



UREĐIVANJE I PRIKAZIVANJE STATISTIČKIH PODATAKA

Statistički niz

- Prikupljene podatke prvo treba urediti kako bi se reducirao velik broj informacija o pojavi koja se velik broj informacija o pojavi koja se želi istražiti.
- Pri uređivanju treba voditi računa da u reduciranim podacima moraju ostati sačuvane glavne značajke pojave.
- Urede li se podaci o jednom obilježju jedinica statističkog skupa će **statistički niz**.
- Naziv niza istovjetan je s vrstom obilježja pa govorimo kvalitativnim i kvantitativnim nizovima.

Statistički niz

- **Nominalni niz** nastaje uređivanjem podataka o modalitetima nominalne varijable.
- **Redoslijedni niz** nastaje uređivanjem podataka o modalitetima rang varijable.
- **Numerički niz** nastaje uređivanjem podataka koji predstavljaju vrijednosti numeričke varijable.
- **Vremenski niz** nastaje kronološkim nizanjem podataka o nekoj pojavi.

Grupiranje podataka

- Među najvažnijim metodama uređivanja podataka je **metoda grupiranja**.
- Grupiranje je postupak raščlanjivanja skupa od N podataka k podskupova podskupova koji se međusobno ne preklapaju.
- Razvrstavaju se prema obilježju i njegovim modalitetima te prema principima:
 - iscrpnosti – svaki se podatak mora razvrstati
 - isključivosti – jedan podatak može biti član samo jedne grupe

Kvalitativni podatci

- Vrlo mali broj podataka navodi se nekim redom odabranim po volji ili prema intenzitetu mjenog svojstva, a u protivnom ih se grupira (N podataka grupira se u k podskupova).
- Broj podataka istog ili sličnog obilježja (broj podataka u jednoj grupi) naziva se **frekvencija**.
- Zbroj frekvencija čini opseg skupa N .
- Frekvencije su brojevi između 0 i N .

Kvalitativni podatci

- **Apsolutna frekvencija** je broj jedinica statističkog skupa koji pripadaju određenoj grupi, tj. jednom modalitetu ili pojavnom obliku promatranog obilježja.

$$f_i, \quad i = 1, 2, \dots, N$$

- **Relativna frekvencija** je jednaka količniku odgovarajuće apsolutne frekvencije modaliteta i opsega statističkoga skupa.

$$p_i = \frac{f_i}{N}, \quad i = 1, 2, \dots, N$$

Statistička tablica

- Grupirani podaci uobičajeno se prikazuju u **statističkoj tablici**.
- Statistička tablica nužno mora imati sljedeće dijelove:
 1. naslov tablice - mora biti jasan i kratak, a istovremeno u sebi mora sadržavati pojamovnu, prostornu i vremensku definiciju statističkog skupa,
 2. tekstualni dio - sastoji se od dva dijela: zaglavlja ili tumača stupaca i predstupca ili tumača redaka,
 3. numerički dio - sastoji se od polja u koja se unose frekvencije, odnosno rezultati grupiranja statističkih podataka,
 4. izvor podataka - navodi ispod tablice, a omogućuje provjeru ispravnosti prikupljenih podataka u tablici kao i eventualnu dopunu podataka.

Statistička tablica

- **Jednostavna statistička tablica** - sadrži jedan statistički niz.

Primjer:

Tablica 1. Osobe prema stupnju naobrazbe stečenom 2000. u Republici Hrvatskoj.

Stupanj naobrazbe	Broj osoba	Udio
osnovna škola	53.839	46,24
srednja škola	49.081	42,15
stručni studij	4.626	3,97
sveučilišni studij	8.884	7,63
Ukupno	116.430	100,00

Izvor: Statistički ljetopis Republike Hrvatske, 2001, str. 426.

Statistička tablica

- **Skupna statistička tablica** – prikazuje dva ili više nizova koji su nastali grupiranjem različitih skupova prema modalitetima istog obilježja.

Primjer:

Tablica 2. Stanovništvo, stanovi i kućanstva odabranih gradova u Republici Hrvatskoj prema popisu 31.03.1991.

Grad	Stanovništvo	Stanovi	Kućanstva
Zagreb	706.770	254.887	250.278
Split	189.388	57.794	58.831
Rijeka	167.964	60.013	59.769
Osijek	104.761	38.383	37.963
Ukupno	1.168.883	411.077	406.841

Izvor: Statistički ljetopis Republike Hrvatske, 1991, str. 15.

Statistička tablica

- **Kombinirana statistička tablica** – prikazuje podatke grupirane prema dva ili više obilježja.

Primjer:

Tablica 3. Stanovništvo Republike Hrvatske po spolu i aktivnosti prema popisu od 31.03.1991.

Aktivnost	Spol		Ukupno
	Muški	Ženski	
aktivno stanovništvo	1.166	874	2.040
osobe s osobnim prihodima	345	404	749
izdržavano stanovništvo	651	1.059	1.710
Ukupno	2.162	2.337	4.499

Izvor: Statistički ljetopis Republike Hrvatske, 1998, str. 89.

Kvantitativni podatci

- Način uređivanja ovisi o broju podataka i o tome da li je numerička varijabla diskretna ili kontinuirana.
- Ako je broj podataka malen broj uređivanje se provodi njihovim nizanjem po veličini od najmanje do najveće (rjeđe obrnuto).
- **Uređen numerički niz** N pojedinačnih vrijednosti varijable X :

$$x_1, x_2, \dots, x_{i-1}, x_i, \dots, x_n, \quad x_{i-1} \leq x_i, \quad i = 1, 2, \dots, N-1$$

Kvantitativni podatci

- Ako diskretna numerička varijabla poprimi manji broj oblika, a broj podataka je velik, uređivanje se provodi grupiranjem podataka.
- **Distribucija frekvencija** se sastoji od parova razreda i pripadajućih frekvencija:

$$(x_i, f_i), \quad i = 1, 2, \dots, N$$

f_i – frekvencija i – tog razreda

x_i – modlitet numeričkog obilježja i – tog razreda

Kvantitativni podatci

- Vrijednosti kontinuirane varijable grupiraju se tako da se raspon između najveće i najmanje vrijednosti podijeli na k intervala (**razreda**).
- U danom razredu naći i će se vrijednosti omeđene donjom L_1 i gornjom L_2 granicom razreda:

$$L_{i,1} \leq x_i \leq L_{2,i}, \quad i = 1, 2, \dots, N$$

- Broj numeričkih podataka razreda predodčava njegovu frekvenciju.
- Poredak razreda je prema rastuđoj vrijednosti numeričke varijable.

Kvantitativni podatci

- Formiranju distribucije frekvencija prethodi određivanje broja razreda k i veličina razreda.
- Broj razreda k za grupiranje N vrijednosti numeričke varijable aproksimira se izrazom:

$$k \approx 1 + 3,3 \log N$$

Kvantitativni podatci

- Razredi mogu biti jednake ili različite veličine.
- Ako su razredi **jednaki**, veličina im se aproksimira tako da se raspon između najveće i najmanje vrijednosti podijeli s brojem razreda:

$$i_i = \frac{X_{\max} - X_{\min}}{k}$$

- Razredi jednakih veličina koriste se za grupiranje približno simetričnih podataka.
- Kad je raspored podataka asimetričan, formiraju se razredi različitih veličina.
- Praktično je pravilo da su razredi uži gdje i gdje je veća koncentracija podataka.

Kvantitativni podatci

- **Granice razreda** su **prave** ako je donja granica tekućeg razreda jednaka gornjoj granici prethodnog razreda.
- U protivnome, granice se nazivaju **nominalnima** i u postupcima se pretvaraju u prave – donja prava granica tekućeg razreda jednaka je poluzbroju donje nominalne granice tekućeg razreda i gornje nominalne granice prethodnog razreda.
- **Veličina i-tog razreda** jednaka je razlici gornje i donje granice toga razreda, pri čemu se pretpostavlja da su granice razreda prave granice.
- **Razredna sredina i-tog razreda** jednaka je poluzbroju njegove donje i gornje granice.

Kvantitativni podaci

Primjer: Pretpostavimo da je 50 studenata Veleučilišta u Šibeniku pristupilo pismenom dijelu ispita iz Statistike. Ostvareni bodovi dani su:

44	22	65	47	43	33	51	79	59	78
66	37	36	31	57	62	37	84	53	76
57	45	43	52	64	74	85	88	46	58
25	41	49	53	66	82	91	54	58	68
77	56	55	52	73	26	78	54	43	62

Ako analiziramo rezultate uočit ćemo da je najmanji broj ostvarenih bodova iznosio 22, a najveći 91.

$$\text{Broj razreda: } i_i = \frac{x_{\max} - x_{\min}}{k} = \frac{91 - 22}{7} = 10$$

$$\text{Širina razreda: } k \approx 1 + 3,3 \log 50 \approx 7$$

Razredi	Frekvencija
22 – 32	3
32 – 42	6
42 – 52	9
52 – 62	13
62 – 72	7
72 – 82	7
82 – 92	5
Ukupno	50

Kumulativni niz

- U statističkoj analizi koriste se i izvedeni nizovi.
- Među takvima je i **kumulativni niz** koji nastaje postupnim zbrajanjem frekvencija ili relativnih frekvencija numeričkog niza.
- Član kumulativnog niza pokazuje koliko podataka ima vrijednost varijable jednaku i manju od numeričke varijable.

Kumulativni niz

Primjer: Nezaposleni u Republici Hrvatskoj prema radnom stažu, stanje 31.12.1996.

Razredi	Frekvencija	Kumulativni niz manje od	Kumulativni niz veće od
0 – 1	123.614	123.614	269.263
1 – 2	19.767	143.381	145.649
2 – 3	12.271	155.652	125.882
3 – 5	16.033	171.685	113.611
5 – 10	29.401	201.086	97.578
10 – 20	43.747	244.833	68.177
20 – 40	24.430	269.263	24.430
Ukupno	269.263		

Grafički prikaz

- Nizovi kvalitativnih podataka grafički se uobičajeno prikazuju **površinskim grafikonima**.
 - Stupci – uspravni, položeni, dvostruki, razdijeljeni, ...
 - Strukturni krug – veličina sektora proporcionalna je frekvencijama niza

$$s_i = \frac{f_i}{N} \cdot 360^\circ$$

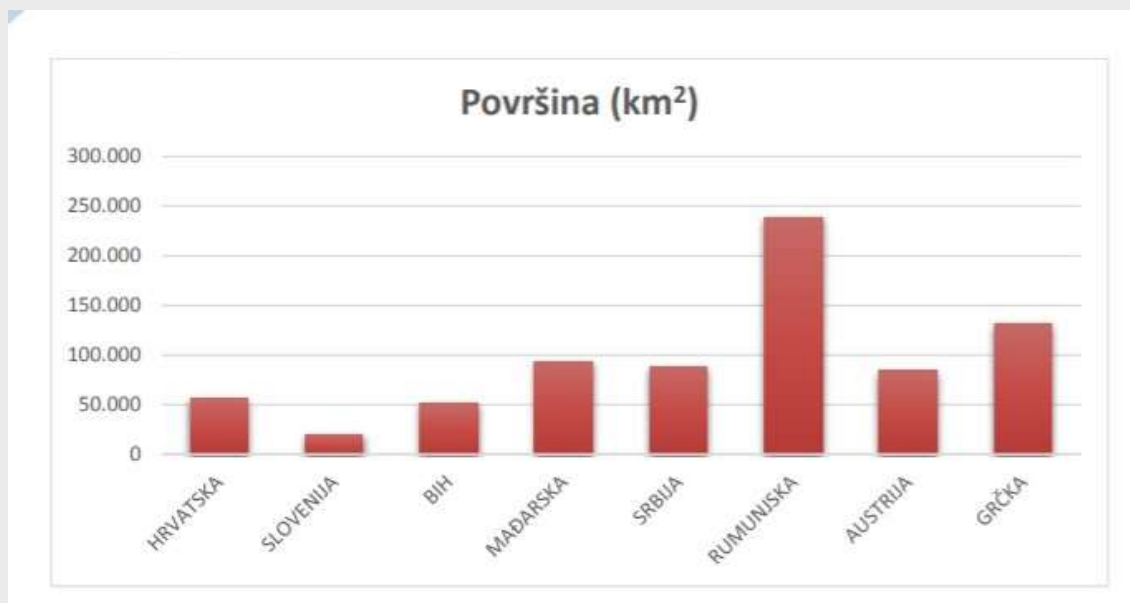
Grafički prikaz

Primjer: Za odabrane zemlje u regiji dani su podaci za površinu, broj stanovnika, broj gradskog stanovništva i BDP (izvor: Wikipedia, 2015.). Prikažite i analizirajte dane podatke različitim grafičkim metodama.

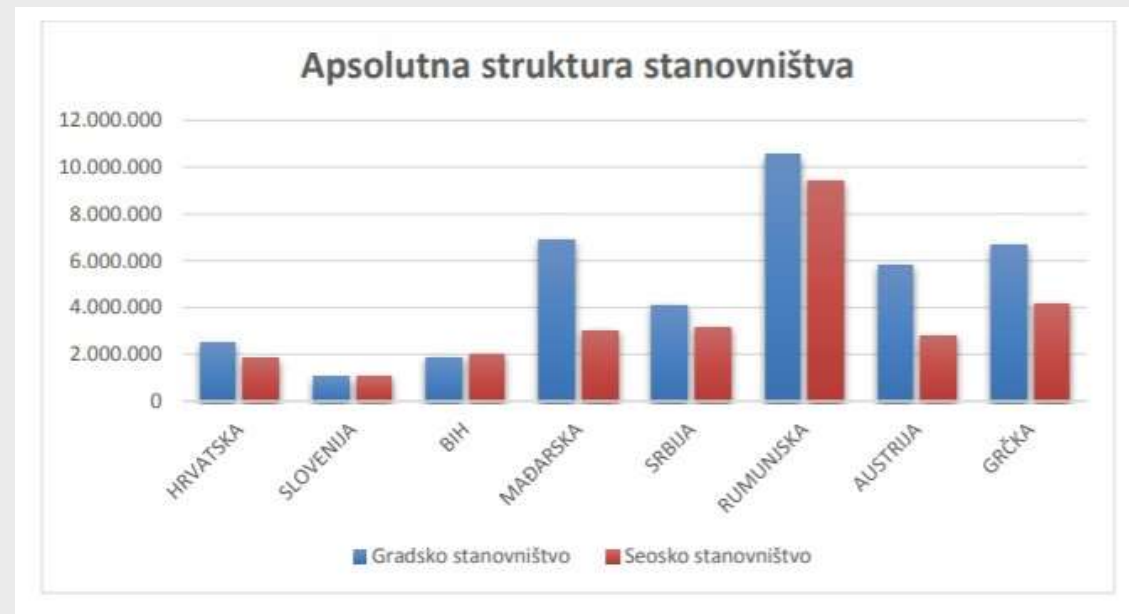
	Površina	Stanovništvo	Gradsko stanovništvo	BDP (milijarde USD)
Hrvatska	56.594	4.284.889	2.476.666	59,9
Slovenija	20.273	2.061.085	1.028.481	43,0
BiH	51.197	3.871.643	1.870.004	19,1
Mađarska	93.030	9.877.365	6.864.769	145,2
Srbija	88.361	7.209.764	4.066.307	42,6
Rumunjska	238.391	19.942.642	10.529.715	215,3
Austrija	83.879	8.602.112	5.823.630	389,6
Grčka	131.957	10.815.197	6.640.531	207,0

Grafički prikaz

Jednostavni stupci



Razdijeljeni stupci

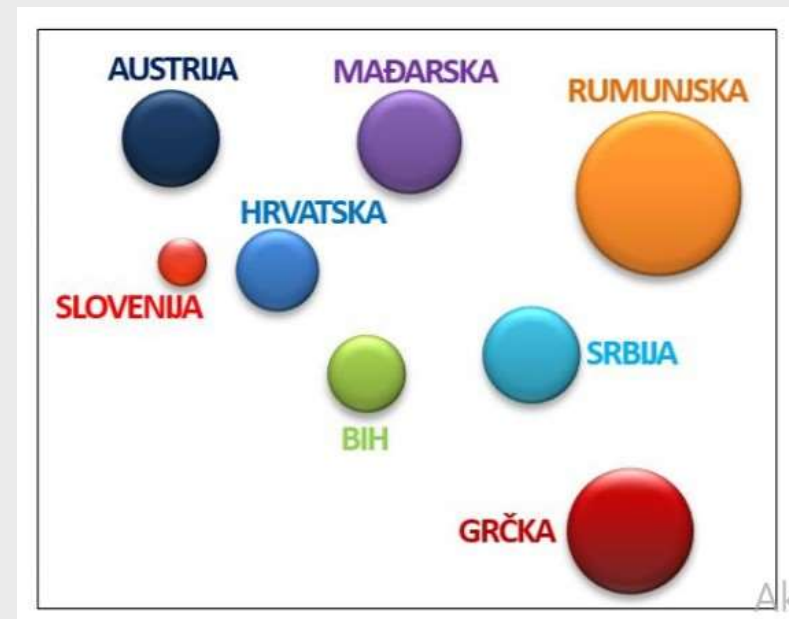


Grafički prikaz

Strukturni krug

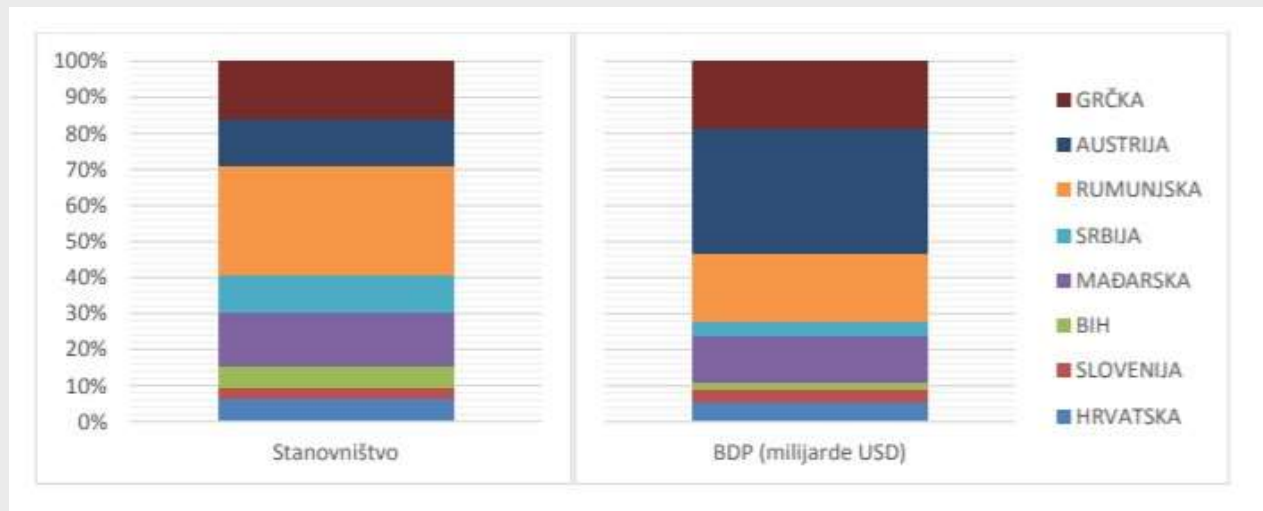


Proporcionalni krugovi

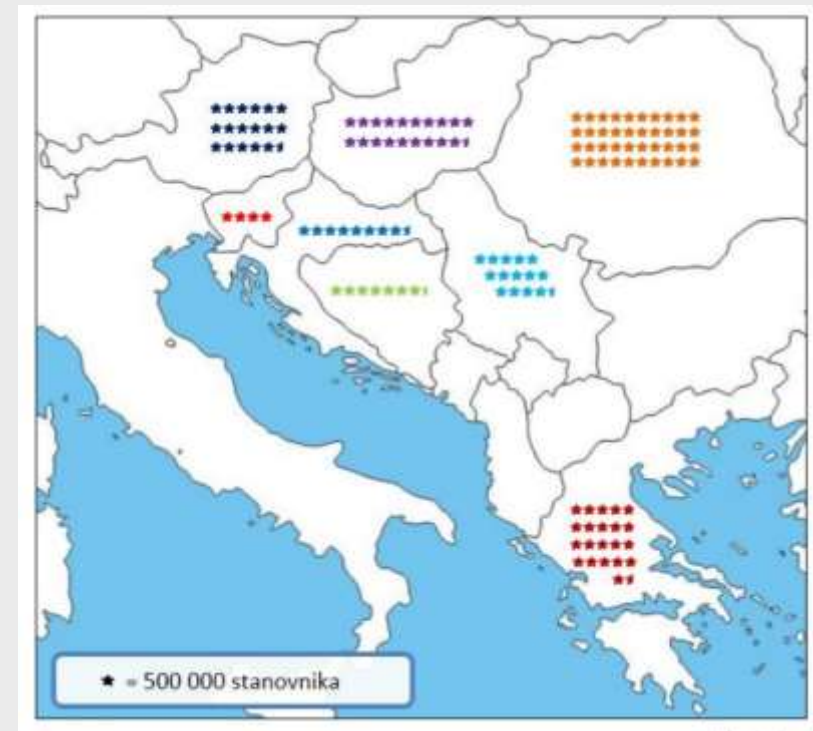


Grafički prikaz

Razdijeljeni pravokutnici



Statistička karta



Grafički prikaz

Piktogram



Grafički prikaz

- Nizovi kvantitativnih podataka grafički se uobičajeno prikazuju **linijskim i površinskim grafikonima**.
- Distribucija frekvencija grafički se prikazuje **histogramom i poligonom frekvencija**.
- Histogram je površinski grafikon koji se sastoji od pravokutnika kojemu je osnovica vrijednost numeričke varijable, a visina ovisi o frekvenciji varijable.

Ako su razredi nejednake veličine, visina ovisi o korigiranoj frekvenciji koja se određuje dijeljenjem frekvencije s veličinom razreda.

- Poligon frekvencija je linijski grafikon koji nastaje spajanjem točaka koje se ucrtavaju iznad razrednih sredina do visine određene frekvencijom ili korigiranom frekvencijom.

Grafički prikaz

Primjer: U tablici je prikazan uspjeh učenika jednog trećeg razreda na kraju školske godine iz predmeta hrvatski jezik i matematika.

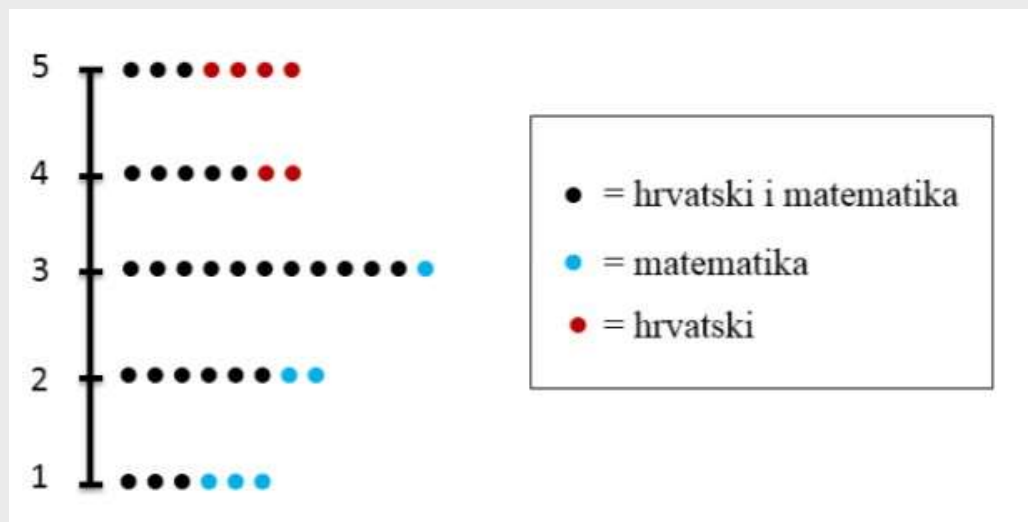
Prikažite i analizirajte dane podatke različitim grafičkim metodama.

R.br.	HRVATSKI		MATEMATIKA	
	Prosjek	Kon.	Prosjek	Kon.
1.	3.4	3	2.2	2
2.	2.5	3	1.3	1
3.	3.2	3	3.0	3
4.	4.1	4	1.9	2
5.	2.1	2	1.0	1
6.	3.9	4	2.4	2
7.	4.8	5	4.1	4
8.	1.4	1	3.2	3
9.	2.8	3	2.7	3
10.	4.5	5	3.4	3
11.	3.3	3	2.0	2
12.	5.0	5	4.0	4
13.	3.0	3	1.4	1
14.	4.4	4	5.0	5
15.	4.9	5	2.4	2
16.	1.2	1	1.8	2
17.	2.8	3	3.2	3

R.br.	HRVATSKI		MATEMATIKA	
	Prosjek	Kon.	Prosjek	Kon.
18.	2.3	2	2.7	3
19.	3.6	4	3.6	4
20.	4.1	4	3.2	3
21.	1.9	2	1.2	1
22.	4.8	5	2.7	3
23.	4.6	5	3.6	4
24.	2.1	2	1.4	1
25.	5.0	5	4.0	4
26.	3.9	4	4.6	5
27.	3.3	3	2.2	2
28.	1.2	1	2.0	2
29.	3.3	3	2.8	3
30.	3.3	3	3.0	3
31.	2.5	3	2.5	3
32.	2.4	2	1.3	1
33.	4.0	4	4.8	5
34.	2.4	2	2.7	3

Grafički prikaz

Točkasti dijagram

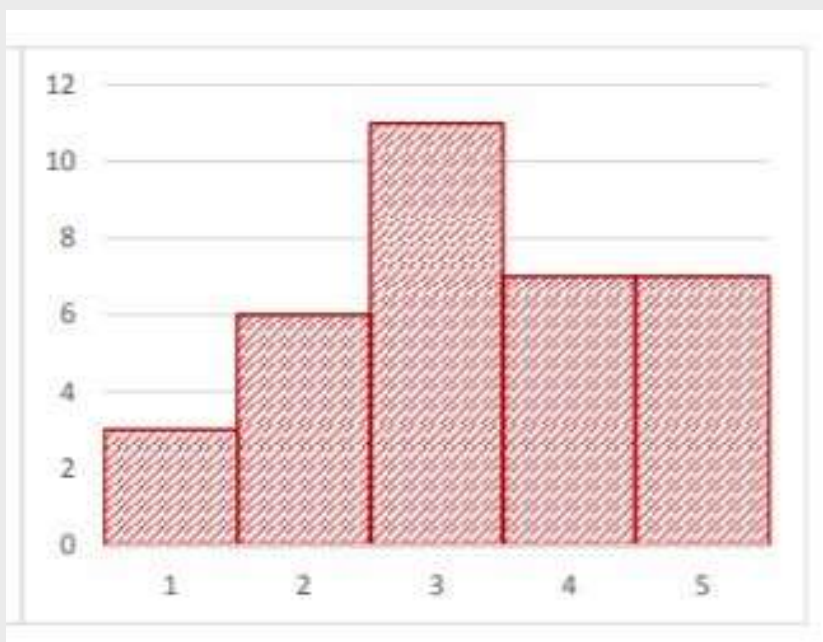


Dijagram stablo-list

Broj pod.	HRVATSKI		MATEMATIKA	Broj pod.
4	9422	1	02334489	8
9	885544311	2	002244577778	12
10	9964333320	3	00222466	8
9	988654110	4	00168	5
2	00	5	0	1
34				34

Grafički prikaz

Histogram



Poligon frekvencija



Zadaci

Zadatak 1. Istraživači s punim radnim vremenom u Republici Hrvatskoj 1999.

Stupanj obrazovanja	Broj istraživača	Struktura u %
s doktoratom znanosti	3.159	
s magisterijem	1.460	
visoka stručna sprema	2.040	
specijalisti	136	
Ukupno	6.795	100

- Kojoj vrsti statističkih varijabli pripada varijabla *stupanj obrazovanja*?
- Izračunajte postotne relativne frekvencije.
- Niz iz tablice prikažite jednostavnim stupcima.

Zadaci

Zadatak 2. Ispituje se kvaliteta rada tipkovnica proizvodaca na temelju reklamacija kupaca. Za isporuku 4.109 tipkovnica reklamacije kupaca tijekom jednogodišnjeg jamstva odnose se na sljedeće kvarove:

Vrsta kvara	blokada tipke	blokada tipkovnice	kontakt tipke	kabel tipkovnice	kontakt razmaknice	gubitak kapice tipke	ostali kvarovi
Broj kvarova	48	33	53	5	8	14	2

- O kojoj je vrsti niza riječ? Niz uredite prema padajućoj vrijednosti frekvencija.
- Izračunajte postotne frekvencije.
- Formirajte kumulativni niz postotnih frekvencija.

Zadaci

Zadatak 3. Studenti sveučilišnih studija u Republici Hrvatskoj školske godine 2000./2001.

Studijsko područje	Broj studenata	Od toga studentice
prirodne znanosti	2.367	1.285
tehničke znanosti	18.398	5.253
medicinske znanosti	4.693	3.079
biotehničke znanosti	3.334	1.416
društvene i humanističke znanosti	40.048	26.445
umjetničke akademije	997	559

- Izračunajte postotak studentica u ukupnom broju studenata po studijskim područjima.
- Niz studenata prema studijskom području prikažite jednostavnim stupcima.
- Strukturu studenata po spolu i studijskim područjima prikažite razdijeljenim stupcima.

Zadaci

Zadatak 4. Odredite prave granice razreda:

Visina objekta u m	Težina u kg
5 - 9	100
10 - 12	150
13 - 15	200
16 - 20	300

Zadaci

Zadatak 5. Rastavljeni brakovi prema dobi muža u Republici Hrvatskoj 1999. godine:

Godine	15 - 24	25 - 29	30 - 34	35 - 39	40 - 44	45 - 49	50 - 54	55 - 59	60 - 79
Broj osoba	105	439	662	683	635	503	305	174	208

- Koliki je opseg promatranog statističkog skupa?
- Odredite prave granice, razredne sredine, veličine razreda i korigirane frekvencije.
- Prikažite distribuciju frekvencija linijskim i površinskim grafikonom.

Zadaci

Zadatak 6. Na jednom zagrebačkom fakultetu izvršeno je mjerenje težine 150 studenata. Rezultati mjerenja dani su u tablici:

- Formirajte distribuciju frekvencija 150 studenata zagrebačkog fakulteta.
- Prikažite poligon frekvencija težine studenata zagrebačkog fakulteta.

64	68	72	76	80	84	88	90	77	73	75	88	70	73	67
91	79	81	79	73	78	83	76	73	81	77	83	71	80	71
85	79	83	75	76	72	79	75	73	69	79	82	84	74	81
78	78	73	66	75	78	67	76	84	87	76	87	85	77	89
72	77	83	76	72	79	79	75	83	74	80	80	78	70	80
75	80	69	76	82	78	87	78	77	77	82	72	80	71	76
82	73	68	76	83	77	79	69	81	78	80	74	81	73	79
75	82	70	78	89	82	72	77	74	77	83	69	77	81	79
77	85	78	82	83	78	74	79	74	82	76	74	86	78	69
79	86	90	86	75	81	75	76	71	72	76	81	91	65	77