

Statistică descriptivă - instrucțiuni:

Calcularea indicatorilor:

Talia sau volumul eșantionului (notat cu N)	Funcție	COUNT
Indicatori de centralitate:		
media aritmetică (\bar{X})	Funcție	AVERAGE
Mediana	Funcție	MEDIAN
Modulul	Funcție	MODE
media geometrică	Funcție	GEOMEAN
valoarea centrală	Formulă	(maxim+minim)/2
Indicatori de dispersie:		
Amplitudinea	Formulă	maxim-minim
media deviației	Funcție	AVEDEV
Varianța	Funcție	VAR
abaterea standard (notată S)	Funcție	STDEV
coeficientul de variație	Formulă	$\frac{S}{\bar{X}}$
Intervalele	Formulă	$\bar{X} - S$
	Formulă	$\bar{X} + S$
	Formulă	$\bar{X} - 2 \cdot S$
	Formulă	$\bar{X} + 2 \cdot S$
	Formulă	$\bar{X} - 3 \cdot S$
	Formulă	$\bar{X} + 3 \cdot S$
eroarea standard	Formulă	$\frac{S}{\sqrt{N}}$
Indicatori de localizare:		
minim	Funcție	QUARTILE sau MIN
prima cvartilă	Funcție	QUARTILE
a doua cvartilă (mediana)	Funcție	QUARTILE sau MEDIAN
a treia cvartilă	Funcție	QUARTILE
maxim	Funcție	QUARTILE sau MAX
percentilele (ex. Percentila de 99%)	Funcție	PERCENTILE
Indicatori de simetrie:		
Asimetria	Funcție	SKEW
Boltirea	Funcție	KURT

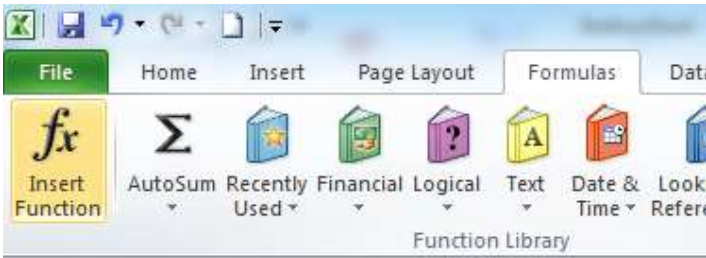
Inserarea unei funcții cu un argument sau cu argumente de același tip (ex. Count):

Selectați celula în care doriți să aveți rezultatul

K	L	M	N	O	
	Denumirea	TAS	TAD	Vârsta	Ta
	talie eșantionului				
	media aritmetică				
	mediana				

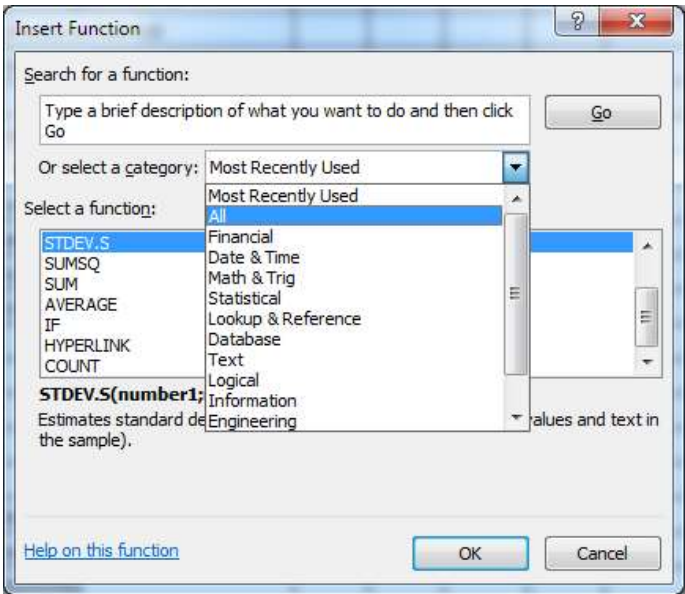
Meniul **Formulas**

Butonul **Insert Function**



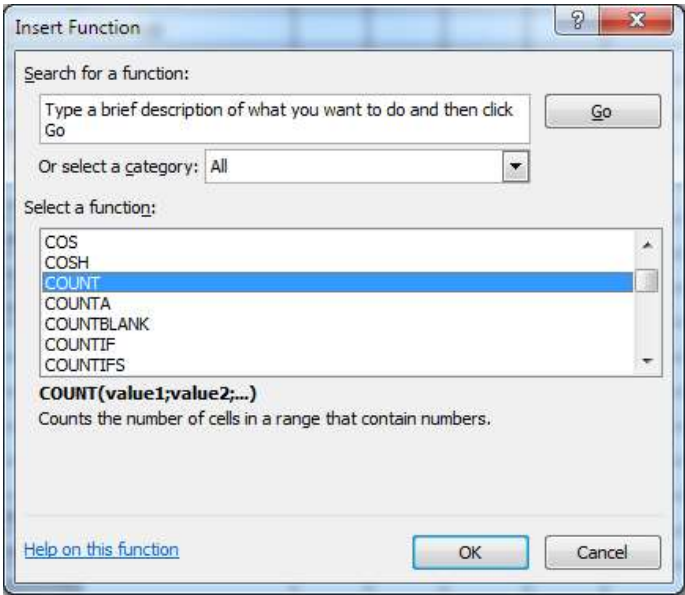
Din lista **Or select a category**

Alegeți **All**

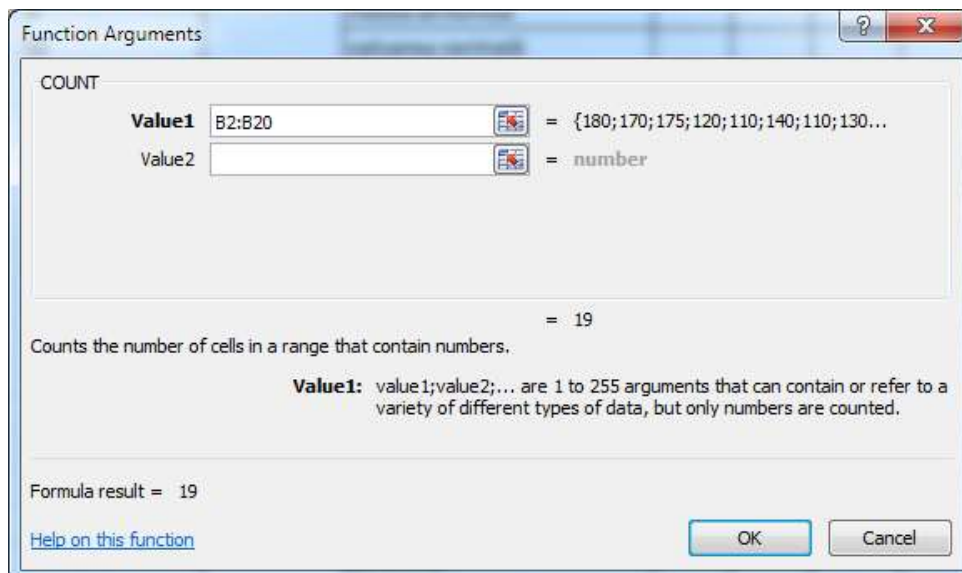


Derulați lista de funcții până la funcția **Count**

Apăsați **OK**



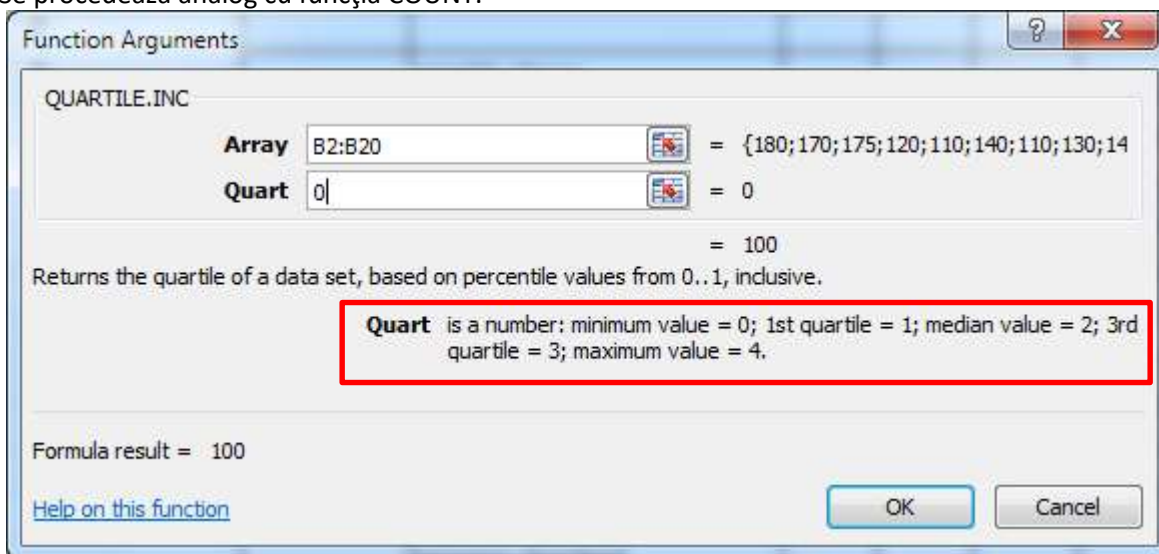
Selectați variabila TAS
sau scieți domeniul
B2:B20



! Observație: Analog se calculează AVERAGE, MEDIAN, MODE, GEOMEAN, HARMEAN, AVEDEV, VAR, STDEV, SKEW, KURT, MIN și MAX.

Pentru următoarele variabile calculăm indicatorul cu **Fill Right**.

Inserarea unei funcții cu mai multe argumente (ex. Quartile.inc):
Se procedează analog cu funcția COUNT.



Observație: Pentru calculul celorlalte cvartile pentru argumentul Quart introduceți valori întregi de la 0 la 4.

Funcția Percentile se va calcula analog cu funcția Quartile. Al doilea argument ia valori de la 0 la 1. Pentru exemplul nostru argumentul va fi 0,99. Calculul percentilelor are sens numai dacă în studiu sunt incluși mai mult de 100 de subiecți.

Interpretarea statistică a rezultatelor:

- **Media aritmetică.** Dacă variabila este normal distribuită, atunci media ne indică mijlocul intervalului dintre minim și maxim (intervalul de distribuție a datelor). Tot în cazul distribuției normale, în jurul mediei (mai precis în intervalul medie-abaterea standard, medie+abaterea standard se găsesc majoritatea datelor).
- **Eroarea standard.** Eroarea standard este implicată în calcularea intervalului de confidență de 95% în jurul mediei (numai pentru o variabilă cu distribuție normală), de asemenea este implicată în inferența statistică.
- **Mediana** este o valoare a seriei astfel încât jumătate dintre observații au valori mai mici (sau egale) și cealaltă jumătate au valori mai mari (sau egale). În cazul distribuției normale media și mediana sunt egale. Astfel că mediana și media aritmetică devin indicatori pentru distribuția normală, cu cât au valorile mai apropiate cu atât este mai probabil că variabila este distribuită normal. Termenul “mai apropiate” se estimează în funcție de mărimea erorii standard.
- **Modulul** este valoarea care are cea mai mare frecvență din serie. În cazul modulului o situație care apare este cea în care seria nu are modul, adică toate valorile apar o singură dată. Atunci va fi afișată valoarea #N/A. O altă situație posibilă este ca seria să fie bimodală sau trimodală. Atunci va fi afișată numai prima valoare în ordinea apariției lor în cadrul seriei. În acest caz pentru determinarea tuturor valorilor modulului se poate face un tabel de frecvență. Modulul este util în cazul unei variabile calitative ordonate, dar și în cazul altor tipuri de variabile, de exemplu în cazul variabilei continue cu distribuție normală modulul este probabil să aibă o valoare apropiată de medie.
- **Deviația standard (Standard Deviation) sau Abaterea standard** se poate calcula și cu STDEV sau pentru deviația standard populațională STDEV.P. Abaterea standard ne arată care este abaterea pătratică medie de la media aritmetică a valorilor variabilei. Dacă are o valoare mică, atunci datele variază puțin în jurul mediei. În cazul în care distribuția este reprezentată prin curba lui Gauss (distribuția este normală) au loc următoarele repartizări ale datelor (conform teoremei lui Cebășev):
 - intervalul $\bar{X} \pm S$ conține circa 68.3 % din observații
 - intervalul $\bar{X} \pm 2 \cdot S$ conține circa 95.5 % din observații
 - intervalul $\bar{X} \pm 3 \cdot S$ conține circa 99.7 % din observații.
- **Varianța** se poate calcula și cu VAR sau pentru variația populațională VAR.P
- **Excesul sau Boltirea** măsoară înălțimea aplatizării sau boltirii unei distribuții în comparație cu o distribuție normală. Excesul este zero pentru o serie de date având o distribuție normală, este pozitiv pentru o serie de date având trena mai înaltă decât cea a unei distribuții normale și este negativ pentru o serie de date a cărei trenă este mai coborâtă decât cea a unei distribuții normale.
Pentru interpretarea boltirii veți folosi următoarele criterii:
 - Dacă **boltirea (kurt) ≈ 0** , distribuția este **mesokurtică**
 - Dacă **boltirea (kurt) > 1** , distribuția este **platicurtică**
 - Dacă **boltirea (kurt) < -1** , distribuția este **leptocurtică**

- **Asimetria** măsoară abaterea de la aspectul simetric și direcția asimetriei (pozitivă sau negativă) față de curba normală. Asimetria este 0 pentru o serie de date având o distribuție normală, este negativă pentru o serie de date asimetrică spre dreapta (seria are mai multe valori mai mici), este pozitivă pentru o serie de date asimetrică spre stânga (seria are mai multe valori mai mari).

Pentru interpretarea asimetriei veți folosi următoarele criterii:

Dacă **asimetria (skew) ≈ 0** , distribuția este aproximativ **simetrică**

Dacă **asimetria (skew) > 1** , distribuția este **asimetrică la stânga**

Dacă **asimetria (skew) < -1** , distribuția este **asimetrică la dreapta**

- **Cvartilele și percentilele** sunt asemănătoare medianei. Astfel, prima cvartilă este o valoare având proprietatea ca 25% dintre datele seriei sunt mai mici sau egale cu ea, iar 75% mai mari sau egale cu prima cvartilă. A doua cvartilă este reprezentată de mediană. A treia cvartilă este o valoare având proprietatea ca 75% dintre datele seriei sunt mai mici sau egale cu ea iar 25% mai mari sau egale cu a treia cvartilă. Percentila de ordinul a este o valoare cu proprietatea că o proporție egală cu a din date sunt mai mici sau egale, iar celelalte sunt mai mari. Cvartilele și percentilele se pot calcula și cu funcția Quartile respectiv Percentile.

Condiții care trebuiesc îndeplinite pentru ca o serie să aibă **distribuție normală**:

- Media aritmetică \approx mediana \approx modul
- Asimetria și boltirea să fie 0 sau cât mai aproape de 0 (se accepta valorile între $[-1; 1]$)

Regulile empirice pentru interpretarea statistică a coeficientului de variație (CV):

- dacă CV este sub 10% atunci populația poate fi considerată omogenă;
- dacă CV este între 10%-20% atunci populația poate fi considerată relativ omogenă;
- dacă CV este între 20%-30% atunci populația poate fi considerată relativ eterogenă;
- dacă CV este peste 30% atunci populația poate fi considerată eterogenă.