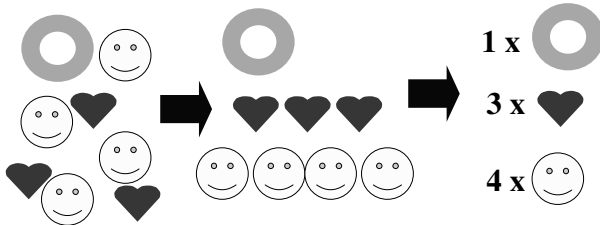


RAZDIOBA OBILJEŽJA



IST - P2

1

Primjer

IST - P2

2

Na jednom čovjeku izvršeno je 50 mjerenja vremena reakcije. Dobiveni su sljedeći podaci (u tisućinkama sekunde):

196	173	186	189	173	165	167	160	140	174
180	151	157	164	154	169	190	180	163	157
169	167	165	160	177	165	157	177	159	175
166	173	185	177	184	183	162	192	174	162
165	172	158	169	146	170	171	169	168	153



IST - P2

3

TABLICA FREKVENCIJA

IST - P2

4

... tablica frekvencija ...

- tablica u kojoj su originalni podaci sažeti u određeni broj kategorija (*razreda*) koje su opisane numerički izraženim granicama

raspon (interval) razreda

- razlika granica razreda

IST - P2

5

... tablica frekvencija ...

sredina razreda

- broj koji najbolje reprezentira dani razred

računanje sredine razreda:

- *diskretne varijable*:

suma granica razreda / 2

- *kontinuirane varijable*:

suma donjih granica razreda / 2

IST - P2

6

apsolutna frekvencija razreda(f)

- broj podataka koji pripadaju intervalu tog razreda

kumulativna frekvencija razreda(cf)

- broj podataka čija je vrijednost manja ili jednaka gornjoj granici razreda

relativna frekvencija razreda(rf)

- apsolutna frekvencija razreda podijeljena s ukupnim brojem podataka

kumulativna relativna frekvencija razreda (crf)

- kumulativna frekvencija razreda podijeljena s ukupnim brojem podataka

za dani razred:● **apsolutna frekvencija:**

- koliko mjerenja ima vrijednosti iz intervala tog razreda

● **apsolutna kumulativna frekvencija:**

- koliko mjerenja ima vrijednost manju ili jednaku gornjoj granici tog razreda

● **relativna frekvencija:**

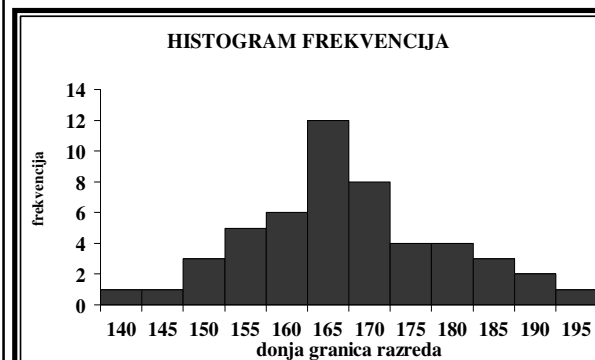
- koliki postotak mjerenja od ukupnog broja mjerenja ima vrijednost iz intervala tog razreda

● **kumulativna relativna frekvencija:**

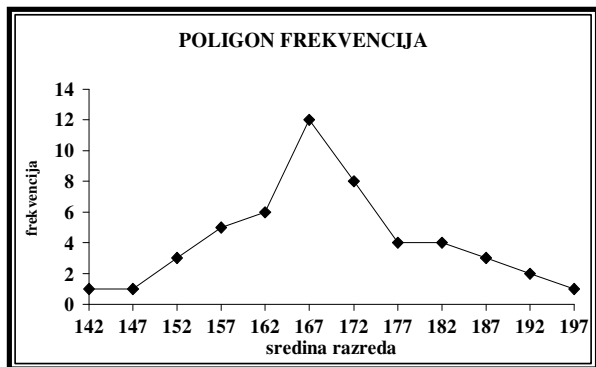
- koliki postotak mjerenja od ukupnog broja mjerenja ima vrijednost manju ili jednaku gornjoj granici tog razreda

Tablica frekvencija podataka iz primjera

min = 140		max = 196				
Broj r.	Granice r.	Sredina r.	f	cf	rf	crf
1	140 - 144	142	1	1	0,02	0,02
2	145 - 149	147	1	2	0,02	0,04
3	150 - 154	152	3	5	0,06	0,10
4	155 - 159	157	5	10	0,10	0,20
5	160 - 164	162	6	16	0,12	0,32
6	165 - 169	167	12	28	0,24	0,56
7	170 - 174	172	8	36	0,16	0,72
8	175 - 179	177	4	40	0,08	0,80
9	180 - 184	182	4	44	0,08	0,88
10	185 - 189	187	3	47	0,06	0,94
11	190 - 194	192	2	49	0,04	0,98
12	195 - 199	197	1	50	0,02	1,00
UKUPNO :			50		1,00	

RAZDIOBA FREKVENCIJA

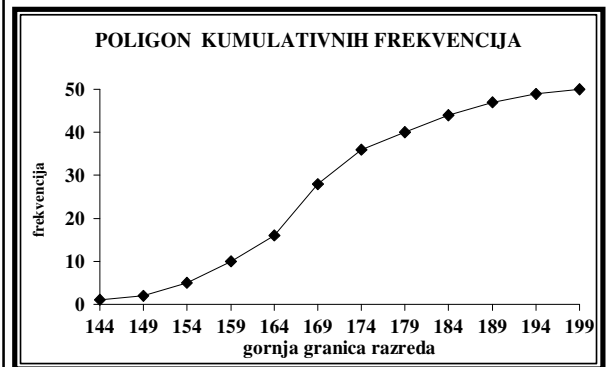
RAZDIJOBA FREKVENCIJA



IST - P2

13

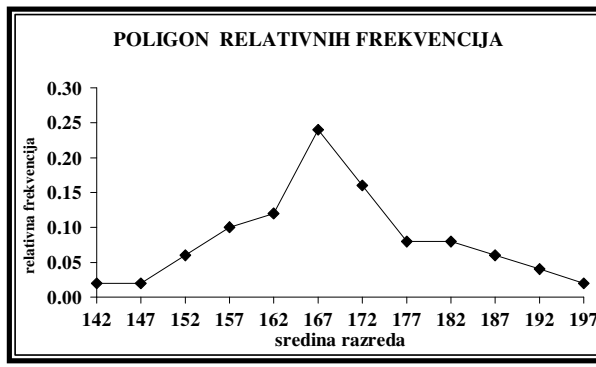
RAZDIJOBA FREKVENCIJA



IST - P2

14

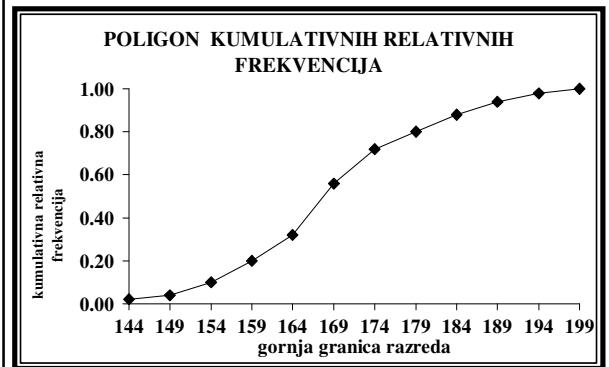
RAZDIJOBA FREKVENCIJA



IST - P2

15

RAZDIJOBA FREKVENCIJA



IST - P2

16

Stablo i list (stem-and-leaf)

f	stablo	list
1.00	14 .	0
1.00	14 .	6
3.00	15 .	134
5.00	15 .	77789
6.00	16 .	002234
12.00	16 .	555567789999
8.00	17 .	01233344
4.00	17 .	5777
4.00	18 .	0034
3.00	18 .	569
2.00	19 .	02
1.00	19 .	6

IST - P2

17

PAZŃJA !

- broj razreda

- preveliki broj razreda => male frekvencije ili prazni razredi
- premali broj razreda => razredi jako saŹeti => izgubljeno puno informacija
- uobičajeno: 10-20 razreda (ovisno o broju i prirodi podataka)
- kod nominalnih varijabli:
broj kategorija=broj razreda

IST - P2

18

PAŽNJA !

- granice razreda

- najmanje na onoj točnosti na kojoj je izvršeno mjerenje
- određene tako da **SVAKI PODATAK PADNE U SAMO JEDAN OD RAZREDA!**

IST - P2

19

OPISIVANJE RAZDIOBE PODATAKA

IST - P2

20

PARAMETAR I STATISTIKA

POPULACIJA



1. UZORAK



2. UZORAK

⋮

⋮



n-ti UZORAK

IST - P2

21

PARAMETAR I STATISTIKA



aritmetička sredina visine populacije
= 175.4



aritmetička sredina visine 1. uzorka
= 172.2



aritmetička sredina visine 2. uzorka
= 178.1



aritmetička sredina visine n-tog uzorka
= 173.7

IST - P2

22

PARAMETAR I STATISTIKA

● parametar:

- vrijednost (obično nepoznata) koja predstavlja neku karakteristiku populacije
- unutar populacije, parametar je nepromjenljiva vrijednost koja **NE VARIRA**

● statistika:

- veličina izračunata iz podataka izmjerenih na uzorku
- vrijednost statistike **MIJENJA SE** od uzorka do uzorka

IST - P2

23

Uobičajene oznake:

	Ocjena parametra (statistika)	Parametar populacije
ARITMETIČKA SREDINA	\bar{x}	μ
STANDARDNA DEVIJACIJA	s	σ
PROPORCIJA	p	π

IST - P2

24

OPISIVANJE RAZDIOBE PODATAKA



- sredina
- varijabilnost
- oblik



IST - P2

25

MJERE SREDINE (centralne tendencije) (srednje vrijednosti, prosjeci, mjere lokacije)

- karakteristične vrijednosti oko kojih se grupiraju mjereni podaci

ARITMETIČKA SREDINA

MEDIJAN (središnja vrijednost)

MOD (dominantna ili tipična vrijednost)

GEOMETRIJSKA SREDINA

HARMONIJSKA SREDINA

IST - P2

26

ARITMETIČKA SREDINA

- oznake: \bar{X} uzorak μ populacija

ARITMETIČKA SREDINA INDIVIDUALNIH PODATAKA

$$\mu = \frac{x_1 + x_2 + \dots + x_N}{N} = \frac{\sum_{i=1}^N x_i}{N}$$

x_i ... vrijednosti mjerenog obilježja
N ... ukupan broj podataka

IST - P2

27

Primjer

IST - P2

28

Kolika je aritmetička sredina niza podataka:

1, 2, 3, 3, 4, 5?

$$\begin{aligned} \bar{x} &= \frac{x_1 + x_2 + x_3 + x_4 + x_5 + x_6}{6} = \\ &= \frac{1 + 2 + 3 + 3 + 4 + 5}{6} = \frac{18}{6} = 3 \end{aligned}$$

Kolika je aritmetička sredina niza podataka:

1, 1, 1, 1, 2, 12?

$$\bar{x} = \frac{1 + 1 + 1 + 1 + 2 + 12}{6} = \frac{18}{6} = 3$$

Imamo dva niza podataka:

A: 1 2 3 3 4 5
B: 1 1 1 1 2 12

aritmetička sredina niza A = 3

aritmetička sredina niza B = 3

- loše opisuje niz B
- veliki utjecaj ekstremne vrijednosti (12)

IST - P2

30

PREDMET	OCJENA		ECTS
	STUDENT 1	STUDENT 2	
Osnove istraživačkog rada u sestrinstvu	2	5	4
Zdravstvena njega odraslih II	5	2	9
Zdravstvena njega psihijatrijskih bolesnika	5	3	8
Klinička medicina III	3	5	4
Σ	15	15	25

$$\text{uspjeh}_{s_1} = \frac{2 \cdot 4 + 5 \cdot 9 + 5 \cdot 8 + 3 \cdot 4}{25} = \frac{105}{25} = 4.20$$

$$\text{uspjeh}_{s_2} = \frac{5 \cdot 4 + 2 \cdot 9 + 3 \cdot 8 + 5 \cdot 4}{25} = \frac{82}{25} = 3.28$$


IST - P2

37

U dva navrata vršeno je mjerenje neke dužine i dobiveni su slijedeći rezultati:
 $\bar{x}_1 = 20\text{cm}$; $N_1 = 15$
 $\bar{x}_2 = 23\text{cm}$; $N_2 = 60$


a) Kolika je zajednička aritmetička sredina?
b) Kolika je zajednička aritmetička sredina za $N_1=60; N_2=15$?

a) $\bar{x}_{zaj} = \frac{20 \cdot 15 + 23 \cdot 60}{15 + 60} = \frac{300 + 1380}{75} = 22.4\text{cm}$
b) $\bar{x}_{zaj} = \frac{20 \cdot 60 + 23 \cdot 15}{15 + 60} = \frac{1200 + 354}{75} = 20.6\text{cm}$

 aritmetička sredina osjetljiva je **ne samo na vrijednost nego i na broj podataka**

IST - P2

38



Aritmetička sredina nema smisla, tj. nije dobar reprezentant podataka ako je:

- razdioba asimetrična
- broj podataka mali, a varijabilnost velika (velike razlike u vrijednostima podataka)

IST - P2

39

Imamo niz podataka:
A: 2 2.5 3.5 3 4


Kolika je suma odstupanja pojedinačnih vrijednosti od aritmetičke sredine? Kolika je suma kvadrata odstupanja pojedinačnih vrijednosti od aritmetičke sredine te od vrijednosti 2; 4; 5?

$(2-3)+(2.5-3)+(3.5-3)+(3-3)+(4-3) = -1-0.5+0.5+0+1 = 0$

x_i	$(x_i-3)^2$	$(x_i-2)^2$	$(x_i-4)^2$	$(x_i-5)^2$
2	1	0	4	9
2.5	0.25	0.25	2.25	6.25
3	0	1	1	4
3.5	0.25	2.25	0.25	2.25
4	1	4	0	1
Σ	2.5	7.5	7.5	22.5

40

SVOJSTVA ARITMETIČKE SREDINE



- $\sum_{i=1}^N (x_i - \mu) = 0$
- $\sum_{i=1}^N (x_i - \mu)^2 < \sum_{i=1}^N (x_i - a)^2, \forall a \neq \mu$

IST - P2

41

MEDIJAN (središnja vrijednost)

- vrijednost koja se u nizu podataka poredanih po veličini nalazi točno u sredini - **središnja vrijednost po položaju**
- vrijednost medijana:
 - za **neparan N**: vrijednost koja se nalazi na $(N+1)/2$ mjestu
 - za **paran N**: sredina vrijednosti podataka koji se nalaze na $N/2$ i $(N+2)/2$ mjestu

IST - P2

42

MEDIJAN (središnja vrijednost)

● prednosti:

- na vrijednost medijana *ne utječu ekstremne vrijednosti*

⇒ pogodan kao mjera centralne tendencije kod **asimetričnih raspodjela**

oznaka: M_e (C , M_d)

IST - P2

43

PRIMJER

Za nizove podataka iz primjera:

A: 1 2 3 3 4 5

B: 1 1 1 1 2 12

niz A: medijan.... $M_e=3$ arit. sred. $\bar{X}=3$

niz B: medijan.... $M_e=1$ arit. sred. $\bar{X}=3$

IST - P2

44

MOD (dominantna vrijednost)

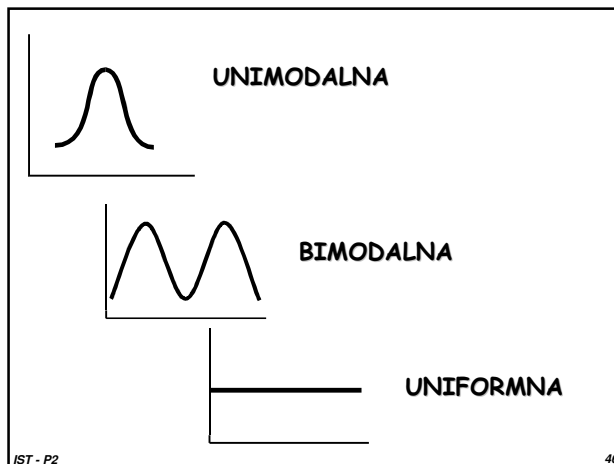
- vrijednost koja se u nizu mjerenja **najčešće** javlja (dominira svojom frekvencijom)

- na mod ne utječu ni broj ni veličina podataka, već **samo frekvencija**

oznaka: M_o

IST - P2

45



IST - P2

46

MJERE RASPRŠENJA (VARIJABILNOSTI)

IST - P2

47

(mjere raspršenja, varijabilnosti, disperzije)

RASPON

KVANTILE

VARIJANCA

STANDARDNA DEVIJACIJA

KOEFICIJENT VARIJABILNOSTI

IST - P2

48

RASPON

$$R = \max - \min$$

● nedostaci:

- uzima u obzir samo dvije ekstremne vrijednosti koje uopće ne moraju biti karakteristične za promatranu varijablu
- ovisi o broju opažanja (veći broj opažanja => veći raspon)

IST - P2

49

KVANTILE

- mjere varijabilnosti po položaju
- kvartile, decile, centile
- donja kvartila (Q_1 ili 25%)
 - vrijednost podatka koji stoji na centralnom mjestu polovice podataka nižih od medijana
- gornja kvartila (Q_3 ili 75%)
 - vrijednost podatka koji stoji na centralnom mjestu polovice podataka viših od medijana
- Q_2 - medijan

IST - P2

50

centila	obuhvat jedinica promatranja	decila	obuhvat jedinica promatranja
prva	1%	prva	10%
druga	2%	druga	20%
treća	3%	treća	30%
....		

IST - P2

51

VARIJANCA

- prosječno kvadratno odstupanje od aritmetičke sredine

$$\sigma^2 = \frac{\sum_{i=1}^N (x_i - \mu)^2}{N}$$

VARIJANCA
POPULACIJE

$$s^2 = \frac{\sum_{i=1}^N (x_i - \bar{x})^2}{N-1}$$

VARIJANCA
UZORKA

IST - P2

52

STANDARDNA DEVIJACIJA ZA POPULACIJU ZA UZORAK

$$\sigma = \sqrt{\frac{\sum_{i=1}^N (x_i - \mu)^2}{N}}$$

$$s = \sqrt{\frac{\sum_{i=1}^N (x_i - \bar{x})^2}{N-1}}$$

- služi za ocjenu pojedinih rezultata oko aritmetičke sredine
- izražava se uz aritmetičku sredinu
- obično je

$$2s < \text{raspon} < 6s$$

IST - P2

53

KOEFICIJENT VARIJABILNOSTI ZA POPULACIJU ZA UZORAK

$$K.V. = \frac{\sigma}{\mu} \cdot 100$$

$$K.V. = \frac{s}{\bar{x}} \cdot 100$$

- relativna standardna devijacija
- govori o HOMOGENOSTI promatranog obilježja
- koristan je ako želimo znati:
 - razlike u varijabilnosti svojstava neke grupe ispitanika
 - razlike u varijabilnosti istog svojstva u različitim grupama ispitanika

IST - P2

54

MJERE ZA OCJENU OBLIKA RAZDIOBE

IST - P2

55

MOMENTI RAZDIOBE

- uzastopne mjere prosječnih odstupanja od aritmetičke sredine nultog, prvog, drugog, trećeg i višeg reda

MOMENT n-tog REDA
(n-ti moment)

$$\mu_n = \frac{\sum_{i=1}^N (x_i - \mu)^n}{N}$$

IST - P2

56

NULTI I PRVI MOMENT

$$\mu_0 = \frac{\sum_{i=1}^N (x_i - \mu)^0}{N} = \frac{\sum_{i=1}^N 1}{N} = \frac{N}{N} = 1$$

prvo svojstvo aritmetičke sredine

$$\mu_1 = \frac{\sum_{i=1}^N (x_i - \mu)^1}{N} = \frac{0}{N} = 0$$

IST - P2

57

DRUGI MOMENT

$$\mu_2 = \frac{\sum_{i=1}^N (x_i - \mu)^2}{N} = \sigma^2 \quad \text{VARIJANCA}$$

IST - P2

58

TREĆI MOMENT

$$\mu_3 = \frac{\sum_{i=1}^N (x_i - \mu)^3}{N}$$

- za simetrične raspodjele $\mu_3 = 0$

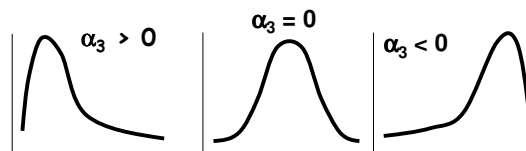
IST - P2

59

KOEFICIJENT ASIMETRIJE (coefficient of skewness)

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

- $\alpha_3 > 0$ asimetrija udesno (pozitivna asimetrija)
- $\alpha_3 < 0$ asimetrija ulijevo (negativna asimetrija)



IST - P2

60

ČETVRTI MOMENT

$$\mu_4 = \frac{\sum_{i=1}^N (x_i - \mu)^4}{N}$$

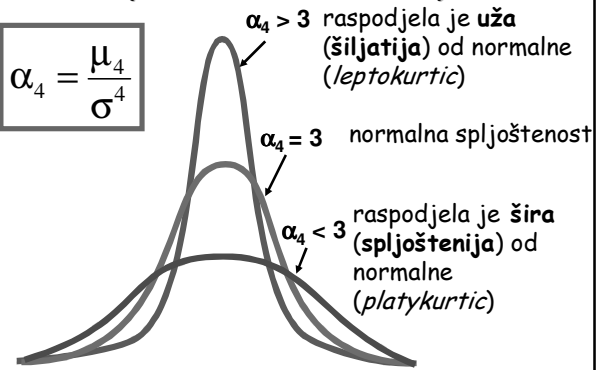
- koristi se za mjeru spljoštenosti

IST - P2

61

KOEFICIJENT SPLJOŠTENOSTI (coefficient of kurtosis)

$$\alpha_4 = \frac{\mu_4}{\sigma^4}$$



IST - P2

62

KOEFICIJENT SPLJOŠTENOSTI (coefficient of kurtosis)

- statistički programi koeficijent spljoštenosti prikazuju kao **eksces spljoštenosti** (kurtosis excess)

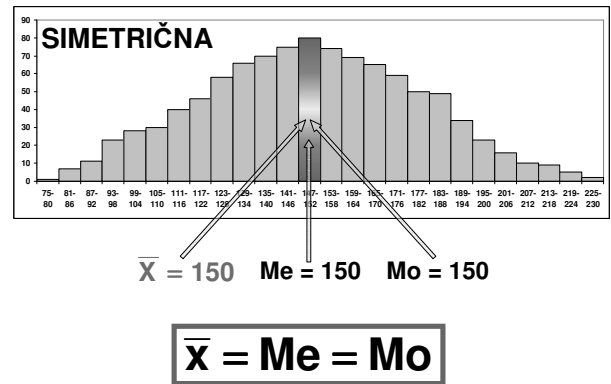
$$\text{eksces spljoštenosti} = \frac{\mu_4}{\sigma^4} - 3$$

- za normalnu raspodjelu eksces spljoštenosti = 0

IST - P2

63

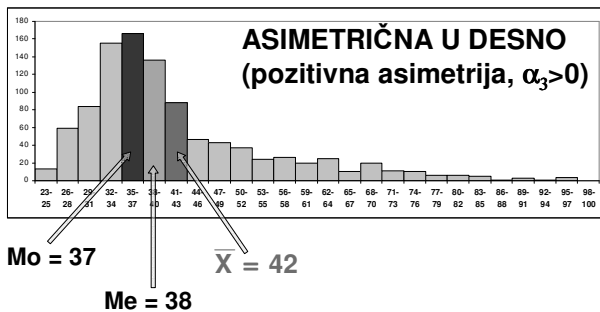
ODNOS MJERA SREDINE I ASIMETRIJE



IST - P2

64

ODNOS MJERA SREDINE I ASIMETRIJE

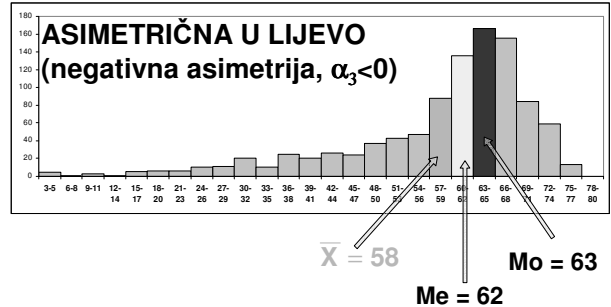


$$Mo < Me < \bar{x}$$

IST - P2

65

ODNOS MJERA SREDINE I ASIMETRIJE



$$\bar{x} < Me < Mo$$

IST - P2

66

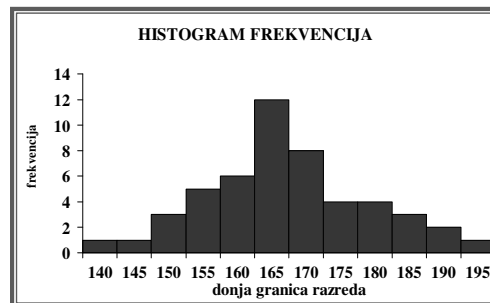
GRAFIČKO PRIKAZIVANJE PODATAKA

IST - P2

67

RAZDIOBA OBILJEŽJA

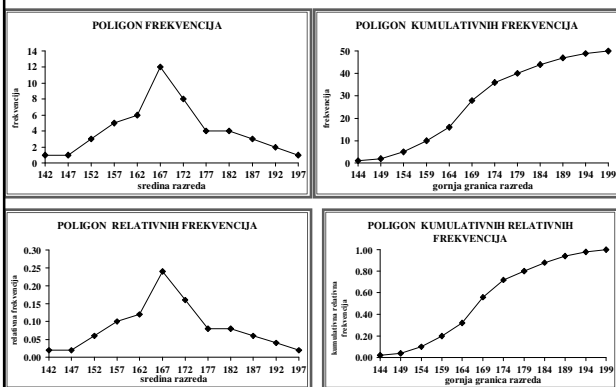
- prikazuje se histogramom (stupičasti grafikon, "column chart") ili poligonom frekvencija (linijski grafikon, "line chart")



IST - P2

68

RAZDIOBA OBILJEŽJA

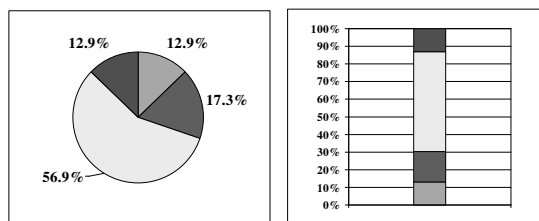


IST - P2

69

STRUKTURA OBILJEŽJA

- pokazuje udio pojedinih kategorija u ukupnom broju promatranja/mjerenja
- prikazuje se kružnim grafikonom ("torta", "pie chart") ili složenim stupičastim grafikonom ("100% stacked column chart")

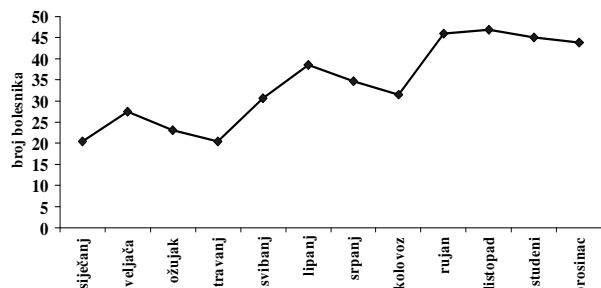


IST - P2

70

PROMJENE U VREMENU

- prikazuju se linijskim grafikonom ("line chart")
- na apscisu se nanose vremenski intervali, a na ordinatu vrijednost promatrane varijable



IST - P2

71

OSNOVNE MJERE SREDINE I RASPRŠENJA VARIJABLE

- "kutija i brkovi" grafikon ("Box-and-Whisker" plot)
- najčešće prikazuje kombinacije:

sredina	kutija	brkovi
aritmetička sredina	standardna devijacija	1.5 SD ili raspon
medijan	25% - 75%	5% - 95% ili raspon

IST - P2

72

