

Oblikovanje tablica, upis i organizacija podataka, šifriranje

Lidija Bilić-Zulle
Mladen Petrovečki



Farmaceutsko-biokemijski fakultet Sveučilišta u Zagrebu Biostatistika

Zapis i pohrana podataka

- zapis u izvornom obliku ("sirovi podatci", engl. *row data*)
 - izvorni ispis instrumenta ⇒ oprez: trajnost zapisa (vrst papira)
 - zapis: "papirnati" i elektronički
- pohrana podataka
 - najmanje 10 godina (nakon objavljivanja)



Farmaceutsko-biokemijski fakultet Sveučilišta u Zagrebu Biostatistika

Zapis/pohrana

- prijenos izvornih podatka u podatke pogodne za obradbu
- najčešće: upis u tablice
- oblikovanje dvodimenzijskih tablica



Farmaceutsko-biokemijski fakultet Sveučilišta u Zagrebu Biostatistika

Zapis/pohrana

id	dob	spol	Hb
1	64	1	126
2	58	1	98
3	49	1	107
4	62	2	107
5	34	2	119



Farmaceutsko-biokemijski fakultet Sveučilišta u Zagrebu Biostatistika

Pohrana podataka

- dvodimenzijaska tablica

ispitanici

varijable

	A	B	C	D
1	Ispitanici	Dob	Spol	Hb
2	1	64	1	126
3	2	58	1	98
4	3	49	1	107
5	4	62	2	107
6	5	34	2	119



Farmaceutsko-biokemijski fakultet Sveučilišta u Zagrebu Biostatistika

Pohrana podataka

- jedno istraživanje ⇒ jedna tablica!

	A	B	C	D	E
1	Ispitanici	Skupina	Dob	Spol	Hb
2	1	1	64	1	126
3	2	2	58	1	98
4	3	2	49	1	107
5	4	1	62	2	107
6	5	3	34	2	119



Farmaceutsko-biokemijski fakultet Sveučilišta u Zagrebu Biostatistika

Pohrana podataka

- kvalitativni vs. kvantitativni podatci
- ⇒ ⇒ ⇒ ⇒ ⇒ šifriranje

A	B	C	D	E
1	Ispitanici	Skupina	Dob	Spol
2	1	64	1	126
3	2	68	1	112
4	3	49	m	98
5	4	62	ž	107
6	kontrola	34	ž	119

A	B	C	D	E
1	Ispitanici	Skupina	Dob	Spol
2	1	1	64	1
3	2	2	68	1
4	3	2	49	1
5	4	1	62	2
6	5	3	34	2



Farmaceutsko-biokemijski fakultet Sveučilišta u Zagrebu Biostatistika

Pohrana podataka

- diskretni vs. kontinuirani
- decimalna mjesta kontinuiranih podataka:
 - ne više od 3 znamenke
 - npr. 123,678; 12,378; 0,239
 - eksponent: $0,0013 = 1,3 \times 10^{-3}$
 - preračunavanje jedinica
 - $0,0013 \text{ g} = 1,3 \text{ mg}$



Farmaceutsko-biokemijski fakultet Sveučilišta u Zagrebu Biostatistika

Pohrana podataka

- jednostavni vs. složeni
- programska potpora može izračunati složene podatke (unos formule)
- upisivati samo potrebne podatke

A	B	C	D
No.	masa (kg)	visina (cm)	BMI
1	1	85	26,8
2	2	97	18,2
3	3	75	23,9
4	4	89	30,8
5	5	69	23,3
6	6	87	24,9



Farmaceutsko-biokemijski fakultet Sveučilišta u Zagrebu Biostatistika

Šifriranje

- šifriranje
- šifarnik

A	B	C
angio	ord	angio
1	0 - 1	1 - 1 (progresivni)
2	1 - 2	2 - 2 (stacionarni)
3	2 - 3	3 - 3 (regresivni)
4		4 - 4 (nepostojanje)
5		
6		
7		
8		
9		
10		



Farmaceutsko-biokemijski fakultet Sveučilišta u Zagrebu Biostatistika

Upitnik – pitanja otvorenog tipa

- Godina rođenja: _____
- Spol: _____
- Koju ste školu završili: _____
- Zaposlenje: _____
- Bračni status: _____
- Koliko dugo imate promjene na koži zbog kojih ste danas došli dermatologu? _____



Farmaceutsko-biokemijski fakultet Sveučilišta u Zagrebu Biostatistika

Upitnik – pitanja zatvorenog tipa

- Godina rođenja: _____ Spol: M Ž
- Stručna sprema:
- osnovna škola i manje
 - srednja škola
 - viša škola
 - visoka škola
- Zaposlenje:
- nezaposlen
 - zaposlen
 - umirovljenik
 - poluprivrednik (domaćin)
- Bračni status:
- udat/oženjen
 - neudat/neoženjen
 - razveden
 - udovac/udovica
- Koliko dugo imate promjene na koži zbog kojih ste danas došli dermatologu?
- manje od 1 mjesec
 - 1-6 mjeseci
 - 6 mjeseci do 1 godine
 - Više od 1 godine

Gdje pripada onaj koji ima promjene 6 mjeseci?



Farmaceutsko-biokemijski fakultet Sveučilišta u Zagrebu Biostatistika

Šifriranje odgovora na upitnik

5. Gdje pretežno upotrebljavate računalo?

- 1 a) ne upotrebljavam računalo
- 2 b) na radnom mjestu
- 3 c) kod kuće
- 4 d) i na radnom mjestu i kod kuće



Farmaceutsko-biokemijski fakultet Sveučilišta u Zagrebu Biostatistika

Odgovori, upitnik

Molimo Vas da iznesete svoje stavove o sljedećim tvrdnjama. Stav izkažete zaokruživanjem broja SAMO JEDNOG od ponuđenih odgovora:

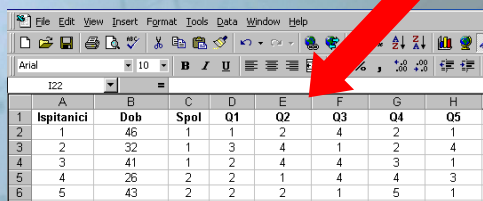
- 1 – uspeće se ne složem
- 2 – uglavnom se ne složem
- 3 – nemam utvrđen stav (niti se ne složem, niti sam suglasan)
- 4 – uglavnom sam suglasan
- 5 – savršeno sam suglasan

1. NE osjećam se ugroženo dok se razgovara o računalima.	1	2	3	4	5
2. Računala su nepouzdana.	1	2	3	4	5
3. Formiranje rada na računalu korisno je vještina.	1	2	3	4	5
4. Volim upotrebljavati računalo.	1	2	3	4	5
5. Način rada računala potpuno mi je nerazumljiv.	1	2	3	4	5
6. Uporaba računala potpuno mi je nerazumljiva.	1	2	3	4	5
7. Pouzdatost podataka o bolesnicima ovisna je o uporabi računala.	1	2	3	4	5



Farmaceutsko-biokemijski fakultet Sveučilišta u Zagrebu Biostatistika

Unos odgovora iz upitnika



	A	B	C	D	E	F	G	H
1	Ispitanici	Dob	Spol	Q1	Q2	Q3	Q4	Q5
2	1	46	1	1	2	4	2	1
3	2	32	1	3	4	1	2	4
4	3	41	1	2	4	4	3	1
5	4	26	2	2	1	4	4	3
6	5	43	2	2	2	1	5	1



Farmaceutsko-biokemijski fakultet Sveučilišta u Zagrebu Biostatistika

Binarni podatci

- 0 – nema
- 1 – ima



Farmaceutsko-biokemijski fakultet Sveučilišta u Zagrebu Biostatistika

Pristranost (iskrivljenje)

- engl. *bias*
- sustavna pogreška tijekom istraživanja
- tijekom prikupljanja podataka:
 - pristranost odziva, prisjećanja, praćenja i gubitka ispitanika
 - pristranosti pri mjerenju, nadziranju, prikupljanju i tumačenju podataka
 - socijalna poželjnost odgovora



Farmaceutsko-biokemijski fakultet Sveučilišta u Zagrebu Biostatistika

Provjera unosa podataka

- uređivanje (sortiranje)
- obročavanje (listića upitnika)
- ponovna provjera 10% unesenih podataka
- neočekivani rezultati obradbe ⇒ moguća grješka unosa podataka

Check, recheck, double check and check again!



Farmaceutsko-biokemijski fakultet Sveučilišta u Zagrebu Biostatistika

Provjera unosa podataka

The screenshot shows an Excel spreadsheet with columns A, B, C, and D. Column A contains patient numbers (1-6), column B contains mass in kg (85, 97, 75, 89, 69, 87), column C contains height in cm (178, 182, 177, 170, 172, 187), and column D contains BMI values (26.6, 29.3, 23.9, 23.1, 23.3, 24.9). A formula bar shows $B6/(C6^2)*1000$. To the right, a 'Variable: dob' summary statistics box is displayed, showing sample size (182), lowest value (34), highest value (630), arithmetic mean (50.8197), and 95% CI for the mean (46.6296 to 55.0098).

Farmaceutsko-biokemijski fakultet Sveučilišta u Zagrebu *Biostatistika*

Povjerljivost podataka

- najviši etički i humani standardi
 - informirani pristanak
 - odobrenje Etičkog povjerenstva
- ne smije se prijetiti, nagovarati, sugerirati niti utjecati na ispitanike
- svi osobni podatci su povjerljivi
 - šifriranje (ispitanici, istraživači, ustanove...)
- nitko neovlašten ne smije pristupiti podacima

Povjerljivost podataka – šifriranje

The screenshot shows two versions of an Excel spreadsheet. The left version, labeled 'F13', has columns A (Ispitanici), B (Dob), C (Spol), and D (Hb) with data for six patients. A large red 'X' is drawn over the data. The right version, labeled 'F16', shows the same data but with the first column (Ispitanici) encrypted with numbers (1-6). A large green checkmark is drawn over the right version.

Farmaceutsko-biokemijski fakultet Sveučilišta u Zagrebu *Biostatistika*

Povjerljivost podataka – pristup

The screenshot shows an Excel spreadsheet with columns A (Ispitanici), B (Dob), C (Spol), and D (Hb). A dialog box titled 'Odobrenje pristupa' (Access Approval) is open, showing a list of users and their access levels. A large green checkmark is drawn over the dialog box.

Farmaceutsko-biokemijski fakultet Sveučilišta u Zagrebu *Biostatistika*

Prikrivanje

- maskiranje (engl. *blinding*)
- osigurava objektivnost istraživanja

Prikrivanje

- Vrst prikrivanja
 - Tko ne zna
- | | | |
|-----------------|---|--|
| – jednostruko | ⇒ | ispitanik |
| – dvostruko | ⇒ | ispitanik i istraživač |
| – trostruko | ⇒ | ispitanik, istraživač i statističar |
| – četverostruko | ⇒ | ispitanik, istraživač, statističar i pisac |

Kada je prikriivanje nemoguće?

- razvidni postupci
- ugrožena sigurnost ispitanika



Farmaceutsko-biokemijski fakultet Sveučilišta u Zagrebu *Biostatistika*

Programi za pohranu podataka

- programi za rad s proračunskim tablicama
 - opći (MS Excel)
 - programi za raščlambu podataka statističkim postupcima (MedCalc, SPSS, Statistica...)
- programi za rad s bazama podataka (MS Access)
- međusobna podudarnost
 - moguće preslikavanje podataka



Farmaceutsko-biokemijski fakultet Sveučilišta u Zagrebu *Biostatistika*

Preslikavanje podataka

Q	R	S	T	U	V
1	3	1	1	1	1
2	3	1	1	1	1
3	2	0	0	0	0
4	2	1	1	1	1
5	3	2	1	2	1
6	3	1	1	1	1
7	2	0	0	0	0
8	3	1	1	1	1
9	3	0	0	0	0
10	3	1	1	1	1
11	3	1	1	1	1
12	2	0	0	0	0
13	3	1	1	1	1
14	2	1	2	1	1
15	1	0	0	0	0
16	3	1	1	1	1
17	3	2	1	2	1

C	D	E	F	G	H	I
1	1	2	3	46	1	46
2	0	1	2	54	1	54
3	0	1	2	55	2	55
4	1	1	2	57	2	57
5	1	2	3	64	1	64
6	1	1	2	48	1	48
7	0	0	1	88	1	88
8	1	1	2	64	1	64
9	0	1	2	51	1	51
10	1	1	2	58	1	58
11	0	2	3	47	1	47
12	1	2	3	64	1	64
13	1	2	3	46	1	46
14	0	2	3	46	1	46
15	1	2	3	38	1	38
16	1	1	2	53	1	53
17	0	1	2	62	1	62



Farmaceutsko-biokemijski fakultet Sveučilišta u Zagrebu *Biostatistika*

Podatci spremni...

- statistička analiza podataka...



Farmaceutsko-biokemijski fakultet Sveučilišta u Zagrebu *Biostatistika*

Pitanja S1

1. Kako vrst podataka određuje vrst istraživanja? Koje vrste izvora podataka koristimo u istraživanjima?
2. Koji su načini očuvanja povjerljivosti podataka?
3. Koji postupak osigurava objektivnost istraživanja i koje su vrste takvog postupka?
4. Kako se pravilno oblikuje tablica za unos podataka? Navedite primjer.
5. Koji su postupci provjere unosa podataka u tablice?
6. Koju sve računalnu programsku potporu rabimo za rad s podacima? Navedite primjere.
7. Što je binariziranje podataka? Navedite primjer.
8. Kako se prikupljaju i pohranjuju izvorni podatci mjerenja. Koliko je dugo obvezno njihovo čuvanje?
9. Koje se pristranosti u istraživanju mogu javiti tijekom prikupljanja podataka?
10. Što su zbunjujući čimbenici i na koji se način izbjegava njihov utjecaj?



Farmaceutsko-biokemijski fakultet Sveučilišta u Zagrebu *Biostatistika*