

INICIJALNI STATISTIČKI POSTUPCI

Provjera ispravnosti za vrijeme unosa podataka

- ⇒ Podatke unositi optimalnom brzinom.
- ⇒ Pri unosu numeričkih vrijednosti obvezno koristiti numerički dio tipkovnice.
- ⇒ Pri unosu nominalnih varijabli obvezno koristiti kodiranje.
- ⇒ Po mogućnosti podatke unositi u dvoje pri čemu jedan čita, a drugi unosi podatke i provjerava ispravnost na zaslonu ekrana.
- ⇒ Izabrati decimalnu oznaku u skladu s korištenim programom.
- ⇒ Koristiti alate za provjeru ispravnosti unosa podataka.

INICIJALNI STATISTIČKI POSTUPCI

Provjera ispravnosti nakon unosa podataka

- ⇒ Numeričke podatke provjeriti pregledom minimalne i maksimalne vrijednosti.
- ⇒ Nominalne varijable provjeriti pregledom tablica frekvencija.
- ⇒ Opcijama za pretraživanje podataka provjeriti ispravnost decimalnih oznaka.
- ⇒ Koristiti alate za naknadnu provjeru ispravnosti podataka.

MICROSOFT EXCEL

Provjera ispravnosti za vrijeme unosa podataka

Uvjeti za provjeru ispravnosti za vrijeme unosa podataka postavljaju se putem dijaloškog okvira *Data Validation* koji se pokreće odabirom opcije *Validation...* padajućeg izbornika *Data*.

Zadatak - U datoteci *Pejcic-318.xls* postavite uvjet da vrijednost u varijabli *ATV* ne može biti izvan raspona od 1200 - 1500!

Zadatak - U datoteci *Pejcic-318.xls* postavite uvjet da se u varijablu *Kraj* ne može upisati ništa osim *Gore* i *More*!

STATISTICA 7

Provjera ispravnosti nakon unosa podataka

Uvjeti za naknadnu provjeru ispravnosti unosa podataka postavljaju se opcijama dijaloškog okvira *Verify Data* koji se pokreće putem padajućeg izbornika *Data*. Uvjete se definira na sljedeći način: npr. *Valid if: spol="z" or spol="m"*.

Zadatak - U datoteci *Pejcic-318.sta* pronađite sve entitete čija je vrijednost u varijabli *ATV* izvan raspona od 1200 - 1500!

Zadatak - U datoteci *Pejcic-318.sta* pronađite sve entitete koji u varijabli *Kraj* imaju oznaku *Gore* i u varijabli *Spol* oznaku *Z*!

INICIJALNI STATISTIČKI POSTUPCI

Kodiranje podataka

U svrhu bržeg i točnijeg unosa podatke je moguće kodirati.

Kodiranje je pridruživanje skraćene oznake ili broja odgovarajućoj kategoriji.

Za prepoznavanje kodova potrebno je izraditi kodnu listu.

Primjer: Matrica s kodiranim podacima

	SPOL	POZ	OKI
AV	M	B	4
EM	M	B	3
KV	M	B	4
MD	M	B	3
MM	M	K	3
NM	M	K	2
NK	M	K	3
SA	M	K	3
SS	M	C	2
VM	M	C	3

INICIJALNI STATISTIČKI POSTUPCI

Primjer - nastavak: Kodna lista za tumačenje kodova

Kratko ime varijable	Dugo ime varijable	Oblici (vrijednosti) varijable	kod
SPOL	Spol	Muškarci	M
		Žene	Z
POZ	Pozicija u igri	Bek	B
		Krilo	K
		Centar	C
OKI	Ocjena kvalitete igrača	Vrlo slaba kvaliteta	1
		Slaba kvaliteta	2
		Dobra kvaliteta	3
		Vrlo dobra kvaliteta	4
		Izvrсна kvaliteta	5

MICROSOFT EXCEL

Pretraživanje i izmjena podataka

Pretraživanje i izmjena podataka vrši se putem dijaloškog okvira *Find and Replace* koji se pokreće kombinacijom tipki *Ctrl+F*, a omogućava brzu zamjenu kodova originalnim vrijednostima.

Zadatak - U varijabli *Spol* datoteke *Pejcic-318.xls* zamijenite sve kodove *Z* s originalnom oznakom *Žensko* te kodove *M* s *Muško*!

STATISTICA 7

Kodiranje podataka

Definiranje pripadajućih kodova za originalne vrijednosti vrši se odabirom opcije *Text Labels...* u dijaloškom okviru za formatiranje varijabli. Ista opcija omogućava trenutnu zamjenu prikazanih kodova originalnim vrijednostima.

Zadatak - U varijabli *Razred* datoteke *Pejcic-318.sta* zamijenite sve kodove *I* s originalnom vrijednošću *prvi*, *II* s *drugi*, *III* s *treći* i *IV* s *četvrti*!

STATISTICA 7

Pretraživanje i izmjena podataka

Pretraživanje i izmjena podataka vrši se putem dijaloškog okvira *Replace* koji se pokreće kombinacijom tipki *Ctrl+H*, a omogućava brzu zamjenu kodova originalnim vrijednostima.

Zadatak - U varijabli *Spol* datoteke *Pejcic-318.sta* zamijenite sve kodove *Z* s originalnom oznakom *Žensko* te kodove *M* s *Muško*!

INICIJALNI STATISTIČKI POSTUPCI

Grupiranje kvalitativnih podataka

Grupiranje podataka je statistički postupak razvrstavanja entiteta s istim oblikom obilježja u određen broj disjunktih podskupova.

Frekvencija je broj entiteta koji imaju isti oblik obilježja, odnosno broj entiteta u određenoj grupi (klasi, kategoriji, razredu).

Relativna frekvencija je omjer između frekvencije određene kategorije i zbroja frekvencija svih kategorija (ukupan broj entiteta).

INICIJALNI STATISTIČKI POSTUPCI

Grupiranje kvalitativnih podataka

Relativna frekvencija može se izraziti u postotku ili proporciji, a izračunava se na sljedeći način:

$$p_g = \frac{f_g}{n} \qquad \%_g = \frac{f_g}{n} \cdot 100$$

gdje je

→ p_g - proporcija grupe g ($g = 1, \dots, k$)

→ f_g - frekvencija grupe g

→ $\%_g$ - postotak grupe g ($g = 1, \dots, k$)

→ n - ukupan broj entiteta

→ k - broj kategorija (grupa)

INICIJALNI STATISTIČKI POSTUPCI

Grupiranje kvalitativnih podataka

Jednodimenzionalno grupiranje je grupiranje entiteta po jednom obilježju (varijabli).

Primjer: Praktičnom dijelu ispita iz KM-a pristupilo je 40 studenata; 15 ih je položilo, a 25 nije.

USPJEH NA ISPITU	FREKVENCIJA	%
NISU POLOŽILI	25	62,5
POLOŽILI	15	37,5
UKUPNO	40	100

INICIJALNI STATISTIČKI POSTUPCI

Grupiranje kvalitativnih podataka

Višedimenzionalno grupiranje je grupiranje entiteta po više obilježja (varijabli).

Primjer: Od 26 studenata koji su pristupili praktičnom dijelu ispita iz KM-a položilo ih je 10, a od 14 studentica položilo ih je 5.

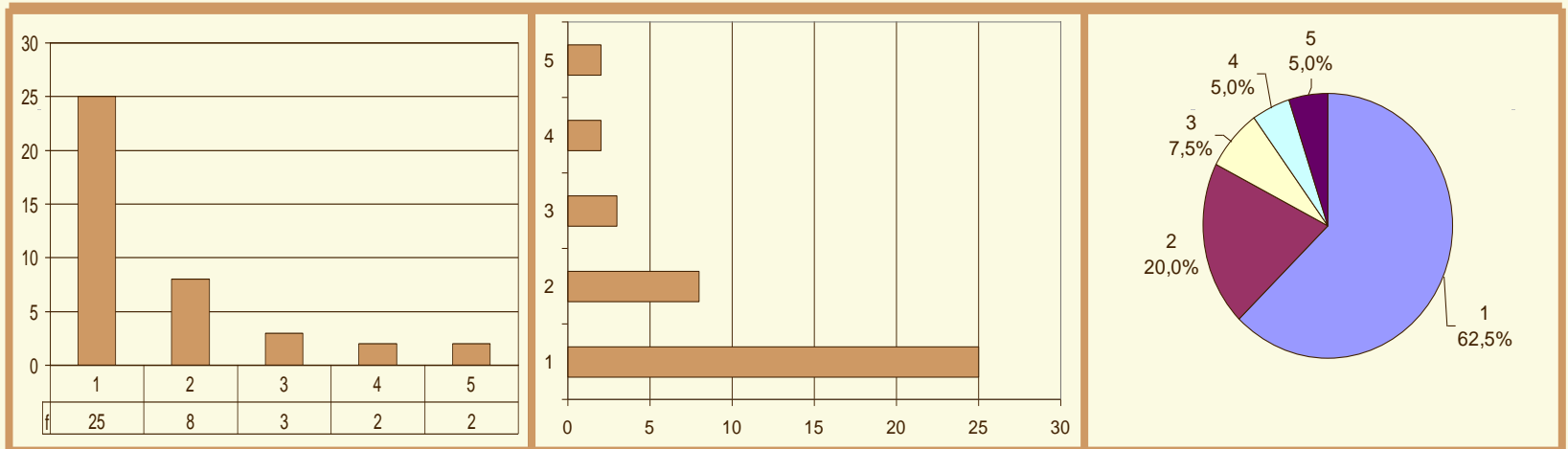
SPOL	NISU POLOŽILI	POLOŽILI	UKUPNO
MUŠKARCI	16	10	26
ŽENE	9	5	14
UKUPNO	25	15	40

INICIJALNI STATISTIČKI POSTUPCI

Grupiranje kvalitativnih podataka

Grupiranje kvalitativnih podataka najčešće se prikazuje pomoću *grafikona stupaca* (vertikalne ili horizontalne orijentacije) i *strukturnog kruga*.

Primjer: Grafikonima je prikazana raspodjela ocjena 40 studenata na praktičnom dijelu ispita iz KM-a.



MICROSOFT EXCEL

Prebrojavanje podataka

Utvrđivanje frekvencije pojedine kategorije vrši se prebrojavanjem pomoću funkcije *Countif*. Funkcija se unosi u označeno polje matrice odabirom opcije *Function...* padajućeg izbornika *Insert*.

Zadatak - U datoteci *Pejcic-318.xls* utvrdite frekvencije i relativne frekvencije (u postotku) za pojedine razrede učenika! Iskoristite funkciju *Countif*!

MICROSOFT EXCEL

Grafikon stupaca i strukturni krug

Iscrtavanje grafikona stupaca vertikalne orijentacije (*Column*), grafikona stupaca horizontalne orijentacije (*Bar*) i strukturnog kruga (*Pie*) vrši se odabirom opcije *Chart...* padajućeg izbornika *Insert*.

Zadatak - Na temelju frekvencija utvrđenih u prethodnom zadatku iscrtajte grafikone stupaca horizontalne i vertikalne orijentacije te strukturni krug!

STATISTICA 7

Jednodimenzionalno grupiranje kvalitativnih podataka

Jednodimenzionalno grupiranje kvalitativnih podataka izvodi se slijedom koraka: padajući izbornik *Statistics* → *Basic Statistics/Tables* → *Frequency Tables* → *Advanced* → *All distinct values*.

Zadatak - U datoteci *Pejcic-318.sta* utvrdite frekvencije i relativne frekvencije (u postotku) za pojedine razrede učenika!

STATISTICA 7

Grafikon stupaca i strukturni krug

Iscrtavanje grafikona stupaca izvodi se slijedom koraka: padajući izbornik *Graphs* → *2D Graphs* → *Bar/Column Plots*. Iscrtavanje strukturnog kruga izvodi se slijedom koraka: padajući izbornik *Graphs* → *2D Graphs* → *Pie Charts*.

Zadatak - Na temelju frekvencija utvrđenih u prethodnom zadatku iscrtajte grafikone stupaca horizontalne i vertikalne orijentacije te strukturni krug!

STATISTICA 7

Višedimenzionalno grupiranje kvalitativnih podataka

Višedimenzionalno grupiranje kvalitativnih podataka izvodi se slijedom koraka: padajući izbornik *Statistics* → *Basic Statistics/Tables* → *Tables and banners*. Odabirom opcije *Specify Tables (select variables)* potrebno je odabrati varijable prema kojima će se izvršiti grupiranje.

Zadatak - U datoteci *Pejcic-318.sta* grupirajte učenike istovremeno prema spolu i razredu! Utvrdite frekvencije i relativne frekvencije kategorija!

INICIJALNI STATISTIČKI POSTUPCI

Sortiranje entiteta

Sortiranje entiteta je nizanje entiteta prema redoslijedu zavisnom o rezultatu u varijabli prema kojoj se vrši sortiranje.

Entitete je moguće sortirati

- ➔ *uzlazno* - ako se podaci nižu od najmanjeg do najvećeg
- ➔ *silazno* - ako se podaci nižu od najvećeg do najmanjeg

INICIJALNI STATISTIČKI POSTUPCI

Sortiranje entiteta

Primjer: Broj osobnih pogrešaka (*OP*) 18 košarkaša (*entiteti*) na jednoj košarkaškoj utakmici - matrica podataka prije i nakon sortiranja

ENTITETI	OP
ANZU-V	4
ERJA-M	2
KRST-V	1
MILA-D	4
MILL-M	3
NORI-M	1
NOVO-K	4
SAMA-A	2
SUBO-S	3
VANJ-M	5
VOLO-D	3
VUJI-I	2
BAZD-M	3
BLAS-M	3
GIRI-G	4
KRUN-D	3
MALI-M	3
MAMI-M	2

ENTITETI	OP
KRST-V	1
NORI-M	1
MAMI-M	2
ERJA-M	2
SAMA-A	2
VUJI-I	2
MILL-M	3
SUBO-S	3
VOLO-D	3
BAZD-M	3
BLAS-M	3
KRUN-D	3
MALI-M	3
ANZU-V	4
NOVO-K	4
GIRI-G	4
MILA-D	4
VANJ-M	5

MICROSOFT EXCEL

Sortiranje

Sortiranje se vrši odabirom opcije *Sort...* padajućeg izbornika *Data* (*Ascending* - uzlazno sortiranje, *Descending* - silazno sortiranje).

Zadatak - U datoteci *Pejcic-318.xls* silazno sortirajte učenike prema rezultatu u testu *MSDM* - *Skok udalj s mjesta!*

STATISTICA 7

Sortiranje entiteta

Sortiranje se vrši odabirom opcije *Sort...* padajućeg izbornika *Data*. Opcija *Ascending* označava uzlazno sortiranje, a opcija *Descending* silazno sortiranje.

Zadatak - U datoteci *Pejcic-318.sta* silazno sortirajte učenike prema rezultatu u testu *ATT - Tjelesna masa!*

INICIJALNI STATISTIČKI POSTUPCI

Grupiranje kvantitativnih podataka

Ako diskretna varijabla ima veliki broj pojava oblika ili se radi o kontinuiranoj varijabli onda se podaci grupiraju u manji broj razreda. Za uspješno grupiranje potrebno je odrediti prikladan *broj razreda* i njihovu veličinu - *interval razreda*.

Primjer: Grupiranje 60 dječaka judaša u 5 razreda prema varijabli *Skok udalj s mjesta*

Interval razreda	f	rf (%)
$120 < x \leq 140$	1	1,67
$140 < x \leq 160$	12	20,00
$160 < x \leq 180$	26	43,33
$180 < x \leq 200$	16	26,67
$200 < x \leq 220$	5	8,33

INICIJALNI STATISTIČKI POSTUPCI

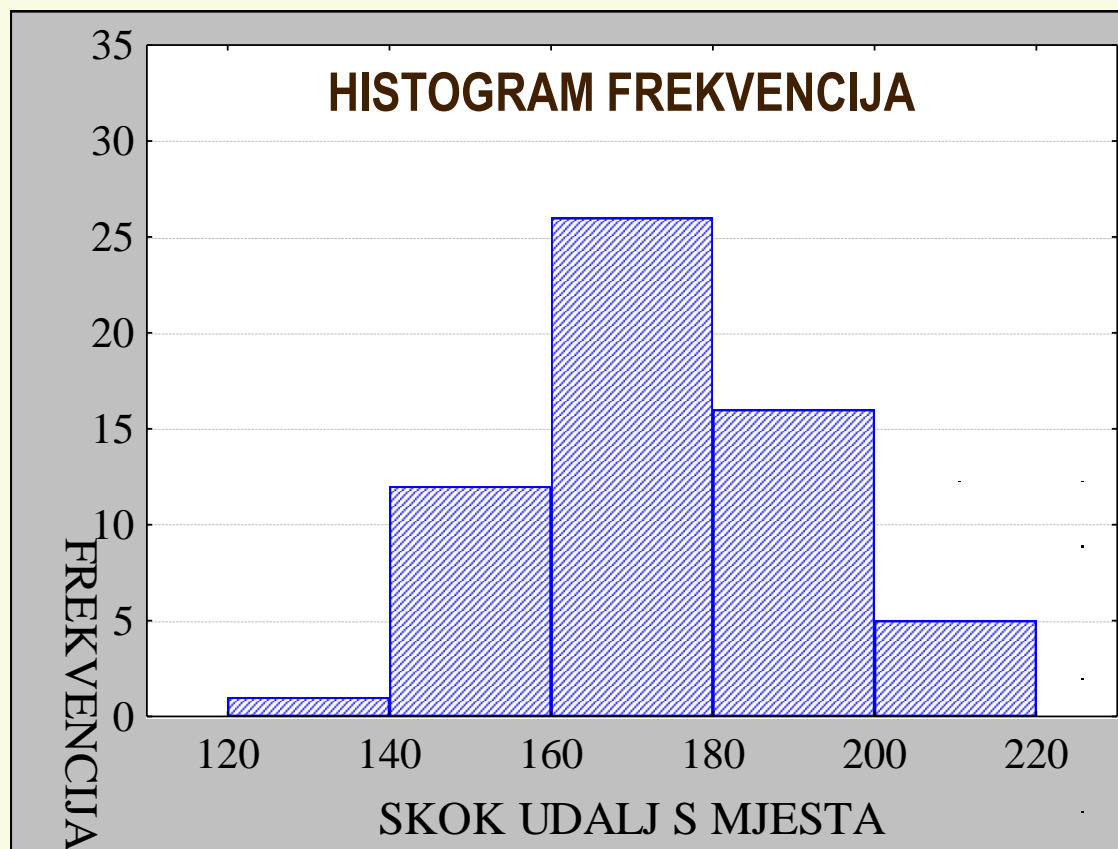
Grupiranje kvantitativnih podataka

Ako se frekvencija (apsolutna ili relativna) razreda zbroji sa sumom frekvencija svih prethodnih razreda dobije se tzv. *kumulativna frekvencija*.

Interval razreda	cf	crf (%)
$120,000 < x \leq 140,000$	1	1,67
$140,000 < x \leq 160,000$	13	21,67
$160,000 < x \leq 180,000$	39	65,00
$180,000 < x \leq 200,000$	55	91,67
$200,000 < x \leq 220,000$	60	100,00

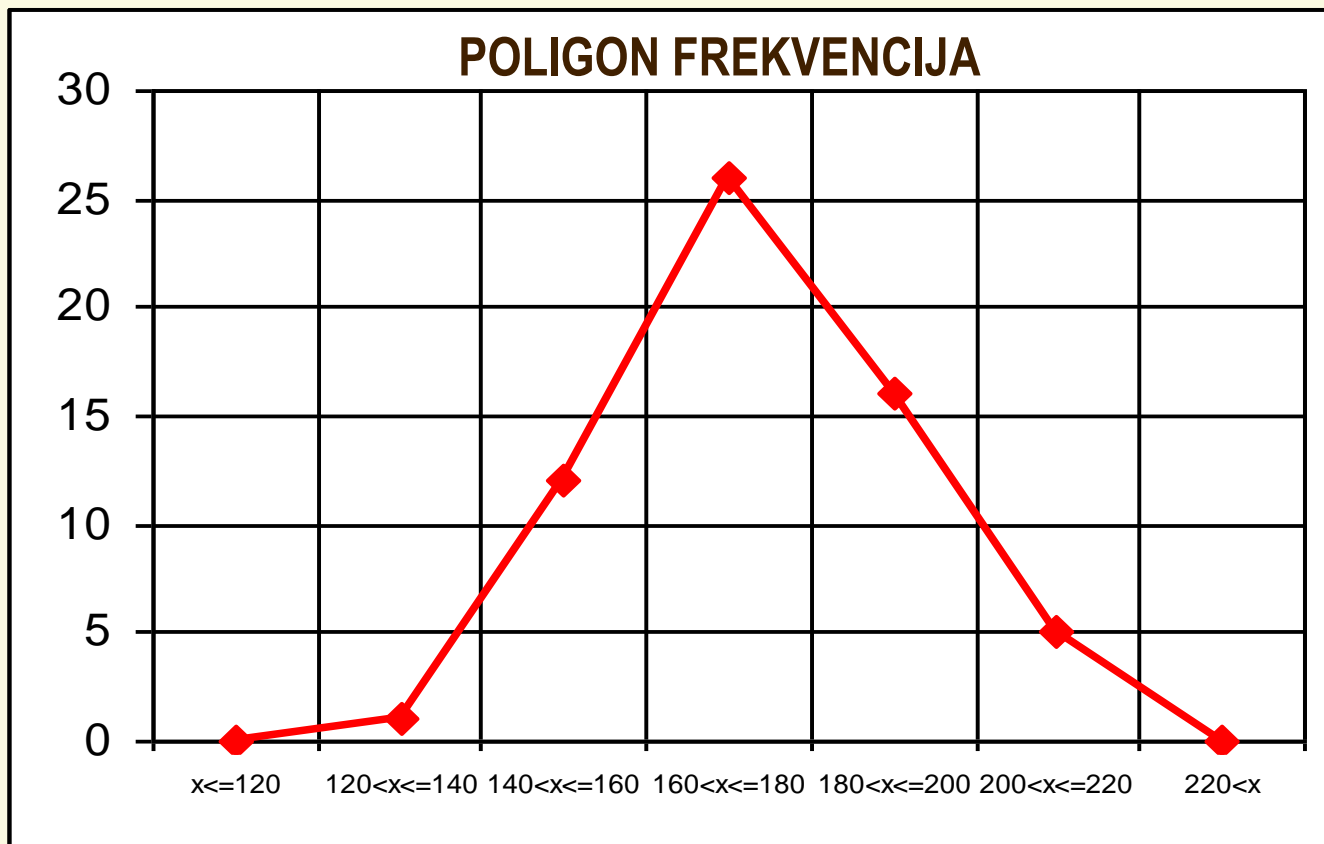
INICIJALNI STATISTIČKI POSTUPCI

Grupiranje kvantitativnih podataka



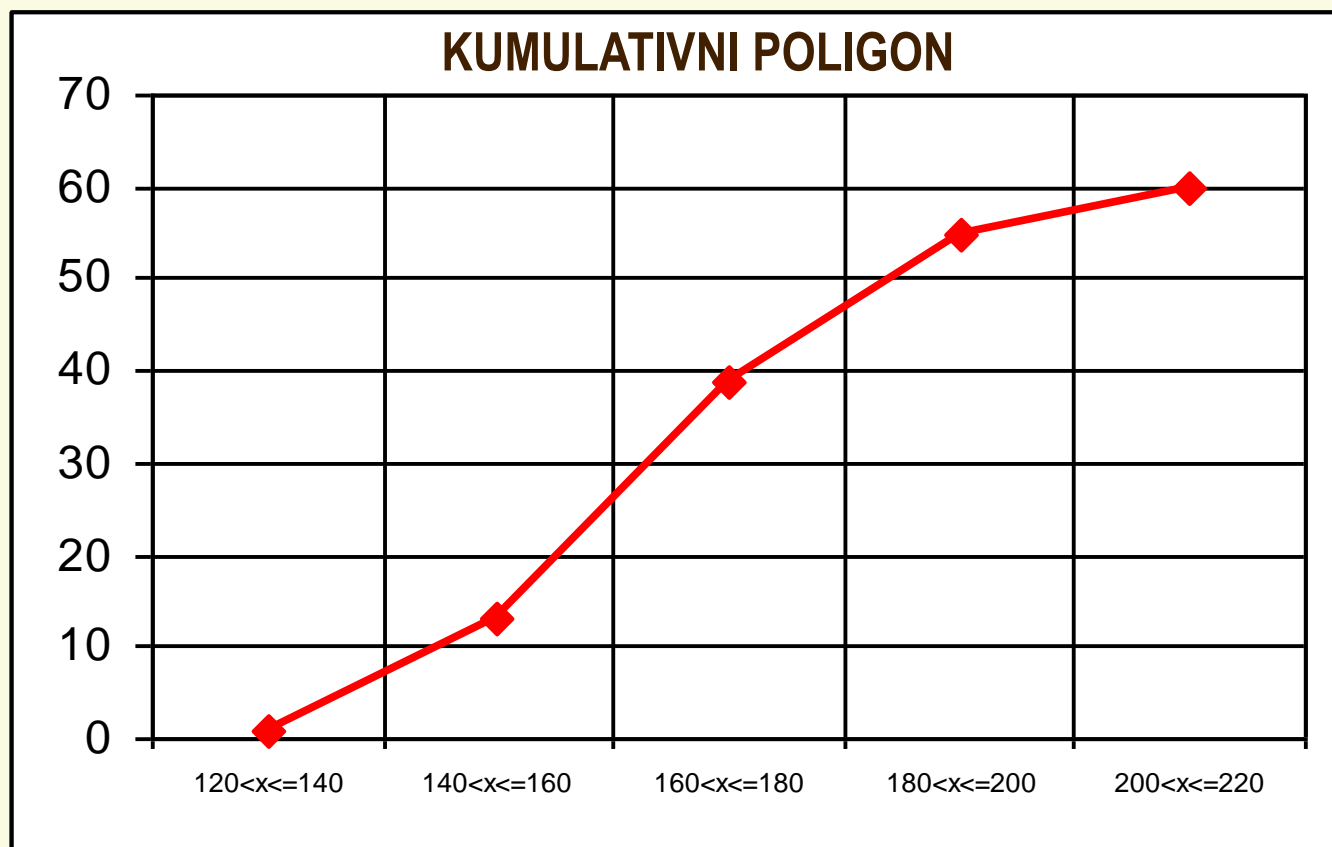
INICIJALNI STATISTIČKI POSTUPCI

Grupiranje kvantitativnih podataka



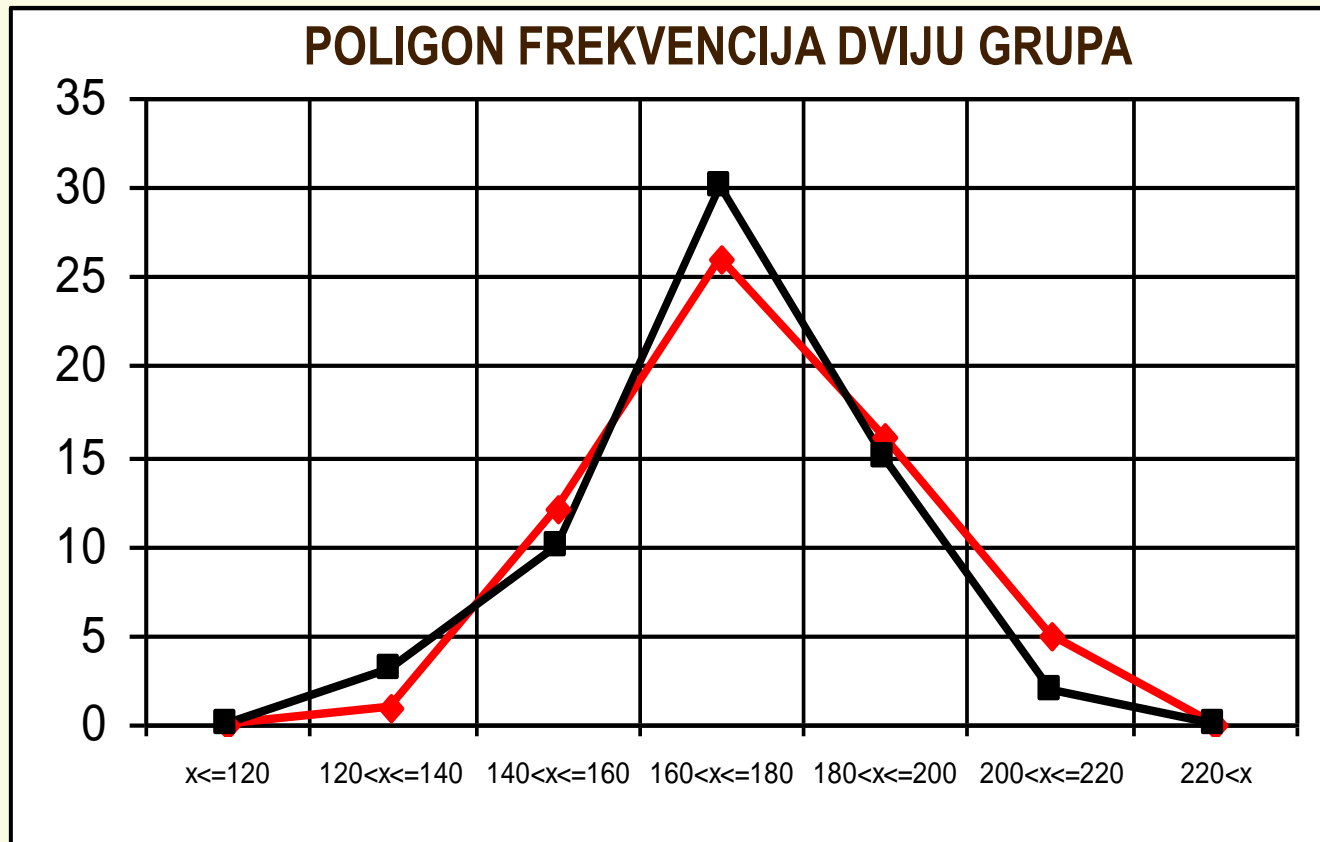
INICIJALNI STATISTIČKI POSTUPCI

Grupiranje kvantitativnih podataka



INICIJALNI STATISTIČKI POSTUPCI

Grupiranje kvantitativnih podataka



STATISTICA 7

Jednodimenzionalno grupiranje kvantitativnih podataka

Izbornik za jednodimenzionalno grupiranje kvantitativnih podataka pokreće se slijedom koraka: padajući izbornik *Statistics* → *Basic Statistics/Tables* → *Frequency Tables* → *Advanced*. Moguće je definirati: točan broj razreda (*No. of exact intervals*), približan broj razreda uz uvjet zaokruženih granica razreda ("*Neat*" *intervals, approximate no.*) i točan interval razreda (*Step size*).

Zadatak - U datoteci *Pejcic-318.sta* utvrdite apsolutne i relativne frekvencije i kumulativne frekvencije razreda prema varijabli *MSDM*! Isprobajte različite opcije!

STATISTICA 7

Histogram frekvencija i poligon frekvencija

Iscrtavanje histograma frekvencija izvodi se slijedom koraka: padajući izbornik *Graphs* → *2D Graphs* → *Histograms*.
Iscrtavanje poligona frekvencija izvodi se slijedom koraka: padajući izbornik *Graphs* → *2D Graphs* → *Line Plots (Variables)*. Pri iscrtavanju poligona frekvencija ulazni podaci moraju biti frekvencije razreda, a ne originalni rezultati entiteta.

Zadatak - Na temelju grupiranja podataka provedenog u prethodnom zadatku iscrtajte histogram frekvencija i poligon frekvencija!

MICROSOFT EXCEL

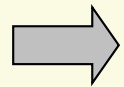
Linijski grafikon

Iscrtavanje linijskog grafikona (*Line*) vrši se odabirom opcije *Chart...* padajućeg izbornika *Insert*.

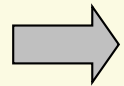
Zadatak - Na temelju frekvencija utvrđenih u prethodnom zadatku iscrtajte poligon frekvencija i poligon kumulativnih frekvencija!

INICIJALNI STATISTIČKI POSTUPCI

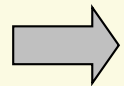
Literatura za pripremanje kolokvija



Dizdar, D. (2006). *Kvantitativne metode*. Zagreb: Kineziološki fakultet, str. 50-62.



Petz, B. (2002). *Osnovne statističke metode za nematematičare*. Jastrebarsko: Naklada Slap, str. 67-78.



Langer, M. (2004). *Brzi vizualni vodič Microsoft Excel 2003 za Windows*. Zagreb: Miš, str. 75-103, 153-181.