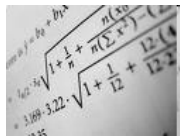


2. vaja:

**OPISNA STATISTIKA OB UPORABI MS EXCEL**



asist. Nejc Horvat, mag. farm.

---

---

---

---

---

---

---

---

**Cilji vaje**



**Načini grafičnega prikaza podatkov:**

- prikaz s stolpci, krogi, trakovi, ..
- histogram, stolpčni diagram, kvantilni diagram - boxplot, histogram s številkami, razsevni diagram

**Načini numeričnega prikaza podatkov:**

- Mere za srednje vrednosti
- Mere za razpršenost podatkov

---

---

---

---

---

---

---

---

**Osnovni pojmi**

- Populacija: skupnost statističnih enot, ki ustrezajo opredeljujočim pogojem
  - pogoji: vsebinski, krajevni, časovni
  - npr. študentje prvega letnika Fakultete za farmacijo, Ljubljana v letu 2009/2010
  - realne (stvarne) ⇔ umišljene (hipotetične)
- Vzorec: del populacije, izbran za študij določenih lastnosti te populacije
- Statistične spremenljivke: niz značilnosti, po katerih se enote v populaciji razlikujejo in jih upoštevamo pri statističnem opazovanju

---

---

---

---

---

---

---

---

## Spremenljivke v statistiki

<b>Atributivne</b> (opisne, kategorične)	<b>Numerične</b> (številске)
<b>Nominalne</b> Ne moremo razporediti po logičnem zaporedju (npr. krvne skupine; spol)	<b>Zvezne</b> Dobimo z merjenjem; niso omejene na cela števila (npr. telesna višina)
<b>Ordinalne</b> Lahko razporedimo v logično zaporedje (npr. majhna, srednja, velika bolečina; učni uspeh)	<b>Nezvezne</b> Dobimo s štetjem; omejene na cela števila (npr. število bolnišničnih dni)

**Dihotomne** (dva možna izida, npr. živ-mrtev)

**Politomne** (več možnih izidov, npr. vzroki nezgod)

---

---

---

---

---

---

---

---

## Prikaz nominalne spremenljivke

- Tabelarično ali z grafikoni
- Frekvenca...št. enot v posamezni kategoriji, skupini, razredu,...
- Relativna frekvenca...delež enot v posamezni kategoriji glede na njihovo celokupno število ( $f\% = F/N$ )

Naloga: prikažite podatke o krvni skupini preiskovancev v obliki:

- frekvenčne tabele,
- krožnega izseka,
- stolpčnega diagrama.



Podatke najdete na domači strani Katedre za socialno farmacijo:

<http://www.ffa.uni-lj.si/o-fakulteti/katedre/katedra-za-socialno-farmacijo/pedagoska-dejavnost/biomedicinska-informatika-uplb/vaje.html>

---

---

---

---

---

---

---

---

## Predstavitev MS Excel

- Možnosti uporabe MS Excel (tabele, računanje, risanje grafov)
- Delovni list sestavljen iz polj – vsako označeno s številko in črko
- Tabele: obrobe, poravnava, barva polnila, barva pisave, slog pisave, ...
- Računanje: uporaba vgrajenih in samo-napisanih funkcij, absolutni in relativni sklic

---

---

---

---

---

---

---

---

### FREKVENČNA TABELA

krvna skupina	frekvenca	relativna frekvenca
A	102	54,3%
B	36	19,1%
AB	45	23,9%
0	5	2,7%
SKUPAJ	188	100,0%

---

---

---

---

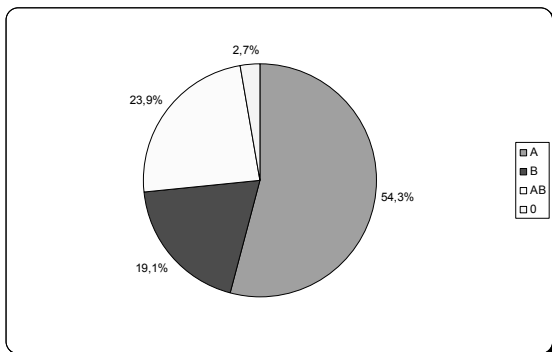
---

---

---

---

### KROŽNI IZSEK



---

---

---

---

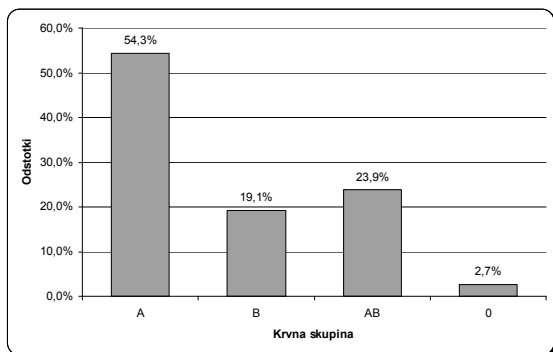
---

---

---

---

### STOLPČNI DIAGRAM



---

---

---

---

---

---

---

---

## Prikaz ordinalne spremenljivke

Kumulativna frekvenca (F): koliko enot ima vrednost pod zgornjo mejo ustreznega razreda

Relativna kumulativna frekvenca (F%): F/število enot

	f	f%	F	F%
nezadosten	3	3,9	3	3,9
zadosten	11	14,5	14	18,4
dober	38	50,0	52	68,4
prav dober	17	22,4	69	90,8
odličen	7	9,2	76	100,0
SKUPAJ	76	100,0	76	100,0

---

---

---

---

---

---

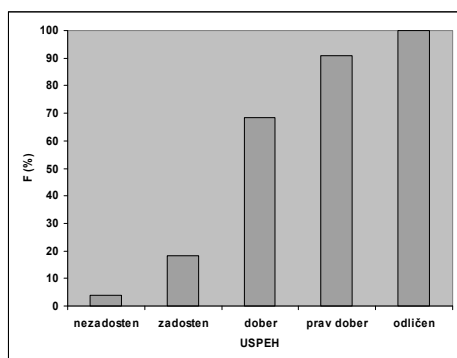
---

---

---

---

## STOLPČNI DIAGRAM KUMULATIVNIH FREKVENC



---

---

---

---

---

---

---

---

---

---

## Urejevanje numeričnih podatkov

- **Ranžirna vrsta:** podatke uredimo po naraščajoči ali padajoči vrednosti → določanje ranga (enake vrednosti, navajanje)

- **Kvantili (območje, znotraj katerega se nahaja del podatkov):**

- **Mediana:** tista vrednost spremenljivke, od katere ima polovica enot manjše, polovica pa večje vrednosti spremenljivke

- št. enot liho: mediana = vrednost srednje enote v ranžirni vrsti  $X_{(N+1)/2}$

- št. enot sodo: mediana = povprečje srednjega para podatkov  $(X_{N/2} + X_{N/2+1})/2$

- **Kvartili:** razdelijo podatke na 4 dele; vsak vsebuje 25% podatkov

- medkvartilni interval

- **Decili, centili**

---

---

---

---

---

---

---

---

---

---

## Urejevanje numeričnih podatkov

### Naloga:

Podatki o koncentraciji glukoze v krvi pri belih laboratorijskih podganah, podani v mg/100ml:  
89, 97, 102, 111, 86, 91, 110, 99, 102, 104, 98, 102, 93, 104, 97

- Uredite v ranžirno vrsto ter določite range
- Določite mediano, zgornji kvartil, spodnji kvartil, minimum, maksimum, medkvartilni interval, variacijski razpon




---

---

---

---

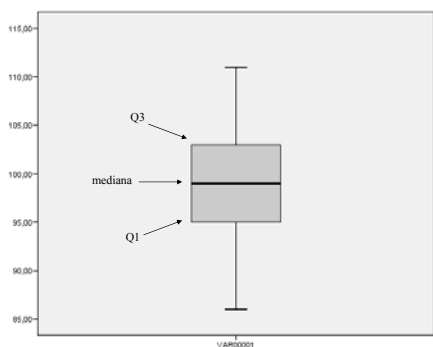
---

---

---

---

### Kvantilni diagram (box plot)




---

---

---

---

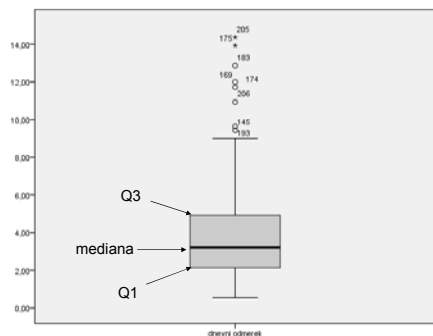
---

---

---

---

### Kvantilni diagram (box plot)




---

---

---

---

---

---

---

---

### Histogram s števkami (Stem and leaf plot)

Za razliko od navadnega histograma ohrani vse vrednosti.

86, 89, 91, 93, 97, 97,  
98, 99, 102, 102, 102,  
104, 104, 110, 111



Koncentracija glukoze	
Stem-and-Leaf Plot	
Frequency	Stem & Leaf
2,00	8 . 69
6,00	9 . 137789
5,00	10 . 22244
2,00	11 . 01
Stem width: 10,00	
Each leaf: 1 case(s)	

---

---

---

---

---

---

---

---

---

---

### Histogram s števkami (Stem and leaf plot)

DOZA/DAN Stem-and-Leaf Plot

Frequency	Stem & Leaf
5,00	0 . 57789
12,00	1 . 000122222333
25,00	1 . 5556777777777777999999999
26,00	2 . 1111111111111233333333333333
10,00	2 . 5555577777
25,00	3 . 000000000022222244444444
16,00	3 . 666668888888888888
9,00	4 . 000222222
14,00	4 . 55777777777799
3,00	5 . 113
3,00	5 . 555
9,00	6 . 000002444
2,00	6 . 88
8,00	7 . 00222222
6,00	7 . 557799
2,00	8 . 13
1,00	8 . 7
4,00	9 . 0000
8,00	Extremes (>=9.4)
Stem width: 1.0	
Each leaf: 1 case(s)	

---

---

---

---

---

---

---

---

---

---

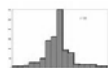
### Izdelava frekvenčne distribucije

- Kadar imamo veliko zveznih podatkov, je pregled nad njimi težko dobiti z ranžirno vrsto => oblikujemo razrede
- Določanje širine oz. števila razredov (5-15)

$$(m-1) \cdot j < (x_{\max} - x_{\min}) < m \cdot j$$

število razredov = m, širina razreda = j

- Določanje meje razredov




---

---

---

---

---

---

---

---

---

---

### Izdelava frekvenčne distribucije

#### NALOGA:

Razvrstite podatke o odmerkih varfarina:

- a) v frekvenčno distribucijo
- b) prikažite v obliki histograma (ni možno v Excel-u).

Podatke najdete na domači strani Katedre za socialno farmacijo:

<http://www.ffa.uni-lj.si/o-fakulteti/katedre/katedra-za-socialno-farmacijo/pedagoska-dejavnost/biomedicinska-informatika-uplb/vaje.html>



---

---

---

---

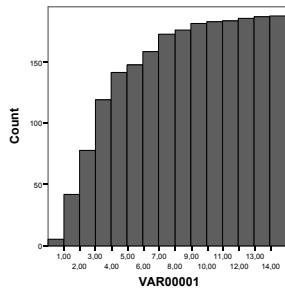
---

---

---

---

### Histogram kumulativnih frekvenc odmerkov varfarina



---

---

---

---

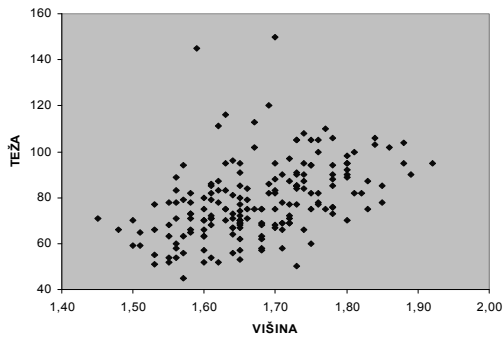
---

---

---

---

### Grafični prikaz v analizi povezanosti (razsevni diagram)



---

---

---

---

---

---

---

---

## Opisovanje vzorca/populacije

1. SREDNJE VREDNOSTI (pokažejo, okrog katerega središča se goste vrednosti spremenljivke posameznih enot => predstavitev vzorca, populacije):
- **Modus** (najpogostejša vrednost => uni-, bi-, polimodalna)
  - **Mediana** (vrednost na sredini po rangiranih vrednosti)  
- nesimetrične razporeditve
  - **Aritmetična sredina**  
- simetrične razporeditve
  - **Tehtana aritmetična sredina**  
- porazdelitve v razredih

---

---

---

---

---

---

---

---

## Opisovanje vzorca/populacije

2. RAZPRŠENOST ENOT (pove, koliko posamezni podatki odstopajo od srednje vrednosti => koliko srednja vrednost res predstavlja statistično maso):
- Variacijski razmik:  $(x_{\max} - x_{\min})$
  - Decilni razmik:  $D_9 - D_1$
  - Kvartilni razmik (medkvartilni interval):  $Q_3 - Q_1$  } nesimetrične razporeditve

---

---

---

---

---

---

---

---

## Opisovanje vzorca/populacije

- **Varianca  $\sigma^2$  in standardni odklon  $\sigma$** : osnovni meri variacije pri simetričnih razporeditvah;
  - varianca = povprečje kvadratov odklonov posameznih vrednosti od aritmetične sredine
  - standardni odklon = koren variance, pogosteje uporablja
- **Koeficient variacije**: primerjava variranja različnih spremenljivk, ki so med seboj v vsebinski zvezi (npr. obseg prsnega koša pri odraslih in dojenčkih)

---

---

---

---

---

---

---

---



## Naloga

1. Izračunajte povprečje odmerka varfarina ter njegovo standardno deviacijo.
2. Naredite ranžirno vrsto odmerka varfarina.  
Razvrstite po vrsti, pripišite rang, enake vrednosti?

---

---

---

---

---

---

---

---