uglavnom na sekundarnim podacima. Riječ je dakle o analizi gotovih teorijskih činjenica i zaključaka drugih autora iz istraživačkih područja sličnim istraživačkom području ovog rada. U ovom radu deduktivna metoda odnosi se na uvažavanja podataka koji su dobiveni studijama slučaja i komparacijom alata, a koji su dostupni u znanstvenim člancima i on-line izvorima.

Eksperimentalna metoda je „postupak promatranja pojave koja se ispituje pod točno određenim uvjetima koji dopuštaju praćenje tijeka pojave i njezino ponavljanje uz iste uvjete“ (Kulenović et.al., 2010). Ovom metodom izvode se i mjere realni pojmovi što prikupljenim primarnim podacima daje iznimnu vrijednost. U kontekstu ovog rada, eksperimentalna metoda koristi se praćenjem korisnika eksperata iz različitih domena u korištenju softverskih alata.

Komparativna metoda je postupak uspoređivanja istih i srodnih funkcionalnosti pojedinog alata, odnosno utvrđivanje njihove sličnosti i razlika u njihovom obliku i intenzitetu. Komparativna metoda temelji se zajedno i na primarnim i sekundarnim podacima.

Metoda anketiranja i intervjuiranja pristavljaju dvije različite metode, ali se u ovom radu provode zajedno. Metoda anketiranja odnosi se na prikupljanje podataka, informacija i stavova o predmetu (korisničkom iskustvu) korištenjem anketnog upitnika. Anketa se zatim upotpunjuje i provođenjem intervjua u kojemu ispitanik dodatno izjašnjava svoje stavove.

1.5 Doprinos istraživanja

Hrvatsko poslovno tržište još nije u cijelosti prepoznalo prednost koju pruža adekvatna obrada podataka generiranih svakodnevnim poslovanjem. S obzirom da cijela grana nije u potpunosti prepoznata, logično je zaključiti kako ni vizualizacija podataka kao jedan od oblika poslovnog izvještavanja nije shvaćen kao sustav koji stvara dodatnu vrijednost. Ovaj rad detaljno će istražiti tržište alata za vizualizaciju podataka, ocijeniti ih prema njihovim funkcionalnostima te doprinijeti jedinstvenom analizom dva najpopularnija rješenja. Objektivnim i subjektivnim rezultatima eksperimentnog istraživanja biti će moguće zaključiti koji alat pruža bolje korisničko iskustvo, bilo da se radi o iskusnim, odnosno novim korisnicima. Ovakvi rezultati, uz kombinaciju detaljne komparacije alata prema njihovim ostalim elementima doprinijet će stvaranju slike o pojedinom alatu. Također će isticanjem pozitivnih i negativnih strana pojedinog alata čitatelju dati da sam procijeni koji alat bi mu bolje odgovarao s obzirom na funkcionalnosti koje su mu od velike važnosti do onih koje su mu u potpunosti nebitne.

1.6 Struktura diplomskog rada

Sadržaj diplomskog rada na temu **Usporedba alata za vizualizaciju podataka** koncipiran je u 5 poglavlja: Uvod, Poslovna inteligencija i vizualizacija podataka, Alati za vizualizaciju podatka, Komparacija alata Tableau i alata Power BI te Zaključak. Na kraju rada nalazi se i sažetak na hrvatskom i engleskom jeziku, popis literature te popis slika i tablica.

Prvo pogljavlje nosi naziv Uvod, a predstavlja uvodni dio rada u kojemu je detaljno opisan problem istraživanja, definiran je predmet istraživanja, iznijeta su istraživačka pitanja je definiran krajnji cilj i doprinos istraživanja. Prvo poglavlje također sadrži i detaljnu strukturu diplomskog rada.

Drugo poglavlje odnosi se na teorijsku analizu poslovne inteligencije kroz kontekste modernog poslovanja. Naziv poglavlja Poslovna inteligencija i vizualizacija podatka upućuje i na detaljnu obradu tematike vizualizacije podataka koja je usko vezana s predmetom istraživanja. Dakle, postavlja se teorijska podloga poslovne inteligencije kao domene i

vizualizacije podataka kao njezine pod domene, te se naglašava važnost korištenja oba sustava. Ovo poglavlje ujedno služi i kao uvod u treće poglavlje

Treće poglavlje obrađuje konkretniju problematiku: Alati za vizualizaciju podataka. U ovom poglavlju navode se najpopularnija rješenja dostupna na tržištu, te je napravljena detaljna analiza pojedinog alata kroz prizmu funkcionalnosti, dostupnosti, cijene te njihovih pozitivnih i negativnih strana. U trećem poglavlju objašnjeno je i zašto su za eksperimentalnu metodu istraživanja izabrani Tableau i Power BI.

Četvrto poglavlje empirijsko je poglavlje diplomskog rada. U ovom poglavlju izneseni su detalji eksperimenta koji se koristi kao empirijska metoda. Prikazani su rezultati istraživanja te se prihvaća jedna od dvije zadanih hipoteza, a oba alata kompariraju se u detalje.

Peto poglavlje diplomskog rada odnosi se na dio u kojem su izneseni zaključci autora. U ovom poglavlju autor se subjektivno osvrće na čitav rad, njegovo istraživanje te komentira dobivene rezultate. Peto poglavlje potpuno je subjektivno napisano s obzirom na problem, predmet i cilj istraživanja.

2. POSLOVNA INTELIGENCIJA I VIZUALIZACIJA PODATAKA

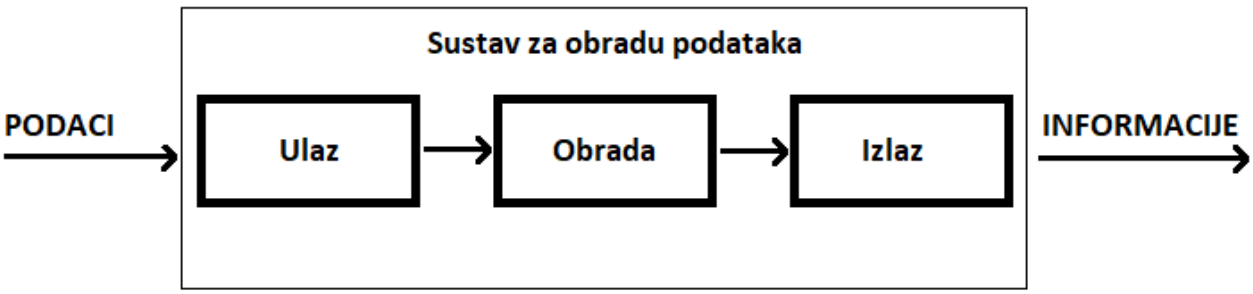
2.1 Poslovna inteligencija

2.1.1 Podatak, informacija i znanje

Informacija se u modernom poslovanju smatra resursom koji je podjednako važan konvencionalnim resursima. Informacija kao resurs osnovni je input u procesu odlučivanja. Kvaliteta donesene odluke u pozitivnom je korelacijskom odnosu sa kvalitetom dostupnih informacija, pod pretpostavkom da donositelj odluke postupa racionalno.

"*Informacija je novost ili vijest. U ovom kontekstu treba je tumačiti kao čovjekov spoznajni doživljaj poruke. Ako poruka za primatelja ima spoznajnu vrijednost ona je informacija*." – Garača, 2007.

Podatak se često krivo navodi kao sinonim informaciji. Podatak označava tek input sustavu njegove obrade koji stvara informaciju. Slikovito prikazano, podatak je moguće objasniti kao sirovinu čijom se obradom dobije gotovi proizvod – informacija.



Slika 1. Transformacija podataka u informacije; Izvor: Garača, 2007.

Informacija kao resurs ima jednu značajku koja je razlikuje od konvencionalnih resursa. Kriza nekog resursa smatra se nedostatkom tog resursa. Tako na primjer naftna kriza označava nedostatak tog resursa iz bilo kojeg razloga. Pod pojmom krize informacija, prije tehnološkog buma sredinom 20. stoljeća vrijedilo je isto pravilo – nedostatak informacija. Ulaskom u digitalno doba kriza informacija poprima novu definiciju – višak dostupnih informacija. Postalo je teško raspoznati kvalitetne informacije od nekvalitetnih, što je otežalo njihovu daljnju obradu. Veliki poslovni sustavi modernih organizacija generiraju ogromne količine informacija što uzrokuje krizu tog resursa i tako otežava odlučivanje.

Osnovne osobine kvalitetne informacije su:

- **Pravovremenost i aktualnost** (vremenska dimenzija informacije – informacija je dostupna u pravom trenutku)
- **Relevantnost** (korisna je korisniku koji je posjeduje)
- **Točnost i pouzdanost** (da su pogreške informacije unutar prihvatljivih okvira.)

Unutar velikih poslovnih organizacija nastala je potreba za filtriranje i analizu sustavom generiranih informacija (i podataka). Organizirani su posebni timovi za analitiku i kreiranje izvještaja što danas nazivamo poslovnom inteligencijom.

2.1.2 Poslovna inteligencija i pojam *intelligence*

Poslovna inteligencija proces je prikupljanja relevantnih internih i dostupnih eksternih podataka, te njihove konverzije u korisne informacije koje mogu pomoći poslovnim korisnicima u odlučivanju. (Panian, 2003.). Poslovna inteligencija sve važniji je segment modernog poslovanja s obzirom na tržišne okolnosti: žestoka konkurencija; zahtjevni kupci; ubrzan tempo života i tehnološkog napretka. Sve navedeno otežava poslovno odlučivanje što poslovnu inteligenciju gura na poziciju osnovnih konkurentskih prednosti. Osnovni cilj poslovne inteligencije je potpora odlučivanju.

Prescott (1999.) nastanak poslovne inteligencije razmatra kroz tri pravca. Sam početak poslovne inteligencije proizlazi od često citiranog stratega Sun Tzua koji smatra da je razumijevanje okoline osnova odlučivanja. Drugi pravac proizlazi iz Drugog svjetskog rata nastankom modernih obavještajnih službi. Treći, najnoviji pravac, koji je bio predmet njegovog istraživanja odnosi se na gospodarske subjekte. Unutar ovog pravca Prescott je analizirao tri razvojne faze poslovne inteligencije. Prvu fazu smješta u 60-e i 70-e godine 20. stoljeća te je naziva „Prikupljanje poslovne inteligencije. 80-ih godina nastupa druga razvojna faza koju označava kao „Industrijsku i konkurentsku analizu“ koja traje do razdoblja pisanja njegovog rada te je zamjenjuje faza pod nazivom „Poslovna inteligencija za strateško odlučivanje“. Prescott u svom radu predviđa i 4. fazu pod nazivom „Poslovna inteligencija kao konkurentska sposobnost“. (Luetić, 2013).

Riječ inteligencija u hrvatskoj verziji pojma Poslovna inteligencija nerijetko se krivo interpretira. Budući da je originalan naziv pojma na engleskom jeziku *business intelligence,* riječ inteligencija je sinonim za obavještavanje ili izvještavanje, a ne za pamet, razum i

snalažljivost. Prema engleskom rječniku hrvatskog izdanja (Filipović, 1991.), pojam *intelligence* prevodi se kao:

1. pamet, razum, inteligencija, razumijevanje, brzo shvaćanje, snalažljivost, prisebnost;
2. obavijest, obavještajna služba.

Pojam *intelligence* uglavnom se koristi u vojnom kontekstu, pa ga Američka vojna enciklopedija tumači kao rezultat prikupljanja, obrade, integracije, analize, procjene i interpretacije dostupnih informacija koje se odnose na druge države ili područja. (Javorović, Bilandžić, 2007.). U poslovnom kontekstu pojma inteligencije, Mohanty (2008.) ga tumači kao sposobnost realizacije poslovnog uspjeha kroz pravovremenu i točnu spoznaju poslovnih okolnosti. Fuld (2010.) objašnjava poslovnu inteligenciju kroz proces odlučivanja u kojemu donositelj odluke djeluje prema prikupljenim informacijama stvaranjem slike koja ne treba biti perfektna. Dakle, Fuld razumije da donositelj odluke ne može imati sve informacije, ali njihovom kvalitetom i kvantitetom stvara sliku približnu realnom sustavu i tako minimizira rizik loše odluke.

2.1.3 Važnost poslovne inteligencije u suvremenom poslovanju

Poslovna inteligencija danas predstavlja jednu od konkurentskih i tržišnih prednosti poslovanja. Pojam komparativna inteligencija nastao je prije pojma poslovna inteligencija u SAD-u, gdje se i danas koristi kao sinonim za poslovnu inteligenciju. Fuld (2010.) objašnjava pojam komparativne inteligencije koristeći dva osnovna kriterija:

- Sposobnost pronalaska prave informacije u pravo vrijeme
- Sposobnost racionalne i nepristrane interpretacije te informacije i povijesnih događaja na tržištu.

Drugi kriterij Fuld smatra kritičnim za komparativnu inteligenciju. On predstavlja svojevrsni *Historia est magistra vitae* (lat. Povijest je učiteljica života) poslovnog svijeta u kontekstu poslovne inteligencije.

Komparativna inteligencija stariji je koncept od poslovne inteligencije jer se odnosio na informiranje o konkurentima na tržištu. Osnova ovog koncepta je prikupljanje znanja, informacija i podataka o konkurenciji poduzeća počevši od korporativnih ciljeva, njihovih snaga i slabosti, pa sve do korporativne kulture te profila menadžera i zaposlenika (Panian et al.., 2007.). Metode prikupljanja ovih podataka mogu prekršiti etičke, a ponekad i zakonske norme poslovanja. Brojni slučajevi korporativne špijunaže primjeri su takvog protuzakonitog i

neetičkog postupanja. Komparativna inteligencija nije sinonim za poslovnu inteligenciju, već je samo jedan od pod skupova tog pojma.
Breeding (2010.) dijeli BI (skraćeno za *business intelligence*/ eng. poslovna inteligencija) na pet podskupina:

- *Competitive intelligence* (Kompratаivna inteligencija)
- *Market intelligence* (Tržišna inteligencija)
- *Customer intelligence* (Korisnička inteligencija)
- *Techincal intelligence* (Tehnička inteligencija)
- *Partner intelligence* (Partnerska inteligencija)

Komparativna inteligencija objašnjena je u prethodnom paragrafu. Tržišna inteligencija je proces prikupljanja podataka, informacija i znanja o pojedinom tržištu. Ovaj podskup BI-a često se naziva i istraživanje tržišta. Korisnička inteligencija (CI) povezana je sa tržišnom inteligencijom, a odnosi se na proces sakupljanja i analiziranja informacija o korisniku. Pojavom društvenih medija, informacije o korisnicima postale su dostupnije nego ikada prije. Korisnikova demografija, njegovi interesi, aktivnosti i ponašanja postala su jeftina za prikupljanje. Korisnička inteligencija postala je česta tema u etičkim raspravama vezanim za korištenje korisnikovih osobnih informacija. Zakonski okviri nisu u stanju pratiti tempo tehnološkog napretka pa je ova sfera u svojevrsnoj „sivoj zoni“. Tehnička inteligencija odnosi se na saznanja o tehnološkim trendovima okruženja, bilo da se radi o industriji, konkurentima, tržištu ili partnerima. Kod tehnologije važno je biti u korak s vremenom, ali je podjednako važno koristiti tehnologiju koja odgovara ustroju i procesima poslovnog subjekta. Najnovija tehnologija u kratkom roku ne znači nužno i napredak. Partnerska inteligencija prikupljanje je podataka informacija i znanja o poslovnim partnerima. Važno je biti upoznam sa strategijama, procesima i kulturom partnerskih poslovnih subjekata u svrhu bolje suradnje i dugoročnosti.

Brojni autori danas promatraju poslovnu inteligenciju isključivo kroz prizmu informatičkih tehnologija. Moderne podatkovne strukture (Skladišta podataka, OLAP sustavi) postali su jezgra poslovne inteligencije. Ovo razmatranje nije krivo, ali je nepotpuno budući da je poslovna inteligencija prvenstveno alat strateškog menadžmenta, dok je IT samo tehnologija na kojoj je sustav izgrađen. Dakle, poslovna inteligencija nije namijenjena IT stručnjacima, već menadžerima i zaposlenicima svih odjela. IT stručnjaci su prvenstveno podrška ovom sustavu,

ali su ujedno i njegovi korisnici (baš kao i ostali zaposlenici). Utjecaj poslovne inteligencije vidljiv je u svim aspektima poslovanja prvenstveno širenjem tog odjela izvan granica IT-a.

U svrhu provođenja poslovne inteligencije unutar organizacije koriste se analitički alati za prikupljanje i analizu podataka. Oni su potrebni za kategoriziranje podataka, njihovu ekstrakciju, konverziju i čišćenje. Kada su podaci spremni, počinje faza njihove analize u svrhu stvaranja informacija i znanja. Ključno je odabrati prave alate kako bi se dobile što bolje informacije. Ovi alati u dijele se na tri osnovne skupine:

- Alati za upite i izvještaje (*Querry and reporting tools*)
- Alati za rudarenje podataka (*Data mining tools*)
- OLAP alati (*Online Analytical Processing*)

Kako bi sustav poslovne inteligencije generirao vrijednost, esencijalno je odabrati odgovarajući sustav koji će se integrirati u trenutni poslovni sustav – naglasak na konektivnost. Moguće je i promijeniti trenutni poslovni sustav u svrhu što bolje integracije sa BI-om, ali takva promjena zahtijeva puno više resursa. Kada je sustav integriran potrebno je razraditi *backup* i *recovery plan* (sigurnosnu povratnu točku) radi zaštite podataka i generiranih informacija. Kako bi sustav bio što jednostavniji za korištenje krajnjem korisniku, potrebno je odabrati *frontend* alat sa jednostavnim i intuitivnim korisničkim sučeljem kako bi se ne-tehničko osoblje što lakše služilo sustavom. (Remy Na, 2016.).

2.1.4 Izvori i organizacija internih podataka

Podaci koji se generiraju svakodnevnim transakcijama unutar poslovnog subjekta mogu biti različitih struktura. U ovom radu usporedit će se sva osnovna koncepta pohrane podataka – baze i skladišta podataka. Izvori podataka unutar nekog poslovnog sustava dijele se na eksterne i interne. Eksterni izvori podataka odnose se na sve podatke iz okolina poduzeća bilo da se radi o tržištu, potrošaču, konkurenciji, zakonskim tijelima itd.

Neki od eksternih podataka su sljedeći:

- Statistika o poslovanju industrije
- Statistika o potrošačima
- Demografska slika tržišta
- Ekonomski indikatori okoline
- Statistika o dohotku
- Novac i kamatne stope
- Proizvodnja i prodaja konkurenata

Ovi podaci prikupljaju se iz nekoliko izvora što smanjuje točnost njihove procjene. Izvori ovih podataka su ankete, upitnici, intervjui, zavodi za statistiku, izvješća o poslovanju konkurenata itd. (Gašpar, 2016.) Kao što je već navedeno, raširenost društvenih medija omogućilo je vrlo efikasno prikupljanje i razumijevanje podataka o potrošačima, pa se taj podskup poslovne inteligencije smatra razvijenijim u odnosu na ostale eksterne izvore podataka. U ovom diplomskom radu, naglasak se ipak stavlja na podatke generirane interno unutar poslovnog subjekta.

Interni podaci generiraju se kroz poslovni informacijski sustav svakodnevnim poslovanjem. Prvenstveno se misli na računovodstveni sustav ili ERP sustav unutar kojega se zapisuju sve transakcije koje je poduzeće u prošlosti izvršilo. U ovom sustavu tako se nalaze sve promjene na stanjima odjela organiziranih u module:

- Financije i računovodstvo
- Ljudski Resursi i plaće
- Marketing i CRM
- Nabava
- Prodaja
- Proizvodnja
- Skladište
- Strateško planiranje

- **Modul Financije i računovodstvo** sastoji se od knjiga ulaznih računa, knjiga izlaznih računa, obračun PDV-a, kartice kupaca i dobavljača, knjiga primitaka i izdataka, knjiženje, salda konta, glavna knjiga, oprema, amortizacija, sitni inventar, matične knjige, obračun plaća, obračun honorara. Ovaj modul također pruža alate za analitiku financijskih pokazatelja.
- **Modul ljudskih resursa** objedinjuje sve podatke o zaposlenicima, njihovim plaćama, statistici radnog vremena, evaluaciji i efikasnosti, stanjima bolovanja i godišnjih odmora. Ovaj sustav također ima cjelokupnu hijerarhijsku strukturu poduzeća.
- **Modul marketing i CRM** objedinjuje sve marketinške procese: aktivnosti prije prodaje, istraživanje tržišta, baza klijenata, povijest kupnje po klijentima, demografija kupaca, status i procesiranje narudžbi, praćenje vraćenih proizvoda, vaučera itd.
- **Modul Nabava** obuhvaća MPR sustav (planiranje potreba za resursima), usklađivanje potreba plana proizvodnje, rezervacije, prioriteti naručivanja, prioriteti proizvoda, strategije naručivanja, automatsko naručivanje (*Just in time* koncept)
- **Modul prodaje** uključuje podatke o veleprodaji, maloprodaji, web shop narudžbama, distribuciji, transportu, upravljanjem poslovnicama, cjenicima, zakonskim evidencijama itd. Ovaj modul povezan je sa CRM modulom, te skupa tvore cjelinu.
- **Modul Proizvodnja** obuhvaća podatke o planu proizvodnje, radnim nalozima, praćenju kapaciteta strojeva, ljudi i ostalih resursa, rasporeda rada, praćenju učinaka po strojevima / ljudima / odjelima / projektima / radnim nalozima, praćenju i kontrola serije proizvoda itd.
- **Modul Skladište** prati skladištenje materijala, poluproizvoda, gotovih proizvoda, alata i inventara, zahtjevnice, narudžbenice, primke, izdatnice, kalkulacije, inventure, kartice, povratnice, otpisi, alokacija proizvoda, pretraživanja, grupiranja, označavanja, nalozi za zaprimanje itd.
- **Modul Strateško planiranje** odnosi se na plan proizvodnje, tehnološku pripremu proizvodnje, terminsko planiranje procesa i radnih naloga, procjene potrebnih resursa, procjene outputa.

2.1.4.1 Relacijske baze podataka

Relacijski model danas je najpraktičniji i najraširenji model podataka. Svoju popularnost duguje matematici i teoriji relacijske algebre na kojoj je utemeljen. Osnovne principe ovog modela postavio je matematičar E.F. Codd 1970 godine. Cijeli model zasnovan je na matematičkom pojmu relacije. Relacijska baza podataka sastoji se od skupa povezanih tablica odnosno relacija. Definicija jedne relacije naziva se relacijska shema i sastoji se od naziva relacije i popisa atributa (obilježja) koji su u sastavu te relacije. Definicija čitave baze podataka naziva se relacijska shema baze podataka, a sastoji se od relacijskih shema svih relacija koje ulaze u sastav baze podataka.

- **Relacija** je dvodimenzionalno područje koje se sastoji od stupaca, odnosno atributa i redaka, odnosno unosa (ili n-torki). Sve vrijednosti jednog atributa moraju biti istog formata, odnosno istog tipa podataka kako bi relacija bila konzistentna. Svaki stupac u tablici ima svoj naziv koji označava pojedini atribut. Svaki entitet relacije skup je atributa koji ga opisuju. (Garača, 2007.)

- **Entitet** relacijske baze podataka je tip objekta realnog ili apstraktnog svijeta čiji se podaci pohranjuju u relacijsku bazu podataka. Tako su u poslovnom svijetu primjeri entiteta: Klijent, Poslovnica, Proizvod itd.

- **Atribut** je svojstvo nekog entiteta. Jedan entitet može imati više atributa koji ga opisuju. Tako na primjer entitet Klijent može imati atribute kao što su: Ime; Prezime; Adresa; Spol i slično. Atribut dakle predstavlja vrijednosti koje opisuju neki entitet.

- **Primarni ključ** osnovni je atribut u relacijskoj bazi podataka koji jedinstveno identificira neki entitet. Primarni ključ mora zadovoljavati pravila identiteta i integriteta kako bi bio validan. Pravilo identiteta nalaže da ne mogu postojati dva unosa ovog atributa sa istom vrijednošću. Pravilo integriteta određuje da atribut primarnog ključa ne može poprimiti *null* ili ništavnu vrijednost. *Null* vrijednost, dakle, nije 0 već izostanak bilo kakve vrijednosti.

- **Strani ključ** je atribut ili skupina atribute neke relacije koja predstavlja primarni ključ u nekoj drugoj relaciji. Osnovna svrha stranog ključa je prikazati odnos dvije tablice.

Drugo pravilo integriteta relacijskog modela nalaže da: *„Baza podataka ne smije sadržavati vrijednost stranog ključa za kojeg ne postoje odgovarajuće vrijednosti primarnog ključa u baznoj relaciji“* – Garača, 2007.

Shema relacijske baze podataka na primjeru Tablica Klijent i Narudžba

Jednostavna relacijska shema u nastavku će biti objašnjena na primjeru baze podataka klijenata imaginarnog poduzeća. Prvoj tablici baze podataka glavni entitet je Klijent. Njegovi atributi određeni su u stupcima tablice: Ime; Prezime; Šifra; Grad; Ulica.

Tablica 1. Tablica entiteta Klijent; Microsoft Access 2016

Ime	Prezime	Šifra Klijenta	Grad	Ulica
Ante	Matić	1001	Split	Dobrilina 2
Mirko	Antić	1002	Zagreb	Varšavska 8
Ivana	Horvat	1003	Varaždin	Zagrebačka 19
Marko	Ivančić	1004	Prelog	Međimurksa 1
Marija	Ivančić	1005	Osijek	

Izvor: izračun autora, Microsoft Access

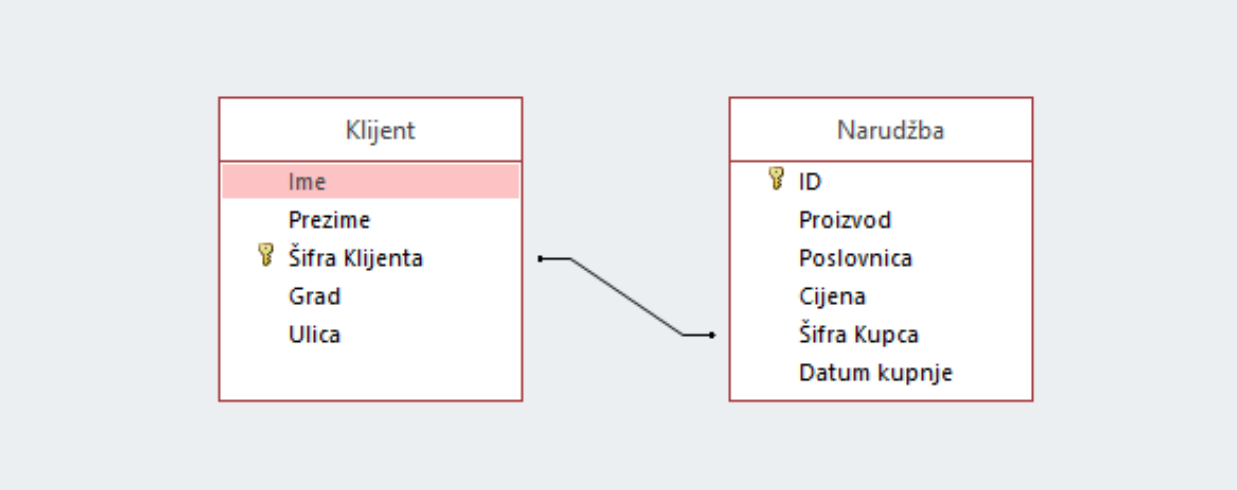
Primarni ključ ove tablice je atribut Šifra. Svaki entitet (klijent) ima unikatan primarni ključ što zadovoljava načelo identiteta. Primarni ključ također ne sadržava *Null* vrijednosti čime je zadovoljeno pravilo integriteta. U tablici je vidljivo da ostali atributi ne moraju zadovoljavati pravila integriteta i identiteta. Tako zadnji i predzadnji klijenti (Marko i Marija) imaju isto prezime Ivančić, dok zadnji klijent Marija Ivančić u atributu Ulica ima *Null* vrijednost. Primarni ključ važan je iz razloga što ni u jednom drugom atributu ne možemo garantirati da će ova dva pravila biti zadovoljena zbog realne prirode atributa (ista imena, prezimena, adrese, nedostatak informacija) pa je važno odrediti atribut bez kojega novi unos nije moguć – u ovom slučaju to je interno generirana šifra klijenta. Moguće je da primarni ključ bude i neki univerzalni (prirodni) atribut kao što su OIB, broj osobne iskaznice, IBAN itd. Za ove atribute univerzalno je određeno načelo identiteta.

Tablica 2. Tablica entiteta Narudžba;

ID	Proizvod	Poslovnica	Cijena	Šifra Kupca	Datum kupn
6	LCD Monitor	Split 2	3000	1002	1/8/2017
4	Laptop	Split 2	11000	1002	1/8/2017
3	Mobilni telefon	Zagreb 2	5000	1001	7/24/2017
1	Laptop	Zagreb 1	6500	1004	8/17/2017
2	LCD TV	Split 2	3500	1003	8/18/2017
5	Ruter	Zagreb 1	300	1005	8/19/2017

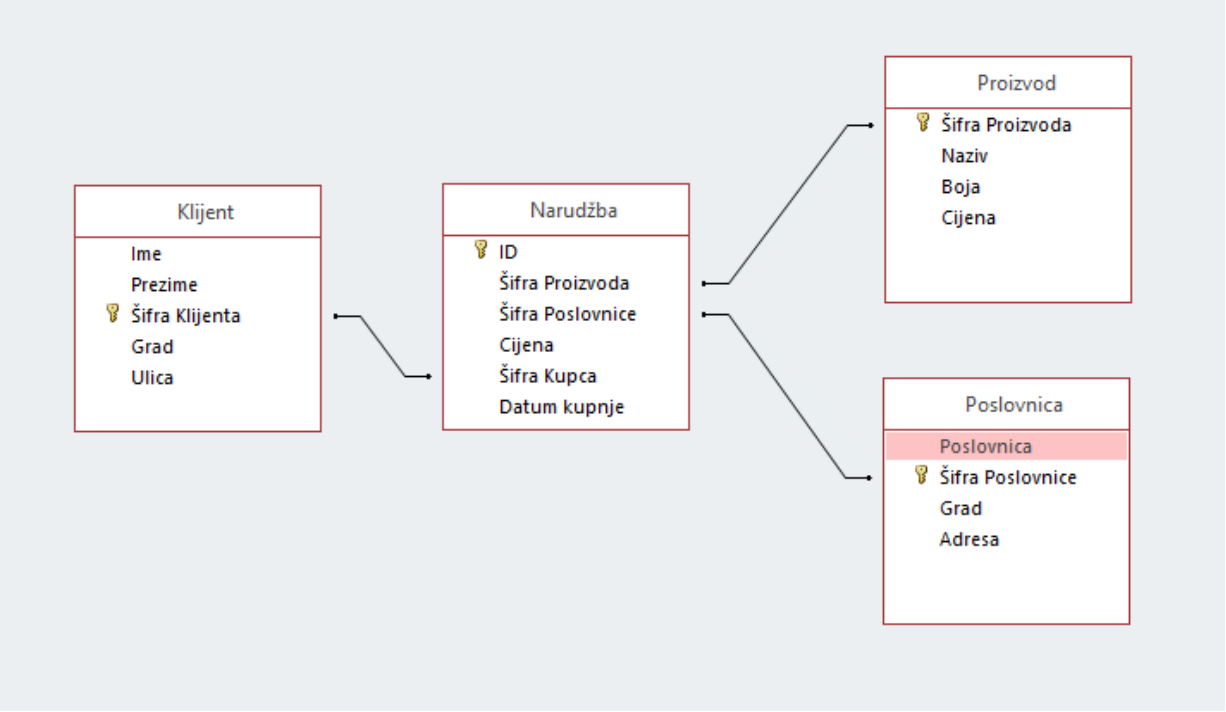
Izvor: Izračun autora, Microsoft Access

Druga tablica u ovoj bazi podataka tablica je narudžbi. U ovoj tablici kronološki su unesene sve narudžbe proizvoda imaginarnog poduzeća. Dakle entitet ove tablice je narudžba dok su ID, Proizvod, Poslovnica; Cijena; Šifra kupca i Datum kupnje njezini atributi. U ovoj tablici moguće je primijetiti isti atribut kao i u prethodnoj tablici – tablici klijenta.



Slika 2. Shema odnosa relacija Klijent i Narudžba. Izvor: izračun autora, Microsoft Access

Šifra klijenta u tablici Klijent predstavlja njezin primarni ključ. Ona je povezana sa stranim ključem relacije Narudžba pod atributom Šifra Kupca. Na ovaj način korisnik dobije puno više informacija o klijentu iz tablice Narudžba do kojih dolazi upitima na bazu – ovo će detaljnije biti objašnjeno u sljedećem poglavlju. U poslovnom scenariju vjerojatno bi i ostali atributi u tablici Narudžba bili relacijski povezani sa pripadajućim im tablicama. Sljedeća shema prikazuje veze tablica Proizvod, Poslovnica i dosadašnje tablice Klijent sa tablicom Narudžba.



Slika 3. Relacijska shema tablica; Izvor: izračun autora, Microsoft Access

Na ovaj način korisnik dobiva proširen pogled na jedan unos u tablici Narudžba. Saznaje više detalja o kupljenom proizvodu i poslovnici u kojoj je proizvod prodan.

2.1.4.2 Pretraživanje baze podataka

Pretraživanje baze podataka vrši se takozvanim upitima na bazu. Upit (eng. *Query*) je objekt baze podataka koji korisniku pruža pogled jedne ili više tablice te baze prema postavljenim kriterija. Upiti se dijele na dvije osnovne skupine:

- **Upit izdvajanja** je upit koji pronalazi i izdvaja elemente (slogove) iz jedne ili više povezanih tablica. Ovaj upit ne kreira novu tablicu u bazi, već kreira samo privremenu tablicu za prikaz postojećih podataka u postojećim tablicama.
- **Akcijski upit** je upit u bazu podataka koji mijenja podatke unutar baze. Ovim upitom moguće je dodavati nove slogove, ali i dodavati kompletno nove tablice. Najčešći akcijski upiti su *Update Query* (ažuriranje podataka); *Delete Query* (brisanje podataka); *Append Query* (dodavanje podataka); *Make Table Query* (kreiranje nove tablice).

Upiti na bazu podataka vrše se u *Query* jezicima. Ovi jezici osmišljeni su isključivo za navigaciju kroz baze podataka, bilo da se radi o upitima izdvajanja ili akcijskim upitima. **SQL** (*Structured Querry Language*) zasigurno je najrašireniji *Querry* jezik. Ovaj jezik datira još iz 1974 pojavom Coddovog relacijskog modela (1970). Od tada pa do danas SQL se koristi u brojnim poslovnim rješenjima za upravljanje bazama podatka.
MySQL *open-source* platforma je za upravljanje relacijskim bazama podataka koji je temeljen na SQL tehnologiji. **MsSQL** Microsoftovo je rješenje za menadžment bazama podataka. Za razliku od MySQL-a, MsSQL je *proprietary* (privatni) softver čije korištenje nije besplatno. Ovaj sustav popularan je zbog cjelokupnog Microsoft nasljedstva unutar poslovnih informacijskih sustava. **Oracle** je također popularno rješenje za upravljanje bazama koje je bazirano na SQL tehnologiji.

Neki primjeri SQL upita:

- **SELECT** (upit će se postaviti na tablicu Klijent)

```
SELECT Ime, Prezime, Grad FROM Klijent;
```

Rješenje:

Tablica 3. Rješenje upita SELECT

Ime	Prezime	Grad
Ante	Matić	Split
Mirko	Antić	Zagreb
Ivana	Horvat	Varaždin
Marko	Ivančić	Prelog
Marija	Ivančić	Osijek
*		

Izvor: izračun autora, Microsoft Access

Važnost relacijskih tablica dolazi do izražaja kod funkcije UNION u kojoj je moguće odabrati atribute iz dvije relacijski povezane tablice koji će se prikazati u jedinstvenoj tablici.

- **UNION SELECT** (upit se postavlja na tablice Narudžba i Klijent)

```
SELECT Proizvod, Cijena FROM Narudžba
UNION
SELECT Ime, Prezime, Grad FROM Klijent;
```

Rješenje:

Tablica 4. Rješenje upita SELECT u kombinaciji s upitom UNION

Proizvod	Cijena	Ime	Prezime	Grad
Laptop	6500	Marko	Ivančić	Prelog
LCD TV	3500	Ivana	Horvat	Varaždin
Mobilni telefon	5000	Ante	Matić	Split
Laptop	11000	Mirko	Antić	Zagreb
Ruter	300	Marija	Ivančić	Osijek
LCD Monitor	3000	Mirko	Antić	Zagreb
*				

Izvor: izračun autora, Microsoft Access

Dobivena tablica kombinira podatke koji se nalaze u dvije odvojene tablice, ali ih prikazuje u novoj tablici upita. Cijela ideja relacijskih baza podataka odnosi se na „razbijanje“ baze podataka u više logički objedinjenih tablica koje su relacijski povezane. Korisnik naknadno svojim upitima može odrediti poglede (tablice u kojima su prikazani upiti) u kojima se nalaze podaci iz bilo koje relacijski povezane tablice.

Relacijska teorija tako definira 5 vrsta relacija koje su vezane za upite (Garača, 2007)

- **Pogled** (*view*) je virtualna relacija izvedena pretraživanjem jedne ili više baznih relacija. Tablice dobivene SQL upitima predstavljaju poglede.
- **Izvedena relacija** (*Snapshot*) je slična pogledu, s tim da osim definicije u bazi podataka ima i svoje podatke
- **Rezultat pretraživanja** je relacija koja je dobivena pretraživanjem baze podataka relacijskim jezicima. Tablice dobivene SQL upitima u prethodnom paragrafu predstavljaju ovu relaciju.
- **Međurezultat pretraživanja** (*intermediate result*) rezultat je pretraživanja koji se koristi kao input novom upitu.

- **Privremena relacija** (*temporary relation)* relacija je kratkog vijeka koja se privremeno upisuje u bazu podataka, ali se briše završetkom transakcije.

2.1.4.3 Skladišta podataka

Moderni poslovni sustavi na dnevnoj bazi generiraju ogromnu količinu podataka koji se pohranjuju u operativne baze podataka. U modernom poslovanju količina generiranih podataka kreće se eksponencijalnim trendom (popularan *buzzword* koji objašnjava taj trend je **eksplozija podataka**). Ovi operativni sustavi dizajnirani su da prikazuju stvarno, najnovije stanje poslovnog sustava. Na taj način podaci nemaju vremensku komponentu, već se novim stanjem nekog elementa zamjenjuje njegovo staro stanje. Zbog eksplozije podataka, poduzeća ulaze u krizu informacija (informacijska kriza objašnjena je na početku drugog poglavlja ovog rada). Ogromne količine nepročišćenih i neprilagođenih podataka imaju vrlo malu informacijsku vrijednost. Liautaud i Hammond (2006.) opisuju pojam anarhije podataka koji nastaje kada organizacijske jedinice unutar poduzeća počnu razvijati informacijske sustave za vlastite transakcijske potrebe koji nisu kompatibilni informacijskim sustavima ostalih jedinica. Nedostatak integracije dovodi do redundantnosti, proturječnosti i netočnosti podataka.

Navedeni fenomeni potakli su razvitak integriranog sustava koji će objediniti sve generirane podatke i pohraniti ih po istom standardu. Skladište podataka podatkovni je sustav koji objedinjava podatke iz više baza podataka nekog poslovnog sustava. Ono je integrirani i centralizirani sustav koji se sastoji od pročišćenih podataka prilagođenim potrebama složenih poslovnih analiza.

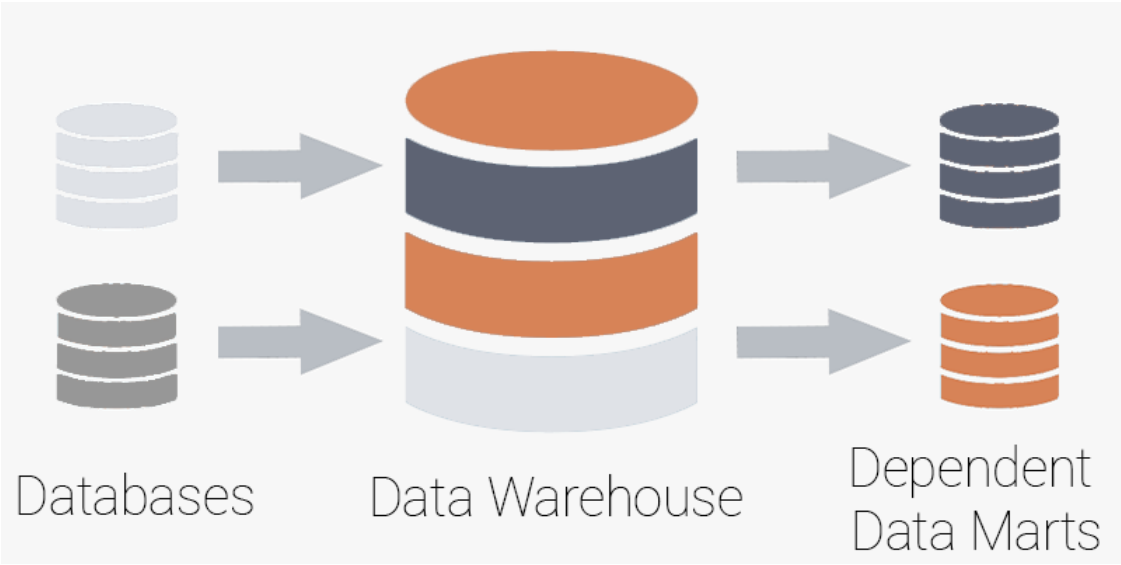
Podaci se izvlače iz operativnih baza, prilagođavaju standardima skladišta, označavaju vremenskim oznakama i pohranjuju u skladište podataka. (Šakić, Bagić, 2013.) Imnon (1992.) postavlja koncept skladišta podataka i dodjeljuje mu četiri osnovna obilježja:

- **Subjektnost** podrazumijeva usmjerenost skladišta na subjekte, odnosno predmete
- **Neporomjenjivost sadržaja** označava da kada su podaci jednom pohranjeni unutar skladišta, oni se ne mogu mijenjati, već se novo staje objekta prikazuje kao novi unos sa novom vremenskom komponentom
- **Integriranost** skladišta znači da su svi podaci pohranjeni po određenim pravilima, uz naglasak na konzistentnost.

- **Povezanost s vremenom** još jednom označava važnost vremena, odnosno povijesnog stanja objekata unutar skladišta podataka. Ovo je najveća vrijednost skladišta podataka u odnosu na operativne baze jer omogućuje vremenske analize.

Upravo zbog svoje povijesne perspektive, skladište podataka namijenjeno je prije svega provođenju analitičkih obrada u svrhu podrške procesima odlučivanja. Postojanje skladišta podataka preduvjet je za moderne sustave poslovne inteligencije.

Laursen i Thorlund (2010) prema Luetiću (2013.) dijele skladište podataka na tehnički i poslovni dio. Tehnički dio je obuhvaća *backend* procese prikupljanja, čišćenja, skladištenja i strukturiranja podataka, dok poslovni dio koristi te sada već pročišćene podatke u kreiranju složenih analiza. Oni smatraju kako je upravo ti tehnički procesi najzahtjevniji i najopsežniji u skladištenju podataka. Takozvani ETL procesi (engl. *Extract, Transform, Load*) imaju zadatak prikupiti, prilagoditi i napuniti podatke u skladište. Podaci se, kao što je već navedeno, ekstrahiraju iz operativnih baza poslovnih sustava (ili podsustava jednog poslovnog sustava). Kako bi podaci ovisno o svom porijeklu i namijeni bili što bolje organizirani unutar skladišta, oni se organiziraju u takozvana spremišta podataka (eng. *Data mart*). *Data mart* podatkovni je podskup skladišta podataka. Podaci su podijeljeni u *data martove* kako bi upiti na skladište podataka bili brži. Određeni *data mart* relevantan je samo za određeni odjel, a s obzirom da je obujam podataka smanjen, performanse *data marta* znatno su bolje od performansi kompletnog skladišta.



Slika 4. Koncept *data martova* unutar skladišta podataka;

Izvor: https://www.betterbuys.com/bi/definitive-guide-bi/data-warehousing

Ovaj koncept najlakše je objasniti na primjeru. Uzmimo da neko proizvodno poduzeće ima skladište podataka u koje se unose podaci iz svih podsustava informacijskog sustava: proizvodnja; prodaja; financije; računovodstvo; marketing i planiranje. Odjel prodaje u svojim analizama želi imati pristup isključivo svim podacima iz podsustava prodaja i određenim podacima iz podsustava financije i planiranja. Za njih se tada kreira *data mart* Analiza prodaje, u koji se unose samo njima relevantni podaci iz cjelokupnog skladišta. Također je važno primijetiti da neki data mart ne mora biti povezan samo sa jednim izvorom podataka (jednom operativnom bazom, u ovom slučaju baza prodaje), već može kombinirati pročišćene podatke iz nekoliko operativnih baza, odnosno podsustava poslovnog informacijskog sustava.

Što se tiče analitičkog dijela skladišta podataka, Banek (2005.) ih dijeli na tri skupine prema složenosti, vremenu potrebnom za izvršenje i znanju korisnika o skladištu:

- **Izvještaj** (eng. *Reporting*) najjednostavnija je metoda korištenja podataka skladišta. Ona podrazumijeva minimalno znanje o strukturi samih podataka. Izvještaj se kreira postavljanjem upita na skladište podataka koji su fiksne strukture i vrlo su slični upitima na transakcijsku bazu (SQL). Razlika u odnosu na transakcijsku bazu, naravno je u kvaliteti podataka i vremenskoj komponenti skladišta.

- **Analitička obrada** (OLAP – eng. *On-Line Analytical Processing*) složeniji je proces analize i obrade podataka skladišta. Upiti se u ovom procesu ne postavljaju direktno na skladište (ili na martove), već se postavljaju na rezultate prethodnih upita čija struktura nije unaprijed poznata. OLAP alati prilagođeni su višedimenzijskim modelu podataka (više u nastavku). Upravo zbog toga, korisnik mora poznavati strukturu skladišta podataka te mora imati dovoljno znanja za rad sa složenijim analitičkim alatima. Analitičkom obradom podataka bave se poslovni analitičari i menadžeri unutar organizacija.

- **Rudarenje podataka** (eng. *Data Mining*) najsloženiji su analitički procesi koji se izvršavaju na podacima iz skladišta. Cilj rudarenja podataka je pronaći obrasce i odnose unutar velike količine podatka koji nisu vidljivi odmah na „prvi pogled“. Alati za rudarenje podataka koriste složene algoritme kako bi se na velikim nizovima podataka identificirali ti „skriveni“ obrasci i odnosi. Rudarenjem podataka u organizaciji se

akumulira sasvim novo znanje koje se može pretvoriti u ogromnu konkurentsku prednost. Korisnici koji provode rudarenje podataka na skladištu podataka moraju razumjeti kompleksne algoritme i strukture u kojima su podaci organizirani. Za ovu analizu potrebno je stručno obrazovanje i iskustvo.

Tablica 5. Razlika Izvještaja, OLAPa i Rudarenja podataka;

	IZVJEŠTAJ	**OLAP**	**RUDARENJE PODATAKA**
Zadatak	Ekstrahija detaljnih i sumiranih podataka	Sumiranje podataka, trendovi i prognoze	Otkrivanje novog znanja
Rezultat	Informacija	Analiza	Uvid i prognoza
Metoda	Dedukcija (postavljanje pitanja, verificiranje sa podacima)	Multidimenzijsko modeliranje podataka, Agregacija i statističke metode	Indukcija (razvijanje složenih algoritamskih modela i njihova upotreba na podacima)
Primjer	Koji korisnici su uplatili životno osiguranje u zadnje 3 godine?	Kojoj prihodnoj skupini pripadaju ti korisnici po regiji, po godini?	Koji korisnici će uplatiti životno osiguranje u sljedećih 6 mjeseci i zašto?

izvor: http://saslearn.blogspot.hr/2012/10/introduction-to-business-intelligence.html

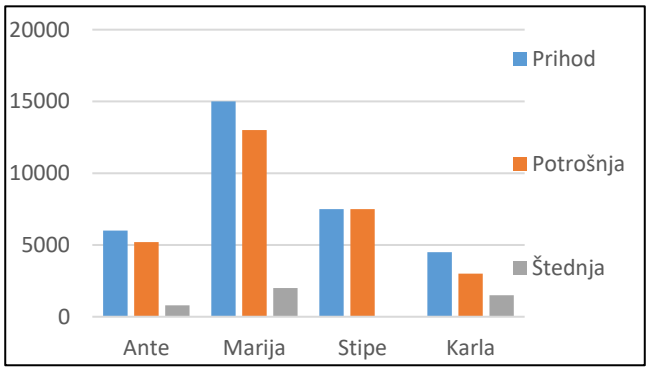
2.1.5 Višedimenzijske strukture i složenije analize podataka

2.1.5.1 Višedimenzijski model podataka

Ljudsko poimanje svijeta temeljeno našim osjetilima je trodimenzionalno. Ljudski mozak, naviknut na trodimenzionalnost može percipirati tri dimenzije ne samo s fizičkog aspekta prostora, nego i s aspekta apstraktnih objekata koji nemaju svoju fizičku verziju. U ovom radu primarno je riječ o podacima, pa tako čovjek može percipirati i trodimenzionalnu strukturu podataka.

Podaci se uglavnom prikazuju u dvije dimenzije te ih kao takve ljudi najlakše percipiraju. Najbolji primjer dvodimenzionalne strukture podataka jest tablica. Tablica sastavljena od redaka (jedna dimenzija) i stupaca (druga dimenzija) sadrži elemente čije je odnose čovjeku jednostavno shvatiti „na prvi pogled“. Vizualni prikaz podataka također se (uglavnom) prikazuje kroz dvije dimenzije.

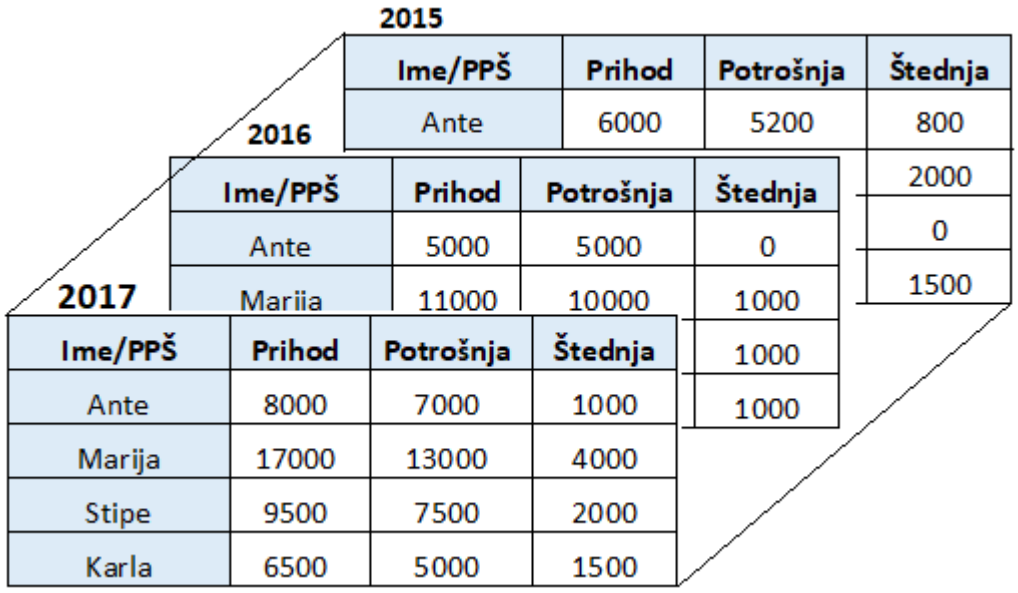
Ime/PPŠ	Prihod	Potrošnja	Štednja
Ante	6000	5200	800
Marija	15000	13000	2000
Stipe	7500	7500	0
Karla	4500	3000	1500



Slika 5. Primjer dvodimenzionalnog prikaza podataka; Izvor: izračun autora, Microsoft Excel

U dvodimenzionalnom prostoru čovjek promatra podatke kroz identifikacijsku i locirajuću dimenziju. Kada se u izračun doda još jedna dimenzija, razumijevanje postaje znatno otežano. Trodimenzionalni prikaz podataka zahtijevan je i neintuitivan, pa je i razumijevanje trodimenzionalnih podatkovnih struktura otežano.
Da se na primjer u gornju tablicu dodaju još tri godine u kojima se pratio odnos prihoda, potrošnje i štednje prikaz i razumijevanje podataka postalo bi znatno kompleksnije. Prikaz takvih podataka najlakše se može zamisliti kroz geometrijski oblik kocke.

Kako bi se najbolje predočila trodimenzionalnost podataka, oni se uglavnom prikazuju poput niza dvodimenzionalnih tablica prikazanih u pravokutnom geometrijskom tijelu. Višedimenzijski podaci (oni koji sadrže više dimenzija) ljudskom mozgu gotovo je nemoguće pojmiti zbog toga što probijaju prostorne dimenzije ljudske percepcije.



2015

Ime/PPŠ	Prihod	Potrošnja	Štednja
Ante	6000	5200	800

2016

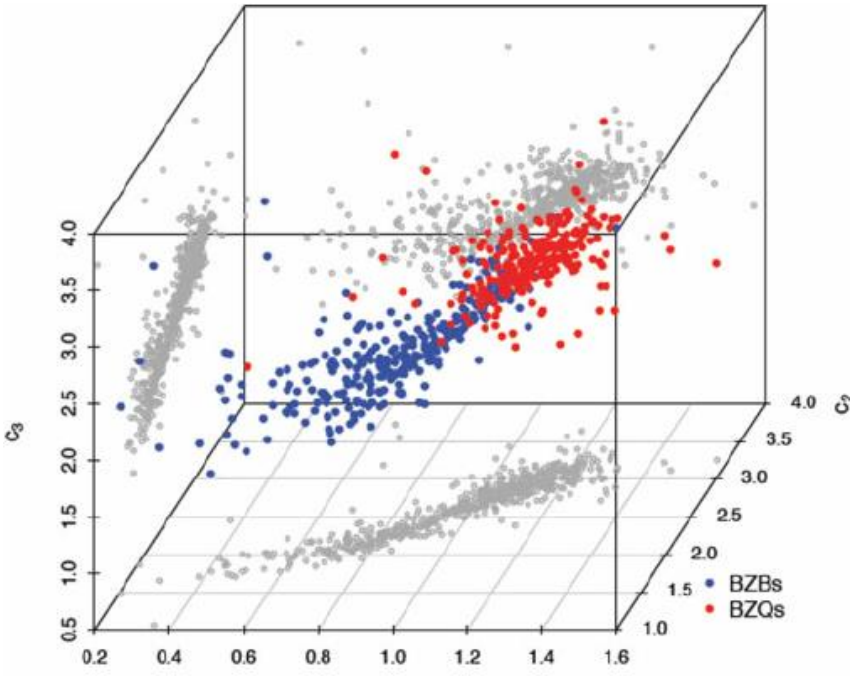
Ime/PPŠ	Prihod	Potrošnja	Štednja
Ante	5000	5000	0
Marija	11000	10000	1000

2017

Ime/PPŠ	Prihod	Potrošnja	Štednja
Ante	8000	7000	1000
Marija	17000	13000	4000
Stipe	9500	7500	2000
Karla	6500	5000	1500

2000
0
1500
1000
1000
1000

Slika 6. Prikaz trodimenzionalne strukture podataka; Izvor: izračun autora



Slika 7. Trodimenzionalni vizualni prikaz podataka; izvor: https://www.researchgate.net/

Iako otežano, trodimenzionalne strukture je moguće percipirati kako se vidi na primjeru kocke sačinjene od tablica i grafičkog trodimenzionalnog prikaza. Višedimenzijski sustavi objašnjavaju se oblikom hiperkocke. Hiperkocka tijelo je sastavljeno od osam 3D kocaka od kojih svaka dodiruje po šest drugih preko svojih dvodimenzionalnih stranica. Prikaz hiperkocke u 3D svijetu vrlo je kompleksan, te ga je najlakše prikazati animacijom pa je zbog toga izostavljen iz ovog rada.

Višedimenzijski model podataka konceptualno je podijeljen na činjenice, mjere, dimenzije i hijerarhije (Banek, 2005) prema (Šakić i Bagić, 2013):

- **Činjenica** (eng. *Fact*) polazna je točka poslovne analize. Ona je istinita informacija koja predstavlja točku interesa u donošenju poslovne odluke

- **Mjera** (eng. *Measure*) je opisna varijabla, odnosno atribut iz neprekidnog skupa vrijednosti, koji se prikazuje numerički, a opisuje činjenicu. Mjere predstavljaju osnovni input u analitičkim metodama koje se provode na podacima iz skladišta. Primjer odnosa činjenice i mjere unutar podatkovne strukture: u skladištu podataka neke banke, određeni bankovni račun je činjenica, a njegovo stanje je mjera te činjenice.

- **Dimenzija** (eng. *Dimension*) nezavisni je parametar koji opisuje neku činjenicu, a sastoji se od atributa diskretnog, odnosno konačnog skupa vrijednosti. Primjer dimenzije u skladištu podataka je vrijeme. Tako bi na prijašnjem primjeru skladišta podataka banke u kojemu je činjenica račun, a mjera stanje računa dimenzija bila vremenska točka u kojoj je taj račun poprimio to stanje (datum).

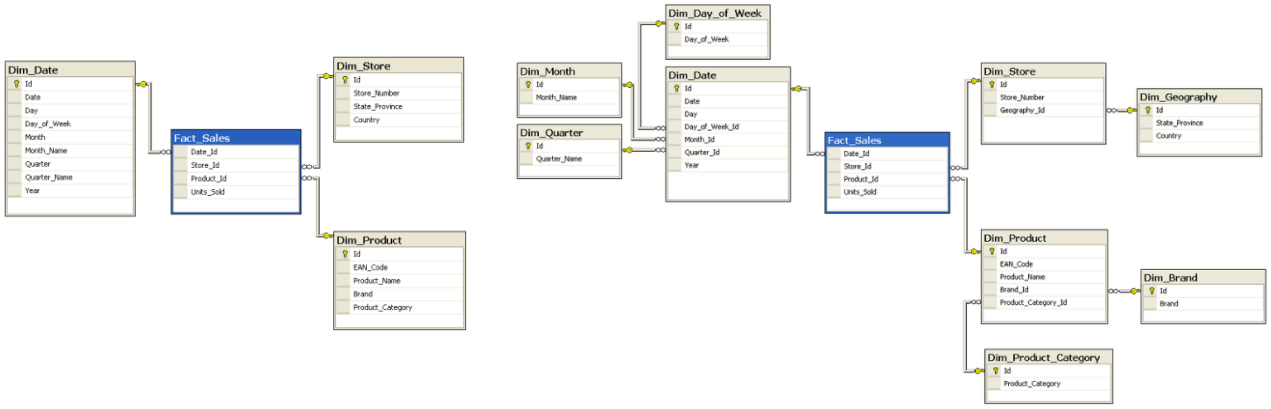
- **Dimenzijska hijerarhija** (eng. *Dimensional hierarchy*) označava hijerarhiju unutar jedne dimenzije i njenih pod skupova. Činjenica može biti opisana mjerama sa različitih stupnjeva granularnosti, odnosno detaljnosti. Granularnost podataka određuje na kojem se hijerarhijskom nivou promatra mjera neke činjenice. Primjer hijerarhija i granularnosti: Godina (najviša granularnost) → Kvartal → Mjesec → Datum → Sat (najmanja gradularnost).

2.1.5.2 OLAP

OLAP (eng. *On-Line Analytical Processing*) softverska je tehnologija koja se koristi u pretraživanju i analizi podataka unutar skladišta. Ova tehnologija prilagođena je ne-tehničkim korisnicima pa ne zahtijeva široko znanje iz struktura podataka i informacijskih tehnologija. OLAP je tako namijenjen menadžerima i analitičarima u svrhu brzog i efikasnog pristupa ogromnim količinama podataka koji su pohranjeni u skladištu. Preduvjet OLAP sustava je višedimenzionalnost podataka koja je detaljnije objašnjena u prethodnom poglavlju.

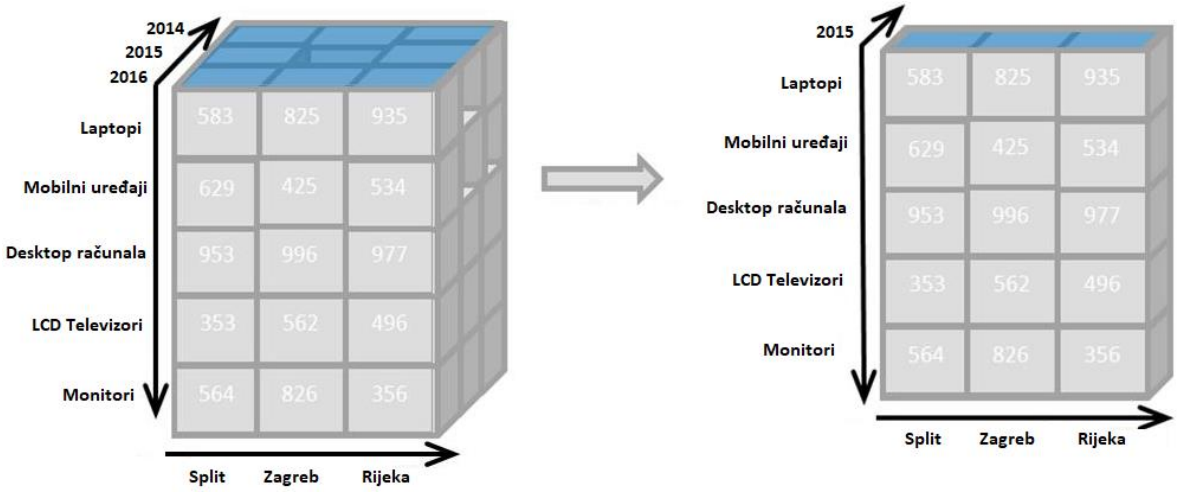
OLAP se razvio iz OLTP sustava koji su se „naslanjali“ na operativne baze poslovnih sustava. Osnovna značajka OLTP-a njegova je brzina koja je potrebna kada se upiti vrše na operativni DB (*Database*). Ovaj sustav namijenjen je svakodnevnim poslovnim transakcijama, ali se s vremenom počeo koristiti i u analitičke svrhe. Pojavom skladišta podataka iz OLTP-a se razvio OLAP koji je postao specijalizirano analitičko rješenje. Dakle osnovna razlika između ova dva sustava je ta što se OLTP koristi transakcijski, a OLAP analitički. Samim time različiti su i izvori podataka ovih sustava; OLTP koristi *snapshot* operativne baze u zadanoj vremenskoj točci, dok OLAP koristi Skladišta podataka. Ovi sustavi razlikuju se i u operacijama koje koriste; OLTP uglavnom koristi brze i jednostavne on-line transakcije (INSERT, UPDATE, DELETE), dok OLAP koristi složenije i sporije analitičke (statističke i agregacijske) operacije.

OLAP Kocka predstavlja način na koji su podaci prikazani unutar OLAP sustava (unutar *datamartova*). Kocka kao objekt se koristi kada podaci imaju tri dimenzije, dok se kod višedimenzijskih podataka koristi shema hiperkocke. Unutar kocke nalaze se Činjenice i mjere organizirane prema dimenzijama. Podaci unutar skladišta uglavnom su organizirani u zvjezdastu (eng. *star*) ili pahuljastu (eng. *snowflake*) shemu.

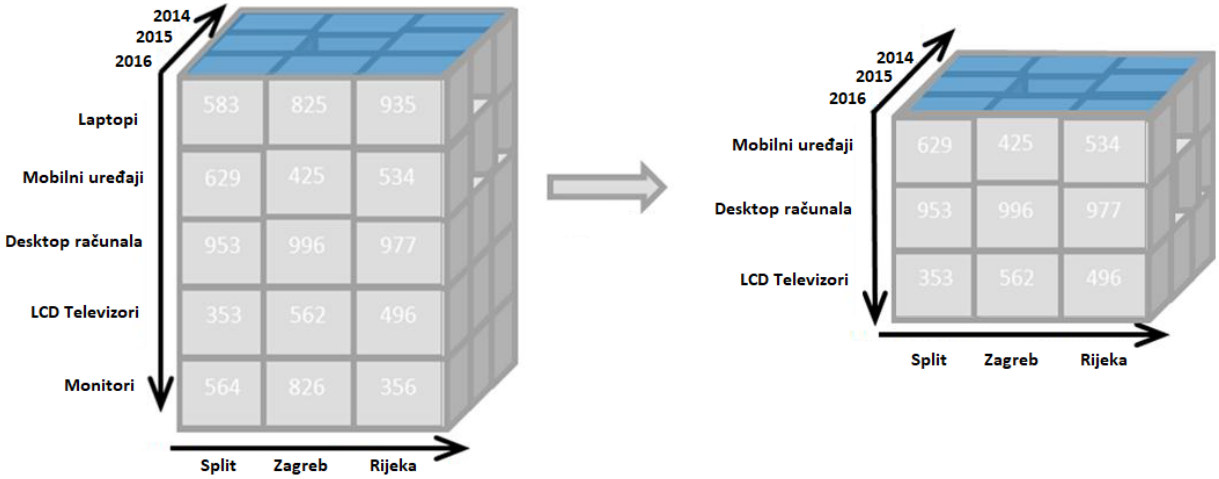


Slika 8. Zvjezdasta vs. Pahuljasta shema podataka; Izvor: izračun autora

S obzirom da su podaci unutar OLAP sustava višedimenzionalni, na njih se primjenjuju neke od osnovnih analitičkih metoda (Vitt et al., 2012). Metoda raslojavanja i presijecanja (engl. *Slice and Dice Method*).



Slika 9. OLAP Slicing; Izvor: izračun autora



Slika 10. OLAP Dicing; Izvor: izračun autora

Uz metode raslojavanja i presijecanja, na OLAP kockama se često primjenjuju i metode pivotiranja i metoda bušenja (*drill down*). Pivotiranje se odnosi na promjenu dimenzija po osima (x,y,z) a *drill down* metoda mijenja pogled unutar dimenzijske hijerarhije (primjer. godina → mjesec).

2.1.5.3 Rudarenje podataka

Rudarenje podataka proces je automatiziranog otkrivanja nepoznatih obrazaca i korelacija među podacima unutar podatkovnih struktura (baza podataka, skladišta podataka, *datamartova*). Luetić (2013., p. 66.) u svojoj doktorskoj disertaciji referencira Kudybai i Hoptroffu (2001) te navodi kako je rudarenje podataka „*skup tehnologija koje ujedinjuju primjenu statističkih tehnika i matematičkih formula kao alata, pokušavajući identificirati značajnost međuveza između povijesnih podataka koje bi potom mogli biti iskorišteni za prognoziranje, izvršenje analize osjetljivosti ili samo identificiranje značajnosti veza između podataka koju su na raspolaganju. Primjenom ovih alata moguće je otkriti do tada neprepoznate matrice ponašanja, lakše i točnije se mogu predvidjeti budući trendovi što bi u konačnici trebalo rezultirati donošenjem poslovnih odluka utemeljenih na znanju*“

Osnova procesa rudarenja podataka složeni su matematički algoritmi kojima se testiraju odnosi među podacima. Panian (2007.) navodi kako su ovi alati vrlo kompleksni za upotrebu, pogotovo iz perspektive zaposlenika koji nisu direktno povezani s IT-om. Proizvođači ovakvih alata pokušavaju pojednostaviti njihovu upotrebu uz istovremeno zadržavanje razine funkcionalnosti. Preduvjet korištenja DM (eng. *data mining*) alata je poznavanje podatkovnih struktura u kojima su podaci pohranjeni. Ne-tehničko osoblje obično ima ograničeno znanje o bazama i skladištima podataka, što znatno ograničava njihovo korištenje naprednih DM alata. Drugi preduvjet za korištenje ovih alata poznavanje je matematičkih algoritama koji će se primjenjivati na podatke u procesu njihove analize. Uspješno provođenje DM procesa zahtijeva stručno obrazovanje u tom području, ali i iskustvo s podatkovnim strukturama.

Luetić (2013) u svojoj disertaciji objedinjuje autore (Kudyba i Hoptroft, 2001; Vit et al., 2002; Panian i Klepac, 2003; Bilandžić, 2008; Garača, 2008; Watson, 2009.) te navodi najčešće korištene metode rudarenja podataka:

- **Metoda potrošačke košarice** otkriva skrivene korelacije kod prodaje roba, odnosno otkriva asocijativna pravila koja prikazuju koji se parovi artikala kupuju zajedno i s kojom vjerojatnošću
- **Klasteriranje** je metoda koja provodi grupiranje objekata (slogova) prema njihovim karakteristikama odnosno atributima. Cilj klasteriranja svrstavanje je objekta sa istim obilježjima unutar grupa kako bi se svakom objektu moglo pristupiti naizgled

individualno (iako se svim elementima neke grupe pristupa jednako). Klasteriranje pojavom društvenih mreža postaje iznimno popularna metoda u analizama potrošača.

- **Neuronske mreže** metoda je koja se koristi za predviđanje trendova i prognoziranje na temelju povijesnih podataka. Svoje ime duguje neuropsihologiji jer se ideja neuronskih mreža javila kroz saznanja o ponašanju živčanih stanica
- **Stablo odlučivanja** metoda je kojom se određuju varijable koje najznačajnije opisuju neki skup. Ova metoda se skupa s klasteriranjem koristi u modernim analizama sklonosti potrošača

Važnost Rudarenja podataka možda je najbolje objasniti primjerom iz poslovnog svijeta – primjena rudarenja podataka u otkivanju bankovnih prevara. Poznato je kako banke ulažu ogromna sredstva u najsofisticiranije sigurnosne sustave, kako fizičke tako i softverske. Unatoč svim sigurnosnim standardima, nemoguće je klijentima garantirati stopostotnu sigurnost. Kako bi se sigurnosni sustavi kvalitetno nosili sa potencijalnim anomalijama potrebno ih je prvenstveno razotkriti. Revizija ogromne količine povijesnih podataka koji sežu unazad nekoliko godina tradicionalnim metodama obrada iziskivalo je visoke resurse, prvenstveno u vidu sati rada, pa nije postojala adekvatna analiza. Danas se tehnikama rudarenja podataka vrlo brzo i efikasno može obraditi velika količina podataka, te je moguće uvidjeti do tada nepoznate uzročno-posljedične veze. Analizom ekstrema (eng. *outliera*) moguće je uvidjeti na probleme u sustavu koji odskaču od prosjeka. **CRISP**, kao jedna od popularnijih metoda DM-a identificira rudarenje podataka kroz različite faze projekta:

Razumijevanje poslovanja inicijalna je faza CRISP metodologije, a u ovom slučaju ona se odnosi na utvrđivanje anomalija i sumnjivijih distribucija unutar cjelokupnog *dataseta* bankovnih transakcija. Iznimno je važno determinirati trenutno stanje sustava kako bi mogli evaluirati model po završetku procesa rudarenja.

Slijedeća faza CRISP-a je **razumijevanje podataka**. Inicijalni podaci koji će biti korišteni za modeliranje trebaju biti prikupljeni i verificirani. Prikupljanje podataka često je proces koji se odvija u dugim vremenskim intervalima. U slučaju banke prikupljanje podataka odvija se na dnevnoj bazi kroz obavljanje transakcija. Važno je i prikupiti podatke o ranijim anomalijama koji sadržavaju atribute vremena, tipa i broja ponavljanja.

Pripremanjem podataka provode se svi nužni procesi koji će podatke sistematizirati u format koji je najprikladniji za softver koji se koristi za samo rudarenje podataka. Na pripremanje podatak otpada velik udio uloženog vremena te je važno da prilikom „krojenja“ podataka oni zadrže, odnosno, povećaju svoju inicijalnu vrijednost. Podaci su uglavnom porijeklom iz

eksternih operativnih baza pa ih je prije modeliranja potrebno „očistiti“ od nerelevantnih atributa. U ovoj fazi se odabiru i algoritmi za analizu podataka.

Modeliranje podataka započinje već u prethodnoj fazi odabirom algoritma za analizu. Ima mnogo tehnika modeliranja, a u ovom slučaju korištena je tehnika stabla odlučivanja (eng. *decision tree*), odnosno korištenje baza podataka za pronalaženje, analizu i validaciju bankovnih prevara prema povijesnim podacima.

Pretposljednja faza odnosi se na **evaluaciju** samog **modela**, odnosno rezultata dobivenih rudarenjem. Uspoređujemo rezultate sa zdravom logikom i racionalnošću kroz prizmu poslovnih ciljeva koji su postavljeni u prvoj fazi procesa.

Last but not least posljednja faza CRISP metodologije je **implementacija dobivenih rezultata**, odnosno novih znanja u postojeći sustav kako bi ga unaprijedili i dodatno zaštitili. Naravno, kako bi sustav bio što bolji potrebno ga je konstantno evaluirati jer čim se otkloni jedan problem slijedi novi, još teži, još skriveniji u sustavu.

S obzirom na kompleksnost procesa, Vitt et al. (2002.) navode kako *Data mining* često otvara više novih pitanja nego što daje odgovora na unaprijed postavljena pitanja pa nije rijedak slučaj da se koristi u kombinaciji s OLAP alatima.

2.1.5.4 OLAM

Korištenjem OLAP alata korisnik traži veze među dimenzijama koje nisu nužno temeljeni na složenim matematičkim algoritmima. Korisnik do rezultata dolazi uglavnom jednostavnim operacijama kao što su već spomenute metoda raslojavanja, presijecanja, pivotiranja i bušenja (eng. *drill down*). Odgovori na postavljena pitanja već su servirani, samo je potrebno pronaći pravi pogled na podatke koji sadrži isključivo relevantne informacije za korisnika. Za razliku od OLAP-a DM podrazumijeva automatiziran proces pronalaženja veza i kauzalnosti među podacima na kojima se vrše složeni matematički algoritmi. U ovom slučaju, algoritam umjesto korisnika provodi analize i istražuje (Silić i Šamanović, 2011)

Watson i Brohman (2006) prema (Silić i Šamanović, 2011) definiraju razliku između OLAP i DM u kontekstu njihovih korisnika. Oni koji rade s OLAP alatima najčešće su ne-tehničko osoblje koje u analizama koristi već predefinirane izvještaje na dostupnim mjerama i dimenzijama, dok napredniji korisnici kreiraju složenije upite na skladište. DM analitičari obično rade sa specijaliziranim softverom za koji trebaju biti stručno i tehnički obrazovani kako bi uspješno primijenili složene algoritme u analizi.

Kombinacija ove dvije metode (OLAP i DM) naziva se OLAP *mining* ili skraćeno OLAM. Na ovaj način moguće je kombinirati kvalitetu podataka unutar OLAP sustava i na njima provesti složenije analize koje OLAP sustav samostalno ne podržava. Han i Kimber (2006) navode nekoliko prednosti korištenja OLAM-a kao integracije OLAP-a i DM-a:

- Alati za rudarenje podataka zahtijevaju rad na integriranim, pročišćenim i konzistentnim podacima što u procesu DM-a obuhvaća relativno velik udio resursa. Čišćenje podataka jedna je od glavnih faza DM-a. OLAP sustavi temeljeni su na skladištima ili *datamartovima* u kojima se nalaze podaci iznimne kvalitete. Spajanje alata za rudarenje podataka na OLAP sustav automatski podrazumijeva pročišćene i kvalitetne podatke, pa se kombinacijom ova dva sustava troši znatno manje resursa potrebnih za njihovo provođenje.

- OLAP alati popularni su zbog svojih eksplorativnih karakteristika u analizama podatkovnih struktura. Interaktivni način pretraživanja podataka unutar OLAP-a temelji se na njegovim osnovnim analitičkim funkcijama: metodama raslojavanja, presijecanja, pivotiranja i bušenja. Analitička vrijednost postaje još veća kada se na te fleksibilne „poglede“ podataka primijene algoritmi rudarenja podataka. Dakle prednost korištenja analitičkih metoda OLAP i DM sustava korisniku pruža detaljnije rezultate.

- Rudarenje podataka se obično provodi na podacima koji nisu dovoljno ažurni. *Online* komponenta OLAP sustava pretpostavlja ažurne podatke, što je još jedna pozitivna strana kombinacije ova dva sustava.

2.2 Vizualizacija podataka

2.2.1 Pojam vizualizacije podataka

U modernom poslovnom okruženju informacija ima iznimnu važnost. Mogućnost i brzina saznavanja određene informacije direktno utječe na poslovne odluke. Višedimenzisjke podatkovne strukture mogu biti iznimno kompleksne za razumijevanje, a rezultati dobiveni analizama često generiraju više novih pitanja nego što daju odgovora. Podaci prikazani numerički u tablicama, ali i rezultati provedenih analiza egzaktni su, ali ih čovjek sporo (i teško) percipira. Kaže se da je čovjek vizualno biće, a u skladu s tim efikasan prikaz informacija je upravo u vazalnom obliku. Vizualizacija podataka i vizualna analiza podataka pomažu

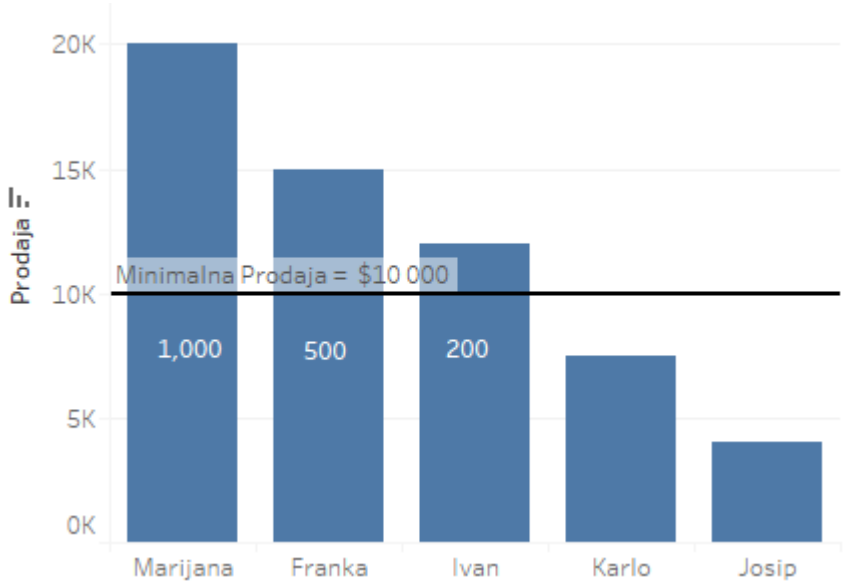
korisniku saznati esenciju nekog podatkovnog niza gotovo momentalno. Vizualizacija podataka ne predstavlja samo statički grafički izvještaj. Boehlen et al. (2008) naglašava važnost na interakciji korisnika sa vizualnim prikazom. Ključno je da korisnik može sam birati dimenzije koje će biti prikazane u grafičkom izvještaju, razinu njihove hijerarhije i osi na kojima su prikazane. Primjenom spomenutih OLAP tehnika korisnik dolazi do grafikona koji točno prikazuje dio *dataseta* koji ga zanima. Jedino kroz interaktivnost, korisnik može prikupiti ključne informacije i generirati znanja koja će na posljetku postati input u donošenju poslovnih odluka.

Tablica 6. Prodaja po zaposlenicima

Zaposlenik	Prodaja	Minimalna prodaja	Bonus
Ivan	$ 12,000.00	$ 10,000.00	$ 200.00
Marijana	$ 20,000.00	$ 10,000.00	$ 1,000.00
Karlo	$ 7,500.00	$ 10,000.00	
Franka	$ 15,000.00	$ 10,000.00	$ 500.00
Josip	$ 4,000.00	$ 10,000.00	

Izvor: izračun autora

U primjeru je prikazana jednostavna tablica koju je čovjeku jednostavno proučiti. Da je na primjer ova tablica bila sastavljena od nekoliko tisuća redaka i stupaca, njezina informacijska vrijednost bila bi znatno manja, odnosno, korisnik ne bi imao gotovo nikakve koristi iz njezinog promatranja. Iako je prikazana tablica pregledna, u sljedećem primjeru će isti podaci biti prikazani vizualno.



Grafikon 1. Podaci prikazani stupčastim grafikonom; Izvor: izračun autora, Tableau

Na grafikonu je prikazan *dataset* iz prijašnje tablice. Odmah je na prvu vidljivo koji zaposlenik je ostvario najvišu prodaju, koji najmanju, te se otprilike može procijeniti i omjer obujma prodaje svakog zaposlenika. Također je vidljivo koji zaposlenici su „probili" liniju minimalno propisane prodaje i za koliki iznos. Dakle, esencija informacija unutar tablice vidljiva je gotovo momentalno u grafičkom prikazu. Čak su i isplaćeni bonusi numerički su izraženi unutar stupaca zaposlenika koji su ga ostvarili.

Promatrani *dataset* sadrži čak 4 dimenzije: Zaposlenik; Prodaja; Minimalna prodaja i Bonus. Na 2D grafikonu uspješno su prikazane sve četiri dimenzije. Vizualna percepcija čovjeku je jača strana, ali ona ima svoja ograničenja. Few (2007) navodi kako čovjek u svom kratkoročnom pamćenju može zadržati samo četiri vizualne informacije. Silić i Šamanović (2011, p.179.) navode sljedeće: „*Numerički podatak u tablici bi predstavljao jednu jedinicu, dio informacije. Ukoliko bi bacili pogled na tablicu mjesečnih prihoda i zatim okrenuli glavu, bili bi sposobni replicirati novčani priljev za otprilike četiri mjeseca. No, 12 mjeseci prihoda prikazanih u formi linije na grafu predstavlja jedinstvenu sliku – jednu informaciju. Tako se odgovarajućim vizualnim kodiranjem podataka dramatično povećava količina informacija koja se može simultano držati u kratkoročnom pamćenju. Dakle, alati za vizualizaciju pomažu pri nadilaženju ograničenja kratkoročne memorije"*

U smislu prikaza podataka u poslovnom okruženju, može se reći da postoje dva ekstrema. Prvi je prikaz podataka kroz tekst, brojeve i tablice, a takav prikaz može biti zamoran i suhoparan. Drugi ekstrem je predizajniranost, prikaz podataka u raznim bojama i oblicima. Kao najbolje rješenje, potrebno je pronaći određenu zlatnu sredinu. Vrlo je važno da se osnova svake vizualizacije, a to je prenošenje informacija, ne izgubi zbog neadekvatnog dizajna.

2.2.2 Važnost vizualizacije u modernom poslovanju

Poslovne odluke svakodnevnica su svih poslovnih subjekata. Donositelji odluka na svim hijerarhijskim razinama kompanije odluke donose temeljem raspoloživih informacije. Kako bi odluka bila pravovremena, a ujedno i kvalitetna ključno je spoznati dostupne informacije u što kraćem vremenu i iz njih razviti određeno znanje. U ovom kontekstu, vizualizacija podataka predstavlja ključan element poslovne inteligencije. Thomsen (2002) prepoznaje tri osnovna načina kojim vizualizacija podataka pojednostavljuje spoznaju informacija i generiranje znanja. Prvi je povećavanje volumena podataka koje analitičar može upamtiti. Samim time

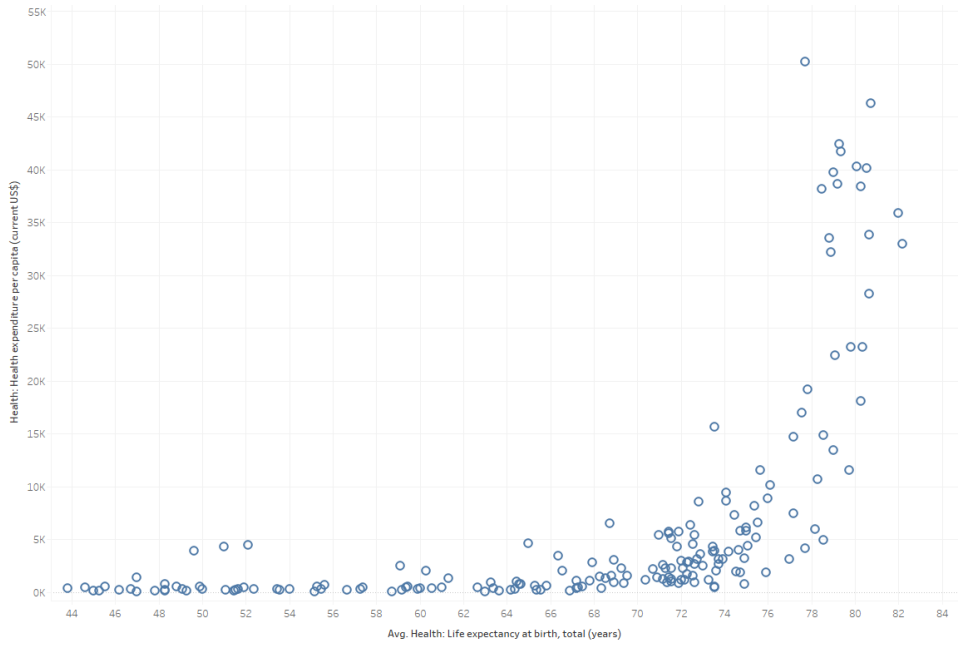
povećava se i opseg informacija za usporedbu. Vizualizacija pruža uvid u podatke iz više perspektiva što povećava njezinu informativnu vrijednost.

Složene procese unutar organizacija najlakše je pratiti vizualnim metodama. Tako su razvijene nadzorne ploče koje korisniku daju uvid u ključne trendove i informacije vezane uz proces. Nadzorne ploče ključne su u proizvodnim poduzećima gdje inženjeri imaju uvid u stanje resursa i repromaterijala. Vizualizacija podataka se koristi i u poslovnim simulacijama. Izgradnja i uporaba simulacijskih modela u pravilu zahtjeva rad tima stručnjaka, kako onih kojima je promatrani sustav predmet interesa tako i stručnjaka informatičara. Simulacijski model realizira se u obliku kompjuterskog programa koji oponaša ponašanje realnog sustava. Taj je program obično napravljen dovoljno općenito i jednostavno kako bi se pomoću njega mogle simulirati različite situacije koje jesu ili bi mogle nastupiti u stvarnosti. Takve situacije simuliraju se pomoću promjenjivih ulaznih podataka koji su analogni uvjetima u koje može biti doveden realni sustav. (Čerić, 2003.) Simulacijski modeli u svom outputu pružaju vizualne prikaze koji se mijenjaju kroz protok vremena. Na ovaj način korisnik može pratiti utjecaje više varijabli na promatrani sustav u vremenskoj perspektivi. Poslovne simulacije viši su oblik predviđanja poslovnih odluka te su ključne za donošenje strateških odluka.
Vizualizacija podataka ključna je u financijskim izvještajima. Kretanje vrijednosti novca ili vrijednosnica najlakše se razumijeva vremenskim nizovima, odnosno njihovim vizualnim trendovima. Na ovaj način moguće je uvidjeti sezonske utjecaje na neki podatkovni niz, uvidjeti obrasce kretanja nekog podatkovnog niza itd. Vremenski nizovi prikazani trendovima koriste se i u predviđanjima. Neki alati za vizualizaciju imaju opcije automatske procjene kretanja trenda, dok neki iziskuju napredno statističko znanje složene preditkivne modele. Redovni izvještaji svih odjela nekog poduzeća koriste vizualizaciju podataka kako bi se što bolje razumjelo trenutno (*AS IS*) stanje, ali kako bi se i predvidjelo buduće (*TO BE*) stanje. U poslovnom svijetu postoje brojni primjeri korištenja vizualizacije podataka, od kojih su trenutno navedeni samo neki. U nastavku rada kroz primjere grafikona biti će navedeno još slučajeva korištenja ove metode prikaza podataka.

2.2.3 Najčešće vrste grafikona i njihovo korištenje u praksi

2.2.3.1 Točkasti grafikon (*Scatter plot*)

Kada se grafički žele prikazati opažanja smještena unutar dvije dimenzije koristi se točkasti grafikon u kojemu svako opažanje predstavlja jednu točku (mjere jedne, odnosno druge dimenzije). Određuju se zavisna i nezavisna varijabla koje se promatraju na grafikonu. Varijable se još nazivaju obilježjima, svojstvima i dimenzijama. Zavisna varijabla je promjenljiva varijabla kojoj se pokušava utvrditi zavisnost s obzirom na nezavisnu varijablu, ali i intenzitet te zavisnosti. Nezavisna varijabla je kontrolna varijabla u modelu, koja se mijenja pod vanjskim utjecajima i jedan je od efekata promjene nezavisne varijable. U realnom svijetu teško je naći varijable koje su 100% zavisne, odnosno nezavisne pa se ovaj model koristi u znanstvenim „laboratorijskim uvjetima“.



Grafikon 2. Prikazuje odnos izvajanja za zdravstvo p.c. (nezavisna varijabla) i očekivane životne (zavisna varijabla); Izvor: izračun autora

Točkasti grafikoni koriste se i prikazu korelacijskih odnosa dviju varijabli. Korelacija je međusobna povezanost dviju varijabli. Povezanost dviju varijabli određuje se korelacijskim koeficijentima od koji određuju intenzitet njihovog međuodnosa. Pearsonov koeficijent korelacije najčešće se koristi u modelima linearne korelacije a kreće se u intervalu $-1 < 0 < 1$. Kod točkastog grafikona fokus je na cjelokupnoj slici odnosa dviju varijabli, a ne na pojedinu vrijednost opažanja. Kada je na grafikonu prikazan velik broj opažanja, teško je pronaći traženo

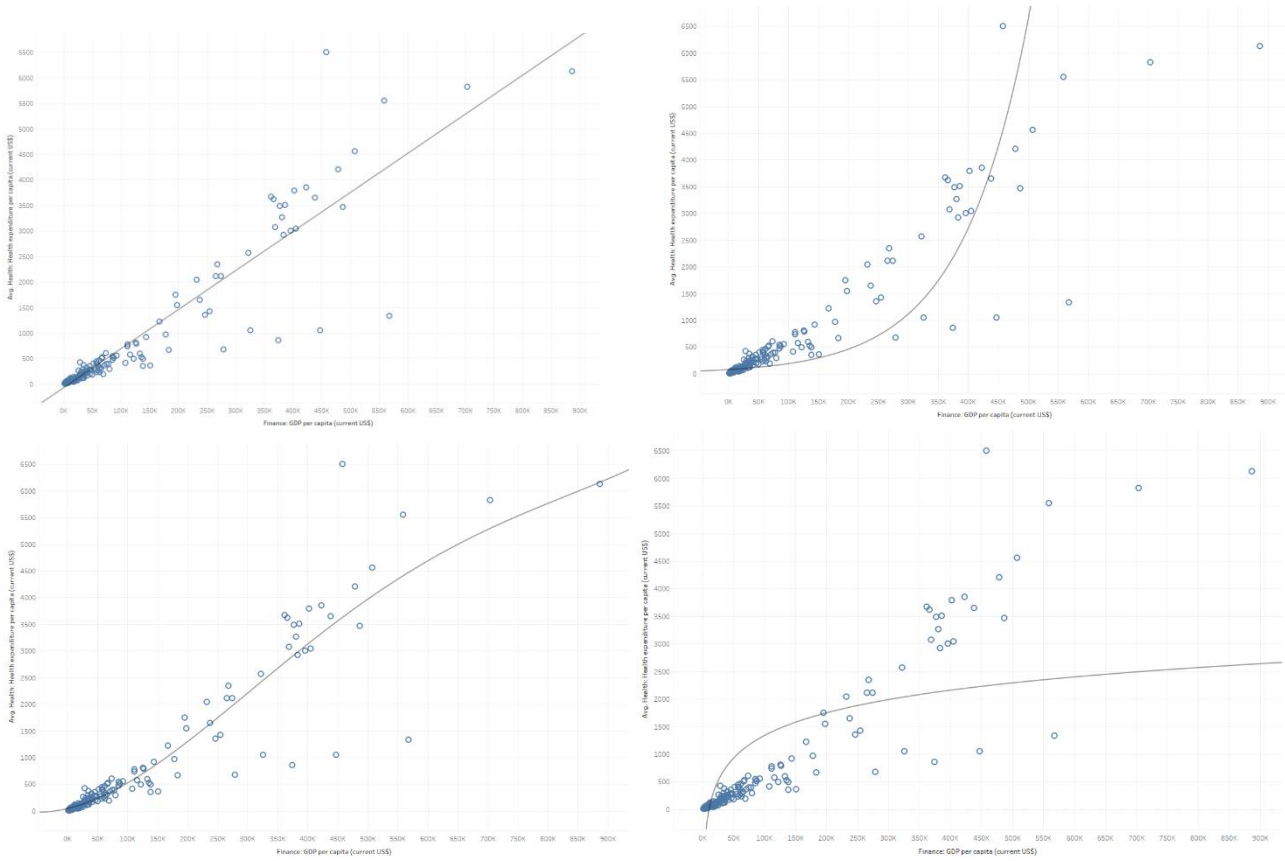
opažanje u „šumi“ ostalih. Legenda se često izostavlja upravo zbog kompleksnosti unikatnog prikaza svakog opažanja. Obično se opažanja podjele u grupe, odnosno klastere kojima se dodaju unikatne boje ili oblici točaka što grafikonu može povećati informacijsku vrijednost. Sephen Few (2012) predlaže korištenje dvije skupine (klastera) opažanja zbog jednostavnosti prikaza.



Grafikon 3. Prikazuje odnos izdvajanja za zdravstvo i BDP-a po glavi stanovnika

Izvor: izračun grafikona, Tableau

Grafikon je dodavanjem boja za svaku regiju dobio dodatnu informacijsku vrijednost, te je moguće uvidjeti da su zemlje iz regije Europa udaljene od ishodišta, dok su Afričke zemlje bliže ishodištu. Na *scatterplot* se često nadoda i linija trenda, odnosno *best-fit* linija. Ova linija prikazuje trend kojim se varijable mijenjaju, a može poprimiti bilo koji linearni ili nelinearni oblik

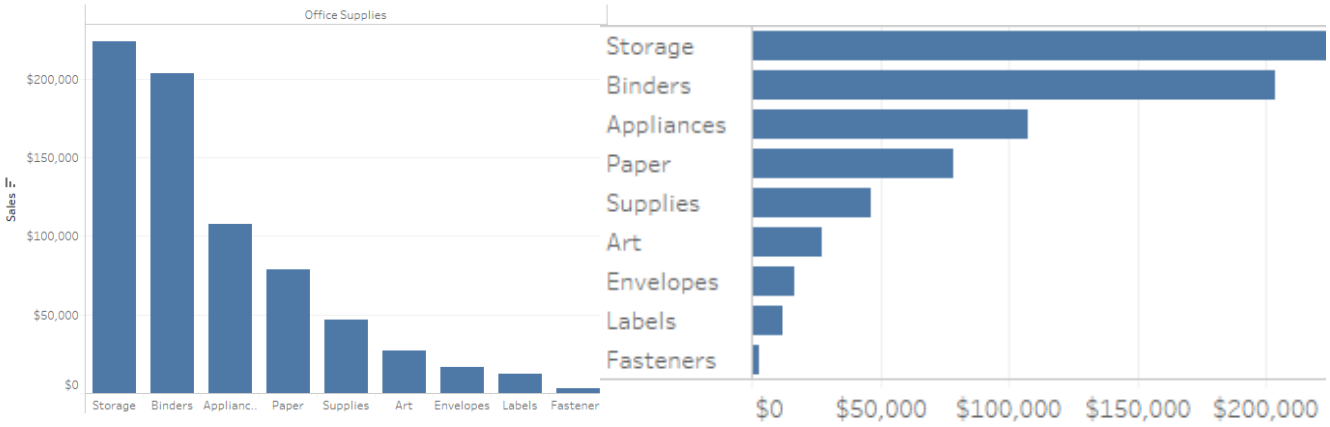


Grafikon 4. Linearni, eksponencijalni, polinomski i logaritamski trend na točkastom dijagramu; Izvor: izračun autora, Tableau

Analitičar sam bira koji trend najviše odgovara podacima. Na prethodnom grafikonu prikazane su četiri moguća trenda od kojih podacima najviše odgovaraju linearni i polinomski. Daljnjim testiranjem, analitičar može doći do odgovora koji trend najbolje pripada podacima (engleski naziv ove metode je *best-fit line)*. Moderni alati za vizualizaciju podataka imaju opciju automatskog generiranja trendova, što ne zahtijeva statističko znanje, ali utječe na točnost modela.

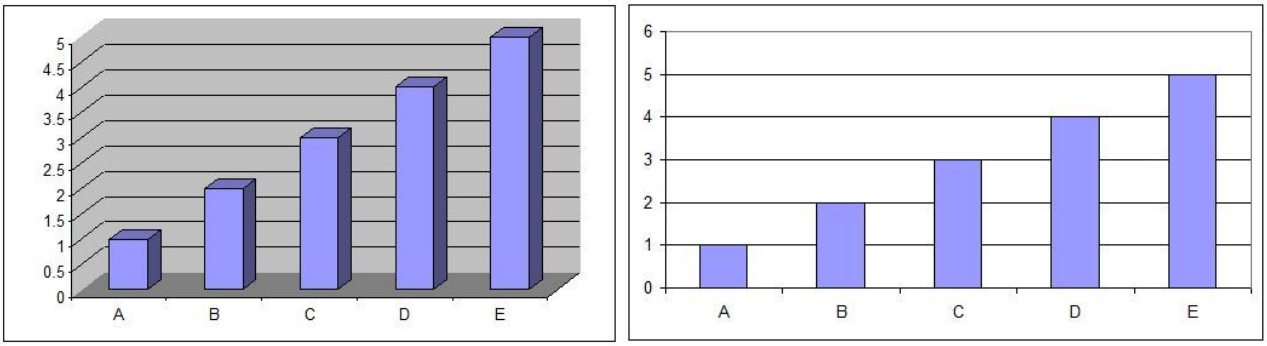
2.2.3.2 Stupčasti grafikon (*Bar chart*)

Kod usporedbe vrijednosti diskretnih podataka ili prikaz trendova kroz vrijeme koriste se stupčasti grafikoni. Stupčasti grafikon prikazuje podatke u obliku pravokutnih stupaca, kao što mu i ime nalaže, koji su postavljeni vertikalno ili horizontalno u odnosu na x os. Riječ je dakle o 2D grafikonima u kojima se na jednoj osi prikazuje vrijednost varijabli (mjera) a na drugoj osi dimenzija koja je hijerarhijski razložena na odgovarajuće pod-dimenzije.



Grafikon 5. Prodaja uredskog materijala po proizvodu (vertikalni i horizontalni); Izvor: izračun autora, Tableau.

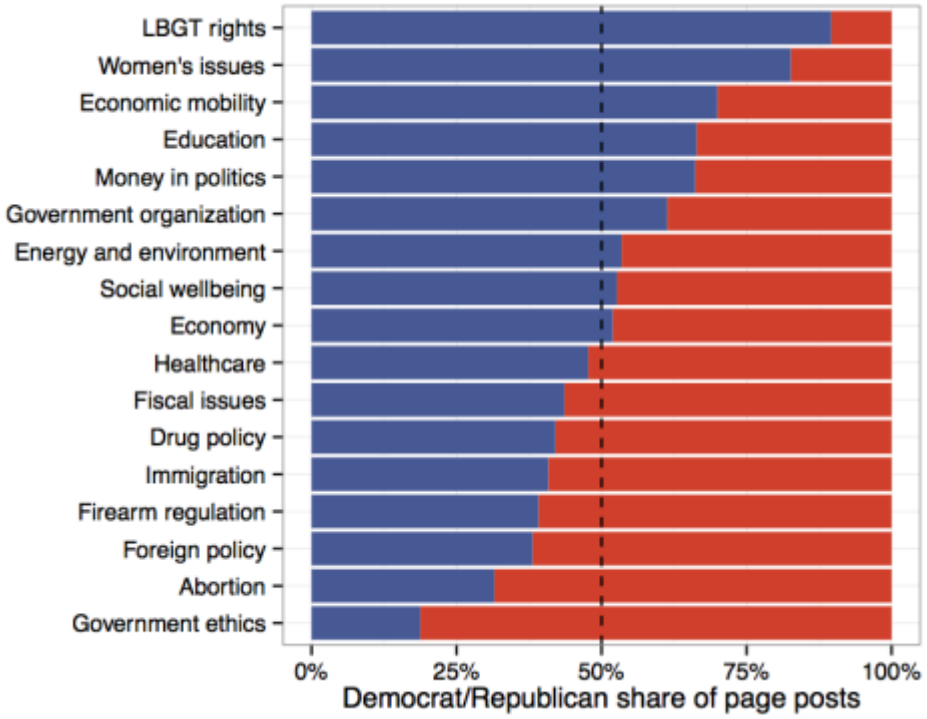
Kod dizajniranja stupčastih grafikona korisnici mogu napraviti greške koje će otežati očitavanje grafikona. 3D grafikoni u kojima treća dimenzija ne predstavlja nikakav odnos prema podacima, već je isključivo u svrsi estetike trebali bi se itekako izbjegavati. Osnova grafičkog prikaza je funkcionalnost, a ne estetika. Takva praksa obično je prisutna u medijima koji su okrenuti široj javnosti (portali, dnevnici, novine), što stvara krivu sliku o funkcionalnosti grafičkih prikaza. (Few, 2012) U sljedećem primjeru prikazani su stupčasti grafikoni na istim podacima. Jedan grafikon prikazan je u dvije, a drugi u tri dimenzije.



Grafikon 6. 3D i 2D stupčasti grafikoni; izvor: Bižaca, Dobrović, 2016

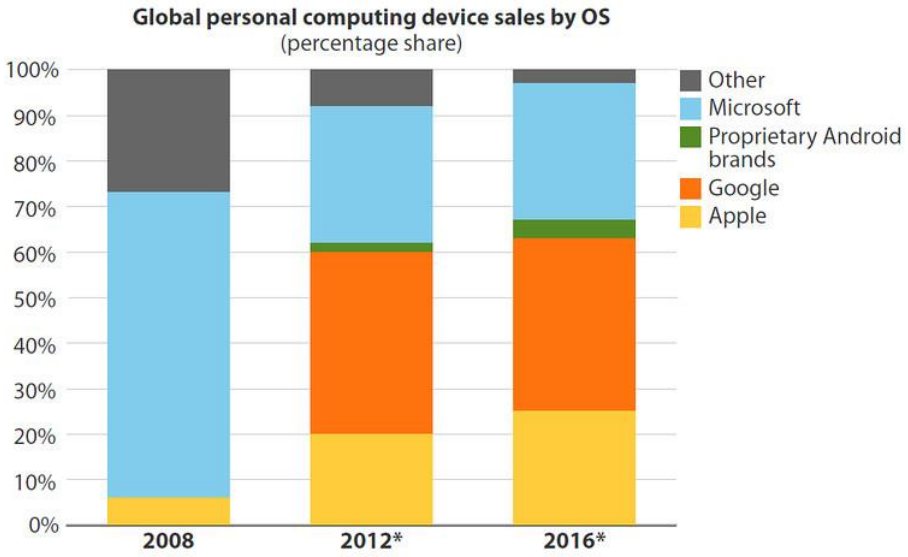
Lijevi grafikon primjer je prakse koja bi se trebala izbjegavati. Prikaz na desnoj slici mnogo je jasniji i razumljiviji. Iako su 3D grafikoni prisutni u gotovo svim bibliotekama BI alata za vizualizaciju, oni bi se u ozbiljnim izvještajima trebali izbjegavati. (Few, 2012)

Stupčasti grafikoni također mogu poprimiti i razdijeljeni oblik. U razdijeljenom obliku moguće je prikazati dvije ili više pod dimenzije u jednom stupcu. Na taj način moguće je uvidjeti omjer tih pod-dimenzija unutar cjelokupne dimenzije koju stupac predstavlja. *Stacked bar chart* ili razdijeljeni stupčasti grafikon ima svoje dobre i loše prakse korištenja.



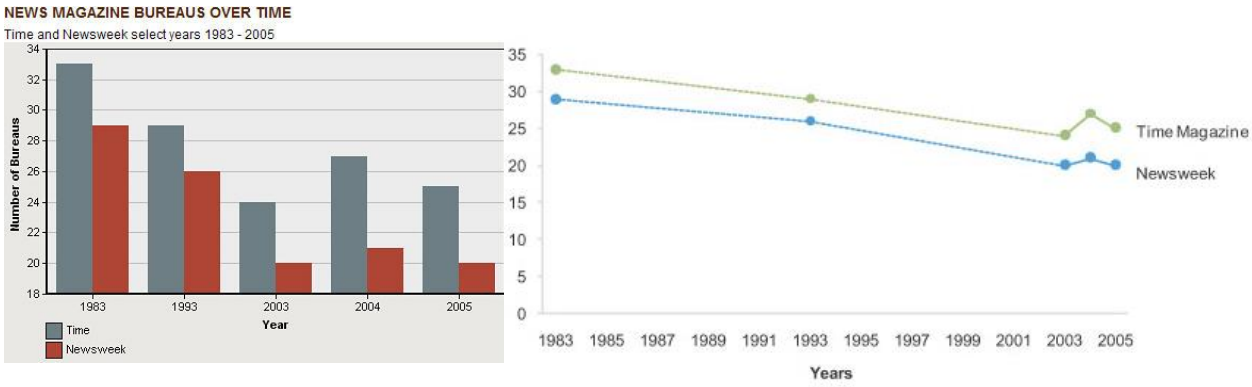
Grafikon 7. Socijalna pitanja u demokratskim odnosno republikanskim objvama na društevnim medijama u SAD-u; Izvor: Bižaca, Dobrović, 2016

Kada postoje samo dvije dimenzije koje se opažaju (Republikanci i Demokrati), tada ovakav prikaz dolazi do izražaja, te je mnogo vrjedniji od običnog stupčanog grafikona. Međutim, ako se broj dimenzija povećava, ovakvi raslojeni prikaz samo dodatno zbunjuje korisnika te mu ne pruža dodatnu informacijsku vrijednost. Vrlo je važno da se kod dizajna grafikona počiva od načela minimalizma, jer se ključno da osoba koja promatra graf može brzo i jasno spoznati esenciju, odnosno „priču“ koju grafikon predstavlja.



Grafikon 8. primjer slojevitog stupčastog grafikona; Izvor: https://www.wsj.com/

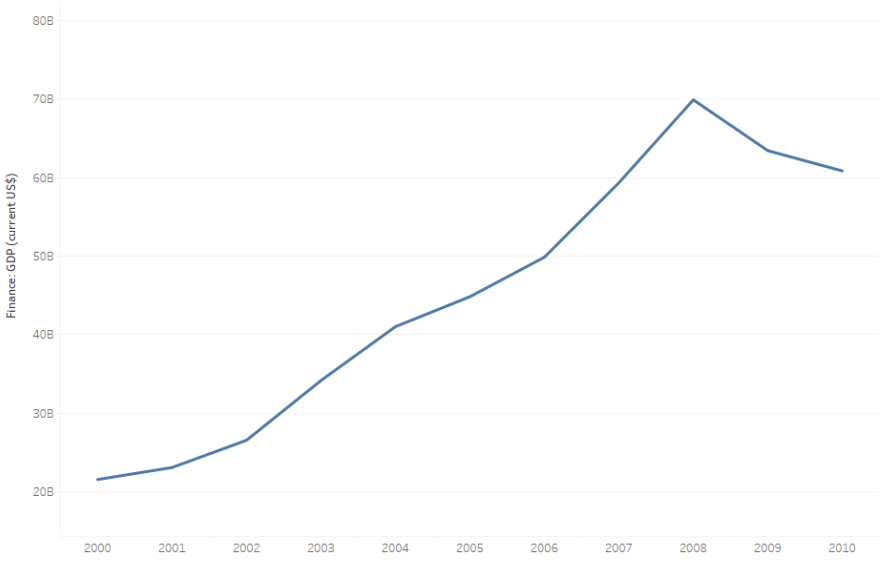
Ovaj primjer, naprotiv, prikazuje lošu praksu korištenja razdijeljenog stupčastog grafikona. Wall Street Journal, jedan od vodećih poslovnih dnevnika, prikazuje odnos tržišnih udjela operacijskih sustava korištenih kroz 3 promatrane godine. Razlozi zbog čega je ovo loša praksa su sljedeći: na ovom grafikonu moguće je iščitati kako se svijetlo plava boja koja predstavlja Microsoft s vremenom smanjuje, dok dvije nijanse narančaste – Google i Apple rastu. U prikazu je nemoguće zaključiti da li je Apple najlošiji, budući da je na dnu, ili poredak u stupcu nema neko značenje već samo određuje udio. Dvosmislena interpretacija ovakvog grafikona nalaže da ga se treba dodatno objasniti – pri čemu se gubi smisao grafikona (Makeovers, 2012.) Česta loša praksa je upotreba stupčanih grafikona kada je jedna varijabla na grafikonu vrijeme. U tom slučaju, puno je bolje koristiti linijski grafikon te linijom prikazivati kretanje vrijednosti u odnosu na vremenski trenutak. (Few, 2012) Na slici 12 je prikazan broj ureda časopisa Time i Newsweek na oba spomenuta načina.



Grafikon 9. Usporedba stupčastog i linijskog grafikona; Izvor: Bižaca, Dobrović, 2016

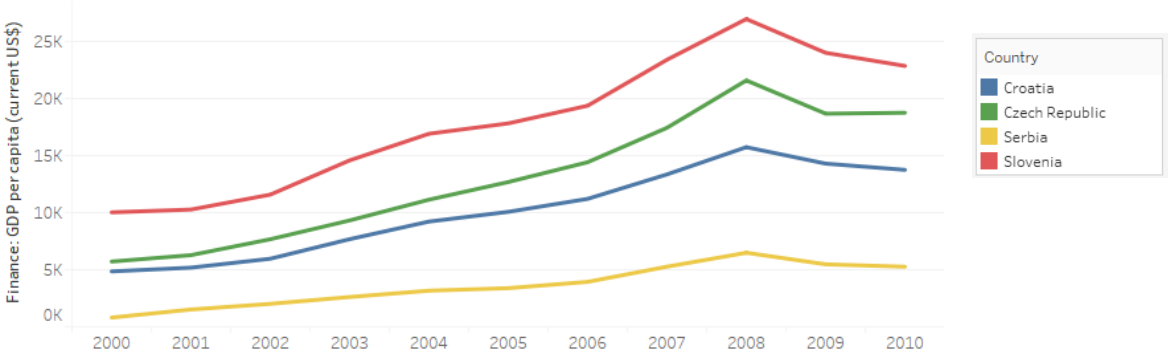
2.2.3.3 Linijski grafikon

Linijski grafikon prikazuje se u prikazivanju vremenskih serija na nekom podatkovnom nizu. Linijski grafikoni mogu se koristiti i za usporedbu više vremenskih serija na istom mjerilu. Linije se uglavnom koriste kao poveznica točaka na grafikonu koje prikazuju vrijednost dimenzije u određenom vremenskom trenutku.



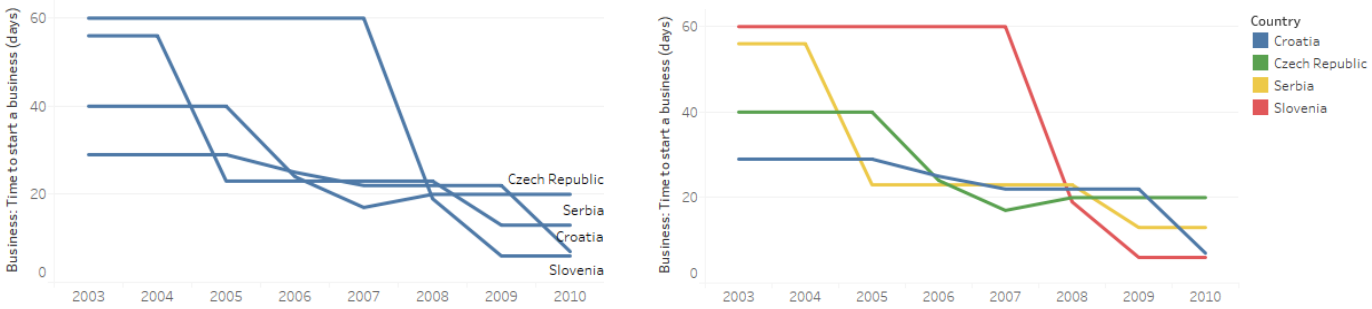
Grafikon 10. BDP p.c. Republike Hrvatske (2000 – 2010). Izvor: izračun autora; Tableau

Linijski grafikon sličan je točkastom grafikonu (*Scatter Plot*-u) . Prikazuje se odnos odabrane varijable (dimenzije) i vremena (koje predstavlja drugu dimenziju), samo što su u ovom slučaju točke opažanja povezane linijama kako bi se što uvjerljivije prikazali vremenski trendovi.
Linijski grafikon koristi se i u usporedbi više varijabli koje se kreću unutar nekog vremenskog intervala. Kada se dvije ili više linija nalazi na istom grafikonu najvažnije ih je prikazati na način da se raspoznaje njihova razlika. Tako je, na primjer, na prethodni grafikon moguće nadodati BDP *per capita*. još nekoliko država.



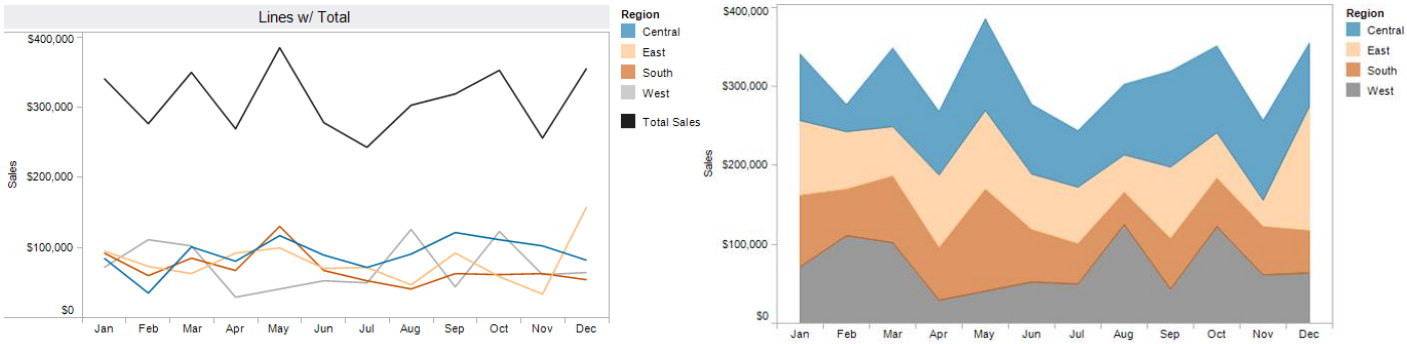
Grafikon 11. Linijski grafikon sa više prikazanih linija; Izvor: izračun autora; Tableau

Kada je na grafikonu prikazano više linija, važno je da se one jasno razlikuju kako bi odmah bilo jasno o kojim opažanjima je riječ (u ovom slučaju riječ je o opažanjima iz dimenzije „*Country*“). Pridruživanje različitih boja linijama najbolji su način njihove distinkcije. Ostale metode distinkcije krivulja mogu zbuniti korisnika, pa Few (2012) preporuča korištenje različitih pastelnih boja.



Grafikon 12. Razlika u prikazu distinkcije različitih linija na istom grafikonu; izvor: izračun autora, Tableau

Jednostavnost je ključ kod izrade linijskih grafikona. Kao i kod ostalih grafikona, cilj je da korisnik odmah razumije trendove i odnose podataka prikazanih na grafikonu. Jedna od tipičnih grešaka je korištenje površinskog (eng *area*) umjesto linijskog grafikona. Površinski grafikon sličan je linijskom, ali odnos i vrijednost varijabli nije prikazan kao točka na Y osi, već kao površina, odnosno kao visina lika koji predstavlja podatke za određenu dimenziju. Ovakva praksa uglavnom zbunjuje korisnika. Površinski grafikon se koristi kada se na istom grafikonu želi prikazati ukupna vrijednost (mjera) promatrane dimenzije (npr. Prodaja) ali i pod-dimenzije koje tvore tu ukupnu vrijednost (npr. Prodaja po regijama). Sljedeći primjer prikazuje razliku između linijskog i površinskog grafikona.

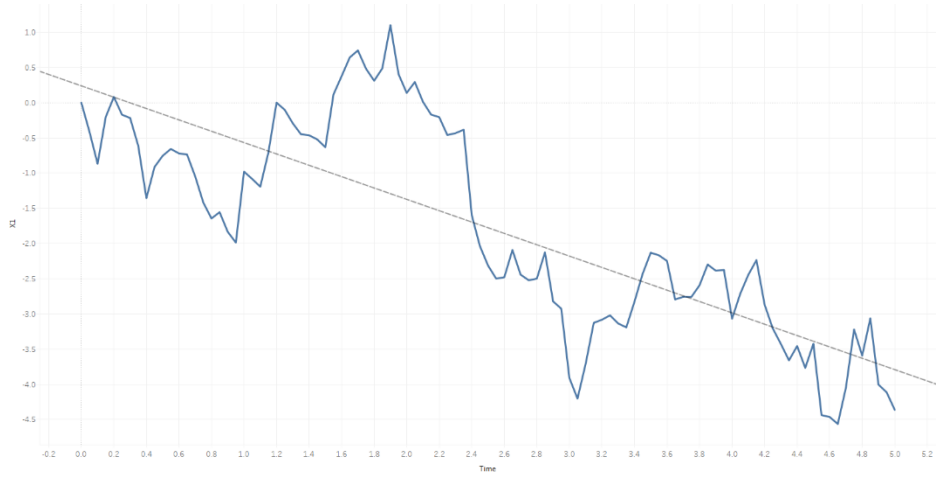


Grafikon 13. Linijski i Površinski grafikon koji prikazuju iste podatke.

Izvor: http://www.vizwiz.com/2012/10/stacked-area-chart-vs-line-chart-great.html

Površinski grafikon odličan je u prikazivanju udjela pojedinih pod-dimenzija unutar određene dimenzije. Međutim, kod prikaza vrijednosti pojedine pod-dimenzije zbunjuje korisnika. Na pitanje „Kolika je bila prodaja u Ožujku za regiju Istok?“ teško je dati odgovor oslanjajući se isključivo na površinski grafikon. Naizgled se može zaključiti da je prodaja bila cca. 200 000$, što je krivo jer se, kao što je već navedeno, vrijednost gleda kao visina površinskog lika u određenoj vremenskoj točki. Na linijskom grafikonu teško je „na prvu“ spoznati udjele pod-dimenzija u Total-u, ali je zato vrlo jednostavno pročitati koja je njihova vrijednost u određenom vremenskom razdoblju. Tako se poprilično brzo može zaključiti da je prodaja za regiju istok u Ožujku iznosila nešto ispod 100 000$.

Ponekad je teško spoznati trendove kretanja niza podataka samo iz promatranja linijskog grafikona. Priroda podataka, pogotovo u financijskom sektoru, može poprimiti tzv.“*random walk*“ oblik, pa se uz prikaz linijskog grafikona prikazuje i linija trenda koja upotpunjuje njegovu informacijsku vrijednost.



Grafikon 14. Linija trenda prikazana na linijskom grafikonu; Izvor: izračun autora, Tableau

2.2.3.4 Ostali grafikoni

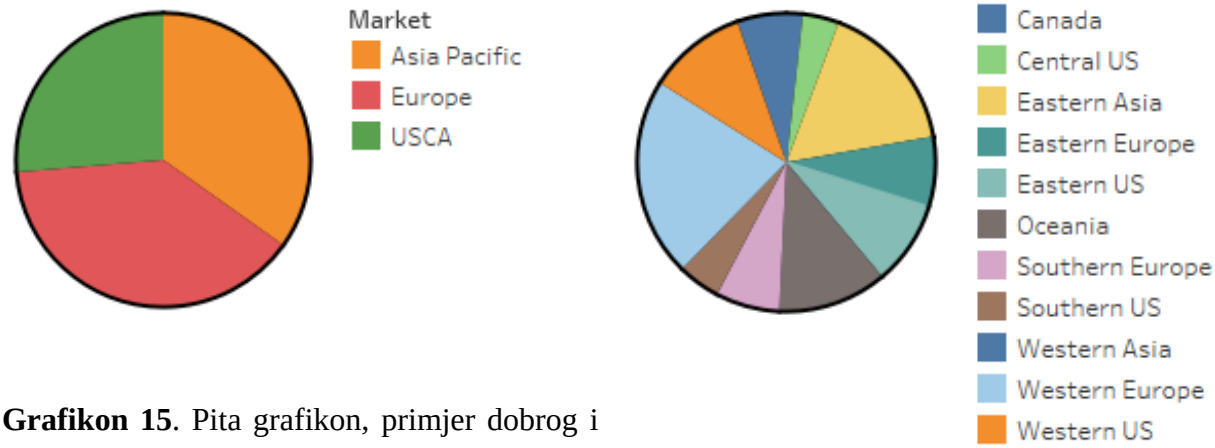
U praksi modernog poslovanja, uz navedene osnovne tipove grafikona, koriste se i mnoge druge varijante, od kojih će u ovom poglavlju biti navedeni:

- Pita grafikon (Pie chart)
- Histogram
- Treemap

Ovi grafikoni smješteni su u isto poglavlje zbog svojeg ograničenog korištenja u poslovnom kontekstu, ali i zbog svoje informacijske vrijednosti.

Pita grafikon (Pie chart)

Pie chart, odnosno pita grafikon, iako omiljen među korisnicma, često je loš primjer prikazivanja podataka. Pita (*pie*) je tip grafikona koji je najpogodniji za jednostavne grafikone na kojima prikazujemo udio svake od vrijednosti u odnosu na sumu. Struka preporučuje korištenje pita grafikona isključivo ako se radi o dva, maksimalno tri promatrana opažanja. Ako se u grafikon uvrsti više opažanja, korisniku je teško percipirati omjere pojedinih opažanja u nekoj cjelini. (Few, 2012)

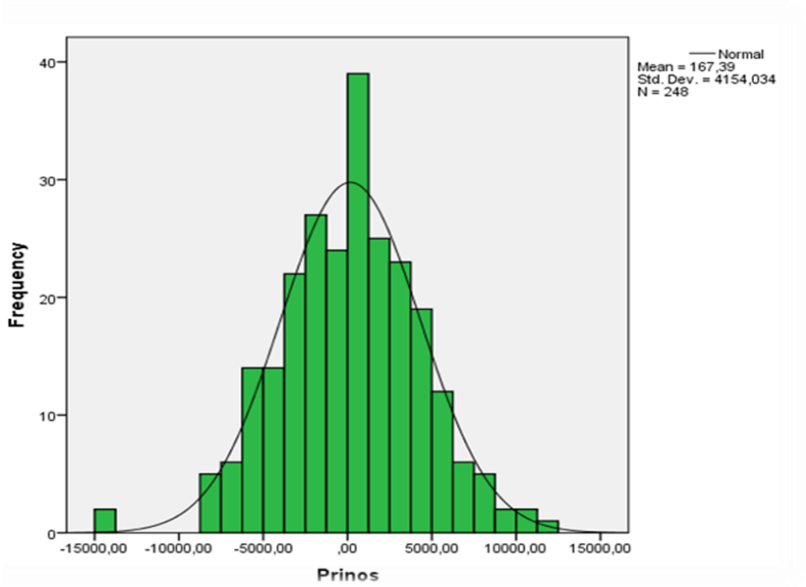


Grafikon 15. Pita grafikon, primjer dobrog i lošeg dizajna; Izvor: izračun autora; Tabelau

Lijevi pita grafikon ima samo tri isječka pa on predstavlja primjer dobrog dizajna. Promatrač brzo i efikasno može razumjeti udjele pojedinog tržišta u ukupnoj prodaji. Desni grafikon primjer je lošeg dizajna istog grafikona. S obzirom da je prikazano čak 11 isječaka, promatrač ima poteškoće u spoznaji udjela navedenih regija unutar ukupne prodaje. Stupčasti grafikon bi u ovom slučaju jednostavnije prikazao vrijednosti prodaje po regijama.

Histogram

Histogram je grafički prikaz distribucije podataka. potražite ovdje. Histogram je podvrsta stupčastog grafikona na kojem se prikazuju podaci o učestalosti pojavljivanja. Na X osi prikazani su svi podaci koji se nalaze u nekom podatkovnom nizu. Na Y osi prikazane je njihova učestalost ponavljanja, odnosno njihova frekvencija. Ukoliko je vrijednost varijable previše raspršena (ima previše različitih vrijednosti), tada se podaci grupiraju u skupine kako bi ih se lakše prikazalo na histogramu.

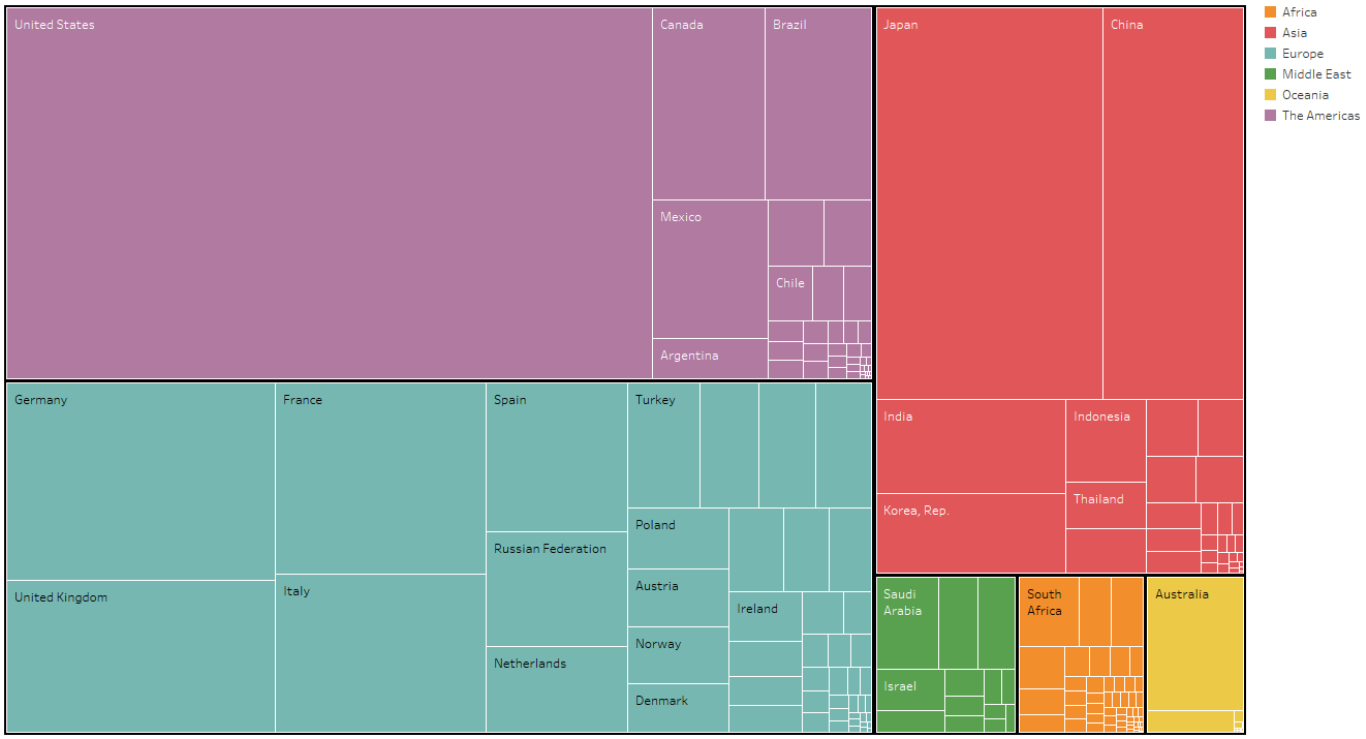


Grafikon 16. Histogram prinosa na CROBEX portfelj za 2013. godinu; Izvor: Bižaca, Erceg, 2015

Na histogram se često dodaje i Gaussova krivulja, koja je prikazana na prethodnom grafikonu. Analizom histograma i odgovarajuće joj Gaussove krivulje, moguće je zaključiti radi li se o normalnoj distribuciji podataka, ili je pak riječ o nekoj od nepravilnih distribucija. Ova krivulja prema Gaussu može poprimiti normalni, šiljasti, tupi, ljevostrani i desnostrani oblik.

Treemap

Treemap ili hijerarhijski ugniježđen grafikona nudi hijerarhijski prikaz podataka i olakšava prepoznavanje uzoraka, npr. koje se stavke u trgovini najbolje prodaju, ili koja država ima najveći BDP s obzirom na BDP čitave Zemaljske kugle, a ujedno podijeljene na kontintete. Hijerarhijski najviše grane (npr. kontineti) prikazuju se kao jednobojni pravokutnici, a svaka pod-grana kao manji pravokutnik (npr. država). Hijerarhijski ugniježđen grafikon prikazuje kategorije po bojama i blizini i može jednostavno prikazati mnogo podataka, što bi bilo teško na drugim vrstama grafikona. (Microsoft, 2016.)



Grafikon 17. *Treemap* koji prikazuje BDP zemalja prikazanim po regijama; izvor: izračun autora, Tableau

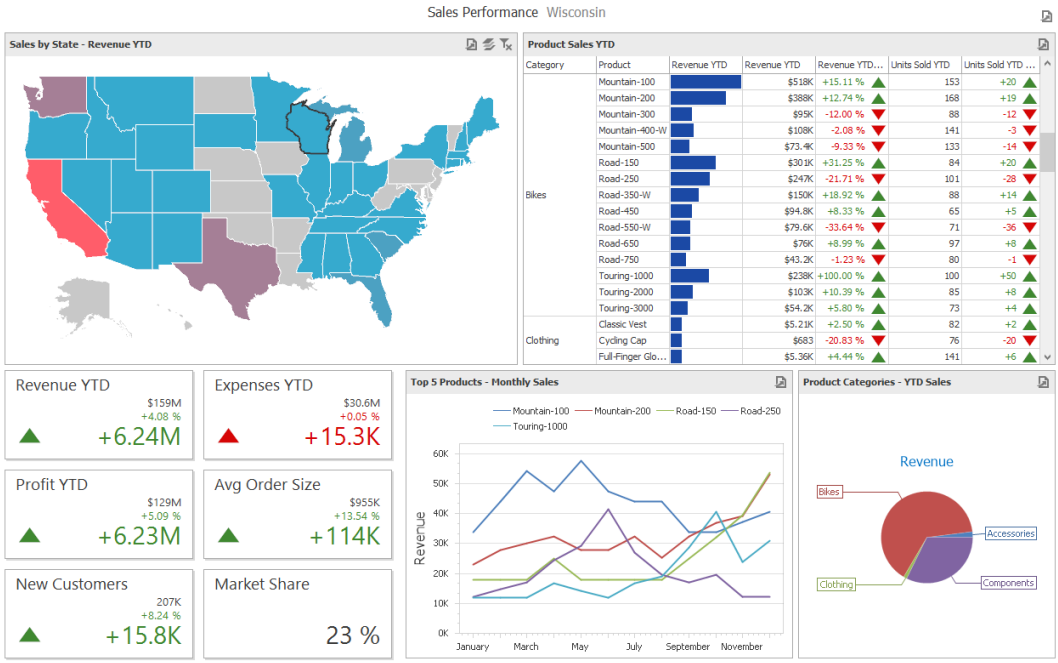
Na primjeru su prikazani podaci koji odlično prikazuju odnos ukupnog BDP-a prema državama ali i prema regijama što bi bilo iznimno teško prikazati na nekom grafikonu druge vrste.

2.2.4 Nadzorne ploče

2.2.4.1 Uvod u nadzorne ploče

Za prikaz podataka u poslovanju često se izrađuju nadzorne ploče (eng. *dashboards*). Njihov izgled inspiriran je takvim pločama koje se nalaze primjerice u automobilima ili avionima. *Dashboard* služi kao vizualni prikaz najvažnijih informacija potrebnih za postizanje jednog ili više ciljeva, konsolidiranih i uređenih na jednom zaslonu tako da se sve informacije mogu pratiti na prvi pogled (Few, 2004). Isto kao što nadzorna ploča automobila pruža ključne informacije potrebne za upravljanje vozilom na prvi pogled, BI nadzorna ploča ima vrlo sličnu svrhu, brz obzira koristi li se za donošenje strateških odluka za veliku korporaciju, izvedbu dnevne operativnih aktivnosti nekog tima, ili za izvršavanje zadataka koji uključuju jednog korisnika (Few, 2006).

Informacije na nadzornoj ploči su prvenstveno prikazane vizualno, a uobičajeno je to kombinacija teksta i grafike, ali s naglaskom na grafičke elemente. Grafički elementi se ne nalaze na ploči radi eventualne ljepote prikaza podataka, već imaju svoju bitnu ulogu u informiranju. Kvalitetno prikazani grafikoni često su efikasniji u prenošenju informacija i imaju bogatije značenje nego goli tekstualni prikaz. Primjer jedne nadzorne ploče koja služi nadgledanju rezultata prodaje prikazan je na slici.



Slika 11. Primjer dashboarda za informacije o prodaji, izvor: https://www.devexpress.com/

2.2.4.2 Prakse u izradi nadzornih ploča

Vizualni prikaz informacija omogućuje korisnicima da svojim očima brzo upiju informaciju, a ljudski mozak je vrlo sposoban u izvlačenju točnog i najvažnijeg zaključka iz takvog prikaza. Za što kvalitetnije dizajniranje nadzorne ploče, potrebno je poznavati karakteristike ljudske vizualne percepcije te znati koje elemente koristiti, a koje izbjegavati.

Nadzorne ploče prikazuju informacije potrebne za ispunjenje određenih ciljeva, a to često zahtijeva pristup informacijama koje nisu direktno povezane jer primjerice dolaze iz različitih dijelova organizacije. Nadzorna ploča može prikazivati informacije koje su potrebne menadžmentu ili bilo kojem drugom zaposleniku u ostvarenju njegovih ciljeva. U skladu s tim, nadzorna ploča može sadržavati i *KPI*-eve (eng. *Key Performance Indicators*), odnosno ključne pokazatelje performansi koji su potrebni za dostizanje cilja.

Nadzorna ploča je obično dizajnirana tako da stane na jedan zaslon računala. Poželjno je da sve informacije stanu na jedan zaslon, kako bi u potpunosti bile dostupne korisniku tako da ih može vidjeti odjednom, na prvi pogled. Ako postoji potreba pomicanja stranice za vidjeti sve informacije na zaslonu, poništava se osnovna namjena ploče s instrumentima. Ako je potrebno gledati nadzornu ploču na više zaslona, tada se smatra da se radi o više ploča. Cilj je da su najvažnije informacije lako i bez napora na raspolaganju korisniku, tako da isti može brzo shvatiti ono što mu je potrebno.

Informacije u nadzornim pločama se mogu prikazivati u zasebnoj aplikaciji ali i u internetskom pregledniku (*browseru*). To je odlično rješenje ako se podacima želi pristupati na jednostavan način s udaljene lokacije. Ovisno o potrebama i ciljevima organizacije, određuje se potreba osvježavanja informacija u realnom vremenu. Primjerice, ako se radi o nadzor zračnog prometa pomoću nadzorne ploče, korisnik morate odmah biti obaviješten kada nešto nije u redu. S druge strane, ako korisnik donosi strateške odluke o tome kako povećati prodaju, bit će mu vjerojatno dovoljna snimka informacija od nekog prijašnjeg trenutka.

Unatoč tome da se informacija o gotovo svemu može na odgovarajući način prikazati na ploči, postoji barem jedna karakteristika koja opisuje gotovo sve informacije u različitim pločama, a to je činjenica da su prikazane informacije sažetak ili iznimka. Razlog je jednostavan, korisnici ne mogu pratiti na prvi pogled sve detalje potrebne za postizanje ciljeva. Nadzorna ploča mora brzo ukazati da se dogodilo nešto što zahtijeva pažnju korisnika i može zahtijevati njegovo djelovanje.

Ne moraju biti prikazani svi detalji potrebni da se poduzmu određene mjere, ali korisnik mora imati mogućnost što jednostavnije doći do informacija. To može uključivati prebacivanje na

drugi zaslon izvan ploče s instrumentima te korištenje funkcionalnosti OLAP sustava, primjerice *drill-down, roll-up,* ili *slice and dice* metode. Primarna namjena nadzorne ploče je da ukaže na potrebu za aktivnošću na prvi pogled. Poželjno je i da ima mogućnost pronalaženja informacija direktno iz svog sučelja.

Nadzorne ploče, dakle, zbirka su *widgeta* koji korisniku pružaju pregled izvješća i mjernih podataka koji su vam najvažniji. Nadzorne ploče omogućuju praćenje brojnih mjernih podataka odjednom, tako da je moguće brzo provjeriti stanje svojih računa ili vidjeti korelacije između različitih izvješća. Nadzorne ploče jednostavno je izraditi, prilagoditi i dijeliti. *Widgete* određuje sam korisnik, što je iznimno važno budući da preferencije pojedinih prikaza ovise o korisniku i industriji. Izrada *dashboarda*, odnosno nadzornih ploča svojevrsni je trend u BI industriji. Svi veći proizvođači BI softvera predstavljaju svoje proizvode za vizualizaciju upravo na primjeru lijepo izrađenih nadzornih ploča.

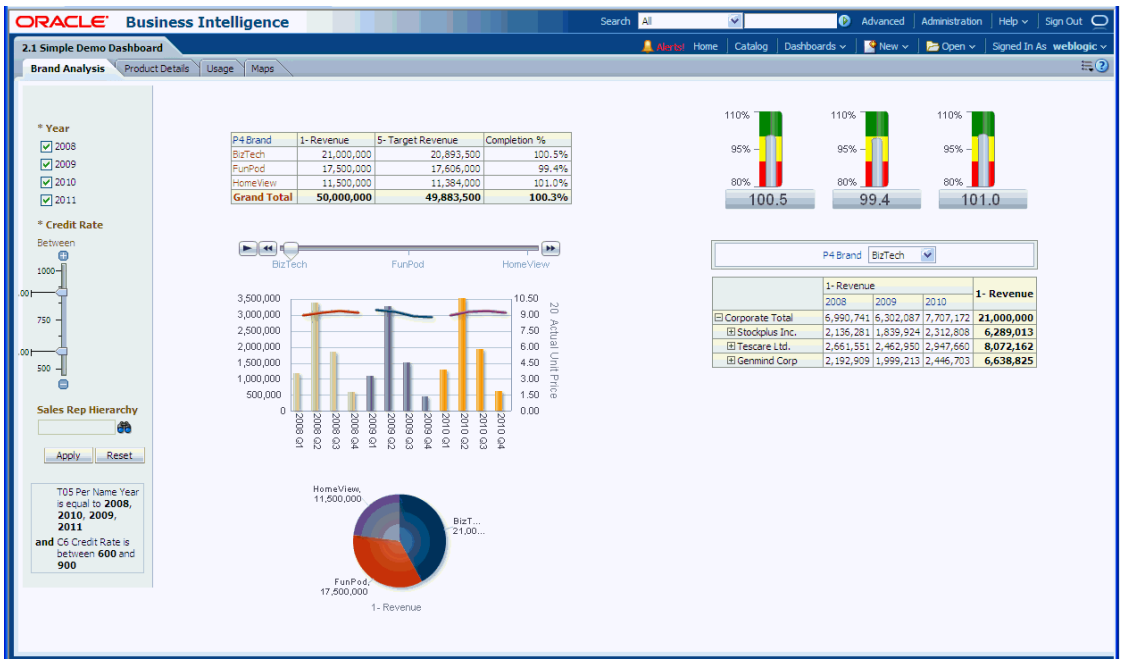


Slika 12. Nadzorna ploča SAP Business Objects softvera;
izvor: https://www.sap.com/products/bi-platform.html

Nadzorne ploče, međutim, nisu univerzalno definirane, pa tako određeni BI proizvođači imaju različito viđenje kad u svojim prezentacijama pričaju o svom softveru koji ima „najbolje“ *dashboardove,* odnosno nadzorne ploče. Business Objects, tako definira nadzorne ploče odnosno *dashboardov*e kao prikaz mjera performansi. Njihove nadzorne ploče nazivaju se „*My KPIs*“, prema tome oni smatraju kako su dashboard i *KPI* praktički sinonimi. (Few, 2006.)

Jednostavan dizajn trebao bi omogućiti korisniku uvid u ključne trendove i informacije vezane za njegov biznis. Međutm, vidimo mnogo nespretnih „brzinomjera“ koji više mogu zbuniti korisnika, nego mu otkriti neko novo znanje. Takvi „brzinomjeri“ (eng. *gauge*) koriste se kada se želi prikazati pojednostavljena slika o razini vrijednosti nekog promatranja (npr. loše, srednje, dobro). Više o brzinomjerima biti će objašnjeno u nastavku. Naravno, Business Objects nadzorne ploče imaju visoku razinu kastomizacije, pa korisnik može definirati dizajn ploče po svojoj volji.

Oracle nadzorne ploče definira kao kolekciju mjera prodaje koje se koriste za analizu performansi pojedine kategorije proizvoda. Sve mjere, naravno, prikazane su grafički, iako su podaci ponekad prikazani i tablično.



Slika 13. Oracle Dashboard – nadzorna ploča; Izvor: https://www.oracle.com/index.html

Tablični prikaz s velikim brojem stupaca i redaka svakako bi se trebao izbjegavati kod izrade nadzornih ploča. U ovom slučaju riječ je o relativno malim tablicama, pa ih je moguće ukomponirati u nadzornu ploču. Filteri (na lijevoj strani ove nadzorne ploče) predstavljaju odličnu funkcionalnost koja korisniku pruža da vrlo brzo i efikasno promijeni postavke (uglavnom je riječ o vremenskom okviru) grafičkog prikaza podataka. Interakcija korisnika s nadzornom pločom u realnom vremenu dodatno obogaćuje vrijednost prikaza korisniku. (Few, 2006.)

3. ALATI ZA VIZUALIZACIJU PODATAKA

3.1 Uvod u tematiku alata za vizualizaciju podataka

Altai za vizualizaciju podataka datiraju još iz 50ih godina 20. stoljeća pojavom prvih tranzistorskih računala. U to vrijeme računala su mogla generirati tablice podataka, ali i jednostavne grafove i slike koji su upotpunjavali rezultate provedenih analiza. (McCormick et al., 1987.) 80ih godina 20. stoljeća, informacijska tehnologija doživljava svojevrstan uzlet. Samim time, vizualizacija podataka u kontekstu informacijskih tehnologija doživljava značajan napredak. Uslijed povećanja računalne snage, počinju se razvijati kompleksniji numerički modeli a količina generiranih podataka raste eksponencijalno iz godine u godinu što dovodi do informacijske krize, koja je objašnjena u drugoj cjelini ovog diplomskog rada. To je stvorilo potrebu za alatima koji su u mogućnosti obraditi ogromne količine podatka i vizualno ih predočiti korisniku (McCormick et al., 1987).

Alati za vizualizaciju podataka prvotno su se koristili u znanstvene svrhe kao dio procesa znanstvenog istraživanja, odnosno računalnog modeliranja i simuliranja u inženjerstvu. Informatizacijom kompletnog poslovanja, alati za vizualizaciju podataka počinju se koristiti i u ostalim granama, prvenstveno u financijama i administraciji, ali i u granama poput medicine i digitalnih medija. Od tada, pa do danas alati za vizualizaciju podataka koriste se u poslovnim, ali i osobnim svrhama. Sve jednostavnija sučelja i jednostavniji načini prikupljanja podataka dovele su vizualizaciju podataka na dnevnu bazu čitave populacije. Primjer vizualizacije podataka u svakodnevnom životu su mobilne aplikacije za *Fitness* (Npr. Apple Health) koji putem brojnih senzora u pametnim telefonima, pametnim satovima ali i obući i odjeći prikupljaju ogromne količine podataka bilo da se radi o treningu, šetnji ili snu korisnika. Dakle, alati za vizualizaciju podatak u današnjem svijetu vrlo su dostupni i bliski ne samo poslovnim korisnicima, nego i potrošačima. S obzirom da se u ovom radu uspoređuju poslovni alati za vizualizaciju podataka, u nastavku će biti riječ isključivo o njima.

U novije vrijeme količina podataka koja se generira u modernim informacijskim sustavima znatno je veća nego ikada prije. Pojam *Big Data* pojavio se kao objašnjenje za skup podataka koji su toliko veliki i kompleksni da ih je nemoguće obraditi konvencionalnim načinima.
Big Data tako izlazi iz okvira baza i skladišta podataka, i obuhvaća doslovno sve moguće informacije koje korisnik može i zna prikupiti. *Big Data* sastoji se od „4 V“:

- *Volume* (Volumen) odnosno količina podataka
- *Velocity* ili brzina obrade podataka
- *Variety* odnosno raznolikost podataka
- *Veracity* ili istinitost podataka.

„Big Data je rezultat razvoja tehnologije zadnjih nekoliko desetljeća pogotovo razvojem MPP-a (*Massive Pararell Processing*), jeftinih servera visoke snage, te *cloud* tehnologije, međutim Big Data kao marketinški pojam je ušao kao *mainstream* tehnologija zadnjih 2-3 godine. Big Data je možda jedan od najrazvikanijih termina u IT-u, ali zato pruža nove prilike tvrtkama da se diferenciraju od konkurencije. Osobno mislim da je Big Data prilika koju apsolutno treba iskoristiti i smatram da će ovo "dijete" još morati rasti kako bi u budućnosti doživilo svoj puni potencijal." – Gabelica, 2013.

3.2 Funkcionalnosti alata za vizualizaciju podataka

Kako bi bilo moguće izvršiti uspješnu analizu, odnosno vizualizaciju podataka, ključno je odabrati alat sa odgovarajućim setom funkcionalnosti. Definicijom pojma *Big Data*, moderni alati za vizualizaciju podataka moraju zadovoljiti 4 V principe kako bi u modernom poslovanju bili što efikasniji. Kvaliteta izvornih podataka često je problem, pa je ključno da korisnik može pročistiti i pripremiti podatke u sučelju alata. Kada je minimalna kvaliteta podataka postignuta, moguće je pristupiti vizualnoj analizi istih. Prema O'relyju (2012) navedene u nastavku su navedene osnovne funkcije modernih analitičkih alata za vizualizaciju.

Konektivnost na različite poslovne baze, skladišta i formate podataka jedna je od osnovnih funkcionalnosti alata. Integriranost alata sa postojećim poslovnim sustavom bez nužnosti za dodatnim ETL procesima odlika je naprednih rješenja za vizualizaciju. Ukoliko je potrebno dodatno obraditi podatke prije analize, korisnik to može obaviti u sučelju samog softvera. Moguće je napraviti „manje zahvate" na skupu podataka kao što su promjena tipa varijable, prebacivanje iz dimenzija u mjere (i obratno).

Dinamički prikaz, odnosno spajanje na *live* izvore podataka također je važna funkcionalnost alata za vizualizaciju. Ovo je ključno kod izrade nadzornih ploča, jer je njihova osnova spajanje na izvore koji se mijenjaju i nadopunjuju kroz vrijeme. Ovakva *on-line* povezanost u početku

je bila odlika samo najmoćnijih alata, dok je danas postala obavezna funkcionalnost bilo kojeg alata.

Brzina konekcija alata na *dataset,* bilo da se radi o *live* ili statičnom izvoru podataka također je vrlo bitna. To se odnosi na *Velocity* kao jednog od elemenata *Big Data* koncepta. Brzina uglavnom ovisi o prvotnoj kvaliteti i strukturiranosti podataka (*raw data*). Većina modernih rješenja gotovo je nemoguće koristiti na nestrukturiranim podacima, pa je tu ostavljen prostor za napredak u budućim iteracijama popularnih alata.

Kada alat uspješno učita podatke, ukoliko je potrebno sa željenim promjenama, korisnik započinje analizu temeljenu na vizualizaciji. Kao što je već navedeno, korisnici ovakvih alata nisu nužno programeri, IT osoblje ili podatkovni stručnjaci. Alati poput Tableau-a i Power BI-a namijenjeni su uglavnom ne-tehničkom osoblju, menadžerima i analitičarima. Zbog toga je važno da se alati mogu uspješno koristiti bez ili sa ograničenim programerskim znanjem. Dakle, uspješna analiza mora biti moguća bez napisane linije koda. Grafičko korisničko sučelje (GUI) funkcionalnost je koja omogućava ovaj kriterij.

Širok izbor grafikona esencijalna je funkcionalnost svakog vizualizacijskog alata. U radu su navedeni najčešće korišteni grafikoni. Važno je da uz njih, alat sadrži i dodatne grafikone, ali i varijacije na one postojeće. Jednostavnost korištenja navedena je u prijašnjem paragrafu, ali isto vrijedi i za izradu grafikona. Uz intuitivno sučelje, važno je i dobro korisničko iskustvo; korištenje provjerenih metoda za izradu i uređivanje grafikona (primjer. *Drag 'n' drop* varijabli i mjera direktno na osi grafikona). Interaktivnost korisnika sa grafikonom još je jedna od funkcionalnosti pomoću koje se povećava informacijska vrijednost grafikona. Odabirom elemenata na grafikonu, korisnik može saznati dodatne informacije koje na statičkom prikazu nisu vidljive.

OLAP analize kao što su *drill-down, roll-up, slice* i *dice* te pivotiranje također su funkcionalnosti alata za vizualizaciju. Filtriranje podataka unutar korisničkog sučelja isto je važna funkcionalnost. Uz OLAP metode analize, alati za vizualizaciju moraju imati i kompleksne analitičke metode. To su na primjer izračuni kvantila, klasteriranje, prognostika, multidimenzijsko skaliranje isl. Ove statističke metode možda nisu namijenjene za prosječnog korisnika, ali su svejedno važne ukoliko se želi napraviti složenija analiza.

Budući da se vizualne analize često kreiraju u prezentacijske svrhe, odnosno u službu komunikacije sa ostalim članovima tima, funkcionalnost dijeljenja (*share*) i kolaboracije isto je vrlo važna. Razvojem *cloud* tehnologije, omogućeno je jednostavno dijeljenje rezultata, njihova interaktivnost sa ostalim članovima, te mogućnost nadogradnje.

3.3 Tržište alata za vizualizaciju podataka

Tržište alata za vizualizaciju podataka mijenja se iz godine u godinu. Poduzeća su prepoznala važnost koncepta poslovne inteligencije i alata za njihovu provedbu, zbog čega je tržište u svojevrsnom porastu. Forbes (2016) predviđa rast prihoda tržišta analitičkih i *Big Data* softvera na više od 50%, odnosno od 122 mlrd. u 2015. na 187 mlrd. dolara u 2019. godini. Samim time, tržište alata za vizualizaciju narast će za 22% unutar pet godina (CAGR) sa 415 milijuna dolara u 2014 na 1.1. mlrd. dolara u 2019. Također se predviđa da će do 2020. tržište *Cloud* rješenja za analitiku rasti 4.5 puta brže od tržišta *On-premise* softvera (desktop verzija softvera). Reuters (2017) također predviđa rast tržišta softvera za vizualizaciju podataka i procjenjuje njegovu vrijednost na 6.99 mlrd. do kraja 2022 godine uz petogodišnji (2017-2022) CAGR rast od 9.2%. Trenutna vrijednost tržišta softvera za vizualizaciju je prema Reutersu procijenjena na 4.12 mlrd. dolara (2017).

Što se tiče tržišta alata za poslovnu analitiku, Oracle je predvodnik sa 16% tržišta, dok su izazivači IBM (10.3%), Microsoft (9.1%) i SAS (6.1%), (Forbes, 2016). Alati za vizualizaciju podataka samo su jedan od segmenata tržišta poslovne analitike, ali ga većina agencija za procjene tržišta ne prikazuje kao zasebnu komponentu. Na popularnim internet forumima i specijaliziranim stranicama za vizualizaciju (PCMag, 2017) najčešće se spominju sljedeći alati: Power BI; Tableau Desktop; Looker; Domo; QlikView i Zoho Reports.

Gartner, marketinška agencija za istraživanje tržišta svake godine u prvom kvartalu izda takozvani „magični kvadrant“, odnosno BCG matricu u kojoj objašnjavaju trenutno stanje na tržištu alata za vizualizaciju podataka. Na osima matrice nalazi se *Ability to execute* (mogućnost softvera), odnosno *Completeness of vision* (sveukupna vizija i njezino ostvarenje). Prema tim varijablama, grafikon je podijeljen u četiri kvadranta: Nišer, Vizionari, Lideri i Izazivači. Na slici se nalaze Gartnerovi „magični kvadranti“ za 2016 i 2017 godinu.



Slika 14. Gartnerov magični kvadrant za 2016. odnosno 2017. godinu,
Izvor: http://www.gartner.com/technology/home.jsp

Gartner stavlja 3 rješenja u kvadrant lidera, Tableau, Microsoft (Power BI i komplementarna rješenja), te Qlik. U 2016. Tableau je istaknut kao rješenje koje ima najveću razinu mogućnosti, dok su Qlik i Microsoft predstavljeni kao proizvodi sa jasnijom vizijom. U 2017. Microsoft i Tableau su se dosta odmakli od ostatka rješenja u smislu mogućnosti, ali je Microsoft određen jasnijom vizijom. Prema tome, Gartner smatra Microsoftova rješenja (Power BI) kao lidera na tržištu, dok je Tableau glavni izazivač za prvo mjesto.

3.4 Najpopularnija rješenja

3.4.1 Tableau

Trenutačno jedno od najpopularnijih rješenja za vizualizaciju podataka. Tableau dolazi u 3 verzije – Desktop, Server i Online. Desktop verzija *On premise* je rješenje koja se instalira na računalu i koristi se za *online* i *offline* konekciju sa izvorima podataka. Ovaj softver osmišljen je za izradu i dijeljenje izvješća iz različitih podatkovnih izvora. Tableau Server omogućava korisnicima obradu podataka na serveru, pa ne zahtijeva kompletnu instalaciju na računalo. Klijent aplikaciju moguće je instalirati ili pokrenuti u pregledniku, što znatno olakšava korištenje u velikim organizacijama. Dijeljenje i skupna obrada podatka olakšana je kada svi korisnici pristupaju istom serveru na kojemu se Tableau nalazi što je jedna od osnovnih prednosti server rješenja. Tableau Online sličan je Tableau serveru, samo što je u ovom slučaju server u vlasništvu kompanije Tableau-u. Prema tome, Tableau Online je SaaS (*Software-as-a-Service*) rješenje koje ne zahtijeva posjedovanje vlastitog servera, već se korisnik jednostavno prijavi na stranicama proizvođača koji mu ustupi čitavo rješenje u oblaku. Tableau nudi različite modele plaćanja, bilo da se radi o godišnjoj pretplati za korisnika ili jednokratnim plaćanjem. Tableau će biti detaljnije objašnjen u četvrtom poglavlju diplomskog rada.

3.4.2 QlikView

QlikView je prema Garneru (2017) jedan od lidera na tržištu alata za vizualizaciju podataka. Ovaj alat specijaliziran je za izradu kontrolnih ploča i interaktivnih grafova sa visokom razinom kastomizacije. Softverska firma Qlik, sa sjedištem u SAD-u proizvođač je ovog vizualizacijskog alata osnovana je u Švedskoj 1993. te je 2016. prodana fondu Thoma Bravo za cca. 3 mlrd. dolara. QlikView je svojevrsni *flagship* (glavni) proizvod tvrtke Qlik. Qlik na svojim korporativnim stranicama tvrdi da trenutno imaju oko 40 000 klijenata, od kojih su izdvojili revizorsku kompaniju Deloitte.

Ovaj alat ima automatske postavke za spajanje na izvore podataka, ali ih je moguće promijeniti ukoliko se korisnik želi spojiti na izvor koji nije naveden u automatskim postavkama. Dakle, QlikView je moguće spojiti na širok spektar izvora podataka, što predstavlja prednost u odnosu

na druge alate. Ovo poslovno rješenje ima i svoje API-je (eng. *application programming interface*, odnosno aplikacijsko programsko sučelje) što mu omogućuje spajanje na bilo koji softverski ekosustav. Što se tiče metoda isporuke, QlikView nudi *on premise* (desktop) i *cloud* rješenje, te je samim time i određen model plačanja – jednokratnom uplatom za *on-premise* rješenje ili mjesečnom pretplatom po korisniku za *cloud* rješenje.

Osnovne karakteristike alata QlikView rješenja su:

- Vizualizacija podataka
- Dinamični BI ekosustav
- Interaktivne kontrolne ploče i grafikoni
- Konektivnost sa izvorima podataka
- *Workflow* i dodjela uloga
- Visoka razina kastomizacije
- Kompatibilnost sa složenim analitičkim modelima
- *Mobile-frendly* pristup

Glavni nedostaci ovog alata su općenita kompleksnost uporabe, visoka razina kastomizacije i mogućih opcija zahtijeva i visoku razinu znanja u području strukture podataka, ali i programiranja.

Što se tiče cijena, QlikView nudi i besplatne opcije za korisnike koji ga ne koriste u komercijalne svrhe:

- Desktop verzija/ licenca
    - Nekomercijalni korisnici – besplatno
    - Komercijalni - 1500$/korisniku
- Cloud verzija
    - Do 5 korisnika besplatno
    - Više od pet pretplata 20$/mjesec
    - Business paket 25$/mjesec (dodatne opcije)

3.4.3 Microsoft Power BI

Power BI relativno je nov alat koji je prema Gartneru unutar godinu dana od izlaska postao jedan od lidera na tržištu softvera za vizualizaciju. Osmišljen kao Microsoftov odgovor na Tabelau, Power BI kombinacija je Microsoftovog *legacy* softvera i intuitivnih alata poput Tableau-a. Power BI ima 74 podatkovnih konektora, odnosno kompatibilan je sa 74 različita izvora podataka bilo da se radi o bazama podataka, skladištima ili običnim formatima datoteka (primjer .xls, .csv). Prednost ovog alata je ta što koristi provjeren Microsoftov UI i UX (eng. *User Interface, User Experience*), odnosno korisničko sučelje i iskustvo. Korisnicima je prema tome lako započeti sa korištenjem alata jer su već do sada vjerojatno naviknuti na standardne Microsoftove proizvode poput Microsoft Office-a. U početku zamišljen kao dodatak na Excel, Microsoft je odlučio Power BI učiniti zasebnim alatom zbog znatnog povećanja funkcija. Opširnije o ovom alatu biti će napisano u 4. poglavlju diplomskog rada.

3.4.4 Looker

Looker je alat, odnosno platforma za vizualizaciju podataka koji ima drugačiji pristup od alata koji su do sada navedeni. Ovaj alat osmišljen je kao rješenje koje objedinjuje više različitih SQL i NoSQL izvora podataka. Looker je, dakle, *Big Data* alat koji integrira više izvora podataka u i prikazuje ih kao jedinstven sustav. Vizualizacija podataka nije u prvom fokusu kod odabira ovog alata, ali je jedna od njegovih osnovnih funkcija. Looker je nišer na tržištu alata za vizualizaciju, što pretpostavlja uži segment tržišta na koji je fokusiran. Na korporativnim web stranicama Looker navodi neke od klijenata koji koriste ovu platformu, a izdvojeni su KickStarter i Yahoo.

Looker poput Qlik-a zahtijeva programerske vještine i znanje SQL-a kako bi se mogao uspješno koristiti. Ove vještine potrebne su kod inicijalnog povezivanja alata na podatkovne izvora. Kada je alat povezan na neki sustav, korisnici koji nemaju navedene vještine mogu bez poteškoća nastaviti koristiti alata u analitičke svrhe.

Osnovne karakteristike ovog alata su:

- Vizualizacija podataka
- *Cloud* tehnologija (SaaS)
- Integracija sa različitim modelima podataka / API integracija
- Visoka razina kastomizacije

- *Workflow* (sigurnost i prava korisnika u pristupu podacima)
- Fleksbilan i jednostavan jezik LookML
- Jednostavno korisničko sučelje
- *Live* pristup podacima

Nedostaci ovog rješenja razina je znanja koja je preduvjet za uspješno korištenje. Međutim, ukoliko organizacija ima stručnjaka u domeni SQL-a ovaj nedostatak postaje prednost. Drugi često navedeni nedostatak odnosi se na cijenu rješenja. Cijena se određuje na mjesečnoj pretplati, te će organizaciju u kojoj 10 ljudi koristi ovaj alat on koštati 3000$ mjesečno. Cijena odabirom dodatnih opcija može narasti i do 5000$ što ovaj alat čini skupim rješenjem u odnosu na konkurenciju.

3.4.5 Domo

Domo je kompletno BI rješenje temeljeno na *Cloud* tehnologiji koje je usmjereno na podatkovne izvore. Domo je alat koji na tržištu alata za vizualizaciju ima najveći broj podatkovnih konektora (više od 450), odnosno najvišu razinu integracije. Alat je napravljen u obliku kontrolne ploče na koju korisnik dodaje grafikone i poglede. Domo omogućava jedinstveni pogled iz više izvora podataka koje korisnik modificira po želji standardnim SQL upitima. Vizualizacija podataka dolazi u paketu sa 85 različitih grafikona koji se dodatno mogu modelirati prema korisnikovim željama. Kao i kod većine modernih alata, Domo koristi provjerene UI/UX prakse koje uključuju *drag-and-drop* u izradi grafikona. Čitav dizajn osmišljen je preko predefiniranih "kartica" koje korisnik dodatno uređuje po želji. Iskusniji korisnici mogu promjene u dizajnu i upite na izvore podataka vršiti pomoću konzole. Domo je prema jednostavnosti korištenja pozicioniran po sredini s obzirom na dosadašnje obrađene alate. Relativno je jednostavan za korištenje ne-tehničkom osoblju, ali pruža i dovoljan broj kompleksnih mogućnosti iskusnijim korisnicima. „Domo ima odličan 80-20 omjer, 80% mogućnosti alata brzo se mogu naučiti korištenjem, dok je za preostalih 20% potrebno iskustvo od nekoliko godina korištenja" PCMag, 2016

Ovaj alat u obliku SaaS pruža mogućnost kolaboracije na projektima što je idealno za timove. FinancesOnline u svom osvrtu na ovaj alat navodi kako je „najdruštvenije rješenje", odnosno rješenje koje ima najbolje razrađenu komponentu kolaboracijskih analiza u obliku svojevrsnog društvenog medija. Samim time, velik dio fokusa nalazi se na *workflowu*, odnosno korisničkim dopuštenjima i toku odobrenja. Domo je također jedan od malobrojnih alata koje ima *mobile-*

first pristup vizualizacijskim alatima, odnosno, fokusiran je na pametne telefone i tablete više nego na osobna računala. Korisnici alatu pristupaju jednostavno preko preglednika ili preko klijent aplikacije. Osnovne karakteristike alatu su:

- Jednostavni *Dashboard* dizajn
- Izbor grafikona i prikaza
- Fleksibilni podatkovni konektori
- Projekt menadžment
- Buzz platforma za komunikaciju
- Socijalni elementi i dijeljenje
- *Mobile-friendly*

Negativne strane ovog alata su pomalo zastarjelo korisničko sučelje i kompleksnost uporabe u smislu dubinske uporabe. Krivulja učenja u početku je eksponencijalna, ali uvođenjem kompleksnijih opcija postaje logaritamska. Što se tiče cjenovnog modela, Domo ima nekoliko skupina. Osnovna verzija koja pruža podatkovni kapacitet od 100GB i 80 podatkvnih konektora u oblaku je besplatna. Profesionalna verzija koja uključuje svih 400 konektora i sustav za administriranje korisnika, a može biti u oblaku ili *on-premise* košta 175$ dolara po korisniku mjesečno. Domo također korisnicima pruža i *Enterprise* model sa dodatnim funkcionalnostima kao što su API-ji, češća ažuriranja i poboljšana sigurnost..

3.4.6 Dodaci za programske jezike

3.4.6.1 D3.js

D3.js (*Data-Driven Documents*) je biblioteka za programski jezik JavaScript a koristi se za stvaranje dinamičkih i interaktivnih grafikona koji se pokreću i prikazuju u web pregledniku. To je alat za vizualizaciju podataka u W3C standardima korištenjem skalabilnih vektor grafika (SVG), JavaScripta, HTML5 i CSS3. D3 je nasljednik ranijeg Protovis frameworka. D3.js namijenjen je za komponiranje unutar web-stranica, pa je prema tome često korišten. Ovaj alat pruža visoku razinu kontrole nad konačnim vizualnim rezultatom, pa je upravo zbog toga jedno od popularnijih rješenja. Najčešći oblik u kojem su podaci prikazani za D3.js je tekst datoteka formata .csv, iako JavaScript može čitati i ostale podatkovne formate. (D3, 2017)

3.4.6.1 Chart.js

Chart.js također je JavaScript biblioteka koja je *open-source* tipa – omogućeno je besplatno korištenje u privatne i poslovne svrhe. Napravljen je na HTML5, CSS3 i JavaScript web tehnologijama, što mu omogućava optimalan prikaz unutar web preglednika. Chart.js nudi relativno mal broj grafikona – samo osam (linijski, stupčasti, točkasti, *donat, area* te nekoliko verzija pita grafikona) ali je zbog interaktivnosti i responzivnosti grafikona često korišteno rješenje u izgradnji web stranica. Ova biblioteka pruža i zanimljive animacije unutar grafikona, koje nemaju informacijsku vrijednost već su isključivo estetske. (Chart.js, 2017)

3.4.6.2 FusionCharts

FusionCharts također je JavaScript biblioteka osmišljena za grafički prikaz podataka iz baze. Kao i prethodna dva alata, FusionCharts napravljen je u web tehnologijama HTML5, CSS3 i JavaScript, te pruža izbor interaktivnih grafikona i animacija. Ovo rješenje nudi i složenije akcije nad podacima kao *drill-down* i pivotiranje. Cijena ovog rješenja je 500$ mjesečno. (FusionCharts, 2017)

3.4.6.3 R

R je statistički programski jezik koji je nastao kao *open-source* i GNU projekt (besplatna licenca), a razvila ga je firma Bell (danas Lucent Technologies) 1993. godine. R se od tada koristi u svrhu razvoja statističkog softvera, analize podataka i projekata rudarenja podataka. R nudi širok spektar statističkih operacija što uključuje klasičnu statistiku, linearno i nelinearno modeliranje, analizu vremenskih serija, klasteriranje i slično. R je objektno orijentiran što ga razlikuje od ostalih statističkih jezika. U vizualizaciji podataka, R korisniku pruža potpunu kontrolu u dizajniranju grafikona. R se također može nadograditi paketima i bibliotekama koje su specijalizirane za grafičke prikaze što ga čini još moćnijim rješenjem. (R, 2017)

3.4.7 Usporedba alata za vizualizaciju

Usporedba navedenih alata napravljena je u obliku tablice. U tablici su samo gotova softverska rješenja, odnosno nema dodataka za programske jezike. U tablici su prikazani neki od osnovnih kriterija prema kojima su alati uspoređeni (PC Mag, 2017).

Tablica 7. Usporedba alata za vizualizaciju podataka:

Alat/Kriterij	**Tableau**	**QlikView**	**PowerBI**	**Looker**	**Domo**
Free-trial	Da	Ne	Da	Da	Da
Implementacija	Cloud/Desktop	Cloud/Desktop	Cloud/Desktop	Cloud/Desktop	Cloud
Podržava API-je	Ne	Da	Da	Da	Da
R kompatibilnost	Da	Ne	Da	Da	Ne
Broj konektora	58	86	74	52	450
Izbor vizualizacija	Visok	Visok	Visok	Srednji	Srednji
Mobilni klijent	Da	Da	Da	Da	Da
Cijena/korisnik/mj.	42$	125$	10$	300$	175$
Preporučena veličina organizacije	Mala Srednja Velika	Mala Srednja	Mala Srednja Velika	Mala Srednja Velika	Mala Srednja Velika
Podržani operacijski sustavi	Macintosh Windows Linux iOS Android	Windows iOS Android	Windows iOS Android	Windows iOS Android	Web-preglednik

Izvor: https://www.pcmag.com/

4. KOMPARACIJA ALATA TABLEAU I POWER BI

4.1 Pregled značajki alata Tableau

Tableau je jedno od najpopularnijih rješenja poslovne inteligencije koje je specijalizirano za vizualizaciju podataka. S obzirom na dugogodišnju poziciju tržišnog leadera Tableau je postao svojevrsno *legacy* rješenje koje je uzor novim proizvođačima ovakvog softvera. Tableau donedavno nije imao ozbiljnijeg konkurenta, ali pojavom novih generacija vizualizacijskih alata vođenih konceptom *Big Data*, mjesto tržišnog lidera postalo je znatno ugroženo. Jedno od takvih novih rješenja zasigurno je Microsoft Power BI koje će biti razrađeno u sljedećem pod-poglavlju. Pod pritiskom SaaS rješenja Tableau je odnedavno sa isključivo *on-premise* poslovnog modela prešao i na modele pretplate čime je učvrstio svoju poljuljanu tržišnu poziciju.

Kao što je navedeno, Tableau je već *legacy* poslovno rješenje koje je godinama dobrog poslovanja stvorilo jaku on-line zajednicu. Ovakva zajednica iznimno je korisna kod učenja korištenja ovog alata jer pruža zavidnu količinu materijala, posebice multimedijskog tipa. Brojni *tutorial* videozapisi, bilo da se radi o službenim sa korporativnih stranica, ili neslužbenih sa raznih kanala, foruma i web stranica. Zavidan je i broj online Tabelau kursova sa platformi kao što su Coursera, Udemy, Lynda i slično. Učenje je važna stavka svakog alata, a dostupnost znanja omogućava brže i efikasnije savladavanje vještina vezanih za alat. (Financesonline, 2016.)

Unatoč lako dostupnim znanjima o ovom alata, krivulja učenja njegova je svojevrsna mana. Tableau svojim korisnicima pretpostavlja određenu razinu predznanja o podacima i analitici. Današnje tržište alata za vizualizaciju sve se više okreće generalnim, odnosno ne-tehničkim korisnicima, dok Tableau navedeni trend ne prati u potpunosti. Tableau je svoju poziciju na tržištu pronašao na način da je ponudio intuitivan softver iskusnim analitičarima koji su tražili najbolji alat u svojoj domeni. Sofisticiranost alata bila je na visokoj razini, što je proizlazilo iz dotadašnjeg iskustva korisnika u toj domeni. Danas, pak ovakav pristup nije praksa, već se rješenja za vizualizaciju pokušavaju što više približiti svim zaposlenicima neke organizacije. Iako sofisticiran, Tableau je i dalje visoko-intuitivno rješenje. Osnovne funkcije lako je savladati u kratkom vremenskom roku, ali za savladavanje kompleksnijih funkcija potrebna je

određena praksa u korištenju. Korisničko sučelje najintuitivnije je u industriji. ETL procesi jednostavni su za provođenje, što je još jedna od prednosti ovog alata. Konekcija na podatke poprilično je brza, što je prednost u spajanju na *live* izvore podataka. Provođenje analiza poprilično je *straight-forward* za iskusnog korisnika, ali za neiskusnog može biti problematično zbog nedostatka savjeta u obliku standardnih praksi kao što su *pop-up* ili pomoćnih linkova. Kada se savlada (nakon nekoliko pokušaja), dizajniranje vizualizacija iznimno je jednostavno. *Drag-and-drop* ukomponiran je u odabir dimenzija i mjera za kreaciju grafikona. Trening videozapisi sa korporativnih stranica omogućavaju savladavanje osnovnih funkcija alata unutar pola sata. (PC Mag, 2017)

Cijena alata:

- Personal Edition – 35$/Korisnik/Mjesec (Desktop verzija)
- Tableau Server – 35$/Korisnik/Mjesec (Server Verzija)
- Professional Edition – 70$/Korisnik/Mjescec (Desktop, Online, Server)
- Tableau Online – 42$/Mjesec (SaaS rješenje za Tableau Server)

Prema ovom pregledu određeni su Za i Protiv argumenti u odabiru Tableau-a:

Za:

- Integracija - velika kolekcija podatkovnih konektora
- Velika kolekcija vizualizacija i grafova
- Iznimno intuitivan dizajn sučelja
- Brzina izvršenja upita
- On-line zajednica
- *Legacy* proizvod
- Ažurna tehnička podrška

Protiv:

- Problematičan za nove korisnike
- Zahtijeva trening prije korištenja
- Relativno skup
- Zahtijeva iskusne korisnike
- Nedovoljna orijentiranost na sigurnost

4.2 Pregled značajki alata Power BI

Power BI relativno je novo rješenje za vizualizaciju podataka, a osmišljeno je kao izravan odgovor Microsoft-a na Tableau. Iako je prva verzija izišla 2015. ovaj alat se već profilirao kao jedan od lidera industrije. Gartner ga je 2017. svrstao kao najkompletniji proizvod u smislu alata za vizualizaciju. Microsoft je strategiju penetracije tržišta na čelu s Tablau-om temeljio na niskim cijenama (9.99$/Korisnik/Mjesec). Poput Tableau-a, Power BI ima velik broj podatkovnih konektora (74) čime je osigurao jednostavnu integraciju sa ostalim poslovnim sustavima.

Krivulja učenja kod ovog alata pojednostavljena je korištenjem dobro znanih Microsoftovog korisničkog iskustva. Korisničko sučelje podsjeća na ono kod alata Microsoftovog proizvoda Office. Pod pretpostavkom da su korisnici koristili makar jedan od alata unutar Office-a, korisničko iskustvo Power BI-a iznimno je intuitivno. ETL procesi vrlo su jednostavni i ne zahtijevaju predznanje o podatkovnim strukturama, osim kod detaljnijeg čišćenja podataka. Željeni *datasetovi* se jednostavno unose principom označivanja tablica (checkbox). Spajanje na izvor podataka odvija se iznimno brzo uključujući i upite koji se naknadno postavljaju na izvor podataka (u sklopu analize). Kada su podaci odabrani, korisnik se prebacuje na kontrolnu ploču unutar koje provodi analizu kreiranjem željenih vizualizacija. Uz jednostavne upite koji su razumljivi i ne-tehničkom osoblju, Power BI pruža i izbor kompleksnijih upita i analiza. Prema toma ovaj alat može biti koristan i početnicima, ali i iskusnim analitičarima i podatkovnim stručnjacima.

Power BI ima ugrađenu *Publish to web* opciju kojom korisnik može podijeliti interaktivnu nadzornu ploču na odabranu web adresu. Power BI Service online je servis koji je Microsoft pokrenuo u svrhu objavljivanja sadržaja kreiranog ovim alatom. Na ovaj način omogućeno je dijeljenje analiza sa poslovnim suradnicima u vidu interaktivnog *dashboard-a* što je jedna od osnovnih funkcija modernih alata za vizualizaciju. Power BI Service također omogućuje spajanje na popularne poslovne aplikacije i prikaz tih podataka pomoću već gotovih rješenja. (Poslovna Inteligencija, 2016.)

Što se tiče vizualizacije, Power BI korisniku nudi raznolik izbor grafikona koji se jednostavno pridodaju odabranim podacima. Svi grafikoni su interaktivni, pa korisnik može dodati dodatne informacije bez narušavanja vizualne jednostavnosti grafikona. Praksa izrade grafikona

nalikuju na Excel pivot tablice, pa će korisnici s iskustvom korištenja Excel-a brzo savladati koncepte dizajna unutar Power BI alata. (PC Mag, 2017.)

Cijena alata:

- Stanndard edition (nekomercijalni korisnici) - besplatno
- Pro (komercijalni korisnici) – 9.99$/Korisnik/Mjesec

Prema ovom pregledu određeni su Za i Protiv argumenti u odabiru Power BI-a:

Za:

- Brzina konekcije i obrade podataka
- Jednostavnost korištenja (UX/UI)
- Microsoft *legacy*
- Besplatna verzija i niska cijena
- Dijeljenje analiza na Power BI Service platformi
- Izbor grafikona

Protiv:

- Mala web zajednica
- Problemi s performansama kod velikih *datasetova*
- Relativno nov alat – povremeni *bugovi*
- Nedostatak prognostičkih modela

4.3 Komparacija funkcionalnosti alata

Tableau i Power BI nude različite prednosti koje se uzimaju u obzir prilikom odabira adekvatnog alata. Ovaj odabir ne može se temeljiti isključivo na komparaciji funkcionalnosti alata, već se mora uzeti u obzir i organizacija odnosno krasnici za koje je alat namijenjen. U ovom pod poglavlju komparirane su isključivo funkcionalnosti, dok će se u zaključku diplomskog rada autor osvrnuti na kompletnu sliku koju korisnik mora uzeti u obzir kod odabira adekvatnog softvera. Funkcionalnosti alata prikazane su u tablici.

Tablica 8. Usporedba Tableau-a i Power BI-a

Kriterij	Tableau	Power BI
Implementacija	Cloud/Desktop	Cloud/Desktop
Broj konektora	58	74
Dostupni grafikoni	Stupčasti, razloženi, linijski, površinski, točkasti , *pie*, *treemap*, zemljopisni, *funnel*, *gauge*, kartice, KPI, *slicer*, *donut*, kombinirani grafikon	Stupčasti, razloženi, linijski, površinski, točkasti , *pie*, *treemap*, zemljopisni, *heatmap, histogram, gantt, bullet, packed bubbles*
Automatske analize (ne zahtijeva znanje iz statistike)	Generiranje trendonve Prognostički linijski grafikon Određivanje korleacije Jednostavna statistika	Određivanje aritmetičke sredine Odstranjivanje outliera Uklanjanje sezonalnosti Određivanje korelacije Jednostavna statistika
Profil korisnika	Iskusni analitičari	Svi zaposlenici
Servis za dijeljenje	Tableau Public	Power BI Service
Dodatne funkcionalnosti	Tableau Story Mobilni klijent Tableau Reader	R kompatibilnost Ad Hoc analiza Mobilni klijent
Operacijski sustav	Macintosh Windows Linux iOS Android	Windows iOS Android
Cijena/Mjesec/Korisnik	35$ - Persona edition 70$ - Pro edition	Besplatno – Personal edition 9.99$ -Pro edition

Izvor: PC Mag, autor

4.4 Testiranje i komparacija korisničkog iskustva na primjeru zadatka

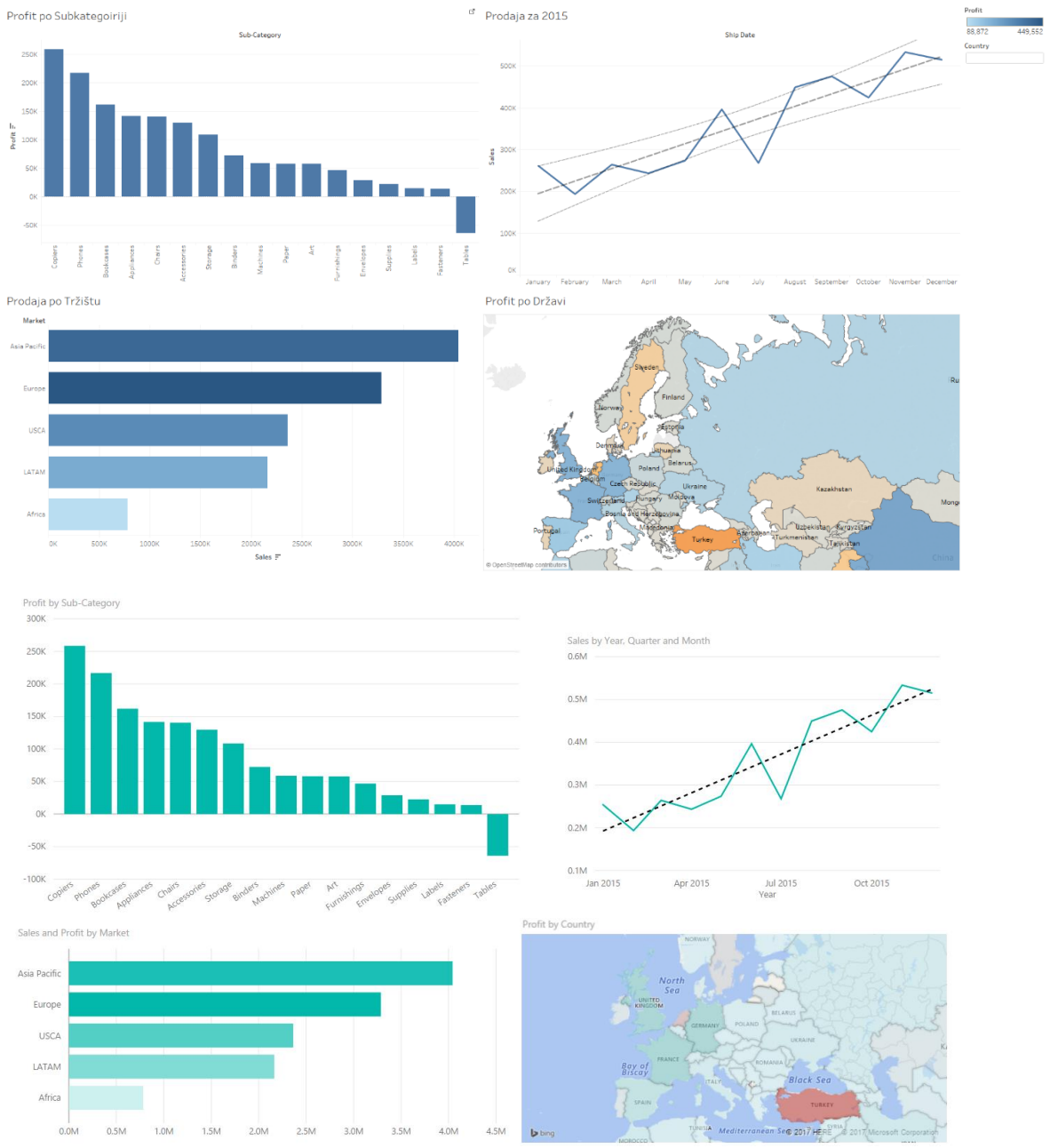
4.4.1 Postavljanje zadatka

U eksperimentalnom dijelu istraživanja napravljena je ekspertna analiza u kojoj je sudjelovalo 8 ispitanika. Ispitanicima je postavljen zadatak koji su trebali riješiti u jednom, odnosno drugom softverskom rješenju – Tableau i Power BI. Svim ispitanicima dodijeljen je isti zadatak za jedan i drugi alat. Četvoro ispitanika riješilo je prvo zadatak za Tableau, a zatim isti za Power BI. Druga četiri ispitanika riješili su zadatak prvo za Power BI, a zatim za Tableau kako bi se uklonio utjecaj redoslijeda rješavanja zadatka na krajnje rezultate. Zadatak je osmišljen kao kratka analiza podataka unutar .xls datoteke pod nazivom Global Store Data a koja je dostupna na linku: https://drive.google.com/drive/u/0/folders/0B3pieZFj_BPXZW4yR1JJVk5kLTA

Unutar datoteke nalaze se podaci fiktivnog poduzeća Global Store koje prodaje tehnološke proizvode, namještaj i uredski inventar a tržište joj se proteže po regijama Europe, Južne i Sjeverne Amerike, Azije, Afrike te Australije i Oceanije. Zadatak vezan za ovaj *dataset* isti je za oba alata a glasi:

1. Konektirajte se na podatke unutar datoteke **Global Store Data** (Excel)
2. Odaberite tablicu **Orders** za analizu
3. Unutar tablice **Orders**, promijenite tip stupca **Row ID** iz tipa **Number** u tip **String (Text)**
4. Spremite promjene i pristupite na **Dashboard**
5. Napravite stupčasti grafikon (**Bar Chart**) koji prikazuje **Profit** po dimenziji **Region**
6. Pivotirajte grafikon tako da **Profit** bude na osi **X**, a dimenzija umjesto dimenzije **Region**, na osi **Y** bude dimenzija **Sub-Category**
7. Na **Dashboard** dodajte novi grafikon (Line Graph) koji sadrži mjeru **Sales** i dimenziju **Ship Date**
8. Na istom grafikonu Drill down metodom prikažite **prodaju po mjesecu**
9. Na istom grafikonu metodom filtriranja prikažite **prodaju po mjesecu samo za 2015. godinu**
10. Na istom grafikonu dodajte liniju trenda (**Add Trendline**)

11. Na istom Dashboard-u dodajte novi grafikon (Bar Chart) koji prikazuje odnos dimenzije **Market** i mjere **Sales** te dodajte dimenziju **Profit** u prikaz boje (**Color Saturation**)
12. Na istom **Dashboard**-u dodajte novi geografski grafikon (**Filled Map**)
13. Na mapi dodajte Profit po državi prikazan nijansom plave boje (**Color Saturation**), a ukoliko je država u minusu neka gubitak bude prikazan nijansom narančaste
14. Dodajte globalni filter po Državi (**Country**) na sve grafikone unutar Dashboard-a
15. Spremite Promjene (**Save**)



Slika 15. Prikaz rješenja zadatka: Prva nadzorna ploča napravljena je u alatu Tableau. Druga nadzorna ploča napravljena je u alatu Power BI; Izvor: Izračun autora

4.4.2 Određivanje sudionika u istraživanju

Za istraživanje je odabrano 8 sudionika. Većina sudionika dolaze iz dvije domene – informacijskih tehnologija i financija, odnosno računovodstva. Jedan sudionik dolazi iz domene električkog inženjerstva, Sudionici imaju različitu razinu iskustva u korištenju alata za analizu podataka, ali svi posjeduju dugogodišnje iskustvo u korištenju Microsoftovog alata Excel. U tablici je prikazan profil svih ispitanika koji su sudjelovali u istraživanju:

Tablica 9. Profil sudionika istraživanja

Ispitanik	**Spol**	**Godine**	**Zanimanje**	**Iskustvo u analitici (alati)**
Ispitanik 1	M	29	Web Develpoer	SQL, Excel, Google Analytics, D3.js
Ispitanik 2	M	26	Programer	SQL, Excel, Baze podataka
Ispitanik 3	Ž	26	Kontroling	SAP Hana, Excel, SAP Lumira
Ispitanik 4	M	25	Student	Excel, Power BI, Tableau, R
Ispitanik 5	Ž	29	Analitičar	Power BI, Tableau, Excel, Pyhton
Ispitanik 6	M	28	Inženjer energetike	Excel, Matlab, Wolfram Alpha
Ispitanik 7	M	38	Sistem inženjer	R2 Server, SAP Hana, Scada, R2 Server, SQL
Ispitanik 8	M	26	Revizija	Excel Pivot Table, Aura Audit software

Izvor: autor

Različito iskustvo i zanimanje ispitanika testirat će i mogućnost korisnosti alata za korisnike različitih profila (a ne samo iskusnih analitičara).

4.4.3 Provođenje eksperimenta

Korisnicima je dodijeljen zadatak i upute za rješavanje u kontekstu pojedinog alata. Četvorica ispitanika prvo su rješavali zadatak u alatu Tableau, a zatim su riješili isti zadataka u alatu Power BI. Druga polovica skupine korisnika prvo je riješila zadatak u alatu Power BI, a zatim u alatu Tableau. Na ovaj način maksimalno se pokušao smanjiti utjecaj redoslijeda rješavanja zadatka na krajnje rezultate. Svi sudionici imali su jednake uvjete rješavanja zadatka, te im je ukoliko je bilo potrebno pružena pomoć autora. Korisnici su rješavali zadatak na računalu (laptopu) tipa HP EliteBook Folio 9470m Ultrabook (Intel i5-3437U; 8GB DDR3; 180GB SSD). Laptop je spojen na eksterni monitor dijagonale 24 inča omjera ekrana 16:9.

4.4.3.1 Praćenje računalnog miša, tipkovnice i vremena

Za vrijeme trajanja zadatka, na računalu je uz alat pokrenuti i softver Mousotron proizvođača Blacksun Software (izvor: http://www.blacksunsoftware.com/mousotron.html). Ovaj alat namijenjen je za praćenje miša, tipkovnice i vremena te se često koristi u analizama korisničkog iskustva softvera. Varijable koje alat pruža, a korištene su u ovom eksperimentu su:

- Vrijeme trajanja (u sekundama i minutama)
- Prijeđena usaljenost kursora (u centimetrima i metrima)
- Broj korištenja tipki na tipkovnici
- Broj korištenja lijeve tipke miša
- Broj korištenja desne tipke miša
- Količina korištenja kotačića miša

Rezultati dobiveni praćenjem korisnika alatom Mousotron prikazani su u sljedećim tablicama:

Tablica 10 . Rezultati za alat Tableau

Alat	Vrijeme	Dužina prijeđena mišem	Tipkovnica	Lijevi klik	Desni klik	Kotačić miša
Tableau	25.01	115.24	1011	321	8	212
Tableau	22.45	95.1	859	398	5	241
Tableau	21.34	112.14	814	312	4	195
Tableau	21.14	83.63	765	353	3	122
Tableau	19.03	67.74	792	241	3	181
Tableau	18.12	78.48	783	253	4	179
Tableau	17.85	69.17	714	260	5	163
Tableau	17.46	79.77	692	261	6	158

Izvor: izračun autora

Tablica 11. Rezultati za alat Power BI

Alat	Vrijeme	Dužina prijeđena mišem	Tipkovnica	Lijevi klik	Desni klik	Kotačić miša
Power BI	20.79	72.04	258	312	12	132
Power BI	19.04	76.25	226	275	8	124
Power BI	17.84	73.48	212	251	9	111
Power BI	16.57	73.11	272	264	6	152
Power BI	16.34	51.45	271	235	6	147
Power BI	12.38	56.24	184	240	4	154
Power BI	12.15	52.75	150	210	6	112
Power BI	11.54	39	146	198	5	131

Izvor: izračun autora

Podaci unutar tablica testirat prikazani su sa sljedećim mjerama:

- Vrijeme je određeno u minutama (min)
- Dužina prijeđena mišem određena je u metrima (m)
- Vrijednost varijable tipkovnica označava koliko je puta korisnik pritisnuo bilo koju tipku na tipkovnici
- Lijevi klik određen je kao broj puta koliko je korisnik koristio lijevu tipku miša
- Desni klik određen je kao broj puta koliko je korisnik koristio desnu tipku miša
- Kotačić miša određen je koliko je broj okretaja kotačića miša

4.4.4 Postavljanje i testiranje Hipoteza

Na dobivenim rezultatima testirana je osnovna hipoteza o Statističkoj značajnosti razlike u korisničkom iskustvu.
H1……...Postoji statistički značajna razlika u korisničkom iskustvu korištenja alata Tableau i Power BI

Ova hipoteza sastoji se od dvije pod-hipoteze, Varijable na kojima su testirane pod-hipoteze su Vrijeme za izvršenje zadatka i dužina prijeđena mišem. Ove dvije varijable određene su kao determinante jednostavnosti korištenja softvera za zadani zadatak kreiranja kontrolne ploče.

H1.1……..Postoji statistički značajna razlika u vremenu potrebnom za izvršavanje zadatka

H1.2…....Postoji statistički značajna razlika u prijeđenoj udaljenosti računalnog miša za izvršenje zadatka

Obje hipoteze testirane su t-testom, odnosno testom za otkrivanje statističke značajnosti razlike između dva uzorka to jest razliku između njihovih aritmetičkih sredina. Tip t-testa koji je korišten u istraživanju je te-test za nezavisne uzorke (eng. *Independent Samples Test*). Ako se ustvrdi da je razlika statistički značajna, to ne znači nužno veliku razliku, nego taj pojam u statistici ima sasvim određen i definiran smisao: Ako se utvrdi da je neka razlika statistički značajna, onda se zapravo utvrdilo da razlika, koja je nađena, bez obzira na veličinu iste, nije slučajna, već da razlika vrlo vjerojatno postoji i među populacijama. Razina signifikantnosti (značajnosti, p vrijednost) postavljena je na znanstveni standard od 5% što označava da je šansa pogreške prihvaćanja krive hipoteze postavljena na 5%. (p = 5%)

Testiranje hipoteze H1.1

H1.1……..Postoji statistički značajna razlika u vremenu potrebnom za izvršavanje zadatka

p = 5%

Tablica 12. Deskriptivna statistika

	Alat	N	Mean	Std. Deviation	Std. Error Mean
Vrijeme	Tableau	8	20.300000	2.6464855	.9356739
	Power BI	8	15.870000	3.4012813	1.2025345

Izvor: Izračun autora; IBM SPSS

Tablica 13. Rezultati T- Testa

		Levene's Test for Equality of Variances		t-test for Equality of Means						
		F	Sig.	t	df	Sig. (2-tailed)	Mean Difference	Std. Error Difference	95% Confidence Interval of the Difference	
									Lower	Upper
Vrijeme	Equal variances assumed	.755	.400	2.907	14	**.011**	4.4300000	1.5236716	1.1620495	7.6979505

Izvor: Izračun autora; IBM SPSS

Izračunata empirijska p vrijednost je 0.011 (1.1%) što je manje od postavljene p vrijednosti 0.05 (5%). S obzirom na rezultat, hipoteza **H1.1** prema kojoj postoji statistički značajna razlika u vremenu potrebnom za izvršavanje zadatka **SE PRIHVAĆA.**

Testiranje hipoteze H1.2

H1.2…....Postoji statistički značajna razlika u prijeđenoj udaljenosti računalnog miša za izvršenje zadatka

p = 5%

Tablica 14. Deskriptivna statistika					
	Alat	N	Mean	Std. Deviation	Std. Error Mean
Duzina prijedena misem	Tableau	8	87.658750	18.1923543	6.4319685
	Power BI	8	61.790000	13.7205602	4.8509506

Izvor: Izračun autora; IBM SPSS

Tablica 15. Rezultati T- Testa										
		Levene's Test for Equality of Variances		t-test for Equality of Means						
		F	Sig.	t	df	Sig. (2-tailed)	Mean Difference	Std. Error Difference	95% Confidence Interval of the Difference	
									Lower	Upper
Duzina prijedena misem	Equal variances assumed	.670	.427	3.211	14	**.006**	25.8687500	8.0561741	8.5899751	43.1475249

Izvor: Izračun autora; IBM SPSS

Empirijska p vrijednost dobijena T-testom iznosi 0.006 (0.6%) što je manje od inicijalno postavljene p vrijednosti od 0.05 (5%). S obzirom na dobiveni rezultat, hipoteza **H1.2** prema kojoj postoji statistički značajna razlika u prijeđenoj udaljenosti računalnog miša za izvršenje zadatka **SE PRIHVAĆA.**

4.4.5 Interpretacija empirijskih rezultata

Empirijski testovi dokazali su točnost postavljene hipoteze H1 koja pretpostavlja statistički značajnu razlika u korisničkom iskustvu korištenja alata Tableau i Power BI s obzirom na zadani zadatak. Ova hipoteza objašnjena je dvjema pod-hipotezama H1.1 i H1.2 koje su prihvaćene kao istinite temeljem rezultata t-testova. Hipoteza H1.1 dokazala je statistički značajnu razliku u vremenu potrebnom za izvršenje zadatka. Prosjek za izvršenje zadatka u alatu Tableau iznosi 20 minuta i 20 sekundi, dok prosjek za izvršenje istog zadatka u alatu Power BI iznosi 15 minuta i 52 sekunde. Očita razlika od gotovo 5 minuta statistički je dokazana. Hipoteza H1.2 pretpostavila je statistički značajnu razliku u prijeđenoj dužini računalnog miša za izvršenje zadatka. Ova razlika također je statistički dokazana t-testom, što se može i zaključiti po prosjecima pojedinog alata. Prosjek u prijeđenoj udaljenosti miša za Tableau iznosi 87 metara i 67 centimetara, dok je za isti zadatak u alatu Power BI prijeđen 61 metar i 79 centimetara. Power BI je u ovom segmentu usporedbe ostvario bolje rezultate od Tableau-a.

4.4.6 Intervjuiranje ispitanika

Po izvršenom zadatku, ispitanici su intervjuirani na temu testiranih alata. Dojmovi ispitanika uzimaju se u obzir u ukupnom rezultatu. Ispitanicima su postavljena četiri temeljna pitanja:

- Koji alat je jednostavniji za korištenje i zašto?
- Koje su glavne prednosti alata Tableau?
- Koje su glavne prednosti alata Power BI?
- Koji alat bi izabrali za daljnje korištenje?

Pet od osam ispitanika odabralo je Power BI kao alat koji je jednostavniji za korištenje. Najčešći argument za potvrdu ovakve izjave bila je činjenica da su korisnici naviknuti na Microsoftovo korisničko sučelje. Microsoftovo nasljeđe odabrano je kao očita prednost u korištenju. Ispitanici su također komentirali manji broj koraka za odrađivanje funkcija koje su bile zadane zadatkom. Troje ispitanika koje je odabralo Tableau kao jednostavnije rješenje imali su prijašnje iskustvo s alatom, te su općenito iskusniji u analizi podataka. Oni su komentirali kako Tableau ima preglednije korisničko sučelje i kako je lakše koristiti *drag-and-*

drop metode na takvo sučelje. Također su primijetili kako su mjere i dimenzije jasno razdijeljene za razliku od onih u alatu Power BI gdje su iste prikazane unutar zajedničke liste.

Na pitanje koja je najbolja prednost Tableau-a, šest ispitanika odgovorilo je preglednije sučelje i ljepši grafikoni. Jedan ispitanik odgovorio je kako je osnovna prednost Tableau-a općenito korisničko iskustvo koje je bolje nego u alatu Power BI. Jedan ispitanik odgovorio je kako je osnovna prednost alata intuitivniji način za filtriranje podataka i dodavanje boja u svrhu prikazivanja više dimenzija. Na pitanje koja je prednost alata Power BI, troje korisnika odabralo je činjenicu da se grafikoni automatski dodavanju na *dashboard,* za razliku od Tableau-a gdje ga korisnik sam mora izraditi. Također su dodali kako je jednostavnije uređivanja grafikona na kontrolnoj ploči. Troje korisnika komentiralo je kako je osnovna prednost Power BI-a intuitivno Microsoftovo sučelje na koje su naviknuli uporabom Excel-a. Jedan ispitanik odabrao je jednostavnije filtriranje podataka na *dashboard*-u (bitno je samo odabrati npr. jednu zemlju na mapi i svi grafikoni se automatski promjene na podatke te zemlje). Jedan ispitanik odabrao je izbor grafikona i njihov izgled kao osnovnu prednost Power BI-a

Na pitanje koji bi alat izabrali za daljnje korištenje, četvoro ispitanika odabralo je Power BI. Riječ je četvoro ispitanika koji su odabrali ovaj alat kao rješenje koje je lakše koristiti. Troje ispitanika koje su na prvo pitanje odabrali Tableau zaključili su kako bi isti alat nastavili koristiti u budućnosti. Jedan ispitanik koji je odabrao Power BI kao lakše rješenje komentirao je kako bi, uz određenu razinu obuke radije nastavio koristiti Tableau jer mu se čini da posjeduje više opcija te da postane iznimno jednostavan i intuitivan kroz duže iskustvo korištenja. Isti taj korisnik izjavio je kako mu Power BI izgleda kao dodatak na Excel, dok je Tableau zasebni alat koji „predivno izgleda“.

Ovim kratkim intervjuom empirijsko statističko istraživanje nadopunjeno je subjektivnim doživljajem korisnika. Zanimljivo je kako su korisnici koji su odabrali Tableau kao bolje rješenje isto imali bolje rezultate u korištenju Power BI-a. Završetkom ovog istraživanja, kombiniranim sa komparacijom funkcionalnosti alata i teorijskom pozadinom dolazi se do zaključka ovog diplomskog rada.

5. ZAKLJUČAK

Vizualizacija podataka kao jedan od koncepata poslovne inteligencije odavno je probila granice poslovnog svijeta te je postala dio svakodnevnice. Ljudski mozak percipira slike puno brže od suhoparnih tablica i brojeva. Donošenje odluka u modernim poslovnim uvjetima znatno je otežano pojavom informacijske krize (koja je objašnjena u drugom poglavlju ovog diplomskog rada). Vizualizacijom podataka, donositelj odluke može puno lakše spoznati veze i uzorke koji se kriju unutar ogromnih podatkovnih struktura, baza i skladišta. Do prije desetak godina za uspješnu vizualizaciju bilo je potrebno izdašno znanje i iskustvo u domeni strukture podataka i njihove analize. Tadašnji alati za vizualizaciju nisu imali grafičko sučelje pa je korisnik morao poznavati neki od programskih jezika kako bi uspješno prikazao podatke grafički. Moderna vizualizacijska rješenja vođena su pretpostavkom da je analiza podataka danas ključan poslovni proces i da više nije dovoljno da isključivo analitičari provode analizu. Ovakvim pristupom alati za vizualizaciju podataka približeni su svim zaposlenicima. Tržište alata za vizualizaciju trenutno je u „mrtvoj utrci“ između dva alata – Tableau i Power BI. Ovaj diplomski rad pokušava odgovoriti na dodatna pitanja vezana uz njihovu komparaciju. Provođenjem analize funkcionalnosti, empirijskim eksperimentom i intervjuiranjem ispitanika došlo se do nekoliko zaključaka.

Prvi zaključak je taj da je nemoguće uspoređivati ova dva alata funkcionalnost po funkcionalnost. To se prvenstveno odnosi na različite profile korisnika – Tableau je namijenjen analitičarima, a Power BI gotovo svima. Power BI je alat koji najbolje funkcionira unutar Microsoftovog ekosustava te je velik dio njegovih funkcionalnosti vezan za Microsoftovu arhitekturu. Tableau s druge strane zasebno rješenje čija je osnovna funkcija vizualizacija podataka. Ovisno o preferencijama korisnik bira rješenje koje mu odgovara. Ukoliko korisnik želi alat koji će se uklopiti u poslovni sustav baziran na Microsoftovoj tehnologiji onda je Power BI očiti izbor. Ako je fokus korisnika na vizualizaciji, naprednim analizama i fleksibilnosti integracije, tada je Tableau bolji izbor. Tableau omogućava i konektivnost sa više izvora podataka istovremeno, na istoj nadzornoj ploći. Power BI tu funkcionalnost nema, ali je kompatibilan sa analitičkim jezikom R što mu znatno širi funkcionalnost.

Cijena je također jedan od elemenata prema kojima su alati uspoređeni. Tableau je sa minimalno 35$ po korisniku skuplji od Power BI-a čiji komercijalni paket iznosi 9.99$ po korisniku. Ukoliko korisnik izbor temelji prema troškovima, Power BI je bolji izbor. Power BI nudi svoj alat besplatno nekomercijalnim korisnicima pa je popularan i među studentima. Tableau, komercijalno je znatno skuplji, ali ima program Tableau for Students preko kojega je jednostavno dobiti studentsku licencu u trajanju od godine dana. Tableau cjenovno targetira specijalizirane analitičke timove, dok Power BI cjenovnom strategijom penetrira tržište, ali ga i širi na ne-tehničke timove. Tableau je rješenje koje je duže na tržištu te je do sada oko njega okupljena veća online zajednica koja olakšava proces učenja korištenja alata raznim člancima, multimedijskim sadržajima i kursovima. Korisnik mora sagledati i ovaj element kod odabira adekvatnog alata.

Provedenim eksperimentom dokazano je kako Power BI efikasnije rješava jednostavne zadatke. Prihvaćene su hipoteze H1.1 i H1.2 kojima se pratilo vrijeme i udaljenost računalnog miša koja je potrebna za izvršavanje zadatka. Zadatak je postavljen heterogenoj skupini ispitanik od kojih su neki imali više a neki manje iskustva u analizi i vizualizaciji podataka. Svi ispitanici bolje rezultate su ostvarili korištenjem alata Power BI. Nakon izvršenog zadatka ispitanici su intervjuirani te je većina odabrala Power BI kao jednostavnije rješenje. Oni iskusniji ispitanici odabrali su Tableau kao jednostavnije rješenje, te je podjednak broj ispitanika izjavio kako bi u budućnosti nastavili koristiti Power BI, odnosno Tableau. Ovo istraživanje dokaz je kako Tableau i Power BI ne targetiraju iste korisnike. Kod odabira ovakvog alata važno je odrediti funkcionalnosti koje on treba sadržavati, pitanja na koje treba dati odgovore i korisnike koji će ga koristiti. Nemoguće je reći koje je bolje rješenje bez konteksta organizacije koja ga implementira. U zaključku su navedeni neki od kriterija koji bi trebali biti sagledani prilikom odabira adekvatnog alata.

SAŽETAK:

Informacija se u modernom poslovanju smatra resursom. Kvaliteta poslovne odluke u pozitivnom je korelacijskom odnosu sa kvalitetom dostupnih informacija. Poslovna inteligencija proces je prikupljanja relevantnih i dostupnih informacija informacija te se danas smatra jednom od osnovnih konkurentskih prednosti. Višedimenzijske podatkovne strukture temelj su moderne poslovne inteligencije. Analize provedene nad takvim podacima često generiraju više novih pitanja nego što daju odgovora. Zbog ove problematike, moderna analitička rješenja usmjerena su na vizualizaciju podataka kako bi korisnik što brže mogao saznati esenciju nekog niza podataka. Tržište je danas puno alata za vizualizaciju podataka, a ovaj diplomski rad komparira najpopularnija rješenja. U nizu alata odabrani su Tableau i Power BI kao tržišni predvodnici te su detaljno uspoređeni prema funkcionalnostima ali i empirijski. Ovaj diplomski rad pokušava dati odgovor na pitanje koje rješenje je najprimjerenije za određenog korisnika, odnosno organizaciju.
Ključne riječi: **Poslovna inteligencija, Vizualizacija podataka, Alati za vizualizaciju podataka, Tableau, Power BI**

SUMMARY:

Information is being considered as a resource in modern business. The quality of a business decision is in a positive correlation with the quality of available information. Business intelligence is a process of collecting relevant and accessible information and is one of the core competitive advantages in the modern market. Multi-dimensional data structures are the foundation of modern business intelligence. Analysis of such data often generates more questions than answers which creates a problem for the decision maker. Therefore, modern analytical solutions are more focused on data visualization than ever before. Using visual techniques, user can find out the essence of a certain data within seconds. There are a lot of tools that support data visualization in the modern market. Tableau and Power BI were chosen as market leaders, and were compared in detail regarding their functionalities. User experience experiment was also carried out to see which of these two performs better. This master thesis is trying to help users choose which visualization tool is the most appropriate for them or their organization.
Keywords: **Business intelligence, Data visualization, Data visualization tools, Tableau, Power BI**

LITERATURA:

1) "Business Analytics vs Business Intelligence?". timoelliott.com. 2011-03-09. Retrieved 2014-06-15.
2) "Frits H. Post, Gregory M. Nielson and Georges-Pierre Bonneau (2002). Data Visualization: The State of the Art.
3) Banek, Marko; Skočir, Zoran; Vrdoljak, Boris(2005): Logical Design of Data Warehouses from XML // Proceedings of the 8th International Conference on Telecommunications (ConTEL 2005)
4) Bilandžić, M. et al.: Business Intelligence u hrvatskom gospodarstvu, Poslovna izvrsnost, Zagreb, god. VI (2012), br. 1.
5) Bilandžić, M.: Poslovno-obavještajno djelovanje: Business Intelligence u praksi, AGM, Zagreb, 2008
6) Bižaca, R., Dobrović, V. (2016): Najbolje i najgore prakse vizualizacije podataka, Seminarski rad, EFST, Split
7) Brath, R., Peters, M. (2004.): Dashboard design: Why design is important. Data Mining Review/Data Mining Direct
8) Breeding, B.: CI and KM Convergence: A Case Study at Shell Services International
9) Bruce H. McCormick, Thomas A. De Fanti, Maxine D. Brown (1987) Visualization in ScientificComputing, The Workshop on Visualization inScientific Computing, 81 str., National Science Foundation
10) Washington D.C., veljaca 9-10 1987, ACM Press, New York USA
11) Codd, E.F. (1970). "A Relational Model of Data for Large Shared Data Banks". Communications of the ACM
12) Choudhury, S. (2013.): The Psychology Behind Information Dashboards, UX Magazine, dostupno na: https://uxmag.com/articles/the-psychology-behind-information-dashboards [11.1.2016.]
13) Ćukušić, M., Garača, Ž.(2013): Više dimenzijski informacijski sustavi, Sveučilište u Splitu, Ekonomski fakultet
14) Eckerson, W. (2009.): Designing Dashboards, TDWI, dostupno na: https://tdwi.org/blogs/tdwi-blog/2009/08/dashboards-to-die-for.aspx [11.1.2016.]
15) Evelson, Boris (21 November 2008). "Topic Overview: Business Intelligence"
16) Few, S. (2004.): Dashboard Confusion, Intelligent Enterprise, SAD
17) Few, S. (2012): Show Me the Numbers, Second Edition, Burligame California
18) Few, S. (2009): Now you see it, Analytics press, Oakland, California
19) Few, S. (2006.): Information dashboard design, O'Reilly, SAD
20) Fuld, L.M.: The secret language of competitive intelligence, 2nd. ed., Dog Ear Publishing, Indianapolis, 2010.
21) Friendly, Michael (2006). "A Brief History of Data Visualization" (PDF). York University. Springer-Verlag. Retrieved 2015-11-22.
22) Garača, Ž. (2008): Poslovni Informacijski Sustavi. Split: Ekonomski Fakultet Sveučilišta u Splitu
23) Garača, Ž (2007) Informatičke tehnologije, Split: Ekonomski Fakultet Sveučilišta u Splitu

24) H P Luhn (1958). "A Business Intelligence System" (PDF). IBM Journal. 2 (4): 314. doi:10.1147/rd.24.0314.
25) Imnon, Bill (1992): Building the Data Warehouse. 1st Edition. Wiley and Sons
26) http://www.storytellingwithdata.com/blog/2012/11/to-stack-or-not-to-stack [11.1.2016.]
27) Friedman (2008) "Data Visualization and Infographics" in: Graphics, Monday Inspiration,
28) Javorović, B., Bilandžić, M. (2007): „Poslovne informacije i business intelligence“, Golden marketing, Zagreb
29) Klepac, G, Mršić, L. (2006.): „Poslovna inteligencija kroz poslovne slučajeve“, Liderpress, Zagreb
30) Kudyba, S., Hoptroff, R.: Data Mining and Business Intelligence: A Guide to
31) Productivity, Idea Group Publishing, Hershey, 2001.
32) Lamza Maronić, M., Glavaš, J.: Poslovna inteligencija, pristupni rad, Sveučilište Josipa Jurja Strossmayera u Osijeku, Ekonomski fakultet u Osijeku, 2008.
33) Laursen, Gert H. N., Thorlund, J.: Business analytics for managers: Taking Business Intelligence Beyond reporting, John Wiley & Sons, New Jersey, 2010.
34) Luetić, A., Šerić, N.: Business Intelligence u funkciji upravljanja nabavnim lancem, Sveučilište u Splitu, Ekonomski fakultet, Split, 2009.
35) Liautaud, B., Hammond, M.: e-poslovna inteligencija, Prudens Consilium, Varaždin, 2006.
36) Ljubetić, V.: Upravljanje znanjem primjenom alata poslovne inteligencije, Sveučilište u Zagrebu, Ekonomski fakultet, Zagreb, 2005.
37) Mohanty, S.: Measuring the Value of Intelligence in Business Intelligence, SM Review, December 2008.,
38) Makeovers (2012.): To stack or not to stack, Makeovers, dostupno na:
39) Negash, S (2004). "Business Intelligence" (PDF). Communications of the Association of Information Systems. 13: 177–195.
40) Panian, Ž.; Klepac, G. (2004): „Poslovna inteligencija“, Masmedia, Zagreb
41) Panian, Ž.: *Odnosi s klijentima u E-poslovanju*, Sinergija nakladništvo d.o.o., Zagreb, 2003.
42) Pfeifer, S., Alpeza, M.: Business Intelligence, Sveučilište Josipa Jurja Strossmayera, Ekonomski fakultet u Osijeku, Osijek, 2011.
43) Pfeifer, S.: Od informacijskog podražaja do znanja, 7. međunarodni simpozij Informacijski sustavi - IS '96, Zbornik radova, Aurer, B., Hunjak, T. (ur.), Fakultet organizacije i informatike, Varaždin, 1996.
44) Pidnyck, R.,S., Rubinfeld, D., L. (2005): Mikroekonomija, Zagreb
45) Phua, C., Lee, V., Smith-Miles, K. and Gayler, R. (2005). A Comprehensive Survey of Data Mining-based Fraud Detection Research. Clayton School of Information Technology, Monash University.
46) Prescott, John E.: The Evolution of Competitive Intelligence – designing a process for action, Proposal Management, spring 1999
47) Rao, R. (2003). "From unstructured data to actionable intelligence" (PDF). IT Professional. 5 (6): 29.

48) Smolić, H. (2011): Kontroverze oko pita-grafa, Qualia, dostupno na: http://www.qualia.hr/kontroverze-oko-pita-grafa [11.1.2016.]
49) Smolić, H. (2014): Loše vizualizacije: zašto BI projekti propadaju?, Skladištenje.com, dostupno na: http://www.skladistenje.com/lose-vizualizacije-zasto-bi-projekti-propadaju/ [11.1.2016.]
50) Stevens, A. (2014.): Color Do's and Don'ts, Dundas, dostupno na: https://www.dundas.com/discover/article/color-dos-and-donts/ [11.1.2016.]
51) Strahonja, V. (1992): Projektiranje informacijskih sustava, Info, Zagreb
52) Tufte, E. R., Graves-Morris, P. R. (1983.): The visual display of quantitative information (Vol. 2, No. 9). Cheshire, CT: Graphics press
53) Vitt, E., Luckevich, M., Misner, S.: Business Intelligence, Microsoft Press, Redmond, 2002.
54) Julian, Taylor (10 January 2010). "Business intelligence implementation according to customer's needs"
55) Vera, D. (2013.): How to Future-Proof Your Dashboards, TDWI, dostupno na: https://tdwi.org/Articles/2013/06/04/Future-Proof-Dashboards.aspx [11.1.2016.]
56) Watson, H.J., Wixom, B.H., Buonamici, J.D., Revak, J.R.: Sherwin-Williams Datamart Strategy: Creating Intelligence Across the Supply Chain, Communications o the Association for Information Systems, 2001., 5 (9), May 2001
57) Zelenika, R. (2000): Metodologija i Tehnologija Izrade Znanstvenog i Stručnog Dijela. Rijeka: Ekonomski Fakultet Sveučilišta u Rijeci.

ONLINE IZVORI:

1. http://www.pcmag.com/roundup/346417/the-best-data-visualization-tools
2. https://www.tableau.com/gartner-magic-quadrant-2015
3. http://thenextweb.com/dd/2015/04/21/the-14-best-data-visualization-tools/
4. https://www.reddit.com/r/dataisbeautiful/
5. https://www.reddit.com/r/visualization/
6. http://blog.inteligencija.com/
7. https://powerbi.microsoft.com/en-us/
8. http://www.qlik.com/us/
9. http://saplumira.com/
10. https://www.quora.com/What-are-the-best-free-data-visualization-tools
11. http://www.tableau.com/compare-tableau-power-bi
12. https://www.reddit.com/r/BusinessIntelligence/comments/3qkpxh/tableau_vs_power_bi/
13. http://www.butleranalytics.com/microsoft-power-bi-vs-tableau/
14. https://www.linkedin.com/pulse/5-reasons-why-power-bi-taking-over-tableau-best-tool-tacoronte
15. https://conferences.oreilly.com/strata/strata2012
16. https://www.pcmag.com/article2/0,2817,2494375,00.asp

17. http://www.techradar.com/news/world-of-tech/microsoft-s-new-power-bi-a-very-different-way-of-doing-business-intelligence-1282485
18. http://www.streetdirectory.com/travel_guide/136735/enterprise_information_systems/the_role_of_business_intelligence_in_modern_business.html
19. http://ef.sve-mo.ba/sites/default/files/nastavni-materijali/DW_DS_07112016_2.pdf
20. http://www.streetdirectory.com/travel_guide/136735/enterprise_information_systems/the_role_of_business_intelligence_in_modern_business.html
21. http://ef.sve-mo.ba/sites/default/files/nastavni-materijali/DW_DS_07112016_2.pdf
22. https://d3js.org/
23. http://www.chartjs.org/
24. http://www.fusioncharts.com/
25. https://www.r-project.org/
26. http://www.reuters.com/
27. http://www.gartner.com/technology/home.jsp
28. https://financesonline.com/

POPIS SLIKA, TABLICA I GRAFIKONA:

Slike:

Slika 1. Transformacija podataka u informaciju 12
Slika 2. Shema odnosa relacija Klijent i Narudžba 21
Slika 3. Relacijska shema tablica 22
Slika 4. Koncept *data martova* unutar skladišta podataka 26
Slika 5. Primjer dvodimenzionalnog prikaza podataka 29
Slika 6. Prikaz trodimenzionalne strukture podataka 30
Slika 7. Trodimenzionalni vizualni prikaz podataka 30
Slika 8. Zvjezdasta vs. Pahuljasta shema podataka 32
Slika 9. OLAP Slicing 33
Slika 10. OLAP Dicing.....33
Slika 11. Primjer dashboarda za informacije o prodaji 53
Slika 12. Nadzorna ploča SAP Business Objects softvera 55
Slika 13. Oracle Dashboard – nadzorna ploča.....56
Slika 14. Gartnerov magični kvadrant za 2016. odnosno 2017. godinu.....61
Slika 15. Prikaz rješenja zadatka 75

Tablice:

Tablica 1. Tablica entiteta Klijent....20
Tablica 2. Tablica entiteta Narudžba....21
Tablica 3. Rješenje upita SELECT....23
Tablica 4. Rješenje upita SELECT u kombinaciji s upitom UNION....24
Tablica 5. Razlika Izvještaja, OLAPa i Rudarenja podataka....28
Tablica 6. Prodaja po zaposlenicima....38
Tablica 7. Usporedba alata za vizualizaciju podataka....68
Tablica 8. Usporedba Tableau-a i Power BI-a....73
Tablica 9. Profil sudionika istraživanja....76
Tablica 10 . Rezultati za alat Tableau....77
Tablica 11. Rezultati za alat Power BI....77
Tablica 12. Deskriptivna statistika / vrijeme....80
Tablica 13. Rezultati T- Testa / vrijeme....80
Tablica 12. Deskriptivna statistika / miš....81
Tablica 13. Rezultati T- Testa / miš....81

Grafikoni:

Grafikon 1. Podaci prikazani stupčastim grafikonom....38
Grafikon 2. Prikazuje odnos izvajanja za zdravstvo p.c. i očekivane životne dobi....41
Grafikon 3. Prikazuje odnos izdvajanja za zdravstvo i BDP-a po glavi stanovnika....42
Grafikon 4. Linearni, eksponencijalni, polinomski i logaritamski trend....43
Grafikon 5. Prodaja uredskog materijala po proizvodu (vertikalni i horizontalni)....44
Grafikon 6. 3D i 2D stupčasti grafikon....44
Grafikon 7. Socijalna pitanja u demokratskim i republikanskim objavama....45
Grafikon 8. Primjer slojevitog stupčastog grafikona....46
Grafikon 9. Usporedba stupčastog i linijskog grafikona....46
Grafikon 10. BDP p.c. Republike Hrvatske (2000 – 2010)....47
Grafikon 11. Linijski grafikon sa više prikazanih linija....47
Grafikon 12. Razlika u prikazu distinkcije različitih linija na istom grafikonu....48
Grafikon 13. Linijski i Površinski grafikon koji prikazuju iste podatke....48
Grafikon 14. Linija trenda prikazana na linijskom grafikonu....49
Grafikon 15. Pita grafikon, primjer dobrog i lošeg dizajna....50
Grafikon 16. Histogram prinosa na CROBEX portfelj za 2013. godinu....51
Grafikon 17. *Treemap* koji prikazuje BDP zemalja prikazanim po regijama....52