ENGINEERING STATISTIC

2nd LECTURE: ORGANIZING AND GRAPHING DATA

1- Qualitative data

a- Frequency distribution for qualitative data : a frequency distribution for qualitative data lists all categories and the number of elements that belong to each categories.

Example (1): A sample of 30 employees from large companies was selected, and these employees were asked how stressful their jobs were. The responses of these employees are recorded. Where very means very stressful , somewhat means somewhat stressful, and none stands for not stressful at all.

Somewhat	none	Somewhat	Very	Very	None
Very	Somewhat	Somewhat	Very	Somewhat	Somewhat
Very	Somewhat	None	very	None	Somewhat
Somewhat	very	Somewhat	Somewhat	Very	None
Somewhat	very	Very	Somewhat	None	Somewhat

For these data:
1-Construct a frequency distribution table.
2- Determine the relative frequency and percentage distributions.

Solution:

1- The variable in this example is *stress on Job*.

2- Classified the variable into three categories : very stressful, somewhat stressful, and not stressful.

3- Relative Frequency of categories= Frequency of category/Sum of all Frequency.

4- The percentage for a category = Relative Frequency of that category *100

Stress on Job	Frequency	Relative Frequency	percentage
Very	10	10/30= 0.333	0.333*100= 33.3
Some what	14	14/30= 0.467	0.467*100= 46.7
None	6	6/30= 0.200	0.2*100= 20
	Sum = 30	Sum = 1	Sum = 100

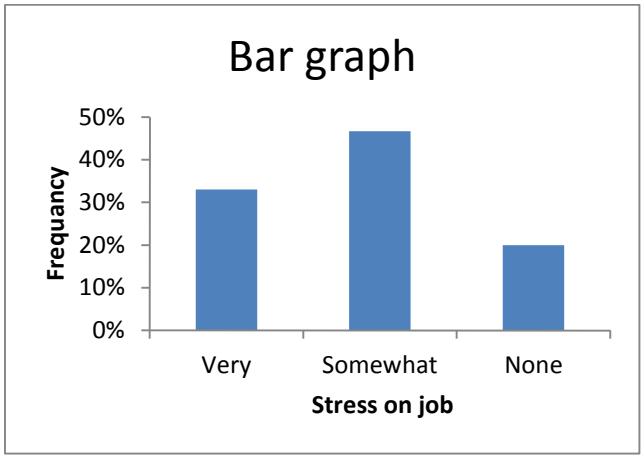
b - Graphical Presentation of Qualitative data:

i- Bar Graph : A graph made of bars whose heights represent the frequencies of respective categories. The graph bar can be also drawn simply for relative frequency and percentage by marking the relative frequency or percentage instead of the class frequencies on the vertical axis.

ii- Pie Chart : it is more commonly used to display percentage, although it can used to display frequencies or relative frequencies. The Whole Pie (or circle) represents the total sample (population). Then we divide the Pie into different portions that represent the different categories.

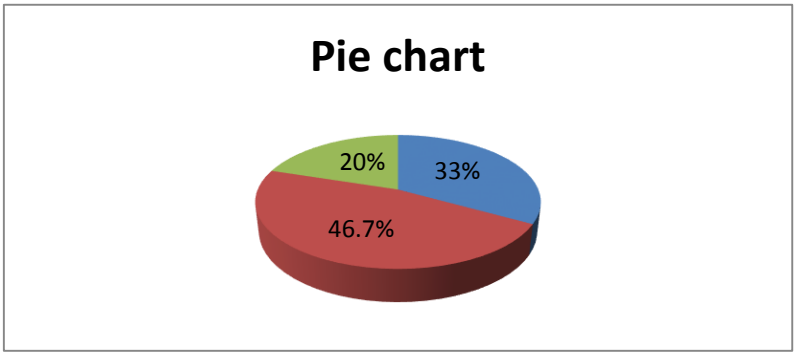
Example (2): Construct bar and pie chart for data in example (1).

- To construct bar graph, we mark the various categories on horizontal axis with same width and marked the frequencies on the vertical axis and then draw one bar for each category and leave a small gap between adjacent bar.



- To construct pie chart, we multiply 360 by relative frequency of each category to obtain the degree measure of angle for the corresponding category.

Stress on Job	Relative Frequency	Angle Size
Very	0.333	0.333*360 = 119.88
Some what	0.467	0.467*360 = 168.12
None	0.200	0.200*360 = 72.00
	Sum = 1	Sum = 360



2- Quantitative data

a- Frequency distribution for quantitative data : a frequency distribution for quantitative data lists all the classes and the values that belong to each class. Data presented in the form of a frequency distribution are called grouped data. When we need to construct the *frequency* distribution, we need to make the following :

- ***Re- arranged data :*** data must be arranged ascension or descending**.**

- ***Number of classes***: usually varies between (5 to 20) and we can use this equation :

$$C = 1 + 3.3 \log n$$

Where : C is the No. of class, and n is the No. of observation in the data.

- Range = largest value – smallest value.

- Approximate Class Width = Range/ Number of class.

Note: usually the approximate class width is rounded to a convenient number, which is then used as the class width.

- Class midpoint = (Lower limit+ upper limit)/2

Example (3): the following data represent the length of (14) concrete beams, measured to the nearest (1cm).(use6classes): :
99,96,94,92,98,89,99,101,104,102,101,106,107,112

Solution:

- Re- arranged data : 89,92,94,96,98,99,99,101,101,102,104,106,107,112

- Range = 112 – 89= 23

- Approximate width of class = 23/6 = 3.8 Say 4.

- Select a starting point for the lowest limit =89.

- The upper limit of first class = 89 +3 = 92. where : class width – accuracy =3

- Lower boundary of first class =89-(accuracy/2)=88.5.

- Upper boundary of first class =92+(accuracy/2)=92.5.

Class limit	Class boundaries	Freq. (f_i)	mid. point (x_m)	cumulative freq. (less than)	cumulative freq. (more than or equal)	Relative Frequency	cumulative Relative freq. (less than)	cumulative Relative freq. (more than or equal)
89-92	88.5-92.5	2	90.5	2	12	0.143	o.857	0.857
93-96	92.5-96.5	2	94.5	4	10	0.143	0.286	0.714
97-100	96.5-100.5	3	98.5	7	7	0.214	0.500	0.500
101-104	100.5-104.5	4	102.5	11	3	0.286	0.786	0.214
105-108	104.5-108.5	2	106.5	13	1	0.143	0.929	0.071
109-112	108.5-112.5	1	110.5	14	0	0.071	1	0
		∑=14				∑=1		

Graphing quantitative data

After the data have been organized into a frequency distribution, they can be presented in graphic forms. The purpose of graphs in statistics is to convey the data to the viewer in pictorial form. Statistical graphs can be used to describe the data set or analyze it.

The three most commonly used graphs are :

1.The histogram.

2. The frequency polygon.

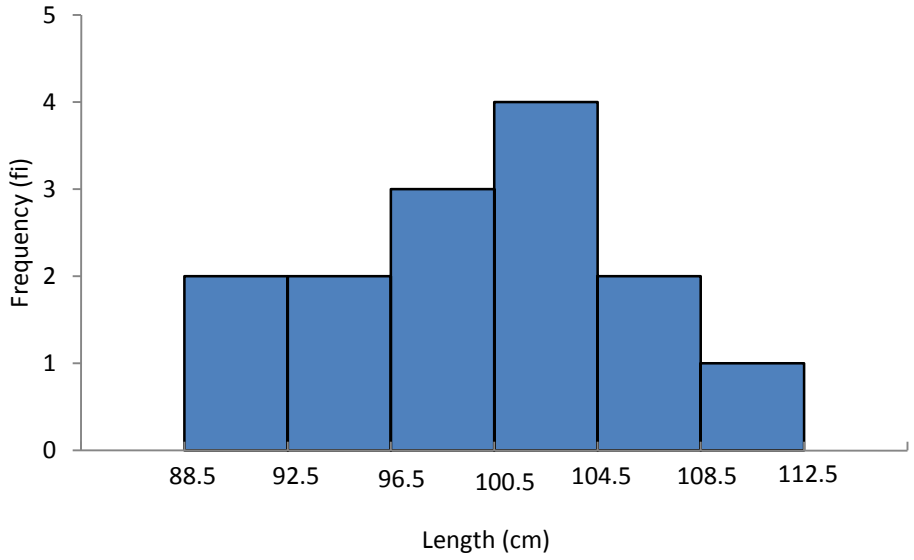
3. The cumulative frequency graph, or ogive.

1. ***Histogram***: It is a graph that can be drawn for frequency distribution, relative frequency distribution, or percentage distribution. To draw a histogram:

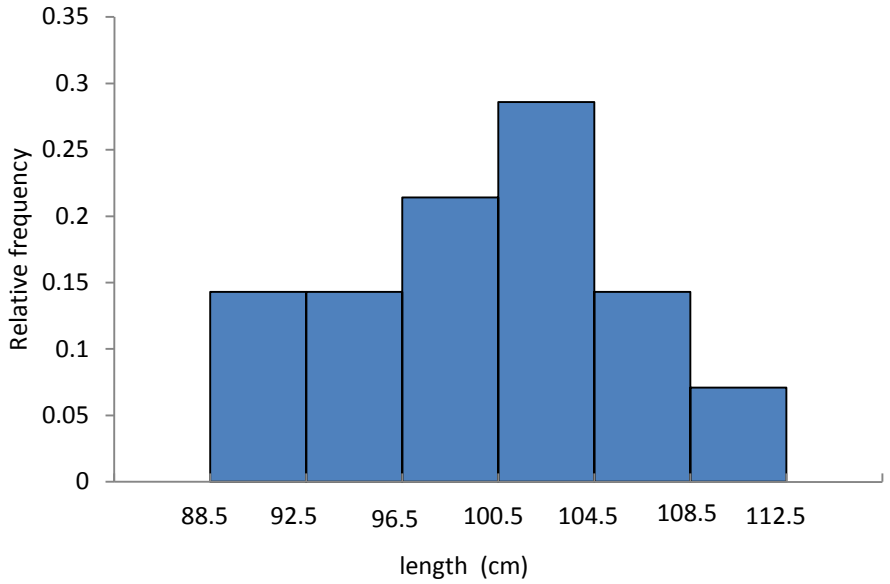
1- Mark classes on the horizontal axis and frequency (or relative frequency or percentages) on vertical axis.

2- Draw a bar for each class so that its height represents the frequency of that class. In histogram the bars are drawn adjacent to each other with no gap between them.

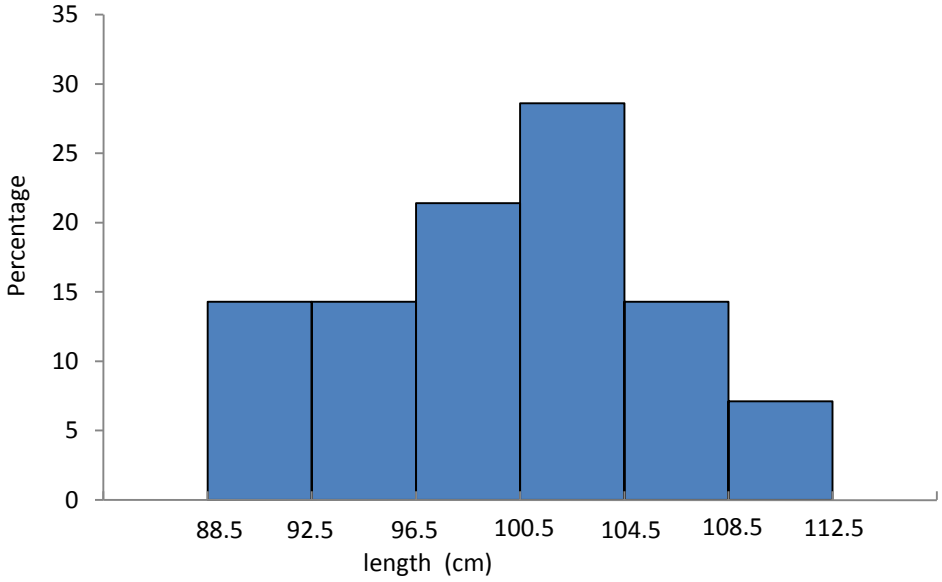
***Construct a frequency histogram, relative frequency histogram, and percentage histogram to represent data shown in example (3)



Frequency histogram



Relative frequency histogram



Percentage histogram

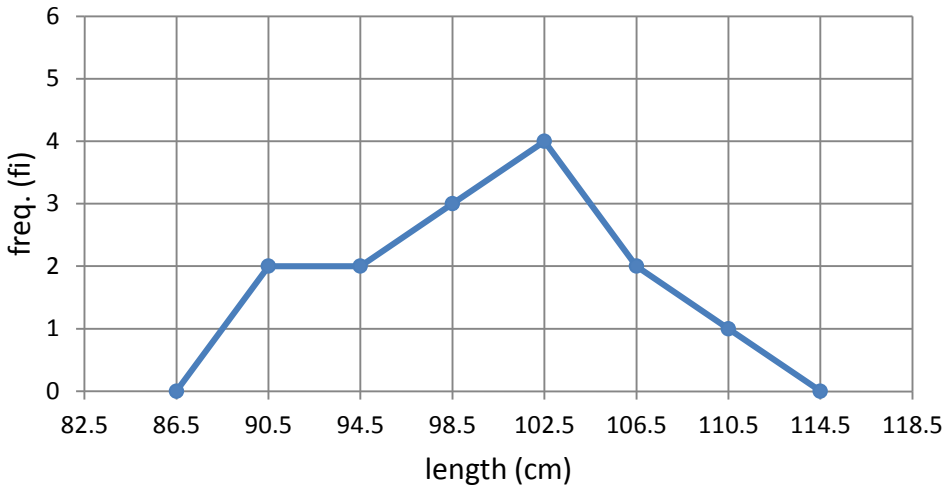
2***- The frequency Polygon***: it is another device that can be used to display the quantitative data in graphic form. To draw the frequency polygon,

1- Mark a dot above the midpoint of each class at a height equal to the frequency of that class

2- Mark two more classes, one at each end, and mark their midpoints. These two classes have zero frequency.

3- Join the adjacent dots with straight lines. The resulting line graph is called frequency polygon.

*** Draw the frequency polygon for above example:



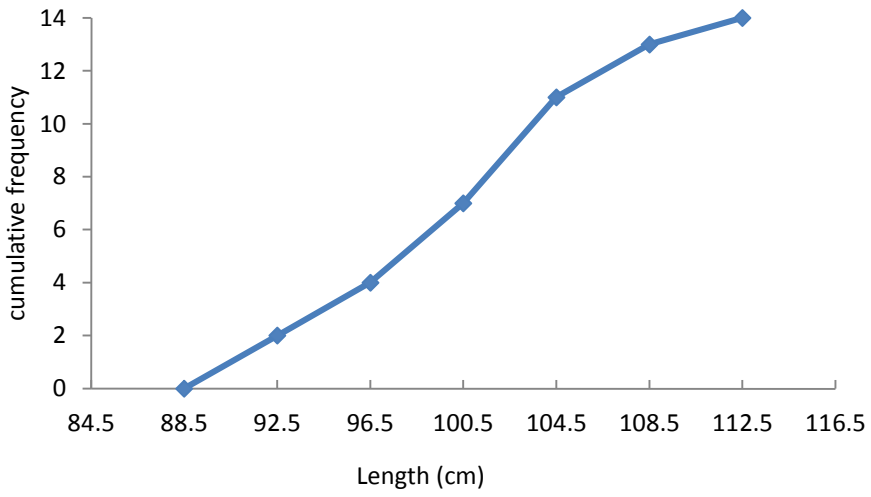
3- Ogives: It is a curve drawn for the cumulative frequency distribution, one advantage of an ogive is that it can be used to approximate the cumulative frequency for any interval. Steps to draw an ogive are:

1- Mark variable on horizontal axis and the cumulative frequencies on the vertical axis.

2- Mark dots above the upper boundaries of various classes at height equal to the cumulative frequency.

3- Joining consecutive points with straight lines. Note that the ogive starts at the lower boundary of the first class and ends at the upper boundary of the last class.

******* Draw an ogive for the cumulative frequency distribution for Exp (3).



Ogive for cumulative frequency distribution