

Peteljka-list dijagram. Medijan i mod



Sanja Varošanec, Zagreb

Citirajmo dio NOK-a koji se odnosi na učeničke ishode vezane uz matematički koncept **Podatci** u ciklusima koji odgovaraju višim razredima osnovne škole:

2. ciklus

Učenici će:

- prikupiti, razvrstati i organizirati podatke te ih na prikidan način prikazati tablicom, tablicom frekvencija, piktogramom, stupčastim i kružnim dijagramom te sustavnom listom
- pročitati i protumačiti podatke prikazane tablicama, slikama, listama te različitim grafovima i dijagramima
- odrediti i primjeniti aritmetičku sredinu, raspon i medijan niza numeričkih podataka.

3. ciklus

Učenici će:

- prikupiti, klasificirati i organizirati podatke te ih na prikidan način, s pomoću računala i bez njega, prikazati sustavnom listom, tablicom, tablicom frekvencija, linijskim, stupčastim i kružnim dijagramom, grafikonom, "brkatom kutijom" (*box and whiskers diagram*) i grafom
- pročitati, tumačiti i analizirati podatke prikazane na različite načine
- odrediti i primjeniti frekvenciju i relativnu frekvenciju za dane podatke te aritmetičku sredinu, medijan, kvartile, mod, raspon i interkvartilni raspon niza numeričkih podataka.

Većina se od ovih ishoda već postiže kroz postojeći nastavni program, ali pojavile su se i neke nove teme. Nekoliko sljedećih članaka posvetit ćemo

ovom dijelu matematike koji se intenzivno koristi u različitim područjima kao što su ekonomija, medicina, tehničke struke itd, zapravo u svim strukama u kojima se prikupljaju i analiziraju podaci različitih vrsta.

Peteljka-list dijagram

Osim prikaza podataka s pomoću tablice frekvencija i različitim grafičkim prikazima koji se već obrađuju u 7. razredu, postoji još jedan vrlo zoran prikaz — peteljka-list dijagram (engl. *stem and leaf diagram*). Ilustrirajmo ga primjerom.

Primjer 1. Učenicima 7.c razreda izmjerena je visina i dobivene su ove visine izražene u centimetrima:

145 147 156 181 163 148 176 156 155 162
155 177 152 165 165 171 174 168 170 163

Ovo je skup **primarnih** podataka. Podatci su numerički (brojčani, kvantitativni), ali skup je nesređen. Iz ovog prikaza skupa podataka teško je odgovoriti na pitanja koja se prirodno javlaju kao što su: koja je visina najveća, koja najmanja, postoji li neka visina koja se javlja mnogo puta ili oko koje se grupiraju visine itd. Stoga ove podatke grupirajmo u ovisnosti o tome kojoj dekadi pripadaju. Tako će podatci 156, 156, 155, 155, 152 pripadati jednoj grupi, a pri zapisu ćemo iskoristiti činjenicu da se broj 15 javlja u svih pet brojeva. Dakle, zapisat ćemo ih ovako:

15 | 66552

odnosno u rastućem poretku:

15 | 25566.

Broj 15 je peteljka zapisa, a znamenke jedinica su lišće. Zapišimo sve podatke ovako.

14	5	7	8
15	2	5	5
16	2	3	3
17	0	1	4
18	1		

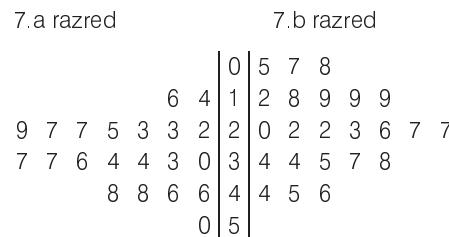
Uz peteljka-list dijagram obavezno zapisujemo i legendu: 14|5 je 145 cm.

Zamijetimo da lišće može, ali i ne mora biti u rastućem poretku. Ovakav je prikaz posebno pogodan za usporedbu dvaju skupova numeričkih podataka.

Primjer 2. Dani su rezultati testa iz hrvatskog jezika u 7.a i 7.b razredu. Nacrtajmo peteljka-list dijagram.

7.a:	39 50 23 25 27 23 22 34 30 46 37 46 16 14 27 29 34 36 37 37 48 48 33
7.b:	34 35 5 38 45 44 12 7 19 19 23 22 46 26 27 27 8 19 22 34 37 18 20

Budući da se u oba razreda rezultati kreću od 0 do 50, napraviti ćemo peteljka-list dijagrame tako da im "peteljka" bude zajednička.



Legenda: 1|2 je 12 bodova.

Kad su podatci ovako zorno prikazani, odmah možemo i ponešto zaključiti o ovim rezultatima: u 7.a razredu najlošiji rezultat je 14 bodova, dok su u 7.b razredu čak tri rezultata ispod 10 bodova. U 7.a postignut je maksimalni broj bodova, što se nije dogodilo u 7.b. Međutim, rezultati u 7.b su manje raspršeni i imaju tendenciju grupiranja u nižim vrijednostima. Kod rezultata 7.a razreda ističe se grupa od 30 do 39 u kojoj se nalazi najveći broj podataka.

Kad imamo numeričke podatke prirodno se naće pitanje kolika je prosječna vrijednost tih rezultata, odnosno kolika je aritmetička sredina skupa podataka. S ovom su mjerom učenici vrlo dobro upoznati jer često računaju prosjek svojih ocjena (iako je sa stanovišta značenja ocjene kao rangiranog podatka to nepravilno). Prosječni rezultat u

metodika

7.a razredu u prethodnom primjeru je

$$\bar{x}_{7.a} = \frac{14+16+22+2 \cdot 23+25+2 \cdot 27+29+30+33}{23}$$
$$+ \frac{2 \cdot 34+36+3 \cdot 37+39+2 \cdot 46+2 \cdot 48+50}{23}$$
$$= 33.1.$$

Analogno, $\bar{x}_{7.b} = 25.5$.

Budući da su podaci cijelobrojni, aritmetička sredina je zaokružena na samo jedno decimalno mjesto. Tu smo primijenili pravilo da se aritmetička sredina daje s jednim decimalnim mestom više od broja decimalnih mesta koje imaju podatci.

Mod i medijan

Osim aritmetičke sredine, zanimljivo je naći podatak koji se javlja najveći broj puta, naziva se **mod**, te podatak koji se nalazi u sredini niza podataka. Zadržimo se i dalje na prethodnom primjeru. Podatak u 7.a razredu koji se javlja najveći broj puta je 37 (javlja se 3 puta, a svi ostali manji broj puta). Dakle, mod niza podataka za 7.a je $M_o = 37$. U nizu podataka za 7.b razred mod je 19, tj. $M_o = 19$. Postoje skupovi podataka koji nemaju mod ili čak imaju više modova.

Primjer 3. Odredimo mod danih skupova podataka:

- a) 2, 4, 6, 7, 8, 9, 10
- b) 2, 3, 3, 3, 3, 4, 5, 5, 5, 5, 8.

U a) zadatku svaki se podatak javlja jednom, pa ovaj niz nema mod. U b) zadatku brojevi 3 i 5javljaju se 4 puta i to je najveća frekvencija, pa ovaj niz ima dva moda: 3 i 5. Pišemo $M_{o1} = 3$, $M_{o2} = 5$. Ovakav niz zovemo bimodalni niz podataka.

Kao što smo primijetili, iz peteljka-list dijagrama Primjera 2, podatci u 7.b razredu kao da su grupirani prema nižim vrijednostima. Pronađimo srednji član niza kako bismo i numerički potkrijepili svoje zapažanje.

Ukupan broj podataka u 7.a je 23, pa je srednji član niza broj na 12. mjestu, tj. broj $x_{12} = 34$, a u

7.b srednji član niza je $x_{12} = 23$. Dakle, polovina rezultata u 7.b nalazi se ispod rezultata 23, dok je za 7.a taj broj 34.

Broj koji dijeli uređeni niz podataka u dvije jednakobrojne skupine naziva se **medijan** i označava s M . Dakle, polovina podataka je manja ili jednaka medijanu, a polovina podataka je veća ili jednaka medijanu. Kad je ukupan broj podataka neparan, tada je medijan srednji član niza. Kad je ukupan broj podataka paran, tada postoje dva srednja člana niza i za medijan se uzima njihova aritmetička sredina. Formalno definicija medijana glasi ovako:

$$M = \begin{cases} \frac{x_{k+1}}{2} & \text{ako je } n = 2k + 1 \\ \frac{x_k + x_{k+1}}{2} & \text{ako je } n = 2k \end{cases}$$

pri čemu je n ukupan broj podataka koji su poslagani u rastućem poretku.

Primjer 4. Odredimo medijan danih skupova podataka:

- a) 2, 4, 6, 7, 8, 9, 10, 11, 17
- b) 2, 3, 13, 12, 3, 24, 5, 15, 15, 17, 8, 25.

U a) zadatku podatci su poredani po veličini i ima ih 9. Srednji član niza je peti podatak ($9 = 2 \cdot 4 + 1$, $k = 4$, $k + 1 = 5$), tj. $M = 8$.

$$\begin{array}{ccccccccccccc} 2 & 4 & 6 & 7 & 8 & 9 & 10 & 11 & 17 \\ & & & & \uparrow & & & & \\ & & & & M & & & & \end{array}$$

U b) zadatku podatci nisu poredani po veličini. Prvo preuređimo skup podataka:

$$2, 3, 3, 5, 8, 12, 13, 15, 15, 17, 24, 25.$$

Ima ih 12, $12 = 2 \cdot 6$, pa su dva člana u sredini šesti i sedmi član, tj.

$$M = \frac{x_6 + x_7}{2} = \frac{12 + 13}{2} = 12.5.$$

$$\begin{array}{ccccccccccccc} 2 & 3 & 3 & 5 & 8 & 12 & 13 & 15 & 15 & 17 & 24 & 25 \\ & & & & & \uparrow & & & & & & \\ & & & & & M & & & & & & \end{array}$$

Primjer 5. Odredimo medijan i aritmetičku sredinu danih skupova podataka:

- a) 2, 3, 3, 4, 5, 7, 8
- b) 2, 3, 3, 4, 5, 7, 235.

Rješenje a) zadatka je

$$M = 4, \bar{x} = \frac{2+3+3+4+5+7+8}{7} = \frac{32}{7} = 4.6.$$

Rješenje b) zadatka je

$$M = 4, \bar{x} = \frac{2+3+3+4+5+7+235}{7} = \frac{259}{7} = 37.$$

Uočimo sličnosti i razlike ovih dvaju nizova. Oba imaju isti broj podataka, čak su im i prvih šest podataka jednaki. Stoga su im srednji članovi niza, tj. medijani jednaki. Dakle, polovina podataka je manja od 4, a polovina je veća od 4 u oba niza. No, u b) nizu se kao zadnji podatak pojavio ekstremno velik broj 235 koji je bitno utjecao na iznos aritmetičke sredine u b) zadatku. Tako se kao prosjek u b) zadatku pojavio broj 37 koji i sam odskače od svih članova niza. Kad bismo se pitali koji pokazatelj: medijan ili aritmetička sredina bolje opisuju ove nizove, odgovor bi za b) zadatak bio: medijan. Dakle, aritmetička je sredina jako osjetljiva na ekstremno male ili velike podatke i u takvim slučajevima taj broj ne daje adekvatnu sliku. Pogledajmo još jedan sličan primjer.

Primjer 6. U tablici je prikazana raspodjela plaća zaposlenika u jednoj tvrtki.

iznos plaće (u kn)	broj zaposlenika koji su dobili tu plaću
25 000	1
14 000	3
4000	20
2500	4

Kolika je prosječna plaća u tom poduzeću? Značući činjenicu da je u tom mjesecu prosječna plaća

u državi bila 5300 kn, možemo li zaključiti da zaposlenici tog poduzeća dobivaju iznadprosječno velike plaće?

Izračunajmo aritmetičku sredinu plaća:

$$\bar{x} = \frac{25\,000 + 3 \cdot 14\,000 + 20 \cdot 4\,000 + 4 \cdot 2\,500}{28} \\ = 5607.1$$

Dakle, prosječna plaća u toj tvrtki je 5607.1 kn. Kad bismo poznavali samo podatak o prosječnoj plaći u poduzeću, bez okolišanja bismo zaključili da zaposlenici tog poduzeća dobivaju plaće koje su iznad državnog prosjeka. Međutim, stvarnost je nešto drugačija: samo 4 zaposlenika dobiva iznadprosječnu plaću, dok ostalih 24 dobiva plaću puno ispod prosjeka. Tu je došlo do izražaja svojstvo aritmetičke sredine na koju jako utječu značajno velike (ili male) vrijednosti podataka. U ovom je primjeru medijan puno pogodnija mjera centralne vrijednosti. Kad bismo plaće poredali po veličini i odredili medijan tog niza, dobili bismo da je $M = 4000$ kn.

Primjena tehnologije. Budući da se pri statističkim istraživanjima pojavljuje obično velik broj podataka, upotreba računala i raznih statističkih programa je potpuno uobičajena pri izračunavanju raznih numeričkih pokazatelja. U nastavi obično nećemo raspolagati specijalnim statističkim programima, no obično džepno računalo i proračunske tablice (kao što je primjerice Excel) su dovoljni za sva potrebna računanja.

LITERATURA

- 1/ A. G. Bluman: *Elementary Statistics*, McGraw Hill, Boston, 2001.
- 2/ J. Gusić: *Petejike i listovi*, Matka, 49 (2004), 8–9.
- 3/ S. Varošanec: *Matematika 1*, udžbenik za prvi razred medicinskih škola, Element, 2012.