***Practical 3***

***Descriptive statistics; Tables and Graphs***

*Descriptive statistics serve as device for organization, summarization of data.*

*{{Measurements that have not been organized summarized or otherwise manipulated are called* ***''raw data''****}}.*

 *Organization and summarization and the displaying of data* عرض البيانات*can be done by the following methods.*

1. ***Tabular (Tables):***
2. ***Graphs (figures)***

*Tables and graphs are often displayed the distributions, which are useful:*

## *(a) To detect any trends in the data*

*(b) To detect “unusual” scores in a set of data*

*(c) To quickly present/summarize a great deal of info*

1. ***Numbers (mathematics)***

*Single number to summarize the whole data*

***Note:*** *The purpose of tables and graphs is to present information in a concise way so that readers can comprehend*  فهم *and remember it more easily.*

 *Tables and Graphs should stand alone in a report.*

* *This means that the reader should not have to refer to the text in order to understand and interpret the information in them.*
* *In practice this means, they require descriptive titles and clear, meaningful labels.*

***1- Tabular (Tables):***

 *Table consists of row(S) & column(S), could be 2x2, 2x3….etc.*

***List*** *is the simplest form of table, consists of two columns only, the first giving an identification of the observation unit and the second giving the value of the variable of that unit.*

*Table must be:*

 *a- As simple as possible (it is better to have 2-3 simple tables than one complicated).*

 *b- Understandable & self explanatory without references to the text. This is done by:*

* *The title should be clear (placed above the table), and answer the questions of: What? Where? And When?*
* *Each row and column should be labeled clearly and concisely.*
* *Specific unit of the measure for the data should be defined.*
* *Total should be placed.*
* *Illustrate symbols, code, and abbreviation by putting a footnote below the table.*

 *c- Source of the table (if not original).*

 *d- Avoid too much over ruling.*

***Tables can display two types of data***

*1- Categorical*

*Ex: 12 subjects were tested for blood type: A, B, O, A, AB, B, O, B, A, A, A, AB*

***2- Numerical***

***Ex: The ages of 15 subjects were: 14, 15,16 20,22, 27, 28, 31, 33, 37, 48, 50, 59, 60, 60***

***Data must be grouped in classes***

***2- Diagrammatic "Graphical","Figuer".***

*Graphs represent pictorial representations of the distribution of data. Graphs should present a clear and accurate picture about the data and convey an easy-to-understand message. Graphs include* ***labels*** *with units of measurement, a title that accurately describes the contents of the graph, and a* ***scale*** *on each axes that neither distorts nor exaggerates the data.* ***Sample size*** *should be included on the graph or legend.*

***Graphs versus Tables***

* *Advantages : Simplicity, clarity*

 *Easily remembered visual image*

 *Picture of complex relationships*

 *Emphasis*

 *Popular*

* *Disadvantages : Lack of precision*

 *Lack of flexibility*

 *Provide for distortion*

* *The form of the diagram varies according to the nature of the data:*

 *a- For* ***categorical data,*** *we have "chart", this can be:*

* *Bar chart; This is a graphical representation of the (relative) frequencies or magnitudes by rectangles of constant width drawn with length proportional to the (relative) frequencies or magnitudes concerned, as the EX. Below:*

**

*Fig: the distribution of patients according to ICUs admission.*

*We have: 1- Clustered bar chart 2- Segmented bar chart*

*Useful for displaying the distribution of more than one categorical variable*



**Bar chart can be plotted horizontally instead of vertically**



* Pie chart; This is a graphical representation of the (relative) frequencies or magnitudes by a circle whose area represent the total frequency and which is divided into segments which represent the proportional composition of total frequency, as the EX. Below:



* Picto-chart ((Picto-gram)). This is a graphical representation of the (relative) frequencies by using symbols (drawing or picture) relevant to the subject matter. Symbols of different size should not be used. A unit value of the data should be represented by standard symbol which may repeat to represent magnitude. As in the Ex.

|  |  |  |
| --- | --- | --- |
|  Parity |  0 | ♀♀♀♀♀ |
| 1-3 | ♀♀♀♀♀♀♀ |
| ≥4 | ♀ |

 Fig. Pictogram, No. of mothers (all ages) in 100, 000.

* Flow chart. Flow-chart. The sequence of series of events is often illustrated by flow chart, as in the Ex.

b-***For numerical data***

* One dimensional dot diagram. This is a diagrammatic representation of the distribution of a variable in which every observation is marked as a dot corresponding to its value on a line (usually horizontal line) calibrated within the range of values of the variable. As in the Ex.
* Arithmetic Scale Line Graph



* Scattered diagram (dot-graph). Each observation is marked as a dote corresponding to its value on each axis (X & Y). The pattern made by the dotes is an indication of relation between the two axes, which may be linear if the follow a straight line or curved if not.
* Histogram: This is a graphical representation of frequency distribution in which rectangle proportional in the area to the frequencies are erected on the horizontal axis. The base lines are continuous (because we are dealing with continues variables). The width of the rectangles should be equal. As in the Ex.



* Frequency polygon: If we join the tops of the rectangles in the histogram→Polygon (total area of histogram= total area of polygon). It is only appropriate when the variables on the horizontal axis are continues and there is only single value on the vertical axis. As in the Ex.





\*Advantage: can present more than one set of data.

Analyze \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_ Descriptive Statisrtics \_\_\_\_\_\_\_\_Frequency distribution \_\_\_\_\_\_ Chart ( we choose the suitible graph )\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_ok