

Statistička analiza podataka koji opisuju interakcije čovjeka i robota

Tkalec, Anabela

Master's thesis / Diplomski rad

2023

Degree Grantor / Ustanova koja je dodijelila akademski / stručni stupanj: **University of Zagreb, Faculty of Mechanical Engineering and Naval Architecture / Sveučilište u Zagrebu, Fakultet strojarstva i brodogradnje**

Permanent link / Trajna poveznica: <https://urn.nsk.hr/urn:nbn:hr:235:779294>

Rights / Prava: [In copyright](#) / [Zaštićeno autorskim pravom.](#)

Download date / Datum preuzimanja: **2024-05-16**

Repository / Repozitorij:

[Repository of Faculty of Mechanical Engineering and Naval Architecture University of Zagreb](#)



SVEUČILIŠTE U ZAGREBU
FAKULTET STROJARSTVA I BRODOGRADNJE

DIPLOMSKI RAD

Anabela Tkalec

Zagreb, 2023.

SVEUČILIŠTE U ZAGREBU
FAKULTET STROJARSTVA I BRODOGRADNJE

DIPLOMSKI RAD

Mentor:

izv. prof. dr. sc. Tomislav Stipančić, dipl. ing.

Student:

Anabela Tkalec

Zagreb, 2023.

Izjavljujem da sam ovaj rad izradila samostalno koristeći znanja stečena tijekom studija i navedenu literaturu.

Zahvaljujem se mentoru izv. prof. dr. sc. Tomislavu Stipančiću na trudu koji je uložio u pružanju konstruktivnih savjeta, odgovaranju na pitanja i sveukupnoj podršci tijekom pisanja ovog rada. Vaša stručnost, ljubaznost i optimizam uvelike su mi olakšali ovaj proces, stoga Vam najviše zahvaljujem na tome.

Hvala i mami, tati, Anamariji, Mateju i svim prijateljima.

Anabela Tkalec



SVEUČILIŠTE U ZAGREBU
FAKULTET STROJARSTVA I BRODOGRADNJE



Središnje povjerenstvo za završne i diplomske ispite
Povjerenstvo za diplomske radove studija strojarstva za smjerove:
proizvodno inženjerstvo, računalno inženjerstvo, industrijsko inženjerstvo i menadžment,
inženjerstvo materijala te mehatronika i robotika

Sveučilište u Zagrebu Fakultet strojarstva i brodogradnje	
Datum:	Prilog:
Klasa: 602 – 04 / 23 – 6 / 1	
Ur. broj: 15 - 1703 - 23 -	

DIPLOMSKI ZADATAK

Student: ANABELA TKALEC

Mat. br.: 0035214252

Naslov rada na hrvatskom jeziku: **Statistička analiza podataka koji opisuju interakcije čovjeka i robota**

Naslov rada na engleskom jeziku: **Statistical analysis of data describing human-robot interactions**

Opis zadatka:

Statistička analiza je vrlo važan alat koji služi za proučavanje trendova, obrazaca te uzročnih veza korištenjem kvantitativnih podataka.

PLEA je afektivni robot razvijen na Fakultetu strojarstva i brodogradnje koji se koristi za proučavanje interakcije ljudi i robota.

U radu je potrebno primijeniti metode statističke analize na podacima prikupljenim tijekom interakcije afektivnog robota PLEA i ljudi na Art & AI festivalu u Leicesteru, Velika Britanija, s ciljem boljeg razumijevanja tih odnosa. Podaci su prikupljeni tijekom tri mjeseca trajanja festivala.

U okviru diplomskog rada je potrebno:

- isplanirati tijek istraživanja
- postaviti početne pretpostavke koje dovode u vezu dominantne i početne emocije s trajanjem interakcije
- podatke organizirati i sažeti koristeći alate opisne statistike
- proučiti vezu između dominantnih i početno izraženih emocija te duljine trajanja interakcije
- interpretirati rezultate analize te donijeti zaključke o testiranim hipotezama.

U radu je potrebno navesti korištenu literaturu te eventualno dobivenu pomoć.

Zadatak zadan:
19. siječnja 2023.

Rok predaje rada:
23. ožujka 2023.

Predviđeni datum obrane:
27. ožujka do 31. ožujka 2023.

Zadatak zadao:
doc. dr. sc. Tomislav Stipančić

Predsjednica Povjerenstva:
prof. dr. sc. Biserka Runje

SADRŽAJ

SADRŽAJ.....	I
POPIS SLIKA.....	III
POPIS TABLICA.....	V
POPIS OZNAKA.....	VI
SAŽETAK.....	VII
SUMMARY.....	VIII
1. UVOD - PLEA.....	1
2. ISTRAŽIVAČKI PROJEKT.....	2
2.1. Hipoteze.....	2
2.2. Strategija istraživanja.....	2
2.2.1. Korelacijsko istraživanje.....	2
2.2.2. Deskriptivno istraživanje.....	3
2.3. Prikupljanje podataka.....	3
2.3.1. Izračun dovoljne veličine uzorka.....	4
2.4. Varijable.....	5
2.4.1. Mjerna ljestvica.....	5
2.4.2. Primjer prikupljenih podataka.....	6
2.4.3. Čišćenje podataka.....	7
2.4.3.1. Ekstremne vrijednosti.....	8
3. OBRADA PRIKUPLJENIH PODATAKA.....	9
3.1. Trajanje interakcija.....	9
3.2. Podaci vezani za emocije.....	12
3.2.1. Ljutnja (angry).....	14
3.2.2. Tuga (sad).....	15
3.2.3. Neutralnost (neutral).....	16
3.2.4. Iznenadenost (surprised).....	17
3.2.5. Strah (fear).....	18
3.2.6. Gađenje (disgust).....	19
3.2.7. Sreća (happy).....	20
3.2.8. Usporedba vrijednosti emocija.....	21
4. TESTIRANJE HIPOTEZA.....	22
4.1. Prva hipoteza.....	22
4.1.1. Odnos trajanja interakcije i dominantnih emocija.....	23
4.1.1.1. Ljutnja (angry).....	28
4.1.1.2. Tuga (sad).....	29
4.1.1.3. Neutralnost (neutral).....	30
4.1.1.4. Iznenadenost (surprised).....	31
4.1.1.5. Strah (fear).....	32
4.1.1.6. Gađenje (disgust).....	33
4.1.1.7. Sreća (happy).....	34
4.1.2. Usporedba kategorija trajanja.....	35
4.2. Mann-Whitneyjev U test.....	36

4.2.1. Testiranje prve hipoteze.....	39
4.2.2. Korelacija.....	42
4.2.3. Problemi korelacijskog istraživanja.....	44
4.3. Druga hipoteza.....	44
4.3.1. Odnos trajanja interakcije i prve izražene emocije tijekom interakcije.....	45
4.3.1.1. Ljutnja (angry).....	49
4.3.1.2. Tuga (sad).....	50
4.3.1.3. Neutralnost (neutral).....	51
4.3.1.4. Iznenadenost (surprised).....	52
4.3.1.5. Strah (fear).....	53
4.3.1.6. Gađenje (disgust).....	54
4.3.1.7. Sreća (happy).....	55
4.3.2. Usporedba kategorija trajanja.....	56
4.3.3. Testiranje druge hipoteze.....	57
5. SAŽETAK REZULTATA.....	61
LITERATURA.....	63
PRILOZI.....	64

POPIS SLIKA

Slika 1.	PLEA izložena na ART AI festivalu.....	1
Slika 2.	Izračun potrebne veličine uzorka.....	5
Slika 3.	Graf vrijednosti emocije kroz interakciju.....	7
Slika 4.	Graf broja interakcija određenog trajanja.....	8
Slika 5.	Prikaz broja interakcija određenog trajanja (grafički prikaz violine).....	10
Slika 6.	Histogram trajanja interakcija.....	10
Slika 7.	Prikaz interkvartilne razlike trajanja.....	11
Slika 8.	Zastupljenost kategorija duljine trajanja interakcija.....	12
Slika 9.	Zastupljenost pojedine emocije u svim interakcijama.....	13
Slika 10.	Grafički prikazi vrijednosti emocije ljutnje.....	14
Slika 11.	Grafički prikazi vrijednosti emocije ljutnje.....	15
Slika 12.	Grafički prikazi vrijednosti emocije neutralnosti.....	16
Slika 13.	Grafički prikazi vrijednosti emocije iznenađenosti.....	17
Slika 14.	Grafički prikazi vrijednosti emocije straha.....	18
Slika 15.	Grafički prikazi vrijednosti emocije straha.....	19
Slika 16.	Grafički prikazi vrijednosti emocije sreće.....	20
Slika 17.	Zastupljenost interakcija sa pojedinom emocijom kao dominantnom.....	23
Slika 18.	Raspodjela prosječnih emocija kod sudionika broj 3 (sreća kao dominantna emocija).....	24
Slika 19.	Vrijednosti emocija kroz interakciju sudionika broj 3 (sreća kao dominantna emocija).....	25
Slika 20.	Raspodjela prosječnih emocija kod sudionika broj 121 (neutralnost kao dominantna emocija).....	26
Slika 21.	Vrijednosti emocija kroz interakciju sudionika broj 121 (neutralnost kao dominantna emocija).....	26
Slika 22.	Trajanje interakcije u ovisnosti sa dominantnom emocijom.....	27
Slika 23.	Grafički prikazi trajanja interakcija sa dominantnom emocijom ljutnje.....	28
Slika 24.	Grafički prikazi trajanja interakcija sa dominantnom emocijom tuge.....	29
Slika 25.	Grafički prikazi trajanja interakcija sa neutralnom dominantnom emocijom.....	30
Slika 26.	Grafički prikazi trajanja interakcija sa dominantnom emocijom iznenađenosti.....	31
Slika 27.	Grafički prikazi trajanja interakcija sa dominantnom emocijom straha.....	32
Slika 28.	Grafički prikazi trajanja interakcija sa dominantnom emocijom gađenja.....	33
Slika 29.	Grafički prikazi trajanja interakcija sa dominantnom emocijom sreće.....	34
Slika 30.	Zastupljenost dominantnih emocija u pojedinoj kategoriji trajanja (I).....	35
Slika 31.	Zastupljenost dominantnih emocija u pojedinoj kategoriji trajanja (II).....	36
Slika 32.	Podaci interakcija u SPSS softveru.....	37
Slika 33.	Odabir neparametrijskog testa.....	38
Slika 34.	Definiranje varijabli statističkog testa.....	38
Slika 35.	Definiranje grupa nezavisne varijable.....	39
Slika 36.	Histogrami trajanja interakcija (prva hipoteza).....	40
Slika 37.	Zastupljenost kategorija trajanja kod pozitivnih i negativnih dominantnih emocija.....	40
Slika 38.	Broj interakcija pojedine grupe (<i>SPSS Statistics</i>).....	42
Slika 39.	Ovisnost duljine trajanja interakcije o udjelu pozitivnih emocija.....	42

Slika 40.	Odabir računanja Kendallovog tau-b u <i>SPSS Statistics</i>	43
Slika 41.	Odabir postavki računanja Kendallovog tau-b.....	43
Slika 42.	Distribucija interakcija sa pojedinom emocijom kao prvom izraženom u interakciji.....	45
Slika 43.	Vrijednosti emocija kroz interakciju sudionika broj 63 (neutralnost kao prva izražena emocija).....	46
Slika 44.	Vrijednosti emocija kroz interakciju sudionika broj 279 (iznenađenost kao prva izražena emocija).....	47
Slika 45.	Trajanje interakcije u ovisnosti sa prvom izraženom emocijom.....	48
Slika 46.	Grafički prikazi trajanja interakcija sa prvom izraženom emocijom ljutnje.....	49
Slika 47.	Grafički prikazi trajanja interakcija sa prvom izraženom emocijom tuge.....	50
Slika 48.	Grafički prikazi trajanja interakcija sa prvom izraženom emocijom neutralnosti.....	51
Slika 49.	Grafički prikazi trajanja interakcija sa prvom izraženom emocijom iznenađenosti.....	52
Slika 50.	Grafički prikazi trajanja interakcija sa prvom izraženom emocijom straha.....	53
Slika 51.	Grafički prikazi trajanja interakcija sa prvom izraženom emocijom gađenja.....	54
Slika 52.	Grafički prikazi trajanja interakcija sa prvom izraženom emocijom sreće.....	55
Slika 53.	Zastupljenost prvih izraženih emocija u pojedinoj kategoriji trajanja (I).....	56
Slika 54.	Zastupljenost prvih izraženih emocija u pojedinoj kategoriji trajanja (II).....	57
Slika 55.	Histogrami trajanja interakcija (druga hipoteza).....	58
Slika 56.	Zastupljenost kategorija trajanja kod pozitivnih i negativnih prvih izraženih emocija.....	58
Slika 57.	Broj pojedine grupe interakcija (druga hipoteza).....	60

POPIS TABLICA

Tablica 1.	Povezanost varijabli.....	3
Tablica 2.	Dio podataka jednog od sudionika.....	6
Tablica 3.	Podaci o trajanju interakcija.....	9
Tablica 4.	Broj interakcija određenog trajanja.....	10
Tablica 5.	Kategorije interakcija.....	12
Tablica 6.	Podaci o vrijednostima emocija.....	12
Tablica 7.	Podaci o prosječnim vrijednostima emocija sudionika broj 3 (sreća kao dominantna emocija).....	24
Tablica 8.	Podaci o prosječnim vrijednostima emocija sudionika broj 121 (neutralnost kao dominantna emocija).....	25
Tablica 9.	Deskriptivna statistika (prva hipoteza).....	39
Tablica 10.	Rangovi statističkog testa (prva hipoteza).....	41
Tablica 11.	Statistika provedenog testa (prva hipoteza).....	41
Tablica 12.	Rezultati provjere Kendallovog tau-b.....	43
Tablica 13.	Prosječne vrijednosti emocija u prve 2 sekunde interakcije sudionika broj 63 (neutralnost kao prva izražena emocija).....	46
Tablica 14.	Prosječne vrijednosti emocija u prve 2 sekunde interakcije sudionika broj 279 (iznenađenost kao prva izražena emocija).....	47
Tablica 15.	Deskriptivna statistika (druga hipoteza).....	57
Tablica 16.	Rangovi statističkog testa (druga hipoteza).....	59
Tablica 17.	Statistika provedenog testa (druga hipoteza).....	59

POPIS OZNAKA

Oznaka	Jedinica	Opis
sk_2	-	Pearsonov koeficijent iskrivljenosti
Q_1	-	donji kvartil
Q_3	-	gornji kvartil
τ_b	-	Kendallov tau-b koeficijent korelacije

SAŽETAK

Statistička analiza je vrlo važan alat koji služi za proučavanje trendova, obrazaca ponašanja te uzročnih veza korištenjem kvantitativnih podataka. U ovom se radu ona koristi na podacima prikupljenim tijekom interakcije afektivnog robota PLEA i ljudi s ciljem boljeg razumijevanja tog odnosa. Za donošenje bilo kakvih ispravnih zaključaka potrebno je pažljivo isplanirati tijek istraživanja, način prikupljanja uzoraka te njihovu količinu. Nakon toga, potrebno je prikupljene podatke organizirati i sažeti uz pomoć opisne statistike. Početne hipoteze i pretpostavke se testiraju za cijelu željenu populaciju. Na kraju, se ti zaključci interpretiraju i generaliziraju. U ovom se radu proučavaju emocije tijekom interakcije te ovisnost njenog trajanja o razvijenim emocijama. Cilj je proučiti vezu između dominantne emocije u cijeloj interakciji i duljine trajanje interakcije te odnos prvog dojma o robotu, tj. prve emocije koju ljudi osjete tijekom interakcije i duljine trajanja. Za obradu podataka korišteni su programski jezik Python i programski softver SPSS Statistics.

SUMMARY

Statistical analysis is a very important tool used to study trends, patterns of behavior and causal relationships using quantitative data. In this paper, it is used on the data collected during interactions between the affective robot PLEA and humans with the aim of better understanding this relationship. In order to draw any correct conclusions, it is necessary to carefully plan the course of the research, the method of collecting samples and their quantity. After that, it is necessary to organize and summarize the collected data with the help of descriptive statistics. Initial hypotheses and assumptions are tested for the entire desired population. Finally, these conclusions are interpreted and generalized. In this paper, the emotions during the interaction and the dependence of the interaction's duration on the developed emotions are studied. The first goal is to study the relationship between the dominant emotions throughout the interaction and the duration of the interaction. The second one is to study the relationship between the first impression of the robot and the duration. Data processing is done using the python programming language and the SPSS Statistics software.

1. UVOD - PLEA

PLEA je afektivna robotska glava koja je razvijena za analizu i primjenu ponašanja koristeći oblik biomimikrije. Prilikom interakcije, PLEA analizira različite izvore društvenih signala i znakova - to uključuje emocije lica, glasnoću u prostoriji i intenzitet pokreta tijela. Također analizira osjećaje u govoru koji čuje. Podaci primljeni od čovjeka, robota i okoline se integriraju u sustavu te se kontekstualne informacije pretvaraju u podatke. To omogućuje donošenje odluka koje se temelje na novostečenim informacijama koje su u trenutku donošenja odluke o načinu ponašanja nepotpune. PLEA reagira i odgovara putem neverbalne ekspresije lica koristeći tehniku vizualizacije podataka. Koriste se principi virtualne i uronjene stvarnosti da bi se na licu robota prikazalo računalom generirano animirano lica virtualnog agenta. [1]

S ciljem prikaza najnovijih razvoja u afektivnoj robotici i komunikaciji robota sa ljudima, PLEA je izložena na ART AI festivalu u Leicesteru, u Velikoj Britaniji 2021. godine [Slika 1]. Tamo su se ujedno i prikupljali podaci interakcija koji su korišteni u ovom radu. Glavni je cilj razumijevanja odnosa, ali i razvoj naprednijih sposobnosti zaključivanja kroz proces kontroliranog učenja.



Slika 1. PLEA izložena na ART AI festivalu

2. ISTRAŽIVAČKI PROJEKT

2.1. Hipoteze

Cilj ovog istraživanja je temeljito istražiti razvoj emocija tijekom interakcije s robotom, uključujući identifikaciju dominantne emocije koju ta interakcija izaziva, prosječno trajanje same interakcije i povezanost između vrste emocija i trajanja. Također želimo istražiti utjecaj "prvog dojma" robota na duljinu interakcije i odnos najizraženije, tj. dominantne emocije i trajanja interakcije. Svi ovi ciljevi istraživanja postavljeni su kao hipoteze koje će biti podvrgnute detaljnoj analizi:

1. nulta hipoteza: Dominantna emocija nema utjecaja na trajanje interakcije.

1. alternativna hipoteza: Dominantna emocija utječe na trajanje interakcije.

2. nulta hipoteza: Prva izražena emocija nema utjecaja na trajanje interakcije.

2. alternativna hipoteza: Prva izražena emocija utječe na trajanje interakcije.

Kako bi se mogle potvrditi ili negirati hipoteze, prvo je potrebno obraditi podatke te odrediti trajanje svake interakcije, odrediti dominantne emocije i emocije izražene na početku interakcija.

2.2. Strategija istraživanja

Za postizanje uspješnih rezultata istraživanja, potrebna je prikladna strategija prikupljanja podataka i njihove analize, kao i odabir odgovarajućih statističkih testova za testiranje hipoteza. U ovom radu se primjenjuje korelacijsko (*correlation design*) i deskriptivno istraživanje (*descriptive design*) jer omogućuju mjerenje varijabli i definiranje njihovih međusobnih odnosa. Svi se sudionici u istraživanju proučavaju na grupnoj razini na isti način.

2.2.1. Korelacijsko istraživanje

Korelacijsko istraživanje služi za proučavanja odnosa između varijabli bez pretpostavki o međusobnoj uzročnosti i bez njihovog kontroliranja ili mijenjanja. Ono uključuje korištenje korelacijskih koeficijenta (*correlation coefficients*) te ispitivanje značajnosti (*significance tests*). Takav pristup prilikom istraživanja koristan je za brzo skupljanje podataka u prirodnom okruženju što pripomaže prilikom generalizacija zaključaka za stvarne situacije [2].

Korelacijsko istraživanje se u ovom se radu koristi prilikom provjeravanja veze između vrste dominantnih emocija tijekom cijele interakcije i trajanja interakcije. Korelacija između varijabli može biti pozitivna, negativna ili nulta [Tablica 1].

Tablica 1. Povezanost varijabli

Vrsta	Opis	Primjer
pozitivna povezanost	obje varijable se mijenjaju u istom smjeru	S povećanjem postotka pozitivnih dominantnih emocija, trajanje interakcije se povećava.
negativna povezanost	varijable se mijenjaju u suprotnom smjeru	S povećanjem postotka pozitivnih dominantnih emocija, trajanje interakcije se smanjuje.
nulta povezanost	ne postoji veza između varijabli	Ne postoji veza između trajanja interakcije i dominantne emocije

2.2.2. Deskriptivno istraživanje

Uz pomoć deskriptivnog istraživanja proučavaju se i opisuju karakteristike populacije, situacije ili pojave koristeći statistička ispitivanja (*statistical test*) za donošenje zaključaka iz ispitnih uzoraka. Ova metoda odgovara na pitanja *tko*, *što*, *gdje* i *kako*, ali ne i *zašto* te je zato korisna kada se ne zna mnogo o temi i problemu koji se istražuje. Kao ni kod korelacijskog pristupa, ne kontroliraju se niti mijenjaju varijable već se samo proučavaju i mjere [3].

U ovom se istraživanju, ovaj pristup koristi za određivanje dominantne emocije sudionika te prosječnog trajanja interakcije.

2.3. Prikupljanje podataka

PLEA je bila izložena na ART-AI festivalu u knjižnici u Leicesteru, u Velikoj Britaniji gdje su joj ljudi mogli slobodno prići te komunicirati s njom. Ona je u razdoblju od tri mjeseca, točnije od 9. kolovoza do 4. listopada 2021., spremala podatke tih interakcija. Podaci su se prikupljali metodom promatranja (*observation*) kako bi se ispitala interakcija u najprirodnijem mogućem okruženju bez bilo kakvog utjecaja na uzorak. Budući da je cilj razviti neko početno razumijevanje u još nedovoljno istraženoj temi, a ne testirati hipotezu na širokoj populaciji, koristila se metoda uzorkovanja bez vjerojatnosti (*non-probability sampling*) [4].

Na temelju činjenice da su se zatekli na festivalu gdje je PLEA bila izložena, ljudi su volontiranjem postali dio uzorka. Radi se o kombinaciji prigodnog uzorka (*convenience sampling*) i dobrovoljnog uzorkovanja (*voluntary response sampling*). Prigodan uzorak

uključuje pojedince koji su istraživaču najpristupačniji. U ovom slučaju, to su oni koji su se nalazili na festivalu. Ovo je jednostavan i jeftin način prikupljanja početnih podataka, ali ne postoji način da se utvrdi je li uzorak reprezentativan za populaciju, tako da ne može proizvesti potpuno generalizirajuće rezultate [5]. Slično prikladnom uzorku, dobrovoljno uzorkovanje uglavnom se temelji na jednostavnosti pristupa. Umjesto da istraživač bira sudionike i izravno ih kontaktira, ljudi sami volontiraju. Ti su uzorci uvijek barem donekle pristrani, jer su neki ljudi skloniji volontiranju od drugih [6].

2.3.1. Izračun dovoljne veličine uzorka

Za računanje minimalne veličine uzorka ljudi korišten je online *sample size* kalkulator [Slika 1] [7]. Odabirom određenih parametara računa se minimalan broj ljudi koji je potreban za istraživanje. Budući da se rade dvije vrste istraživanja test se provodi za obje vrste te se kao minimalni broj sudionika odabire veći broj. Potrebno je odabrati porodicu ispitivanja (*test family*), broj grupa (*sample groups*), tip testa (je li jednostrani ili dvostrani) (*number of tails*), razinu značaja (α) (*significance level*), statističku moć (*power*), korelacijski koeficijent (*correlation coefficient*) i veličinu učinka (*effect size*) [Slika 2] .

Prvo se računa minimalna veličina uzorka za korelacijsko istraživanje – zato je za porodicu ispitivanja odabran *correlation* (način ispitivanja odnosa između dvije varijable). Hoće li tip testa biti jednostrani ili dvostrani ovisi o tome ima li hipoteza implicirani smjer. Budući da ga obje hipoteze koje se provjeravaju nemaju, odabrana je opcija *two tail*. Korelacijski koeficijent označava pretpostavljenu relativnu mjeru korelacije dvije veličine te je odbrana vrijednost 0.4 jer se pretpostavlja da će odnos varijabli biti pozitivan. Razina značaja (α) je vjerojatnost lažnog odbacivanja nulte hipoteze, a najčešća vrijednost je 0,05. Statistička moć je sposobnost proučavanja da otkrije rezultat koji postoji u prirodi. Općenito je poželjno se da snaga bude što veća, međutim, postavljanje previsoke vrijednosti može rezultirati veličinom uzorka koja nije praktična. U praksi se često koristi vrijednost od 0,8. Za te je parametre minimalna potrebna veličina uzorka 40 sudionika.

Nakon toga, potrebno je izračunati broj sudionika za deskriptivno istraživanje. U tom se slučaju za porodicu testa odabire *t-test*, za broj grupa *same subjects* te se za vrijednost veličine učinka odabire vrijednost od 0,5 (srednji učinak). Sa tim parametrima dobiven je broj sudionika 34 [8][9].

To znači da je za istraživanje potrebno minimalno 40 sudionika što je i zadovoljeno budući da se u ovom radu obrađuju podaci njih 290.

The total number of participants: 40

Test family: Correlation

Sample groups: Independent groups

Number of tails: Two

Correlation co-efficient: 0.4

Significance level (α): 0.05

Power: 0.8

Submit

The total number of participants: 34

Test family: t-test

Sample groups: Same subjects

Number of tails: Two

Effect size: 0.5

Significance level (α): 0.05

Power: 0.8

Submit

Slika 2. Izračun potrebne veličine uzorka

2.4. Varijable

Cijelo se istraživanje temelji na proučavanju emocija koje su prisutne tijekom interakcija. Emocija je za razliku od visine ili godina apstraktni pojam te se ne može na jednostavan način izmjeriti. Zbog toga se koristi proces operacionalizacije (*operationalization*) kojim se apstraktni pojmovi pretvaraju u mjerljive [8]. To je postignuto kvantifikacijom emocija putem skale sa vrijednostima od 0 do 10000. U istraživanju sustav robota mjeri sedam različitih emocija. To su ljutnja (*angry*), tuga (*sad*), neutralno (*neutral*), iznenađenost (*surprised*), strah (*fear*), gađenje (*disgust*) i sreća (*happy*). U svakom se trenutku mjeri brojčana vrijednost svake pojedine emocije u obliku vrijednosti između 0 i 10000. 0 predstavlja pravu nultu vrijednost, odnosno potpuni nedostatak određene emocije. Tijekom interakcije se spremaju i varijable *frame_id* i *timestamp*. *Frame_id* se koristi kako bi se prikazala kontinuiranost mjerenja interakcije, a iz varijable *timestamp* se određuje trajanje. Emocije su prediktorske varijable (*predictors*), a vrijeme je varijabla ishoda (*outcome*). To su svi ulazni podaci koji su korišteni tijekom daljnje analize [10].

2.4.1. Mjerna ljestvica

Za statističku je analizu bitno da su mjerene varijable operativne, odnosno da su dovoljno precizno izmjerene kako bi se na njima mogli provoditi svi potrebni testovi. Prema mjernoj ljestvici, u ovom su istraživanju podaci mjereni prema zadnjoj razini - skali omjera (*ratio level*). Prikupljeni se podaci mogu kategorizirati, rangirati, može se odrediti jednaki razmak između susjednih točaka podataka te postoji prava nulta vrijednost. Nulta vrijednost znači apsolutni nedostatak varijable [10].

Uz pomoć deskriptivne statistike dobiva se predodžba o izmjerenim podacima kroz mjerenje centralne tendencije (*central tendency*) i varijabilnosti (*variability*). Ovisno o razini na mjernoj ljestvici, mogu se koristiti različite matematičke operacije i metode. Budući da su podaci na najvišoj razini, mogućnosti su najveće. Podaci se mogu obrađivati uz pomoć matematičkih operacija jednakosti, usporedbe, zbrajanja, oduzimanja, množenja i dijeljenja. Metode mjerenja centralne tendencije koje će se koristiti su mod (*mode*), medijan (*median*) i aritmetička sredina (*arithmetic mean*). Mjerenje varijabilnosti provodi se sa računanjem raspona podataka (*range*), varijancije (*variance*), standardne devijacije (*standard deviation*) i interkvartilne razlike (*interquartile rang*) [9][10].

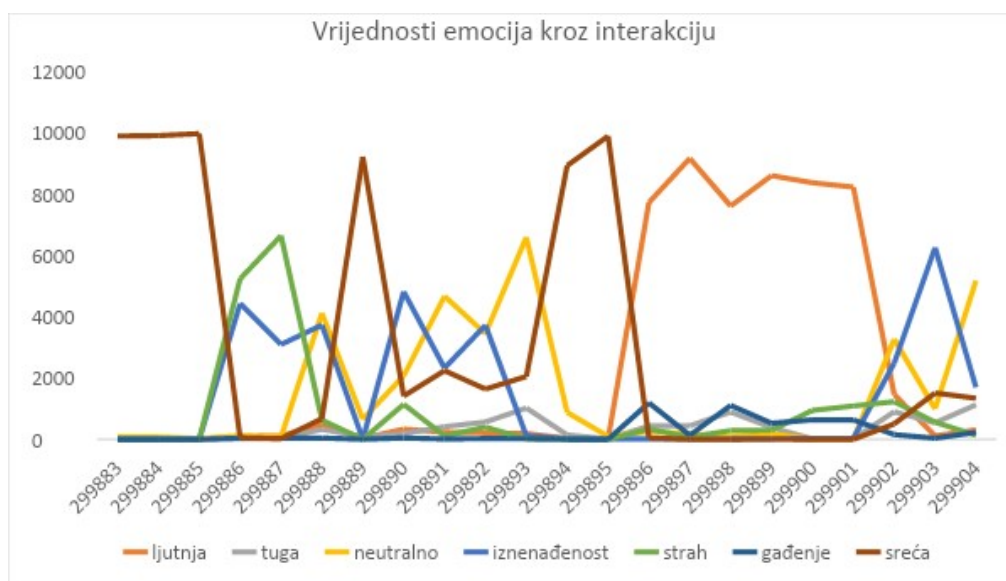
2.4.2. Primjer prikupljenih podataka

Radi lakšeg razumijevanja, mjerenje varijabli će biti detaljnije objašnjeno na primjeru jednog od ispitanika. Prikazan je samo dio interakcije ispitanika i robota [Tablica 2]. Puna interakcija traje od 12:35:23 do 12:39:01 te *frame_id* varijabla ima vrijednosti od 7472 sve do 7690.

Tablica prikazuje ranije spomenute mjerene varijable emocija. Isti se podaci mogu pratiti kroz graf gdje je na apscisi varijabla *frame_id*, a na ordinati su emocije [Slika 3].

Tablica 2. Dio podataka jednog od sudionika

<i>frame_id</i>	ljutnja (<i>angry</i>)	tuga (<i>sad</i>)	neutralno (<i>neutral</i>)	iznenađeno st (<i>surprised</i>)	strah (<i>fear</i>)	gađenje (<i>disgust</i>)	sreća (<i>happy</i>)	
299883	8	10	89	6	0	1	9886	19.08.2021 15:59:26
299884	6	7	81	5	0	0	9901	19.08.2021 15:59:27
299885	1	1	26	5	0	0	9965	19.08.2021 15:59:28
299886	100	30	120	4425	5233	37	54	19.08.2021 15:59:29
299887	113	26	74	3091	6638	35	23	19.08.2021 15:59:30
299888	486	328	4115	3726	629	44	673	19.08.2021 15:59:31
299889	38	87	657	3	0	1	9215	19.08.2021 15:59:32
299890	333	185	2066	4816	1143	45	1412	19.08.2021 15:59:33
299891	233	413	4666	2325	113	10	2240	19.08.2021 15:59:34
299892	208	564	3446	3719	400	25	1639	19.08.2021 15:59:35
299893	190	1022	6587	115	16	21	2049	19.08.2021 15:59:36
299894	31	143	878	19	1	3	8923	19.08.2021 15:59:37
299895	2	11	97	5	0	0	9884	19.08.2021 15:59:38
299896	7721	438	252	17	335	1190	48	19.08.2021 15:59:39
299897	9161	442	185	3	73	133	2	19.08.2021 15:59:40
299898	7607	892	95	9	288	1106	2	19.08.2021 15:59:41
299899	8597	384	176	16	295	525	7	19.08.2021 15:59:42
299900	8364	17	6	31	945	637	0	19.08.2021 15:59:43
299901	8223	17	5	31	1087	635	0	19.08.2021 15:59:44
299902	1461	900	3268	2485	1228	154	504	19.08.2021 15:59:45
299903	114	540	991	6246	560	37	1511	19.08.2021 15:59:46
299904	310	1126	5173	1694	131	225	1340	19.08.2021 15:59:47



Slika 3. Graf vrijednosti emocije kroz interakciju

Prateći podatke na tablici i grafu može se vidjeti kako je na samom početku interakcije do *frame_id* 299886, tj. u prvih 3 sekunde, samo prisutna emocija sreće sa gotovo maksimalnom vrijednosti od 10000. Nakon toga do 299896 se javljaju emocije straha, iznenađenosti i neutralnosti istovremeno. Zatim se izražava emocija ljutnje. Nakon toga se ponovo javljaju i ostale emocije u većim postocima.

2.4.3. Čišćenje podataka

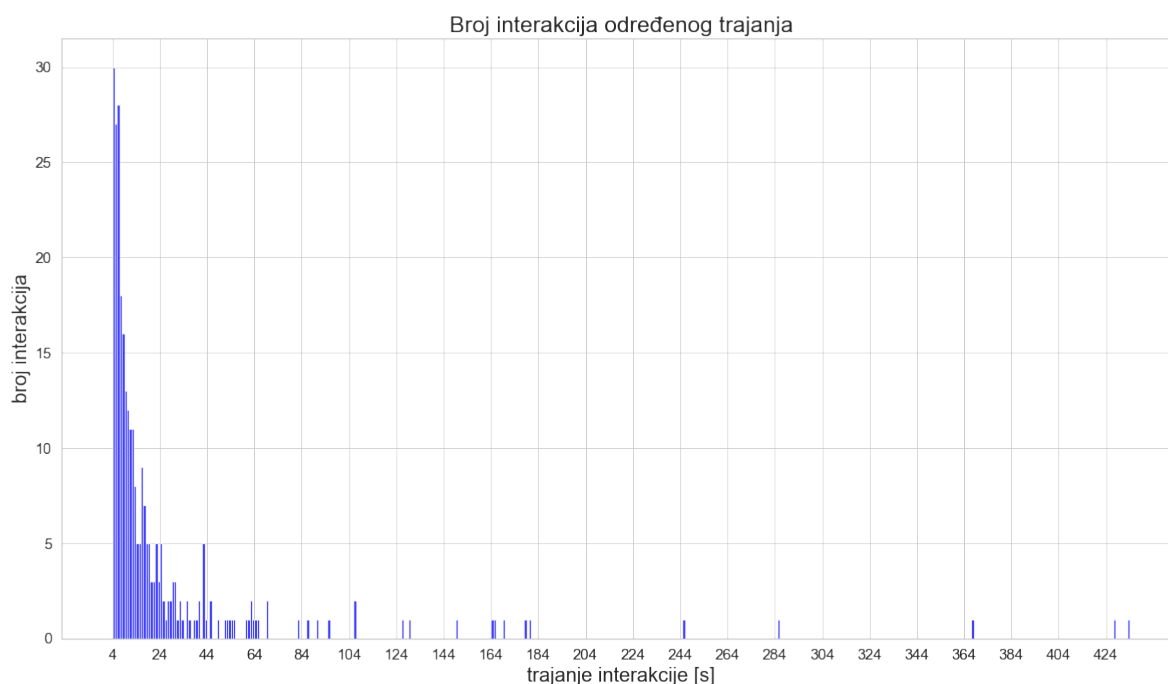
Čišćenje podataka služi za minimizaciju potencijalnih pogrešaka koje mogu dovesti do lažnog ili nevažećeg zaključka. S netočnim ili nevažećim podacima moguće je napraviti grešku tipa I ili II. Pogreška tipa I je lažno pozitivan zaključak, dok je pogreška tipa II lažno negativan zaključak [10]. Čišćenje podataka uključuje uočavanje i rješavanje potencijalnih nedosljednosti ili pogrešaka podataka radi poboljšanja kvalitete podataka. Svi skupljeni podaci moraju biti točni, potpuni, konzistentni, jedinstveni i ujednačeni.

Budući da prikupljanje podataka u ovom istraživanju nije imalo neku čvrstu strukturu određenu minimalnim ili maksimalnim trajanjem, trajanja interakcija se razlikuju. Podaci su prvo očišćeni na način na su sve interakcija kraće od 4 sekunde isključene iz daljnje obrade. To je napravljeno iz razloga što su četiri sekunde nedovoljno duge da bi se interakcija promatrala kao takva. Na taj su se način odvojili podaci 295 interakcija koje su konzistentne, bez prekida i traju najmanje 4 sekunde.

2.4.3.1. Ekstremne vrijednosti

Ekstremne vrijednosti (*outliers*) su vrijednosti koje se razlikuju od većine drugih podataka u skupu. One mogu biti istinite vrijednosti ili pogreške. Imaju veliki utjecaj na pojedine matematičke metode koje se koriste prilikom analize podataka, primjerice računanje prosjeka [10].

Od 295 sudionika koji su imali interakciju dulju od 4 sekunde, petero njih je imalo ekstremno duge interakcije s obzirom na ostatak. Njihove su interakcije bile dulje od 245 sekundi. To je moguće vidjeti na histogramu koji prikazuje broj interakcija određenog trajanja [Slika 4]. Budući da tih 5 sudionika čini svega 2.1% ukupnih sudionika, a utjecaj na rezultate je velik, njih se smatra ekstremnim vrijednostima i ne uključuju se u daljnju analizu. Time je ukupan broj ispitanika u istraživanju 290.



Slika 4. Graf broja interakcija određenog trajanja

3. OBRADA PRIKUPLJENIH PODATAKA

Sljedeći korak je prikupljene podatke organizirati i opisati njihove važne karakteristike uz pomoć opisne statistike (*descriptive statistics*). Nakon toga, uz pomoć statističkog zaključivanja (*inferential statistics*) treba donijeti zaključke i napraviti predviđanja na temelju zadanih podataka. Što se tiče opisne statistike postoje tri glavna područja obrade koja se koriste: distribucija (*distribution*) za opisivanje učestalosti svake vrijednosti, centralna tendencija (*central tendency*) za prosjeke vrijednosti i varijabilnost (*variability*) za opisivanje raspoređenosti, tj. disperzije vrijednosti.

3.1. Trajanje interakcija

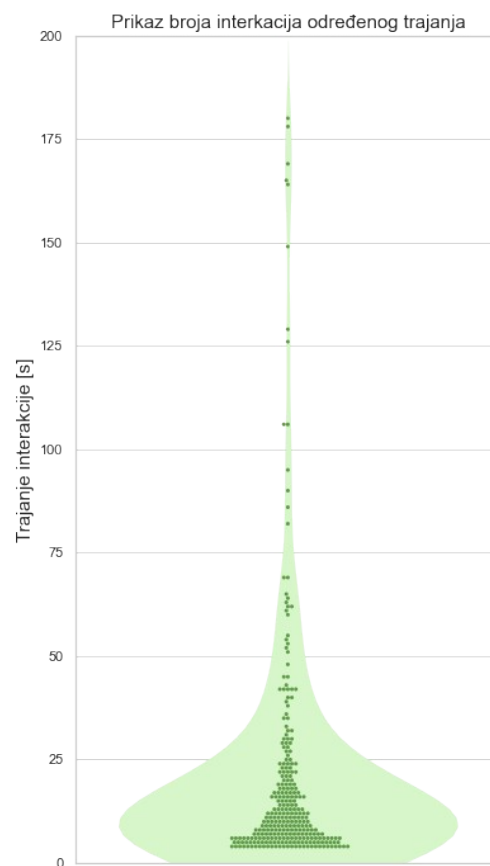
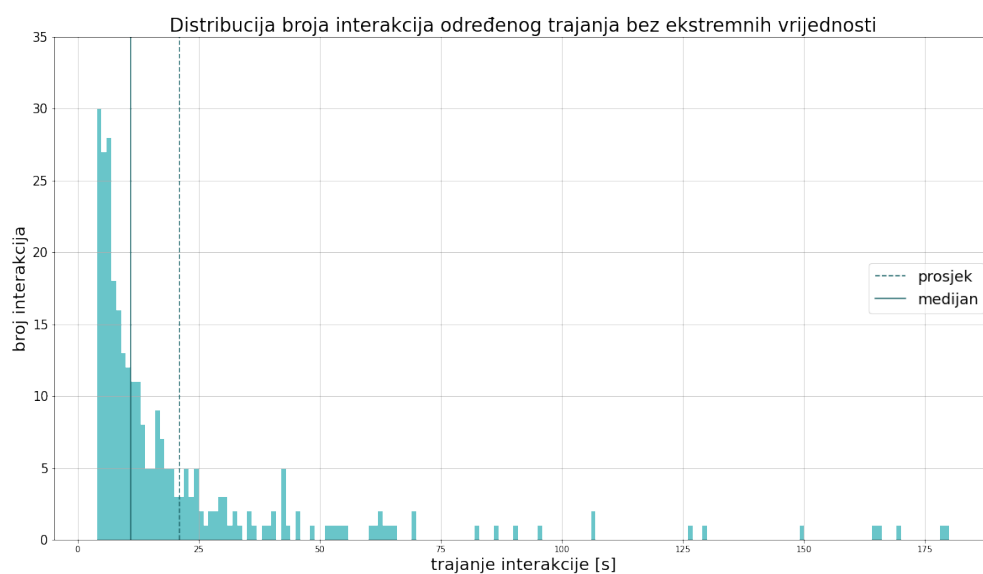
Tablica 3. Podaci o trajanju interakcija

Maksimalno trajanje	180
Minimalno trajanje	4
Mod (najčešća vrijednost)	4
Medijan (srednja vrijednost)	11
Prosjek	21,02069
Raspon	176
Interkvartilna razlika	16
Standardna devijacija	29,35276
Varijanca	861,5843

Najviše zabilježenih interakcija ima trajanje od 4 sekunde (vrijednost moda je 4) što je ujedno i najkraće moguće prihvatljivo trajanje. Od sveukupno 290, takvih interakcija je 30. S povećanjem duljine trajanja smanjuje broj interakcija, tj. više je kraćih nego duljih interakcija. Na histogramu trajanja interakcija moguće je vidjeti tu pozitivno asimetričnu raspodjelu. Kao i kod većine pozitivnih raspodjela, prosjek trajanja je veći od medijana zbog velikih vrijednosti (vrijednosti na desnom 'repu' grafa). Medijan iznosi 11, a prosječna vrijednost je 21,02 sekunde. Da je raspodjela vrlo iskrivljena potvrđuje i Pearsonov koeficijent iskrivljenosti koji iznosi 1.024. On se računa prema formuli $sk_2 = \frac{3 \cdot (\text{prosijek} - \text{medijan})}{\text{standardna devijacija}}$ te govori koliko standardnih devijacija razdvaja prosječnu vrijednost i medijan. Kod normalne razdiobe on iznosi 0 ili vrlo blizu te vrijednosti.

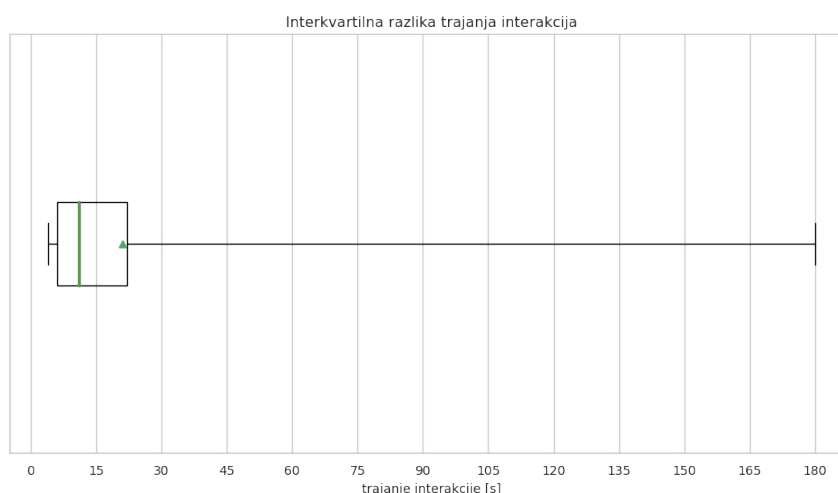
Tablica 4. Broj interakcija određenog trajanja

trajanje [s]	broj interakcija tog trajanja	trajanje [s]	broj interakcija tog trajanja	trajanje [s]	broj interakcija tog trajanja	
4	30	27	2	62	2	
5	27	28	2	63	1	
6	28	29	3	64	1	
7	18	30	3	65	1	
8	16	31	1	69	2	
9	13	32	2	82	1	
10	12	33	1	86	1	
11	11	35	2	90	1	
12	11	36	1	95	1	
13	8	38	1	106	2	
14	5	39	1	126	1	
15	5	40	2	129	1	
16	9	42	5	149	1	
17	7	43	1	164	1	
18	5	45	2	165	1	
19	5	48	1	169	1	
20	3	51	1	178	1	
21	3	52	1	180	1	
22	5	53	1	245	1	outlier
23	3	54	1	285	1	outlier
24	5	55	1	367	1	outlier
25	2	60	1	427	1	outlier
26	1	61	1	434	1	outlier

**Slika 5. Prikaz broja interakcija određenog trajanja (grafički prikaz violine)****Slika 6. Histogram trajanja interakcija**

Središnja tendencija govori gdje se većina točaka podataka nalazi, a varijabilnost koliko su točke udaljene jedna od druge. To je važno jer količina varijabilnosti određuje koliko se dobro mogu generalizirati podaci uzorka na željenu populaciju. Idealna varijabilnost je niska, jer to znači da se bolje mogu predvidjeti informacije za populaciju. Visoka varijabilnost označava da vrijednosti nisu dosljedne pa je predviđanje teže [9][10].

Kao što je već spomenuto, najkraće zabilježeno trajanje je 4, a najdulje 180 sekundi iz čega se zaključuje da je raspon trajanja 176. Dakle, sve interakcije traju unutar 3 minute. Interkvartilni raspon iznosi 16 [Slika 7]. Na Slici 7. 'kutijom' je označeno područje interkvartilne razlike koje se proteže od 6 do 22. To je područje gdje se nalazi 50% od svih podataka koji su u samoj sredini cjelokupnog raspona. Rubovi označavaju najkraće i najdulje trajanje, zelenom je linijom označen prosjek, a trokutićem medijan.



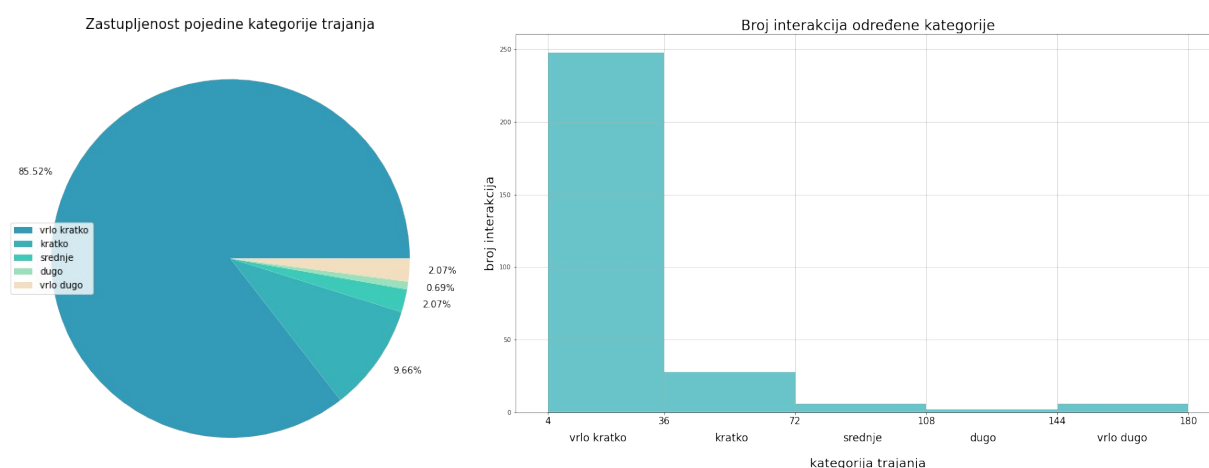
Slika 7. Prikaz interkvartilne razlike trajanja

Budući da nije riječ o normalnoj razdiobi već o pozitivno asimetričnoj, standardna devijacija i varijanca nisu najbolje mjera varijabilnosti. Standardna devijacija govori koliko je, u prosjeku, svaki rezultat udaljen od srednje vrijednosti. Što je ona veća, skup podataka je varijabilniji, tj. izmjerene vrijednosti su udaljenije od prosjeka. Ona za trajanje iznosi 29,35. Varijanca daje informaciju o stupnju širenja u vašem skupu podataka. Što su podaci rašireniji, veća je varijanca u odnosu na srednju vrijednost [9][10]. Za trajanje ona iznosi 861,5843. [Tablica 3]

Trajanje interakcija podijeljeno je u pet kategorija: vrlo kratke, kratke, srednje, duge i vrlo duge interakcije [Tablica 5]. Najviše ima vrlo kratkih interakcija i to čak 248 njih, tj. 85,52% svih interakcija je kraće od 36 sekundi. Kratkih interakcija ima 28, srednjih 6, dugih 2 i vrlo dugih 6 [Slika 8].

Tablica 5. Kategorije interakcija

kategorija	trajanje (s)
vrlo kratko	$4 \leq t < 36$
kratko	$36 \leq t < 72$
srednje	$72 \leq t < 108$
dugo	$108 \leq t < 144$
vrlo dugo	$144 \leq t \leq 180$



Slika 8. Zastupljenost kategorija duljine trajanja interakcija

3.2. Podaci vezani za emocije

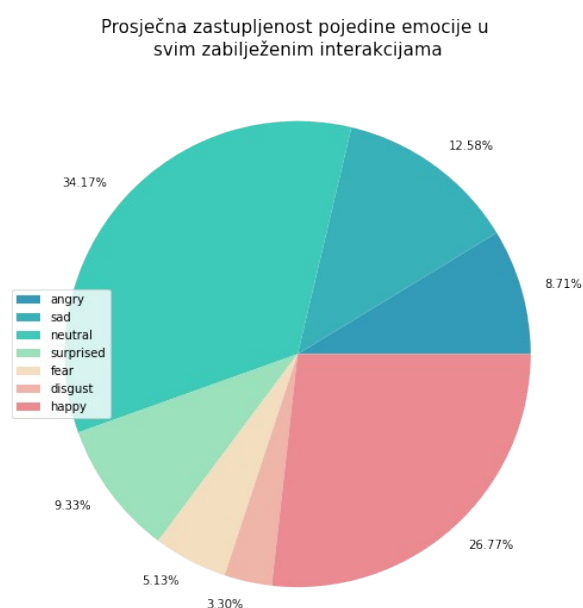
Kao što je već spomenuto na početku, sedam emocija koje se registriraju prilikom interakcija su ljutnja (*angry*), tuga (*sad*), neutralnost (*neutral*), iznenađenost (*surprised*), strah (*fear*), gađenje (*disgust*) te sreća (*happy*). Tablica u nastavku prikazuje sve važne podatke o vrijednostima emocija. Za svaku je emociju napisana njena najčešća vrijednost, srednja vrijednost, prosječna vrijednost, broj interakcija s tom emocijom kao dominantnom i broj interakcija kod koje je ta emocija prva izražena emocija u interakciji.

Tablica 6. Podaci o vrijednostima emocija

Emocija	Mod	Medijan	Prosjek	Broj interakcija sa tom emocijom kao dominantnom	Broj interakcija kod koje je ta emocija prva izražena emocija u interakciji
ljutnja (<i>angry</i>)	0	240	876,7444	8	11
tuga (<i>sad</i>)	0	491	1108,375	10	14
neutralnost (<i>neutral</i>)	1	2756	3172,875	122	137
iznenađenost	1	225	987,7439	18	25

(surprised)					
strah (fear)	0	65,5	580,4982	5	3
gađenje (disgust)	0	44	381,0265	3	5
sreća (happy)	0	1202,5	2892,724	124	95

Promatrajući sve zabilježene interakcije, u prosjeku je najviše bila prisutna neutralna emocija s 34,20%. Zatim su redom po zastupljenosti bile prisutne emocije sreće s 28,80%, tuga s 12,57%, iznenađenost s 9,31%, ljutnja s 8,69%, strah s 5,12% i zadnja, najmanje zastupljena emocija je bila gađenje sa svega 3,30% [Slika 9].

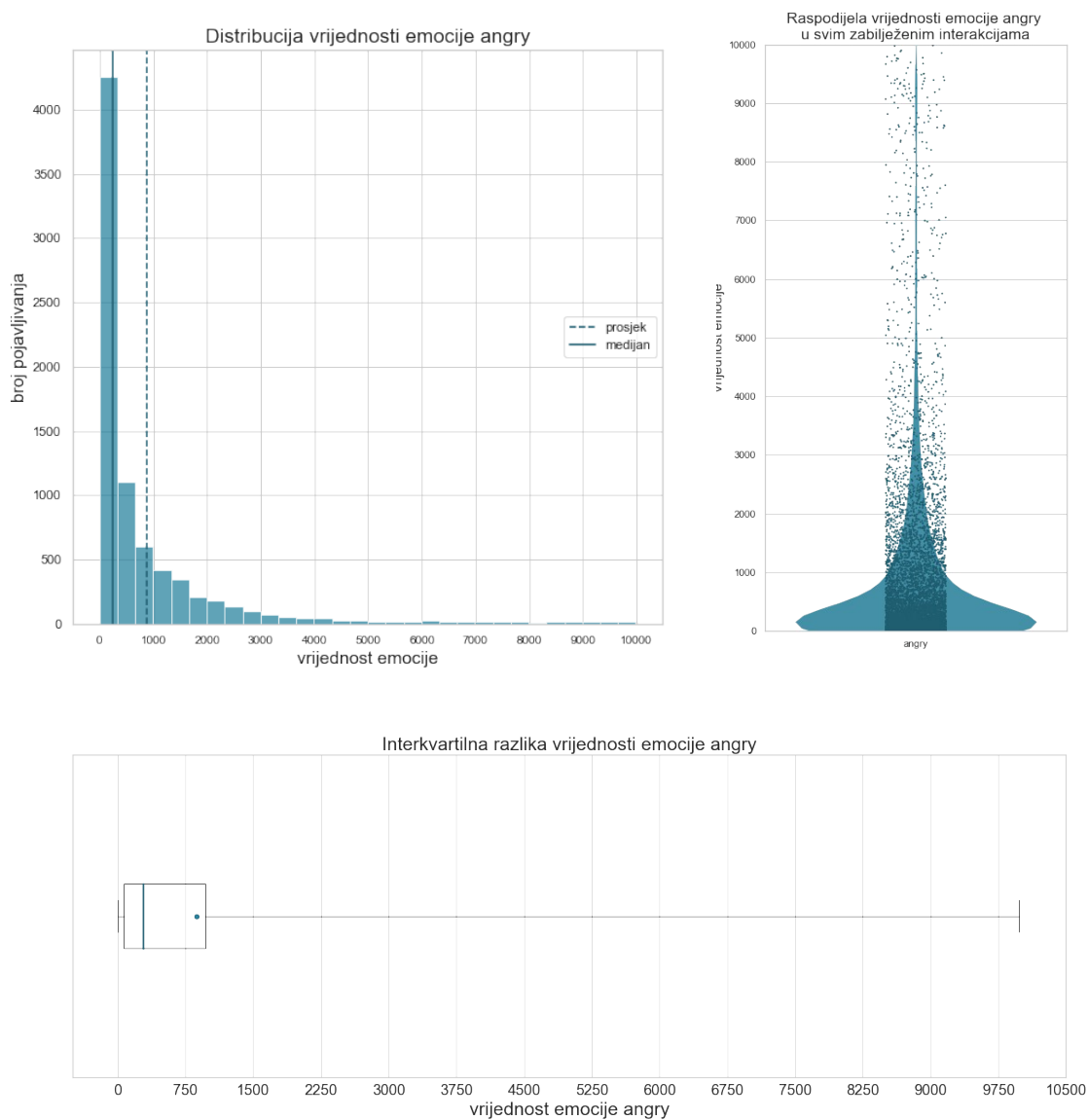


Slika 9. Zastupljenost pojedine emocije u svim interakcijama

U nastavku će za svaku pojedinu emociju biti prikazani histogram distribucije vrijednosti emocije, grafički prikaz u obliku violine koji prikazuje vrijednosti, te interkvartilna razlika vrijednosti.

3.2.1. Ljutnja (angry)

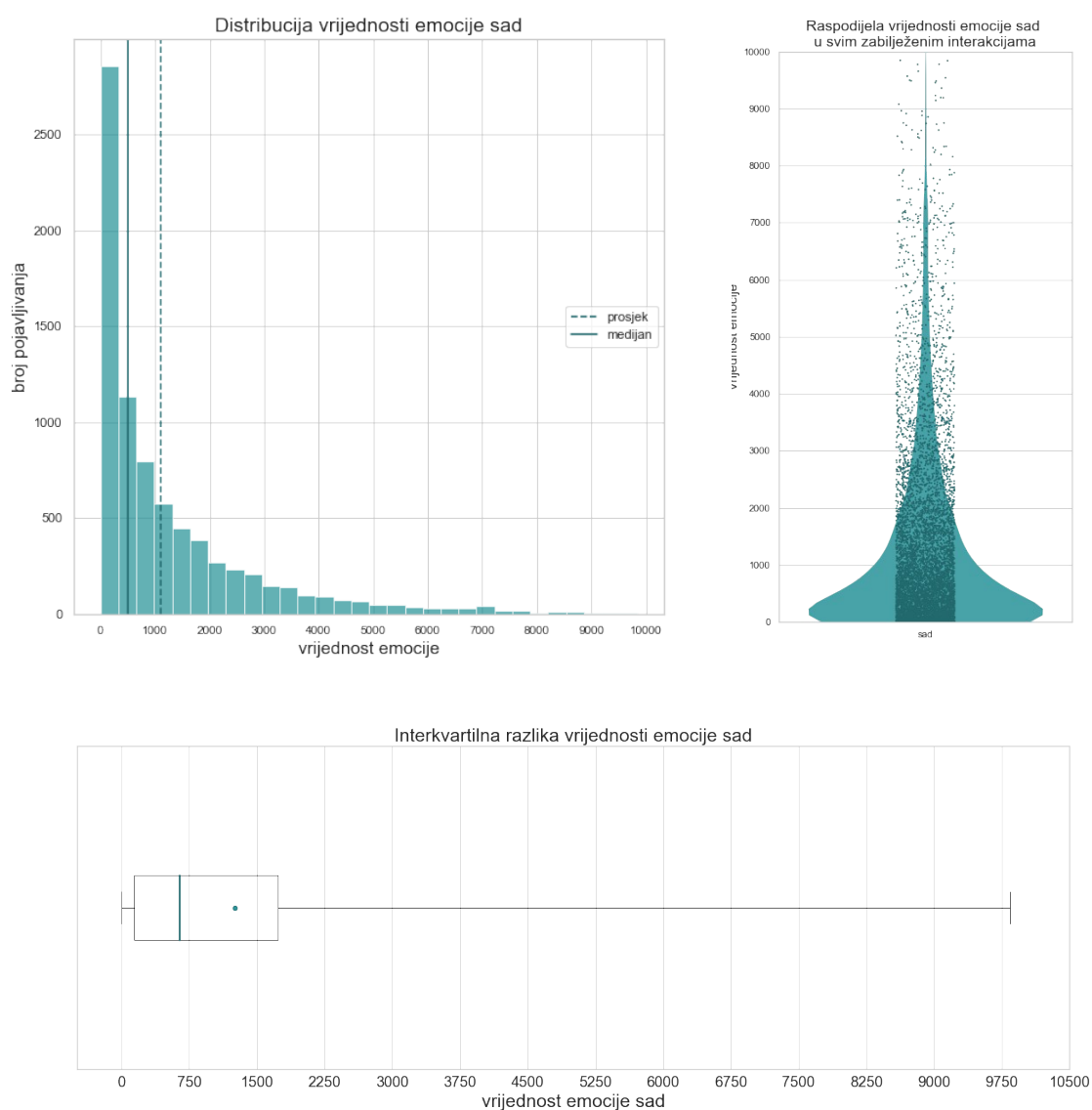
Kod emocije ljutnje maksimalna zabilježena vrijednost je 9984, a minimalna 0. Samim time, raspon vrijednosti je 9984. Srednja vrijednost (medijan) je 279, a prosječna vrijednost 870,95. Interkvartilna razlika je 901, gdje je $Q_1=70$, a $Q_3=971$ [Slika 10].



Slika 10. Grafički prikazi vrijednosti emocije ljutnje

3.2.2. Tuga (sad)

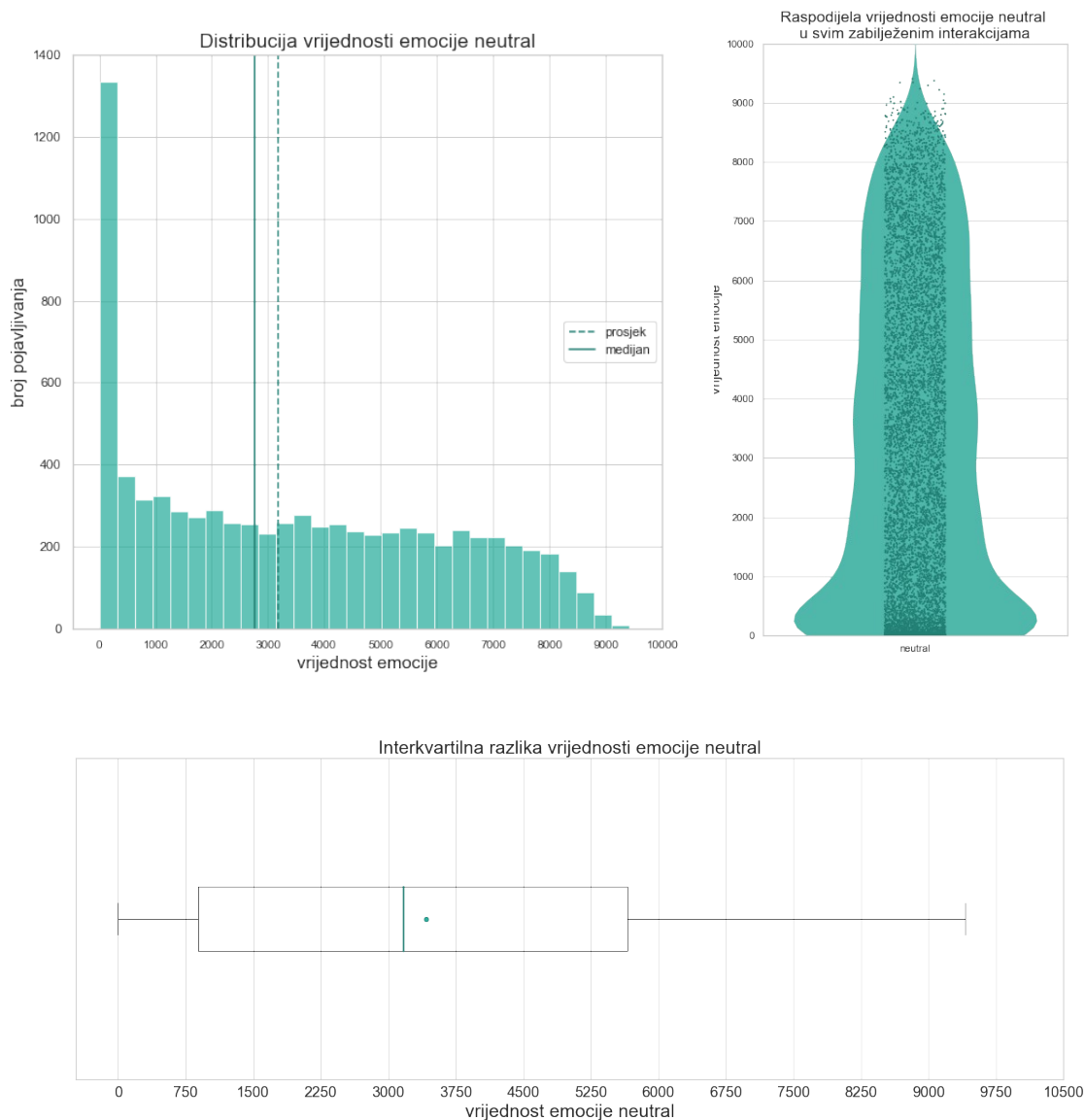
Maksimalna zabilježena vrijednost emocije tuge je 9845, a minimalna 0. To znači da je raspon vrijednosti 9845. Srednja vrijednost (medijan) je 644, a prosječna 1258,24. Interkvartilna razlika je 1591, gdje je $Q_1=138$, a $Q_3=1729$ [Slika 11].



Slika 11. Grafički prikazi vrijednosti emocije ljutnje

3.2.3. Neutralnost (neutral)

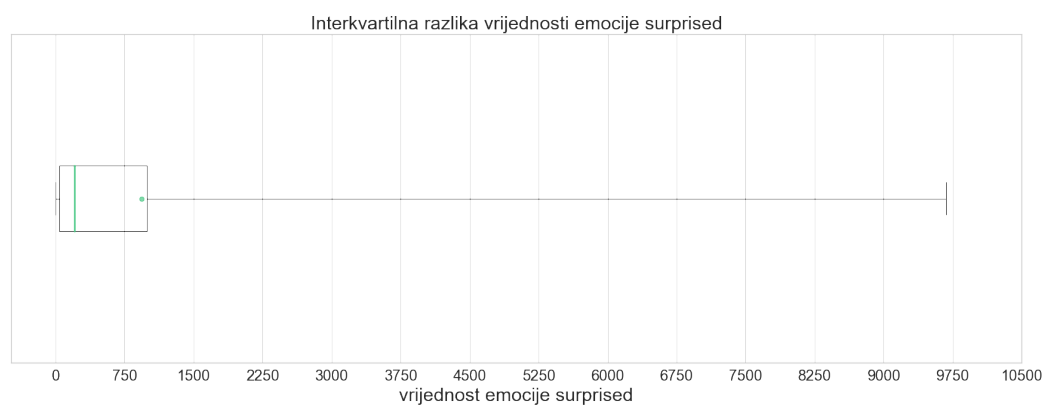
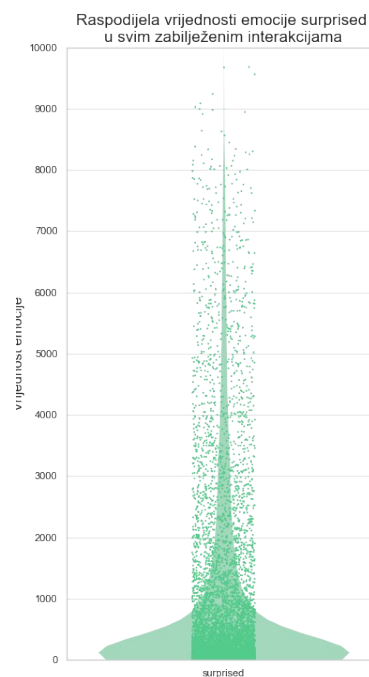
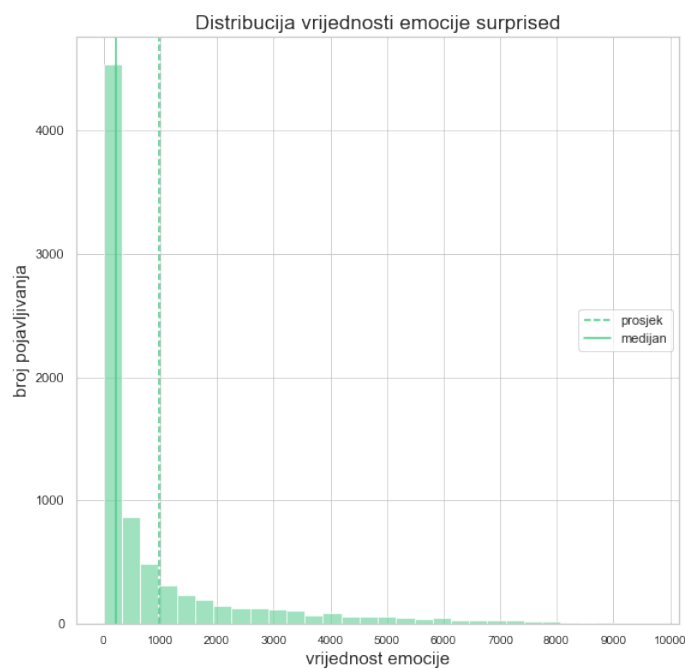
Za emociju neutralnog stava maksimalna zabilježena vrijednost je 9405, a minimalna 0, što znači da je raspon vrijednosti 9405. Srednja vrijednost (medijan) je 3166, a prosječna 3417,31. Interkvartilna razlika je 4771, gdje je $Q_1=8860$, a $Q_3=5657$ [Slika 12].



Slika 12. Grafički prikazi vrijednosti emocije neutralnosti

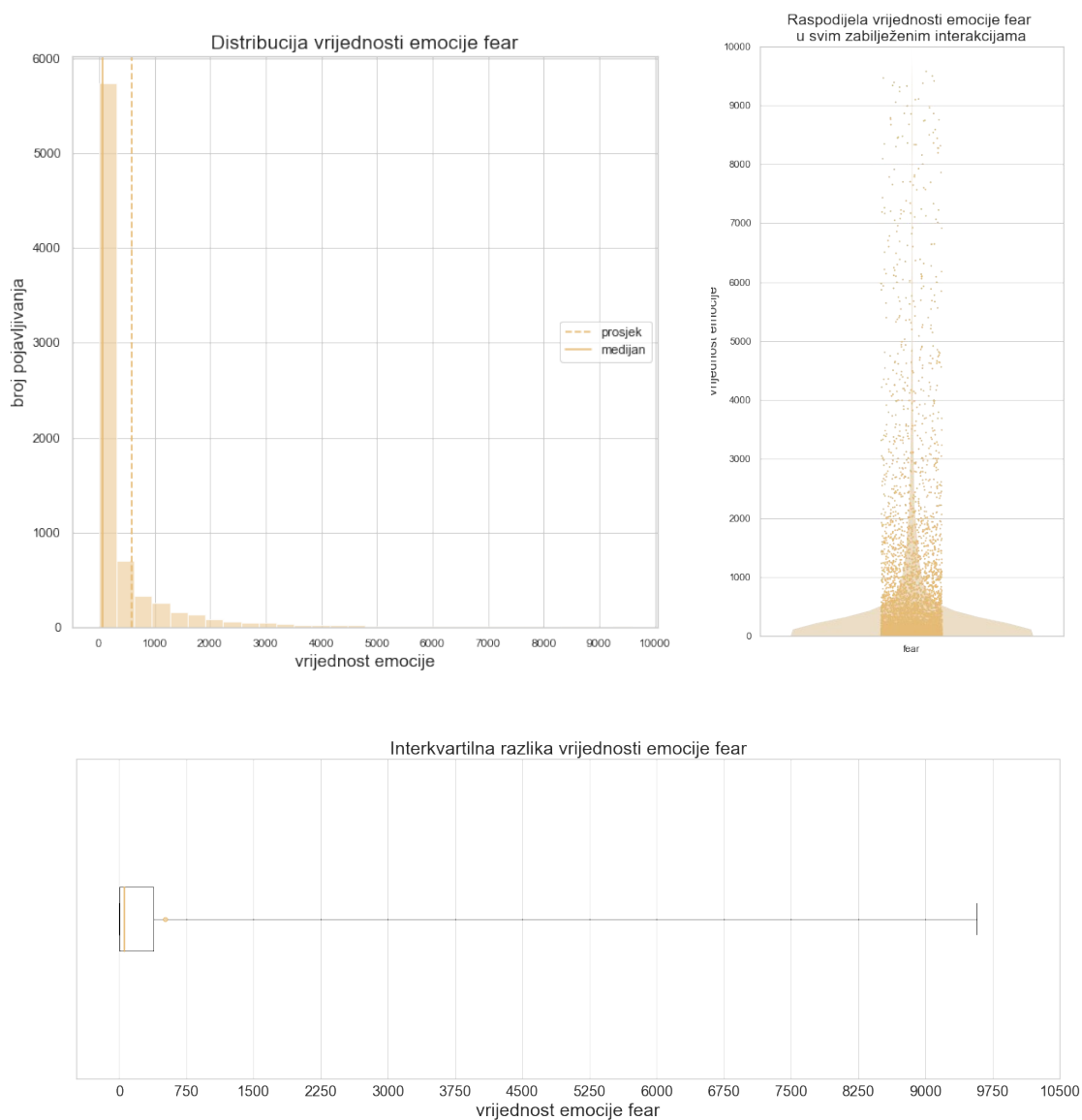
3.2.4. Izenađenost (surprised)

Kod emocije izenađenosti je maksimalna zabilježena vrijednost 9679, a minimalna 0. Samim time, raspon vrijednosti je 9679. Srednja vrijednost (medijan) je 201, a prosječna 932,909. Interkvartilna razlika je 950, gdje je $Q_1=43$, a $Q_3=993$ [Error: Reference source not found].



3.2.5. Strah (fear)

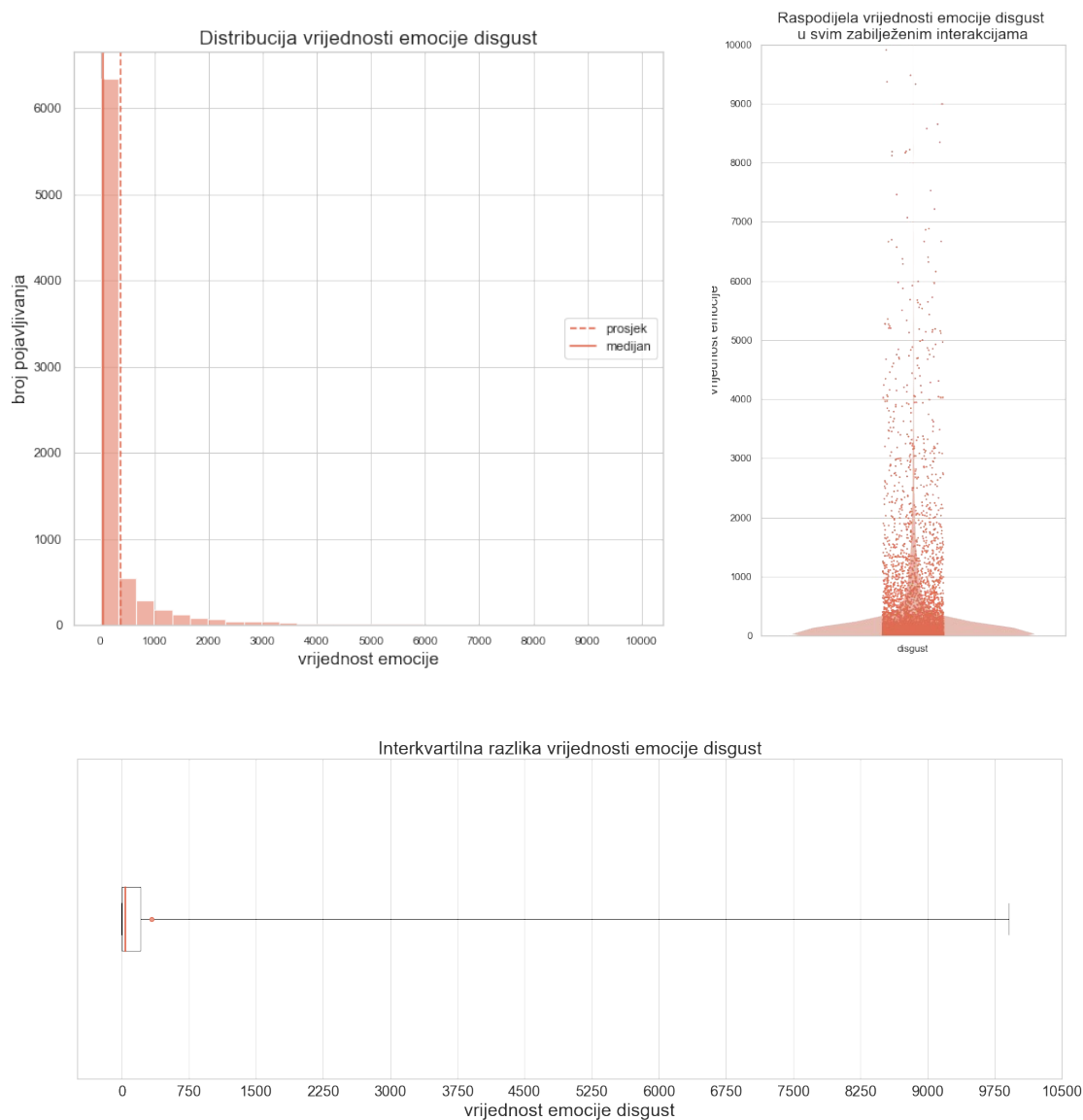
Za emociju straha maksimalna zabilježena vrijednost je 9573, a minimalna 0, što čini raspon vrijednosti 9573. Srednja vrijednost (medijan) je 54, a prosječna 513,116. Interkvartilna razlika je 374, gdje je $Q_1=8$, a $Q_3=382$ [Slika 13].



Slika 13. Grafički prikazi vrijednosti emocije straha

3.2.6. Gađenje (disgust)

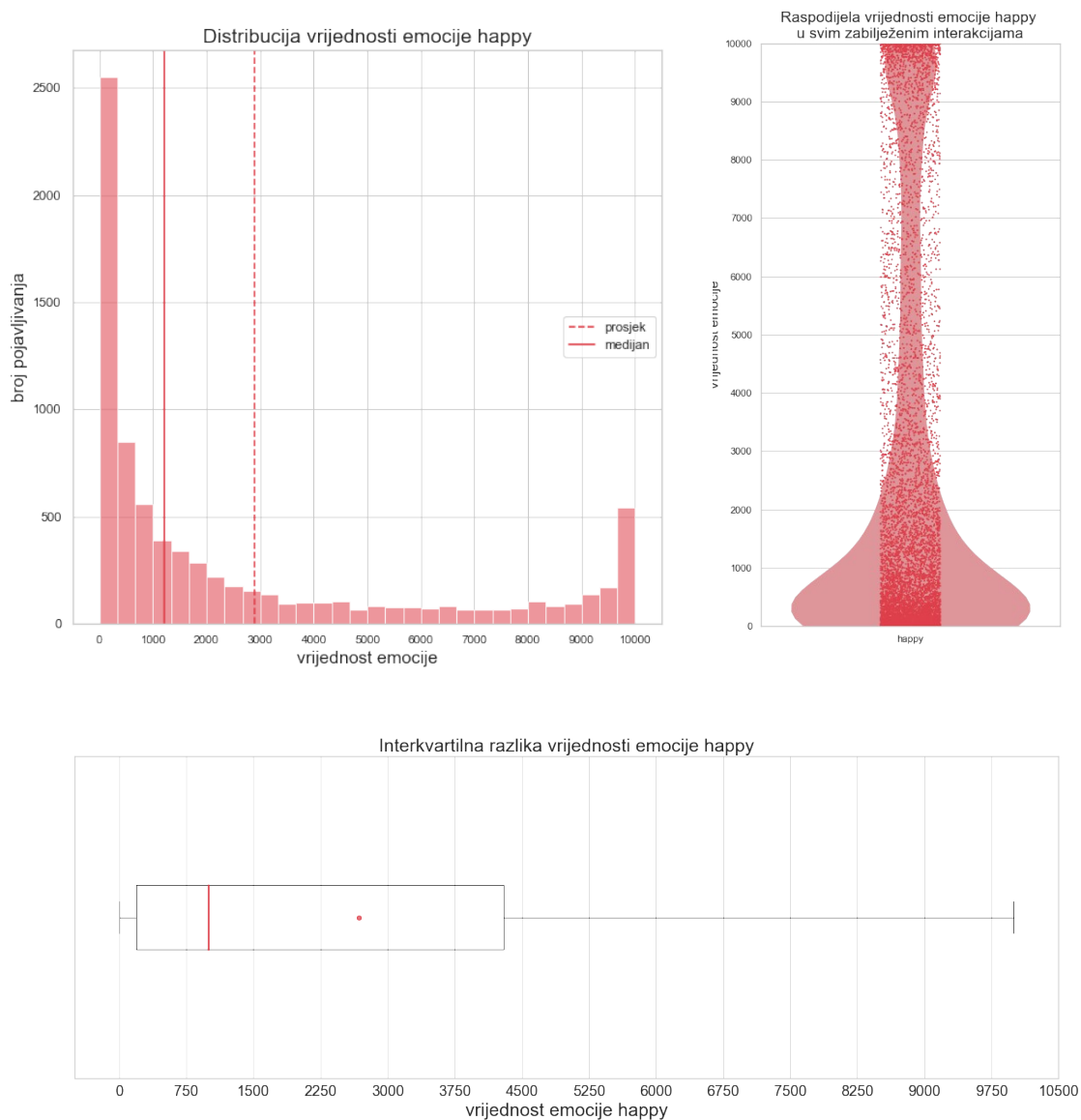
Maksimalna zabilježena vrijednost emocije gađenja je 9906, a minimalna 0, stoga je raspon vrijednosti 9906. Medijan iznosi 37, a prosječna vrijednost 330,117. Interkvartilna razlika je 203, gdje je $Q_1=6$, a $Q_3=209$ [Slika 14].



Slika 14. Grafički prikazi vrijednosti emocije straha

3.2.7. Sreća (happy)

Sreća je jedina emocija kod koje je kao maksimalna vrijednost zabilježena i maksimalna moguća od 10000. Minimalna zabilježena je 0, što čini raspon vrijednosti 10000. Srednja vrijednost (medijan) je 995, a prosječna je vrijednosti 2677,29. Interkvartilna razlika je 4108, gdje je $Q_1=191$, a $Q_3=4299$ [Slika 15].



Slika 15. Grafički prikazi vrijednosti emocije sreće

3.2.8. Usporedba vrijednosti emocija

Vrijednosti većine emocija imaju pozitivno asimetričnu razdiobu što znači da je najviše zabilježenih slučajeva malih vrijednosti, a najmanje onih visokih. Tu se razlikuje sreća čija je razdioba vrlo slična ostalima no prema samom kraju velikih vrijednosti, one se povećavaju. Neutralnost s druge strane, kao i sve ostale emocije, ima najviše onih sa malim vrijednostima, no kroz cijeli raspon vrijednosti, od otprilike vrijednosti od 1000 do 8000, broj vrijednosti se ne smanjuje već je otprilike isti. Sve emocije imaju najmanju zabilježenu vrijednost 0 koja je apsolutna nula, a svima je maksimalna vrijednost oko 10000 što je i maksimalno moguće. Time sve emocije imaju otprilike jednaki raspon. Emocija sreće i neutralnosti imaju najveće područje interkvartilne razlike. Kod ostalih je emocija ono dosta malo i smješteno kod vrlo malih vrijednosti. To znači da emocije sreće i neutralnosti imaju najveću raznolikost dok su kod ostalih vrijednosti većinom male. To potvrđuju i prosjeci i medijani pojedinih vrijednosti.

4. TESTIRANJE HIPOTEZA

Za testiranje hipoteze koristi se statistički test Mann-Whitney U test kojim se provjerava razina značajnosti (*significance testing*) u *IBM SPSS Statistics* softveru. Općenito, nulta hipoteza predviđa da nema nikakvog značajnog učinka – nema odnosa između varijabli ni razlika između grupa. Alternativna hipoteza iznosi glavno predviđanje pravog učinka – postojanje odnosa između varijabli ili razlika između skupina. Statističkim testom se pokazuje koliko se podudaraju prikupljeni podaci i nulta hipoteza. Dobivena p-vrijednost pokazuje vjerojatnost dobivanja tog rezultata ukoliko je nulta hipoteza istinita. Ona određuje i statističku značajnost (*statistical significance*). Niska p-vrijednost znači visoku statističku značajnost, dok visoka p-vrijednost nisku ili nikakvu statističku značajnost. Na početku istraživanja određena je razina značaja (*significance level*), tj. alfa (α). Ona predstavlja prag statističke značajnosti, tj. najveći rizik donošenja lažno pozitivnog zaključka (pogreška tipa I). Ona je postavljena na vrijednost od 0,05 odnosno 5%. Prilikom testiranja hipoteze, p-vrijednost se uspoređuje s njom kako bi se odlučilo odbacuje li se nulta hipoteza. Ukoliko je p-vrijednost veća od razine značaja, nulta hipoteza nije opovrgnuta te rezultati nisu statistički značajni, a ako je ona manja od razine značaja, nulta hipoteza se odbija te se rezultati smatraju statistički značajnima. Treba imati na umu da p-vrijednost govori samo je li nulta hipoteza podržana ili ne, a ne je li alternativna hipoteza istinita te zašto. Osim statističkog značaja važan je i praktični značaj (*practical significance*). On pokazuje je li rezultat istraživanja dovoljno važan da bude značajan u stvarnom svijetu. Određuje se uz pomoć veličine učinka (*effect size*) [11][12].

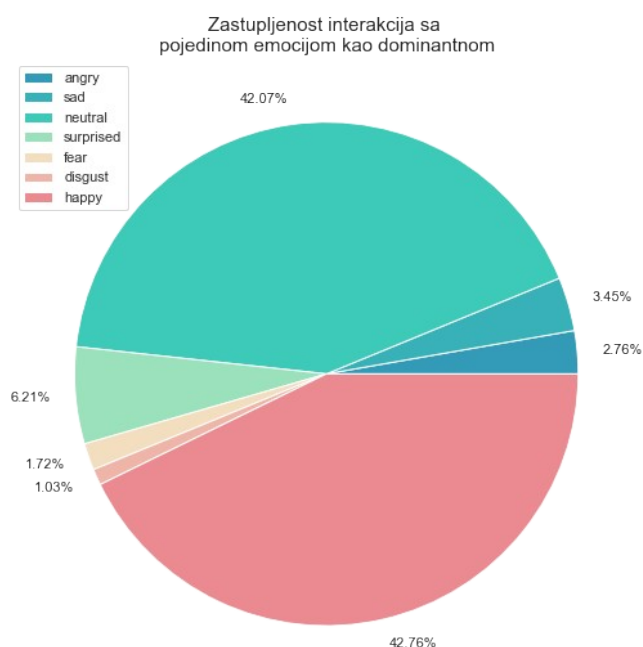
4.1. Prva hipoteza

Na početku rada definirana je prva alternativna hipoteza koja se razmatra - *dominantna emocija utječe na trajanje interakcije*. Nulta hipoteza je - *dominantna emocija nema utjecaja na trajanje interakcije*. Osjećaji koji su dominantni u interakciji trebali bi utjecati na duljinu interakcije. Ukoliko osoba osjeća negativnu emociju tijekom interakcije, primjerice strah, osoba će se htjeti čim prije udaljiti od izvora negativne emocije i imati što kraću interakcije. S druge strane, ukoliko je ono što osjeća pozitivno (npr. sreća), tu će pozitivnu emociju željeti zadržati što dulje te će time i interakcija biti dulja. Stoga, alternativna hipoteza tvrdi da je

duljina interakcije sudionika s robotom definirana dominantnom emocijom koju on osjeća, odnosno je li ona pozitivna ili negativna.

4.1.1. Odnos trajanja interakcije i dominantnih emocija

U svakoj interakciji postoji pojedina emocija koja je najviše zastupljena tijekom te interakcije, tj. koja je dominantna. Od ukupno 290 ukupnih interakcija, u 124 njih dominantna je emocija sreće. Odmah iza toga, u 122 interakcija dominantna emocija je neutralna. Te su dvije emocije dominantne u vrlo velikoj većini interakcija, točnije u 84,83% njih. Emocija koja je dominantna u najmanjem broju interakcija jest gađenje (u svega tri interakcije). [Slika 16]

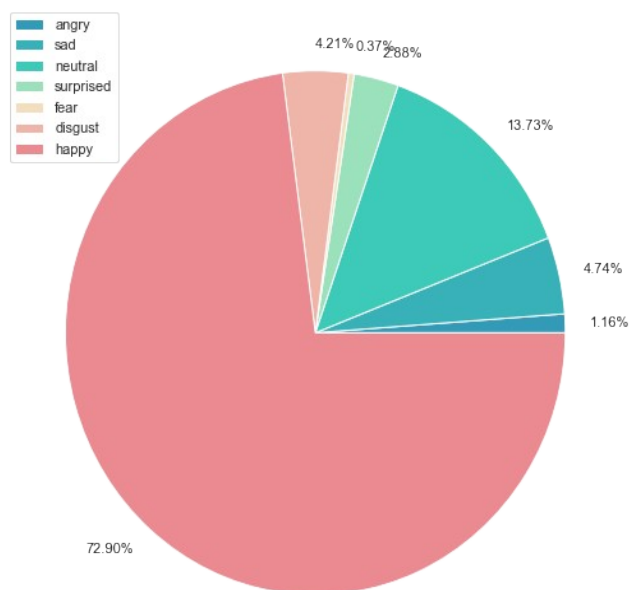


Slika 16. Zastupljenost interakcija sa pojedinom emocijom kao dominantnom

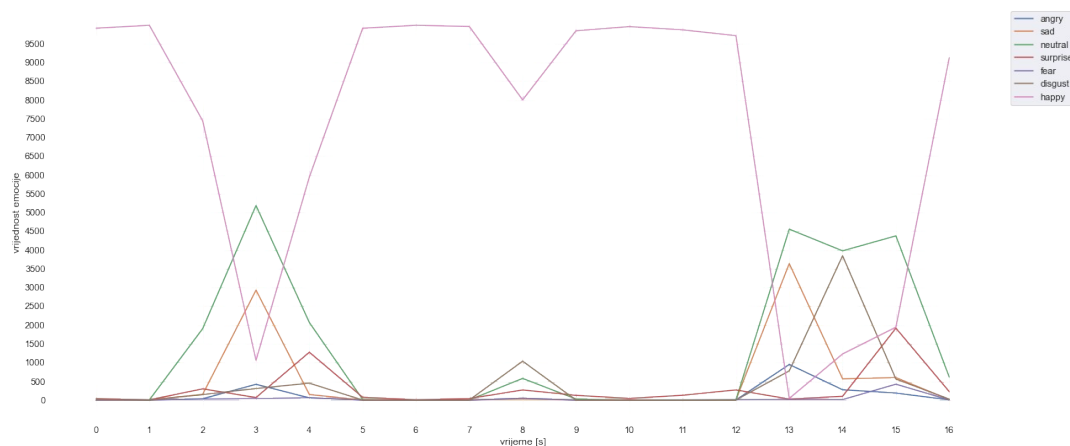
Radi lakšeg razumijevanja pojma dominantne emocije u interakciji u nastavku su prikazani podaci o dvije različite interakcije. Prvo je prikazana interakcija sudionika gdje je dominantna emocija sreća. Riječ je o sudioniku broj 3 čija je interakcija trajala 17 sekundi. Prosječna vrijednost emocije sreće iznosila 7289,824 te je sudionik u prosjeku od 100% interakcije, 72,9% nje bio sretan. Najmanje je osjećao emociju straha koja je u prosjeku iznosila 37,41, odnosno od 100% interakcije, 0,37% nje je bio tužan. [[Slika 17][Slika 18]

Tablica 7. Podaci o prosječnim vrijednostima emocija sudionika broj 3 (sreća kao dominantna emocija)

Emocija	Prosječna vrijednost emocije
ljutnja (angry)	115,7647
tuga (sad)	474,4706
neutralnost (neutral)	1372,941
iznenađenost (surprised)	288
strah (fear)	37,41176
gađenje (disgust)	421,4118
sreća (happy)	7289,824



Slika 17. Raspodjela prosječnih emocija kod sudionika broj 3 (sreća kao dominantna emocija)

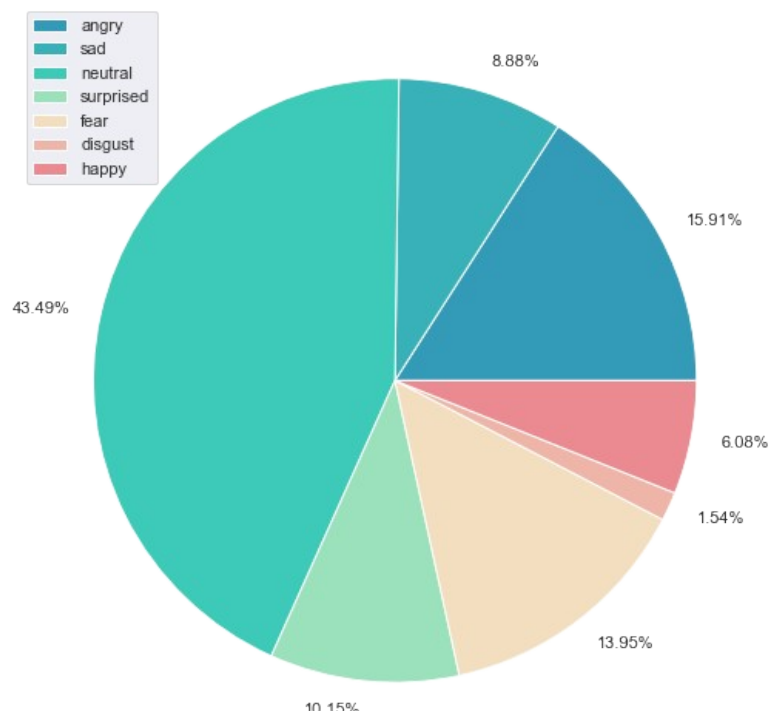


Slika 18. Vrijednosti emocija kroz interakciju sudionika broj 3 (sreća kao dominantna emocija)

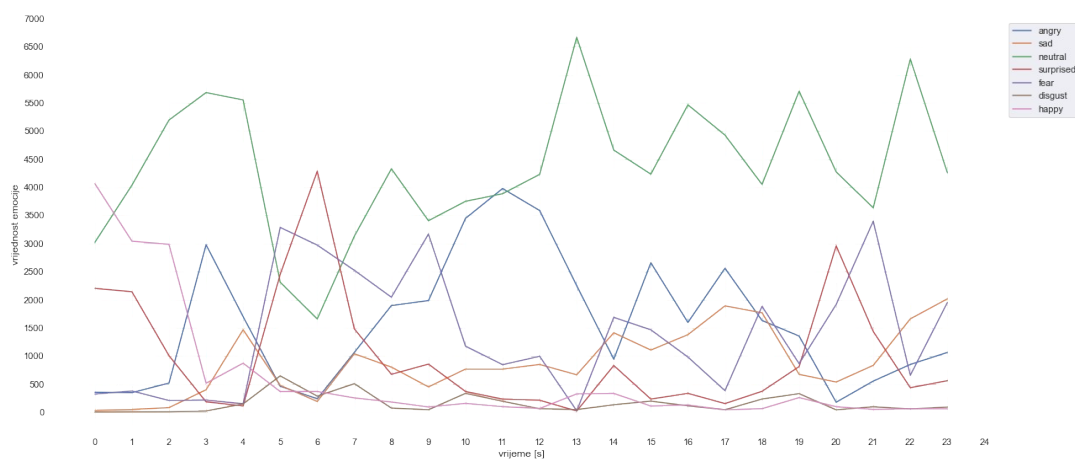
Sljedeći primjer je interakcija s neutralnom emocijom kao dominantnom. Radi se o sudioniku broj 121 gdje je interakcija trajala 23 sekunde. Prosječna vrijednost neutralne emocije je iznosila 4348,83 te je sudionik od 100% interakcije, 43,49% nje bio neutralan. Najmanje je osjećao emociju gađenja koja je u prosjeku iznosila 887,75, odnosno od 100% interakcije, u prosjeku je 1,54% nje osjećao gađenje. [Slika 19][Slika 20][ablica 8]

Tablica 8. Podaci o prosječnim vrijednostima emocija sudionika broj 121 (neutralnost kao dominantna emocija)

Emocija	Prosječna vrijednost emocije
ljutnja (angry)	1591,25
tuga (sad)	887,75
neutralnost (neutral)	4348,833
iznenađenost (surprised)	1014,667
strah (fear)	1395,375
gađenje (disgust)	154,4583
sreća (happy)	607,875

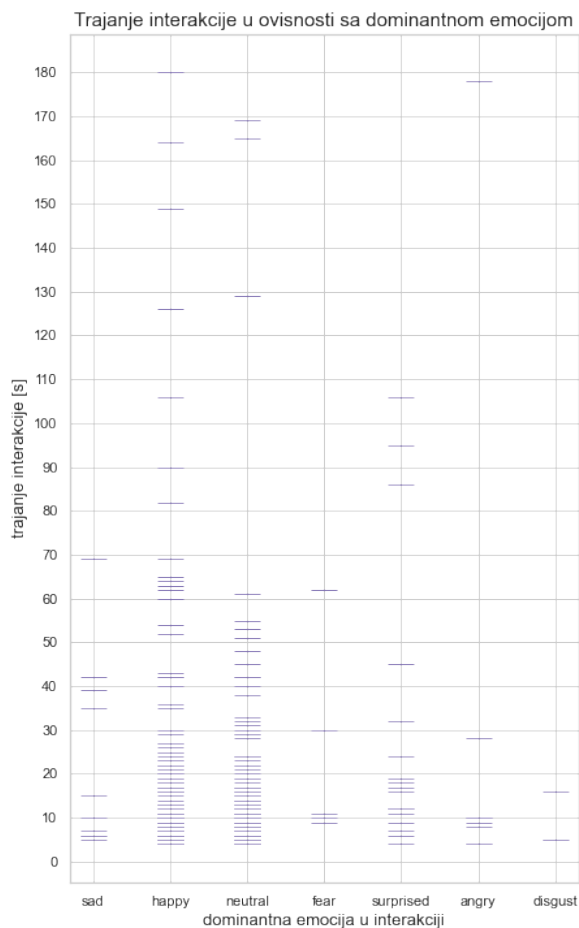


Slika 19. Raspodjela prosječnih emocija kod sudionika broj 121 (neutralnost kao dominantna emocija)



Slika 20. Vrijednosti emocija kroz interakciju sudionika broj 121 (neutralnost kao dominantna emocija)

Od sedam emocija čije se vrijednosti mjere i 5 kategorija trajanja interakcija, sve emocije nisu prisutne kao dominantne u svih 5 kategorija. Samo emocija sreće postoji kao dominantna u interakcijama svih kategorija trajanja. Emocija gađenja, upravo suprotno, postoji kao dominantna samo u najkraćim interakcijama kategorije „vrlo kratko” [Slika 21].

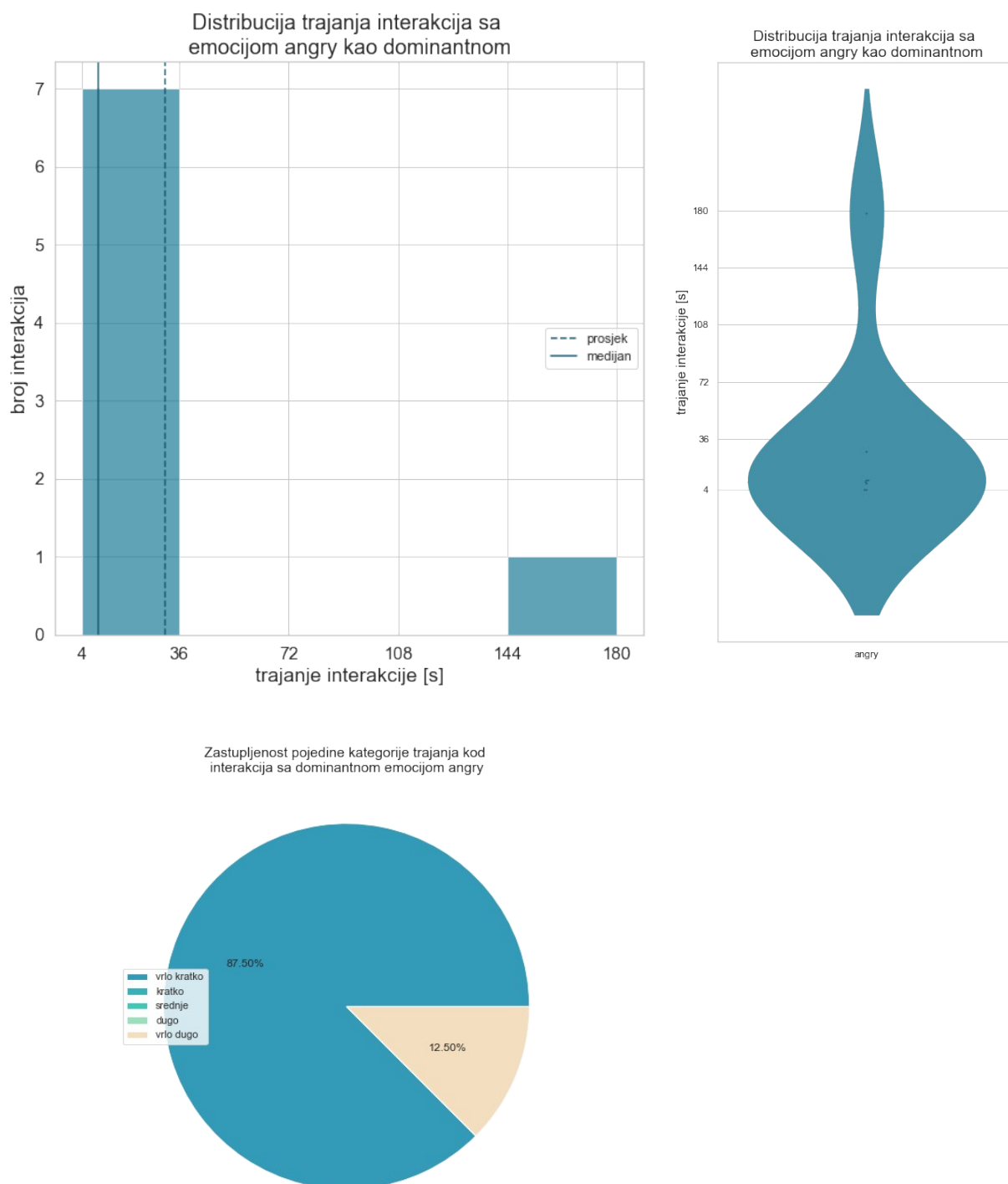


Slika 21. Trajanje interakcije u ovisnosti sa dominantnom emocijom

U nastavku su prikazani grafovi na kojima je moguće za svaku emociju vidjeti u kojem postotku se nalazi kao dominantna u kojoj kategoriji trajanja.

4.1.1.1. Ljutnja (angry)

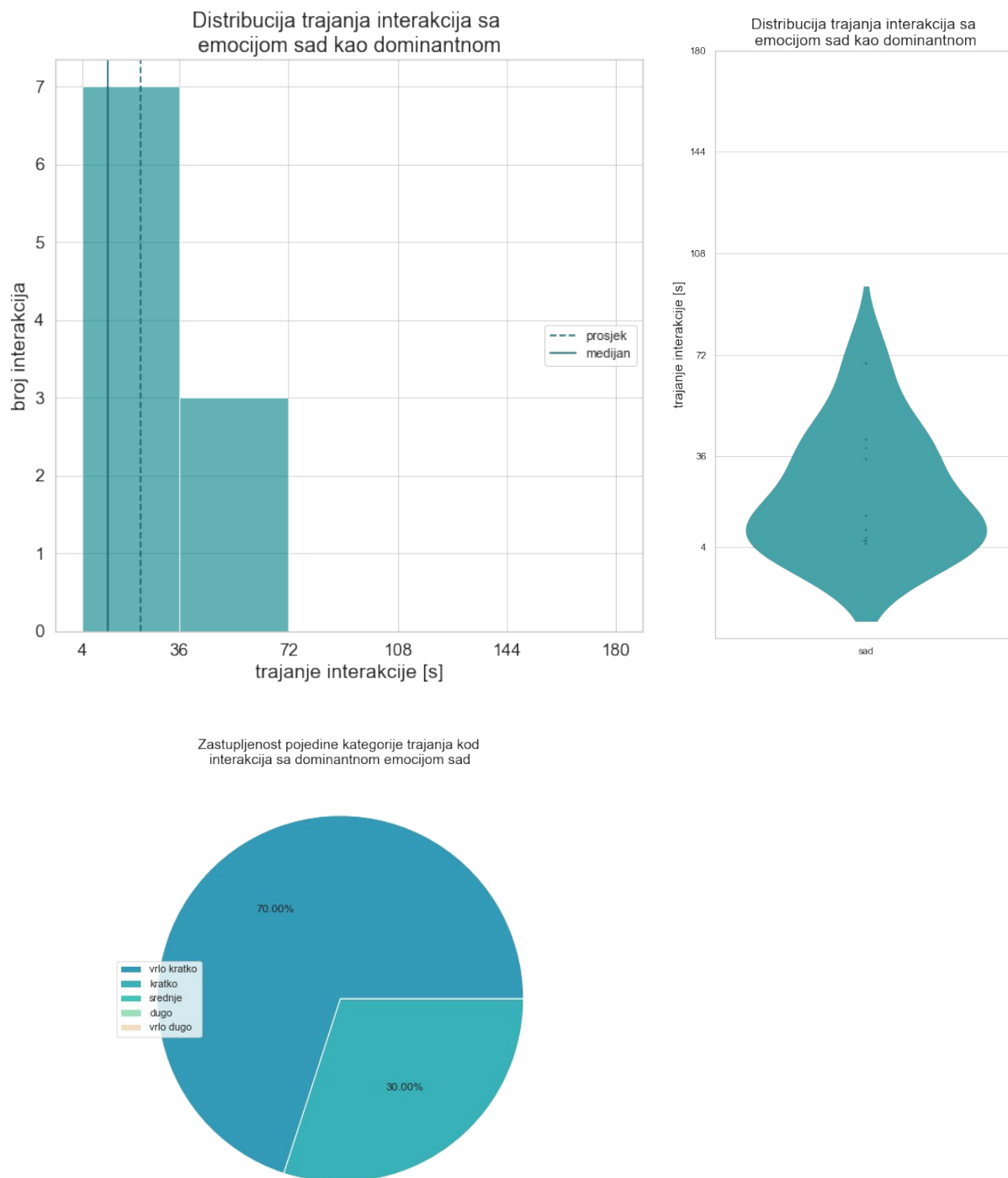
Što se tiče interakcija u kojima je emocija ljutnje dominantna, njih je 8. Prosjek njihovog trajanja je 31,375 sekundi, a medijan trajanja iznosi 9,5. 7 interakcija (87,5%) spada u grupu vrlo kratkih interakcija (kraćih od 36 sekunde), a samo jedna interakcija, tj. 12,5% je vrlo duga [Slika 22].



Slika 22. Grafički prikazi trajanja interakcija sa dominantnom emocijom ljutnje

4.1.1.2. Tuga (sad)

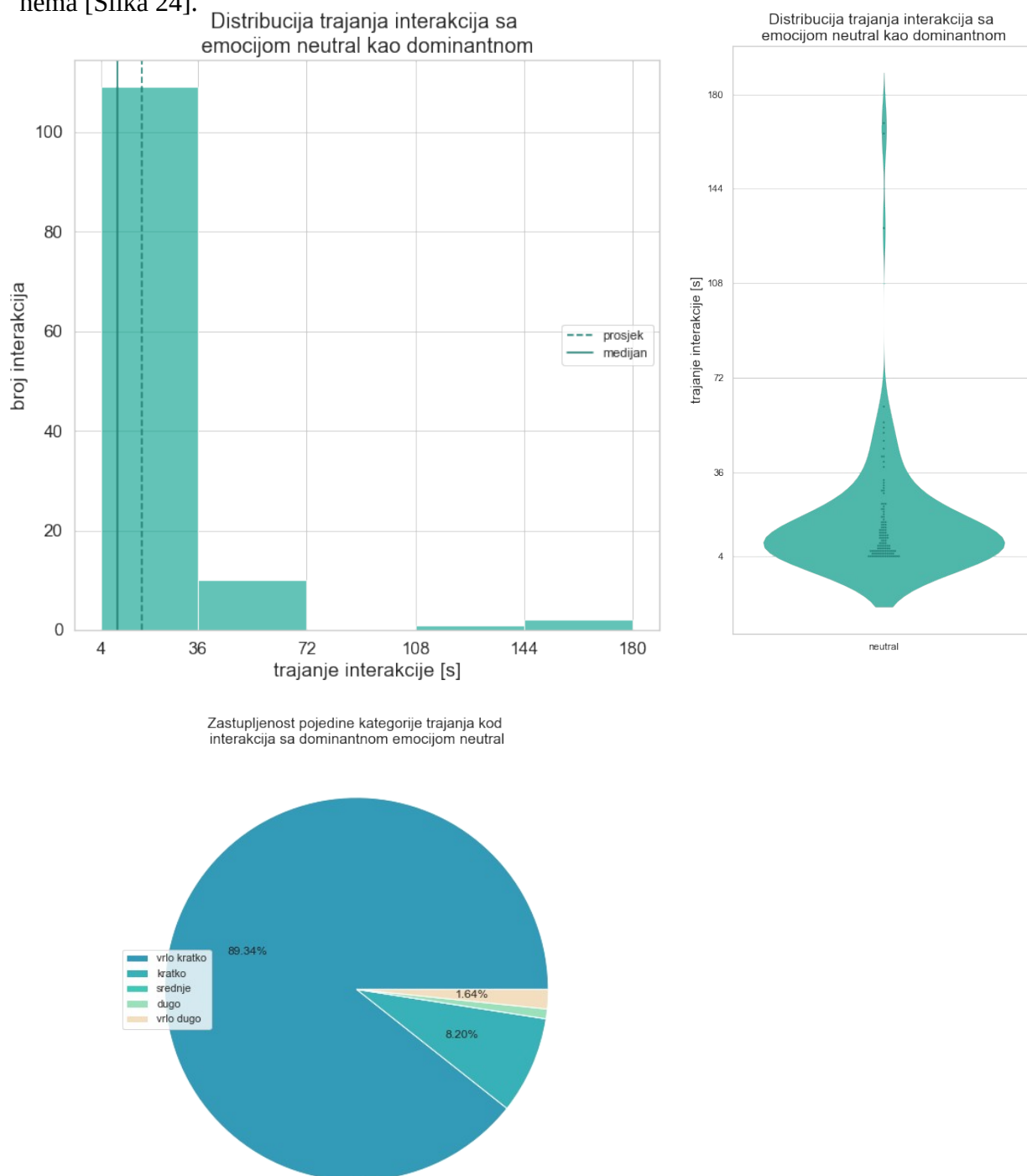
Interakcija u kojima je emocija tuge dominantna ima sveukupno 10. Prosjek njihovog trajanja je 23,4 sekundi, a medijan trajanja iznosi 12,5 sekundi. 7 (70%) interakcija je vrlo kratko (kraće od 36 sekundi), a 3 (30%) su kratke [Slika 23].



Slika 23. Grafički prikazi trajanja interakcija sa dominantnom emocijom tuge

4.1.1.3. Neutralnost (neutral)

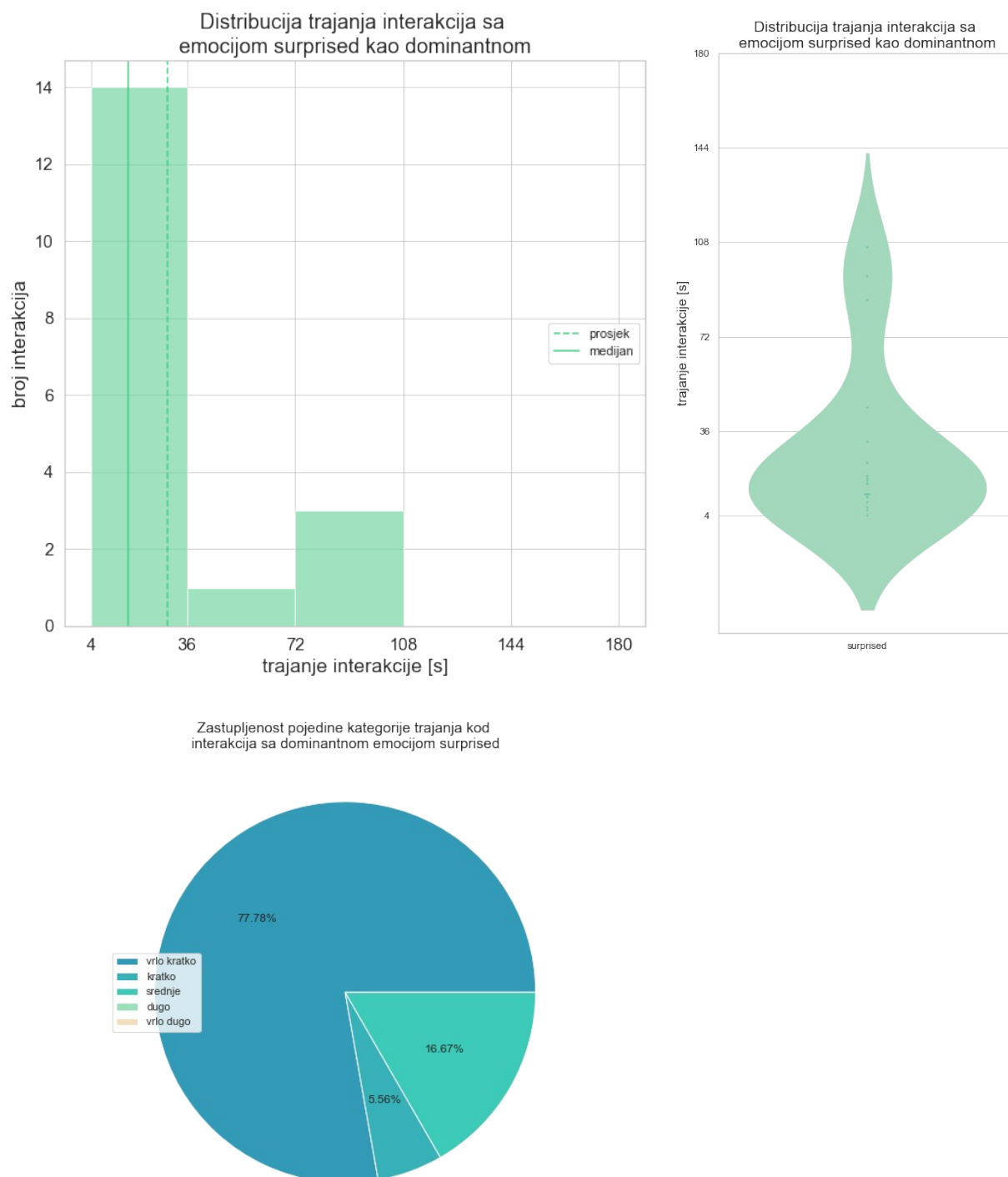
Što se tiče interakcija u kojima je emocija ljutnje dominantna, njih je 122. Prosjek njihovog trajanja je 17,5 sekundi, a medijan trajanja iznosi 9,5. Velika većina (89,34%) pripada kategoriji vrlo kratkih interakcija, no postoje i one drugih kategorija. 8,20% je kratko, 1,64% je kratko, 1,64% vrlo dugo, a ostatak dugo. Srednjih interakcija u kojima je emocija neutralnosti dominantna nema [Slika 24].



Slika 24. Grafički prikazi trajanja interakcija sa neutralnom dominantnom emocijom

4.1.1.4. Iznenadenost (surprised)

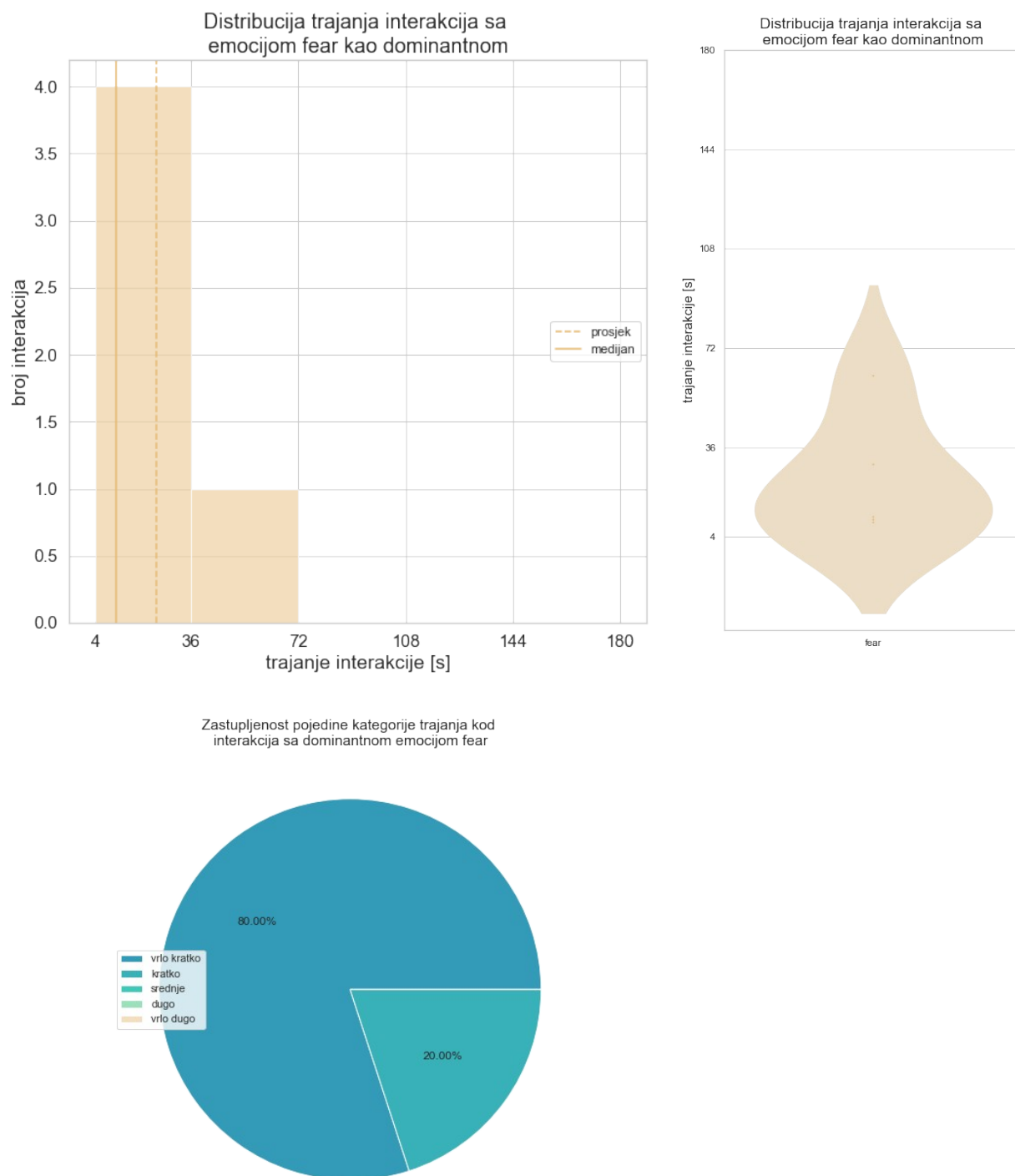
Interakcija u kojima je emocija iznenadenosti dominantna ima 18. Prosjek njihovog trajanja je 29,5, a medijan trajanja iznosi 16,5 sekundi. Iznenadenost postoji kao dominantna emocija kod interakcija koje spadaju u kategorije vrlo kratko, kratko i srednje. Od 100% interakcija u kojima je ona dominantna, tj. najizraženija, 77,78% njih je vrlo kratko, 16,67% srednje, a 5,56% kratko [Slika 25].



Slika 25. Grafički prikazi trajanja interakcija sa dominantnom emocijom iznenadenosti

4.1.1.5. Strah (*fear*)

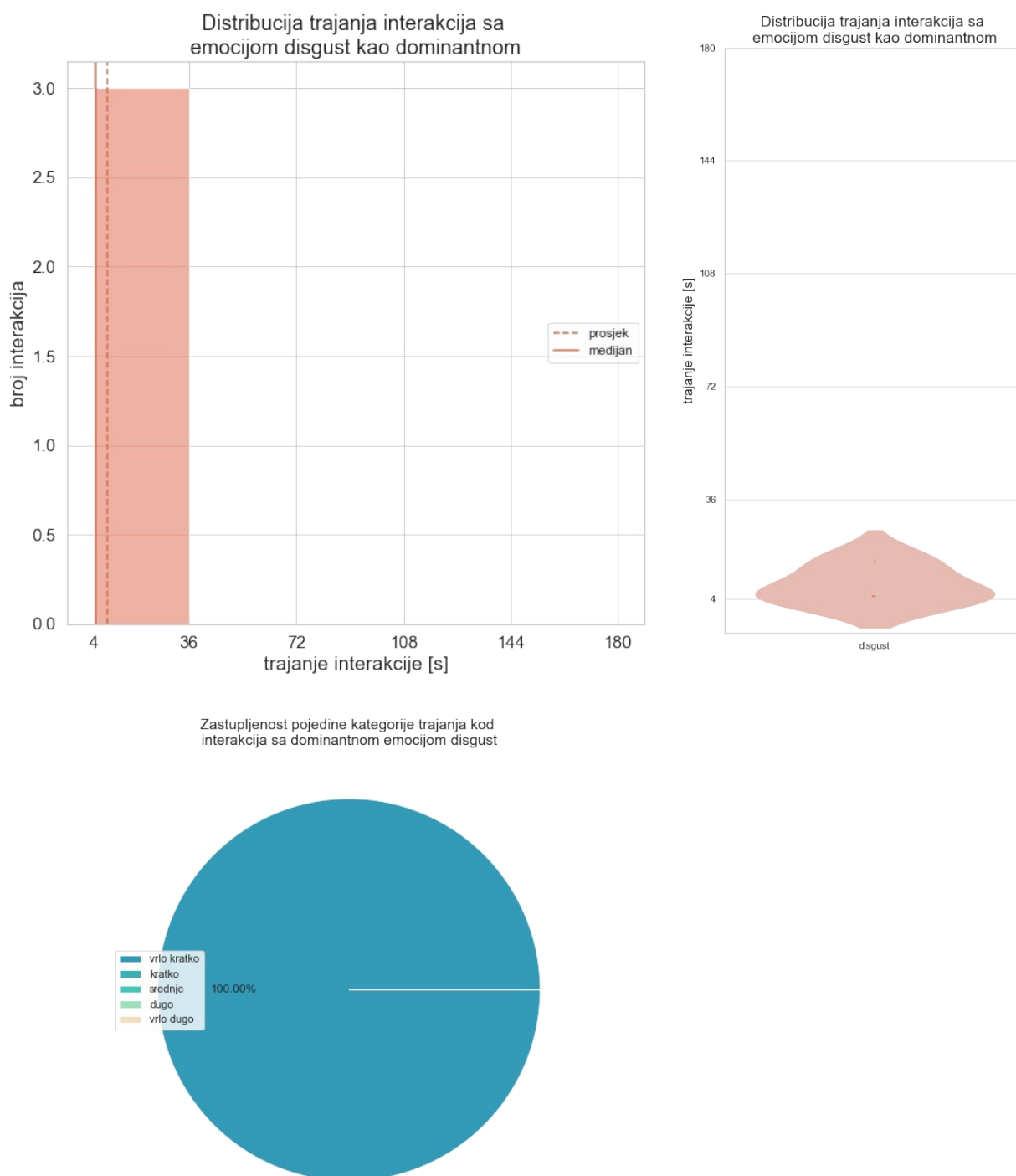
Interakcija u kojima je strah dominantan ima 5. Prosjek njihovog trajanja je 24,2 sekundi, a medijan trajanja iznosi 11. Sve su interakcija (80%) vrlo kratke, osim jedne (20%) koja je kratka [Slika 26].



Slika 26. Grafički prikazi trajanja interakcija sa dominantnom emocijom straha

4.1.1.6. Gađenje (disgust)

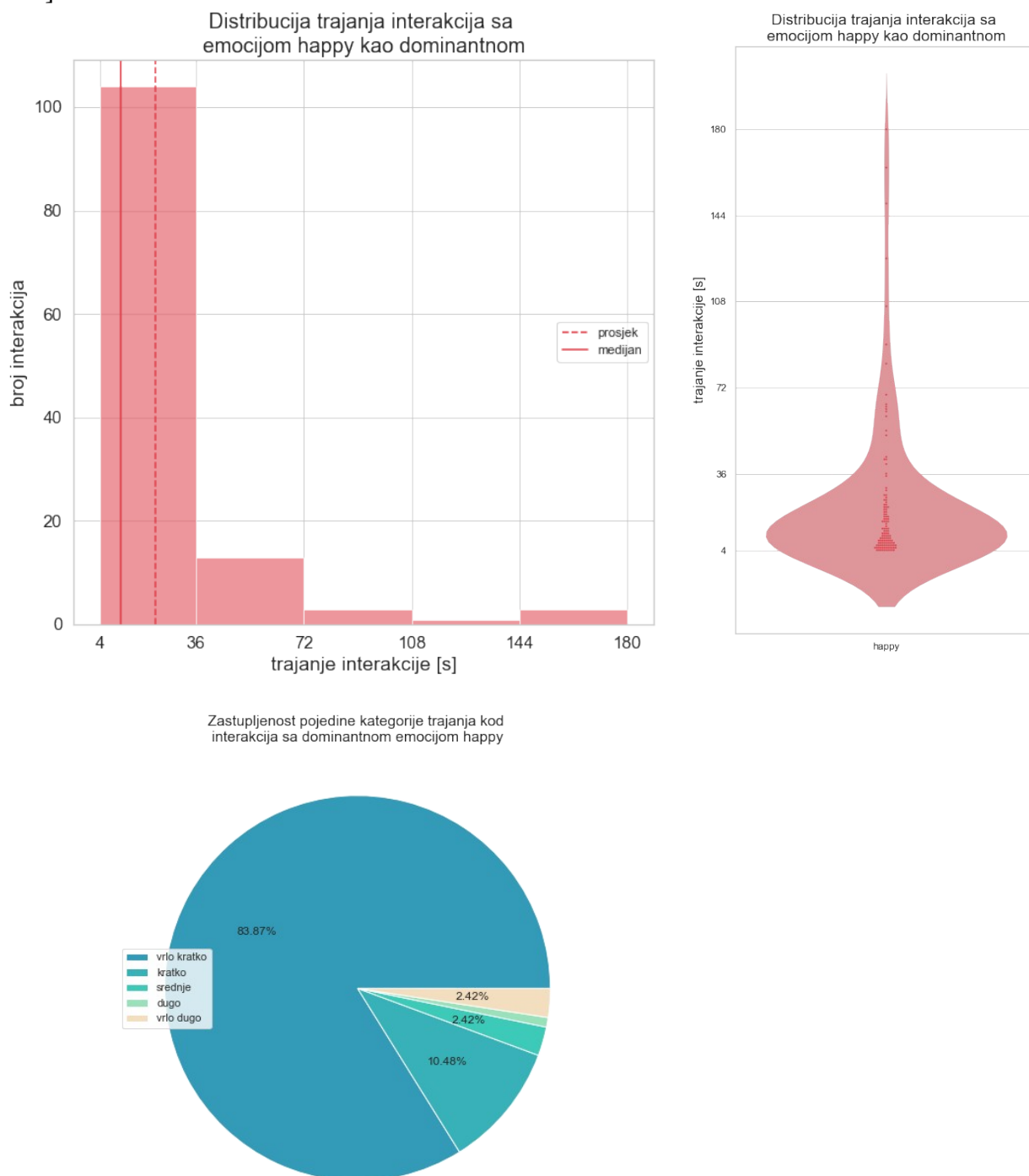
Postoje svega 3 interakcije u kojima je emocija gađenja dominantna te sve spadaju u kategoriju vrlo kratkih interakcija. Prosjek njihovog trajanja je 8,67 sekundi, a medijan 5 [Slika 27].



Slika 27. Grafički prikazi trajanja interakcija sa dominantnom emocijom gađenja

4.1.1.7. Sreća (happy)

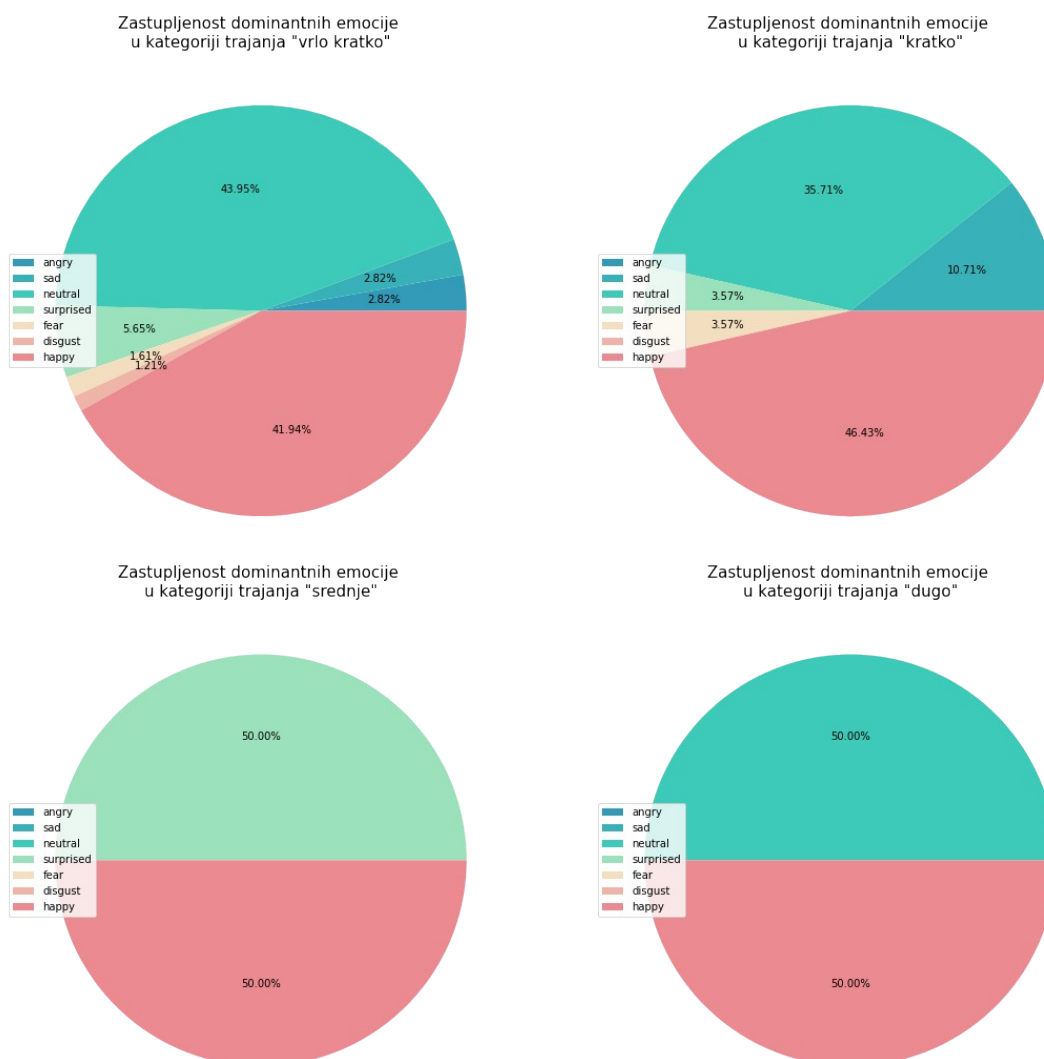
Emocija sreće je dominantna u 124 interakcije. Prosjek njihovog trajanja je 22,55 sekundi, a medijan trajanja iznosi 11. Velika većina interakcija (njih 100, tj. 83,87%) pripada grupi vrlo kratkih interakcija (kraćih od 36 sekunde), ali osim te postoje interakcije i drugih kategorija. 10,48% od svih interakcija je kratko, 2,42% srednje, 2,42% vrlo dugo, a ostatak dugo [Slika 28].



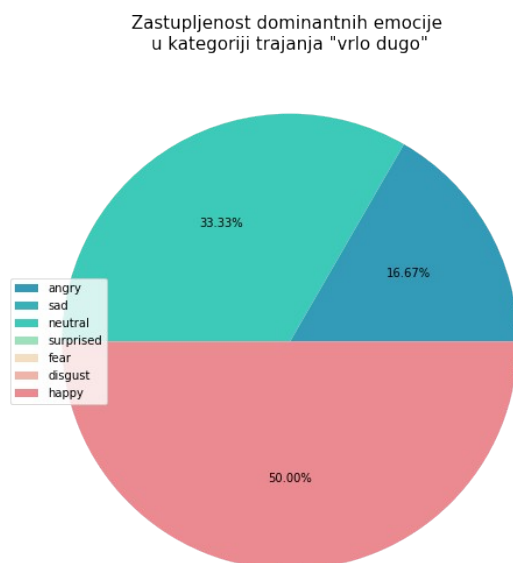
Slika 28. Grafički prikazi trajanja interakcija sa dominantnom emocijom sreće

4.1.2. Usporedba kategorija trajanja

Isti su podaci prikazani i kroz grafove pojedinih kategorija koji prikazuju koje se sve emocije i u kojem postotku pojavljuju kao dominantne u njima. Od svih interakcija koje su vrlo duge, u 50% njih je dominantna emocija sreće, 33,33% dominantna emocija neutralnosti, a 16,67% emocija ljutnje. Kod interakcija koje pripadaju kategoriji dugih interakcija pola njih ima neutralnu emociju kao dominantno, a druga polovica sreću. Kod srednjih interakcija, kod 50% njih je iznenađenje dominantno, a 50% sreća. Od 100% interakcija koje su kratke, najviše (46,43%) ih ima sa srećom kao dominantnom emocijom, 35,71% ih je sa neutralnom emocijom, 10,71% sa ljutnjom i po 3,57% sa strahom i iznenađenošću. Najviše je vrlo kratkih interakcija, njih 23,95% ima neutralnu emociju kao dominantnu, 41,94% sreću, 5,65% iznenađenje, 2,82% njih je sa ljutnjom te tugom, kod 1,61% od svih interakcija je strah dominantan, te najmanje njih (1,21%) ima gađenje kao dominantnu emociju [Slika 29][Slika 30].



Slika 29. Zastupljenost dominantnih emocija u pojedinoj kategoriji trajanja (I)



Slika 30. Zastupljenost dominantnih emocija u pojedinoj kategoriji trajanja (II)

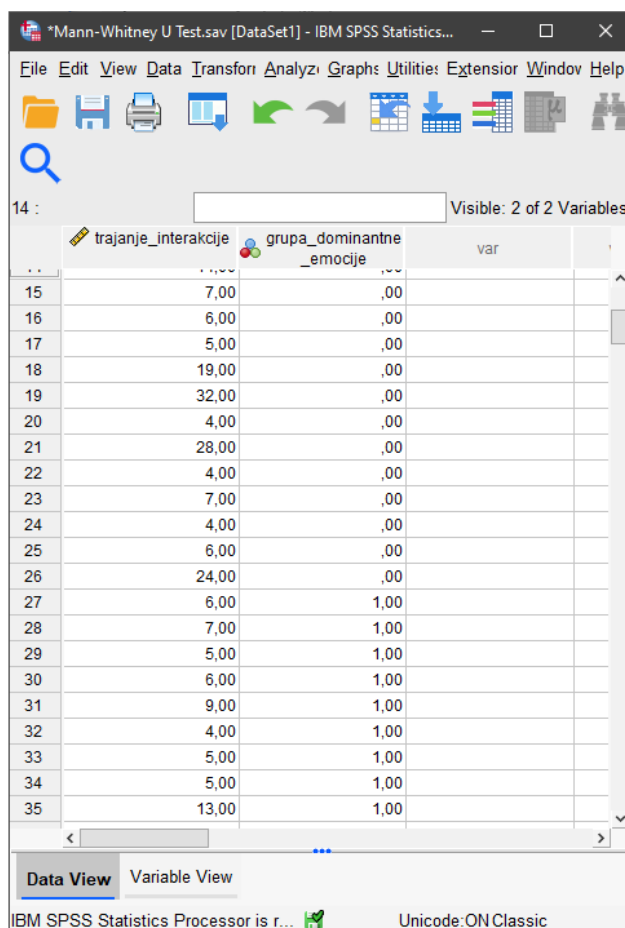
4.2. Mann-Whitneyjev U test

Mann-Whitneyjev U test je neparametrijski statistički test koji se koristi za usporedbu razlika između dviju neovisnih skupina kada je zavisna varijabla ili ordinalna ili kontinuirana, ali nije normalno distribuirana. U ovom slučaju, koristi se za uočavanje razlike u trajanju interakcija (koja se mjeri kontinuiranom skalom) ovisno o vrsti dominantne emocije i ovisno o prvoj izraženoj emociji. U provjeri obje hipoteze zavisna varijabla je *trajanje interakcije*. U prvoj provjeri je nezavisna varijabla *dominantna emocija*, a u drugoj *prva izražena emocija*. U oba slučaja, nezavisna varijabla ima dvije grupe: *pozitivna* i *negativna emocija*. Mann-Whitneyjev U test često se smatra neparametrijskom alternativom nezavisnom t-testu. Neparametrijski testovi koriste se u slučajevima kada distribucija podataka nije normalna te u nekim drugim specifičnim slučajevima kada nisu zadovoljeni svi preduvjeti za korištenje odgovarajućeg parametrijskog testa.

Da bi se koristio Mann-Whitneyjev U test potrebno je zadovoljiti 4 pretpostavke:

1. Zavisna varijabla mora biti mjerena na ordinalnoj ili kontinuiranoj skali
2. Nezavisna varijabla se mora sastojati od dvije nezavisne skupine
3. Treba biti zadovoljena neovisnost promatranja (*independence of observations*), tj. u svakoj grupi moraju biti različiti sudionici, pri čemu niti jedan sudionik nije u više od jedne grupe
4. Obje distribucije moraju imati isti oblik (npr. obje su pozitivno asimetrične)

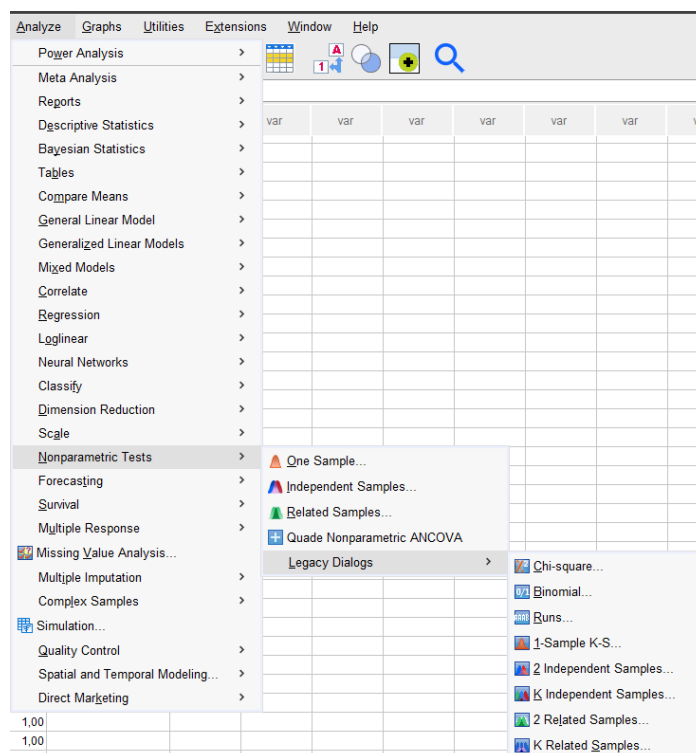
Sve četiri pretpostavke su zadovoljne te je Mann-Whitneyjev test proveden u *SPSS Statistics* softveru. Provođenje testa u softveru je vrlo jednostavno. U nastavku je prikazan način provođenja za provjeru prve hipoteze [13]. Prvo je potrebno u softver unijeti podatke istraživanja.



	trajanje_interakcije	grupa_dominantne_emocije
15	7,00	,00
16	6,00	,00
17	5,00	,00
18	19,00	,00
19	32,00	,00
20	4,00	,00
21	28,00	,00
22	4,00	,00
23	7,00	,00
24	4,00	,00
25	6,00	,00
26	24,00	,00
27	6,00	1,00
28	7,00	1,00
29	5,00	1,00
30	6,00	1,00
31	9,00	1,00
32	4,00	1,00
33	5,00	1,00
34	5,00	1,00
35	13,00	1,00

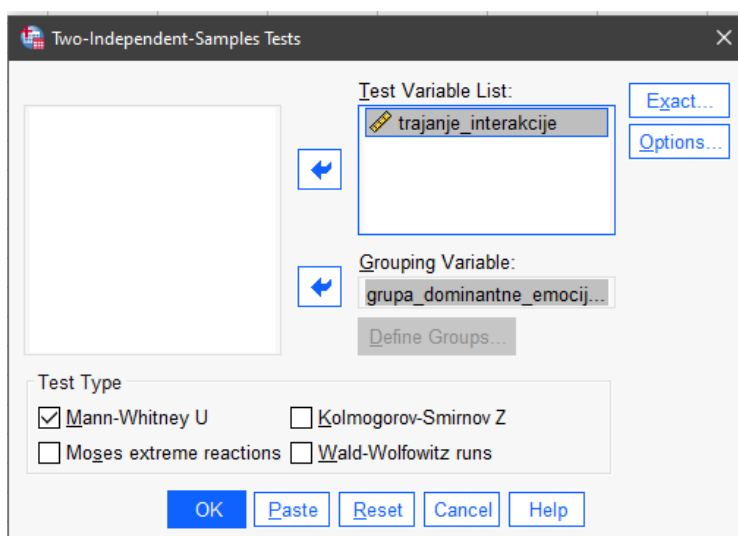
Slika 31. Podaci interakcija u SPSS softveru

Zatim je potrebno kliknuti na *Analyze > Nonparametric Tests > Legacy Dialogs > 2 Independent Samples...* [Slika 32]

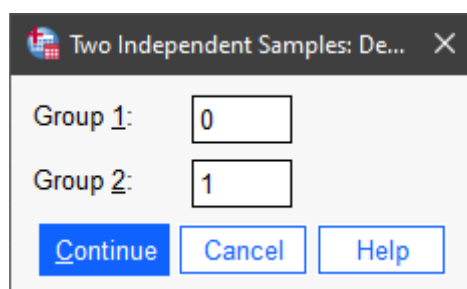


Slika 32. Odabir neparametrijskog testa

Zatim se otvara *Two-Independent-Samples Tests* dijaloški prozor na kojem je potrebno postaviti varijable, njihove odnose te odrediti grupe nezavisne varijable [Slika 33]. U oba testiranja grupe su postavljene kao 0 za negativne emocije, a 1 za pozitivne [Slika 34]. Nakon toga se generiraju podaci za provedeni test.



Slika 33. Definiranje varijabli statističkog testa



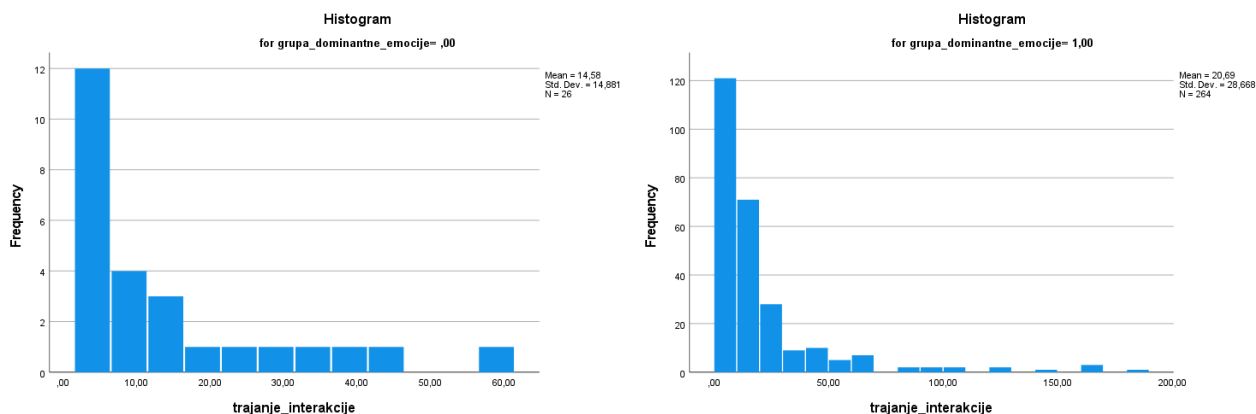
Slika 34. Definiranje grupa nezavisne varijable

4.2.1. Testiranje prve hipoteze

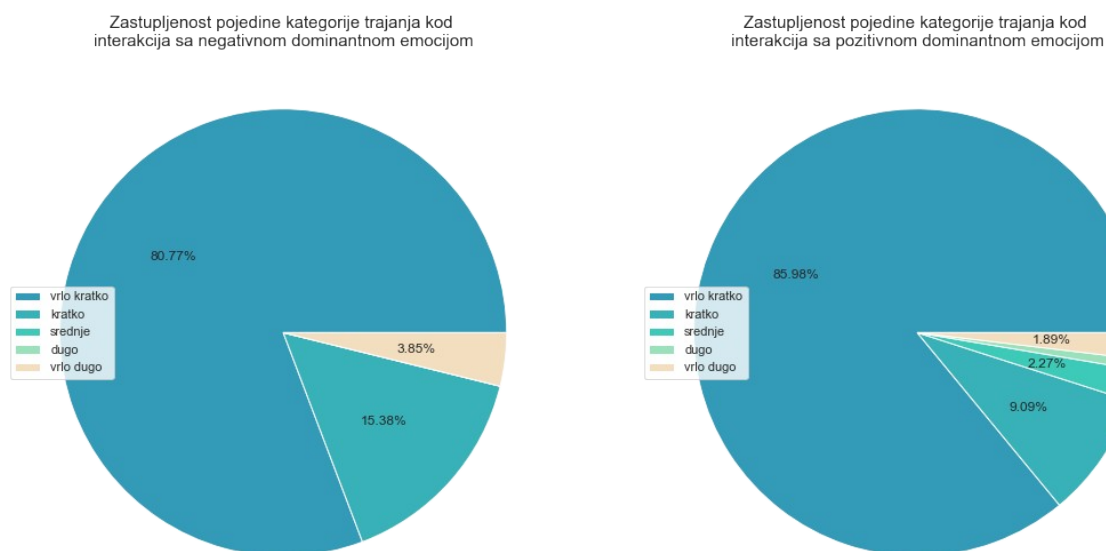
Za potrebe testiranja emocije su podijeljene u dvije kategorije: pozitivne emocije i negativne. Pozitivne su sreća, neutralnost i iznenađenost, a negativne strah, tuga, gađenje i ljutnja. Zavisna varijabla nazvana je *trajanje_interakcije*, a nezavisna *grupa_dominantne_emocije*. Budući da nezavisna varijabla ima dvije grupe – *pozitivna* i *negativna*, pozitivna je grupa označena sa vrijednošću 1, a negativna sa 0. Uz pomoć *SPSS Statistics* programa moguće je prikazati i vrijednosti deskriptivne statistike. [Tablica 9]

Tablica 9. Deskriptivna statistika (prva hipoteza)

Descriptives				
		grupa_dominantne_emocije		
			Statistic	Std. Error
trajanje_interakcije	,00	Mean	14,5769	2,91847
		95% Confidence Interval for Mean		
		Lower Bound	8,5662	
		Upper Bound	20,5876	
		5% Trimmed Mean	12,7906	
		Median	7,0000	
		Variance	221,454	
		Std. Deviation	14,88133	
		Minimum	4,00	
		Maximum	61,00	
		Range	57,00	
		Interquartile Range	15,25	
		Skewness	1,807	,456
		Kurtosis	2,880	,887
	1,00	Mean	20,6932	1,76442
		95% Confidence Interval for Mean		
		Lower Bound	17,2190	
		Upper Bound	24,1674	
		5% Trimmed Mean	15,7803	
		Median	11,0000	
		Variance	821,879	
		Std. Deviation	28,66843	
		Minimum	4,00	
		Maximum	180,00	
		Range	176,00	
		Interquartile Range	15,75	
		Skewness	3,370	,150
		Kurtosis	12,776	,299



Slika 35. Histogrami trajanja interakcija (prva hipoteza)



Slika 36. Zastupljenost kategorija trajanja kod pozitivnih i negativnih dominantnih emocija

Tablica rangova prva je tablica koja pruža informacije o rezultatu Mann-Whitney U testa [Tablica 10]. Prikazuje rang prosjeka (*mean rank*) i sumu rangova (*sum of ranks*) za dvije testirane skupine. Dvije testirane skupine su one kod kojih je dominantna negativna emocija (označena sa ,00) i one kod kojih je ona pozitivna (označeno sa 1,00). Tablica pokazuje koja grupa ima dulje interakcije (ona sa većim prosjekom ranga). Ovdje je to grupa s pozitivnom dominantnom interakcijom. Njen je prosjek ranga 147,59, što je veće od prosjeka ranga interakcija sa negativnom dominantnom emocijom koji je 124,27.

Tablica 10. Rangovi statističkog testa (prva hipoteza)**Mann-Whitney Test**

		Ranks		
	grupa_dominantne_emocije	N	Mean Rank	Sum of Ranks
trajanje_interakcije	,00	26	124,27	3231,00
	1,00	264	147,59	38964,00
	Total	290		

Tablica statistike provedenog testa prikazuje značajnost testa [Tablica 11]. Najvažnija vrijednost u ovoj tablici je p-vrijednost koja je prikazana u zadnjem redu tablica (*asympt. Sig (2-tailed)*) koja iznosi 0,175. Budući da je ta vrijednost veća od početno postavljene vrijednosti značaja koja iznosi 0,05 rezultat istraživanja nije statistički značajan. Odnosno, nulta hipoteza (*Dominantna emocija ne utječe na trajanje interakcije*) ne može biti opovrgnuta.

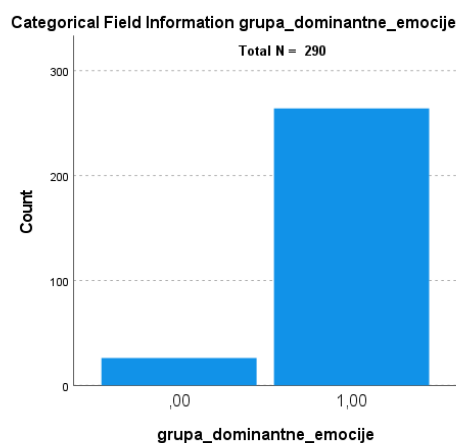
Tablica 11. Statistika provedenog testa (prva hipoteza)

Test Statistics ^a	
	trajanje_interakcije
Mann-Whitney U	2880,000
Wilcoxon W	3231,000
Z	-1,356
Asymp. Sig. (2-tailed)	,175

a. Grouping Variable:
grupa_dominantne_emocije

Iako je Mann-Whitneyjev test moguće provesti upravo u ovakvom slučaju gdje broj prikupljenih podataka za dvije različite skupine nije jednak, u ovom slučaju to može predstavljati problem zbog velike razlike.

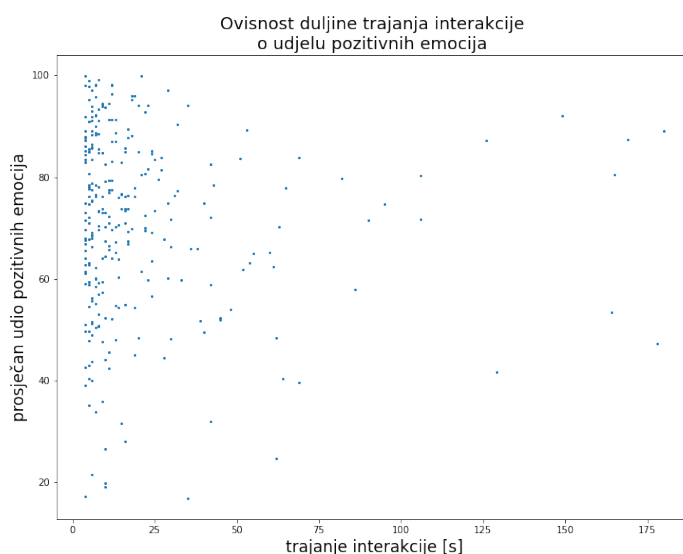
Moguće je vidjeti na grafu ispod da je kod 264 od 290 ispitanika dominantna emocija tijekom interakcije pozitivna [Slika 37]. Postoji mogućnost da bi rezultati testiranja hipoteze bili drugačiji ukoliko bi broj interakcija obje grupa bio podjednak. Također je moguće da na rezultat utječe to što je većina reakcija vrlo kratka te bi se povećanjem broja duljih interakcija rezultat testiranja hipoteze promijenio.



Slika 37. Broj interakcija pojedine grupe (SPSS Statistics)

4.2.2. Korelacija

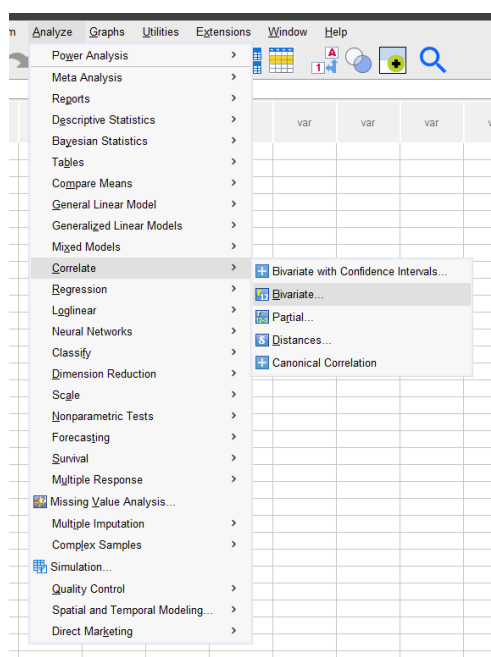
U nastavku je promatrana korelacija između trajanja interakcije i postotka pozitivnih emocija tijekom nje. Za analizu korelacije potrebno je izračunati koeficijent korelacije - vrijednost između -1 i 1 koja govori o jačini i smjeru odnosa između varijabli. Vrijednost 1 označava savršeno pozitivnu korelaciju, -1 savršeno negativnu, a 0 znači da nema korelacije. Na slici ispod prikazan je graf na kojem je na x osi trajanje interakcija, a na y prosječan udio pozitivnih emocija u interakciji [Slika 38]. Na grafu je moguće vidjeti da ne postoji korelacija između te dvije varijable, što je provjereno i računanjem Kendallovog tau-b u softveru *SPSS Statistics*.



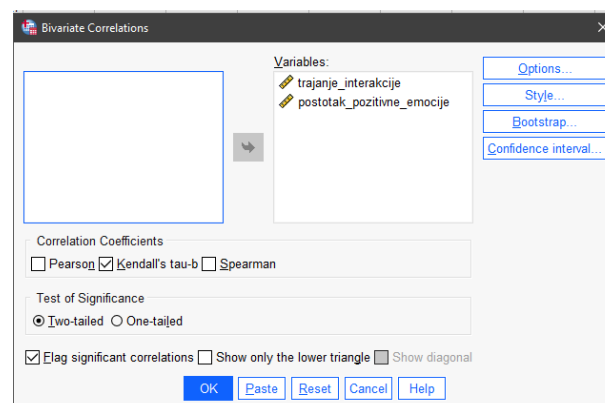
Slika 38. Ovisnost duljine trajanja interakcije o udjelu pozitivnih emocija

Kendallov tau-b (τ_b) koeficijent korelacije (skraćeno Kendallov tau-b) neparametarijska je mjera snage i smjera povezanosti koja postoji između dviju varijabli mjerenih na barem ordinalnoj ljestvici. Glavni uvjet koji se mora zadovoljiti kako bi se provelo računanje je to da su obje varijable mjerene na ordinalnoj ili kontinuiranoj skali (što je u ovom slučaju zadovoljeno) [14].

Princip provođenja testa u *SPSS Statistics* softveru je vrlo jednostavan [15]. Potrebno je kliknuti na *Analyze > Correlate > Bivariate...* [Slika 39]. Nakon toga treba na prikazanom dijaloškom okviru odabrati željene varijable i označiti računanje Kendallovog tau [Slika 40] poslije čega se prikazuju rezultati [Tablica 12].



Slika 39. Odabir računanja Kendallovog tau-b u SPSS Statistics



Slika 40. Odabir postavki računanja Kendallovog tau-b

U danoj tablici prikazani su rezultati koji prikazuju kako nema korelacije između dvije zadane varijable. Budući da p-vrijednost iznosi 0,386 što je veće od 0,05 alfa vrijednosti, rezultat nije statistički značajan te se zaključuje da korelacije nema.

Tablica 12. Rezultati provjere Kendallovog tau-b

			trajanje_interakcije	postotak_pozitivne_emocije
Kendall's tau_b	trajanje_interakcije	Correlation Coefficient	1,000	-,035
		Sig. (2-tailed)	.	,386
		N	290	290
	postotak_pozitivne_emocije	Correlation Coefficient	-,035	1,000
		Sig. (2-tailed)	,386	.
		N	290	290

Ti rezultati dodatno potvrđuju rezultate provjere prve hipoteze provedene putem Mann-Whitneyevog testa.

4.2.3. Problemi korelacijskog istraživanja

Treba imati na umu da je ovdje riječ o korelacijskom istraživanju, a ne uzročnom. Korelacija opisuje povezanost između varijabli (kada se jedna varijabla promijeni, mijenja se i druga). Ona je statistički pokazatelj odnosa između varijabli gdje se one mijenjaju zajedno (kovariiraju) što nije nužno posljedica izravne ili neizravne uzročne veze. Uzročnost znači da promjene u jednoj varijabli izazivaju promjene u drugoj, tj. između varijabli postoji uzročno-posljedična veza. Korelacija ne implicira uzročnost, no uzročnost uvijek implicira korelaciju.

Glavni razlog zbog kojeg korelacija ne podrazumijeva uzročnost u ovom istraživanju je problem treće varijable. Zbog kojeg su obje varijable prividno kauzalno povezane. Na primjer, prodaja sladoleda i stopa nasilnog kriminala su u uskoj korelaciji, ali nisu međusobno uzročno povezane. Treća varijabla, visoka temperatura je ta koja utječe na obje varijable zasebno. Povećanjem temperature ljudi kupuju više sladoleda te se češće upuštaju u kriminalne aktivnosti. Prodaja sladoleda i kriminal međusobno nisu u ni u kakvoj ovisnosti. U slučaju ovog istraživanja, moguće je postojanje treće varijable poput odbojnosti prema tehnologiji i robotima ili nedostatka slobodnog vremena za sudjelovanje u istraživanju, što utječe na to kako se ljudi osjećaju tijekom interakcije i koliko vremena "potroše" na nju. Ukoliko netko osjeća vrlo veliku odbojnost prema tehnologiji, vrlo je vjerojatno da će se tijekom korištenja iste osjećati negativno te će manje vremena željeti uložiti u njeno korištenje.

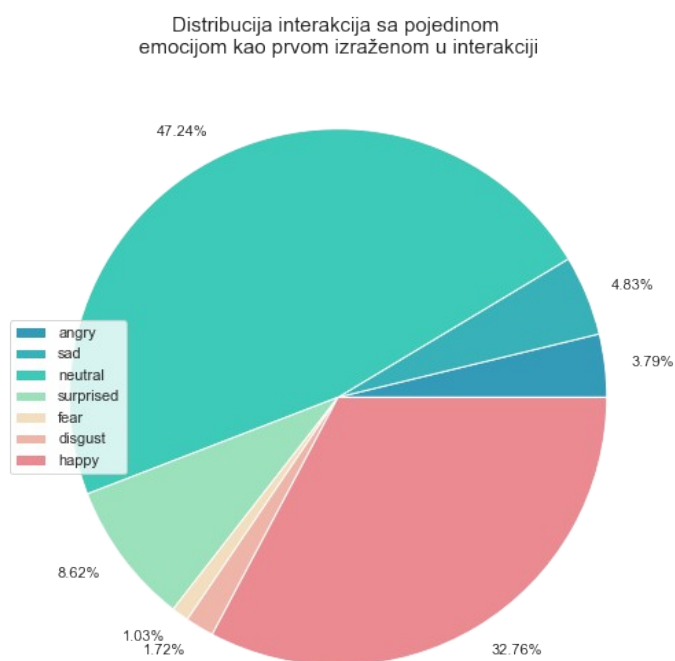
4.3. Druga hipoteza

Druga alternativna hipoteza koja se testira glasi: Prva izražena emocija utječe na trajanje interakcije, a njena nulta hipoteza: Prva izražena emocija nema utjecaja na trajanje interakcije. Prvi dojmovi su trenutni zaključci koje ljudi donose kada se susretnu s nekim ili nečim. To je mišljenje koje se stvara brzim prikupljanjem informacija. Ako se radi o susretu s osobom, ono uključuje informacije lica, odjeće, držanja i tona glasa. Sa stvaranjem prvog dojma razvijaju se i trenutni osjećaji. Ukoliko je prvi stečeni dojam pozitivan, emocija koju osoba osjeća je pozitivna. Isto vrijedi i za prvi dojam koji je negativan. Ukoliko je emocija pozitivna, ljudi prirodno žele zadržati tu emociju, a ako je negativna, žele je se čim prije " riješiti". U ovom

slučaju interakcije čovjeka i robota, hipoteza govori da je početna izražena emocija koja se razvija tijekom prvog dojma o robotu povezana sa duljinom interakcije koja se ostvaruje.

4.3.1. Odnos trajanja interakcije i prve izražene emocije tijekom interakcije

Prva emocija koju su sudionici najčešće razvili tijekom interakcije je neutralna emocija. Najmanje je interakcija u kojima je prva izražena emocija strah (svega 3 interakcije).



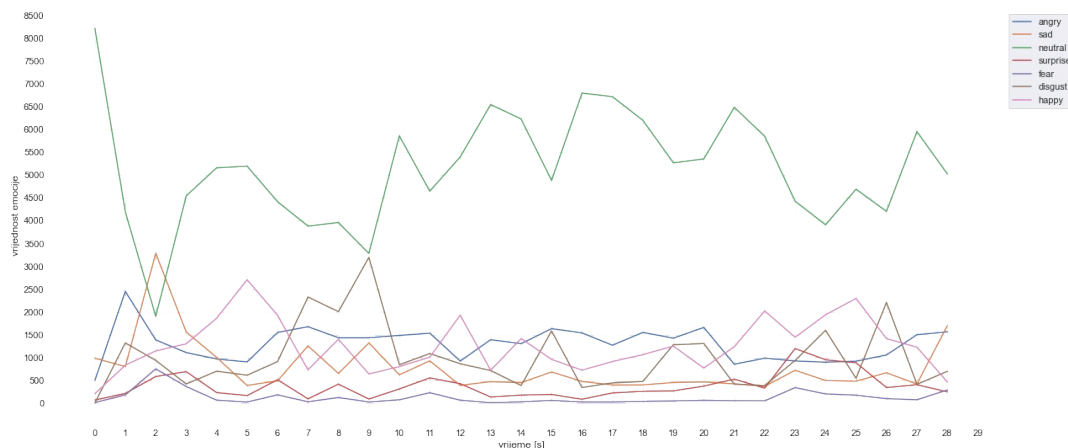
Slika 41. Distribucija interakcija sa pojednom emocijom kao prvom izraženom u interakciji

U nastavku će radi lakšeg razumijevanja biti prikazani podaci dvije interakcije sa različitim prvim emocijama.

Prvo je prikazana interakcija sudionika broja 63 čija je prva izražena emocija neutralna. Njegova je ukupna interakcija sa robotom trajala 28 sekundi. Na grafu vrijednosti emocija kroz interakciju moguće je vidjeti neutralnost kao prvu emociju [Slika 42]. To potvrđuju i prosjeci vrijednosti emocije u prve dvije sekunde interakcije prikazane u tablici ispod [Tablica 13]. Prosječna vrijednost neutralne emocije u prve dvije se sekunde iznosi 4773 što je veće od prosječnih vrijednosti svih ostalih emocija.

Tablica 13. Prosječne vrijednosti emocija u prve 2 sekunde interakcije sudionika broj 63 (neutralnost kao prva izražena emocija)

Emocija	Prosječna vrijednost emocije u prve dvije sekunde interakcije
ljutnja (angry)	1445,667
tuga (sad)	1692,667
neutralnost (neutral)	4773
iznenađenost (surprised)	290
strah (fear)	315,3333
gađenje (disgust)	754,6667
sreća (happy)	728,6667

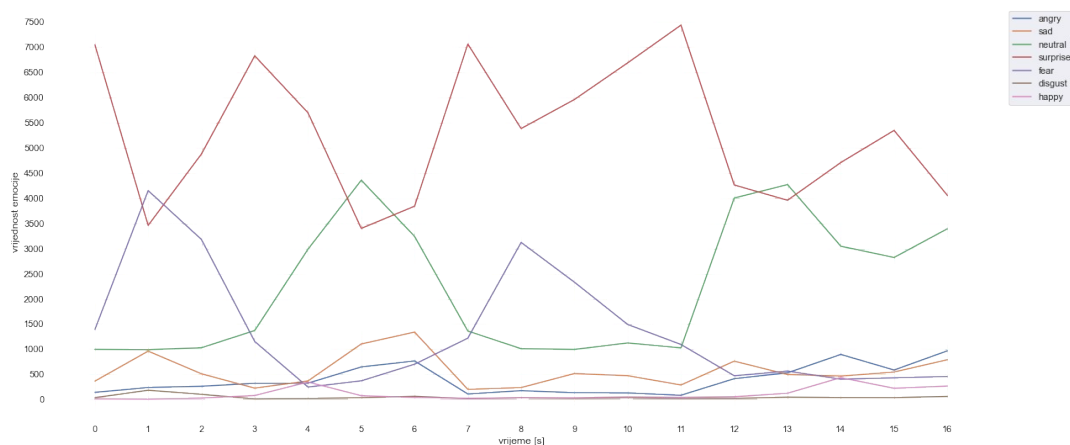


Slika 42. Vrijednosti emocija kroz interakciju sudionika broj 63 (neutralnost kao prva izražena emocija)

Drugi primjer je interakcija sudionika broj 279. Njegova je ukupna interakcija trajala 17 sekundi te je prva izražena emocija iznenađenost. Njezina vrijednost u prve dvije sekunde interakcije iznosi 5127,333, što je najviše od svih emocija. Interakciju je moguće vidjeti na grafu ispod [Slika 43], a vrijednosti u danoj tablici [Tablica 14].

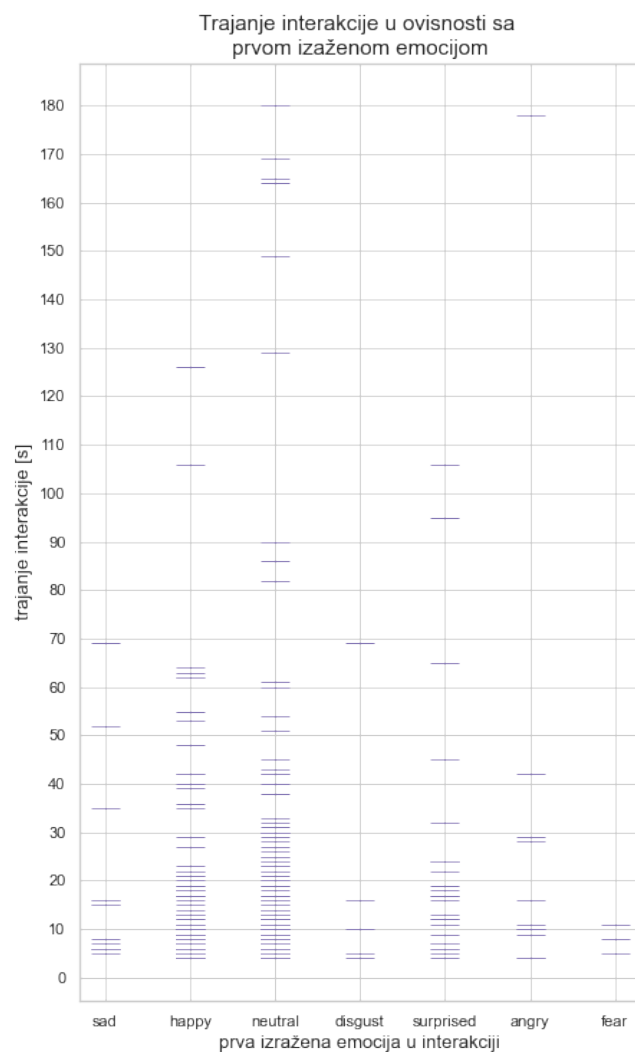
Tablica 14. Prosječne vrijednosti emocija u prve 2 sekunde interakcije sudionika broj 279 (iznenađenost kao prva izražena emocija)

Emocija	Prosječna vrijednost emocije u prve dvije sekunde interakcije
ljutnja (angry)	1445,667
tuga (sad)	1692,667
neutralnost (neutral)	4773
iznenađenost (surprised)	290
strah (fear)	315,3333
gađenje (disgust)	754,6667
sreća (happy)	728,6667



Slika 43. Vrijednosti emocija kroz interakciju sudionika broj 279 (iznenađenost kao prva izražena emocija)

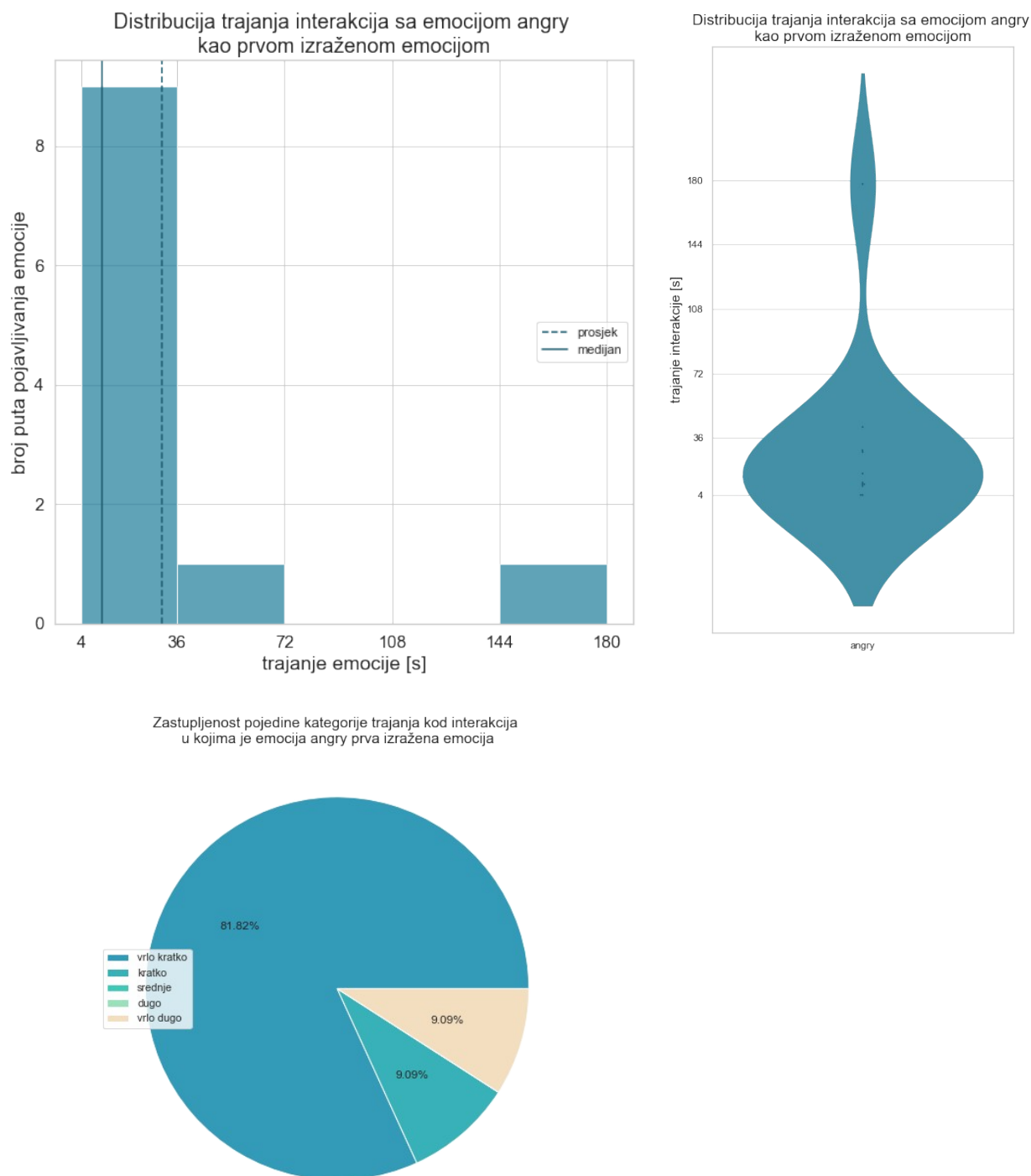
Isto i kao sa dominantnim emocijama u interakciji, ne postoje interakcije svih kategorije trajanja za svaku pojedinu emociju [Slika 44]. Primjerice, emocija straha prisutna je kao prva izražena emocija samo u kategoriji vrlo kratkih interakcija.



Slika 44. Trajanje interakcije u ovisnosti sa prvom izraženom emocijom

4.3.1.1. Ljutnja (angry)

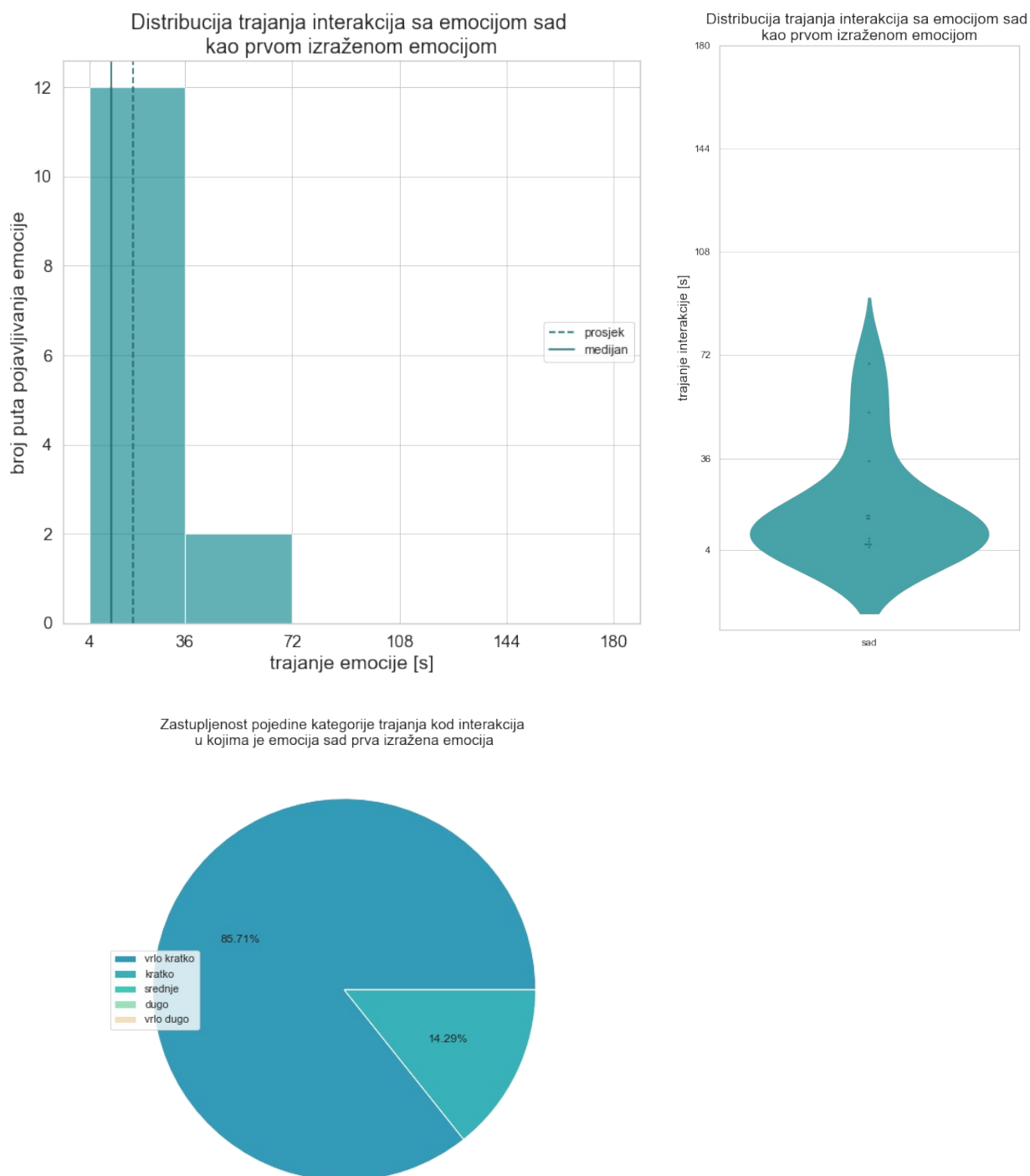
Interakcija u kojima je emocija ljutnje prva izražena emocija ima 11. Prosjek njihovog trajanja je 31 sekundi, a medijan trajanja iznosi 11. 9 interakcija (81,82%) spada u grupu vrlo kratkih interakcija (kraćih od 36 sekunde), dok po jedna interakcija (9,09% od svih interakcija) pripada vrlo dugim i kratkim interakcijama [Slika 45].



Slika 45. Grafički prikazi trajanja interakcija sa prvom izraženom emocijom ljutnje

4.3.1.2. Tuga (sad)

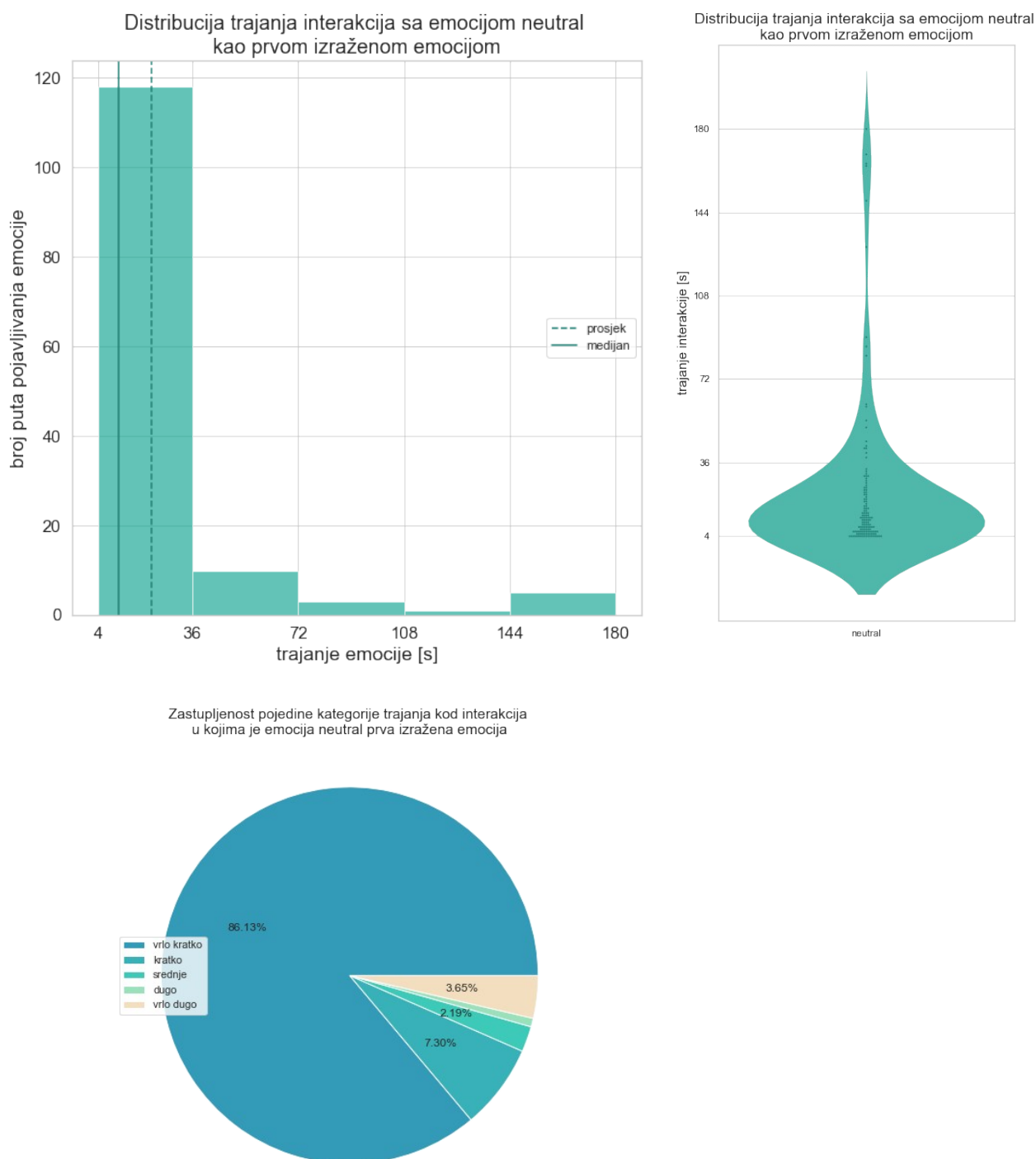
Interakcija u kojima je emocija tuge prva izražena ima sveukupno 14. Prosjek njihovog trajanja je 18,7 sekundi, a medijan trajanja iznosi 11,5 sekundi. 12 (85,71%) interakcija je vrlo kratko (kraće od 36 sekundi), a 2 (14,29%) su kratke [Slika 46].



Slika 46. Grafički prikazi trajanja interakcija sa prvom izraženom emocijom tuge

4.3.1.3. Neutralnost (neutral)

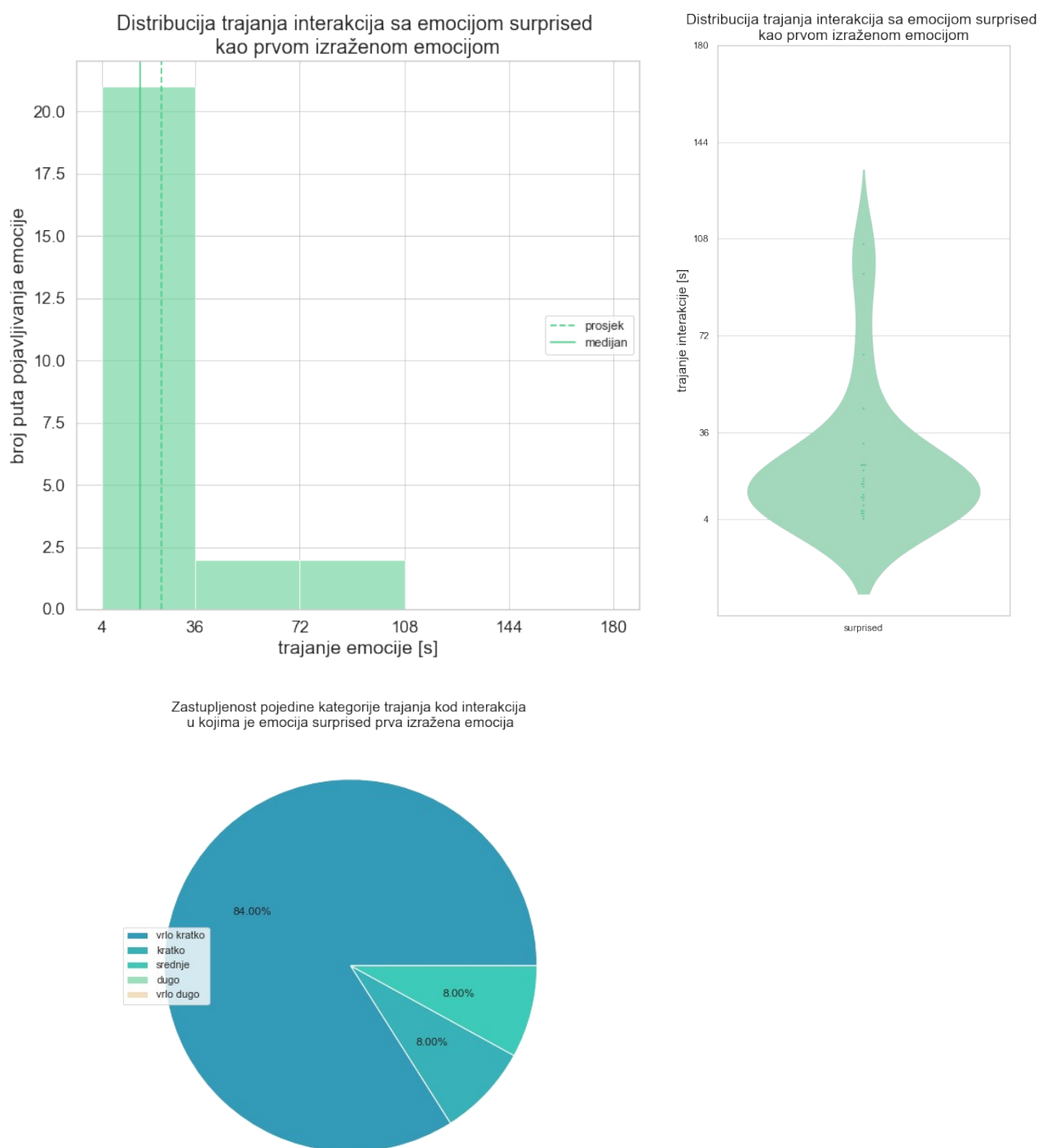
Što se tiče interakcija u kojima je emocija ljutnje prva izražena, njih je 137. Prosjek njihovog trajanja je 22,34 sekundi, a medijan trajanja iznosi 11. Velika većina interakcija (86,13%) pripada kategoriji vrlo kratkih interakcija, ali postoje i interakcije svih drugih kategorija. 7,30% je kratko, 3,65% vrlo dugo, 2,19% srednje, a ostatak dugo [Slika 47].



Slika 47. Grafički prikazi trajanja interakcija sa prvom izraženom emocijom neutralnosti

4.3.1.4. Iznenadenost (surprised)

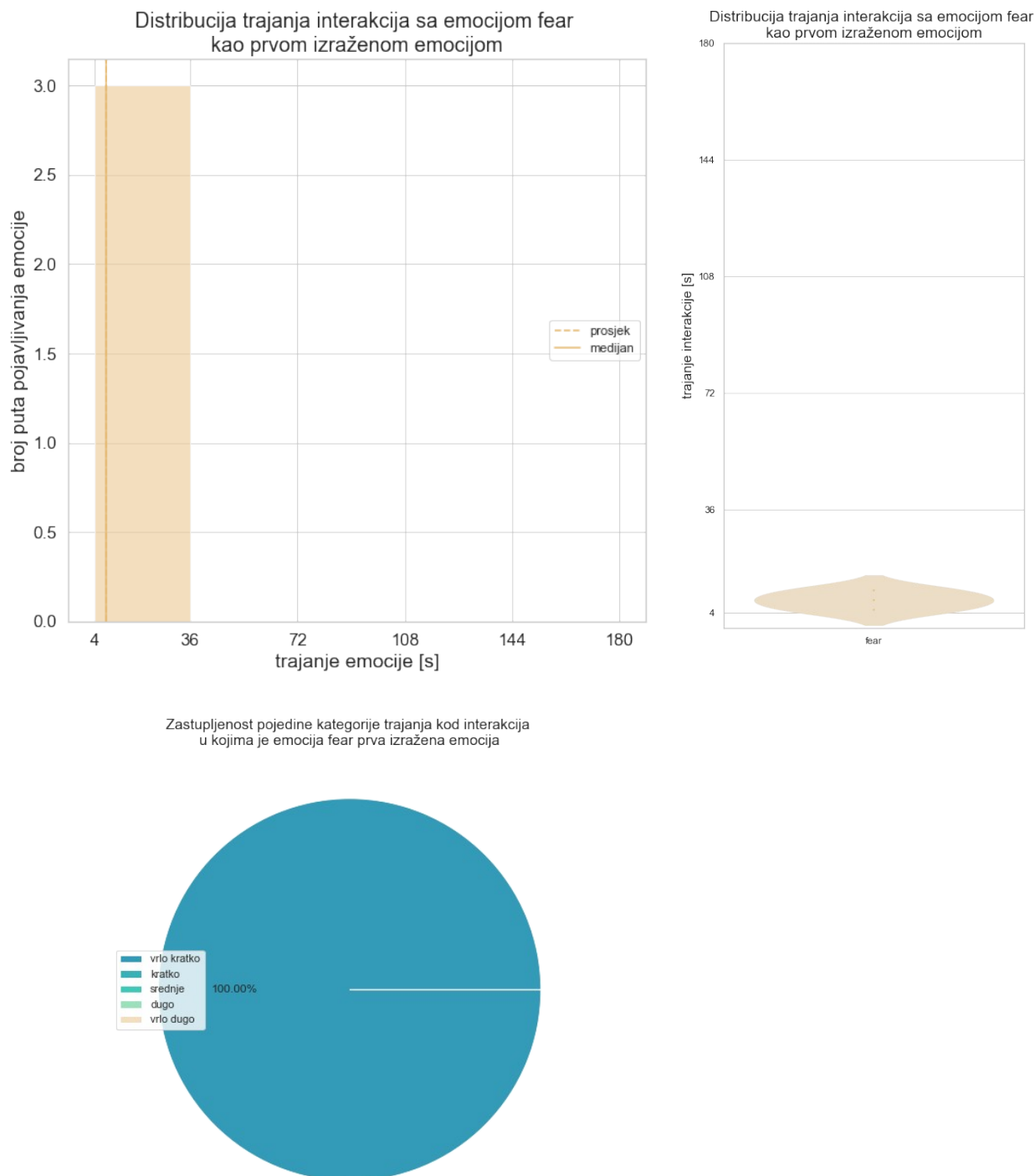
Interakcija u kojima je emocija iznenadenosti prva izražena ima 25. Prosjek njihovog trajanja je 24,64, a medijan trajanja iznosi 17 sekundi. Iznenadenost postoji kao prva izražena emocija kod interakcija koje spadaju u kategorije vrlo kratko, kratko i srednje. 21 interakcija (84%) je vrlo kratko, a po dvije su u skupini kratkih i srednjih interakcija [Slika 48].



Slika 48. Grafički prikazi trajanja interakcija sa prvom izraženom emocijom iznenadenosti

4.3.1.5. Strah (fear)

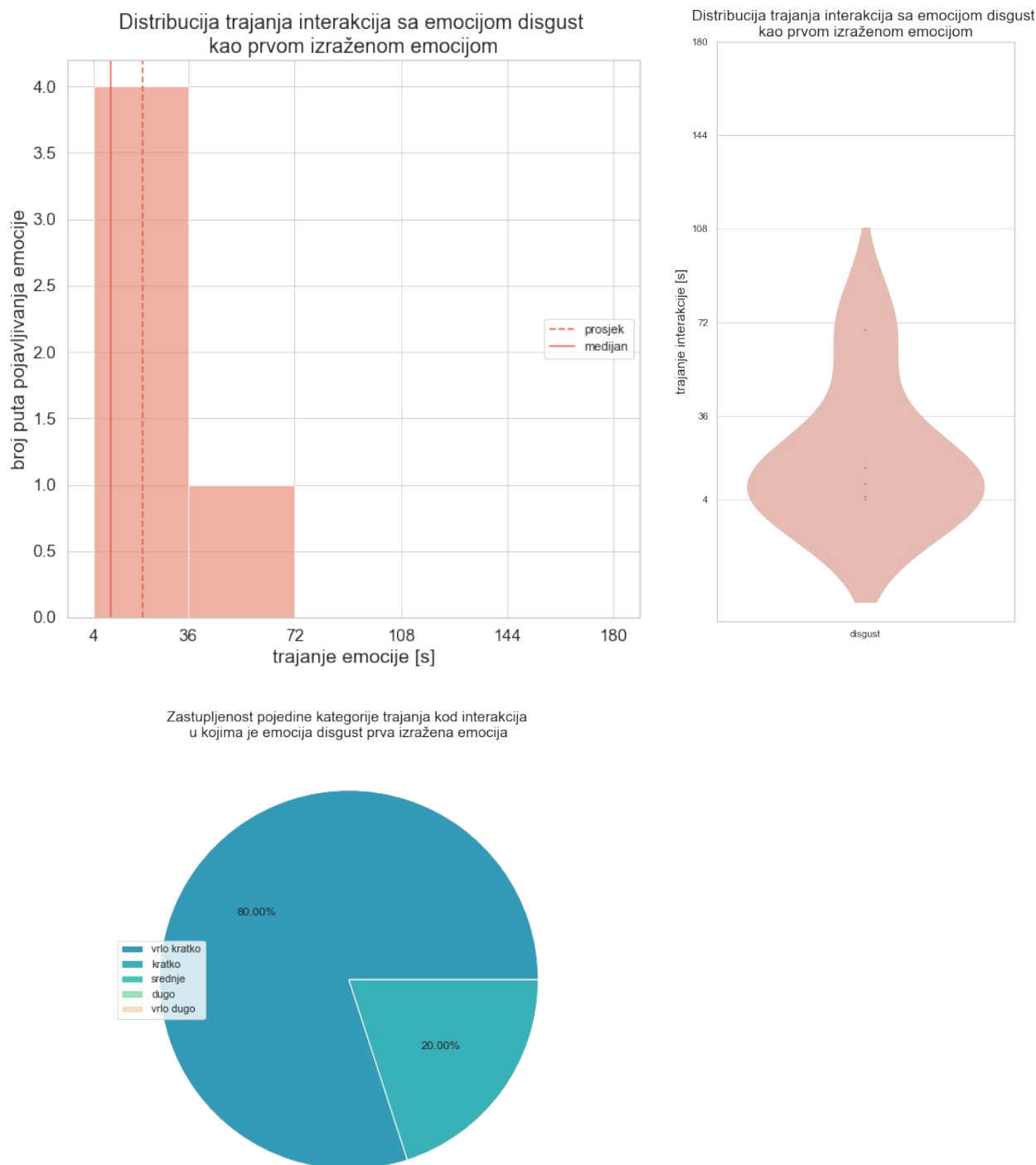
Strah je prva izražena emocija u samo 3 interakcije. Prosjek njihovog trajanja je 8 sekundi, a medijan trajanja iznosi 8. Sve 3 interakcije su vrlo kratke [Slika 49].



Slika 49. Grafički prikazi trajanja interakcija sa prvom izraženom emocijom straha

4.3.1.6. Gađenje (disgust)

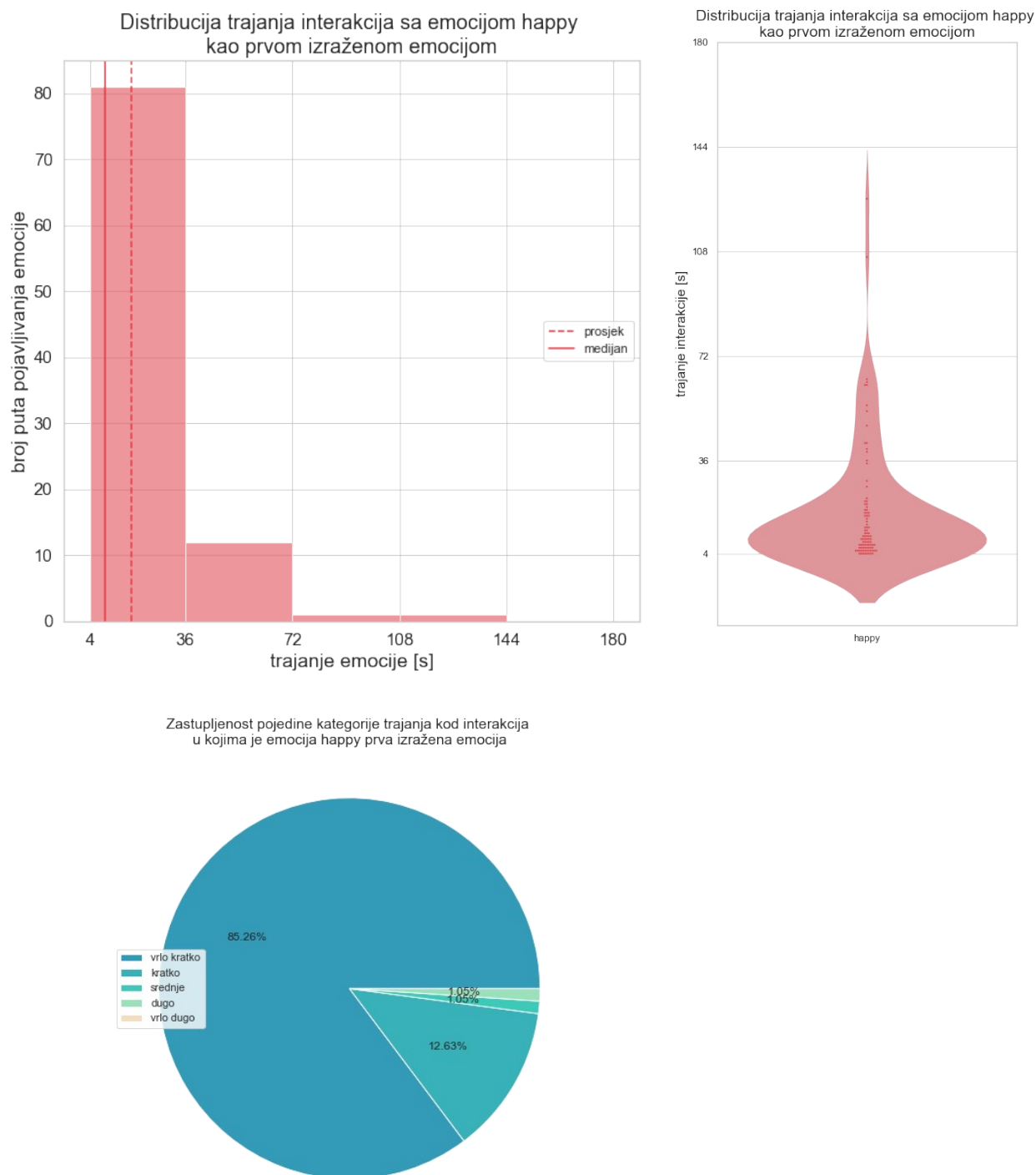
Postoji 5 interakcija u kojima je emocija gađenja prva izražena emocija. Prosjek njihovog trajanja je 20,8 sekundi, a medijan 10. 4 interakcije (80%) su vrlo kratke, a 1 (20%) je kratka [Slika 50].



Slika 50. Grafički prikazi trajanja interakcija sa prvom izraženom emocijom gađenja

4.3.1.7. Sreća (happy)

Emocija sreće je prva izražena emocija u 95 interakcija. Prosjek njihovog trajanja je 17,78 sekundi, a medijan trajanja iznosi 9. Velika većina interakcija (njih 80, tj. 85,26%) pripada grupi vrlo kratkih interakcija (kraćih od 36 sekunde). Osim tih, postoje i interakcije drugih kategorija. 12,63% ih je kratko, a 1,05% njih srednje i 1,05% dugo [Slika 51].



Slika 51. Grafički prikazi trajanja interakcija sa prvom izraženom emocijom sreće

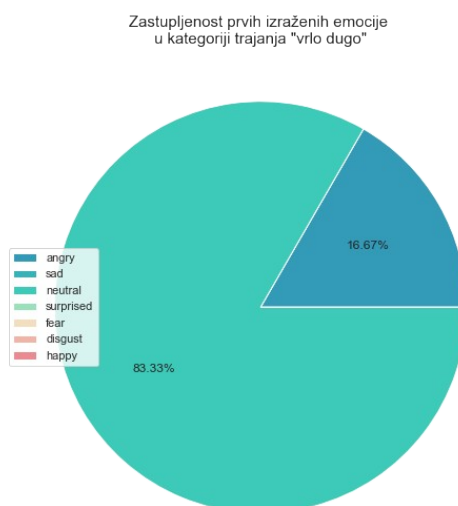
4.3.2. Usporedba kategorija trajanja

Ponovo su isti podaci prikazani i kroz grafove pojedinih kategorija koji prikazuju koje se sve emocije i u kojem postotku pojavljuju kao prve izražene u interakcijama. Od svih interakcija koje su vrlo duge, u 83,33% njih je prva izražena emocija neutralna, a 16,67% ljutnja. Kod interakcija koje pripadaju kategoriji dugih interakcija pola njih ima neutralnu emociju kao prvu izraženu, a druga polovica sreću. Kod interakcija srednjeg trajanja 50% njih ima neutralnost kao prvu izraženu emociju, 33,33% iznenađenost, a 16,67% sreću. Od 100% interakcija koje su kratke, najviše (42,86%) ih ima emociju sreće za prvu izraženu emociju, 35,71% neutralnu, po 7,14% iznenađenost i tugu, a po 3,57% gađenje i tugu. Najviše je vrlo kratkih interakcija. Njih 47,58% ima neutralnu emociju kao prvu izraženu tijekom interakcije, 32,66% sreću, 8,47% iznenađenost, 4,84% tugu, 3,63% gađenje i tugu. Najviše je vrlo kratkih interakcija. Njih 47,58% ima neutralnu emociju kao prvu izraženu tijekom interakcije, 32,66% sreću, 8,47% iznenađenost, 4,84% tugu, 3,63% gađenje i tugu, 1,61% gađenje te 1,21% strah

[Slika 52][Slika 53]



Slika 52. Zastupljenost prvih izraženih emocija u pojedinoj kategoriji trajanja (I)



Slika 53. Zastupljenost prvih izraženih emocija u pojedinoj kategoriji trajanja (II)

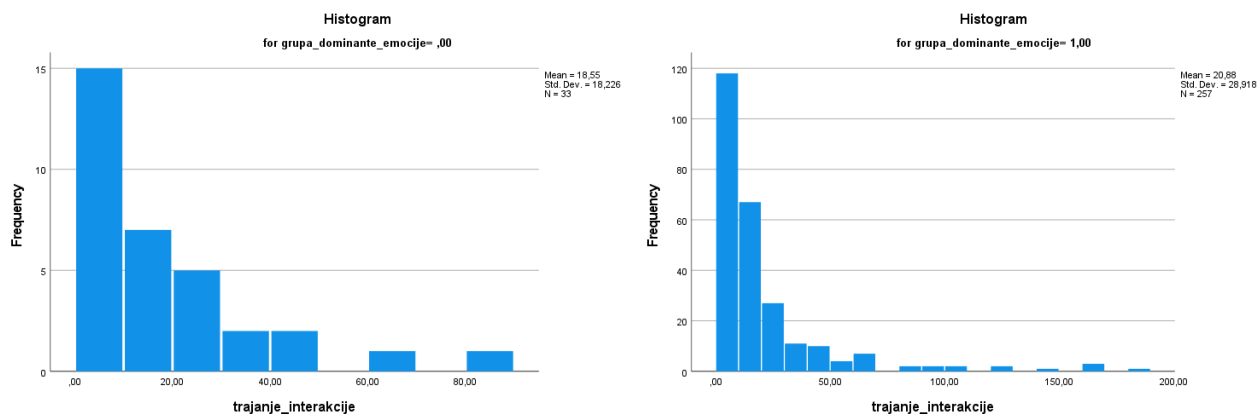
4.3.3. Testiranje druge hipoteze

Način testiranja druge hipoteze jednak je testiranju prve. Koristi se Mann-Whitneyjev U test u *SPSS Statistics* softveru. Za 290 sudionika prikupljene su duljine trajanja te podaci koja je emocija prva izražena tijekom interakcije. Za potrebe testa, prva izražena emocija je definirana kao najdominantnija emocija tijekom prve dvije sekunde interakcije. Emocije su podijeljene u dvije grupe kao i u prošlom testiranju. Pozitivne emocije su sreća, neutralnost i iznenađenost, a negativne strah, gađenje, ljutnja i tuga. Dvije varijable koje se koriste su *trajanje_interakcije* te *grupa_dominantne_emocije* gdje je sa 0 označena *negativna*, a sa 1 *pozitivna* grupa. I u ovom su slučaju zadovoljena sva 4 kriterija Mann-Whitneyevog testa.

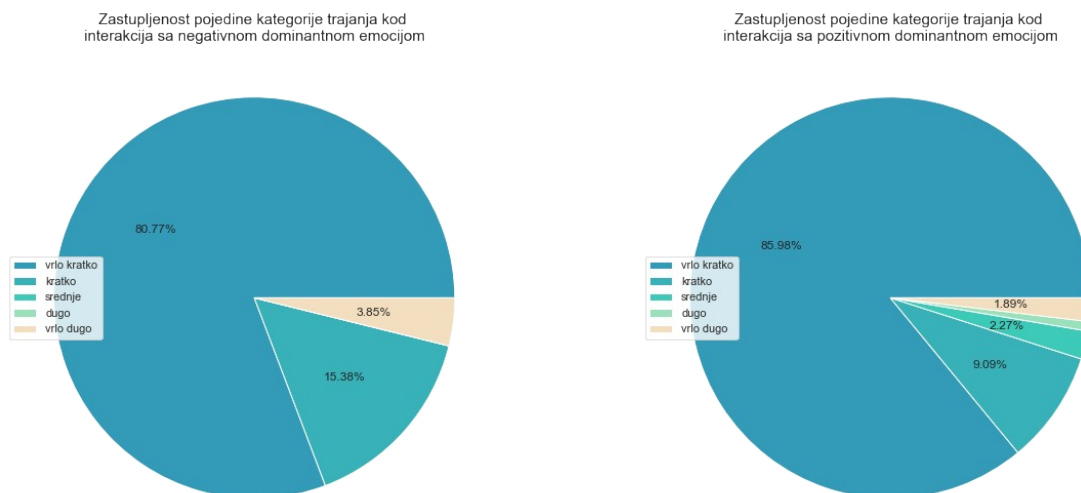
Tablica 15. Deskriptivna statistika (druga hipoteza)

Descriptives				Statistic	Std. Error
trajanje_interakcije	grupa_dominantne_emocije				
,00	Mean			18,5455	3,17277
	95% Confidence Interval for Mean	Lower Bound		12,0827	
		Upper Bound		25,0082	
	5% Trimmed Mean			16,2879	
	Median			12,0000	
	Variance			332,193	
	Std. Deviation			18,22617	
	Minimum			4,00	
	Maximum			82,00	
	Range			78,00	
	Interquartile Range			20,50	
	Skewness			1,874	,409
	Kurtosis			3,836	,798
	1,00	Mean			20,8755
95% Confidence Interval for Mean		Lower Bound		17,3232	
		Upper Bound		24,4278	
5% Trimmed Mean				15,9053	
Median				11,0000	
Variance				836,242	
Std. Deviation				28,91785	
Minimum				4,00	
Maximum				180,00	
Range				176,00	
Interquartile Range				16,00	
Skewness				3,346	,152
Kurtosis				12,557	,303

Sljedeći grafovi prikazuju histograme trajanja pojedine grupe i tablicu podataka izračunatih deskriptivnom statistikom.



Slika 54. Histogrami trajanja interakcija (druga hipoteza)



Slika 55. Zastupljenost kategorija trajanja kod pozitivnih i negativnih prvih izraženih emocija

Kao i kod prve hipoteze, tablica rangova prikazuje iste informacije: rang prosjeka i sumu rangova za dvije testirane skupine (one kod kojih je prva izražena negativna (označena sa ,00) i one kod kojih je ona pozitivna (označeno sa 1,00)). Prosjeci rangova u ovom su slučaju vrlo blizu - 146,35 za grupu sa interakcijama kod kojih je prva izražena emocija pozitivna te 145,39 za grupu sa negativnim prvom izraženom emocijom [Tablica 16].

Tablica statistike provedenog testa prikazuje značajnost testa [Tablica 17]. U ovom slučaju p-vrijednost (*asympt. Sig. (2-tailed)* u tablici) iznosi 0,951 što je veće od početno postavljene vrijednosti značaja koja iznosi 0,05. Na temelju toga se zaključuje da rezultat istraživanja nije statistički značajan, odnosno nulta hipoteza (*Prva izražena emocija nema utjecaja na trajanje interakcije*) ne može biti opovrgnuta.

Tablica 16. Rangovi statističkog testa (druga hipoteza)

Mann-Whitney Test

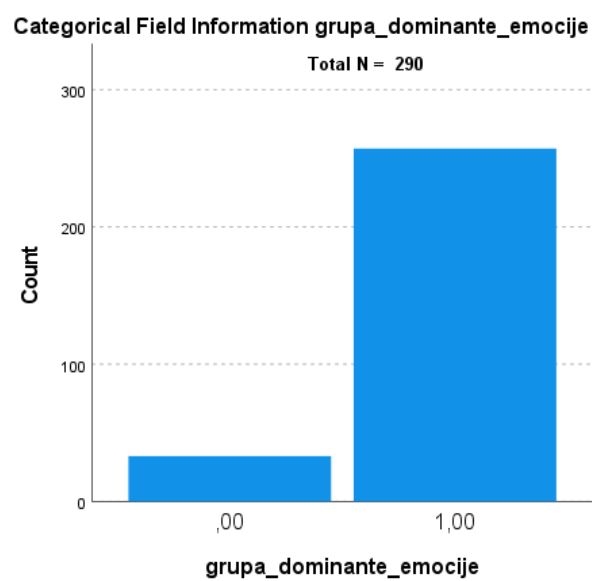
Ranks				
	grupa_dominante_emocije	N	Mean Rank	Sum of Ranks
trajanje_interakcije	,00	33	146,35	4829,50
	1,00	257	145,39	37365,50
	Total	290		

Kao i kod prvog testiranja, razlika u broju interakcije pojedine grupe je velika. Od 290 interakcija, u 257 njih je prva izražena emocija pozitivna, a u 33 njih negativna [Slika 56]. Stoga ponovo postoji mogućnost da bi rezultati testiranja hipoteze bili drugačiji ukoliko bi broj interakcija obaju grupa bio podjednak. Također postoji mogućnost da bi se rezultat provjere hipoteze promijenio ukoliko bi postojao veći broj duljih interakcija.

Tablica 17. Statistika provedenog testa (druga hipoteza)

Test Statistics ^a	
	trajanje_interakcije
Mann-Whitney U	4212,500
Wilcoxon W	37365,500
Z	-,062
Asymp. Sig. (2-tailed)	,951

a. Grouping Variable:
grupa_dominante_emocije



Slika 56. Broj pojedine grupe interakcija (druga hipoteza)

5. SAŽETAK REZULTATA

S ciljem razumijevanja odnosa između ljudi i afektivnog robota proučavale su se i analizirale njihove interakcije. Zabilježene interakcije koje su se proučavale uglavnom su bile vrlo kratke, u prosjeku oko 20 sekundi. U globalu su u interakcijama prevladavale pozitivne emocije nad onim negativnim. U prosjeku, emocije koje su ljudi najviše osjećali tijekom svojih interakcija su bile neutralnost i sreća, a najrjeđe gađenje. Što se tiče dominantnih emocija u interakciji, one su u većini interakcija također bile pozitivne. Isto vrijedi i za prve emocije koje su se razvile tijekom interakcije. Što se tiče veze trajanja interakcije sa emocijama, rezultati testiranja hipoteza rezultati nisu bili statistički značajni. Zbog toga se alternativne hipoteze nisu mogle prihvatiti, tj. hipoteze koje govore da postoji povezanost između trajanja interakcije i dominantne emocije te trajanja interakcije sa emocijom koja je prva izražena nisu podržane rezultatima istraživanja.

Postoje dvije mogućnosti vezane uz ove rezultate koji su statistički neznčajni. Prvo je da su nulte hipoteze stvarno istinite te da nema nikakvo stvarnog učinka emocija na trajanje interakcija. S druge strane, hipoteze istraživanja mogu biti istinite, ali nema dovoljno dokaza kako bi se poduprle. Ova druga mogućnost je pogreška tipa II. U tom se slučaju ne bi trebalo zaključiti da su nulte hipoteze istinite već postoji mogućnost da nema dovoljno dokaza da ih se odbaci. Potencijalni problem veličine uzorka je već spomenut u ranijim poglavljima. Riječ je o podskupinama vrsta emocija. U oba se testiranja hipoteza proučavaju dvije skupine emocija – pozitivne i negativne. Interakcija sa pozitivnim emocijama je puno više nego onih negativnih. Prikladnije bi za testiranje bilo skupiti informacije o podjednakom broju interakcija sa negativnim i pozitivnim emocijama.

Podrška HRZZ

Ovaj rad je sufinancirala Hrvatska zaklada za znanost projektom “AMICORC – Afektivna multimodalna interakcija temeljena na konstruiranoj robotskoj spoznaji” (UIP-2020-02-7184).

LITERATURA

- [1] <https://www.art-ai.io/programme/plea/> (20.2.2023.)
- [2] CURTIS, Elizabeth A.; COMISKEY, Catherine; DEMPSEY, Orla. Importance and use of correlational research. *Nurse researcher*, 2016, 23.6.
- [3] LAMBERT, Vickie A.; LAMBERT, Clinton E. Qualitative descriptive research: An acceptable design. *Pacific Rim international journal of nursing research*, 2012, 16.4: 255-256.
- [4] VEHOVAR, Vasja; TOEPOEL, Vera; STEINMETZ, Stephanie. *Non-probability sampling*. The Sage handbook of survey methods, 2016.
- [5] STRATTON, Samuel J. Population research: convenience sampling strategies. *Prehospital and disaster Medicine*, 2021, 36.4: 373-374.
- [6] MURAIRWA, Stanley. Voluntary sampling design. *International Journal of Advanced Research in Management and Social Sciences*, 2015, 4.2: 185-200.
- [7] <https://www.ai-therapy.com/psychology-statistics/sample-size-calculator> (20.2.2023.)
- [8] DIXON, Wilfrid J.; MASSEY JR, Frank J. Introduction to statistical analysis. 1951.
- [9] <https://www.scribbr.com/category/statistics/>
- [10] Nastavni materijali - Uvod u deskriptivnu i inferencijalnu statistiku, Odjel za psihologiju, Sveučilište u Zadru
<https://www.unizd.hr/Portals/12/Nastavnici/Irena%20Buric/Uvod%20u%20deskriptivnu%20i%20inferencijalnu%20statistiku.pdf> (21.3.2023.)
- [11] MCKNIGHT, Patrick E.; NAJAB, Julius. Mann-Whitney U Test. *The Corsini encyclopedia of psychology*, 2010, 1-1.
- [12] <https://www.statisticshowto.com/mann-whitney-u-test/> (21.3.2023.)
- [13] <https://statistics.laerd.com/spss-tutorials/mann-whitney-u-test-using-spss-statistics.php> (21.3.2023.)
- [14] NEWSON, Roger. Parameters behind "nonparametric" statistics: Kendall's tau, Somers' D and median differences. *The Stata Journal*, 2002, 2.1: 45-64.
- [15] <https://statistics.laerd.com/spss-tutorials/kendalls-tau-b-using-spss-statistics.php> (21.3.2023.)
- [16] Visentin, Denis C et al. "The earnestness of being important: Reporting non-significant statistical results." *Journal of advanced nursing* vol. 76,4 (2020): 917-919.
doi:10.1111/jan.14283

PRILOZI

- I. Python kod
- II. Excel dokument sa podacima

Selekcija podataka - selekcija_podataka_2023.py

```
from openpyxl import Workbook, load_workbook
from openpyxl.utils import get_column_letter

wb = load_workbook('dataset_plea_2023.xlsx', data_only=True)
ws = wb.active

max_row = ws.max_row

#PRVA SELEKCIJA
#određivanje novih interakcija na temelju datuma
for row in range(3, max_row + 1):
    dan_1 = ws['O' + str(row)].value
    dan_2 = ws['O' + str(row+1)].value
    mjesec_1 = ws['P' + str(row)].value
    mjesec_2 = ws['P' + str(row+1)].value

    if mjesec_1 == mjesec_2:
        if dan_1 == dan_2:
            val_1 = ws['N' + str(row)].value
            val_2 = ws['N' + str(row+1)].value
            raz = val_2 - val_1
            if raz >= 3:
                ws['A' + str(row+1)].value = 'new'
            else: ws['A' + str(row+1)].value = 'new'
        else: ws['A' + str(row+1)].value = 'new'
    else: ws['A' + str(row+1)].value = 'new'

#određivanje važećih interakcija na temelju trajanja timestampa
n1 = ws['N2'].value #inicijalna vrijednost
red = 'B2' #inicijalni red
for row in range(3, max_row + 1):
    read_new = ws['A' + str(row)].value
    if read_new == 'new':
        n2 = ws['N' + str(row-1)].value #zadnji timestamp u interakciji
        usp = n2-n1
        n1 = ws['N' + str(row)].value #prvi timestamp u interakciji
        if usp <=3:
            ws[red].value = 'nevazece'
        else: ws[red].value = 'vazece'
        red = 'B' + str(row)

#određivanje broja važećih interakcija
br=1
for row in range(2, max_row + 1):
    read_A = ws['A' + str(row)].value
    read_B = ws['B' + str(row)].value
    if read_A == 'new':
        if read_B != 'nevazece':
            ws['C' + str(row)].value = (br)
            br=br+1
print("broj vežećih interakcija: ", br-1 )
ws['R2'].value = 'broj važećih interakcija (kraćih od 4 s)'
ws['S2'].value = br-1

print('done - selekcija 1')

#DRUGA SELEKCIJA
#označavanje redova kao važećih za lakšu daljnju obradu
for row in range(2, max_row + 1):
    read_A = ws['A' + str(row)].value
    if read_A == 'new':
        ref = 'B' + str(row)
    if read_A != 'new':
        ws['B' + str(row)].value=ws[ref].value

print('done - selekcija 2')

#TREĆA SELEKCIJA
#brisanje nevažećih interakcija
for row in range(max_row - 1, 1, -1):
    read_B = ws['B' + str(row)].value
    if read_B == 'nevazece':
```

```
ws.delete_rows(row)

print('done - selekcija 3')

wb.save('selekcija_podataka_2023.xlsx')
print('done - cijelokupna selekcija')
```

Postavke dokumenta - postavke_usera_2023.py

```
from openpyxl import Workbook, load_workbook

wb = load_workbook('podaci_zao_bradu_2023.xlsx', data_only=True)
ws = wb["dataset_plea"]

#Citiranje broja pojedinog sudionika i stvaranje posebnog data sheeta za svakog
max_row = ws.max_row

for row in range(2, max_row-1):
    read_A = ws['A' + str(row)].value
    if read_A == 'end':
        break
    if read_A != None:
        br_usera = read_A

for k in range(1, br_usera+1):
    wb.create_sheet('user_' + str(k))

#stvaranje izgleda tablice za svakog usera
for z in range(1, br_usera+1):
    sheet = "user_" + str(z)
    ws_new = wb[sheet]
    ws_new['B1'] = 'angry'
    ws_new['C1'] = 'sad'
    ws_new['D1'] = 'neutral'
    ws_new['E1'] = 'surprised'
    ws_new['F1'] = 'fear'
    ws_new['G1'] = 'disgust'
    ws_new['H1'] = 'happy'
    ws_new['I1'] = 'date_and_time'
    ws_new['J1'] = 'time_in_seconds'

print('done - napravljeni sheetovi usera')

#kopiranje podataka iz glavnog data sheeta u sheetove usera
ws = wb["dataset_plea"]
#br_redova = 0
lista_br_redova = []

for row in range(2, max_row+2):
    read_A = ws['A' + str(row)].value

    if read_A != None:
        lista_br_redova.append(br_redova)
        br_redova = 1

    if read_A == None:
        br_redova = br_redova + 1

lista_br_redova.pop(0)

broj_reda_p = []
oznaka_reda = 0
for row in range(1, max_row+2):
    oznaka_reda = oznaka_reda + 1
    read_A = ws['A' + str(row)].value
    if read_A != None:
        broj_reda_p.append(oznaka_reda)

main = wb["dataset_plea"]
col_s = 2
col_e = 11
no_of_users = len(broj_reda_p)
for h in range(0, no_of_users):
```

```

sit = "user_" + str(h+1)
if sit == 'user_' + str(no_of_users):
    break
user = wb[sit]
row_s = (broj_reda_p[h])
row_e = (broj_reda_p[h+1]-1)
for r in range(row_s, row_e + 1):
    for c in range(col_s, col_e + 1):
        user.cell(r,c).value = main.cell(r,c).value
print('done - prepisani podaci')

```

```

##pomicanje redova u sheetovima usera
ws = wb["dataset_plea"]
mr = user.max_row

```

```

no_of_users = len(broj_reda_p)
for h in range(0, no_of_users+1):
    sit = "user_" + str(h+1)
    if sit == 'user_' + str(no_of_users):
        break
    user = wb[sit]

    polje_1 = 'A' + str(broj_reda_p[h])
    polje_2 = 'J' + str(broj_reda_p[h+1]-1)
    polje = polje_1 + ':' + polje_2

    red = -(broj_reda_p[h]-2)
    stupac = 0
    user.move_range(polje, red, stupac)

```

```

print('done - pomkanuti podaci u sheetovima usera')

```

```

wb.save('podaci_za_obradu_definirani_useri_2023.xlsx')
print('done')

```

Obrada podataka - obrada_podataka_2023.py

```

from openpyxl import Workbook, load_workbook
import openpyxl
from openpyxl.utils import get_column_letter
from openpyxl.styles import Font
import matplotlib.pyplot as plt
import numpy as np
import statistics
from scipy import stats
import pandas as pd
import seaborn as sns
from collections import Counter

#definiranje boja grafova
output_dir = "users_graphs"
colors_graph = ['#339AB8', '#38B2B8', '#3DCAB8', '#9BE1BC', '#F3DFBF',
                '#EFB5A8', '#EB8A90']
colors_graph_dark = ['#2C85A0', '#30979C', '#30B09F', '#81D9AA', '#EFD4A9',
                    '#E99886', '#E7737B']
colors_graph_dark2 = ['#1F5D70', '#22696D', '#238074', '#52CB8B', '#E6BB75',
                    '#DF6C53', '#DE3F4A']

#definiranje sheetova u excel dokumentu
wb = load_workbook('podaci_za_obradu_definirani_useri_2023.xlsx', data_only=True)
wb.create_sheet("emotions") #sheet za prikaz obrađenih podataka emocija
wb.create_sheet("graphs_duration") #sheet za prikaz obrađenih podataka trajanja
wb.create_sheet("dominant_emotion") #sheet za prikaz obrađenih podataka dominantnih
emocija
wb.create_sheet("first_emotion") #sheet za prikaz obrađenih podataka prvih izraženih
emocija
wb.create_sheet("mann_whitney_u_test") #sheet za prikaz podataka potrebnih za
provođenje testiranja hipoteza
ws = wb["dataset_plea"]
de = wb["dominant_emotion"]
fe = wb["first_emotion"]

```

```
#ukupni broj sudionika u interakciji
no_of_users = ws['M2'].value + 1

#TRAJANJE INTERAKCIJE
lista_trajanja_all = []
for g in range(0, no_of_users+1):
    sheet = "user_" + str(g+1)
    if sheet == 'user_'+str(no_of_users):
        break
    user = wb[sheet]
    mr = user.max_row

#računanje trajanja interakcije za svakog usera
    read_time_start = user['J2'].value
    read_time_end = user['J'+str(mr)].value
    trajanje = read_time_end-read_time_start

    user['L2'].value = 'trajanje interakcije'
    user['M2'].value = trajanje
    lista_trajanja_all.append(trajanje)

print('done - izračunate vrijednosti trajanja pojedine interakcije')

#histogram trajanja interakcija - Distribucija broja interakcija određenog trajanja
graphs = wb["graphs_duration"]

plt.figure(dpi=120)
plt.figure(figsize=(20, 11))
w=1
n, bins, patches = plt.hist(lista_trajanja_all, bins=np.arange(min(lista_trajanja_all),
max(lista_trajanja_all) + w, w), facecolor='blue', alpha=0.75)
plt.title('Distribucija broja interakcija određenog trajanja', fontsize=22.5)
plt.grid(True)
plt.yticks(fontsize=15)
plt.xticks(np.arange(4, max(lista_trajanja_all)+1, 20.0), fontsize=15)
plt.xlabel('trajanje interakcije [s]', fontsize=20)
plt.ylabel('broj interakcija', fontsize=20)
plt.savefig("broj_interakcija_određenog_trajanja".format(output_dir), transparent=True)
plt.show()

img = openpyxl.drawing.image.Image("broj_interakcija_određenog_trajanja.png")
img.anchor = 'A1'
graphs.add_image(img)

#TRAŽENJE EKSTREMNIH VRIJEDNOSTI (OUTLIERS)
lista_trajanja=[]
outliers=[]

for g in range(0, no_of_users+1):
    sheet = "user_"+str(g+1)
    if sheet == 'user_'+str(no_of_users):
        break
    user = wb[sheet]
    mr = user.max_row

#čitanje vrijednosti trajanja
    read_trajanje = user['M2'].value
    if read_trajanje > 240:
        #print(sheet, "- trajanje:", read_trajanje)
        wb.remove(wb[sheet])
        outliers.append(g+1)
    else: lista_trajanja.append(read_trajanje)

print('done - obrisani outliersi- sheetovi')

#preimenovanje sheetova usera
br=1
for worksheets in wb.sheetnames[1:291]:
    worksheet = wb[worksheets]
    worksheet.title= "user_" + str(br)
    br=br+1

print('done - preimenovanje sheetova')
```

```

#brisanje ekstremnih vrijednosti iz glavnog data sheeta
max_row=ws.max_row
ws = wb["dataset_plea"]
br_redova = 0
lista_br_redova = []

for row in range(2,max_row+2):
    read_A = ws['A' + str(row)].value
    if read_A != None:
        lista_br_redova.append(br_redova)
        br_redova = 1
    if read_A == None:
        br_redova = br_redova + 1
lista_br_redova.pop(0)

broj_reda_p=[]
oznaka_reda=0
for row in range(1,max_row+2):
    oznaka_reda=oznaka_reda+1
    read_A = ws['A' + str(row)].value
    if read_A != None:
        broj_reda_p.append(oznaka_reda)

max_row=ws.max_row

for i in reversed(outliers):
    ws.delete_rows(broj_reda_p[i-1], lista_br_redova[i-1])

n=1
for row in range(2,max_row+2):
    read_A = ws['A' + str(row)].value
    if read_A != None and read_A != 'end':
        ws['A' + str(row)].value=n
        n=n+1

print('done - obrisani outliersi- glavni datasheet')

#popis trajanja interakcija
sort_lista_trajanja_all = sorted(lista_trajanja_all)
broj_puta_all = Counter(sort_lista_trajanja_all)

br=6
graphs['W5'].value = "trajanje [s]"
graphs['X5'].value = "broj interakcija tog trajanja"
for i in list(broj_puta_all):
    graphs['W'+str(br)].value =i
    graphs['X'+str(br)].value = broj_puta_all[i]
    br=br+1

ws['L3'].value = 'konačan broj interakcija'
ws['M3'].value = len(lista_trajanja)

for i in range(70,75):
    graphs['Y'+str(i)].value = 'outlier'

print('done - trajanje interakcija')

#####
ws = wb["dataset_plea"]
no_of_users = ws['M3'].value + 1

##MEAN, MODE I MEDIAN - trajanje
##measures of central tendency
mean_trajanje = statistics.mean(lista_trajanja)
mode_trajanje = statistics.mode(lista_trajanja)
median_trajanje = statistics.median(lista_trajanja)

ws = wb["graphs_duration"]
ws['AB5'].value = 'MODE - najčešća vrijednost trajanja interakcije'
ws['AC5'].value = mode_trajanje
ws['AB6'].value = 'MEDIAN - srednja vrijednost trajanja interakcije'
ws['AC6'].value = median_trajanje
ws['AB7'].value = 'MEAN - prosječna vrijednost trajanja interakcije'
ws['AC7'].value = mean_trajanje

```

```

#measures of variability
ws = wb["graphs_duration"]
ws['AB9'].value = 'MAX trajanje interakcije'
ws['AC9'].value = max(lista_trajanja)
ws['AB10'].value = 'MIN trajanje interakcije'
ws['AC10'].value = min(lista_trajanja)
ws['AB12'].value = 'RANGE - raspon trajanja interakcije'
ws['AC12'].value = max(lista_trajanja) - min(lista_trajanja)
ws['AB13'].value = 'INTERQUARTILE RANGE - interkvartalni raspon trajanja interakcije'
ws['AC13'].value = stats.iqr(lista_trajanja, interpolation = 'midpoint')
ws['AB14'].value = 'STDEV - standardna devijacija'
ws['AC14'].value = statistics.stdev(lista_trajanja)
ws['AB15'].value = 'VARIANCE - odstupanje'
ws['AC15'].value = statistics.variance(lista_trajanja)

zakrivljenost = (3 *(mean_trajanje - median_trajanje))/statistics.stdev(lista_trajanja)

ws['AB17'].value = "Pearsonov koeficijent zakrivljenosti"
ws['AC17'].value = statistics.variance(lista_trajanja)
print('done - trajanje racunica')

#crtanje grafa - Distribucija broja interakcija određenog trajanja bez ekstremnih
vrijednosti
plt.figure(dpi=120)
plt.figure(figsize=(20, 11))
w=1
n, bins, patches = plt.hist(lista_trajanja, bins=np.arange(min(lista_trajanja),
max(lista_trajanja) + w, w), facecolor='#38B2B8', alpha=0.75)
plt.title('Distribucija broja interakcija određenog trajanja bez ekstremnih
vrijednosti', fontsize=22.5)
plt.grid(True)
plt.yticks(np.arange(0, 40, 5.0), fontsize=15)
plt.xlabel('trajanje interakcije [s]', fontsize=20)
plt.ylabel('broj interakcija', fontsize=20)
plt.axvline(mean_trajanje, color = '#22696D', linestyle='dashed')
plt.axvline(median_trajanje, color = '#22696D', linestyle='solid')
plt.legend(labels = ['prosjek', 'medijan'], loc="center right", fontsize = 18)
plt.savefig("histogram_interakcija", transparent=True)
plt.show()

img = openpyxl.drawing.image.Image("histogram_interakcija.png")
img.anchor = 'A39'
graphs.add_image(img)

#grafovi trajanja interakcija - Distribucija broja interakcija određenog trajanja
graphs = wb["graphs_duration"]

#1
fig, ax = plt.subplots()
sns.set(rc={'figure.figsize':(6,12)})
sns.set_style('whitegrid')
ax = sns.violinplot(data=lista_trajanja, color="#D4FFC2", linewidth=0)
ax = sns.swarmplot(data=lista_trajanja, color="#5E9A47", size = 3)
ax.set_ylabel('trajanje interakcije [s]', fontsize=15)
ax.set_ylim(0, max(lista_trajanja)+20)
ax.set_title('Distribucija broja interakcija određenog trajanja', fontsize=15)
plt.yticks(np.arange(0, max(lista_trajanja)+30, 25))
plt.xticks([0], [''])
plt.savefig('Prikaz_broja_interkacija_odredenog_trajanja_1.png', transparent=True)
plt.show()

img = openpyxl.drawing.image.Image("Prikaz_broja_interkacija_odredenog_trajanja_1.png")
img.anchor = 'C80'
graphs.add_image(img)

print('done - racunica trajanja grafovi')

#interquartile range
q3, q1 = np.percentile(lista_trajanja, [75, 25])
iqr = q3 - q1

ws = wb["graphs_duration"]
ws['AE5'].value = 'intekrvartilna razlika'

```

```
ws['AE6'].value = 'Q1'
ws['AF6'].value = q1
ws['AE7'].value = 'Q3'
ws['AF7'].value = q3

#crtanje boxplot grafa - Interkvartilne razlika trajanja interakcija
plt.figure(dpi=120)
plt.figure(figsize=(12, 6))
medianprops = dict( linewidth=2.5, color='#5E9A47')
bp = plt.boxplot(lista_trajanja, showmeans=True, whis=99, vert = False, widths = 0.2,
medianprops=medianprops)
plt.title('Interkvartilne razlika trajanja interakcija', fontsize=11.5)
plt.yticks([])
plt.xticks(np.arange(0, max(lista_trajanja)+1, 15.0), fontsize=10)
plt.xlabel('trajanje interakcije [s]', fontsize=10)
plt.savefig("Interkvartilna_razlika.png", transparent=True)
plt.show()

img = openpyxl.drawing.image.Image("Interkvartilna_razlika.png")
img.anchor = 'AA58'
graphs.add_image(img)

#histogram kategorija - Distribucija interakcija pojedine kategorije trajanja
plt.figure(dpi=120)
plt.figure(figsize=(20, 11))
n, bins, patches = plt.hist(lista_trajanja, bins=[4, 36, 72, 108,
144, 180], facecolor='#38B2B8', alpha=0.75)
plt.title('Distribucija interakcija pojedine kategorije trajanja', fontsize=22.5)
plt.grid(True)
plt.text(10, -20, 'vrlo kratko', fontsize = 17)
plt.text(50, -20, 'kratko', fontsize = 17)
plt.text(85, -20, 'srednje', fontsize = 17)
plt.text(122, -20, 'dugo', fontsize = 17)
plt.text(156, -20, 'vrlo dugo', fontsize = 17)
plt.xticks([4, 36, 72, 108, 144, 180], fontsize=15)
plt.xlabel('kategorija trajanja', fontsize=20, labelpad=50)
plt.ylabel('broj interakcija', fontsize=20)
plt.savefig("histogram_kategorija.png", transparent=True)
plt.show()

img = openpyxl.drawing.image.Image("histogram_kategorija.png")
img.anchor = 'Z18'
graphs.add_image(img)

vrlo_kratko=[]
kratko=[]
srednje=[]
dugo=[]
vrlo_dugo=[]

for i in lista_trajanja:
    if i<36:
        vrlo_kratko.append(i)
    if i>=36 and i<72:
        kratko.append(i)
    if i >=72 and i<108:
        srednje.append(i)
    if i >=108 and i<144:
        dugo.append(i)
    if i>=144 and i<=180:
        vrlo_dugo.append(i)

kategorije_trajanja = [vrlo_kratko, kratko, srednje, dugo, vrlo_dugo]
kategorije_trajanja_ime = ['vrlo kratko', 'kratko', 'srednje', 'dugo', 'vrlo dugo']
br_int_kategorije= []
for i in range(0, len(kategorije_trajanja)):
    br_int_kategorije.append(len(kategorije_trajanja[i]))

ws['AH5'].value = "kategorije trajanja"
ws['AI5'].value = "broj interakcija tog trajanja"
br=6
for i in range(0, len(kategorije_trajanja_ime)):
    ws['AH'+str(br)].value = kategorije_trajanja_ime[i]
```



```

ws['AI'+str(br)].value = br_int_kategorije[i]
br=br+1

#pie chart kategorija- Zastupljenost pojedine kategorije trajanja
br_int_kategorije_format = []
for i in range(0, len( br_int_kategorije)):
    read_br_int_kategorije = br_int_kategorije[i]/290*100
    br_int_kategorije_format_n= "{:.2f}".format(read_br_int_kategorije)+"%"
    br_int_kategorije_format.append(br_int_kategorije_format_n)

plt.figure(dpi=1200)
plt.figure(figsize=(9, 9))
plt.pie(br_int_kategorije, labels = br_int_kategorije_format, colors=colors_graph)
plt.title('Zastupljenost pojedine kategorije trajanja', fontsize = 15)
plt.legend(labels = kategorije_trajanja_ime, loc="center left")
plt.savefig("pie_chart_kategorija.png", transparent=True)
plt.show()

img = openpyxl.drawing.image.Image("pie_chart_kategorija.png")
img.anchor = 'AN57'
graphs.add_image(img)

#####
#####
#EMOCIJE

path = r"C:\Users\DELL\Desktop\work\n_podaci_za_obradu_definirani_useri.xlsx"
df = pd.read_excel(path)

emocije= ['angry', 'sad', 'neutral', 'surprised', 'fear', 'disgust', 'happy']

list_angry = list(df.angry)
list_sad = list(df.sad)
list_neutral = list(df.neutral)
list_surprised = list(df.surprised)
list_fear = list(df.fear)
list_disgust = list(df.disgust)
list_happy = list(df.happy)

sve_emocije = [list_angry, list_sad, list_neutral, list_surprised, list_fear,
list_disgust, list_happy]

for i in range (len(sve_emocije)):
    ready_pop = sve_emocije[i]
    duljina_liste = len(ready_pop)-1
    ready_pop.pop(duljina_liste)

mean_angry = statistics.mean(list_angry)
mean_sad = statistics.mean(list_sad)
mean_neutral = statistics.mean(list_neutral)
mean_surprised = statistics.mean(list_surprised)
mean_fear = statistics.mean(list_fear)
mean_disgust = statistics.mean(list_disgust)
mean_happy = statistics.mean(list_happy)

means = [mean_angry, mean_sad, mean_neutral, mean_surprised,
mean_fear,mean_disgust,mean_happy]

#pie chart zastupljenosti emocija u svim interakcijama - Prosječna zastupljenost
pojedine emocije u svim zabilježenim interakcijama
means_format = []
for i in range(0, len(means)):
    read_mean = means[i]/100
    means_format_n= "{:.2f}".format(read_mean)+"%"
    means_format.append(means_format_n)

plt.figure(dpi=1200)
plt.figure(figsize=(9, 9))
plt.pie(means, labels = means_format, colors=colors_graph)
plt.title('Prosječna zastupljenost pojedine emocije u \nsvim zabilježenim
interakcijama', fontsize = 15)
plt.legend(labels = emocije, loc="center left")
plt.savefig("postotak_em_svih_interakcija", transparent=True)

```

```

plt.show()

emotions_ds = wb["emotions"]
img = openpyxl.drawing.image.Image("postotak_em_svih_interakcija.png")
img.anchor = 'A9'
emotions_ds.add_image(img)

print("done - piechart raspodjele svih emocija")

#prosječne emocije pojedinog usera
emocija = []
b=0

for g in range(0,no_of_users+1):
    sheet = "user_"+str(g+1)
    if sheet == 'user_'+str(no_of_users):
        break
    user = wb[sheet]
    nameCH = "user_"+str(g+1)+"_interakcija.png"
    M_red = 5

    plt.show()
    plt.figure(dpi=120)
    plt.figure(figsize=(20, 9))

    for col in range(2, 9):
        for row in range(2, mr+1):
            char = get_column_letter(col)
            read_cell = user[char+str(row)].value
            if read_cell != None: emocija.append(read_cell)

    mean_emocija = statistics.mean(emocija)
    user['M'+str(M_red)].value = mean_emocija
    M_red = M_red+1
    x= list(range(0,(len(emocija))))

    plt.yticks(np.arange(0, 10000, 500))
    plt.xticks(np.arange(0, max(x)+2, 1))
    plt.grid(True)
    plt.xlabel('vrijeme [s]')
    plt.ylabel('vrijednost emocije')
    plt.plot (x,emocija)

    b = b +1
    plt.legend(labels = emocije, loc = "upper right", bbox_to_anchor=(1.101, 1))
    if b == 7: b=0
    emocija = []
    plt.savefig (nameCH, transparent = True)

    user['L4'].value = 'prosječna vrijednost emocije'
    user['M4'].value = 'tijekom cijele interakcije'
    for i in range(0,len(emocije)):
        L_red = i+5
        user['L'+str(L_red)].value = emocije[i]

print('done - napravljeni grafovi interakcija pojedinih usera')

for g in range(0, no_of_users+1):
    sheet = "user_"+str(g+1)
    if sheet == 'user_'+str(no_of_users):
        break
    user = wb[sheet]
    name = "user_"+str(g+1)+".png"
    name2 = "user_"+str(g+1)+"_interakcija.png"
    user_emo_for_pch = []
    legenda_postoci =[]
    for row in range(5,12):
        read_M = user['M'+str(row)].value/100
        user_emo_for_pch.append(read_M)
        format_read_M = "{:.2f}".format(read_M)+"%"
        legenda_postoci.append(format_read_M)
        if len(user_emo_for_pch) == 7:

            plt.figure(dpi=1200)

```

```

plt.figure(figsize=(9, 9))
plt.pie(user_emo_for_pch, labels = legenda_postoci, colors=colors_graph)
plt.legend(labels = emocije, loc="upper left")
plt.savefig(name.format(output_dir), transparent=True)
plt.show()

img = openpyxl.drawing.image.Image(name)
img.anchor = 'O1'
user.add_image(img)

img = openpyxl.drawing.image.Image(name2)
img.anchor = 'Y3'
user.add_image(img)

print('done - napravljeni pie chartovi emocija pojedinih usera')

#određivanje dominantne emocije pojedinog usera
for g in range(0, no_of_users+1):
    sheet = "user_"+str(g+1)
    if sheet == 'user_'+str(no_of_users):
        break
    user = wb[sheet]

    user['L13'].value = 'dominantna emocija'
    dominantna = 0
    dom_em = 'error'
    for row in range(5,12):
        read_M = user['M'+str(row)].value
        read_emocija = user['L'+str(row)].value
        if read_M > dominantna:
            dominantna = read_M
            dom_em = read_emocija
    user['M13'].value = dom_em

#broj dominantnih emocija kod usera
lista_dom_em = []
for g in range(0, no_of_users+1):
    sheet = "user_"+str(g+1)
    if sheet == 'user_'+str(no_of_users):
        break
    user = wb[sheet]
    read_dom_em = user['M13'].value
    lista_dom_em.append(read_dom_em)

broj_dom_em = []
for i in range(0, len(emocije)):
    count = lista_dom_em.count(emocije[i])
    broj_dom_em.append(count)

#pie chart broja ispitanika sa pojedinom emocijom kao dominantnom
broj_dom_em_format = []
for i in range(0, len(broj_dom_em)):
    read_broj_dom_em = broj_dom_em[i]/290*100
    broj_dom_em_format_n = "{:.2f}".format(read_broj_dom_em)+"%"
    broj_dom_em_format.append(broj_dom_em_format_n)

plt.figure(dpi=1200)
plt.figure(figsize=(9, 9))
plt.pie(broj_dom_em, labels = broj_dom_em_format, colors=colors_graph)
plt.title('Zastupljenost interakcija sa \npojedinom emocijom kao dominantnom', fontsize
= 15)
plt.legend(labels = emocije, loc="upper left")
plt.savefig("broj_usera_sa_pojedinom_em_kao_dom", transparent=True)
plt.show()

img = openpyxl.drawing.image.Image("broj_usera_sa_pojedinom_em_kao_dom.png")
img.anchor = 'A1'
de.add_image(img)

#graf trajanja interakcija u ovisnosti sa dominantnom emocijom
plt.figure(dpi=1200)
plt.figure(figsize=(7, 12))
plt.plot(lista_dom_em, lista_trajanja, 'm_', markersize=20)
plt.title('Trajanje interakcije u ovisnosti sa dominantnom emocijom', fontsize = 15)

```

```

plt.yticks(np.arange(0, max(lista_trajanja)+1, 10.0))
plt.grid(True)
plt.xlabel('dominantna emocija u interakciji', fontsize= 13)
plt.ylabel('trajanje interakcije [s]', fontsize= 13)
plt.savefig("Trajanje_interakcije_u_ovisnosti_sa_dom_em", transparent=True)
plt.show()

img = openpyxl.drawing.image.Image("Trajanje_interakcije_u_ovisnosti_sa_dom_em.png")
img.anchor = 'B29'
de.add_image(img)

print('done - grafovi vezani uz dominantne emocije')

#histogram vrijednosti pojedine emocije
emocije= ['angry', 'sad', 'neutral', 'surprised', 'fear', 'disgust', 'happy']
ws = wb["dataset_plea"]
mr = ws.max_row

#violin graf vrijednosti pojedine emocije
for i in range (len(sve_emocije)):
    fig, ax = plt.subplots()
    sns.set(rc={'figure.figsize':(6,12)})
    sns.set_style('whitegrid')
    ax= sns.violinplot(data=sve_emocije[i],color= colors_graph[i], linewidth=0)
    ax = sns.stripplot(data=sve_emocije[i],color= colors_graph_dark2[i], size = 2)
    ax.set_ylabel( "Vrijednost emocije", fontsize=15)
    ax.set_ylim(0, 10000)
    ax.set_title('Raspodijela vrijednosti emocije '+emocije[i]+' \nu svim zabilježenim
interakcijama', fontsize=17)
    plt.yticks(np.arange(0, 11000, 1000))
    plt.xticks([0], [emocije[i]])
    plt.savefig('Vrijednosti_emocije_'+emocije[i]+'_u_svim_interakcijama',
transparent=True)
    plt.show()

#crtanje boxplot grafa
for i in range (len(sve_emocije)):
    plt.figure(dpi=120)
    plt.figure(figsize=(24, 8))
    medianprops = dict( linewidth=2.5, color=colors_graph_dark2[i])
    meanpointprops = dict(marker='o', markeredgecolor=colors_graph_dark2[i],
markerfacecolor=colors_graph_dark[i])
    bp = plt.boxplot(sve_emocije[i],showmeans=True, whis=99 , vert = False , widths =
0.2, medianprops=medianprops,meanprops=meanpointprops)
    plt.title('Interkvartilna razlika vrijednosti emocije '+emocije[i], fontsize=25)
    plt.yticks([])
    plt.xticks(np.arange(0,10550, 750.0), fontsize=20)
    plt.xlabel('vrijednost emocije '+emocije[i], fontsize=25)
    plt.savefig("Interkvartilna_razlika_emocije_"+emocije[i]+".png", transparent=True)
    plt.show()

    q3, q1 = np.percentile(sve_emocije[i], [75 ,25])
    iqr = q3 - q1
    print(emocije[i], 'q1', q1, 'q3', q3)

print("done - crtanje grafova pojedine emocije")

##MEAN, MODE I MEDIAN - emocije
ws = wb["emotions"]
for i in range (0, len(emocije)):
    B_red = 3+i
    ws['B'+str(B_red)] = emocije[i]

##MODE
ws = wb["dataset_plea"]
mr = ws.max_row
vrijednosti = []
mode_py = []
n=0

for col in range (2, 9):
    #print('nova emocija')
    for row in range(2, mr+1):
        char = get_column_letter(col)

```

```

        read_cell = ws[char+str(row)].value
        if read_cell != None: vrijednosti.append(read_cell)

    mode_py.append(statistics.mode(vrijednosti))
    vrijednosti= []

ws = wb["emotions"]
ws['C2']='MODE - najčešća vrijednost trajanja interakcije'
for i in range (0, len(emocije)):
    C_red = 3+i
    ws['C'+str(C_red)] = mode_py[i]

##MEDIAN
ws = wb["dataset_plea"]
mr = ws.max_row
vrijednosti = []
median_py =[]
n=0

for col in range (2, 9):
    #print('nova emocija')
    for row in range(2, mr+1):
        char = get_column_letter(col)
        read_cell = ws[char+str(row)].value
        if read_cell != None: vrijednosti.append(read_cell)

    median_py.append(statistics.median(vrijednosti))
    vrijednosti= []

ws = wb["emotions"]
ws['D2']='MEDIAN - srednja vrijednost pojedine emocije'
for i in range (0, len(emocije)):
    D_red = 3+i
    ws['D'+str(D_red)] = median_py[i]

##MEAN
ws = wb["dataset_plea"]
mr = ws.max_row
vrijednosti = []
mean_py =[]
n=0

for col in range (2, 9):
    #print('nova emocija')
    for row in range(2, mr+1):
        char = get_column_letter(col)
        read_cell = ws[char+str(row)].value
        if read_cell != None: vrijednosti.append(read_cell)

    mean_py.append(statistics.mean(vrijednosti))
    vrijednosti= []

ws = wb["emotions"]
ws['E2']='MEAN - prosječna vrijednost pojedine emocije'
for i in range (0, len(emocije)):
    E_red = 3+i
    ws['E'+str(E_red)] = mean_py[i]

for i in range (len(sve_emocije)):
    plt.figure(figsize=(9, 9))
    num_bins =30
    n, bins, patches = plt.hist(sve_emocije[i], num_bins, facecolor=colors_graph_dark[i],
alpha=0.75, rwidth=2)
    plt.title('Distribucija vrijednosti emocije '+emocije[i], fontsize=17)
    plt.grid(True)
    plt.axvline(mean_py[i], color = colors_graph_dark2[i], linestyle='dashed')
    plt.axvline(median_py[i], color = colors_graph_dark2[i], linestyle='solid')
    plt.xticks(np.arange(0, 11000, 1000), fontsize=10)
    plt.xlabel('vrijednost emocije', fontsize=15)
    plt.ylabel('broj pojavljivanja', fontsize=15)
    plt.legend(labels = ['prosjeak', 'medijan'], loc="center right")
    plt.savefig('Distribucija_vrijednosti_'+emocije[i]+'_emocije', transparent=True)
    plt.show()

```

```

##measures of central tendency
ws = wb["emotions"]

red=39
for i in range (len(sve_emocije)):
    ws['B'+str(red)].value = emocije[i]
    ws['B'+str(red)].font = Font(bold=True)

    red=red+2
    ws['B'+str(red)].value = 'MODE - najčešća vrijednost'
    ws['C'+str(red)].value = statistics.mode(sve_emocije[i])
    ws['E'+str(red)].value = 'MAX vrijednost'
    ws['F'+str(red)].value = max(sve_emocije[i])
    ws['H'+str(red)].value = 'RANGE - raspon vrijednosti'
    ws['I'+str(red)].value = max(sve_emocije[i])-min(sve_emocije[i])
    ws['K'+str(red)].value = 'STDEV - standardna devijacija'
    ws['L'+str(red)].value = statistics.stdev(sve_emocije[i])
    ws['N'+str(red)].value = 'broj usera sa dominantnom emocijom'
    ws['O'+str(red)].value = broj_dom_em[i]

    red=red+1
    ws['B'+str(red)].value = 'MEDIAN - srednja vrijednost'
    ws['C'+str(red)].value = statistics.median(sve_emocije[i])
    ws['E'+str(red)].value = 'MIN vrijednost'
    ws['F'+str(red)].value = min(sve_emocije[i])
    ws['H'+str(red)].value = 'INTERQUARTILE RANGE - interkvartalni raspon'
    ws['I'+str(red)].value = stats.iqr(sve_emocije[i], interpolation = 'midpoint')
    ws['K'+str(red)].value = 'VARIANCE - odstupanje'
    ws['L'+str(red)].value = statistics.variance(sve_emocije[i])

    red=red+1
    ws['B'+str(red)].value = 'MEAN - prosječna vrijednost'
    ws['C'+str(red)].value = statistics.mean(sve_emocije[i])

    red=red+38

#broj usera s dom em
ws = wb["emotions"]
ws['F2']='broj usera sa dominantnom emocijom'
for i in range (0, len(broj_dom_em)):
    F_red = 3+i
    ws['F'+str(F_red)] = broj_dom_em[i]

#usporedba trajanja i dom em
vrlo_kratko_dom=[]
kratko_dom=[]
srednje_dom=[]
dugo_dom=[]
vrlo_dugo_dom=[]

for g in range(0, no_of_users+1):
    sheet = "user_"+str(g+1)
    if sheet == 'user_'+str(no_of_users):
        break
    user = wb[sheet]

    read_M2 = user['M2'].value
    read_M13 = user['M13'].value
    if read_M2<36:
        user['N2'].value = kategorije_trajanja_ime[0]
        vrlo_kratko_dom.append(read_M13)

    if read_M2>=36 and read_M2<72:
        user['N2'].value = kategorije_trajanja_ime[1]
        kratko_dom.append(read_M13)

    if read_M2 >=72 and read_M2<108:
        user['N2'].value = kategorije_trajanja_ime[2]
        srednje_dom.append(read_M13)

    if read_M2 >=108 and read_M2<144:
        user['N2'].value = kategorije_trajanja_ime[3]
        dugod_dom.append(read_M13)

```

```

        if read_M2>=144 and read_M2<=180:
            user['N2'].value = kategorije_trajanja_ime[4]
            vrlo_dugo_dom.append(read_M13)

br_vrlo_kratko_dom = []
for i in range(0,len(emocije)):
    count = vrlo_kratko_dom.count(emocije[i])
    br_vrlo_kratko_dom.append(count)

br_kratko_dom = []
for i in range(0,len(emocije)):
    count = kratko_dom.count(emocije[i])
    br_kratko_dom.append(count)

br_srednje_dom = []
for i in range(0,len(emocije)):
    count = srednje_dom.count(emocije[i])
    br_srednje_dom.append(count)

br_dugo_dom = []
for i in range(0,len(emocije)):
    count = dugo_dom.count(emocije[i])
    br_dugo_dom.append(count)

br_vrlo_dugo_dom = []
for i in range(0,len(emocije)):
    count = vrlo_dugo_dom.count(emocije[i])
    br_vrlo_dugo_dom.append(count)

#pie chart - Zastupljenost dominantnih emocije u određenoj kategoriji trajanja
lista_kat_trajanja_dom_em= [br_vrlo_kratko_dom, br_kratko_dom,
br_srednje_dom,br_dugo_dom,br_vrlo_dugo_dom]

def autopct_more_than_1(pct):
    return ('%.2f%%' % pct) if pct >= 1 else ''

for i in range (len(kategorije_trajanja_ime)):
    plt.figure(dpi=1200)
    plt.figure(figsize=(9, 9))
    plt.pie(lista_kat_trajanja_dom_em[i], colors=colors_graph,
autopct=autopct_more_than_1) #labels = emocije
    plt.title('Zastupljenost dominantnih emocije \nu kategoriji trajanja
'+kategorije_trajanja_ime[i]+'', fontsize = 15)
    plt.legend(labels = emocije, loc="center left")
    plt.savefig('Zastupljenost_dominantne_emocije_kategorije_'+kategorije_trajanja_ime[i]
+'.png', transparent=True)
    plt.show()

#####
angry_kat_tr=[]
sad_kat_tr=[]
neutral_kat_tr=[]
surprised_kat_tr=[]
fear_kat_tr=[]
disgust_kat_tr=[]
happy_kat_tr=[]
lista_em_kat_tr=[angry_kat_tr,sad_kat_tr,neutral_kat_tr,surprised_kat_tr,fear_kat_tr,dis
gust_kat_tr,happy_kat_tr]

for i in range(0,len(emocije)):
    for g in range(0, no_of_users+1):
        sheet ="user_"+str(g+1)
        if sheet == 'user_'+str(no_of_users):
            break
        user = wb[sheet]

        read_M13 =user['M13'].value
        read_N2 =user['N2'].value
        if str(read_M13) == emocije[i]:
            lista_em_kat_tr[i].append(read_N2)

emocija_trajanje_pie = []

for k in range(0,len(emocije)):

```

```

for i in range(0, len(kategorije_trajanja_ime)):
    count = lista_em_kat_tr[k].count(kategorije_trajanja_ime[i])
    emocija_trajanje_pie.append(count)

plt.figure(dpi=1200)
plt.figure(figsize=(9, 9))
plt.pie(emocija_trajanje_pie, colors=colors_graph, autopct=autopct_more_than_1)
plt.title("Zastupljenost pojedine kategorije trajanja kod \ninterakcija sa
dominantnom emocijom "+emocije[k], fontsize = 15)
plt.legend(labels = kategorije_trajanja_ime, loc="center left")
plt.savefig("Zastupljenost_pojedine_kategorije_trajanja_kod_emocije_"+emocije[k]
+'.png', transparent=True)
plt.show()

emocija_trajanje_pie = []

hipoteza_imena=["pozitivnom", "negativnom"]
pozitivne_emocije_kat_tr = neutral_kat_tr + surprised_kat_tr + happy_kat_tr
negativne_emocije_kat_tr = angry_kat_tr + sad_kat_tr + fear_kat_tr + disgust_kat_tr
hipoteza_1=[pozitivne_emocije_kat_tr, negativne_emocije_kat_tr]

hipoteza_1_pie = []
for k in range(0, len(hipoteza_1)):
    for i in range(0, len(kategorije_trajanja_ime)):
        count = hipoteza_1[k].count(kategorije_trajanja_ime[i])
        hipoteza_1_pie.append(count)

plt.figure(dpi=1200)
plt.figure(figsize=(9, 9))
plt.pie(hipoteza_1_pie, colors=colors_graph, autopct=autopct_more_than_1,
wedgeprops={"edgecolor": "white", 'linewidth': 1, 'antialiased': True})
plt.title("Zastupljenost pojedine kategorije trajanja kod \ninterakcija sa
"+hipoteza_imena[k]+" dominantnom emocijom", fontsize = 15)
plt.legend(labels = kategorije_trajanja_ime, loc="center left")
plt.savefig("Zastupljenost_pojedine_kategorije_trajanja_kod_"+hipoteza_imena[k]
+'.emocija.png', transparent=True)
plt.show()
hipoteza_1_pie = []

emocija_trajanje_hist=[]
red=3
for i in range(0, len(emocije)):
    for g in range(0, no_of_users+1):
        sheet = "user_"+str(g+1)
        if sheet == 'user_'+str(no_of_users):
            break
        user = wb[sheet]

        read_M13 =user['M13'].value
        read_M2 =user['M2'].value
        if str(read_M13) == emocije[i]:
            emocija_trajanje_hist.append(read_M2)

plt.figure(figsize=(9, 9))

num_bins =10
n, bins, patches = plt.hist(emocija_trajanje_hist, bins=[4, 36, 72, 108, 144,180],
facecolor=colors_graph_dark[i], alpha=0.75, rwidth=2)

plt.axvline(statistics.mean(emocija_trajanje_hist), color = colors_graph_dark2[i],
linestyle='dashed')
plt.axvline(statistics.median(emocija_trajanje_hist), color = colors_graph_dark2[i],
linestyle='solid')

print('broj interakcija', len(emocija_trajanje_hist))
de['AG'+str(red)].value = len(emocija_trajanje_hist)
red=red+1
print('mean', statistics.mean(emocija_trajanje_hist))
de['AG'+str(red)].value = statistics.mean(emocija_trajanje_hist)
red=red+1
print('median', statistics.median(emocija_trajanje_hist))
de['AG'+str(red)].value = statistics.median(emocija_trajanje_hist)

plt.xticks([4, 36, 72, 108, 144,180], fontsize=15)

```



```

plt.yticks(fontsize=15)
plt.title('Distribucija trajanja interakcija sa \nemocijom '+emocije[i]+' kao
dominantnom', fontsize=18)
plt.xlabel('trajanje interakcije [s]', fontsize=17)
plt.ylabel('broj interakcija', fontsize=17)
plt.legend(labels = ['prosjeak', 'medijan'], loc="center right")
plt.savefig('Distribucija_trajanja_dom_em_'+emocije[i]+'.png', transparent=True)
plt.show()
red=red+40

#violin graf vrijednosti pojedine emocije
fig, ax = plt.subplots()
sns.set(rc={'figure.figsize':(6,12)})
sns.set_style('whitegrid')
ax= sns.violinplot(data=emocija_trajanje_hist,color= colors_graph[i], linewidth=0)
ax = sns.swarmplot(data=emocija_trajanje_hist,color= colors_graph_dark2[i], size = 2)
ax.set_ylabel( "trajanje interakcije [s]", fontsize=15)
ax.set_title('Distribucija trajanja interakcija sa \nemocijom '+emocije[i]+' kao
dominantnom', fontsize=17)
plt.yticks([4, 36, 72, 108, 144,180])
plt.xticks([0], [emocije[i]])
plt.savefig('Violin_plot_trajanja_interakcije_emocije_'+emocije[i]+'.png',
transparent=True)
plt.show()

emocija_trajanje_hist=[]

ws = wb["emotions"]

broj=42
dodatak=42
for i in range(0, len(emocije)):
    img = openpyxl.drawing.image.Image('Distribucija_vrijednosti_'+emocije[i]
+'_emocije.png')
    img.anchor = 'A'+str(broj)
    ws.add_image(img)
    img = openpyxl.drawing.image.Image('Vrijednosti_emocije_'+emocije[i]
+'_u_svim_interakcijama.png')
    img.anchor = 'K'+str(broj-1)
    ws.add_image(img)
    img = openpyxl.drawing.image.Image("Interkvartilna_razlika_emocije_"+emocije[i]
+".png")
    img.anchor = 'O'+str(broj)
    ws.add_image(img)
    broj=broj+dodatak

pozitivne_emocije=['happy','neutral','surprised']
negativne_emocije=['sad','angry','fear','disgust']

emocija_trajanje_poz=[]
emocija_trajanje_neg=[]

for i in range(0, len(emocije)):
    for g in range(0, no_of_users+1):
        sheet = "user_"+str(g+1)
        if sheet == 'user_'+str(no_of_users):
            break
        user = wb[sheet]

        read_M13 =user['M13'].value
        read_M2 =user['M2'].value

        if str(read_M13) == emocije[i]:
            if emocije[i] in pozitivne_emocije:
                emocija_trajanje_poz.append(read_M2)
            if emocije[i] in negativne_emocije:
                emocija_trajanje_neg.append(read_M2)

mw = wb["mann_whitney_u_test"]
mw['A1'].value = 'Podaci za testiranje prve hipoteze'
mw['A2'].value = 'trajanje interakcije'
mw['B2'].value = 'grupa emocija (0 - negativna, 1 - pozitivna)'
br=3
brr=267

```

```

for i in range(0, len(emocija_trajanje_poz)):
    mw['B'+str(br)].value = '1'
    mw['A'+str(br)].value = emocija_trajanje_poz[i]
    br=br+1

for i in range(0, len(emocija_trajanje_neg)):
    mw['B'+str(brr)].value = '0'
    mw['A'+str(brr)].value = emocija_trajanje_poz[i]
    brr=brr+1

#druga hipoteza
emocija = []
b=0
mr = ws.max_row

for g in range(0, no_of_users+1):
    sheet = "user_"+str(g+1)
    if sheet == 'user_'+str(no_of_users):
        break
    user = wb[sheet]
    nameCH = "user_"+str(g+1)+"_interakcija.png"
    N_red = 5
    read_J_s = user['J2'].value
    J_e = read_J_s+2

    plt.show()
    plt.figure(dpi=120)
    plt.figure(figsize=(20, 9))

    for col in range(2, 9):
        for row in range(2, mr+1):
            char = get_column_letter(col)
            read_cell = user[char+str(row)].value
            read_J = user['J'+str(row)].value
            if read_cell != None and read_J <= J_e:
                emocija.append(read_cell)

    mean_emocija = statistics.mean(emocija)
    user['N'+str(N_red)].value = mean_emocija
    N_red = N_red+1
    x= list(range(0, (len(emocija))))

    plt.yticks(np.arange(0, 10000, 500))
    plt.xticks(np.arange(0, max(x)+2, 1))
    plt.grid(True)
    plt.xlabel('vrijeme [s]')
    plt.ylabel('vrijednost emocije')
    plt.plot(x, emocija)

    b = b +1
    plt.legend(labels = emocije, loc="upper right", bbox_to_anchor=(1.101, 1))
    if b == 7: b=0
    emocija = []
    plt.savefig(nameCH, transparent = True)

    user['N4'].value = 'tijekom prve dvije sekunde'
    for i in range(0, len(emocije)):
        L_red = i+5
        user['L'+str(L_red)].value = emocije[i]

print('done - napravljeni grafovi interakcija pojedinih usera')

#određivanje dominantne emocije na početku interakcije
for g in range(0, no_of_users+1):
    sheet = "user_"+str(g+1)
    if sheet == 'user_'+str(no_of_users):
        break
    user = wb[sheet]

    dominantna = 0
    dom_em = 'error'
    for row in range(5, 12):
        read_N = user['N'+str(row)].value
        read_emocija = user['L'+str(row)].value

```

```

        if read_N > dominantna:
            dominantna = read_N
            dom_em = read_emocija
        user['N13'].value = dom_em

#broj dominantnih emocija kod usera
lista_dom_em_2 = []
for g in range(0, no_of_users+1):
    sheet = "user_"+str(g+1)
    if sheet == 'user_'+str(no_of_users):
        break
    user = wb[sheet]

    read_dom_em_2 = user['N13'].value
    lista_dom_em_2.append(read_dom_em_2)

broj_dom_em_2 = []
for i in range(0, len(emocije)):
    count_2 = lista_dom_em_2.count(emocije[i])
    broj_dom_em_2.append(count_2)

#pie chart broja ispitanika sa pojedinom emocijom kao dominantnom
broj_dom_em_2_format = []
for i in range(0, len(broj_dom_em_2)):
    read_broj_dom_em_2 = broj_dom_em_2[i]/290*100
    broj_dom_em_2_format_n= "{:.2f}".format(read_broj_dom_em_2)+"%"
    broj_dom_em_2_format.append(broj_dom_em_2_format_n)

plt.figure(dpi=1200)
plt.figure(figsize=(9, 9))
plt.pie(broj_dom_em_2, labels = broj_dom_em_2_format ,colors=colors_graph)
plt.title('Distribucija interakcija sa pojedinom \nemocijom kao prvom izraženom u
interakciji', fontsize = 15)
plt.legend(labels = emocije, loc="center left")
plt.savefig("broj_usera_sa_pojedinom_em_kao_dom_prve_2_sekunde", transparent=True)
plt.show()

img =
openpyxl.drawing.image.Image("broj_usera_sa_pojedinom_em_kao_dom_prve_2_sekunde.png")
img.anchor = 'A1'
fe.add_image(img)

#graf trajanja interakcija u ovisnosti sa dominantnom emocijom
plt.figure(dpi=1200)
plt.figure(figsize=(7, 12))
plt.plot(lista_dom_em_2, lista_trajanja, 'm_', markersize=20)
plt.title('Trajanje interakcije u ovisnosti sa \nprvom izraženom emocijom', fontsize =
15)
plt.yticks(np.arange(0, max(lista_trajanja)+1, 10.0))
plt.grid(True)
plt.xlabel('prva izražena emocija u interakciji', fontsize= 13)
plt.ylabel('trajanje interakcije [s]', fontsize= 13)
plt.savefig("Trajanje_interakcije_u_ovisnosti_sa_dom_em_prve_2_sekunde",
transparent=True)
plt.show()

img =
openpyxl.drawing.image.Image("Trajanje_interakcije_u_ovisnosti_sa_dom_em_prve_2_sekunde.
png")
img.anchor = 'B29'
fe.add_image(img)

print('done - grafovi vezani uz dominantne emocije')

#broj usera s dom em
ws = wb["emotions"]
ws['G2']='broj usera sa dom em(prve 2 sekunde interakcije)'
for i in range (0, len(broj_dom_em_2)):
    G_red = 3+i
    ws['G'+str(G_red)] = broj_dom_em_2[i]

pozitivne_emocije=['happy','neutral','surprised']
negativne_emocije=['sad','angry','fear','disgust']
emocija_trajanje_poz_2=[]

```

```

emocija_trajanje_neg_2=[]

for i in range(0,len(emocije)): #
    print(emocije[i])
    for g in range(0, no_of_users+1):
        sheet ="user_"+str(g+1)
        if sheet == 'user_'+str(no_of_users):
            break
        user = wb[sheet]
        read_N13 =user['N13'].value
        read_M2 =user['M2'].value
        if str(read_N13) == emocije[i]:
            if emocije[i] in pozitivne_emocije:
                print('poz', sheet)
                emocija_trajanje_poz_2.append(read_M2)

            if emocije[i] in negativne_emocije:
                print('neg', sheet)
                emocija_trajanje_neg_2.append(read_M2)

mw = wb["mann_whitney_u_test"]
mw['G1'].value = 'Podaci za testiranje druge hipoteze'
mw['G2'].value = 'trajanje interakcije'
mw['H2'].value = 'grupa emocija (0 - negativna, 1 - pozitivna)'
br=3
brr=260
for i in range(0,len(emocija_trajanje_poz_2)):
    mw['H'+str(br)].value = '1'
    mw['G'+str(br)].value = emocija_trajanje_poz_2[i]
    br=br+1

for i in range(0,len(emocija_trajanje_neg_2)):
    mw['H'+str(brr)].value = '0'
    mw['G'+str(brr)].value = emocija_trajanje_poz_2[i]
    brr=brr+1

#####DRUGA HIPOTEZA
emocija_trajanje_hist_2=[]

for i in range(0,len(emocije)): #
    #print('EMOCIJAAAAAAAAA', emocije[i])
    for g in range(0, no_of_users+1):
        sheet ="user_"+str(g+1)
        if sheet == 'user_'+str(no_of_users):
            break
        user = wb[sheet]

        read_N13 =user['N13'].value
        read_M2 =user['M2'].value
        #print(read_M13)
        ##print(read_M13)
        if str(read_N13) == emocije[i]:
            #print('nutra')
            emocija_trajanje_hist_2.append(read_M2)

plt.figure(figsize=(9, 9))

num_bins =10
n, bins, patches = plt.hist(emocija_trajanje_hist_2, bins=[4, 36, 72, 108, 144,180],
facecolor=colors_graph_dark[i], alpha=0.75, rwidth=2)

plt.axvline(statistics.mean(emocija_trajanje_hist_2), color = colors_graph_dark2[i],
linestyle='dashed')
plt.axvline(statistics.median(emocija_trajanje_hist_2), color =
colors_graph_dark2[i], linestyle='solid')
print('mean', statistics.mean(emocija_trajanje_hist_2))
print('median', statistics.median(emocija_trajanje_hist_2))
print('broj interakcija', len(emocija_trajanje_hist_2))

plt.xticks([4, 36, 72, 108, 144,180], fontsize=15)
plt.yticks(fontsize=15)
plt.title('Distribucija trajanja interakcija sa emocijom '+emocije[i]+' \nkao prvom
izrazenom emocijom', fontsize=18)
plt.xlabel('trajanje emocije [s]', fontsize=17)

```

```

plt.ylabel('broj puta pojavljivanja emocije', fontsize=17)
plt.legend(labels = ['prosjeak', 'medijan'], loc="center right")
plt.grid(True)
plt.savefig('Distribucija_trajanja_'+emocije[i]+'_kao_prva_izrazena.png',
transparent=True)
plt.show()

#violin graf vrijednosti pojedine emocije
fig, ax = plt.subplots()
sns.set(rc={'figure.figsize':(6,12)})
sns.set_style('whitegrid')
ax= sns.violinplot(data=emocija_trajanje_hist_2,color= colors_graph[i], linewidth=0)
ax = sns.swarmplot(data=emocija_trajanje_hist_2,color= colors_graph_dark2[i], size =
2)
ax.set_ylabel( "trajanje interakcije [s]", fontsize=15)
ax.set_title('Distribucija trajanja interakcija sa emocijom '+emocije[i]+' \nkao
prvom izrazenom emocijom', fontsize=17)
plt.yticks([4, 36, 72, 108, 144,180])
plt.xticks([0], [emocije[i]])
plt.savefig('Violin_plot_trajanja_sa_'+emocije[i]+'_kao_prvom_izr.png',
transparent=True)
plt.show()

emocija_trajanje_hist_2=[]

#####
angry_kat_tr_2=[]
sad_kat_tr_2=[]
neutral_kat_tr_2=[]
surprised_kat_tr_2=[]
fear_kat_tr_2=[]
disgust_kat_tr_2=[]
happy_kat_tr_2=[]
lista_em_kat_tr_2=[angry_kat_tr_2,sad_kat_tr_2,neutral_kat_tr_2,surprised_kat_tr_2,fear_
kat_tr_2,disgust_kat_tr_2,happy_kat_tr_2]

for i in range(0,len(emocije)):
    for g in range(0, no_of_users+1):
        sheet ="user_"+str(g+1)
        if sheet == 'user_'+str(no_of_users):
            break
        user = wb[sheet]

        read_N13 =user['N13'].value
        read_N2 =user['N2'].value

        if str(read_N13) == emocije[i]:
            lista_em_kat_tr_2[i].append(read_N2)

emocija_trajanje_pie_2 = []

for k in range(0,len(emocije)):
    for i in range(0,len(kategorije_trajanja_ime)):
        count = lista_em_kat_tr_2[k].count(kategorije_trajanja_ime[i])
        emocija_trajanje_pie_2.append(count)

plt.figure(dpi=1200)
plt.figure(figsize=(9, 9))
plt.pie(emocija_trajanje_pie_2, colors=colors_graph, autopct=autopct_more_than_1)
plt.title("Zastupljenost pojedine kategorije trajanja kod interakcija \nu kojima je
emocija "+emocije[k]+" prva izražena emocija", fontsize = 15)
plt.legend(labels = kategorije_trajanja_ime, loc="center left")
plt.savefig("Zastupljenost_pojedine_kategorije_trajanja_kod_emocije_"+emocije[k]
+'_2_hipo.png', transparent=True)
plt.show()

emocija_trajanje_pie_2 = []

hipoteza_imena_2=["pozitivna","negativna"]
pozitivne_emocije_kat_tr_2 = neutral_kat_tr_2 + surprised_kat_tr_2 + happy_kat_tr_2
negativne_emocije_kat_tr_2 = angry_kat_tr_2 + sad_kat_tr_2 + fear_kat_tr_2 +
disgust_kat_tr_2
hipoteza_2=[pozitivne_emocije_kat_tr_2,negativne_emocije_kat_tr_2]

```

```

hipoteza_2_pie = []

for k in range(0, len(hipoteza_2)):
    for i in range(0, len(kategorije_trajanja_ime)):
        count = hipoteza_2[k].count(kategorije_trajanja_ime[i])
        hipoteza_2_pie.append(count)

    plt.figure(dpi=1200)
    plt.figure(figsize=(9, 9))
    plt.pie(hipoteza_2_pie, colors=colors_graph, autopct=autopct_more_than_1,
    wedgeprops={"edgecolor": "white", 'linewidth': 1, 'antialiased': True})
    plt.title("Zastupljenost pojedine kategorije trajanja kod \ninterakcija u kojima je
    prva izražena emocija "+hipoteza_imen_2[k], fontsize = 15)
    plt.legend(labels = kategorije_trajanja_ime, loc="center left")
    plt.savefig("Zastupljenost_pojedine_kategorije_trajanja_kod_"+hipoteza_imen_2[k]
    +'_emocija_2_hipo.png', transparent=True)
    plt.show()
    hipoteza_2_pie = []

#usporedba trajanja i prve izražene emocije
vrlo_kratko_dom_2=[]
kratko_dom_2=[]
srednje_dom_2=[]
dugo_dom_2=[]
vrlo_dugo_dom_2=[]

for g in range(0, no_of_users+1):
    sheet = "user_"+str(g+1)
    if sheet == 'user_'+str(no_of_users):
        break
    user = wb[sheet]

    read_M2 = user['M2'].value
    read_M13 = user['N13'].value
    if read_M2<36:
        user['N2'].value = kategorije_trajanja_ime[0]
        vrlo_kratko_dom_2.append(read_M13)

    if read_M2>=36 and read_M2<72:
        user['N2'].value = kategorije_trajanja_ime[1]
        kratko_dom_2.append(read_M13)

    if read_M2 >=72 and read_M2<108:
        user['N2'].value = kategorije_trajanja_ime[2]
        srednje_dom_2.append(read_M13)

    if read_M2 >=108 and read_M2<144:
        user['N2'].value = kategorije_trajanja_ime[3]
        duغو_dom_2.append(read_M13)

    if read_M2>=144 and read_M2<=180:
        user['N2'].value = kategorije_trajanja_ime[4]
        vrlo_dugo_dom_2.append(read_M13)

br_vrlo_kratko_dom_2 = []
br_kratko_dom_2 = []
br_srednje_dom_2 = []
br_dugo_dom_2 = []
br_vrlo_dugo_dom_2 = []

for i in range(0, len(emocije)):
    count = vrlo_kratko_dom_2.count(emocije[i])
    br_vrlo_kratko_dom_2.append(count)

for i in range(0, len(emocije)):
    count = kratko_dom_2.count(emocije[i])
    br_kratko_dom_2.append(count)

for i in range(0, len(emocije)):
    count = srednje_dom_2.count(emocije[i])
    br_srednje_dom_2.append(count)

for i in range(0, len(emocije)):
    count = duغو_dom_2.count(emocije[i])

```

```

    br_dugo_dom_2.append(count)

for i in range(0, len(emocije)):
    count = vrlo_dugo_dom_2.count(emocije[i])
    br_vrlo_dugo_dom_2.append(count)

#pie chart - Zastupljenost prvih izraženih emocije u pojedinoj kategoriji
lista_kat_trajanja_dom_em_2= [br_vrlo_kratko_dom_2, br_kratko_dom_2,
br_srednje_dom_2, br_dugo_dom_2, br_vrlo_dugo_dom_2]

def autopct_more_than_1(pct):
    return ('%.2f%%' % pct) if pct >= 1 else ''

for i in range (len(kategorije_trajanja_ime)):
    plt.figure(dpi=1200)
    plt.figure(figsize=(9, 9))
    plt.pie(lista_kat_trajanja_dom_em_2[i], colors=colors_graph,
    autopct=autopct_more_than_1) #labels = emocije
    plt.title('Zastupljenost prvih izraženih emocije \nu kategoriji trajanja
    '+'kategorije_trajanja_ime[i]+'"', fontsize = 15)
    plt.legend(labels = emocije, loc="center left")
    plt.savefig('Zastupljenost_dominantne_emocije_kategorije_'+kategorije_trajanja_ime[i]
    +'_2_hipo.png', transparent=True)
    plt.show()

hipoteze_sheets=['dominant_emotion', 'first_emotion']
for i in range(0, len(hipoteze_sheets)):
    ws = wb[hipoteze_sheets[i]]
    red=1
    for i in range (len(sve_emocije)):
        ws['AF'+str(red)].value = emocije[i]
        ws['AF'+str(red)].font = Font(bold=True)
        red=red+2
        ws['AF'+str(red)].value = 'broj interakcija'
        ws['AG'+str(red)].value = statistics.mode(sve_emocije[i])#
        red=red+1
        ws['AF'+str(red)].value = 'MEAN - prosječna vrijednost'
        ws['AG'+str(red)].value = statistics.median(sve_emocije[i])
        red=red+1
        ws['AF'+str(red)].value = 'MEDIAN - srednja vrijednost'
        ws['AG'+str(red)].value = statistics.mean(sve_emocije[i])
        red=red+38

    broj=30
    dodatak=25
    for k in range(0, len(emocije)):
        img =
        openpyxl.drawing.image.Image("Zastupljenost_pojedine_kategorije_trajanja_kod_emocije_"+e
        mocije[k]+'_2_hipo.png')
        img.anchor = 'J'+str(broj)
        de.add_image(img)
        img =
        openpyxl.drawing.image.Image("Zastupljenost_pojedine_kategorije_trajanja_kod_emocije_"+e
        mocije[k]+'_2_hipo.png')
        img.anchor = 'J'+str(broj)
        fe.add_image(img)
        broj=broj+dodatak

    broj=29
    dodatak=25
    for i in range(0, len(kategorije_trajanja_ime)):
        img =
        openpyxl.drawing.image.Image('Zastupljenost_dominantne_emocije_kategorije_'+kategorije_t
        rajanja_ime[i]+'_2_hipo.png')
        img.anchor = 'T'+str(broj)
        de.add_image(img)
        img =
        openpyxl.drawing.image.Image('Zastupljenost_dominantne_emocije_kategorije_'+kategorije_t
        rajanja_ime[i]+'_2_hipo.png')
        img.anchor = 'T'+str(broj+1)
        fe.add_image(img)
        broj=broj+dodatak

    broj=6

```

```

dodatak=42
for i in range(0, len(emocije)):
    img = openpyxl.drawing.image.Image('Distribucija_trajanja_dom_em_'+emocije[i]+'.png')
    img.anchor = 'AE'+str(broj)
    de.add_image(img)
    img =
openpyxl.drawing.image.Image('Violin_plot_trajanja_interakcije_emocije_'+emocije[i]
+'.png')
    img.anchor = 'A0'+str(broj-1)
    de.add_image(img)
    img = openpyxl.drawing.image.Image('Distribucija_trajanja_'+emocije[i]
+'.kao_prva_izrazena.png')
    img.anchor = 'AE'+str(broj-1)
    fe.add_image(img)
    img = openpyxl.drawing.image.Image('Violin_plot_trajanja_sa_'+emocije[i]
+'.kao_prvom_izr.png')
    img.anchor = 'A0'+str(broj-2)
    fe.add_image(img)
    broj=broj+dodatak

img =
openpyxl.drawing.image.Image("Zastupljenost_pojedine_kategorije_trajanja_kod_pozitivnom_
emocija.png")
img.anchor = 'J1'
de.add_image(img)
img =
openpyxl.drawing.image.Image("Zastupljenost_pojedine_kategorije_trajanja_kod_negativnom_
emocija.png")
img.anchor = 'T1'
de.add_image(img)

img =
openpyxl.drawing.image.Image("Zastupljenost_pojedine_kategorije_trajanja_kod_pozitivna_e
mocija_2_hipo.png")
img.anchor = 'J1'
fe.add_image(img)
img =
openpyxl.drawing.image.Image("Zastupljenost_pojedine_kategorije_trajanja_kod_negativna_e
mocija_2_hipo.png")
img.anchor = 'T1'
fe.add_image(img)

#spearman test
for g in range(0,no_of_users+1): #
    sheet ="user_"+str(g+1)
    if sheet == 'user_'+str(no_of_users):
        break
    user = wb[sheet]
    nameCH = "user_"+str(g+1)+"_interakcija.png"
    print(nameCH)

    poz = user['M7'].value+user['M8'].value+user['M11'].value
    neg = user['M5'].value+user['M6'].value+user['M9'].value+user['M10'].value

    user['L15'].value = 'prosječan udio '
    user['L16'].value = 'prosječan udio pozitivnih emocija tijekom interakcije'
    user['L17'].value = 'prosječan udio negativnih emocija tijekom interakcije'
    user['M16'].value = poz/10000*100
    user['M17'].value = neg/10000*100

    pie_podaci=[poz, neg]
    plt.pie(pie_podaci, colors=colors_graph, autopct=autopct_more_than_1) #labels =
emocije
    plt.show()

    emocija_p = []
    emocija_n = []

    udio_poz_em=[]
    udio_neg_em=[]
    for g in range(0,no_of_users+1):
        sheet ="user_"+str(g+1)
        if sheet == 'user_'+str(no_of_users):
            break

```



```
user = wb[sheet]
read_M16 = user['M16'].value
udio_poz_em.append(read_M16)
read_M17 = user['M16'].value
udio_neg_em.append(read_M17)

plt.figure(dpi=1200)
plt.figure(figsize=(12, 9))
plt.title('Ovisnost duljine trajanja interakcije \no udjelu pozitivnih emocija ',
fontSize=18)
plt.xlabel('trajanje interakcije [s]', fontsize=17)
plt.ylabel('prosječan udio pozitivnih emocija', fontsize=17)
plt.plot(lista_trajanja, udio_poz_em, 'o', markersize=2)
plt.savefig('Ovisnost_duljine_o_udio_poz_em.png', transparent=True)
plt.show()

img = openpyxl.drawing.image.Image("Ovisnost_duljine_o_udio_poz_em.png")
img.anchor = 'AU1'
de.add_image(img)

mw = wb["mann_whitney_u_test"]
mw['M1'].value = 'Podaci za provođenje Spearmanovog testa'
mw['M2'].value = 'trajanje interakcije'
mw['N2'].value = 'postotak prosječne pozitivne emocije'
br=3

for i in range(0, len(lista_trajanja)):
    mw['M'+str(br)].value = lista_trajanja[i]
    mw['N'+str(br)].value = udio_poz_em[i]
    br=br+1

wb.save("rezultati_obrade_2023.xlsx")
print('final_done')
```