

## OBDELAVA PODATKOV

### Razpršenost podatkov in škatla z brki.

**O tem si preberi tudi v učbeniku na strani 208 do 209 (oglej si tudi rešene primer na stani 210 in 211).**

*Tu je razlaga in povzetek tega, kar moraš vedeti o razpršenosti podatkov in njihovi predstavitvi.*

V gradivu o srednjih vrednostih naborov podatkov smo videli, da eno samo število – **srednja vrednost** – **včasih težko povzame značilnosti celega nabora**. Če temu podatku **dodamo še podatek o razpršenosti**, je slika že **jasnejša**. V tem primeru vemo, kolikšna je srednja vrednost nabora in ali lahko podatke iz nabora pričakujemo blizu te srednje vrednosti ali daleč stran.

#### ZATO JE POMEMBNO POZNATI V NABORU TUDI:

- Variacijski razmik** nekega nabora je **razlika med največjo in najmanjšo vrednostjo v naboru** (to nam zadošča za nabore z majhnim številom podatkov).

#### KAKO DOLOČIMO VARIACIJSKI RAZMIK IN KAJ NAM POVE. PRIMER:

Pri matematiki je Grega dobil ocene 4, 2, 4, 5, 2, 1, 4, Špela pa 4, 5, 4, 4, 4, 5, 5.

Izračunajmo najprej **mediano obeh naborov** ocen. Pri Gregu dobimo (saj veš, tisti postopek z urejanjem podatkov in odčitavanjem srednjega podatka) 4, pri Špeli pa prav tako 4.

Izračunajmo še **variacijski razmik**. V primeru nabora Gregovih ocen je najvišja ocena 5, najnižja pa 1, zato je **variacijski razmik enak  $5-1 = 4$** . V primeru Špelinih ocen je ta enak  **$5-4 = 1$** .

Gregove ocene so bistveno spremenljivejše in bolj razpršene ... Oba učenca imata enako mediano ocen, variacijska razmika pa se zelo razlikujeta,

	Mediana	Variacijski razmik
Grega	4	4
Špela	4	1

Četudi ne bi poznali začetnega nabora podatkov, bi z informacijami iz zgornje tabele dobili kar dobro predstavo o dogajanju. **KAJ LAHKO SKLEPAMO?**

Pri Gregu lahko pričakujemo oceno 4, vendar pa nas tudi nobena druga ocena ne sme presenetiti.

Pri Špeli bi bila ocena, bistveno različna od 4, veliko presenečenje.

#### 2. Medčetrtnski razmik

Če imamo v naboru malo podatkov, nam variacijski razmik morda še zadošča. **Pri velikem številu** (recimo milijon) **podatkov** pa nas zmoti dejstvo, da je variacijski razmik odvisen samo od dveh ekstremnih podatkov: največjega in najmanjšega; vsi preostali v izračunu ne sodelujejo. Želeli bi si mero razpršenosti podatkov, ki bi upoštevala več podatkov.

Nabor podatkov uredimo od najmanjšega do največjega in z mediano nabora razdelimo v dva nova nabora (levo od naše mediane je prvi nabor in desno od nje drugi nabor podatkov).

Če je število podatkov sodo, v prvi nabor vzamemo spodnjo polovico podatkov, v drugi nabor pa zgornjo polovico. Če je število podatkov liho, postopamo podobno, le da tokrat osrednji podatek – mediano – vzamemo dvakrat, tako da nastopa tako v prvem kot v drugem naboru.

V vsakem novem naboru ponovimo postopek določanja mediane. **Mediano prvega nabora označimo s  $Q_1$ . Mediano drugega nabora označimo s  $Q_3$ . Število  $Q_1$  imenujemo prvi kvartil, število  $Q_3$  pa tretji kvartil.**

Razliko  $Q_3 - Q_1$  imenujemo **medčetrtnski razmik**.

#### KAKO DOLOČIMO MEDČETRRTINSKI RAZMIK. PRIMER:

Oglejmo si primer nabora z lihim številom podatkov: 1, 2, 2, 4, 5, 6, 8, 9, 9.

- Spet ga razdelimo na dva dela. Mediana celotnega nabora je 5.
- Ker je podatkov 9, osrednje število – mediano – vzamemo tako k prvi kot drugi polovici podatkov.
- Dobimo dva nabora: 1, 2, 2, 4, 5 in 5, 6, 8, 9, 9.
- Mediana prvega nabora je 2, mediana drugega 8, zato je
- **medčetrtnski razmik enak njuni rali - 6.**

*Še kratek komentar o imenih kvartil in medčetrtnski razmik.*

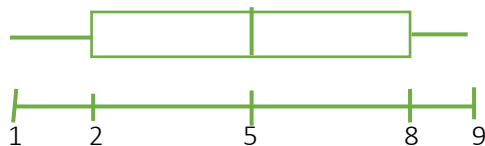
*Spomnimo se, kako pridemo do prvega kvartila - števila  $Q_1$ . To je mediana spodnje polovice nabora, zato leži nekako na polovici prve polovice nabora, torej nekako na četrtnini celotnega nabora. (Beseda kvart nakazuje četrtno, prvi kvartil pa prvo četrtno.) Podobno leži tretji kvartil  $Q_3$  nekako na treh četrtninah nabora. Razmik med tema dvema vrednostma zato imenujemo medčetrtnski razmik.*

#### IN KJE JE ŠKATLA Z BRKI 😊

Škatla z brki (škatlasti diagram) nastane kot pravokotnik, ki zajame celotne podatke med 1. in 3. kvartilom. V pravokotniku postavimo navpičnico, ki označuje mediano nabora. Vodoravni daljici, ki »štrlita« levo in desno iz našega pravokotnika, pa segata na najmanjšega podatka – levo in največjega podatka – desno.

#### PA NARIŠIMO TAK DIAGRAM IZ ZGORNJEGA PRIMERA:

Oglejmo si primer nabora z lihim številom podatkov: 1, 2, 2, 4, 5, 6, 8, 9, 9.



V NALOGI izpiši urejene podatke in :

- Izračunaj vse tri srednje vrednosti.
- Izračunaj variacijski razmik.
- Določi 1. in 3. kvartil.
- Izračunaj mečetrtinski razmik.
- Nariši škatlasti diagram.
- Zapiši komentar.

V košarkarskem moštvu merijo košarkarji 185, 191, 195, 212, 204, 182, 194, 196, 185, 202, 209, 203 in 215 cm.

Za vsa vprašanja in dodatna pojasnila se lahko obrneš na svojo učiteljico matematike:

- Polona Boldin: [polona.boldin@ucitelj.oskm.si](mailto:polona.boldin@ucitelj.oskm.si)
- Tina Klavs Kožuh: [tina.klavskozuh@ucitelj.oskm.si](mailto:tina.klavskozuh@ucitelj.oskm.si)
- Vojka Vilfan: [vojka.vilfan@ucitelj.oskm.si](mailto:vojka.vilfan@ucitelj.oskm.si)