

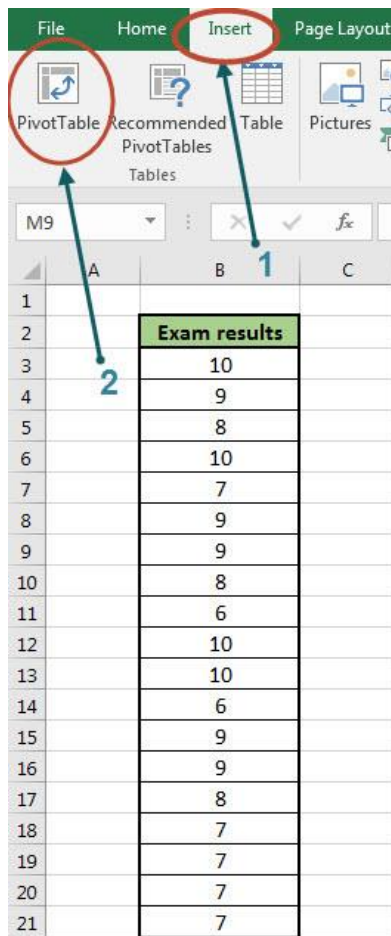
Cuprins

1. Variabile cantitative – grafic tip coloane	2
2. Variabile calitative – grafic pie (diagramă circulară)	9
3. Variabile cantitative continue – histogramă	12
4. Deviația standard și intervalul de încredere pentru 95 % – statistica descriptivă	18

1. Variabile cantitative – grafic tip coloane

Având în vedere notele din tabelul de mai jos (cu valori de la 1 la 10) le vom procesa ca fiind variabile cantitative. Vom folosi opțiunea **Pivot Table** pentru a determina de câte ori apare fiecare din note (frecvența) și apoi opțiunea **Column Graph** pentru a efectua reprezentarea grafică a rezultatelor obținute.

Este posibilă determinarea frecvenței și prin folosirea funcției **COUNTIF** sau unor variații ale acesteia. Această funcție va fi folosită în exemple ce vor fi prezentate ulterior, alături de alte funcții similare; în continuare ne concentrăm pe detalierea opțiunilor funcției **Pivot Table**.



	A	B	C
1			
2		Exam results	
3		10	
4		9	
5		8	
6		10	
7		7	
8		9	
9		9	
10		8	
11		6	
12		10	
13		10	
14		6	
15		9	
16		9	
17		8	
18		7	
19		7	
20		7	
21		7	

Figura 1 – Variabile cantitative (rezultate examene - exam results): Insert – Pivot table

- ✓ Pasul 1 – Selectați opțiunea **Insert**
- ✓ Pasul 2 - Selectați **Pivot Table**

La crearea unui Pivot Table trebuie luate în considerare mai multe aspecte. În primul rând trebuie definit **Table/Range** (selectarea datelor de procesat pentru Pivot Table). În cazul de față vom selecta toate datele, inclusive capul de tabel (de la B2 la B22 – figura 2). Zona selectată va apărea cu linie punctată. Trebuie selectată și locația în care va fi creat noul Pivot Table – figura 2.

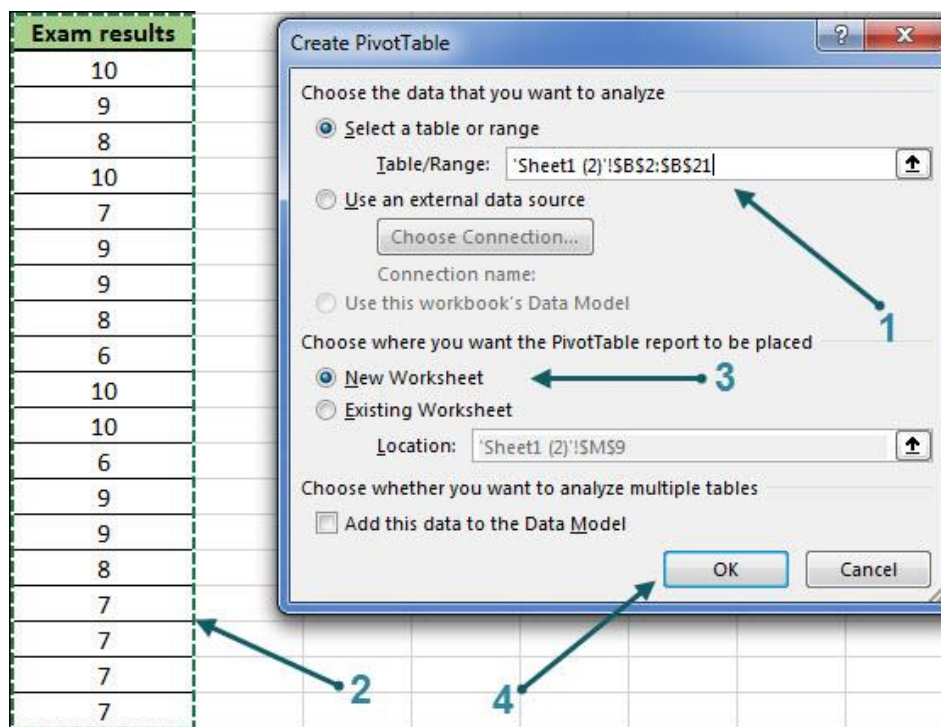


Figura 2 – Crearea unui Pivot Table

- ✓ Pasul 1 și 2 – Selectăm **Range** – datele ce urmează a fi procesate (în jurul datelor selectate va apărea o linie punctată)
- ✓ Pasul 3 – Selectăm unde dorim să fie afișat noul Pivot Table ce urmează a fi creat
- ✓ Pasul 4 – Apăsăm butonul OK

În figura 3 sunt detaliate aspectele legate de crearea unui Pivot Table (în acest caz pentru rezultatele obținute la examen – exam results). În acest caz lucrăm cu o singură variabilă – această trebuie să apară în opțiunea pentru rows (rânduri) și values (valori). Este important de reținut faptul că trebuie definită tipul de operație matematică potrivit pentru tipul de date ce urmează a fi procesat **Value Field Settings** – figurile 4 și 5.

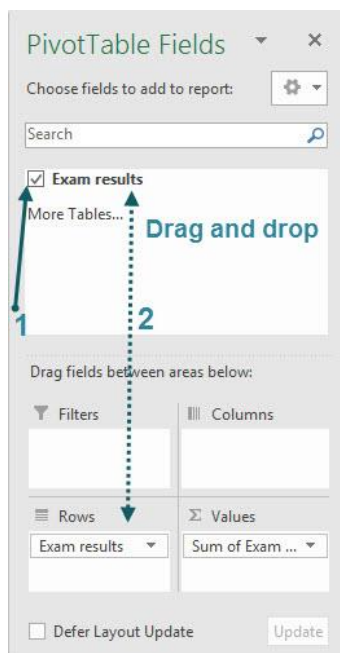


Figura 3 – Selectarea opțiunilor pentru crearea unui Pivot Table

- ✓ Pasul 1 – Bifăm opțiunea pentru Exam results (rezultate examen)
- ✓ Pasul 2 – Facem drag and pentru Exam results în zona pentru rows (rânduri)

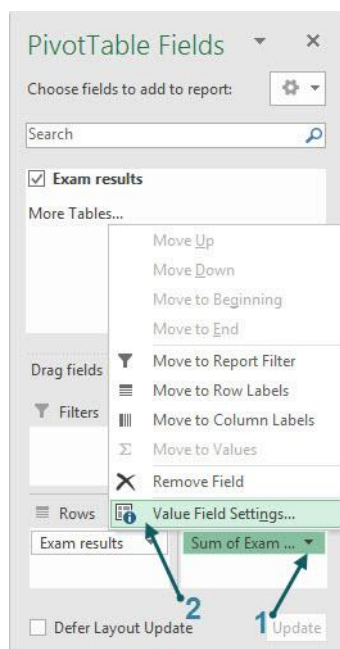


Figura 4 – Opțiunile pentru **Value Field Settings**

- ✓ Pasul 1 – click aici pentru a prezenta opțiunile
- ✓ Pasul 2 – selectăm opțiunea Value Field Settings

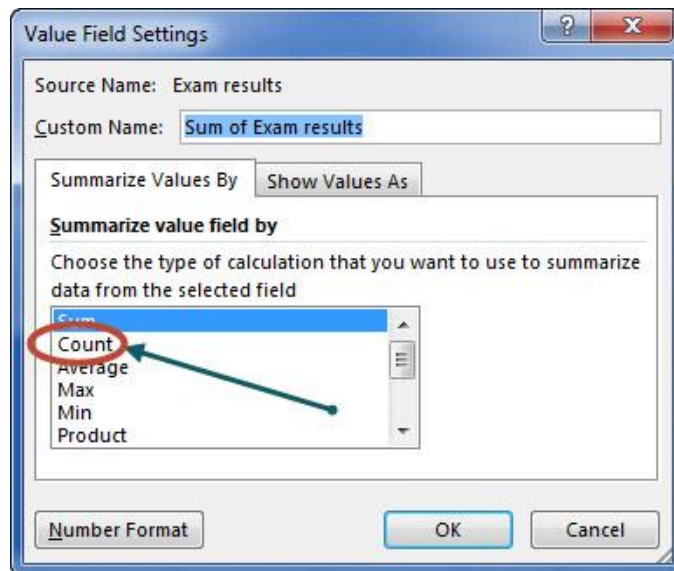


Figura 5 – Opțiunile pentru Value Field Settings: selectăm **Count**

După selectarea opțiunii **Count**, rezultatele obținute vor fi cele prezentate în figura 6. Pentru a crea un grafic tip coloane pentru rezultatele obținute trebuie în primul rând să selectăm datele din care dorim să creăm tabelul (datele sunt selectate după cum este detaliat în figura 6). După selectarea datelor folosim opțiunea **Insert** și apoi alegem grafic de tip **2D column** – figura 7.

După crearea graficului sunt disponibile două opțiuni prezentate în figura 8: **Chart Elements** și **Chart Styles**.

Row Labels	Count of Exam results
6	2
7	5
8	3
9	5
10	4
Grand Total	19

Figura 6 – Selectarea rezultatelor obținute după finalizarea Pivot Table

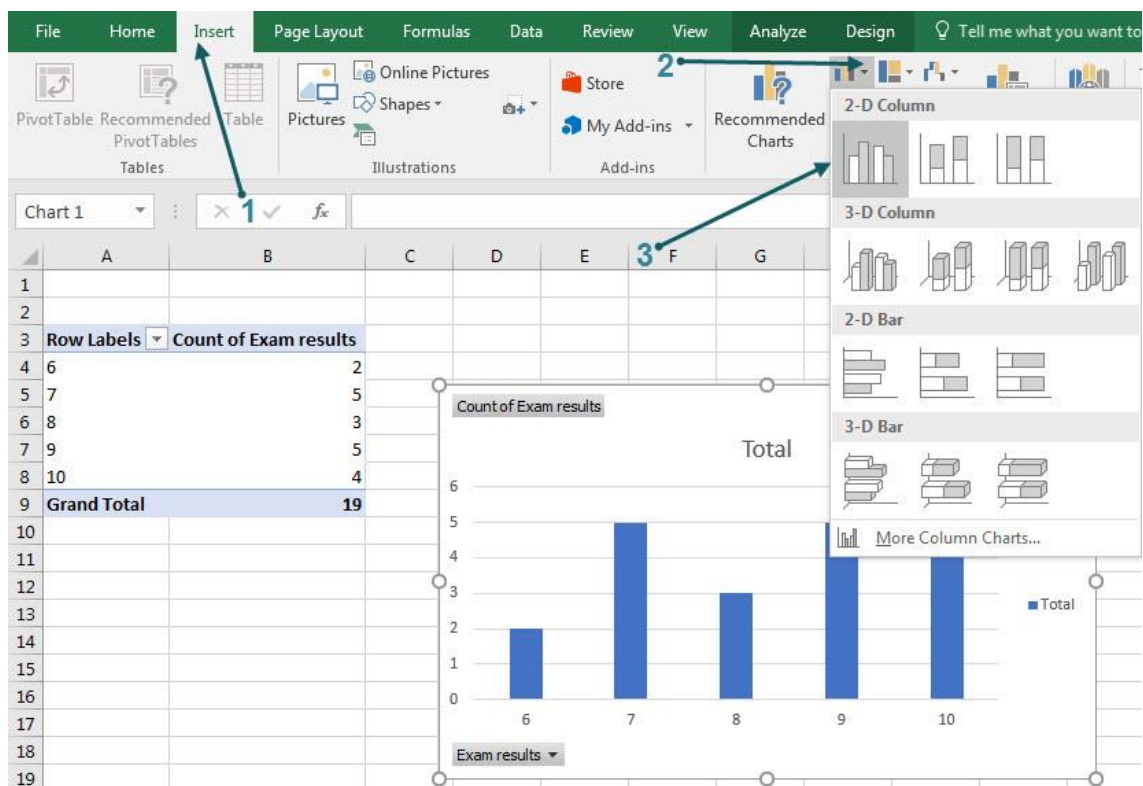


Figura 7 – Crearea unui grafic tip coloane

- ✓ Pasul 1 – Selectăm opțiunea **Insert**
- ✓ Pasul 2 – Extindem selecția pentru tipurile de grafic coloane
- ✓ Pasul 3 – Selectăm opțiunea de grafic **2D column**

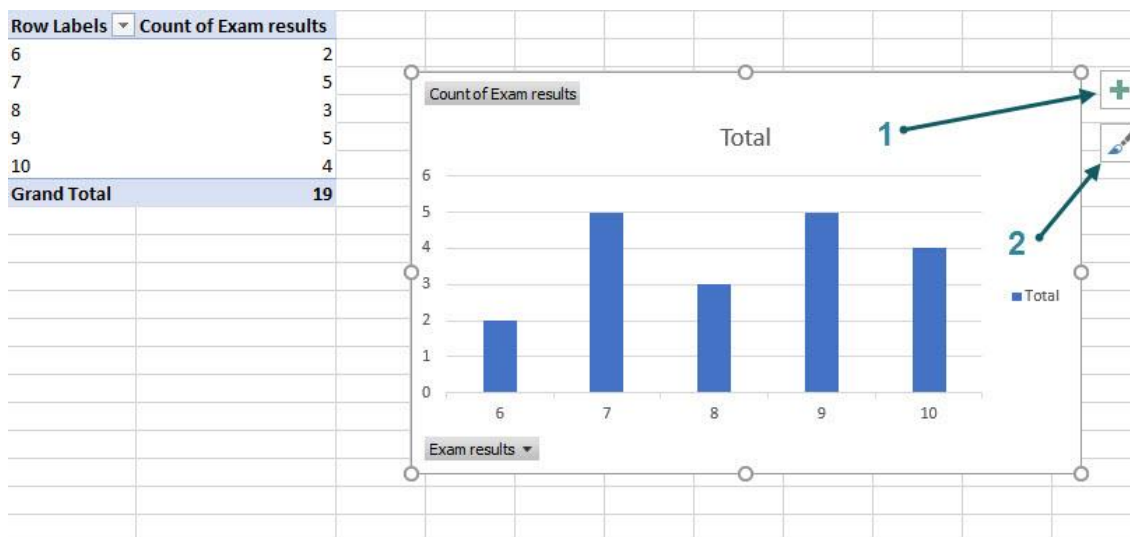


Figura 8 – Opțiunile Chart Elements (1) și Chart style (2)

Pentru a face graficul obținut mai clar vom folosi opțiunea Chart Elements. Astfel vom bifa opțiunile pentru **Axis Titles** și **Gridlines** conform detalierii din figura 9. În figura 10 este prezentată modalitatea de a ascunde butoanele din grafic **Hide All Field Buttons on Chart**.

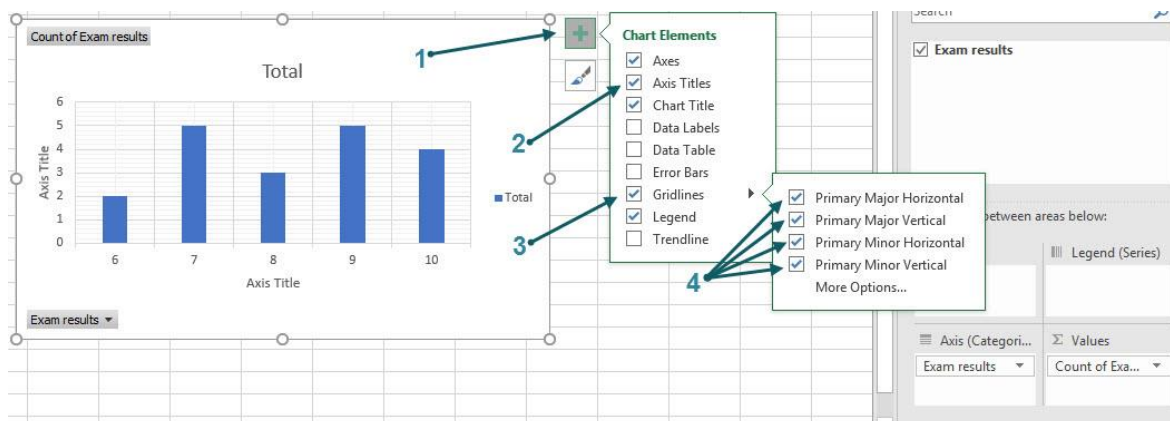


Figura 9 – Opțiunile pentru Chart Elements

- ✓ Pasul 1 – Selectăm opțiunea **Chart Elements**
- ✓ Pasul 2 – Selectăm opțiunea **Axis Titles**
- ✓ Pasul 3 – Extindem opțiunile pentru **Gridlines**
- ✓ Pasul 4 – Bifăm toate opțiunile pentru axe orizontale și verticale

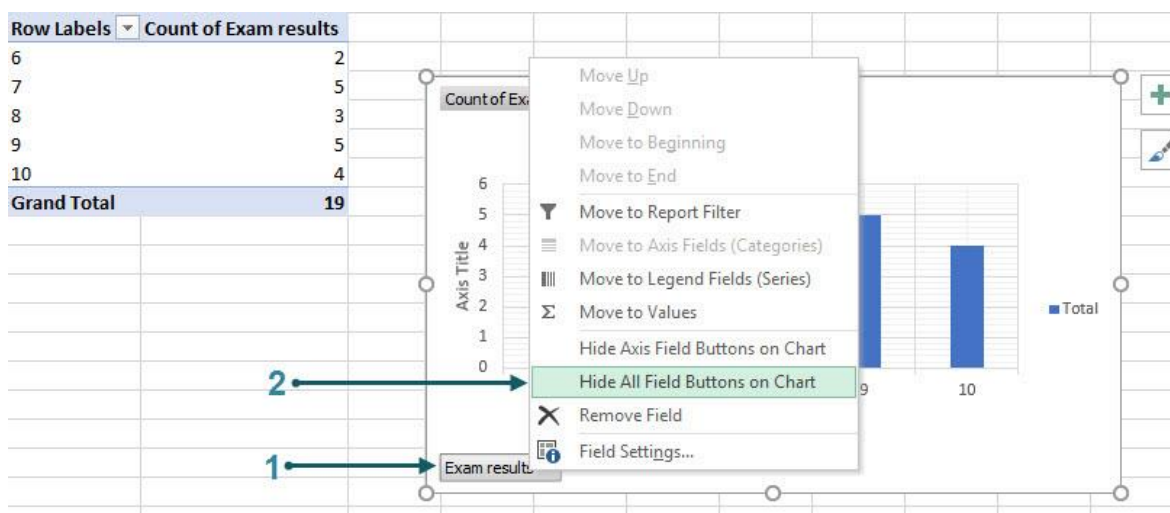


Figura 10 – Opțiunea **Hide All Field Buttons on Chart**

- ✓ Pasul 1 – Click dreapta pe Exam results
- ✓ Pasul 2 – Selectăm **Hide All Field Buttons on Chart**

În final vom edita axa orizontală, verticală și titlul graficului conform detalierii din figura 11. În acest caz legenda pentru grafic nu este necesară și poate fi ștearsă – figura 12.

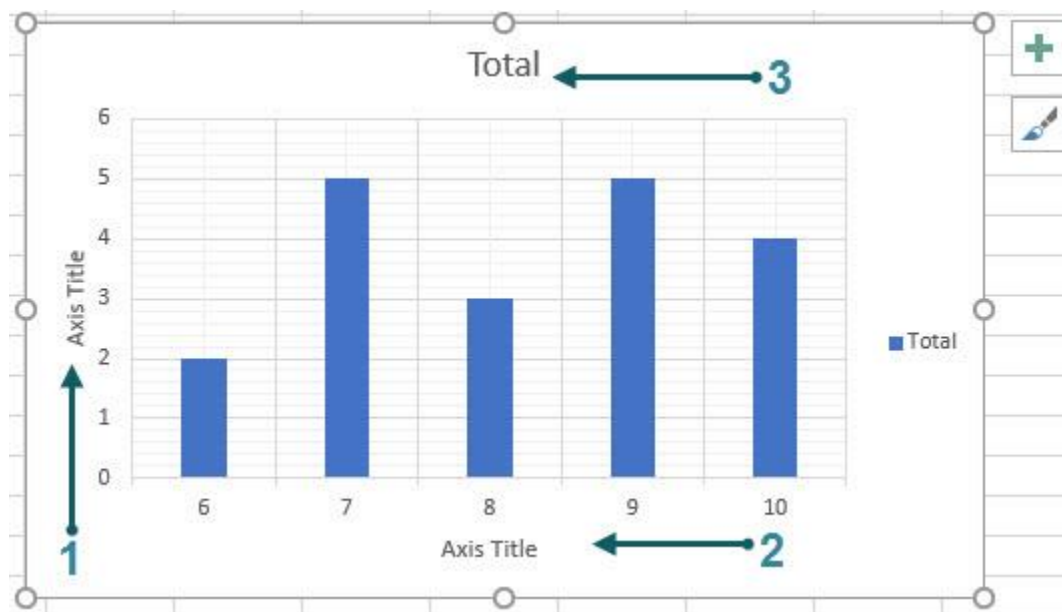


Figura 11 – Editarea axelor și a titlului graficului

- ✓ Pasul 1 – Schimbăm titlul axei verticale (axa y)
- ✓ Pasul 2 – Schimbăm titlul axei orizontale (axa x)
- ✓ Pasul 3 – Redenumim graficul

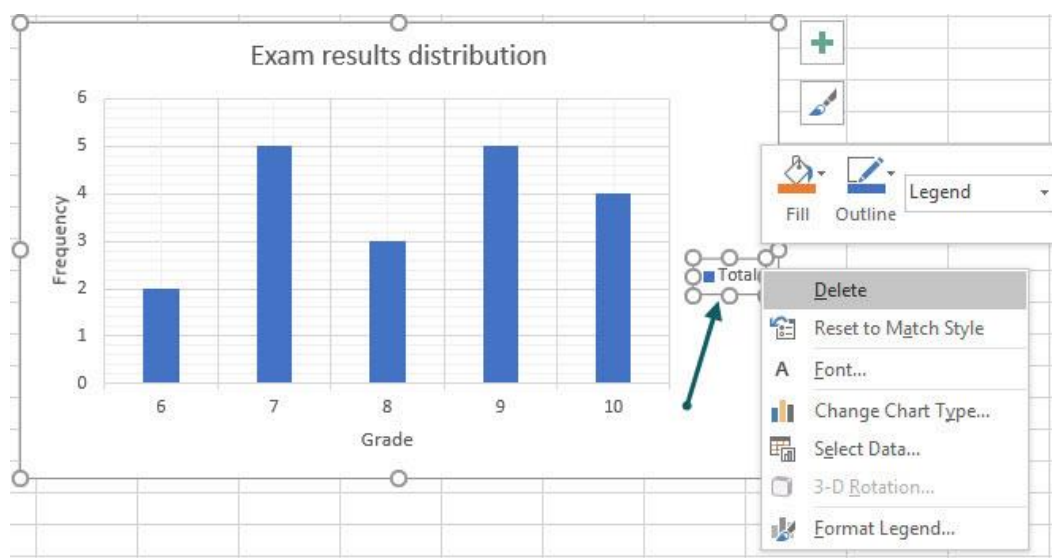


Figura 12 – Ștergerea legendei graficului

2. Variabile calitative – grafic pie (diagramă circulară)

În figura 13 este prezentat un exemplu pentru variabile calitative. Pentru procesarea acestui tip de date vom folosi opțiunea **Pivot Table**, similar cu exemplul anterior, dar de această dată pentru reprezentarea grafică vom folosi diagrama circulară (pie chart), care este recomandată în cazul reprezentărilor ce includ procente.

Folosind Pivot Table trebuie să obținem o numărare a diferitelor tipuri/producători de telefoane (phone brands). Pașii pentru realizarea unui **Pie Chart** (diagramă circulară) sunt detaliați în figura 14 – selectăm opțiunea **Insert** și apoi la tipul de grafic selectăm 3D Pie Chart. Editarea graficului astfel obținut este prezentată în figura 15, iar opțiunile de afișare **Label Options** sunt detaliate în figura 16.

Phone brand sales
Samsung
Iphone
Lg
Huawei
Allview
Iphone
Samsung
Samsung
Samsung
Samsung
Lg
Huawei
Lg
Lg
Iphone
Iphone
Iphone
Samsung
Samsung

Figura 13 – Variabile cantitative (phone brand sales – vânzări telefoane producători)

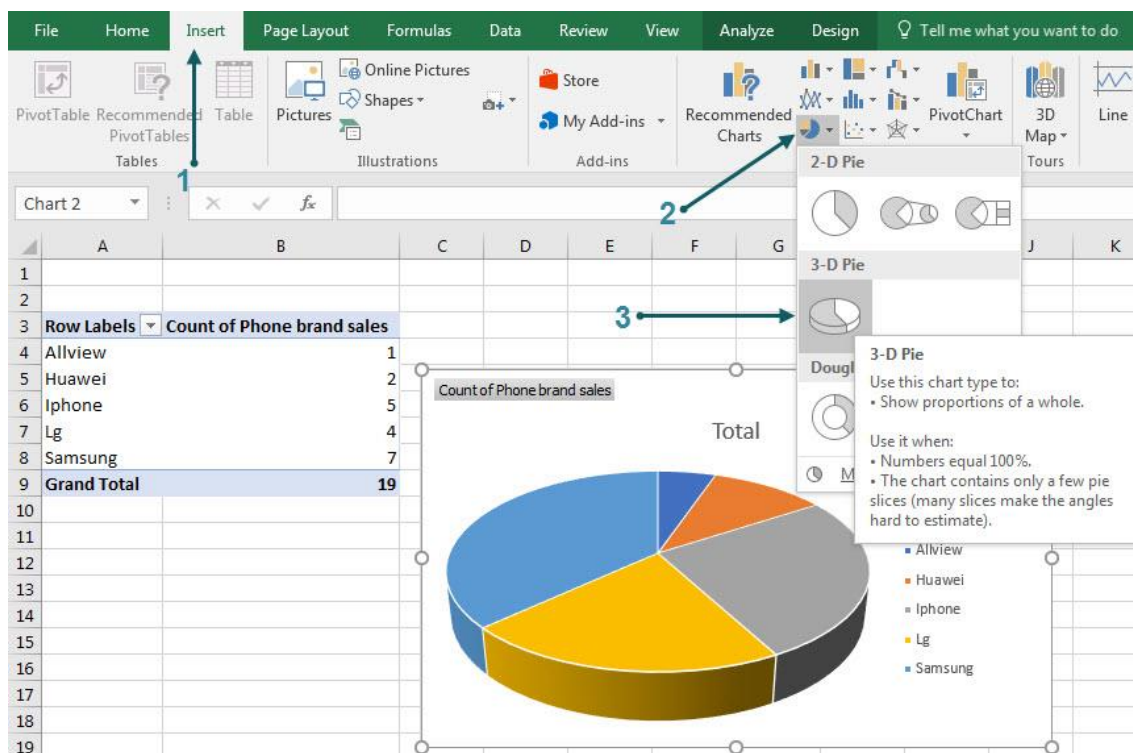


Figura 14 – Realizarea unui grafic tip 3D Pie chart

- ✓ Pasul 1 – Selectăm opțiunea Insert
- ✓ Pasul 2 – Extindem opțiunile pentru tipurile de grafic Pie
- ✓ Pasul 3 – Alegem opțiunea 3D Pie

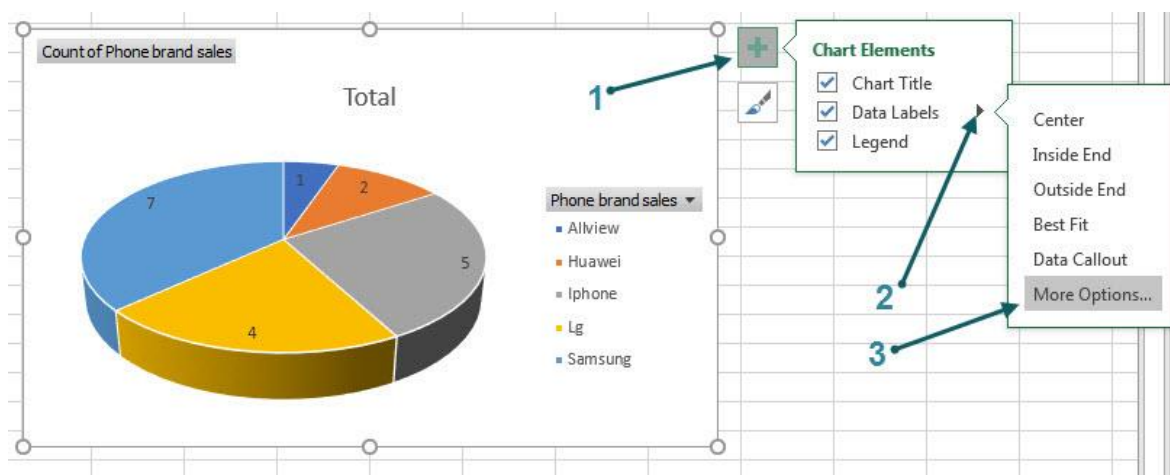


Figure 15 – Editarea unui grafic de tip Pie

- ✓ Pasul 1 – Apăsăm pe butonul Chart Elements
- ✓ Pasul 2 – Extindem opțiunile pentru **Data Labels**
- ✓ Pasul 3 – Selectăm opțiunea **More Options**

Pentru opțiunea **Format Data Labels** elementele ce trebuie completate sunt detaliate în figura 16, iar forma finală a graficului pentru variabilele calitative procesate este prezentată în figura 17.

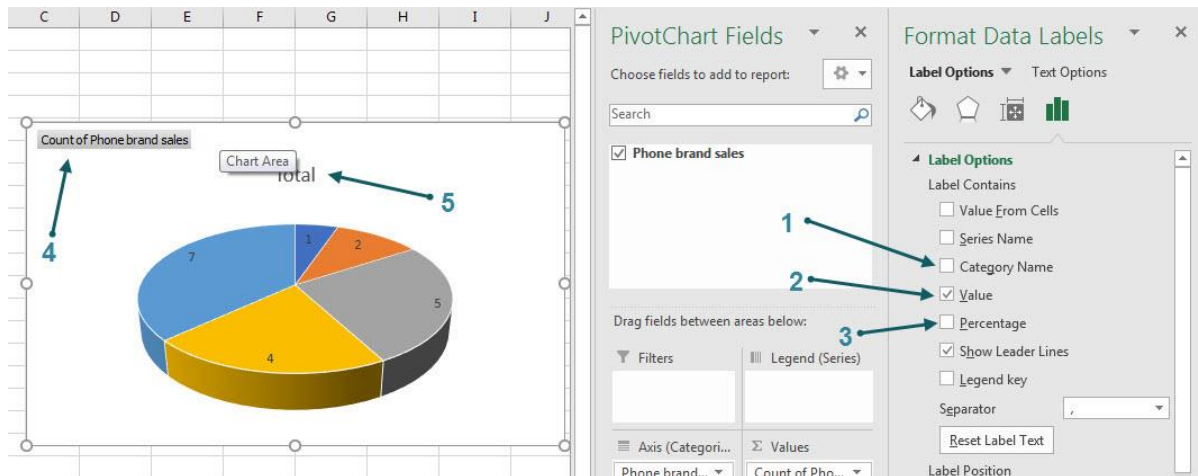


Figura 16 – Opțiunea **Format Data Labels**

- ✓ Pasul 1 – Bifați opțiunea pentru **Category Name**
 - ✓ Pasul 2 – Debifați opțiunea pentru **Values**
 - ✓ Pasul 3 – Bifați opțiunea pentru **Percentage**
- ✓ Pasul 4 – Click dreapta pe Count of Phone brand sales și alegeți Hide all Field Buttons on Chart
- ✓ Pasul 5 – Redenumiți graficul ca Phone sales distribution

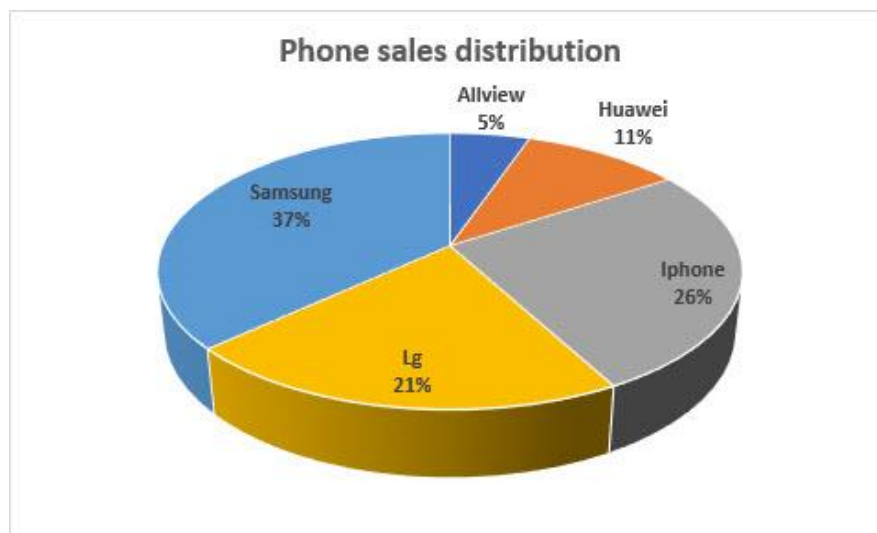


Figure 17 – Forma finală a graficului pentru variabilele calitative procesate

3. Variabile cantitative continue – histogramă

Pentru a folosi opțiunea **Histogram**, este necesar să instalăm **Analysis Tool Pack** pentru MS Excel. Mergem la **File** și alegem **Options**, așa cum este prezentat în figurile 18, 19 și 20.



Figura 18 – Din meniul **File** selectăm **Options**

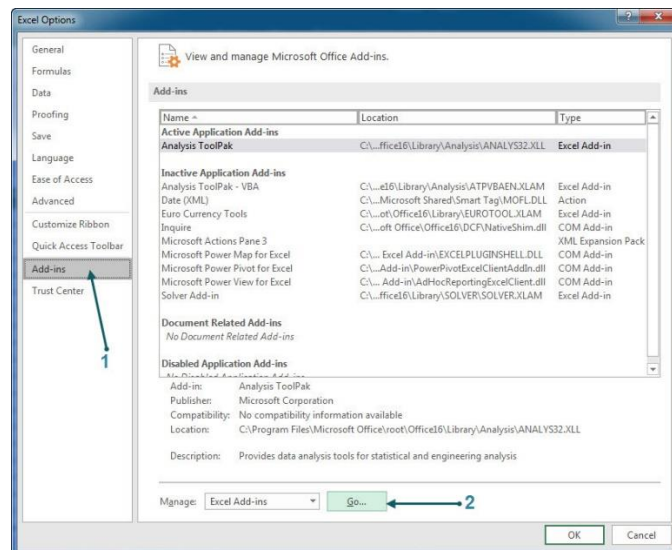


Figura 20 – Selectăm **Add Ins** – **Excel Add Ins** – **Go**

- ✓ Pasul 1 – Secetăm opțiunea **Add Ins**
- ✓ Pasul 2 – Apăsăm butonul **Go** pentru **Excel Add Ins**

În figura 21 avem datele pentru greutatea la naștere (weight at birth) și intervalele de lucru (Bin Range) – aceste valori reprezintă capetele intervalelor ce vor fi folosite pentru procesarea datelor. Alegem din meniul **Data – Data Analysis – Histogram**.

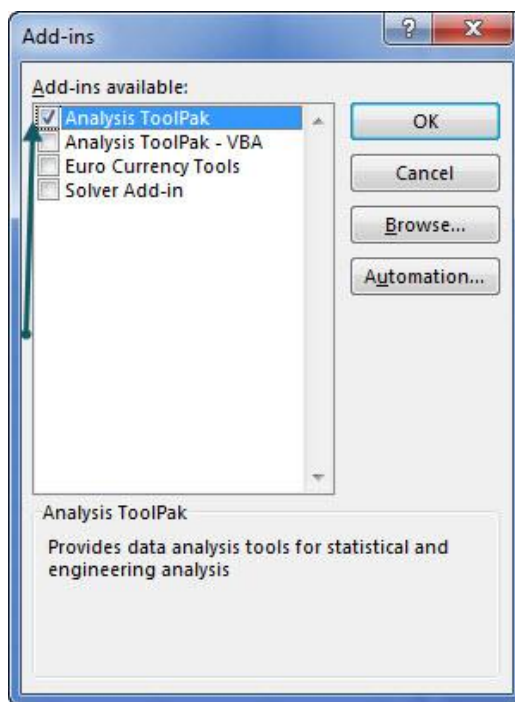


Figura 20 – Bifăm opțiunea pentru **Analysis Tool Pack**

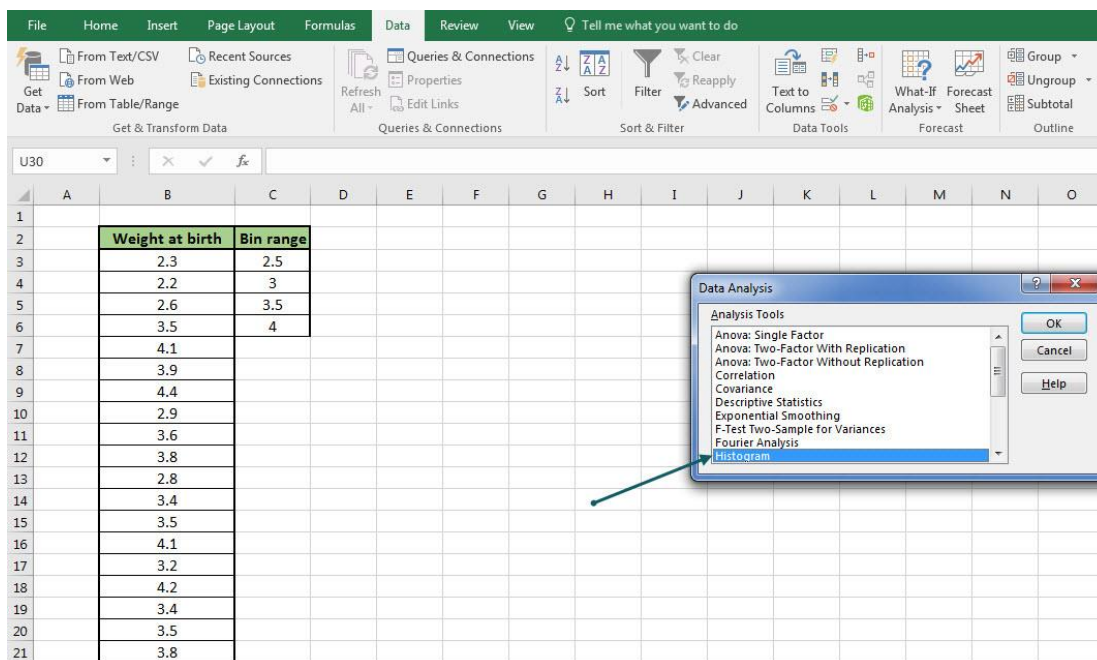


Figura 21 – Alegerea opțiunii Histogram (din meniul Data – Data Analysis)

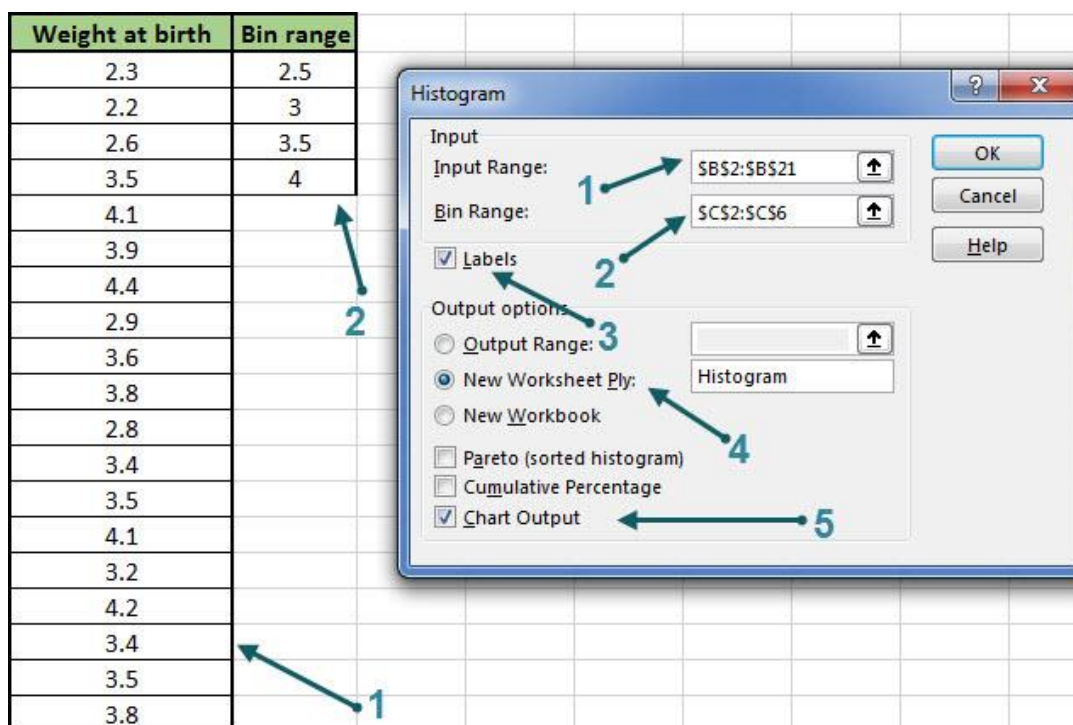


Figura 22 – Selecțiile necesare pentru definirea Histogram

- ✓ Pasul 1 – Selectăm **Input Range**
- ✓ Pasul 2 – Selectăm **Bin Range**
- ✓ Pasul 3 – Bifăm opțiunea **Labels**
- ✓ Pasul 4 – Selectăm **New Worksheet Play** și o numim **Histogram**
- ✓ Pasul 5 – Bifăm opțiunea **Chart Output**

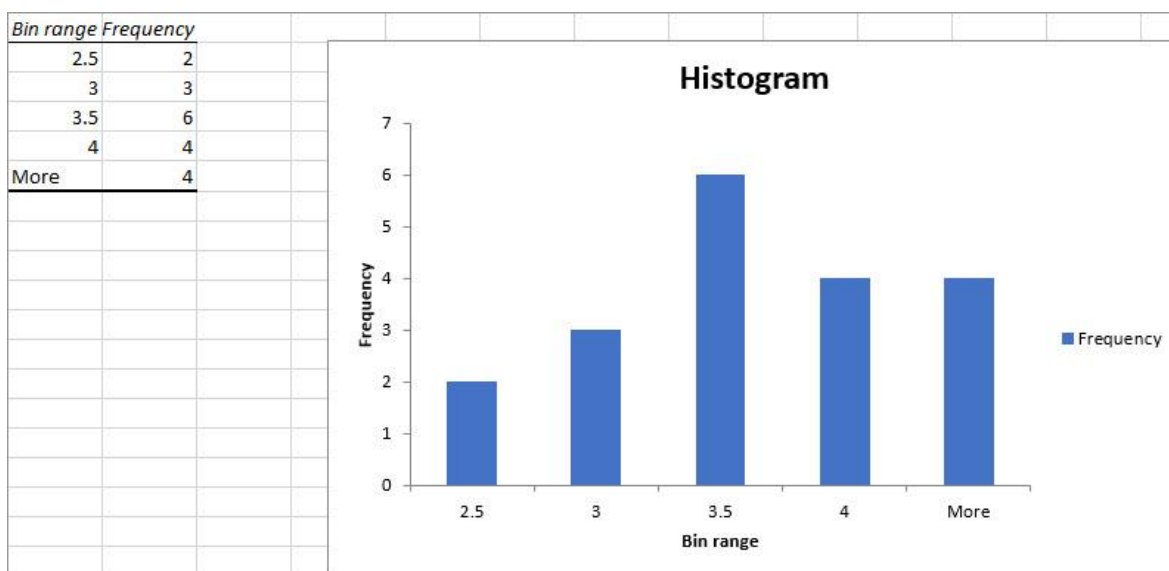


Figura 23 – Rezultatul obținut pentru Histogram: grafic și valori

Rezultatele obținute în histogramă sunt pentru intervalele ce au fost definite inițial (bin range). Pentru definirea clară a intervalelor vom folosi valorile prezentate în figura 24. Pentru a înlocui valorile din grafic cu valorile nou definite vom alege opțiunea **Select Data** - figura 25.

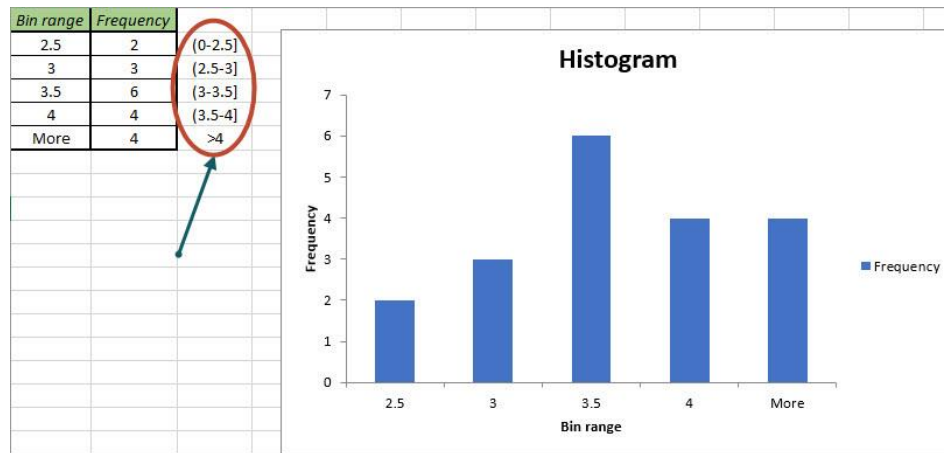


Figura 24 – Definirea intervalelor ce vor fi afișate pe axa x

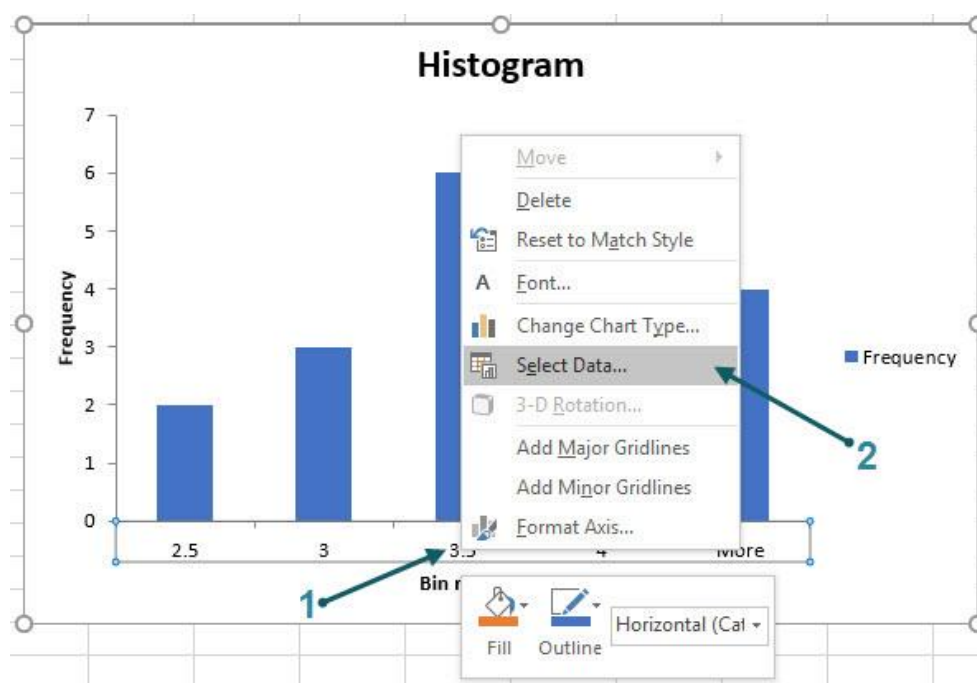


Figura 25 – Editarea elementelor ce sunt afișate pe axa x

- ✓ Pasul 1 – Click dreapta pe oricare din valorile afișate pe axa x
- ✓ Pasul 2 – Selectăm opțiunea **Select Data**

Pentru a modifica datele afișate vom alege butonul **Edit** din zona **Horizontal Axis Labels** după cum este prezentat în figura 26 și apoi vom alege datele scrise anterior așa cum este detaliat în figura 27.

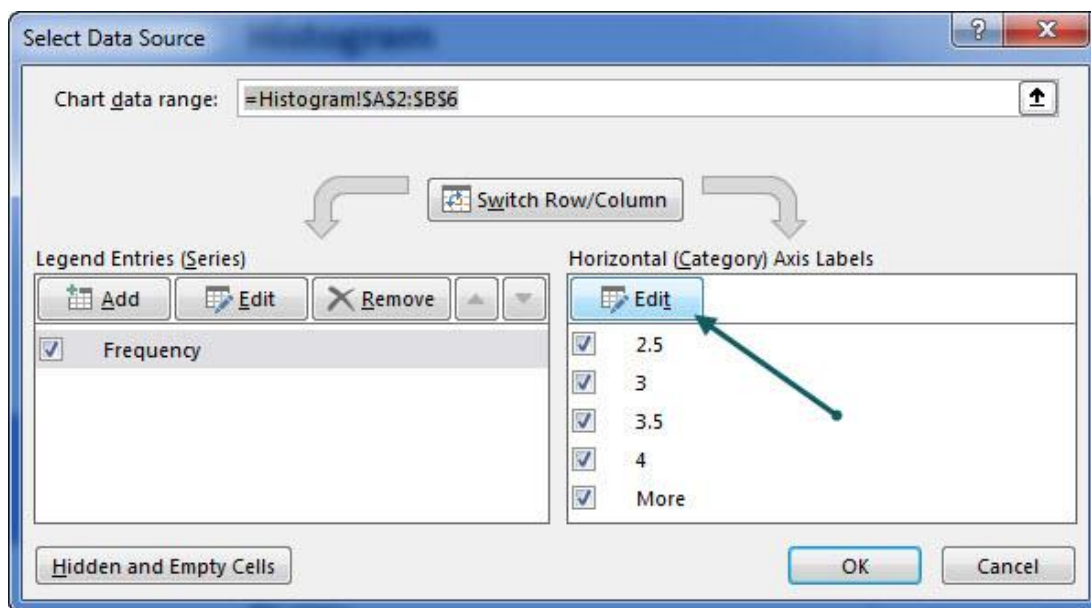


Figura 26 – Editarea datelor pentru axa x (Horizontal Axis Labels)

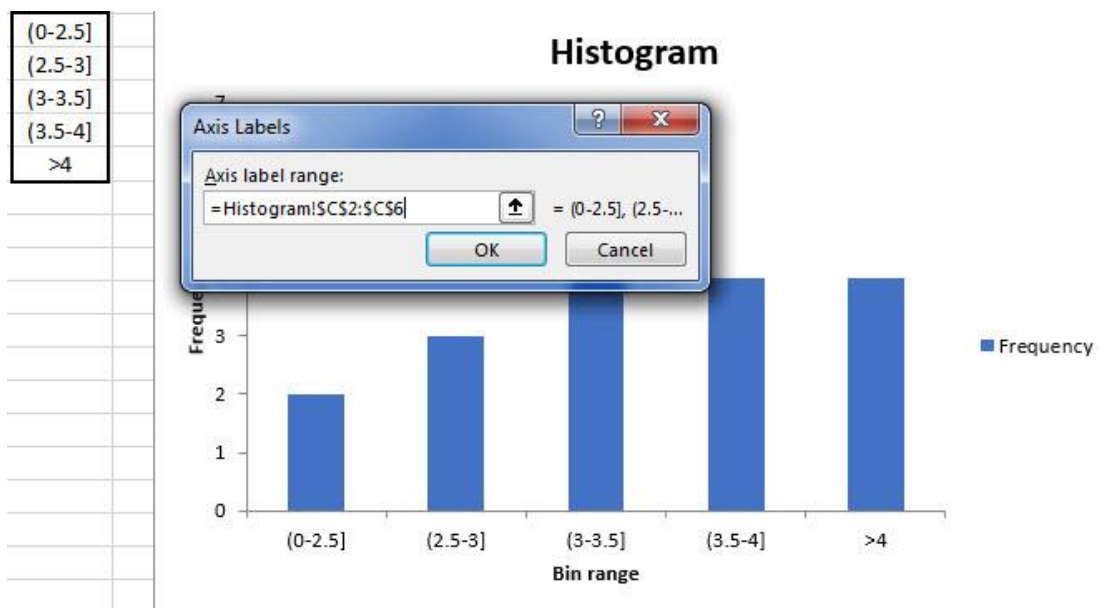


Figura 27 – Selectarea noilor date ce trebuie afișate pe axa x

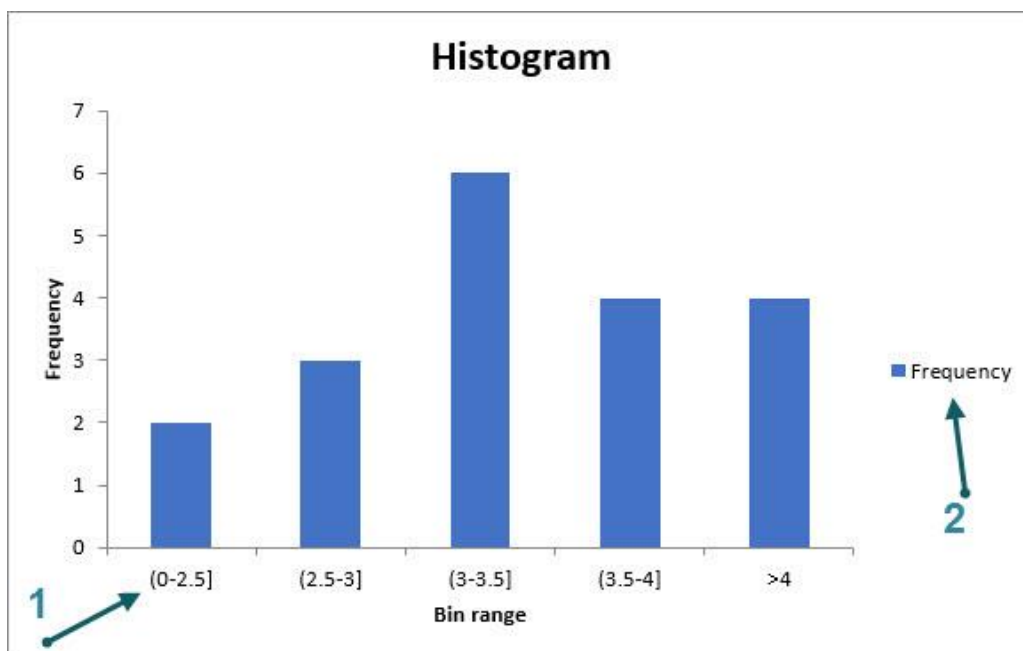


Figura 28 – Finalizarea editării histogramei

- ✓ Pasul 1 – Verificăm corectitudinea datelor ce sunt acum afișate pe axa x
- ✓ Pasul 2 – Șterfem legenda graficului

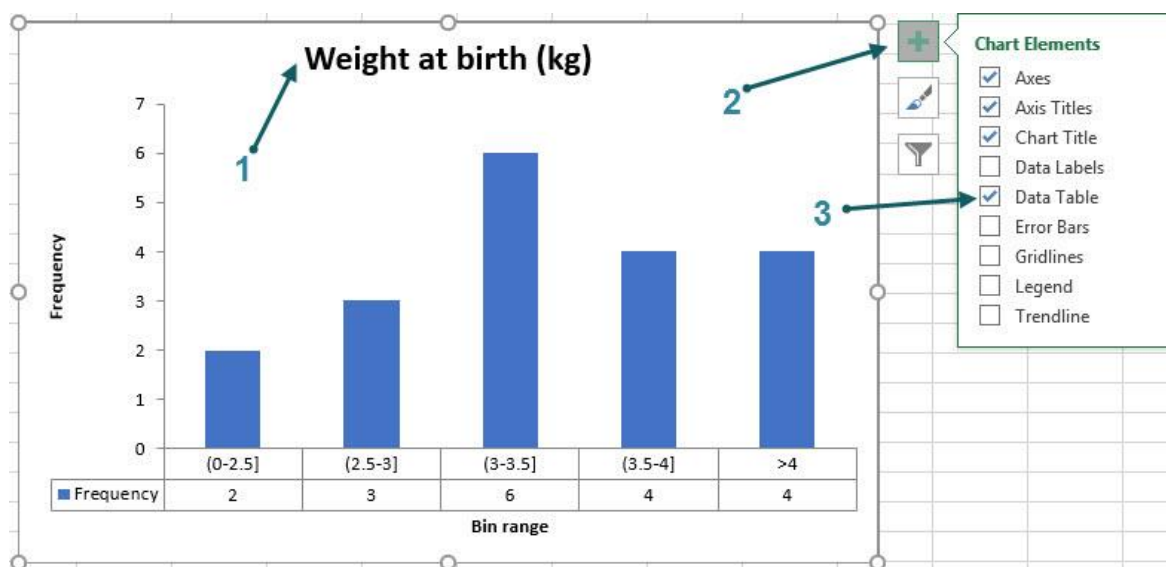


Figura 29 – Forma finală a histogramei

- ✓ Pasul 1 – Scimbăm titlul histogramei
- ✓ Pasul 2 – Apăsăm pe butonul Chart Elements
- ✓ Pasul 3 – Bifăm opțiunea pentru **Data Table**

4. Deviația standard și intervalul de încredere pentru 95 % – statistica descriptivă

Considerând datele prezentate în figura 30 vom calcula deviația standard și intervalul de încredere pentru 95 % folosind **Data analysis – Descriptive statistics**. Se vor putea determina astfel și mediana, moda, media aritmetică și alte elemente statistice. În final vom verifica și unul din rezultate folosind calcule matematice.

Age	Weight	BMI	Sys	Dya	Blood Sugar	Cholesterol
59	95	33	140	100	100	210
68	85	35	150	100	103	327
70	54	22	160	80	99	281
29	74	26	110	60	84	174
29	61	24	120	70	82	223
52	82	23	120	80	72	183
43	67	25	130	80	89	183
47	86	29	140	100	80	211
30	69	28	110	50	76	179
47	107	33	130	90	108	244
41	84	25	110	80	85	195
41	104	34	110	70	122	262
60	60	24	120	70	80	277
67	74	28	160	90	93	177
73	61	26	160	80	95	223
68	77	26	140	80	104	185
49	109	38	160	100	89	171
50	88	32	130	90	123	169
40	64	20	120	80	68	148
48	78	26	140	80	93	133
38	60	22	90	40	73	179
44	108	37	140	100	89	120
26	75	25	110	60	89	118
47	87	27	120	80	87	209
26	96	31	130	90	81	210
29	83	27	120	70	85	202
33	83	27	100	70	71	174
41	81	29	120	90	90	183
52	73	26	140	100	88	141
43	90	34	100	70	82	210

Figura 30 – Datele inițiale ce urmează a fi procesate

Din meniul **Data – Data Analysis** vom alege **Descriptive Statistics** (figura 31). La **Input Range** vom selecta tot setul de date și vom bifa **Labels in First Row, New Worksheet Ply, Summary Statistics** și **Confidence Level for Mean at 95 %** (figura 32).

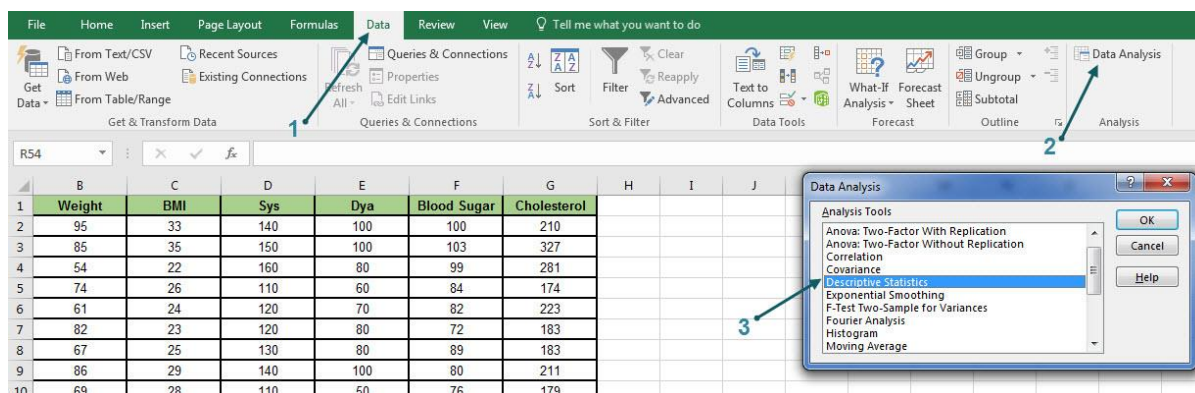


Figura 31 – Selectarea **Data Analysis – Descriptive Statistics**

- ✓ Pasul 1 – Selectăm **Data**
- ✓ Pasul 2 – Alegem **Data Analysis**
- ✓ Pasul 3 – Alegem **Descriptive Statistics**

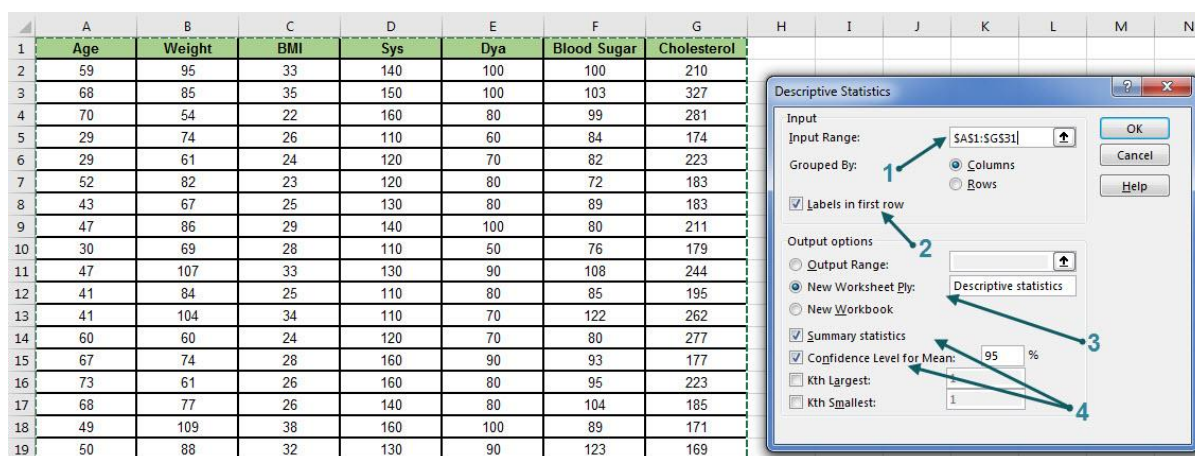


Figura 32 – Selecțiile necesare pentru **Descriptive Statistics**

- ✓ Pasul 1 – La **Input Range** selectăm toate datele
- ✓ Pasul 2 – Bifăm opțiunea **Labels in first row**
- ✓ Pasul 3 – Selectăm **New Worksheet Ply** și pentru nume scriem **Descriptive statistics**
- ✓ Pasul 4 – Bifăm opțiunile pentru **Summary statistics** și **Confidence Level for Mean 95 %**

Age		Weight		BMI	
Mean	46.33333	Mean	80.5	Mean	28.06667
Standard Error	2.507155	Standard Error	2.758373	Standard Error	0.845452
Median	45.5	Median	81.5	Median	27
Mode	29	Mode	74	Mode	26
Standard Deviation	13.73225	Standard Deviation	15.10823	Standard Deviation	4.630732
Sample Variance	188.5747	Sample Variance	228.2586	Sample Variance	21.44368
Range	47	Range	55	Range	18
Minimum	26	Minimum	54	Minimum	20
Maximum	73	Maximum	109	Maximum	38
Sum	1390	Sum	2415	Sum	842
Count	30	Count	30	Count	30
Confidence Level(95.0%)	5.127707	Confidence Level(95.0%)	5.641506	Confidence Level(95.0%)	1.729144

Sys		Dya	
Mean	127.6667	Mean	80
Standard Error	3.480102	Standard Error	2.795727
Median	125	Median	80
Mode	120	Mode	80
Standard Deviation	19.0613	Standard Deviation	15.31283
Sample Variance	363.3333	Sample Variance	234.4828
Range	70	Range	60
Minimum	90	Minimum	40
Maximum	160	Maximum	100
Sum	3830	Sum	2400
Count	30	Count	30
Confidence Level(95.0%)	7.117608	Confidence Level(95.0%)	5.717905

Blood Sugar		Cholesterol	
Mean	89.33333	Mean	196.7
Standard Error	2.439773	Standard Error	8.651377
Median	88.5	Median	184
Mode	89	Mode	210
Standard Deviation	13.36318	Standard Deviation	47.38554
Sample Variance	178.5747	Sample Variance	2245.39
Range	55	Range	209
Minimum	68	Minimum	118
Maximum	123	Maximum	327
Sum	2680	Sum	5901
Count	30	Count	30
Confidence Level(95.0%)	4.989895	Confidence Level(95.0%)	17.69405

Figura 33 – Descriptive statistics – rezultatele obținute

Vom verifica rezultatele obținute pentru parametrul Age (vârstă). Vom calcula media aritmetică, deviația standard și intervalul de încredere pentru 95 %. Formula pentru calculul deviației standard este prezentată mai jos (1).

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

$$\begin{aligned} \bar{X} = & (59 + 68 + 70 + 29 + 29 + 52 + 43 + 47 + 30 + 47 + 41 + 41 + 60 + 67 + 73 + 68 \\ & + 49 + 50 + 40 + 48 + 38 + 44 + 26 + 47 + 26 + 29 + 33 + 41 + 52 \\ & + 43)/30 = 46.33 \end{aligned}$$

$$\begin{aligned} S = \sqrt{((59 - 46.33)^2 + (68 - 46.33)^2 + (70 - 46.33)^2 + (29 - 46.33)^2 + (29 - 46.33)^2 \\ + (52 - 46.33)^2 + (43 - 46.33)^2 + (47 - 46.33)^2 + (30 - 46.33)^2 \\ + (47 - 46.33)^2 + (41 - 46.33)^2 + (41 - 46.33)^2 + (60 - 46.33)^2 \\ + (67 - 46.33)^2 + (73 - 46.33)^2 + (68 - 46.33)^2 + (49 - 46.33)^2 \\ + (50 - 46.33)^2 + (40 - 46.33)^2 + (48 - 46.33)^2 + (38 - 46.33)^2 \\ + (44 - 46.33)^2 + (26 - 46.33)^2 + (47 - 46.33)^2 + (26 - 46.33)^2 \\ + (29 - 46.33)^2 + (33 - 46.33)^2 + (41 - 46.33)^2 + (52 - 46.33)^2 \\ + (43 - 46.33)^2)/29)} = 13.73 \end{aligned}$$

Formula pentru calculul intervalului de încredere la 95 % este (2):

$$\bar{X} \pm Z \frac{S}{\sqrt{n}} \quad (2)$$

Valorile pentru z sunt prezentate în tabelul 1. Este indicat pentru estimarea constantei să se țină cont și de eșantion (tabel 2)

, astfel putem spune cu 95 % încredere ca a 31-a persoană, considerând datele din tabelul inițial, va avea o vârstă de:

$$46.33 \mp 1.960 \frac{13.73}{\sqrt{30}}$$

$$46.33 \mp 4.91 \text{ years}$$

În MS Excel intervalul de încredere pentru 95 % este estimat cu funcția CONFIDENCE.T.

Confidence interval	z
0.8	1.282
0.85	1.44
0.9	1.645
0.95	1.96
0.99	2.576
0.995	2.807
0.999	3.291

Tabel 1 – Intervale de încredere pentru z

DF	A	0.8	0.9	0.95	0.98	0.99	0.995	0.998	0.999
	P	0.2	0.1	0.05	0.02	0.01	0.005	0.002	0.001
1		3.078	6.314	12.706	31.82	63.657	127.321	318.309	636.619
2		1.886	2.92	4.303	6.965	9.925	14.089	22.327	31.599
3		1.638	2.353	3.182	4.541	5.841	7.453	10.215	12.924
4		1.533	2.132	2.776	3.747	4.604	5.598	7.173	8.61
5		1.476	2.015	2.571	3.365	4.032	4.773	5.893	6.869
6		1.44	1.943	2.447	3.143	3.707	4.317	5.208	5.959
7		1.415	1.895	2.365	2.998	3.499	4.029	4.785	5.408
8		1.397	1.86	2.306	2.897	3.355	3.833	4.501	5.041
9		1.383	1.833	2.262	2.821	3.25	3.69	4.297	4.781
10		1.372	1.812	2.228	2.764	3.169	3.581	4.144	4.587
11		1.363	1.796	2.201	2.718	3.106	3.497	4.025	4.437
12		1.356	1.782	2.179	2.681	3.055	3.428	3.93	4.318
13		1.35	1.771	2.16	2.65	3.012	3.372	3.852	4.221
14		1.345	1.761	2.145	2.625	2.977	3.326	3.787	4.14
15		1.341	1.753	2.131	2.602	2.947	3.286	3.733	4.073
16		1.337	1.746	2.12	2.584	2.921	3.252	3.686	4.015
17		1.333	1.74	2.11	2.567	2.898	3.222	3.646	3.965
18		1.33	1.734	2.101	2.552	2.878	3.197	3.61	3.922
19		1.328	1.729	2.093	2.539	2.861	3.174	3.579	3.883
20		1.325	1.725	2.086	2.528	2.845	3.153	3.552	3.85
21		1.323	1.721	2.08	2.518	2.831	3.135	3.527	3.819
22		1.321	1.717	2.074	2.508	2.819	3.119	3.505	3.792
23		1.319	1.714	2.069	2.5	2.807	3.104	3.485	3.768
24		1.318	1.711	2.064	2.492	2.797	3.09	3.467	3.745
25		1.316	1.708	2.06	2.485	2.787	3.078	3.45	3.725
26		1.315	1.706	2.056	2.479	2.779	3.067	3.435	3.707
27		1.314	1.703	2.052	2.473	2.771	3.057	3.421	3.69
28		1.313	1.701	2.048	2.467	2.763	3.047	3.408	3.674
29		1.311	1.699	2.045	2.462	2.756	3.038	3.396	3.659
30		1.31	1.697	2.042	2.457	2.75	3.03	3.385	3.646
31		1.309	1.695	2.04	2.453	2.744	3.022	3.375	3.633
32		1.309	1.694	2.037	2.449	2.738	3.015	3.365	3.622
33		1.308	1.692	2.035	2.445	2.733	3.008	3.356	3.611
34		1.307	1.691	2.032	2.441	2.728	3.002	3.348	3.601
35		1.306	1.69	2.03	2.438	2.724	2.996	3.34	3.591
36		1.306	1.688	2.028	2.434	2.719	2.991	3.333	3.582
37		1.305	1.687	2.026	2.431	2.715	2.985	3.326	3.574
38		1.304	1.686	2.024	2.429	2.712	2.98	3.319	3.566
39		1.304	1.685	2.023	2.426	2.708	2.976	3.313	3.558
40		1.303	1.684	2.021	2.423	2.704	2.971	3.307	3.551
42		1.302	1.682	2.018	2.418	2.698	2.963	3.296	3.538
44		1.301	1.68	2.015	2.414	2.692	2.956	3.286	3.526
46		1.3	1.679	2.013	2.41	2.687	2.949	3.277	3.515
48		1.299	1.677	2.011	2.407	2.682	2.943	3.269	3.505
50		1.299	1.676	2.009	2.403	2.678	2.937	3.261	3.496
60		1.296	1.671	2	2.39	2.66	2.915	3.232	3.46
70		1.294	1.667	1.994	2.381	2.648	2.899	3.211	3.435
80		1.292	1.664	1.99	2.374	2.639	2.887	3.195	3.416
90		1.291	1.662	1.987	2.369	2.632	2.878	3.183	3.402
100		1.29	1.66	1.984	2.364	2.626	2.871	3.174	3.391
120		1.289	1.658	1.98	2.358	2.617	2.86	3.16	3.373
150		1.287	1.655	1.976	2.351	2.609	2.849	3.145	3.357
200		1.286	1.652	1.972	2.345	2.601	2.839	3.131	3.34
300		1.284	1.65	1.968	2.339	2.592	2.828	3.118	3.323
500		1.283	1.648	1.965	2.334	2.586	2.82	3.107	3.31
∞		1.282	1.645	1.96	2.326	2.576	2.807	3.09	3.291

Tabel 2 – Intervale de încredere pentru t