

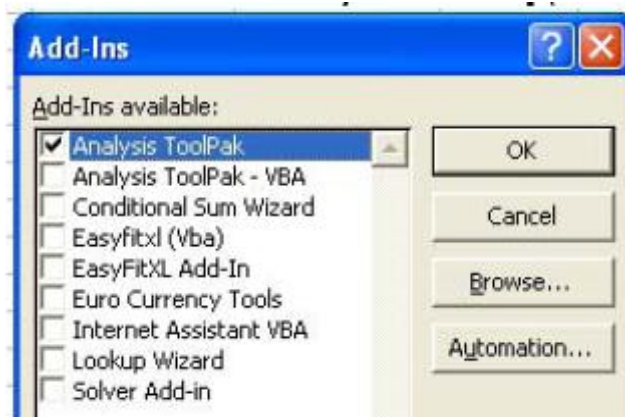
MICROSOFT EXCEL: STATISTICĂ DESCRIPTIVĂ, INDICAȚII 3

Pentru a crea o histogramă (o reprezentare grafică a frecvenței distribuției):

- Creați în foaia numită Histogramă tabelul de frecvențe și limitele maxime pentru fiecare interval:

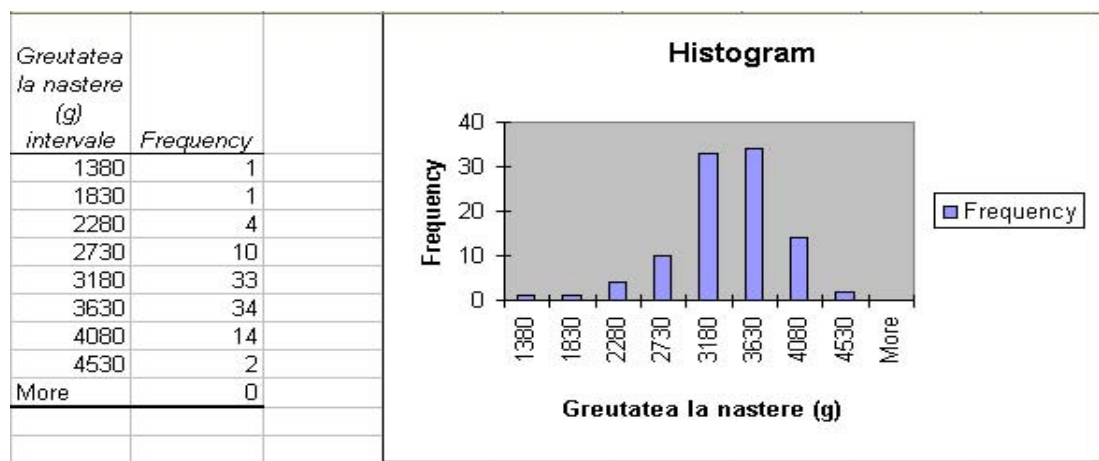
F	G	H
Greutatea la naștere (g) intervale	Hemoglobină (mg/dl) la 6 luni intervale	Hemoglobină (mg/dl) la 12 luni intervale
1380	8,8	8,9
1830	9,6	9,8
2280	10,4	10,7
2730	11,2	11,6
3180	12	12,5
3630	12,8	13,4
4080	13,6	14,3
4530		

- **[Tools – Data Analysis - Histogram]**
- Pentru a avea opțiunea **Data Analysis**, verificați că este validată opțiunea **[Tools – Add Ins.. – Analysis ToolPak]** (vezi imaginea de mai jos)



- Creați histograma: **[Tools – Data Analysis - Histogram]**. În fereastra de dialog a histogramei completați câmpurile ca mai jos:
 - Input Range: selectați seria de date pentru care se calculează histograma (deci pentru Greutatea la naștere selectați \$A\$1:\$A\$100)
 - Bin Range: selectați celulele în care ați completat limita maximă a intervalelor de frecvențe (deci pentru Greutatea la naștere selectați \$F\$1:\$F\$9)
 - Validați opțiunea Labels (deoarece ați selectat și numele variabilei)
 - Output range: selectați celula unde va fi colțul din stânga-sus al tabelului (de ex. \$J\$1).
 - Validați opțiunea Chart Output (pentru a genera și un graphic pe lângă tabelul cu frecvențe).
 - Ca rezultat veți avea ceva similar cu imaginea de mai jos.

Observație. Graficul de mai jos este de tip “coloane”, în timp ce într-o Histogramă nu există spații între coloane.

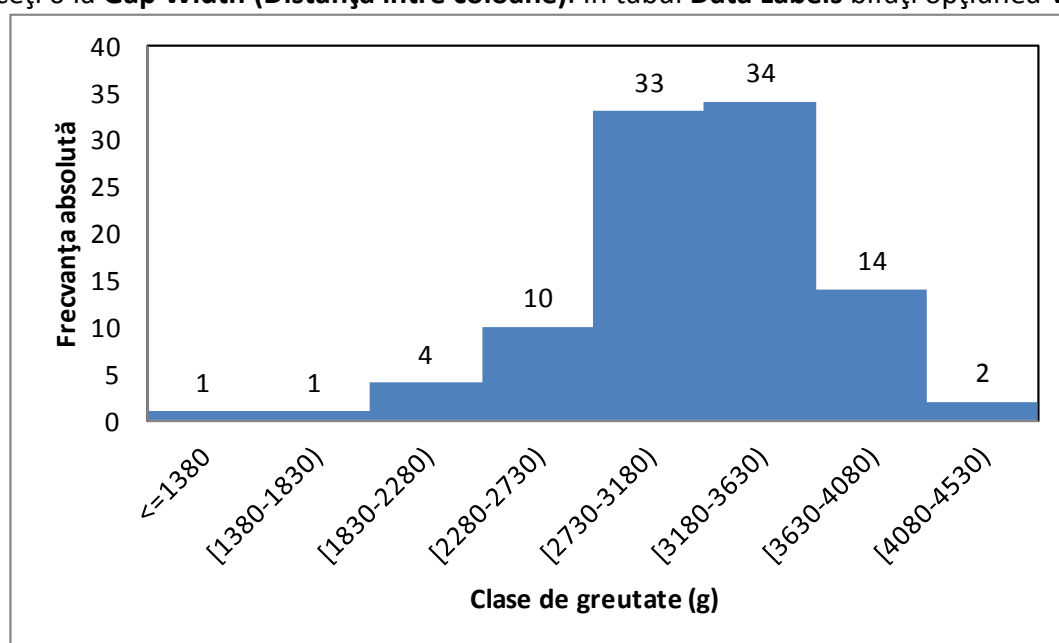


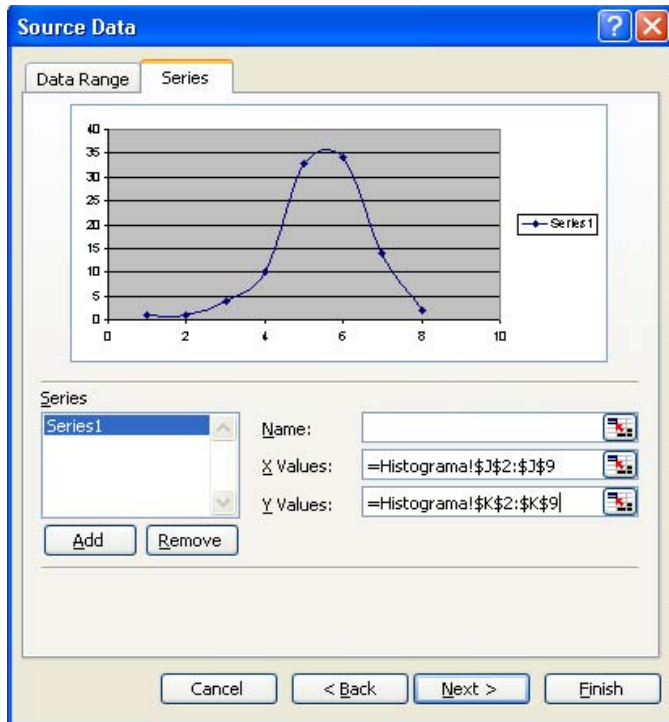
Modificați coloana Greutatea la naștere scriind intervalele între care se încadrează greutatea la naștere, ca mai jos:

Birth weight (g)	Frequency
<=1380	1
[1380-1830)	1
[1830-2280)	4
[2280-2730)	10
[2730-3180)	33
[3180-3630)	34
[3630-4080)	14
[4080-4530)	2

Modificați graficul pentru a arăta ca și în imaginea de mai jos.

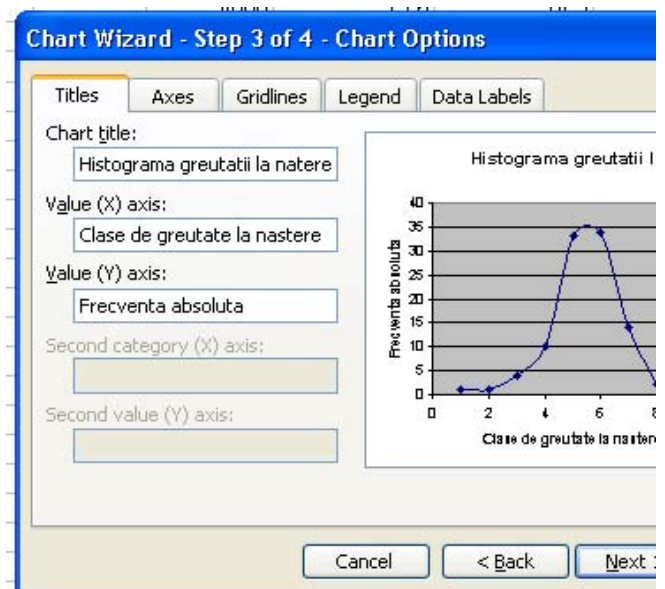
Pentru aceasta, dublu-clic pe o coloană, și în fereastra **Format Data Series**, tabul **Options**, introduceți 0 la **Gap Width (Distanța între coloane)**. În tabul **Data Labels** bifați opțiunea **Values**.



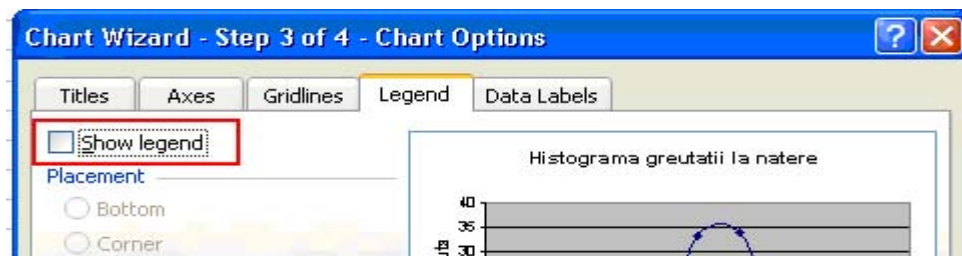


Dacă doriți să obțineți un grafic similar cu cel gaussian, (dacă vă este mai ușor să-l interpretați) creați-l folosind tabelul de frecvențe, urmând pașii următori:

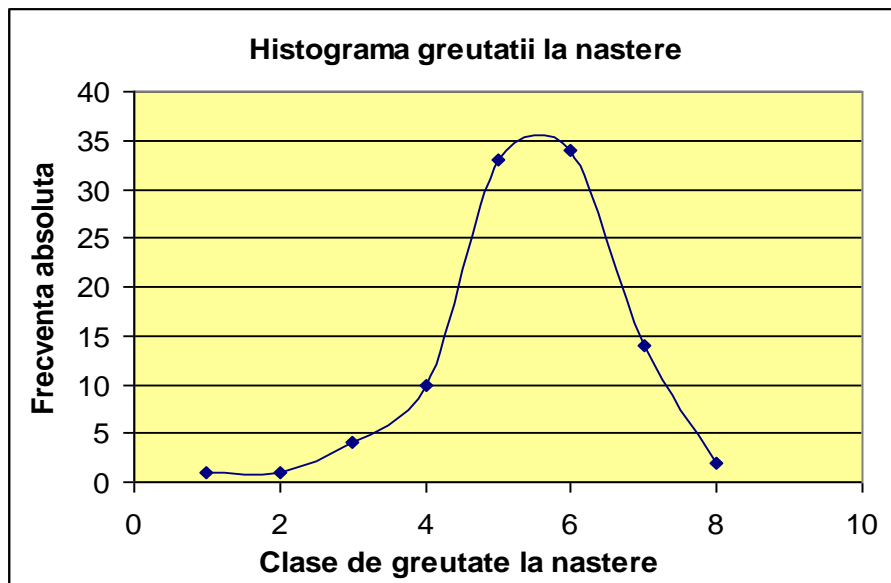
- **[Insert – Chart – Scatter – Second sub-Type]**
- La pasul 2, **Chart Source Data**, în tabul **Series**, secțiunea **Series** clic pe **Add**
- Completați **X Value** – valorile pe axa Ox, cu limitele maxime ale intervalelor și **Y Values** – valorile pe axa Oy, cu frecvențele care se încadrează între limitele respective ca în exemplul de mai jos (exemplul este pentru variabila Greutatea la naștere):
- Introduceți titlul și denumirea axelor



- Ștergeți legenda



- Graficul obținut va fi similar cu cel din imaginea de mai jos:



Observație. Este necesar ca în partea de jos a graficului să precizați semnificația claselor de greutate la naștere.

Punctul 6. Pentru a calcula parametrii de statistică descriptivă:

- **[Tools – Data Analysis – Descriptive statistics]**
- În fereastra de dialog **Descriptive Statistics**:
 - 1. La **Input Range**: selectați aria în care se află datele, inclusive capul de table (pentru cerința noastră, datele cerute sunt în coloanele Greutatea la naștere și Hemoglobina la 6 și la 12 luni).
 - 2. Precizați că aveți capul de tabel (**Labels**) în primul rând al coloanelor.
 - 3. La secțiunea **Output options, Output range**: precizați prima celulă în care se va afișa tabelul rezultat (lăsați tabelul în aceeași foaie cu datele).
 - 4. Validați opțiunile **Summary Statistics** pentru a obține funcțiile de statistică descriptivă și **Confidence Level for Mean**, pentru a obține nivelul de încredere pentru medie de 95%.

The screenshot shows an Excel spreadsheet with the following data:

Greutatea la naștere (g)	Hemoglobină (mg/dl) la 6 luni	Hemoglobină (mg/dl) la 12 luni
3650	12,10	13,20
2600	13,20	13,20
3000	10,10	13,30
3150	9,20	12,10
3700	10,60	12,20
3800	12,30	13,00
3000	12,30	12,90
3000	11,40	12,90
2600	10,50	11,20
2200	13,10	12,70
3200	10,20	12,70
3400	13,00	13,40
3200	12,10	12,90
3300	13,10	13,60
3000	11,50	12,90
3000	13,10	12,70
3200	11,50	12,00

The 'Descriptive Statistics' dialog box is open with the following settings:

- Input Range:** \$B\$1:\$D\$100
- Grouped By:** Columns
- Labels in first row:**
- Output Range:** \$G\$1
- Summary statistics:**
- Confidence Level for Mean:** 95%

- Tabelul rezultat va fi similar cu cel de mai jos, și conține o parte dintre funcțiile calculate la laboratorul 2.

G	H	I	J	K	L
		<i>Hemoglo bina (mg/dl) la 6 luni</i>		<i>Hemoglo bina (mg/dl) la 12 luni</i>	
<i>Greutatea la nastere (g)</i>					
Mean	3143,636	Mean	11,01515	Mean	11,79798
Standard Error	53,66345	Standard E	0,108373	Standard E	0,116686
Median	3200	Median	11,1	Median	11,9
Mode	3000	Mode	11,4	Mode	11,8
Standard Deviation	533,9446	Standard D	1,078298	Standard D	1,161015
Sample Variance	285096,8	Sample Va	1,162727	Sample Va	1,347955
Kurtosis	2,553056	Kurtosis	-0,4873	Kurtosis	1,15706
Skewness	-0,85169	Skewness	-0,07276	Skewness	-0,84112
Range	3470	Range	4,9	Range	6,3
Minimum	930	Minimum	8,3	Minimum	7,8
Maximum	4400	Maximum	13,2	Maximum	14,1
Sum	311220	Sum	1090,5	Sum	1168
Count	99	Count	99	Count	99
Confidence Level(95,0%)	106,4934	Confidence	0,215063	Confidence	0,23156

- Mutați numele variabilelor din capul de tabel cu o coloană spre dreapta și ștergeți coloanele I și K (conțin aceeași informație ca și coloana G). Formatați celulele care conțin numere cu zecimale să arate 2 zecimale. Rezultatul va arăta ca în imaginea de mai jos:

G	H	I	J
	Greutatea la naștere (g)	Hemoglobina (mg/dl) la 6 luni	Hemoglobina (mg/dl) la 12 luni
Mean	3143.64	11.02	11.80
Standard Error	53.66	0.11	0.12
Median	3200.00	11.10	11.90
Mode	3000.00	11.40	11.80
Standard Deviation	533.94	1.08	1.16
Sample Variance	285096.85	1.16	1.35
Kurtosis	2.55	-0.49	1.16
Skewness	-0.85	-0.07	-0.84
Range	3470.00	4.90	6.30
Minimum	930.00	8.30	7.80
Maximum	4400.00	13.20	14.10
Sum	311220.00	1090.50	1168.00
Count	99.00	99.00	99.00
Confidence Level(95.0%)	106.49	0.22	0.23

- Interpretarea prin exemple (pentru variabila Greutate la naștere):

<i>Media</i>	Media aritmetică pentru cei 99 copii nou-născuți incluși în studiu este de 3143.63 grame.
<i>Eroarea Standard</i>	Eroarea Standard pentru medie, pentru greutatea la naștere este de 53.66 g.
<i>Mediana</i>	Valoarea greutății la naștere din mijlocul distribuției greutății este de 3200 grame.
<i>Valoarea modală (Modulul)</i>	Valoarea greutății la naștere care are cea mai mare frecvență în studiul nostru este de 3000 grame.
<i>Deviația Standard</i>	Deviația Standard a populației pentru greutatea la naștere este egală cu 533.94.
<i>Varianța</i>	Pătratul deviației standard pentru greutatea la naștere este egală cu 285096.85.
<i>Excesul de boltire (Kurtosis)</i>	Distribuția greutății la naștere are o boltire mai mare (graficul distribuției are un vârf mai înalt comparativ cu distribuția normală) comparativ cu distribuția normală, kurtosis fiind egal cu 2,55. Observație Dacă valoarea kurtosis-ului este în intervalul [-0.5, +0.5], se poate considera că datele au o distribuție normală a boltirii, apropiată de cea gaussiană.
<i>Simetria (Skewness)</i>	Valoarea negativă de -0.85 a simetriei, pentru datele pacienților din studiul nostru, indică că distribuția greutății la naștere este asimetrică în sens negativ. Asimetria negativă indică că o graficul are o prelungire mai accentuată în direcția valorilor scăzute ale distribuției. Observație Dacă asimetria se încadrează în intervalul [-1, 1] se poate considera că datele urmează o simetrie normală.
<i>Amplitudinea (Range)</i>	Amplitudinea pentru distribuția noastră se obține scăzând minimul 930 din maximumul 4400, și se obține o amplitudine egală cu 3470.
<i>Minimumul</i>	Cea mai mică valoare a greutății la naștere la copiii nou-născuți cuprinși în studiu este de 930 g.
<i>Maximumul</i>	Cea mare valoare a greutății la naștere la copiii nou-născuți cuprinși în studiu este de 4400 g.
<i>Suma</i>	Suma valorilor greutăților din distribuție pentru copiii cuprinși în studiu este de 311220 g.
<i>n (Count)</i>	Numărul de observații în distribuția greutății la naștere din eșantionul studiat este n = 99
<i>Nivelul de încredere (95.0%)</i>	Valoarea obținută reprezintă reprezentă estimarea erorii care se scade și se adaugă mediei eșantionului când se calculează intervalul de încredere pentru media populației.

Pentru problema noastră, intervalul de încredere de 95% este: $3037.14 \leq \mu \leq 3250.13$

Punctul 7. Pentru a calcula intervalul de încredere de 95% pentru medie:

- Introduceți următoarele informații mai jos de rezultatele statisticii descriptive:

Count	99	99	99
Confidence Level(95,0%)	106,493372	0,215062803	0,231560287

95% limita superioară			
95% limita inferioară			

- Definiția formulelor:
 - Limita inferioară este egală cu media aritmetică din care se scade Nivelul de încredere (95%)
 - Limita superioară este egală cu media aritmetică la care se adaugă Nivelul de încredere (95%)

Punctele 7 și 8. Pentru a calcula parametrii de statistică descriptivă pentru diferite grupuri:

- Ordonăți datele crescător **după** variabila **Schema de tratament: [Data – Sort – Sort by – Ascending]**. La Sort by verificați să aveți variabila respectivă.
- Calculați parametrii de statistică descriptivă pentru pacienți tratați bi-săptămânal **și din mediul rural**.
- Ordonăți datele descrescător **după** variabila **Schema de tratament: [Data – Sort – Sort by – Descending]**.
- Calculați parametrii de statistică descriptivă pentru pacienți tratați zilnic.
- Calculați limita superioară și cea inferioară pentru pentru intervalele de încredere 95% pentru toate cazurile.
- Rezultatele vor arăta ca în imaginea de mai jos:

	<i>Greutatea la nastere (g)</i>		<i>Hemoglobina (mg/dl) la 6 luni</i>		<i>Hemoglobina (mg/dl) la 12 luni</i>	
	<i>bisaptaminal</i>	<i>zilnic</i>	<i>bisaptaminal</i>	<i>zilnic</i>	<i>bisaptaminal</i>	<i>zilnic</i>
Mean	3202	3084,08	11,21	10,82	11,57	12,03
Standard Error	69,53	81,83	0,12	0,18	0,18	0,14
Median	3200	3150	11,3	10,5	11,8	12
Mode	3000	3000	10,8	10,2	11,8	12,9
Standard Deviation	491,66	572,83	0,85	1,25	1,28	0,99
Sample Variance	241730,61	328137,16	0,72	1,56	1,63	0,99
Kurtosis	0,03	3,52	-0,15	-0,72	0,96	-0,33
Skewness	0,05	-1,39	-0,30	0,26	-0,99	-0,19
Range	2400	3070	3,7	4,9	5,6	4,4
Minimum	2000	930	9,3	8,3	7,8	9,7
Maximum	4400	4000	13	13,2	13,4	14,1
Sum	160100	151120	560,4	530,1	578,7	589,3
Count	50	49	50	49	50	49
Confidence Level(95,0%)	139,7285287	164,536724	0,240796065	0,359074	0,362397778	0,285299

Pentru a crea fișierul PowerPoint:

- Pentru a crea o prezentare PowerPoint: **[Start – Programs – Microsoft Office – Microsoft PowerPoint]**

- Pentru a adăuga un design predefinit prezentării: [**Format – Slide Design ...**]
- Pentru a modifica design-ul: [**View – Master – Slide Master**]
- Pentru a adăuga un slide nou: [**Insert – New Slide**]
- Pentru a șterge un slide: selectați slide-ul pe care vreți să-l ștergeți în partea stângă, la Slides, și apăsați **tasta Delete**.
- Pentru a ascunde un slide: clic dreapta pe slide-ul pe care vreți să-l ascundeți și alegeți opțiunea **HIDE**.
- Pentru a adăuga o poză în prezentare: [**Insert - Picture**]
- Pentru a anula o prezentare: [**Slide Show – Slide Transition**] pentru a alege cum să apară slide-ul; [**Slide Show – Custom Animations**] pentru a anima text și /sau poze;
- Pentru a vedea prezentarea: [**Slide Show – View Show**].
- Pentru a salva prezentarea: [**File – Save as – Save as type: PowerPoint Show**].