

# Primjena algoritama rudarenja podataka i umjetne inteligencije

---

Mihajlović, Marin

Master's thesis / Diplomski rad

2018

*Degree Grantor / Ustanova koja je dodijelila akademski / stručni stupanj:* **University of Zagreb, Faculty of Organization and Informatics / Sveučilište u Zagrebu, Fakultet organizacije i informatike**

*Permanent link / Trajna poveznica:* <https://urn.nsk.hr/urn:nbn:hr:211:109492>

*Rights / Prava:* [Attribution 3.0 Unported/Imenovanje 3.0](#)

*Download date / Datum preuzimanja:* **2024-09-04**



*Repository / Repozitorij:*

[Faculty of Organization and Informatics - Digital Repository](#)



**SVEUČILIŠTE U ZAGREBU**  
**FAKULTET ORGANIZACIJE I INFORMATIKE**  
**V A R A Ž D I N**

**Marin Mihajlović**

**Primjena algoritama rudarenja podataka i umjetne  
inteligencije**  
*Diplomski rad*

**Varaždin, 2018.**

**SVEUČILIŠTE U ZAGREBU**  
**FAKULTET ORGANIZACIJE I INFORMATIKE**  
**V A R A Ž D I N**

**Marin Mihajlović, 45272/16-R**

**Primjena algoritama rudarenja podataka i umjetne  
inteligencije**  
*Diplomski rad*

**Mentor:**

Doc. dr. sc., Dijana Oreški

**Varaždin, rujan 2018.**

# Sadržaj

1. Uvod.....	4
2. Metodologija rada.....	5
2.1. Korelacija.....	5
2.1.1. Pearsonov koeficijent korelacije .....	6
2.1.2. Spermanov koeficijent korelacije.....	7
2.2. Stablo odlučivanja.....	8
3. Prethodna istraživanja .....	10
4. Analiza navika kupaca .....	13
4.1. Ukupan iznos računa po grupama .....	15
4.2. Sumirani izvještaji iznosa računa po grupi .....	19
4.3. Ukupni iznos računa po danima u tjednu .....	24
4.4. Ukupan iznos računa po zaposleniku.....	28
4.5. Upozorenja o kupcima .....	32
4.6. Veza između ukupnog iznosa računa i ukupne uštede po računu.....	33
4.7. Određivanje vjerojatnosti budućeg iznosa računa .....	34
Zaključak .....	36
Literatura .....	38
Popis slika.....	40

# 1. Uvod

Podaci u današnje vrijeme donose sve veću vrijednost. Iz tog razloga se pristupa detaljnijoj obradi podataka s ciljem veće konkurentnosti na tržištu. Podaci dobiveni iz podataka mogu unaprijediti prodaju, ali i proizvodnju, te sveukupni dojam koji se ostavlja na kupce i slično. Velika količina podataka omogućuje različite analize i predviđanje budućeg kretanja prodaje i/ili potražnje.

Za analizu podataka postoji više metoda:[1]

1. Metode za utvrđivanje razlike između grupa entiteta
2. Metode za utvrđivanje relacija između varijabli
3. Faktorska analiza
4. Metode za grupiranje entiteta i varijabli
5. Određivanje efekta i snage testa

Kao za analizu postoji i veliki broj metoda za predviđanje i kreiranje buduće slike iz postojećih podataka. Neke od metoda predviđanja su:

1. Stablo odlučivanja
2. Neuronske mreže

U nastavku rada opisat će se metodologije rudarenja podataka i umjetne inteligencije s posebnim naglaskom na korelaciju i stablo odlučivanja. Nakon toga će se analizirati prethodna istraživanja navika kupaca, koja su na neki način povezani s analizom autora. Na kraju će biti prezentiran kreirani sustav i dobiveni rezultati analize i predviđanja kupovnih navika. Za predviđanje budućih aktivnosti kupaca u radu će se koristiti stablo odlučivanja.

## 2. Metodologija rada

U ovom dijelu rada autor dublje opisuje metode koje će koristiti u nastavku rada. Naravno postoje i druge metodologije, kao što je neuronska mreža. Općenito metoda je planiran ili unaprijed smišljen postupak za postizanje određenog ili praktičnog cilja.[2] Metode se razlikuju ovisno o području primjene, podacima nad kojima se izvršavaju metode, vrijednostima koje želimo dobiti od metode. Metode koje će u daljnjem tekstu biti opisane su:

- Korelacija
- Stablo odlučivanja

Treba staviti napomenu na važnost neuronske mreže kao metode predviđanja iz podataka. Umjetne neuronske mreže, koje se koriste u obradi podataka, služe rješavanju problema na području umjetne inteligencije.

### 2.1. Korelacija

Korelacija predstavlja stupanj međusobne povezanosti različitih varijabli. Vrste korelacije su: [3]

- Linearna
- Nelinearna

Koeficijent korelacije bila bi vrijednost koja pokazuje vrijednost i smjer dvije varijable. Koeficijent može biti u rasponu od -1 do 1. Korelacija može biti pozitivna maksimalna korelacija, pozitivna nepotpuna korelacija, nepostojeća korelacija, negativna nepotpuna korelacija i negativna maksimalna korelacija.[4]

Pozitivna maksimalna korelacija je ako linearnom porastu jedne varijable odgovara linearni porast druge varijable tako da je jedna određena vrijednost jedne varijable uvijek povezana s jednom korespondentnom vrijednošću druge varijable. Vrijednost korelacije je +1.

Pozitivna nepotpuna korelacija je ako linearnom porastu jedne varijable uglavnom odgovara porastu druge varijable. Vrijednost korelacije je između 0 i +1.

Nepostojanje korelacije je ako iz određene vrijednosti jedne varijable ne možemo ništa zaključiti na vrijednost druge varijable. Vrijednost korelacije je 0.

Negativna nepotpuna korelacija je ako linearnom porastu jedne varijable odgovara linearno opadanje druge varijable, ali je povezanost takva da je jedna vrijednost jedne varijable povezano s više vrijednosti druge varijable. Vrijednost korelacije je između -1 i 0.

Negativna potpuna korelacija je ako linearnom porastu jedne varijable odgovara linearni pad druge varijable i to tako da je jedna određena vrijednost jedne varijable povezana s jednom korespondentnom vrijednosti druge varijable.

Postoji više vrsta koeficijenata korelacije, a neki od njih su:

- Pearsonov koeficijent korelacije
- Spermanov koeficijent korelacije
- Parcijalna korelacija
- Fi koeficijent

Najpoznatiji i najčešće korišteni koeficijenti korelacije su Pearsonov i Spermanov koeficijent korelacije.

### 2.1.1. Pearsonov koeficijent korelacije

Personov koeficijent korelacije se koristi kada između varijabli promatranog modela postoji linearna povezanost i neprekidna normalna distribucija, a kao što je definirano za koeficijent korelacije, Pearsonov koeficijent korelacije može imati vrijednosti između -1 i 1.

Funkcije prema kojoj se računa Personov koeficijent korelacije glasi:

$$r_{xy} = \frac{\sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y})}{\left[ \sum_{i=1}^n (x_i - \bar{x})^2 \sum_{i=1}^n (y_i - \bar{y})^2 \right]^{1/2}} [5]$$

U brojniku nalazi se suma razlike umnoška svake varijable x i srednje vrijednosti varijable x, te svake varijable y i srednje vrijednosti varijable y. U nazivniku nalazi se umnožak kvadrata razlike varijable x i varijable y.

Primjer pozitivne maksimalne korelacije računata Personovim koeficijentom korelacije je potrošnja vode prema broju stanovništva nekog mjesta. Ako se linearno povećava broj stanovnika nekog mjesta pretpostavljamo da linearno raste i ukupna potrošnja vode na području tog mjesta.

Nepostojana korelacija računata Personovim koeficijentom korelacije dobro bi bila prikazana računanjem koeficijenta za potrošnju vode i broj dućana u mjestu. Možemo pretpostaviti da koliko god se troši vode u naselju neće utjecati na broj dućana u istom mjestu.

Negativna maksimalna korelacija računata Personovim koeficijentom korelacije jednostavno bi bilo shvatiti proučavanjem broja stanovnika i broja slobodnih stanova na nekom području. Što je veći broj stanovnika na nekom području pretpostavka je da je manji broj slobodnih stanova na tom području, i obrnuto.

### 2.1.2. Spermanov koeficijent korelacije

Spermanov koeficijent korelacije je neparametrizirani Personov koeficijent korelacije, odnosno mjeri snagu i smjer dvije povezane rangirane varijable. Za razliku od Personovog koeficijenta korelacije, Spermanov koeficijent omogućuje računanje monotonog odnosa između dvije varijable.

Monotonost može podrazumijevati više vrsta funkcija[6]

- Za funkciju  $y = f(x)$  kažemo da je monotonno rastuća u intervalu od  $a$  do  $b$  podskup domene funkcije  $f$  ako vrijedi:  $f(x_1) \leq f(x_2)$  za svaki  $x_1, x_2$  u intervalu od  $a$  do  $b$  i  $x_1 < x_2$
- Za funkciju  $y = f(x)$  kažemo da je monotonno padajuća u intervalu  $a, b$  iz domene  $f$  ako vrijedi:  $f(x_1) \geq f(x_2)$  za svaki  $x_1, x_2$  element skupa  $a, b$  i  $x_1 < x_2$

Za računanje Spermanovog koeficijenta korelacije potrebno je dodijeliti rang svakoj vrijednosti skupa od manje vrijednosti skupa prema manjoj. Rang vrijednosti poprima vrijednosti od 1 do broja elemenata skupa za koji se računa koeficijent korelacije. Nakon što se svakoj vrijednosti dodijeli rang, računamo kvadrat ranga. Zatim prema slijedećoj formuli računamo korelaciju između varijabli:

$$\rho = 1 - \frac{6 \sum d_i^2}{n(n^2-1)} [7]$$

- $n$  – broj redova za koje se računa rang
- $d_i$  – rang
- $d_i^2$  – kvadrat ranga



## 2.2. Stablo odlučivanja

Stablo odlučivanja je jedan od najčešće korištenih metoda rudarenja podataka koja se koristi za analizu i modeliranje. Primjenjuje se za razvrstavanje, predviđanje, procjenu vrijednosti, grupiranje, opisivanje podataka i vizualizaciju.[8]

Stablo odlučivanja koristi se u:[9]

- Analizi odlučivanja
- Strojnom učenju

Stablo odlučivanja u analizi odlučivanja koristi se kako bi vizualno predstavila način donošenja odluka od strane eksperata. Takvi sustavi omogućuju prijenos znanja s eksperata na ostatak organizacije, kako bi ostali dijelovi organizacije imali mogućnost donošenja odluka.

Stablo odlučivanja u strojnom učenju su prediktivni modeli koji na temelju podataka izvode njihove veze u cilju dobivanja izlaznih vrijednosti. Stabla odlučivanja u strojnom učenju koriste se za pronalaženje skrivenih veza među podacima.

Kod kreiranja stabla odlučivanja koriste se samo obilježja koja se mogu kvantificirati. Što znači da se odabiru samo numeričke vrijednosti ili ne numeričke vrijednosti koje se mogu kvantificirati.

Analiza stabla odlučivanja temelji se na metodi povratne indukcije. [8] Povratna indukcija, u teoriji igara, je iterativni proces razmišljanja u vremenu, od kraja problema ili situacije, za rješavanje konačnih opsežnih oblika i sekvencijalnih igara i zaključuje niz optimalnih akcija. [10] Za analizu stabla ova analiza predstavlja kretanje od kraja stabla, te logičko zaključivanje prethodnog čvora stabla odlučivanja. Izvršavanjem analize dobiva se vrijednost stabla odlučivanja i njegova vjerodostojnost.

Prednosti stabla odlučivanja:[8]

- Rad s nedostajućim vrijednostima
- Pronalaženje zamjene za svaku značajku

Nedostatak stabla odlučivanja je njegova nestabilnost.[8] Nestabilnost znači da se malom promjenom ulaznih vrijednosti, stablo odlučivanja daje drugačije vrijednosti, odnosno postojeće stablo odlučivanja je potrebno promijeniti.

Jedan od najpopularnijih algoritama za kreiranje stabla odlučivanja je ID3. Slijedi postupak koji definira algoritam:[11]

- Testira se svaki atribut da se ocijeni kako dobro klasificira primjere
- Najbolji se odabire kao čvor, a njegove vrijednosti su silazne grane
- Primjeri za učenje se sortiraju prema odgovarajućem silaznom čvoru
- Cijeli postupak se ponavlja koristeći primjere koji su dodijeljeni silaznom čvoru

Algoritam C4.5 je proširenje algoritma ID3, a postupci njegovog provođenja su:[11]

- Nauči stablo odluke iz skupa za učenje sve dok svi podaci ne pristaju dobro
- Pretvori stablo u ekvivalentni skup pravila stvarajući jedno pravilo za svaku stazu od korijena do lista
- Podrezuj pravila uklanjajući bilo koji preduvjet koji rezultira u poboljšanju procijenjene točnosti
- Složi podrezana pravila po procijenjenoj točnosti i razmataj ih u tom nizu kod klasificiranja primjera

### 3. Prethodna istraživanja

Kao uvod u istraživanje proučeno je nekoliko sličnih istraživanja. U jednom u kojem se proučava utjecaj emocija na kupovne navike kupaca, gotovo 28% kupaca ide u kupovinu jednom dnevno, dok još 22% kupuje i više od jednom dnevno. Najveći dio kupaca kupuje 2-6 puta tjedno i to 47% ispitanika. Impulzivna kupovina je nagonska i neplanirana kupovina koja je rezultat emocije i određenog impulsa.[11] Samo 9% ispitanika sebe smatra impulzivnim kupcem, dok 22% ispitanima izrazito negira svoje kupovine kao impulzivne. 50% ispitanika potvrdilo je da su barem jednom kupili nešto iz razloga jer im je ponuđena dobra prilika u prodajnom procesu. Zanimljivo je da čak 31% ispitanika nije kupilo na promidžbi, odnosno da ih takav pristup odbija. Bez obzira na deklaraciju s početka o tome jesu li ili nisu impulzivni kupci, 89% ispitanika reklo je da barem jednom godišnje kupe proizvod na temelju emocija. Također, među ispitanima je jako malo kupaca koje privlače retro izdanja nekog proizvoda, a najviše je onih koje takva izdanja proizvoda odbijaju. U istraživanju je otkriveno da kupci više biraju proizvod koji je snižen, nego proizvod koji ima istu cijenu, ali je veće pakiranje. Skoro svi ispitanici ovog istraživanja osjećaju se bolje nakon kupovine, odnosno odlaze u kupovinu ako se loše osjećaju da si poprave raspoloženje. U ukupnoj procjeni kupci nisu sigurni je li glazba u prodavaonici utječe na njihovo raspoloženje i na konačni ishod kupovine. Najviše kupaca kupuje sa srednjih polica dućana. Većem broju kupaca bitniji je proizvođač nekog proizvoda i imaju omiljene proizvođače čije proizvode kupuju bez obzira na cijenu proizvoda u odnosu na konkurentne proizvođače. Ako je prethodno spomenuti proizvod na sniženju, 86% ispitanika kupuje više od jednog takvog proizvoda. U ispitanjoj grupi, najveću emociju budi kupovina namještaja, a najmanju kupovina hrane i pića. Istraživanje je dokazalo da iako se kupci većinom ne definiraju kao impulzivnim, velika većina je impulzivna barem u nekim slučajevima.[13]

U drugom istraživanju[14] kupci su podijeljeni u tri skupine: racionalni kupci, dijelom racionalni kupci i impulzivni kupci. Impulzivni kupci su većinom kupci iz mlađe dobne skupine, dok u druge dvije skupine spadaju stariji kupci. Studenti su skloniji impulzivnoj kupovini, dok su umirovljenici najrjeđe u statusu impulzivnog kupca. Impulzivnu skupinu kupaca većinom čine žene, dok su u ostale dvije skupine kupaca podjednako raspoređene osobe i ženskog i muškog spola. Obrazovanje i prihodi u nikakvoj mjeri nisu povezani s impulzivnošću. Što dijelom začinjuje

jer bi mogli zaključiti da kupci s većim prihodima kupovini pristupaju impulzivnije, a kupci s manjim prihodima racionalnije.

Treće proučavano istraživanje bavi se utjecajem oglasa na internetu da kupovne navike kupaca.[15] Veći broj ispitanika mišljenja je da bi oglasi trebali biti van glavnog sadržaja na internetu (60%), dok ostali misle da bi isti oglasi trebali biti na glavnom sadržaju (40%). Oglasi na internetu u 75% slučajeva ispitanicima odvlače pažnju sa glavnog sadržaja stranice na kojoj se prikazuje oglas. Većina osoba koje vide reklamu na internetu, ne reagiraju na istu i ona ne izaziva kupovinu reklamiranog proizvoda. Oni kupci koji uoče oglas, u 66% slučajeva traže dodatne informacije o oglasu i podacima prikazanim u oglasu. Zanimljivo je da 50% ispitanika smatra da je najbolje mjesto za oglas u sredini web lokacije, a s obzirom da su u početku izjavili da ne treba biti oglas u glavnom sadržaju, ova dva rezultata su upitna. Također, cijelo istraživanje treba uzeti s mogućnosti velike pogreške jer je rađeno na samo 54 ispitanika.

Slijedeće istraživanje vezano je uz kupovne navike na koje utječu promocije u supermarketima.[16] Veću sklonost kupovini na promociji pokazali su kupci koji su planirali kupiti barem jedan od promoviranih proizvoda. Veću sklonost kupovini na promociji pokazali su lojalni kupci, od ne lojalnih kupaca. S obzirom da ovo istraživanje donosi određene zakluče vezane za lojalnost, zanimljivo je istraživanje za rezultate u nastavku rada.

Posljednje proučavano istraživanje odnosi se na čimbenike ponašanja potrošača prilikom donošenja odluke o kupnji luksuznih kozmetičkih proizvoda.[17] Pola ispitanika je u starosti od 18-24 godine, a još 37% ispitanika je u dobi od 25-34 godine. Također, što je i logično prema temi, veći dio ispitanika su ženskog spola 60%. Više od pola ispitanika raspoloživi dohodak ima između 2400 kn i 5000 kn. Skoro svi ispitanici najčešće kupuju u fizičkim prodavaonicama, dok samo mali dio ispitanika, oko 10%, kupuje preko interneta i kataloga. Skoro 80% potrošača ponekad se raspita o proizvodu prije donošenja odluke o kupnji, što je veliki broj koji predstavlja opreznost prije kupnje pojedinog proizvoda. Iako se raspituju o proizvodnji, najveću ulogu u donošenju odluke o kupovini ima cijena. Skoro svi kupci, odnosno njih 98% nakon kupovine se osjeća zadovoljno, što znači da je ova skupina ljudi prema klasifikaciji iz prethodnih ispitivanja, skupina impulzivnih kupaca. Skoro 70% ispitanika ne prati ili rijetko prati trendove u kozmetičkoj industriji. S obzirom i na rezultate prethodnog pitanja, samo 10% ispitanika kupuje luksuzne kozmetičke proizvode. Prilikom kupovine kozmetičkih proizvoda kupcima najveću važnost

predstavlja da je proizvod pouzdan i ne izaziva reakcije na koži, dok im manju važnost predstavlja da iste marke moraju biti i šampon za kosu i gel za tijelo. Skoro svi ispitanici ne troše velike iznose na luksuzne kozmetičke proizvode. S obzirom da ispitivanje obuhvaća mali segment ispitanika koji kupuju luksuzne kozmetičke proizvode, relevantnost istraživanja dovodi se u pitanje jer u njemu nisu sudjelovali kupci koji kupuju i konzumiraju takve proizvode.

## 4. Analiza navika kupaca

Za analizu kupovnih navika koriste se računi jednog lanca trgovina robe široke potrošnje. Svaki račun sadržava slijedeće podatke:

- id – jedinstveni identifikator računa u bazi podataka
- invoiceId – jedinstveni identifikator računa koji se ispisuje na svakom računu
- employeeCode – šifra prodavača koji je izdao račun
- datetime – datum i vrijeme kada je izdan račun
- payment – način plaćanja računa (gotovina ili kartica)
- cardNumber – broj kartice kupca
- merchantTaxNumber – oib prodavača
- branchDescription – naziv i adresa poslovnice
- totalAmount – ukupan iznos računa
- totalDiscount – ukupan iznos uštede po računu

Kupci su podijeljeni u 4 grupe:

- Kupci u starosti od 18 do 29 godina
- Kupci u starosti od 30 do 45 godina
- Kupci u starosti od 46 do 60 godina
- Ostali (kupci mlađi od 18 godina i stariji od 60 godina)

Osnovne analize kupovnih navika grupa su:

- Ukupan iznos računa po grupama
- Sumirani izvještaj iznos računa pregled po grupi
- Ukupan iznos računa po zaposleniku
- Ukupan iznos računa svake grupe po danu u tjednu
- Ukupan iznos uštede po grupama
- Obavijesti o kupcima

Predviđanje kupovnih navika grupa:

- Povezanost ukupnog iznosa računa i ukupne uštede po računu
- Procjena vrijednost budućeg iznosa računa

Za sve izvještaje računa se minimum, medijan, aritmetička sredina i maksimum.

Minimum je najmanja vrijednost skupa.

Medijan je vrijednost središnjeg podatka koja podatke poredane po veličini dijeli dijelu u dva jednaka brojna dijela. Ako je broj podataka neparan medijan je vrijednost središnjeg dijela, a ako je broj podataka paran medijan predstavlja srednju vrijednost dva središnja podatka.

Aritmetička sredina ili srednja vrijednost je omjer zbroja svih vrijednosti i broj vrijednosti numeričke varijable. Jednostavna aritmetička sredina se primjenjuje kod negrupiranih podataka. Vagana (ponderirana) aritmetička sredina se koristi kod grupiranih podataka. U radu se koristi jednostavna aritmetička sredina.

Maksimum je najveća vrijednost skupa.

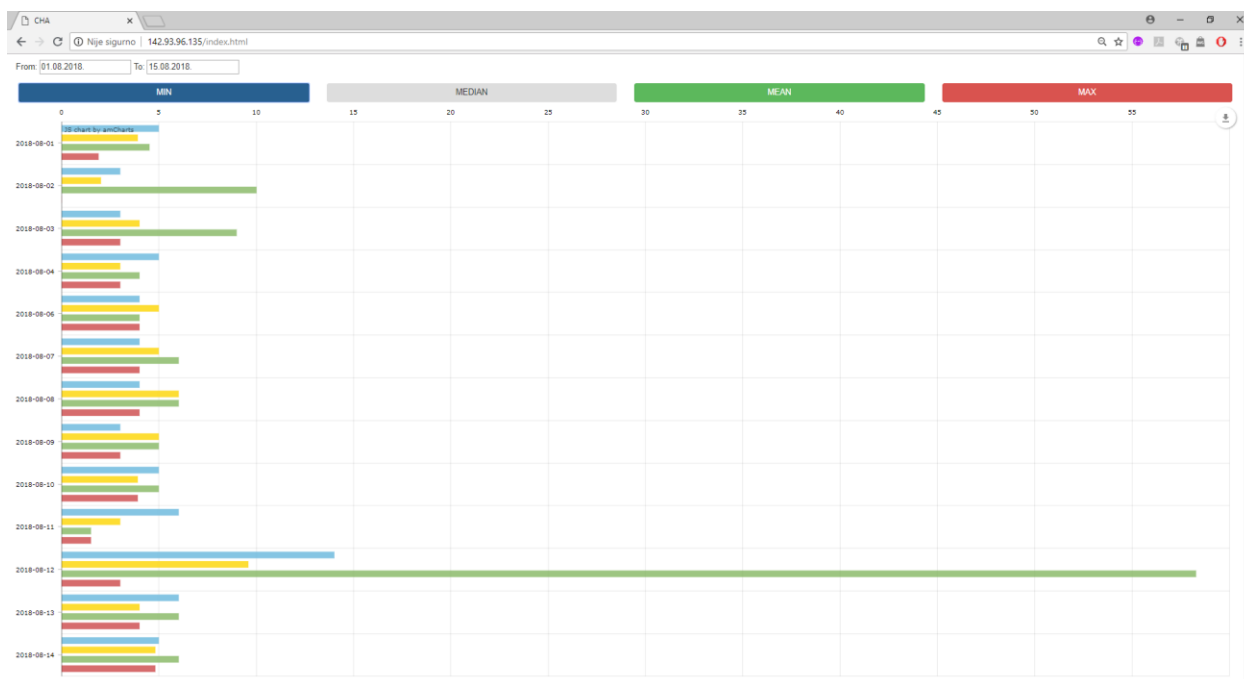
Kao pokazatelji kupovnih navika pojedine grupe kupaca veći značaj imaju mjere medijan i aritmetička sredina, od mjera minimuma i maksimuma. Mjere minimum i maksimum značajne su za otkrivanje ekstrema prema kojima se želi težiti ili koje se želi izbjegavati.

## 4.1. Ukupan iznos računa po grupama

Ukupan iznos računa po grupama prikazuje se za odabrani vremenski period. Unutar odabranog vremenskog perioda za svaki pojedini dan se prikazuju četiri vrijednosti, po jedna vrijednost za svaku grupu kupaca. Analiza u periodu od 01.08.2018. do 14.08.2018.

Legenda:

- Plavo – od 18 do 29 godina
- Žuto – od 30 do 45 godina
- Zeleno – od 46 do 60 godina
- Crveno – mlađi od 18 ili stariji od 60 godina



Slika 1: Minimalan iznos računa

U najvećoj mjeri minimalni iznosi se kreću do 5 kuna za svaku grupu korisnika. Na slici 1 vidljiv je veliko izdvajanje od prosjeka na dan 12.08.2018. da minimalna vrijednost zelene grupe je puno veća od bilo koje druge vrijednosti, a iznosi 58,3 kune. Zanimljivo je da su i plava i žuta grupa na isti dan veće nego ostali dani.

Gotovo svakog dana najmanju vrijednost minimalnog iznosa računa ima crvena grupa, dok su ostale grupe približno jednake.





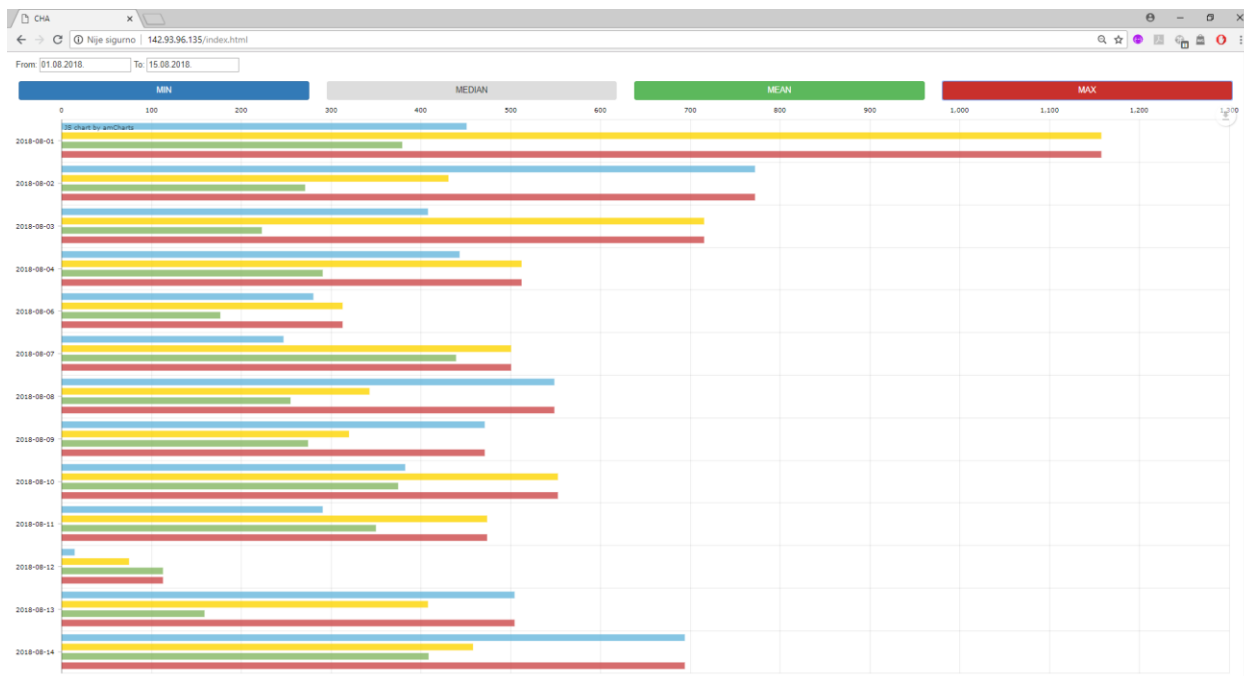
Slika 2: Medijan iznos računa

Pod pretpostavkom da bi medijan zelene grupe mogao biti najveći dana 12.08.2018. jer je tada bi i najveći minimum zelene grupe, na slici 2 potvrđuje se pretpostavka. Medijan zelene grupe iznosi 63,5 kune. Zanimljivo je da iako je dana 12.08.2018. minimum plave grupe bio iznadprosječan, medijan plave grupe istog dana je ispodprosječne vrijednosti i iznosi 14 kuna. U ostalim danima plava grupa im najveći mediana praktički u svim danima osim 04.08.2018. kada s 62,8 kuna dominira medijan iznosa računa žute grupe.



Slika 3:Aritmetička sredina iznosa računa

Prosječne dnevne vrijednosti aritmetičkih sredina iznosa računa grafički su slične medijanu iznosa računa. Objašnjenje je da u tim skupovima nema ekstremno malih niti ekstremno velikih vrijednosti koje utječu na aritmetičku sredinu skupa. Ponovno dana 12.08.2018. zelena grupa ima najveću prosječnu vrijednost iznosa računa, ali ta vrijednost niti nije najveća u cijelom periodu, niti izudara iz skupa kao kod minimuma.



Slika 4: Maksimum iznosa računa

Vidljivo je da maksimum sadrži puno više ekstremnih vrijednosti od minimuma, skoro svaki dan po vrijednosti iskače barem jedna grupa. Žuta grupa iskače čak dva puta, u tri dana, a najznačajniji maksimum je 1157,5 kuna na dan 01.08.2018. Jednak maksimum istog dana ima i crvena grupa. To je vrijednost maksimuma koja je daleko najveća u cijelom odabranom periodu. Najmanju minimalnu vrijednost ima plava grupa na dan 12.08.2018. s iznosom od 14 kuna.

Zelena grupa koja je imala daleko najveću vrijednost minimuma dana 12.08.2018., maksimum joj je najveći za taj dan, ali daleko od ostalih ekstremnih maksimalnih vrijednosti. Maksimum zelene grupe iznosi 112,4 kune isto kao i crvene grupe taj dan.

## 4.2. Sumirani izvještaji iznosa računa po grupi

Za odabrani vremenski period (od 01.08.2018. do 15.08.2018.) prikazuju se deskriptivni podaci vrijednosti za odabranu grupu kupaca. Deskriptivni podaci koji se prikazuju su:

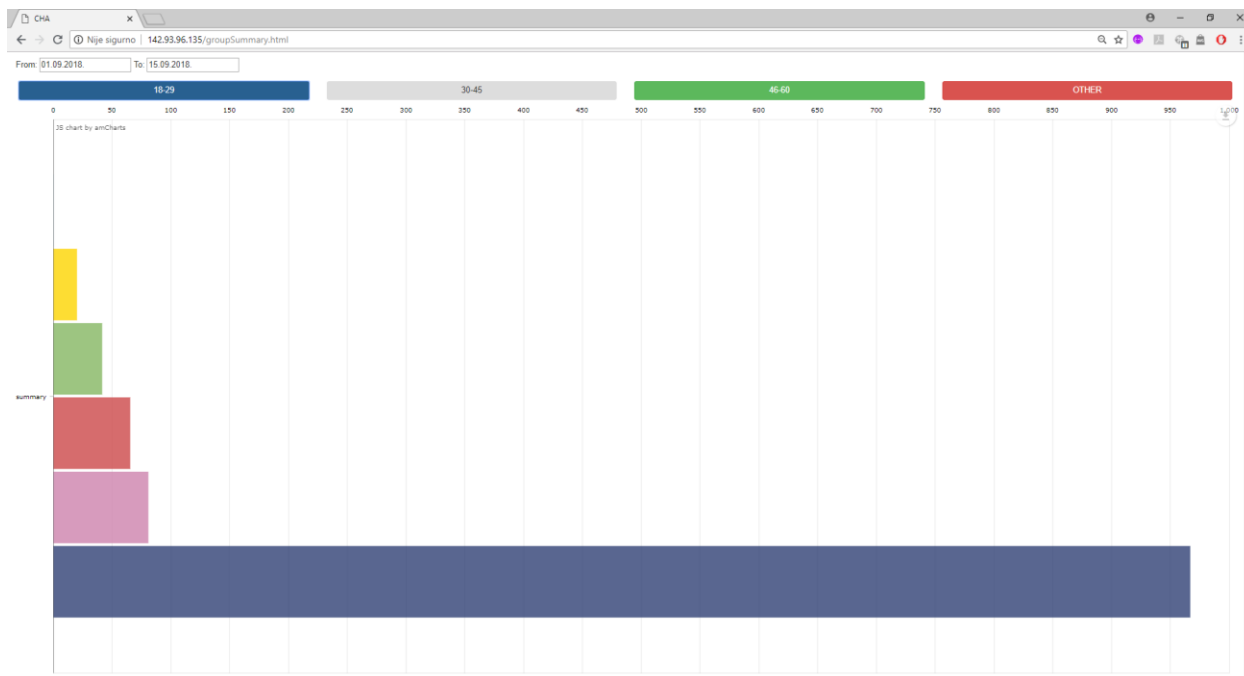
- Svjetlo plavo – minimalna vrijednost iznosa računa
- Žuto – donji kvartil vrijednosti iznosa računa
- Zeleno – medijan vrijednosti iznosa računa
- Crveno – srednja vrijednost iznosa računa
- Roza – gornji kvartil vrijednosti iznosa računa
- Tamnoplavo – maksimalna vrijednost iznosa računa

Kvartili su vrijednosti statističkog obilježja koje statistički niz dijele na četiri jednaka dijela. Mogu se podijeliti na:

- Donji kvartil  $Q_1$
- Gornji kvartil  $Q_3$

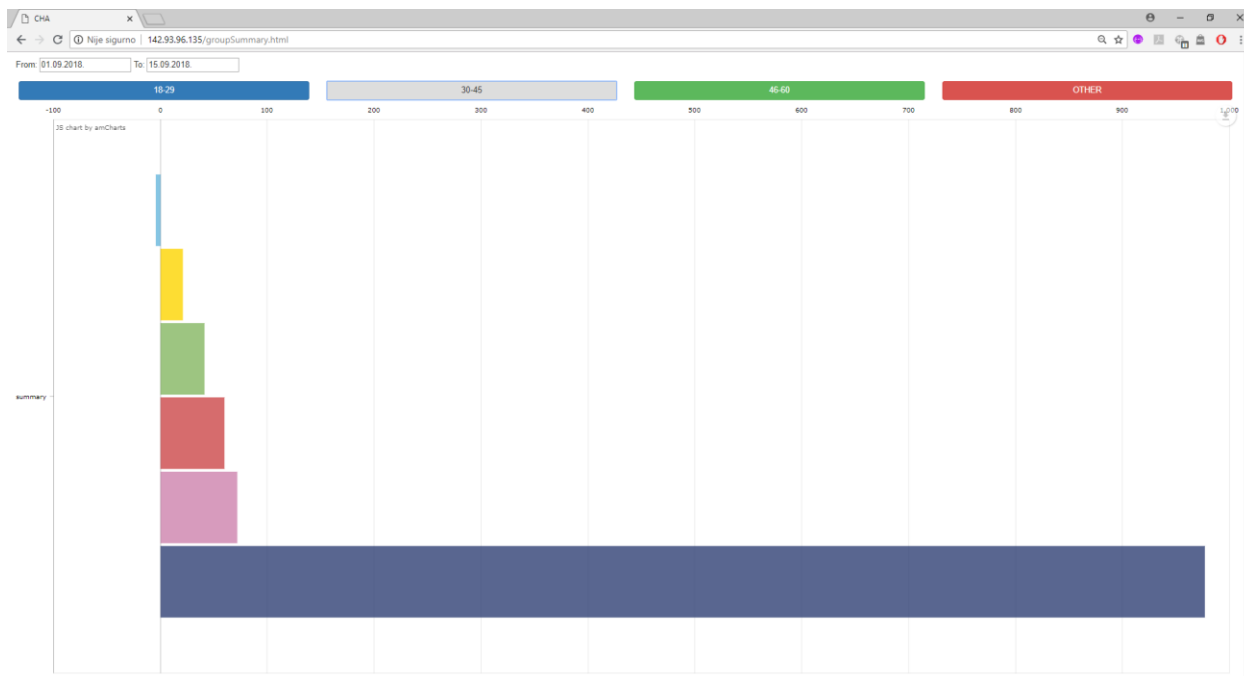
Donji kvartil dijeli statistički niz na četiri jednaka dijela u omjeru 1:3, odnosno 25% elemenata statističkog skupa ima vrijednost obilježja manju od donjeg kvartila, 75% elemenata statističkog skupa ima vrijednost obilježja veću od donjeg kvartila.

Gornji kvartil dijeli statistički niz na četiri jednaka dijela u omjeru 3:1, odnosno 75% elemenata statističkog skupa ima vrijednost obilježja manju od gornjeg kvartila, 25% elemenata statističkog skupa ima vrijednost obilježja veću od gornjeg kvartila.



Slika 5:Grupni izvještaj grupe 1

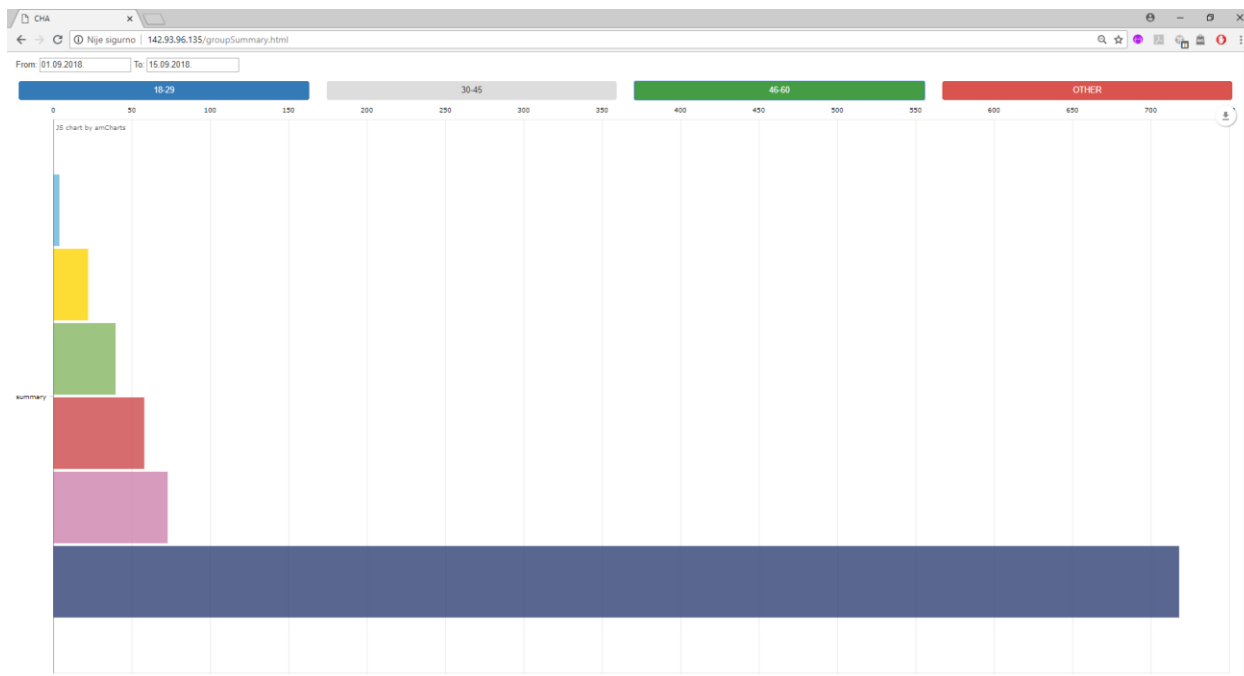
Grupa kupaca od 18 do 29 godina, u odabranom periodu nema minimalnu vrijednost računa, odnosno minimalna vrijednost računa je 0, a taj iznos na ništa ne govori. Vrijednost prvog kvartila je 20 kuna, a mediana iznosa računa 41,3 kune. Aritmetička sredina vrijednosti računa za prvu grupu je 65 kuna. Može se reći da svaka prosječna kupovina kupaca iz ovog skupa je oko 50 kuna. Treći kvartil iznosa je 80 kuna što i je većina računa. Odnosno, rijetke će biti vrijednosti računa veće od 80 kuna, za ovu grupu korisnika. Maksimalna vrijednost iznosi 966,64 kune što po grafu i predstavlja izoliranu maksimalnu vrijednost.



Slika 6:Grupni izvještaj grupe 2

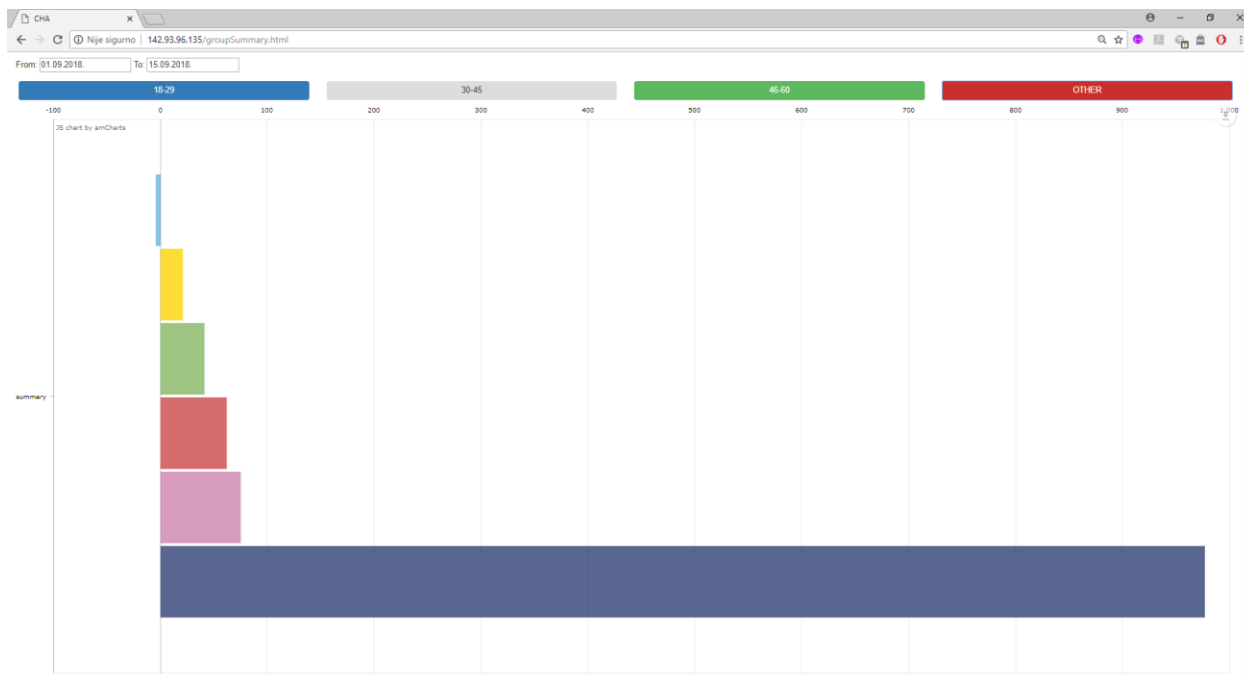
Minimalna vrijednost ove grupe je -4,5 kune. Negativna vrijednost predstavlja storno računa, te negativne vrijednosti treba ignorirati. Storno računa nije kupovina, nego povrat robe iz nekog razloga.

Iznos prvog kvartila je 21 kuna, medianana 41, a aritmetičke sredine 59 kuna. Iznosi ove tri mjere prilično su slične vrijednostima grupe kupaca od 18 do 29 godina. Treći kvartil iznosi 71 kunu, što je za 9 kuna manje od prethodne grupe, tako da može se reći da je prosječna kupovina kupaca iz grupe 18 do 29 godina veća od kupovine kupaca od 30 do 45 godina. Izolirana maksimalna vrijednost iznosi 976 kuna.



Slika 7:Grupni izvještaj grupe 3

Minimalna kupovina grupe kupaca od 46 do 60 godina je 3,9 kuna, što možemo pretpostaviti da je vrijednost jednog artikla. Vrijednost prvog kvartila je 22 kune, medijana 40, a aritmetičke sredine 58 kuna. I ova grupa predstavlja slabiju grupu kupaca od grupe kupaca od 18 do 29 godina. Ova grupa je približno jednaka kao grupa kupaca od 30 do 45 godina. Treći kvartil grupe je na 72,85 što je također slično vrijednosti prethodne grupe. Iznos izolirane maksimalne vrijednosti je 718 kuna.



Slika 8: Grupni izvještaj grupe 4

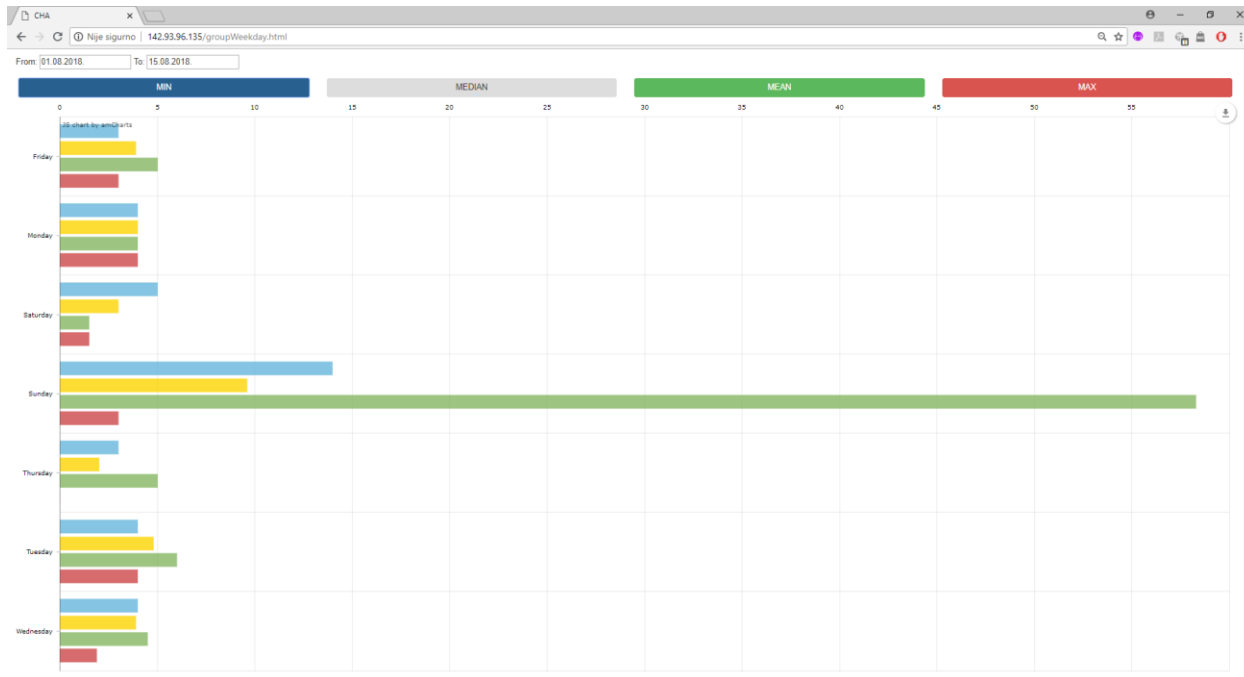
Na slici 8 prikazan je grupni izvještaj za grupu korisnika mlađih od 18 godina i starijih od 60 godina. Prvi kvartil ima vrijednost 21 kunu, medijan 41 kunu, a aritmetička sredina 61,62 kune. I ova grupa korisnika prilično je po kupovnim navikama slična grupi 2 i 3. Treći kvartil iznosi 75 kuna, a izolirani maksimum 976 kuna. Iako je maksimum za ovu grupu najveći, ona predstavlja prosječnu grupu.

Kao najbolja grupa u odabranom vremenskom periodu pokazuje se grupa kupaca od 18 do 29 godina, koja otprilike vrijednost trećeg kvartila ima veću od 5 do 10 kuna od ostalih grupa kupaca.



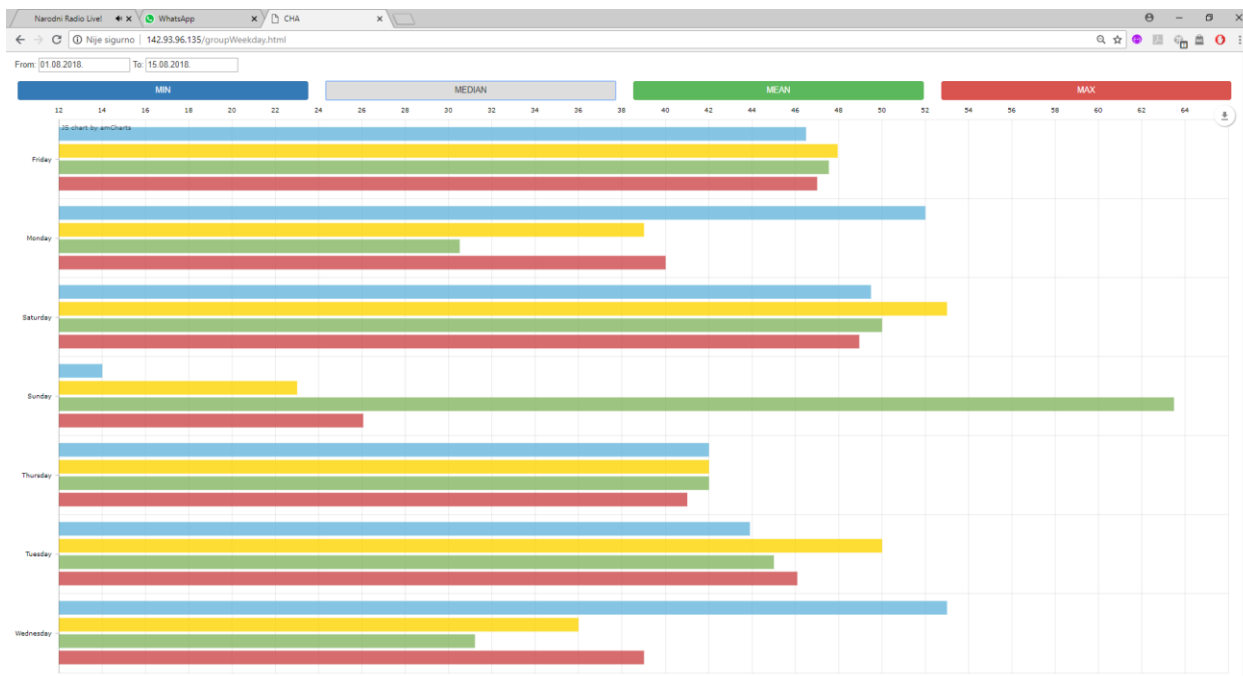
### 4.3. Ukupni iznos računa po danima u tjednu

Ovom analizom prikazuje se ukupan iznos računa u periodu o 01.08.2018. do 15.08.2018. za pojedini dan. Na grafikonu su prikazane po danu u tjednu sve grupe.



Slika 9: Minimalni iznosi po danu

Minimalna vrijednost iznosa računa po svakom radnom danu u zadanom vremenskom periodu u prosjeku je do 5 kuna za svaku grupu kupaca. Minimalna potrošnja od prosjeka iskače nedjeljom. Nedjeljom za grupu kupaca od 18 do 29 godina starosti (plavi stupac) iznosi 14 kuna, za grupu kupaca od 30 do 45 godina starosti (žuti stupac) 9,6 kuna, a za grupu kupaca od 46 do 60 godina starosti (zeleni stupac) 58,3 kuna. Prema minimalnim iznosima možemo očekivati za ovu skupinu kupaca i u ostalim kategorijama iskakanje u vrijednosti nedjeljom.



Slika 10: Medijan iznosi po danu

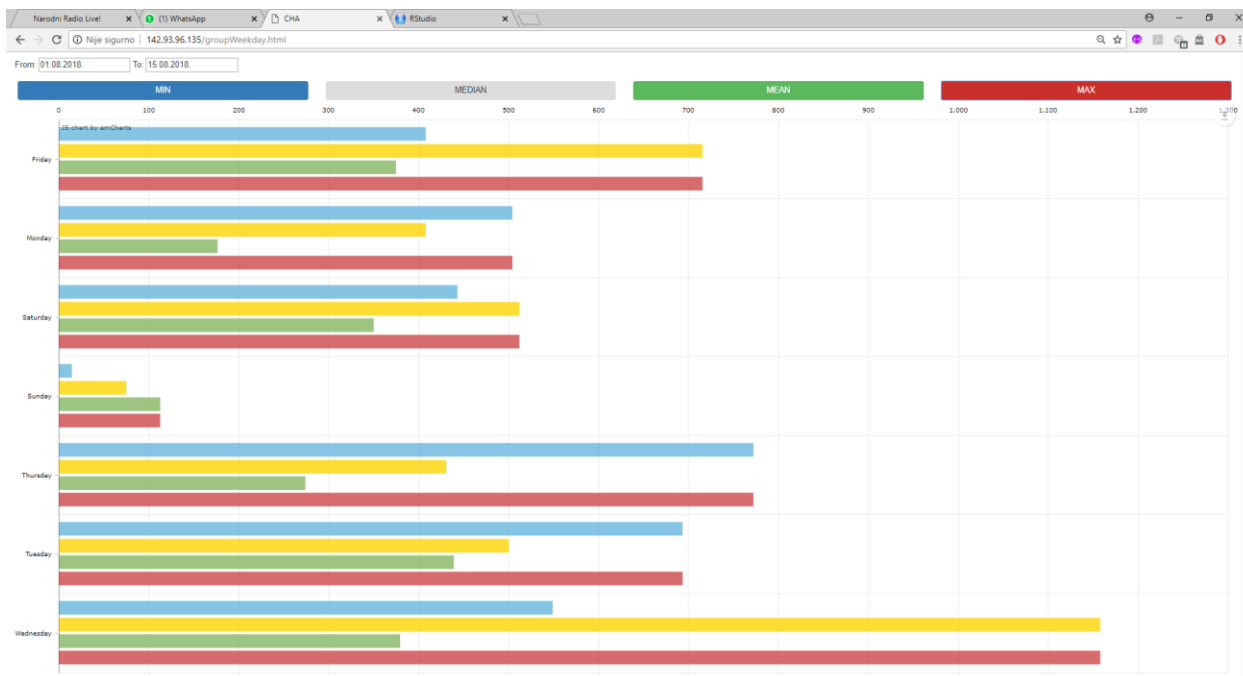
Kao što je i očekivano u nedjelju medijan vrijednosti iznosa računa iskače po iznosu od ostalih vrijednosti za grupu korisnika od 45 do 60 godina starosti, s iznosom od 63,5 kuna. Međutim ostale grupa za ostale kategorije su ispod ukupnog prosjeka. Za grupu kupaca od 18 do 29 godina starosti iznosi 14 kuna, za grupu kupaca od 30 do 45 godina starosti iznosi 23 kune, a za ostale iznosi 26 kuna. Prosječna medijan vrijednost iznosa računa po danu je otprilike 42 kune. Grupa kupaca od 18 do 29 godina starosti najveću medijan vrijednost iznosa računa ima srijedom u iznosu od 53 kune. Žuta grupa kupaca ima najveću medijan vrijednost iznosa računa subotom i ona iznosi 52 kune, a ostali također subotom u iznosu 49 kuna.



Slika 11: Aritmetička sredina vrijednosti po danu

Vrijednosti aritmetičkih sredina iznosa računa u velikoj mjeri se razlikuje od iznosa medijalnih vrijednosti što ukazuje da po danima postoje velike ekstremne vrijednosti iznosa računa. Ponovno po iznosu je najveća vrijednost za zelenu grupu koja za nedjelju iznosi 78 kuna, međutim to nije jedina najveća vrijednost, nego jednaki iznos ima i žuta grupa utorkom. Prosječna vrijednost aritmetičkih sredina za sve grupe je otprilike 55 kuna. Ispod prosječne vrijednosti nedjeljom ima plava grupa u iznosu 14 kuna, žuta grupa 29 kuna i crvena grupa 40 kuna. Ostalih dana sve su grupe poprilično jednake.

Prema dobivenim rezultatima zaključujemo da sve grupe jednako kupuju svakoga dana, izuzevši grupu kupaca od 46 do 60 godina koja najviše kupuje nedjeljom.

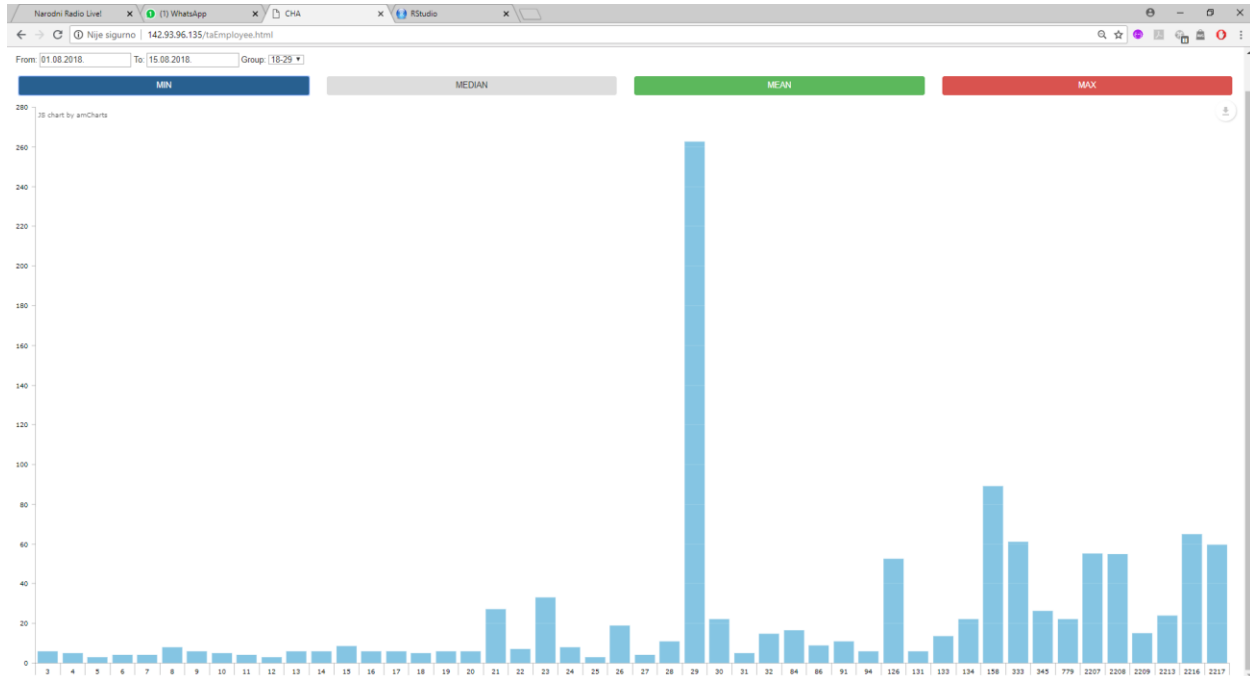


Slika 12: Maksimalne vrijednosti iznosa po danu

Izolirane maksimalne vrijednosti po iznosu iskaču za srijedu u grupi kupaca od 30 do 45 godina starosti i za grupu ostalih kupaca u iznosu od 157 kuna. Vidljivo je da zelena grupa ima malu vrijednost maksimuma u nedjelju što znači da srednja vrijednost i medijalna vrijednost kupovina ove grupe u nedjelju nisu zbog određene ekstremne vrijednosti, nego da navedena grupa stvarno najviše kupuje nedjeljom. Prosječna maksimalna vrijednost je otprilike 300 kuna.

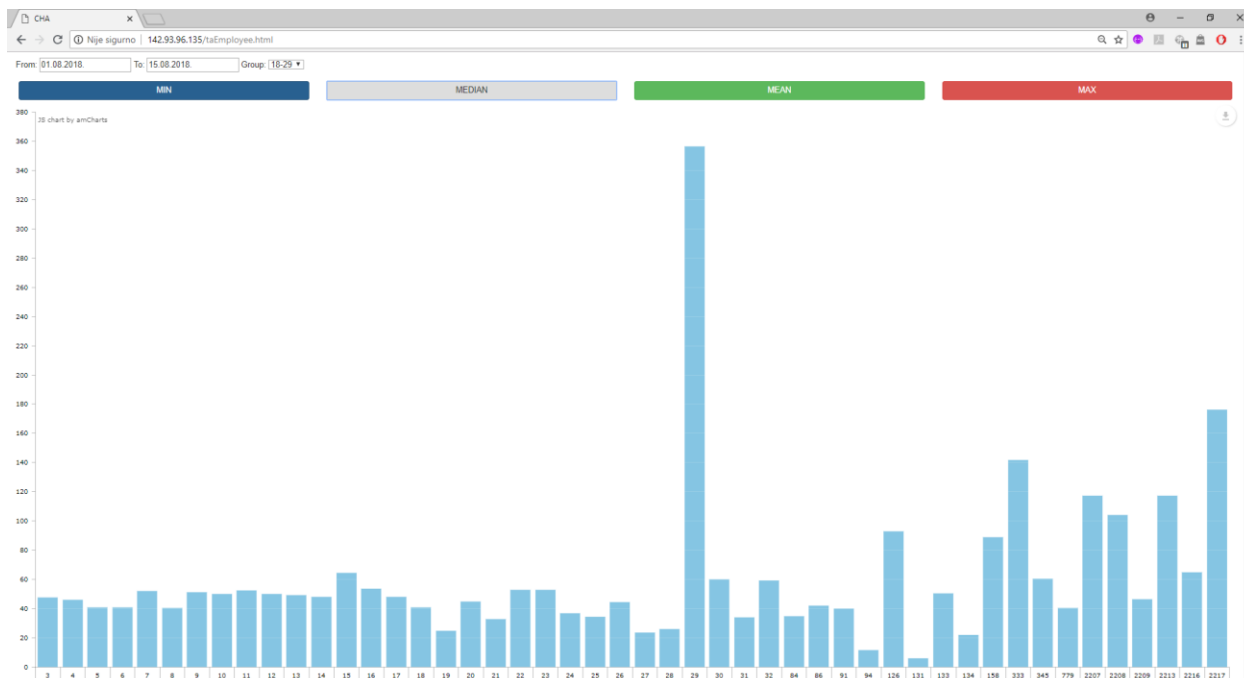
## 4.4. Ukupan iznos računa po zaposleniku

Dosadašnje analize prikazivale su potrošačke navike kupaca. U ovoj analizi prikazuje se efikasnost prodavača po grupama kupaca. U primjeru koji će se prikazati u radu koristit će se vremenski raspon od 01.08.2018. do 15.08.2018. i grupa kupaca od 18 do 29 godina starosti. Za svaku grupu se prikazuje minimalna vrijednost, medijalna vrijednost, aritmetička sredina i maksimalna vrijednost.



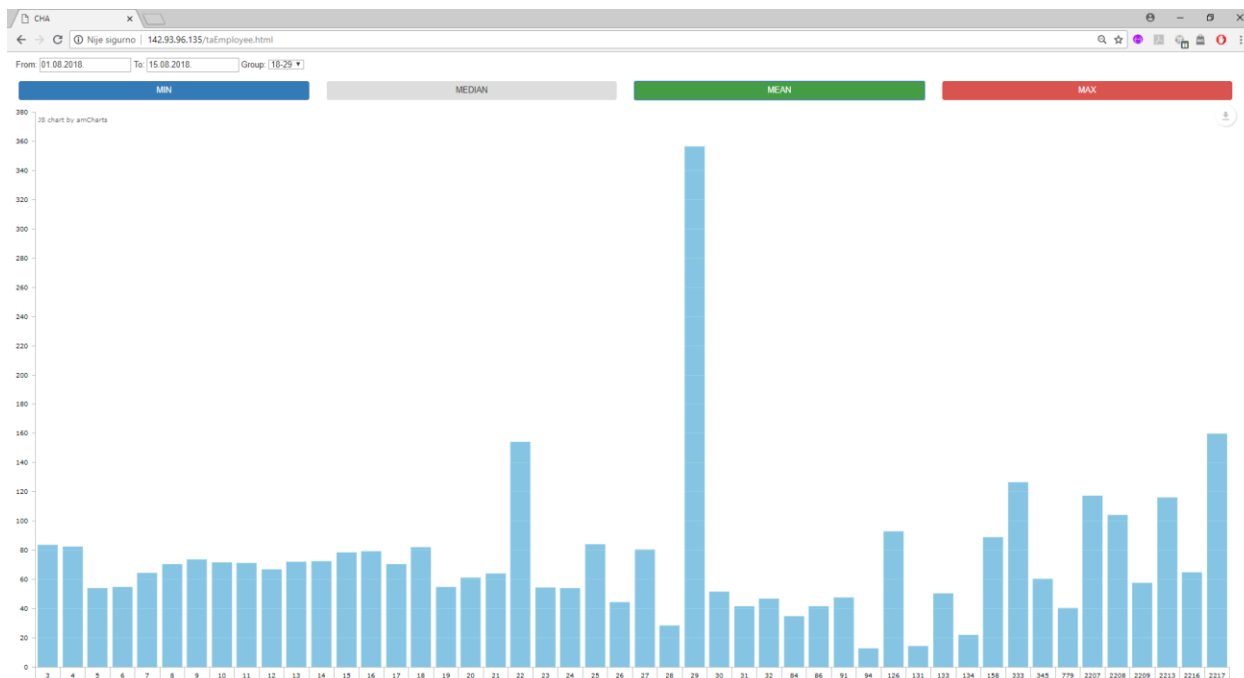
Slika 13: Minimalna vrijednost po zaposleniku

Minimalna vrijednost iznosa računa najveća je za zaposlenika 29 i iznosi 262,5 kune. Pretpostavlja se da je to i jedina vrijednost računa koju je izdao zaposlenik 29. Prosječna minimalna vrijednost iznosa računa je 7 kuna. Najmanje minimalne vrijednosti računa imaju zaposlenici 5 i 25 s iznosom 3 kune.



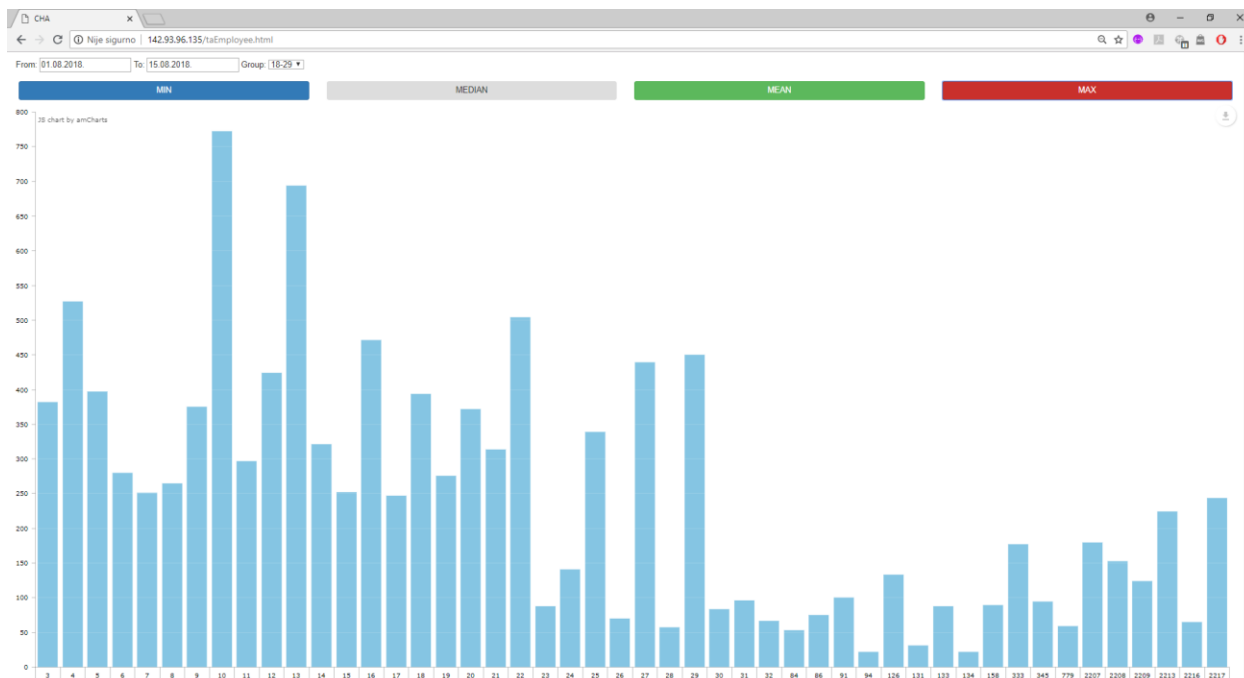
Slika 14: Medijan vrijednosti po zaposleniku

Prethodna pretpostavka da je minimalni račun zaposlenika 29 i jedini račun nije točna, jer kao što se vidi na grafičkom prikazu, medijalna vrijednost iznosa računa zaposlenika 29 ponovno iskače od ostalih s iznosom 356,45 kuna. Prosječna medijan vrijednost iznosa računa je 50 kuna, a ispod tog prosjeka su zaposlenici 94 s 11,8 kuna i 131 s 6 kuna.



*Slika 15: Aritmetička sredina vrijednost po zaposleniku*

Kao i u svim kategorijama tako i u aritmetičkoj sredini vrijednost zaposlenika 29 iskače od prosjeka s iznosom 356,45 kuna. S tim se zaključuje također da je zaposlenik 29 najučinkovitiji u prodaji kupcima starosti od 18 do 29 godina. Aritmetička vrijednost u prosjeku je 60 kuna. Vrijednosti koje odskaku od prosjeka su za zaposlenika 22 u iznosu 154 kune, 333 u iznosu 126 kuna i 2217 u iznosu 160 kuna. Zaposlenici s izrazito malom vrijednosti su zaposlenik 94 s iznosom od 12 kuna i 131 sa 14,5 kuna.



Slika 16: Maksimalna vrijednost po zaposleniku

Maksimalne vrijednosti za zaposlenike su u prosjeku 250 kuna. Po maksimalnim vrijednostima se ističu zaposlenici 4, 10 i 13 koji u prosjeku nisu bolji od ostalih zaposlenika. Prema tome zaključujemo da su prikazane maksimalne vrijednosti zapravo izolirani ekstremi za navedene prodavače. Prodavač 29 koji se ispostavi kao najbolji prodavač maksimalnu vrijednost iznosa računa ima 450 kuna.



## 4.5. Upozorenja o kupcima

U ovom dijelu prikazuju se kupci koji zahtijevaju veću pozornost prodavača. To su kupci koji su u proteklom mjesecu (u periodu od 01.08.2018. do 31.08.2018.) poduzeli određene akcije u kupovini na koje bi prodavač trebao obratiti pozornost.

U prvu skupinu pripadaju kupci koji su u prvom dijelu mjeseca izvršili najmanje 5 kupovina, a u drugom dijelu mjeseca nisu izvršili niti jednu kupovinu. To su kupci koje treba analizirati i pokušati shvatiti zašto su prestali dolaziti u prodavaonice, te kako ih se može vratiti u prodavaonice.

U drugu skupinu kupaca spadaju kupci koji su u prvom dijelu mjeseca izvršili najmanje 5 kupovina, a u drugom dijelu mjeseca su također izvršili najmanje 5 kupovina. To su kupci koji zaslužuju nagradu za svoju vjernost i prodavač treba odrediti kako se u budućnosti ponašati prema tim kupcima.

Customer habits analyse   Total amount by group   Total amount group summary   Total amount by employee   Total amount by weekday   Total discount   Customer alerts

Customers who stopped shopping (in first part of the month they made 5 or more purchase and in second part of the month they made 0 purchase)

- 60595510850080832
- 60595510850005110
- 60595510850000590
- 60595510850001978

Customers who shop constantly (in first and in second part of the month they made 5 or more purchase)

- 60595510850003826
- 605955108500060774
- 60595510850000483
- 60595510850000566
- 60595510850004022
- 60595510850044671
- 60595510850046965
- 60595510850000659
- 60595510850002455
- 60595510850006282
- 60595510850003529
- 60595510850004998
- 60595510850003230
- 60595510850009666
- 60595510850069942
- 60595510850001226
- 60595510850001648
- 60595510850004097
- 60595510850009200
- 60595510850001945
- 60595510850000418
- 60595510850000988

Slika 17: Upozorenje o kupcima

Kupaca koji su prestali kupovati u drugom dijelu mjeseca ima 4, dok kupaca koji su konstanto kupovali kroz cijeli mjesec i zaslužuju nagradu ima 22.

## 4.6. Veza između ukupnog iznosa računa i ukupne uštede po računu

Korelacija znači povezanost, a može biti linearna i ne linearna. Koeficijent korelacije izražava veličinu povezanosti između dvije varijable. U programskom jeziku R koeficijent korelacije podrazumijeva Personov koeficijent korelacije. Uz Personov koeficijent korelacije postoji još mnogo koeficijenata korelacije, a neki od njih su: Spearmanov koeficijent korelacije, parcijalna korelacija, koeficijent multiple korelacije i mnogi drugi.

Koeficijent korelacije prikazuje se Scatter-dijagramom. Na os x ider nezavisna varijabla, a na os y zavisna varijabla.

U primjeru kupci su ponovno podijeljeni u 4 grupe kao do sada:

1. Kupci starosti od 18 do 29 godina
2. Kupci starosti od 30 do 45 godina
3. Kupci starosti od 46 do 60 godina
4. Kupci mlađi od 18 godina i stariji od 60 godina

Računa se koeficijent korelacije za ukupan iznos računa i ukupan popust po računu. Koeficijent korelacije se računa za vremenski period od 01.08.2018. do 31.08.2018.

From:  To:

Correlation between total amount and total discount

- G1(18-29): 0.2474 => Correlation is easy between two values
- G2(30-45): 0.0936 => Correlation doesn't exist.
- G3(46-60): 0.0729 => Correlation doesn't exist.
- G4(other): 0.1708 => Correlation doesn't exist.

**ALL: 0.1927 => Correlation doesn't exist.**

*Slika 18: Korelacija ukupnog iznosa računa i ukupne uštede, izvor: Autor*

Grupa kupaca u starosti od 18 do 29 godina ima pozitivnu nepotpunu korelaciju između ukupnog iznosa računa i ukupne uštede po računu. Ostale grupe imaju nepotpunu korelaciju između te dvije vrijednosti.

Prema dobivenim rezultatima zaključujemo da samo grupa kupaca od 18 do 29 godina u kupovini razmišlja o uštedi prilikom kupovine. Kod drugih grupa kupaca takva povezanost ne postoji i ne postoji ovisnost iznosa računa i iznosa uštede po računu.

#### 4.7. Određivanje vjerojatnosti budućeg iznosa računa

Za određivanje vjerojatnosti iznosa ukupnog iznosa računa koristi se stablo odlučivanja. Vrijednost za koju se određuje vjerojatnost je 50 kuna po računu. Vjerojatnost se računa za svaku pojedinu grupu kupaca koje su već prethodno definirane, a one su:

- Kupci starosti od 18 do 29 godina (G1)
- Kupci starosti od 30 do 45 godina (G2)
- Kupci starosti od 46 do 60 godina (G3)
- Kupci mlađi od 18 godina i stariji od 60 godina (G4)

Vremenski period za koji se računa vjerojatnost iznosa veća od 50 kuna je od 15.08.2018. do 15.09.2018. U konačnosti prikazuje se i vjerojatnost za sve skupa kupce, generalno.

##### Probability to invoice total amount been more then 50 kn

- Group G1 (18-29):
  - YES = 44.629999999999995%
  - NO = 55.37%
- Group G2 (30-45):
  - YES = 45.540000000000006%
  - NO = 54.459999999999994%
- Group G3 (46-60):
  - YES = 39.93%
  - NO = 60.07%
- Group G4 (other):
  - YES = 44.45%
  - NO = 55.55%
- **ALL:**
  - **YES = 44.190000000000005%**
  - **NO = 55.81%**

Slika 19: Vjerojatnost iznosa računa

Vjerojatnost da će račun biti veći od 50 kuna za grupu kupaca G1 iznosi 44,63%, odnosno da neće biti veća od 50 kuna iznosi 55,37%. Dobiveni rezultati daju malu prednost da će slijedeći

račun grupe G1 biti manji od 50 kuna. Slični rezultati su dobiveni i za grupe G2 i G4. U slučaju grupe G3 dobiveni rezultati su isti, odnosno daju prednost računu manjem od 50 kuna. S tim da je za ovu grupu negativna vjerojatnost 60,07%, a pozitivna vjerojatnost 39,93%.

Ukupna vjerojatnost naravno odgovara i vrijednostima po grupama. Da će račun biti veći od 50 kuna je 44,19%, a da će biti manji od 50 kuna je 55,81%.

Kako bi analizirali vjerodostojnost, ali i navike kupaca u različitim vremenskim periodima, analizirali smo i stablo odlučivanja u periodu od 10.06.2018. do 10.07.2018.

#### Probability to invoice total amount been more then 50 kn

- Group G1 (18-29):
  - YES = 41.72%
  - NO = 58.28%
- Group G2 (30-45):
  - YES = 40.63%
  - NO = 59.37%
- Group G3 (46-60):
  - YES = 44.78%
  - NO = 55.22%
- Group G4 (other):
  - YES = 41.589999999999996%
  - NO = 58.41%
- **ALL:**
  - **YES = 41.510000000000005%**
  - **NO = 58.489999999999995%**

*Slika 20: Vjerojatnost iznosa računa iz prošlosti*

Dobiveni rezultati su slični s tim da je zanimljivo vidjeti da je u grupama kupac G1, G2 i G4 u ovom periodu vjerojatnost da će račun biti veći od 50 kuna puno manji nego u prethodno proučavanom periodu. U grupi G3 je vjerojatnost da će račun biti veći od 50 kuna je nešto veći od prosjeka što u prvom periodu nije bio primjer.

Ovakvi rezultati predstavljaju različite kupovne navike kupaca svih grupa kupaca u različitim vremenskim periodima. Takvi rezultati znače da je u različitim periodima teško predvidjeti ponašanje kupaca u nadolazećim periodima, ali da je moguće predvidjeti kupovne navike kupaca u aktualnom periodu.

## Zaključak

Metode za određivanje znanja u podacima pružaju veliku mogućnost otkrivanja različitih znanja iz velike količine podataka. Svaka metodologija donosi svoje koristi u pojedinim područjima i za vrstu te raspodjelu podataka. Također, otkriveno je da svaka metoda nije idealna za svaku vrstu podatka. Iz tog razloga odlučeno je da za predviđanje podataka u istraživanju se koristi korelacija i stablo odlučivanja.

Prethodna istraživanja ponašanja kupca otkriva da su bez obzira kako se kupci personalno klasificiraju, barem u nekom slučaju impulzivni. Iako se u autorovom istraživanju nije određivala impulzivnost kupaca, podatak o mogućnosti impulzivnosti u kupovini otvaraju mogućnosti za nadolazeća istraživanja kupaca u trgovini čiji se podaci koriste u autorovom istraživanju.

U istraživanju je otkriveno da su navike kupaca slične ne ovisno o grupi u kojoj su kupci. Malo veću potrošnju pokazali su kupci mlađe dobne skupine, odnosno starosti od 18 do 29 godina. Većina kupaca najviše troši subotom, dok mlađi kupci troše najviše srijedom. Taj rezultat predstavlja zanimljivost koja će biti predstavljena prodavaonici s ciljem otkrivanja razloga zašto mlađe dobne skupine najviše troše srijedom. Možda je ta činjenica rezultat što mlađe dobne skupine su još uvijek učenici/studenti pa imaju više vremena za kupovinu u tjednu, ali taj podatak u ovom istraživanju nije bio dostupan, tako da može biti samo pretpostavka za buduća istraživanja.

Ono što više začuđuje je rezultat da su za sve dobne skupine potrošača jako slabe veze između potrošenog iznosa po računu i uštedenog iznosa po računu. Očigledno je da je ovdje problem koji povlači dvije mogućnosti za sobom. Ili prodavaonica pogrešne proizvode stavlja na akciju ili kupci kupuju samo one stvari po koje su došli, a vezane proizvode na akciji uopće ne uočavaju. Ovdje je zanimljivo i prethodno analizirano istraživanje vezano za kupovne navike s obzirom na promociju proizvoda u trgovini. Dobiveni rezultati prodavaonici predstavljaju preporuku da povećaju svoje promocijske akcije kako bi povećali prodaju proizvoda na akciji.

Što se tiče predviđanja iznosa računa, u trgovini se može predvidjeti da će kupci iz bilo koje grupe češće kupovati za manje od 50 kuna. Iznos od 50 kuna određen je u dogovoru s prodavaonicom jer bi oni voljeli da većina kupovina bude u iznosu većem od 50 kuna. Iz proučavanih rezultata vidljivo je da je za sada taj iznos teško dohvatljiv, ali u slučaju bolje promocije moguće je i povećati mogućnost za ostvarivanje zadane pretpostavke. Iz rezultata se

iščitava da u različitim periodima su i različite vjerojatnosti da će kupci kupovati za više od 50 kuna. To znači da je na temelju jednog perioda teško odrediti ponašanje kupaca u drugom periodu, ali se može očekivati slično ponašanje kupaca u istom periodu svake slijedeće godine.

## Literatura

- 1) Metode analize podataka, dostupno na: [http://krneta.in.rs/images/Fajlovi/METODE\\_ANALIZE\\_PODATAKA\\_-\\_kurikulum.pdf](http://krneta.in.rs/images/Fajlovi/METODE_ANALIZE_PODATAKA_-_kurikulum.pdf), zadnji pristup: 14.09.2018.
- 2) Enciklopedija, dostupno na: <http://www.enciklopedija.hr/natuknica.aspx?id=40437>, zadnji pristup: 14.09.2018.
- 3) Biochemia Medica: Što treba znati kada izračunavamo koeficijent korelacije?, dostupno na: <http://www.biochemia-medica.com/content/s-treba-znati-kada-izracunavamo-koeficijent-korelacije>, zadnji pristup: 14.09.2018.
- 4) Zdravstveno veleučilište: Korelacija, dostupno na: [https://ldap.zvu.hr/~oliverap/MetodeIstrazivanjaFT/11\\_Korelacija.pdf](https://ldap.zvu.hr/~oliverap/MetodeIstrazivanjaFT/11_Korelacija.pdf), zadnji pristup: 14.09.2018.
- 5) Veleučilište Rijeka: Korelacijska i regresijska analiza, dostupno na: [https://www.veleri.hr/files/datotekep/nastavni\\_materijali/k\\_poduzetnistvo\\_s1/Kvantitativne\\_za\\_poduzetnike\\_Pr2\\_Izv.pdf](https://www.veleri.hr/files/datotekep/nastavni_materijali/k_poduzetnistvo_s1/Kvantitativne_za_poduzetnike_Pr2_Izv.pdf), zadnji pristup: 14.09.2018.
- 6) FESB: Monotonost, dostupno na: <http://lavica.fesb.hr/mat1/predavanja/node118.html>, zadnji pristup: 14.09.2018.
- 7) Sveučilište u Zadru: Regresija i korelacija, dostupno na: [www.unizd.hr/portals/4/nastavni\\_mat/2\\_godina/statistika/statistika\\_09.ppt](http://www.unizd.hr/portals/4/nastavni_mat/2_godina/statistika/statistika_09.ppt), zadnji pristup: 14.09.2018.
- 8) Stabla odlučivanja, dostupno na: <http://www.skladistenje.com/stabla-odlucivanja/>, zadnji pristup: 14.09.2018.
- 9) Sveučilište u Osijeku: Stabla odlučivanja, dostupno na: [http://www.efos.unios.hr/sustavi-poslovne-inteligencije/wp-content/uploads/sites/192/2017/10/P4\\_Stabla-odlucivanja-2017.pdf](http://www.efos.unios.hr/sustavi-poslovne-inteligencije/wp-content/uploads/sites/192/2017/10/P4_Stabla-odlucivanja-2017.pdf), zadnji pristup: 14.09.2018.

- 10) Ekonomski fakultet u Zagrebu: Stabla odlučivanja, dostupno na: <http://www.efzg.unizg.hr/UserDocsImages/OIM/mdarabos/4-Stablo%20odlu%C4%8Divanja.pdf>, zadnji pristup: 14.09.2018.
- 11) Investopedia: Backward Induction, dostupno na: <https://www.investopedia.com/terms/b/backward-induction.asp>, zadnji pristup: 14.09.2018.
- 12) FER: Stabla odlučivanja, dostupno na: <http://www.zemris.fer.hr/education/ml/nastava/ag20022003/3-stabla-odluke.pdf>, zadnji pristup: 14.09.2018.
- 13) Kocijan Samanta: Emocija kao čimbenik ponašanja potrošača, veljača 2017.
- 14) Mihić Mirela, Kursan Ivana: TEMELJU IMPULZIVNOSTI U KUPOVNOM PONAŠANJU POTROŠAČA
- 15) Anić Ivan-Damir, Radas Sonja: Analiza ponašanja potrošača s obzirom na njihovu sklonost prema promociji u hipermarketu u Hrvatskoj
- 16) Ravlić Mila: UTJECAJ OGLAŠAVANJA NA INTERENTU NA PONAŠANJE POTROŠAČA
- 17) Gal Martina: VREDNUJUĆI ČIMBENICI PONAŠANJA POTROŠAČA PRILIKOM DONOŠENJA ODLUKE O KUPNJI LUKSUZNIH KOZMETIČKIH PROIZVODA



## Popis slika

Slika 1: Minimalan iznos računa.....	15
Slika 2: Medijan iznos računa.....	16
Slika 3: Aritmetička sredina iznosa računa .....	17
Slika 4: Maksimum iznosa računa .....	18
Slika 5: Grupni izvještaj grupe 1.....	20
Slika 6: Grupni izvještaj grupe 2.....	21
Slika 7: Grupni izvještaj grupe 3.....	22
Slika 8: Grupni izvještaj grupe 4.....	23
Slika 9: Minimalni iznosi po danu.....	24
Slika 10: Medijan iznosi po danu .....	25
Slika 11: Aritmetička sredina vrijednosti po danu .....	26
Slika 12: Maksimalne vrijednosti iznosa po danu .....	27
Slika 13: Minimalna vrijednost po zaposleniku.....	28
Slika 14: Medijan vrijednosti po zaposleniku.....	29
Slika 15: Aritmetička sredina vrijednost po zaposleniku .....	30
Slika 16: Maksimalna vrijednost po zaposleniku .....	31
Slika 17: Upozorenje o kupcima.....	32
Slika 18: Korelacija ukupnog iznosa računa i ukupne uštede, izvor: Autor.....	33
Slika 19: Vjerojatnost iznosa računa .....	34
Slika 20: Vjerojatnost iznosa računa iz prošlosti.....	35