

Tema: Transformacije podataka



Transformacije podataka

z - vrijednosti

Postupak standardizacije provodi se pomoću formule

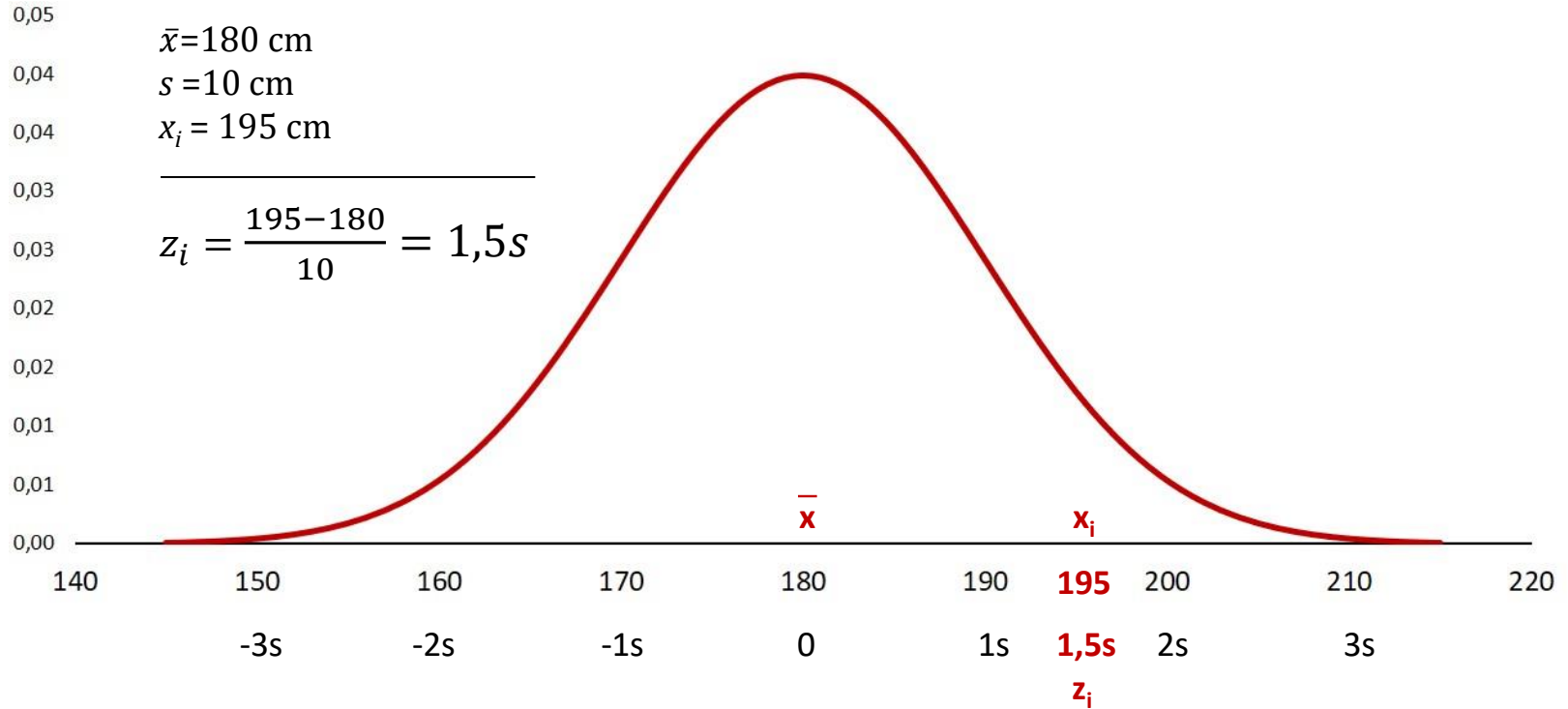
$$z_i = \frac{x_i - \bar{x}}{s}$$

gdje je:

- z_i – standardizirani rezultat entiteta i
- x_i – originalna vrijednost entiteta i
- \bar{x} – aritmetička sredina
- s – standardna devijacija.

Transformacije podataka

z - vrijednosti



Transformacije podataka

z - vrijednosti

Primjer: Deset učenika natjecalo se u tri atletske discipline:

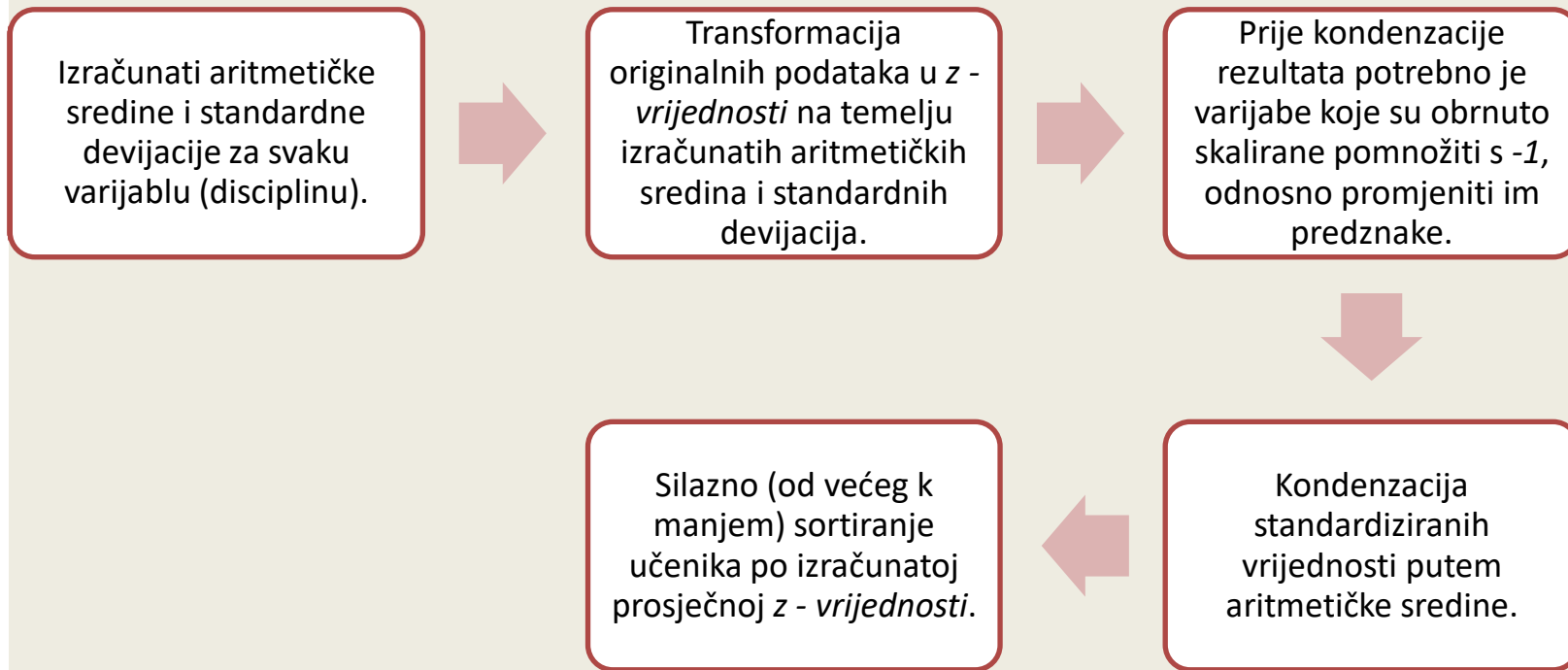
- skok udalj (SD),
- trčanje na 100 metara (T100m) i
- bacanje kugle (BK) i postiglo rezultate navedene u tablici.

Potrebno je utvrditi ukupan poredak (rang) ovog natjecanja ?

Učenik	SD	T100m	BK
AB	359	13,6	561
DF	321	13,9	550
JG	346	13,7	538
KL	332	14,0	490
DD	450	12,2	518
ED	314	14,1	551
TB	410	12,5	589
ZN	425	12,3	602
RG	369	13,5	547
EN	378	13,8	510

Transformacije podataka

z - vrijednosti



Transformacije podataka

z - vrijednosti

Prvi korak: Izračunati aritmetičke sredine i standardne devijacije za svaku varijablu (disciplinu).

	SD	T100m	BK
\bar{x}	370,4	13,36	545,6
s	45,66	0,73	34,21

Učenik	SD	T100m	BK
AB	359	13,6	561
DF	321	13,9	550
JG	346	13,7	538
KL	332	14,0	490
DD	450	12,2	518
ED	314	14,1	551
TB	410	12,5	589
ZN	425	12,3	602
RG	369	13,5	547
EN	378	13,8	510

Transformacije podataka

z - vrijednosti

Drugi korak: Transformacija originalnih podataka u *z - vrijednosti* na temelju izračunatih aritmetičkih sredina i standardnih devijacija.

	SD	T100m	BK
\bar{x}	370,4	13,36	545,6
s	45,66	0,73	34,21

$$z_{AB,SD} = \frac{359 - 370,4}{45,66} = -0,25s$$

Učenik	SD	T100m	BK
AB	359	13,6	561
DF	321	13,9	550
JG	346	13,7	538
KL	332	14,0	490
DD	450	12,2	518
ED	314	14,1	551
TB	410	12,5	589
ZN	425	12,3	602
RG	369	13,5	547
EN	378	13,8	510

Transformacije podataka

z - vrijednosti

Drugi korak: Transformacija originalnih podataka u *z - vrijednosti* na temelju izračunatih aritmetičkih sredina i standardnih devijacija.

	SD	T100m	BK
\bar{x}	370,4	13,36	545,6
s	45,66	0,73	34,21

$$z_{AB,SD} = \frac{359 - 370,4}{45,66} = -0,25s$$

Učenik	SD	T100m	BK
AB	-0,25	0,33	0,45
DF	-1,08	0,74	0,13
JG	-0,53	0,46	-0,22
KL	-0,84	0,87	-1,63
DD	1,74	-1,58	-0,81
ED	-1,24	1,01	0,16
TB	0,87	-1,17	1,27
ZN	1,20	-1,44	1,65
RG	-0,03	0,19	0,04
EN	0,17	0,60	-1,04

Transformacije podataka

z - vrijednosti

Treći korak: Prije kondenzacije rezultata varijable koje su obrnuto skalirane pomnožiti s -1 , odnosno rezultatima promijeniti predznake.

Učenik	SD	T100m	BK
AB	-0,25	0,33	0,45
DF	-1,08	0,74	0,13
JG	-0,53	0,46	-0,22
KL	-0,84	0,87	-1,63
DD	1,74	-1,58	-0,81
ED	-1,24	1,01	0,16
TB	0,87	-1,17	1,27
ZN	1,20	-1,44	1,65
RG	-0,03	0,19	0,04
EN	0,17	0,60	-1,04

Transformacije podataka

z - vrijednosti

Treći korak: Prije kondenzacije rezultata varijable koje su obrnuto skalirane pomnožiti s -1 , odnosno rezultatima promijeniti predznake.

Učenik	SD	T100m	BK
AB	-0,25	-0,33	0,45
DF	-1,08	-0,74	0,13
JG	-0,53	-0,46	-0,22
KL	-0,84	-0,87	-1,63
DD	1,74	1,58	-0,81
ED	-1,24	-1,01	0,16
TB	0,87	1,17	1,27
ZN	1,20	1,44	1,65
RG	-0,03	-0,19	0,04
EN	0,17	-0,60	-1,04

Transformacije podataka

z - vrijednosti

Četvrti korak: Kondenzacija standardiziranih vrijednosti putem aritmetičke sredine.

$$\begin{aligned}\bar{z}_{AB} &= \frac{z_{AB,SD} + z_{AB,T100m} + z_{AB,BK}}{3} = \\ &= \frac{-0,25 + (-0,33) + 0,45}{3} = -0,04\end{aligned}$$

Učenik	SD	T100m	BK	\bar{z}
AB	-0,25	-0,33	0,45	-0,04
DF	-1,08	-0,74	0,13	
JG	-0,53	-0,46	-0,22	
KL	-0,84	-0,87	-1,63	
DD	1,74	1,58	-0,81	
ED	-1,24	-1,01	0,16	
TB	0,87	1,17	1,27	
ZN	1,20	1,44	1,65	
RG	-0,03	-0,19	0,04	
EN	0,17	-0,60	-1,04	

Transformacije podataka

z - vrijednosti

Četvrti korak: Kondenzacija standardiziranih vrijednosti putemaritmetičke sredine.

$$\begin{aligned}\bar{z}_{AB} &= \frac{z_{AB,SD} + z_{AB,T100m} + z_{AB,BK}}{3} = \\ &= \frac{-0,25 + (-0,33) + 0,45}{3} = -0,04\end{aligned}$$

Učenik	SD	T100m	BK	\bar{z}
AB	-0,25	-0,33	0,45	-0,04
DF	-1,08	-0,74	0,13	-0,56
JG	-0,53	-0,46	-0,22	-0,41
KL	-0,84	-0,87	-1,63	-1,11
DD	1,74	1,58	-0,81	0,84
ED	-1,24	-1,01	0,16	-0,70
TB	0,87	1,17	1,27	1,10
ZN	1,20	1,44	1,65	1,43
RG	-0,03	-0,19	0,04	-0,06
EN	0,17	-0,60	-1,04	-0,49

Transformacije podataka

z - vrijednosti

Peti korak: Silazno (od većeg k manjem) sortiranje učenika po izračunatoj prosječnoj z - vrijednosti.

Učenik	RANG	\bar{z}
ZN	1	1,43
TB	2	1,10
DD	3	0,84
AB	4	-0,04
RG	5	-0,06
JG	6	-0,41
EN	7	-0,49
DF	8	-0,56
ED	9	-0,70
KL	10	-1,11

Transformacije podataka

= **NORM.S.DIST (z; TRUE)**

Primjer: 257 dječaka je izmjereno testom za procjenu eksplozivne snage *Skok udalj s mjesta*. Aritmetička sredina iznosi 215 cm, a standardna devijacija 12 cm. Učenik XY postigao je rezultat 230 cm. Potrebno je procijeniti postotak (%) i broj učenika koji postižu lošije i bolje rezultate od učenika XY.

$$z_{XY} = \frac{230-215}{12} = 1,25s$$

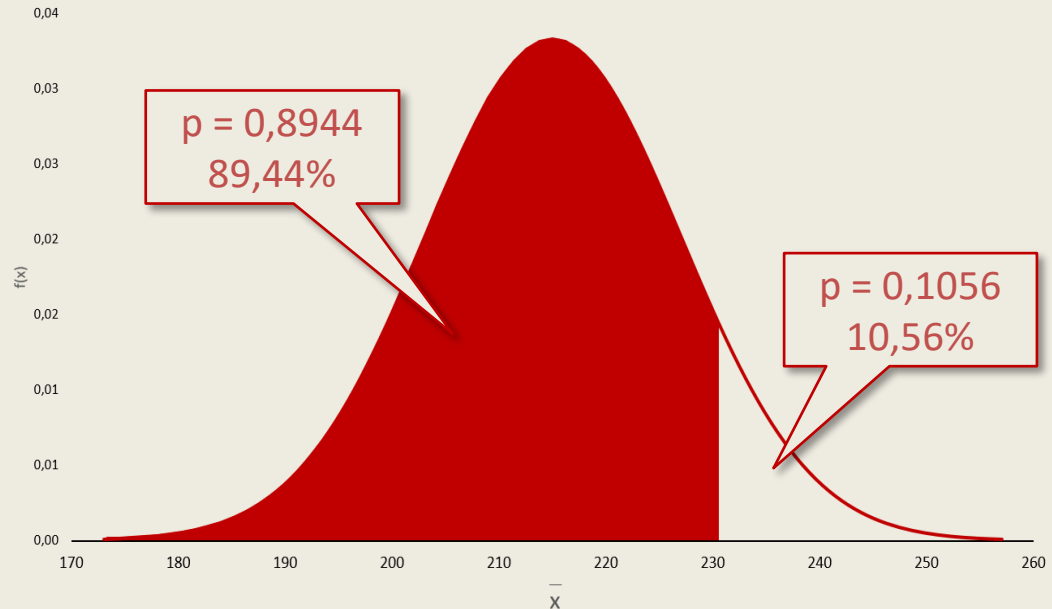
$$p = \text{NORM.S.DIST}(1,25; \text{TRUE})$$

$$p = 0,8944 - \text{lošijih}$$

$$1 - p = 0,1056 - \text{boljih}$$

$$p \cdot n = 0,1056 \cdot 257 \approx 27 - \text{boljih}$$

$$257 - 27 = 230 \text{ lošijih}$$



Transformacije podataka

z - vrijednosti

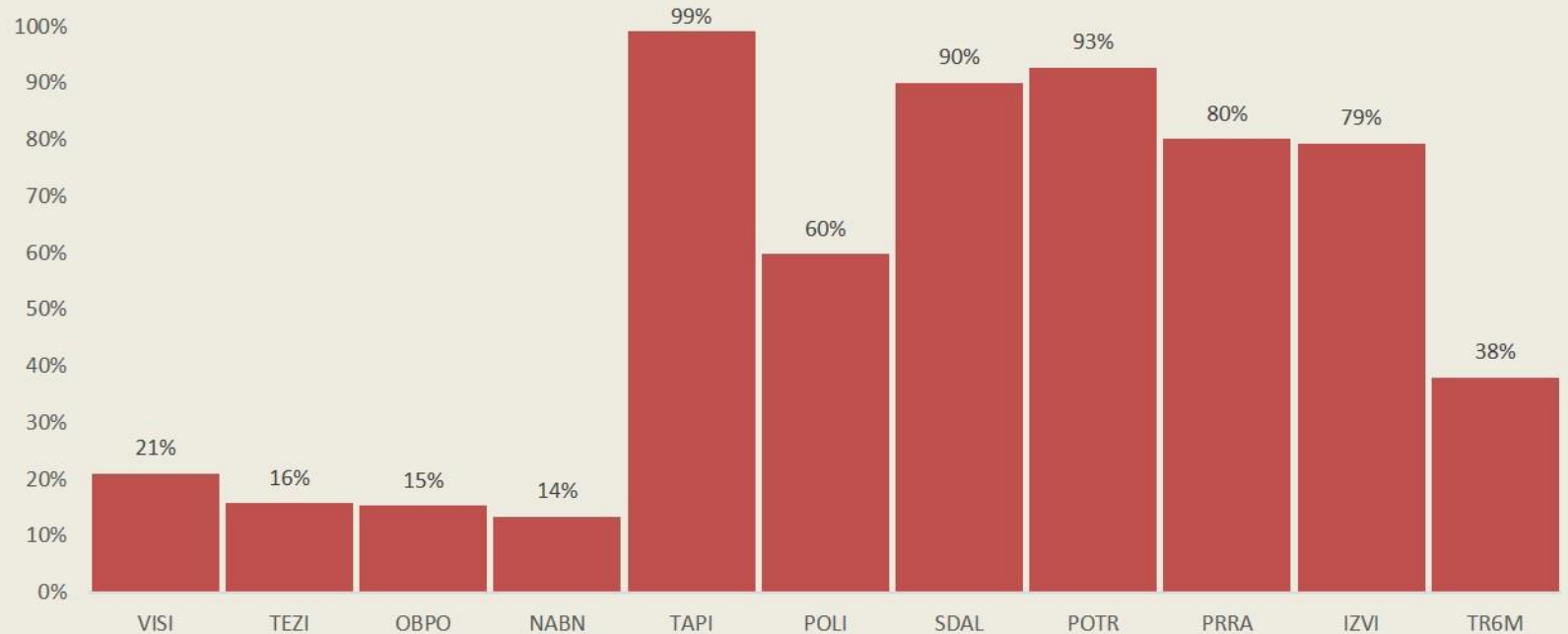
Praktična korist od standardizacije rezultata ogleda se i u mogućnosti grafičkog prikazivanja rezultata entiteta u većem broju varijabli koje opisuju njegov antropološki profil.



Transformacije podataka

z - vrijednosti

Praktična korist od standardizacije rezultata ogleda se i u mogućnosti grafičkog prikazivanja rezultata entiteta u većem broju varijabli koje opisuju njegov antropološki profil.



Transformacije podataka

Ostale transformacije podataka

Standardizirane rezultate moguće je transformirati na različite načine, zavisno o potrebi. Najčešće se transformiraju u vrijednosti na sljedećim skalama:

- skala školskih ocjena (1-5): $ocjena_i = 3 + 0,83 \cdot z_i$
- skala T - skorova (20-80): $T - skor_i = 50 + 10 \cdot z_i$

Osim navedenih transformacija podataka u kineziološkoj praksi moguće je koristiti i sljedeće transformacije za:

- za normalno skalirane varijable: $p_i = \frac{x_i - x_{min}}{R}$
- za obrnuto skalirane varijable: $p_i = \frac{x_{max} - x_i}{R}$; gdje je $R = x_{max} - x_{min}$

Opisanim se postupcima rezultati ispitanika transformiraju u relativne vrijednosti u odnosu na totalni raspon pri čemu 0 predstavlja najlošiji, a 1 najbolji rezultat.

Ovako dobivene vrijednosti moguće je transformirati u skalu s proizvoljno određenim maksimalnim brojem bodova formulom: $bod_i = bod_{max} \cdot p_i$