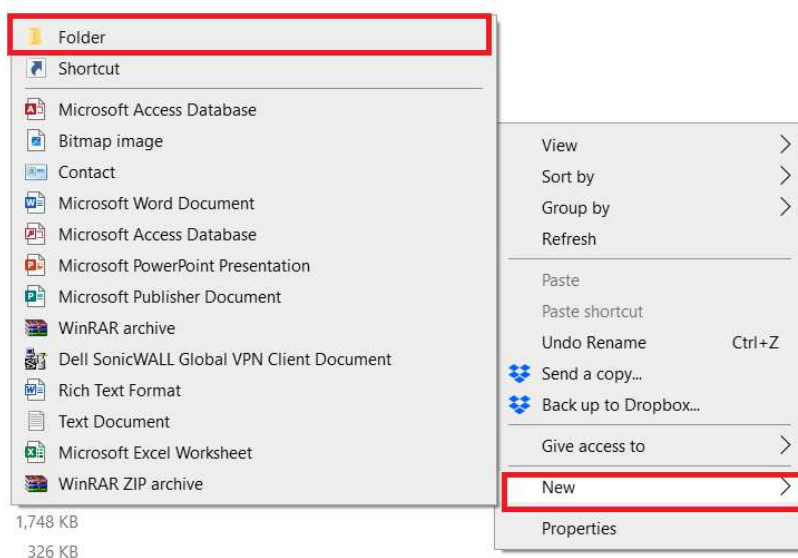


De reținut:

- Denumirile fișierelor și folderelor trebuie să fie în conformitate cu cerințele.

Crearea unui nou folder pe Desktop

- Poziționați-vă pe Desktop
- Creați un nou folder [click dreapta de mouse și utilizați opțiunea Folder-> New].

**Salvarea unui fișier Excel anterior creat într-un anumit folder**

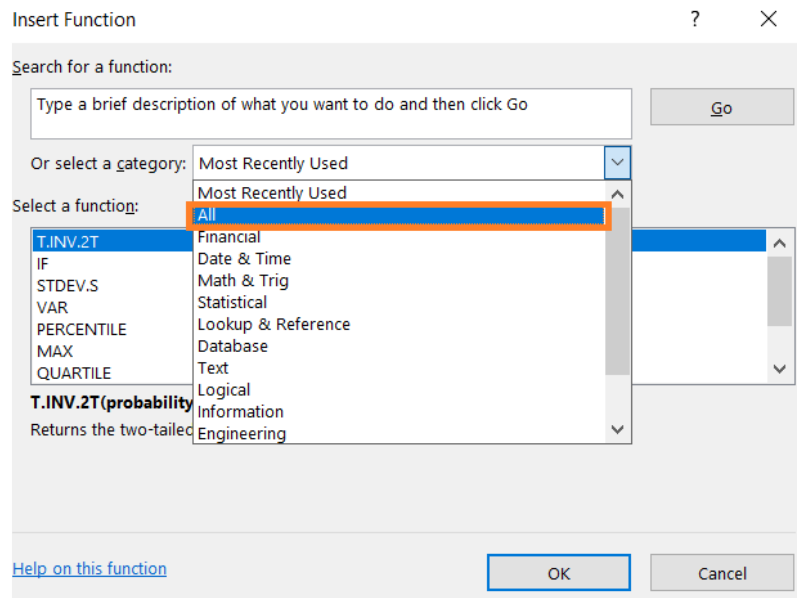
- Descărcați fișierul **BD_LP06_AMG** de pe site (el se va salva automat în folderul Downloads)
- Deschideți fișierul **BD_LP06_AMG** și apoi resalvați fișierul în folderul specificat în exercițiu (Lab04) utilizând opțiunea: File- Save As-Desktop-folder **Lab06**

Calculul statisticilor descriptive (medie, mediană, modul, deviație standard) prin funcții predefinite

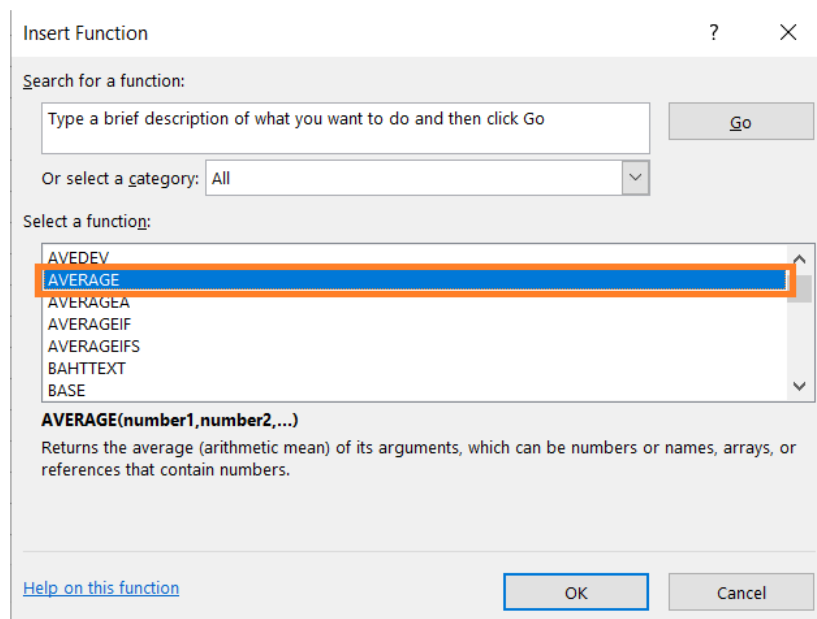
- la dreapta bazei de date, în aceeași foaie de calcul (de ex. in coloanele M si N), aveți dat un tabel de forma:

M	N
	Vârsta (ani)
Medie	
Mediana	
Modul	
Deviație standard	
Cuartila 1 (Q1)	
Cuartila 3 (Q3)	

- Poziționați-vă în acest tabel, pe celula unde doriți să obțineți rezultatul mediei și folosiți opțiunea **Formulas->Insert Functions**



- Se selectează funcția dorită (de exemplu, pentru media aritmetică se va alege funcția AVERAGE)



- În fereastra funcției AVERAGE pentru argumentul Number1 se vor selecta celulele cu valorile vârstei pacienților (de ex. A2: A101).

Function Arguments

AVERAGE

Number1 = {38;49;44;52;49;54;51;39;41;49;48;48;55;56;55;48;...}

Number2 = number

= 50.26

Returns the average (arithmetic mean) of its arguments, which can be numbers or names, arrays, or references that contain numbers.

Number1: number1,number2,... are 1 to 255 numeric arguments for which you want the average.

Formula result = 50.26

[Help on this function](#) OK Cancel

M	N
	Vârsta (ani)
Medie	50.26
Mediana	
Modul	
Deviație standard	
Cuartila 1 (Q1)	
Cuartila 3 (Q3)	

- Pentru calculul mediei se va folosi funcția **AVERAGE**
- Pentru calculul mediei se va folosi funcția **AVERAGE**
- Pentru calculul mediei se va folosi funcția **AVERAGE**

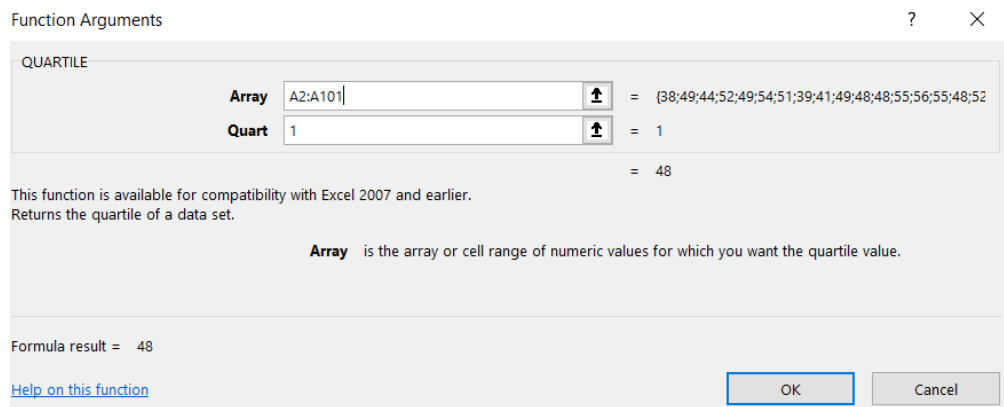
Calcul cuartile (Q1 si Q3) utilizând funcția QUARTILE

- Pozitionare pe celula goala de langa cuartila 1->Se cauta din Formulas->Insert Functions-> QUARTILE

Sintaxa functiei se va completa astfel:

Array: valorile vârstei pentru pe tot esantionul

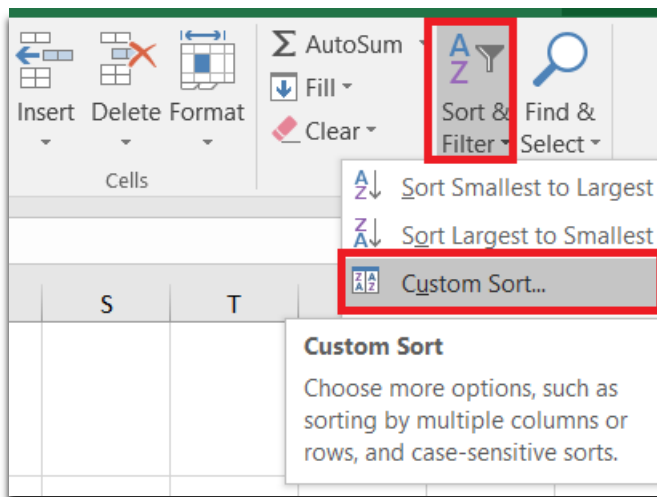
Quart: ordinul cvartilei



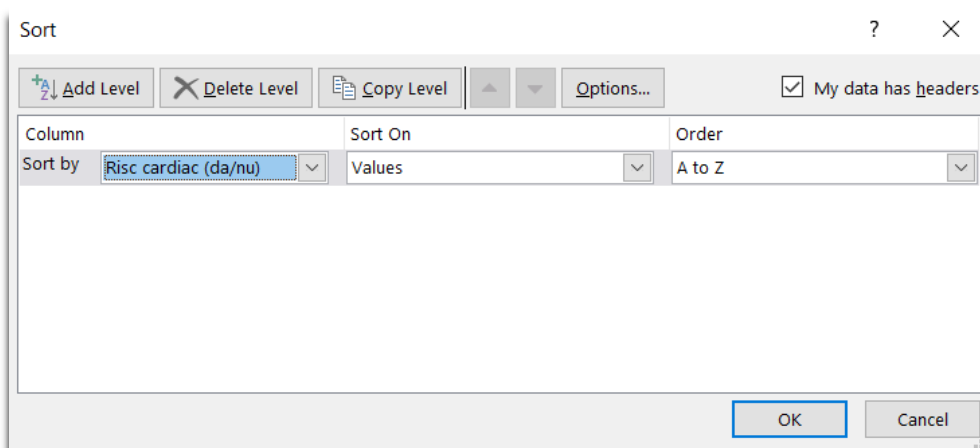
- Calculul cuartilei 3 (Q3) se va realiza similar;

Sortarea datelor in Excel în funcție de un criteriu/variabilă

- Selecția întregii baze de date
- Home-> Sort&Filter->Custom Sort:



- Selecția coloanei Risc cardiac, astfel:



- Validare prin OK.

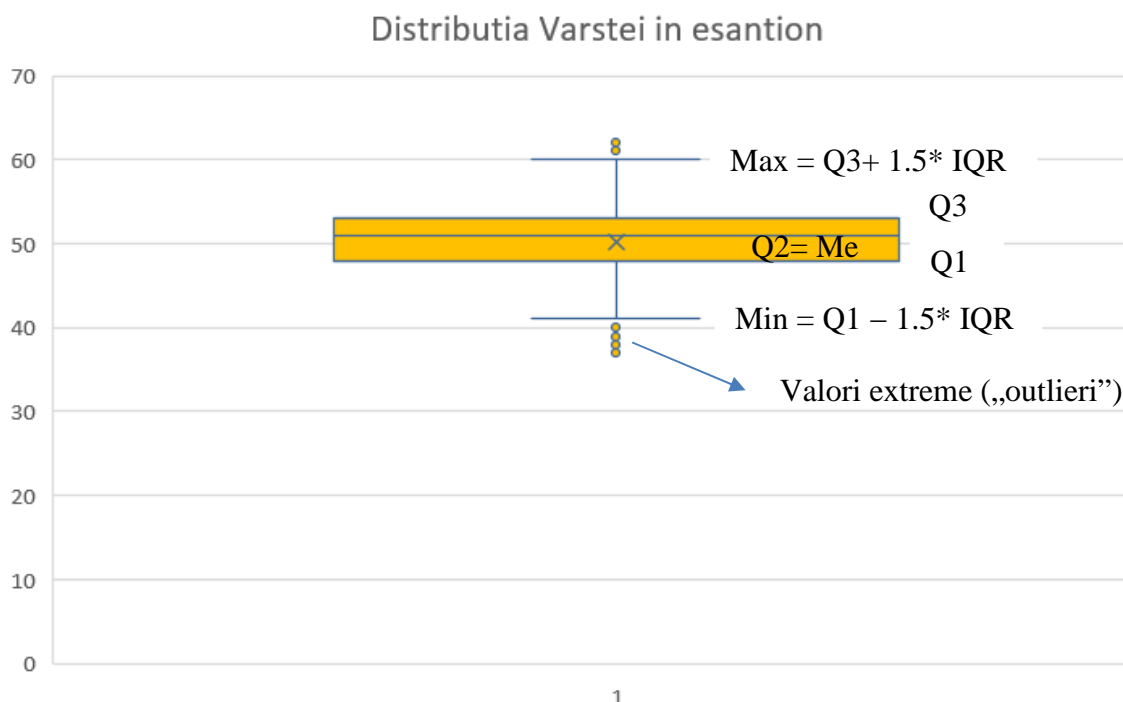
Realizarea unui grafic de tip box-plot

- Selectam valorile Vârstei din foaia de calcul Date
- Alegem opțiunea Insert→ Chart

The screenshot shows the Excel interface with the 'Insert' tab selected. In the 'Charts' group, the 'Box & Whisker' icon is highlighted. The 'Insert Chart' task pane is open on the right, with 'All Charts' selected. The 'Box & Whisker' chart type is chosen from the list. A preview of the box plot is shown in the task pane.

	A	B	C	D	E	F	G
	Varsta (ani)	Gen (M/F)	Mediu (U/R)	Greutate (kg)	Înălțime (cm)	Glicemie (mg/dL)	Colesterol (mg/dL)
1							
2	38	M	U	92	175	98	198
3	49	M	R	57	170	98	211
4	44	F	R	89	151	73	203
5	52	F	R	108	185	91	231
6	49	M	U	107	139	81	206
7	54	M	U	68	182	98	285
8	51	F	R	68	173	77	210
9	39	F	U	92	154	85	233
10	41	M	R	60	166	77	216
11	49	F	U	73	148	76	216
12	48	M	R	87	172	82	216
13	48	F	R	87	178	92	210
14	55	M	R	63	162	90	206
15	56	M	R	76	161	92	219
16	55	F	R	64	176	92	156
17	48	F	U	69	159	64	209

- Graficul obținut va fi de forma



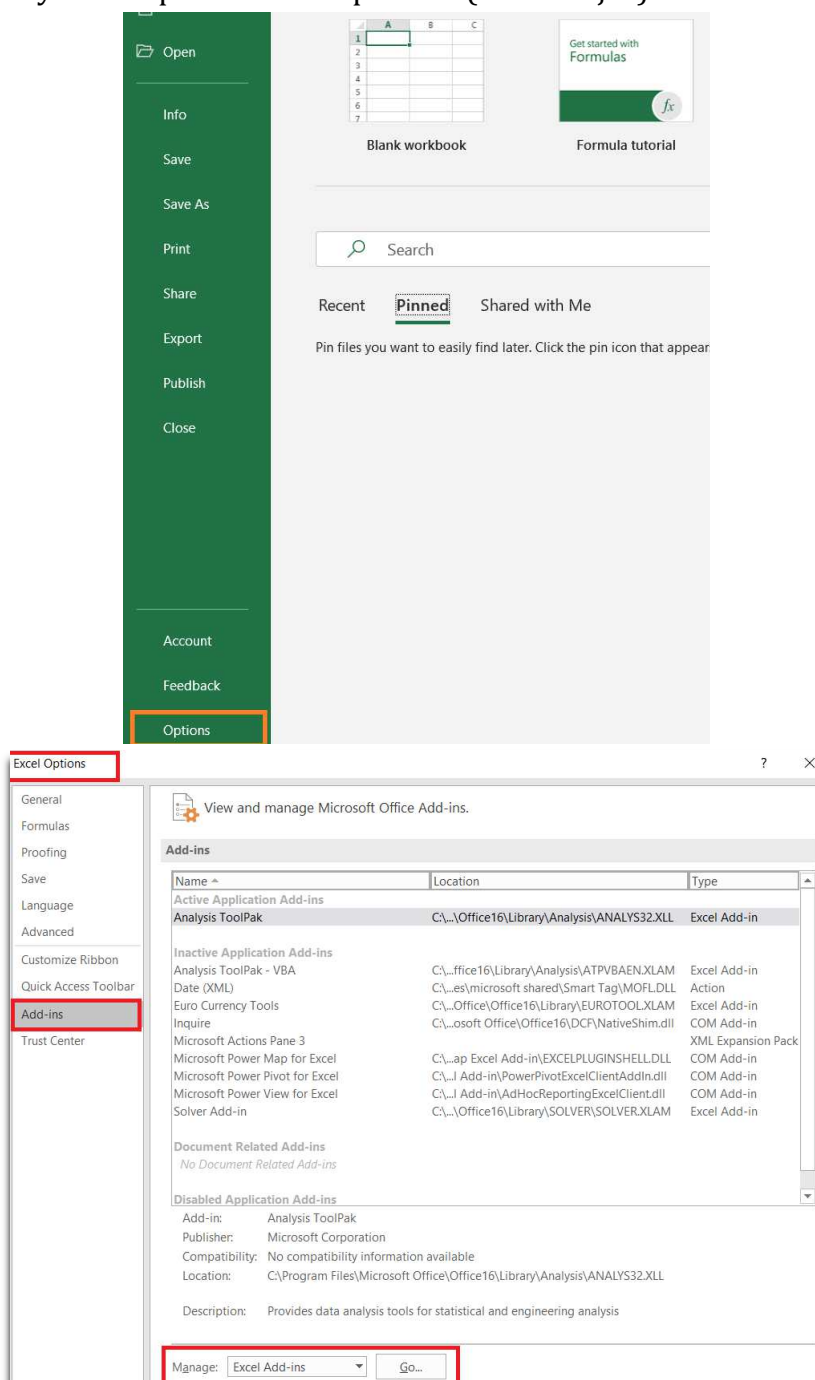
1

Calcul statisticilor descriptive utilizând pachetul de funcții Data Analysis

Pas 1. Verificați dacă în meniul Data există instalat pachetul Data Analysis.

Dacă nu îl găsiți atunci parcurgeți următoarele etape pentru instalarea lui:

- Buton File- Options-Add-Ins-click buton Go de la opțiunea Manage-Excel-Add-Ins-analysis Toolpack-validare prin OK (vezi mai jos)



Pas 2. In tab-ul **DATA** se cauta pachetul **Data Analysis**, se caută opțiunea **Descriptive Statistics**

- Dublu click pe aceasta opțiune;
- În căsuța corespunzătoare la **Input Range** selectați toate celulele care conțin valorile variabilelor cantitative corespunzătoare **pacienților cu risc cardiac!!!**-vezi imaginea de mai jos:

Descriptive Statistics

Input
Input Range:
Grouped By: ☒ Columns ☐ Rows
☒ Labels in first row

Output options
☒ Output Range:
☐ New Worksheet Ply:
☐ New Workbook
☒ Summary statistics
☐ Confidence Level for Mean: %
☐ Kth Largest:
☐ Kth Smallest:

OK Cancel Help

- Specificați că aveți denumiri de coloană în primul rând (bifați **Labels in first row**).
- În secțiunea **Output options**: alegeți să afișați rezultatele începând cu celula K1 și explicați programului să calculeze toate statisticile descriptive (bifați **Summary statistics**).
- **Rezultatele obținute vor fi de forma:**

Varsta (ani)		Greutate (kg)		Înălțime (cm)		Glicemie (mg/dL)		Colesterol (mg/dL)		Trigliceride (mg/dL)		HDL colesterol (mg/dL)	
Mean	50.22917	Mean	81.14583	Mean	171.1458	Mean	87.5625	Mean	211.4792	Mean	202	Mean	43.9375
Standard Error	0.741524	Standard Error	2.038414	Standard Error	1.955169	Standard Error	1.72536	Standard Error	4.164289	Standard Error	7.193315948	Standard Error	0.915492
Median	51	Median	82	Median	172	Median	87.5	Median	215.5	Median	198	Median	43.5
Mode	51	Mode	92	Mode	160	Mode	98	Mode	221	Mode	233	Mode	43
Standard Deviation	5.137429	Standard Deviation	14.12255	Standard Deviation	13.54581	Standard Deviation	11.95364	Standard Deviation	28.85104	Standard Deviation	49.83675479	Standard Deviation	6.342717
Sample Variance	26.39317	Sample Variance	199.4464	Sample Variance	183.4889	Sample Variance	142.8896	Sample Variance	832.3825	Sample Variance	2483.702128	Sample Variance	40.23005
Kurtosis	0.154518	Kurtosis	-0.98485	Kurtosis	-0.30453	Kurtosis	1.469124	Kurtosis	0.437838	Kurtosis	-0.35924712	Kurtosis	-0.60084
Skewness	-0.51322	Skewness	0.058909	Skewness	-0.23366	Skewness	-0.50192	Skewness	-0.09993	Skewness	0.278789338	Skewness	0.065564
Range	22	Range	54	Range	57	Range	65	Range	139	Range	203	Range	24
Minimum	38	Minimum	55	Minimum	139	Minimum	51	Minimum	146	Minimum	105	Minimum	32
Maximum	60	Maximum	109	Maximum	196	Maximum	116	Maximum	285	Maximum	308	Maximum	56
Sum	2411	Sum	3895	Sum	8215	Sum	4203	Sum	10151	Sum	9696	Sum	2109
Count	48	Count	48	Count	48	Count	48	Count	48	Count	48	Count	48

- ✓ **Identificați (colorați) în tabelul d-voastră acele statistici descriptive specificate în enunțul problemei:**
- ✓ Parametrii de centralitate: medie (engl. Mean), mediana (engl. Median), modul (engl. Mode)
- ✓ Parametrii de dispersie: deviație standard, eroare standard
- ✓ Parametrii de asimetrie: asimetrie (engl. Skewness); boltire (engl. Kurtosis)

Calcul coeficient de variație

- Se aplica formula dată, ținând cont că deviația standard a variabilei Varsta se află în celula L7 iar media aritmetică a variabilei Varsta se află în celula L3, deci formula devine: $=L7/L3*100$
- Extindeți formula la dreapta pentru restul variabilelor.

Distributie simetrică de date

- Dacă media aritmetică a valorilor variabilei de studiu este aproximativ egală cu mediana și modulul atunci vom spune că distribuția variabilei este simetrică ($M_a \approx M_e \approx M_o$)

Distribuție Gauss (normală) de date

- Dacă media aritmetică a valorilor variabilei de studiu este aproximativ egală cu mediana și modulul și în plus coeficienții de asimetrie și boltire sunt în jurul lui 0 (intervalul $[-1; 1]$) vom spune că distribuția variabilei este normală (Gaussiană)

Interpretare coeficient de variație

- Dacă $CV < 10\%$ atunci valorile colesterolului pot fi considerate omogene (în jurul mediei)
- Dacă $10 \leq CV < 20\%$ atunci valorile colesterolului pot fi considerate **relativ** omogene (în jurul mediei)
- Dacă $20 \leq CV < 30\%$ atunci valorile colesterolului pot fi considerate **relativ** eterogene (în jurul mediei)
- Dacă $CV \geq 30\%$ atunci valorile colesterolului pot fi considerate eterogene (în jurul mediei)