



Univerzitet Singidunum

Departman za posleddiplomske studije

Master studijski program

Savremene informacione tehnologije

Master rad

Alati za statističku obradu podataka

- Primena IBM SPSS programa -

Mentor:

Prof.dr. Aleksandar Jevremović

Student:

Nikola Draganović

(410258/2014)

Beograd, 2017 god.

SADRŽAJ

Uvod.....	- 4 -
Predmet istraživanja.....	- 5 -
Ciljevi istraživanja	- 5 -
Struktura rada.....	- 5 -
Hipoteze istraživanja.....	- 5 -
1. Alati za obradu statističkih podataka	- 6 -
1.1. MATLAB.....	- 7 -
1.1.1. Nastanak MATLAB-a.....	- 7 -
1.1.2. Primena i okruženje MATLAB-a	- 7 -
1.1.3. Radno okruženje MATLAB-a.....	- 8 -
1.1.4. Toolbox-ovi MATLAB-a.....	- 9 -
1.1.5. Prednosti i mane MATLAB-a.....	- 9 -
1.2. Programiranje u R-u.....	- 10 -
1.2.1. Nastanak R-a	- 10 -
1.2.2. Editori R koda	- 10 -
1.2.3. Okruženje programskog jezika R.....	- 11 -
1.2.4. Prednosti i mane R programa.....	- 11 -
1.3. IBM SPSS	- 12 -
1.3.1. Nastanak SPSS paketa	- 12 -
1.3.2. Primena i okruženje SPSS paketa	- 12 -
1.3.3. Analitički proces	- 15 -
1.3.5. Prednosti i mane SPSS-a.....	- 17 -
2. Inteligentni poslovni sistemi	- 18 -
3. Primena SPSS-a	- 19 -
3.1. Editor podataka - Data Editor	- 20 -
3.2. Prozor izlaznih rezultata - Viewer	- 21 -
3.3. Sintaksni editor - Syntax Editor	- 22 -
3.4. Padajući meniji u programu SPSS	- 23 -
3.4.1. File meni	- 23 -
3.4.2. Edit meni	- 24 -
3.4.3. View meni	- 24 -
3.4.4. Data meni	- 25 -

3.4.5.	Transform meni.....	- 25 -
3.4.6.	Analyze meni	- 26 -
3.4.7.	Direct Marketing meni	- 26 -
3.4.8.	Graphs meni	- 27 -
3.4.9.	Utilities meni.....	- 28 -
3.4.10.	Add-ons meni	- 28 -
3.4.11.	Window i Help meniji.....	- 29 -
3.5.	Paleta alatki i statusna traka u SPSS programu.....	- 30 -
4.	Analiza podataka.....	- 31 -
4.1.	Deskriptivna statistika.....	- 33 -
4.2.	Deskriptivne statističke mere	- 34 -
4.3.	Testovi značajnosti razlika.....	- 37 -
Zaključak		- 41 -
Literatura.....		- 42 -

Uvod

Statistika je danas alat koji je neizostavan u velikom broju naučnih istaživanja, a i u mnogim drugim oblastima gde je velika potreba za analizom i obradom većeg broja podataka.

Marketing stručnjaci istražuju tržište, sociolozi i psiholozi proučavaju ljudsko ponašanje, finansijski analitičari se bave vremenskim serijama, medicinari prate pacijente, biolozi analiziraju DNK, menadžeri primenjuju kontrolu kvaliteta, javna uprava koristi informacije za upravljanje naseljima i regionima, itd. Ono što je zajedničko za sve njih jeste da se u svojoj profesiji moraju oslanjati na statistiku. U današnje vreme moderne tehnologije omogućavaju svima da relativno lako prikupe i manipulišu velikim brojem podataka.¹

U statistici je najbitnije prikupiti što preciznije i kvalitetne podatke, da bi se mogli kasnije isto tako kvalitetno analizirati i izvući što precizniji rezultati i informacije radi donošenja adekvatnih odluka. Jer ukoliko podaci ne budu prikupljeni na pravi način i sa profesionalnošću obrađeni, dobijeni rezultati mogu dovesti do pogrešnih odluka i zaključaka.

Statistika je postala jako bitna u poslednjih tridesetak godina kako tehnologija napreduje, pojavljuju se nova i poboljšavaju stara programska rešenja za analizu podataka. Tako su nastali paketi kao što su R, QlikSense, Minitab, MATLAB, SPSS i drugi.

U radu je prikazana primena statističkog paketa SPSS. Opisan je rad i osnovne funkcije softvera i sama primena u statističkoj analizi kroz određeni set podataka.

Set podataka preuzet je sa sajta Centra za istraživanje po imenu Pew, sa sedištem u Vašingtonu. Podaci u ovom izveštaju su prikupljeni od strane GfK grupe, na engleskom i španskom jeziku, na nacionalno reprezentativnom uzorku od 1060 tinejdžera, uzrasta od 13 do 17 godina i roditelja ili staratelja. Podaci su prikupljeni u periodu od 25. septembra do 9. oktobra 2014. godine i 10. februara do 16. marta, 2015.godine. U prikupljanju podataka na jesen, 1016 parova (roditelj-tinejdžer) su intervjuisani. Istraživanje je ponovo otvoreno u proleće i 44 para su dodata u uzorak.

Napomena: Set podataka koji je korišćen za istraživanje i analizu je preuzet i obrađivan u originalnom formatu na engleskom jeziku.

¹ http://stat.uns.ac.rs/LLLprogramme/NS/Katalog_kurseva.pdf, stranica posećena 20. marta 2015.

Predmet istraživanja

Predmet istraživanja su statistički paketi i njihovo korišćenje u analizi. U radu su teorijski opisani statistički alati Matlab, R i SPSS, a i praktično je prikazan rad i određene statističke analize u SPSS programskom paketu kroz određeni set podataka.

Ciljevi istraživanja

Cilj istraživanja je da se prikaže primena statističkih alata, kao i njihove mogućnosti, nedostaci i prednosti u obradi podataka.

Naučni cilj je definisanje kombinacija deskriptivnih analiza, zajedno sa neparametarskim i parametarskim funkcijama, koje se mogu prihvatiti kao osnova za istraživanje sprovedeno radi iznošenja detaljnih i osnovanih zaključaka.

Struktura rada

Prvi deo rada posvećen je statistici i statističkim alatima za analizu, njihov nastanak i razvijanje tokom godina.

Drugi deo rada je vezan za okruženje i korišćenje statističkog programa SPSS, opis alata, funkcija i analiza koje se koriste za obradu statističkih podataka.

Treći deo se odnosi na primenu funkcija i analiza uz pomoć SPSS programskog alata i uz to su analizirani podaci o korišćenju interneta i društvenih mreža kod tinejdžera.

Četvrti deo se odnosi na tumačenje rezultata i izvođenje zaključaka iz istih.

Hipoteze istraživanja

Glavna hipoteza ovog istraživanja ima za cilj da pokaže da su alati za statističku obradu podataka pogodni za istraživanje i donošenje zaključaka na osnovu deskriptivnih i grafičkih prikaza, koji nisu na raspolaganju prilikom korišćenja nekih drugih programa za obradu.

Posebne hipoteze ovog istraživanja su:

- Izbor adekvatnog alata za statističku obradu utiče na rezultat istraživanja
- Duža i konkretnija analiza pruža više rezultata koji detaljnije objašnjavaju problem
- Primena različitih funkcija dovodi do različitih zaključaka koji zajedno utiču na celokupno istraživanje

1. Alati za obradu statističkih podataka

Primena programskih paketa danas znatno olakšava statističku obradu i smanjuje mogućnost greške na minimum ukoliko se kvalitetno rukuje sa njima. Danas postoji mnoštvo paketa za statističku obradu podataka. Postoje programi i paketi koji su jednostavniji za korišćenje i primena tih alata je namenjena široj populaciji, a postoje i oni koji su namenjeni profesionalnom bavljenju statistikom.

Neki od alata za obradu statističkih podataka su: SPSS, R, MATLAB, STATS, STATISTICA, QlikSense, Minitab i sl.

R i Matlab sadrže impresivnu kolekciju alata u svojim bibliotekama, koja je svakim danom sve veća. Čak i ako se želi specifični model koji nije deo standardnog paketa, možete implementirati model sami, jer R i Matlab su programski jezici sa relativno jednostavnom sintaksom. Što se tiče savremenih primenjenih statističkih alata, R biblioteka je nešto bogatija od Matlab biblioteke. Sa druge strane, MATLAB ima mnogo bolju grafiku.

Na drugoj strani spektra je paket poput SPSS programa. SPSS je prilično uzak u svojim mogućnostima i omogućava vam da uradite samo oko polovinu glavne statistike. Ipak, SPSS je veoma popularan među korisnicima, jer ne zahteva gotovo nikakav trening i program poseduje širok spektar funkcija koje su lake za korišćenje, a opet veoma korisne u analizi. SPSS je veoma detaljan program, koji pruža odlične vizuelne rezultate. Program sadrži sve glavne testove i dijagnostičke alate u vezi sa postupkom i omogućava da se napiše informativni statistički deo empirijske analize. SPSS se često koristi za standardne zahteve klijenata, kao što su linearne regresije, deskriptivna statistika, ANOVA ili analiza glavnih komponenti.

1.1. MATLAB

1.1.1. Nastanak MATLAB-a

Cleve Moler, predsednik odeljenja informatike na Univerzitetu u Novom Meksiku, počeo je da razvija MATLAB u kasnim 1970-im. On ga je dizajnirao da bi omogućio svojim studentima pristup LINPACK i EISPACK, jer bi bez njih morali da uče Fortran. MATLAB se vrlo brzo proširio i na druge univerzitete gde je pronašao veliku publiku u okviru primenjene matematičke zajednice. Džek Litl, inženjer, upoznao se sa programom tokom Molerove posete Stanford Univerzitetu, 1983.godine. Prepoznavši komercijalni potencijal programa, pridružio se Moleru i Stiv Bangertu. Oni su prekodirali MATLAB u C i osnovali Mathworks 1984.godine, koji je nastavio njegov razvoj. Ove prekodirane biblioteke su poznate kao JACKPAC. U 2000-oj godini, MATLAB je ponovo prekodiran kako bi koristio noviji skup biblioteke za manipulaciju matrice, LAPACK.²

MATLAB je prvi usvojen od strane istraživača i praktičara u kontrolnom inženjerstvu, što je bila uža specijalnost Jack Litla, ali se brzo proširila i na mnoge druge domene. Sada se takođe koristi i u obrazovanju, posebno u učenju linearne algebre i numeričke analize.

U 2004-oj godini, MATLAB je imao oko milion korisnika širom industrije i akademskih zajednica. Korisnici dolaze iz različitih sredina inženjerstva, nauka i ekonomije. Koristi se u naučno istraživačkim ustanovama, kao i industrijama.

1.1.2. Primena i okruženje MATLAB-a

MATLAB je jezik visokih performansi i interaktivnog je okruženja za numeričku upotrebu, vizuelizaciju i programiranje. Koristeći ga mogu se analizirati podaci, razvijati algoritmi, stvoriti modeli i aplikacije. Jezik, alati i ugrađene matematičke funkcije nam omogućavaju da istražimo više pristupe i postignemo rešenja brže nego sa tabelama ili tradicionalnim programskim jezicima, kao što su C/C++, Java®, Fortran i Python.

MATLAB možemo koristiti za niz aplikacija, uključujući obradu signala i komunikacije, slike i video obradu, upravljačkih sistema, testovima i merenjima, kompjuterskih finansija i računarsku biologiju.

Više od milion inženjera i naučnika u industriji i akademskim krugovima koriste MATLAB.

² Preuzeto i prevedeno sa, <http://en.wikipedia.org/wiki/MATLAB>

1.1.3. Radno okruženje MATLAB-a

Funkcije koje se mogu obraditi MATLAB-om:

- **Jezičke osnove:** Sintaksa, operatori, tipovi podataka, niz indeksiranja i manipulacija.
- **Matematika:** Linearna algebra, osnovna statistika, diferencijacija i integrali, fourier transformacije i druga matematika.
- **Grafika:** Dvo i trodimenzionalni dijagrami, slike, animacije, vizualizacije.
- **Programiranje skripte i funkcije:** Programski fajlovi, kontrola protoka, uređivanje, ispravljanje grešaka.
- **Uvoz i izvoz podataka:** Tekstualni fajlovi, tabele i drugi formati podataka.
- **GUI Building:** Razvoj aplikacija pomoću uputstva i Callbacks-a.
- **Razvoj naprednih softvera:** Objektno orijentisano programiranje, kod performanse, jedinica za testiranje, spoljni interfejsi Java, C/C ++, .NET i drugi jezici.
- **Desktop okruženje:** Podešavanja, različite platforme.
- **Podržani hardver:** Podrška za treće strane hardvera, kao što su web kamere, Raspberri Pi TM i Arduino® hardver.

➤ Operativni sistemi koji podržavaju MATLAB:

- Microsoft Windows
- Mac OS X
- Linux

Poslednja verzija koja se koristi je MATLAB 9.2.

1.1.4. Toolbox-ovi MATLAB-a

Toolbox-ovi su pomoćni alati koji su razvijeni za razne potrebe korisnika u rešavanju problema, i oni obezbeđuju više mogućnosti MATLAB-a u određenim obalastima.

Postoje pomoćni alati (*toolbox*) za oblasti:

- Baza podataka
- Statistike
- Sistema za obradu zvuka
- Modela predviđanja
- Globalne optimizacije
- Robotičkih sistema
- Finansija
- Trgovanja i drugi

Postojeći pomoćni alati se usavršavaju, a nastaju i novi samim usavršavanjem programa.

1.1.5. Prednosti i mane MATLAB-a

MATLAB se konstantno usavršava i proširuje svoje primene. Uz alate koje sadrži ga čini vrlo pogodnim za analize u različitim okruženjima i teško da postoji kvalitetniji programski alat. Dosta se koristi u savremenoj edukaciji. Manom ga čini to što je opširan i spor, pa zbog toga nije praktičan za korišćenje bez duže obuke i posvećenosti rada u njemu.



Slika 1.1 - Logo i naziv Matlab-a³

³ Slika preuzeta sa sajta, <https://engineering.usu.edu/students/matlab>

1.2. Programiranje u R-u

1.2.1. Nastanak R-a

R program je bio razvijen u Bell laboratorijima od strane Ricka Beckera, Johna Chambersa i Allana Wilksa. R se može promatrati kao implementacija S jezika, koji takođe predstavlja osnovicu S-PLUS-a.

Program su kreirali Ross Ihaka i Robert Gentleman na Univerzitetu u Oklandu, Novi Zeland, i trenutno se razvio R razvojni glavni štab, gde je John Chambers bio član veća. R je nazvan delimično po inicijalima imena autora R-a, a delom kao igra na ime S jezik je često vozilo izbora za istraživanja u statističkoj metodologiji, i R pruža Open Source put do učešća u toj delatnosti.

Jedna od prednosti R-a je lakoća sa kojom se mogu dizajnirati i proizvesti kvalitetni dijagrami, uključujući matematičke simbole i formule, gde su potrebne. Veliku pažnju preuzimaju osnovna podešavanja i rešenja manjih dizajna u grafici, ali korisnik zadržava potpunu kontrolu.

R je dostupan kao slobodan softver pod uslovima Free Software Foundation GNU General Public License u obliku izvornog koda. Može se prevoditi i izvršavati na najrazličitijim UNIXS platformama i sličnim sistemima (uključujući FreeBSD i Linux), Windows i MacOS.

Veći komercijalni softverski sistemi podržavaju veze ili integracija sa R-om, uključujući: JMP, Mathematica, MATLAB, Spotfire, SPSS, STATISTICA, Platforma Simphoni, SAS i Tableau.

1.2.2. Editori R koda

Nekoliko odličnih editora koda su na raspolaganju da obezbede funkcije kao što su podvlačenje sintakse R koda, auto izjednačavanje koda i komande koje šalju kod ili funkcije u R konzolu.

- Osnovni editor koda od Rguis
- RStudio: Integrated Disk Electronic, bazirana na grafičkom korisničkom interfejsu za R
- Vim-R-Tmux: radno okruženje R-a zasnovano na vim-u i tmux-u
- Emacs (ESS dodati paket)
- gedit i Rgedit
- RKward
- Eclipse
- Tinn-R
- Notepad++ (NppToR) ⁴

Poslednja verzija R programa je 3.4.0.

⁴ <http://science.webhostinggeeks.com/programiranje-u-r>

1.2.3. Okruženje programskog jezika R

R je integrisani paket softverskih objekata za upravljanje podacima, obračuna i grafičkog prikaza.

Sadrži:

- efikasno rukovanje podacima i skladište,
- paket operatera za proračun o nizovima, posebno matricama,
- veliku, koherentnu i integrisanu kolekciju posrednih alata za analizu podataka,
- grafički objekti za analizu podataka i prikaz ili na ekranu ili na papiru, i
- razvijen, jednostavan i efikasan programski jezik koji uključuje kondicionale, petlje, definisane korisničke rekurzivnih funkcije i ulazne i izlazne objekte.

Izraz "okruženje" upotrebljen je da istakne da je R dobro planiran i konzistentan alat, a ne alat koji se postepeno dopunjava sa specifičnim i nefleksibilnim programskim alatima, što je često slučaj kod drugih programa za analizu podataka.⁵

1.2.4. Prednosti i mane R programa

Prednost R programa je to što je besplatan i ima mnoštvo statističkih mogućnosti za obradu podataka. A mana je ta da je težak za učenje, i više je namenjen je programerima, nego statičarima koji se prvi put susreću sa programskim jezikom.



Slika 1.2 - Logo R programa⁶

⁵ <https://www.r-project.org/about.html>

⁶ Preuzeto sa [https://en.wikipedia.org/wiki/R_\(programming_language\)](https://en.wikipedia.org/wiki/R_(programming_language))

1.3. IBM SPSS

1.3.1. Nastanak SPSS paketa

SPSS Inc. je softverska kuća sa sedištem u Čikagu, nazvan kao i vlasnički softver istog imena SPSS. Kompanija je osnovana 1968. godine, kada su Norman Nie, Dale Bent i Hadlai "Tex" Hull razvili i počeli da prodaju SPSS softver. Kao legalna kompanija, društvo je osnovano 1975. godine, i Norman je bio na mestu generalnog direktora od 1975. do 1992.godine. Od 1992. do sticanja SPSS Inc. 2009. od strane IBM-a, Jack Noonan bio je generalni direktor.

U 2008. SPSS Inc. ostvario je prodaju od 302.9 miliona dolara i imao je preko 250.000 klijenata. Pored softvera koji deli ime, SPSS Inc. prodao je širok spektar softvera za istraživanje tržišta, anketnog istraživanja i statističku analizu. Među njima je i SPSS Modeler (ranije poznat kao Clementine ili PASW Modeler) za data mining i mrInterview za CATI i onlajn ankete.

SPSS je sklopio partnerstva sa nekoliko drugih kompanija visokog profila kao što su Oracle Corporation i sada učestvuje u brojnim vladinim programima. Dana 28. jula 2009.godine, IBM je objavio kupovinu SPSS Inc., za 1,2 milijarde dolara u gotovini. Zbog spora oko vlasništva nad imenom "SPSS", između 2009. i 2010. godine, softver je nazvan PASW (Predictive Analytics SoftWare). U januaru 2010. godine, kompanija je postala "SPSS: An IBM Company". Kompletan prenos poslovanja na IBM završen je 1. oktobra 2010.⁷

Sadašnja verzija softvera v.24 (2016) nosi naziv „IBM SPSS Statistics“.

1.3.2. Primena i okruženje SPSS paketa

Statistički paket IBM SPSS Statistic (Statistical Package for the Social Sciences „SPSS“) pripada redu najprimenjivijih programa za statističku analizu. Osim toga što značajno doprinosi rešavanju poslovnih i istraživačkih problema, ujedno i nudi širok spektar mogućnosti za celokupan analitički proces.

Korišćenje statističkih procedura SPSS programa obezbeđuje informacije koje će poslužiti u kasnijem donošenju odluka. Rezultati analize, prezentovani u tabelarnim i grafičkim outputima, doprinose efikasnijem i bržem odlučivanju kroz otkrivanje elementarnih činjenica i obrazaca.

⁷ Preuzeto i prevedeno sa sajta http://en.wikipedia.org/wiki/SPSS_Inc

➤ IBM SPSS moguće je koristiti u različitim oblastima, na primer:

- Istraživanje tržišta i direktni marketing,
- Visokoškolsko obrazovanje,
- Administrativna istraživanja, ljudski resursi i planiranje resursa,
- Oblasti medicinskog, kliničkog i socijalnog istraživanja,
- Planiranje i predviđanje,
- Unapređenje kvaliteta,
- Izveštavanje u ad hoc odlučivanju.

Statistika koja je podržana IBM SPSS softverom:

Korelacija, t-test, analiza varijanse (ANOVA), trend analiza, kombinovana analiza (SPANOVA), analiza kovarijanse (ANCOVA), analiza pouzdanosti, faktorska analiza, višestruka regresija, logistička regresija, multivarijaciona analiza varijanse (MANOVA), neparametarske tehnike, analiza višestrukih odgovora i višestruka dihotomna analiza, višedimenzionalno skaliranje.⁸ Mnogi karakteristike SPSS Statistics su dostupne preko padajućih menija ili se mogu programirati sa vlasničkim 4GL sintaksom jezika. Komanda sintaksa programiranja ima koristi od reproducibilnosti, pojednostavljenje ponovljenih zadataka, i složenije rukovanje manipulacije podacima i analizama. Pored toga, neke složene aplikacije mogu se programirati samo u sintaksi i nisu dostupni preko strukture menija. Takođe se može ubaciti u sintakstički fajl koristeći "paste" dugme koje je prisutno u svakom meniju. Programi mogu da se izvršavaju interaktivno ili bez nadzora, pomoću priloženog Production Job Facility.

Pored toga "makro" jezik može da se koristi za pisanje komandnog jezika potprograma. Python programski jezik kao dodatka može da pristupi informacijama u bazi podataka i dinamički izgradi sintaksu programa. Python dodatak je uveden u SPSS 14, zamenio je manje funkcionalni SAX Basic "skripte" u više svrha, iako SAX Basic ostaje na raspolaganju. Pored toga, Python dodatak omogućava SPSS-u za pokretanje bilo koje statistike u slobodan softverski paketa R. Od verzije 14, pa nadalje, SPSS može biti predvođen od strane Python ili VB.NET programa koristeći priloženi "plug-in". (Od verzije 20., ovi dva programska jezika, kao i mnoge skripte, pripadaju osnovnoj instalaciji paketa.)

⁸ Sheridan J. Coakes, prevod Soldić Aleksić dr Jasna, (2013), SPSS 20: Analiza bez muke, Kompjuter biblioteka, Beograd

SPSS stavlja ograničenja na unutrašnjoj strukturi datoteke, tipove podataka, obradu podataka i podudaranja fajlova, koji zajedno znatno pojednostavljaju programiranje. SPSS datasets imaju dvodimenzionalnu strukturu tabele, gde redovi obično predstavljaju slučajeve (kao što su pojedinci ili domaćinstva), a kolone predstavljaju mere (kao što su godine, pol, ili prihodi domaćinstva). Samo dva tipa podataka su definisani: numerički i tekst (ili "string"). Sva obrada podataka javlja redom od slučaja do slučaja kroz fajl. Fajlovi se mogu upariti jedan-najedan i jedan-na-više, ali ne više-prema-više.

Korisnički interfejs ima dva pogleda koji se mogu menjati klikom na jednu od dve kartice u donjem levom uglu prozora SPSS-a. *Data view* pokazuje unakrsne tabele slučajeva (redova) i promenljive (kolone). Čelije mogu da sadrže samo brojeve ili tekst, ali ne i formule. *Variable View* prikazuje podatke gde svaki red predstavlja promenljivu i pokazuje ime promenljive, promenljive oznake, vrednost oznake, širina štampe, vrstu merenja, i razne druge karakteristike. Čelije se mogu u oba slučaja ručno promeniti, definisati strukturu datoteke i omogućavajući unos podataka bez korišćenja komandne sintakse. Ovo može biti dovoljno za manji set podataka. Veći set podataka poput statističkih istraživanja se češće prave u samom softveru.

SPSS može pokrenuti podatke sa ASCII tekstualnih fajlova (uključujući hijerarhijske fajlove), druge statističke pakete, sa tabelama i bazama podataka. Takođe može pokrenuti eksterne tabele baza podataka preko ODBC i SQL.

Statistički izlaz je u sopstvenom formatu datoteke (*.spv file). Fajl se može izvoziti u tekstu ili Microsoft Word-u, PDF-u, Excel-u i drugim formatima. Alternativno, izlaz se može izvući kao podatak (koristeći komandu OMS), kao tekst, PDF, XLS, HTML, XML, SPSS datasets ili različitim grafičkim formatima slika (JPEG, PNG, BMP i EMF).⁹

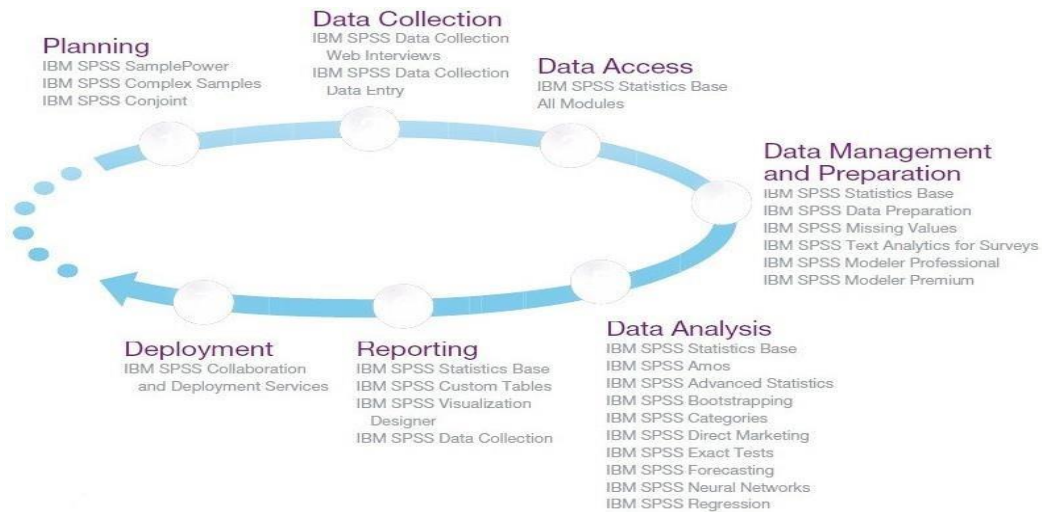
⁹ Preuzeto i prevedeno sa sajta, <http://en.wikipedia.org/wiki/SPSS#Overview>

1.3.3. Analitički proces

Analiza nije jednostavan proces. Potrebno je prikupiti podatke, uraditi analizu, izvući rezultate i portumačiti i predstaviti što razumljivije. Ako se koristi neki analitički paket sa ograničenim mogućnostima, sprovođenje analize bi podrazumevalo korišćenje dodatnih softvera. IBM SPSS nudi sve što je potrebno za proces analize, od planiranja i obrade podataka, do analiziranja izlaznih rezultata.

SPSS sa svojim dodatnim paketima, olakšava rad u svakoj oblasti analitike:

- **Planiranje** (*Planning*) - ušteda vremena i novca,
- **Prikupljanje podataka** (*Data collection*) - efikasnije prikupljanje i unošenje podataka,
- **Pristup podacima** (*Data access*) – brža analiza uz lakši pristup podacima,
- **Priprema i upravljanje podacima** (*Data management and data preparation*) – kroz efikasniju pripremu i upravljanje podacima, brže se stiže i do procesa analize,
- **Analiza podataka** (*Data analysis*) - analiza podataka kroz korišćenje različitih statističkih procedura u cilju dobijanja preciznijih modela,
- **Izveštavanje** (*Reporting*) - prezentovanje rezultata korisnicima,
- **Razvoj** (*Deployment*) - razmena rezultata preko Web ili drugih prezentacija i publikacija.



Slika 1.3 - Proces analize podataka

1.3.4. Svi trenutni IBM SPSS paketi:

- IBM SPSS Statistics Base
- IBM SPSS Advanced Statistics
- IBM SPSS Bootstrapping
- IBM SPSS Categories
- IBM SPSS Complex Samples
- IBM SPSS Conjoint
- IBM SPSS Custom Tables
- IBM SPSS Data Preparation
- IBM SPSS Decision Trees
- IBM SPSS Direct Marketing
- IBM SPSS Exact Tests
- IBM SPSS Forecasting
- IBM SPSS Missing Values
- IBM SPSS Neural Networks
- IBM SPSS Regression
- IBM SPSS Amos
- SPSS Statistics GradPack¹⁰
- IBM SPSS SamplePower
- IBM SPSS Statistics Developer
- IBM SPSS Statistics for Linux on System z
- IBM SPSS Statistics Programmability Extension
- IBM SPSS Statistics Server
- IBM SPSS Text Analytics for Surveys
- IBM SPSS Visualization Designer
- IBM SPSS Statistics Academic Products
- IBM SPSS Statistics Campus
- IBM SPSS Statistics Faculty Pack
- IBM

¹⁰ <http://www-01.ibm.com/software/analytics/spss/products/statistics>

1.3.5. Prednosti i mane SPSS-a

SPSS prednosti:

Sveobuhvatan paket za statističku analizu sa statističkim tehnikama od vrlo jednostavnih do vrlo sofisticiranih, vrlo zahvalno i prijatno grafičko okruženje sa opisnim menijima i jednostavnim padajućim menijima, pokriva većinu rutinske analize, više funkcionalnosti sa dodatnim modulima i dobar pomoćni interfejs.

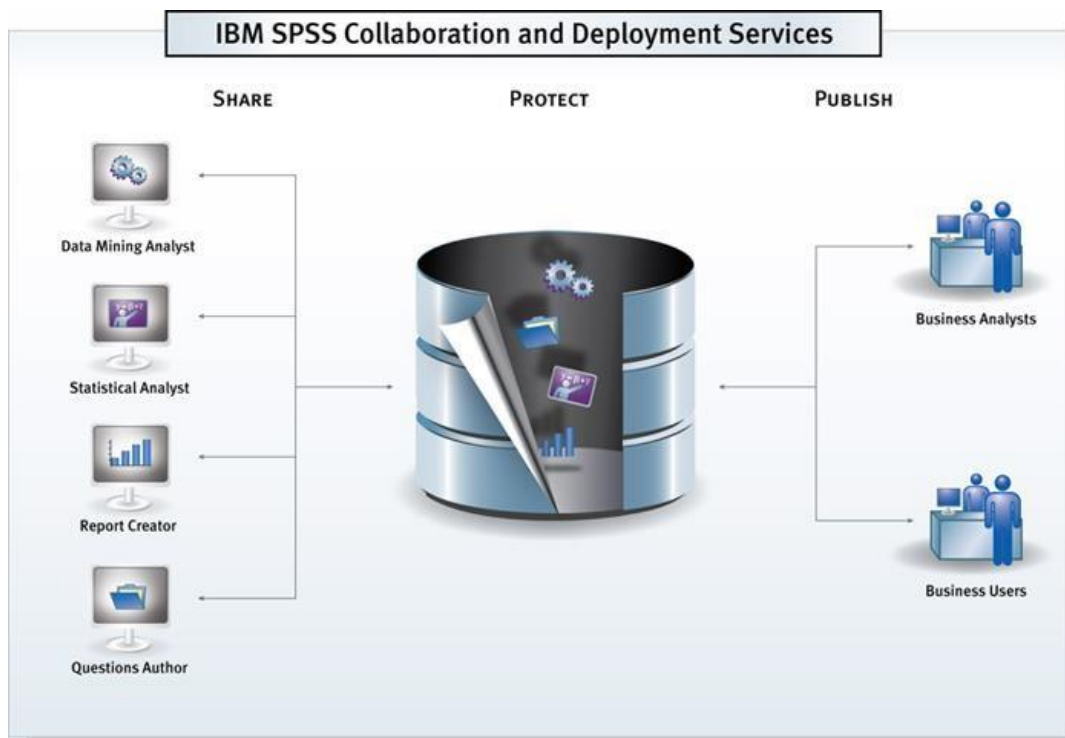
SPSS mane:

Grafikonski prikaz je manje poželjan, neiskusni korisnici greškom mogu da primenjuju pogrešne ili neadekvatne metode, teško se menja izlaz, teškoće pokretanja SPSS preko interneta zbog velikih privremenih podataka koje stvara, potrebno vam je znanje statističke terminologije za upravljanje softverom i neke koje su specifične za SPSS.

2. Inteligentni poslovni sistemi

Inteligentni sistemi su predmet izučavanja računarskih nauka od samih početaka, i može se reći da predstavljaju jednu od veoma značajnih oblasti. Savremena civilizacija ima velike potrebe za inteligentnim sistemima, pre svega zbog velike složenosti raznih sistema kojima treba na efikasan način racionalno upravljati. Razvoj inteligentnih sistema predstavlja jednu od najnovijih “revolucija” posle “industrijske” i “informatičke” revolucije. Industrijska revolucija je oslobodila čoveka od napornog fizičkog rada, informatička revolucija je omogućila automatsku obradu ogromnog broja podataka, velikom brzinom, bez grešaka, prenos podataka, skladištenje, pretraživanje. Inteligentni sistemi idu korak dalje i obavljaju intelektualne aktivnosti vezane za prepoznavanje, klasifikaciju, učenje, rešavanje problema, odlučivanje, upravljanje i slično. Zbog svoje velike složenosti, inteligentni sistemi zahtevaju poznavanje raznih oblasti koje su od značaja za razvoj ovih sistema.

Poslovni inteligentni alati su vrsta softvera koji su dizajnirani da preuzmu, analiziraju i izveštavaju podatke. Ova široka definicija obuhvata sve od tabela, vizuelnih analitika i upita softvera za data mining, skladištenja i donošenja rešenja, SPSS možemo svrstati u jedan takav alat, tj sistem.

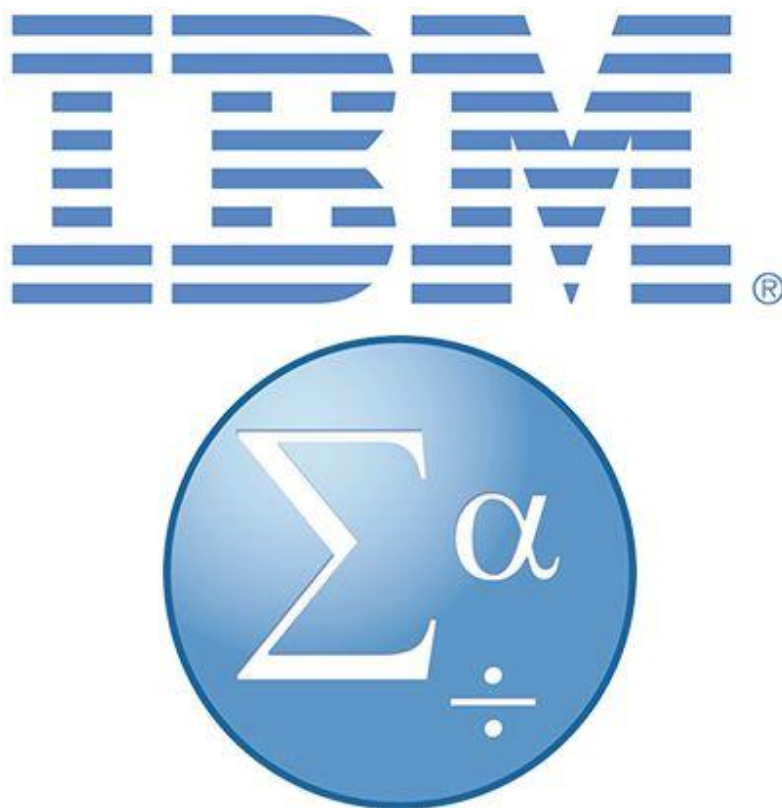


Slika 2.1 - Karakteristike SPSS-a kao inteligentnog sistema

3. Primena SPSS-a

SPSS koristi više vrsta prozora, ili ekrana, u svojim osnovnim operacijama. Svaki prozor je povezan sa specifičnim zadacima i vrstama SPSS datoteka. Prozori uključuju Editor podataka (*Data Editor*), Prozor izlaznih rezultata (*Viewer*), Sintaksni editor (*Syntax Editor*), Editor izvedenih tabela (*Pivot Table Editor*), Editor dijagrama (*Chart Editor*) i Editor tekstualnog izlaza (*Text Output Editor*).¹¹

Najbitniji prozori za korišćenje SPSS alata su *Data Editor*, *Viewer* i *Syntax Editor*, jer ova tri prozora su osnova za većinu akcija izvedenih u program. Ostali prozori su relevantni za više specifičnih tipova zadataka.



Slika 3.1 - Logo SPSS programa

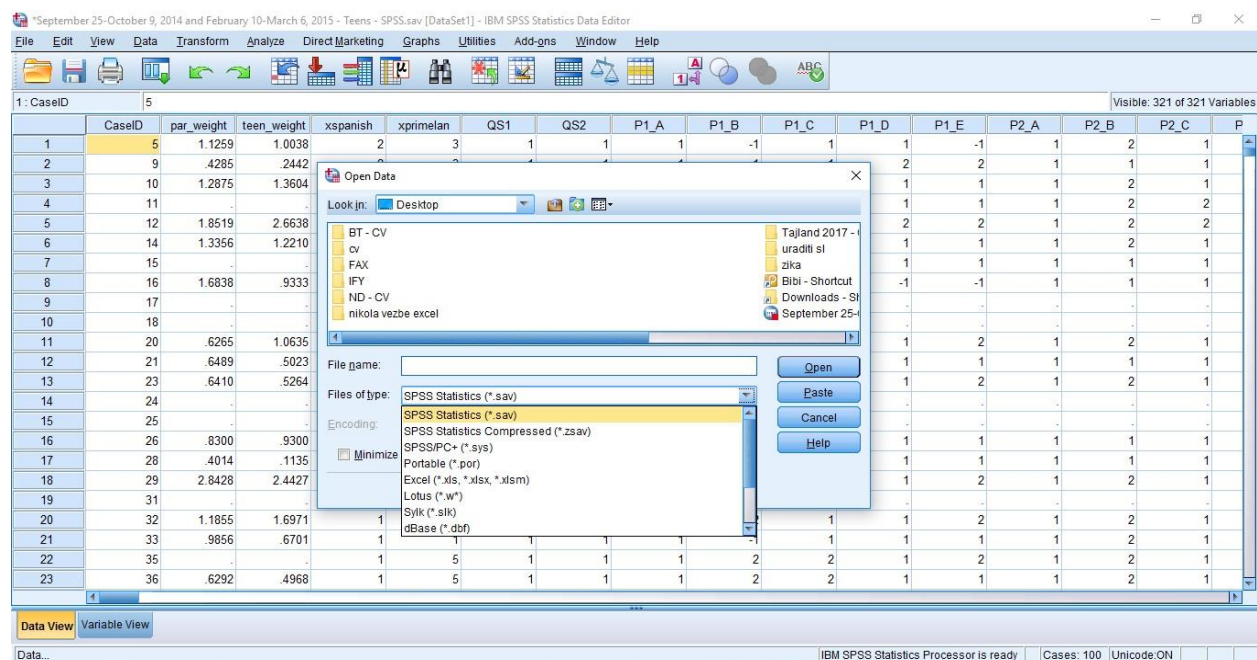
¹¹ Preuzeto i prevedeno sa sajta <http://libguides.library.kent.edu/SPSS/Environment>

3.1. Editor podataka - Data Editor

Prozor *Data Editor* se otvara kada se SPSS pokrene. Ovaj prozor prikazuje sadržaj otvorenih datoteka podataka i pruža izbor padajućih menija koji omogućuju da izmenite i analizirate podatke. Podaci su prikazani u obliku pravougaone tabele, gde kolone predstavljaju promenljive, a redovi opservacije. Format tabele sadrži dve kartice na dnu s oznakom pregled podataka (*Data View*) i pregled varijabli (*Variable View*).

Data View kartica prikazuje otvoreni skup podataka, u svakom preseku reda i kolone ćelija sadrži podatak za jednu promenljivu i jednu opservaciju. Ova kartica služi za unošenje i manipulaciju podataka. Podaci se mogu unositi ručno ili ubacivanje eksternih fajlova koji se prethodno koriguje da bi se lakše i preciznije manipuliralo podacima nakon ubacivanja u SPSS. Formati za unos su:

- SPSS Statistics (*.sav)
- SPSS Statistics Compressed (*.zsav)
- SPSS/PC+ (*.sys)
- Systat (*.syd, *.sys)
- Portable (*.por)
- Excel (*.xls, *.xlsx, *.xslm, *.xlsb)
- Lotus (*.w*)
- Sylk (*.slk)
- SAS (*.sas7bdat, *.sd7, *.sd2, *.ssd01, *.ssd04, *.xpt)
- Stata (*.dta)
- Text (*.txt, *.dat, *.csv)
- dBASE Files (*.dbf, *.ndx, *.mdx)
- MS Access Database (*.mdb, *.accdb)

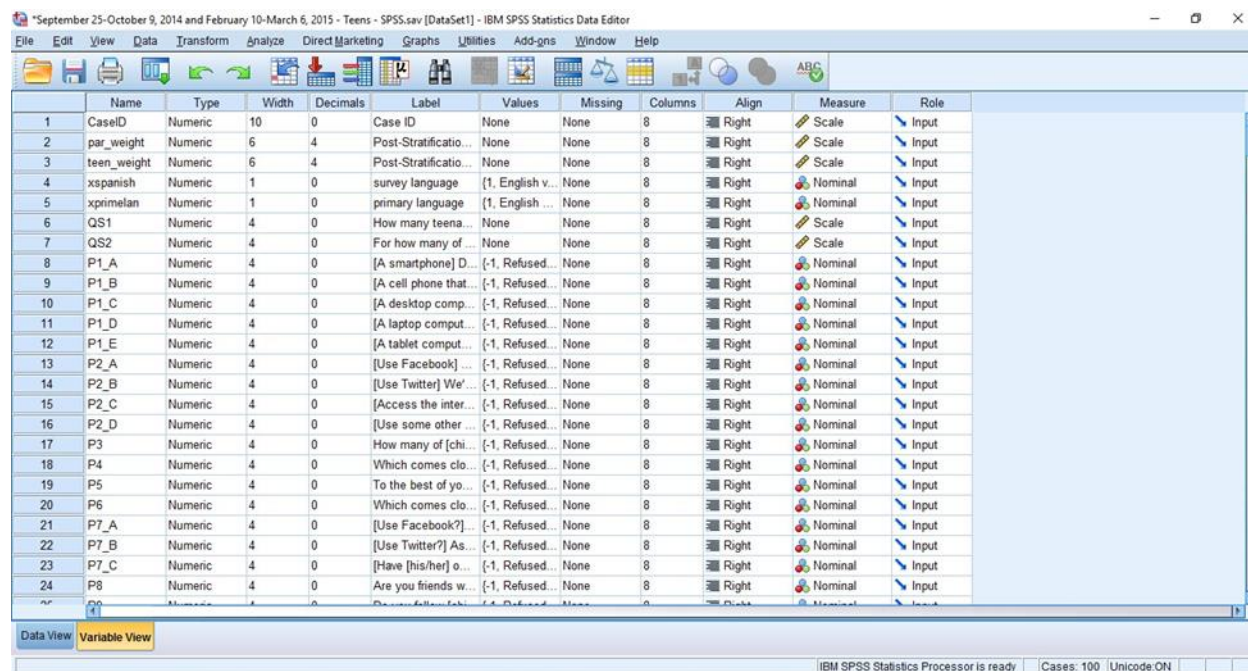


Slika 3.2 - Pregled podataka (Data View) i prozor za ubacivanje eksternih podataka

Variable View kartica prikazuje informacije o varijablama koje su unešene u *Data View*, ali ne i same podatke, ova kartica služi za definisanje varijabli.

Osnovni atributi svake promeljive su:

- Ime (Name)
- Tip (Type)
- Širina (Width)
- Broj decimala (Decimals)
- Oznaka (Label)
- Vrednost (Values)
- Nedostajući podaci (Missing)
- Širina kolone u editoru podataka (Columns)
- Način poravnanja sadržaja u ćeliji (Align)
- Način meranje promenljivih (Measure)
- Uloga promenljive (Role)



Slika 3.3 - Pregled varijabli (Variable View)

3.2. Prozor izlaznih rezultata - Viewer

Kad izvodite bilo koju naredbu u SPSS-u, *Viewer* automatski otvara i prikazuje zapisnik o preduzetim radnjama. Prvenstveno, *Viewer* je mesto gde se prikazuju rezultati statističke analize, ali svaka komanda pozvana putem padajućih menija i sintakse će se ispisati u *Viewer*-u.

To uključuje otvaranje, zatvaranje ili snimanje datoteke podataka. Ako *Viewer* prozor nije otvoren kada je naredba pokrenuta, novi prozor će automatski biti kreiran. *Viewer* je podeljen u dva dela, ili okvira. Levi okvir sadrži prikaz sadržaja. Ovaj prikaz je posebno koristan kada ste pokrenuli više SPSS naredba i morate pronaći određeni deo rezultata. Desni okvir sadrži stvarni izlaz, klikom na stavku u levom okviru će skočiti na taj sadržaj u desnom okviru. Predmeti koji su odabrani u desnom okviru označene su crvenom strelicom i okvirom oko sadržaja.

U *Viewer*-u selekcijom sadržaja na levoj ili desnoj strani okvira možete izvršiti kopiranje, lepljenje ili brisanje sadržaja. Za uklanjanje sadržaja iz *Viewer*-a, kliknite na njegovo ime u levom okviru ili kliknite na sam sadržaj u desnom okviru, a zatim pritisnite dugme Delete na tastaturi.

Viewer prozor može biti usnimljen kao fajl (*.spv), tako da ga možete ponovno otvoriti, bez potrebe da se ponovno pokreću iste naredbe u SPSS-u.

T-Test

One-Sample Statistics

	N	Mean	Std. Deviation	Std. Error Mean
SMEAN(OS1)	1638	1.235	1.0873	.0269

One-Sample Test

	t	df	Sig. (2-tailed)	Mean Difference	95% Confidence Interval of the Difference	
					Lower	Upper
SMEAN(OS1)	8.758	1637	.000	.2353	.183	.288

T-TEST GROUPS=PPGENDER(1 2)
/MISSING=ANALYSIS
/VARIABLES=R4
/CRITERIA=CI (.95).

T-Test

Group Statistics

	Gender	N	Mean	Std. Deviation	Std. Error Mean
On an average day, about how many text messages do you send and receive	Male	410	62.28	273.425	13.504
	Female	1228	1.18	1.0873	.0269

Slika 3.4 - Prozor izlaznih rezultata (Viewer)

3.3. Sintaksni editor - Syntax Editor

SPSS sintaksa je programski jezik jedinstven za SPSS koji se može koristiti kao alternativa padajućim menijima za manipulaciju podataka i statističke analize. *Syntax Editor* je mesto gde korisnici mogu pisati, ispravljati, i izvršiti SPSS sintaksu. Da biste otvorili novi prozor *Syntax Editor*, kliknite na *File>New>Syntax*.

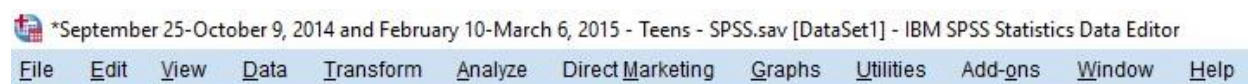
U desnom okviru *Syntax Editor*-a je mesto gde se upisuju sintakse. Leva strana okvira *Syntax Editor*-a prikazuje pregled naredbi ispisane u sintaksi, a može se koristiti za navigaciju unutar koda. Možete skočiti na određeni deo vašeg koda klikom na naredbu u levom okviru. Ova prečica je korisna za prikaz početaka i kraja naredbe, pogotovo ako je naredba je duža od jednog reda.

Iskusniji korisnici SPSS programa mogu da koriste ovaj prozor da direktno nose komande za izveštanje SPSS procedura, čime postižu veću fleksibilnost u radu i otvaraju dodatne mogućnosti programa, koje nisu dostupne sistemom menija i ikona.¹²

3.4 Padajući meniji u programu SPSS

SPSS verzija 23 nudi 12 glavnih padajućih menija, a to su: File meni, Edit meni, View meni, Data meni, Transform meni, Analyze meni, Direct Marketing meni, Graphs meni, Utilities meni, Add-ons meni, Window meni i Help meni.

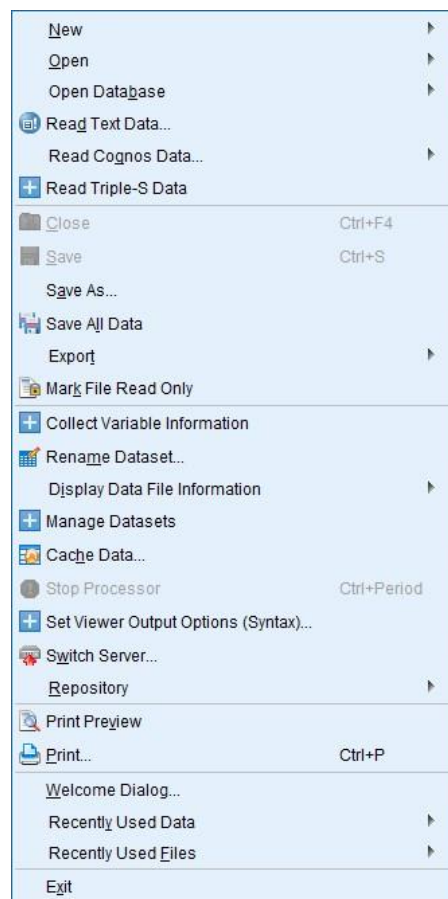
Data, Transform i Analyze su tri najbitnija menija i zajedno služe za obradu naprednih statističkih zadataka.



Slika 3.5 - Padajući meniji

3.4.1. File meni

Namena File menija je rad sa datotekama. Korišćenjem ponuđenih opcija na ovom meniju, pravite nove datoteke, otvarate postojeće ili snimate datoteke u različitim formatima, prikazujete informacije o datoteci, štampate datoteku i izlazite iz programa SPSS. Meni File takođe navodi spisak nedavno korišćenih datoteka sa podacima (*Recently Used Data*) i ostalih nedavno korišćenih datoteka (*Recently Used Files*), tako da se brzo možete vratiti na prethodni dokument.¹³



Slika 3.6 - File meni

¹² Soldić-Aleksić dr Jasana, Praktikum iz primenjene analize podataka (Beograd, Ekonomski fakultet u Beogradu, 2013), str 10.

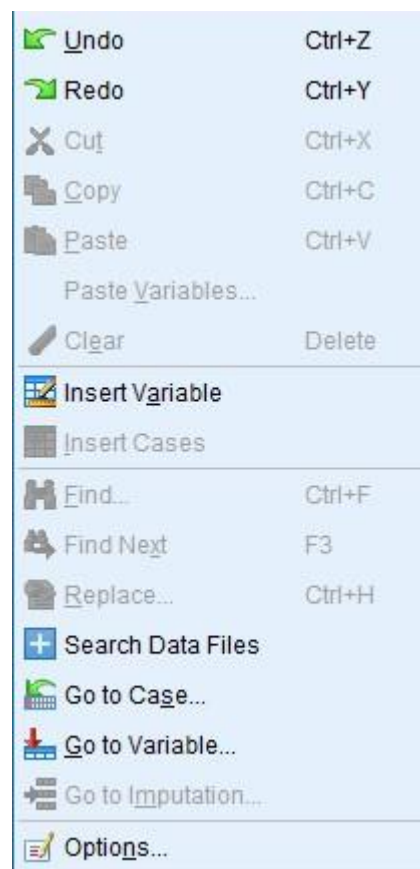
¹³ http://cet.rs/media/k2/attachments/spss_pog01.pdf, sstranica posećena 16. marta 2017.

3.4.2. Edit meni

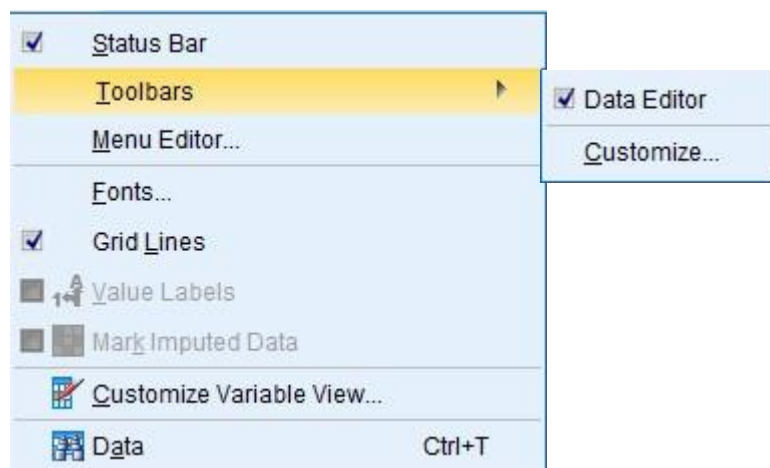
Ovaj meni se koristi kada je potrebno da kopirate ili isečete podatke i umetnete ih na drugo mesto u tekućoj ili nekoj drugoj datoteci sa podacima. Opcije na Edit meniju se koristite još i da biste pronašli neki podatak ili tekst, zamenili tekst ili prilagodili podešavanja programa SPSS.¹⁴

3.4.3. View meni

View meni je sličan kao i kod mnogih drugih programskih okruženja. Komande koje se pojavljuju u ovoj padajućoj listi koriste se za podešavanje radnog okruženja u pozoru editora podataka. To podrazumeva da se izabere prikazivanje (ili ne) statusne linije, editora menija, palete sa ikonama, okvira ćelija, imena ili labela (naziva) promenljivih i slično. Uglavnom se prihvataju opcije koje su već unapred programski postavljene, tako da se kasnije retko vrši njihova promena.¹⁵



Slika 3.7 - Edit meni



Slika 3.8 - View meni

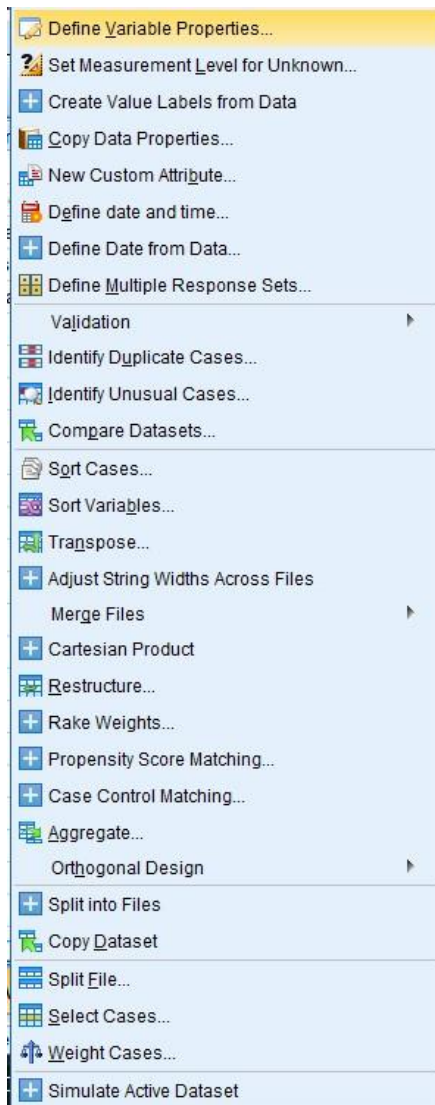
¹⁴ http://cet.rs/media/k2/attachments/spss_pog01.pdf

¹⁵ Soldić-Aleksić dr Jasana, Praktikum iz primenjene analize podataka (Beograd, Ekonomski fakultet u Beogradu, 2013), str 23.

3.4.4. Data meni

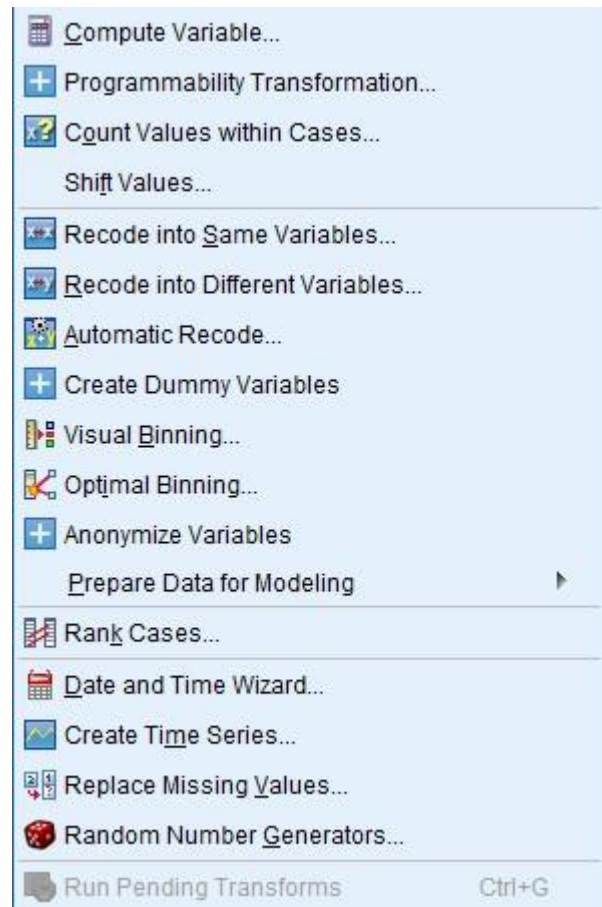
Glavni elementi u bilo kojoj SPSS analizi su promenljive i njihove vrednosti, a za to su potrebne moćne alatke za rad koje Data meni

pruža. Kao što možete da vidite sa slike 2.2, na meniju Data postoji skup komandi koje služe za: definisanje promenljivih, sortiranje slučajeva, spajanje i grupisanje datoteka, podelu podataka u editoru podataka, filtriranje podataka i još drugih funkcija.



3.4.5. Transform meni

Podaci koji su uneti u program, nisu ponekad uvek pogodni za dalju analizu. Dešava se da neka varijabla nije adekvatno numerički kodirana, ili je potrebno dodatno da se definiše, izračuna i slično. Nakon unošenja podataka u program, a pre statističkih analiza i dodatnih kalkulacija, poželjno je izvršiti transformaciju podataka, ukoliko je potrebna.



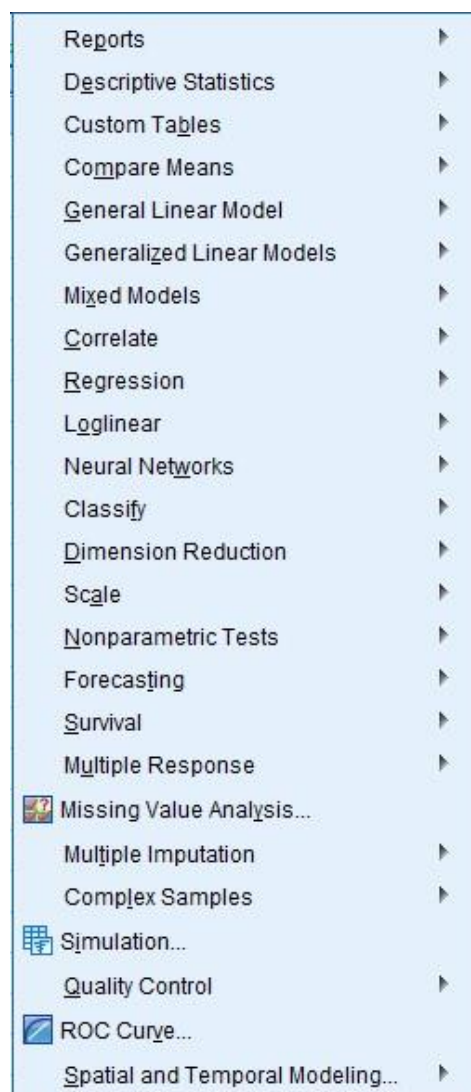
Slika 3.9 - Data meni (levo) i Transform meni (desno)

3.4.6. Analyze meni

U Analyze meniju postoji 24 različitih opcija (i brojni podmeniji) koji vode do skoro svih statističkih tehnika za analizu koje biste poželili da koristite. Izbor se kreće od jednostavnog izračunavanja srednje vrednosti i standardne devijacije do analize vremenskih nizova i višestruke regresije, kao i drugih veoma složenih analiza.¹⁶

3.4.7. Direct Marketing meni

Direct Marketing meni je usko specijalizovani meni koji SPSS korisnicima omogućava da naprave opcije menija za neke veoma posebne slučajeve.¹⁷



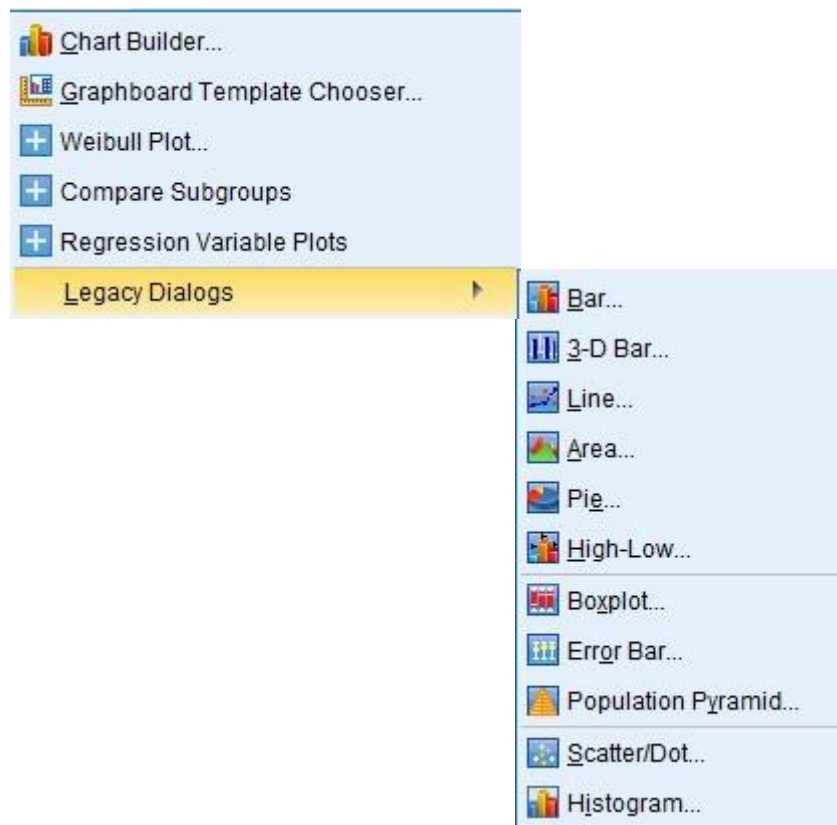
Slika 3.10- Analyze meni

¹⁶ http://cet.rs/media/k2/attachments/spss_pog01.pdf

¹⁷ Isto

3.4.8. Graphs meni

Ovaj meni služi da, ukoliko želite, vidite kako dobijeni rezultati izgledaju. Grafikonima oživljavate brojeve, a možete da napravite stubičasti, linijski, površinski dijagram kao i druge vrste grafikona. Od verzije 21 imate mogućnost da koristite komandu menija *Chart Builder* (pri čemu će vas SPSS provesti kroz izradu grafikona) ili komandu menija *Legacy Dialogs* preko koje je dostupan SPSS interfejs iz njegovih ranijih verzija.



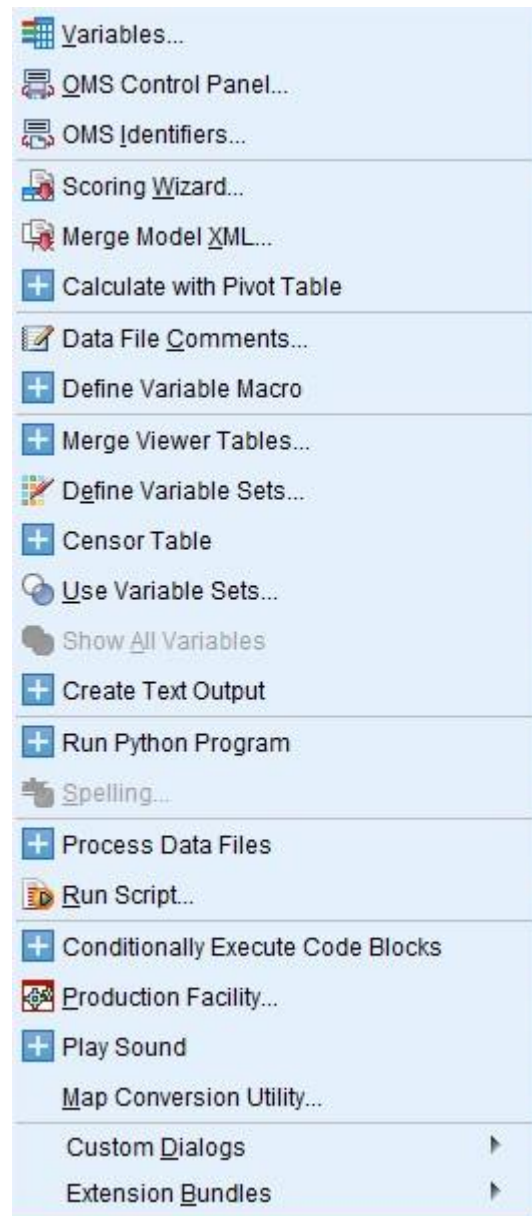
Slika 3.11 - Graph meni, sa pod menijem i izborom dizajna dijagrama

3.4.9. Utilities meni

Ovde možete da pronađete informacije o promenljivama i datotekama, a možete i da definišete i koristite skup promenljivih.

3.4.10. Add-ons meni

Ovde su smeštene informacije sa savetima i uputstvima za rad sa SPSS programom.

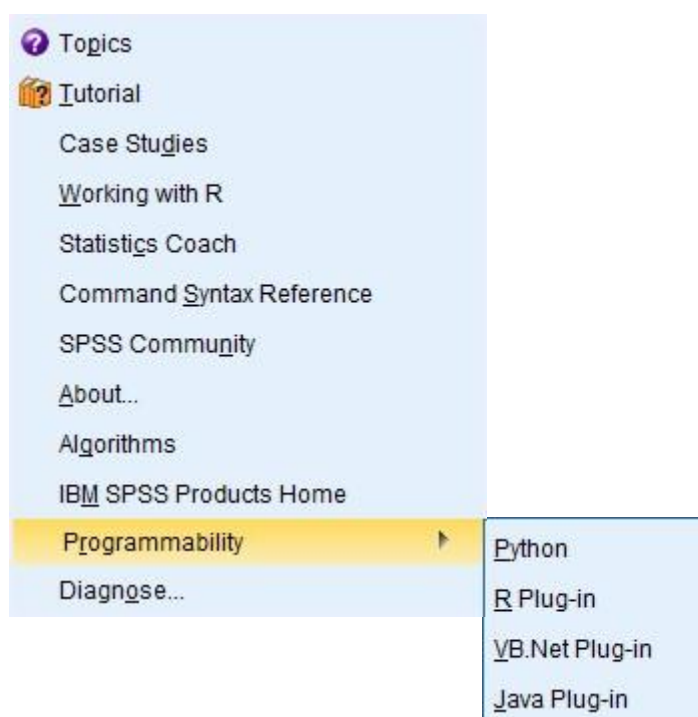


Slika 3.12 - Utilities meni

3.4.11. Window i Help meniji

Ova dva menija se ponašaju na sličan način kao i u bilo kojoj drugoj Windows aplikaciji. Window meni pomaže da prelazite sa jednog prozora na drugi i da umanjite prozore *Data Editor* ili *Viewer*. Help meni nudi onlajn pomoć.¹⁸

Help meni takođe nudi onlajn objašnjenja za korišćenje opcija kao što su *Case Studies*, *Working with R* i *Algorithms*, *About* (Informacije o SPSS paketu), a tu je i opcija *Programmability* u kojoj je detaljno objašnjeno korišćenje programskih jezika unutar SPSS programa.



Slika 3.13– Help meni i podmeni Programmability

¹⁸ http://cet.rs/media/k2/attachments/spss_pog01.pdf

3.5. Paleta alatki i statusna traka u SPSS programu

Paleta alatki je najlakši način za pokretanje nekih od najkorišćenijih alata u radu sa SPSS-om, skup ikona koje se nalaze ispod menija. Na slici 3.14 možete da vidite paletu alatki u prozoru *Data View* i opis šta svaka ikona predstavlja. One su, takoreći, prečice nekih komandi iz padajućih menija. Za svaki prozor u SPSS-u se razlikuju ponuđeni alati u paleti, takođe se mogu i izmeniti po sopstvenoj potrebi, uz pomoć, *View>Toolbars>Customize*

Tabela 1 *Ikone na paleti alatki*

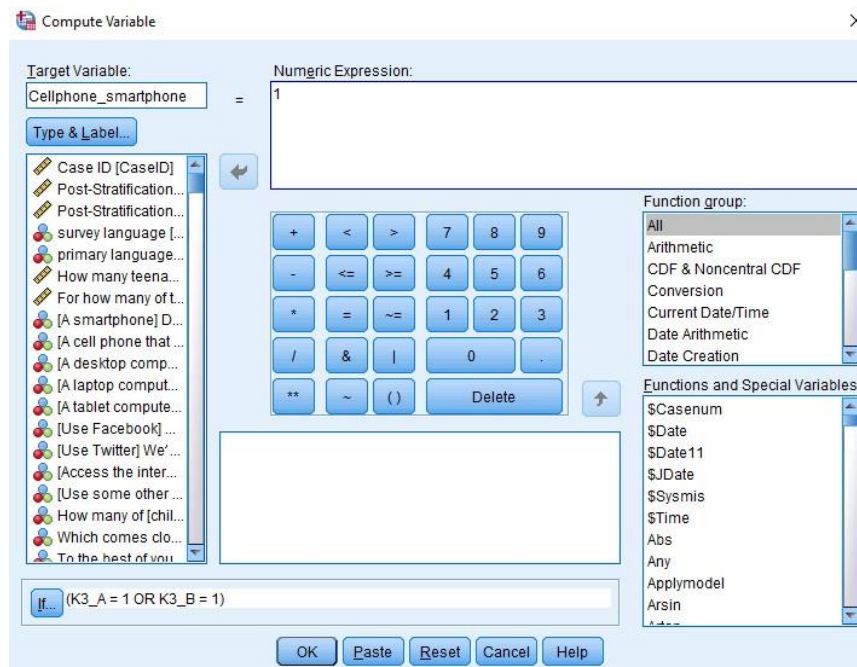
Ikona	Naziv	Šta obavlja
	Open data document	Otvora ranije napravljenu datoteku
	Save this document	Snima novu ili upravo napravljenu datoteku
	Print	Štampa datoteku
	Recall recently used dialogs	Ponovo poziva poslednje korišćene okvire za dijalog
	Undo a user action	Poništava izmene u formatiranju ili unosu podataka
	Redo a user action	Ponovo obavlja prethodnu izmenu
	Go To Case	Ide na brožčani slučaj
	Go To Variable	Ide na imenovanu promenljivu
	Variables	Pruža informacije o određenoj promenljivoj
	Run descriptive statistics	Izračunava deskriptivnu statistiku
	Find	Pronalazi zapis
	Insert Cases	Umeće slučaj u datoteku sa podacima
	Insert Variables	Umeće novu promenljivu u datoteku sa podacima
	Split File	Deli datoteku između definisanih promenljivih
	Weight Cases	Određuje značaj slučaja
	Select Cases	Bira skup slučajeva prema nekom kriterijumu
	Value Labels	Uključuje i isključuje oznake
	Use Variable Sets	Pravi skup promenljivih
	Show All Variables	Prikazuje sve promenljive u skupu podataka
	Spell Chek	Proverava pravopis

Slika 3.14 – Paleta alata

Statusna traka je još jedna korisna alatka koja se nalazi pri dnu SPSS prozora. Na njoj možete da vidite jednoređni izveštaj o tome koju aktivnost SPSS trenutno obavlja.

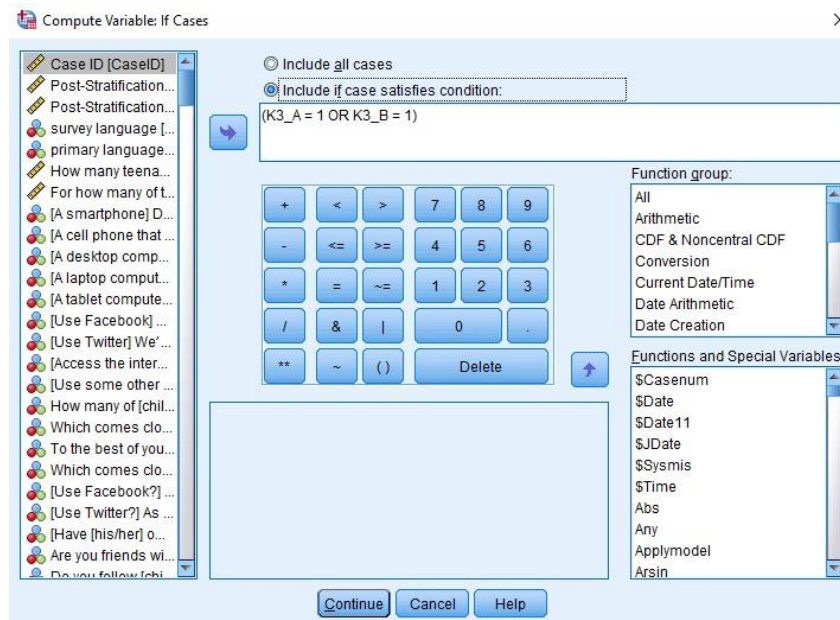
4. Analiza podataka

Pretpostavimo da imamo dve varijable, jednu koja nam govori o tome da li ispitanik poseduje pametni telefon *“Do you, personally, have or have access to each of the following items, or not. Do you have a smartphone?”*, a druga da li poseduje mobilni telefon koji nije pametni telefon *“Do you, personally, have or have access to each of the following items, or not. Do you have a cell phone that is not a smartphone?”*. Nama je za dalju analizu u programu SPSS potrebna promenljiva koja će da objedini ove dve varijable, zapravo, da predstavi sve osobe koje imaju bar jedan od dva ponuđena telefona. Opcija *„Transform/Compute Variable“* će nam otvoriti Dialog Box, slika 4.1, u kojem podešavamo kreiranje nove varijable na sledeći način :



Slika 4.1 - Dialog box u kom kreiramo novu varijablu

U polje *Target Variable* unosimo naziv nove promenljive, u našem slučaju *„Cellphone_smartphone“* . U polje *Numeric Expression* unosimo vrednost koju želimo da promenljiva ima ukoliko budu zadovoljeni uslovi. Pritiskom na dugme *„If“* otvara nam se nov prozor u kom definišemo uslove. Opcija *IF* omogućava da se precizira uslov pod kojima će se analiza primenjivati.



Slika 4.2 - Compute variable : opcija IF

Nakon što selektujemo opciju *Include if case satisfies condition*, u prazno polje ispod toga unosimo uslov da od dve varijable koje ispitujeemo u analizu uđu samo vrednosti koje su jednake 1, odnosno, samo pozitivni odgovori na pitanje.

$$(K3_A = 1 \text{ OR } K3_B = 1)$$

Na desnoj strani slike 4.2 ponuđene su i već unapred pripremljene matematičke funkcije za brže i lakše preračunavanje, koje nam mogu olakšati i ubrzati transformaciju varijabli. Tu je grupa aritmetičkih i finansijskih funkcija, funkcija koje se odnose na računanje vremena i slično. Nama to nije bilo potrebno u ovom slučaju, pa pritiskom na dugme *Continue*, vraćamo se na prethodni prozor i završavamo definisanje varijable.

Na slici 4.3 možemo videti izlazni rezultat funkcije *Compute Variable*. Predstavljen je u vidu kolone pod nazivom „*Cellphone_smartphone*“, koja nam je potrebna za dalju analizu i korišćenje nam u deskriptivnoj statistici.

The screenshot shows the IBM SPSS Statistics Data Editor window. The title bar indicates the file is 'September 25-October 9, 2014 and February 10-March 6, 2015 - Teens - SPSS.sav [DataSet1] - IBM SPSS Statistics Data Editor'. The menu bar includes File, Edit, View, Data, Transform, Analyze, Direct Marketing, Graphs, Utilities, Add-ons, Window, and Help. The toolbar contains various icons for file operations, data manipulation, and analysis. The main data grid shows 23 rows of data. The columns are labeled: PPNET, QS1_1, K4_1, RK4_1, and Cellphone_smartphone. The 'Cellphone_smartphone' column contains binary values (0 or 1). The status bar at the bottom indicates 'Data View' and 'Variable View' tabs.

	PPNET	QS1_1	K4_1	RK4_1	Cellphone_smartphone
1	1	1.0	15.0	342.000	1
2	1	1.0	45.0	592.500	1
3	1	1.0	10.0	279.000	1
4	1	1.0	58.9	1057.000	1
5	1	2.0	-98.0	51.500	1
6	1	1.0	50.0	648.500	1
7	1	2.0	58.9	1057.000	1
8	1	1.0	20.0	411.000	1
9	1	1.2	58.9	1057.000	1
10	1	1.2	58.9	1057.000	1
11	1	2.0	50.0	648.500	1
12	1	1.0	500.0	1622.000	1
13	1	1.0	50.0	648.500	1
14	1	1.2	58.9	1057.000	1
15	1	1.2	58.9	1057.000	1
16	1	1.0	7.0	230.000	1
17	1	1.0	80.0	1435.500	1
18	1	1.0	100.0	1489.000	1
19	1	0	58.9	1057.000	1
20	1	2.0	58.9	1057.000	1
21	1	1.0	10.0	279.000	1
22	1	1.0	58.9	1057.000	1
23	1	1.0	500.0	1622.000	1

Slika 4.3 - Nova varijabla izračunata pomoću *Transform / Compute Variable*

4.1. Deskriptivna statistika

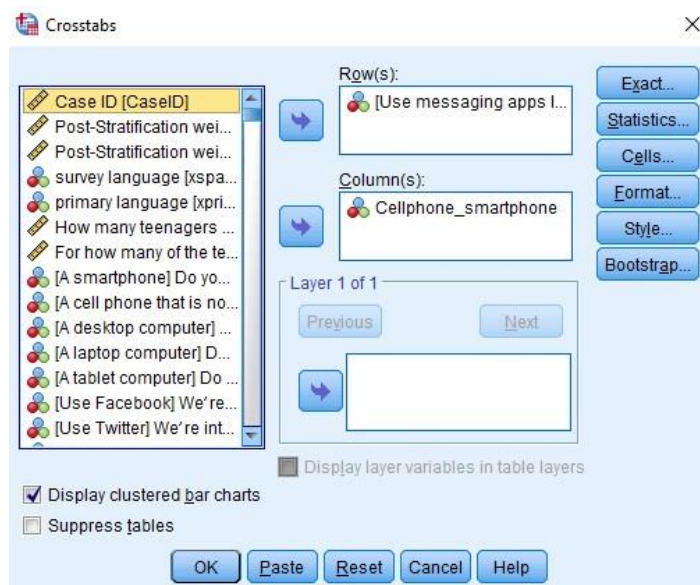
Deskriptivna statistička analiza predstavlja skup metoda kojima se vrši izračunavanje, prikazivanje i opisivanje osnovnih karakteristika statističkih serija.

Deskriptivna statistička analiza ima sledeće zadatke:

1. grupisanje i sređivanje statističkih podataka
2. prikazivanje statističkih podataka
3. određivanje osnovnih pokazatelja statističkih serija.

4.2. Deskriptivne statističke mere

Deskriptivne statističke mere moguće je izračunati kod numeričkih varijabli i one su veoma korisne za svako istraživanje. Komunikacija između ljudi preko poruka je pretrpela promene u poslednjih nekoliko godina sa pojavom aplikacija na mobilnim telefonima koje omogućavaju brže i lakše komuniciranje, kao i smanjenje troškova. Prema analizi koju smo sprovedi birajući opciju „*Analyze/Descriptive Statistics/Crosstab*“, 91% tinejdžera, koji poseduju pametni ili običan mobilni telefon, koristi tekstualne poruke - ili direktno preko svojih mobilnih telefona ili putem aplikacije ili veb sajta.



Slika 4.4 – Crosstab analiza

U polje *Row(s)* ubacujemo promenljivu “*Do you do any of the following online or on your cellphone: Use messaging apps like WhatsApp or Kik?*”, a u polje *Column(s)* promenljivu koju smo u prethodnom primeru izračunali pomoću opcije *Transform/Compute Variable*. Kako bismo imali i grafički prikaz rezultata, štikliramo i opciju *Display clustered bar charts*, kao što se može videti na slici 4.4 u donjem levom uglu.

Kao rezultat odabranih statistika, u prozoru Viewer dobijamo izlazne podatke koje tumačimo na sledeći način:

U tabeli koja se zove *Case Processing Summary*, slika 4.5, nalaze se tri kolone: *Valid*, *Missing* i *Total*. Od ukupno ispitanih 1638 roditelja i tinejdžera, validan odgovor dalo je 929 ljudi, dok se pod *Missing values* podrazumeva broj nevalidnih ili nepruženih odgovora i u ovom slučaju taj broj iznosi 709. Kolona *Total* je suma prethodne dve kolone.

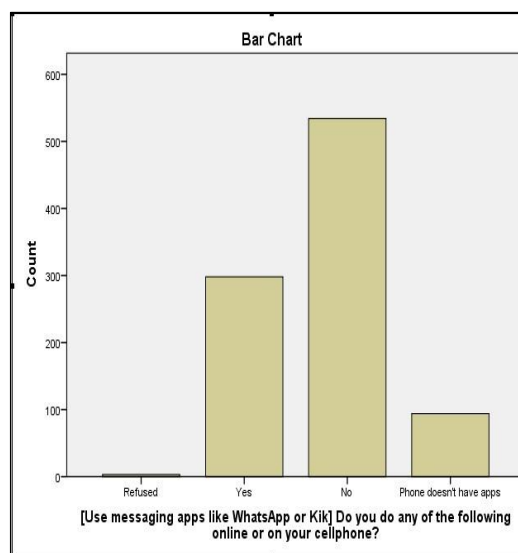
U drugoj tabeli prikazano je unakrsno tabeliranje (*Cross-tabulation*) dve promenljive koje smo prethodno definisali i objasnili. Promenljiva “*Do you do any of the following online or on your cellphone: Use messaging apps like WhatsApp or Kik?*” ima četiri vrednosti: *Refused* (bez odgovora), *Yes* (da), *No* (ne), *Phone doesn't have apps* (telefon ne podržava aplikacije). Promenljiva „*Cellphone “smartphone“*“ nalazi se u prvoj koloni, pored kolone *Total* koja prikazuje sumiran prikaz za svaki red. Kao što se može primetiti na slici 4.5, 32.1% tinejdžera koji poseduju mobilne telefone koristi aplikacije za slanje poruka. Čak 57.5% tinejdžera i dalje komunicira isključivo preko osnovne aplikacije za slanje poruka, koja podržava svaki tip telefona.

➔ Crosstabs

Case Processing Summary						
	Cases					
	Valid		Missing		Total	
	N	Percent	N	Percent	N	Percent
[Use messaging apps like WhatsApp or Kik] Do you do any of the following online or on your cellphone? * Cellphone_smartphone	929	56.7%	709	43.3%	1638	100.0%

[Use messaging apps like WhatsApp or Kik] Do you do any of the following online or on your cellphone? * Cellphone_smartphone Crosstabulation

			Cellphone_s marphone	
			1	Total
[Use messaging apps like WhatsApp or Kik] Do you do any of the following online or on your cellphone?	Refused	Count	3	3
		% of Total	0.3%	0.3%
	Yes	Count	298	298
		% of Total	32.1%	32.1%
	No	Count	534	534
		% of Total	57.5%	57.5%
	Phone doesn't have apps	Count	94	94
		% of Total	10.1%	10.1%
Total	Count	929	929	
	% of Total	100.0%	100.0%	



Slika 4.5 – Rezultati dobijeni funkcijom unakrsnog tabeliranja

Koristeći opet unakrsno tabeliranje na isti način kao i za prethodni primer, ali ovaj put za promenljivu “*Do you do any of the following online or on your cellphone: use social media?*” (Da li koristite društvene mreže na vašim telefonima?), saznajemo da je čak 77.8 % ispitanih tinejdžera aktivno na društvenim mrežama preko svojih telefona. Od preostalih 22.2%, njih 22.1% je odgovorilo da ne koriste telefon za pristup društvenim mrežama, dok 0.1% nije pružilo odgovor na postavljeno pitanje.

[Use social media] Do you do any of the following online or on your cellphone?
^ Cellphone_smartphone Crosstabulation

			Cellphone_s martphone	
			1	Total
[Use social media] Do you do any of the following online or on your cellphone?	Refused	Count	1	1
		% of Total	0.1%	0.1%
	Yes	Count	723	723
		% of Total	77.8%	77.8%
	No	Count	205	205
		% of Total	22.1%	22.1%
Total		Count	929	929
		% of Total	100.0%	100.0%

Slika 4.6 – Rezultati unakrsnog tabeliranja

4.3. Testovi značajnosti razlika

Kada se statističko posmatranje organizuje sa ciljem da se analizira uzorak, javlja se potreba da se opis upoređuje bilo sa nekom zamišljenom teorijskom veličinom, bilo s nekim drugim određenim uzorkom. Razlike ili istovetnosti mogu biti slučajne ili organizovane. Slučajne razlike su posledica slučajnih variranja i po pravcu, odnosno smislu i ne mogu se predvideti. Organizovane promene vode poreklo od delovanja faktora koji se mogu, ili je poželjno da se identifikuju.

Kod poređenja testovima pripadništva uzorka skupu, odnosno razlikama među njima, istraživač postavlja nultu i radnu hipotezu: H_0 – obično za nultu hipotezu uzimamo pretpostavku da ne postoji statistički značajna razlika između uzorka i populacije (ili između dva uzorka); H_1 – radna ili alternativna hipoteza: kod testiranja obično pretpostavljamo da postoji statistički značajna razlika između uzorka i populacije (ili između dva uzorka). Takođe, postavlja se i nivo značajnosti sa kojim ćemo raditi test, odnosno nivo sigurnosti sa kojom možemo tvrditi da je nešto onako kako tvrdimo. Kao standardan nivo pouzdanosti se obično uzima nivo $p=0.05$, što znači da sa sigurnošću od 95% možemo tvrditi da je nešto onako kako tvrdimo. Kao veći stepen pouzdanosti se uzima $p=0.01$, odnosno da smo 99% sigurni u ono što tvrdimo.¹⁹

T-test nezavisnih uzoraka, *Independent-Samples T-test*, upotrebljavamo za poređenje srednje vrednosti neke neprekidne promeljive u dve različite grupe subjekata²⁰. Reč je o parametarskom statističkom testu, za koji se vezuju pretpostavke da je posmatrano obeležje (promenljiva) merena bar na intervalnoj skali (numerička promenljiva) i da su uzorci izvučeni iz populacije sa normalnom raspodelom posmatranog obeležja. Ovaj test se primenjuje kada je nepoznata varijansa osnovnih skupova, pa se ona ocenjuje na osnovu uzoračke varijanse. Za testiranje hipoteze o jednakosti varijansi najčešće se koristi *Levene test*, koji je od svih testova o homogenosti varijanse najmanje osetljiv na pretpostavku o normalnoj raspodeli posmatranog obeležja.

Za dublju analizu naše datoteke podatka, pretpostavićemo da želimo da saznamo da li se na nivou značajnosti od 5% može prihvatiti tvrđenje da postoje razlike između devojak i momaka u korišćenju društvenih mreža: Facebook, Twitter, Instagram, Google +, Snapchat i ostalih društvenih mreža.

Biramo opciju *Analyze/Compare Means/Independent-Samples T Test*. Promenljive koje testiramo su “[Facebook] Which of the following social media do you use?”, “[Twitter] Which of the following social media do you use?”, “[Instagram] Which of the following social media do you use?”, “[Google+] Which of the following social media do you use?”, “[Snapchat] Which of the following social media do you use?” i “[Other social media not listed here] Which of the following social media do you use?”, koje su unete u polje *Test Variable(s)* (slika 4.7, levo), a promenljiva koja formira dve grupe jeste “Child_gender”, koja se unosi u polje *Grouping Variable*. Grupe se definišu u dijalogu prozoru *Define Groups* (slika 4.7, desno), gde se u polja

¹⁹ <http://www.ftn.kg.ac.rs/download/SIR/SIR%20Cuslovic%20Milos.pdf>

²⁰ Julie Pallant, SPSS Survival Manual, Third Edition, (Sydney, Allen & Unwin, 2007), str. 235.

Group 1 i Group 2 unose šifre za dve posmatrane grupe. U našem slučaju 1 predstavlja muški pol, a 2 ženski pol.



Slika 4.7 – Okvir za dijalog procedure (levo) i okvir za definisanje grupa (desno)

Na osnovu podataka koje možemo videti u tabeli na slici 4.8, o grupnim statistikama zaključujemo da je intervjuisano je 523 momka i 537 devojaka. Prosečan stav momaka prema korišćenju: Facebook-a je 1.3, Twittera je 1.7, Instagram-a 1.59, Google+ je 1.67, Snapchat-a je 1.7 i ostalih društvenih mreža oko 1.88 (sve je mereno na intervalnoj skali koja uzima vrednosti od 1-5). Prosečan stav žena prema korišćenju: Facebook-a je 1.3, Twitter-a je 1.62, Instagram-a je 1.39, Google + je 1.67, Snapchat-a je 1.49 i ostalih društvenih mreža je 1.85. Postavljamo pitanje da li su ove razlike statistički značajne, odnosno da li možemo prihvatiti tvrđenje da postoje razlike između devojaka i momaka u korišćenju prethodno navedenih društvenih mreža?

Group Statistics					
	DATA-ONLY: Gender of child selected	N	Mean	Std. Deviation	Std. Error Mean
[Facebook] Which of the following social media do you use?	Male	523	1.30	.469	.021
	Female	537	1.30	.471	.020
[Twitter] Which of the following social media do you use?	Male	523	1.70	.482	.021
	Female	537	1.62	.509	.022
[Instagram] Which of the following social media do you use?	Male	523	1.59	.516	.023
	Female	537	1.39	.500	.022
[Google+] Which of the following social media do you use?	Male	523	1.67	.484	.021
	Female	537	1.67	.482	.021
[Snapchat] Which of the following social media do you use?	Male	523	1.70	.484	.021
	Female	537	1.49	.511	.022
[Other social media not listed here] Which of the following social media do you use?	Male	523	1.88	.375	.016
	Female	537	1.85	.372	.016

Slika 4.8 – Prikaz rezultata dobijenih upoređivanjem dve grupe korisnika društvenih mreža

Na slici 4.9 vidimo da je prvo primenjen *Levene test* za proveru tvrđenja o homogenosti varijanse posmatranih obeležja u dva skupa. U našem primeru to znači da se testira hipoteza o jednakosti varijansi korišćenja društvenih mreža ispitanika – muškaraca i žena. Na osnovu realizovanog nivoa značajnosti *Levene test*-a na slici 4.9, kolona Sig. iznosi za:

- Facebook: $0.742 \geq 0.05$
- Twitter: $0.000 \leq 0.05$
- Instagram: $0.245 \geq 0.05$
- Google +: $0.677 \geq 0.05$
- Snapchat: $0.000 \leq 0.05$
- Ostale društvene mreže: $0.025 \geq 0.05$

Independent Samples Test										
		Levene's Test for Equality of Variances		t-test for Equality of Means						
		F	Sig.	t	df	Sig. (2-tailed)	Mean Difference	Std. Error Difference	95% Confidence Interval of the Difference	
									Lower	Upper
[Facebook] Which of the following social media do you use?	Equal variances assumed	.109	.742	-.184	1058	.854	-.005	.029	-.062	.051
	Equal variances not assumed			-.184	1057.473	.854	-.005	.029	-.062	.051
[Twitter] Which of the following social media do you use?	Equal variances assumed	18.953	.000	2.740	1058	.006	.083	.030	.024	.143
	Equal variances not assumed			2.742	1057.242	.006	.083	.030	.024	.143
[Instagram] Which of the following social media do you use?	Equal variances assumed	1.351	.245	6.282	1058	.000	.196	.031	.135	.257
	Equal variances not assumed			6.280	1054.492	.000	.196	.031	.135	.257
[Google+] Which of the following social media do you use?	Equal variances assumed	.174	.677	-.231	1058	.817	-.007	.030	-.065	.051
	Equal variances not assumed			-.231	1056.912	.817	-.007	.030	-.065	.051
[Snapchat] Which of the following social media do you use?	Equal variances assumed	53.721	.000	6.741	1058	.000	.206	.031	.146	.266
	Equal variances not assumed			6.746	1057.125	.000	.206	.031	.146	.266
[Other social media not listed here] Which of the following social media do you use?	Equal variances assumed	5.067	.025	1.243	1058	.214	.029	.023	-.016	.074
	Equal variances not assumed			1.243	1056.696	.214	.029	.023	-.016	.074

1

Slika 4.9 - Test značajnosti – *Levene Test*

Kako ja za promenljive koje se tiču Facebook-a, Instagrama, Google +, i ostalih društvenih mreža vrednost kolone Sig. veća od 0.05, možemo da zaključimo da se za te promenljive može prihvatiti tvrđenje da su jednake varijanse u dve grupe. U tom slučaju dalje posmatramo kolonu Sig.(2-tailed) sa slike 4.9, i to prvi red gde se podrazumeva jednakost varijansi. Kako je za promenljivu “[Facebook] Which of the following social media do you use?” realizovani nivo značajnosti t-testa 0.854 što je veće od rizika greške prve vrste 0.05, ne odbacujemo nultu hipotezu o jednakosti korišćenja Facebook-a između momaka i devojaka.

Kad je u pitanju promenljiva, “[Instagram] Which of the following social media do you use?”, realizovani nivo značajnosti t-testa iznosi 0.245, što je veće od rizika greške prve vrste 0.05, pa zato ne odbacujemo nultu hipotezu o jednakosti korišćenja Instagrama između momaka i devojaka. Što se tiče promenljive, “[Google+] Which of the following social media do you use?”,

realizovani nivo značajnosti t-testa iznosi 0.677 što je veće od rizika greške prve vrste 0.05, pa zato ne odbacujemo nultu hipotezu o jednakosti korišćenja Google + između momaka i devojaka.

Kada su u pitanju promenljive “[Twitter] Which of the following social media do you use?”, “[Snapchat] Which of the following social media do you use?” i “[Other social media not listed here] Which of the following social media do you use?” realizovan nivo značajnosti Levene test-a manji je od 0.05, kao što je i realizovani nivo značajnosti t-testa za svaku od njih, redom, iznosi : 0.000, 0.000 i 0.025, što je manje od rizika greške prve vrste 0,05, i prema tome odbacujemo nultu hipotezu o jednakosti korišćenja Twitter-a, Snapchat-a i ostalih društvenih mreža između muškaraca i žena.

Zaključak

Tehnologija zasnovana na statističkoj analizi podataka omogućava brzo uočavanje trendova i zakonitosti u velikim količinama podataka i predviđanje budućnosti na osnovu njih. Poslovni programi koji koriste tehnologiju zasnovanu na statističkoj analizi podataka daju veću sigurnost u poslovanju, jer istovremeno postaju esencijalno značajan alat koji se koristi pri donošenju poslovnih odluka.

Dobra i precizna analiza podataka je od velikog značaja za većinu organizacija. Međutim, da bi se dobila kvalitetna analiza, kao i izveštaji potrebni za donošenje poslovnih odluka, neophodno je mnogo znanja i veština kako bi se isplanirao sam postupak analize, prikupili podaci, upravljalo i pristupilo podacima, izvršila analiza i, na kraju, putem raznih izveštaja konačni rezultati prezentovali svim zainteresovanim subjektima. Dobro rešenje je izbor jednog analitičkog alata koji se sastoji od potrebnih modula ili samostalnog softvera koji ima različite specijalizovane funkcije.

Statistički softver IBM SPSS znatno povećava produktivnost jednog analitičkog procesa, jer se njegovom primenom značajno smanjuje obim rada i olakšava se prolazak kroz sve etape procesa analize korišćenjem jedinstvenog softverskog rešenja. Prezentovanjem svoje moći na jednom mestu, lako se zaključuje da se rad sa velikim brojem podataka i ne bi mogao zamisliti bez ovog ili nekog sličnog softvera.

U godinama koje dolaze, može se očekivati da će se softverski paketi sve više usavršavati kako bi održali korak sa savremenim svetom, istovremeno udovoljavajući potrebama modernog tržišnog poslovanja.

U današnje doba internet je dostupan skoro svuda u svetu i to sa različitih tipova uređaja, konstantno se razvijaju nove tehnologije i pojavljuju se različiti načini komunikacije uz pomoć interneta. Tinejdžeri danas internet koriste kao deo svakodnevnog načina života. Najdostupniji način pristupa internetu su prenosivi uređaji koji mogu da se konektuju preko Wi-Fi i/ili mobilne mreže.

Prema istraživanju koje je sprovedeno kroz ovaj rad, došli smo do zaključka da velika većina tinejdžera koristi mobilne telefone, bili to pametni ili obični mobilni telefoni. Takođe koriste tekstualne poruke u svakodnevnoj komunikaciji, direktno preko svojih mobilnih telefona ili putem aplikacije ili web sajta. Na osnovu podataka saznajemo da tinejdžeri i dalje koriste u malo većem procentu standardne SMS poruke od mobilnih aplikacija koje se koriste uz pomoć interneta. Saznajemo i da tinejdžeri koji poseduju pametne telefone u velikoj meri, nešto više od $\frac{3}{4}$ ispitanika, aktivno koriste internet za socijalne društvene mreže.

Tendencija za inovacijom u modernom svetu dovodi do toga da informacione tehnologije napreduju konstantno, a pritom je tinejdžerima sve to lako dostupno i dostižno.

Danas se skoro sve odvija preko interneta, tako da će se ovaj trend korišćenja interneta u budućnosti konstantno povećavati kod mlađe populacije, a društvene mreže postaće sve učestaliji vid druženja i komunikacije među mladima.

Literatura

Knjige:

- [1] Sheridan J. Coakes, prevod Soldić Aleksić dr Jasna, (2013), SPSS 20: Analiza bez muke, Kompjuter biblioteka, Beograd
- [2] Soldić-Aleksić dr Jasana, “Praktikum iz primenjene analize podataka”, Ekonomski fakultet u Beogradu, 2013.
- [3] Julie Pallant, “SPSS Survival Manual, Third Edition”, Sydney, Allen & Unwin, 2007.

Internet izvori:

- 1. http://stat.uns.ac.rs/LLLprogramme/NS/Katalog_kurseva.pdf
- 2. <http://en.wikipedia.org/wiki/MATLAB>
- 3. <http://www.r-project.org/about.html>
- 4. [https://en.wikipedia.org/wiki/R_\(programming_language\)](https://en.wikipedia.org/wiki/R_(programming_language))
- 5. http://en.wikipedia.org/wiki/SPSS_Inc
- 6. <http://en.wikipedia.org/wiki/SPSS#Overview>
- 7. <http://www-01.ibm.com/software/analytics/spss/products/statistics>
- 8. <http://www.pewinternet.org/2015/04/09/teens-social-media-technology-2015>
- 9. http://cet.rs/media/k2/attachments/spss_pog01.pdf
- 10. <http://libguides.library.kent.edu/SPSS/Environment>
- 11. <http://www.ftn.kg.ac.rs/download/SIR/SIR%20Cuslovic%20Milos.pdf>