

## Descriptive Statistics

### Summarization and Presentation of Data

Presentation of data can be:

- Tabular: using tables.
- Graphical: using graphs, pictures or charts.
- Pictorial: using pictures or charts

Summarization of data by calculation of:

- a. Measures of Central Location
- b. Measures of Variation (dispersion)

**Tabular presentation:** Construction of a Frequency Distribution table.

**Steps of construction a Frequency Distribution table for a continuous quantitative variable include:**

- Count the number of classes (N) by using sturge's law
- $K=1+3.322 \text{ Log } n$  (k=no. of class, n=sample size)
- Estimation of the width of class intervals.(W)

$W = R / N$ . R range of readings = largest value (L) – smallest value (S), N no. of classes.

- Determination of the lower class limit for the first class by selecting a convenient number that is smaller than the lowest data value.
- Determination of other class limits by repeatedly adding the class

width (*from Step 2*) to the prior class limits.

- Tallying can be used to calculate numbers
- **Example: of continuous quantitative data**
- **weights of malignant tumors removed from the abdomen of 57 subjects:**  
68, 63, 42, 27, 30, 36, 28, 32, 79, 27, 22, 23, 24, 25, 44, 65, 43, 25, 74, 51,  
36, 42, 28, 31, 28, 25, 45, 12, 57, 51, 12, 32, 49, 38, 42, 27, 31, 50, 38, 21,  
16, 24, 69, 47, 23, 22, 43, 27, 49, 28, 23, 19, 46, 30, 43, 49, 12.

10 -	5	5	8.77	8.77
20 -	19	24	33.33	42.1
30 -	10	34	17.54	59.64
40 -	13	47	22.81	82.45
50 -	4	51	7.02	89.47
60 -	4	55	7.02	96.49
70 -80	2	57	3.51	100.00
<b>Totals</b>	<b>57</b>	<b>---</b>	<b>100.00</b>	<b>-----</b>

### Relative Frequency

**The Relative frequency** : is the percentage of number of observations in each class out of the total number of observations.

**Number of observations in each class interval**

$$R.F = \frac{\text{-----}}{\text{-----}} \times 100$$

**Total number of observations**

It is important for the comparison between two distributions having different totals.

**Cumulative Frequency:** it is the number of observations in each class plus the total number of observations in the preceding classes

▣ Example of qualitative data

▣ 12 subjects were tested for blood type:

▣ A,B,O,A,AB,O,B,A,A ,O ,AB,B.

▣ Category	Tallying	Frequency	R.F %
▣ A	IIII	4	33.333
▣ B	III	3	25.00
▣ AB	II	2	16.666
▣ O	III	3	25.00

**The class interval :**

- Classes should be continuous to each other.
- Classes should not be overlapped, i.e. not 0-10, 10-20,20-30 .
- Each class interval has the same “class width”.

**The class interval :**

- Classes should be continuous to each other.
- Classes should not be overlapped, i.e. not 0-10, 10-20,20-30 .
- Each class interval has the same “class width”.
- Graphical Presentations of data

**Graphical presentations of data may aid the reader to pick up the most important idea by just looking to the graph.**

**1.Histogram:** It is presented as rectangles, the width represents the class interval , its height represents the frequency. The rectangles are continuous adjacent to

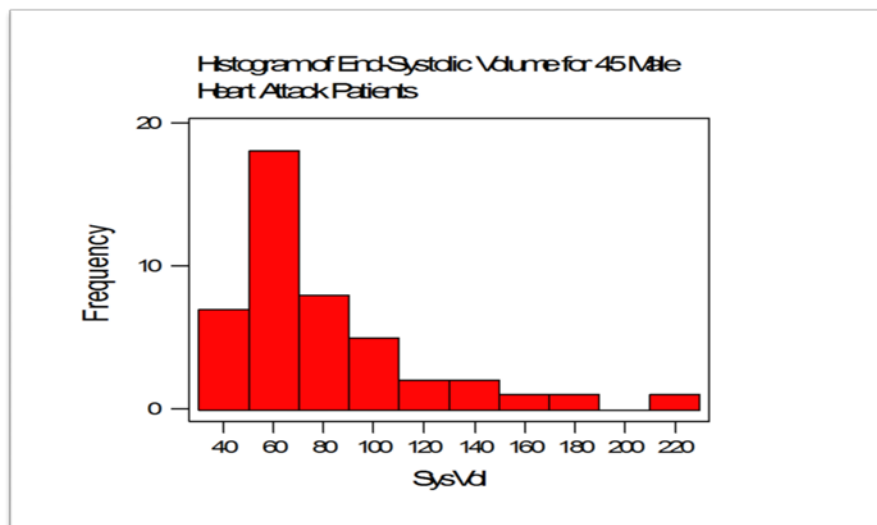
each other . since intervals are usually equal, the widths are equal. If widths are changed then heights are altered such that the area under the histogram is constant (unchanged). Histogram is used for continuous quantitative variable & for only one set of data.

**2. Frequency polygon:** it is similar to histogram in its use for quantitative variable but polygon can be used for 2 or more sets of data & this is an advantage of polygon in facilitating comparisons. It can be

### **•Graphical Presentations of data**

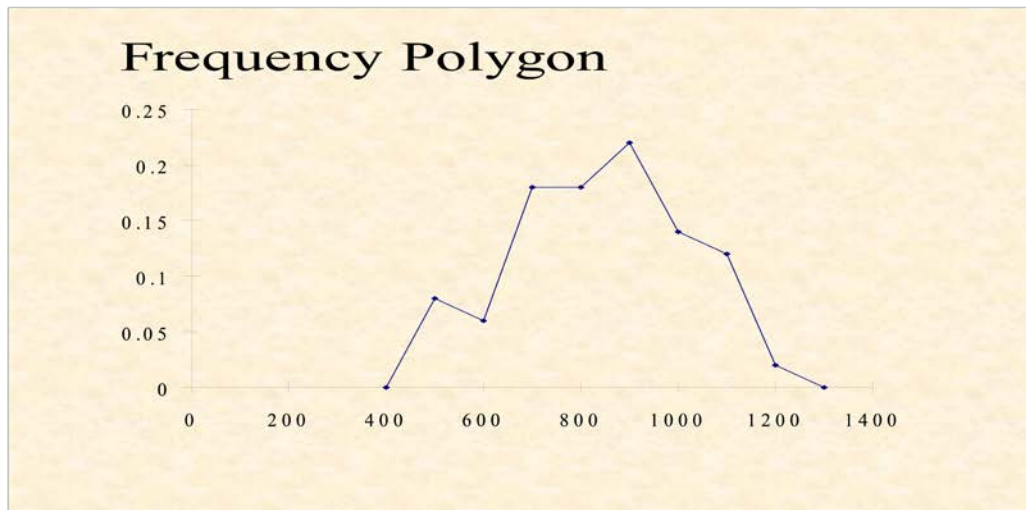
**Graphical presentations of data may aid the reader to pick up the most important idea by just looking to the graph.**

**Histogram:** It is presented as rectangles, the width represents the class interval , its height represents the frequency. The rectangles are continuous adjacent to each other . since intervals are usually equal, the widths are equal. If widths are changed then heights are altered such that the area under the histogram is constant (unchanged). Histogram is used for continuous quantitative variable & for only one set of data.



constructed from histogram by taking the midpoint dot of each rectangle (class interval).

**2. Frequency polygon:** it is similar to histogram in its use for quantitative variable but polygon can be used for 2 or more sets of data & this is an advantage of polygon in facilitating comparisons. It can be constructed from histogram by taking the midpoint dot of each rectangle (class interval).



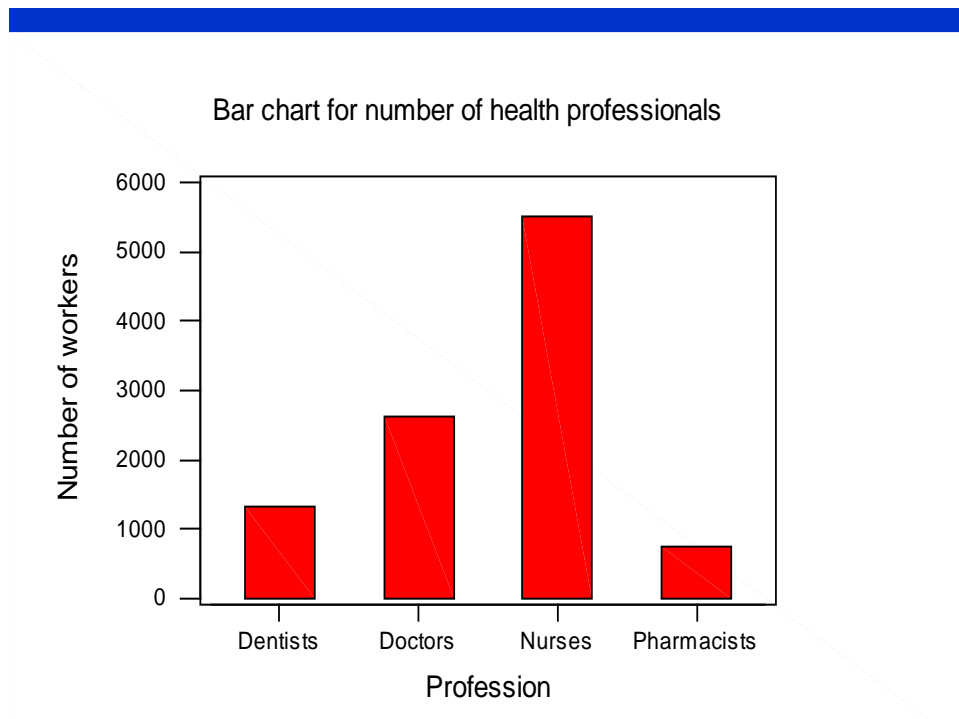
### 3. Frequency curve

Its similar to frequency polygon but the straight line is replace by smooth curve.

**4. Bar chart:** it is used for discrete quantitative variables and qualitative variable. The bars are constructed to show the frequency or relative frequency for each category of the variable on Y-axis, while X-axis is for qualitative & discrete values. It is important that Y axis should start at zero.

- Bar chart is represented as separated rectangles. Width of bars , the horizontal spaces between bars ,and the ordering of the bars are chosen for convenience
- Only heights of bars are important

Bar chart can be used for more than 1 set of data



**Things to keep in mind when completing the results section of your paper:**

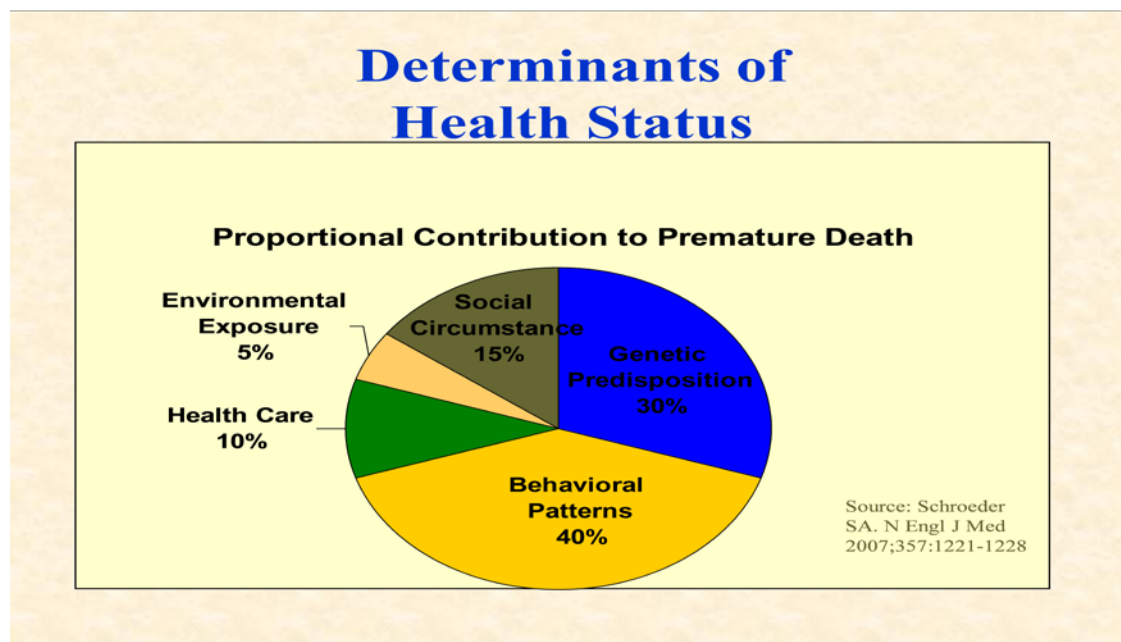
1. Any table or graph should include a title that clearly states what is included in the table.

2. In any graph , it is essential to clearly label the axes so that the reader knows how to read the data being presented.
3. Tables and graphs should include only information that is relevant for seeing (this is the information the researcher wants to convey).
4. If you make adjustment with your data, you should explain it in the text. It is important to document why you may exclude outliers in the text of your paper
5. The table shouldn't include too much information.
6. The important thing is that the tables and graphs are clear and easy to the reader .The table should be well organized.

**6. Pie Diagram (chart):** it is a circle divided into sectors with areas proportional to the frequencies or the relative frequencies of the categories of the variable. It is used for one set of data.

To represent the data as pie chart we must :

1. Find the relative frequency distribution of each category (i.e. % of each variable).
2. Multiply the relative frequency distribution by  $360^\circ$  to find the degree of each category.





# Clustered Bar Chart



**Component bar chart (stacked bar chart) or segmental bar chart:**

