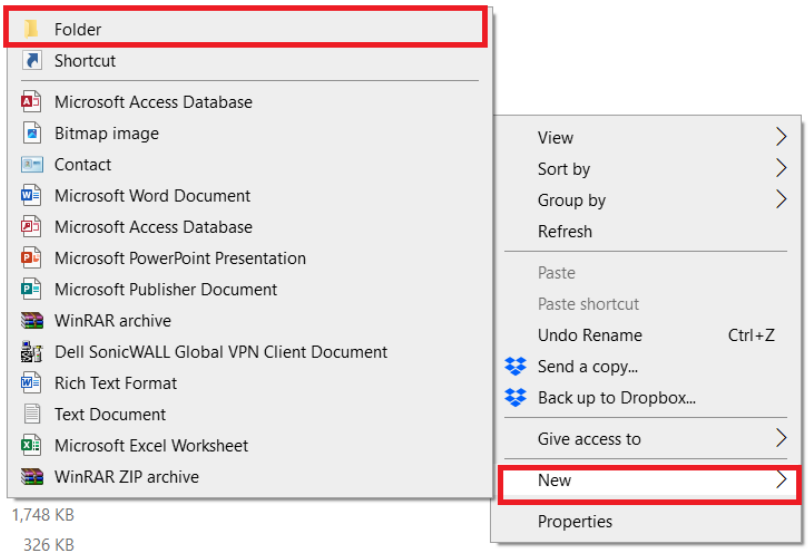
## De reţinut:

- Denumirile fişierelor şi folderelor trebuie să fie în conformitate cu cerinţele.

## Crearea unui nou folder pe Desktop

- Poziţionaţi-vă pe Desktop
- Creaţi un nou folder [click dreapta de mouse şi utilizaţi opţiunea Folder-> New].




## Salvarea unui fişier Excel anterior creat într-un anumit folder

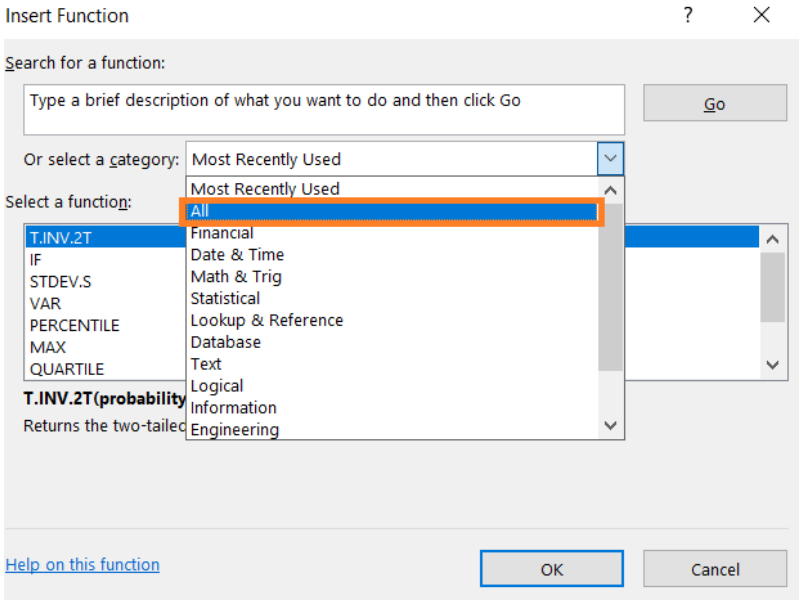
- Descărcaţi fişierul **BD_LP05_TD** de pe site (el se va salva automat în folderul Downloads)
- Deschideţi fişierul **BD_LP05_TD** şi apoi resalvaţi fişierul în folderul specificat în exerciţiu (Lab04) utilizând opţiunea: File- Save As-Desktop-folder **Lab05**

## Calculul statisticilor descriptive (medie, mediană, modul, deviaţie standard) prin funcţii predefinite

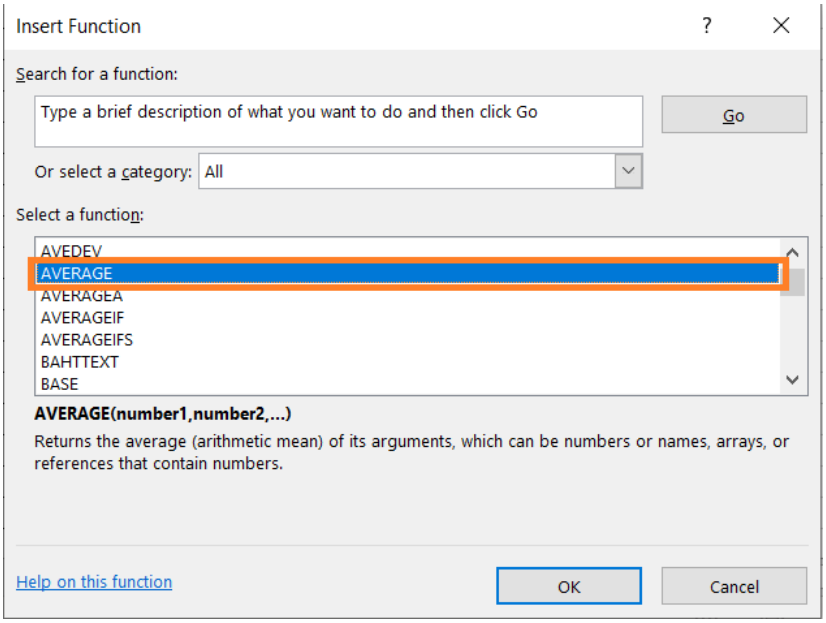
- la dreapta bazei de date, în aceeaşi foaie de calcul (de ex. in coloanele M si N), aveţi dat un tabel de forma:

| M | N |
|---|---|
| **TABEL 1** | |
| | **Adaptarea marginala (micrometri)** |
| **Medie** | |
| **Mediana** | |
| **Modul** | |
| **Deviație standard (DS)** | |
| **Cuartila 1 (Q1)** | |
| **Cuartila 3 (Q3)** | |
| **Coeficient de variatie (CV)** | |

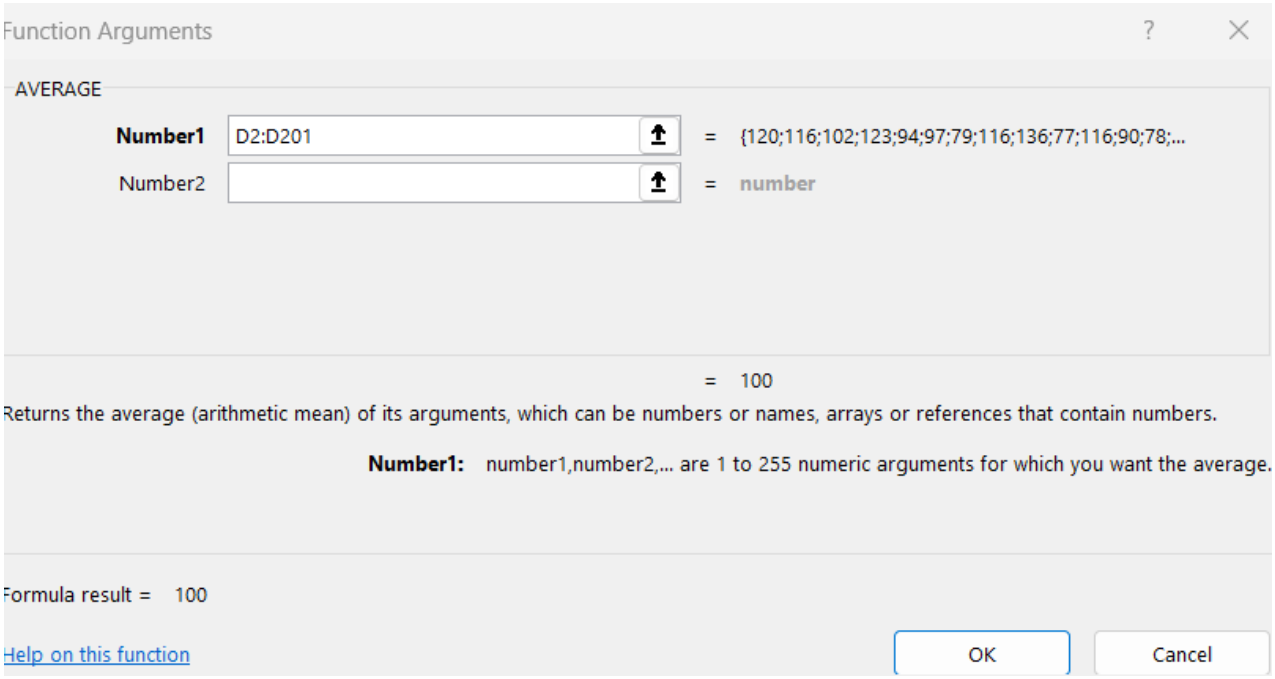
- Poziționați-vă în acest tabel, pe celula unde doriți sa obțineți rezultatul mediei și folosiți opțiunea **Formulas->Insert Functions**




- 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 adaptarii marginale petnru ceel 200 lucrari protetice (de ex. D2: D201).



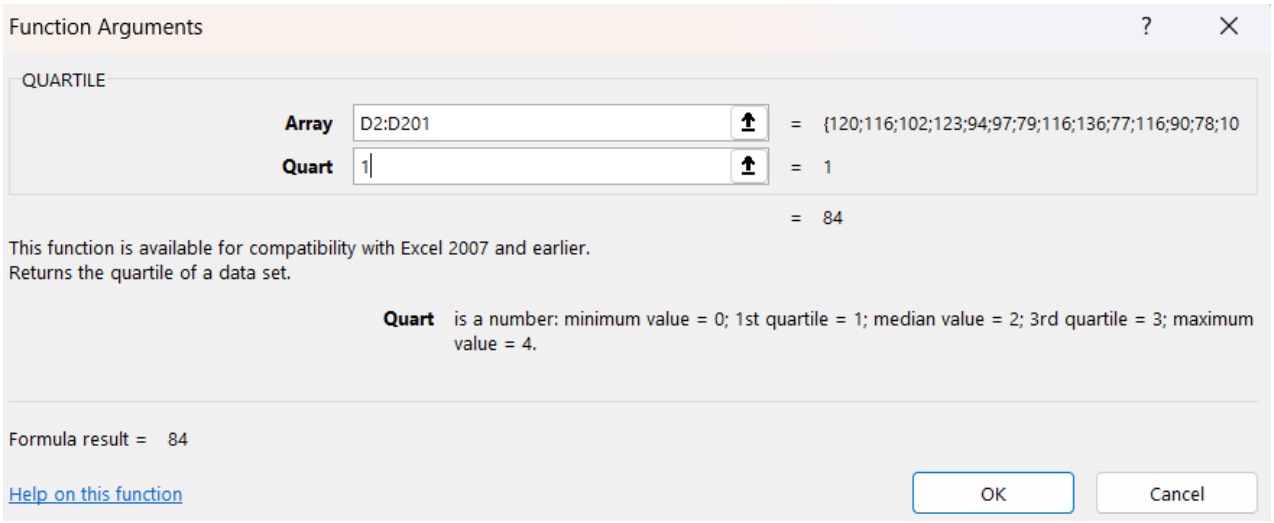

| | L | M |
|---|---|---|
| | **TABEL 1** | |
| | | **Adaptarea marginala (micrometri)** |
| | **Medie** | **100** |
| | **Mediana** | |
| | **Modul** | |
| | **Deviație standard (DS)** | |
| | **Cuartila 1 (Q1)** | |
| | **Cuartila 3 (Q3)** | |
| | **Coeficient de variatie (CV)** | |

- **Pentru calculul medianei se va folosi funcția MEDIAN**
- **Pentru calculul modulului se va folosi funcția MODE**
- **Pentru calculul deviației standard se va folosi funcția STDEV.**

**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;**

## Calcul coeficient de variație

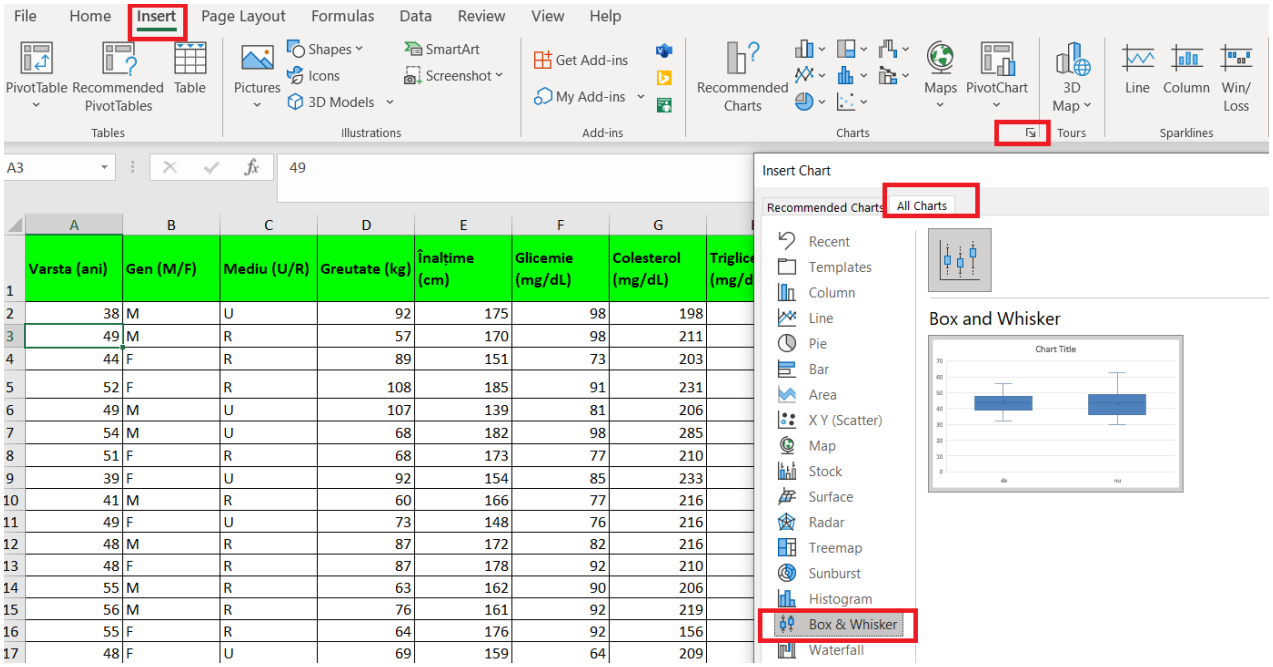
- Se aplica formula dată, ținând cont ca deviația standard a variabilei Adaptare marginala se află în celula N7 iar media aritmetica a variabilei Vârsta se află în celula N4, deci formula devine: =N7/N4*100

## Interpretare coeficient de variație

- Daca CV<10% atunci valorile adaptarii marginale pot fi considerate omogene (in jurul mediei)
- Daca 10≤CV<20% atunci valorile adaptarii marginale pot fi considerate **relativ** omogene (in jurul mediei)
- Daca 20≤CV<30% atunci valorile adaptarii marginale pot fi considerate **relativ** eterogene (in jurul mediei)
- Daca CV≥30% atunci valorile adaptarii marginale pot fi considerate eterogene (in jurul mediei)

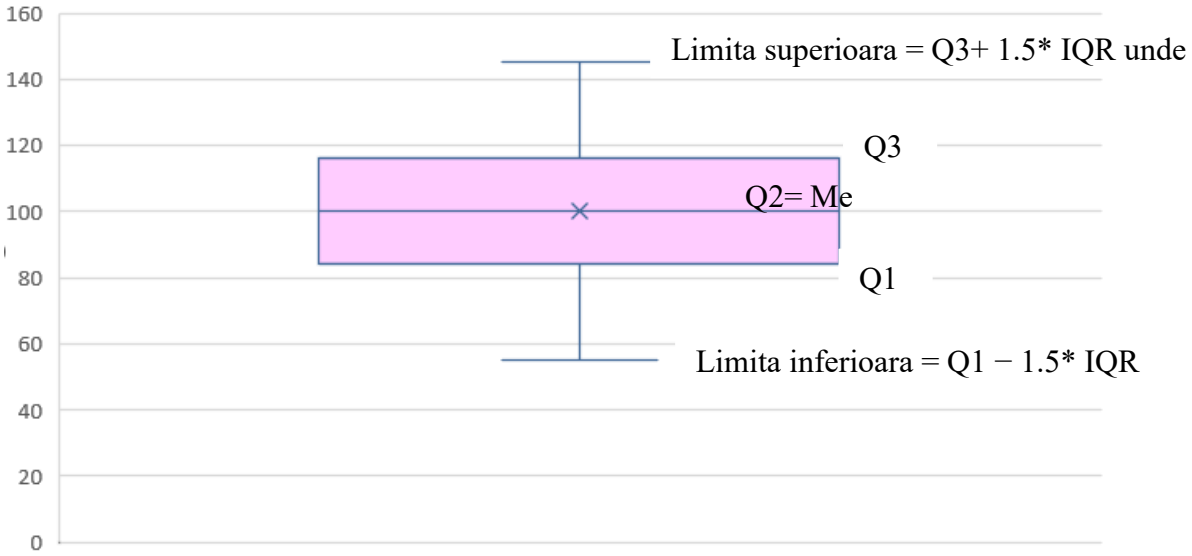
## Realizarea unui grafic de tip box-plot pentru o variabilă CANTITATIVA

- **Selectam valorile Adaptarii marginale din foaia de calcul Date**
- **Alegem opțiunea Insert→ Chart → ALL charts→ Box and Whisker**




| | A | B | C | D | E | F | G |
|---|---|---|---|---|---|---|---|
| 1 | Varsta (ani) | Gen (M/F) | Mediu (U/R) | Greutate (kg) | Înaltime (cm) | Glicemie (mg/dL) | Colesterol (mg/dL) |
| 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**




**IQR = distanța intercuartilică =Q3-Q1**

**Mustățile reprezintă intervalul de valori care sunt considerate normale pentru variabila de studiu (intervalul de „normalitate" este dat de limita inferioara si limita superioara).**

**Punctele di afara mustăților (oultlieri): valori neobisnuite**

**<u>Interpretare grafic:</u>**

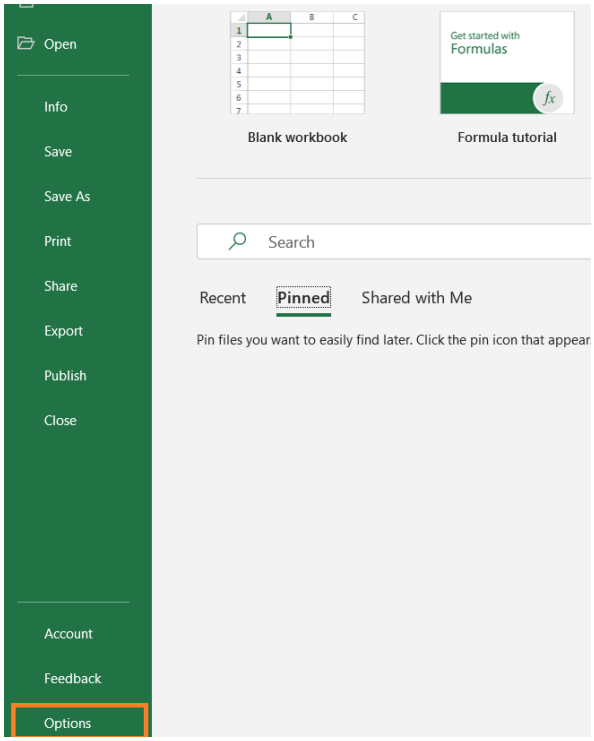
- **Dacă mediana este centrată în cutie (≈ media) iar mustățile sunt aproximativ egale → distribuție normală de probabilitate**
- **Dacă mediana este mai aproape de Q1 și mustața de sus este mai mare (lungă) → distribuție cu asimetrie pozitivă**
- **Dacă mediana este mai aproape de Q3 și mustața de jos este mai mare (lungă) → distribuție cu asimetrie negativă**

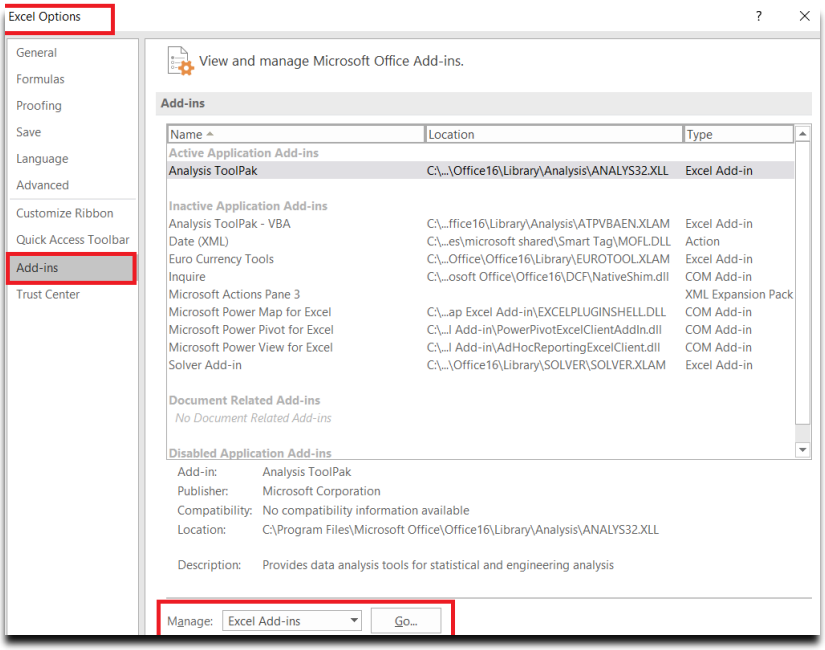
## Calcul statisticilor descriptive utilizând pachetul de funcţii DATA ANALYSIS:

**Pas 1.** Verificaţi dacă în meniul Data există instalat pachetul Data Analysis.

**Dacă nu îl găsiţi atunci parcurgeţi următoarele etape pentru instalarea lui:**

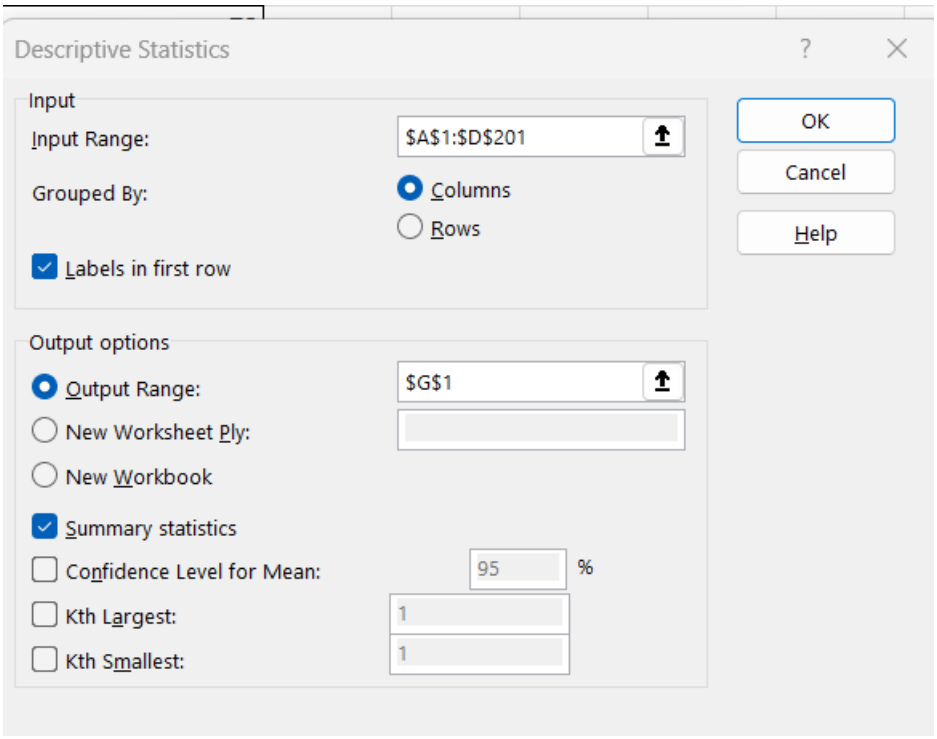
- Buton File- Options-Add-Ins-click buton Go de la opţiunea Manage-Excel-Add-Ins-analysis Toolpack-validare prin OK (vezi mai jos)




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

- Dublu click pe aceasta optiune;
- În căsuţa corespunzătoare la **Input Range** selectaţi toate celulele care conţin valorile variabilelor cantitative!!!-vezi imaginea de mai jos:




- Specificaţi că aveţi denumiri de coloană în primul rând (bifaţi **Labels in first row**).
- În secţiunea **Output** options: alegeţi să afişaţi rezultatele începând cu celula K1 şi explicaţi programului să calculeze toate statisticile descriptive (bifaţi **Summary statistics**).
- **Rezultatele obţinute vor fi de forma:**

| G | H | I | J | K | L | M | N |
|---|---|---|---|---|---|---|---|
| *Timp de modelare in ceara (min)* | | *Contractie la sinterizare (%)* | | *Adaptare marginala (µm)* | | *Timp de polimerizare rasina acrilica (min)* | |
| | | | | | | | |
| Mean | 88.64 | Mean | 25.69 | Mean | 100.00 | Mean | 94.38 |
| Standard Error | 1.47 | Standard Error | 0.25 | Standard Error | 1.44 | Standard Error | 1.01 |
| Median | 90.00 | Median | 25.46 | Median | 100.00 | Median | 94.00 |
| Mode | 120.00 | Mode | 20.40 | Mode | 116.00 | Mode | 87.00 |
| Standard Deviation | 20.72 | Standard Deviation | 3.48 | Standard Deviation | 20.30 | Standard Deviation | 14.24 |
| Sample Variance | 429.50 | Sample Variance | 12.14 | Sample Variance | 411.99 | Sample Variance | 202.80 |
| Kurtosis | -1.25 | Kurtosis | -0.40 | Kurtosis | -0.54 | Kurtosis | -1.14 |
| Skewness | -0.13 | Skewness | 0.29 | Skewness | 0.00 | Skewness | 0.10 |
| Range | 66.00 | Range | 16.14 | Range | 90.00 | Range | 50.00 |
| Minimum | 54.00 | Minimum | 18.83 | Minimum | 55.00 | Minimum | 70.00 |
| Maximum | 120.00 | Maximum | 34.98 | Maximum | 145.00 | Maximum | 120.00 |
| Sum | 17728.00 | Sum | 5138.48 | Sum | 20000.00 | Sum | 18875.00 |
| Count | 200.00 | Count | 200.00 | Count | 200.00 | Count | 200.00 |

- ✓ **Identificati (colorati) in tabelul d-voastra acele statistici descriptive specificate în enunţul problemei:**
- ✓ Parametrii de centralitate: medie (engl. Mean), mediana (engl. Median), modul (engl. Mode)
- ✓ Parametrii de dispersie: deviaţie standard, eroare standard
- ✓ Parametrii de asimetrie: asimetrie (engl. Skewness); boltire (engl. Kurtosis)

## Distributie simetrică a variabilei de interes

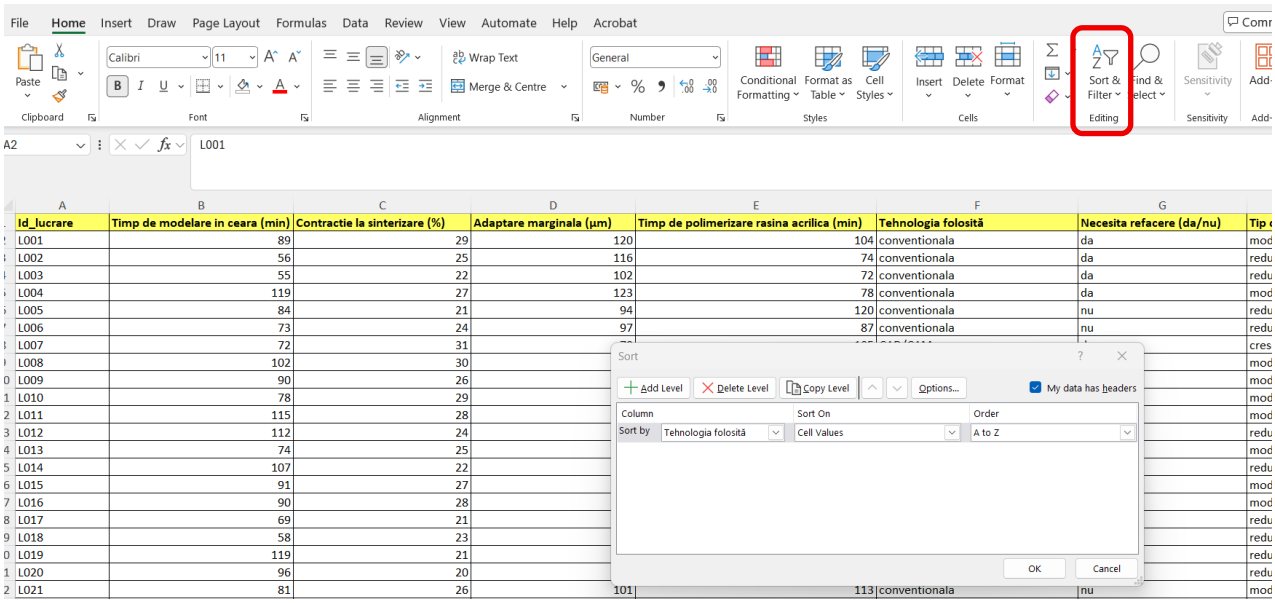
- Daca media aritmetica a valorilor variabilei de studiu este aproximativ egala cu mediana si modulul atunci vom spune ca distributia variabilei este simetrica (Ma≈Me≈Mo)

## Distributie Gaussiană (normală) a variabilei de interes

- Daca media aritmetica a valorilor variabilei de studiu este aproximativ egala cu mediana si modulul ŞI în plus coeficientii de asimetrie si boltire sunt in jurul lui 0 (intervalul [-1; 1]) vom spune ca distributia variabilei este normala (Gaussiană)

## Sortarea datelor in raport cu un criteriu

- Daca dorim sa sortam pacienţii in raport cu genul pacienţilor
- Se selectează toata baza de date
- Se alege din meniu optiunea **Sort & filter → CUSTOM SORT**
- **Se selecteaza criteriul (variabila) în funcţie de care se realizează sortarea datelor**

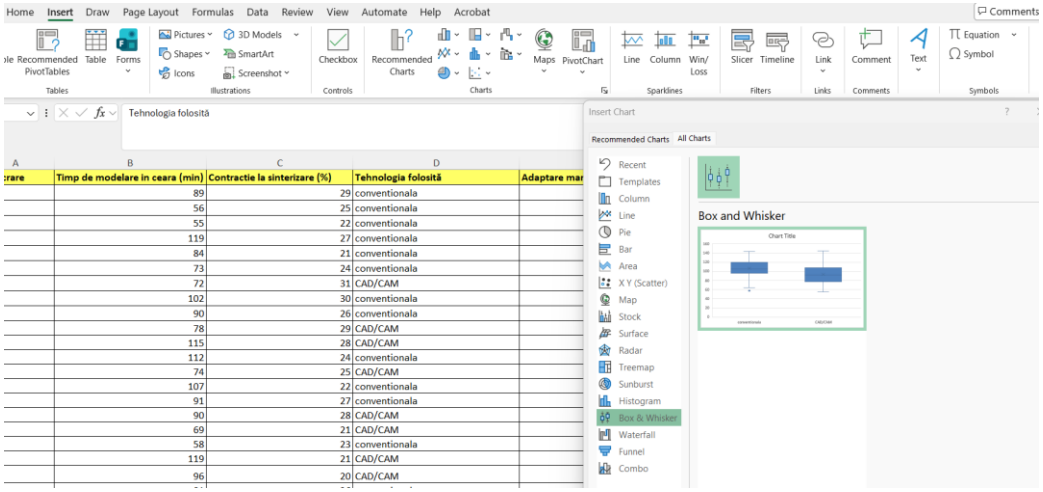


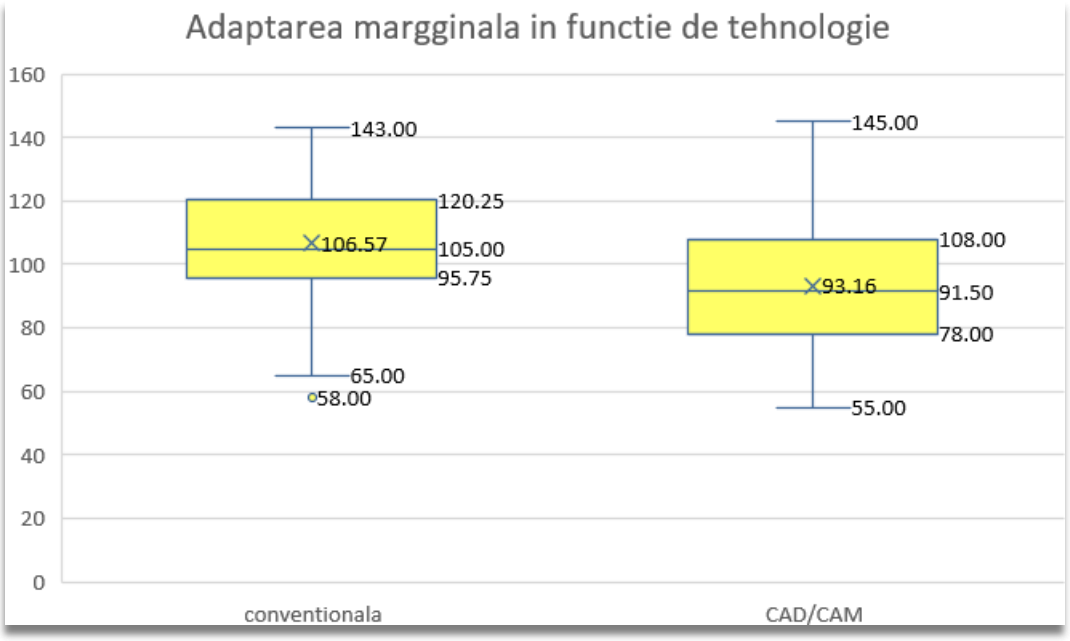

| Id_lucrare | Timp de modelare in ceara (min) | Contractie la sinterizare (%) | Adaptare marginala (µm) | Timp de polimerizare rasina acrilica (min) | Tehnologia folosită | Necesita refacere (da/nu) |
|---|---|---|---|---|---|---|
| L001 | 89 | 29 | 120 | 104 | conventionala | da |
| L002 | 56 | 25 | 116 | 74 | conventionala | da |
| L003 | 55 | 22 | 102 | 72 | conventionala | da |
| L004 | 119 | 27 | 123 | 78 | conventionala | da |
| L005 | 84 | 21 | 94 | 120 | conventionala | nu |
| L006 | 73 | 24 | 97 | 87 | conventionala | nu |
| L007 | 72 | 31 | | | | |
| L008 | 102 | 30 | | | | |
| L009 | 90 | 26 | | | | |
| L010 | 78 | 29 | | | | |
| L011 | 115 | 28 | | | | |
| L012 | 112 | 24 | | | | |
| L013 | 74 | 25 | | | | |
| L014 | 107 | 22 | | | | |
| L015 | 91 | 27 | | | | |
| L016 | 90 | 28 | | | | |
| L017 | 69 | 21 | | | | |
| L018 | 58 | 23 | | | | |
| L019 | 119 | 21 | | | | |
| L020 | 96 | 20 | | | | |
| L021 | 81 | 26 | | 101 | 113 conventionala | nu |

- Validam prin OK.

## Realizarea unui grafic de tip box-plot pentru o variabilă CANTITATIVA pe subgrupuri

- **Selectam coloana Sex şi Vârsta din foaia de calcul Date**
- **Alegem opţiunea Insert→ Chart→ ALL charts→ Box and Whisker**




- Grafic final: