Statistical Tables

A statistical table is an orderly and systematic presentation of numerical data in rows and columns. Rows (stubs) are horizontal and columns (captions) are vertical arrangements. The use of tables for organizing data involves grouping the data into mutually exclusive categories of the variables and counting the number of occurrences (frequency) to each category.

These mutually exclusive categories, for qualitative variables, are naturally occurring groupings. For example, Sex (Male, Female), Marital status (Single, Married, Divorced, Widowed, etc.), Blood group (A, B, AB, O), Method of Delivery (Normal, forceps, Cesarean section, etc.), etc. are some qualitative variables with exclusive categories.

In the case of large size quantitative variables like weight, height, etc. measurements, the groups are formed by amalgamating continuous values into classes of intervals. There are, however, variables which have frequently used standard classes. One of such variables, which have wider applications in demographic surveys, is age. The age distribution of a population is described based on the following intervals:

< 1	20 – 24	45 -49
1 – 4	25 – 29	50 – 54
5 – 9	30 – 34	55 – 59
10 – 14	35 – 39	60 – 64
15 – 19	40 – 44	65 +

Construction of tables

Although there are no hard and fast rules to follow, the following general principles should be addressed in constructing tables.

1. Tables should be as simple as possible.

2. Tables should be self-explanatory. For that purpose

- Title should be clear and to the point(a good title answers: what? when? where? how classified ?) and it be placed above the table.
- Each row and column should be labeled.
- Numerical entities of zero should be explicitly written rather than indicated by a dash. Dashed are reserved for missing or unobserved data.
- Totals should be shown either in the top row and the first column or in the last row and last column.

3. If data are not original, their source should be given in a footnote.

Examples

A) Simple or one-way table: The simple frequency table is used when the individual observations involve only to a single variable. In addition to the frequency counts, the relative frequency is used to clearly depict the distributional pattern of data. It shows the percentages of a given frequency count.

Table 1: Overall immunization status of children in Babel city.

Immunization status	Number	Percent
Not immunized	75	35.7
Partially immunized	57	27.1
Fully immunized	78	37.2
Total	210	100.0

B. Two-way table: This table shows two characteristics and is formed when either the caption or the stub is divided into two or more parts.

In cross tabulated frequency distributions where there are row and column totals, the decision for the denominator is based on the variable of interest to be compared over the subset of the other variable. For example, in Table 2 the interest is to compare the immunization status of mothers in different marital status group.

Hence, the denominators for the computation of proportion of mothers under each marital status group will be the total numbers of mothers in each marital status category, i.e. row total.

Table 2: Immunization by marital status of the women of childbearing age, Al-Najaf city.

Marital Status	Immunization Status				Total
	Immunized		Non Immunized		
	No.	%	No.	%	
Single	58	24.7	177	75.3	**235**
Married	156	34.7	294	65.3	**450**
Divorced	10	35.7	18	64.3	**28**
Widowed	7	50.0	7	50.0	**14**
Total	**231**	**31.8**	**496**	**68.2**	**727**

C. Higher Order Table: When it is desired to represent three or more characteristics in a single table. Thus, if it is desired to represent the `Profession,' `sex' and `Residence,' of the study individuals, the table would take the form as shown in table 3 below and would be called higher order table**.**

Example: A study was carried out on the degree of job satisfaction among doctors and nurses in rural and urban areas. To describe the sample a cross-tabulation was constructed which included the sex and the residence (rural urban) of the doctors and nurses interviewed.

Table 3: Distribution of Health Professional by Sex and Residence

Profession/Sex		Residence		Total
		Urban	Rural	
Doctors	Male	8 (10.0)	35 (21.0)	43 (17.7)
	Female	2 (3.0)	16 (10.0)	18 (7.4)
Nurses	Male	46 (58.0)	36 (22.0)	82 (33.7)
	Female	23 (29.0)	77 (47.0)	100 (41.2)
Total		79 (100.0)	164 (100.0)	243 (100.0)

Diagrammatic Representation of Data

Appropriately drawn **graph** allows readers to obtain rapidly an overall grasp of the data presented. The relationship between numbers of various magnitudes can usually be seen more quickly and easily from a graph than from a table.

Figures are not always interesting, and as their size and number increase they become confusing and uninteresting to such an extent that no one (unless he is specifically interested) would care to study them. Their study is a greater strain upon the mind without, in most cases, any scientific result. The aim of statistical methods, is to reduce the size of statistical data and to render them easily intelligible. To attain this objective the methods of classification, tabulation, averages and percentages are generally used. But the method of diagrammatic

representation (visual aids) is probably simpler and more easily understandable. It consists in presenting statistical material in **geometric figures**, **pictures**, **maps** and **lines** or **curves.**

Construction of graphs

The choice of the particular form among the different possibilities will depend on personal choices and/or the type of the data.

1. **Bar charts** and **pie chart** are commonly used for qualitative or quantitative *discrete data*.
2. **Histograms**, **frequency polygons** are used for quantitative *continuous data.*

Examples of diagrams:

1. Bar Chart

Bar diagrams are used to represent and compare the frequency distribution of *discrete variables* and *attributes* or *categorical series*. When we represent data using bar diagram, all the bars must have equal width and the distance between bars must be equal.

There are different types of bar diagrams, the most important ones are:

A. Simple bar chart: It is a one-dimensional diagram in which the bar represents the whole of the magnitude. The height or length of each bar indicates the size (frequency) of the figure represented.

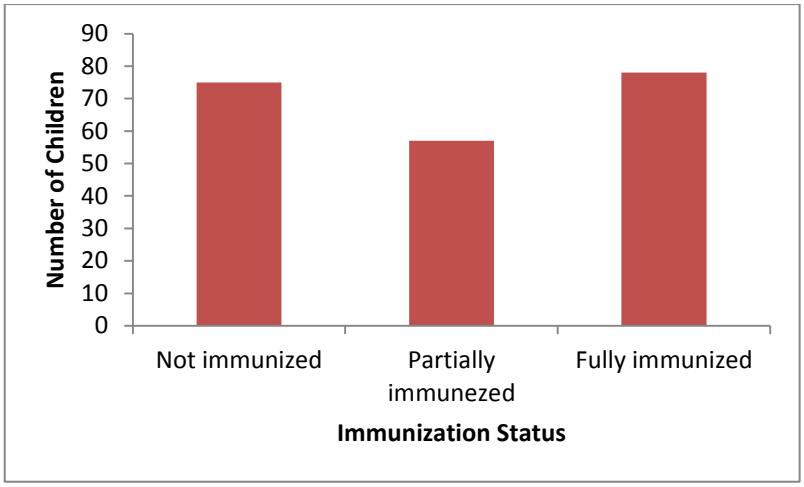

Fig. 1. Immunization status of Children in Babel city

B. Multiple bar chart: In this type of chart the component figures are shown as separate bars adjoining each other. The height of each bar represents the actual value of the component figure. It depicts distributional pattern of more than one variable.

Example of multiple bar diagrams: consider that data on immunization status of women by marital status.

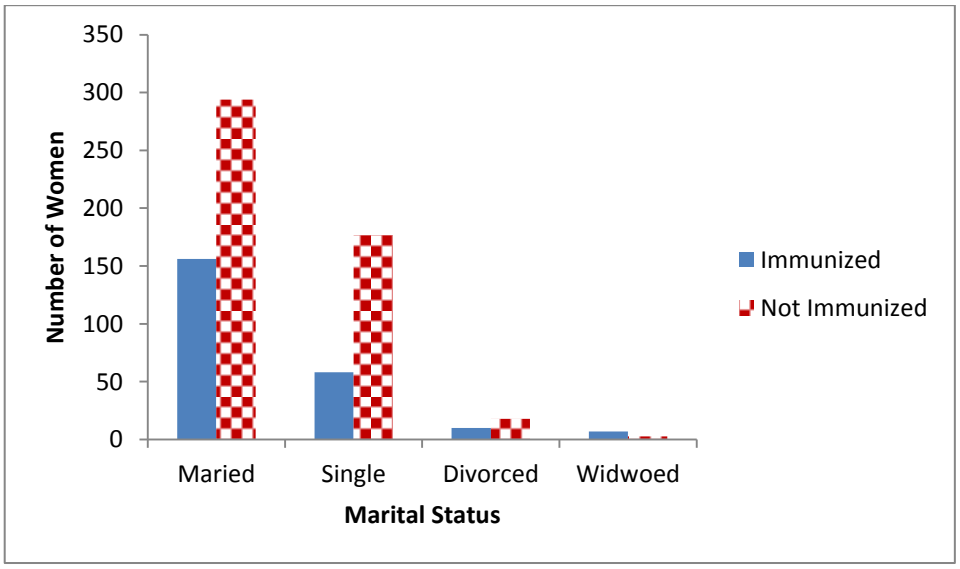

Fig. 2 Immunization status by marital status of women 15-49 years, Al-Najaf.

2. Pie-chart (qualitative or quantitative discrete data)

It is a circle divided into sectors so that the areas of the sectors are proportional to the frequencies.

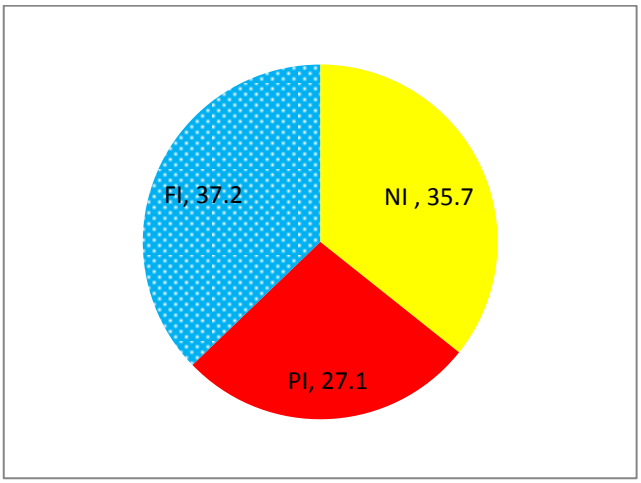

Fig. 3. Immunization status of children in Babel city

3. Histograms (quantitative continuous data)

A histogram is the graph of the frequency distribution of continuous measurement variables. It is constructed on the basis of the following principles:

a) The horizontal axis is a continuous scale running from one extreme end of the distribution to the other. It should be labeled with the name of the variable and the units of measurement.

b) For each class in the distribution a vertical rectangle is drawn with (i) its base on the horizontal axis extending from one class boundary of the class to the other class boundary, there will never be any gap between the histogram rectangles. (ii) the bases of all rectangles will be determined by the width of the class intervals. If a distribution with unequal class-interval is to be presented by means of a histogram, it is necessary to make adjustment for varying magnitudes of the class intervals.

Example: Consider the data on time (in hours) that 80 college students devoted to leisure activities during a typical school week:

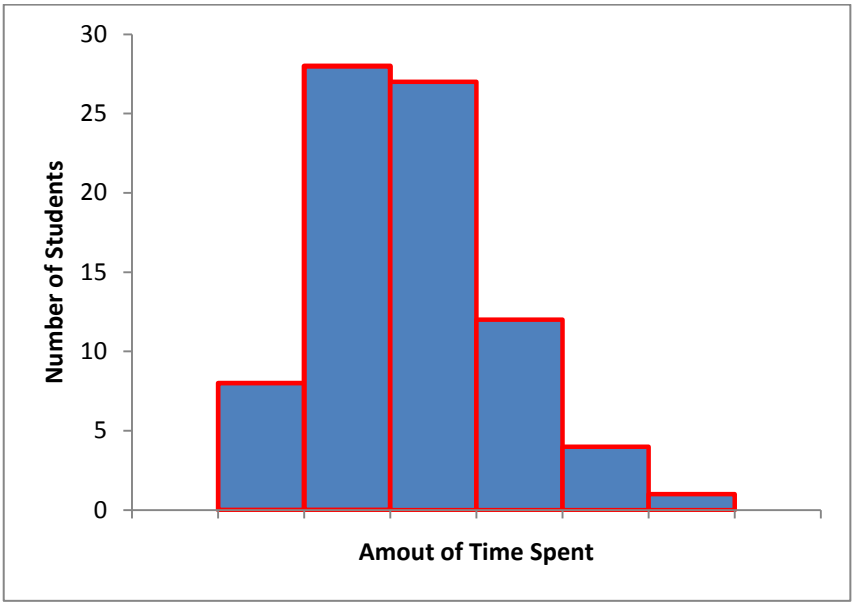

Fig: Histogram for amount of time college students devoted to leisure activities

4. Frequency Polygon

If we join the midpoints of the tops of the adjacent rectangles of the histogram with line segments a frequency polygon is obtained. When the polygon is continued to the X-axis just outside the range of the lengths the total area under the polygon will be equal to the total area under the histogram.

Note that it is not essential to draw histogram in order to obtain frequency polygon. It can be drawn without erecting rectangles of histogram as follows:

1) The scale should be marked in the numerical values of the midpoints of intervals.

2) Erect ordinates on the midpoints of the interval - the length or altitude of an ordinate representing the frequency of the class on whose mid-point it is erected.

3) Join the tops of the ordinates and extend the connecting lines to the scale of sizes.

Example: Consider the above data on time spend on leisure activities

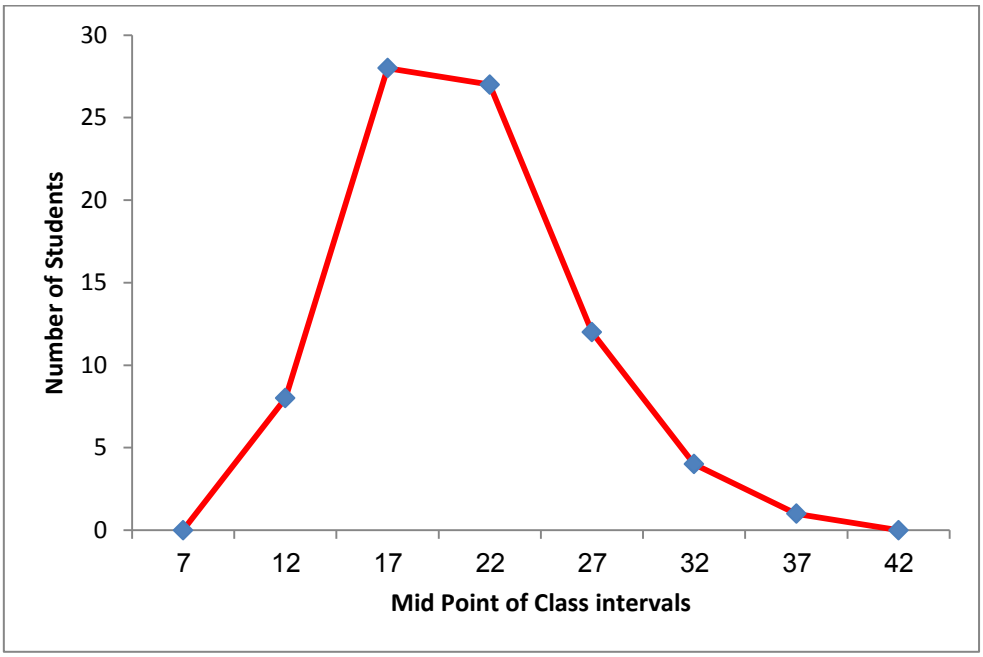

Fig: Frequency polygon curve on time spent for leisure activities by students