Kazalo

[1. Uvod (Boris) 2](#_Toc518049015)

[2. Zasnova raziskave 2](#_Toc518049016)

[2.1. Predstavitev problema 2](#_Toc518049017)

[2.2. Sestava vprašalnika 3](#_Toc518049018)

[2.3. Opredelitev populacije in vzorca 5](#_Toc518049019)

[2.4. Izvedba raziskave 5](#_Toc518049020)

[3. Analiza podatkov (Bojan) 6](#_Toc518049021)

[3.1. Analiza vzorca 6](#_Toc518049022)

[3.2. Analiza mnenjskega dela anketnega vprašalnika 12](#_Toc518049023)

[4. Bivariatni testi (deljeno Boris in Jure) 15](#_Toc518049024)

[4.1. Predstavitev obeh raziskovalnih vprašanj 15](#_Toc518049025)

[4.2. Analiza 1. raziskovalnega vprašanja 15](#_Toc518049026)

[4.3. Analiza 2. raziskovalnega vprašanja 20](#_Toc518049027)

[5. Multivariatne metode (Jure in Bojan) 20](#_Toc518049028)

[5.1. Predstavitev problema in raziskovalnih vprašanj 20](#_Toc518049029)

[5.2. Prva multivariatna metoda (navedete ime metode) 20](#_Toc518049030)

[5.3. Druga multivariatna metoda (navedete ime metode) 24](#_Toc518049031)

[6. Kritična analiza zasnove in izvedbe raziskave (Jure) 29](#_Toc518049032)

[7. Zaključek (skupaj) 29](#_Toc518049033)

[8. Literatura 29](#_Toc518049034)

[Priloge 29](#_Toc518049035)

[A. Vprašalnik 29](#_Toc518049036)

# Uvod (Boris)

Današnji svet v katerem danes živimo se razvija hitro, tako znanstveno, kot družbeno. Da bi lahko sledili tem trendom in napovedovali prihodnost rabimo neko orodje/znanstveno vedo za lažje odločanje ali razlago nekih določenih pojavov. Ena izmed orodji/znanstvenih ved je tudi statistika.

Statístika je znanost in veščina o razvoju znanja z uporabo izkustvenih podatkov. Njena osnova je matematična statistika, ki je veja uporabne matematike. V statistični teoriji naključnost in negotovost opišemo v okviru teorije verjetnosti. Ker je naloga statistike izluščiti »najboljšo« informacijo iz razpoložljivih podatkov, jo nekateri uvrščajo v teorijo odločanja (Wiki: https://sl.wikipedia.org/wiki/Statistika).

Aktualno področje v tem spomladanskem obdobje so dopustovanja. Zato smo se odločili, da bomo analizirali to temo. Tema je zanimiva ne samo za nas ampak tudi za slovenski turizem, saj bi lahko s tem napovedovali kašne trende lahko pričakujejo gostinci in s tem posledično prilagodili svojo ponudbo.

V okviru našega raziskovanja nas je zanimalo zadovoljstvo ljudi na zadnji dopustovanjih. Kako so bili zadovoljni z nastanitvijo, ponudbo različnih aktivnosti, različnih storitev in na koncu s ceno nastanitve/storitev.

Želja je, da bi z našo raziskovalno nalogo analizirali anketo in prišli do uspešnega zaključka (predvidevanja za prihodnost).

Na začetku bomo postavili temelje za sestavo raziskovanega vprašanja to je zasnova spletne ankete. Nato bomo pridobljene podatke analizirali, jih z bivariatn testi in mulitvariantnimi metodami analizirali in na koncu potrdili ali zavrnili raziskovalna vprašanja.

# Zasnova raziskave

Celotna raziskovana naloga temelji na uspešni izvedbi anketiranja na predvideni populaciji. Želimo čim večjo populacijo.

## Predstavitev problema

Problem zasnove raziskave je:

* kako bi našo idejo realizirali.
* Kaj bo naša anketa vsebovala(katera vprašanja)
* Katero populacijo vzeti
* Način anektiranja (izvedba raziskave)

## Sestava vprašalnika

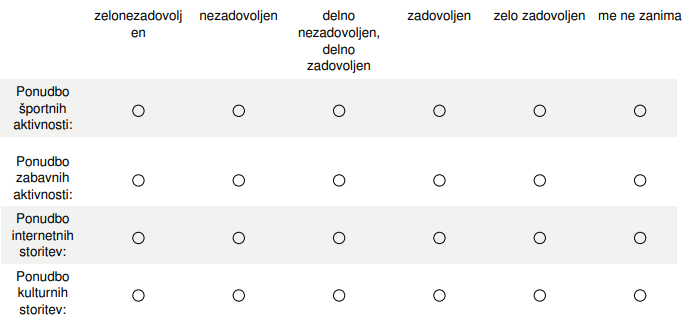
Odločili smo se da bomo naša anketa vsebovala naslednja vprašanja:

1. DOPUSTOVAL SEM:
   1. SAM
   2. S PARTNERJEM
   3. Z DRUŽINO
   4. S PRIJATELJI
   5. DRUGO (Prosimo, dopišite):
2. Q2 - S KATERIM PREVOZNIM SREDSTVOM STE POTOVALI (Možnih je več odgovorov)
   1. OSEBNI AVTO
   2. AVTODOM
   3. AVTOBUS
   4. VLAK
   5. LETALO
   6. KOLO
   7. LADJA
   8. MOTORNO KOLO
   9. DRUGO (Prosimo, dopišite):
3. KJE STE PRENOČEVALI
   1. HOTEL
   2. KAMP
   3. APARTMA
   4. NA PLOVILU
   5. DRUGO (Prosimo, dopišite):
4. ŠTEVILO NOČITEV
   1. 1 - 3
   2. 4 - 7
   3. 8 - 14
   4. Več
5. ZADOVOLJSTVO S POČITNIŠKO NASTANITVIJO



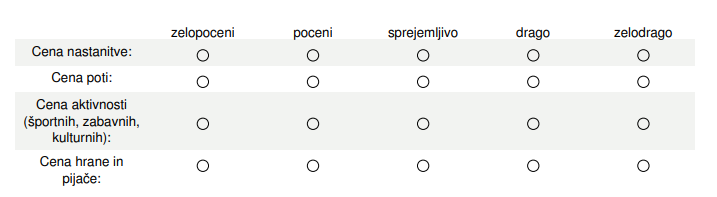
Slika 1 Zadovolstvo s počitniško nastanitvijo

1. ZADOVOLJSTVO S PONUDBO AKTIVNOSTI IN STORITEV



Slika 2 Zadovoljstvo s ponudbo aktivnosti in storitev

1. CENA NASTANITVE IN STORITEV



Slika 3 Cena nastanitve in storitev

1. SPOL
   1. M
   2. Ž
2. STAROST (Prosimo vpišite število let).
   1. Vpis let

## Opredelitev populacije in vzorca

Za populacijo smo vzeli zaposlene v našem podjetju. Predvidena populacija je cca. 300 vprašanih.

Število vzorcev/uspešno odgovorjenih je cca. 40.

## Izvedba raziskave

Anketo smo naložili na spletni portali [www.1ka.si](http://www.1ka.si). Nato razposlali vsem zaposlenim prošnjo za izpolnitev ankete s naslednjim besedilom:

Pozdravljeni, za potrebe izdelave raziskovalne naloge z naslovom Zadovoljstvo z zadnjim počitnikovanjem na morju smo pripravili slednjo anketo. Namen ankete je ugotavljanje vašega zadovoljstva z vašo zadnjo počitniško nastanitvijo, ponudbo aktivnosti in storitev ter ceno storitev in nastanitve. Vaše sodelovanje je za raziskavo ključno, saj le z vašimi odgovori lahko dobimo neodvisen vpogled v zadovoljstvo s počitniškimi nastanitvami in storitvami. Anketa je anonimna, sestavljena je iz 9 vprašanj. Za izpolnjevanje boste potrebovali približno 5 minut časa. Zbrani podatki bodo uporabljeni izključno za pripravo te raziskovalne naloge. Za vaše obdelovanje se vam prijazno zahvaljujemo. Bojan, Boris, Jure

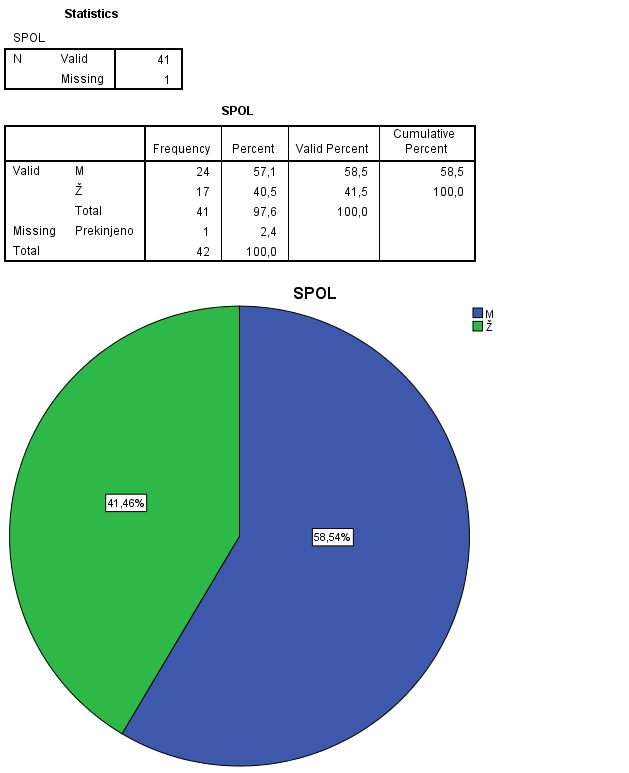
Anketa je bila aktivna 14 dni.

# Analiza podatkov

## Analiza vzorca

### VPRAŠANJE: SPOL

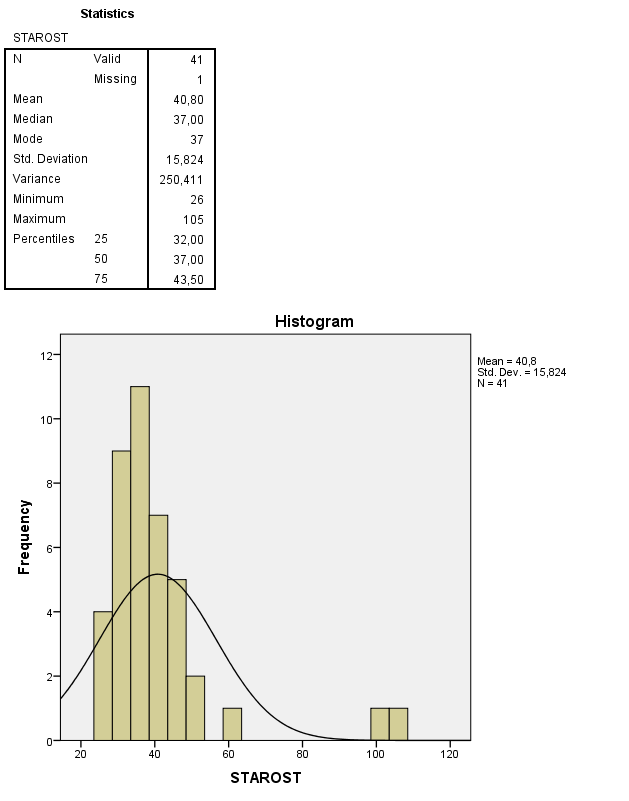
Rezultati so podani v tabelah in grafično prikazani s tortnim diagramom na sliki 1. Iz prve tabele lahko razberemo, da imamo 42 odgovorov anketirancev za spremenljivko »Spol«. Od tega 41 veljavnih in 1 manjkajoč odgovor. Iz druge frekvenčne tabele lahko razberemo, da je v naši anketi sodelovalo 24 moških (58,5%) in 17 žensk (41,5%)



Slika 4: Rezultati za spremenljivko Spol

### VPRAŠANJE: STAROST

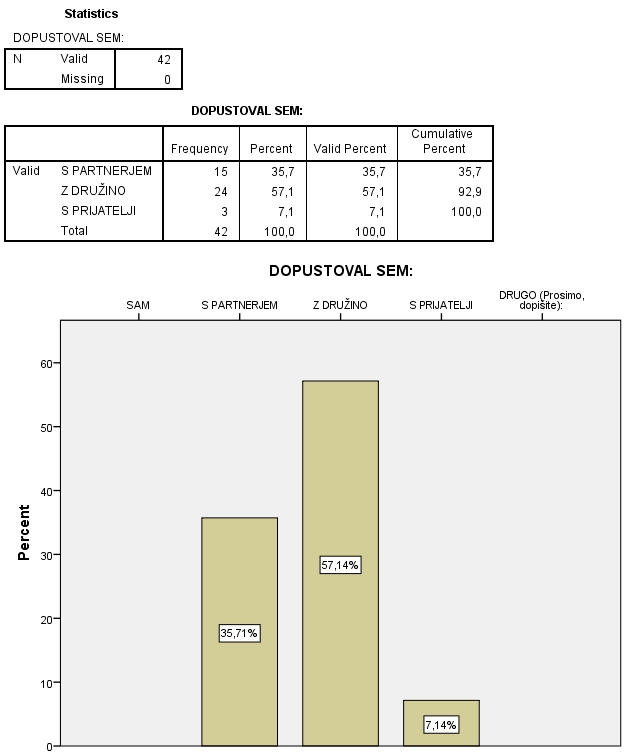
Rezultati so podani v tabeli z opisnimi statistikami in histogramom na sliki 2. Iz tabele razberemo, da je na anketo odgovarjalo 42 anketirancev, vendar jih je za spremenljivko »Starost« ustreznih 41, en odgovor pa je manjkajoč. Najmlajši anketiranec je star 26 let, najstarejši pa 105 let. Povprečna starost anketirancev je 41 let, standardni odklon pa je 15,8 let. Polovica anketirancev je mlajših od 37 let, polovica pa je starejših. Četrtina (25%) anketirancev je mlajših od 32 let, četrtina anketirancev pa je starejših od 43,5 let. Iz histograma razberemo, da imamo dva ekstrema »outlier«-ja eden izmed njiju je najstarejši anketiranec s 105 leti. Verjetno je take vrednosti smiselno popraviti (oceniti) ali jih odstraniti, saj vplivajo na podatke.



Slika 5: Rezultati za spremenljivko Starost

### VPRAŠANJE: DOPUSTOVAL SEM

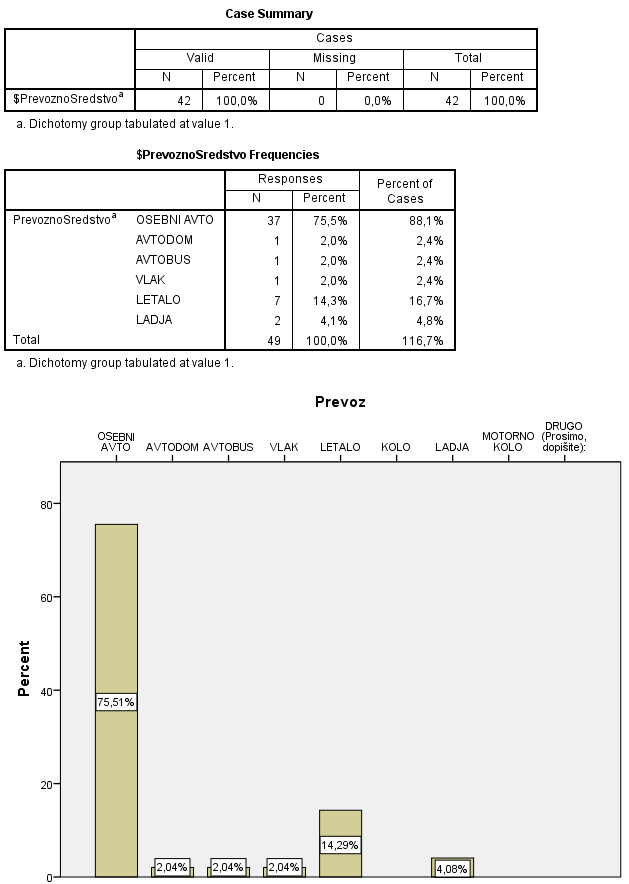
Rezultati so podani v tabelah in grafično prikazani s paličnim diagramom na sliki 3. Iz prve tabele lahko razberemo, da imamo za spremenljivko »Dopustoval sem« 42 odgovorov ter nobenega manjkajočega odgovora. Iz druge frekvenčne tabele lahko razberemo, da je 24 anketirancev (57,1%), kar je več kot polovica, dopustovalo z družino, 15 anketirancev (35,7%) je dopustovalo s partnerjem, trije anketiranci (7,1%) pa so dopustovali s prijatelji.



Slika 6: Rezultati za spremenljivko Dopustoval sem

### VPRAŠANJE: S KATERIM PREVOZNIM SREDSTVOM

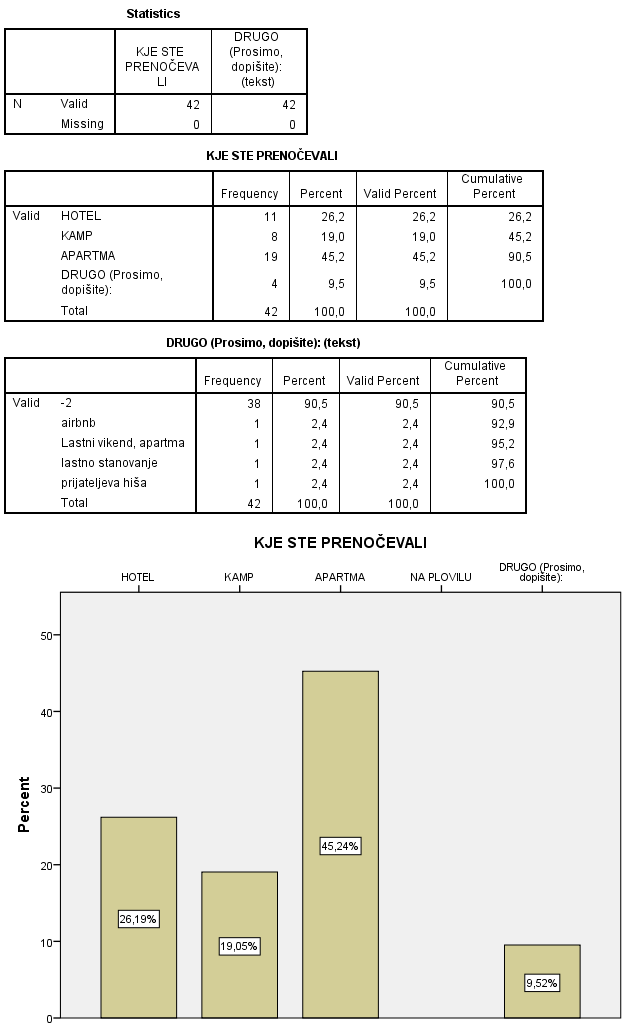
Rezultati so podani v tabelah in grafično prikazani s paličnim diagramom na sliki 4. Iz prve tabele lahko razberemo, da je na vprašanje oziroma spremenljivko »S katerim prevoznim sredstvom« odgovorilo 42 anketirancev. Manjkajočih odgovorov ni. Je pa nekaj anketirancev izbralo več prevoznih sredstev hkrati, kar lahko razberemo, če upoštevamo prvo in drugo tabelo. Razberemo lahko tudi, da je večina kar 37 anketirancev (75%) potovala z osebnim avtomobilom, sledi potovanje z letalom, kar je izbralo 7 anketiranih (14,3%) ter nato ostala prevozna sredstva. Razlika in prevladovanje potovanja z avtomobilom je prikazana na paličnem diagramu



Slika 7: Rezultati za spremenljivko S katerim prevoznim sredstvom

### VPRAŠANJE: KJE STE PRENOČEVALI

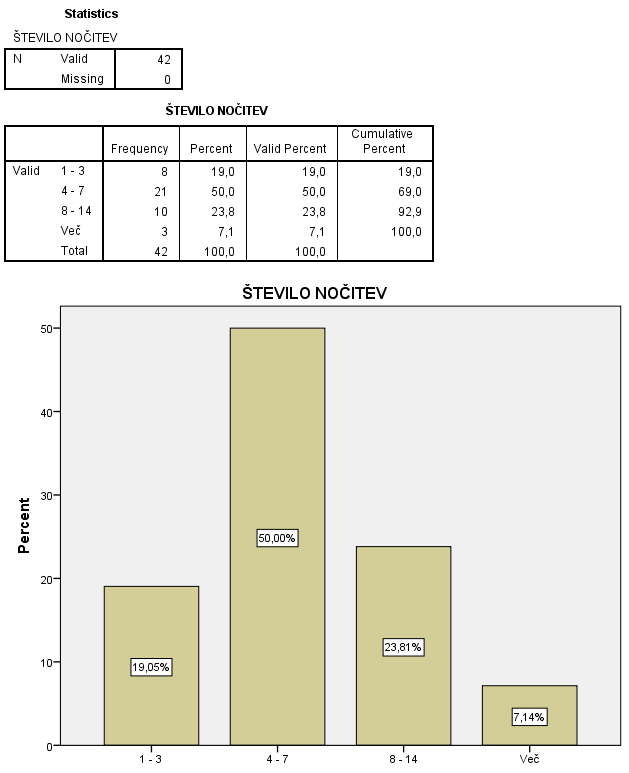
Rezultati so podani v tabelah in grafično prikazani s paličnim diagramom na sliki 5. Iz prve tabele lahko razberemo, da imamo za spremenljivko »Kje ste prenočevali« 42 odgovorov ter nobenega manjkajočega odgovora. Največ oziroma 19 anketiranih (45,2%) je prenočevalo v apartmaju, 11 (26.2%) v hotelu, 8 (19%) v kampu, 4 anketiranci (9,5%) pa so izbrali možnost drugo in navedli: Airbnb, lastni vikend, apartma, prijateljeva hiša.



Slika 8: Rezultati za spremenljivko Kje ste prenočevali

### VPRAŠANJE: ŠTEVILO NOČITEV

Rezultati so podani v tabelah in grafično prikazani s paličnim diagramom na sliki 6. Iz prve tabele lahko razberemo, da imamo 42 odgovorov anketirancev za spremenljivko »število nočitev« ter nobenega manjkajočega. Iz tabele frekvenc in diagrama najprej razberemo, da največ, kar 21 anketirancev (50%) oziroma polovica anketiranih počitnikuje od 4-7 dni, 10 anketirancev (23,8%) počitnikuje od 8-14 dni, 8 anketirancev (19%) počitnikuje od 1-3 dni, 3-je anketiranci pa na počitnicah preživijo več kot 14 dni.

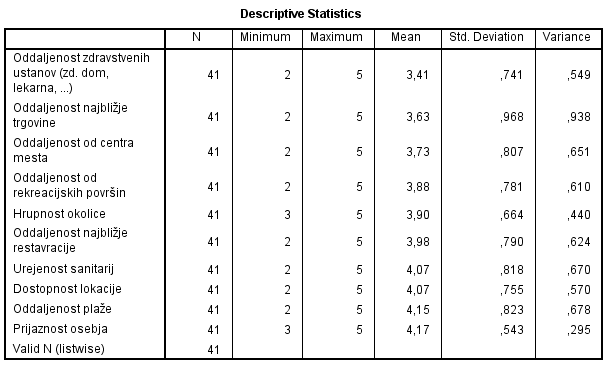


Slika 9 Rezultati za spremenljivko Število nočitev

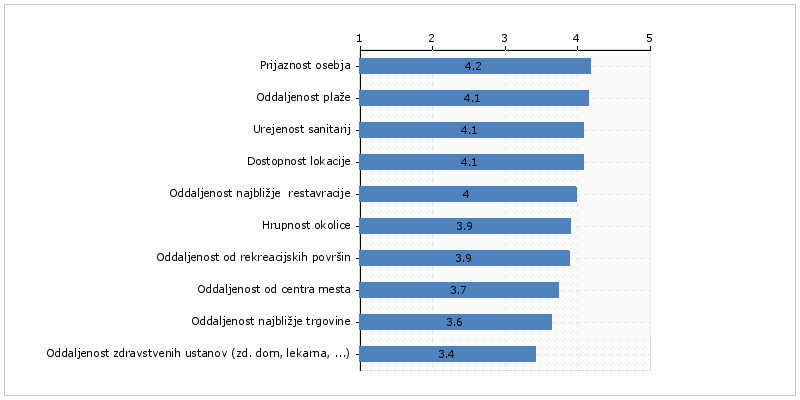
## Analiza mnenjskega dela anketnega vprašalnika

### VPRAŠANJE: ZADOVOLJSTVO S POČITNIŠKO NASTANITVIJO

Rezultati so prikazani s tabelo opisnih statistik na sliki 7 in grafom na sliki 8, ki prikazuje povprečja spremenljivk. Razberemo lahko, da so anketiranci na vprašanje o zadovoljstvu s prijaznostjo osebja označili 3 ali več na pet-stopenjski lestvici. Anketiranci so v povprečju najbolj zadovoljni s prijaznostjo osebja (, sledi oddaljenost plaže ( ter dostopnost lokacije in urejenost sanitarij z (. Najmanj zadovoljni so anketiranci z oddaljenostjo zdravstvenih ustanov ( in najbližje trgovine (. Podatki so najmanj razpršeni pri vprašanju o zadovoljstvu s prijaznostjo osebja ( ter najbolj razpršeni pri vprašanju o oddaljenosti najbližje trgovine (.



Slika 10: Opisne statistike Zadovoljstvo s počitniško nastanitvijo



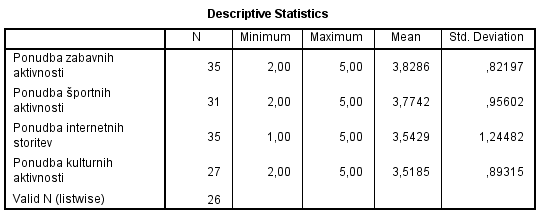
Slika 11 Grafična predstavitev povprečij posameznih spremenljivk za vprašanje Zadovoljstvo s počitniško nastanitvijo

### VPRAŠANJE: ZADOVOLJSTVO S PONUDBO AKTIVNOSTI IN STORITEV

Rezultati so prikazani s tabelo opisnih statistik na sliki 9 in poljubno tabelo. Vprašanje smo postavili v obliki šest-stopenjske lestvice, kjer vrednost 1 predstavlja »zelo nezadovoljen« vrednost 6 pa »me ne zanima«. V kolikor anketiranec izbere »me ne zanima« oziroma vrednost 6 se ta upošteva pri izračunu statistik, kar ni pravilno. Saj bi vrednost 6 v tem primeru predstavljala veliko zadovoljstvo, dejansko pa anketiranca nima mnenja.

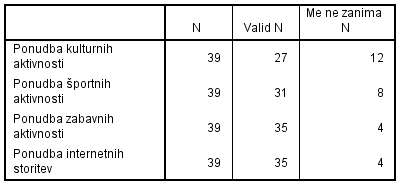
Da bi se izognili vplivu odgovora »me ne zanima« na izračune, smo vprašanje s pomočjo programa SPSS transformirali v pet-stopensko lestvico, kjer je vrednost 1 »zelo nezadovoljen« vrednost 5 pa »zelo zadovoljen«. Vrednost »me ne zanima« smo prikazali kot manjkajočo (missing). Na ta način smo lahko izračunali statistike za katere so anketiranci podali oceno.

Anketiranci so v povprečju najbolj zadovoljni s ponudbo zabavnih aktivnosti ( sledi jim ponudba športnih aktivnosti (. Najmanj pa so anketiranci zadovoljni z ponudbo internetnih storitev ( in ponudbo kulturnih aktivnosti (. Podatki so najmanj razpršeni pri vprašanju o zadovoljstvu s ponudbo zabavnih aktivnosti ( ter najbolj razpršeni pri vprašanju o ponudbi internetnih storitev (s .



Slika 12: Opisne statistike Zadovoljstvo s ponudbo aktivnosti in storitev

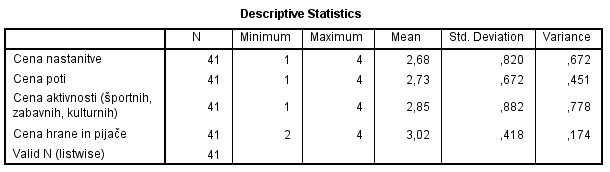
Na sliki 10 lahko iz poljubne tabele razberemo, da kar 12 anketirancev (30,8 %) ponudba kulturnih aktivnosti ne zanima. Sledi nezanimanje za športne aktivnosti, kar je izbralo 8 anketirancev (20,5 %). Izbira odgovora »me ne zanima« je najmanjša pri ponudbi športnih aktivnosti in internetnih storitvah s po 4 odgovori (10,3%).



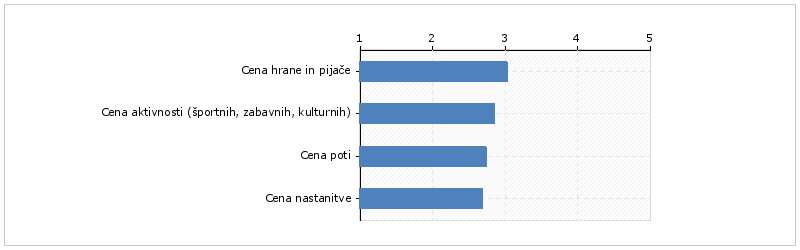
Slika 13 Poljubna tabela prikaz veljavnih in manjkajočih odgovorov

### VPRAŠANJE: CENA NASTANITVE IN STORITEV

Rezultati so prikazani s tabelo opisnih statistik na sliki 11 in grafom na sliki 12, ki prikazuje povprečja spremenljivk. Iz tabele opisnih statistik lahko razberemo, da so cene v povprečju za anketirance sprejemljive (2,68 Najdražje se anketirancem zdijo cene hrane in pijače ( sledijo cene aktivnosti (. V povprečju so anketiranci najbolj zadovoljni s cenami nastanitev (. Podatki so najmanj razpršeni pri vprašanju o cenah hrane in pijače ( ter najbolj razpršeni pri vprašanju o cenah aktivnosti (.



Slika 14 Opisne statistike Cena nastanitve in storitev



Slika 15 Grafična predstavitev povprečij posameznih spremenljivk za vprašanje Cena nastanitve in storitev

### UGOTOVITVE:

Pri analizi vzorca natančneje pri analizi vprašanja »STAROST« smo zaznali, da imamo v podatkih dve ekstremni vrednosti. To sta enoti oziroma anketiranca, ki sta pri dotičnem vprašanju vnesli vrednosti 105 let druga pa 99 let. Zaznali smo tudi, da imamo anketiranca, ki na pet vprašanj od devetih ni podal odgovora oziroma je anketo predčasno zaključil. Da bi čim bolj izničili vpliv takih vrednosti na rezultate, smo pri nadaljnjih analizah podatkov, bivariatnih testih in multivariatnih metodah podatke teh enot odstranili.

Po odstranitvi ekstremov pri vprašanju starost se vrednost povprečne starosti anketirancev zmanjša iz 41 let na 37 let, standardni odklon se zmanjša iz 15,8 let na 7 let, vrednost variance pa se zmanjša iz 250,4 na 55,9.

# Bivariatni testi

## Predstavitev obeh raziskovalnih vprašanj

Pri prvem raziskovalnem vprašanju nas zanima ali dolžina počitkovanja vpliva na izbor počitniške nastanitve. Za ugotovitev naše teze bom uporabili bivarnatno metodo χ2test.

Durgo raziskovalno vprašanje....

## Analiza 1. raziskovalnega vprašanja

Za analizo 1. raziskovalnega vprašanja smo uporabili bivariatni χ2test. Poznamo dve vrsti tega testa, χ2 test za preizkus hipoteze neodvisnosti in χ2 test za preizkus hipoteze enake verjetnosti. Prvi, ki se ga sicer uporablja najpogosteje, se uporablja, ko želimo preveriti soodvisnost dveh spremenljivk oz. njuno povezanost. Drugi pa se uporablja, ko imamo samo eno spremenljivko in želimo preveriti, ali se rezultati opazovane spremenljivke razlikujejo od rezultatov, ki bi nam jih prineslo golo naključje.

Pri našem 1. raziskovalnem vprašanju nas je zanimalo, ali dolžina počitnikovanja kakorkoli vpliva na izbiro počitniške nastanitve. Torej, primerjali smo povezanost spremenljivk »ŠTEVILO NOČITEV« in »KJE\_STE\_PRENOČEVALI«. Podatek o številu nočitev je v našem primeru predstavljen na ordinalni (urejenostni) lestvici, kjer smo določili 4 časovne intervale, ki si sledijo od manjšega časovnega obdobja k večjemu:

* 1 do 3 dni,
* 4 do 7 dni,
* 8 do 14 dni in
* Več kot 14 dni.

Podatek o počitniški nastanitvi smo beležili na imenski lestvici, kjer pa enote med seboj niso povezane. Kandidati so lahko izbirali med štirimi možnostmi, in sicer:

* hotel,
* kamp,
* apartma in
* Drugo.

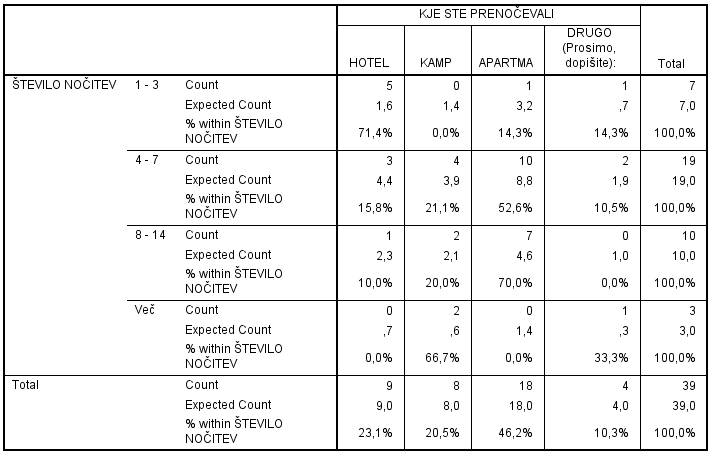
Za izvedbo χ2 testa nad omenjenima spremenljivkama smo najprej postavili ustrezni statistični hipotezi:

*H0*: Spremenljivki *Število\_nočitev*  in *Kje\_ste\_prenočevali* nista povezani.

*H1*: Spremenljivki *Število\_nočitev*  in *Kje\_ste\_prenočevali* sta povezani.

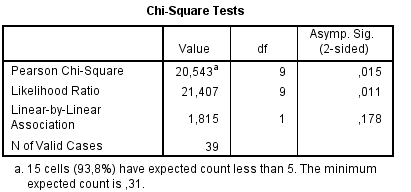
S pomočjo orodja SPSS smo izvedli hi-kvadrat test nad našimi podatki, izpisali pa smo tudi kontingenčno tabelo za boljšo predstavo o povezanosti spremenljivk.

Na sliki 1 je prikazana kontingenčna tabela za spremenljivki *Število\_nočitev* in  *Kje ste prenočevali.* Tabeli smo dodali še izpis za število opazovanih enot in pa teoretične frekvence. Iz tabele lahko razberemo, katero vrsto nastanitve so anketirance izbrali glede na dolžino njihovega počitnikovanja. Že na prvi pogled opazimo, kako se z večanjem števila nočitev vse manj anketirancev odloča za počitnikovanje v hotelu in se posledično odloča za alternativne tipe nastanitve (apartma, kamp).



Slika 16: Kontingenčna tabela spremenljivk Število\_nočitev in Kje\_ste\_prenočevali.

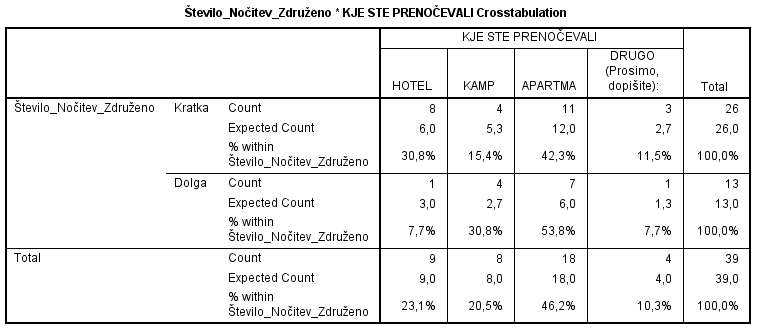
Na sliki 2 je prikazana tabela χ2 testa, ki nam prikazuje stopnjo povezanosti izbranih dveh spremenljivk. Program SPSS nas je ob izpisu tabele opozoril, da je kar 15 celic v kontingenčni tabeli z pričakovanimi frekvencami, ki so manjše od 5, kar je v nasprotju z zahtevami χ2 testa. Pri χ2 testu je potrebno za kakovostne rezultate zadostiti pogoju, in sicer da pričakovane frekvence ne smejo biti manjše od 5 (avtor pogoja je statistik William G. Cochran). Seveda je ta pogoj zelo strog in velikokrat onemogoči nadaljevanje raziskovanja, zato se v praksi uporablja nekoliko milejša oblika pogoja, in sicer da je najvišji dovoljeni delež pričakovanih frekvenc manjših od 5 enak 20% in da nobena med njimi ne sme biti manjša od 1 (Benstat).



Slika 17: Tabela χ2 testa za spremenljivki Število\_nočitev in Kje\_ste\_prenočevali.

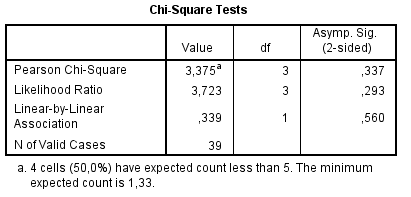
Z namenom, da bi zadostili osnovnemu pogoju χ2 testa, smo se odločili, da združimo kategorije znotraj spremenljivke *Število\_nočitev.* Tako smoiz štirih kategorij, ki opisujejo dolžino počitnikovanja, naredili samo dve kategoriji. Kategoriji nam podata informacijo o tem, ali je bila dolžina počitnikovanja kratka (od 1 do 7 dni) ali dolga (8 dni ali več).

Slika 3 prikazuje kontingenčno tabelo, kjer smo 4 kategorije spremenljivke *Število\_nočitev* združili v 2 kategoriji z namenom zadostitve pogoju χ2 testa. Vidimo, da se z izbiro daljšega obdobja počitnikovanja spremeni tudi izbira tipa nastanitve, vendar pa ta razlika ni tako očitna kot pri izvedbi testa na spremenljivke brez združevanja kategorij.



Slika 18: Kontingenčna tabela z združevanjem kategorij spremenljivke Število\_nočitev.

Kljub združevanju kategorij v spremenljivki nam ni uspelo zadostiti pogojem χ2 testa, saj so še vedno 4 celice v tabeli z pričakovanimi vrednostmi, ki so manjše od 5. Kljub temu smo izvedli test in dobili rezultate, ki so prikazani na sliki 4.



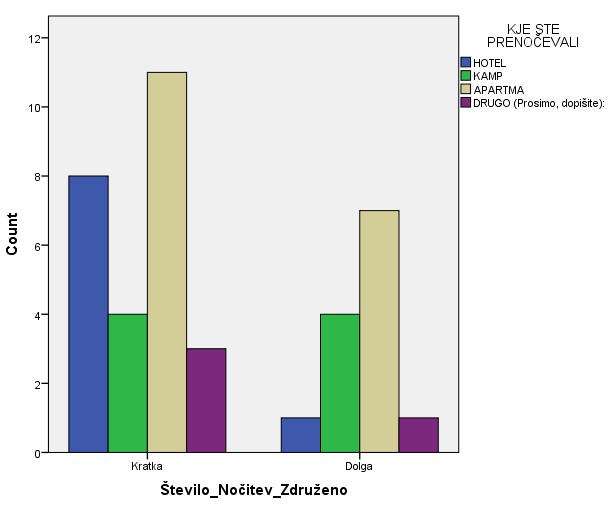
Slika 19: Tabela χ2 testa z združenimi kategorijami spremenljivke Število\_nočitev.

Glede na dobljene rezultate obeh hi-kvadrat testov lahko sklepamo, da sta spremenljivki v primeru, ko nismo združevali kategorij v spremenljivki *Število\_nočitev*,precej povezani. To lahko razberemo iz tabele na sliki 1, kjer vidimo da je p-vrednost χ2 testa enaka 0.015 < 0.05, kar pomeni, da lahko v prvem primeru ničelno hipotezo zavrnemo pri 5% stopnji značilnosti.

Drugačen rezultat pa smo dobili v drugem primeru, ko smo kategorije v spremenljivki *Število\_nočitev*  združevali. Na sliki 4 vidimo, da je p-vrednost χ2 v primeru združevanja kategorij enaka 0.337 < 0.05, kar pa pomeni, da v drugem primeru ničelne hipoteze ne moremo zavrniti. Rezultat nam torej pove, da spremenljivki v tem primeru nista povezani.

Omenimo lahko še, da sta oba testa nerelevantna, saj noben od njiju ne zadošča osnovnemu pogoju χ2 testa. Sklepamo, da je prišlo do kršitve tega pogoja zaradi premajhne količine podatkov.

Na sliki 5 je prikazana izbira nastanitve še s pomočjo stolpičnega grafa, kjer je bolj očitno razvidno, da se z izbiro daljšega obdobja počitnikovanja gostje raje odločijo za bivanje v apartmajih in kampih, kot pa v hotelih.



Slika 20: Stolpični graf izbire tipa nastanitve glede na trajanje počitnikovanja.

## Analiza 2. raziskovalnega vprašanja

# Multivariatne metode

## Predstavitev problema in raziskovalnih vprašanj

Problem razdelitve anketirancev oziroma enot v statistično značilne skupine bomo izvedli z drugo multivariatno metodo razvrščanje v skupine. Anketirance bomo razvrstili glede na spremenljivke o zadovoljstvu z nastanitvijo in glede na ceno nastanitve.

Zanimalo nas je zadovoljstvo anketirancev razvrščenih v zgoraj omenjene skupine glede na neodvisni spremenljivki in sicer na spol in vrsta počitniške nastanitve.

## Prva multivariatna metoda

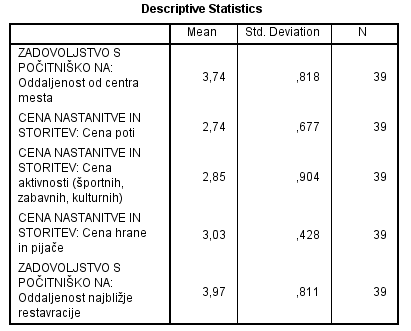
Pri analizi podatkov nas je zanimalo, kako bi lahko s podatki o cenah in zadovoljstvu o nekaterih storitvah napovedali oddaljenost lokacije od samega mestnega centra. Za preučevanje povezav med spremenljivkami smo uporabili multivariatno metodo, ki se imenuje multipla (večkratna) regresija.

Z multiplo regresijo napovedujemo vrednost kriterija (odvisne spremenljivke) na osnovi dveh ali več prediktorjev (neodvisnih spremenljivk). Z linearno regresijo torej iščemo linearno kombinacijo prediktorjev, katere vrednosti bodo čim bliže vrednostim kriterija (ta linearna kombinacija bo imela najvišjo možno korelacijo s kriterijem).

Za začetek smo za napovedovanje oddaljenosti od centra mesta uporabili vse spremenljivke, ki izražajo zadovoljstvo anketirancev s ceno ali ponudbo storitev na njihovem počitnikovanju. Podatke o ceni storitev smo merili na lestvici od 1 do 5, kjer je 1 najbolj poceni, 5 pa najdražje. Podatke o zadovoljstvu s storitvami in ponudbo smo prav tako merili na lestvici od 1 do 5, kjer 1 predstavlja nezadovoljstvo, 5 pa popolno zadovoljstvo.

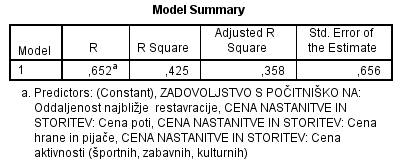
S pomočjo orodja SPSS smo izvedli linearno regresijo nad podatki in poizkušali poiskati najbolj vplivne spremenljivke, ki bi nam pomagale napovedati oddaljenost od centra mesta. Ugotovili smo, da na model najbolj vplivajo spremenljivke *Cena\_poti*, *Cena\_aktivnosti*, *Cena\_hrane\_in\_pijače* in *Oddaljenost\_najbližje\_restavracije*.

Na sliki 6 vidimo tabelo opisnih statistik, kjer vidimo osnovne podatke za odvisno in neodvisne spremenljivke. Povprečno zadovoljstvo z oddaljenostjo od centra mesta je 3.74, anketiranci pa so bili precej zadovoljni tudi z oddaljenostjo najbližje restavracije. Opazimo, da so bili anketiranci malce manj zadovoljni s ceno hrane in pijače, ceno zabavnih in kulturnih storitev in ceno poti.



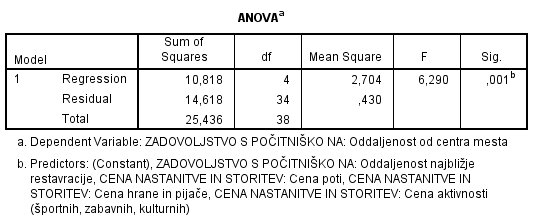
Slika 21: Tabela opisnih statistik.

Na sliki 7 je prikazana tabela, ki opisuje lastnosti našega modela. Vrednost *R*2 nam pove, da lahko izbrane neodvisne spremenljivke pojasnijo 42,5% variance spremenljivke *Oddaljenost\_od\_centra\_mesta*, kar je precej slabo, celo slabše od od golega ugibanja, kjer imamo 50% možnosti za pravilno napoved dogodka.



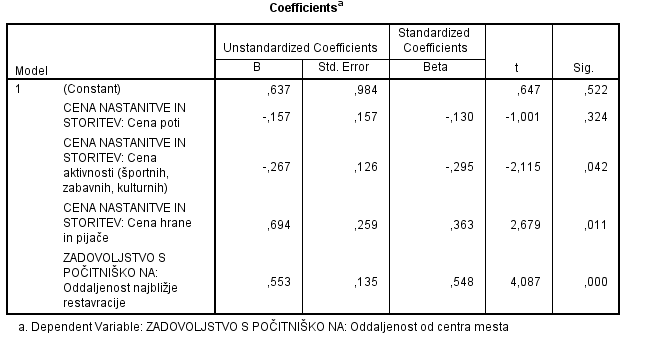
Slika 22: Lastnosti modela z linearno regresijo.

Tabela ANOVA, ki je prikazana na sliki 8 nam pove, ali naš model dobro opisuje podatke ali ne. Ker je p-vrednost enaka 0.001 < 0.05 lahko pri 5% tveganju trdimo, da je naš model precej dober in da statistično značilno bolje napove oddaljenost od centra mesta kot če bi vzeli le povprečno zadovoljstvo z oddaljenostjo od mesta.



Slika 23: Tabela ANOVA za linearno regresijo.

Najpomembnejša tabela pri linearni regresiji je prikazana na sliki 9. To je tabela koeficientov regresijskega modela, ki kažejo, kako in za koliko posamezna neodvisna spremenljivka vpliva na določite odvisne spremenljivke. Vidimo, da nam model izpostavi 2 spremenljivki, ki imata pomemben vpliv na določanje oddaljenosti od centra mesta. Višja kot je cena hrane in pijače in manjša kot je oddaljenost do najbližje restavracije, bližje smo centru mesta.



Slika 24: Tabela koeficientov pri linearni regresiji.

S pomočjo regresijskega modela in pridobljenih koeficientov lahko zapišemo tudi regresijsko enačbo, ki se v našem primeru glasi:

*Y1 = 0.637 – 0.157\*X1 – 0.267\*X2 + 0.694\*X3 + 0.553\*X4,* kjer so

*Y1....oddaljenost\_od\_centra\_mesta*

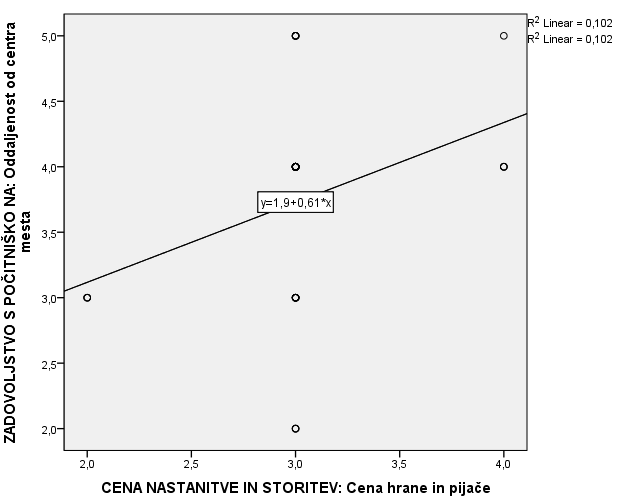
*X1….cena\_poti*

*X2…cena\_aktivnosti*

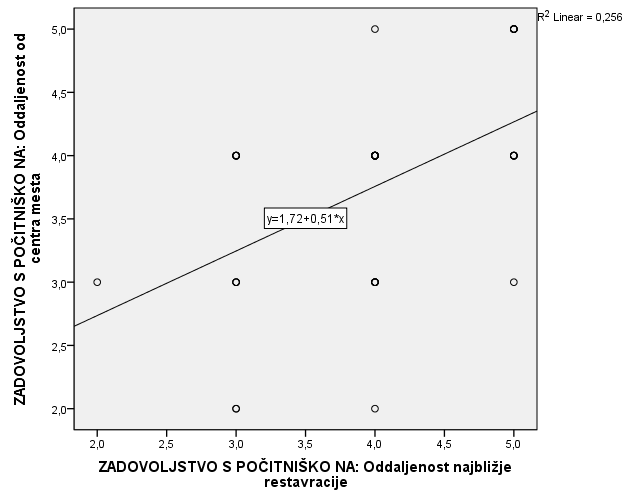
*X3…cena\_hrane\_in\_pijače*

*X4…oddaljenost\_najbližje\_restavracije*

Na slikah 10 in 11 je linearna odvisnost med neodvisnima spremenljivkama *Cena\_hrane\_in\_pijače* in *Oddaljenost\_od\_najbližje\_restavracije* v primerjavi z odvisno spremenljivko *Oddaljenost\_od\_centra\_mesta* prikazana še grafično.



Slika 25: Vpliv cene hrane in pijače na oddaljenost od mesta.



Slika 26:Vpliv oddaljenosti restavracije na oddaljenost od centra mesta.

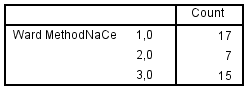
## Druga multivariatna metoda

Uporabili smo multivariatno metodo, ki se imenuje razvrščanje v skupine. Metodo razvrščanje v skupine uporabljamo za proučevanje odvisnosti med statističnimi enotami. Po v naprej določenem kriteriju, med seboj podobne enote združimo v skupine, ki niso v naprej znane. Med skupinami pa so si različne. Vsaka enota pripada le eni skupini, skupine se ne prekrivajo. (Žnidaršič, 2018)

Poznamo dve skupini metod razvrščanja. Hierarhični in Nehierarhični skupini metod. Pri hierarhičnih metodah skupine oblikujemo postopoma in njihovo število ni vnaprej znano. Pri nehierarhičnih metodah število skupin izberemo vnaprej in vsako enoto iz množice uvrstimo v tisto skupino, ki je glede na izbrano mero različnosti najbližja. (Žnidaršič, 2018)

Odločili smo se za uporabo hierarhičnih metod bolj natančno za razvrščanje z Wardovo metodo. Wardova metoda oblikuje skupine na podlagi vsote kvadratov razlik, tako, da so enote znotraj skupin kar najbolj homogene. (Žnidaršič, 2018)

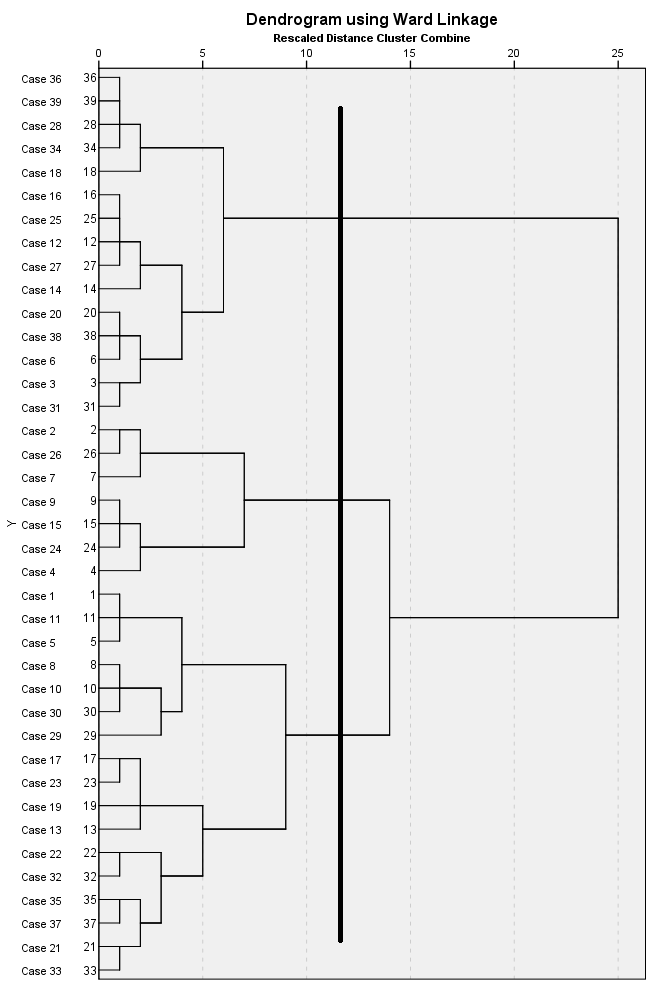
Z uporabo programa SPSS in hierarhičnega razvrščanja ter Wardove metode smo glede na spremenljivke *Oddaljenost najbližje trgovine, Oddaljenost najbližje restavracije, Oddaljenost od centra mesta, Oddaljenost od rekreacijskih površin in Ceno nastanitve* enoterazvrstili v 3 skupine. V prvo skupino se je razvrstilo 17 enot, v drugo skupino 7 enot in v tretjo skupino 15 enot. Število razvrščenih enot v posamezno skupino je prikazano na sliki 13 prav tako pa je prikazano na dendogramu.



Slika 27 Število enot v skupinah

Potek združevanja lahko grafično ponazorimo z drevesom združevanja oziroma dendogramom. Listi tega drevesa so enote, točke združitve pa sestavljene skupine: levi in desni naslednik vsake točke sta skupini, iz katerih je nastala. Višina točke, ki jo imenujemo nivo združevanja, je sorazmerna meri različnosti med skupinama. (Ferligoj, 1989).

Dendogram za naš primer oziroma drevo združevanja z Wardovo metodo je viden na sliki 14. Rez, ki loči skupine, je prikazan z odebeljeno črno črto. Razdaljo med enotami smo merili z kvadratom Evklidske razdalje. Spremenljivke smo standardizirali.



Slika 28 Dendogram

Razdelitev v tri skupine je jasno vidna na drevesu razvrščanja, ki je prikazan na sliki 14.

V prvo skupino , so se razvrstile enote, ki so delno zadovoljne, delno nezadovoljne z oddaljenostjo najbližje trgovine, so zadovoljne z oddaljenostjo najbližje restavracije in oddaljenostjo nastanitve od centra mesta. Z oddaljenostjo rekreacijskih površin pa so anketiranci zadovoljni in zelo zadovoljni. Cena nastanitve je za enote v prvi skupini poceni oziroma sprejemljiva. Skupino smo poimenovali »zadovoljni«.

(N=17)

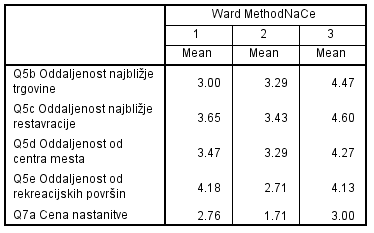
V drugo skupino so se razvrstile enote, ki so delno zadovoljne, delno nezadovoljne z oddaljenostjo najbližje trgovine, nekaj enot pa je z oddaljenostjo najbližje trgovine zadovoljnih. Podobno zadovoljstvo imajo enote za vprašanji oddaljenost najbližje restavracije in oddaljenost od centra mesta. Enote v tej skupini so nezadovoljne in delno zadovoljne z oddaljenostjo rekreacijskih površin. Cena nastanitve pa se jim zdi zelo ugodna. Skupino smo poimenovali »manj zadovoljni«.

(N=7)

V tretjo skupino so se razvrstile enote, ki so zadovoljne in zelo zadovoljne z oddaljenostjo najbližje trgovine, so zelo zadovoljne z oddaljenostjo najbližje restavracije. Zadovoljne z oddaljenostjo od centra mesta, prav tako z oddaljenostjo rekreacijskih površin. Cena nastanitve pa se jih zdi sprejemljiva. Skupino smo poimenovali »zelo zadovoljni«.

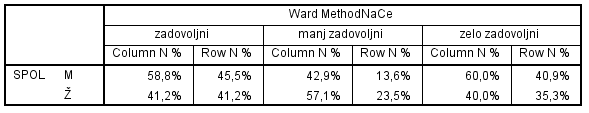
(N=15)

Zgoraj opisane vrednosti so prikazane v tabeli na sliki 15. Iz razvrstitve skupin lahko sklepamo, da cene nastanitve vpliva na zadovoljstvo o nastanitvi. Cenejša kot je nastanitev bolj je oddaljena ostala infrastruktura (rekreacijske površine, center mesta, trgovine).



Slika 29 povprečja skupin glede na zadovoljstvo z nastanitvijo in ceno nastanitve

Zanima nas zadovoljstvo anketirancev glede na spol. Zato primerjamo spremenljivko spol med skupinami. Na anketo je odgovarjalo 17 žensk in 22 moških. Iz tabele prikazane na sliki 16 lahko razberemo, da so v skupini »zadovoljni« v večjem deležu moški prav tako v skupini »zelo zadovoljni«, v skupini »manj zadovoljni« pa prevladujejo ženske. Zgornje trditve veljajo tudi, če izračunamo deleže po skupinah na celotno populacijo žensk in moških. Ne glede na dejstvo, da je na anketo odgovarjalo več moških kot žensk. Deleži so prikazani v tabeli na sliki 16.



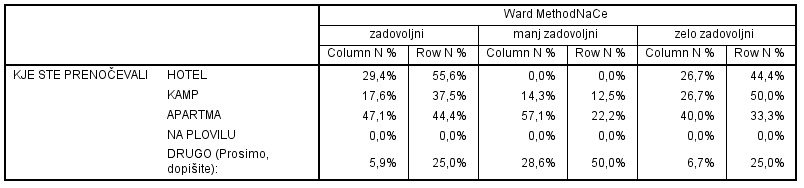
Slika 30 Zadovoljstvo po skupinah glede na spol

Prav tako nas zanima zadovoljstvo anketirancev glede na to kje so prenočevali. Če deleže izračunamo glede na število vseh odgovorov za posamezno vrsto nastanitve, lahko iz tabele, ki je prikazana na sliki 17 razberemo, da je več kot polovica, kar (56%) anketirancev v skupini »zadovoljni« prenočevala v hotelu. Sledijo prenočevanje v apartmaju z (44%) in prenočevanje kampu z (38%).

Anketiranci v skupini »manj zadovoljni« so v polovici primerov (50%) prenočevali preko sistema »airbnb«, v lastni ali prijateljevi nastanitvi, (22%) jih je prenočevalo v apartmaju, (13%) pa jih je prenočevalo v kampu. V dotični skupini prenočitev v hotelu ni bilo.

V skupini, ki smo jo poimenovali »zelo zadovoljni« so v polovici primerov (50%) anketiranci prenočevali v kampu. Na drugem mestu je s (44%) prenočevanje v hotelu. Sledita prenočevanje v apartmaju z (33%) in prenočevanje preko sistema »airbnb«, v lastni ali prijateljevi nastanitvi z (25%).

Anketiranci niso prenočevali na plovilu.



Slika 31 Zadovoljstvo glede na vrsto nastanitve

# Kritična analiza zasnove in izvedbe raziskave (Jure)

# Zaključek (skupaj)

# Literatura

Benstat, Hi-kvadrat test

Dostopno na:

<http://www.benstat.si/blog/hi-kvadrat-test>

Statistik.si, Hi-kvadrat test

Dostopno na:

<http://www.statistik.si/storitve/statisticne-analize-testi/hi-kvadrat-test/>

Žnidaršič, A (2018). Prosojnice iz predavanj *RAZVRŠČANJE V SKUPINE Statistika v raziskovanju (MAG).* Kranj: Fakulteta za organizacijske vede, Kranj

Ferligoj, A. (1989). *RAZVRŠČANJE V SKUPINE*. Ljubljana: Raziskovalni inštitut, Fakulteta za sociologijo, politične vede in novinarstvo, Kardeljeva pl. 5, Ljubljana

# Priloge

# A. Vprašalnik