

MO 34: ŠTATISTIKA

MO 34:

ŠTATISTIKA**ŠTATISTIKA**

- matematická disciplína, ktorá skúma a vyhodnocuje spoločenské javy

Základné pojmy:

- **Štatistický súbor**
 - neprázdna konečná množina M
 - (napr. žiaci triedy 1.A)
- **Rozsah štatistického súboru**
 - počet prvkov množiny M
 - označujeme $|M| = n$
 - (napr. 20 žiakov)
- **Štatistický znak**
 - funkcia, ktorá každému prvku množiny M pridelí práve 1 reálne číslo = kvantitatívny znak alebo kvalitatívny znak
 - kvantitatívny znak: vek, výška, váha
 - kvalitatívny znak: farba očí, pohlavie, národnosť
 - hodnoty znaku označujeme: x_1, x_2, \dots, x_n
- **Absolútna početnosť**
 - číslo, ktoré udáva, koľkokrát sa v súbore M vyskytuje hodnota znaku x_i
 - označujeme n_i
- **Relatívna početnosť**
 - podiel $\frac{n_i}{n}$,
 - kde n_i je absolútna početnosť hodnoty znaku x_i , n je rozsah súboru M

Grafy:

- spracovanie štatistického súboru vykonávame pomocou grafov a tabuliek
- napr. V 4.B boli namerané tieto hodnoty výšok žiakov:

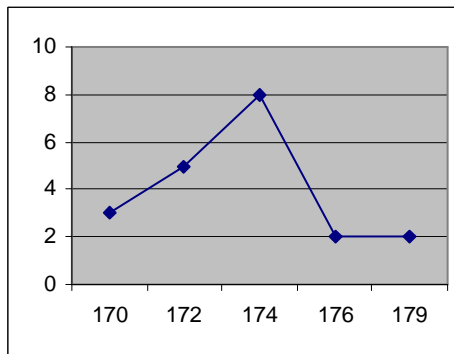
x_i	n_i	n_i/n
170	3	0,15
172	5	0,25
174	8	0,4
176	2	0,1
179	2	0,1

MO 34: ŠTATISTIKA

- **Polygón početnosti** = spojnicový diagram

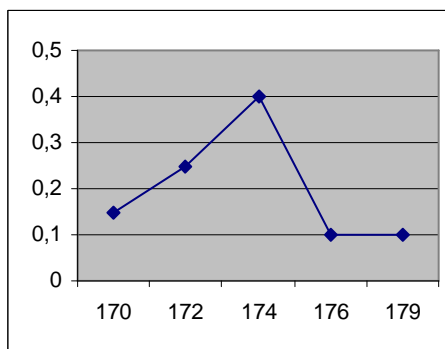
- absolútny:

OS x = x_i
OS y = n_i



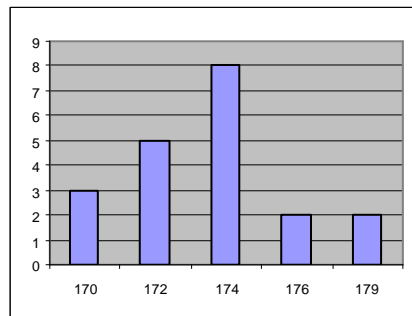
- relatívny:

OS x = x
OS y = $\frac{n_i}{n}$

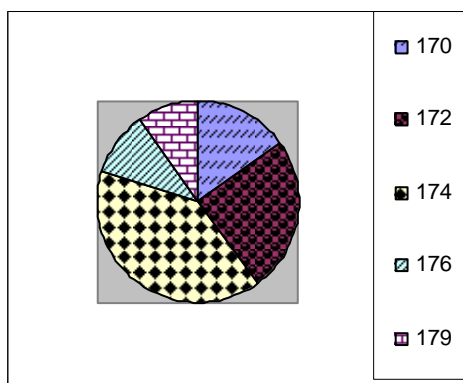


- **Histogram** = stĺpcový diagram

- existuje absolútny alebo relatívny



- **Kruhový diagram** (s legendou)



MO 34: ŠTATISTIKA**Charakteristiky:**

- štatistický súbor nám opisujú charakteristiky

- **Aritmetický priemer**

- označujeme \bar{x}
- aritmetický priemer hodnôt x_1, x_2, \dots, x_n
- $$\bar{x} = \frac{x_1 + x_2 + \dots + x_n}{n} = \frac{1}{n} \sum_{i=1}^n x_i$$
- ak majú hodnoty x_i početnosť n_i , hovoríme o váženom aritmetickom priemere
- $$\bar{x} = \frac{x_1 \cdot n_1 + x_2 \cdot n_2 + \dots + x_n \cdot n_k}{n_1 + n_2 + \dots + n_k}$$

- **Geometrický priemer**

- označujeme \bar{X}_G
- $$\bar{X}_G = \sqrt[n]{x_1 \cdot x_2 \cdot x_3 \cdot \dots \cdot x_n} = \sqrt[n]{\prod_{i=1}^n x_i}$$
- v praxi sa používa pri určovaní priemerného tempa výroby

- **Harmonický priemer**

- označujeme \bar{X}_H
- $$\bar{X}_H = \frac{n}{\frac{1}{x_1} + \frac{1}{x_2} + \dots + \frac{1}{x_n}} = \frac{n}{\sum_{i=1}^n \frac{1}{x_i}}$$

- **Medián**

- označujeme $\text{med}(x)$
- je to prostredný člen spomedzi hodnôt x_i , ak sú usporiadané podľa veľkosti
- ak je rozsahom súboru párne číslo n , $\text{med}(x)$ určíme:

$$\text{med}(x) = \frac{x_i + x_{i+1}}{2}; i = \frac{n}{2}$$

- **Modus**

- označujeme $\text{mod}(x)$
- je to najčastejšie vyskytujúca sa hodnota v súbore M

- **Smerodajná odchýľka s**

- $$s = \sqrt{\frac{1}{n} \cdot \sum (x_i - \bar{x})^2}$$

- **Rozptyl s^2**

citlivejšie charakterizujú variabilitu
hodnôt súboru, zvyrazňujú váhu odchýliek
hodnôt od \bar{x} , dávajú informáciu

MO 34: ŠTATISTIKA

- $s^2 = \frac{1}{n} \cdot \sum (x_i - \bar{x})^2$ o rozptýlení hodnôt okolo \bar{x}

- **Koeficient korelácie r**

- $r = \frac{k}{s_x \cdot s_y}$

- vyjadruje závislosť 2 znakov x, y

- $r < 0,3$ slabá korelácia (slabo súvisí)

- $0,3 \leq r < 0,8$ stredná (mierna) korelácia

- $0,8 \leq r \leq 1$ silná korelácia (veľmi úzko spolu súvisia)

- k = kovariancia

- $k = \frac{1}{n} \cdot \sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y})$

- $s_x = \sqrt{\frac{1}{n} \cdot \sum (x_i - \bar{x})^2}$

- $s_y = \sqrt{\frac{1}{n} \cdot \sum (y_i - \bar{y})^2}$