

# 7. Statistička obrada podataka SPSS

Do sada ste već stekli temeljno razumijevanje statistike, manipulacije podacima, uspostavljanja simulacije, modeliranja i analize unutar logističkih opskrbnih lanaca, zajedno s izravnim metodama linearne regresije. Dok statistika nudi raznolik raspon modela i tehnika za

poboljšanje vaših napora optimizacije, provođenje analize i prepoznavanje potencijalnih poboljšanja, možda ste primijetili da kako složenost analiziranih podataka i izračuna raste, tradicionalni pristupi mogu postati sve zamršeniji i izazovniji za računanje. Kako zamršenost vaših podataka i

izračuna raste, konvencionalne metode mogu zastarjeti i, u nekim slučajevima, ugroziti pouzdanost vaših rezultata. Kako bi se to prepoznalo, statistika koristi razne softverske programe koji automatiziraju analizu i interpretaciju prikupljenih podataka, a istovremeno pružaju mnoštvo modela i funkcija za osiguranje pouzdanih rezultata. Jedan takav softver je IBM-ov SPSS, koji će biti ključni alat u ovom poglavlju. U ovom poglavlju pružit ćemo sažeti uvod u primarnu upotrebu softvera SPSS, istražujući njegove funkcionalnosti i praktične primjene. Nakon početnog uvoda slijedi praktična primjena programa kroz četiri temeljna testa za izračun rezultata: T-test, korelacije, Hi-kvadrat i ANOVA. Kako bismo vam olakšali učenje, predstavit ćemo vam jednostavne probleme i njihova rješenja kako biste se lakše upoznali s ovim testovima.

# 7.1 Osnove IBM-ovog SPSS-a

Možda ste već imali iskustva sa SPSS softverom. Međutim, ako ipak trebate, dopustite da vam ponudimo kratki uvod. SPSS, poput svog općepriznatijeg pandana Excela, olakšava manipulaciju podacima, analizu i vizualizaciju. Ipak, za razliku od Excela, koji ponekad može biti naporan i složen u programiranju funkcija, SPSS nudi korisničko sučelje za statističku analizu (IBM, 2021.). Nudi niz funkcija i metodologija za učinkovito rukovanje vašim podacima. Dok se SPSS softver ističe u pružanju opsežnih mogućnosti statističke analize, upravljanje manipulacijom podacima i konfiguriranje početnih postavki za analizu ponekad može biti izazovno (IBM, 2021.). Stoga ćemo se pozabaviti osnovama uvoza podataka i pripreme podataka za naknadne statističke testove.



S obzirom na široku upotrebu Excela za rukovanje numeričkim podacima, vaši se podaci mogu dobiti ili pripremiti u proračunskoj tablici programa Excel. Srećom, SPSS može uvesti podatke iz različitih formata datoteka u svoje proračunske tablice. Nakon što pripremite finalizirane Excel proračunske tablice, otvorite softver SPSS. Na početnom zaslonu idite na karticu "Datoteka" i odaberite "Uvezi podatke". U sljedećem prozoru možete odabrati format podataka koji namjeravate uvesti (pogledajte sliku 7.1). Slijedeći ovaj korak, pronađite pripremljenu datoteku, odaberite je i prijeđite na sljedeći prozor. Ovaj prozor će od vas tražiti da konfigurirate dodatne postavke. Ako ste već uključili nazive stupaca u prvi red vaših podataka, odaberite opciju "Pročitaj nazive varijabli iz prvog retka podataka" (pogledajte sliku 7.1), a zatim kliknite "Završi" da bi se podaci pojavili u proračunskoj tablici (IBM, 2021).

File Edit View Data Iransform Analyz	e Graphs Unities Egensions	Windo	v Bela			Range:	inhia annat i	ham first som	of data		
©pen Impot (jata	Database Excel GSV Data Text Data			ndde 0 of 0 vit	Variables	Percenta Percenta Ignore hi Remove	ige of values t dden rows an leading space trailing space	that determined columns es from string es from string	e data type: values	95	
See On+5     Spe Ac.     Spe Ac.     See A/Das     See A/Das     See Restar Part.     Evet     May Ere Rest Only     Rest to See Fig.	SAS Stata Cognos Business Intelligence Cognos T <u>M</u> 1					1 2 3 4 5 6 7 7	Anne Jacob Marcus Anne John Mana Simon Peter data type is ret 200 data a	Gender M F M F M based on all of based on all of based on all of	Age 21 34 24 38 42 19 25 fata and can fata and can	Workh	m the preview, which is based on 00 columns

#### Slika 4.29SPSS postavke uvoza podataka.

Sada kada imamo podatke u našoj proračunskoj tablici, primijetit ćete jasnu razliku u prezentaciji u usporedbi s Excelom. SPSS kategorizira podatke u dvije primarne vrste, svaka s

dvije dodatne podvrste. Kao što je prikazano na slici 7.2, podaci se mogu klasificirati kao numerički ili kategorički. Numerički podaci sastoje se od brojeva i mogu se kategorizirati kao diskretni (s ograničenim opcijama) ili kontinuirani (nude beskonačne mogućnosti). S druge strane, kategorički podaci



sastoje se od riječi i mogu se dalje razlikovati kao redni (imaju hijerarhiju) ili nominalni (bez hijerarhije). Ovisno o prirodi vaših podataka, možda ćete morati konfigurirati varijable kako bi se uskladile sa željenom analizom. U većini slučajeva, SPSS će automatski prikladno kategorizirati varijable. Pretpostavimo da želite izvršiti daljnju manipulaciju tipovima podataka. U tom slučaju možete pristupiti opciji "Prikaz" i pod "Prikaz varijable" prilagoditi varijabilne informacije kao što su naziv, vrsta, širina, mjera i više (IBM, 2021.).



At View Data Transform Analyze Graph Ubities Extensions Wiedow Jacob Atama A 21 120 A 1410 A 24 145 A 24 145 A 24 145 A 24 145 A 24 145 A 24 145 A 25 125 A 24 145 A 24 145 A 25 125 A 26 125 A 26 125 A 27 125 A 28 125 A	w Belp	Elle 1 2 3 4 5 6 7 8	Edit View Name Name Gender Age Workhours	Data Type String Numeric Numeric	Iransform Width 9 1 2 3	Analyze Decimals 0 0 0	Graphs L Label	Values None None	Missing None None	indow He Columns 9 10	Align	Q Search applica Measure	abon Ri
Image: Constraint of the second se	Var	1 2 3 4 5 6 7 8	Name Name Gender Age Workhouts	Type String String Numeric Numeric	Width 9 1 2 3	Decimals 0 0 0	Label	Values None None	Missing None None	Columns 9 10	Align	Q Search applics Measure	R
Jacob         Verthours         var         var         var           Atame         da Geröer         Age         Worthours         var         var         var           Atame         da Geröer         Worthours         var         var         var         var           Cruss         M         34         145               nn         M         38         125 <td< th=""><th>var</th><th>1 2 3 4 5 6 7 8</th><th>Name Name Gender Age Workhours</th><th>Type Seting String Numeric Numeric</th><th>Width 9 1 2 3</th><th>Decimals 0 0 0</th><th>Label</th><th>Values None None</th><th>Missing None None</th><th>Columns 9 10</th><th>Align</th><th>Measure &amp; Nominal</th><th>R</th></td<>	var	1 2 3 4 5 6 7 8	Name Name Gender Age Workhours	Type Seting String Numeric Numeric	Width 9 1 2 3	Decimals 0 0 0	Label	Values None None	Missing None None	Columns 9 10	Align	Measure & Nominal	R
A fame da cender Age Workhours var var var war var var var kasen af 200 M 34 145 nn M 34 145 nn M 34 155 non M 19 170 for M 25 125 anna F 23 50 sama F 20 150	Vär	1 2 3 4 5 6 7 8	Gender Age Workhours	String Numeric Numeric	1 2 3	0		None	None	10	an Let	Promisings	
Com         M         21         120           Com         M         34         145           Ana         P         24         145           Stat         P         42         155           Stat         P         42         155           Stat         P         25         125           Status         P         23         06           Status         P         20         150		3 4 5 6 7 8	Age Workhours	Numeric Numeric	2	0					m Let	Nominal	> ing
ICUS M 24 145 ne F 24 146 no M 30 125 Tab F 42 155 non M 19 170 Fer M 25 125 anna F 20 150 ristina F 20 150		4 5 6 7 8	Workhours	Numeric	3			None	None	12	Right	A Scale	> in
ne μ 22 4 415 No M 23 25 Ala F 42 155 Non M 19 170 Fer M 25 125 Anna F 23 06 Anna F 20 150		5 6 7 8				0	Work hours	None	None	12	I Right	/ Scale	> In
vi m 230 125 mon M 19 170 fer M 25 125 anna F 23 90 ristina F 20 150		6 7 8								-			_
nan P 42 100 hon M 10 100 ter M 25 125 nana P 23 80 nstra P 20 100		7 8											
non m r 75 ter M 25 125 anna F 23 90 ristina F 20 150		8											
even # 23 80 nsBra F 23 80 nsBra F 20 120													
restria F 20 100													
		10											
		.11											
		12											
		13											
		:14											
		15											
		15											
		17											
		18											
		19											
		20											
		21											
		22											
		23											
		24											
		25											
		26											
		27											
		28											
		29											
		30											
		31											
			6										

### Slika 4.30Prozori za prikaz podataka i varijabli.

Nakon što ispravno postavite svoje podatke, možete ih istraživati unutar SPSS-a. SPSS omogućuje korisnicima izvođenje temeljne statističke analize bez oslanjanja na unaprijed definirane funkcije. Na početnom zaslonu (pogledajte sliku 7.3), idite na "Analiziraj", nakon čega slijedi "Deskriptivna statistika", a zatim odaberite "Istraži". U odjeljku "Istraži" pronaći ćete različite opcije ovisno o karakteristikama podataka koje ste unijeli. U ovom načinu rada SPSS će vam pružiti informacije o "deskriptivnoj statistici" o vašim podacima. Iako je ovo vrijedno za početnu analizu podataka, ono nudi samo temeljne uvide i ne ulazi u detaljniju statističku analizu, koja će biti obrađena u narednim poglavljima. Prije nego što nastavimo dalje, također ćemo istražiti još jednu funkciju u SPSS-u — vizualizaciju grafikona (IBM, 2021).



ile	Edit Vie	ew Data	Iransform	Analyze Graphs Utilities Extensi Power Analysis	ons Window Help	Search applica	ition								
Na	me	Jacob		Reports	,									Visible:	4 of 4 Variab
	a Name	🚜 Gender	A Age	Descriptive Statistics	> Frequencies	var	var	var	var	var	var	var	var	var	var
2	Marcus	M		Bayesian Statistics	> Descriptives										
3	Anne	F		Tables											
4	John	м		Compare Means and Proportions	Population Descriptives										
5	Maria	F		General Linear Model	A Explore										
6	Simon	M		Generalized Linear Medels	Crosstabs										
7	Peter	M E		Mined Medele	TURF Analysis										
9	Christina	F		Mixed models	Ratio										
10				Correlate	Proportion Confidence Interv	als									
11				Regression	D D Diate										
12				Loglinear	> E F Flots										
13				Neural Networks	> 🛃 Q-Q Plots								1		
14				Classify	>	000	piore				^ -				
16				Dimension Reduction	>	-		_ [	Dependent List:		tatistics				
17		-		Scale	>	-	Name Gandar		& Age		Plots				
18				Nonparametric Tests	>	/	Work hours [Wor.				Options				
19				Forecasting	>				Eactor List:	B	ootstrap	Explore: St	atistics		×
20				Survival	>	_		\$				_			
21				- Multiple Response	>				Label Cases by			[ _ Descripti Confide	ince Interval for 1	Mean Inc	~
23				Missing Value Analysis				-	Laber Sases by.			Mastim	ators	1400. [35	~
24				Multiple Instantian		Dis	play				-	Outliers	actor a		
25				Complex Complex		۲	Both O Statistics					Percentil	les		
26				Complex Samples	·					1					
27				Simulation		_	OK	Paste	Reset Cancel	Help	-	Conti	nue Cancel	Help	
28				Quality Control	>										_
28	1			Spatial and Temporal Modeling	>										
	C]			Direct Marketing	>							Acti	vate Win	dows	>

### Slika 4.31Postavke deskriptivne statistike.

SPSS nudi niz opcija za vizualizaciju podataka, uključujući histograme, dijagrame pravokutnika, stupčaste grafikone, raspršene dijagrame, linijske grafikone, tortne grafikone i još mnogo toga. Do ove točke trebali biste imati osnovno razumijevanje o tome što svaka vrsta grafikona predstavlja i kako tumačiti rezultate koje oni pružaju. Stoga ćemo se usredotočiti na to kako izraditi te grafikone unutar softvera SPSS. Da biste izradili grafikone, odaberite karticu "Grafikoni" na početnom ekranu, nakon čega slijedi "Izrada grafikona". U novom prozoru možete odabrati vrstu grafikona koju želite izraditi i odabrati varijable koje želite uključiti. Nakon odabira "Završi", pojavit će se novi prozor s rezultatima vizualiziranim u odabranom formatu grafikona. U ovom novom prozoru možete aktivno komunicirati s grafikonom, što vam omogućuje izmjenu varijabilnih boja i fontova, istraživanje distribucija varijabli na grafikonu, i više (IBM, 2021).



		ew Data	Iransform Analyze	<u>G</u> raphs Utilities Extensions W	(Indone Date Chart Builder	view uses example data	Element Properties Chart Annexance Ontions	
: Nar	he	Jacob		Graphboard Template Chooser	A Name		Edit Properties of	Visible: 4 of 4 Variable
	& Name	& Gender	Age & Wor	Relationship Map	da Gender	hart count of Gender	Polar-Interval1	luar var
1	Jacob	M	21	H Weibull Plot	# Age		Angle-Axis1 (Polar-Interval1)	Tar Tan
2	Marcus	M	34	Compare Subgroups	Work hours [Wor		GroupColor (Polar-Interval1)	
3	Anne	F	24			5	Statistics	
4	John	м	38	Regression Variable Plots	Filter by	-	Vanable. Statistic	
5	Maria	F	42	1 Bar		Lund	Count	
6	Simon	М	19	11 3-D Bar	Category 1		Set Parameters	
7	Peter	м	25	I inc	Category 2	Set color?		
8	Joanna	F	23	Ene	-	Gender		
9	Christina	F	20	Area				
10		<u> </u>		Ne	Gallery Basic Elements Groups/Point ID	Titles/Footnotes		
11		-		High-Low	Choose from:	🏠 *Output2 (Document2) - IBM	SPSS Statistics Viewer	- 0 × -
14								
13				Bavalat	Favorites	Elle Edit View Data	Iransform Insert Format Analyze Graphs Utilities Extra	ensions Window Help
13 14		-		Boxplot	Bar Bar	Ele Est Yew Data	Janshom joset Fyrnat Analyze Graphs (Milles Exp 🖗 🛄 🕼 🐄 🖼 🧱 🚣 🗐 😒 📬	ensions Window Help
13 14 15				Boxplot Errgr Bar	Favorites Bar Line Area	Elle Edit View Data	Janston josef Fyrnat gradyce gradys (pilles Ep 🕘 🛄 🕼 🖘 🎬 🎇 🏪 🗐 😒 🧃	ensions Window Help
13 14 15 16				Bogplot Errgr Bar Population Pyramid	Favorites Bar Line Area Pie/Polar	Elle Edit View Data	Janston part Fyrnat grafys grafs (piltes Ep	ansions Wedow Belp
13 14 15 16 17				Boxplot Error Bar Population Pyramid Scatter/Dot	Favorites Bar Line Area Pin/Polar Seatter/Der	Ele Edit Xiew Data	Jantorn pret Fignat Analyze Graphs Ubless Ep	ensions Window Help
13 14 15 16 17 18				Bogplot Errgr Bar Population Pyramid Scatter/Dot Histocram	Parontes Bar Line Area PeripPotan Scatter/Ove Histogram High-Low	Ele Edit View Data Control Control Control Control Control Control Control Control Control Control Control Control Control Control Control Control Control Control Control Control Control Control Control Control Control Control Control Control Control Control Control Control Control Control Control Control Control Control Control Control Control Control Control Control Control Control Control Control Control Control Control Control Control Control Control Control Control Control Control Control Control Control Control Control Control Control Control Control Control Control Control Control Control Control Control Control Control Control Control Control Control Control Control Control Control Control Control Control Control Control Control Control Control Control Control Control Control Control Control Control Control Control Control Control Control Control Control Control Control Control Control Control Control Control Control Control Control Control Control Control Control Control Control Control Control Control Control Control Control Control Control Control Control Control Control Control Control Control Control Control Control Control Control Control Control Control Control Control Control Control Control Control Control Control Control Control Control Control Control Control Control Control Control Control Control Control Control Control Control Control Control Control Control Control Control Control Control Control Control Control Control Control Control Control Control Control Control Control Control Control Control Control Control Control Control Control Control Control Control Control Control Control Control Control Control Control Control Control Control Control Control Control Control Control Control Control Control Control Control Control Control Control Control Control Control Con	Jander past Fynnt drahyse Grafe Dillers for Diller The The The The States for Goraph Pis Chart Count of Gender	ensons Window Help
13 14 15 16 17 18 19				Boxplot  Grgr Bar  Constant Population Pyramid  Constant Population Pyramid  Figure Pystogram	Parontes Bar Une Area PeliPitian Gesete-Over Histogram High-Low Bocylot	Elle Est Ver Data Chart spirature Chart spirature B Sonne B Sonne Chart spirature Chart spirature Cha	Denterror prest Format doubles Copies Datase Datase	Gender
13 14 15 16 17 18 19 20				Boxplot Errgr Bar Population Pyramid Scatter/Dot Histogram	Farotes Bar Line Area Bouter-Dav High-Dow Bostat Dava Area	Elle Est Vev Des Charles application Biornal Biornal Biornal Biornal Biornal Biornal Biornal Biornal Biornal Biornal Biornal Biornal Biornal Biornal Biornal Biornal Biornal Biornal Biornal Biornal Biornal Biornal Biornal Biornal Biornal Biornal Biornal Biornal Biornal Biornal Biornal Biornal Biornal Biornal Biornal Biornal Biornal Biornal Biornal Biornal Biornal Biornal Biornal Biornal Biornal Biornal Biornal Biornal Biornal Biornal Biornal Biornal Biornal Biornal Biornal Biornal Biornal Biornal Biornal Biornal Biornal Biornal Biornal Biornal Biornal Biornal Biornal Biornal Biornal Biornal Biornal Biornal Biornal Biornal Biornal Biornal Biornal Biornal Biornal Biornal Biornal Biornal Biornal Biornal Biornal Biornal Biornal Biornal Biornal Biornal Biornal Biornal Biornal Biornal Biornal Biornal Biornal Biornal Biornal Biornal Biornal Biornal Biornal Biornal Biornal Biornal Biornal Biornal Biornal Biornal Biornal Biornal Biornal Biornal Biornal Biornal Biornal Biornal Biornal Biornal Biornal Biornal Biornal Biornal Biornal Biornal Biornal Biornal Biornal Biornal Biornal Biornal Biornal Biornal Biornal Biornal Biornal Biornal Biornal Biornal Biornal Biornal Biornal Biornal Biornal Biornal Biornal Biornal Biornal Biornal Biornal Biornal Biornal Biornal Biornal Biornal Biornal Biornal Biornal Biornal Biornal Biornal Biornal Biornal Biornal Biornal Biornal Biornal Biornal Biornal Biornal Biornal Biornal Biornal Biornal Biornal Biornal Biornal Biornal Biornal Biornal Biornal Biornal Biornal Biornal Biornal Biornal Biornal Biornal Biornal Biornal Biornal Biornal Biornal Biornal Biornal Biornal Biornal Biornal Biornal Biornal Biornal Biornal Biornal Biornal Biornal Biornal Biornal Biornal Biornal Biornal Biornal Biornal Biornal Biornal Biornal Biornal Biornal Biornal Biornal Biornal Biornal Biornal	Denter peet Format doubter could be to the second s	Gender
13 14 15 16 17 18 19 20 21				Bogplot       III Error Bar       III Population Pyramid       III Scatter/Dot       III Histogram	Barotes Barotes Line Area PietROdat Bento-Con Histopam Boylot Doublaces		Jankar Jank Fyrar (2007) Gold (2007)	Gender
13 14 15 16 17 18 19 20 21 22				Bogplot  Description  Population Pyramid  Scatter/Dot  Histogram	Bar Dar Ares Ares Personal Helipson Boolid Dad Ares CK Epuis E		Jankar Janet Fyner (2007) Greis (2008) (2008) Gersen (2007) Gorspin Pir Chart Count of Gender	Gender
13 14 15 16 17 18 19 20 21 22 23 24				Begglot Comparison of the second	Bartis Dan Aras Refinite Bartis Bartis Bartis Bartis Bartis Bartis Bartis Bartis Bartis Bartis Bartis Bartis Bartis Bartis Bartis Bartis Bartis Bartis Bartis Bartis Bartis Bartis Bartis Bartis Bartis Bartis Bartis Bartis Bartis Bartis Bartis Bartis Bartis Bartis Bartis Bartis Bartis Bartis Bartis Bartis Bartis Bartis Bartis Bartis Bartis Bartis Bartis Bartis Bartis Bartis Bartis Bartis Bartis Bartis Bartis Bartis Bartis Bartis Bartis Bartis Bartis Bartis Bartis Bartis Bartis Bartis Bartis Bartis Bartis Bartis Bartis Bartis Bartis Bartis Bartis Bartis Bartis Bartis Bartis Bartis Bartis Bartis Bartis Bartis Bartis Bartis Bartis Bartis Bartis Bartis Bartis Bartis Bartis Bartis Bartis Bartis Bartis Bartis Bartis Bartis Bartis Bartis Bartis Bartis Bartis Bartis Bartis Bartis Bartis Bartis Bartis Bartis Bartis Bartis Bartis Bartis Bartis Bartis Bartis Bartis Bartis Bartis Bartis Bartis Bartis Bartis Bartis Bartis Bartis Bartis Bartis Bartis Bartis Bartis Bartis Bartis Bartis Bartis Bartis Bartis Bartis Bartis Bartis Bartis Bartis Bartis Bartis Bartis Bartis Bartis Bartis Bartis Bartis Bartis Bartis Bartis Bartis Bartis Bartis Bartis Bartis Bartis Bartis Bartis Bartis Bartis Bartis Bartis Bartis Bartis Bartis Bartis Bartis Bartis Bartis Bartis Bartis Bartis Bartis Bartis Bartis Bartis Bartis Bartis Bartis Bartis Bartis Bartis Bartis Bartis Bartis Bartis Bartis Bartis Bartis Bartis Bartis Bartis Bartis Bartis Bartis Bartis Bartis Bartis Bartis Bartis Bartis Bartis Bartis Bartis Bartis Bartis Bartis Bartis Bartis Bartis Bartis Bartis Bartis Bartis Bartis Bartis Bartis Bartis Bartis Bartis Bartis Bartis Bartis Bartis Bartis Bartis Bartis Bartis Bartis Bartis Bartis Bartis Bartis Bartis Bartis Bartis Bartis Bartis Bartis Bartis Bartis Bartis Bartis Bartis Bartis Bartis Bartis Bartis Bartis Bartis Bartis Bartis Bartis Bartis Bartis Bartis Bartis Bartis Bartis Bartis Bartis Bartis Bartis Bartis Bartis Bartis Bartis Bartis Bartis Bartis Bartis Bartis Bartis Bartis Bartis Bartis Bartis Bartis Bartis Bartis Bartis Bartis Barti	Elle Entre Jeres Gran	Jankar Jank Fyrar (2007) Golds (2007)	Gender
13 14 15 16 17 18 19 20 21 22 23 24 23				Bogplot         Engr Bar         Population Pyramid         Scatter/Dot         Histogram	Bar Bar Bar Benne Benne Benne Hestyn Hestyn Befyl ow Bogh ow Doal Aves OK Eyste B	Chi Cancel Hot	Jankar Jank Frank (2007) Grafs (2008) (2008) Grafs (2007) Gorgen Pis Chart Count of Gender	Gender
13 14 15 16 17 18 19 20 21 22 23 24 25 26				Bogplot  Description  Population Pyramid  Scatter/Dot  Histogram	Bar Dar Aras Berthilde Heitiggen Begelar Dad Aras (K. Easts E	Erest Cancer Here	Jankar Janet Fyner (2007) Grefe (2007) Gerefe (2007) Gorsph Pie Chart Court of Gender	Gender
13 14 15 16 17 18 19 20 21 22 23 24 25 26 27				Begglot     EngrBar     Population Pyramid     ScatterDot     Histogram	Bar Bar Share Shar	Ethe Early Spee Gran	Jankar junt frank (varia Grafia Grafi	Gender
13 14 15 16 17 18 19 20 21 22 23 24 25 26 27 28				Bogplot  Constraints Description Descr	Bar Dar Ani Ani Hatayan Boptor Dod Asss OK. Pasis (B		Jankar Janet Franz (2007) Graha (2007) Garage (2007) Goragen Pis Chart Count of Gender	Gender
13 14 15 16 17 18 19 20 21 22 23 24 25 26 27 28 29				Begoptet Comparison of the second	Bar	Et En Les yers Here Charter Sentences Charter Sentences B Groups B Gr	Denker part France (2007) Grade (2007)	Gender

#### Slika 4.32Postavke izrade grafikona u SPSS-u.

Do ove točke pokrili smo tri od četiri pravila za "Istraživanje podataka", koja uključuju gledanje podataka (istraživanje neobrađenih podataka), identifikaciju podataka (određivanje tipova podataka) i, u određenoj mjeri, grafičko prikazivanje i opisivanje podataka putem deskriptivne statistike i izrade grafikona. Posljednje pravilo je "Formulacija pitanja", gdje se pitamo što želimo postići analizom podataka i u skladu s tim postavljamo grafikone i deskriptivnu statistiku kako bismo dobili odgovore na naša specifična pitanja. Na primjer, u našem trenutnom primjeru, pitanje bi moglo biti: "Je li naša analizirana populacija pretežno ženska?" Korištenjem i grafikona i deskriptivne statistike možemo zaključiti da se naša populacija sastoji uglavnom od muških osoba. Kada formulirate svoja pitanja, uvijek uzmite u obzir dostupne podataka i varijable koje ste identificirali (Garth, 2008). Ovime je završen prvi dio SPSS analize podataka, a sada ćemo nastaviti s pripremom testa.

# 7.2 Upravljanje podacima

Kada se bavite ključnim podacima u SPSS softveru, postaje ključno razumjeti tehnike za manipuliranje informacijama u pojedinačnim aktivnim skupovima podataka. SPSS pruža funkcionalnosti koje olakšavaju manipulaciju postojećim podacima sadržanim u aktivnim skupovima podataka. Povremeno možete naići na dvije baze podataka odvojeno uvezene u



skupove podataka, no prednost je da ih spojite radi poboljšane analize. Razmotrimo logističku tvrtku s dvije podružnice, od kojih svaka daje podatke o troškovima i prijevozu tereta u kilogramima. Cilj menadžera je analizirati ukupnu učinkovitost poduzeća. U SPSS-u to uključuje navigaciju "Podaci", odabir "Spoji datoteke" i dvije različite opcije. Jedan uključuje odabir "Slučajevi" i određivanje varijable za spajanje, uklanjanje te varijable dok spaja ostale. Alternativno, odabirom opcije "Varijabla" zadržava se varijabla u novom skupu podataka. Praktična primjena očita je u našem scenariju logistike, gdje spajanje skupova podataka pojednostavljuje sveobuhvatnu analizu učinka tvrtke.



### Slika 4.33Prozor za spajanje datoteka.

Dok funkcije spajanja i razdvajanja omogućuju određenu manipulaciju podacima, opcija "Odaberi slučajeve" nudi različite prednosti. Zamislite da imate podatke za prodavaonice B, C i D u jednoj bazi podataka, a fokus je isključivo na usporedbi prodavaonice A i prodavaonice C. Odabirom "Podataka" i "Odaberi slučajeve" možete odrediti varijable od interesa, učinkovito filtrirajući izbaciti neželjene podatke. Na primjer, postavljanje Prodavaonica C kao 2 upućuje softver da se koncentrira isključivo na Prodavaonicu C, generirajući izlaz koji je zatim dostupan za naknadne analize, kao što je deskriptivna statistika, fokusirajući se isključivo na odabrane slučajeve. Takav pristup također omogućuje komparativnu analizu samo između vrijednosti Prodavaonica A i Prodavaonica C.



Contitled	5 [DataSet4]	- IBM SPSS Statistics Data Editor									1.0	ntitled5	DataSet	4] - IBM 9	PSS Statistics D	ata Editor				
Ele Edit	View	Data Iransform Analyze Graphs	Utilities	Extensions	Window	Help		_			Eile	Edit	View	Data	Iransform	Analyze	Graphs	Utilities	Egter	nsions
		Define Variable Properties		A			٠ (	Q Sear	ch application			H			53	1 10	*==		n	111
15 :		3 Set Measurement Level for Unknown								_			-	-			00000 0			an
	10	Copy Data Properties		Var		var	var	var	107	Var			n		hoodractor		A Trac	action		
1	1	Atribute		900										9a 9	opconsactor		4 Trans	sacaron	000	
2	2	Bogfine date and time	ta s	olit File						×	2		-	5.4					400	-
3	3	Pane Multiple Response Sets									3		5	A					407	1
4	4	Validation	- 1	n.		O Ana	hore all o	ases do	not create grou	05	4		1	1 A					355	3
5	5	vajidation	2	Transaction		0.014	nyze an c	ases, uu	not create grou	Po .	5		12	2 A					329	1
7	0	Identify Dyplicate Cases		Transaction		O Cor	npare gro	ups			6		16	5 A					1200	1
8	8	Identify Unusual Cases				⊕ Qrg	anize out	put by gr	oups		7		17	7 A					100	F
9	9	R Compare Datasets					Groups	Based or	n:		8		20	AC					1000	1
10	10	B Sort Cases				-	Q. Sub	contracto			9		2.	3 A					604	-
11	11	Cash Mariahian				*	ad one	contracto		- 10	10		24	4 A					571	
12	12	Son vanages				-					11			5 A					438	-
13	13	Transpose				Sor	t the file l	v aroupir	o variables		12		-	0					400	-
14	14	Hadjust String Widths Across Files				OEla	is alward	u sodad			14			78					200	-
10	15	Merge Files				Office	is areas	y soneu			15		-	8.8					900	1
17	17	Restructure	Cur	rrent Status: An	alysis	by groups	s is off.				16		15	5.8					430	1
18	18	Descentity Press Metabing									17		18	8 8					500	1
19	19	Propensity Score watching									18		19	9 B					1500	1
20	20	Rake Weights		OK	P	R	eset C	ancel	Help		19		22	2 B					657	1
21	21	Propensity Score Matching		Gitt	100				(the p		20		25	58					538	J
22	22	Case Control Matching		657							21		25	98					405	5
23	23	R Agregate		604							22		30	D B					372	t
24	24	Cyprogeter.		571							23			3 C					500	f
20	25	Split into Files		538							24		4	4 C					300	1
20	20	Orthogonal Design	>	504							25		10	D C					381	1
20	20	Compare Datasets		439							26		1	3 C					302	1
29	29	R. Conv Dataset		405							27		14	4 C					650	-
**		The state of the s	_								28		2	10					200	-
		Dut Fre									29	-		10					174	-
Data Vier	w Variabl	Select Cases									_	<								
Split File		4 Weight Cases									Dat	a View	Variat	de View						

Slika 4.34Prozor za dijeljenje datoteke.

Dok funkcije spajanja i razdvajanja omogućuju određene manipulacije podacima, postoji i opcija "Odaberi slučajeve". Zamislite da pouzdano znamo da prodavaonica A ima u prosjeku 120 € dobiti i želimo to usporediti s prodavaonicom C. Nažalost, u našoj bazi podataka imamo podatke za prodavaonice B, C i D u jednoj bazi podataka i analiza bi uključivala podatke iz sve tri prodavaonice. Klikom na "Podaci" i "Odaberi slučajeve" možemo odabrati koju varijablu želimo fokusirati. U našim smo slučajevima postavili da prodavaonica C treba biti postavljena kao 2, a zatim smo stvorili funkciju za softver da se fokusira samo na prodavaonicu C. Izlaz se zatim može koristiti za naknadnu analizu odabirom ovog novog stupca (npr. deskriptivna statistika).



Slika 4.350dabir slučaja.



Povremeno skupovi podataka mogu već sadržavati varijable, ali ipak postoji potreba za uvođenjem novih varijabli na temelju postojećih. Uzmimo, na primjer, menadžera logističke tvrtke koji posjeduje podatke o težini i prijeđenoj udaljenosti za razne proizvode, ali zahtijeva vrijeme isporuke za optimizaciju ruta. U SPSS-u, postizanje toga uključuje klik na "Transform", a zatim na "Compute Variables". Nova varijabla, DeliveryTime, stvara se unutar novog prozora postavljanjem numeričkih izraza. U ovom slučaju, dodjeljivanje ljestvice od 0,8 za udaljenost i 0,2 za težinu rezultira novom varijablom koja predstavlja vrijeme isporuke, što je ključni dodatak skupu podataka. Postoji fleksibilnost izračunavanja dodatnih varijabli, kreiranih za potrebe statističkih testova.

1 ·U	ntitled2 [DataSet]	1] - IBM S	PSS Statistics Data Editor									1 - C	Untitled2	[DataSet1]	] - IBM SPSS Stati	stics Data Editor					
Ele	Edit View	Data	Iransform Analyze Grap	ohs Utilities Extens	sions Wi	ndow H	elp					Eile	Edit	View	Data Iransfe	orm Analyze	Graphs Utilities	Extensi	ons Window	Help	
-	日合		Compute Variable					. Q Se	arch applicati	on						-	±	44			
-		-	Programmability Transform	nation	ELLER 1									0						[1]ed	
		11	Count Values within Cases	s										o ductio	A Maight	Ø Distance	A Daluan Time				
	Producau	-	Shift Values		var	var	A34	var	var	var	Var		at Pl	100000	Vveight 20	Distance     26	22.00	ABL	var.	var	Var
2	2		Recode into Same Variabl	las								2		2	30	45	42,00				
3	3		Contraction Date value									3		3	40	30	32,00				
4	4		Becode into Different Vana	acres								4		4	20	20	20,00				
5	5		Automatic Recode									5		5	30	30	30,00				
6	6		Create Dummy Variables									6		6	40	55	52,00				
8	8		H Visual Binning									8		8	10	41	34.80				
9	9		Coptimal Binning									9		9	20	100	84,00				
10	10	)	Prepare Data for Modelin	Compute Variable									×	10	60	30	36,00				
11			M Rank Cases																		
12			gej nang cases	Target Variable:		Numeric	Expression														
13			Date and Time Wizard	DeliveryTime		(Weight)	*0.2) + (Dista	ince*0.8)													
15			Create Time Series	Type & Label																	
16			Replace Missing Values.	Libbe of Faces																	
17			GRandom Number General	& ProductID																	
18			Run Pending Transforms	Weight Distance						Funct	tion group:										
19				A DelveryTime		+	< >	7 8	9	All			^								
20				· Sentery mile		H	88			Arith	R Noncastro	N CDE									
22							<= >=	4 0	6	Conv	ersion										
23								1 2	3	Curre	ent Date/Tim	10									
24						F			iFi	Date	Arithmetic										
25										Date	Creation										
26						-	~ ()	Delet	e 1	P Eunci	ions and Sp	ecial val	nables:								
2/						-				SDat	e		iii -								
29										\$Dat	e11										
30										SJDa	te										
31										SSys	e										
32										Abs											
33										Any											
34										Appl	ymodel		~								
36										Film	he Direch	da darci	riction								
37				(optional case s	election co	ndition)				Fiter	ox. Clinca	INE JUSC	mpandi								
38																					
39						OK	Paste	Beset Ca	ncel Help	3											
40	. 22						1		1				1.00								
-	¢											-	<	_							
Data	View Variab	le View										Dat	a View	Variable	e View						
-												_									

### Slika 4.36Postupak izračunavanja varijabli.

Ovo zaključuje mali pregled funkcija upravljanja podacima koje pokriva SPSS, a koje bi mogle biti korisne tijekom sljedećih testova modela koji su obuhvaćeni u ovom poglavlju. Nastavit ćemo s fazama koje su potrebne prije nego što možemo provesti statistički test u softveru SPSS.

# 7.3 Priprema testa

Prije nego što nastavite sa statističkim testovima, bitno je pridržavati se standardnog tijeka procesa analize podataka, koji uključuje istraživanje podataka (kao što je objašnjeno u



poglavljima 7.1 i 7.2), analizu podataka i interpretaciju rezultata (Garth, 2008; George i Mallery, 2022). U ovom poglavlju naš fokus je na analizi podataka pomoću softvera SPSS. Budući da smo hipoteze već obradili u prethodnim poglavljima, naš primarni fokus bit će na provođenju testova normalnosti unutar SPSS-a. Postoje tri metode za procjenu normalnosti: histogram, QQ-grafikon i test normalnosti. Preporučljivo je upotrijebiti najmanje dvije, ako ne i sve tri ove opcije, budući da svaka pruža različite informacije (Ghasemi i Zadesiasl, 2012.). Za izradu histograma idite na " Graphs ", a zatim na "Chart Builder". U novom prozoru odaberite "Histogram". Ako imate više varijabli, morate ponoviti ovaj postupak za svaku kako biste dobili rezultate. Histogram potvrđuje test za normalnu distribuciju ako stupci koji predstavljaju varijable vrijednosti nalikuju zvonolikoj krivulji. Ako su stupci više nagnuti u lijevu ili desnu stranu, to može značiti eksponencijalnu distribuciju. Na primjer, generirali smo bazu podataka od 100 ID-ova, svaki s varijablom koja predstavlja težinu u kilogramima. Prateći upute izradili smo histogram, kao što je prikazano na slici 7.9. Kao što je vidljivo sa slike, stupci su raspoređeni po grafikonu i iako možda ne odražavaju savršeno krivulju, ipak sugeriraju normalnu distribuciju i pozitivan rezultat testa (George i Mallery, 2022; Goeman i Solari, 2021).



Slika 4.37Histogram rezultata testa normalnosti.

Još jedna opcija za provođenje testova normalnosti je QQ-grafikon, koji se može pokrenuti klikom na "Analyze" (hrv. Analiziraj), nakon čega slijedi "Descriptive Statistics", a zatim odabirom "Q-Q Plots". Prednost ovog pristupa je što omogućuje procjenu više varijabli istovremeno (Williamson, bd). Test se smatra uspješnim kada se točke na dijagramu grupiraju usko oko ravne crte, što predstavlja normalnu distribuciju. Ako točke formiraju "repove", to ukazuje na neuspješan test normalnosti (Andersen i Dennison, 2018). Koristeći istu bazu



podataka iz testa histogramskog grafikona, proveli smo QQ grafikon test. Na slici 7.10 u nastavku možete primijetiti da je većina točaka klastera za našu varijablu poravnata s ravnom linijom, što ukazuje na normalnu distribuciju naših podataka. Iako smo već u ovoj fazi mogli zaključiti da je test normalnosti pozitivan, odlučili smo tražiti potvrdu iz sva tri testa.



### Slika 4.38QQ grafikon test normalnosti - postavke i rezultati.

Konačna opcija za provođenje testa normalnosti je takozvani Test normalnosti, koji se smatra statističkim testom. Obično se koristi Kolmogorov-Smirnov test, ali za male veličine uzorka može se koristiti Shapiro-Wilkov test (Goeaman i Solari, 2021.). U SPSS-u možete izvršiti ovaj

test klikom na "Analyze", nakon čega slijedi "Descriptive Statistics", a zatim "Explore". Morate postaviti varijable koje želite provjeriti ispod okvira "Dependent List" (hrv. Zavisna lista). Zatim pod "Plots" odaberite "Normality Plots with Tests" (hrv. Grafike normalnosti s testovima). Test se smatra uspješnim

ako je stupac Sig (p -vrijednost) u rezultatima veći od 0,05, što ukazuje na normalnu distribuciju. Ako je p -vrijednost manja od 0,05, to ukazuje da distribucija nije normalna i test se smatra neuspješnim. Ovaj smo test još jednom proveli koristeći istu bazu podataka kao i u prethodnim testovima. Iz rezultata možemo zaključiti da je prema standardu Kolmogorov-Smirnov test pozitivan jer je p -vrijednost veća od 0,05. Međutim, za Shapiro-Wilkov test, p - vrijednost je niža, što ukazuje na negativan rezultat testa. Do ovih različitih rezultata dolazi jer



oba pristupa imaju različite postavke osjetljivosti i snagu u otkrivanju odstupanja (Ghasemi i Zahediasl, 2012.). Budući da smo već proveli testove QQ dijagrama i histogramskog grafikona, Test normalnosti može se općenito smatrati pozitivnim. Uz potvrđene testove normalnosti, možemo provesti glavne testove, kao što je test jednog uzorka.



# Slika 4.39Postavke i rezultati testa normalnosti.

# 7.4 T-test jednog uzorka

Već ste pokrili teoriju iza T-testa jednog uzorka u prethodnim poglavljima; stoga ćemo se prvenstveno usredotočiti na provođenje testa sa softverom SPSS. Za naš T-test jednog uzorka pripremili smo bazu podataka s uzorkom od 200 ispitanika, koji uključuje 1 kategoričku varijablu (ID studenta) i 2 numeričke varijable (težinu i dob) (Kim, 2015.). Prateći upute iz prethodnih potpoglavlja provodimo sljedeće korake:

- Istražite podatke, točnije naše varijable i deskriptivnu statistiku i postavite naše pitanje.
- Provjerite normalnost, budući da bi samo jedan varijabilni histogram i QQ grafikon trebali biti dovoljni.
- Postavite hipotezu, gdje se za nultu varijabla ne razlikuje od određene vrijednosti i alternativu gdje je drugačija.



# • Provedite Studentov T-test .

• Tumačite rezultate, fokusirajući se na to je lit **nulta hipoteza odbijena** ili **ne**, odgovorite na pitanje i napišite izvješće o našem testu.

U našem slučaju odlučili smo da naše pitanje bude: Je li prosječna težina učenika veća od 74 kilograma? Nakon pitanja postavljamo našu hipotezu za pitanje, a to je "Nulta = nema razlike" i "Alternativa = postoji razlika". Proveli smo histogram i QQ grafikone kako bismo provjerili testove normalnosti, a nakon njihovog završetka, slijedio je T-test. Da bismo pokrenuli T-test, kliknemo "Analyze" i nastavimo s "Compare Means" (hrv. Usporedi srednje vrijednosti) i "One-Sample T-test" (hrv. T-test jednog uzorka). U okvir s varijablama testa stavljamo studentski ID, postavljamo vrijednost testa na 74 i započinjemo test (pogledajte sliku 7.12).

Elle	Edit View	Data Transform	Analyze Graphs Utilities Extens	ions Wind	low H	slp											
		🛄 🖛 🗅	Power Analysis	> III		۰ 💽	Q Sear	ch applicatio	0								
88 :			niteta Analysis													Visible:	3 of 3 Variable
	StudentD	Neight 🖉	Reports		var	Var	var	var	var	var	Var	var	VBr .	var	Var	Var	197
181	181	78	Descriptive Statistics														
182	182	64	Bayesian Statistics	'													
183	163	64	Tables	>													
184	184	63	Compare Means and Proportions	> 🖬 M	eans												
186	185	69	General Linear Model	· 🖬 o	ne-Sampl	e T Test		-									
187	187	80	Generalized Linear Models	> 131 In	lenender	t-Samples T Te	-1	_									
188	188	72	Mixed Models	,	repender	Coampies 1 14											
189	189	60	Correlate	2 51	immary I	ndependent-Sa	mples T Te	st									
190	190	71	Permanetion	, BB	aired-San	ples T Test											
191	191	62	Degression		te-Way A	NOVA											
192	192	79	Loginear	1 10	te-Samo	e Proportions											
193	193	69	Neural Networks	· ·													
194	194	64	Classify	> Es in	sepender	t-Samples Pilop	sortions										
100	195	61	Dimension Reduction	> SP	ired-San	nples Proportion	15 🛃	One-Samp	le T Test								×
197	197	71	Scale	>			_										
198	198	70	Nonparametric Tests	>								Test	Variable(s):			0	otions
199	199	77	Forecasting	>				Student	ID [Student]	D]		11	Neight				
200	200	64	Superal	,				Age								Be	otstrap
201			Multiple Deserves														
			Multiple Response	-													
20.3			Missing Value Analysis														
204			Multiple Imputation	>								-					
206			Complex Samples	>													
207			Simulation														
208			Quality Control	>													
209			Spatial and Temporal Modeling														
	6		Direct Marketing														
Data	View Variable	le View	Desce wardened							_	_	Test	Value:	<u>⊡ E</u> st	limate effect	sizes	
-	-										OK Ea	ste Re	set Cance	H Help			

# Slika 4.40Postavke T-testa jednog uzorka.

Nakon potvrde testa, pojavit će se drugi prozor s rezultatima naše analize (pogledajte sliku 7.13). Ovaj prozor pruža nekoliko informacija u vezi s našom analizom. U ovom su slučaju obje p-vrijednosti niže od 0,05, što ukazuje na značajnost testa. Dodatno, provjeravamo vrijednosti t i df, koje su u našem slučaju -9,806 odnosno 199. Iz ovih rezultata možemo zaključiti da je naša nulta hipoteza odbačena. Stoga je cjelovito izvješće o rezultatima sljedeće: "Prosječna težina studenta značajno je niža (srednja vrijednost = 69,63) od vrijednosti od 74 kg (t-test jednog uzorka, t = -9,806, df = 199, p-vrijednost < 0,001)".







# 7.5 Korelacija

Prijeđimo sada na drugi test, a to je test korelacije. Provest ćemo ga koristeći istu bazu podataka kao u primjeru t-testa s jednim uzorkom. Slično t-testu s jednim uzorkom, slijedit ćemo postupak uz nekoliko izmjena. Kada se vrši korelacija između dvije varijable, važno je odrediti koja je zavisna, a koja nezavisna varijabla (Janse i dr., 2021.; Mishra i dr., 2019). Ovaj odabir možete napraviti na temelju vašeg istraživačkog pitanja. U našem slučaju želimo istražiti "Postoji li korelacija između dobi studenta i njegove težine?". Nakon pitanja, težinu smatramo zavisnom varijablom, a dob nezavisnom varijablom, jer želimo istražiti jesu li varijacije u dobi povezane s varijacijama u težini. Definiramo naše nulte i alternativne hipoteze (vidi 7.3 i 7.4),

a zatim pokrećemo test klikom na "Analyze", nakon čega slijede "Correlate" (hrv. Koreliraj) i "Bivariate" (hrv. Bivarijantno). Obje varijable treba staviti u polje "Variable". Provjerite jesu li odabrani ili postavljeni "Pearson", "Two-Tailed" i "Flag Significant" (pogledajte sliku 7.14). U ovom smo slučaju odabrali



"Pearson" jer naši podaci pokazuju normalnu distribuciju i mogu se analizirati pomoću parametarskih metoda. Ako normalna distribucija nije naznačena, treba koristiti neparametarske metode (u ovom slučaju, odabrali biste Spearmana umjesto Pearsona) (George i Mallery, 2022; McClure, 2005).





Slika 4.42Postavke testa korelacije.

Još jednom dobivamo rezultate u novom prozoru (pogledajte sliku 7.15). Iz rezultata možemo vidjeti da je naša Pearsonova korelacija -0,038, a p -vrijednost 0,596. U korelacijskoj analizi, što je vrijednost korelacije bliža nuli, to je korelacija između varijabli slabija. U našem slučaju, korelacija je vrlo blizu nule, što ukazuje da nema značajne korelacije između dvije varijable (McClure, 2005). Dodatno, visoka p -vrijednost (0,596) sugerira da nema značajnih dokaza za zaključak da postoji značajna korelacija između dviju odabranih varijabli (Williamson, bd). Kao rezultat toga, naša nulta hipoteza nije odbačena. Na temelju toga možemo izvijestiti da "nije bilo korelacije između dobi i težine studenta".

	laons		
	Correlatio	ons	
		Weight	Age
Weight	Pearson Correlation	1	-,038
	Sig. (2-tailed)		,596
	N	200	200
Age	Pearson Correlation	-,038	1
	Sig. (2-tailed)	,596	
	N	200	200

Slika 4.43Rezultati testa korelacije.

# 7.6 Hi-kvadrat

Treći test koji ćemo izvesti u SPSS softveru je Hi-kvadrat test. Za razliku od prethodna dva testa, Hi-kvadrat test uspoređuje dvije kategoričke varijable, a ne numeričke varijable (Turhan, 2020). Kao i postupak u odjeljcima 7.4 i 7.5, počinjemo istraživanjem podataka i formuliranjem



istraživačkog pitanja. U našem primjeru imamo logističku tvrtku s 200 kupaca, te imamo podatke o vrsti plaćanja i vrsti prijevoza koju je svaki kupac odabrao. Pitanje na koje želimo odgovoriti je: "Pokazuju li različite vrste plaćanja različite preferencije za vrste prijevoza?" Budući da se radi samo o kategoričkim varijablama, nema potrebe za testom normalnosti. Postavljamo našu nultu hipotezu (preferencije za vrste prijevoza iste su za sve vrste plaćanja) i alternativnu hipotezu. Za provođenje hi-kvadrat analize kliknite na "Analyze", nakon čega slijedi "Descriptive Statistics", i odaberite "Crosstabs" (hrv. Unakrsne analize). Ključno je smjestiti varijable na temelju vašeg istraživačkog pitanja u okvir stupca ili retka (pogledajte sliku 7.16) (Garth, 2008.).



#### Slika 4.44Postavke Hi-kvadrat testa.

Nakon analize, novi prozor prikazuje rezultate (pogledajte sliku 7.17). U ovom prozoru možete primijetiti da Pearsonova hi-kvadrat vrijednost iznosi 11,614, *df* vrijednost 12, a *p* -vrijednost (asimptotska značajnost) 0,477. Na temelju ovih rezultata možemo zaključiti da ne postoji značajna povezanost između dviju varijabli, a nulta hipoteza nije odbačena. Stoga izvješće slijedi: "Nema otkrivenih značajnih preferencija između različitih vrsta plaćanja za različite vrste prijevoza (dvostrani Chi-Square test, chi-sq = 11,614, *df* = 12, *p* -vrijednost = 0,477)."



+ Crosstabs



#### Slika 4.45Rezultati Hi-kvadrat testa.

# **7.7 ANOVA**

Posljednji test koji ćemo pokriti je ANOVA test, koji se posebno fokusira na jednostavniji model poznat kao jednosmjerna ANOVA, koji uključuje kategoričku varijablu i numeričku varijablu (Goeman i Solari, 2021). Kao i kod T-testa, slijedit ćemo isti postupak: istražiti podatke, formulirati istraživačko pitanje, provesti test normalnosti i postaviti hipoteze. Razmotrimo studiju slučaja transportnog dispečera koji radi za logističku tvrtku. Dispečer blisko surađuje s partnerskom tvrtkom i redovito planira tri različite rute kamionima za isporuku robe. Zbog politike "Just-in-time" koja naglašava brže isporuke, postavlja se pitanje: "Utječe li izbor rute dostave na vrijeme isporuke za tvrtku?" Da biste pokrenuli ANOVA test u SPSS-u, idite na "Analyze" nakon čega slijedi "Compare Means..." i zatim "One-way ANOVA".

Postavite zavisnu varijablu u okvir "Dependent List" (hrv. Popis zavisnih), a varijablu Faktor u okvir "Factor" (pogledajte sliku 7.18). Za temeljitu analizu uključili smo i Post Hoc postavku. Važno je napomenuti da se Post Hoc analiza



treba provesti samo ako je početni ANOVA test pozitivan. Primjenom Post Hoc analize možemo identificirati optimalan izbor (u našem slučaju rutu). Najpouzdanije metode koje se koriste za Post Hoc analizu su ili Bonferronijeva korekcija ili Tukeyjeva HSD metoda (Goeman i Solari, 2021.).



Ea	nt view vata Iransform	Power Analyze Graphics Extension		_		
	- 🖶 🛄 🖛 1	Meta Analysis	Search application	www.abictiva		×
		Decede				le: 2 of 2 Variabl
	💦 Route 🛷 DeliveryTir	ne	zar var var		Dependent List:	Cogtrasts var va
	1	2 Descriptive Statistics	·		Delivery Time [DeliveryTime]	Post Hoc
	2	2 Bayesian Statistics	>			Ontions
	3	2 Tables	>			Spinner
	1	<sup>2</sup> Compare Means and Proportions	Means		*	Bootstrap
	2	General Linear Model	> Dna-Samola T Tast			
	1	2 Generalized Linear Models				
	2	2 Mixed Models	Mindependent-Samples Tiest			
	3	2 Correlate	Summary Independent-Samples T Test			
	1	2	Paired-Samples T Test		Eactor.	
	2	2 Regression	One-Way ANOVA		Koute	
	3	2 Loglinear	2 One Sample Proportions		Estimate effect size for overall tests	
	1	2 Neural Networks	> Cone-Sample Proportions		OK Paste Reset Cancel Help	
	2	2 Classify	Independent-Samples Proportions			
	3	2 Dimension Reduction	> Same Paired-Samples Proportions		One-Way ANOVA: Post Hoc Multiple Comparisons	×
	2	2 Scale	>		Frual Variances Assumed	
	3	2 Nonparametric Tests	>		LSD S-N-K Waller-Duncan	
	1	2 Executing			Bonferroni Tukey Type I/Type II Em	or Ratio: 100
	2	2 Sustal			Sidak Tukey's-b Dunnett	
	3	2			Scheffe Duncan Control Category	Track .
	1	2 Multiple Response	5		D.E.G.W.E DHochbards GT2 Test	and an
	2	2 Missing Value Analysis			DECWO Cobriel @2-sided @c	Control On Control
	3	2 Multiple Imputation	>		Live out 7 Diamigi	odiana A - colliga
	1	Complex Samples	>		Equal Variances Not Assumed	
-	3	2 B Simulation			Tamhane's T2 Dunnett's T3 Games-Howell	Dynnett's C
	1	2 Quality Control				
	2	2 Costial and Taxanal Med 1			Null Hypothesis test	
<	•	Spatial and remporal Modeling	·		O Use the same significance level [algha] as the setting in C	Options
		Direct Marketing	· ···		O Specify the significance level [alpha] for the post hoc test	·
ta Vie	w Variable View				Legel: 0.05	

#### Slika 4.46Postavke za ANOVA analizu.

Rezultati naše analize pokazuju da je naša F-statistička vrijednost 11,173 (više vrijednosti ukazuju na više varijacija između skupina) i p-vrijednost <0,001, što znači da je naša nulta hipoteza odbačena (vidi sliku 7.19). Budući da postoji značajna razlika između tri rute (<0,001), post hoc test je također valjan u našem slučaju (George & Mallery, 2022). Nakon provođenja Bonferronijevog testa korekcije, možemo vidjeti da su najbolje p-vrijednosti zabilježene u slučaju rute 2 (pogledajte sliku 7.19). U izvješću možemo zaključiti da je "postojala značajna razlika u odabiru rute isporuke u korelaciji s vremenom isporuke (1-way ANOVA, F = 11,173, df = 47, p-vrijednost = <0,001). Ruta 2 imala je najbolje rezultate u vremenu isporuke."

						Post Ho	c Tests					
								Mult	iple Comp	arisons		
						Depende	nt Variable:	Delivery Time				
						Bonferron	ni				05% Confid	aneo Intorval
		ANOVA				(I) Route	(J) Route	Mean Difference (I-J)	Std. Error	Sig.	Lower Bound	Upper Bound
Delivery Time						1	2	-1,4250	,3040	<,001	-2,181	-,669
	Sum of						3	-,5500	,3040	,231	-1,306	,206
	Squares	df	Mean Square	F	Sig.	2	1	1,4250	,3040	<,001	,669	2,181
Between Groups	16,527	2	8,263	11,173	<,001		3	,8750	,3040	,018	,119	1,631
Nithin Groups	33,280	45	,740			3	1	,5500	,3040	,231	-,206	1,306
Cotal	49 807	47					2	- 8750	3040	.018	-1.631	- 119

#### Slika 4.47Početni rezultati ANOVA analize i rezultati post hoc testa.



Zaključujemo ovo poglavlje knjige uz razumijevanje da smo u ovom poglavlju pokrili neke od uobičajenih testova. Postoje i drugi testovi, kao što je ANOVA ponovljenih mjerenja, testovi pouzdanosti i testovi osjetljivosti, koji se također mogu modelirati i analizirati pomoću softvera SPSS. Ovi dodatni testovi pružaju širi raspon alata za analizu podataka i daju smislene uvide u različita istraživanja i praktične primjene.

# Literatura 7. poglavlja

- Andersen, A.J. & Dennison, J.R. (2018). An Introduction to Quantile-Quantile Plots for the Experimental Physicist. Journal Articles, 51.
- Garth, A. (2008). Analysing data using SPSS [available at: https://students.shu.ac.uk/lits/it/documents/pdf/analysing\_data\_using\_spss.pdf, access October 26, 2023]
- George, D. & Mallery, P. (2022). IBM SPSS Statistics 27 Stet by Step: A Simple Guide and Reference, 17<sup>TH</sup> edition, Abingdon: Routledge
- Ghasemi, A. & Zahediasl, S. (2012). Normality Tests for Statistical Analysis: A Guide for Non-Statisticians. International Journal of Endocrinology and Metabolism, 10(2), pp. 486-489.
- Goeman, J.J. & Solari, A. (2021). Comparing Three Groups. The American Statistician, 76(2), pp. 168-176
- IBM (2021). IBM SPSS Statistics 28 Brief [available at: https://www.ibm.com/docs/en/SSLVMB\_28.0.0/pdf/IBM\_SPSS\_Statistics\_Brief\_Guide. pdf, access October 26, 2023]
- Janse, R.J., Hoekstra, T., Jager, K.J., Zoccali, C., Tripepi, G., Dekker, F.W. & van Diepen, M. (2021). Conducting correlation analysis: important limitations and pitfalls, 14(11), pp. 2332-2337.
- Kim, T.K. (2015). T test as a parametric statistic. Korean Journal of Anesthesiology, 68(6), pp. 540-546
- Landau, S. & Everitt, B.S. (2004). A Handbook of Statistical Analyses using SPSS, 1<sup>st</sup> edition, London: Chapman & Hall/CRC
- McClure, P. (2005). Correlation Statistics Review of the Basics and Some Common Pitfalls. Journal of Hand Therapy, 18(3), pp. 378-380



- Mishra, P., Singh, U., Pandey, C.M., Mishra, P. & Pandey, G. (2019). Application of Student's t-test, Analysis of Variance, and Covariance. Annals of Cardiac Anesthesia, 22(4), pp. 407-411
- Turhan, N.S. (2020). Karl Pearson's chi-square tests. Educational Research and Reviews, 15(9), pp. 575-580
- Williamson, M. (b.d.). Data Analysis using SPSS [available at: https://med.und.edu/research/daccota/\_files/pdfs/berdc\_resource\_pdfs/data\_analysis \_using\_spss.pdf, access October 26, 2023]