Uvod i pregled

Ova knjiga je namenjena studentima koji pohađaju nastavu iz projektovanja i statističke obrade istraživanja, kao i onima koji samostalno planiraju i obavljaju istraživanja. Trebalo bi da vam ulije samopouzdanje da statističkim analizama pristupate spokojno i pribrano – ili barem bez velikog stresa!

Dobar deo poteškoća koje studenti iskuse u statističkoj analizi potiču od anksioznosti i zbunjenosti zbog nepoznatog žargona, složenih teorija i prevelikog broja odluka koje moraju doneti. Nažalost, većina statističkih predmeta i udžbenika podstiču oba ta osećanja! Pokušala sam da u ovoj knjizi prevedem statistiku na jezik koji je lakše razumeti i prihvatiti.

Knjiga *SPSS: priručnik za preživljavanje* veoma je pregledna zahvaljujući svojoj strukturi; u njoj se postupno – korak po korak – izlaže ono što vam treba za pripremu i analizu podataka. Posmatrajte podatke kao sirove sastojke kulinarskog recepta. Sastojke možete pripremiti na različite načine: kao predjelo, glavno jelo ili desert. U zavisnosti od sastojaka koje imate, razne moguće tehnike obrade mogu biti prikladne ili neprikladne. (Nema smisla planirati pripremanje lovačkih šnicli ako imate samo piletinu.) Planiranje i priprema su važan deo postupka (i u kuvanju i u analizi podataka). Trebalo bi da razmislite o sledećim pitanjima:

- Imate li odgovarajuće sastojke u odgovarajućim količinama?
- Kakva priprema je potrebna da bi sastojci bili pogodni za kuvanje?
- Koju vrstu termičke obrade ćete upotrebiti (kuvanje, pečenje, dinstanje, prženje)?
- Imate li predstavu kako rezultat (recimo, kolač s čokoladnim prelivom) treba da izgleda?
- Kako ćete znati da je jelo gotovo?
- Kada kulinarska obrada bude gotova, kako ćete ga poslužiti a da izgleda privlačno?

Ista pitanja su jednako relevantna i za postupak analize podataka. Eksperiment ili anketu morate isplanirati tako da pruži potrebne informacije u odgovarajućem formatu. Datoteku morate pripremiti kako treba i podatke pažljivo uneti u nju. Morate imati jasnu predstavu o pitanjima na koje treba dobiti odgovore i načinu na koji ćete ih formulisati. Morate znati koje statističke tehnike postoje i koje vrste

promenljivih su za njih prikladne a koje ne. Morate umeti da ispravno primenite odabranu statističku tehniku (npr. t-test) i da protumačite njen rezultat. Najzad, taj rezultat treba umešno da dovedete u vezu s prvobitnim istraživačkim pitanjima i predstavite u svom izveštaju. (Kada se prebacimo na kulinarstvo, tome bi odgovarala pitanja treba li kolač s čokoladnim prelivom da poslužite sa šlagom ili sa sladoledom, ili možda s jagodama/malinama i šećernom glazurom?).

Ni u kulinarstvu ni u analizi podataka ne možete samo izmešati sve sastojke, gurnuti ih u rernu (odnosno IBM SPSS) i nadati se da će rezultat biti dobar. Nadam se da će vam ova knjiga pomoći da malo bolje razumete postupak analize podataka i dati vam samopouzdanje i veštine potrebne dobrom „kuvaru".

Struktura knjige

SPSS: priručnik za preživljavanje ima 22 poglavlja, koja obuhvataju ceo postupak istraživanja – od projektovanja studije do analize podataka i predstavljanja rezultata. Podeljena je na pet delova. Prvi deo (*Kako početi*) obuhvata uvodne teme: projektovanje studije, pripremu šifarnika (engl. *codebook*) i upoznavanje IBM SPSS-a. U drugom delu (*Priprema datoteke s podacima*) naučićete kako se priprema datoteka, unose podaci i pronalaze greške u njoj. Preliminarne analize obrađene su u trećem delu knjige, čija su poglavlja posvećena primeni opisnih (deskriptivnih) statističkih pokazatelja (engl. *descriptive statistics*) i dijagrama (grafikona), doradi podataka i postupcima provere pouzdanosti mernih skala. Osim toga, provešću vas korak po korak kroz katkada težak zadatak izbora statističkih tehnika prikladnih za date podatke.

U četvrtom delu su predstavljene statističke tehnike za pronalaženje međuzavisnosti (npr. korelacija, delimična korelacija, višestruka regresija, logistička regresija i faktorska analiza). U tim poglavljima sažeto je opisano sledeće: namena svake tehnike, njene polazne pretpostavke, kako dobiti rezultate, kako tumačiti rezultate i kako ih predstaviti u naučnom radu ili izveštaju.

U petom delu se razmatraju statističke tehnike za poređenje grupa. Među njima su neparametarske tehnike, t-testovi, analiza varijanse (ANOVA), analiza varijanse više od dve promenljive (multivarijaciona analiza – MANOVA) i analiza kovarijanse (ANCOVA).

Kako koristiti ovu knjigu

Da biste ovu knjigu mogli delotvorno koristiti kao vodič kroz IBM SPSS, morate poznavati osnovne veštine rada na računaru. U uputstvima i primerima polazim od pretpostavke da već umete da koristite PC računar, naročito Windowsove funkcije. U nastavku je spisak potrebnih veština. Ukoliko s bilo kojom od narednih operacija imate poteškoća, potražite nečiju pomoć:

- upotreba Windowsovih padajućih menija;
- upotreba levog i desnog tastera miša;
- biranje teksta tako što ga pritisnete i prevučete mišem preko njega;

- spuštanje i podizanje, to jest minimiziranje i maksimiranje prozora;
- pokretanje i zatvaranje programa;
- prelazak iz jednog programa u drugi kada su istovremeno otvoreni;
- otvaranje, snimanje, preimenovanje, premeštanje i zatvaranje datoteka;
- rad s više datoteka istovremeno i prelazak iz jedne u drugu istovremeno otvorenu datoteku;
- kopiranje datoteka sa USB memorije na čvrsti disk ili obrnuto;
- pravljenje direktorijuma (foldera, fascikle, omotnice) i premeštanje datoteka iz jednog direktorijuma u drugi.

Ova knjiga nije sveobuhvatna. Pretpostavlja se da ste već naučili osnove statistike i da imate odgovarajući udžbenik. Treba da imate predstavu o tome šta SPSS radi ispod površine. IBM SPSS je izuzetno moćan softverski paket za analizu podataka, koji ume da obavlja veoma složene statističke procedure. U ovom priručniku se ne objašnjavaju sve statističke tehnike koje taj program podržava, već samo one koje se najčešće koriste. Knjiga bi trebalo da vam omogući da samostalno započnete statističku analizu i da steknete samopouzdanje u korišćenju SPSS-a.

Zavisno od podataka koje imate i istraživačkih pitanja (hipoteza) na koja tražite odgovor, možda ćete morati da upotrebite neku od složenijih analiza u SPSS-u. Ima mnogo dobrih knjiga o raznim statističkim tehnikama. Čitajte sve što vam padne šaka. Prelistajte knjige na policama biblioteka u koje ste učlanjeni, potražite one koje statistiku objašnjavaju (barem delimično) razumljivim jezikom! Prikupite sav taj materijal u celinu koju ćete koristiti u svim statističkim predmetima i svom istraživačkom projektu. Umesno je prikupiti i članke iz stručnih časopisa u kojima se objašnjavaju statističke analize i predstavljaju njihovi rezultati. Oni vam mogu poslužiti kao uzor za pisanje sopstvenog izveštaja.

SPSS: priručnik za preživljavanje prikladan je i kao udžbenik za rad s predavačem koji objašnjava razne vidove istraživačkog postupka, i kao knjiga za samostalno učenje, za one koji treba samostalno da realizuju neki istraživački projekat. Kada budete učili, ne propustite da vežbate korišćenje IBM SPSS-a za analizu podataka koji se mogu preuzeti s prateće veb lokacije ove knjige. (Pojedinosti o tome date su na str. ix.) Najbolje se uči uz rad, a ne samo čitajući kako to drugi rade. Pre nego što počnete da obrađujete sopstvene datoteke s podacima, poigrajte se onima iz kojih su uzeti primeri u ovoj knjizi. Tako ćete steći veće samopouzdanje i dobiti priliku da proverite da li ispravno radite analize.

Ponekad će vam se desiti da dobijete drugačiji rezultat od onoga u knjizi. To će se verovatno dogoditi ako budete upotrebljavali drugu verziju SPSS-a od one koja je korišćena u ovoj knjizi (IBM SPSS Statistics v23). IBM SPSS se stalno ažurira i to je odlično sa aspekta poboljšavanja programa, ali zbunjuje studente koji na ekranu svog računara ne vide rezultat naveden u knjizi. Razlika obično nije velika, pa ostanite pribrani i poigrajte se detektiva. Možda je informacija zapravo pred vama, samo u drugom obliku. Informacije o promenama u IBM SPSS-u potražite na veb lokaciji programa (www.spss.com).

Saveti za sprovođenje istraživanja

Evo još nekoliko saveta za one koji knjigu koriste kao vodič kroz sopstvene istraživačke projekte.

- **Pažljivo isplanirajte projekat.** Neka vam postojeće teorije i istraživanja posluže kao uzor prilikom planiranja sopstvenog projekta. Morate shvatiti šta pokušavate da dobijete, i zašto.
- **Razmišljajte unapred.** Očekujte i predvidite potencijalne probleme, pošto ih svaki projekat ima! Upoznajte statistički postupak koji nameravate da primenite i pomoću tih informacija oblikujte materijale za prikupljanje podataka. Postarajte se da, kada dođe vreme za primenu statističkih metoda analize, imate vrstu podataka koji se njima *mogu* analizirati.
- **Organizujte se.** Pažljivo vodite beleške o svim relevantnim istraživanjima, referencama itd. Smislite delotvoran sistem arhiviranja gomile stručnih članaka koje ćete prikupiti i, kasnije, rezultata IBM SPSS-a. Uopšte nije teško izgubiti se u gomili papira i brojki.
- **Vodite dobre beleške.** Kada pomoću IBM SPSS-a budete analizirali podatke, pažljivo pišite šta radite. Svim studentima preporučujem da u svesci sa spiralnim povezom vode beleške o svakoj sesiji korišćenja SPSS-a. Zapišite datum, imena promenljivih koje ste napravili, analize koje ste sproveli i imena datoteka u koje ste snimili izlaz iz SPSS-a. Kada naiđete na problem ili se datoteka s rezultatima na misteriozan način ošteti, instruktor će vas pomoću tih vaših beležaka spasiti!
- **Ostanite pribrani!** Ako prvi put analizirate podatke pomoću IBM SPSS-a, desiće se da se osetite bespomoćno. Udahnite nekoliko puta duboko i sami sebi recite da sve ide dobro i da će sve biti dobro. Mislite samo o koraku koji neposredno predstoji – dozvolite sebi da ponekad pogrešite i da ne znate sami kako da nastavite. Kada vam bude svega dosta i previše, prekinite, prošetajte i razbistrite glavu pre nego što pokušate da nastavite rad. Većina studenata lako koristi SPSS kada ga upoznaju. Kao sa svakim drugim učenjem, treba samo prevazići početnu zbunjenost i nedostatak samopouzdanja.
- **Dajte sebi dovoljno vremena.** Istraživački postupak uvek traje duže nego što se očekivalo, pogotovo unošenje i analiza podataka, pa za te aktivnosti planirajte dosta vremena.
- **Radite s prijateljem.** Tokom analize podataka tražite emocionalnu i praktičnu pomoć od drugih studenata. Podrška drugih je odlična zaštita od stresa!

Dodatni resursi

U ovoj knjizi obrađeno je više oblasti – od početnog projektovanja studije, sastavljanja upitnika, osnovnih statističkih tehnika (t-testovi, korelacija), pa sve do težih statističkih alatki (analiza varijanse više promenljivih, faktorska analiza). Po celoj knjizi posejane su preporuke za dodatnu literaturu. Pokušajte da čitate što više možete, naročito ako se budete bavili složenijim statističkim postupcima.

Sadržaj

Predgovor vii
Datoteke s podacima i prateća veb lokacija ix
Uvod i pregled xi

Deo I Kako početi **1**
1 Projektovanje studije 3
2 Priprema šifarnika 11
3 Upoznajte IBM SPSS 15

Deo II Priprema datoteke s podacima **25**
4 Pravljenje datoteke za podatke i unošenje podataka 27
5 Pronalaženje i uklanjanje pogrešnih podataka 43

Deo III Preliminarne analize **51**
6 Opisni statistički pokazatelji 53
7 Upotreba dijagrama za opisivanje i analizu podataka 67
8 Dorada podataka 83
9 Provera pouzdanosti merne skale 97
10 Izbor prikladnih statističkih tehnika 103

Deo IV Statističke tehnike za istraživanje veza između promenljivih **121**
11 Korelacija 129
12 Delimična korelacija 141
13 Višestruka regresija 145
14 Logistička regresija 165
15 Faktorska analiza 177

Deo V Statističke tehnike za poređenje grupa **197**
16 Neparametarske tehnike 205
17 T-testovi 233
18 Jednofaktorska analiza varijanse 243

19 Dvofaktorska analiza varijanse različitih grupa 259
20 Kombinovana analiza varijanse 269
21 Multivarijaciona analiza varijanse 279
22 Analiza kovarijanse 291

Dodatak: Pojedinosti o datotekama s podacima 311
Preporučena literatura 327
Bibliografija 331
Spisak termina korišćenih u knjizi 335
Indeks 337

DEO II
Priprema datoteke s podacima

Priprema datoteke s podacima za analizu obuhvata više koraka. Prvo se napravi prazna datoteka i u nju unesu podaci dobijeni u istraživanju, u obliku definisanom u šifarniku (objašnjeno u poglavlju 2). Zatim u datoteci s podacima treba pronaći eventualne greške i ukloniti ih. U drugom delu knjige razmatraju se upravo ta dva koraka. U poglavlju 4 obrađeni su postupci za pravljenje datoteke i unošenje podataka u nju. U poglavlju 5 objašnjen je postupak uklanjanja grešaka iz datoteke s podacima.



Pravljenje datoteke za podatke i unošenje podataka

Postupak izrade datoteke s podacima i njihove analize sastoji se od više koraka. Dijagram toka prikazan na sledećoj stranici sadrži osnovne potrebne korake. U ovom poglavlju provešću vas kroz postupak pravljenja datoteke i unošenja podataka u nju.

Objasnićemo tri ključna koraka u pripremi datoteke s podacima:

- *Korak 1.* Prvi korak je pregled i, po potrebi, izmena opcija koje IBM SPSS koristi za prikazivanje podataka i dobijenih rezultata.
- *Korak 2.* Sledeći korak je izgradnja strukture datoteke s podacima definisanjem promenljivih.
- *Korak 3.* Poslednji korak je unošenje podataka, tj. vrednosti koju je svakoj promenljivoj dodelio svaki učesnik ili ispitanik.

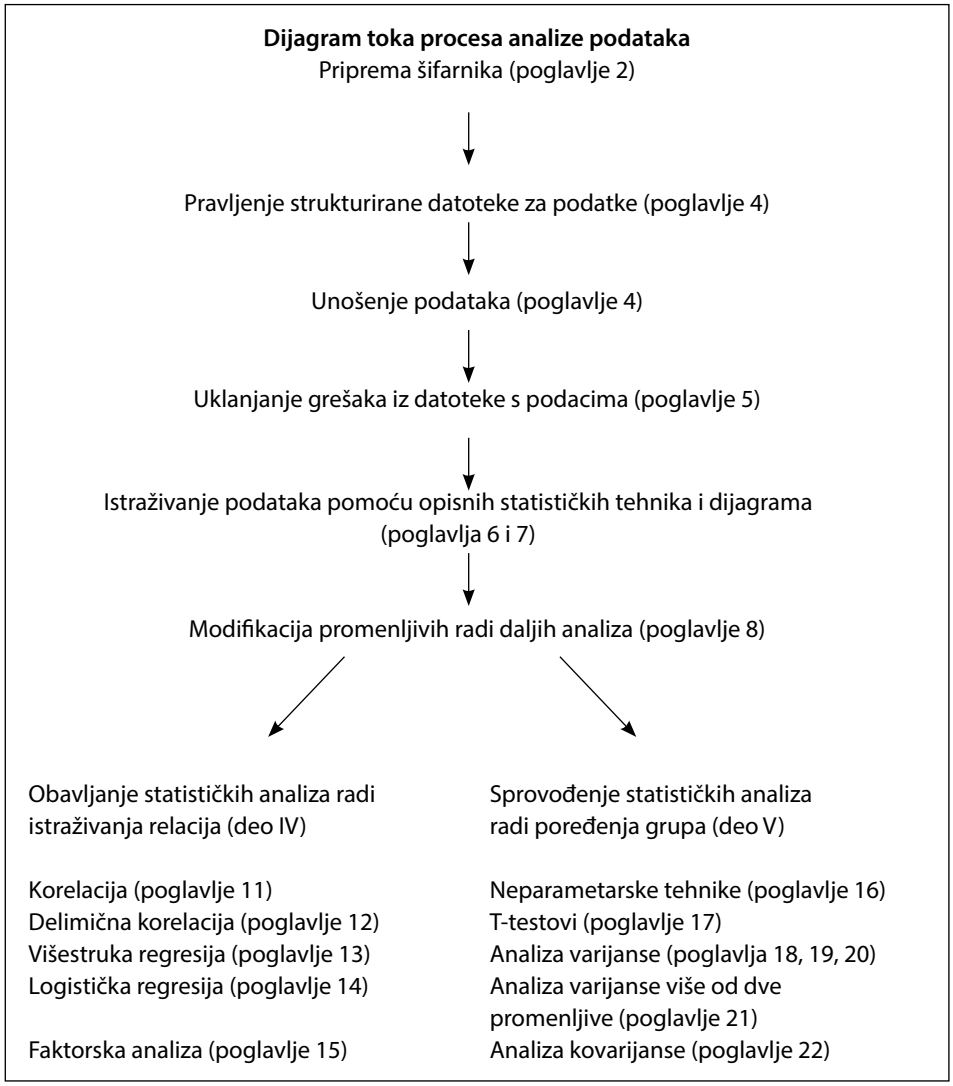
Navedene postupke ćemo ilustrovati na primeru datoteke **survey.sav**, opisanoj u dodatku knjige. U dodatku je i šifarnik pomoću kojeg su dobijeni numerički podaci za analizu rezultata.

Datoteke s podacima mogu se uvesti i iz drugih programa za tabelarne proračune (npr. iz Excela). To može umnogome da olakša postupak unošenja podataka, naročito studentima koji na svojim računarima nemaju IBM SPSS. U Excelu korisnik može da napravi praznu datoteku i da u nju unese podatke na svom kućnom računaru. Kada to završi, datoteku uveze u IBM SPSS i obradu i analizu podataka nastavi u njemu. Pri kraju poglavlja biće navedena uputstva za unošenje podataka u Excelu.

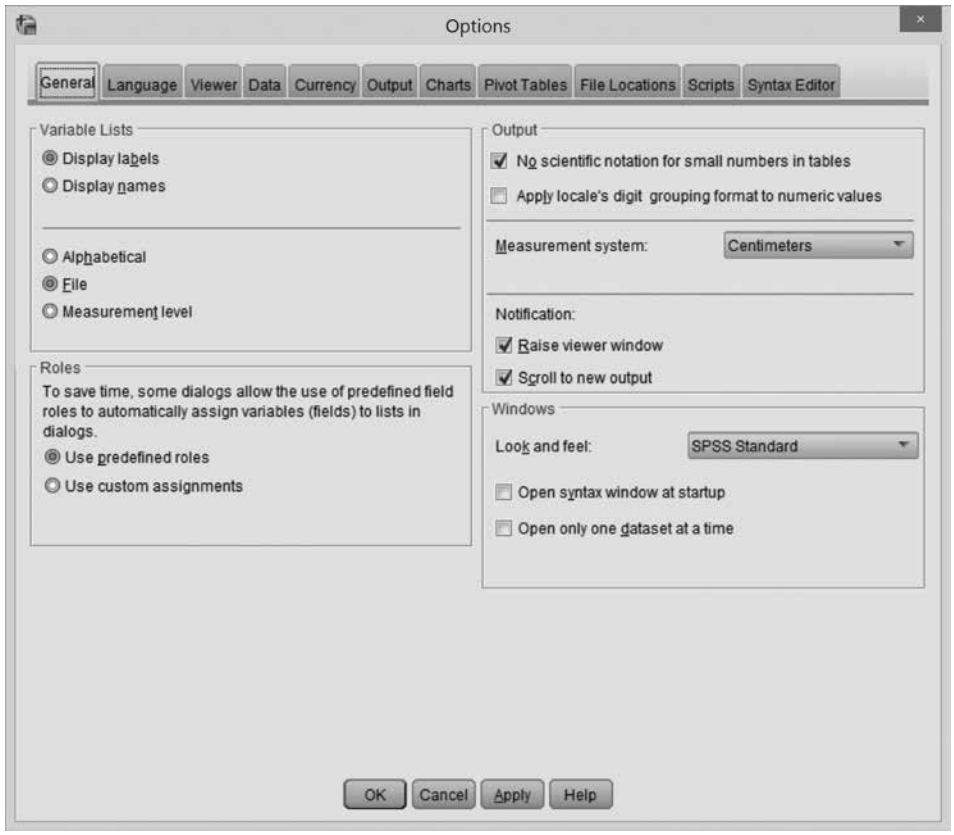
Izmena opcija u IBM SPSS-u

Pre nego što napravite datoteku s podacima, bilo bi dobro da proučite opcije (postavke) IBM SPSS-a od kojih zavisi način prikazivanja podataka i rezultata. To su opcije za definisanje načina prikazivanja promenljivih, vrste tabela u kojima će rezultati biti prikazani i mnogih drugih aspekata programa. Možda će vam spočetka nešto od ovoga biti nerazumljivo, ali čim počnete da unosite i analizirate podatke u IBM SPSS-u, možete se vratiti na ovaj odeljak.

Ukoliko računar na kome je instaliran IBM SPSS delite s drugim korisnicima (npr. u računarskoj laboratoriji), morate znati ove postavke zato što ih drugi studenti mogu promeniti, a to će uveliko izmeniti izgled programa. Trebalo bi da znate kako da vratite opcije u staro, poželjno stanje.



Da biste otvorili prozor **Options**, pritisnite **Edit** u glavnom meniju na vrhu ekrana, pa izaberite stavku **Options**. Otvoriće se prozor prikazan na slici 4.1. Tu je navedeno mnogo opcija; većinu nećete morati da menjate. U nastavku ću opisati samo ključne opcije, podeljene po karticama na kojima se nalaze. (S kartice na karticu prelazite pritiskom na jezičak one na koju hoćete da pređete.) Nemojte pritiskati **OK** dok na svim karticama ne izmenite sve opcije koje nameravate da izmenite.



Slika 4.1 Prozor Options.

Kartica General

Kada dođe vreme za obavljanje analiza, promenljive mogu biti poređane abecedno ili redom kojim se pojavljuju u datoteci. Ja uvek koristim redosled pojavljivanja u datoteci, zato što je on jednak redosledu stavki u upitniku i u šifarniku. Da bi se promenljive prikazivale po redosledu pojavljivanja u datoteci, samo se uverite da je uključena opcija **File** u odeljku **Variable Lists** (gore levo).

U odeljku **Output** (gore desno), treba da potvrdite polje **No scientific notation for small numbers in tables**. Tako ćete sprečiti pojavljivanje nekih veoma čudnih brojeva među rezultatima vaših statističkih analiza.

Kartica Data

Pritisnite karticu **Data** da biste podesili način prikazivanja datoteke s podacima. Ako promenljive s kojima radite nisu decimalne vrednosti, mogli biste promeniti format prikazivanja svih promenljivih. U odeljku **Display Format for New Numeric Variables**, za broj decimalnih mesta zadajte 0. To znači da se decimalna mesta neće prikazivati ni za jednu novu promenljivu. Tako smanjujete datoteku s podacima i pojednostavljujete njen izgled.

Kartica Output

Opcije na ovoj kartici služe za prilagođavanje načina prikazivanja običnih imena i dužih, opisnih imena promenljivih u rezultatima. Na samom dnu, u odeljku **Variable values in labels are shown as**, u padajućoj listi izaberite **Values and Labels**. Tako ćete u tabelama generisanim u prozoru **Viewer** moći da vidite i numeričke vrednosti i duža, opisna imena promenljivih.

U odeljku **Output Display** izaberite **Pivot tables and charts**. To je posebno važno ako nameravate da koristite postupke za neparametarske testove u SPSS-u.

Kartica Pivot Tables

Većinu rezultata statističkih analiza IBM SPSS prikazuje u izvedenim tabelama. Na kartici **Pivot Tables** možete sa opsežnih lista izabrati format tih tabela. Stil koji vam najviše odgovara naći ćete eksperimentisanjem. Kada tek počnem s nekom analizom, upotrebljavam stil **CompactBoxed**, čime štedim prostor (i papir prilikom štampanja).

Možda će vam biti korisna i opcija koja se nalazi na dnu kartice **Pivot Tables** – označena sa **Copying wide tables to the clipboard in rich text form**. Otvorite padajuću listu i izaberite **Shrink width to fit**. To je korisno kada komandom **Paste** prenosite rezultate iz IBM SPSS-a u Microsoft Word a oni su preširoki za datu stranicu (čest problem u nekim statističkim postupcima predstavljenim kasnije u ovoj knjizi).

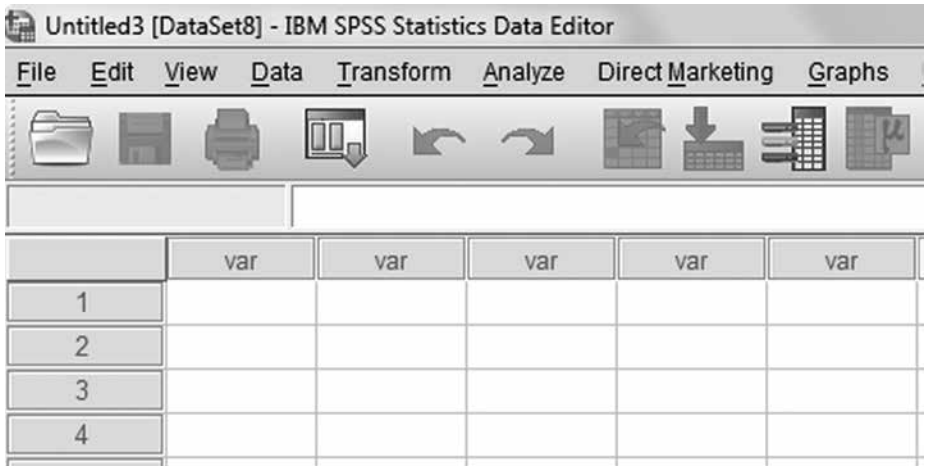
Stilove tabela možete menjati koliko god puta hoćete, samo ne zaboravite da stil treba promeniti *pre* analize. Kada se tabele već pojave u rezultatima, više im se ne može menjati stil, ali se mnoga njihova svojstva (recimo, veličina fontova, širina kolona) ipak mogu menjati u prozoru **Pivot Table Editor**. Otvoraćete ga kada dvaput pritisnete tabelu koju treba izmeniti.

Nakon što izmenite sve što želite na raznim karticama iz okvira za dijalog **Options**, pritisnite **OK**. Zatim možete preći na definisanje promenljivih i unošenje podataka.

Definisanje promenljivih

Pre nego što počnete da unosite podatke, IBM SPSS-u morate saopštiti kako vam se promenljive zovu i dati mu uputstva za šifrovanje (kodiranje), tj. pretvaranje vrednosti promenljivih u brojeve. To se naziva definisanje promenljivih, a obavlja se u prozoru **Data Editor** (slika 4.2). Prozor **Data Editor** se sastoji od dva različita prikaza: **Data View** i **Variable View**. Iz jednog u drugi prelazite tako što pritiskate jezičke njihovih kartica na dnu ekrana, s leve strane.

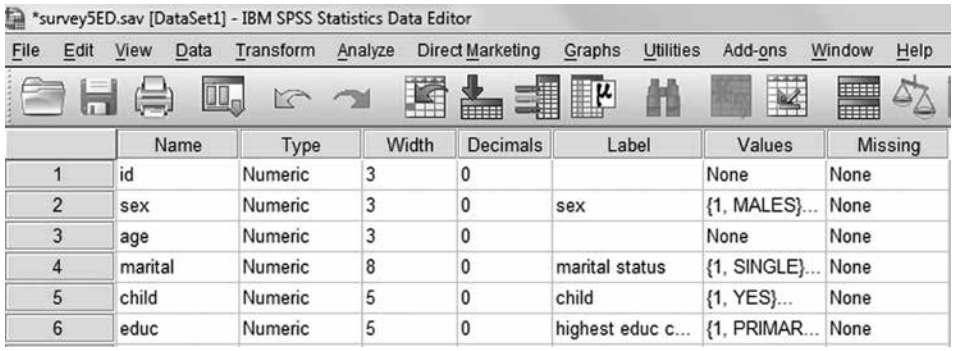
Svaka kolona u prozoru **Data View** odgovara jednoj promenljivoj (engl. *variable*) i nosi ime **var** (slika 4.2). Ta imena će biti zamenjena imenima promenljivih koje ste naveli u svom šifarniku. Duž leve strane su brojevi **1**, **2**, **3** itd. To su redni brojevi analiziranih slučajeva, koje IBM SPSS dodeljuje svakom redu s podacima. To *nisu* vaši brojevi za identifikaciju slučajeva; osim toga, brojevi redova se menjaju kada sortirate datoteku ili je podelite da biste podatke analizirali u podskupovima.



Slika 4.2 Prozor Data Editor.

Postupak definisanja promenljivih

Da biste definisali sve promenljive od kojih se vaši podaci sastoje, najpre pritisnite jezičak kartice **Variable View** na dnu ekrana, levo. U tom prikazu (slika 4.3), duž leve strane su navedeni redni brojevi promenljivih, a duž vrha njihovi atributi (ime, tip, širina, broj decimala, duže opisno ime itd.).



	Name	Type	Width	Decimals	Label	Values	Missing
1	id	Numeric	3	0		None	None
2	sex	Numeric	3	0	sex	{1, MALES}...	None
3	age	Numeric	3	0		None	None
4	marital	Numeric	8	0	marital status	{1, SINGLE}...	None
5	child	Numeric	5	0	child	{1, YES}...	None
6	educ	Numeric	5	0	highest educ c...	{1, PRIMAR...	None

Slika 4.3 Variable View.

Vaš zadatak je da sve promenljive definišete tako što ćete u šifarniku navesti potrebne informacije za svaku od njih. Deo tih informacija morate sami da obezbedite (npr. ime); ostale automatski dodeljuje IBM SPSS, koristeći podrazumevane vrednosti. Naravno, te vrednosti se mogu menjati po potrebi. U nastavku ćemo navesti osnovne neophodne informacije. Upotrebila sam zaglavlja koja odgovaraju zaglavljima kolona prikazanim u prozoru **Variable View**. Sledi opis jednostavnih koraka od kojih se postupak sastoji; međutim, postoje i brojne prečice koje ćete moći da koristite kada upoznate postupak. Njih navodim kasnije, u odeljku „Opcione prečice". Najpre treba upoznati osnovne tehnike.

Name

U ovu kolonu upišite kratko ime svake promenljive navedene u šifarniku. Pomoću imena će svaka promenljiva u datoteci s podacima biti identifikovana. Imena treba da su što kraća i da ne premašuju 64 znaka. Uz to, moraju zadovoljiti IBM SPSS-ovu konvenciju o imenovanju (objašnjenu u poglavlju 2). Imena moraju biti jedinstvena, početi slovom, i ne smeju sadržati razmake (beline) niti specijalne znakove. Primere dobrih imena za promenljive pročitajte u šifarniku u dodatku. Tamo su navedena imena promenljivih koje su korišćene u datotekama s podacima napravljenim za ovu knjigu. (Pojedinosti o tim datotekama videti na str. ix.)

Type

Kada upišete prvo ime promenljive, u koloni **Type** će se automatski pojaviti podrazumevana vrednost tipa promenljive – **Numeric**. To će vam odgovarati za numeričke promenljive, ali za tekstualne informacije (recimo, prezime) tip treba promeniti u **String**. Ukoliko među vašim podacima ima i datuma, tu je opcija **Date**. Tip promenljive menjate tako što pritisnete unutar ćelije; trebalo bi da se pojavi okvir s tri tačke i da vam se prikažu dostupne opcije. U istom prozoru može se podešavati širina promenljive i broj njenih decimalnih mesta.

Width

Podrazumevana vrednost širine je 8 (osim ako je promenjena korišćenjem instrukcija iz prozora **Options** predstavljenim ranije u ovom poglavlju). To je obično dovoljno za većinu podataka. Ukoliko vaša promenljiva može imati veoma velike vrednosti ili je u pitanju tekstualna promenljiva tipa *string*, promenite podrazumevanu vrednost. U protivnom, ostavite je na 8.

Decimals

Podrazumevana vrednost za broj decimalnih mesta obično je 2 (to se može izmeniti u prozoru **Options**, opisanom ranije u ovom poglavlju). Ako vaša promenljiva ima decimale, promenite ovu postavku kako god vam odgovara.

Label

Kolona **Label** služi da promenljivoj date duže ime od onog u koloni **Name**. Duža, opisna imena promenljivih navode se u rezultatima svih analiza koje obavljate u IBM SPSS-u. Na primer, promenljivoj kratkog imena UOPTIM mogli biste dodeliti duže opisno ime Ukupan optimizam.

Values

U koloni **Values** definišete značenje brojeva kojima ste šifrovali tekstualni sadržaj promenljivih, tj. zamenili ga. Pokazaću taj postupak za promenljivu pol (sex) s prethodne slike.

1. Pritisnite unutar ćelije pod naslovom **Values** za promenljivu koju želite da zadate. Trebalo bi da se na desnoj strani ćelije pojavi okvir s tri tačke. Otvoriće se okvir za dijalog **Value Labels**.
2. Pritisnite polje **Value** i u njega upišite *1*.
3. Pritisnite polje **Label** i u njega upišite „muški".
4. Pritisnite dugme **Add**. U zbirnom polju pisaće 1=muški.
5. Ponovite za ženski pol: za **Value** upišite *2*, za **Label** upišite „ženski". Pritisnite dugme **Add**.
6. Kada definišete sve brojeve navedene u šifarniku, pritisnite dugme **OK**.

Missing

Ponekad istraživači koriste specifične vrednosti kako bi pokazali da podatak nedostaje, tj. ne postoji. To ne morate da radite, pošto IBM SPSS svaku praznu ćeliju tumači kao nepostojanje podatka. Dakle, ukoliko nameravate da ostavite prazno mesto kada neki delić informacija nije dostupan, u kolonu **Missing** prikaza **Variable View** ne treba ništa da upisujete.

S druge strane, ako nepostojanje vrednosti nameravate da kodirate određenim brojem (npr. 99=nema vrednost), *morate* ga upisati u odeljak **Missing**, inače će ga IBM SPSS u svim statističkim analizama smatrati za legitimnu vrednost (s potencijalno katastrofalnim posledicama!). Pritisnite unutar ćelije a zatim zasenčen okvir s tri tačke koji će se pojaviti. Izaberite stavku **Discrete missing values** i upišite taj broj (npr. 99) u predviđeno polje. Mogu se zadati do tri vrednosti. Pritisnite **OK**.

Ako koristite te posebne kodove, preporučljivo je da se vratite i dodelite duža, opisna imena odgovarajućim vrednostima u koloni **Values**.

Columns

Podrazumevana širina kolona obično je 8, što je dovoljno za većinu namena. Izmenite je samo ukoliko očekujete ogromne vrednosti ili dugačka imena promenljivih.

Align

Vrednosti u kolonama obično su poravnate kako treba. Nema potrebe da to menjate.

Measure

Zaglavlje kolone **Measure** odnosi se na način merenja svake promenljive pojedinačno. Podrazumevana vrednost je **Scale**, što znači da se koristi intervalna merna skala, pogodna za neprekidne (kontinualne) podatke. Kada se promenljiva sastoji od kategorija (npr. pol, bračno stanje…), pritisnite tu ćeliju i zatim dugme sa strelicom koje će se pojaviti. Za kategorijske podatke izaberite nominalnu skalu (**Nominal**), a za one za koje je važan poredak mogućih vrednosti (npr. nivo obrazovanja), izaberite ordinalnu skalu (**Ordinal**). Važno je da ispravno zadate merne nivoe promenljivih jer će inače SPSS sprečiti da koristite neke postupke (npr. izradu dijagrama).

Role

Nema potrebe da išta menjate u ovom odeljku; samo ostavite podrazumevanu vrednost, **Input**.

Opcione prečice

Kada datoteka s podacima ima mnogo promenljivih, opisani postupak ume da bude zametan. Ima više prečica pomoću kojih ga možete ubrzati. Ukoliko vise promenljivih ima iste atribute (npr. tip, širina kolone, broj decimala), definišite prvu promenljivu a zatim njene atribute kopirajte na ostale promenljive.

Kopiranje atributa iz definicije jedne promenljive na drugu promenljivu

1. U prikazu **Variable View** pritisnite ćeliju sa atributom koji treba kopirati (npr. širinu, Width).
2. Otvorite meni **Edit** i u njemu izaberite **Copy**.
3. Pritisnite ćeliju istog atributa koja pripada drugoj promenljivoj (onoj na koju treba primeniti taj atribut).
4. Otvorite meni **Edit** i u njemu izaberite **Paste**.

Kopiranje atributa iz definicije jedne promenljive na više drugih promenljivih

1. U prikazu **Variable View** pritisnite ćeliju sa atributom koji treba kopirati (npr. širinu, Width).
2. Otvorite meni **Edit** i u njemu izaberite **Copy**.
3. Pritisnite ćeliju istog atributa koja pripada prvoj promenljivoj na koju ga treba primeniti, zadržite levi taster miša pritisnut i prevucite kursor niz kolonu birajući tako sve promenljive na koje taj atribut treba primeniti.
4. Otvorite meni **Edit** i u njemu izaberite **Paste**.

Zadavanje istih atributa nizu novih promenljivih

Kada podaci obuhvataju skale koje se sastoje od više pojedinačnih stavki, odjednom napravite sve te nove promenljive i definišite njihove atribute. Kao primer, sledi opis tog postupka za šest stavki (od optim1 do optim6) skale optimizma. Ako želite da uradite ovo kao vežbu, otvorite novu datoteku za podatke (**File**, **New**, **Data**).

1. U prikazu **Variable View** definišite atribute prve promenljive (optim1) prema prethodnim uputstvima. Trebalo bi da zadate i duža opisna imena koja odgovaraju mogućim vrednostima te promenljive: 1=nimalo se ne slažem, 2=ne slažem se, 3=neutralno, 4=slažem se, 5=potpuno se slažem.
2. Dok je izabran prikaz **Variable View**, pritisnite broj reda u kome se promenljiva nalazi (trebalo bi da se time izabere ceo taj red).



Korelacija

Korelacija opisuje jačinu i smer linearne veze između dve promenljive. Pomoću IBM SPSS-a može se izračunati više statističkih pokazatelja, koje korisnik bira u zavisnosti od nivoa merenja promenljivih i prirode podataka. U ovom poglavlju predstaviću postupke dobijanja i tumačenja koeficijenta Pirsonove linearne korelacije (r) i koeficijenta Spirmanove korelacije ranga (ro). Pirsonov r je prikladan za intervalne (neprekidne) promenljive. Može se upotrebiti i kada imate jednu neprekidnu promenljivu (npr. rezultate merenja samopoštovanja) i jednu dihotomnu promenljivu (npr. pol: M/Ž). Spirmanov ro je prikladan za ordinalne veličine ili veličine koje se mogu rangirati, i naročito je koristan kada podaci ne zadovoljavaju kriterijume za Pirsonovu korelaciju.

IBM SPSS izračunava obe vrste korelacije. Prvo ćete dobiti jednostavnu bivarijantnu korelaciju (tj. korelaciju između dve promenljive), poznatu i kao korelacija nultog reda (engl. *zero-order correlation*). U IBM SPSS-u takođe možete istraživati odnos između dve promenljive, uz kontrolu treće promenljive. Ta tehnika je poznata kao delimična (parcijalna) korelacija. U ovom poglavlju, predstavljena je procedura dobijanja koeficijenta Pirsonove korelacije proizvod/momenat i neparametarskog Spirmanovog koeficijenta ro. Parcijalna korelacija je obrađena u poglavlju 12.

Koeficijenti Pirsonove korelacije (r) mogu poprimiti samo vrednosti od 1 do +1. Predznak pokazuje da li je korelacija pozitivna (obe promenljive zajedno i opadaju i rastu) ili negativna (jedna promenljiva opada kada druga raste i obrnuto). Apsolutna vrednost tog koeficijenta (kada zanemarimo njegov predznak) pokazuje jačinu veze. Savršena korelacija, koja iznosi 1 ili -1, pokazuje da se vrednost jedne promenljive može tačno utvrditi kada znamo vrednost druge. Dijagram rasturanja koji ilustruje takvu vezu prava je linija. S druge strane, korelacija jednaka nuli pokazuje da između te dve promenljive ne postoji nikakva veza. Poznavanje vrednosti jedne promenljive tada nimalo ne pomaže u predviđanju vrednosti druge. Dijagram rasturanja tada izgleda kao oblak nasumično raspoređenih tačaka.

Za tumačenje korelacije potrebno je znati odgovore na više pitanja. To su uticaj nelinearnih veza, netipičnih tačaka, ograničenosti opsega uzorka, odnos korelacije i kauzalnosti, i odnos statističke i praktične značajnosti izračunatih koeficijenata korelacije. Te teme su razmotrene u uvodu u četvrti deo knjige. Pre nego što nastavite čitanje ovog poglavlja, trebalo bi da pročitate taj uvod.

Objašnjenje primera

Upotrebu korelacije pokazaću na primeru istraživanja uzajamnih veza nekih promenljivih sadržanih u datoteci s podacima **survey.sav**, s prateće veb lokacije ove knjige. Ta anketa je projektovana da istraži činioce koji utiču na psihološku prilagođenost i opšte raspoloženje ispitanika (celovit opis studije videti u dodatku). U ovom primeru, hoću da ocenim korelaciju između osećanja samokontrole ispitanika i njihovog nivoa subjektivno doživljenog stresa. Dve upotrebljene promenljive detaljno su opisane u nastavku teksta. Ukoliko želite da pratite i odmah primenjujete navedena uputstva, pokrenite IBM SPSS i otvorite datoteku **survey.sav**.

Primer istraživačkog pitanja: Postoji li veza između nivoa samokontrole ljudi nad njihovim unutrašnjim stanjima i njihovog nivoa subjektivno doživljenog stresa? Da li osobe sa visokim nivoom samokontrole doživljavaju manje stresa?

Šta vam treba: Dve promenljive, obe neprekidne; ili jedna neprekidna, a druga dihotomna (može imati samo dve vrednosti).

Šta se postiže: Korelacija opisuje vezu između dve neprekidne promenljive – i jačinu te veze i njen smer.

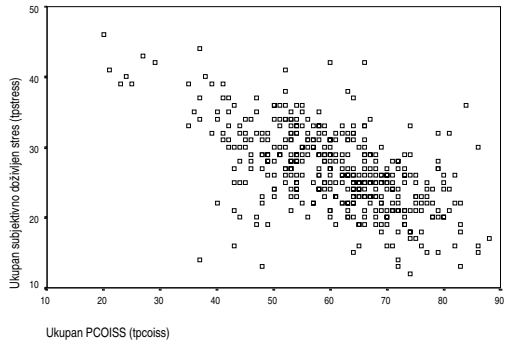
Pretpostavke: Videti uvod u četvrti deo knjige.

Neparametarska alternativa: Spirmanova korelacija ranga (ro).

Preliminarne analize za korelaciju

Pre izračunavanja korelacije treba nacrtati dijagram rasturanja. Na njemu proverite da li su zadovoljene pretpostavke o linearnosti i homogenosti varijanse (videti uvod u četvrti deo knjige). Pregledom dijagrama rasturanja steći ćete i bolju predstavu o prirodi veze između ispitivanih promenljivih.

Da biste nacrtali dijagram rasturanja koji prikazuje zavisnost između vaše nezavisne promenljive (Total PCOISS) i zavisne promenljive (Ukupan doživljen stres, tpstress), pratite detaljno uputstvo iz poglavlja 7. Ovde je prikazan rezultat generisan pomoću ovog postupka:



Tumačenje dijagrama rasturanja

Na dijagramu se može sagledati više svojstava raspodele dve ispitivane promenljive.

Korak 1: traženje netipičnih tačaka

Pogledajte ima li na dijagramu rasturanja netipičnih tačaka, tj. tačaka izdvojenih od ostalih, izuzetaka, eventualno mnogo većih ili manjih od ostalih. Kada netipične tačke imaju ekstremne vrednosti, treba proveriti da li su informacije uopšte bile ispravno unete? Da to nisu greške? Netipične tačke znatno utiču na neke analize, pa to vredi ispitati. Pojedini udžbenici statistike preporučuju da se ekstremne netipične tačke izbace iz skupa podataka. Drugi predlažu da im se promene šifre (kodovi) tako da im vrednost ne bude toliko ekstremna (videti poglavlje 6).

Kada prepoznate netipičnu tačku i poželite da je identifikujete pomoću ID broja datog slučaja (opservacije), upotrebite ikonicu **Data Label Mode** u prozoru **Chart Editor**. Dvaput pritisnite dijagram da biste aktivirali prozor **Chart Editor**. Pritisnite ikonicu koja pomalo liči na centar streljačke mete (ili izaberite **Data Label Mode** u meniju **Elements**) i dovedite kursor na tačku dijagrama koju želite da identifikujete. Pritisnite tačku jednom i prikazaće se njen ID broj (ako ste u koraku 5 prethodnog postupka izabrali ID) odnosno broj slučaja (engl. *case number*) koji mu je dodelio IBM SPSS. Prikazivanje brojeva isključićete ako ponovo pritisnete istu ikonicu.

Korak 2: pregled raspodele vrednosti

Raspodela vrednosti kazuje više stvari:

- Da li su tačke raspršene posvuda? To bi značilo da je korelacija veoma mala.
- Da li su tačke lepo raspoređene u obliku tanke cigare? To bi značilo da je korelacija veoma velika.
- Može li se povući prava linija kroz glavnu gomilu tačaka ili bi tačke bolje predstavila neka kriva linija? Ako je kriva linija očigledna (što sugeriše krivolinijsku vezu između promenljivih), ne treba koristiti Pirsonovu korelaciju pošto ona počiva na pretpostavci da je veza između promenljivih linearna (pravolinijska).
- Koji je oblik glavne gomile tačaka? Da li je jednake širine od jednog kraja do drugog? Ili je na početku uska pa postaje sve šira? U tom slučaju, podaci možda ne zadovoljavaju pretpostavku o homogenosti varijanse (engl. *homoscedasticity, homogenity of variance*).

Korak 3: određivanje smera veze između promenljivih

Dijagram rasturanja kazuje da li je veza između dve promenljive pozitivna ili negativna. Kada bi se kroz tačke nacrtala linija, koji smer bi pokazivala – sleva nadesno, da li bi pokazivala naviše ili naniže? Trend naviše prikazuje pozitivnu vezu, tj. da velikim vrednostima X odgovaraju velike vrednosti Y. Linija naniže kazuje da se radi o negativnoj korelaciji, tj. da malim vrednostima X odgovaraju velike vrednosti Y. U ovom primeru, izgleda da je korelacija negativna i umereno jaka.

Pošto ste istražili raspodelu rezultata na dijagramu rasturanja i utvrdili da je veza između promenljivih približno linearna i da su rezultati ravnomerno raspoređeni u obliku cigare, možete nastaviti izračunavanje koeficijenta Pirsonove ili Spirmanove korelacije.

Pre nego što počnete sledeći postupak, u glavnom meniju izaberite **Edit**, zatim **Options**, pa proverite da li je na kartici **General**, u odeljku **Output**, potvrđena opcija **No scientific notation for small numbers in tables**.

Postupak izračunavanja Pirsonovog koeficijenta r ili Spirmanovog ro

1. U meniju na vrhu prozora pritisnite **Analyze**, zatim **Correlate**, pa **Bivariate**.
2. Izaberite svoje dve promenljive i prebacite ih u polje **Variables** (npr. ukupan subjektivno doživljen stres: tpstress, ukupan PCOISS:tpcoiss). Ovde se može zadati cela lista promenljivih, ne samo dve. U rezultujućoj matrici (tabeli) biće navedene korelacije svih mogućih parova promenljivih. Tabela će biti velika ako u listi zadate više od nekoliko promenljivih.
3. U odeljku **Correlation Coefficients**, podrazumevana opcija je **Pearson** (što znači da se podrazumevano izračunava ta korelacija). Ako hoćete da izračunate i Spirmanovu korelaciju ranga (neparametarsku alternativu), potvrdite i polje **Spearman rho**.
4. Pritisnite dugme **Options**. Za nedostajuće vrednosti (odeljak **Missing Values**), potvrdite opciju **Exclude cases pairwise**. Na kartici **Options** može se zadati i izračunavanje srednje vrednosti i standardnog odstupanja.
5. Pritisnite **Continue** i zatim **OK** (odnosno dugme **Paste** da biste komandu snimili u **Syntax Editor**).

Opisani postupak generiše sledeće komande:

```
CORRELATIONS  /VARIABLES=tpstress tpcoiss
  /PRINT=TWOTAIL NOSIG
  /MISSING=PAIRWISE .
NONPAR CORR  /VARIABLES=tpstress tpcoiss
  /PRINT=SPEARMAN TWOTAIL NOSIG
  /MISSING=PAIRWISE .
```

Rezultat prethodnog postupka (sa obe izračunate korelacije – i Pirsonovom i Spirmanovom) prikazan je u tabelama.

Correlations

		Total perceived stress	Total PCOISS
Total perceived stress	Pearson Correlation	1	-.581**
	Sig. (2-tailed)		.000
	N	433	426
Total PCOISS	Pearson Correlation	-.581**	1
	Sig. (2-tailed)	.000	
	N	426	430

** Correlation is significant at the 0.01 level (2-tailed).

Neparametarske korelacije

Correlations

			Total perceived stress	Total PCOISS
Spearman's rho	Total perceived stress	Correlation Coefficient	1.000	-.556**
		Sig. (2-tailed)	.	.000
		N	433	426
	Total PCOISS	Correlation Coefficient	-.556**	1.000
		Sig. (2-tailed)	.000	.
		N	426	430

** Correlation is significant at the 0.01 level (2-tailed).

Tumačenje rezultata korelacije

Rezultate obe korelacije – i Pirsonove i Spirmanove – IBM SPSS prikazuje u tabeli koeficijenata korelacije između svakog para navedenih promenljivih, uz nivo značajnosti i broj analiziranih slučajeva (opservacija) u uzorku. Rezultati Pirsonove korelacije dati su u odeljku **Correlation**. Kada zatražite da se izračuna Spirmanov koeficijent ro (*rho*), ti rezultati se prikazuju u odeljku **Nonparametric Correlations**. Rezultati obe vrste korelacije, i parametarske i neparametarske, tumače se jednako.

Korak 1: provera informacija o uzorku

U tabeli **Correlations** najpre treba pogledati N, tj. broj slučajeva (opservacija) u uzorku. Da li je taj broj tačan? Ako mnogo podataka nedostaje, treba da utvrdite zašto je tako. Da niste zaboravili da potvrdite opciju **Exclude cases pairwise** u odeljku **Missing Data** na kartici **Options**? Druga opcija (**Exclude cases listwise**) znači da su iz svih analiza isključeni svi slučajevi kojima nedostaje vrednost bilo koje promenljive. Time se N često preoštro (i nepotrebno) smanjuje. U gornjem primeru imamo 426 slučajeva sa vrednostima na obe skale upotrebljene u ovoj analizi. Naravno, iz nje su morali da se izbace svi slučajevi kojima nedostaje podatak za bilo koju od te dve promenljive.

Korak 2: utvrđivanje smera veze

Drugo što treba pogledati jeste smer veze između promenljivih. Da li je ispred vrednosti koeficijenta korelacije negativan predznak? Ako jeste, to znači da između te dve promenljive postoji negativna korelacija (tj. da velike vrednosti na jednoj skali prate male vrednosti na drugoj). Tumačenje ovoga se menja u zavisnosti od načina dodeljivanja vrednosti promenljivama. Uvek proverite šta piše u upitniku i ne zaboravite da mnoge skale imaju negativno formulisane stavke koje treba obrnuti pre dodeljivanja vrednosti. Šta veliki brojevi (vrednosti) zapravo znače? To često zbunjuje studente, pa vam mora biti savršeno jasno pre nego što se usudite da tumačite rezultate.

U ovom primeru, koeficijent Pirsonove korelacije (-0,58) i Spirmanov koeficijent ro (-0,56) negativni su, što pokazuje negativnu korelaciju između subjektivno doživljene samokontrole i stresa. Što *više* samokontrole osoba oseća da ima, to *manje* stresa misli da doživljava.

Korak 3: utvrđivanje jačine veze

Treće što treba pogledati u rezultatima jeste veličina koeficijenta korelacije. On može imati vrednost od -1 do 1. Taj broj pokazuje jačinu veze između te dve promenljive. Vrednost 0 pokazuje da nikakva veza (korelacija) ne postoji; vrednost 1 pokazuje da je korelacija potpuna i pozitivna, dok vrednost -1 pokazuje da je korelacija potpuna i negativna.

Kako tumačimo korelacije čije su vrednosti između 0 i 1? Razni autori daju različita tumačenja; međutim, Cohen (1988, str. 79–81) daje sledeće smernice za veličinu korelacije:

mala	r=0,10 do 0,29
srednja	r=0,30 do 0,49
velika	r=0,50 do 1,0

Navedene smernice važe bez obzira na to da li ispred koeficijenta r stoji negativan predznak. Ne zaboravite, negativan predznak pokazuje smer veze, a ne njenu jačinu. *Jačina* korelacije je jednaka kada su koeficijenti r=0,5 odnosno r=-0,5. Razlikuje se samo njen *smer*.

U gornjem primeru, između dve promenljive postoji jaka korelacija (iznad 0,5), što govori u prilog tvrdnje da je veza između subjektivno doživljene samokontrole i stresa jaka.

Korak 4: izračunavanje koeficijenta determinacije

Koeficijent determinacije pokazuje koliki je deo varijanse dve promenljive zajednički; još se kaže „koliki je deo varijanse jedne promenljive objašnjen (ili prouzrokovan) varijansom druge". Zvuči impresivno, a izračunava se jednostavno; samo kvadrirajte vrednost r (pomnožite r sa samim sobom). Da biste to pretvorili u 'procenat varijanse', dobijeni proizvod pomnožite sa 100 (pomerite decimalni zarez za dva mesta udesno). Na primer, dve promenljive s koeficijentom korelacije r=0,2 imaju samo $0,2 \times 0,2 = 0,04 = 4$ procenta zajedničke varijanse. Te dve promenljive nemaju mnogo veze. Međutim, korelacija r=0,5 znači već 25 procenata zajedničke varijanse ($0,5 \times 0,5 = 0,25$).

U našem primeru, Pirsonova korelacija iznosi 0,581; kada se digne na kvadrat, dobija se 33,76 procenata zajedničke varijanse. Subjektivno doživljena samokontrola objašnjava gotovo 34 procenta varijanse u odgovorima ispitanika na skali subjektivno doživljenog stresa. To je sasvim pristojan deo objašnjene varijanse, kada se uporedi s mnogim istraživanjima sprovedenim u društvenim naukama.

Korak 5: ocena nivoa značajnosti

Sledeće što treba pogledati jeste nivo značajnosti (označen sa **Sig. 2 tailed**). Ovaj podatak se često pogrešno tumači, pa obratite pažnju. Nivo statističke značajnosti ne pokazuje jačinu veze između dve promenljive (nju pokazuje r odnosno ro), nego s koliko poverenja treba posmatrati dobijene rezultate. Na značajnost dobijenog iznosa koeficijenta r odnosno ro jako utiče veličina uzorka. Iz malih uzoraka (npr. n=30) mogu se izračunati umerene korelacije koje nisu statistički značajne na uobičajenom nivou p<0,05. Međutim, već i vrlo male korelacije (npr. r=0,2) izračunate iz velikih uzoraka (N=100+) mogu biti statistički značajne. Naravno da statističku značajnost koeficijenta korelacije treba navesti u izveštaju o istraživanju, ali se usredsredite na jačinu veze i veličinu zajedničke varijanse (videti korak 4).

Pri objavljivanju radova iz nekih oblasti (posebno zdravstva i medicine), možda će se od vas tražiti da navedete intervale poverenja za svoje koeficijente korelacije. Drugim rečima, opseg vrednosti u kome smo 95 posto sigurni da se nalazi 'prava vrednost' (ukoliko smo mogli da je izmerimo!). Nažalost, IBM SPSS to ne omogućava; međutim, postoje veb lokacije s odgovarajućim kalkulatorima. Ako vam trebaju te vrednosti, predlažem da posetite lokaciju http://vassarstats.net/rho.html. Možete pregledati i celu lokaciju VassarStats (http://vassarstats.net/), koja sadrži niz alatki za obavljanje statističkih izračunavanja.

Predstavljanje rezultata korelacije

Navešćemo kako bi se rezultati Pirsonove korelacije iz prethodno navedenog primera mogli predstaviti u izveštaju o istraživanju. (Ukoliko treba da predstavite rezultate Spirmanove korelacije, samo zamenite r vrednošću ro datom u drugoj tabeli.)

> Veza između subjektivno doživljene samokontrole unutrašnjih stanja (kako je meri skala PCOISS) i subjektivno doživljenog stresa (merenog na skali Perceived Stress) istražena je pomoću koeficijenta Pirsonove linearne korelacije. Obavljene su preliminarne analize da bi se dokazalo zadovoljenje pretpostavki o normalnosti, linearnosti i homogenosti varijanse. Izračunata je jaka negativna korelacija između te dve promenljive, r = -0,58, n = 426, p < ,001, pri čemu visoke nivoe subjektivno doživljene samokontrole prate niski nivoi subjektivno doživljenog stresa.

Korelacijom se često istražuju veze unutar cele grupe promenljivih, a ne samo između dve promenljive kao u gornjem primeru. U takvim slučajevima, bilo bi nezgrapno u jednom pasusu navoditi sve izračunate koeficijente korelacije; bolje je prikazati ih u tabeli. Evo kako se to može uraditi:

DEO V
Statističke tehnike za poređenje grupa

U petom delu ove knjige istražujemo neke IBM SPSS-ove tehnike za ocenjivanje razlika između grupa ili okolnosti. Te tehnike su vrlo složene i oslanjaju se na obimnu teoriju i statistička načela. Pre nego što započnete analizu u IBM SPSS-u, neophodno je da steknete barem elementarno znanje o statističkim tehnikama koje nameravate da upotrebite. U tome vam može pomoći mnogo dostupnih udžbenika iz statistike. (Na kraju ove knjige, u odeljku Preporučena literatura, dat je spisak nekih prikladnih naslova.) Bilo bi dobro da tu građu pregledate odmah. Tako biste lakše shvatili šta IBM SPSS izračunava, šta to znači i kako protumačiti složene nizove brojeva u njegovim rezultatima. U narednim poglavljima pretpostavljam da već imate osnovna znanja o statistici i da poznajete njenu terminologiju.

Tehnike obrađene u petom delu knjige

Za ispitivanje značajnih razlika između grupa postoji ceo niz tehnika. Programski paket IBM SPSS sadrži mnogo statističkih tehnika, a ovde ću obraditi samo one glavne – i to i parametarske i neparametarske. Parametarske tehnike počivaju na više pretpostavki o populaciji iz koje je izvučen uzorak (npr. da su rezultati normalno raspodeljeni) i prirodi tih podataka (da su mereni na intervalnim skalama). Neparametarske tehnike nemaju tako stroge pretpostavke i često su prikladnije za male uzorke ili kada su prikupljeni podaci izmereni samo na ordinalnim skalama (čiji se iznosi mogu rangirati). U narednoj tabeli dat je spisak svih tehnika obrađenih u ovom poglavlju.

Lista parametarskih tehnika i njihovih neparametarskih ekvivalenata obrađenih u petom delu

Parametarska tehnika	Neparametarska tehnika
Ne postoji	hi-kvadrat za kvalitet aproksimacije (fitovanja)
Ne postoji	hi-kvadrat za nezavisnost
Ne postoji	Maknemarov test
Ne postoji	Kohranov Q test
Ne postoji	Kappa mera slaganja
T-test nezavisnih uzoraka	Man-Vitnijev U test
T-test uparenih uzoraka	Vilkoksonov test ranga
Jednofaktorska ANOVA različitih grupa	Kruskal-Volisov test
Jednofaktorska ANOVA ponovljenih merenja	Fridmanov test
Dvofaktorska analiza varijanse (različitih grupa)	Ne postoji
Kombinovana ANOVA	Ne postoji
Multivarijaciona analiza varijanse (MANOVA)	Ne postoji
Analiza kovarijanse	Ne postoji

U poglavlju 10, provela sam vas kroz proces odlučivanja o tome koja je statistička tehnika prikladna za konkretno istraživačko pitanje. Odgovor se menja u zavisnosti od prirode istraživačkog pitanja, vrste podataka na raspolaganju i broja promenljivih i grupa. (Ako dosad niste pročitali to poglavlje, trebalo bi da to uradite pre nego što nastavite.) Evo ključnih tačaka koje treba imati u vidu kada se traži prikladna statistička tehnika:

- T-testovi se upotrebljavaju kada imate samo dve grupe (npr. muškarci/žene) ili dve tačke u vremenu (npr. pre intervencije, posle intervencije).
- Tehnike analize varijanse koriste se kada imate *dve ili više* grupa ili tačaka u vremenu.
- Tehnike uparenih uzoraka ili ponovljenih merenja upotrebljavaju se za testiranje *istih ljudi u više navrata*, ili kada imate uparene uzorke.
- Tehnike analize različitih grupa i nezavisnih uzoraka primenjuju se kada su subjekti u svim grupama *različiti ljudi* (ili nezavisni).
- Jednofaktorska analiza varijanse upotrebljava se kada imate samo jednu nezavisnu promenljivu (npr. pol).
- Dvofaktorska analiza varijanse primenjuje se kada imate dve *nezavisne* promenljive (pol, starosna grupa).
- Analiza varijanse više zavisnih promenljivih (multivarijaciona analiza) upotrebljava se kada imate više *zavisnih* promenljivih (anksioznost, depresija).
- Analiza kovarijanse (ANCOVA) upotrebljava se kada treba statistički kontrolisati (ukloniti) uticaj dodatne, remetilačke promenljive koja utiče na vezu između nezavisne i zavisne promenljive.

Pre nego što se upustimo u istraživanje nekih dostupnih tehnika, treba razmotriti više zajedničkih pitanja. Te teme će biti relevantne za mnoga poglavlja iz ovog dela knjige, pa ćete se vraćati ovde kako budete prolazili kroz ostatak knjige.

Pretpostavke

Na nekim opštim pretpostavkama počivaju sve ovde razmotrene parametarske tehnike (npr. t-testovi, analiza varijanse), a pojedine tehnike se zasnivaju i na dodatnim pretpostavkama. Opšte pretpostavke predstavljamo u ovom poglavlju, a specifičnije pretpostavke u narednim poglavljima, po potrebi. Trebalo bi da se vraćate na ovaj uvod kada budete primenjivali neku od tehnika predstavljenih u petom delu knjige. Razmatranje postupaka provere zadovoljenja pretpostavki videti u Tabachnick i Fidell (2013, 4. poglavlje).

Nivo merenja

Za sve parametarske pristupe pretpostavka je da se zavisna promenljiva meri na intervalnoj skali (pa rastojanja između brojčanih vrednosti odgovaraju rastojanjima između obeležja koja se mere). Dakle, upotrebljava se neprekidna skala, a ne diskretne kategorije. Kad god je to moguće prilikom projektovanja istraživanja, pokušajte da upotrebite neprekidne mere zavisne promenljive, a ne kategorijske. Tako stičete mogućnost upotrebe većeg broja tehnika analize podataka.

Slučajnost uzorkovanja

Parametarske tehnike obrađene u petom delu knjige zasnivaju se na pretpostavci da su rezultati dobijeni iz slučajnog uzorka populacije. U stvarnim istraživanjima ta pretpostavka često nije zadovoljena.

Nezavisnost opservacija

Opservacije od kojih se sastoje podaci moraju biti uzajamno nezavisne, tj. ni na jednu opservaciju ili merenje ne sme uticati nijedna druga opservacija ili merenje. Kršenje ove pretpostavke ima vrlo ozbiljne posledice; raspravu videti u knjizi Stevensa (1996, str. 238). U brojnim istraživačkim situacijama, krši se pretpostavka o nezavisnosti. U nastavku dajem primere takvih studija (preuzete iz knjiga koje su napisali Stevens, 1996, str. 239; i Gravetter & Wallnau, 2004, str. 251):

- Istraživanje učinka studenata koji rade u parovima ili malim grupama. Ponašanje svakog člana grupe utiče na sve ostale članove i tako se krši pretpostavka o nezavisnosti.
- Istraživanje navika i preferencija dece u vezi s gledanjem TV-a, kada su deca iz iste porodice. Ponašanje jednog deteta u porodici (koje, recimo, gleda program A) najčešće utiče na svu ostalu decu iz te porodice; zato opservacije nisu nezavisne.

- Istraživanje metoda poučavanja u učionici i njihovog uticaja na ponašanje i učinak studenata. Prisustvo malog broja problematičnih studenata može uticati na sve ostale, pa merenja ponašanja i učinka pojedinaca nisu nezavisna.

Za svaku situaciju u kojoj se opservacije ili merenja prikupljaju u grupnom okruženju ili su učesnici podvrgnuti nekom obliku interakcije, mogu biti potrebne specifičnije tehnike, kao što je modelovanje u više nivoa (hijerarhijsko). Taj pristup je sada uobičajen u istraživanjima koja obuhvataju decu u učionici, u školama, u gradovima; ili studije s pacijentima, različitim medicinskim specijalistima, u ordinaciji/kancelariji, u gradu ili zemlji. Više o tome videti u 15. poglavlju knjige autorki Tabachnick i Fidell (2013).

Normalnost raspodele

Parametarske tehnike daju tačne rezultate samo kada su populacije iz kojih su uzorci uzeti normalno raspodeljene. U mnogim istraživanjima (naročito u društvenim naukama), vrednosti zavisne promenljive nisu normalno raspodeljene. Srećom, većina tehnika je prilično robusna, tj. narušavanje ove pretpostavke prouzrokuje malu netačnost rezultata. Kada su uzorci dovoljno veliki (npr. preko 30 opservacija), kršenje ove pretpostavke ne bi trebalo da prouzrokuje veće probleme. Raspodela rezultata za svaku grupu može se proveriti pregledom histograma koji se crtaju pomoću IBM SPSS-ovog menija **Graphs** (videti 7. poglavlje). Podrobniji opis tog procesa videti u knjizi koju su napisale Tabachnick i Fidell (2013, 4. poglavlje).

Homogenost varijanse

Parametarske tehnike iz ovog dela knjige temelje se na pretpostavci da su uzorci dobijeni iz populacija jednakih varijansi. To znači da je promenljivost rezultata u svim grupama jednaka. Za ispitivanje te homogenosti u sklopu t-testova i analiza varijansi, IBM SPSS obavlja Leveneov test jednakosti varijansi. Njegovi rezultati su deo IBM SPSS-ovog izlaza za te tehnike. Budite pažljivi pri tumačenju rezultata tog testa: nadate se da pokazatelj *nije* značajan (tj. da mu je značajnost *veća* od 0,05). Kada dobijete značajnost manju od 0,05, to znači da varijanse dveju grupa nisu jednake i da ne važi pretpostavka o homogenosti varijanse. Ne paničite kada vam se to dogodi. Analiza varijanse je prilično neosetljiva na narušavanje te pretpostavke ukoliko su veličine grupa slične (npr. najveća/najmanja = 1,5; Stevens, 1996, str. 249). U t-testovima dobijate dva niza rezultata: jedan za situacije u kojima pretpostavka nije narušena, a drugi za one kada jeste. U tom slučaju, samo upotrebite onaj skup rezultata koji odgovara podacima.

Greška prve vrste, greška druge vrste i moć testa

T-testovi i analize varijansi služe za ispitivanje hipoteza. U toj vrsti analiza uvek je moguće doneti pogrešan zaključak. Mogu se napraviti dve različite greške. Kada

odbacimo nultu hipotezu – koja je u stvari tačna – radi se o grešci prve vrste. To se dešava kada zaključimo da između grupa postoji razlika, a ona zapravo ne postoji. Verovatnoću te greške minimizujemo tako što izaberemo malu vrednost alfa; najčešće se koriste vrednosti 0,05 ili 0,01, tj. svesno rizikujemo da jedanput u 20 odnosno 100 slučajeva odbacimo nultu hipotezu (o nepostojanju razlike) kada je tačna.

Pri ispitivanju (verifikaciji) hipoteza može se napraviti i druga vrsta greške: kada ne odbacimo pogrešnu nultu hipotezu (tj. kada zaključimo da između grupa ne postoji razlika, a ona zapravo postoji). Nažalost, te dve greške su obrnuto srazmerne. Što više smanjujemo verovatnoću greške I vrste, povećavamo verovatnoću da ćemo napraviti grešku II vrste.

Idealno bi bilo da se pomoću upotrebljenih testova tačno utvrdi postoji li zaista razlika između grupa. To je tzv. *moć* testa, tj. verovatnoća da se otkrije postojeća razlika, odnosno da se ne napravi greška II vrste. Testovi se razlikuju po svojoj moći; recimo, kada su njihove osnovne pretpostavke zadovoljene, parametarski testovi (kao što su t-testovi, analiza varijanse itd.) potencijalno su moćniji od neparametarskih testova. Međutim, na moć testa u datoj situaciji utiču i drugi činioci:

- veličina uzorka;
- veličina uticaja istraživane razlike između grupa, tj. uticaja nezavisne promenljive; i
- alfa nivo (rizik greške I vrste) koji je zadao istraživač (npr. 0,05/0,01).

Moć testa se jako menja u zavisnosti od veličine uzorka upotrebljenog u studiji. Stevens (1996) tvrdi da 'moć ne predstavlja problem' (str. 6) kada je uzorak velik (npr. 100 ili više učesnika). Međutim, u istraživanjima na malim uzorcima (npr. n=20) morate biti svesni mogućnosti da neznačajan rezultat može biti posledica nedovoljne moći testa. Stevens (1996) predlaže da se za male grupe po potrebi poveća alfa (rizik greške I vrste), npr. na 0,10 ili 0,15 umesto uobičajenih 0,05.

Postoje i tabele (videti Cohen, 1988) u kojima se može očitati veličina uzorka potrebna za određenu (dovoljnu) moć testa, za datu veličinu uticaja razlike koju želite da otkrijete. Sve više je i softverskih programa koji to umeju da izračunaju (jedan od najboljih je G*Power, dostupan na adresi www.gpower.hhu.de).

I u nekim IBM SPSS procedurama izračunava se moć obavljenog testa, pri čemu se vodi računa o veličini uticaja razlike i veličini uzorka. Idealno bi bilo da imate 80 procenata šanse da otkrijete postojanje veze. Kada dobijete neznačajan rezultat i imate veoma mali uzorak, trebalo bi da pogledate moć upotrebljenog testa. Uz moć testa manju od 0,80 (80 procenata šanse da otkrijete postojeću razliku), treba pažljivo protumačiti razlog neznačajnosti rezultata. Ona može biti posledica nedovoljne moći testa, a ne nepostojanja razlike između grupa. Analiza moći pokazuje koliko poverenja treba imati u rezultate kada se ne odbaci nulta hipoteza o jednakosti grupa. Što je veća moć testa, to više tada treba biti uveren da stvarna razlika između grupa ne postoji.

Planirana poređenja / naknadne analize

Analizom varijanse utvrđujete da li postoje značajne razlike između raznih grupa ili okolnosti. Katkada će vas zanimati da li se grupe kao celina razlikuju (da li nezavisna promenljiva na neki način utiče na vrednosti zavisne promenljive). U drugim istraživačkim kontekstima usredsredićete se više na ispitivanje razlika između pojedinih – a ne svih mogućih grupa. Vodite računa o toj razlici, pošto se za svaku od tih namena koristi drugačija analiza.

Planirana (ili *a priori*) poređenja služe za ispitivanje konkretnih hipoteza (obično izvučenih iz teorije ili prethodnih istraživanja) u vezi s razlikama unutar određenog podskupa grupa (npr. da li se grupe 1 i 3 značajno razlikuju?). Ta poređenja treba specificirati (isplanirati) pre nego što analizirate podatke, umesto da pecate po rezultatima u nadi da ćete naći nešto zanimljivo!

Ukoliko nameravate da specificirate mnogo različitih, ali istovremenih poređenja, moraćete postupati pažljivo. Planirana poređenja ne uklanjaju povećani rizik od grešaka I vrste, koji je posledica velikog broja paralelnih hipoteza koje se ispituju. Greška I vrste znači odbaciti nultu hipotezu (npr. da nema razlike između grupa) koja je u stvari tačna. Drugim rečima, povećan je rizik da ćete misliti da ste otkrili značajan rezultat (razliku), a on je zapravo sasvim slučajan. Kada istražujete veliki broj razlika, bezbedniji je drugačiji pristup, tj. naknadna (post-hoc ili *a posteriori*) poređenja, koja štite od grešaka I vrste.

Treća mogućnost je da na alfa nivo (rizik greške I vrste) koji ćete upotrebiti za procenu statističke značajnosti, primenite tzv. Bonferonijevo prilagođenje. To znači zadati stroži alfa nivo za svako poređenje, da bi alfa u svim testovima zajedno ostao na razumnom nivou. To se postiže deljenjem alfa nivoa (najčešće 0,05) brojem poređenja koje nameravate da obavite; zatim se ta nova vrednost koristi kao zahtevani alfa nivo. Primera radi, za tri nameravana poređenja novi alfa nivo bio bi 0,05 podeljeno s 3, što je jednako 0,017. Raspravu o varijaciji ove tehnike videti u Tabachnick i Fidell (2013, str. 52).

Naknadna (*post-hoc* ili *a posteriori*) poređenja upotrebljavaju se kada želite da obavite ceo niz poređenja, tj. istražite razlike između svih mogućih grupa ili uslova u studiji. Ukoliko odaberete taj pristup, analiza treba da se sastoji od dva koraka. Prvo se izračuna ukupan F pokazatelj koji kazuje ima li značajnih razlika između grupa u projektu. Ako je ukupan F pokazatelj značajan (što ukazuje na to da postoji razlika između grupa), možete nastaviti i obaviti dodatne testove za identifikaciju tih razlika (npr. da li se Grupa 1 razlikuje od Grupe 2 ili Grupe 3, da li se razlikuju Grupa 2 i Grupa 3).

Naknadna poređenja štite od moguće greške I vrste kao posledice velikog broja različitih poređenja. To se postiže zadavanjem strožih kriterijuma za značajnost, koju je utoliko teže postići. Kod malih uzoraka to ume da bude problem, zato što je ponekad vrlo teško dobiti značajan rezultat, čak i kada je vidljiva razlika u rezultatima između grupa veoma velika.

Postoje i brojni naknadni testovi koji se razlikuju po svojoj prirodi i strogosti. Razlikuju se i pretpostavke na kojima se zasnivaju. U nekima se pretpostavlja da

su varijanse dve grupe jednake (npr. Tukey); u drugima se ne pretpostavlja jednakost varijansi (npr. Dunnettov C test). Među najčešće upotrebljavanim naknadnim testovima su Tukejev test „zaista značajne različitosti" (engl. *Honestly Significant Different*, HSD) i Šefeov (Scheffe) test. Od ta dva, Šefeov test je bezbednija metoda za smanjenje rizika od greške I vrste. Međutim, to se plaća u moći. Tim testom je teže (manje verovatno) otkriti stvarno postojeću razliku između grupa.

Veličina uticaja

Sve tehnike razmotrene u ovom delu knjige pokazuju da li je razlika između grupa statistički značajna (tj. neslučajna). Za većinu istraživača i studenata nastupi trenutak uzbuđenja kada utvrde da su njihovi rezultati statistički značajni! Međutim, istraživanje znači više od dobijanja statističke značajnosti. Verovatnoća ne pokazuje stepen povezanosti promenljivih (jačinu veze). Kada su izračunate za velike uzorke, čak i vrlo male razlike između grupa postaju statistički značajne. To ne znači da je ta razlika dovoljno velika da bi imala ikakvu praktičnu ili teorijsku važnost.

Jedan od načina da ocenite važnost svojih rezultata jeste da izračunate veličinu uticaja (engl. *effect size*), tj. jačinu veze između promenljivih. To je skup pokazatelja koji pokazuje relativnu veličinu razlika između srednjih vrednosti ili iznos ukupne varijanse u zavisnoj promenljivoj koji se može predvideti na osnovu poznavanja vrednosti nezavisne promenljive (Tabachnick & Fidell, 2013, str. 54).

Ima više pokazatelja veličine uticaja. Za poređenje grupa najčešće se upotrebljavaju pokazatelji *parcijalni eta kvadrat* i *Koenov d*. IBM SPSS izračunava parcijalni eta kvadrat u sklopu analize varijanse, ali ne i u sklopu t-testova; no, to je lako izračunati na osnovu drugih njegovih rezultata.

Pokazatelj veličine uticaja *parcijalni eta kvadrat* srazmeran je delu varijanse zavisne promenljive koji je objašnjen nezavisnom promenljivom. Može imati vrednosti u opsegu od 0 do 1. S druge strane, *Koenov d* predstavlja razliku između grupa izraženu brojem standardnih odstupanja. Pazite da ne pobrkate te pokazatelje kada budete tumačili jačinu veze.

Cohen (1988, str. 22) predložio je sledeće smernice za tumačenje veličine uticaja (kada se ocenjuje istraživanje koje obuhvata poređenje grupa). [Koenova preporuka se odnosi na pokazatelj eta kvadrat, ali se može primeniti i na tumačenje pokazatelja parcijalni eta kvadrat. Formula za parcijalni ima neznatno drugačiji imenilac. Više o tome pročitajte u knjizi autorki Tabachnick i Fidell, 2013, str. 55.]

Veličina uticaja	Eta kvadrat (% objašnjene varijanse)	Koenov d (jedinica standardnog odstupanja)
Mali	0,01 tj. 1%	0,2
Srednji	0,06 tj. 6%	0,5
Veliki	0,138 tj. 13,8%	0,8

Vodite računa o tome da Koen daje drugačije smernice za korelacione projekte (obrađene u četvrtom delu knjige). Gornje vrednosti važe za poređenja grupa.

Nedostajući podaci

Kada radite istraživanje, naročito ono sa ljudima, vrlo retko ćete u svakoj opservaciji tj. od svakog ispitanika dobiti kompletne podatke. Zato pregledajte šta nedostaje u datoteci s podacima. Pokrenite proceduru **Descriptives** i utvrdite procentualni udeo nedostajućih podataka za svaku promenljivu. Kada određenoj promenljivoj neočekivano nedostaje mnogo podataka, zapitajte se zašto je tako? Trebalo bi da razmislite i o tome da li su nedostajući podaci raspodeljeni nasumično po slučajevima ili tu postoji neka pravilnost (npr. mnoge žene nisu odgovorile na pitanje o svojoj starosti). IBM SPSS ima proceduru **Missing Value Analysis** koja olakšava pronalaženje pravilnosti u nedostajućim podacima. Više o tome pročitajte u 4. poglavlju knjige autorki Tabachnick i Fidell, 2013.

Takođe razmotrite šta ćete s nedostajućim podacima kada dođe vreme da uradite statističke analize. U mnogim IBM SPSS-ovim statističkim procedurama, preko dugmeta **Options** birate način na koji će IBM SPSS tretirati nedostajuće podatke. Birajte pažljivo, pošto time znatno utičete na statističke rezultate. To je posebno važno kada prilažete spisak promenljivih i za sve njih ponavljate iste analize (npr. korelacije između grupe promenljivih, t-testovi niza zavisnih promenljivih).

- Opcija **Exclude cases listwise** znači da će biti analizirani samo slučajevi u kojima za *sve promenljive* navedene u polju **Variables** postoje *svi* podaci. Svaki slučaj kome nedostaje makar i delić podataka uopšte neće biti analiziran. Time biste znatno i nepotrebno ograničili veličinu uzorka.
- Opcija **Exclude cases pairwise** (katkada piše *Exclude cases analysis by analysis*) znači da će slučaj (osoba) biti isključen samo iz onih analiza za koje mu nedostaje neki od neophodnih podataka. Dakle, i takvi slučajevi će se analizirati kad god je to moguće, tj. kad god imaju podatke potrebne za određenu analizu.
- Opcija **Replace with mean**, dostupna u nekim IBM SPSS-ovim statističkim postupcima (npr. višestrukoj regresiji), znači da će biti izračunata srednja vrednost svih promenljivih i da će njome biti zamenjeni nedostajući podaci. Ovu opciju *nikada* ne bi trebalo da koristite, pošto može znatno da iskrivi rezultate analize – naročito onda kada nedostaje mnogo podataka.

Kad god radite neki statistički postupak, pritisnite dugme **Options** i proverite koja je od navedenih opcija potvrđena (jer se podrazumevana opcija menja u zavisnosti od postupka). Ako nemate jak razlog da postupite drugačije, preporučila bih vam da slučajeve isključite samo iz onih analiza za koje im nedostaju podaci (**pairwise**). Jedina situacija kada bi vam moglo zatrebati da analize ograničite samo na slučajeve koji imaju podatke za sve promenljive (**listwise**) jeste ona kada treba razmotriti samo podskup slučajeva koji daje potpun skup rezultata.

Brojevi koji čudno izgledaju

Izlazni rezultat iz IBM SPSS-a ponekad sadrži čudne brojeve oblika 1,24E-02. To su mali brojevi prikazani u tzv. naučnoj notaciji. Da vam se to ne bi dešavalo, u glavnom meniju izaberite **Edit**, zatim **Options**, i potvrdite opciju **No scientific notation for small numbers in tables** na kartici **General**.

Dodatak: Pojedinosti o datotekama s podacima

Ovaj dodatak sadrži informacije o šest datoteka s podacima, dostupnih na pratećoj veb lokaciji knjige (pojedinosti o tome videti na str. ix):

1. survey.sav
2. error.sav
3. experim.sav
4. manipulate.sav
5. depress.sav
6. sleep.sav
7. staffsurvey.sav

Datoteke služe za praćenje postupaka i vežbi opisanih u poglavljima ove knjige. Da biste mogli koristiti te datoteke s podacima, morate posetiti navedenu veb lokaciju i – prateći uputstva na ekranu – preuzeti ih na čvrsti disk, CD ili USB. Zatim bi trebalo da pokrenete IBM SPSS i otvorite onu datoteku s podacima koju nameravate da koristite. Takve datoteke mogu da se otvore samo u IBM SPSS-u. U ovom dodatku je za svaku datoteku dat šifarnik, sa svim pojedinostima o promenljivama i pratećim uputstvima za šifrovanje.

survey.sav

Ovo su stvarni podaci, izdvojeni iz jedne studije koju su sproveli moji postdiplomci psihologije obrazovanja. Studija je trebalo da istraži činioce koji utiču na psihološko prilagođenje i opšte raspoloženje ispitanika. Pripadajuća anketa je obuhvatala razne validirane skale za merenje konstrukata, za koje bi se po obimnoj literaturi o stresu i njegovom prevazilaženju moglo smatrati da utiču na subjektian doživljaj stresa. Pomoću tih skala mere se samopoštovanje, optimizam, subjektivno doživljen stepen samokontrole, tj. upravljanja sopstvenim životom, subjektivno doživljen stres, pozitivna i negativna osećanja i zadovoljstvo životom. Korišćena je i jedna skala za merenje sklonosti ljudi da se predstave u povoljnom ili društveno poželjnom svetlu. Anketa je podeljena pripadnicima opšte populacije u Melburnu i okolnim okruzima Australije. Konačno je dobijen uzorak sa 439 osoba, od čega 42 procenta muškaraca i 58 procenata žena, starosti od 18 do 82 godine (srednja vrednost = 37,4).

error.sav
Podaci u ovoj datoteci dobijeni su modifikovanjem datoteke **survey.sav,** tako što su u nju namerno ubačene greške koje treba otkriti postupcima objašnjenim u 5. poglavlju. Više o njenim promenljivama i drugim detaljima videti u opisu datoteke **survey.sav**.

experim.sav
Ovo je skup izmišljenih podataka, napravljen da bi se dobilo nešto prikladno za ilustraciju statističkih tehnika kao što su t-test za ponovljena merenja i jednofaktorska ANOVA ponovljenih merenja. Taj skup podataka upućuje na fiktivno istraživanje uticaja dve različite intervencije koje pomažu studentima da se izbore sa sopstvenom anksioznošću u vezi s predstojećim kursom iz statistike. Podeljeni u dve jednake grupe, studenti su (u trenutku 1) popunili nekoliko testova. To su bili testovi za ispitivanje straha od statistike (Fear of Statistics), samopouzdanja u vezi s polaganjem ispita iz statistike (Confidence in Coping with Statistics) i depresije (Depression). Potom je jedna od tih grupa (grupa 1) pohađala višečasovni kurs matematičkih veština; druga grupa (grupa 2) učestvovala je u programu izgradnje samopouzdanja u polaganje ispita iz statistike. Nakon završetka oba programa (u trenutku 2) opet su popunjavali iste testove kao pre programa, i ponovo tri meseca kasnije (u trenutku 3). Bile su izmerene i njihove performanse na ispitu iz statistike.

manipulate.sav
Ova datoteka je dodata da bi se mogli vežbati neki od postupaka razmotrenih u poglavlju 8. Ona sadrži podatke unete u obliku teksta (males – muškarci, females – žene), koji se moraju pretvoriti u numeričke podatke radi statističkih analiza. U njoj se nalaze i datumi koji se mogu koristiti pri radu s čarobnjakom Date and Time Wizard. Ti podaci potiču iz bolničke baze podataka o prijemu i otpustu pacijenata.

depress.sav
Ova datoteka je priložena radi prikazivanja određenih tehnika u 16. poglavlju. Obuhvata nekoliko ključnih promenljivih iz jedne stvarne studije, koju je jedan od mojih postdiplomaca sproveo istražujući činioce koji utiču na opšte raspoloženje prvorotki. Sadrži rezultate dobijene na više različitih psiholoških skala projektovanih za merenje depresije (pojedinosti videti u poglavlju 16, o meri podudaranja kapa).

sleep.sav
Stvarni podaci iz studije sprovedene radi istraživanja rasprostranjenosti i uticaja problema sa spavanjem na razne aspekte života. Univerzitetsko osoblje iz Melburna u Australiji popunilo je upitnik s pitanjima o broju sati noćnog sna koje uspevaju da ostvare, problemima sa spavanjem (da li teško zaspu, da li se lako bude) i uticaju tih problema na razne aspekte njihovih života (rad, vožnju, odnose s životnim partnerima). Uzorak je imao 271 ispitanika (55 procenata žena, 45 procenata muškaraca) starosti od 18 do 84 godine (srednja vrednost = 44 god).

staffsurvey.sav

Stvarni podaci, izdvojeni iz studije sprovedene radi istraživanja stepena zadovoljstva osoblja jedne obrazovne ustanove sa ograncima u više mesta po Australiji. Osoblje je popunilo kratak anoniman upitnik (dat kasnije u ovom dodatku), u kojem su dali svoja mišljenja o raznim aspektima organizacije i tretmana koji su imali kao zaposleni.

Deo A: materijali za survey.sav

Pojedinosti o skalama iz datoteke survey.sav

Skale su navedene redom kojim se pojavljuju u anketi.

Skala	Referenca
Life Orientation Test Skala optimizma ili životne orijentacije (*šest stavki*)	Scheier, M.F. & Carver, C.S. (1985). Optimism, coping and health: An assessment and implications of generalized outcome expectancies. *Health Psychology*, 4, 219–47. Scheier, M.F., Carver, C.S. & Bridges, M.W. (1994). Distinguishing optimism from neuroticism (and trait anxiety, self-mastery and self-esteem): A re-evaluation of the Life Orientation Test. *Journal of Personality and Social Psychology*, 67, 6, 1063–78.
Mastery Scale Skala subjektivnog osećaja upravljanja sopstvenim životom (*sedam stavki*)	Pearlin, L. & Schooler, C. (1978). The structure of coping. *Journal of Health and Social Behavior*, 19, 2–21.
Positive and Negative Affect Scale Skala pozitivnih i negativnih osećanja (*dvadeset stavki*)	Watson, D., Clark, L.A. & Tellegen, A. (1988). Development and validation of brief measures of positive and negative affect: the PANAS scales. *Journal of Personality and Social Psychology*, 54, 1063–70.
Satisfaction with Life Scale Skala zadovoljstva životom (*pet stavki*)	Diener, E., Emmons, R.A., Larson, R.J. & Griffin, S. (1985). The Satisfaction with Life scale. *Journal of Personality Assessment*, 49, 71–6.
Perceived Stress Scale Skala subjektivno doživljenog stresa (*deset stavki*)	Cohen, S., Kamarck, T. & Mermelstein, R. (1983). A global measure of perceived stress. *Journal of Health and Social Behavior*, 24, 385–96.
Self-esteem Scale Skala samopoštovanja (*deset stavki*)	Rosenberg, M. (1965). *Society and the adolescent selfimage*. Princeton, NJ: Princeton University Press.
Social Desirability Scale Skala za merenje sklonosti ljudi da se predstave u društveno poželjnom svetlu (*deset stavki*)	Crowne, D.P. & Marlowe, P. (1960). A new scale of social desirability independent of psychochopathology. *Journal of Consulting Psychology*, 24, 349–54. Strahan, R. & Gerbasi, K. (1972). Short, homogeneous version of the Marlowe-Crowne Social Desirability Scale. *Journal of Clinical Psychology*, 28, 191–3.
Perceived Control of Internal States Scale (PCOISS) Skala subjektivno doživljene samokontrole unutrašnjih stanja (*osamnaest stavki*)	Pallant, J. (2000). Development and validation of a scale to measure perceived control of internal states. *Journal of Personality Assessment*, 75, 2, 308–37.

Šifarnik za survey.sav

Ime promenljive u SPSS-u	Puno ime promenljive	Uputstvo za šifrovanje	Merna skala
id	Identifikacioni broj	identifikacioni broj subjekta	scale
sex	Pol subjekta	1=muškarci, 2=žene	nominal
age	Starost subjekta	u godinama	scale
marital	Bračno stanje	1=nevenčan(a), 2=stalan partner, 3=živi s partnerom, 4=prvi put venčan(a), 5=ponovo venčan(a), 6=živi odvojeno od supružnika, 7=razveden(a), 8=udovac/udovica	nominal
child	Dete	1=da, 2=ne	nominal
educ	Najviša završena škola	1=osnovna škola, 2=deo srednje škole 3=cela srednja škola, 4=viša škola ili dodatna obuka, 5=dodiplomske studije na univerzitetu, 6=postdiplomske studije na univerzitetu	ordinal
source	Glavni izvor stresa	1= radno mesto, 2=supružnik iil partner, 3=ljubavni odnosi, 4=deca, 5=porodica, 6=zdravstveno stanje/ bolest, 7=život uopšte, 8=finansijsko stanje, 9=nedostatak slobodnog vremena	nominal
smoke	Pušač	1=da, 2=ne	nominal
smokenum	Broj popušenih cigareta sedmično	broj popušenih cigareta sedmično	scale
op1 do op6	Skala optimizma	1=nimalo se ne slažem, 5=potpuno se slažem	scale
mast1 do mast7	Skala subjektivnog osećaja upravljanja sopstvenim životom	1=nimalo se ne slažem, 4=potpuno se slažem	scale
pn1 do pn20	Skala pozitivnih i negativnih osećanja (PANAS)	1=vrlo malo, 5=izuzetno mnogo	scale
lifsat1 do lifsat5	Skala zadovoljstva životom	1=nimalo se ne slažem, 7=potpuno se slažem	scale
pss1 do pss10	Skala subjektivno doživljenog stresa	1=nikada, 5=vrlo često	scale
sest1 do sest10	Skala samopouzdanja	1=nimalo se ne slažem, 4=potpuno se slažem	scale
m1 do m10	Marlowe-Crowne skala za merenje sklonosti ljudi da se predstave u društveno poželjnom svetlu	1=tačno, 2=netačno	nominal
pc1 do pc18	Skala subjektivno doživljene samokontrole unutrašnjih stanja (PCOISS)	1=nimalo se ne slažem, 5=potpuno se slažem	scale

Ukupni rezultati na skalama koje su deo datoteke survey.sav

Puno ime promenljive	Ime promenljive u SPSS-u	Uputstvo za šifrovanje
Ukupan optimizam	toptim	obrnuti stavke op2, op4, op6 sabrati vrednosti od op1 do op6 raspon od 6 do 30
Ukupan subjektivni osećaj upravljanja sopstvenim životom	tmast	obrnuti stavke mast1, mast3, mast4, mast6, mast7 sabrati vrednosti od mast1 do mast7 raspon od 7 do 28
Ukupna pozitivna osećanja	tposaff	sabrati vrednosti pn1, pn4, pn6, pn7, pn9, pn12, pn13, pn15, pn17, pn18 raspon od 10 do 50
Ukupna negativna osećanja	tnegaff	sabrati vrednosti pn2, pn3, pn5, pn8, pn10, pn11, pn14, pn16, pn19, pn20 raspon od 10 do 50
Ukupno zadovoljstvo životom	tlifesat	sabrati sve vrednosti od lifsat1 do lifsat5 raspon 5 do 35
Ukupan subjektivno doživljen stres	tpstress	obrnuti stavke psp4, pss5, pss7, pss8 sabrati sve vrednosti od pss1 do pss10 raspon od 10 do 50
Ukupno samopouzdanje	tslfest	obrnuti stavke sest3, sest5, sest7, sest9, sest10 sabrati sve vrednosti od sest1 do sest10 raspon od 10 do 40
Ukupna sklonost osobe da se predstavi u društveno poželjnom svetlu	tmarlow	obrnuti stavke od m6 do m10 (rešifrovati na tačno=1, netačno=0) sabrati sve vrednosti od m1 do m10 raspon od 0 do 10
Ukupna subjektivno doživljena samokontrola sopstvenih unutrašnjih stanja	tpcoiss	obrnuti stavke pc1, pc2, pc7, pc11, pc15, pc16 sabrati sve vrednosti od pc1 do pc18 raspon od 18 do 90
Nove obrazovne kategorije	educ2	zbog malog broja slučajeva u svakoj od tih grupa, kategorije Osnovna škola i Deo srednje škole rešifrovane su u jednu grupu: 1=osnovna ili deo srednje škole 2=cela srednja škola 3=viša škola ili dodatna obuka 4=dodiplomske studije na univerzitetu 5=postdiplomske studije na univerzitetu
Starosna grupa, 3 kategorije	agegp3	1=18–29 god, 2=30–44 god, 3=45+ god.
Starosna grupa, 5 kategorija	agegp5	1=18–24 god, 2=25–32 god, 3=33–40 god, 4=41–49 god, 5=50+ god.