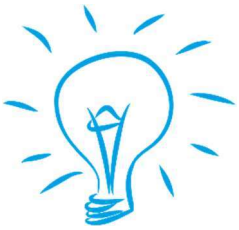
7. Statistička obrada podataka SPSS

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

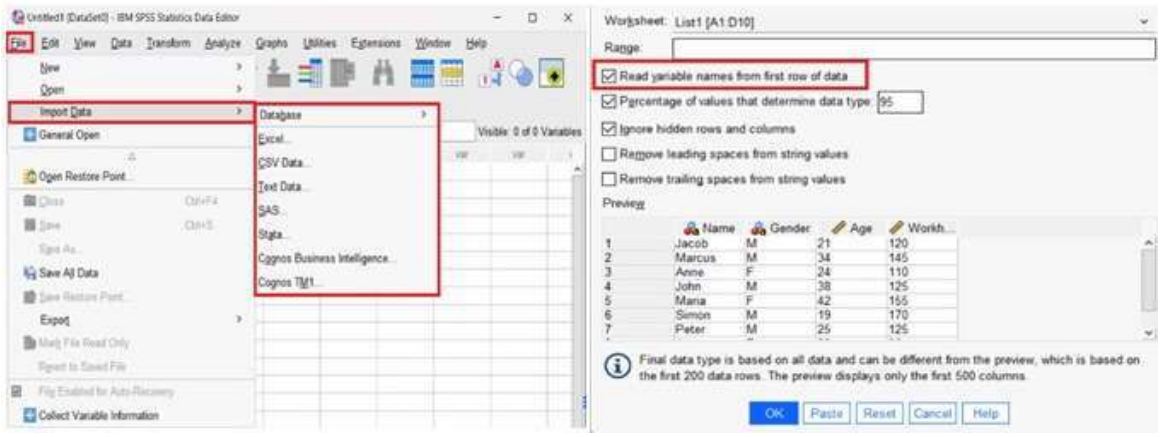


poboljšanje vaših napora optimizacije, provođenje analize i prepoznavanje potencijalnih poboljšanja, možda ste primijetili da kako složenost analiziranih podataka i izračuna raste, tradicionalni pristupi mogu postati sve zamršeniji i izazovniji za računanje. Kako zamršenost vaših podataka i izračuna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

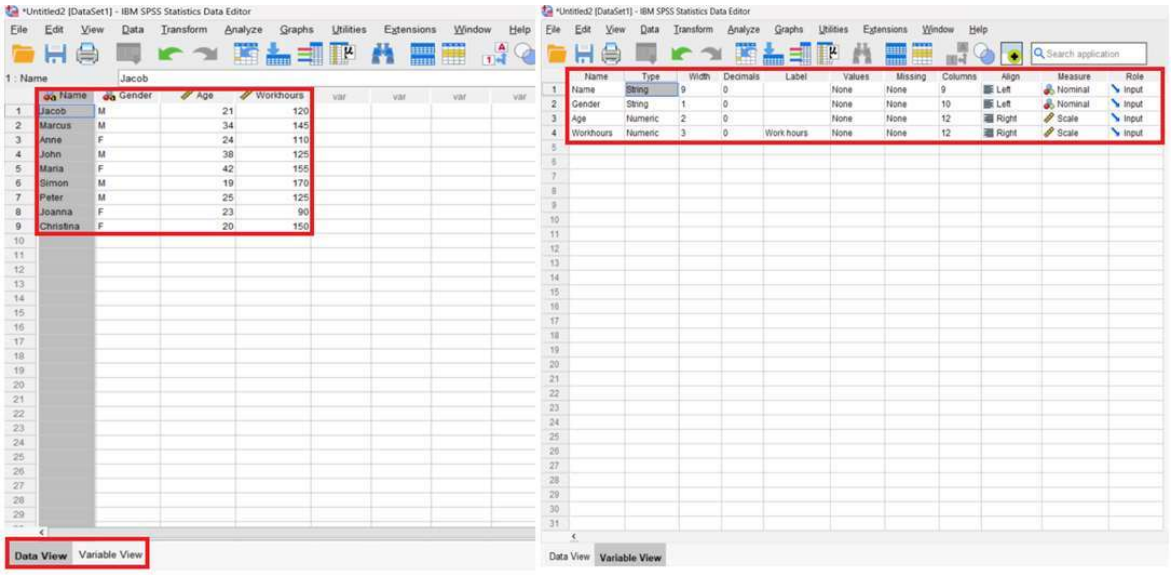
Možda ste već imali iskustva sa SPSS softverom. Međutim, ako ipak trebate, dopustite da vam ponudimo kratki uvod. SPSS, poput svog općepriznatijeg pandana Excela, olakšava manipulaciju podacima, analizu i vizualizaciju. Ipak, za razliku od Excela, koji ponekad može biti naporan i složen u programiranju 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).



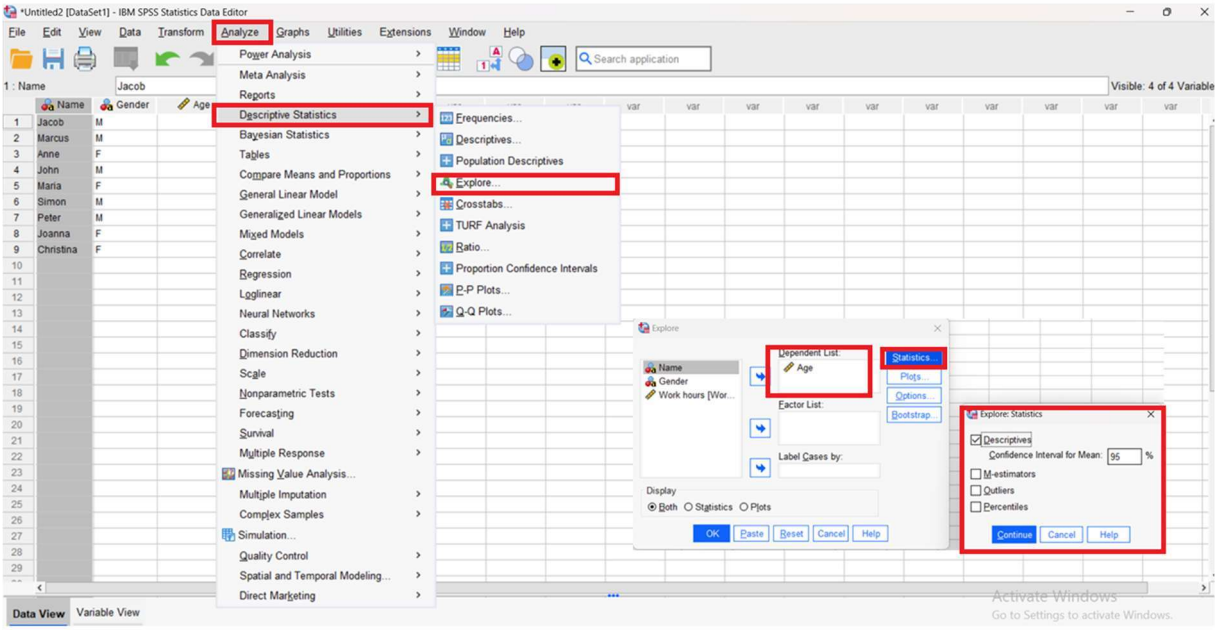
Slika 4.29SPSS postavke uvoza podataka.

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.).



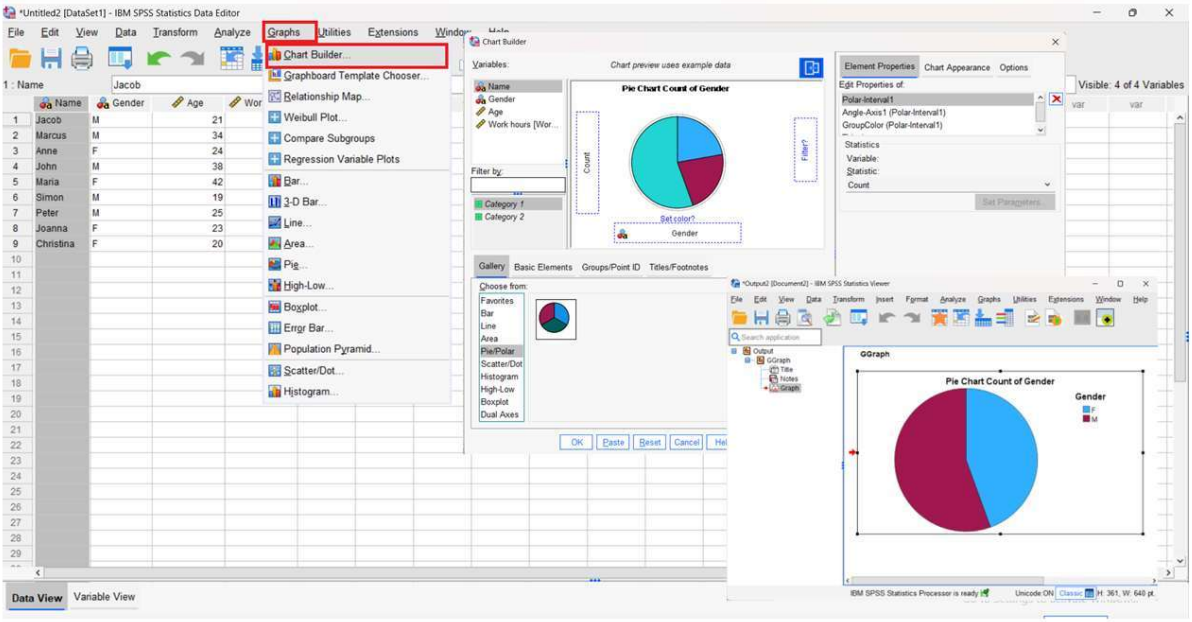
Slika 4.30Prozori za prikaz podataka i varijabli.

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).



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).



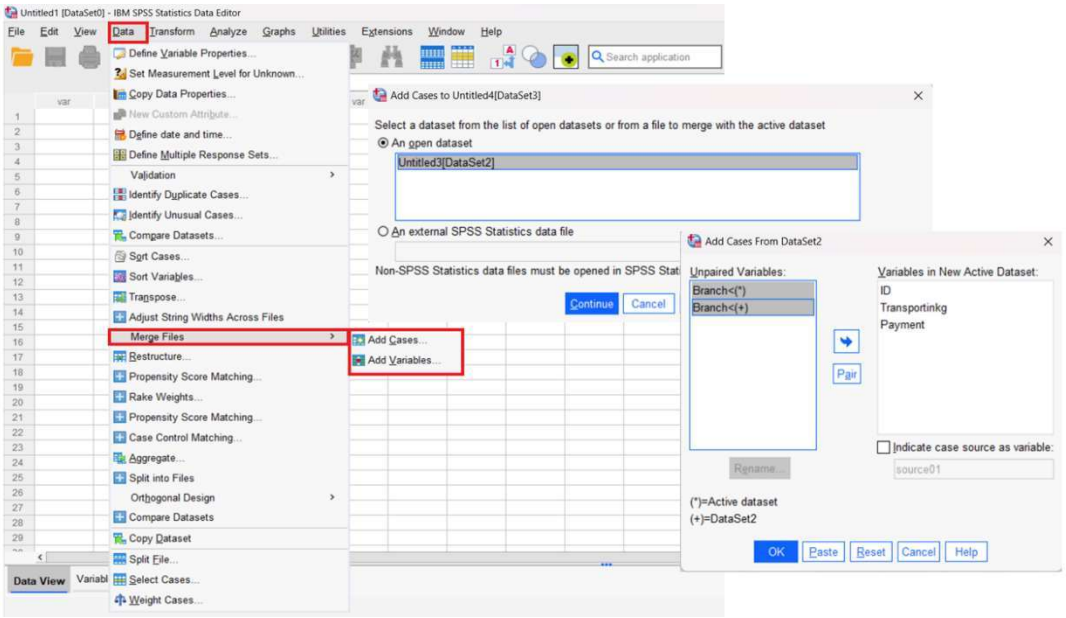
Slika 4.32Postavke izrade grafikona u SPSS-u.

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 podatke 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

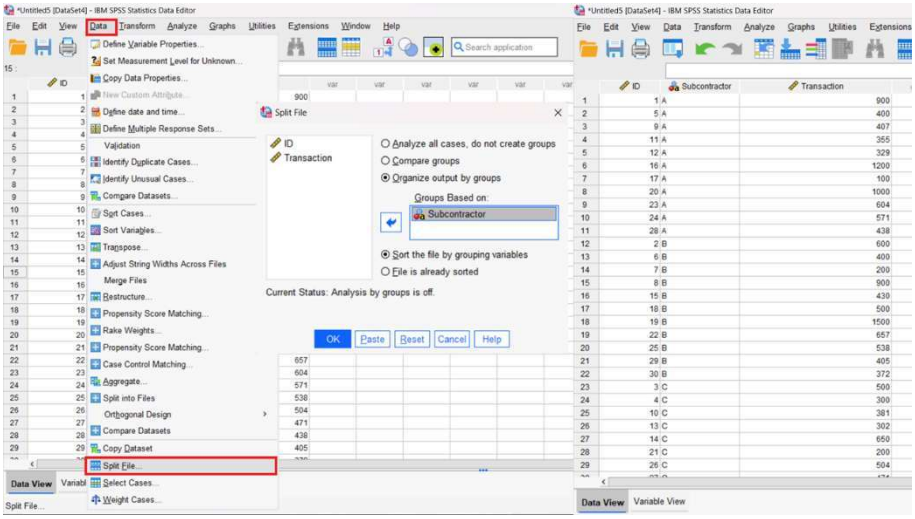
Kada se bavite ključnim podacima u SPSS softveru, postaje ključno razumjeti tehnike za manipuliranje informacijama u pojedinačnim aktivnim skupovima podataka. SPSS pruža funkcionalnosti koje olakšavaju manipulaciju postojećim podacima sadržanim u aktivnim skupovima podataka. Povremeno možete naići na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



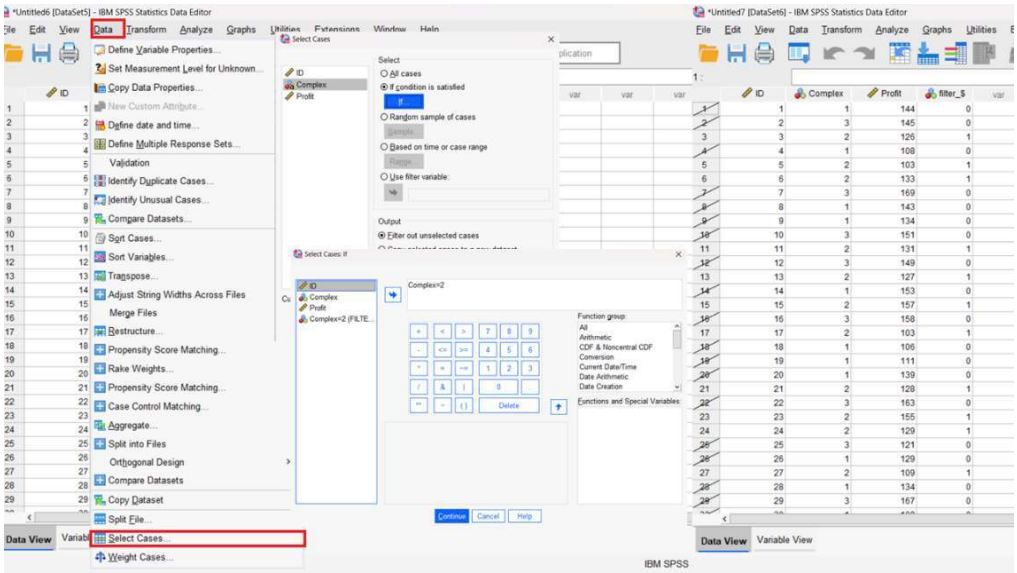
Slika 4.33Prozor za spajanje datoteka.

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.



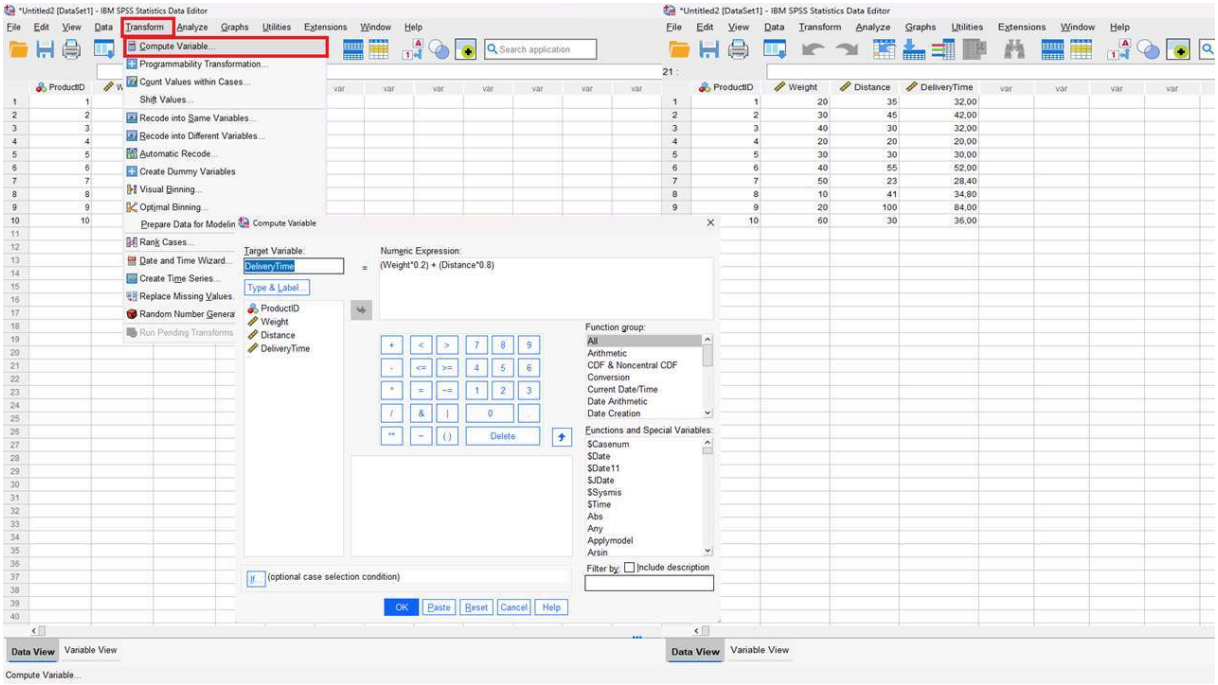
Slika 4.34Prozor za dijeljenje datoteke.

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 4.35Odabir 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.



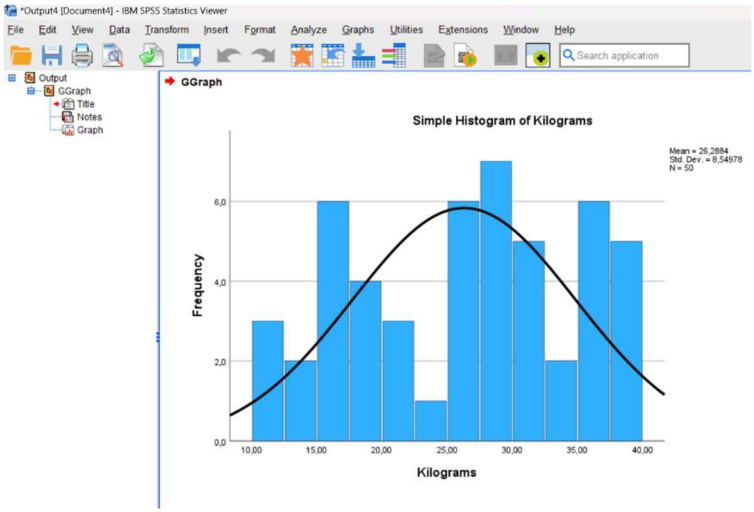
Slika 4.36Postupak izračunavanja varijabli.

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

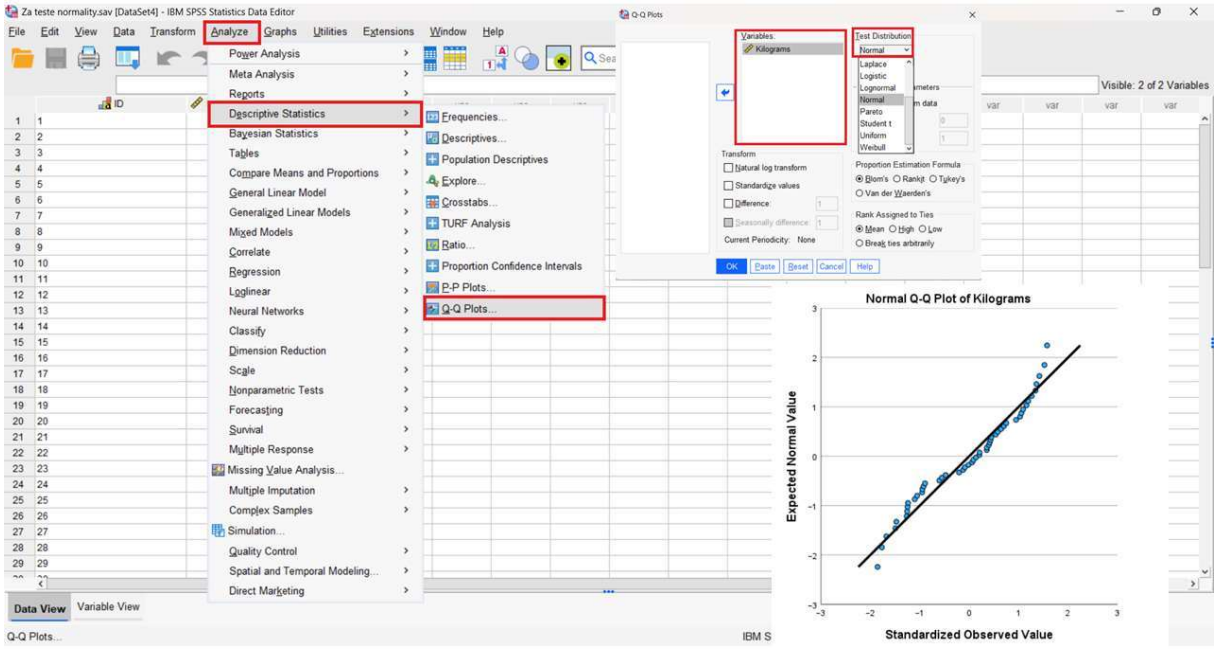
poglavljima 7.1 i 7.2), analizu podataka i interpretaciju rezultata (Garth, 2008; George i Mallery, 2022). U ovom poglavlju naš fokus je na analizi podataka pomoću softvera SPSS. Budući da smo hipoteze već obradili u prethodnim poglavljima, naš primarni fokus bit će na provođenju testova normalnosti 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

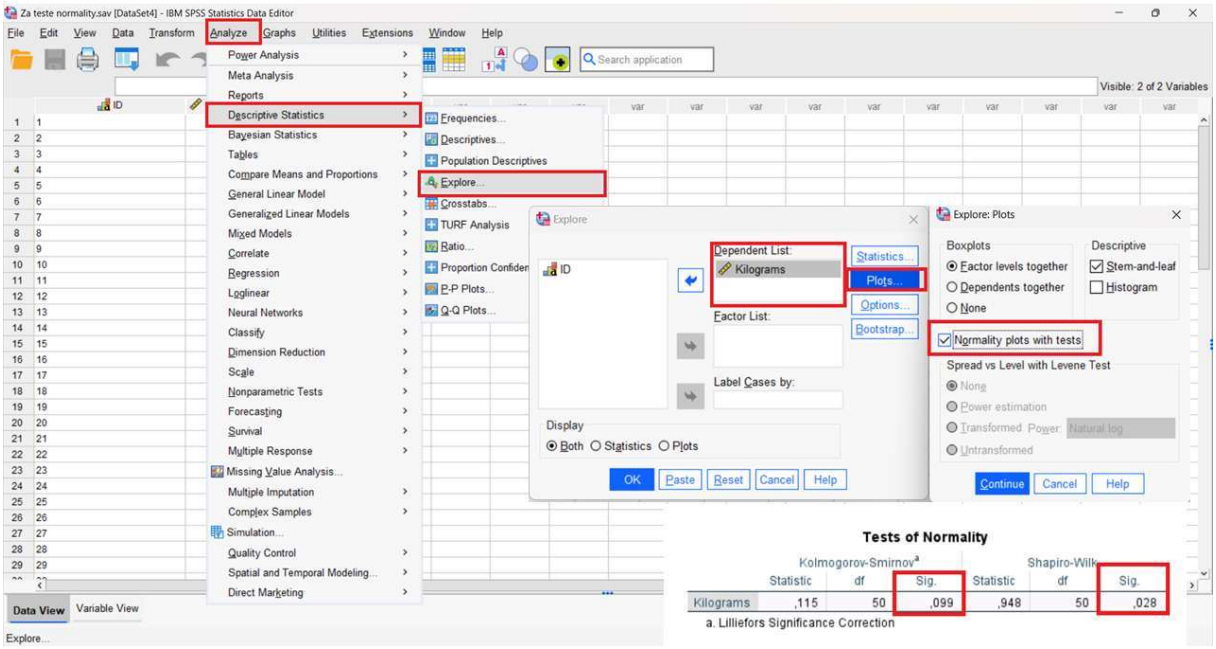
podataka iz testa histogramskog grafikona, proveli smo QQ grafikon test. Na slici 7.10 u nastavku možete primijetiti da je većina točaka klastera za našu varijablu poravnata s ravnom linijom, što ukazuje na normalnu distribuciju naših podataka. Iako smo već u ovoj fazi mogli zaključiti da je test normalnosti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

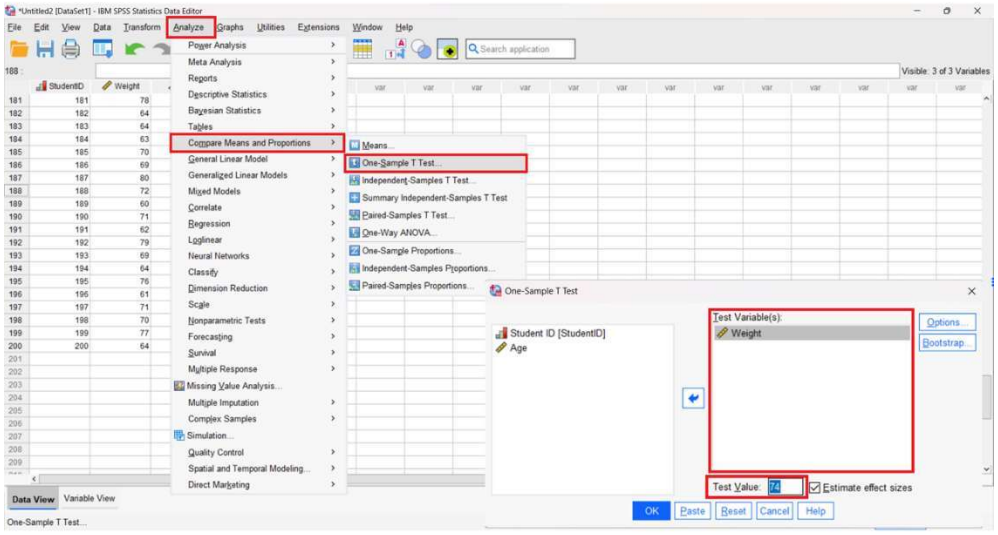
7.4 T-test jednog uzorka

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).



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)".

T-Test

One-Sample Statistics

	N	Mean	Std. Deviation	Std. Error Mean
Weight	200	69,63	6,303	,446

One-Sample Test

Test Value = 74

	t	df	Significance One-Sided p	Significance Two-Sided p	Mean Difference	95% Confidence Interval of the Difference Lower	95% Confidence Interval of the Difference Upper
Weight	-9,806	199	<,001	<,001	-4,370	-5,25	-3,49

One-Sample Effect Sizes

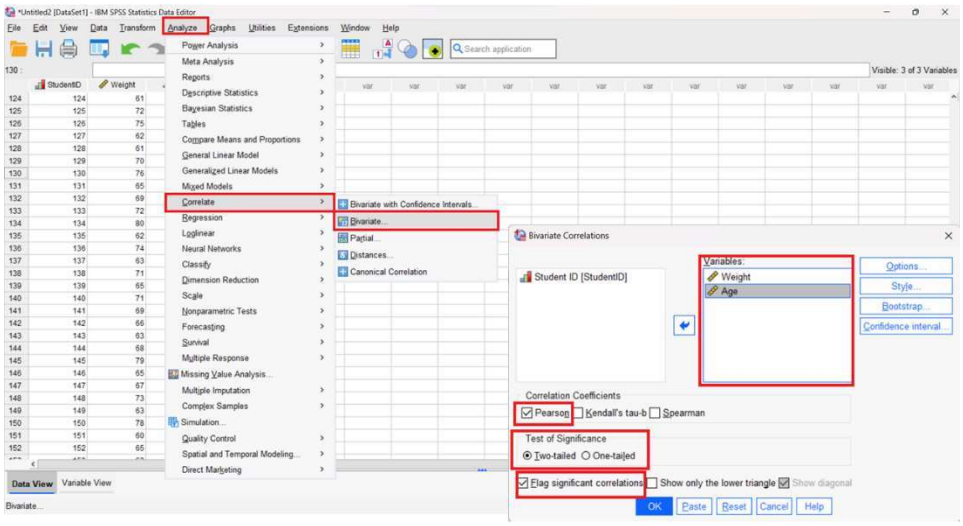
		Standardizer[a]	Point Estimate	95% Confidence Interval Lower	95% Confidence Interval Upper
Weight	Cohen's d	6,303	-,693	-,847	-,538
	Hedges' correction	6,326	-,691	-,844	-,536

a. The denominator used in estimating the effect sizes.
Cohen's d uses the sample standard deviation.
Hedges' correction uses the sample standard deviation, plus a correction factor.

Slika 4.41Rezultati T-testa jednog uzorka.

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".

Correlations

Correlations

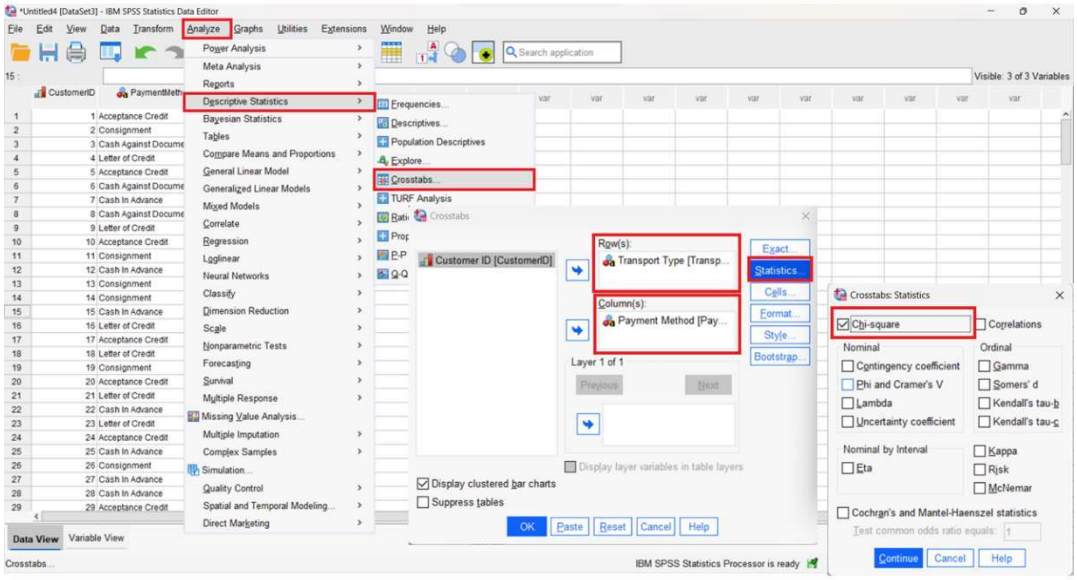
		Weight	Age
Weight	Pearson Correlation	1	-,038
	Sig. (2-tailed)		,596
	N	200	200
Age	Pearson Correlation	-,038	1
	Sig. (2-tailed)	,596	
	N	200	200

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

Case Processing Summary

	Cases					
	Valid		Missing		Total	
	N	Percent	N	Percent	N	Percent
Transport Type * Payment Method	200	100,0%	0	0,0%	200	100,0%

Transport Type * Payment Method Crosstabulation

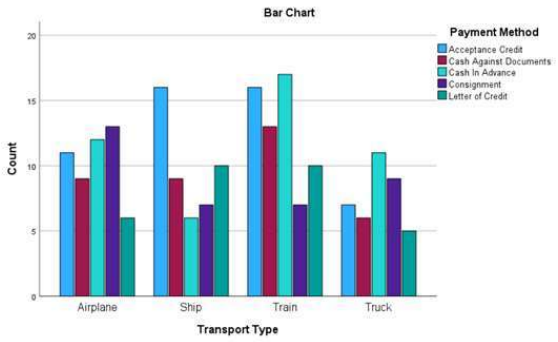
Count

		Payment Method					
		Acceptance Credit	Cash Against Documents	Cash In Advance	Consignment	Letter of Credit	Total
Transport Type	Airplane	11	9	12	13	6	51
	Ship	16	9	6	7	10	48
	Train	16	13	17	7	10	63
	Truck	7	6	11	9	5	38
Total		50	37	46	36	31	200

Chi-Square Tests

	Value	df	Asymptotic Significance (2-sided)
Pearson Chi-Square	11,614[a]	12	,477
Likelihood Ratio	11,965	12	,448
N of Valid Cases	200		

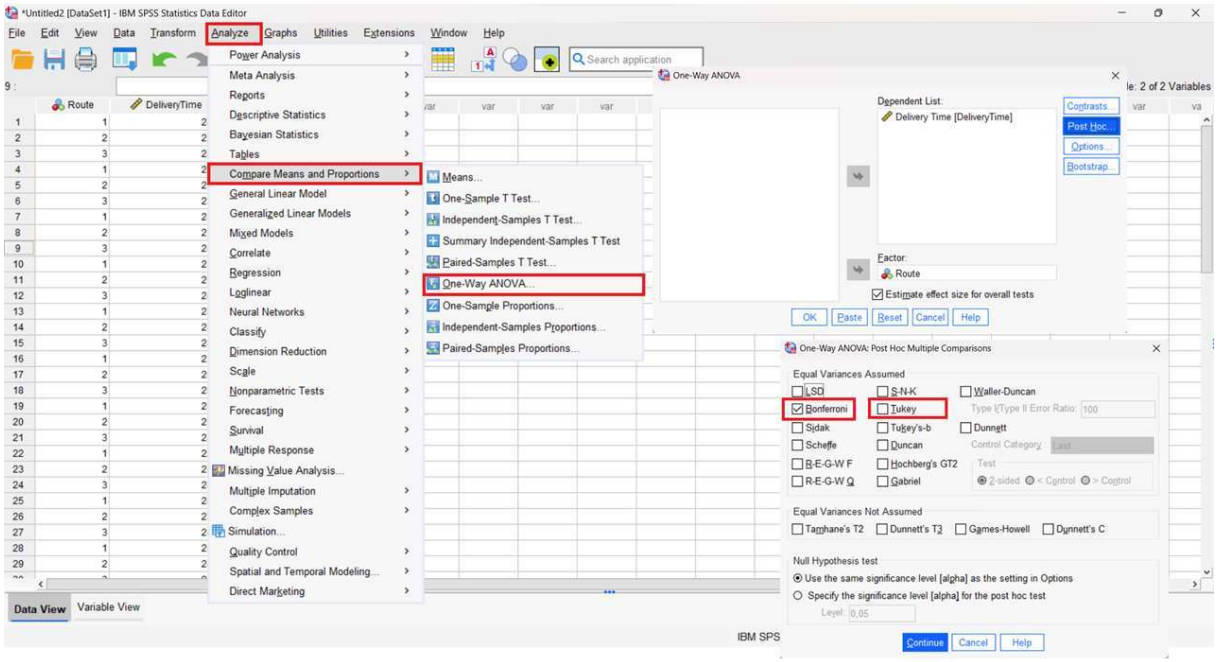
a. 0 cells (0,0%) have expected count less than 5. The minimum expected count is 5,89.



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.).



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."

ANOVA

Delivery Time

	Sum of Squares	df	Mean Square	F	Sig.
Between Groups	16,527	2	8,263	11,173	<,001
Within Groups	33,280	45	,740		
Total	49,807	47			

Post Hoc Tests

Multiple Comparisons

Dependent Variable: Delivery Time

Bonferroni

(I) Route	(J) Route	Mean Difference (I-J)	Std. Error	Sig.	95% Confidence Interval Lower Bound	95% Confidence Interval Upper Bound
1	2	-1,4250*	,3040	<,001	-2,181	-,669
	3	-,5500	,3040	,231	-1,306	,206
2	1	1,4250*	,3040	<,001	,669	2,181
	3	,8750*	,3040	,018	,119	1,631
3	1	,5500	,3040	,231	-,206	1,306
	2	-,8750*	,3040	,018	-1,631	-,119

*. The mean difference is significant at the 0.05 level.

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, 17TH 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, 1st 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]