

4.

Kreiranje i rekodiranje varijabli

TEME

1. Uvodne napomene

2. Kreiranje varijabli procedurom
Add Computed Column

3. Rekodiranje varijabli funkcijom ifElse

4.1. Uvodne napomene

JASP je analitički orijentiran statistički paket, što znači da JASP posjeduje ograničene opcije za uređivanje, odnosno za kreiranje, rekodiranje i transformiranje podataka. Prema riječima razvojnog tima, više se pažnje posvećuje jednostavnosti i prilagodljivosti (složenih) analitičkih procedura nego rudimentarnim opcijama uređivanja podataka, koje su odveć pokrivene klasičnim statističkim paketima ili, još jednostavnije, Excelom. Drugim riječima, za konkretnu suverenost u obradi i analizi podataka u JASP-u, vrlo će često biti potrebno koristiti još jedan softver, prvenstveno za uređivanje i transformaciju podataka (jer i besplatnost ima svoju cijenu). JASP doduše nudi opciju sinkronizacije podatka s nekim od često korištenih softvera za obradu podataka (SPSS, Excel), što znači da ono što napravite s podacima u Excelu može biti odmah vidljivo u JASP-u i spremno za analizu.

Što to znači u praksi? JASP-ove opcije za uređivanje podataka poprilično su rudimentarne, ali svejedno omogućivanju jednostavnije varijante kreiranja i rekodiranja varijabli.

- a) Kreiranje varijable podrazumijeva izradu nove varijable na temelje jedne i više drugih varijabli koristeći različite matematičke operacije (primjerice, aritmetičko operiranje vrijednostima: zbrajanje, dijeljenje, množenje itd.).

Izrada aditivne varijable na temelju vrijednosti više varijabli je jedan od oblika kreiranja varijabli. Uzmemo li za primjer već korišteni instrument za mjerenje klerikalizma:

9. Koliko se slažete sa svakom od sljedećih tvrdnji?		UOPĆE SE NE SLAŽEM	UGLAVNOM SE NE SLAŽEM	NITI SE SLAŽEM NITI SE NE SLAŽEM	UGLAVNOM SE SLAŽEM	U POTPUNOSTI SE SLAŽEM
1	Kada bi Crkva imala veći politički utjecaj, društvo bi bilo mnogo zdravije.	1	2	3	4	5
2	Zakoni bi se trebali temeljiti na naučavanju Crkve.	1	2	3	4	5
3	Katolička Crkva je jedini istinski moralni autoritet u ovoj zemlji.	1	2	3	4	5
4	Katolička Crkva treba imati predstavnike u Saboru.	1	2	3	4	5

Odgovori li neki ispitanik na pitanja kako je prikazano za slici iznad, njegov će rezultat na novoj aditivnoj varijabli (nazovimo je 'kler') iznositi 10. Drugim riječima, varijablu 'kler' čini zbroj vrijednosti na varijablama p9_1, p9_2, p9_3 i p9_4. Na istom principu možemo kreirati druge varijable raznim matematičkim operacijama.

b) Rekodirati varijablu znači promijeniti jednu ili više vrijednosti te varijable. Primjer toga jest objedinjavanje određenih kategorija neke nominalne varijable.

p1 Regija:

1 – Zagreb i okolica
2 – Sjeverna Hrvatska
3 – Slavonija
4 – Lika i Banovina
5 – Istra, Primorje i Gorski Kotar
6 – Dalmacija



p1r Regija (rekodirano):

1 – Središnja Hrvatska
2 – Istočna Hrvatska
3 – Lika i Dalmacija
4 – Istra, Primorje i Gorski Kotar

Nadalje, često je potrebno rekodirati određene čestice u instrumentu jer im je smjer značenja suprotan od drugih čestica, odnosno od predmeta mjerenja. Na primjer, imamo tri čestice koje mjere klerikalizam:

Neprimjereno je da se svećenici bave politikom.

Zakoni bi se trebali temeljiti na naučavanju Crkve.

Kad bi Crkva imala veći politički utjecaj, društvo bi bilo mnogo zdravije.

Tim je česticama pridružena skala procjene slaganja od pet stupnjeva. Da bismo napravili smislenu kompozitnu mjeru (aditivnu varijablu) klerikalizma, potrebno je rekodirati vrijednosti skale procjene za prvu česticu. Drugim riječima, očekuje se da će ispitanici koji su izrazito skloni klerikalizmu na prvoj čestici odabrati "1 = Uopće se ne slažem", a na druge dvije "5 = U potpunosti se slažem". Zbog toga je za prvu česticu potrebno sve "petice" rekodirati u "jedinice", sve "četvorke" u "dvojke", sve "dvojke" u "četvorke" i sve "jedinice" u "petice", dok "trojke" ostaju nepromijenjene. Takvu vrstu rekodiranja nazivamo refleksijom skale:

1 = 5

2 = 4

3 = 3

4 = 2

5 = 1

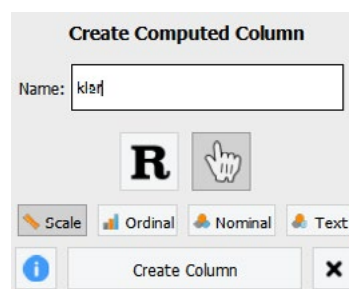
4.2. Kreiranje varijabli procedurom Add Computed Column

Primjer na stranicama JASP-a jednostavnog kreiranja varijable kao razlike vrijednosti na dvjema postojećim varijablama:
https://jasp-stats.org/wp-content/uploads/2018/12/comp_col_1.gif

Radimo na primjerima u datoteci pod nazivom:
[04_kreiranje_rekodiranje_varijabli.sav](#)

U nastavku je prikazana procedura kreiranja aditivne varijabla zbrajanjem odgovara na česticama p9_1, p9_2, p9_3 i p9_4. Varijabla predstavlja skalu (kompozitnu mjeru) klerikalizma, a u podacima je nazvana 'kler'.

Procedura *Add Computed Column* nalazi se u gornjem desnom kutu JASP sučelja.



Aditivna varijabla koju kreirate mora biti predefinicirana kao kvantitativna ('Scale'), ali i varijable od kojih kreirate novu varijablu također moraju biti definirane kao kontinuirane (varijable p9_1 do p9_4 treba definirati kao 'scale').


Nakon što vam se otvori procedura *Add Computed Column*, s lijeve strane nalaze se varijable, iznad je niz aritmetičkih i logičkih operatora, a s desne je strane niz drugih matematičkih i logičkih funkcija. Korištenje ove procedure nalik je korištenju procedure za odabira ispitanika (filter).

U sljedećem koraku – jednu po jednu – uvodimo varijable p9_1 do p9_4, a nakon svake uvedene varijable dodajemo operator +.

Sintaksa za kreiranje varijable 'kler' izgleda ovako:

$$((p9_1 + p9_2) + p9_3) + p9_4$$

Pritiskom na tipku *Compute column* pokrećete proceduru koja rezultira novom varijablom 'kler' na samom kraju niza vaših varijabli.

 f _x kler
4
13
7

Na identičan način, koristeći druge aritmetičke operatore, možemo kreirati nove varijable, koje nisu nužno aditivne varijable (primjerice, vrijednosti neke varijable podijeliti s određenim brojem ili vrijednosti neke varijable preračunati u z-vrijednosti).

4.3. Rekodiranje varijabli funkcijom *ifElse* unutar procedure *Add Computed Column*

Radimo na primjerima u datoteci pod nazivom:
04_kreiranje_rekodiranje_varijabli.sav

Logičku funkciju *ifElse* možemo iskoristiti za rekodiranje varijabli. Funkcija se sastoji od tri uvjeta:

ifElse(test,then,else)

test = ovdje uvodimo polazišnu varijablu i definiramo je kao logički izraz (npr. *kler = 4*)

then = ovdje navodimo vrijednost (ili izraz) koju će ishodišna varijabla zaprimiti ako je prethodni logički izraz (npr. *kler = 4*) **istinit/točan/prisutan**.

else = ovdje navodimo vrijednost (ili izraz) koju će ishodišna varijabla zaprimiti ako prethodni logički izraz (npr. *kler = 4*) **NIJE istinit/točan/prisutan**.

U nastavku je prikazana procedura rekodiranja varijable učestalosti odlazaka na vjerske obrede ('p6') u dihotomnu varijablu 'p6r'.

p6 Ako se izuzmu izuzetne prilike kao vjenčanja, pogrebi, krštenja, ponoćna misa za Božić i slično, koliko često idete na obrede karakteristične za Vašu vjeru?

- 1 – Nisam vjernik
- 2 – Nikada
- 3 – Jednom u nekoliko godina
- 4 – Nekoliko puta godišnje
- 5 – Jednom mjesečno
- 6 – Jednom tjedno
- 7 – Gotovo svakodnevno



p6r Praktična religioznost (rekodirano):

- 1 – Nisu religiozni
- 2 – Religiozni

Rekodiranje možemo napraviti tako da sve vrijednosti varijable 'p6' koje su jednake ili manje od vrijednosti 2 definiramo kao vrijednost 1 u ishodišnoj varijabli 'p6r'.

Koraci:

- 1) Operiramo numeričkim vrijednostima varijable 'p6' i zbog toga je trebamo definirati kao 'scale'.
- 2) Pokrećemo proceduru *Add Computed Column*, novoj (ishodišnoj) varijabli dodjeljujemo naziv 'p6r' te je odmah definiramo kao 'nominal'.
- 3) Unutar procedure, u desnom izborniku pronađemo funkciju *ifElse* i aktiviramo je.
- 4) Odaberemo operator \leq , koji se automatski uvede u *ifElse* funkciju (u uvjet *test*).
- 5) Uvedemo varijablu 'p6' u funkciju *ifElse* (lijevo od operatora \leq), a desno od operatora \leq navedemo vrijednost 2.
- 6) U uvjet *then* upišemo 1, a u uvjet *else* upišemo 2. Sintaksa sada izgleda:

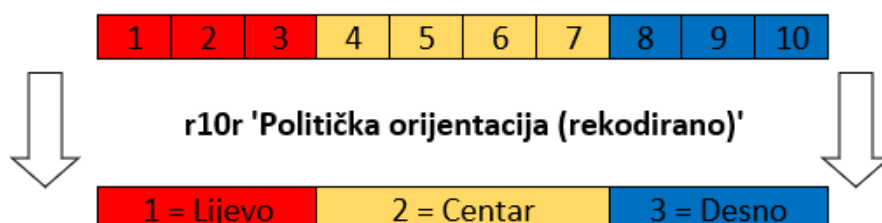
ifElse(p6 ≤ 2,1,2)

Značenje ove sintakse jest: Ako je vrijednost varijable 'p6' manja ili jednaka od 2, varijable p6r zaprimat će vrijednost 1. Ako vrijednost varijable 'p6' **nije** manja ili jednaka od 2, varijabla 'p6r' zaprimat će vrijednost 2.

- 7) Pokrenemo proceduru kojom ćemo dobiti željenu varijablu 'p6r'.
 8) Na kraju varijabli 'p6r' labeliramo vrijednosti (1 = Nisu religiozni i 2 = Religiozni).

$f_{x:p6r}$	p6
1	2
2	4
2	6
1	2
2	4

Moguće je kombinirati više *ifElse* funkcija. Primjerice, ako bismo varijablu 'p10' (Politička orijentacija) željeli rekodirati u varijablu 'p10r' (Politička orijentacija – Lijevo-Desno-Centar)' na način da su određene kategorije agregirane (označeno bojama):



...koristili bismo sljedeću sintaksu *ifElse* funkcije:

```
ifElse(  $\color{red}{\blacktriangleright}$  p10  $\leq$  3, 1, ifElse(  $\color{red}{\blacktriangleright}$  p10  $\geq$  8, 3, 2))
```

U ovoj smo sintaksi u prvu *ifElse* funkciju ugnjezdili drugu *ifElse* funkciju tako da smo u uvjet *else* prve funkcije uvedli dodatnu *ifElse* funkciju. Prva *ifElse* funkcija rekodira 1 = Lijevo, ugnježdjena *ifElse* funkcija rekodira 3 = Desno, a sve što je ostalo (*else* u ugnježdjenoj funkciji) rekodirano je kao 2 = Centar.

Literatura

- Goss-Sampson, M. A. (2019). *Statistička analiza u JASP programu: vodič za studente*. Poglavlje Transformacija podataka.
 URL: http://static.jasp-stats.org/Manuals/Statistic%cc%8cka_analiza_u_JASP_programu_v0.10.2.pdf