UNIT 18 GRAPHICAL PRESENTATIONS OF DATA

Structure
18.0 Objectives
18.1 Introduction
18.2 Charts and Its Types
18.3 Preparing Your Data
18.4 Transforming Your Data into Charts
18.5 Cross Tabulation and Charting
18.6 Let Us Sum up
18.7 Key Words
18.8 Terminal Questions

18.0 OBJECTIVES

After completing this unit, you should be able to:
- understand the basic features for charts generation using Excel;
- create worksheets and computation of data;
- understand the data handling through Excel;
- prepare data; and
- visualise data using charts.

18.1 INTRODUCTION

Every transaction related to any activity of a process generates relevant data. Once this data is properly recorded using Microsoft Excel, the same shall be visualized through various options available in charts. Different stages related to data preparation and creations of charts are discussed in this unit.

18.2 CHARTS AND ITS TYPES

Charts are a set of display tools used to present datasets varied in their size depending upon the problem at hand. Its objective is to create easy to understand chart formats that helps in easily inferring of data. There are various chart types available to visualize the data collected for a better and meaningful understanding.
You can create charts using MS Excel by following steps:

- Click on dataset prepared for data visualization.
- Select option insert from tabs and than choose recommended charts for a set of auto chart.
- Select a chart from the recommended charts and preview the same.
- Choose a best fit as per your requirements for data presentation.

If you want to create a trendline then following steps are to be taken:

- Click on dataset prepared for creating trend line.
- Select design to add the chart element.
- Select a trendline.
- Choose the type for trendline from its different types like linear, exponential, moving average etc.

User can choose from following types of charts for data visualization as per the requirement:

1) **Column Charts:** There are many kinds available in MS Excel like clustered column, stacked column and 3-D visualisation available for the same kind of charts.

2) **Line Charts:** There are various examples of Line charts like stacked line, marked line charts and 3-D visualisation available for the same.

3) **Pie Charts:** Variety of pie charts are available in MS Excel such as Pie of pie, Doughnut and 3-D visualisation for the same.

4) **Bar Charts:** Similar to column charts, these charts are horizontally spreaded. Charts like clustered, stacked and 3-D visualisation of these charts are available.

5) **Area Charts:** Various area charts are provided in MS Excel. Like Stacked Area charts, 3D visualisations of the same are also available.
6) **X Y Scatter Charts**: Scatter charts are a nice addition to the pool of charts available in MS Excel. Charts such as Smooth lines, or with markers, straight lines and Bubble charts.

7) **Stock Charts**: These charts come into use when we need to interpret data related to Stocks and Securities. Various kinds are available like, High-Low-Close, Open-High-Low-Close, Volume-High-Low-Close and Volume-Open-High-Low-Close.

8) **Surface Charts**: These charts can be used to showcase a dataset in a more visual way. Charts such as 3D surface, wireframe, contour and Wireframe Contour are available in MS Excel.

### 18.3 PREPARING YOUR DATA

Data generated through various activities need to be prepared for presentation using charts. Although simple processes lead to the preparation of data, which transforms it into graphical representation used towards further analysis and related generation of inference. It requires further understanding of points to prepare datasets such as variables used along with transactional outcomes and finally the data generated. These points help towards the exchange of data, which in turn generates related information in planning, and outcomes of the process. Importance of the data identification emerges and leads towards a clear description of data using charts and graphs. The correct identification of data types and their respective tabulation and preparation influence the inference through its graphical representation.

Data collection shall happen by means of systematic observations of problem statement related tests. A set of correct data collected supports to perform better statistical analysis. Data collection happens through planned observations for carrying out measurements of one or more variables. These variables constituted by data bear measurable characteristics or attributes. These variables are divided into two large groups of categorical (qualitative) and numerical (quantitative) variables.

1) Categorical or qualitative variables further divides in dichotomous, nominal and (cardinal or ordinal) variables.
   a) Dichotomous variables are binary in nature means having 2 states to record.
   b) Cardinal or Ordinal variables are having 3 or more states or categories but they are in order.
   c) Nominal variables are also having 3 or more categories but no order like color coding of eye-lens black, brown, blue, green, grey or blood groups A, B, AB, O

2) Numerical or quantitative variables further divides into continuous and discrete variables.
a) Continuous variables measured through continuous scale; they have decimal places depending upon the measuring capacity of the instrument like weighing scale, length measuring tapes, and stopwatch.

b) Discrete variables measured through numerical values stored against some mathematical functions used for aggregation of some whole numbers as if count of students attended the class.

Column charts, Bar charts or Line charts for single or multiple data sets are drawn using two axis charts where x-axis constitutes of independent variables and y-axis reflects the data constituted with dependent variable. In case of multiple data sets column or bar charts shown as clustered charts mainly used to display the comparison amongst competitive items with the same range for measure. These types of charts used for comparative analysis for different datasets belong to same category.

18.4 TRANSFORMING YOUR DATA INTO CHARTS

As discussed in previous section, data collection happens through variables. Data collection through planned observations is stored as datasets via one or more variables. These datasets usually displayed with the help of charts or graphs depending upon the type of data used for display.

You are required to transform the required data before displaying it as chart or graph. For example, the dataset having a survey about type of food most liked by people while dining out. The same shall be stored as frequency or count of opted food choice, bar chart and pie chart are most suitable to display as information.

Table 18.1: Aggregate Data for Display

<table>
<thead>
<tr>
<th></th>
<th>Chinese</th>
<th>Spanish</th>
<th>Korean</th>
<th>Mughlai</th>
<th>Italian</th>
<th>Indian</th>
</tr>
</thead>
<tbody>
<tr>
<td>27</td>
<td>5</td>
<td>12</td>
<td>35</td>
<td>15</td>
<td>45</td>
<td></td>
</tr>
</tbody>
</table>
Furthermore, dichotomous variables are having two states to record and hence, the same are displayed using pie charts in a better way as compared to bar chart.

In case of cardinal or ordinal dataset, the outcome displayed using either using bar charts or histogram depending upon the interval of range. Continuous data like collection of student’s height displayed using histogram and discrete data as counted strength of students in a class acquired in bits of data range or categories displayed using bar chart or column chart.

While displaying nominal data like count of people having different blood groups shall also be displayed using bar chart or pie chart. Pie chart provides best results when data is calculated in percentages.

While transforming the available data for creating charts following steps shall provide better results.

1) **Identification of Variables**: Identify all the variables under which the data is classified for a better presentation of data and inference generation.
2) **Missing Values**: Treatment of various cells with missing values.

3) **Outliers**: Identification of outliers with respective treatment leads to better inference.

Following steps are used for creating a chart or graph:

1) Data acquisition
2) Variable mapping
3) Select variables for chart preparation
4) Click Insert ➔ Charts ➔ Recommended Charts
5) Customize Chart Title, Legends, Charts axis as per requirement

The dataset below is comparing the prices of commodities at Shop-A and Shop-B

<table>
<thead>
<tr>
<th>Commodity</th>
<th>Price A</th>
<th>Price B</th>
<th>Status</th>
</tr>
</thead>
<tbody>
<tr>
<td>Rice (per Kg.)</td>
<td>100.00</td>
<td>120.00</td>
<td>Shop-B</td>
</tr>
<tr>
<td>Refined Oil (Per Litre)</td>
<td>85.00</td>
<td>92.00</td>
<td>Shop-B</td>
</tr>
<tr>
<td>Butter (500 Gms.)</td>
<td>180.00</td>
<td>169.00</td>
<td>Shop-A</td>
</tr>
<tr>
<td>Butter (100 Gms.)</td>
<td>45.00</td>
<td>43.00</td>
<td>Shop-A</td>
</tr>
</tbody>
</table>

The charts for above data set shall be displayed as follows:

![Commodity Price Comparison](image)

**Fig. 18.4: Commodity Price Comparison**

While customizing the chart, there are following provisions available:

1) Layout options helps in choosing the layout of chart and its interaction with the text around it.
2) Title, legends, grid lines and data labels can be added or removed from existing charts using chart elements section of customization.
3) While displaying the chart, there are style options available to choose.
4) There is an option of filtering amongst all the elements and data labels available for display.

18.5 CROSS TABULATION AND CHARTING

Pivot tables are used for cross tabulation of data, and pivot charts are used in charting the cross tabulated data. It is a process to validate the relationship amongst various variables; it helps in setting up correlations amongst them. Pivot table are used to rearrange different variables having data in respective columns; furthermore, they are analyzed to generate required report.

Steps for using pivot table:

- Arrange data in rows and columns in a worksheet.
- Insert → Pivot Table → Select the rows and column to be used

Fig. 18.5: Cross Tab - Pivot Table

- A new worksheet generated with interface of generating Pivot Table.
Fig. 18.6: Interface of Pivot Table Generation

- Select the fields for drag and drop as filters, columns or rows.
- Fine-tune your calculations by choosing aggregation functions in values.
- Automatically, the pivot table starts getting shaped up.

Detailed outcomes for selected criteria are shown in Figure 18.7.

Check Your Progress A

1) What are charts in MS Excel?

...................................................................................
...................................................................................
...................................................................................
2) What are the two basic types of data variables?

3) While displaying percentages as data, which kind of chart is the best possible way to visualize the dataset?

4) What is the role of Pivot Tables?

18.6 LET US SUM UP

Graphical representation is important for any kind of presentation. It makes the dataset look more readable and understandable for the audience. Pivot tables and Pivot charts are a necessity in today’s Analytics Industry. These are the basic requirements from any person who wants to enter this kind of industry. They help interpreting the data in a more effective and efficient way. Check and complete the below case study to understand the topic, better:
Now insert a pivot table for the given dataset and a pivot chart for the same. Then try to analyze your findings.

18.7 KEY WORDS

Charts: Charts are the set of display tools used to present datasets varied in their size depending upon the problem and its objective is to easy-to-understand chart formats to reflect the inference.

Continuous Variables: These variables are measured through continuous scale; they have decimal places depending upon the measuring capacity of the instrument like weighing scale, length measuring tapes, and stopwatch.

Discrete Variables: These variables are measured through numerical values stored against some mathematical functions used for aggregation of some whole numbers as if the count of students attended the class.

Dichotomous Variables: A dichotomous variable is one that takes on one of only two possible values when observed or measured. The value is most often a representation for a measured variable. These variables are binary in nature means having two states to record.

Pivot Table: Pivot table is used to rearrange different variables having data in respective columns; furthermore, they are analyzed to generate required report.

18.8 TERMINAL QUESTIONS

1) Explain the usage of different charts in provision of information using data sets in Excel.
2) How does a cross tab query can be generated in Excel?
3) How do Pivot charts help in understanding the outcome of cross tabulation of data set?
4) Explain the following types of charts:
   1) Pie Charts
   2) Histogram
   3) Stock Charts

Note: These questions are helpful to understand this unit. Do efforts for writing the answer of these questions but do not send your answer to university. It is only for yours practice.