

Statisztika

"I only believe in statistics that I doctored myself"

Winston S. Churchill



Egy kisváros tűzoltó parancsnoka összesítést készített, hogy az előző hónapban naponta hány riasztáshoz vonultak ki:

A napi riasztások száma	0	1	2	4	6	14
Hány napon fordult elő?	1	11	6	8	4	1

- 93 eset 31 nap alatt, ez naponta átlagosan 3 riasztás – elemezte a táblázatot a parancsnok.
- Csak azért, mert volt egy nap, amikor szokatlanul sok dolgunk akadt! – válaszolta az egyik társa. – Valójában jobb a helyzet, hiszen a hónap több, mint a felében maximum kétszer voltunk, én inkább ezt gondolnám tipikusnak.
- Szerintem meg a legjellemzőbb esteszám az egy, hiszen ilyen napból volt a legtöbb – szólt közbe egy harmadik.

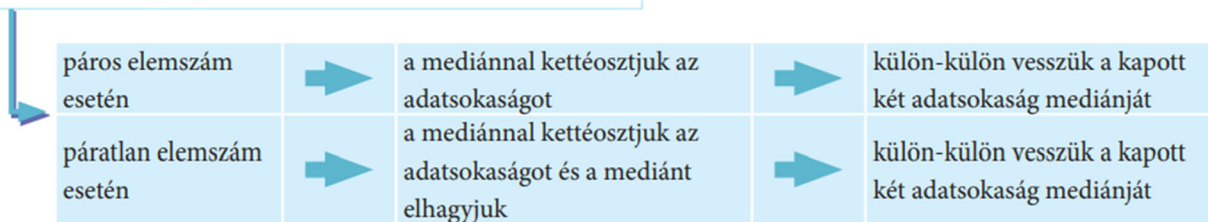
Kinek van igaza?

Statisztikai jellemzők:

Név	Definíció	Képlet
Átlag (\bar{x})	Az adatok összege osztva az adatok számával.	$A = \bar{x} = \frac{x_1 + x_2 + \dots + x_n}{n}$
Súlyozott átlag (x')	A súlyozott adatok átlaga.	$A^* = x' = \frac{s_1x_1 + s_2x_2 + \dots + s_nx_n}{s_1 + s_2 + \dots + s_n}$
Módusz (M_o)	A leggyakoribb adat, ha több ilyen is van, akkor módusz halmazról beszélünk.	-
Medián (M_e vagy Q_2)	Az a szám, aminél az adatok közelítőleg fele kisebb vagy egyenlő.	A sorba rendezett adatok közül a középső, páros elemszám esetén a középső kettő átlaga.
Alsó kvartilis (Q_1)	Az a szám, aminél az adatok közelítőleg negyede kisebb vagy egyenlő.	Lásd lent a rajzokon.
Felső kvartilis (Q_3)	Az a szám, aminél az adatok közelítőleg negyede nagyobb vagy egyenlő.	
Minimum (Min vagy Q_0)	A legkisebb adat.	-
Maximum (Max vagy Q_4)	A legnagyobb adat.	-
Terjedelem	A legnagyobb és a legkisebb adat különbsége.	Terjedelem = $Q_4 - Q_0$
Fél terjedelem	A felső kvartilis és az alsó kvartilis különbsége.	Fél terjedelem = $Q_3 - Q_1$
Szórás (σ)	Az átlagtól való eltérést mutatja meg.	$\sigma = \sqrt{\frac{(x_1 - \bar{x})^2 + (x_2 - \bar{x})^2 + \dots + (x_n - \bar{x})^2}{n}}$

Hogyan számolunk kvartiliseket?

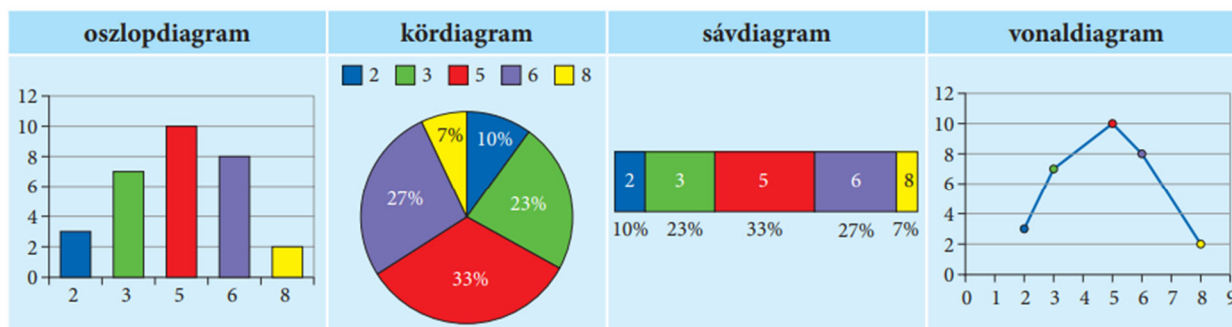
Rendezzük nagyság szerint növekvő sorrendbe az adatokat!



Például:

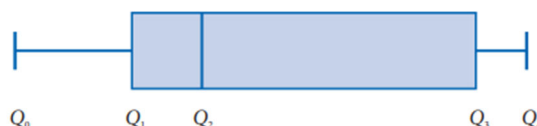
	adatok	alsó kvartilis	medián	felső kvartilis
8 adat	1 2 3 3 4 5 6 6	2,5	3,5	5,5
9 adat	1 2 3 3 4 5 5 6 6 6	2,5	4	5,5
10 adat	1 2 3 3 4 5 5 6 6 7	3	4,5	6
11 adat	1 2 3 3 4 5 6 6 7 8	3	5	6

Diagramtípusok



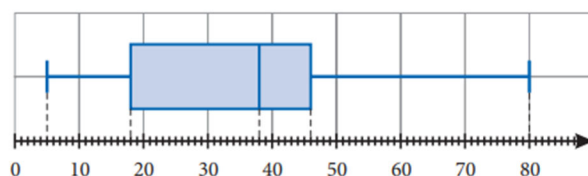
A dobozdiagram

A kvartilisek segítségével készül a **dobozdiagram** (más néven **box-plot diagram** vagy **sodrófadiagram**).



A doboz két széle az alsó, illetve a felső kvartilisnél, azaz Q_1 -nél és Q_3 -nál jelenik meg. A dobozon belül egy vonal jelzi a medián, azaz Q_2 helyét. A dobozból kiinduló vonalak a minimumig (Q_0), illetve a maximumig (Q_4) tartanak.

Például



minimum	alsó kvartilis	medián	felső kvartilis	maximum
5	18	38	46	80

terjedelem	félterjedelem
$80 - 5 = 75$	$46 - 18 = 28$