Predavanja:
Dr.sc.Suzana Marković,docent
E-mail: suzanam@hika.hr

Vježbe:
Ana Vučak, dipl.oec.
E-mail: vucaka@yahoo.com

STATISTIKA

Temeljni pojmovi

1. Što je statistika?
2. Što je predmet izučavanja statistike?
Statistika je znanstvena disciplina koja se bavi prikupljanjem,. analizom i tumačenjem podataka masovnih pojava.
3. Što je deskriptivna, a što inferencijalna statistika?
Statistika se kao znanstvena metoda dijeli na:
Opisnu (deskriptivnu) i Inferencijalnu (induktivnu, analitičku)

Deskriptivna statistika sastoji se od primjene postupaka kojima se podaci uređuju, te tabelarno i grafički prikazuju.
U sklopu DS provodi se i raznovrsna brojčana analiza.
Rezultati statističke analize dobiveni primjenom metoda DS služe za donošenje sudova o pojavi koju oni predočuju.

Inferencijalna statistika
Temelj IS je **uzorak**, tj. podaci o dijelu jedinica statističkog skupa. Na temelju podataka o dijelu zaključuje se o svojstvima cjeline.

Primjenom metoda IS, a na temelju tih podataka iz realnog ili zamišljenog skupa, donose se sudovi o cjelini (skupu).

5.Statistički skup
Statistički skup je temeljni pojam u statistici.
Statistički skup sastoji se od jedinica kojima se ispituje (mjeri) jedno svojstvo ili više svojstava (varijabli, obilježja) koja od jedinice do jedinice očituju statističku promjenjivost.

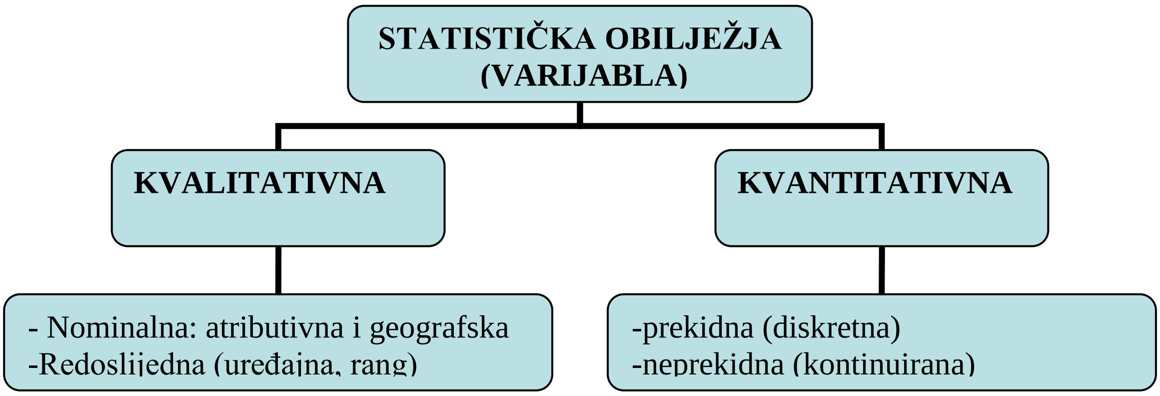
Broj jedinica naziva se **opsegom skupa.**

Prema opsegu se statistički skupovi dijele na:
konačne i beskonačne, te stvarne i zamišljene.

Podaci o danoj varijabli za svaki element statističkog skupa tvore skup podataka koji se naziva statističkom populacijom odnosno osnovnim skupom.

Podskup statističkog skupa naziva se **uzorkom. Uzorkom se također smatraju podaci o dijelu statističke populacije.**

7.Statistička obilježja (varijable) dijele se na:



Svojstva jedinica po kojima se članovi statističkog skupa razlikuju ili jedni drugima nalikuju nazivaju se **statističkim obilježjima**, odnosno **statističkim varijablama**.

KVALITATIVNO obilježje se naziva **kategorijalnim.**
Nominalno obilježje označava svojstvo jedinice skupa koje se izražava opisno. Podvrste nominalnih obilježja su **atributivna i geografska obilježja**.
Atributivnim obilježjem izražava se opisom svojstva (atribut) elementa statističkog skupa (npr. pismenost, narodnost, vrsta proizvoda, način plaćanja računa i sl.)
Geografsko ili prostorno obilježje je nominalno obilježje kojim se označava mjesto s kojim je jedinica statističkog skupa u nekoj vezi (npr. mjesto regulacije tvrtke, mjesto rođenja, mjesto porijekla turista i sl.)
Obilježje kojim se izražava stupanj nekog svojstva naziva se **redoslijednim**, odnosno **uređajnim obilježjem, ili obilježjem ranga** (npr. ocjena).

KVANTITATIVNA ili numerička obilježja izražavaju se brojčano. Ako obilježje poprima konačan broj vrijednosti, naziva se **prekidnim** (diskontinuiranim, diskretnim) numeričkim obilježjem (npr. broj učenika u razredu). Obilježje koje može poprimiti bilo koju vrijednost iz nekog intervala jest **neprekidno (kontinuirano) obilježje** (npr. visina, težina, stopa nezaposlenosti, potrošnja električne energije i sl.).

Vrijednosti numeričkog obilježja tvore mjerne skale. Vrste mjernih skala: nominalna, ordinalna (rang), intervalna i omjerna.

7.Vrste i izvori statističkih podataka

Podaci će se smatrati **statističkim** ako su prikupljeni prema određenom planu prikupljanja, odnosno poznatom nacrtu eksperimenta, ako su varijabilni i ako ih je dovoljan broj.

Izvori podataka mogu biti: **primarni i sekundarni.**

Primarni podaci se prikupljaju na temelju plana promatranja ili eksperimenta(reprezentativniji), a njihov opseg i struktura ovise o zadacima određenog istraživanja (skuplji).

Sekundarni podaci rezultat su djelatnosti drugih institucija, odnosno subjekata (pojavljuje se greška). Suvremeni je pristup podacima preko informacijskih sustava (internet i intranet), zato su jeftiniji.

10. Kako nastaje statistički niz?

UREĐIVANJE I PRIKAZIVANJE PODATAKA

Uređivanje statističkih podataka provodi se nakon prikupljanja podataka.

Da bi se mogli koristiti prikupljenim podacima potrebno ih je urediti i prikazati u odgovarajućem obliku.

Uređivanje podataka provodi se na različite načine. Podaci se **navode prema nekom pravilu ili se grupiraju.**

Ako se podaci grupiraju istodobno prema modalitetima dvaju ili više obilježja, riječ je o **dvodimenzionalnom ili višedimenzionalnom grupiranju.**

Uređivanjem podataka nastaje **statistički niz.** Statistički nizovi pregledno se prikazuju u **tabelama i grafikonima.**

FORMIRANJE STATISTIČKIH NIZOVA

Grupiranje je postupak kojim se statistički skup rasčlanjuje u disjunktne podskupove prema oblicima obilježja,i to tako da se u svaki podskup rasporede jedinice s jednakim, odnosno jednakim i sličnim oblikom obilježja.

Broj jedinica statističkog skupa koje imaju isti oblik obilježja naziva se **frekvencijom.** Zbroj frekvencija čini **opseg skupa.**

Grupiranjem se dobiva statistički niz. **Statistički niz** grupiranih podataka je skup parova različitih oblika obilježja s pripadajućim frekvencijama.

<u>Disjunktni podskupovi nemaju nijedan zajednički element.</u>

VRSTE STATISTIČKIH NIZOVA

Statističkih nizova ima onoliko vrsta i podvrsta koliko ima vrsta i podvrsta obilježja.

Grupiranjem kvalitativnih podataka, te nizanjem grupa s pripadajućim frekvencijama nastat će **kvalitativni statistički niz.** Urede li se podaci o modalitetima nominalne varijable doći će do **nominalnog niza. Redoslijedni niz** nastaje uređenjem podataka o rang varijabli. Nominalni i redoslijedni niz ubrajaju se u kvalitativne statističke nizove.

Nizanjem numeričkih grupa nastaje **numerički ili kvantitativni niz.**
Kronološko uređivanje podataka čini posebnu vrstu niza koji se zove **vremenski niz.**

11. Što je tabeliranje i čemu služi?
12. Nabrojite vrste statističkih tabela?

TABELIRANJE

Tabeliranje je postupak svrstavanja podataka u sheme, redove i stupce - tabele, prema određenom pravilu.
Tabelarnim načinom prikazivanja olakšava se praćenje statističkih podataka, a time i zaključci o pojavama koje oni predočuju.
Postoji više **vrsta** tabela.

13. Koje se tabele nazivaju izvještajnim?
One u kojima se navode svi prikupljeni podaci zovu se **izvještajnim tabelama.**

Analitička tabela sadrži dio uređenih podataka izdvojenih za određenu analizu. Analitičke su tabele manjih dimenzija.

Tabele se dalje dijele na: **jednostavne, skupne i kombinirane.**

Jednostavna statistička tabela sadrži samo jedan statistički niz.

U **skupnoj statističkoj tabeli** nalaze se dva statistička niza ili više njih. Nizovi prikazani u skupnoj tabeli odnose se na podatke različitih skupova, uređenih prema oblicima **istog obilježja.**

U kombiniranoj tabeli prikazani su podaci grupirani istodobno prema **dva ili više obilježja.**

Tablica se sastoji od zaglavlja (preko cijele tablice vodoravno, prvi okomiti stupac je PRETKOLONA, najšira sredina zove se BROJČANI DIO TABLICE (POLJE TABLICE), krajnji desni stupac zove se ZBIRNI STUPAC. U dnu vodoravno nalazi se preko cijele tablice ZBIRNI RED (SUME STUPCA).

Naslov tablice: …

	ZAGLAVLJE	∑
PRETKOLONA	BROJČANI DIO (POLJA TABLICE)	ZBIRNI STUPAC
∑	ZBIRNI RED (SUME STUPACA)	

Izvor: …

14. Čemu služi grafičko prikazivanje u statistici? Nabrojite vrste grafikona.

GRAFIČKO PRIKAZIVANJE

-iz njega možemo zaključiti nešto o pojavi koju promatramo

VRSTE GRAFIKONA

POVRŠINSKI

- stupci (jednostavni, dvostruki, višestruki, razdijeljeni)
- strukturni krug i polukrug
- kvadrat (Varzarov znak)
- Histogram (numerički niz)
- Kartogrami, a to su dijagramska karta, statistička karta, piktogram (geografski niz).

LINIJSKI

- linijski grafikon (numerički i vremenski niz)
- poligon frekvencija (numerički niz)

Grafikon stupaca

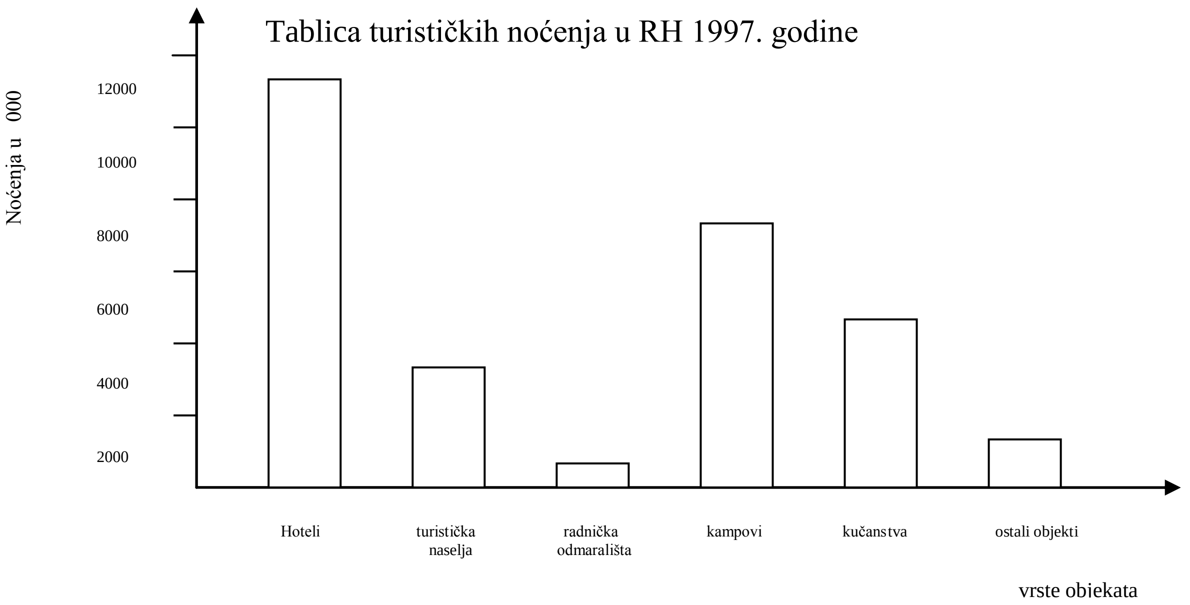
Grafikon stupaca je površinski grafikon statističkog niza koji se crta u pravokutnom koordiniranom sustavu. Frekvencije se prikazuju pravokutnicima (stupcima) jednakih osnovica.

Tablica turističkih noćenja u RH 1997. godine

VRSTA OBJEKTA	NOĆENJA U 000 (Fi)
Hoteli	11.247
Turistička naselja	3.791
Radnička odmarališta	685
Kampovi	7.857
Kućanstva (privatne osobe, stanovi i sl.)	5.660
Ostali objekti	1.075
UKUPNO	30.314

Izvor: Mjesečno izvješće, broj 10,1998.,str. 59

STUPAC MORA BITI OBAVEZNO JEDAN CENTIMETAR, KOLIKA JE ŠIRINA STUPCA TOLIKI JE I RAZMAK IZMEĐU STUPACA (kod crtanja grafičkih prikaza); CRTATI VEĆI GRAFIKON



Izvor: Mjesečno izvješće, broj 10,1998.,str. 59

→ Shema jednostavnih stupaca

Dvostruki stupci

Usporedba više nizova se provodi dvostrukim odnosno višestrukim stupcima.

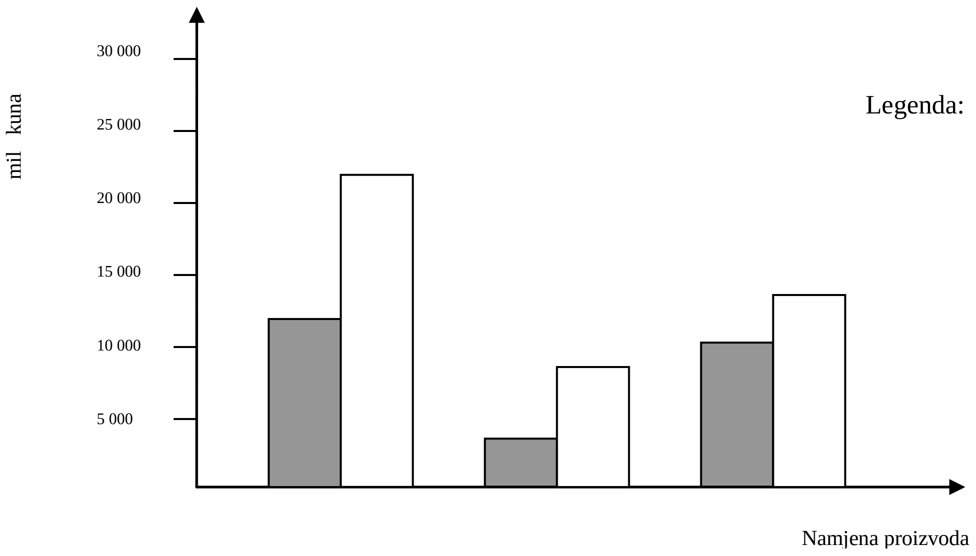
Tabela: Vanjskotrgovinska razmjena RH u 1996. prema ekonomskoj namjeni proizvoda.

NAMJENA PROIZVODA	IZVOZ (u mil.kuna)	UVOZ (u mil.kuna)
Proizvodi za reprodukciju	12.087	22.280
Proizvodi za investicije	3.734	8.164
Proizvodi za potrošnju	8.657	11.593
UKUPNO	24.478	42.037

Izvor: Statistički ljetopis RH,1997.str 316

Razmak stupaca obavezno 1 cm.Svaki grafikon i tabela moraju imati naslov obavezno.

Tablica: Vanjskotrgovinska razmjena RH u 1996. prema ekonomskoj namjeni proizvoda.



Izvor: Statistički ljetopis RH,1997.str 316

→ primjer dvostrukih stupaca

Razdijeljeni stupci

Razdijeljenim se stupcima prikazuje statistički niz kod kojega se frekvencije rastavljaju na dva ili više dijela. Mogu se crtati na osnovu apsolutnih i relativnih frekvencija.

Grafikon s krugovima

Krugom se mogu prikazati nominalni nizovi tako da se istakne struktura skupa, usporede opsezi dvaju ili više statističkih nizova, te usporedi opseg i strukturu više statističkih nizova.

Ako je svrha grafikona prikazati strukturu skupa, osim **strukturnog stupca,** koristi se i **strukturni krug.** Polumjer strukturnog kruga određuje se proizvoljno. Dijelovi kruga, isječci (sektori), proporcionalni su frekvencijama niza. Za njegovo crtanje, treba izračunati broj stupnjeva sektora kruga. Krug ima **360** stupnjeva, a stupnjevi sektora jesu:

fi-zadano

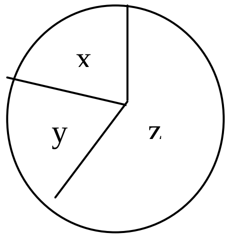
$$S_i = \frac{f_i}{N} * 180; \quad \text{N-ukupno}$$

Tabela: Zaposleno osoblje u trgovini prema djelatnostima poslovnih subjekata u RH 1997.

DJELATNOST POSLOVNIH SUBJEKATA	Broj zaposlenih fi	SEKTORI KRUGA $S_i = \frac{f_i}{N} * 360$
Trgovina na malo	45.674	258,46
Trgovina na veliko	7.719	43,68
Ostale djelatnosti	10.224	57,86
UKUPNO	63.617	360,00

Izvor:Statistički ljetopis RH,1998.,str. 347

Zaposleno osoblje u trgovini prema djelatnostima poslovnih subjekata u RH 1997.



x-trgovina na veliko
y-ostale djelatnosti
z-trgovina na malo

Strukturni krugovi primjenjuju se i za usporedbu više nominalnih nizova. Uspoređivati se može i njihov opseg.

Grafikon kojim se uspoređuje struktura više nominalnih nizova s istim nominalnim obilježjem zove se **grafikon strukturnih krugova.**
Uspoređuje li se istodobno opseg skupova i njihova struktura, riječ je o **grafikonu proporcionalnih strukturnih krugova.** Tim se grafikonom mogu uspoređivati frekvencije istog niza.
Grafikon više strukturnih krugova konstruira se ovako: najprije se nacrtaju krugovi jednakih polumjera s ishodištem na istom zamišljenom pravcu, a zatim se za svaki krug pomoću frekvencija izračunaju stupnjevi kruga.

U **grafikonu proporcionalnih strukturnih krugova** razlike u površinama krugova razmjerne su razlikama opsega skupova. Usporedbom isječaka krugova uočavaju se razlike u veličinama frekvencija istih nizova. Za crtanje grafikona potrebno je odrediti **površinu kruga (P=r^2Π) i polumjer kruga:** $r = \sqrt{\frac{P}{\Pi}}$

Grafikon strukturnih polukrugova

Struktura dvaju nominalnih nizova uspoređuje se i **strukturnim polukrugovima.** Ako se istodobno uspoređuju **opseg i struktura nizova,** mogu se primijeniti **proporcionalni strukturni polukrugovi.**

Za crtanje proporcionalnih strukturnih polukrugova, osim sektora polukruga potrebno je odrediti i **radijuse i strukturne isječke** primjenom slijedećih izraza:

$r = \sqrt{\frac{2P}{\Pi}}$;P= ukupno iz tabele $\qquad S_i = \frac{f_i}{N}$*180

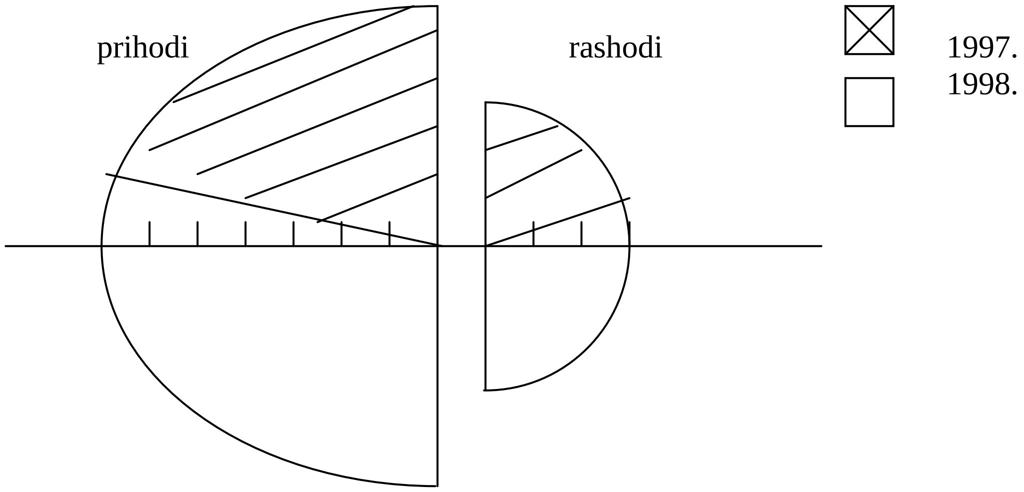
Tabela: Prihodi i rashodi od putovanja u mil USD

GODINA	PRIHODI	RASHODI	ISJEČAK ZA PRIHODE (D/C*180)	ISJEČAK ZA RASHODE (D/C*180)
1997.	2.529,1	521,4	86,62	83,67
1998.	2.726,3	600,3	93,38	96,33
UKUPNO	5.255,4	1.121,7	180,00	180,00

Izvor:Statistički ljetopis RH,1998.,str. 347

$$r_p = \sqrt{\frac{2P}{\prod}} = \sqrt{\frac{2\cdot 5255,4}{3,14}} = 57,86 \qquad r_R = \sqrt{\frac{2\cdot 1121,7}{3,14}} = 26,7$$

Prihodi i rashodi od putovanja u mil USD



Izvor:Statistički ljetopis RH,1998.,str. 347

Kartogrami

Kartogrami služe za prikazivanje geografskih statističkih nizova. Osnova za crtanje kartograma

15. Koje grafikone nazivamo površinskima?
16. Kako se crta strukturni krug, a kako strukturni stupci? Što je po vašem mišljenju prikladnije?
17. Kada koristimo dvostruke, a kada višestruke stupce?
18. Kojom vrstom grafikona prikazujemo numerički niz?

RELATIVNI BROJEVI

Svaki relativni broj nastaje dijeljenjem dviju veličina. Veličina s kojom se dijeli zove se osnovom relativnog broja. Relativni se brojevi razlikuju jedan od drugog ovisno o tome što im je osnova.

Vrste relativnih brojeva
1. RELATIVNI BROJEVI STRUKTURE (DIO/CJELINA)
- postoci (dio/cjelina x 100),
- promili (dio/cjelina x 100).

2. **RELATIVNI BROJEVI DINAMIKE (INDEKSI)**
- INDIVIDUALNI INDEKSI (BAZNI I VERIŽNI INDEKSI)
- SKUPNI INDEKSI (KOLIČINE, CIJENE, VRIJEDNOSTI)

3. RELATIVNI BROJEVI KOORDINACIJE

Relativne frekvencije

Relativne frekvencije danog oblika obilježja (grupe) statističkog niza računaju se **dijeljenjem frekvencije tog oblika obilježja sa zbrojem vrijednosti.**

Pomnožimo li spomenuti omjer sa 100, dobiti ćemo **postotnu relativnu frekvenciju**

Postotak

Postotak (P) je omjer postotnog iznosa (D) i postotne osnovice ©, pomnožen sa 100, tj.

$$P = \frac{D}{C} \cdot 100$$

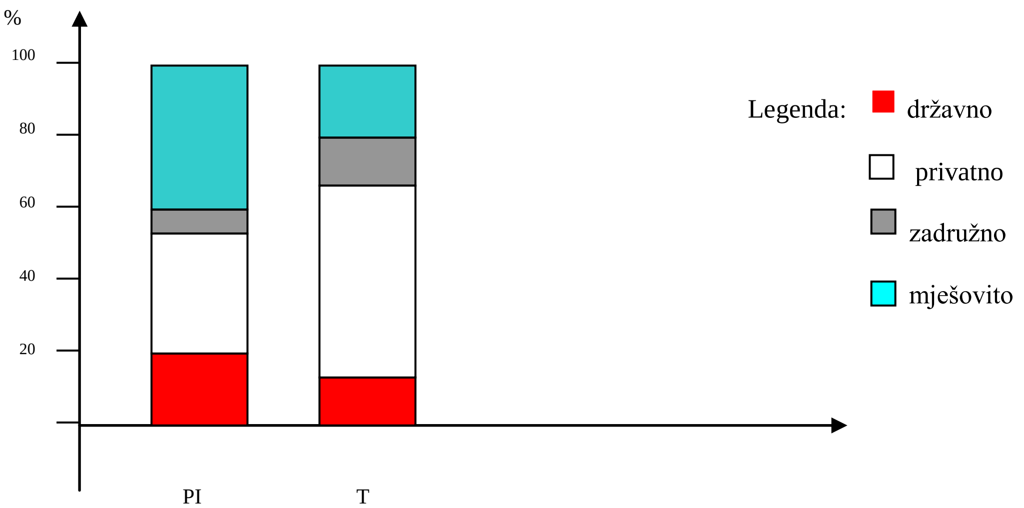
Postotak se grafički prikazuje **razdijeljenim stupcima.**

Tabela: Zaposleni u području prerađivačke industrije i trgovini na veliko i malo u RH prema oblicima vlasništva, stanje 31.ožujka 1997.

OBLICI VLASNIŠTVA	BROJ ZAPOSLENIH		STRUKTURA U %	
	Prerađivačka Industrija Fi1	Trgovina Fi2	Prerađivačka Industrija Pi1	Trgovina Pi2
Državno	61,037	15.658	21,77	13,54
Privatno	96,419	70,785	34,39	61,23
Zadružno	519	1.613	0,19	1,40
Mješovito	122,428	27,544	43,65	23,83
UKUPNO	280,403	115,600	100,00	100,00

Izvor: Statistički ljetopis RH, str.112

Tablica: Zaposleni u području prerađivačke industrije i trgovini na veliko i malo u RH prema oblicima vlasništva, stanje 31.ožujka 1997.



Izvor: Statistički ljetopis RH,1997.str 112

Za grafikon baza razmaka između razdijeljenih stupaca od 1 cm.

Indeksi (individualni indeksi)

Indeksi niza kvalitativnih podataka su relativni brojevi koji nastaju dijeljenjem vrijednosti članova niza s baznom veličinom i množenjem omjera sa 100.

Bazna veličina je odabrana vrijednost niza ili neka druga prikladna veličina.

Simbolički se taj omjer predočuje izrazom:

$$I_i = \frac{f_i}{B} \cdot 100$$

Indeksi niza kvalitativnih podataka grafički se prikazuju **jednostavnim stupcima.**

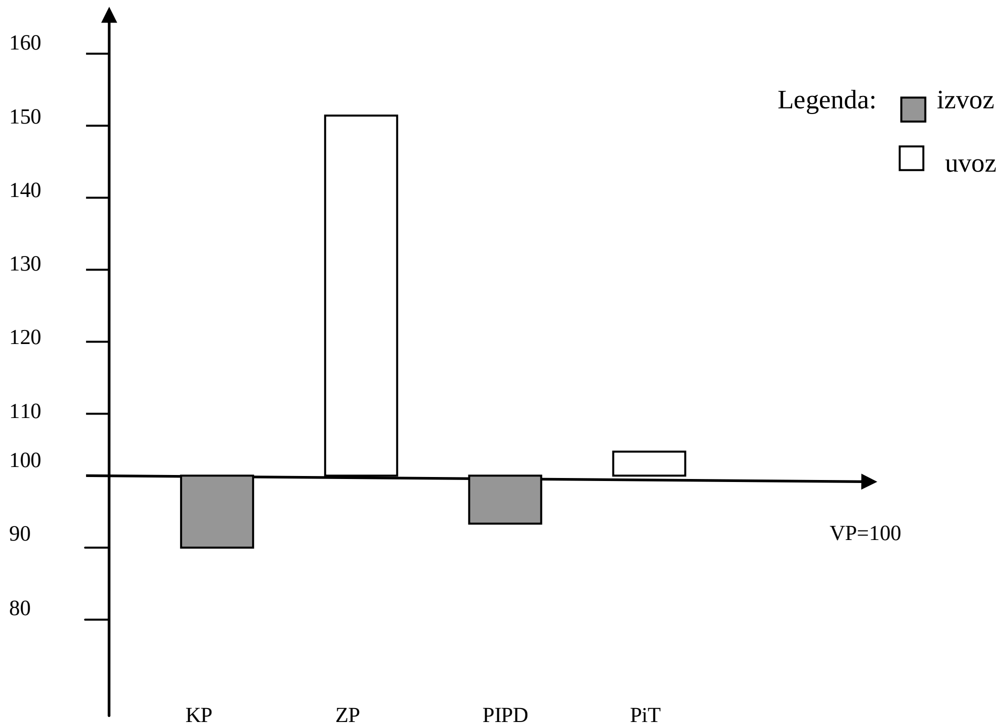
Primjer: Bazni indeksi

Tabela: Neto plaća po zaposlenome u području djelatnosti prijevoza, skladištenja i veza u RH za IX. Mjesec 1998.

VRSTA DJELATNOSTI	PROSJEČNA MJESEČNA PLAĆA U KUNAMA	INDEKSI (PROSJEČNA PLAĆA U DJELATNOSTI VODENOG PROMETA =100)
Kopneni prijevoz i Cjevovodni transport	2.816	88,77
Vodeni prijevoz	2.947	100,00
Zračni prijevoz	4.499	152,66
Prateća i pomoćne djelatnosti u prijevozu	2.819	95,66
Pošta i telekomunikacije	3.076	104,38

Izvor: Mjesečno statističko izvješće broj: 11,1998, str.38

Tabela: Neto plaća po zaposlenome u području djelatnosti prijevoza, skladištenja i veza u RH za IX. Mjesec 1998.



Izvor: Mjesečno statističko izvješće broj: 11,1998, str.38

Grafikon mora imati naslov isto kao i tablica.

Relativni brojevi koordinacije

Relativni brojevi koordinacije (statistički pokazatelji, omjerni brojevi) nastaju diobom dviju veličina koje ima smisla uspoređivati.

Npr. uspoređivanje uvoza i izvoza (ukoliko jedinica izvoza dolazi na jedinicu uvoza), gustoća napučenosti na određenom području (odnos broja stanovnika i površine područja), stupanj gospodarske razvijenosti (odnos narodnog dohotka i broja stanovnika), itd

Računanju RBK prethodi određivanje pojava koje se uspoređuju, njihovih brojčanih vrijednosti, te osnovice usporedbe. Ako se s **R** označi statistički omjerni broj, **v** njego brojnik, a **B** njegov nazivnik, izraz omjernog broja može se napisati:

$$R_i = \frac{v_i}{B_i}$$

Relativni brojevi koordinacije (statistički koeficijenti) se grafički često prikazuju **jednostavnim stupcima** ili **pravokutnicima (Varzarov znak)** kojima su baze ovisne o nazivniku relativnog broja, a ovise o njegovoj vrijednosti.

Kada želimo istaknuti **samo** veličine RBK, nacrtat ćemo **grafikon jednostavnih stupaca.** Stupci su jednakih osnovica, a visina im je određena vrijednošću koeficijenta u aritmetičkom mjerilu osi ordinata. Visine stupaca govore o veličini koeficijenata, a razlike visina o razlikama vrijednosti koeficijenata.

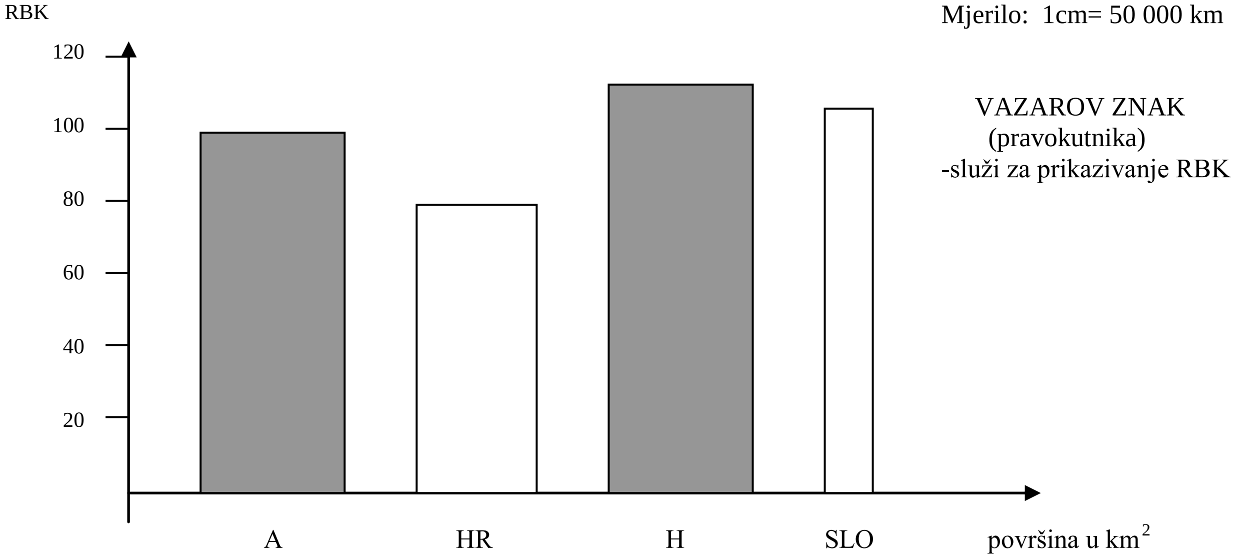
Kada se želi istaknuti i veličinu osnovice RBK, onda se za prikaz upotrebljavaju **pravokutnici (Varzarov znak),** kojima su osnovice proporcionalne bazama relativnog broja, a visine relativnim brojevima koordinacije. Visina pravokutnika predstavlja izračunatu vrijednost RBK, osnovica njegovu bazu (nazivnik RBK), a površine veličinu u brojniku RBK.

Primjer:
Tabela: Stanovništvo i površina odabranih europskih zemalja

ZEMLJA	BROJ STANOVNIKA U 000 (V_I)	POVRŠNA U KM^2 (BI)	BROJ STAN. NA KM^2 (RBK)
Austrija	7.987	83.858	7987000/83858=95,2
Hrvatska	4.776	56.610	4776000/56610=84,4
Mađarska	10.372	93.032	10372000/93032=111,5
Slovenija	2.052	20.251	2052000/20251=101,3

Izvor: Statistički ljetopis RH, 1996, str.20-821.

Tabela: Stanovništvo i površina odabranih europskih zemalja

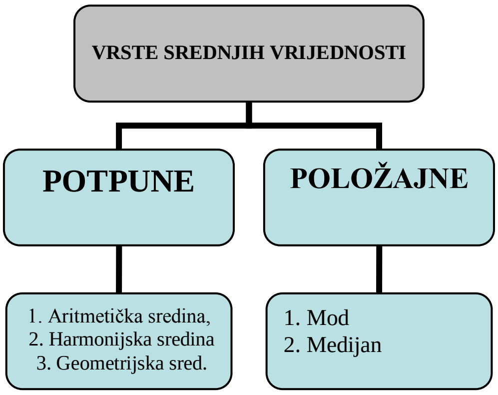


Izvor: Statistički ljetopis RH,1997.str 316

SREDNJE VRIJEDNOSTI STATISTIČKOG NIZA

Definicija srednje vrijednosti

Srednja vrijednost je konstanta kojom se predstavlja niz varijabilnih podataka.



.**Potpuna SV** određuje se na temelju svih podataka.

Položajna SV po pravilu je jedan modalitet statističke varijable, koji se identificira sukladno definiciji SV ili se aproksimira pomoću manjeg broja podataka.

ARITMETIČKA SREDINA

AS je najvažnija i najraširenija SV. Određuje se tako da se **zbroje vrijednosti numeričke varijable i podijele s njihovim brojem.**

Zbroj vrijednosti numeričke varijable naziva se **total,** pa je AS jednaki dio totala po jedinici.

Jednostavna (neponderirana) AS

Ako su dane pojedinačne vrijednosti numeričke varijable Xi: **X1, X2, X3,…, Xi, …, Xn,** njihova je aritmetička sredina, (to je suma svih Xi kroz N)ovo u zagradi je dovoljno znati za ispit)

$$AS = \bar{x} = X_1, X_2, X_3, ..., X_i, ..., x_N$$

$$\bar{x} = \frac{x_1 + x_2 + ... + x_N}{N} = \frac{\sum_{i=1}^{N} x_i}{N}$$

Vagana (ponderirana) AS

Ponderi su veličine kojima se množe (važu) vrijednosti numeričke varijable Xi. Postavljaju se pitanja:
Čime je vagana ponderirana AS? -frekvencijama

Što se važe pri izračunavanju vagane AS? -vrijednost numeričkog obilježja (frekvencije, relativne frekvencije ili njima proporcionalne veličine).

$$\bar{x} = \frac{\sum_{i=1}^{k} f_i x_i}{N}$$

$$\bar{x} = \frac{\sum_{i=1}^{k} p_i x_i}{100}$$

$$\bar{x} = \sum_{i=1}^{k} P_i x_i, \bar{x} = m$$

Izračunavanje AS pomoću apsolutnih frekvencija

Negrupirane jedinice	Grupirane jedinice
$\bar{x} = \frac{\sum_{i=1}^{N} x_i}{N}$	$\bar{x} = \frac{\sum_{i=1}^{k} f_i x_i}{\sum_{i=1}^{k} f_i}$

(rijetko se javlja na pismenom ispitu).

Oko a	$\bar{x} = a + \frac{\sum_{i=1}^{N} d_i}{N}$ $d_i = x_i - a$	$\bar{x} = a + \frac{\sum_{i=1}^{k} f_i d_i}{\sum_{i=1}^{k} f_i}$ $d_i = x_i - a$
Oko a uz b	$\bar{x} = a + b\frac{\sum_{i=1}^{N} d_i}{N}$ $d_i = \frac{x_i - a}{b}$	$\bar{x} = a + b\frac{\sum_{i=1}^{k} f_i d_i}{\sum_{i=1}^{k} f_i}$ $d_i = \frac{x_i - a}{b}$

(ovo se rijetko javlja na ispitu)

Izračunavanje AS pomoću relativnih frekvencija

Grupirane jedinice - Vagana (ponderirana) AS

Oko nule	$\bar{x} = \sum_{i=1}^{k} p_i x_i$

Linearna transformacija ili kodiranje

Oko a	$\bar{x} = a + \sum p_i d_i, d_i = x_i - a$
Oko a uz b (b≠0)	$\bar{x} = a + b\sum p_i d_i^{'}, d_i^{'} = \frac{x_i - a}{b}$

Primjer 1. Izračunavanje jednostavne AS (negrupirani podaci)

Koliko iznosi AS za navedeni numerički niz?
X_i: 17,17,21,34,35,40,41,42,50,50,53,55

Rješenje:

$$\bar{x} = \frac{1}{N}\sum x_i \qquad \sum x_i = 455$$

$$\bar{x} = \frac{1}{12} 455 = 37{,}91667$$

(zbroj svih x kroz ukupan broj brojeva)

Primjer 2: Izračunavanje vagane AS (grupirani podaci)

Koliko iznosi AS za navedeni numerički niz?

x_i:	500	550	600	700	750	800	$\sum$
f_i:	35	78	22	15	10	4	**164**
$f_i x_i$:	17 500	42 900	13 200	10 500	7 500	3 200	**94 800**

(pomoćni stupac koji treba izračunati)

Rješenje:

$$\bar{x} = \frac{\sum f_i x_i}{\sum f_i} = \frac{94800}{164} = 578{,}04878$$

Primjer 3: Izračunavanje AS distribucije frekvencija s razredima

Razredi	f_i	(razredna sredina) x_{si}	$f_i\ x_i$
0-5	123	2,5	307,5
5-10	157	7,5	1177,5
10-15	20	12,5	250,0
15-20	10	17,5	175,0
ukupno	310	-	1910

$$x_{si} = \frac{l_{i1} + l_{i2}}{2}$$

Rješenje

Vagana AS distribucije frekvencija iznosi:

$$\bar{x} = \frac{\sum f_i x_i}{\sum f_i} = \frac{1910}{310} = 6{,}1612903$$

3.ARITMETIČKA SREDINA ARITMETIČKIH SREDINA

Aritmetička sredina aritmetičkih sredina izračunava se kao vagana sredina u kojoj se za pondere uzima broj podataka za koje su izračunate pojedine sredine ili tom broju proporcionalne veličine.

Izračunavanje ASAS

Izračunavanje AS na temelju već izračunatih AS

- Kada je osnovni skup podijeljen u k-podskupova:
N1,N2,…, Nk

- U svakom od k-podskupova je izračunata AS:

$\overline{x}_1, \overline{x}_2, ..., \overline{x}_k$

$$\overline{\overline{x}} = \frac{\sum_{i=1}^{k} N_i \overline{x}_i}{\sum_{i=1}^{k} N_i} \qquad i=1,2,\ldots,k.$$

Primjer 5: Aritmetička sredina aritmetičkih sredina

Za podatke o prosječnoj ocjeni studenata EF u Osijeku, na pojedinim studijskim godinama, potrebno je izračunati prosječnu ocjenu za sve četiri studijske godine.

Studijska godina	Broj studenata Ni	Prosječna ocjena	NiXi
I	600	2,7	1620
II	400	3,0	1200
III	350	3,1	1085
IV	200	2,8	560
Ukupno	1550		4465

(ovaj četvrti stupac je pomoćni)

Rješenje:

$$\overline{\overline{x}} = \frac{\sum_{i=1}^{k} N_i \overline{x}_i}{\sum_{i=1}^{k} N_i} = \frac{4465}{1550} = 2{,}88$$

Prosječna ocjena na sve četiri studijske godine iznosila je 2,88.

HARMONIJSKA SREDINA

Definicija harmonijske sredine

HS je recipročna vrijednost AS njezinih recipročnih vrijednosti.

Uporaba HS (rjeđa od AS)

Za izračunavanje prosječnog vremena za izradu jedinice proizvoda,
- srednjeg vremena obrtaja kapitala,
- prosječnog vremena prijeđene jedinice puta i sl.,
- izračunavanje sredine relativnih brojeva s istim brojnicima

Izračunavanje HS

Jednostavna HS (negrupirani podaci) $$H = \frac{N}{\sum_{i=1}^{N} \frac{1}{x_i}}$$

Vagana HS (grupirani podaci) $$H = \frac{\sum_{i=1}^{k} f_i}{\sum_{i=1}^{k} \frac{f_i}{x_i}}$$

Primjer 7: Vagana HS

Prosječna prodajna cijena proizvoda 2002.godine, te struktura vrijednosti prodaje prema prodajnim područjima:

Područje	Prosječna prodajna cijena u kunama	Struktura vrijednosti prodaje u %
Sjever	490	35,0
Središnja regija	500	40,0
Jug	494	25,0

Odredite kolika je prosječna prodajna cijena za sva tri područja zajedno.

Rješenje:

$$H = \frac{\sum f_i}{\sum \frac{f_i}{x_i}} = \frac{35 + 40 + 25}{\frac{35}{490} + \frac{40}{500} + \frac{25}{494}} = \frac{100}{0{,}2020359} = 494{,}9615$$

Prosječna prodajna cijena za sva tri područja zajedno je (zaokruženo) 495 kuna.

5. GEOMETRIJSKA SREDINA

Geometrijska sredina je potpuna sredina vrijednosti numeričke varijable.

Izračunava se za niz pojedinačnih vrijednosti (jednostavna) i za grupirane podatke (vagana).

Koristi se u analizi vremenskih nizova, za izračunavanje prosječne stope promjene pojave.

Izračunavanje GS

Jednostavna GS:

$$G = \sqrt[N]{x_1 x_2,\ldots x_i,\ldots . x_N}$$

$$\log G = \frac{1}{N}\sum_{i=1}^{N}\log x_i$$

Vagana GS

$$G = \sqrt[N]{x_1^{f1} x_2^{f2},\ldots x_i^{fi},\ldots x_k^{fk}}$$

$$\log G = \frac{1}{N}\sum_{i=1}^{k} f_i \log x_i$$

Primjer 8: Jednostavna GS

Zadane su pojedinačne vrijednosti numeričke varijable:
x_i: 115 120 98 117 134 100 104 95 125 130 116

Kolika je geometrijska sredina? Odredite i aritmetičku sredinu.

broj brojeva u nizu

$$G = \sqrt[11]{115 \cdot 120 \cdot 98 \cdot \ldots \cdot 130 \cdot 116} = 112{,}997$$

$$\bar{x} = 113{,}727$$

Primjer: 9: Vagana GS

Distribucija anketiranih prema broju članova je:

x_i:	1	2	3	4	5	6
f_i:	3	6	26	15	6	4

Odredite vrijednost geometrijske i aritmetičke sredine.

$$G = \sqrt[N]{x_1^{f1} x_2^{f2},\ldots x_i^{fi},\ldots x_k^{fk}}$$

suma svih f_i

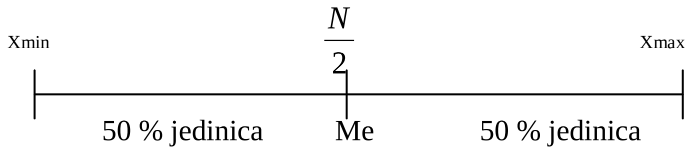
$$G = \sqrt[60]{1^3 \cdot 2^6 \cdot 3^{26} \cdot 4^{15} \cdot 5^6 \cdot 6^4} = 3{,}23$$

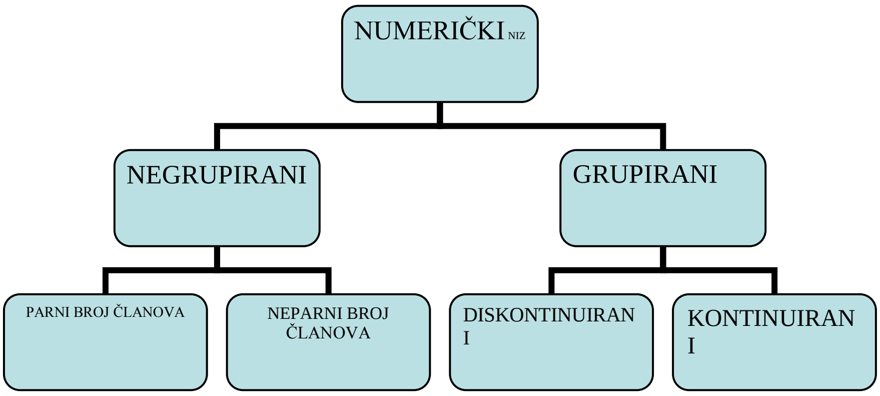
Aritmetička sredina distribucije je 3,45 članova.

6. MEDIJAN
(Me)

Medijan je položajna SV koja numerički niz uređen po veličini dijeli na dva jednaka dijela.

U jednom se dijelu numeričkog niza nalaze elementi koji imaju vrijednost numeričkog obilježja jednaku ili manju od medijana, dok se u drugom dijelu nalaze oni elementi koji imaju vrijednost numeričkog obilježja jednaku ili veću od medijana.





Medijan
za negrupirane statističke nizove

Ako je broj podataka **neparan,** medijan je vrijednost **varijable središnjeg člana niza uređenog po veličini.**
$Me = x_r$
Ako niz ima **parni** broj članova, medijan je jednak **poluzbroju vrijednosti varijable središnjih dvaju članova** uređenog niza.

$$Me = \left\{ \frac{X_r + X_{r+1}}{2} \right.$$

Primjer 10: Izračunavanje Me za NEPARNI broj podataka:
X_i: 2 12 3 4 2 5 2 7 8
Koliko iznosi medijan?

Uređeni podaci su:
2 2 2 3 4 5 7 8 12

N=9, X_r =5, Me=4

Medijan je 4, što znači da 50% elemenata niza ima vrijednost numeričkog obilježja 4 i manju od 4, a 50% elemenata niza ima vrijednost 4 i veću od 4.

Primjer 11: Izračunavanje Me za PARNI broj podataka

X_i: 2 12 3 4 2 5 2 7 8 2 9 2
Koliko iznosi medijan?

Uređeni niz podataka:
2 2 2 2 2 3 4 5 7 8 9 12

$$Me = \frac{x_r + x_{r+1}}{2} = \frac{3+4}{2} = 3{,}5$$

Me je srednja vrijednost dvaju sredšnjih članova (X_r, X_{r+1})
Me=3,5
Medijan je 3,5, što znači da 50% elemenata niza ima vrijednost numeričkog obilježja 3,5 i manju, a 50% elemenata niza ima vrijednost obilježja 3,5 i veću od 3,5.

MEDIJAN
za grupirane statističke nizove u razrede (bitno)

$$Me = l + \frac{\frac{N}{2} - \sum f_1}{f_{med}} i$$ -samo kad imamo zadanu distribuciju s razredima !

N - broj frekvencija (apsolutnih ili relativnih): f_{med} -frekvencija medijalnog razreda; i - veličina medijalnog razreda; l_1 - donja granica medijalnog razreda; $\sum f_1$ - frekvencija kumulativnog niza ″manje od″ ispred medijalnog razreda.

Primjer 12: Izračunavanje Me za distribuciju frekvencija s razredima

Naslov:Osobe prijavljene u Hrvatskom zavodu za zapošljavanje, stanje potkraj 1999.godine:

Godine života	broj osoba	Kumulativni niz ″manji od″	Veličina razreda (i)
15-20	67170	67170	5
20-25	48482	115652	5
25-30	119819	235471	5
30-40	82263	317734	10
40-50	10604	32338	10
50-(65)	13392	341730	(15)
ukupno	341730	-	-

N

Veličina rasta (i) predstavlja razliku između donje i gornje granice razreda,

RJEŠENJE:

Medijalni razred $\left(\frac{N}{2}\right)$:

$$Me = l_1 + \frac{\frac{N}{2} - \sum f_1}{f_{med}} i = 25 + \frac{170865 - 115652}{119819} 5 = 27.304402$$

$$\frac{N}{2} = \frac{341730}{2} = 170865$$

kumulativni niz ″manje od″
(ista ili prva veća vrijednost)

primjer 25-30

Medijan iznosi (zaokruženo) 27 godina. Prema njemu, dob prve polovice osoba koje su bile prijavljene u zavodu za zapošljavanje iznosila je 27 i manje godina, a druga polovica osoba je starija od 27 godina.

Primjer 13: Izračunavanje Me za diskontinuirani numerički niz.
Test sadrži pet zadataka.Broj riješenih zadataka 43 studenta bio je ovakav:

BROJ RJEŠENIH ZADATAKA	BROJ STUDENATA f_i	KUMULATIVNI NIZ ″MANJE OD″
0	3	3
1	7	10
2	12	22
3	16	38
4	3	41
5	2	43
UKUPNO	43	-

Koliki je medijalni broj riješenih zadataka?

$$Merazred = \frac{N}{2} = \frac{\sum f_i}{2} = \frac{43}{2} = 21{,}5 \longrightarrow$$ 22; 2 = Medijan (u istom razredu kao mediijalni razred)

MOD

Mod je položajna SV.
To je vrijednost ili modalitet varijabli koji se najčešće pojavljuje u nizu. Mod postoji ako su u nizu bar dva jednaka podatka.
Prema tome, mod je modalitet nominalne varijable, rang - varijable, ili numeričke varijable, sa najvećom frekvencijom.
Mod dijeli distribuciju na rastuću i padajuću stranu.

Izračunavanje moda distribucije frekvencija sa razredima

$$M_o = l_1 + \frac{(b-a)}{(b-a)+(b-c)} i$$

b-najveća (korigirana) frekvencija;
a-frekvencija ispred nje;
c-frekvencija iza najveće korigirane frekvencije;
l-donja granica modalnog razreda;
i-veličina modalnog razreda

Kada su razredi jednakih veličina koristiti će se originalne frekvencije. Ako su razredi nejednakih veličina, vrši se "korigiranje frekvencije" ($f_{ci} = \frac{f_i}{i}$ **).**
Modalni je razred onaj sa najvećom frekvencijom.

Primjer 14: Izračunavanje moda za niz negrupiranih podataka (ispit)
Za statistički niz:
X_i: 2 12 3 4 2 5 2 7 8 2 9 2

Odredite koliko iznosi mod za navedeni niz podataka

Rješenje: Na temelju podataka iz tabele, može se zaključiti da je najčešća vrijednost 2. Prema tome, mod je 2.

Primjer 15:Izračunavanje moda za niz kvalitativnih podataka.
Rezultati prvog kolokvija iz statistike održanog u zimskom semestru 2002.g. na FTHM Opatija (grupa zadataka B), prikazani u tabeli:

ocjena	izvrstan	Vrlo dobar	dobar	dovoljan	nedovoljan
Br. studenata	2	9	20	23	57

Odredite MOD za niz podataka u tabeli.

Varijabla ocjena je kvalitativna (rang-varijabla). Mod je ocjena koju je postigao najveći broj studenata. U primjeru je modalna ocjena - nedovoljan.

Mod nije najveća frekvencija, već je mod obilježje (kvalitativno, kvantitativno) sa najvećom frekvencijom.

Primjer 16. Izračunavanje moda distribucije rekvencija sa razredima.
Broj prometnih nezgoda prema godinama starosti:

God. starosti	Br. sudionika u prom. nezgodama f_i	Precizne granice razreda	Razredna sredina x_{si}	Veličina razreda i
0-4	12	0-5	,5	5
5-9	20a	5-10	7,5	5
10-14	28b	10-15	12,5	5
15-19	19c	15-20	17,5	5
20-24	11	20-25	22,5	5
ukupno	90	-	-	

Modalni razred s najvećom frekvencijom

(x_i) $x_{si}=\frac{l_1+l_2}{2}$

$f_{ci}=\frac{f_i}{i}$ - nije sada primjer

Precizne granice razreda (stupac 3) određujemo tako da donja granica razreda kojeg promatramo mora biti jednaka gornjoj granici prethodnog razreda.

Rješenje:

$$M_o = 10 + \frac{28-20}{(28-20)+(28-19)}5 = 12{,}35 god$$

Dobna skupina koja najčešće stradava u prometnim nezgodama je stara 12, 35 godina.

MJERE DISPERZIJE (RASPRŠNOSTI)

MJERE DISPERZIJE (RASPRŠNOSTI)	
MJERE DISPERZIJE (RASPRŠNOSTI)	RELATIVNE MJERE DISPERZIJE
1. Raspon varijacije (R) 2.Interkvartil (I_Q) 3.Varijanca(δ^2) 4.Standardna devijacija (δ)	1.Koeficijent varijacije(V) 2.Koeficijent kvartalne devijacije (I_Q)

1. Raspon varijacije je najjednostavnija mjera disperizije.

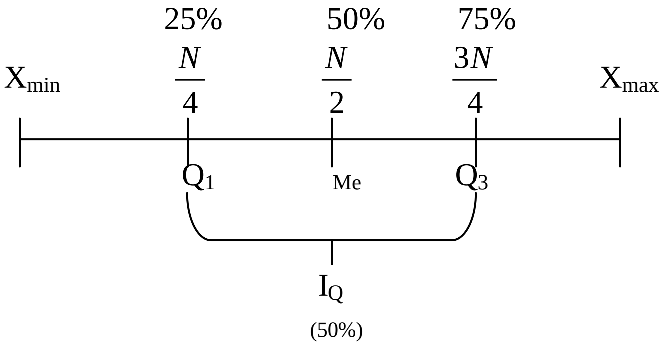
Ako su svi podaci jednaki, vrijednost mu je jednaka nuli, a povećava se sa povećanjem stupnja varijabilnosti podataka.

Raspon varijacija niza kvalitativnih podataka je <u>razlika između najveće i najmanje vrijednosti u nizu, tj.</u>

$$R_x = X_{max} - X_{min}.$$

2. Interkvartil je apsolutna mjera disperzije. To je raspon varijacije središnjih 50% članova niza uređenih podataka. Pripadajuća relativna mjera je <u>koeficijent kvartilne devijacije.</u>

Interkvartil je razlika između gornjeg i donjeg kvartila, a koeficijent kvartilne devijacije je omjer interkvartila i zbroja kvartila



Interkvartili

Koeficijent kvartne devijacije $I_{Q=} Q_3 \;\; Q_1$

$$V_Q = \frac{Q_3 - Q_1}{Q_3 + Q_1}$$

I_Q

$0 \leq V_Q \leq 1$

KVARTILI za negrupirane podatke

Donji kvartili (Q_1)

Gornji kvartili (Q_3)

$$Q_1 = \begin{cases} X_r, \frac{N}{4} \neq INT \\ \frac{X_r + X_{r+1}}{2}, \frac{N}{4} = INT \end{cases}$$

$$Q_3 = \begin{cases} X_r, \frac{3N}{4} \neq INT \\ \frac{X_r + X_{r+1}}{2}, \frac{3N}{4} = INT \end{cases}$$

Izračunavanje kvartila u distribucijama frekvencija s razredima

Donji kvartil (Q_1)

Gornji kvartil (Q_3)

$$Q_1 = l_1 + \frac{\frac{N}{4} - \sum f_1}{f_{Q1}} i$$

$$Q_3 = l_1 + \frac{\frac{3N}{4} - \sum f_1}{f_{Q3}} i$$

Opis simbola iz formule:

Donji (prvi) kvartil, Q_1:

$\frac{N}{4}$ - četvrtina elemenata niza (određivanje razreda Q_1)

L_1- donja granica razreda Q_1

$\sum f_1$- frekvencija kumulativnog niza «manje od) ispred razreda Q_1

f_{Q1} - originalna frekvencija Q_1

i - veličina razreda Q_1

Gornji (treći) kvartil Q_3:

$\frac{3N}{4}$ - tri četvrtine elemenata niza (određivanje razreda Q_3)

L1 - donja graniza razreda Q_3

$\sum f_1$- frekvencija kumulativnog niza « manje od « ispred razreda Q_3

f_{Q3} -originalna frekvencija Q_3

i - veličina razreda Q_3

Primjer - . Izračunavanje kvartila za distribuciju frekvencija sa razrednim

Tabela. Nezaposlenost prijavljene Hrvatskom zavodu za zapošljavanje, stanje 31.12.1997.!!!

	Sati nastave X_i	Studenti f_i	Sredina razreda X_{si}	Veličina razreda i	Kumulativni niz «manje od»
	118-126	3	122	9	3
	127-135	5	131	9	8
Q_1	136-144	9	140	9	17
	144-153	12	149	9	29
Q_3	154-162	5	158	9	34
	163-171	4	167	9	38
	171-180	2	176	9	40
	ukupno	40	-	-	-

Zadano treba izračunati

Donji kvartil Q_1

frekvencija ispred razreda

$$Q_1 = l_1 + \frac{\frac{N}{4} - \sum f_1}{f_{Q1}} i$$

$$Q_1 = 136 + \frac{\frac{40}{4} - 8}{9} 9 = 138$$

Gornji kvartil Q_3

$$Q_3 = l_1 + \frac{\frac{3N}{4} - \sum f_1}{f_{Q3}} i$$

$$Q_3 = 154 + \frac{\frac{3 \cdot 40}{4} - 29}{5} 9 = 158{,}8$$

Momenti distribucije frekvencija

MOMENTI OKO SREDINE

Neprupirani podaci

$$\mu_r = \frac{\sum (X_i - \overline{X})^r}{N}$$

$$\mu_2 = \frac{\sum (X_i - \overline{X})^2}{N}$$

$\mu_0 = 1; \mu_{1=0}$ $\longrightarrow$ const.

MOMENTI OKO SREDINE

Grupirani podaci

$$\mu_r = \frac{\sum f_i (X_i - \overline{X})^r}{\sum f_i}$$

$$\mu_2 = \frac{\sum f_i (X_i - \overline{X})^2}{\sum f_i}$$

MOMENTI OKO NULE
Negrupirani podaci

$$m_r = \frac{\sum (x_i - 0)^r}{N} = \frac{\sum x_i^r}{N}$$

$$m_1 = \overline{x}$$

MOMENTI OKO NULE
Grupirani podaci

$$m_r = \frac{\sum f_i (x_i - 0)^r}{\sum f_i} = \frac{\sum f_i x_i^r}{\sum f_i}$$

VARIJANCA (δ²)

Varijanca je prosječno kvadratno odstupanje od prosjeka, tj. od aritmetičke sredine. Varijanca (sigma) je jednaka drugom momentu oko sredine (Mi2).

Negrupirani podaci

$$\delta^2 = \frac{\sum (x_i - \overline{x})^2}{N} = \mu^2$$

Grupirani podaci

$$\delta^2 = \frac{\sum f_i (x_i - \overline{x})^2}{\sum f_i} = \mu^2$$

STANDARDNA DEVIJACIJA (δ)

Standardna devijacija je prosječno odstupanje originalnih vrijednosti od aritmetičke sredine, tj. pozitivni drugi korijen iz varijance:

$$\delta = \sqrt{\delta^2}$$ varijanca

KOEFICIJENT VARIJACIJE

Koeficijent varijacije je relativna mjera disperzije, a izražena je omjerom standardne devijacije i aritmetičke sredine pomnoženim sa 100.

$$V = \frac{\delta}{\overline{x}} 100$$

Koeficijent varijacije je postotak standardne devijacije od aritmetičke sredine.

MJERE ASIMETRIJE I MJERE ZAOBLJENOSTI

Mjera asimetrije je brojčana karakteristika **načina rasporeda** podataka.
Najvažnije su mjere asimetrije: **koeficijent asimetrije (znak) Pearsonova i Bowleyjeva mjera.**
Koeficijent asimetrije alfa tri omjer je trećeg momenta oko sredine i standardne devijacije podignute na treću potenciju, tj.

$$\alpha_3 = \frac{\mu_3}{\delta^3}$$ koeficijent asimetrije

Temelj mjere je treći moment oko sredine (alfa tri). U simetričnom rasporedu alfa tri jednak je nuli, u pozitivno asimetričnom je pozitivan, a u negativno asimetričnom negativan.

Koeficijent asimetrije alfa tri uobičajeno poprima vrijednosti iz intervala ± 2, a ponekad i veće vrijednosti.

Treći moment oko sredine (μ_3)

Negrupirani podaci

$$\mu_3 = \frac{\sum \left(x_i - \bar{x}\right)^3}{N}$$

Grupirani podaci

$$\mu_3 = \frac{\sum f_i \left(x_i - \bar{x}\right)^3}{f_i}$$

PEARSONOVA MJERA ASIMETRIJE

Ta je mjera standardizirano odstupanje vrijednosti medijana ili moda od aritmetičke sredine.
U simetričnim distribucijama kontinuirane varijable sve su tri vrijednosti jednake, pa je razlika moda ili medijana i aritmetičke sredine jednaka nuli.

U pozitivno asimetričnim distribucijama ta je razlika pozitivna, a u negativno asimetričnim razlika je negativna.

Pearsonova mjera uobičajeno poprima vrijednosti iz intervala ± 3, ali može biti i izvan tog intervala.

Pearsonova mjera asimetrije:

$$S_{k1} = \frac{\left(\bar{x} - Mo\right)}{\delta}$$

$$S_{k1} = \frac{3\left(\bar{x} - Me\right)}{\delta}$$

BOWLEYEVA MJERA ASIMETRIJE

Bowleyeva mjera asimetrije se temelji na odnosima kvartila i medijana.

U simetričnom rasporedu vrijednosti razlika gornjeg kvartila i medijana, jednaka je razlici medijana i donjeg kvartila tj. $Q_1 + Q_3 - 2Me = 0$

U pozitivno asimetričnom rasporedu razlika gornjeg kvartila i medijana veća je od razlike medijana i donjeg kvartila, a u negativno asimetričnom razlika gornjeg kvartila i medijana manja je od razlike medijana i donjeg kvartila.

Mjera poprima vrijednosti iz intervala $\pm$ 1.

Bowleyeva mjera asimetrije: $$S_{KQ} = \frac{Q_1 + Q_3 - 2Me}{Q_3 - Q_1}$$

Primjer: Izračunavanje mjera asimetrije
Tabela: Sudionici prometnih nezgoda prema godinama starosti

Godina starosti	Broj sudionika f_i	Precizne granice	Sredina razreda X_{si}	Veličina razreda	$f_i X_{si}$	$f_i (X_i - X)^3$
0-4	12	0-5	2,5	5	3,0	-11294,394
5-9	20	5-10	7,5	5	150,0	-2211,84
10-14	28	10-15	12,5	5	350,0	0,224
15-19	19	15-20	17,5	5	332,5	2871,552
20-24	11	20-25	22,5	5	247,5	11673,288
ukupno	90	-	-	-	1110,0	838,919

Koeficijent asimetrije (α_3)
Aritmetička sredina = 12,3 godine
Standardna devijacija = 6,03

$$\mu_3 = \frac{\sum f_i\left(x_i - \overline{x}\right)^3}{\sum f_i} = \frac{838,92}{90} = 0,043$$

$$\alpha_3 = \frac{\mu_3}{\delta^3} = \frac{9,32}{6,03^3} = \frac{9,32}{219,26} = 0,043$$

Pearsonove mjere asimetrije
Aritmetička sredina = 12,3 godine
Mod = 12,35 godina
Medijan = 12,3 godine

$$S_{k1} = \frac{\overline{x} - Mo}{\delta} = \frac{12,3 - 12,35}{6,03} = 0,017$$

$$S_{k2} = \frac{3\left(\overline{x} - Me\right)}{\delta} = \frac{3(12,3 - 12,3)}{6,03} = \frac{0}{6,03} = 0$$

Bowleyeva mjera asimetrije

Medijan = 12,3

Donji kvartil = 7,63

$$S_{KQ} = \frac{Q_1 + Q_3 - 2Me}{Q_3 - Q_1} = \frac{7,63 + 16,97 - 2 \cdot 12,3}{16,97 - 7,63} = 0$$

Gornji kvartil : 16,97

MJERE ZAOBLJENOSTI

Zaobljenost modalnog vrha distribucije mjeri se **koeficijentom zaobljenosti (α_4).**
Koeficijent zaobljenosti α_4-omjer četvrtog momenta oko sredine i standardne devijacije na četvrtu potenciju, tj. alfa četiri: $\alpha_4 = \frac{\mu_4}{\delta_4}$**.**
Četvrti moment oko sredine μ_4:

Negrupirani podaci

$$\mu_4 = \frac{\sum\left(x_i - \overline{x}\right)^4}{N}$$

Grupirani podaci

$$\mu_4 = \frac{\sum f_i\left(x_i - \overline{x}\right)^4}{\sum f_i}$$

Ako je u empirijskoj distribuciji frekvencija:
- α_4 =3, **distribucija ima zaobljenost normalne krivulje,**
- α_4 **>3, distribucija je oblikom šiljastija od normalne,**
- α_4 <3, **distribucija je oblikom plosnatija od normalne,**
- α_4 = 1,8, **disribucija je pravokutnog oblika,**
- α_4 < 1,8, **riječ je o U-distribuciji.**

Primjer: Koeficijent zaobljenosti
Tabela: Sudionici prometnih nezgoda prema godinama starosti

Godina starosti	Broj sudionika f_i	Precizne granice	Sredina razreda X_{si}	Veličina razreda	$f_i X_{si}$	$f_i (X_i - X)^3$	$f_i (X_i - X)^4$
0-4	12	0-5	2,5	5	3,0	-11294,394	110684,17
5-9	20	5-10	7,5	5	150,0	-2211,84	10616,832
10-14	28	10-15	12,5	5	350,0	0,224	0,0448
15-19	19	15-20	17,5	5	332,5	2871,552	13872,07
20-24	11	20-25	22,5	5	247,5	11673,288	119067,53
ukupno	90	-	-	-	1110,0	838,919	254260,64

Četvrti moment oko sredine:

$$\mu_4 = \frac{\sum f_i\left(x_i - \bar{x}\right)^4}{\sum f_i} = \frac{254260,64}{90} = 2825,12$$

Koeficijent zaobljenosti:

$$\alpha_4 = \frac{\mu_4}{\delta_4} = \frac{2825,12}{1322,12} = 2,1368$$

Koeficijent zaobljenosti je manji od 3, pa se može zaključiti da se radi o distribuciji koja je plosnatija od normalne.
Primjer:
Vježba 13/98 Numerički niz

x_i	p_i	f_i	Kumulativni niz ″manje od»	$f_i x_i$	$\frac{f_i}{x_i}$	$f_i (x_i - \bar{x})^2$	$f_i (x_i - \bar{x})^3$	$f_i (x_i - \bar{x})^4$
0	0,05	10	10	0	0	99,225	-312,56	984,56
1	0,1	20	30	20	20	92,45	-198,77	427,35
2	0,15	30	60	60	15	39,675	-45,63	52,47
3,3	0,3	60	120	180	20	1,35	-0,20	0,03
4	0,25	50	170	200	12,5	36,125	30,71	26,10
5	0,05	10	180	50	2	34,225	63,32	117,14
6	0,1	20	200	120	3,3	162,45	462,98	1319,50
$\sum$		200		630	72,8	465,5	-0,15	3027,15

$f_i = N * p$

Numerički niz zadan je svojim relativnim frekvencijama : p(x=0)=0,05, p(k=1)=0,1, p(x=2)=0,15, p(x=3)=0,3 p(x=4)=0,25, p(x=5)=0,05, p(x=6)=0,1. Ako je N=200 odredite frekvencije tog niza, zatim mod, medijan, aritmetičku sredinu, harmonijsku sredinu, koeficijent asimetrije i koeficijent zaobljenosti.
0 0,05 10 10 0 0 99,225 -312,56 924,

1. Mod
Mo=3

2. Medijan

$$Me = \frac{N}{2} = \frac{200}{2} = 100 \qquad Me=3$$

3. Aritmetička sredina

$$x = \frac{\sum f_i x_i}{\sum f_i} = \frac{630}{200} = 3{,}15$$

4. Harmonijska sredina

$$H = \frac{\sum f_i}{\sum \frac{f_i}{x_i}} = \frac{200}{72{,}8} = 2{,}75$$

5. Koeficijent asimetrije

$$\alpha_3 = \frac{\mu_3}{\delta^3}$$

$$\alpha_3 = -\frac{0{,}00075}{1{,}533} = \mathbf{0{,}0002}$$

$$\delta = \sqrt{\mu_2} \longrightarrow \mu_2 = \frac{\sum f_i \left(x_i - \bar{x}\right)^2}{\sum f_i} = \frac{465{,}5}{200} = 2{,}33$$

$$\delta = \sqrt{2{,}33} = 1{,}53$$

$$\mu_3 = \frac{\sum f_i \left(x - \bar{x}\right)^3}{\sum f_i} = -\frac{0{,}15}{200} = 0{,}00075$$

6. Koeficijent zaobljenosti

$$\alpha_4 = \frac{\mu_4}{\delta^4} \longrightarrow \mu_4 = \frac{\sum f_i \left(x_i - \bar{x}\right)^4}{\sum f_i} = \frac{3027{,}15}{200} = 15{,}14$$

$$\alpha_4 = \frac{15{,}14}{15{,}34^4} = 2{,}76$$