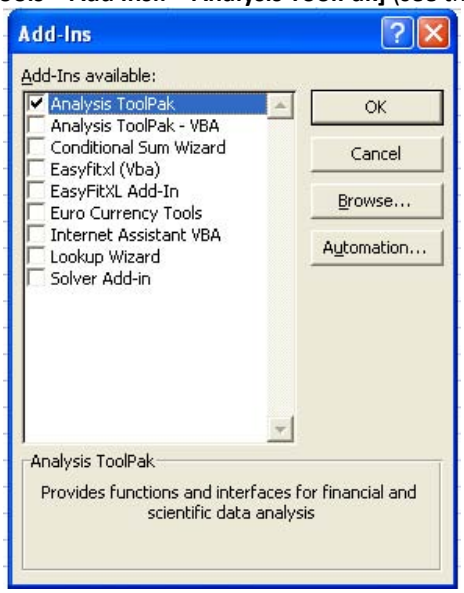
MICROSOFT EXCEL: DESCRIPTIVE STATISTICS: HINTS

To create a Histogram (graphical representation of frequency distribution):

- Create first in the sheet named Histograms the frequency table and include here the upper bound of each class:

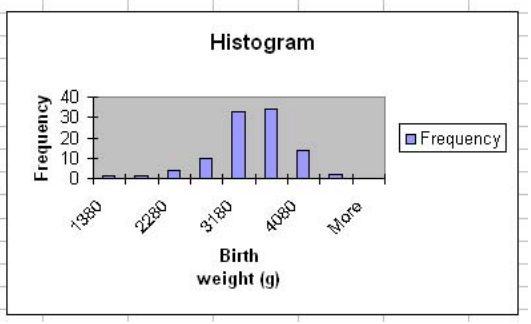
F	G	H
BIN Birth weight (g)	**BIN Hemoglobin (mg/dl) 6 month**	**BIN Hemoglobin (mg/dl) 12 month**
1380	8.8	8.9
1830	9.6	9.8
2280	10.4	10.7
2730	11.2	11.6
3180	12.0	12.5
3630	12.8	13.4
4080	13.6	14.3
4530		

- **[Tools – Data Analysis - Histogram]**
- To activate Data Analysis: **[Tools – Add Ins.. – Analysis ToolPak]** (see the image bellow)



- Create the histogram: **[Tools – Data Analysis - Histogram].** In Histogram dialog box filed the fields as follows:
  - Input Range: select the range were your data are placed (e.g. for Birth weight select $A$1:$A$100)
  - Bin Range: select the cells were you put the upper bounder of interval (e.g. for Birth weight select $F$1:$F$9)
  - Specify that you have Labels in first row (the name of variables)
  - Output range: click on a cell where you want to start the output of Histogram function (e.g. $O$1).
  - Click on Chart Output (beside the frequency table you will also have the graphical representation).
  - As a results you will have something as in the image bellow:

Birth weight (g)	Frequency
1380	1
1830	1
2280	4
2730	10
3180	33
3630	34
4080	14
4530	2
More	0

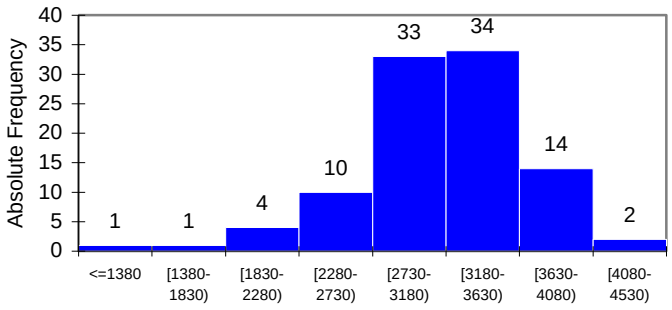


N.B. The above graphical representation is a column chart. In a Histogram there are no spaces between columns.

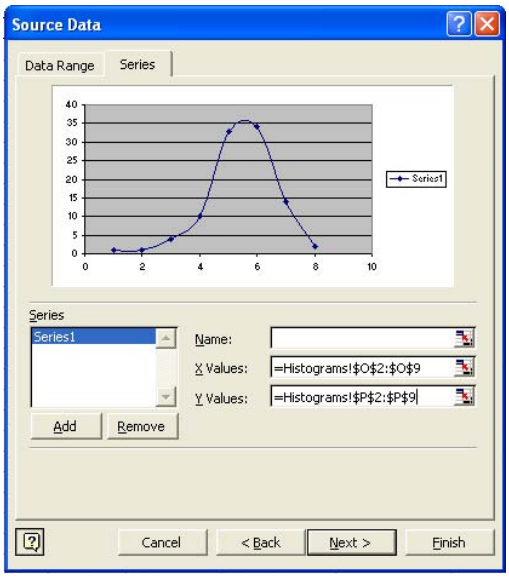
- Modify the Birth weight as follows:

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

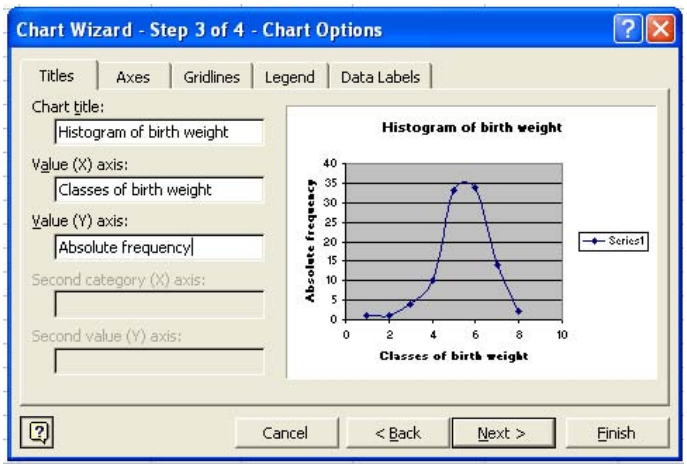
- Work on graph design in order to look like in the image bellow:



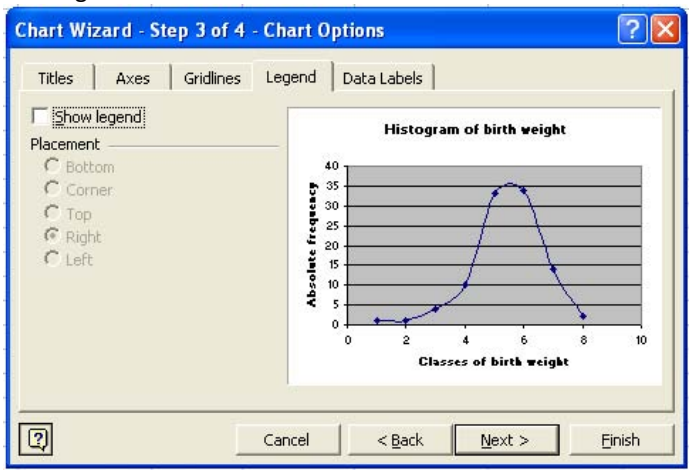
- If you want a bell graph (if is easiest for you to interpret it) create using the frequency table a Scatter following the steps:
  - **[Insert – Chart – Scatter – Second sub-Type]**
  - To the Data range window select Series and Add
  - Fill the X Value (classes of variable without label) and Y Values (frequency of variable without label) as in the example bellow (the example is for Birth weight variable):



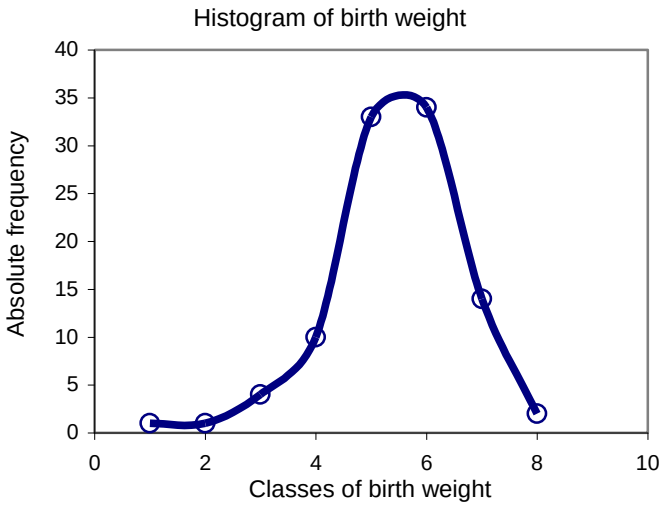
- Include a title and the name for axes:



- Remove the legend:



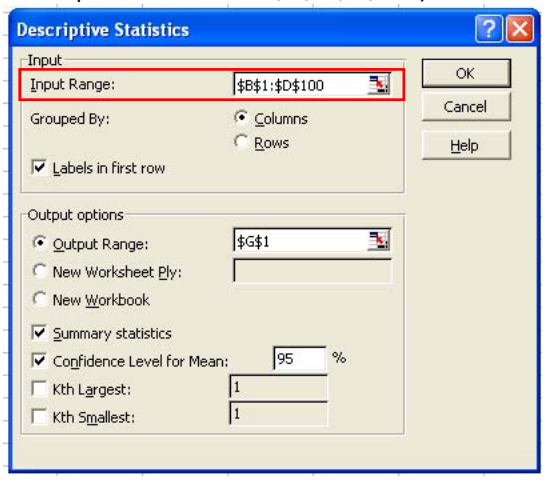
- Your graphical representation will look like in the image bellow:



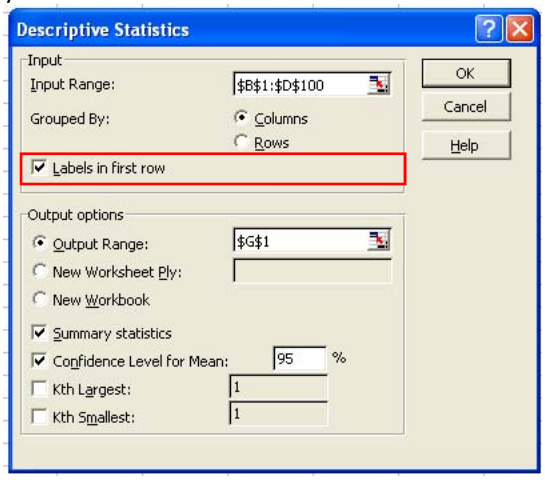
N.B. A disadvantage of the above graph is that is necessary to specify at the bottom of the graph the significance of classes of birth weight.

To compute descriptive statistics parameters:

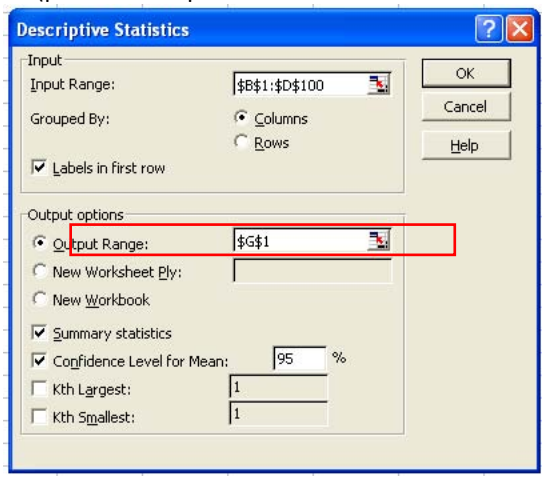
- **[Tools – Data Analysis – Descriptive statistics]**
- Descriptive Statistics dialog box:
  - Input Range: select the range where the data (including the label of variable) are (e.g. for our request the data are $B$1:$D$100).



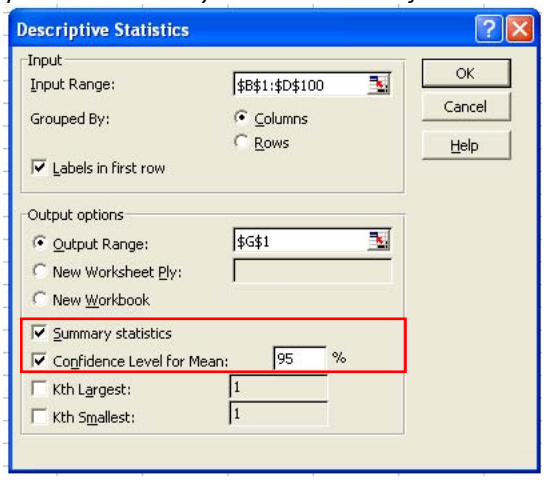
  - Specify that you have Labels in first row.



- Output options: Output range: specify the first cell from which the output will be displayed (place the output in the same worksheet as the date).



- Specify that you want *Summary Statistics* and *Confidence Level for Mean*.



- The result will be like in the image bellow:

G	H	I	J	K	L
Birth weight (g)		*Hemoglobin (mg/dl) 6 month*		*Hemoglobin (mg/dl) 12 month*	
Mean	3143.636	Mean	11.01515	Mean	11.79798
Standard E	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 D	533.9446	Standard D	1.078298	Standard D	1.161015
Sample Va	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	106.4934	Confidence	0.215063	Confidence	0.23156

- Move the name of variables one cell to the right and delete columns I and K (contain the same information as column G). Display the decimal numbers with 2 decimals. Your results will look like those in the image bellow:

G	H	I	J
	Birth weight (g)	*Hemoglobin (mg/dl) 6 month*	*Hemoglobin (mg/dl) 12 month*
Mean	3143.64	11.02	11.80
Standard Error	53.66	0.11	0.12
Median	3200	11.10	11.90
Mode	3000	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	4.90	6.30
Minimum	930	8.30	7.80
Maximum	4400	13.20	14.10
Sum	311220	1090.50	1168
Count	99	99	99
Confidence Level(95.0%)	106.49	0.22	0.23

- Interpretation by example (birth weight variable):

Mean	The arithmetic average of the 99 newborn child included into the study was equal with 3143.63 gram.
Standard Error	The standard error of the mean for the birth weight was of 53.66.
Median	The observation that split the distribution of birth weight in half was equal with 3200 gram.
Mode	The observation value associated with the highest frequency is equal for our study with 3000 gram.
Standard deviation	The population standard deviation for birth weight is equal with 533.94.
Variance	The standard deviation squared for birth weight was equal with 285096.85.
Kurtosis	The distribution of the birth weight is peakedness distribution comparing with normal distribution, kurtosis being equal with 2.55. N.B. If the value belong to the interval [-0.5, +0.5] could be considered that the data follow a normal peakedness distribution.
Skewness	The negative value of the -0.85 for our sample research problem indicates that the distribution of the birth weight is negatively skewed. The negative skew indicates that the longer tail extends in the direction of low values in the distribution. N.B. If the value belong to the interval [-1, 1] could be considered that the data follows a normal skewness.
Range	The range for our distribution is found by subtracting 930 from 4400, producing a range equal to 3470.
Minimum	The lowest value of birth weight by newborn in the studies sample was of 930.
Maximum	The highest value of birth weight by newborn in the studied sample was of 4400.
Sum	The sum of the values in the distribution in the studied sample was of 311220.
Count	The number of observations in the birth weight distribution the studied sample, n = 99
Confidence level (95.0%)	The value obtained represent the amount of error subtracted from and added to the sample mean when constructing the confidence interval fro the population mean. For our problem, the 95% confidence interval is: $3037.14 \leq \mu \leq 3250.13$

To calculate 95% confidence interval for means:

- Insert the following information bellow to the results of descriptive statistics:

G	H	I	J
	Birth weight (g)	*Hemoglobin (mg/dl) 6 month*	*Hemoglobin (mg/dl) 12 month*
Mean	3143.64	11.02	11.80
Standard Error	53.66	0.11	0.12
Median	3200	11.10	11.90
Mode	3000	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	4.90	6.30
Minimum	930	8.30	7.80
Maximum	4400	13.20	14.10
Sum	311220	1090.50	1168
Count	99	99	99
Confidence Level(95.0%)	106.49	0.22	0.23
95% lower bounder			
95% upper bounder			

- Defined the formulas:
  - The lower limit is equal to arithmetic mean (average) minus Confidence Level(95%)
  - The upper limit is equal to arithmetic mean (average) plus Confidence Level(95%)

To compute descriptive statistics for different groups:

- Sort the data ascending by Treatment schema variable: [Data – Sort – Sort by - Ascending].
- Compute descriptive statistics parameter for patients fro Be-weekly patients Rural.
- Sort the data descending by Treatment schema variable: [Data – Sort – Sort by - Ascending].
- Compute descriptive statistics parameter for patients for Daily patients.
- Calculate the upper and lower limit for 95% confidence interval for all cases.
- Your results will look like the ones in the image bellow:

N	O	P	Q	R
	Hemoglobin (mg/dl) 6 month		**Hemoglobin (mg/dl) 12 month**	
	Bi-weekly	**Daily**	**Bi-weekly**	**Daily**
Mean	11.21	10.82	11.57	12.03
Standard Error	0.12	0.18	0.18	0.14
Median	11.30	10.50	11.80	12.00
Mode	10.80	10.20	11.80	12.90
Standard Deviation	0.85	1.25	1.28	0.99
Sample Variance	0.72	1.56	1.63	0.99
Kurtosis	-0.15	-0.72	0.96	-0.33
Skewness	-0.30	0.26	-0.99	-0.19
Range	3.70	4.90	5.60	4.40
Minimum	9.30	8.30	7.80	9.70
Maximum	13.00	13.20	13.40	14.10
Sum	560.40	530.10	578.70	589.30
Count	50	49	50	49
Confidence Level(95.0%)	0.24	0.36	0.36	0.29
95% lower limit	10.97	10.46	11.21	11.74
95% upper limit	11.45	11.18	11.94	12.31

Working to PowerPoint:

- To create a PowerPoint presentation: **[Start – Programs – Microsoft Office – Microsoft PowerPoint]**
- To add a predefined design to your presentation: **[Format – Slide Design …]**
- To modify a design: **[View – Master – Slide Master]**
- To add a new slide: **[Insert – New Slide]**
- To delete a slide: select the slide that you want to delete it in Slides view and use Delete key.
- To hide a slide: right click on the slide that you want to hide and choose HIDE option.
- To add a Picture to the Presentation: **[Insert - Picture]**
- To animate a Presentation: **[Slide Show – Slide Transition]** to impose how a slide appear; **[Slide Show – Custom Animations]** to animate text and/or pictures;
- To view your presentation: **[Slide Show – View Show]**.

- To save the presentation: **[File – Save as – Save as type: PowerPoint Show]**.