

An Introduction to Microsoft Word 2007

Microsoft offers a bundle of software programs within MS Office, with each offering a different function. Microsoft Word is a word processing program similar (but better than) the antiquated (out of date) Word Perfect or the Microsoft Works Word processor. Just like any other word processing program, you use MS Word to view and edit formatted text within a document. A formatted text document is defined as a document that contains text which may be in bold, italics, underlined, a different color, or a different font. All word processing programs can do all of this, but Microsoft Word allows you to do a lot more. For that reason, Microsoft Word is the most used word processing program in existence today.

Microsoft word is a software that helps you to create neat and attractive documents easily and quickly. You will find its multi-facilitated capability quite handy in drafting letter casual and official letter, preparing mailing lists and reports.

Various Features of MS WORD

The word processor has many unique features as follows: -

- 1) **Typing easy:** In MS WORD typing is so easy because we need not click enter button after end of the line as the type writer. The word processor itself takes matter to the next line of the document. This facility is called word wrapping. There is no limit for typing the matter. You can easily correct mistakes as the typing matter appears on the screen.
- 2) **Saving:** The document type can be stored for future use. We can preserve the document for any number of years in word processing.
- 3) **Adding, removing and applying text:** Documents can be modified easily in MS office we can easily place a new word in place of existing one. The new words will automatically be adjusted in the place of deleted or modified text. We can also copy a part or whole of the matter from one document to another document.
- 4) **Spell check of words:** The spelling of the words in the document can be rectified automatically we can also find alternative words to our typed words. The grammatical errors can also be rectified in word processor.
- 5) **Change the style and shape of characters and paragraph:** The documents in word processor can be made attractive because the shape and style of characters in the documents can be changed according to our requirement.
- 6) **Bullets and Numbering:** Bullets are special symbols which can be put for different points, paragraphs and documents where as numbers are ascending figure while 1,2,3 etc.
- 7) **Headers and footers:** A header is the text appearing above the documents and footer is the text appearing below the documents.
- 8) **Creation of tables:** If you give directions for required number of rows and columns word provides a readymade table to work.
- 9) **Inserting pictures and objects:** In the word processor we can insert the pictures in the document to make it more attractive. These pictures can be obtained from clip art which is available in word processor.

10) **Preview facility:** Before getting a print out of our document we can get a preview of document if mistakes are there we can change the document.

11) **Mail merge facility:** If you want to send information to different persons there is a facility in word processor known as mail merge the main body of the letters can be typed once and we can send the same to different addresses.

12) **Macros:** With the help of macros we can avoid certain type of repetitive works. This saves our time and efforts.

Starting Microsoft Word 2007

You open Microsoft Word by clicking on the icon on your desktop (if you have one there) or in the program bar. The icon for Microsoft Word 2007



When you click on the icon, a blank document will open. This is a new document for which the default name is Document1. For each additional new document that you open, the name increases by one digit: Document2, Document3, etc. If you start MS Word by clicking on an already existing document on your computer, it will open automatically and your document will be displayed in the MS Word window.

Virtual Tour of MS Word 2007

There are main components of the MS Word 2007 window that you need to be aware of before we even get into features and functions.



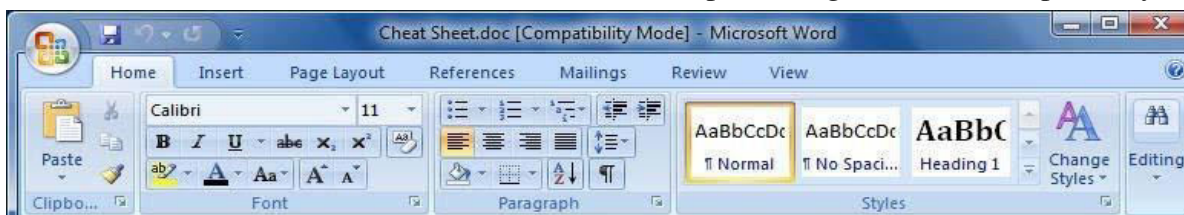
As with any software program or web page that you look at, the line allows you to minimize the page, the box allows you to maximize, and the X closes out the page on your computer. Don't worry. If you should accidentally click the X,

MS Word will prompt you to save the document.

The Ribbon in Word 2007

Once you get used to it, you'll find the Ribbon is far easier to use than the file menu interface used in previous versions of word.

By default the Ribbon is divided into seven tabs, with an optional eighth tab (Developer) they are:



Home: This contains the most-used Word features, such as changing fonts and font attributes, customizing paragraphs, using styles, and finding and replacing text.

Insert: As you might guess, this one handles anything you might want to insert into a document, such as tables, pictures, charts, hyperlinks, bookmarks, headers and footers, WordArt ... etc.

Page Layout: This is where you'll change margins, page size and orientation, set up columns, align objects and add effects. There is some overlap between this tab and the Home tab. E.G. On the Page Layout tab you set paragraph spacing and indents, while on the Home tab you set paragraph alignment and can also set spacing between lines.

References: This tab handles tables of contents, footnotes, bibliographies, indexes and similar material. It also lets you insert a "Table of Authorities," which like a security setting but is in fact is a list of references in a legal document.

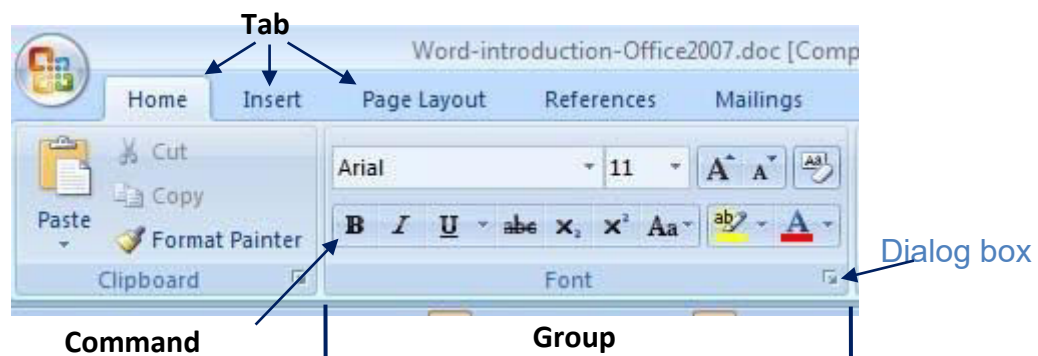
Mailings: As the name says, this is where you'll go for anything to do with mailings, from something as simple as creating labels to the more intimidating task of mail merges.

Review: To check spelling and grammar use the thesaurus, track changes, review other people's changes or compare documents? This is the tab to use.

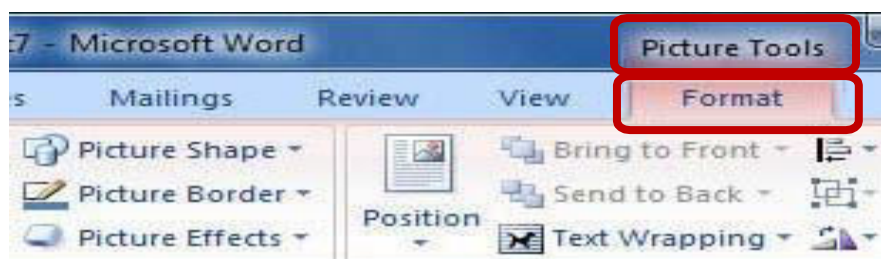
View: This tab allows you to change the view in any way, including displaying a ruler and gridlines, zooming in and out, splitting a window and so on.

Developer: To display this tab, click the Office Button and choose Word Options > Popular > Show Developer tab in the Ribbon. Most users will only need this tab to record Macros but if you write code or want to create forms and applications for Word, this is your tab.

Each tab along the Ribbon is organized to make it easy to get your work done. As you can see below, each tab is organized into a series of groups that contain related commands for getting something done. Inside each group is a set of command buttons, (these buttons as the name suggests, carry out commands, display menus etc, in the example below, the “ B ” command button changes the font to bold. There's also a small diagonal arrow in the bottom right corner of some groups that are called dialog box launchers. Click it to display more options related to the group.



The part of the Ribbon that takes the most getting used to is that it changes according to what you're doing. Depending on the task you're engaged in, it adds or removes tabs and sub tabs. For instance, when you insert and highlight a picture, an entirely new tab appears -- the Format tab, with a "Picture Tools" supertitle on top, as you can see below.

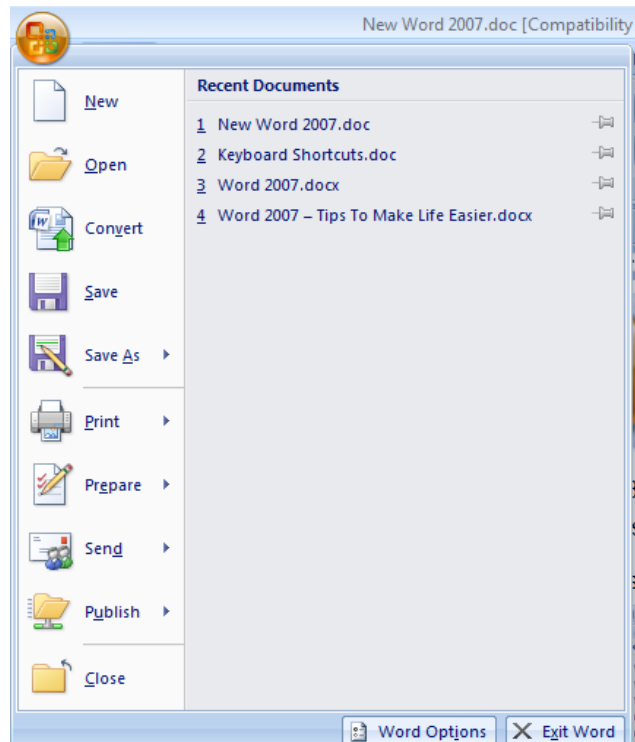


The Office Button and the Quick Access Toolbar

Two tools that make life easier in Office 2007 are the Office Button and the Quick Access toolbar.

The Office Button

The Office Button is like a greatly expanded File menu from Word 2003.



As you can see from the screen shot below, it's where to go for the various Open, Save, New, Print and related options, it also includes a list of all your recently opened files.

Some of the more useful new features accessed via the Office Button are:

Prepare –when you've finished writing or editing your document and you're ready to send it to someone else. There are plenty of options here, such as marking a document as final or read-only; encrypting the document; inspecting it for hidden metadata and information you'd prefer remain private also here is where you can edit a document's properties, such as title, keywords and author.

Publish –gives options for publishing a document. You'll be able to publish your document as a blog to a variety of blogging services, including Blogger, Windows Live Spaces, etc. If your company uses a document management server or SharePoint, you can publish it there as well.

Convert – this feature lets you convert documents saved in older formats (.doc etc) to the new Microsoft Office Open XML format (.docx, the new Office standard).

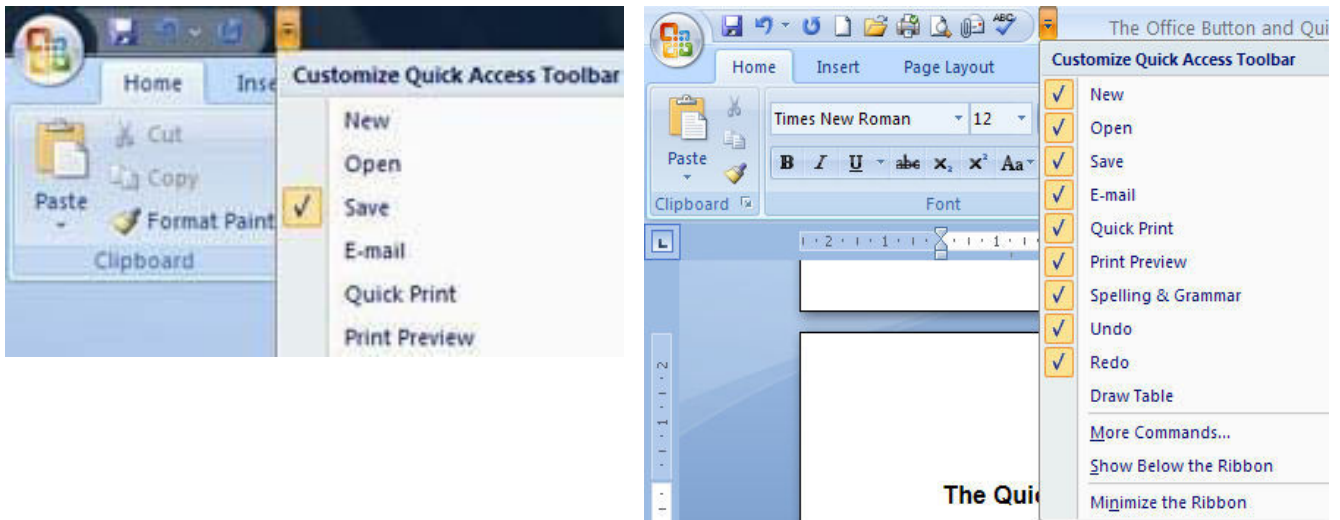
Word Options – For those who like to fiddle with the Word interface and how it works, the Word Options button, located at the bottom of the Office Button's box, lets you customise Word in many ways, including its display and editing options. It has many of the features that you accessed via Tools > Options in previous versions of Word.

The Quick Access Toolbar

Sitting just to the right of the Office Button, the Quick Access toolbar seems a minor addition, but after a while you'll see it's a very handy tool.

The three default buttons Save, Undo and Redo are good to have, but the almost invisible Down arrow to the right of them is the key to the toolbar. Click it, and you'll be able to add and remove toolbar buttons for a preset list of commands and customize the Quick Access toolbar to your liking.

By adding your own shortcuts to the Quick Access toolbar you will greatly reduce the time it takes



to do all the tasks you do on a regular basis.

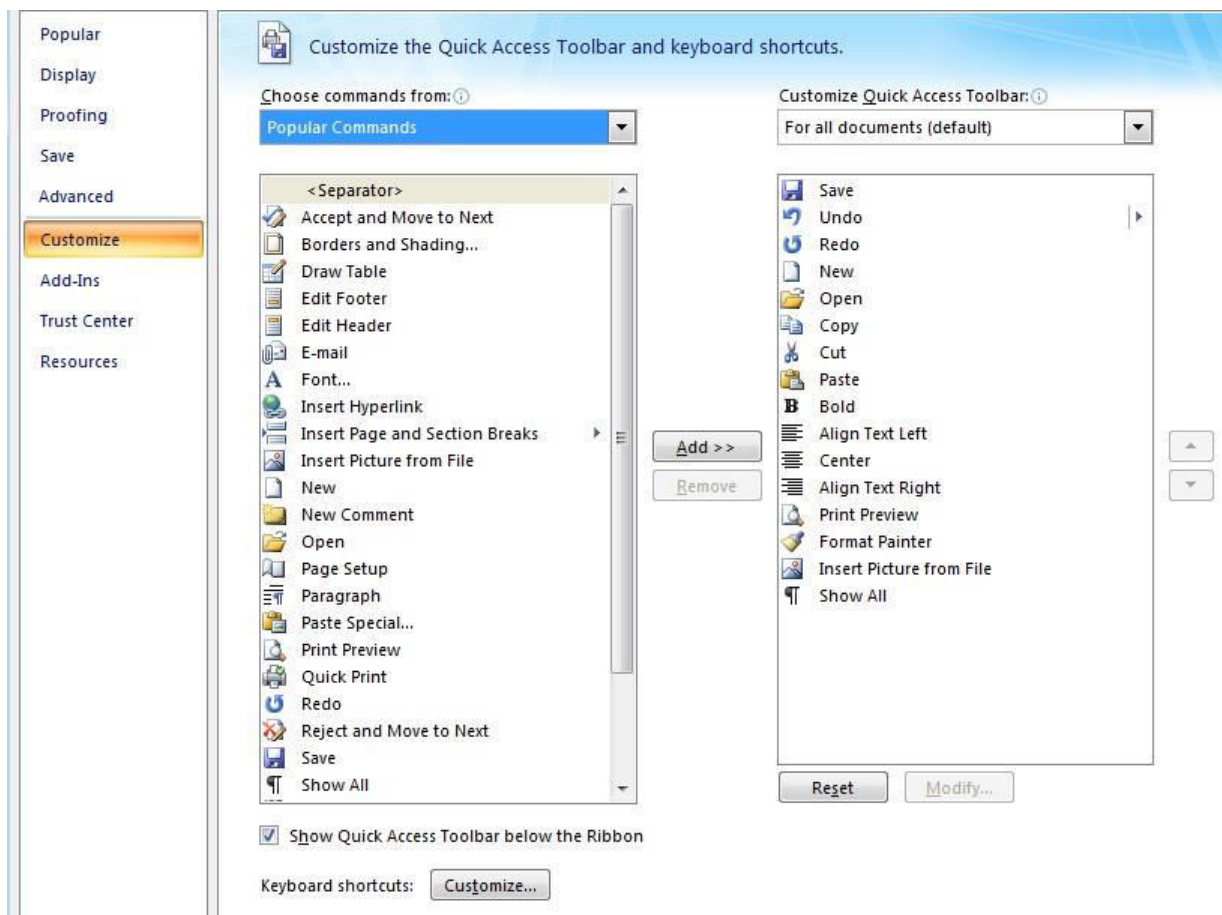
Another way to add commands to the quick access toolbar is to click the **Office Button**, choose **Word Options > Customize**.

The Customize the Quick Access Toolbar and keyboard shortcuts screen is shown.

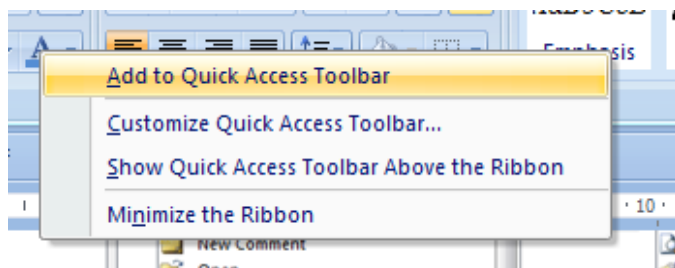
Choose a command from the left-hand side of the screen that you want to add to the Quick Access toolbar and click Add.

You can change the order of the buttons by highlighting a button on the right side of the screen and using the “Up” and “Down” arrows to move it.

The list of commands you see on the left may seem somewhat limited at first. That's because Word is showing you only the most popular commands. There are plenty of others you can add. Click the drop-down menu under "Choose commands from" at the top of the screen, and you'll see other lists of commands. (All Commands, Home Tab and so on) Select any option, and there will be plenty of commands you can add.

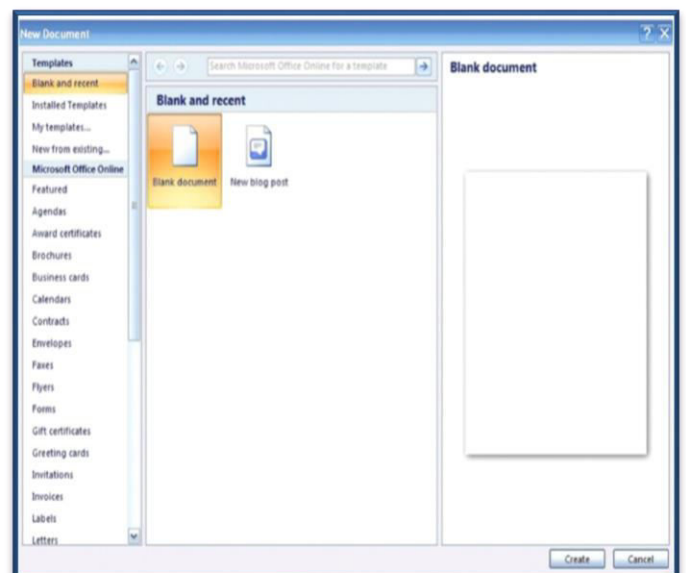


Finally, and probably the easiest is to right-click any object on the Ribbon and choose "Add to Quick Access Toolbar." You can add not only individual commands in this way, but also entire groups. (E.g. the Font group).

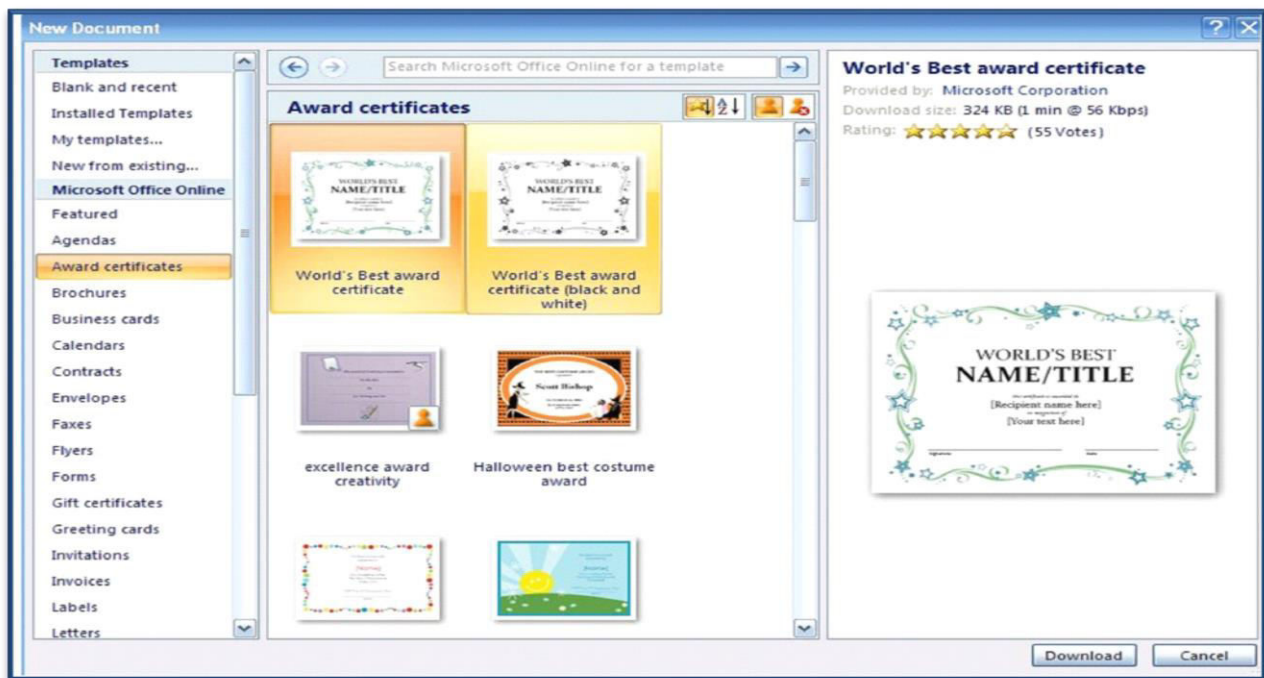


Opening a New Document:- Opening a new document can be accomplished two ways. The first way is to click on the MS Word 2007 icon on your computer. If you already have MS Word open and wish to start a new document, simply click on the Window Office button and select 'New'. You will be presented with the screen pictured.

On the left, you have templates that you may choose from. If you would select 'Award Certificates,' you would then see a screen that



would allow you to choose a certificate template as your new document.



Saving a Document

Let's say that you've created a new document and wish to save it to your computer to use again later. You can do this quickly and easily in MS Word 2007 in one of two ways. The first way is to use the Microsoft Office button. Click on it, then scroll down to select either 'Save' or 'Save As'. Clicking 'Save' will enable you to save the file under its current name. Keep in mind that if this is a new document, it will save the file by the default name of Document1. When you click 'Save,' if another file of the same name exists, MS Word will prompt you to either enter a new file name or to replace the existing copy with the new version you are currently saving.

Clicking 'Save As' gives you a lot more options to saving your work. First of all, when you click 'Save As,' you must specify a file name. You must also specify the format that you want to save the file in. MS Word's default file format is .doc or Word Document. This is an acceptable and much-used format that should be satisfactory for most MS Word users, but you can select the format that you need depending on the work you need to save. You can also save your work as a template or in MS Word 97-2003 recognizable format if you'll be forwarding the file to users who may not have MS Word 2007. When you click on 'Save As,' it will present all these options to you in a slide-out window to the right of 'Save As.'

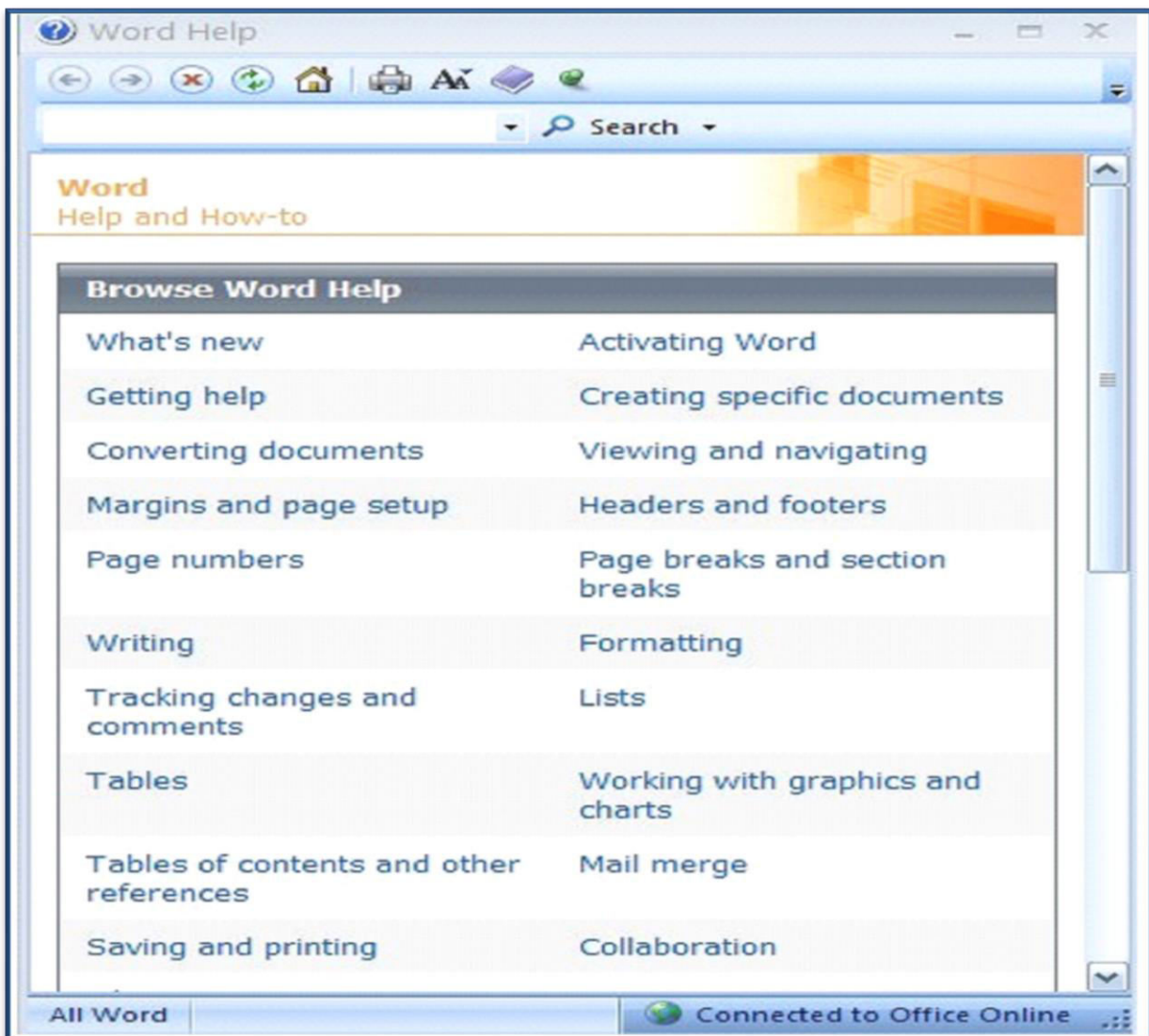
You can also save a document by clicking the picture of the floppy disk that is located to the left of the Quick Access button. However, this will save the file under the current name. You will not be able to name the file or select the format. It's a good idea to click this button every so often while working in a document to save it in case of a power outage, computer freeze, or anything else that may cause you to lose your work.

Warning: Do NOT close MS Word without saving your document. You MUST save your document using one of the methods listed above or your work will be lost forever.

Getting Help with MS Word 2007

If at any time you have a question or need to know exactly how to find or accomplish a task, you can also ask MS Word for help.

In the upper right hand corner of the screen, you'll see a dark blue circle with a question mark inside. This is the help button. You can click on this and a menu will pop up so that you may find the answer to your question. You can also bring up this help window by simply clicking F1 on your keyboard.



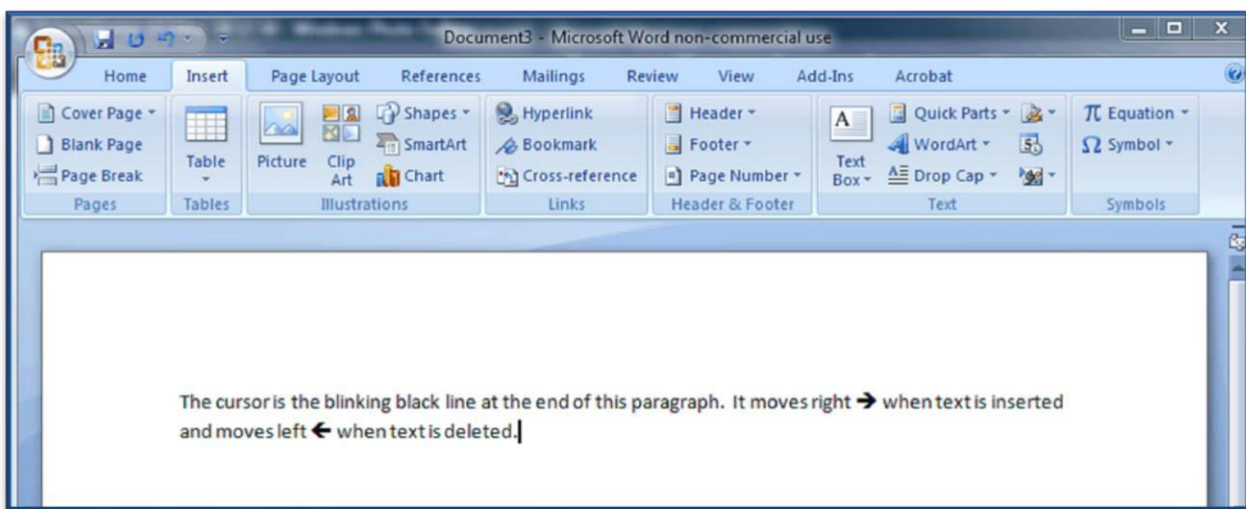
You can use MS Word help to remind you how to complete certain tasks or to refresh your memory. It is always there as a reference tool and is free to use.

The Cursor

If you've spent any time at all on the computer, you should be familiar with the cursor. In Word, it is a thick blinking black line. It is the line that lets you know where you are in the document and where characters will appear as you type. You can move the cursor anywhere in a document by using the arrow keys on the keyboard or by moving the mouse pointer to the desired location and clicking the left mouse button.

The cursor moves to the right as you type and to the left as you backspace.

The cursor can be quickly repositioned to certain preprogrammed places in the document by using keyboard shortcuts. To move the cursor to the beginning of a line, use the **HOME** key. To move it to the end of a line, use the **END** key. To move the cursor to the beginning of the document, use **CTRL + HOME** and to move it to the end of the document, use **CTRL + END**.



Inserting Text

You can **insert text** anywhere in a document simply by **moving the cursor to the desired location and typing**. Word automatically moves all text to the right of the cursor as you type.

Deleting Text


You can delete text by using the **Backspace** and **Delete** keys. But you must remember that Backspace removes characters to the left of the cursor, and Delete removes characters to the right. To delete blocks of text, select the text to be deleted and press either the Backspace or Delete key.


Moving Text

You can move selected text by holding the left mouse button and moving the cursor to the desired location in the document. When moving text, a small box appears near the mouse pointer and the cursor turns into a broken black line. You position the cursor in the place you want the text and release the left mouse button.

Undo and Redo

But let's say you accidentally delete something or deleted it and then decided that you want it back. You grit (patience) your teeth and start to grumble (complain), trying to remember the exact wording. It's a lost cause, right? Wrong. The makers of Word anticipated (likely) this problem and supplied an easy solution. The Undo button!

The Undo button can be found in the upper left corner of the program window in what Microsoft calls the "Quick Access Toolbar" The Undo button is the blue arrow shaped  like a comma. If you are not sure you've got the right button, you can move your mouse over it and wait and a small box that reads "Undo Typing (Ctrl-z)" will appear. Word allows you to undo up to 100 actions.

The Redo button is to the right of the Undo button. It looks like the recycle symbol you might see on trashcans at the airport. The redo button allows you redo an action that you just undid or to repeat the last action. If Word cannot redo the last action, the button will be faded. 

MS Word's basic Commands

New :- "New" command is use to open a new document. To open a "New" document click on "Office button" then click on "New" and lastly click on "Blank document" and then click on "Create" button

Open:- From "Open" command we can open saved files. For example we have saved a file by the name of "Microsoft word" and now we want to open it so firstly click on "Office button" and then click on "Open" and lastly select your file and click open and we can also open a file by the help of Shortcut key. To open a file by the help of shortcut key click "Ctrl+O".

Save:- This command is use to save a document. To save a file click "Office button" then click "Save" command and "type the name" that in which name you want to save this file. For example I want to save a file by the name of "Microsoft notes" and then I click "Save". The shortcut key for "Save" command is "Ctrl+S".

Save As:-"Save As" command is use to save a file by two or more than two names. What is difference between "Save" and "Save As" command. "Save" command is use to save a document by only one name but In "Save As" command we can save a file by two or more than two names. For example we have already saved a file by the name of "Microsoft Word" and now we want to save it by two names or more than two names as given below. The shortcut key for "Save As" is "F12".

Print Preview:- Print preview command is use when you want to see your document that how it look when I will print it. To see your document click on "Office button" then bring your mouse pointer to "Print" and lastly click on "Print Preview" command.

Print:- This command is use for print a file. If you want to print a file so First of all, click on “Office button” then click on “Print” command or click “Ctrl+P” from keyboard.

Close:- To close a document in Word, click the File tab and choose the Close command or use the keyboard shortcut Ctrl+W. Word banishes the document from its window, but then the program sits there and waits for you to do something else, such as start working on a new document or open a document you previously saved.

Copy:- To copy a piece of data to a temporary location. In word processing, for example, copying refers to duplicating a section of a document and placing it in a buffer (sometimes called a clipboard). The term copy differs from cut, which refers to actually removing a section of a document and placing it in a buffer.

Cut:- Cut is a command that allows you to "cut out" a selection of text or other data and save it to the clipboard. It is similar to the Copy command, but instead of just copying the data to the clipboard, it removes the selected data at the same time.

Paste:- Paste is a command that allows you to insert data from the clipboard into an application. In order to use the Paste command, you must first use either the Copy or Cut command to save data to the clipboard. Once the clipboard contains data, you can paste the saved data into any supporting program.

Find:- Find command is used to find specific text in a document you can also use Ctrl+F shortcut key.

Replace:- Replace command is used to replace a name or word in a document.

Undo:- Reverses the last change made in the text. This feature restores moved text to its original location.

Redo:- Reverses the last Undo action.

Bold:- Prints text darker than other copy as it is keyed. Bold may be added after text has been keyed by first selecting the text.

Borders:- Adds a border to any or all sides of a page, paragraph, or column, as well as to a table or a cell within a table. Page border options may include small pictures. Borders not only enhance appearance, but also make text easier to read by emphasizing certain passages. Borders are most effective when used sparingly.

Bullets:- Highlights each item in a list with a heavy dot or other character, as demonstrated in this list of features. Bullets add visual interest and emphasis.

Change Case:- Changes capitalization. The lowercase option changes all selected text to lowercase; the UPPERCASE option changes selected text to all capitals. The Sentence case option capitalizes the first letter of the first word, and the Title Case option capitalizes the first letter of each selected word.

Drop Cap:- Formats paragraphs to begin with a large dropped capital letter. Drop caps are objects (pictures) that can be formatted and sized.

Font:- Consists of the typeface, style, size, and any effects used. Font features may be changed before or after text is keyed. The number and size of fonts available depends on the software and printer used.

Size:- Measured in points such as 10 point or 12 point. One point is about 1/72 of an inch.

Effect:- Added to give text a special look. Examples of effects are shadow, emboss, and small caps. Effects should be used infrequently.

Italic :- Prints letters that slope up toward the right. Italic may be added after text has been keyed by first selecting the text.

Numbering :- Shows the proper order of a series of steps or items. Use numbers instead of bullets whenever the order of items is important.

Shading:- Adds color or patterns to paragraphs or table cells to emphasize them and focus the reader's attention on the contents.

Superscript:- Places text slightly higher than other text on a line. Superscript is commonly used for footnotes and endnotes, and for mathematical formulas and equations.

Underline:- Underlines text as it is keyed. An underline may be added after text has been keyed by first selecting the text.

Watermark:- Prints any desired text or graphics behind the main text on document pages.

Formatting Features :- Formatting features change the arrangement, or layout, of pages. These features include aligning text vertically and horizontally, indenting and hyphenating text, and changing margins and line spacing.

Alignment:- Refers to the horizontal position of a line of text (also called justification).

- Use left alignment to start text at the left margin (even left margin).

- Use right alignment to align text at the right margin (even right margin).
- Use center alignment to center text between the left and right margins.
- Use justify alignment for even left and right margins.

Center Page :- Centers text between top and bottom margins of a page. This feature leaves an equal (or nearly equal) amount of white space above and below text. Inserting two hard returns below the last keyed line gives centered documents a better appearance.

Hyphenation:- Automatically divides (hyphenates) words that would normally wrap to the next line. When used with left-aligned text, hyphenation makes the right margin less ragged, making text more attractive.

Indent:- Moves text away from the left or right margin.

Left indent:- (paragraph indent) Moves the text one tab stop to the right (usually 0.5"), away from the left margin.

Hanging Indent :- Moves all but the first line of a paragraph 0.5" (one tab stop) to the right. Hanging indent is commonly used on bibliographies and reference pages.

Line Spacing:- Changes the amount of blank space between lines of text.

Margins:- Changes the amount of blank space at the top, bottom, right, and/or left edges of a page. Default margin settings are not the same for all software.

Page Break:- Inserts a soft page break automatically when the current page is full. A hard page break can be inserted manually to start a new page before the current page is full.

Tab:- Aligns text according to the type of tab set. By default, most word processing software programs have left tabs already set at half-inch (0.5") intervals from the left margin. These preset tabs can be cleared and reset.

Left Tab:- Commonly used to align words, align text evenly at the left by placing the keyed text to the right of the tab setting.

Right Tab:- Commonly used to align whole numbers, align text evenly at the right by placing the keyed text to the left of the tab setting.

Decimal Tab:- Aligns numbers at the decimal point, regardless of the number of places before or after the decimal point.

Dot Leader Tab:- Automatically place dot leaders (. . .) between columns of designated text. The leaders lead the reader's eyes from text in the left column to text in the right column.

Text Wrapping:- Positions text in relation to an inserted object (graphic). Text may appear above and below, around, behind, or in front of an object.

Insert Features:- Various commands on the Insert menu are used to add page numbers, dates, headers, footers, and footnotes and to insert symbols, clip art, files, pictures, shapes, text boxes, and word art.

Shapes:- Adds a variety of shapes (rectangles, stars, banners, arrows, flow chart symbols, etc.) to a document.

Clip Art:- Drawings, pictures, sounds, and video clips can be inserted into documents. A collection of clip art files is provided with word processing software; additional clip art can be purchased or downloaded from the Internet.

Date :- Automatically inserts the date into a document. Some software contains an Update option that automatically replaces the previous date with the current date each time the document is opened or printed.

Footnote and Endnote:- Identifies sources quoted or paraphrased in the text or gives extra information about the main text. Word processing software automatically positions and prints each footnote at the bottom of the same page as the reference to it. It prints endnotes on a separate page at the end of a report. When footnotes or endnotes are edited, added, or deleted, the software feature automatically makes the necessary changes in numbering, formatting, and page breaks.

Header and Footer:- Adds text (such as a chapter title, date, filename, or name of a person or company) or graphic (a company logo, for example) in the top margin (header) or bottom margin (footer) of a page. Headers or footers often include page numbers.

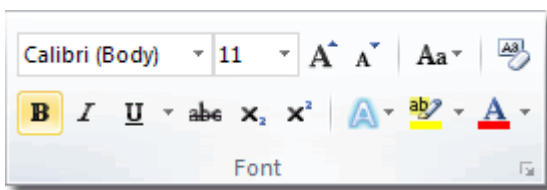
Page Numbers:- Places page numbers in a specified location on printed pages. Most software contains a variety of numbering styles from which to choose: Arabic numerals (1, 2, 3), lowercase Roman numerals (i, ii, iii), uppercase Roman numerals (I, II, III), uppercase letters (A, B, C), and lowercase letters (a,b,c). Numbers can be placed at the top or bottom of the page, and aligned at the left margin, center, or right margin. The Hide or Suppress option keeps the page number from appearing on a specified page.

Text Boxes:- Frequently used for labels or callouts in a document. Once a text box is inserted in a document, it can be formatted, resized, and moved.

Word Art:- Changes text into a graphic object. Most word processing software programs have a word art gallery that contains predefined styles such as curved or stretched text.

Formatting-Text and Document

Formatting text in Microsoft Word involves tasks like bolding the text, italicizing it, and changing the font and size. The commands to perform all of these formatting tasks are found on the Home tab in the Font group. Select your text and then click on the required formatting button to see the effects.



Alternatively, you can use the keyboard shortcuts for those commands. Here are the shortcuts for some of the more commonly used formatting commands:

- ❖ Bold: ctrl-b
- ❖ Italic: ctrl-I
- ❖ Underline: ctrl-u
- ❖ Font style.
- ❖ Font size.
- ❖ Font color.
- ❖ Use super script and subscript option to change in text.
- ❖ Fill color behind text using text highlighter color.
- ❖ Using change case we can change the text style like sentence case, uppercase, lowercase & toggle case.

In this way a lot of changing of text can be done in MS Word.

Document formatting refers to the way a document is laid out on the page—the way it looks and is visually organized—and it addresses things like font selection, font size and presentation (like bold or italics), spacing, margins, alignment, columns, indentation, and lists.

Word has two basic types of Word document formatting styles: Paragraph Styles - They apply to a minimum of an entire paragraph, and contain paragraph formatting (alignment, indents, etc.) and character formatting (font, color, etc.).

There are other formatting by which formatting can be done in a document in the MS Word, they are as follows:-

Page Setup and Margin Formatting

The settings of output page when it is printed on a printer are called page formatting. These include page size, page orientation, page margins, etc. Page formatting is defined in Page Setup under File menu. When Page Setup command is applied.

Formatting Spacing

The standard spacing format for a paper is double-spacing. Double-space the entire document, including the reference page.

1. Select the Home tab.
2. From the Paragraph menu, you may adjust the line spacing from the Line Spacing dropdown menu.

Hanging Indent

Select the text that requires a hanging indent OR create a hanging indent before typing the text. From the Home menu, click on the Paragraph command. This will bring up the Indents and Spacing tab. Under Indentation in the Special box, click on Hanging and OK.

Styles

A style is a pre-defined or user defined group of formatting options united under one common name which can be applied to a piece of text. Most of the time you use the 'normal' style to format text.

In a style, there are stored all formatting options of a paragraph or of a piece of text such as:

Font family:-The form of letters

Font face:- Additional visual characteristics of letters such as bold, italic, underlined etc.

Font size:- The size of letters defined in so-called points (pt). Normal body text has most of the time the size of 10 or 12 pt.

Paragraph margins:- The distances to the left and right paper edge, to the previous and next paragraph, between the lines in the paragraph.

Headings

The best idea to get a structure in a text is to divide it into chapters, sections, subsections etc. Each of the sectioning parts starts with its heading - title.

There are 9 levels of pre-defined headings in Word. Each of them has a slightly different size and appearance. You should format the chapters with the highest level (heading 1), sections with the heading 2, subsections with the heading 3 etc. This way you have a clear hierarchical structure in your text which allows you to:

- Change in a uniform way the appearance of chapter or section headings in the whole document
- Effectively use the View option Outline
- Automatically produce the Table of Contents etc.

Lists

Use lists to format the enumeration of similar elements such as:

- Concepts
- Key words
- Presentation notes etc.

Word supports two kinds of lists:

1. Bulleted lists and
2. Enumerated lists.

Tables

Beside lists, tables are a very powerful way of visual presentation of data sets. For example, all spreadsheet programs work with tables.

Tables consist of cells organized in rows and columns. Each cell contains a data field. A row means a horizontal, a column a vertical grouping of cells.

Tables are a more compact way of visualizing various types of information about several objects of the same kind.

Automatic Table of Contents

After you made proper structuring of your document using Headings, Word has the ability to make a Table of Contents automatically. The entries in a Table of Contents are at the same time hyperlinks to the referred headings. That means that a left mouse button click on an entry in the Table of Contents will put the insertion mark exactly on the beginning of the corresponding section heading in the text.

Comments

If you put specific information about any word, table, picture or a heading for this you can use the comment option in MS Word.

Sorting and Tables

Sorting Information Word's sorting tool allows you to sort text a number of different ways in your document. You can sort tables, dates, paragraphs, and more with just a few clicks in Word.

To sort a table in Word, click into the table to sort. Then click the “Layout” tab of the “Table Tools” contextual tab in the Ribbon. Then click the “Sort” button in the “Data” button group to open the “Sort” dialog box. You use this dialog box to sort the table information.

You can sort a table in Word that is used for storing and organizing data. It is possible to sort a table in Word by one or more columns of data in the table. You can sort a table in Word both alphabetically or numerically. You can also sort column data in either ascending (A-Z, 1-9) or descending (Z-A, 9-1) order.

To sort a table in Word, click into the table to sort. Then click the “Layout” tab of the “Table Tools” contextual tab in the Ribbon. Then click the “Sort” button in the “Data” button group to open the “Sort” dialog box. You use this dialog box to sort the table information.

If your table has column headers, or titles, for each data column, you can sort the table using those names. To show the field names in the “Sort by” drop-down menus, select the “Header row” option. This option appears in the “My list has” section at the bottom of this dialog box. If you don't have column headers, instead select the column number by which you want to sort the data. Select either the column names or numbers from the “Sort by” drop-down menus.

After choosing by which columns to sort, you then select the type of data in the column. Select the type of data contained in the column from the “Type:” and “Using:” drop-down menus. Your data type choices are text, number, or date. Then select in what order you want to sort the column's data: “Ascending” or “Descending.”

Sort a Table in Word

1. To sort a table in Word, click into the table to sort.
2. Then click the “Layout” tab of the “Table Tools” contextual tab in the Ribbon.
3. Then click the “Sort” button in the “Data” button group to open the “Sort” dialog box.
4. To show field names in the “Sort by” drop-downs for tables with column headers, select the “Header row” option.
5. This option appears in the “My list has” section at the bottom of this dialog box.
6. If you don't have column headers, instead select the column number by which you want to sort the data.
7. Select either the column names or numbers from the “Sort by” drop-down menus.
8. After choosing by which columns to sort, you then select the type of data in the column.
9. Select the type of data contained in the column from the “Type:” and “Using:” drop-down menus. Your data type choices are text, number, or date.
10. Then select in what order you want to sort the column's data. You can choose either “Ascending” or “Descending” order.

11. To sort a table in Word by multiple columns, repeat steps 7 through 10 within the “Then by” sections.
12. To perform the sort after deciding how to sort the table, click the “OK” button.

Working with graphics in MS Word

Microsoft Word allows to simply insert graphics in a document. Graphics are give life to the dull and boring text documents. Following types of graphics are used in Microsoft Word to make the document beautiful and attractive:-

Text Box

A text box is used to add a piece of text at any place in a document. For example, a text box may be used to add text in margins of a page, in the center of the page, etc. To enter a text box:

- On the Drawing toolbar, click Text Box.
- Click or drag in the document where text box is to be inserted.

Types of Graphics in MS Word

There are two types of graphics that are usually added in a word processor. These are bitmap graphics and vector graphics. **Bitmap graphics** are made up of pixels. These can be thought of as painted pictures. **Vector graphics** are made up of lines and curves. These can be thought of as drawings.

Bitmap Graphics

Bitmap graphics are added to a document by selecting Insert/Picture option from the menu bar. Pictures can be inserted from a file, Clip Art, scanner. Charts and Word Art can also be added as pictures. When a picture is selected, the picture tool bar appears. This toolbar can also be accessed by selecting View>Toolbars>Picture. The picture toolbar is used to change shape, color, brightness, size, etc. of the selected picture.

Vector Graphics

Vector graphics consist of lines, curves, etc. These are drawn using the Drawing toolbar. Clicking the Drawing icon W in the standard toolbar displays the Drawing Toolbar.



The Drawing toolbar contains buttons for creating and inserting different graphics, like lines, circles, rectangles, text boxes, WordArt and Auto shapes. It also has tools for editing and formatting these objects. Various tools for editing include fill tool, line-style tool, shadow and 3-D options. The Draw button at the leftmost corner of the drawing toolbar is used to select and draw drawing objects. The Drawing toolbar also provides a collection of predefined objects called Auto shapes that are commonly used in documents. Following figure shows some of these shapes. To draw a vector graphics:

1. Click drawing icon to display drawing toolbar.
2. Click on Select Object icon.
3. The mouse pointer will change.

4. Click the mouse on a drawing object, for example click on line object to select it. The mouse pointer will change to a cross-hair shape.
5. Click and draw the object.

WordArt Graphics:

WordArt is a small program that runs in the Word environment. It is used to turn ordinary words into graphics objects. It can convert text into a variety of shapes, stretch letters, rotate words, add shading, colors, borders and shadows to text. A WordArt image is just like any other graphic object in Word.



Once created, it can be edited by using the standard graphics editing tools like move, copy, re-size, crop, etc.

Inserting a Word Art Object

To insert a WordArt object:

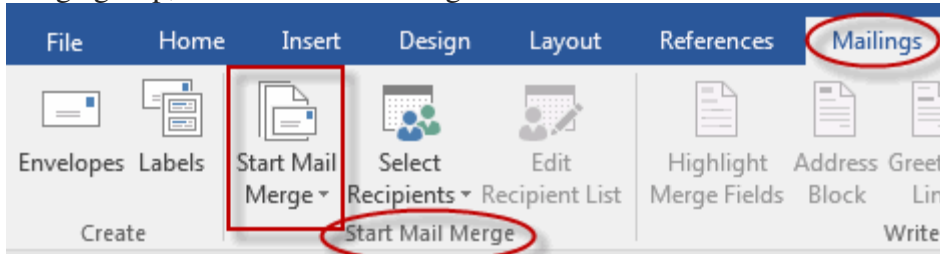
1. Click the WordArt icon in the Drawing Toolbar, or choose WordArt from the Picture submenu in the Insert menu. The WordArt Gallery dialog box will appear.
2. Choose the text effect that is to be applied and click OK. This will display Edit WordArt Text window.
3. Choose the options for Font type, style, and size.
4. Enter the text to which graphic effects are to be applied.
5. Click OK and WordArt will create the text effect and place it at the insertion point.

Introduction to mail merge

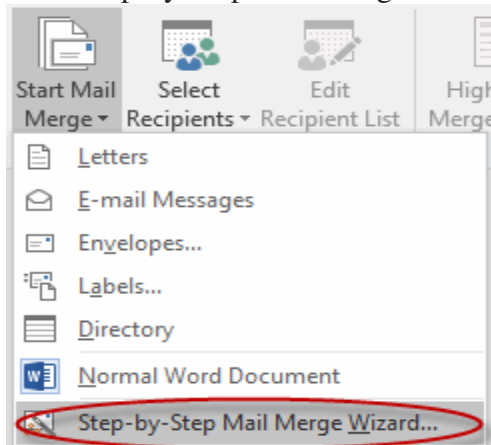
Mail merge is a feature within most data processing applications that enables users to send a similar letter or document to multiple recipients. It enables connecting a single form template with a data source that contains information about the recipient's name, address and other predefined and supported data.

Steps to use mail merge

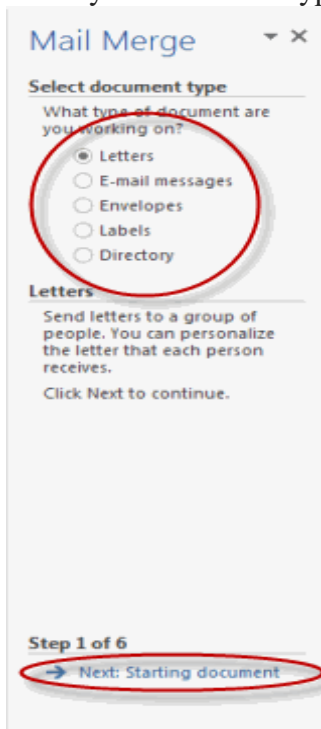
1. In a blank Microsoft Word document, click on the Mailings tab, and in the Start Mail Merge group, click Start Mail Merge.



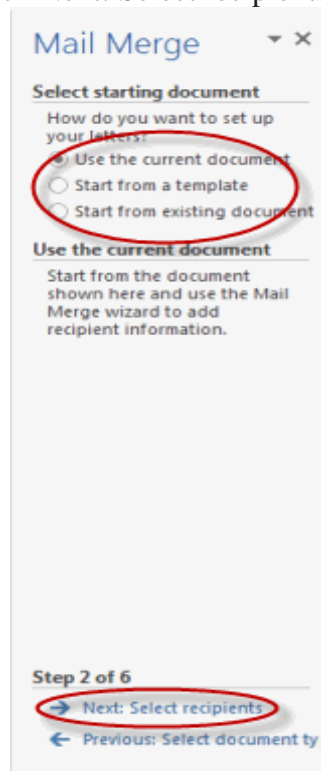
2. Click Step-by-Step Mail Merge Wizard.



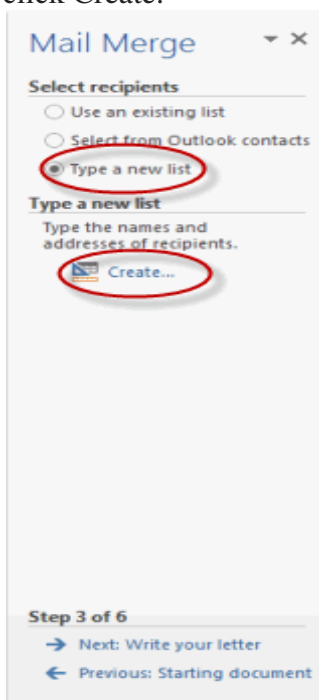
3. Select your document type. In this demo we will select **Letters**. Click Next: Starting document.



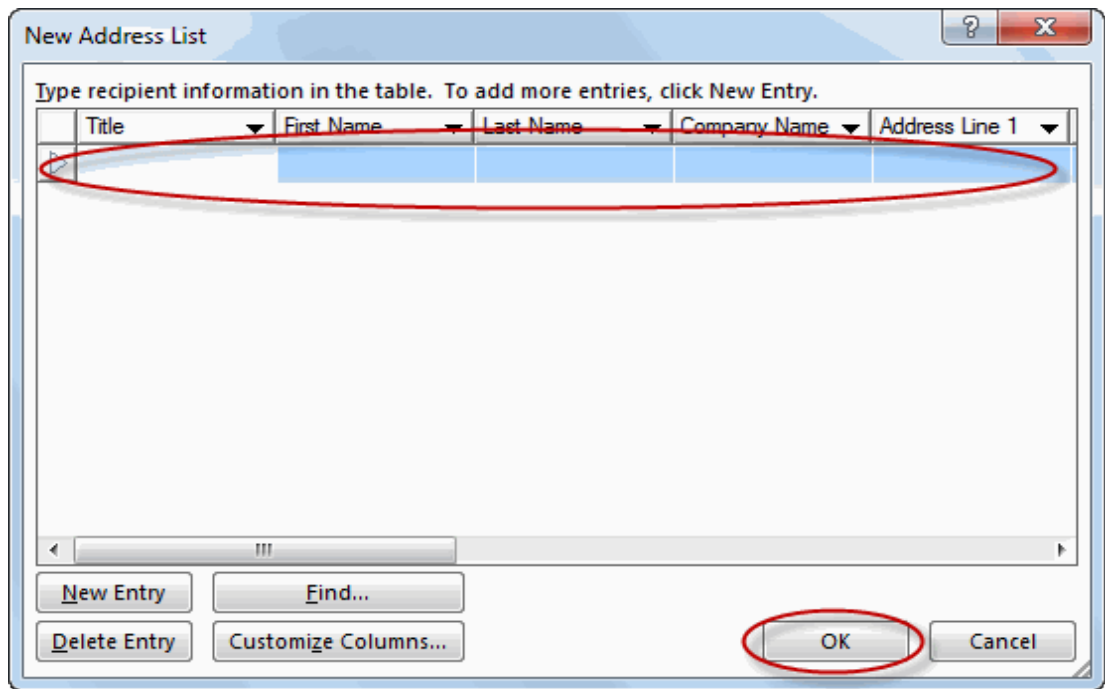
4. Select the starting document. In this demo we will use the current (blank) document. Select **Use the current document** and then click Next: Select recipients.



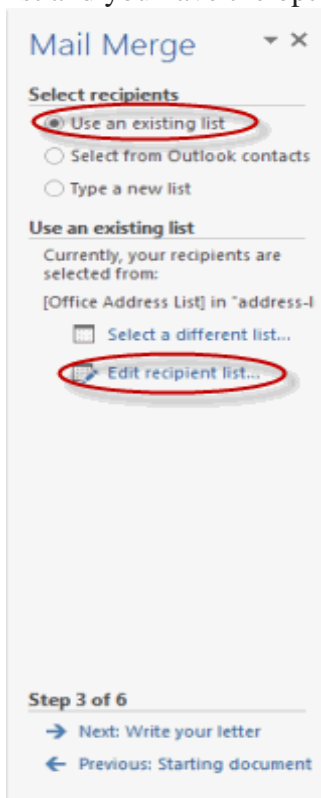
- Note that selecting Start from existing document (which we are not doing in this demo) changes the view and gives you the option to choose your document. After you choose it, the Mail Merge Wizard reverts to Use the current document.
5. Select recipients. In this demo we will create a new list, so select **Type a new list** and then click Create.



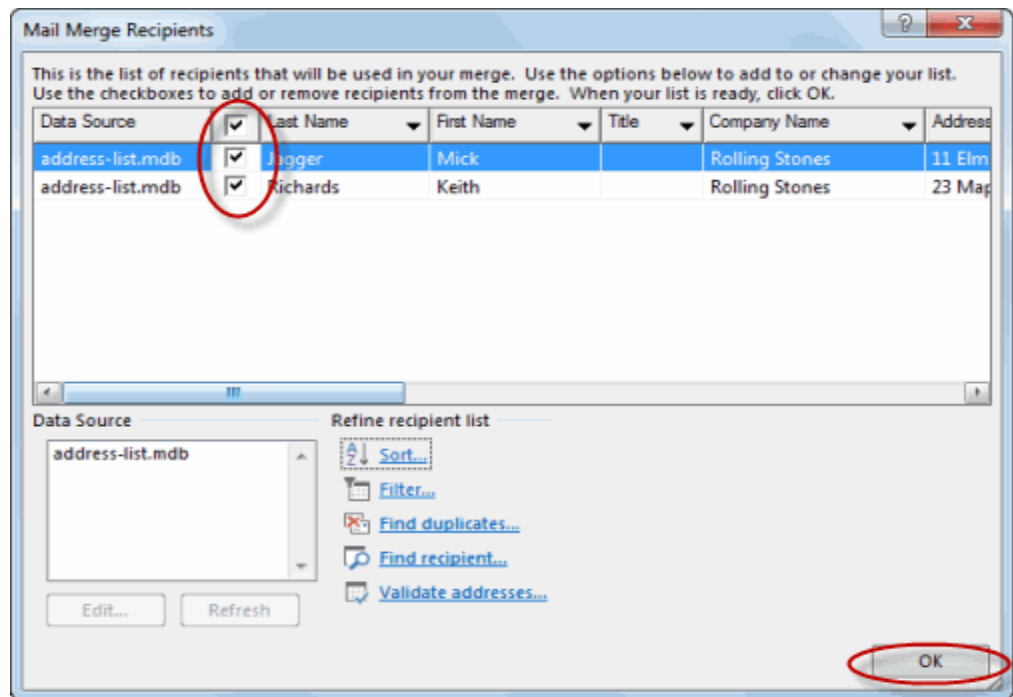
- Create a list by adding data in the New Address List dialog box and clicking OK.



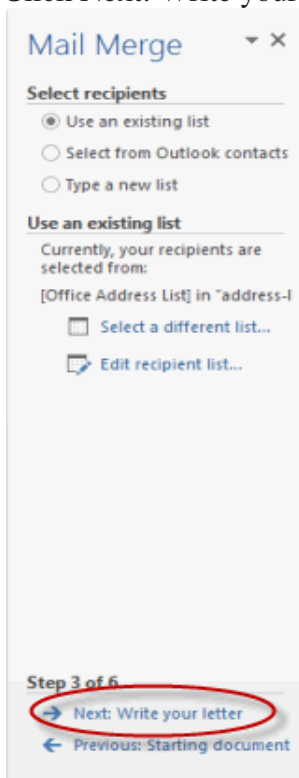
- Save the list.
- Note that now that a list has been created, the Mail Merge Wizard reverts to Use an existing list and you have the option to edit the recipient list.



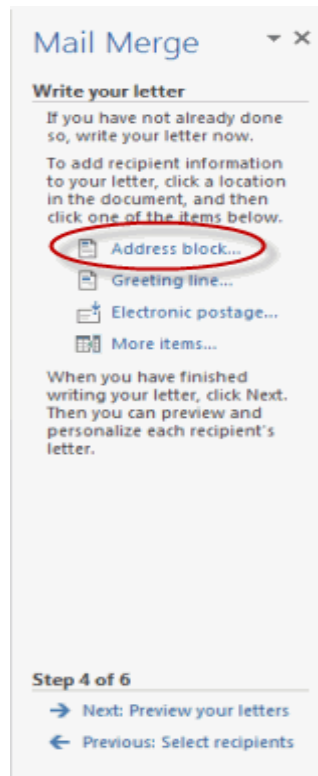
- Selecting Edit recipient list opens up the Mail Merge Recipients dialog box, where you can edit the list and select or unselect records. Click OK to accept the list as is.



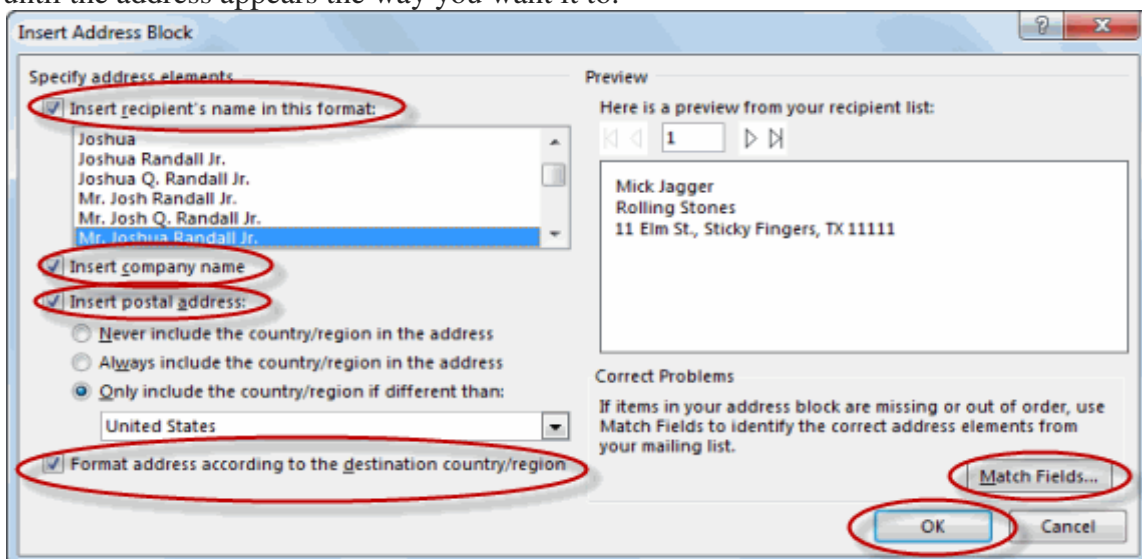
- Click Next: Write your letter.



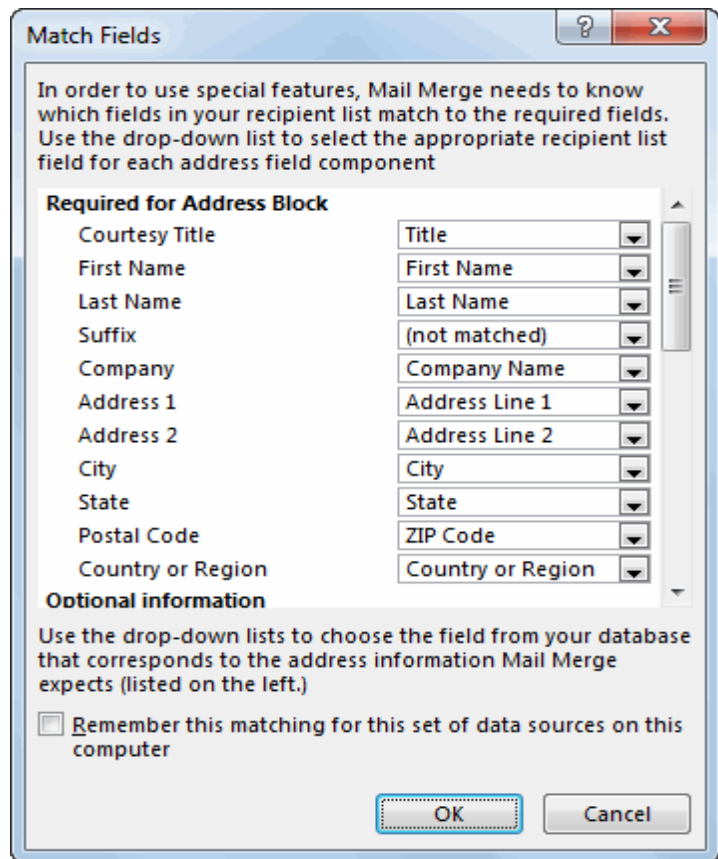
6. Write the letter and add custom fields.
- Click Address block to add the recipients' addresses at the top of the document.



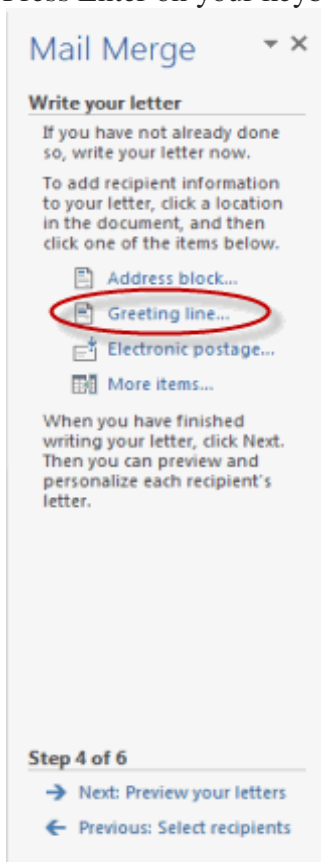
- In the Insert Address Block dialog box, check or uncheck boxes and select options on the left until the address appears the way you want it to.



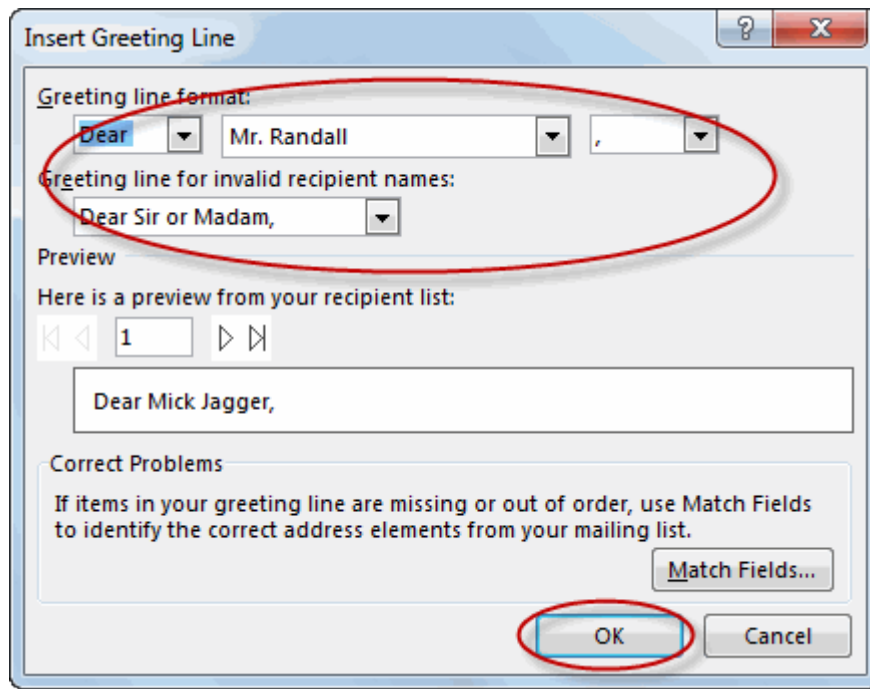
- Note that you can use Match Fields to correct any problems. Clicking Match Fields opens up the Match Fields dialog box, in which you can associate the fields from your list with the fields required by the wizard.



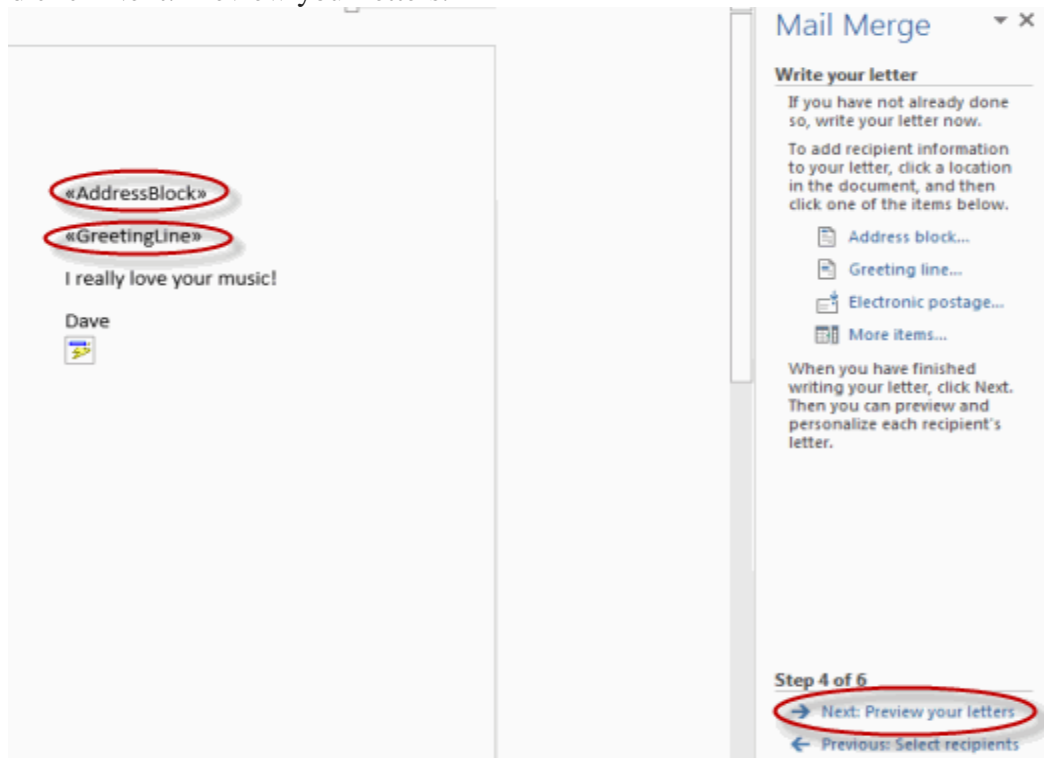
7. Press Enter on your keyboard and click Greeting line... to enter a greeting.



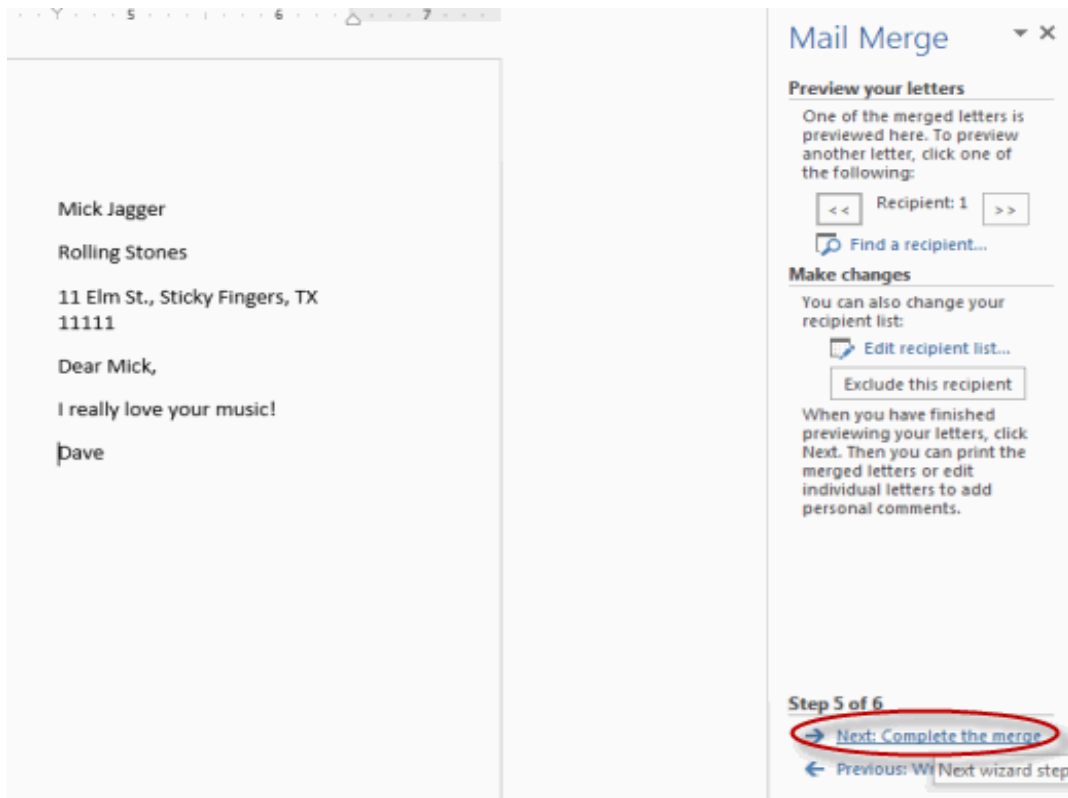
8. In the Insert Greeting Line dialog box, choose the greeting line format by clicking the drop-down arrows and selecting the options of your choice, and then click OK.



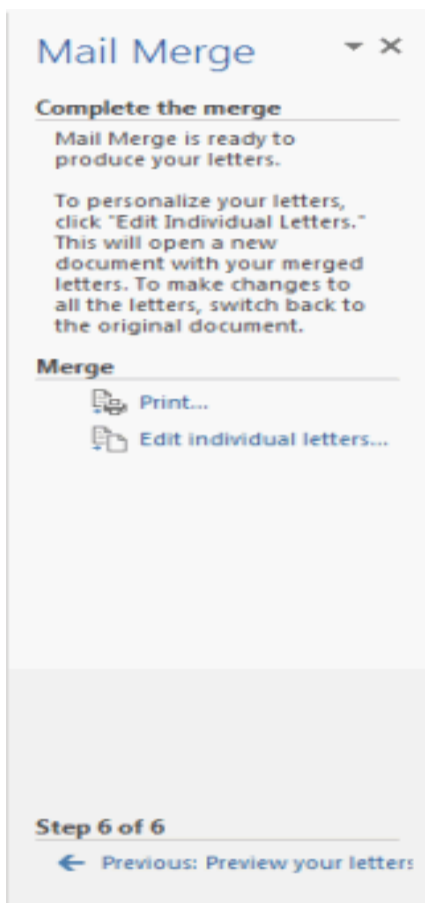
9. Note that the address block and greeting line are surrounded by chevrons (« »). Write a short letter and click Next: Preview your letters.



- Preview your letter and click Next: Complete the merge.



- Click Print to print your letters or Edit individual letters to further personalize some or all of the letters.



What is a Citation?

A citation is how you quote the source of the ideas within the body of the research paper. A citation is usually put at the end of the sentence within brackets. Generally, this citation includes information such as the author's surname, date of publication or the page number in which the part you have taken the idea appears in the original book. This citation method also changes according to the format you are following. See the following examples.

APA – 'Her past would not let her be at peace (Martin, 2014).'

MLA – 'Her past would not let her be at peace (Martin 251).'

The purpose of citation is intellectual honesty. You want to give due credit to a particular author for the quotation you have picked from his work in support of your work. Citation occurs at various places wherever there are relevant quotations.

What is a Bibliography?

Bibliography is very easy to understand. It stands out from the other contents of the thesis or the dissertation as it appears as a list towards the end of the paper. The main purpose of compiling the bibliography is to let the reader know the books and journals you have referred to as a researcher in writing your thesis or dissertation. A bibliography is a list that contains all the sources the writer has used when he or she was writing the paper. When we say all the sources, the sources that were actually quoted or paraphrased in the body of the paper and the ones that were just consulted but not used in the body of the paper are both included. So, bibliography is not just a list of sources the writer has used. It is a complete list of sources as it even includes the ones that the writer only read to have an idea of his or her topic. A bibliography is in the alphabetical order. That is usually decided by the writer's surname. The format of the bibliography is as same as the format of the paper. Say that you are writing the paper in APA format. Then, the bibliography is also in the APA format. If the format is MLA, then, the bibliography is also MLA. Here are some examples.

APA:

Bibliography

Martin, G. B. (2014). *The Stormy Night*. New York: Cyclone Publishers.

What is APA?

APA stands for American Psychological Association. The general format of the APA style of research paper writing consists of,

- a title page
- page numbers
- abstract
- headings

What is MLA?

MLA stands for Modern Language Association. The general format of a MLA paper consists of,

- thesis statement
- page numbers
- evidence

What is the difference between Bibliography and Citation?

- A bibliography is a list of all the sources the writer has used when writing the paper. This includes the sources that were actually cited or paraphrased in the text as well as the ones that the writer just read to have an idea of the topic.
- A citation is how you quote the source of the ideas within the body of the research paper.
- Both bibliography and citation differ from each other in terms of their purpose too. The main purpose of compiling the bibliography is to let the reader know the books and journals you have referred to as a researcher in writing your thesis or dissertation. On the other hand, the purpose of citation is intellectual honesty. That is why you include the citation wherever you have quoted someone directly or paraphrased their ideas. This is another main difference between bibliography and citation.
- It is very important to know that both bibliography and citation jointly contribute to the formation of a thesis or a well-written dissertation.
- Bibliography usually appears towards the end of the thesis. On the other hand, citation occurs at various places wherever there are relevant quotations. Normally, the citation is placed at the end of the sentence.

These are the very important differences between bibliography and citation that a researcher has to know. As you can see, all the citations that you used in the body of the paper appear at the end of the paper, under the name bibliography, together with the sources that were not cited.

Definition: - Power Point is a Microsoft presentation program that creates a slide show of important information, charts, and images to display during a presentation. It is most often used for business and school presentations.

Benefits of Power Point

Power Point provides multiple benefits to users, including:

- It is widely used, and considered the "standard" for presentation software. If you create a Power Point presentation, it's more likely it will be easier for others to open and view.

- It includes many optional presentation features, including slide transitions, animations, layouts, templates, and more.
- It offers the option to export its slides to alternative file formats, including GIF and JPG images, MPEG-4 video, PDF, RTF (rich text format), WMV (Windows Media Video), and Power Point XML.

The Ribbon in Power Point



Once you get used to it, you'll find the Ribbon is far easier to use than the file menu interface used in previous versions of word. By default the Ribbon in Power Point is divided into seven tabs, with an optional eighth tab (Developer)

They are:

- **Home:** This contains commonly used Power Point features, such as creating new slides, changing layouts, deleting slides, formatting text and paragraphs, inserting shapes, and creating drawings.
- **Insert:** As you might guess, this one handles anything you might want to insert into a presentation, such as pictures, clip art, charts, tables, movies, audio clips, photo albums (a new feature), headers and footers, text boxes, WordArt ... etc.
- **Design:** From this tab you can apply different themes to your presentation; change colours, fonts and effects; select background styles; change the slide orientation and page setup...etc.
- **Animations:** A more accurate name for this tab might be transitions, because it lets you control the animations used as transitions between slides. Creating transitions is now very easy, just move your mouse cursor over an icon of a particular transition on this tab and you'll see a live preview of that transition. Click on that transition and you have added it to your slide. On this tab you can also create custom animations, change transition sounds and speeds, and make other changes to animations.
- **Slide Show:** On this tab you can do anything you need to with your slide show -- preview it, create a custom slide show, change the set up or rehearse your narration and timings.
- **Review:** To check spelling and grammar use the thesaurus, track changes, review other people's changes or compare documents? This is the tab to use.
- **View:** This tab allows you change the view in any way, including changing presentation views -- such as slide master, handout master, notes master and slide sorter. You can also show or hide a ruler and gridlines; zoom in and out; choose from colour, gray scale, or black and white; and display windows as either split or cascaded.
- **Developer:** To display this tab, click the **Office Button > Power Point Options > Popular > Show Developer tab in the Ribbon**. Most users will only need this tab to record Macros but it also lets you insert controls such as buttons and text boxes, if you have written your own code for them.

New features in Power Point 2007

The new look of Power Point is not the only change to the application, there are plenty of new

features as well.

No Right Hand Pane

In Power Point 2007, the right-hand pane from previous versions has been banished; its functions have been replaced by the Ribbon. When you create a new slide and want to apply a new layout to it, you now use the Design group in the Ribbon.

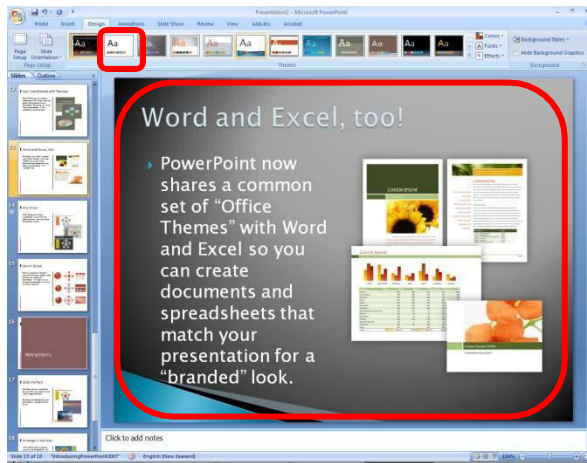
Themes

As in all of the Office 2007 applications, Themes have been introduced to allow users to quickly apply common colours, backgrounds and fonts to all elements of a presentation for a unified look. In previous Power Point versions, you had to make those changes separately for charts, tables and slides.

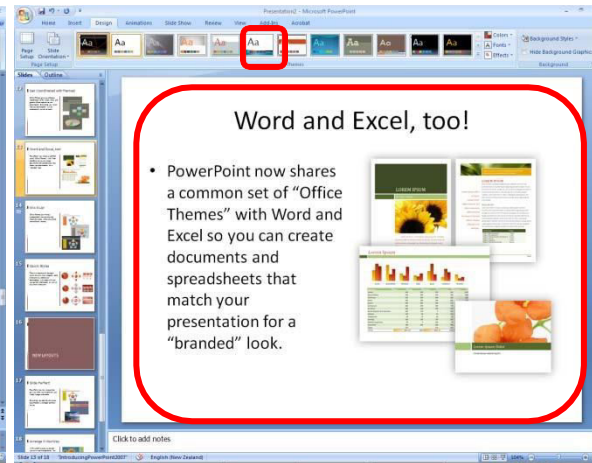
To apply a theme to a presentation, click the Design tab and move your mouse cursor over each theme to preview them, when you find a theme you like, click it to keep it.

Shown below is the same slide in a show shown with different themes selected for it.

Office theme



Concourse theme



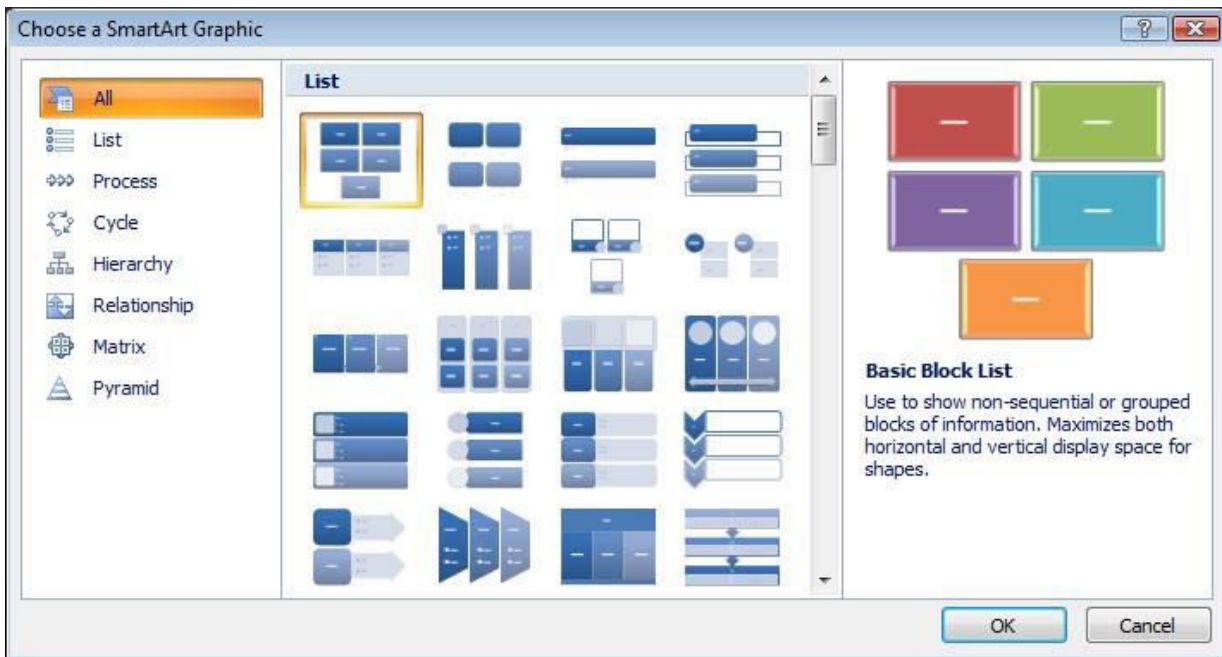
Be aware that Themes only work if you're using Power Point's new Office XML format; they won't work on old-style .ppt files.

Custom Slide Layouts

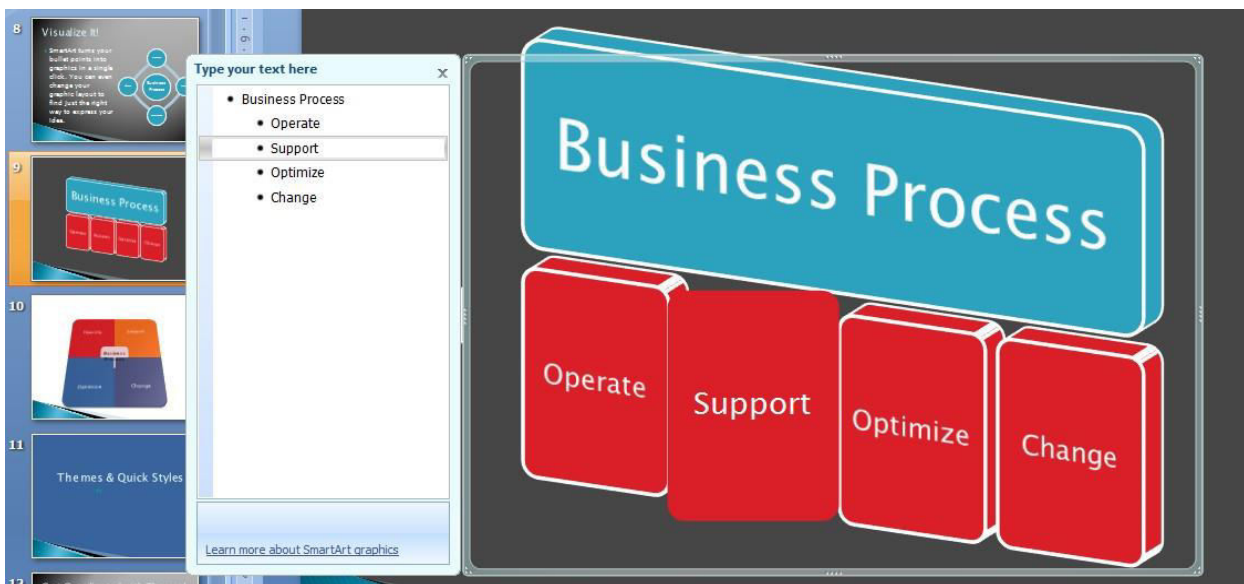
From the **View tab > Slide Master**, you can now create and save custom layouts for your presentations, including charts, movies, tables, pictures, Smart Art, clip art and other elements. You can also create and save multiple slide master sets that contain tailored layouts for different topics.

Smart Art

In Office 2007, you get far better prebuilt diagrams and charts via Smart Art. Choose the **Insert tab, > Smart Art**, and you'll be able to insert a wide variety of diagrams and charts, as you can see in the figure below. Click any of the available diagrams or charts, and you'll get a description of what it's best used for and how to use it. Smart Art offers many new types of diagrams and charts.



In addition, you can add a variety of effects, such as gradients and 3-D, to Smart Art and other Power Point graphics. You can also convert existing content to Smart Art, by right-clicking (on a bulleted list or similar) and selecting **Convert to Smart Art** you will be able to port the list into a prebuilt Smart Art template.



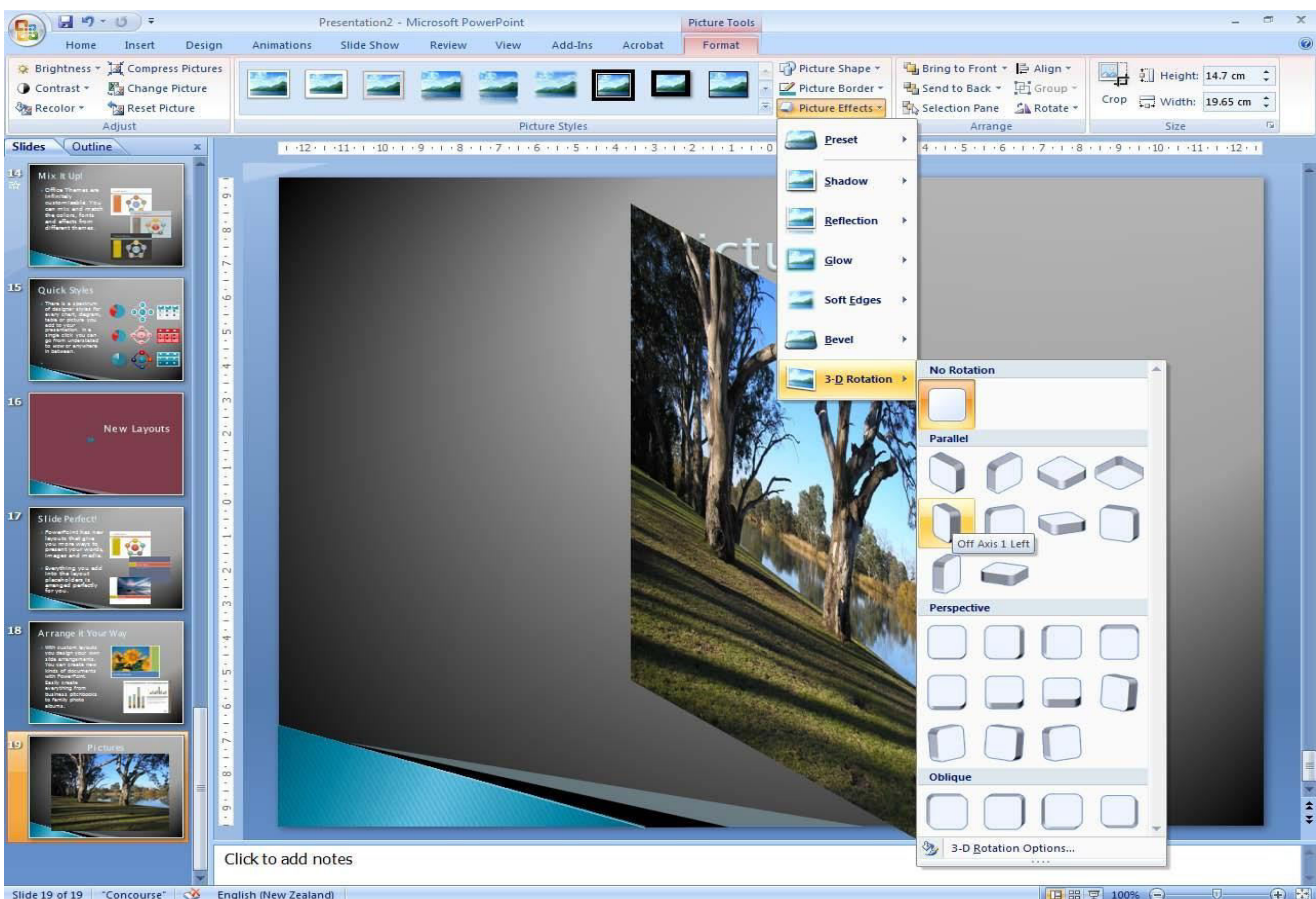
Better Multimedia Integration

It's now easier to incorporate multimedia into your presentations. On the **Insert tab**, choose either **Movie** or **Sound** from the Media Clips group, and insert your multimedia object. From here, you can also play a CD audio track during the presentation. And you can even record audio directly from within Power Point and embed it in your presentation.

New effects

There are a host of new and better effects in Power Point 2007, including warp, bevel, shadow, reflection, soft edges and rotating 3-D shapes. You can get to these effects in a variety of ways, depending upon where you're using the effect.

If you want to use an effect on a picture, put your cursor on the picture, select the **Format tab** > **Picture Tools** (this appears only when you've selected a picture). > **Picture Effects** drop-down list, select the effect you want to use.



Discontinuous text selection

This change is an example of a very small change that can save lots of time. You can now select multiple pieces of discontinuous text and then make changes to them. This works the same way as it does in Word, hold down the **Ctrl** key as you select your text, and then apply changes to all the text simultaneously.

Some More Basic Power Point Features

Power Point is the presentation software of the Microsoft Office software suite. One of the most widely used office programs; Power Point has applications for personal use, academics and business. It is highly customizable; you can edit Power Point presentations to be as personal or professional as you want. Power Point has a relatively simple user interface which you can easily pick up. However, mastering each of its functions will allow you to create impressive presentations

Design

The design features of Power Point allow you to customize the appearance and format of the slides. Power Point typically comes with a set of preloaded themes for you to choose from. These can range from simple color changes to complete format layouts with accompanying font text. Themes can be applied through the whole presentation or a single slide. Using the page setup allows you to optimize the presentation for the display size; for instance, you should use a larger screen ratio when displaying on a projector compared to a computer screen.

Animation

Power Point animation is divided between slide transitions and element animation. Using slide transition adds an effect when switching slides during a slide show. You can edit the transition effect and timing, as well as opt for an on-click or automatic transition between slides. Element animation adds movement and sounds to the objects within the slide. For example, if you're constructing a photo gallery as a slide show, you can choose which pictures enter the slide first, how they enter and add a sound as they enter.

Presentation

The presentation function of Power Point is largely designed to accommodate public speaking. Power Point comes with a built-in notes function; when printing out presentation slides, you can add presenter notes beside each slide as accompanying content. This is useful to clarify points in the slide without sacrificing the slide's readability. As of the 2007 version of Power Point, you can pre-record narration for a presentation. Power Point also has a rehearsal function as well, allowing you or your team to practice your timing and monitor the length of your presentation.

Integration

Power Point is compatible with all other software in the Microsoft Office suite; you can export slides into Word documents or use Excel charts within your presentation. In addition to image and audio support, Power Point 2007 also has video-integration functionality; you can embed videos within a presentation for easy playback without exiting the program. You can also export presentation files to an online interface for multi-user remote editing and presentation practice.

Creating a Power Point Presentations using Easy Way

Step 1:

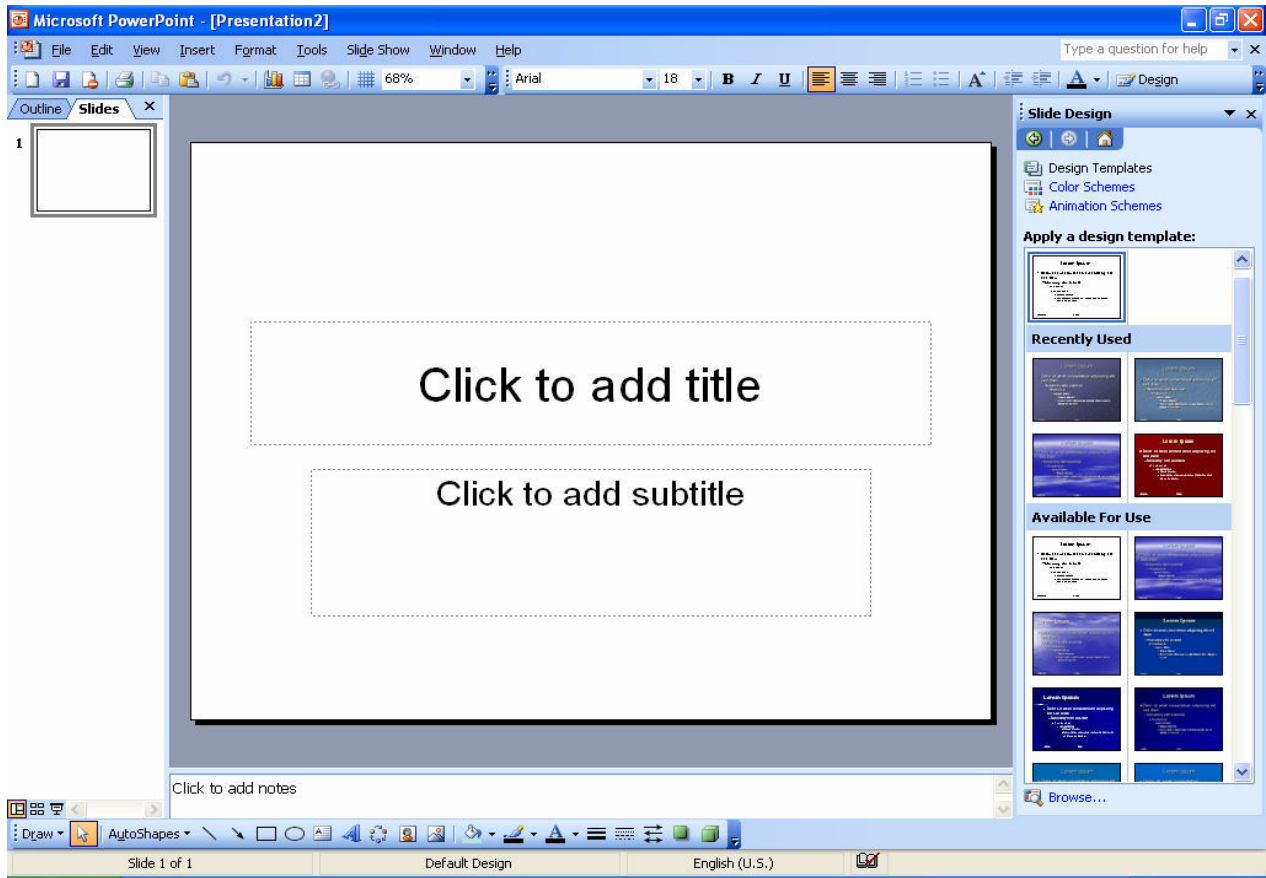
Open Microsoft Power Point.

Step 2:

Go to File at the top of the screen and click New. A box that says “New Presentation” should appear on the right side of your screen.

Step 3:

In the “New Presentation” dialog box, click on “From Design Template.” You may then scan through design templates and choose one that you like.

**Step 4: Slide Design**

Select a design template by clicking on the template you like. You may choose a different color for your template by clicking on “Color Schemes” in the “New Presentation” dialog box.

Step 5: Slide Layout

Change the Slide Layout. You may change the slide layout (how information is presented in the slide) by going to the top of the screen and clicking on “Format” – “Slide Layout.” A box will appear on the right side of your screen (where “New Presentation” appeared) labeled “Slide Layout.” You may select a design by clicking on it.

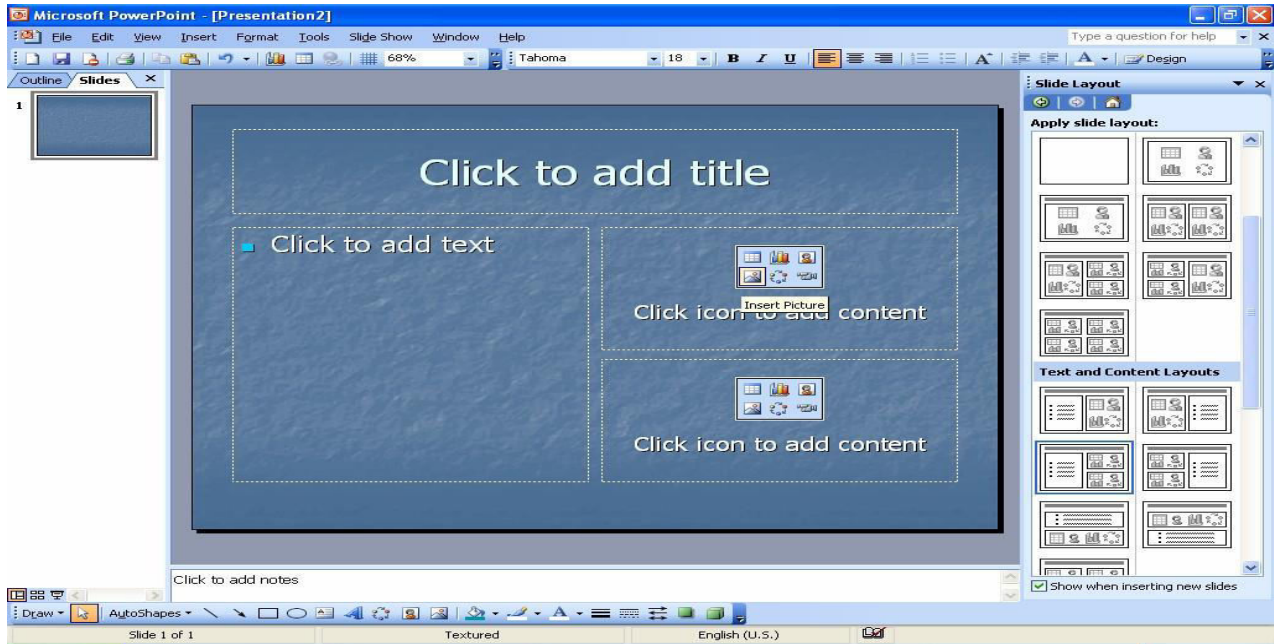
Step 6: Adding Text

Enter your text by clicking and then typing in the box titled “Click to Add Text” or “Click to Add Title.”

Step 7: Adding Pictures

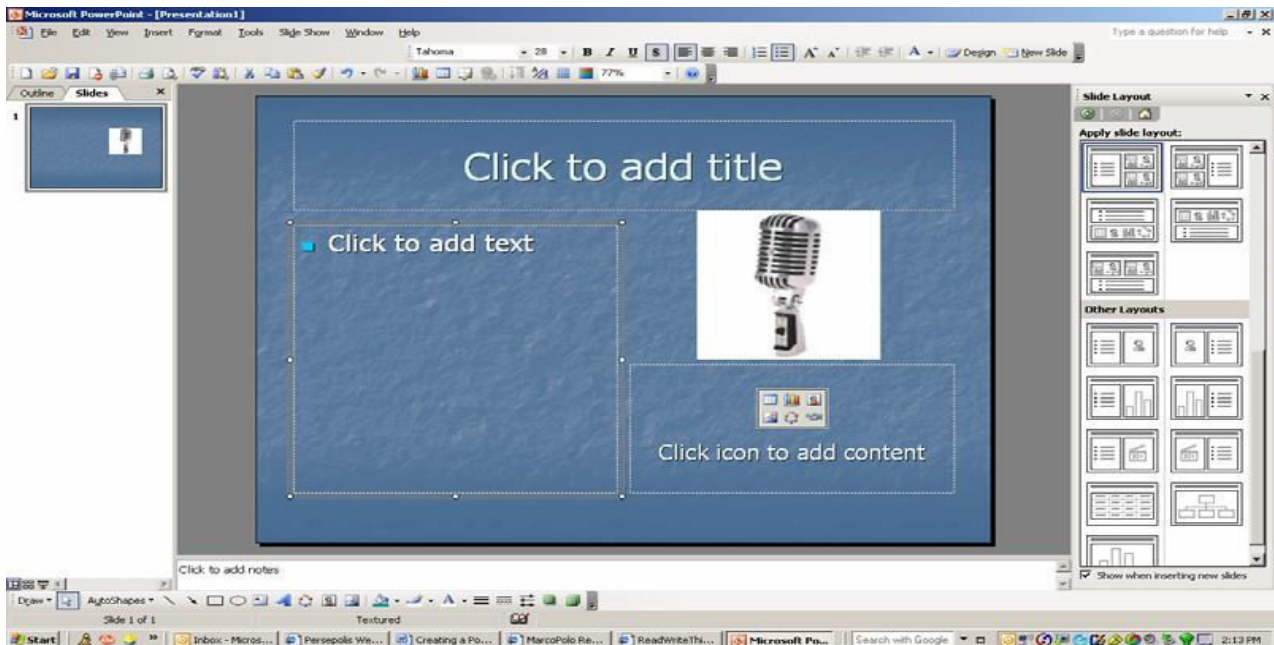
You may add pictures by clicking on the box that says “Click to add content.” Inside that box, there will be a smaller box with six icons. Click on the icon that looks like a photograph of a mountain. A

new window will open, allowing you to browse for a picture on your computer or a CD. Once you find your picture, click on it and then click “Insert.”



Step 8: Resizing Pictures

You may change the size of your picture by clicking on the picture. The picture will then have black lines around it with small bubbles or boxes in the corners. Place your mouse over the bubbles or boxes and click. Holding the mouse pointer down, drag the picture to the size you want.



Step 9: Now You are Done

Now you completely have done your Presentation using nine steps.

Working with graphics in Power Point

A graphic in Power Point as any image that is not a picture. This includes Smart Art, which is a type of native Power Point graphic and has special attributes. A Power Point graphic most typically refers to a grouping of shapes into an image commonly referred to as vectors.

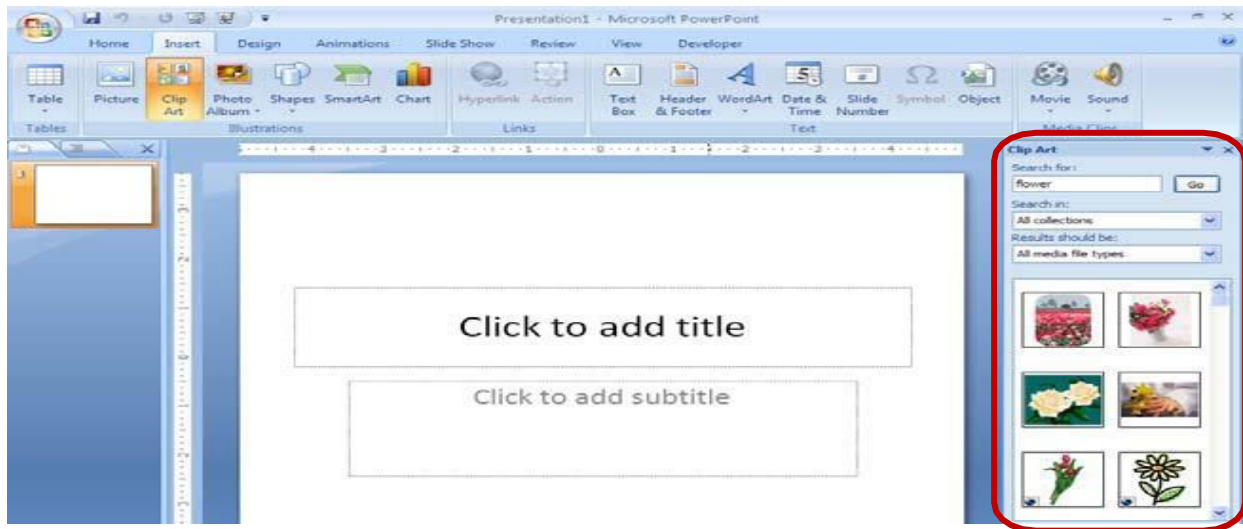
Working with various types of graphics in Microsoft Power Point. A design ‘rule of thumb’ is to include some sort of graphic on at least every third slide to aid in keeping the audience interested and focused.

Inserting Clip Art

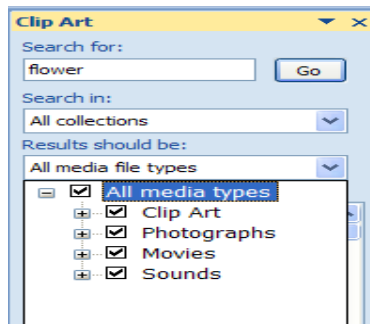
A gallery of clip art is available both within Power Point as well as on the Microsoft website. These graphics are free for use in Microsoft products.

Sound design theory recommends using clip art judiciously; it should not be used as space filler. Clip art was innovative a decade ago; now it is old news.

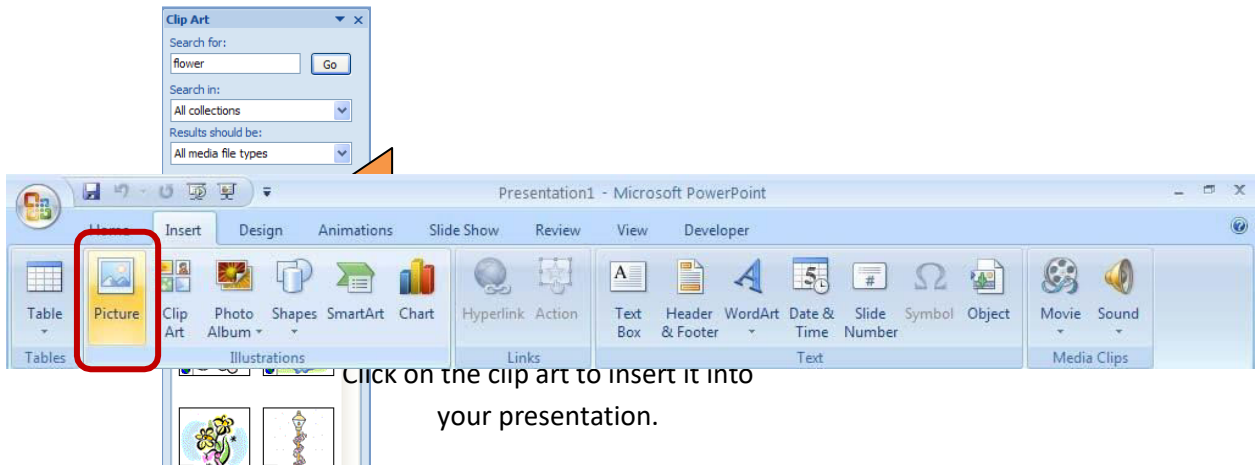
1. To insert clip art, go to the Insert tab and click Clip Art.
2. This brings up the Clip Art pane on the right side of the screen.
3. Enter a keyword and click Go to begin your search.



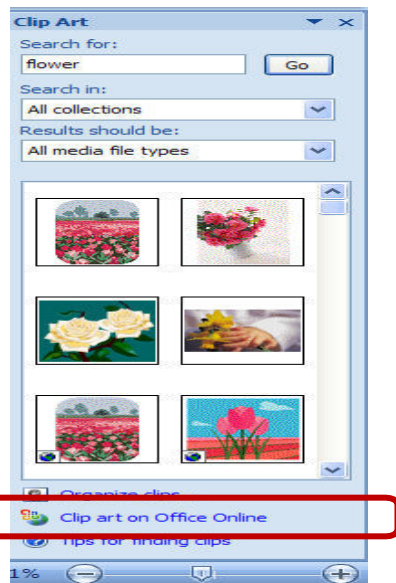
- a. You may limit the type of clip art for which you are searching – clip art, photographs, movies, or sound.



- b. To insert the graphic into your presentation, simply click on it.



4. If you are connected to the Internet, you can also search Office Online for additional clip art. In the Clip Art pane on the right side of the window, click Clip art on Office Online. This opens up Microsoft Office Online in a browser, and you can now search for additional clip art to download.

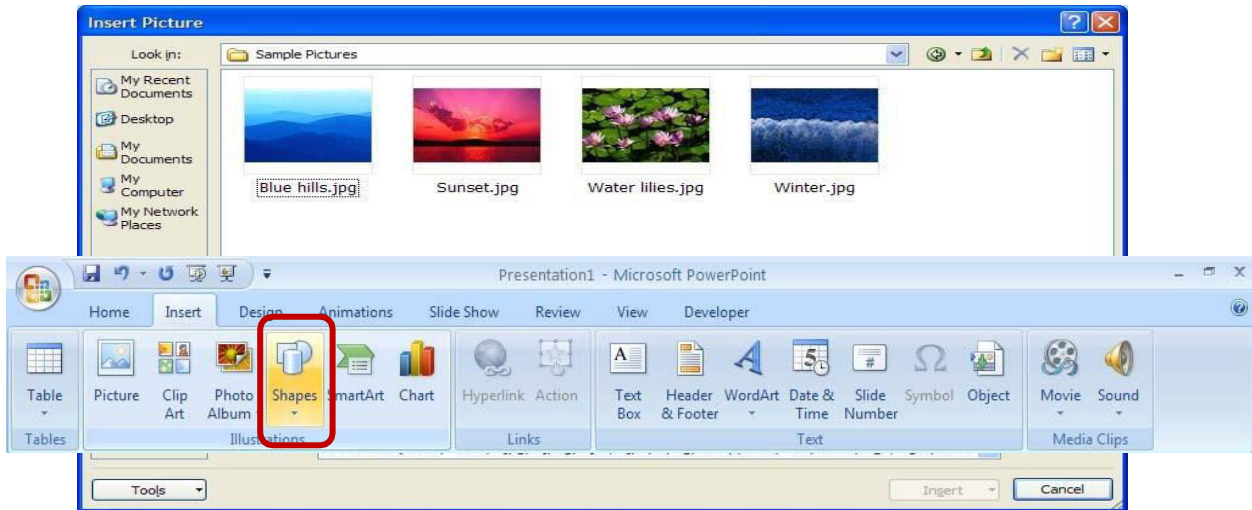


Inserting Pictures

You can also insert pictures that you have stored (for example, a photograph from a digital camera or a scanned image).

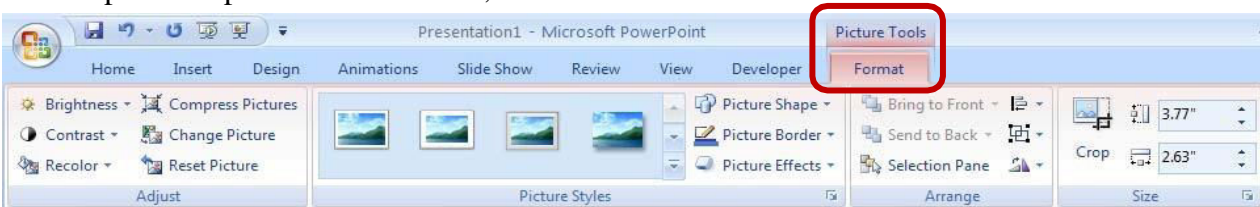
1. To insert a picture, go to the **Insert** tab and click **Picture**.

2. Browse for the location of your picture. Select it and then click **Insert**.



Working with Clip Art and Pictures

When clip art and pictures are selected, additional Picture Tools are available.



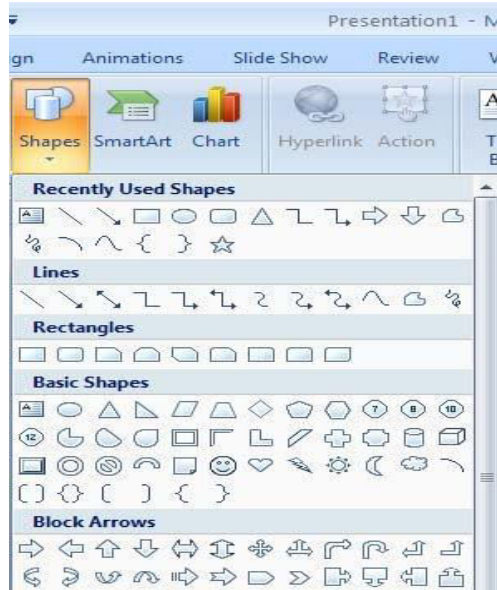
1. The **Compress Pictures** feature is a helpful tool to reduce the size of the presentation.
2. The features available in the **Picture Styles** group provide tools for enhancing your graphic.
3. The features available in the **Arrange** group allow you to select how you want to align, group, rotate, or layer the graphic.
4. The features available in the **Size** group allow you to crop a graphic as well to specifically indicate its height and width.
5. If these **Picture Tools** are not available, check to make sure the graphic is selected.
6. Microsoft Power Point was not designed to be an image editor. If significant alterations to a graphic are necessary, you should use dedicated image software, like Photoshop.

Drawing Shapes

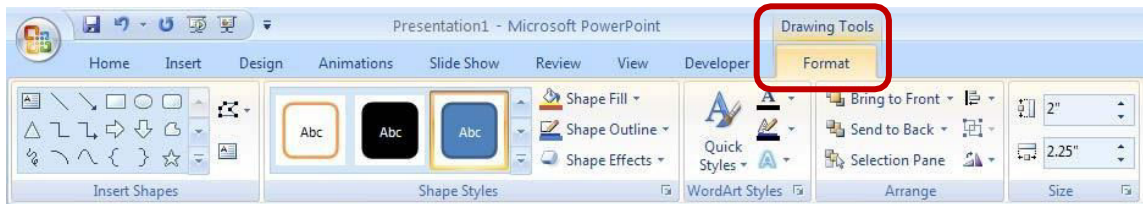
Microsoft Power Point comes with a variety of shapes, lines, arrows, callouts, and more that can be drawn in your presentation.

1. To draw a shape, go to the **Insert** tab and click **Shapes**.

2. Select a shape.



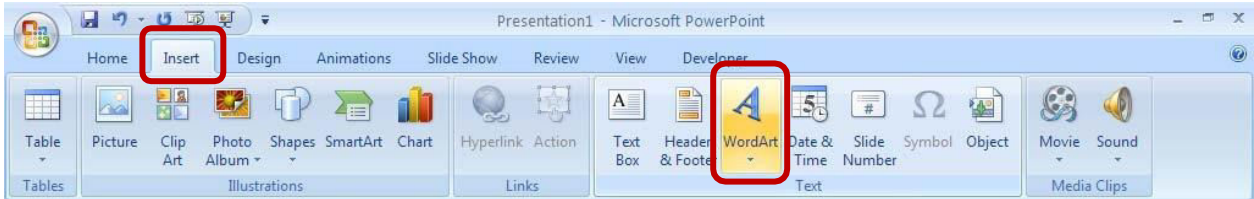
3. Move your mouse where you would like to draw the object, and while holding down the mouse button, draw the shape. The pointer will look like a plus sign (+) while drawing shapes.
4. Shapes can be moved by clicking and dragging.
5. When the shape is selected, additional features are available on the **Drawing Tools** tab.



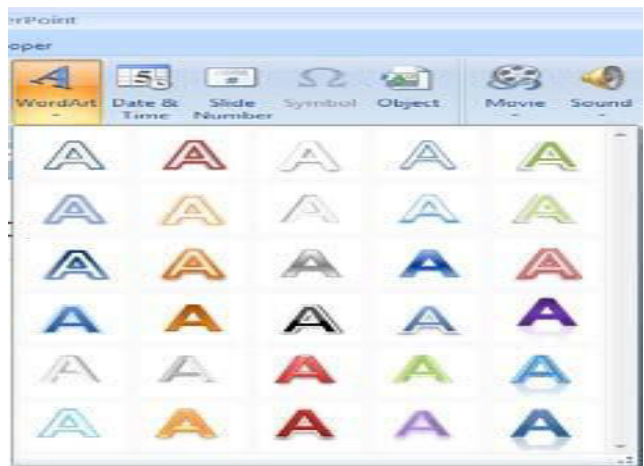
Inserting WordArt

WordArt is a gallery of text styles that you can add to your presentation to create decorative effects, such as shadowed or mirrored (reflected) text. You can also convert existing text into WordArt.

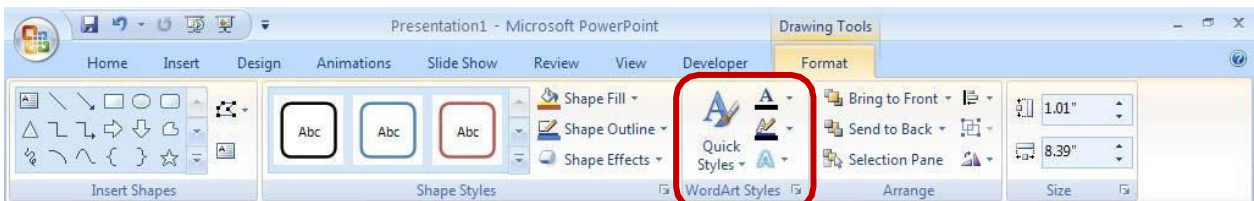
1. To insert WordArt, go to the **Insert** tab and click **WordArt**.



2. From the gallery, select the style of WordArt you wish to use.



3. Type the text into the placeholder.
4. WordArt can be moved by clicking and dragging, just as any other graphic.
5. When WordArt is selected, an additional **Drawing Tools** tab is available with formatting utilities to adjust the style, color, outline, and effect. Many of the effects are new to Office 2007, including reflection and 3-D rotation effects.



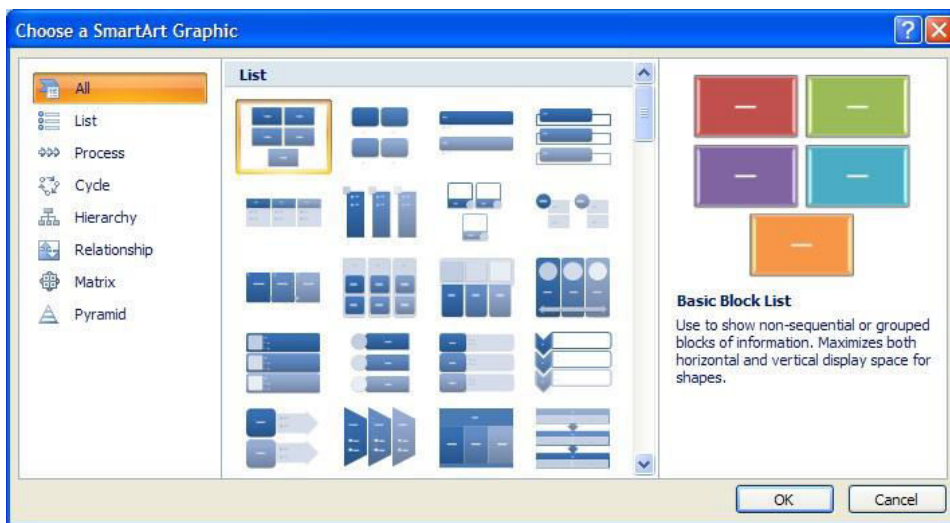
Inserting Smart Art

Smart Art, a visual representation of your message, is new to Office 2007. Smart Art allows you to create editable illustrations simply and without the aid of a professional designer. You can easily create dynamic workflow, relationship, or hierarchy diagrams.

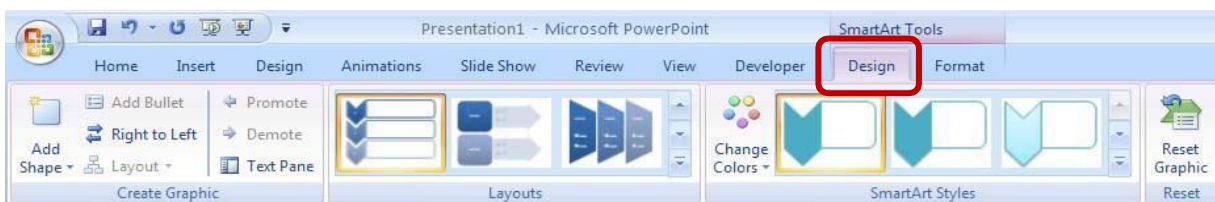
1. To insert Smart Art, go to the **Insert** tab and click **Smart Art**.



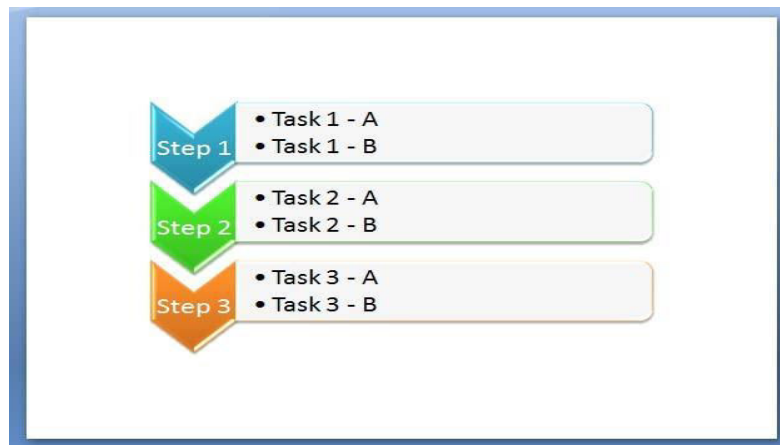
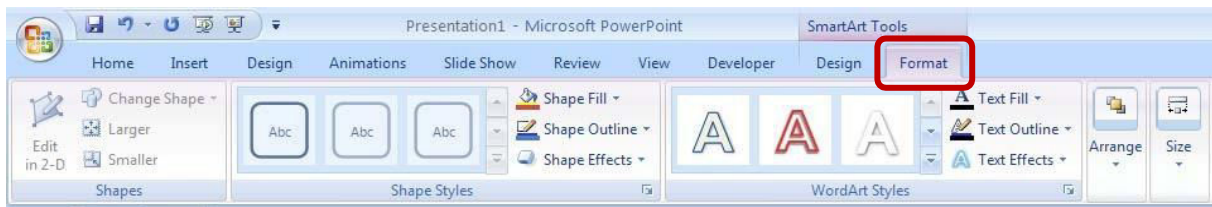
2. From the gallery, select the style of Smart Art you wish to add and click OK. The styles are organized into categories, from which you can browse via the list on the left. Or you can browse through all the styles by selecting the all category on the left. You will be able to customize the Smart Art in the next step.



3. Once you have selected the Smart Art style, you can begin customizing the graphic to meet your needs.
 - a. Click on the text placeholders to enter text.
 - b. Use the features available on the **Design** tab to add shapes, rearrange the shapes, adjust the layout, change colors, change the style, or reset the graphic to its original look.



- c. Use the features on the **Format** tab to adjust the formatting of the shapes.



Show time and Rehearse timing in power point

The scheduled or actual time at which a show of slides or linking of slides begins within a specific time this is set for them.

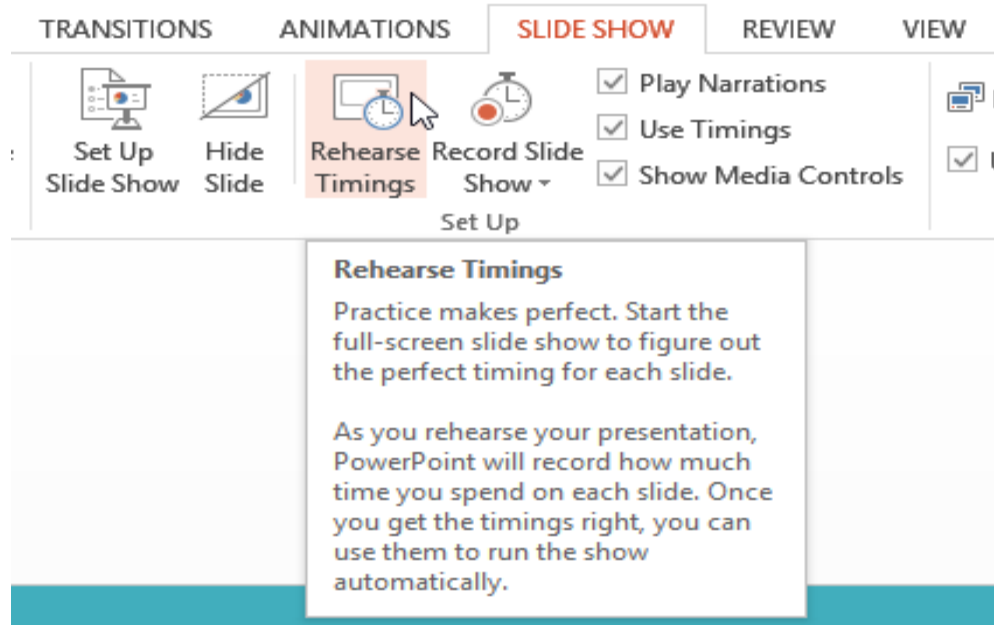
Rehearse timing is a feature which allows one to record the time of slide preparation. It automatically records the time and presents the slides in the orderly manner and keeps them on screen for the set time.

Rehearsing timings can be useful if you want to set up a presentation to play at a certain speed without having to click through the slides to present it. Think of it as a tool to help you practice presenting your slide show. Using this feature, you can save timings for each slide and animation. Power Point will then play back the presentation with the same timings when you present it.

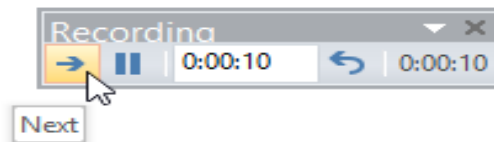
Rehearsing is often one of the last steps in creating a well-polished presentation. If you've never presented a slide show, you may want to review our lesson on Presenting Your Slide Show to become familiar with the basics.

To rehearse timings:

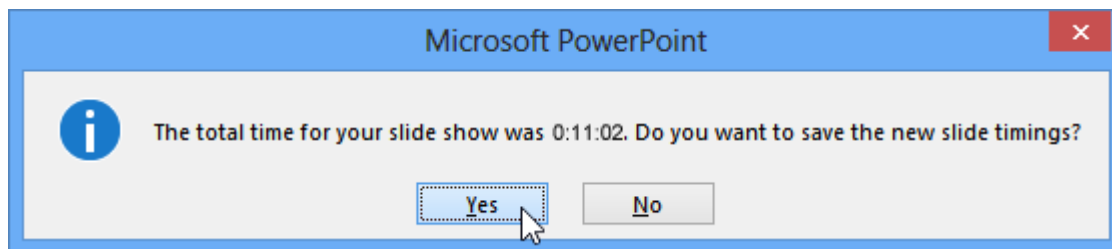
1. Select the **Slide Show** tab, and then locate the **Set Up** group.
2. Click the **Rehearse Timings** command. You will be taken to a full-screen view of your presentation.



- Practice presenting your slide show. When you are ready to move to the next slide, click the **Next** button on the **Recording toolbar** in the top-left corner. You also can also use the right arrow key.



- When you have reached the end of the show, a dialog box will appear with the total time of your presentation.



- If you are satisfied with your timings, click yes.

If you need more than one try to get the timings just right, the **Recording toolbar** has options to let you take a break or start over on a slide. To pause the timer, click the **pause** button on the toolbar. No actions taken while the timer is paused will be included in the timings. To re-record the timings on the current slide, click the **repeat** button.

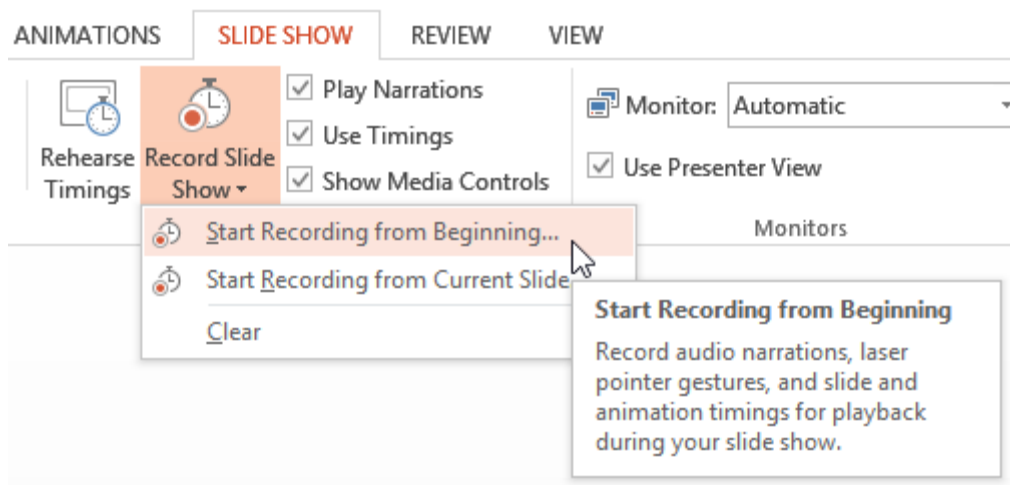


Recording your slide show

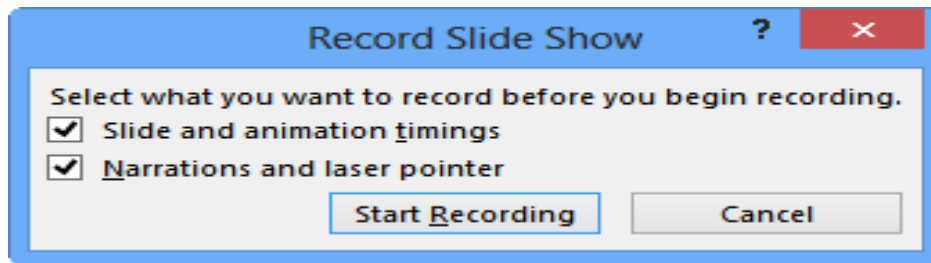
The **Record Slide Show** feature is similar to the **Rehearse Timings** feature, but it's a bit more comprehensive. If you have a microphone for your computer, you can even **record voiceover narration** for the entire presentation. This is useful if you plan on using your slide show for a self-running presentation or a video. Your mouse won't show up on screen in recorded slide shows, so if you want to **point out details on screen** you can use Power Point's **laser pointer** option.

To record a slide show:

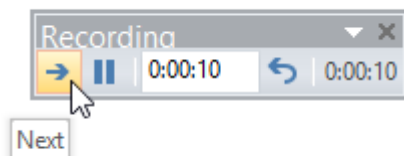
1. Click the Slide Show tab, then locate the Set Up group.
2. Click the Record Slide Show drop-down arrow. Select either Start Recording from Beginning or Start Recording from Current Slide.



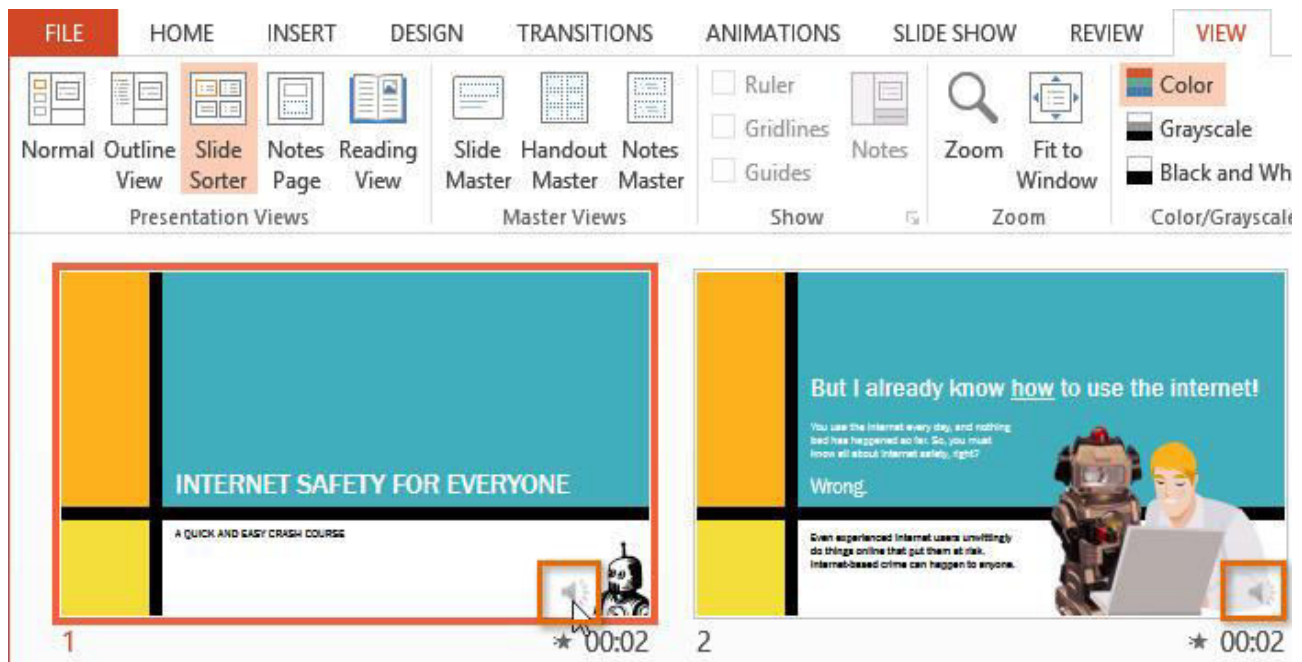
3. A dialog box will appear. Select the desired options. Remember, you can only record narration if you have a microphone attached to your computer.



4. Click Start Recording. Your presentation will open to a full screen.
5. Perform your slide show. Make sure to speak clearly into the microphone if you are recording narration. When you are ready to move to the next slide, click the Next button on the Recording toolbar in the top-left corner or use the right arrow key.

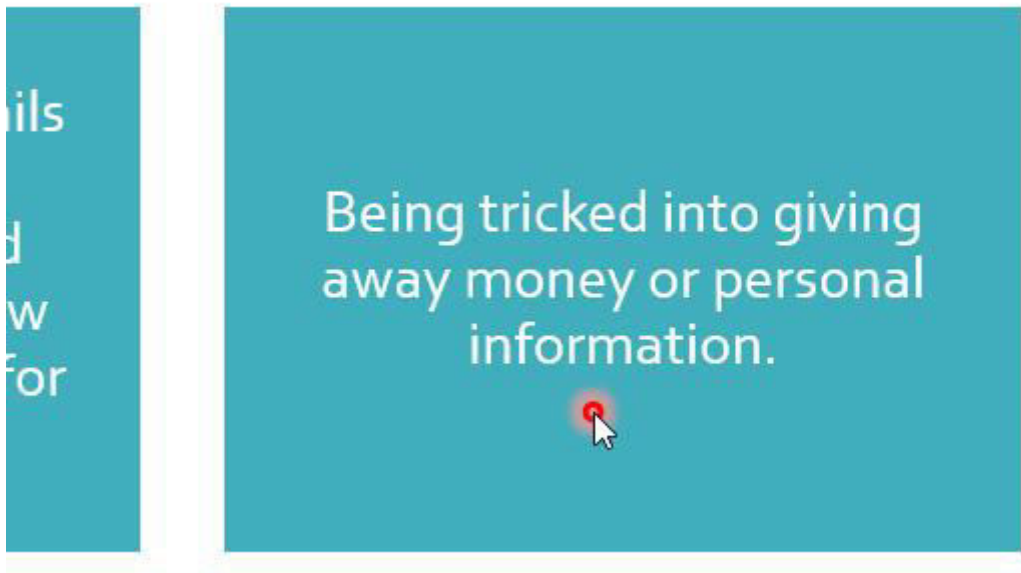


6. When you reach the end of the show, Power Point will close the full-screen view.
7. Your slide show timings and narration are now included in your presentation. The slides with narration will be marked with a speaker icon in the bottom-right corner.



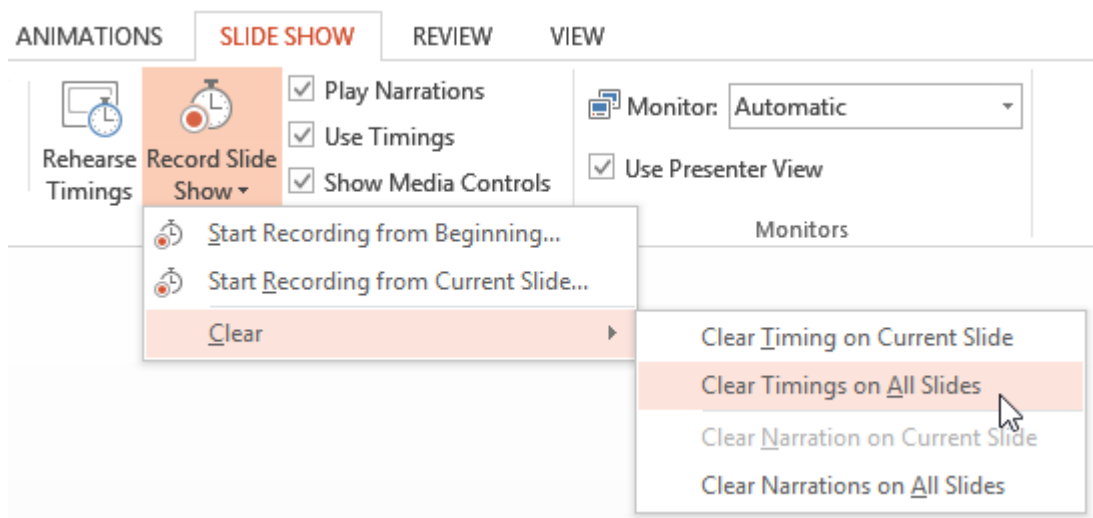
To point out details during your recording, press and hold the **Ctrl** key and the **left** mouse button. Your cursor will show up as a **laser pointer**. Simply move your mouse to indicate the

desired details. Release the **Ctrl** key when you are finished pointing out things on screen.



To remove narration or timings from a recorded slide show:

1. Click the Slide Show tab, then locate the **Set Up** group.
2. Click the **Record Slide Show** drop-down arrow.
3. Hover the mouse over **Clear**, then select the desired option.



Sharing presentation options

Power Point offers several options to enhance or even totally change the way you deliver presentations. Instead of presenting your slide show normally, you can choose to present it as

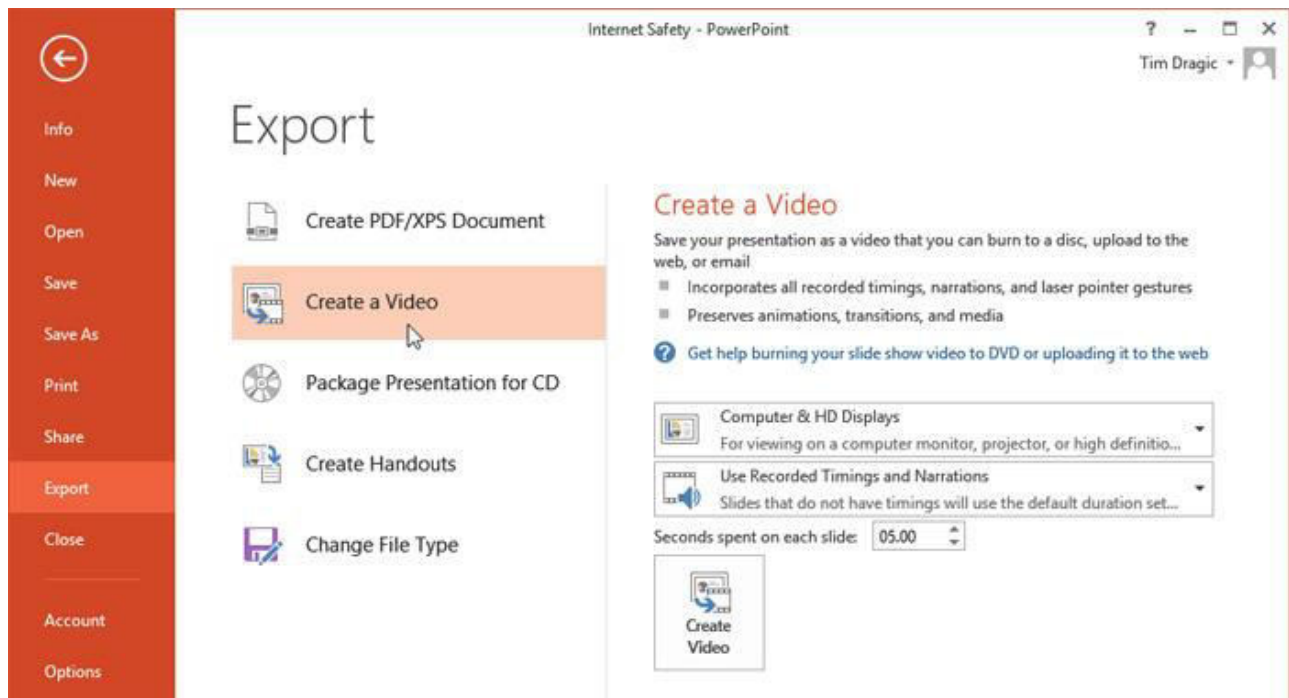
a **video**, or even **present it live online** so others can view it remotely. No matter how you choose to give your presentation, you can enhance it by **customizing** your slide show to remove or reorder slides. All of these options can help you give a polished and professional presentation.

Exporting a presentation as a video

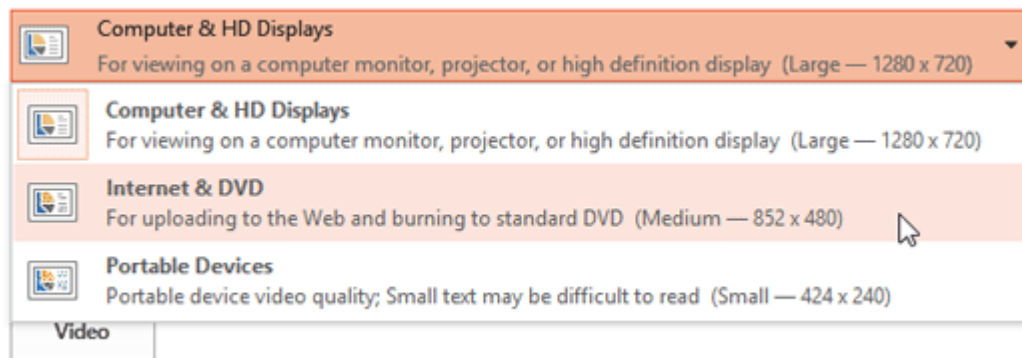
The **Create a Video** feature allows you to save your presentation as a video. This can be useful because it lets your viewers watch the presentation whenever they want. To make sure your viewers have enough time to view each slide, you might want to **rehearse** the timings or **record** your slide show before using this feature.

To create a video:

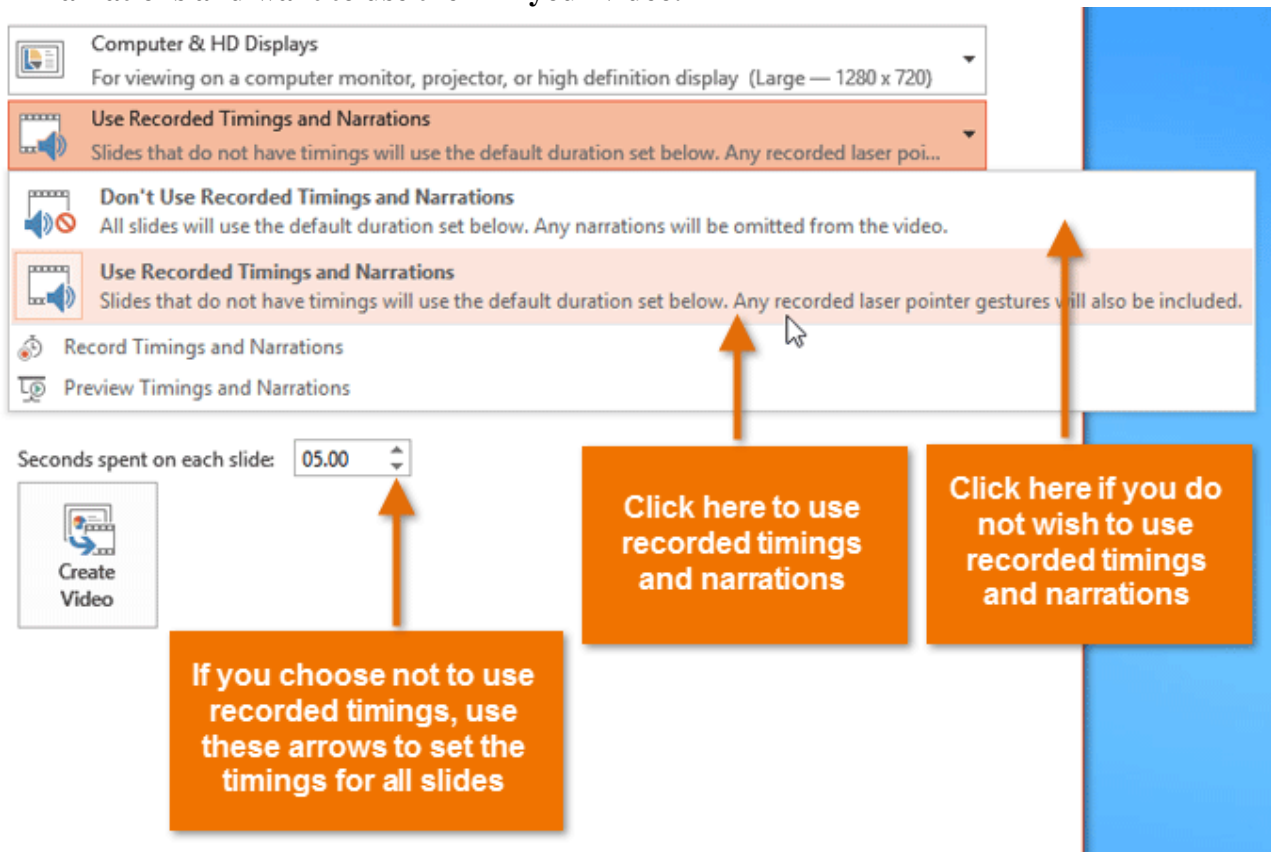
1. Click the **File** tab to access Backstage view.
2. Select **Export**, then click **Create a Video**. **Video export options will appear on the right.**



3. Click the drop-down arrow next to **Computer and HD Displays** to select the size and quality of your video.

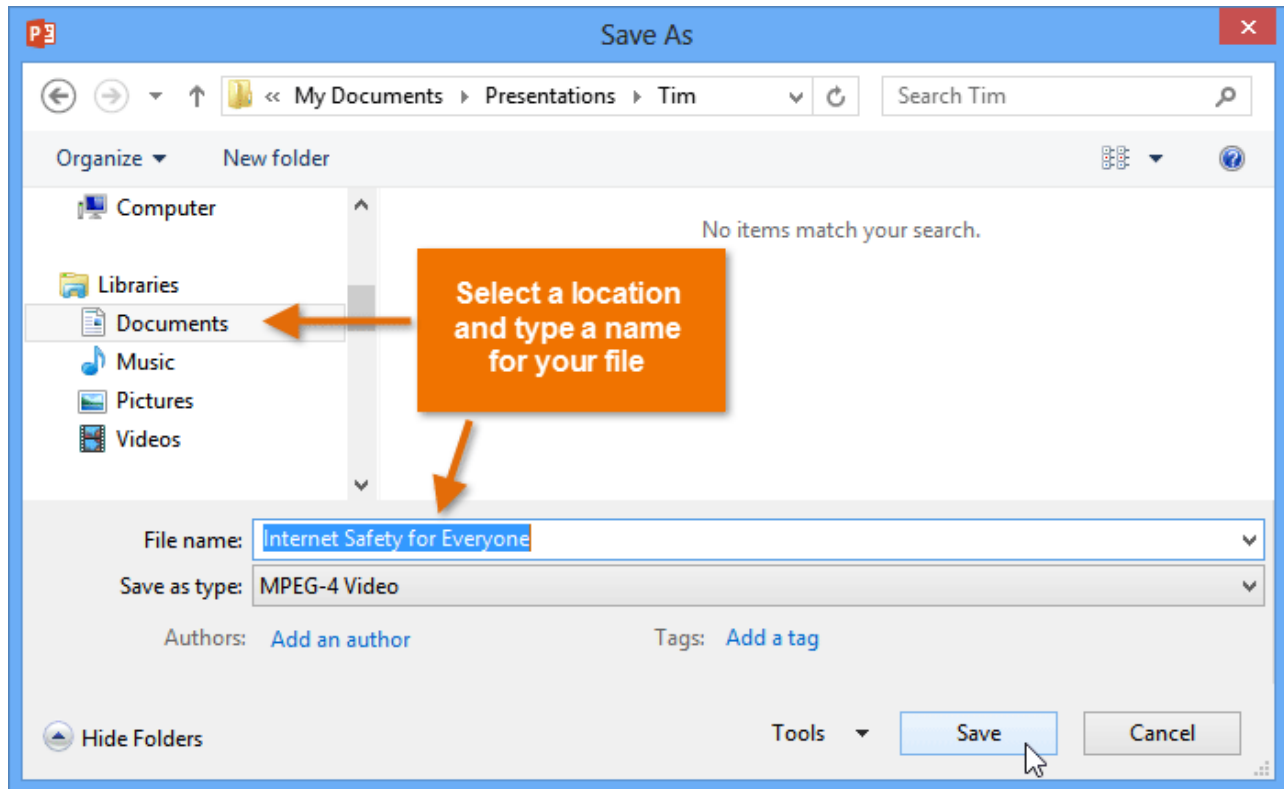


4. Select the drop-down arrow next to **Use Recorded Timings and Narrations**.
 - Choose **Don't Use Recorded Timings and Narrations** if you don't have or don't want to use recorded timings. You can adjust the default **Seconds to spend on each slide**: in the box below the drop-down menu.
 - **Choose Use Recorded Timings and Narrations if you have already recorded timings and narrations and want to use them in your video.**



5. Click **Create Video**. The **Save As** dialog box will appear.
6. Select the location where you want to save the presentation, then enter a name for the

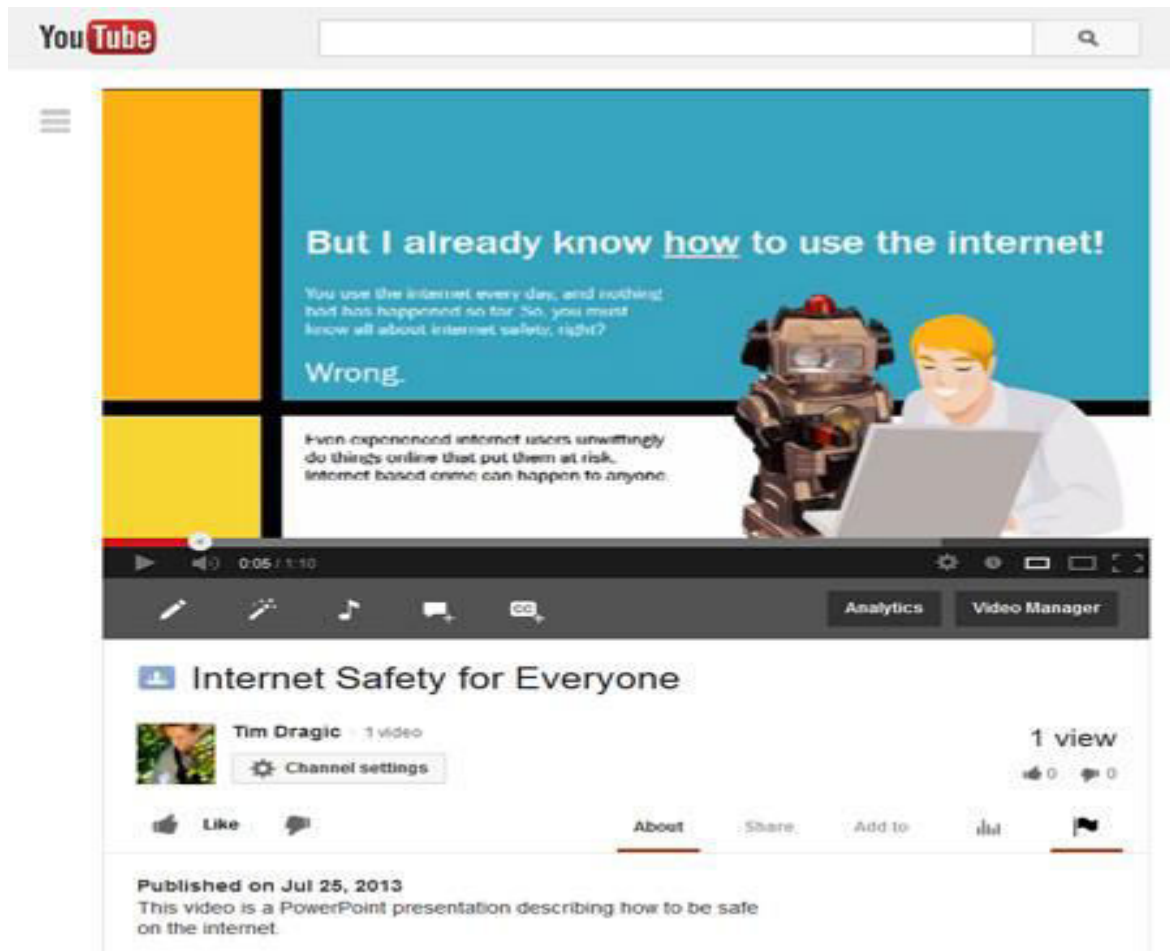
presentation.



7. Click **Save**. As Power Point creates your video, a status bar will appear at the bottom of the Power Point window. When the bar is complete, your video is ready to view, send, or upload.



In addition to emailing your video, you can **upload** your video to free **video-sharing sites** like YouTube. YouTube offers guides to creating an account and uploading videos. Once your video is uploaded, you can give the link to anyone you want to see it.



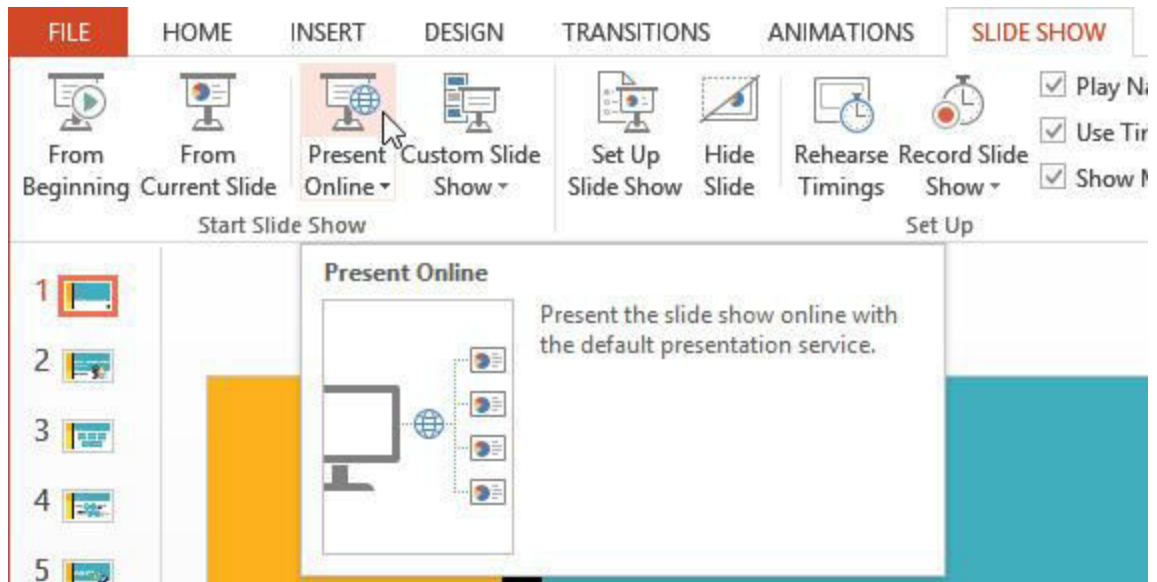
Presenting your slide show online

Presenting a slide show **online** is surprisingly easy. All you and your viewers need is an Internet connection—they don't even need Power Point. Once your viewers are connected, you can start the presentation as you normally would.

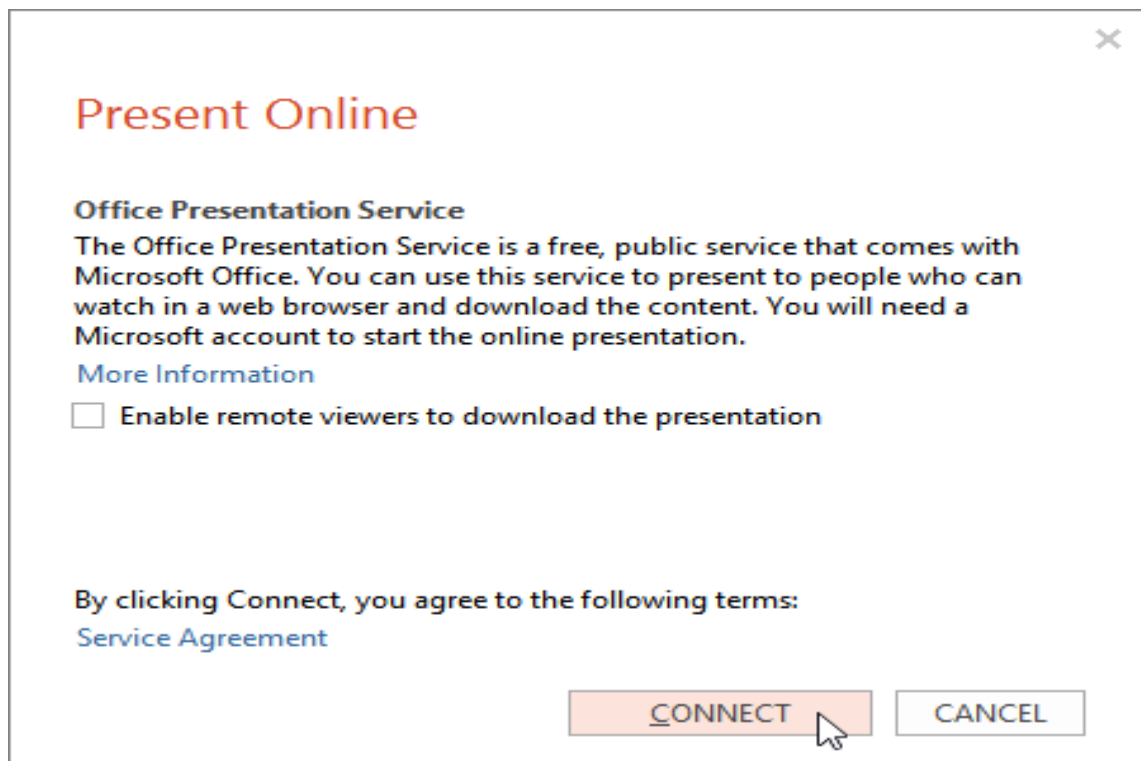
Please note that you cannot edit your presentation or mark it with a highlighter or pen while you are presenting a slide show online. You also cannot use Power Point to speak to your audience. Plan to communicate with your viewers through teleconferencing, or pre-record your narration.

To present online:

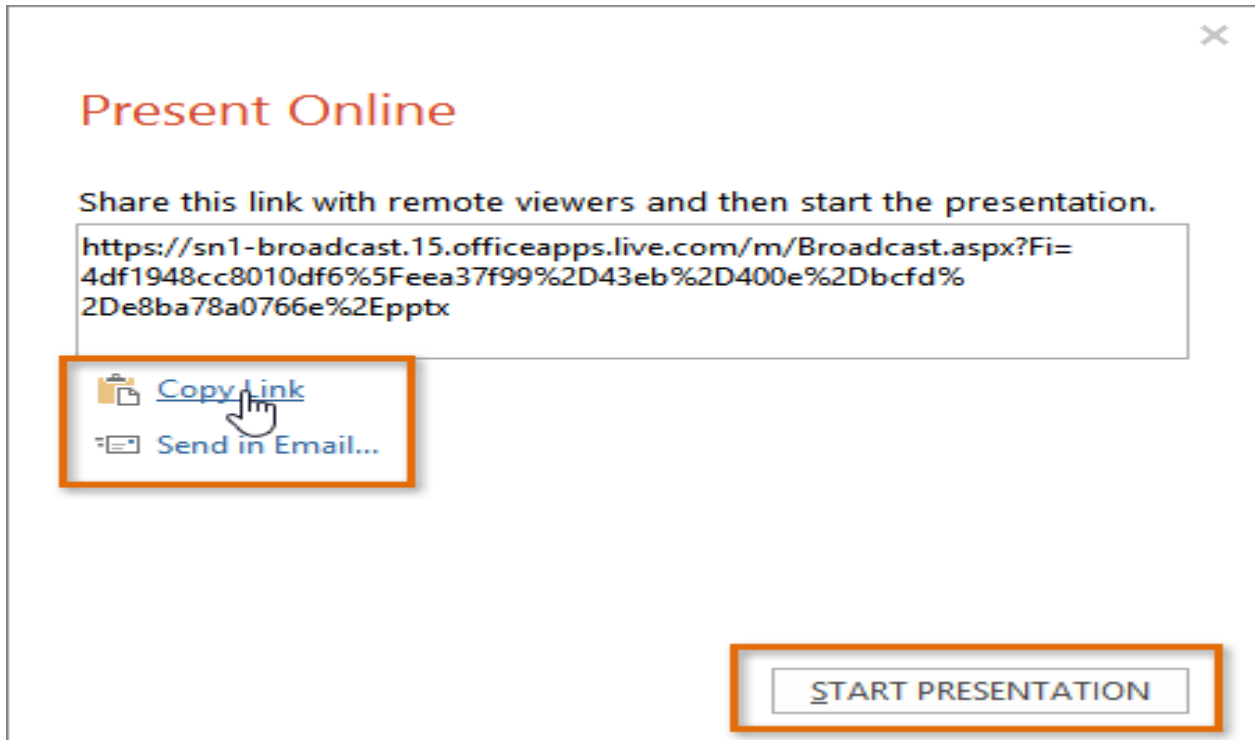
1. Select the **Slide Show** tab, then locate the **Start Slide Show** group.
2. Click the **Present Online** command.



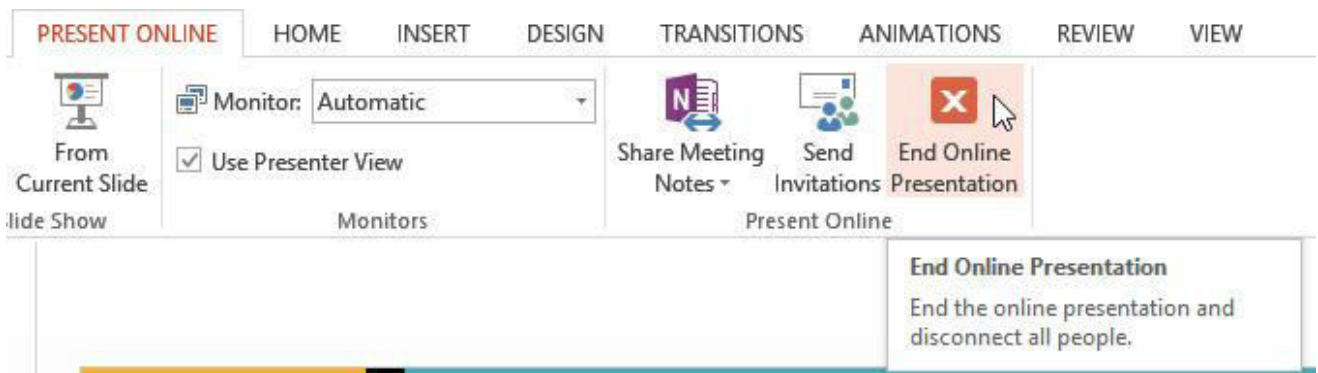
3. A dialog box will appear. Click **Connect**. A status message will appear as Power Point prepares your online presentation.



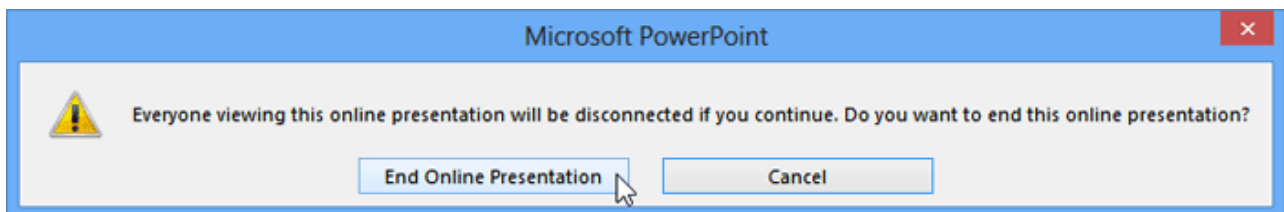
4. A link will appear. If it's not already selected, select the link.
5. Click **Copy Link** to make a copy of the link or **Send in Email** to send an email with the link to your viewers.
6. Click **Start Presentation**.



7. Present the slide show as you normally would, using the mouse or keyboard to advance the slides.
8. When you are finished, click **End Online Presentation**.



9. A dialog box will appear to confirm that you want to end the presentation. Click **End Online Presentation**.

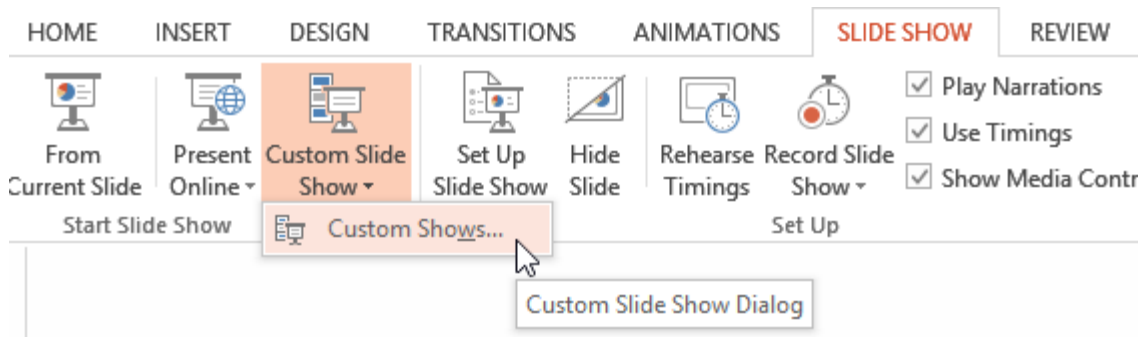


Customizing your slide show

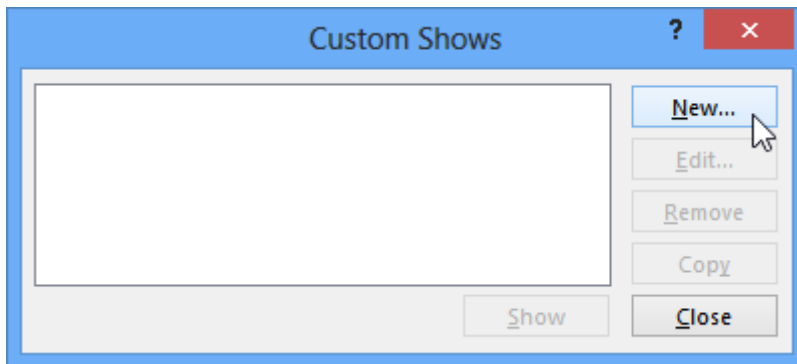
Sometimes you might want to **hide** a slide while still keeping it in your presentation. For instance, if you are presenting a slide show to more than one group of people, hiding or even rearranging certain slides could help you tailor your slide show to each group you present it to. You could also choose to create a shortened version of your slide show to present when you're short on time. The **Custom Slide Show** feature allows you to create and name different versions of your slide show with hidden or rearranged slides.

To create a custom show:

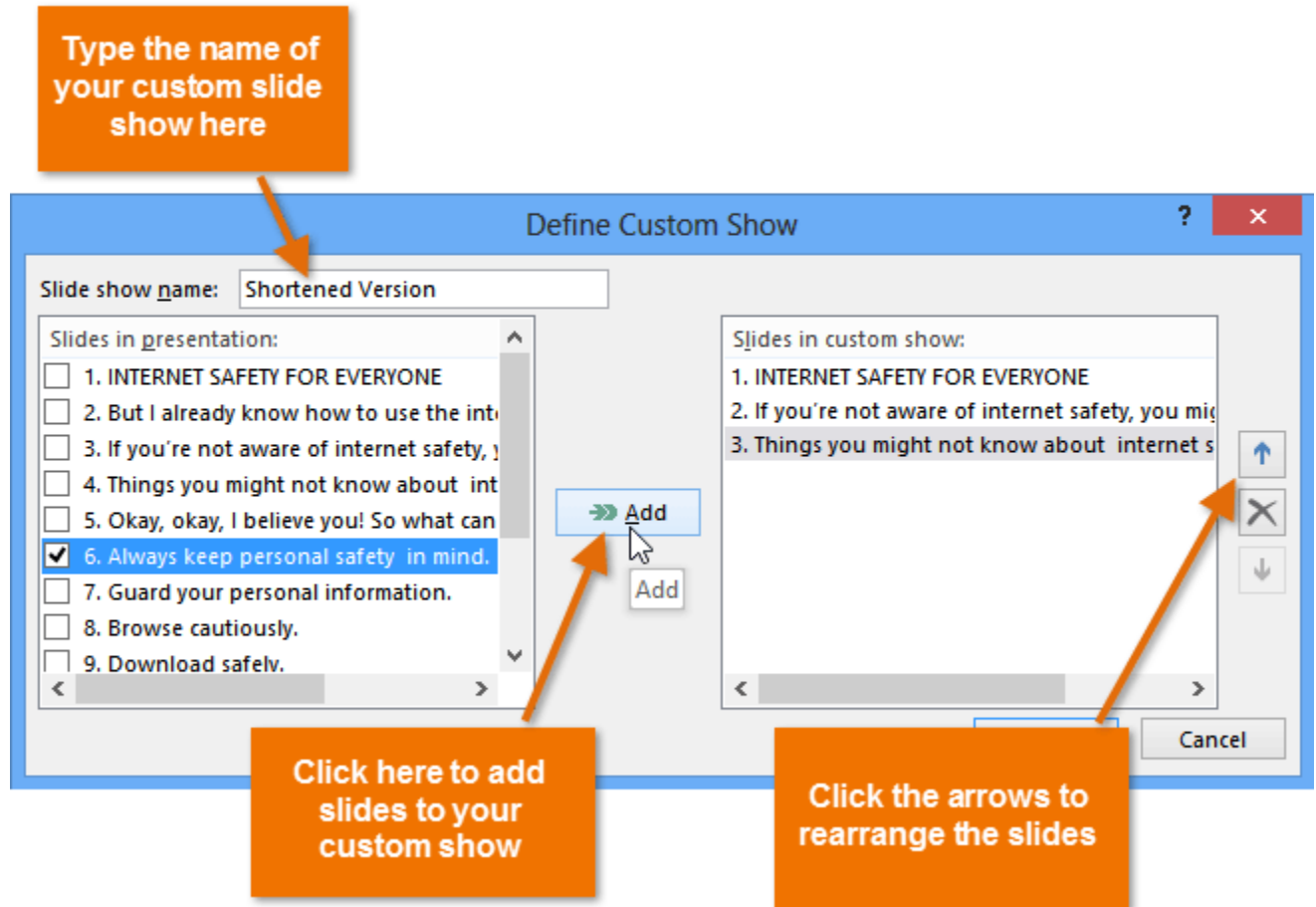
1. Select the **Slide Show** tab, then locate the **Start Slide Show** group.
2. Click the **Custom Slide Show** command, then select **Custom Shows**.



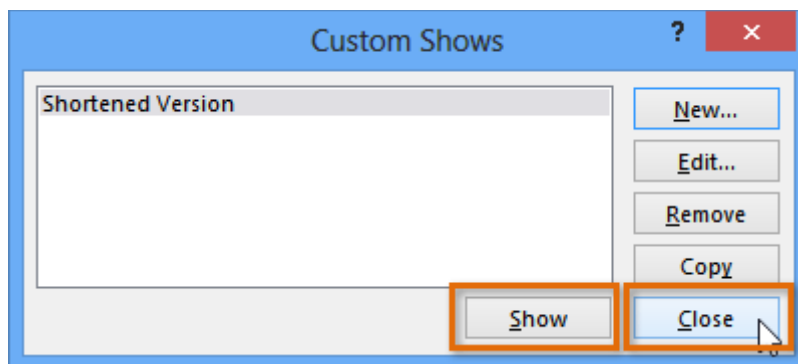
3. A dialog box will appear. Click **New**.



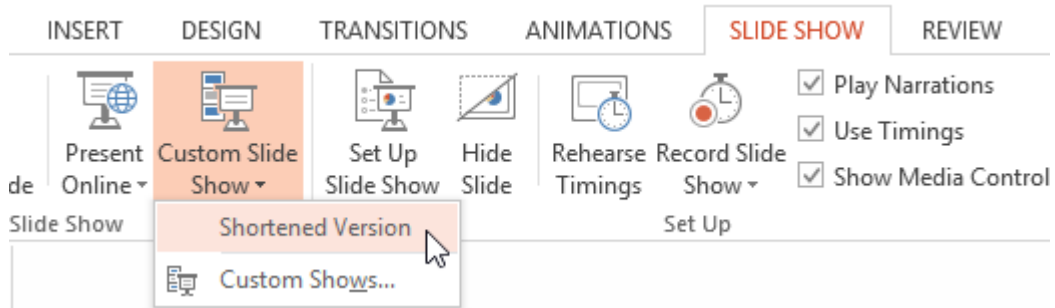
4. The **Define Custom Show** dialog box will appear. Locate the **Slide show name** field, and type in a name for your custom show.
5. In the **Slides in presentation:** area, check the box next to the slides you want to include in your custom show.
6. Click **Add** to add the slides to the **Slides in custom show:** box. If necessary, use the **up** and **down arrows** to reorder the added slides.



7. Click **OK**.
8. Select **Close** to exit or **Show** to view your custom show.

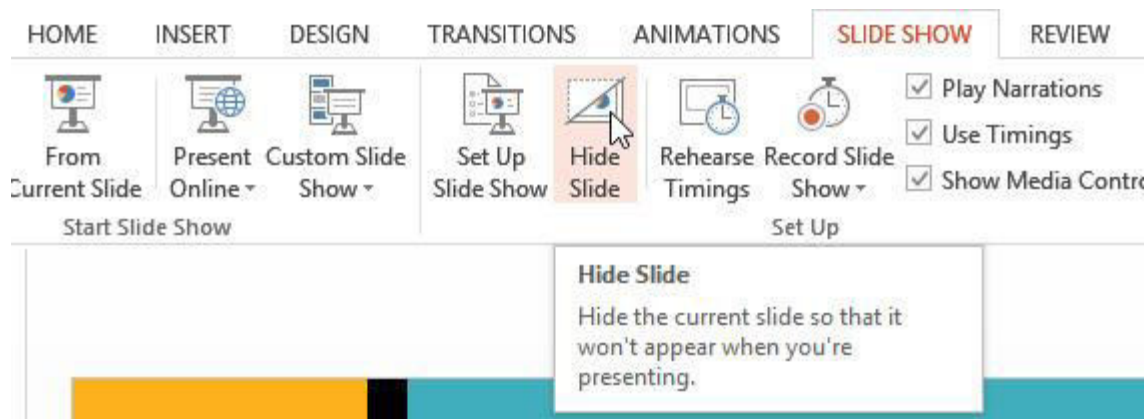


9. To play the custom show at any time, click the **Custom Slide Show** command, then select the desired presentation from the drop-down menu.



Hiding slides

Sometimes you may just need to **hide** a few slides without rearranging the entire slide show. You can hide slides by selecting the desired slide and then clicking the **Hide Slide** command on the **Slide Show** tab. To unhide a slide, simply click the **Hide Slide** command again.



Challenge!

1. Open an existing Power Point presentation. If you want, you can use our [practice presentation](#).
2. Use the **Rehearse Timings** feature to record slide and animation timings for the presentation. If you have a microphone attached to your computer, use the **Record Slide Show** feature.
3. Create a **custom slide show** from your presentation.
4. Create a **video** of your presentation.

Animation effects in Power Point

An animation effect is a special visual or sound effect added to a text or an object on a slide or chart. It is also possible to animate the text and the other objects using the buttons on the Animation Effects toolbar.

You can have organization charts appear or you can have the bullet points appearing one at a time.

Manual Animation

It is possible to automatically animate almost everything in your presentation. The problem with this is that you are tied to the timings of the slide show when you will probably need to have the flexibility to allow questions and comments. For this reason you should consider triggering some of the animation manually.

- Right click the object, select Custom Animation.
- Click on the Effects tab and choose the animation you want to use.
- Now click on Timings and select Animate and the On Mouse click radio button.

Action Button toolbar

Slide Transitions (Process of Change)

Slide Transitions are not technically animations because they do not involve movement of the individual items on a slide. A transition is how you get from one slide to another. The normal way to move from one slide to another is to jump to the next slide. There are over 50 types of slide transitions that can be used a transition happens whenever a presentation slide changes to the next slide. The first decision to make is whether the slide transitions will occur manually or automatically. With manual transitions the mouse button or a key on the keyboard must be pressed for the next slide to be displayed. Manual transitions are essential if there is a chance that someone might ask a question during the presentation.

Shortcut Keys

(Page Up) - Perform the previous animation or move to the previous slide.

(Page Down) - Perform the next animation or move to the next slide.

Options

(Edit, new animation effects) - Disables the new animation effects. Presentations with the new animation effects will not look the same when opened in earlier versions of Power Point.

Animating Text

The most common reason for animating text is draw attention to it typically one paragraph at a time. One way to do this is to create an entrance effect for the text placeholder; then adjust the effect settings so that the entrance is applied one paragraph at a time. When you do this the slide will initially appear empty except for the title.

Click the mouse once and the first line will appear.


Click the mouse again for the next paragraph to appear.

Another approach is to use an emphasis instead of an entrance effect. This allows all the paragraphs to be displayed on the slide initially. When you click the mouse the emphasis will be on the relevant paragraph. In either case you must first add the effect for the text placeholder and then display the Effect Settings dialog box by clicking the down arrow next to the effect and selecting Effect Options. The Group text drop-down box on the Text Animation tab of the Effect Options dialog box controls how the paragraphs appear.



Add Animations

You can animate the objects on your Power Point slides. Power Point provides four types of animations: Entrance, Emphasis, Exit, and Motion Paths. An Entrance animation determines the manner in which an object appears on a slide; for example, an object can move onto a slide. An Emphasis animation does something to draw attention to an object; for example, the object can become larger. An Exit animation determines the manner in which an object leaves a slide; for example, an object can move off a slide. A Motion Paths animation determines how an object moves around a slide; for example, an object can move from left to right.

After you add an animation, you can use the Custom Animation pane to modify it by choosing an effect. Choosing an effect enables you to define what starts the animation, its properties (such the direction from which an object moves onto the slide), and control the speed of the animation. In addition, you can have an animation start when you click the mouse, start along with the previous animation, or start at a specified time after the previous animation.


If the Auto Preview box is checked on the Custom Animation pane, Power Point provides you with preview of your animation after you create it and each time you modify it. You can also use the Play button  Play on the Custom Animation pane to preview an animation.

To choose an effect:

1. Select the object you want to animate.
2. Choose the Animations tab.
3. Click the Custom Animation button  Custom Animation . The Custom Animation pane appears.
4. Click the Add Effect button  Add Effect . A menu appears.
5. Choose the type of effect you want. A submenu appears.
6. Click the effect you want. Power Point applies the effect.

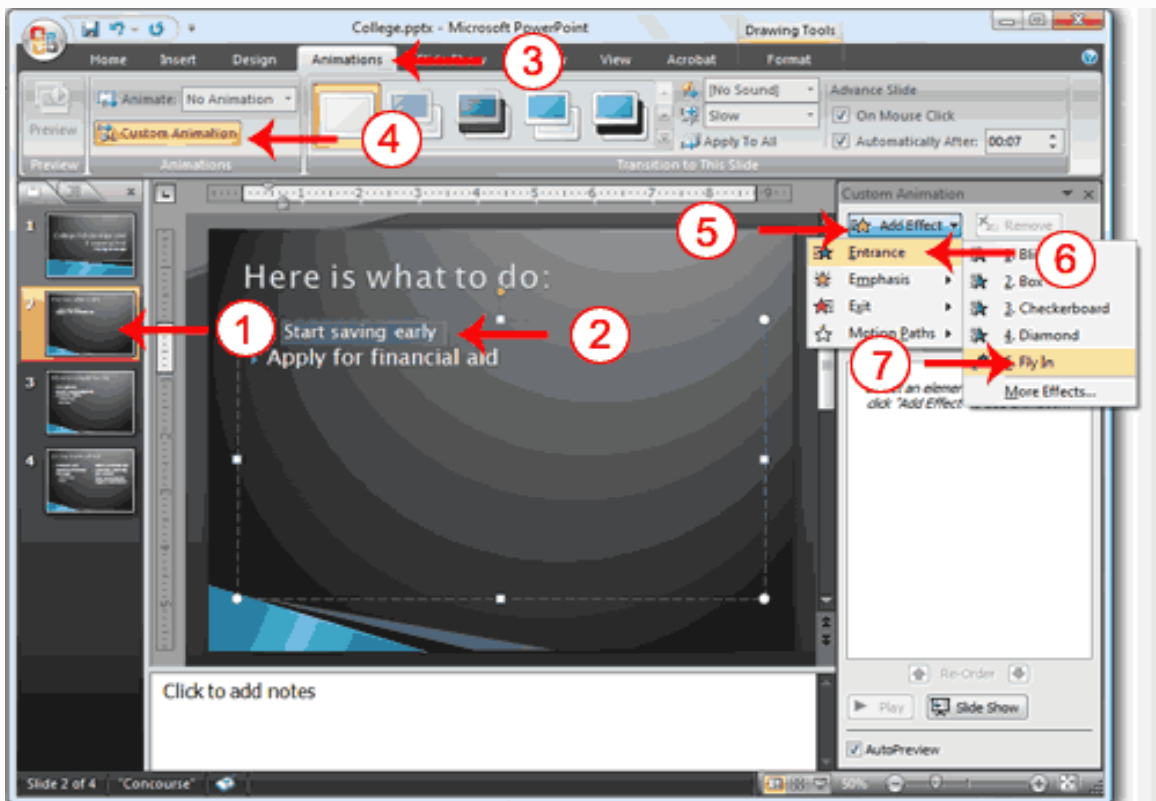
To modify an effect:



1. Click the down arrow next to the Start field on the Custom Animations pane and then select the start method you want.

2. Click the down arrow next to the Property field on the Custom Animations pane and then select the property you want. The Property field might be labeled Direction, Size, or some other property.
3. Click the down arrow next to the Speed field on the Custom Animations pane and then select the speed you want to apply to your animation.
4. To preview the animation, click the Play button  on the Custom Animations pane.

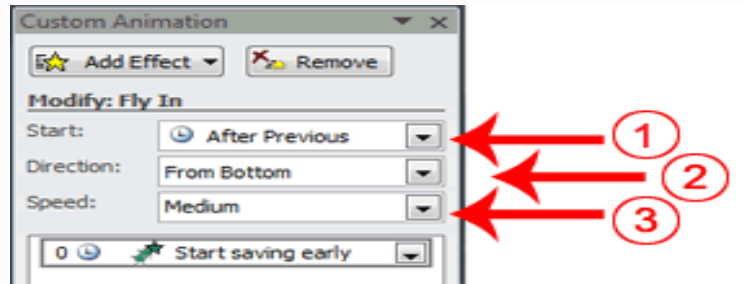
EXERCISE 1

Add an Animation to a Slide



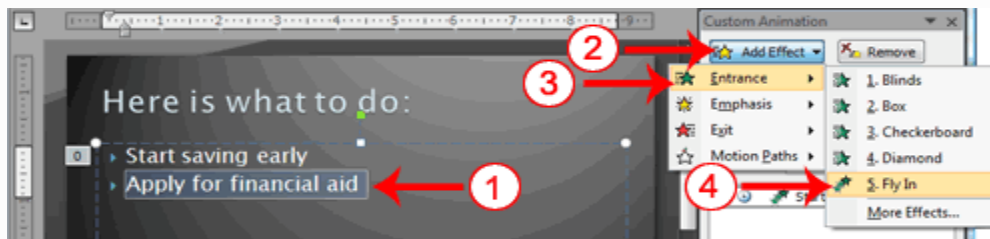
1. Click Slide 2 on the Slides tab.
2. Select "Start saving early."
3. Choose the Animations tab.
4. Click the Custom Animation button . The Custom Animation pane appears.
5. Click the Add Effect button . A menu appears.
6. Choose Entrance. A submenu appears.
7. Click Fly In. Power Point applies the effect. If the Auto preview box is checked, Power Point automatically provides you with a preview of the animation.


Modify the Effect



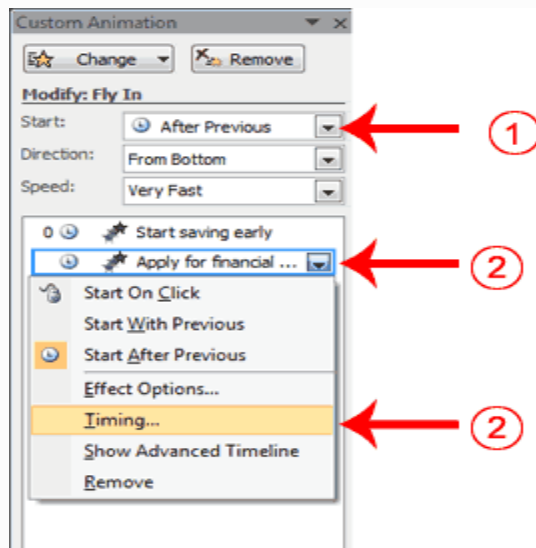
1. Click the down arrow next to the Start field and then select After Previous.
2. Click the down arrow next to the Direction field and then select From Bottom.
3. Click the down arrow next to the Speed field and then select Medium.

Add Another Animation

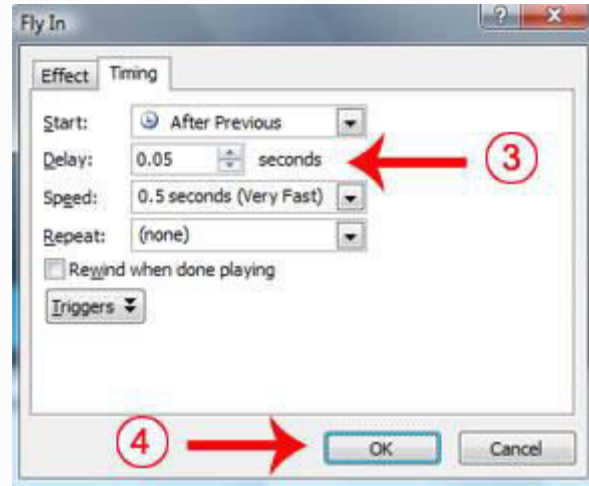


1. Select "Apply for financial aid."
2. Click the Add Effect button  Add Effect. A menu appears.
3. Choose Entrance. A submenu appears.
4. Click Fly In. Power Point applies the effect. If the Auto preview box is checked, Power Point automatically provides you with a preview of the animation.

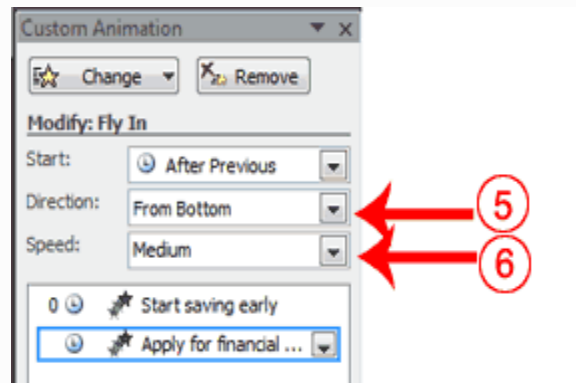
Modify the Animation




1. Click the down arrow next to the Start field and then select After Previous. The Apply for Financial Aid field appears in the center of the Custom Animation pane.
2. Click the down arrow next to the Apply for Financial Aid field and then click Timing. The Fly In dialog box appears.



3. Type **0.05** in the Delay text box.
4. Click OK.




5. Click the down arrow next to the Direction field and then select From Bottom.
6. Click the down arrow next to the Speed field and then select Medium. If the Auto preview box is checked, Power Point automatically provides you with a preview of the animation. You can click the Play button  Play on the Custom Animation pane at anytime to preview an animation.

Add Transitions



Transitions determine how your presentations move from one slide to the next. For example, a slide can move up onto the screen and replace the previous slide. Power Point provides several transition methods. You can add sound to a transition and you can control its speed. You can apply a transition to selected slides or to all of the slides in your presentation.

A transition can occur when the presenter clicks the mouse or after the amount of time you specify.

To apply a transition to selected slides:

1. On the Slides tab, hold down the Ctrl key and then click the slides to which you want to apply the transition.
2. Choose the Animations tab.
3. Click the More button  in the Transition to this Slide group. A menu of transitions appears.
4. Click the transition you want to apply. Power Point applies the transition. As you roll your pointer over each transition, Power Point provides you with a live preview of the transition.

To apply a transition to all slides:

1. Choose the Animations tab.
2. Click the More button  in the Transition to this Slide group. A menu of transitions appears.
3. Click the transition you want to apply. As you roll your pointer over each transition, Power Point provides you with a live preview of the transition.
4. Click the Apply to All button  in the Transition to This Slide group.

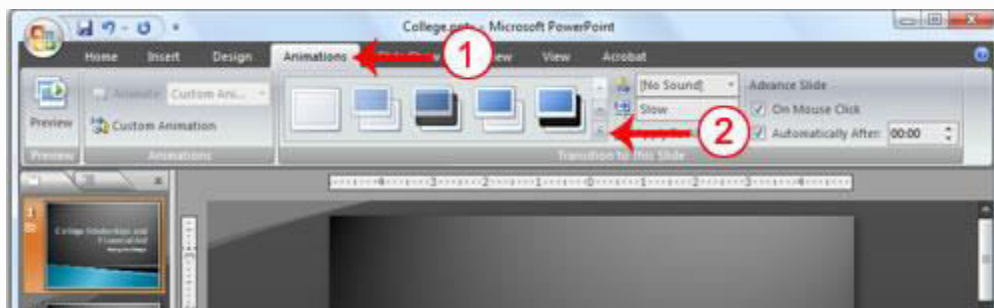
To add a sound to a transition:


1. Choose the Animations tab.
2. Click the down arrow next to the Transition Sound field and then click the sound you want. As you roll your pointer over each sound, Power Point plays the sound.

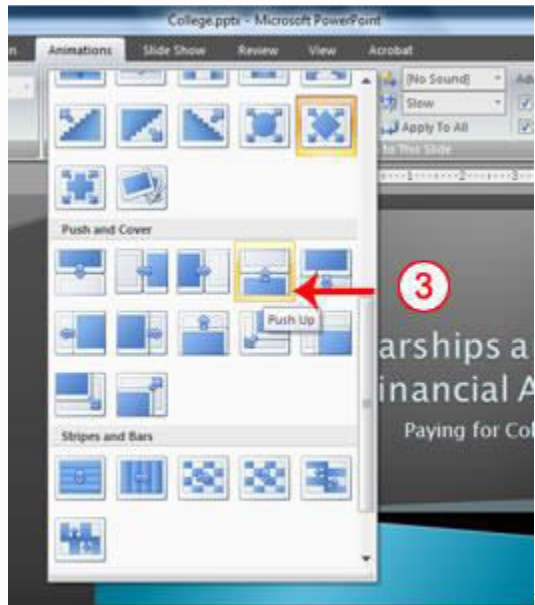
To set the speed of a transition:

1. Choose the Animations tab.
2. Click the down arrow next to the Transition Speed field and then click the speed you want.

If you want the transition to occur after the presenter clicks the mouse, check the On Mouse Click check box. If you want a transition to occur after a specified period of time, check the Automatically After check box and then specify the amount of time you want to elapse before the transition occurs. The On Mouse Click check box and the Automatically After check box are both located on the Animations tab in the Transition to This Slide group.

EXERCISE 2**Add Transitions**

1. Choose the Animations tab.
2. Click the More button  in the Transition to this Slide group. A menu of transitions appears.



3. Click the Push Up transition. As you roll your pointer over each transition, Power Point provides you with a live preview of the transition.

Use of sound and its effects in Power Point

The main purpose of [adding animation](#) to any slide object is to draw the attention of the audience to some movement. Once you set an [animation event](#), and set the [speed of the animation](#), you may also want some sound to play along with the animation -- one aspect that you should always remember is that although you can add sound to an animation, it is not always necessary to do so. We suggest you only add sound sparingly -- and even then, you must make sure that the sound adds some value to the animation.

Sound effects in PPT

Power Point comes with sound files you can insert into your Power Point slides. Windows comes with some useful sounds, too. If you have access to the Web, you have access to an unlimited supply of sounds. Sound files consume large amounts of hard drive space. Even a few seconds of sound can take 100K or more.

Power Point allows you to use two basic varieties of sound files:

- **Wave files:** Wave files contain digitized recordings of real sounds. These sounds can be sound effects, such as cars screeching, guns firing, or drums rolling; music; or even quotes from movies or your favorite TV shows.
- **Wave files come in several formats:**
 - **WAV:** Windows and Power Point come with a collection of WAV files that provide simple sound effects such as swooshes, blips, applause, and drum rolls.
 - **MP3 and WMA:** For longer sound clips, such as complete songs; the popular formats to use include MP3, a compressed format that is popular for sounds obtained from the Internet, and WMA, a newer audio format developed by Microsoft for newer versions of Windows. You can tell the format of a sound file by the filename's extension (.MP3 or .WMA).
- **MIDI files:** MIDI files contain music stored in a form that the sound card's synthesizer can play. Windows comes with several MIDI files, and you can download many more from the Internet. MIDI files have the file extension .mid.

You're more likely to use wave files than MIDI files in a Power Point presentation. MIDI files are great for playing music, but the wave files enable you to add a wider variety of sounds to a presentation.

How to add a sound to a presentation

Adding sounds to your Power Point presentation represents a great chance to further engage your audience. In Microsoft Office 2007, the process has been streamlined, and you will find many important options in a contextual tab follow these steps to create a truly "multimedia" presentation, tactfully.

Navigate to the "Insert" tab and select "Sound" from the "Media" grouping. To add a specific sound file, select to "Insert sound from file."

Browse to where your file is located, and click "Ok" to drop the sound file onto the slide, in which you are working, as an object.

Choose whether you would like the sound to play automatically, or to be started upon clicking on the sound object icon within your slide.

Assess your options. You have inserted a sound into your presentation, but the choices as to how you present it are open to your whimsy and creativity.

- If you would like to change your choice of whether to play the sound automatically, or after a mouse click, access the drop down menu in the "Sound Tools - Options" contextual tab in the Ribbon.

By default, Power Point 2007 leaves the sound object as a visible object in your slide. This can be toggled by clicking the box labeled "Hide during show," located in the "Sound Tools - Options" contextual tab. This will help you have a cleaner presentation when playing a sound automatically, but if you have selected to start the sound playback upon clicking the object (see Step 3), you will not be able to do so if you've elected to hide the object during the presentation.

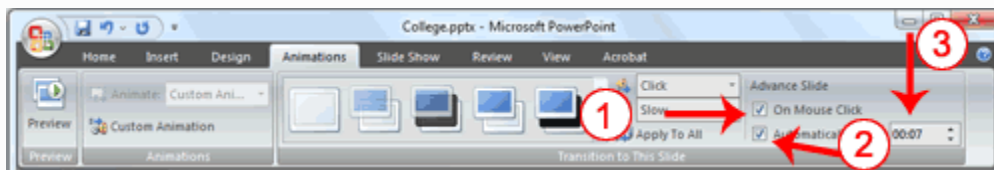
Increasing the file size limit for embedded sounds will absorb the sound file, if it is under the limit, into your presentation's single ".pptx" file. If the sound(s) you have included in your presentation are above the limit you specify, you will need to include the sound file(s) in the same location as your ".pptx" file, wherever you save it or share it (e.g., you would attach the ".pptx" file *and* the accompanying ".mp3" file in email if you were to be sending your deck to a colleague).

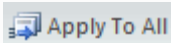
Add Sound and Set the Speed



1. Click the down arrow next to the Transition Sound field and then click Click.
2. Click the down arrow next to the Transition Speed field and then click Slow.

Advance Slide



1. Check the On Mouse Click check box.
2. Click the Automatically After check box.
3. Type **00:07** in the Automatically After text box.
4. Click the Apply to All button . Power Point applies all of your changes to all of the slides.
5. Click Slide 1 on the Slides tab.
6. Type **00:03** in the Automatically After text box. Power Point changes the timing for Slide 1.

Definition - What does Microsoft Excel mean?

Def: - Microsoft Excel is a software program produced by Microsoft that allows users to organize, format and calculate data with formulas using a spreadsheet system. This software is part of the Microsoft Office suite and is compatible with other applications in the Office suite.

Def: - Excel is handy software that can be used to store and organize many data sets. Using its features and formulas, you can also use the tool to make sense of your data. For example, you could use a spreadsheet to track data and automatically see sums averages and totals.

Excel Basics

If you're just starting out with Excel, there are a few basic commands that we suggest you become familiar with. These are things like:

- Creating a new spreadsheet.
- Executing basic computations in a spreadsheet, like adding, subtracting, multiplying, and dividing in a spreadsheet.
- Writing and formatting column text and titles.
- Excel's auto-fill features.
- Adding or deleting single columns, rows, and spreadsheets. Below, we'll get into how to add things like multiple columns and rows.
- Keeping column and row titles visible as you scroll past them in a spreadsheet, so that you know what data you're filling as you move further down the document.

Feature's of MS Excel 2007

In Excel 2007, more is truly better. Microsoft has increased the number of columns per spreadsheet (and per PivotTable) to 16,384 (up from 256) and the number of rows to 1,048,576 (up from 65,536). Other limits have been also expanded: Text cells can now contain more than 32,000 characters (up from 255).

Chances are you'll never reach other new limits: PivotTables can manipulate more than 16,000 fields (up from an already generous 255), and formulas can now refer to up to 8,000 cells (memory permitting), so it's fortunate that Excel 2007 lets you drag the corner of the formula bar to expand it. Excel 2007's memory manager can handle 2GB (double the amount in Excel 2003), so calculations execute faster. The new version also takes advantage of dual-core processors and multithreaded chip sets, so if you're lucky enough to be running it on a machine with either feature, expect a noticeable speed boost.

New Visualization Tools

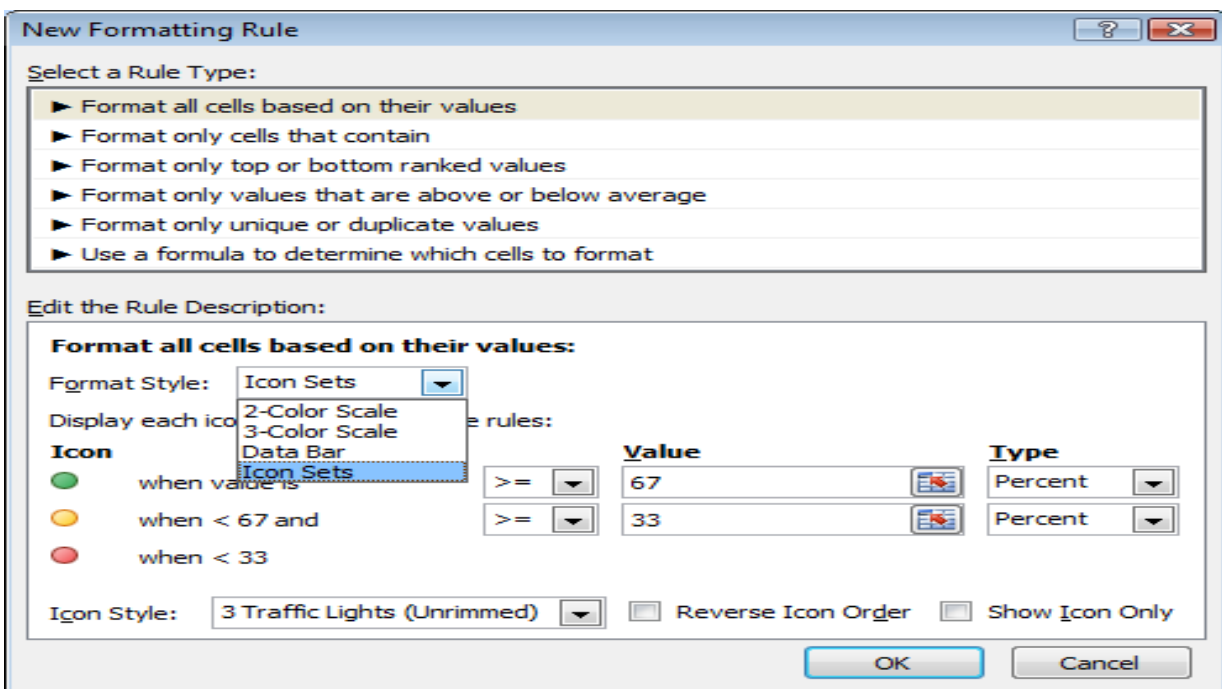
Charts and graphs now support 16 million colors, and improved color support is evident throughout this version, especially in several new visual tools for highlighting data. For example, in Excel 2007 you can use conditional formatting to set the background color of a cell or use a colored bar (called a data bar) -- the length corresponds to the cell's value.

You can also add icons to cells based on their value, giving your worksheet a dashboard-like quality. For example, assigning traffic-light icons to a range of cells is a snap, and Excel's built-in logic assigns colored circles based on the value of the cell: green for the highest third, yellow for the middle third and red for the bottom third.

1	Student	Exam Mark	1	Student	Exam Mark	1	Student	Exam Mark
2	Mary Parker	56	2	Mary Parker	56	2	Mary Parker	56
3	Peter Hollis	44	3	Peter Hollis	44	3	Peter Hollis	44
4	Alex Grant	92	4	Alex Grant	92	4	Alex Grant	92
5	Matthew Johnston	76	5	Matthew Johnston	76	5	Matthew Johnston	76
6	Freya Jones	83	6	Freya Jones	83	6	Freya Jones	83
7	John Wilson	76	7	John Wilson	76	7	John Wilson	76
8	Zoe Townsend	88	8	Zoe Townsend	88	8	Zoe Townsend	88
9	John Baker	72	9	John Baker	72	9	John Baker	72
10	Karen Ellis	89	10	Karen Ellis	89	10	Karen Ellis	89
11	Jane Carson	79	11	Jane Carson	79	11	Jane Carson	79
12	Hamish Taylor	92	12	Hamish Taylor	92	12	Hamish Taylor	92
13	Thomas Cotton	82	13	Thomas Cotton	82	13	Thomas Cotton	82
14	Nicholas Mercer	66	14	Nicholas Mercer	66	14	Nicholas Mercer	66
15	Robert Potter	75	15	Robert Potter	75	15	Robert Potter	75
16	Laura Renson	81	16	Laura Renson	81	16	Laura Renson	81
17	Mark Sherwood	69	17	Mark Sherwood	69	17	Mark Sherwood	69
18	Colin Farrer	57	18	Colin Farrer	57	18	Colin Farrer	57
19	Christopher Denton	49	19	Christopher Denton	49	19	Christopher Denton	49
20	Graham Aberforce	84	20	Graham Aberforce	84	20	Graham Aberforce	84

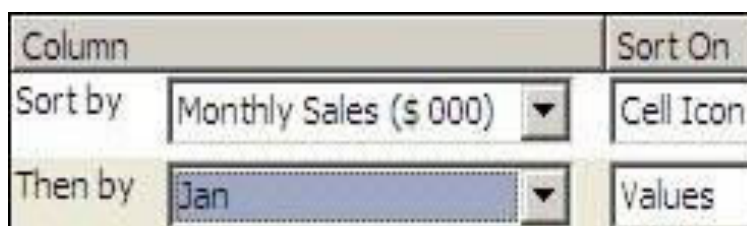
You can add colored bars to indicate the value in a cell (left) or apply a three-icon set to indicate which third data falls into (right); a red traffic light indicates the cell is in the lowest third of all values in the range.

There are in fact many variations on these built in to Excel 2007, but if you can't find one that suits you can always set up your own rules. **(Conditional Formatting > New Rule)**



Better Sorting and Filtering

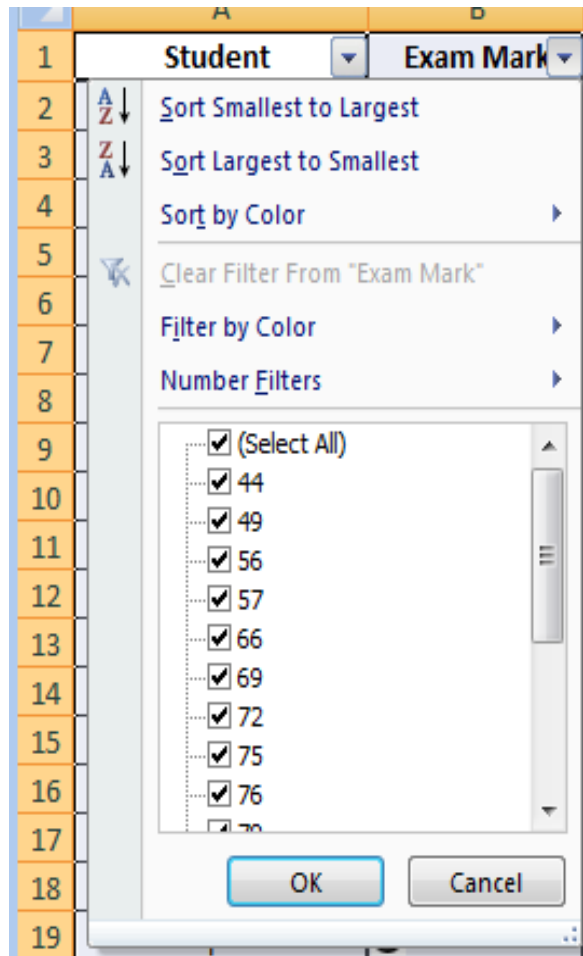
Sorting data -- previously limited to three levels -- has been expanded to 64 levels, while you can still sort data based on values (Top to Bottom or A - Z, etc) you can also sort by font, colour or icon



used with conditional formatting. Thus, you can display all your green traffic lit cells together, followed by the yellow lights, then the red.

Other visualization tools remove the need for complicated macros or formulas. New conditional formatting options let you highlight duplicates, unique values, the top/bottom 10%, values above or below the average, cells less than or greater than a specified value, or cells within a range, highlighting cells containing values between 1 and 10, for instance).

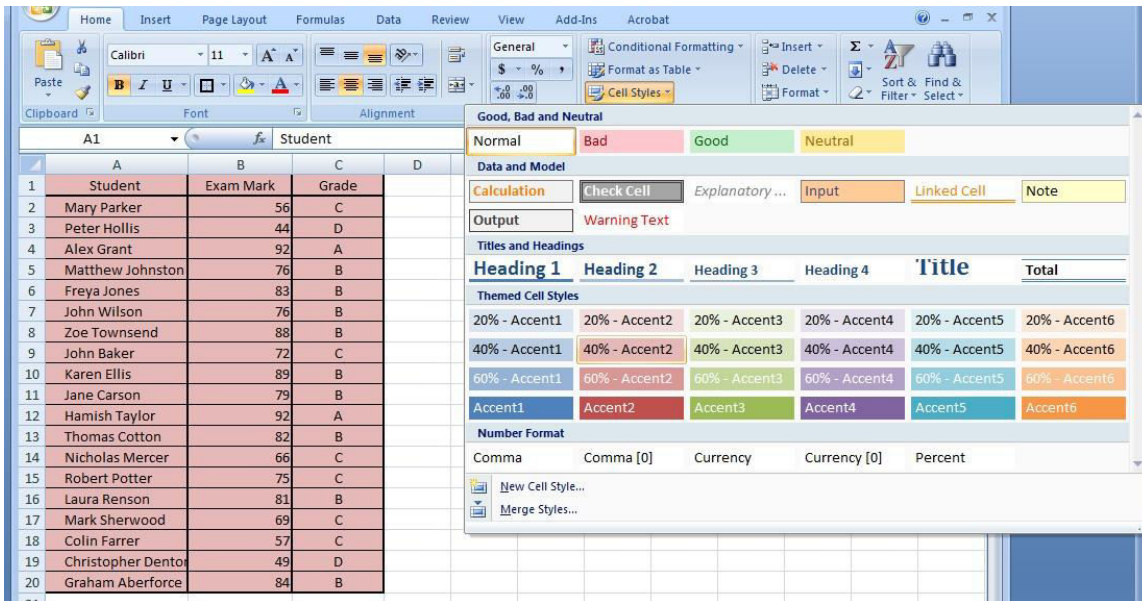
If you don't need to see all values, the vastly improved Filter feature puts check boxes (for up to 1,000 values) in a pull-down list, allowing you to easily pick multiple values to display. Likewise, the new Remove Duplicates feature hides rows based on the duplicate values in columns you specify.



Styles and Themes

Styles

Styles is a formatting tool from previous versions of Excel that is now available using a "gallery" interface introduced in Office 2007. You can quickly apply a collection of settings, from the font used to the background color and border style to cells, tables and PivotTables. As you mouse over the choices, Excel 2007 applies each style to your selection so you can preview the effect without making the change permanent.



One particularly noteworthy improvement to formatting is how Styles now respond to changes within your worksheet. In Excel 2003, you could apply a "green bar" effect so that the background color in rows alternated between green and white. However, once you added a row, the pattern was interrupted, and you needed to reapply the AutoFormat. In Excel 2007, that same pattern is adjusted whenever you add one or more rows. (Styles are equally smart when you add columns, for patterns that alternate between columns.) Styles will even adjust when you filter or hide rows or columns.

Themes

Themes are a feature new to Office 2007, they are style collections that include a color scheme, font, fill effects and more. Shared by several Office 2007 applications, themes can be applied to charts, tables and PivotTables in Excel, giving your work a consistent look and feel. That's especially useful when you're creating a chart that you want to copy to PowerPoint or Word.

	A	B
1	Monthly Sales (\$ 000)	Month
2	Cabbage, Green	Jan
3	Cabbage, Green	Feb
4	Cabbage, Green	Mar
5	Cabbage, Red	Jan
6	Cabbage, Red	Feb
7	Cabbage, Red	Mar
8	Lettuce, Green	Jan

To use Themes, select the Page Layout tab and click the Themes button to choose a new theme. You can also customize any theme or create new ones. One important limitation is that Themes only work if you're using the new Office XML format; they won't work on old-style .xls files.

Charts

The Ribbon interface also makes it more enjoyable to work with charts. Excel's charts have a whole new look, thanks to the new graphics engine in all Office 2007 applications. The layouts use different color palettes and fonts, but the important difference is the ability to more easily apply graphical effects, such as bevels and shadows, to individual elements (such as columns or pie slices).

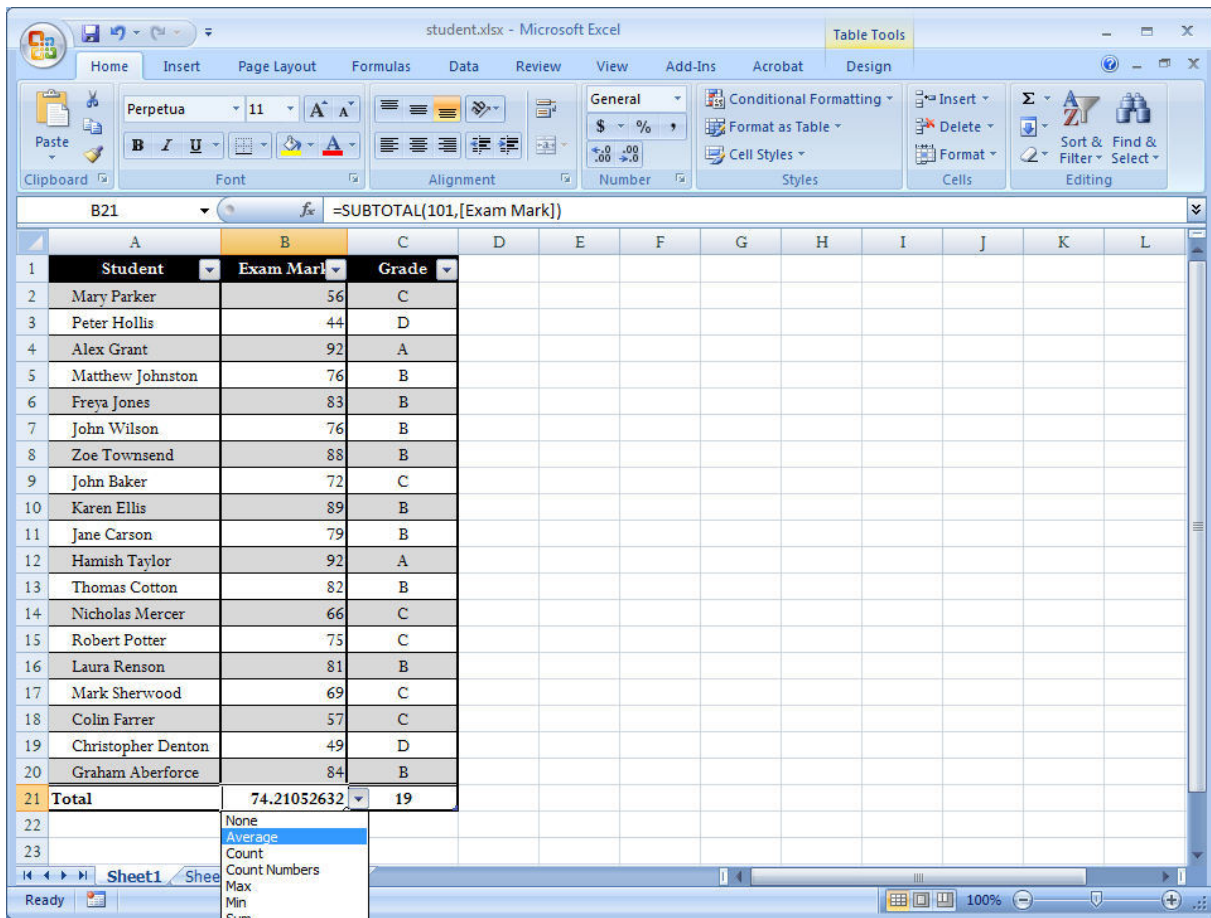


The Ribbon interface has a Chart Tools group (with tabs for Design, Layout and Format) to put more charting options at your fingertips and eliminate most of the right-clicking you had to do to adjust charts in previous versions: switching rows and columns, controlling gridlines and axes, and adding trend lines.

Table Tools

Excel's new table features make it less likely you'll have inconsistent formulas. Once you identify a contiguous range of cells as a table, Excel provides calculated columns. For example, if you add a column to the right of your table and enter a formula in any row, the formula will be copied to all cells in that new column, saving the time of executing a copy/paste command.

Even smarter, add a row and Excel is sure to include it in a total on the bottom row. (In previous versions of Excel, adding a row at the top or bottom of a range meant you risked omitting cells in that row from the sum formula.)



Furthermore, options on the Table Tools Design context-sensitive tab let you toggle the formatting of the first column or the first row. One click and you can add a Total row (though Excel lacks a similar command to add a Total column), then change what each column in that row computes (total, average, minimum and so on).

In addition, as you scroll down through a lengthy table, Excel replaces the column headings (the gray boxes with A/B/C above the columns) with values from the table's header row -- a subtle improvement, to be sure, but it's a more efficient technique than having to freeze rows to see column headings.

Finally, the new Table Gallery makes it easy to select and apply a sophisticated look.

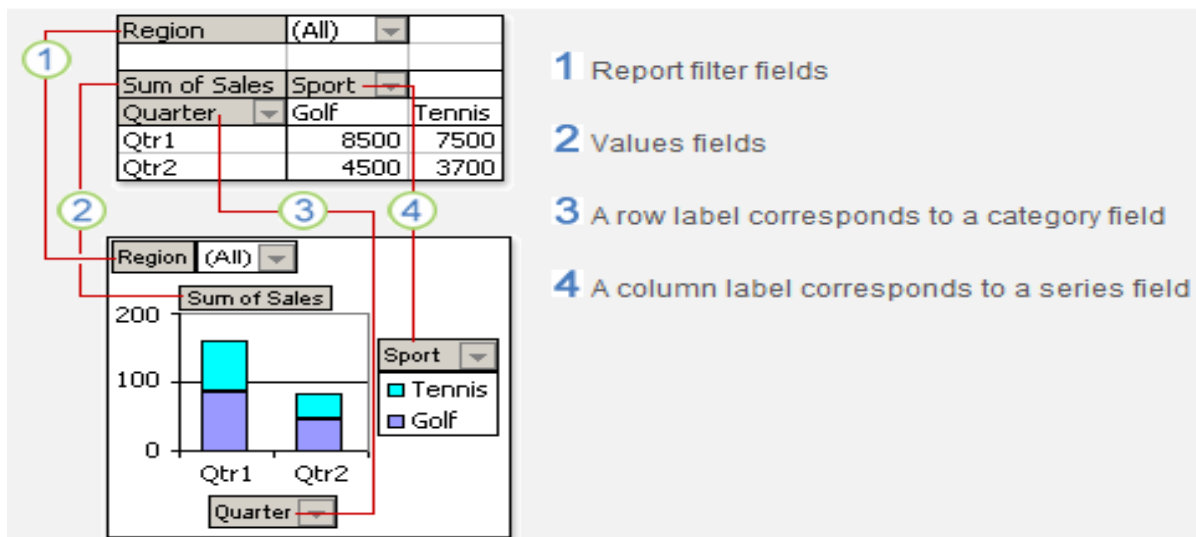
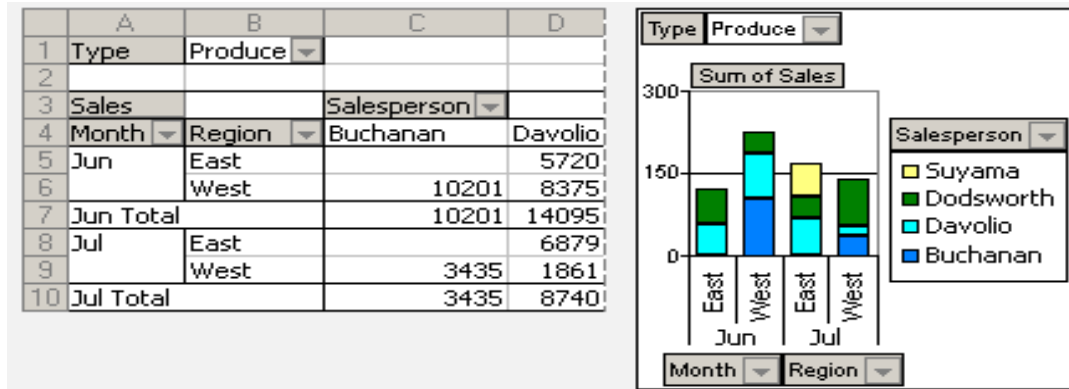
Pivot Tables

Among the notable improvements in Excel are tools to make existing features easier to use. Take PivotTables, for example. (For the uninitiated, PivotTables allow you to view your data

differently -- think "slice and dice." For example, you can summarize sales by agent by month or, with a simple drag-and-drop motion, summarize sales by month and within month by agent.)

In Excel 2007, you still set up PivotTables using a wizard, which is slightly changed from Excel 2003. However, once you have a PivotTable defined, manipulating it is considerably easier.

Instead of dragging and dropping elements within the table itself, you can use the wizard to make choices -- checking boxes to select which fields to display or choose sorting options, for example. Excel 2007 makes it easier to switch columns and rows, filter values, and use or hide



field names. In addition, conditional formatting (those data bars or traffic lights we mentioned) can be applied to cells displayed in PivotTables.

The Most Common Formatting Commands

There are many types of formatting that can be applied to Microsoft Excel worksheets. The most commonly used formatting commands show up on the Home tab in three groups:

1. **The Font Group.** The font group commands change the appearance of text within a cell or of the cell itself.
2. **The Alignment Group.** The alignment group commands change the position of text within a cell or cells.

3. **The Number Group.** The number group commands change the format of numbers and dates within a cell.

Formatting changes can be applied to a whole worksheet, a range of cells within a worksheet, individual cells, and sometimes even text within a cell.

Excel Formatting

As well as giving your spreadsheet a professional look, the use of Excel formatting can provide essential information that influences the way a user interprets the data in the spreadsheet.

This is particularly the case with Excel numbers.

As dates, time, percentages and currency values are all stored as numeric values in Excel, the user often has to rely of the formatting of these numbers to understand what they represent.

For example, a cell containing the data value 0.5 could represent any of the following:

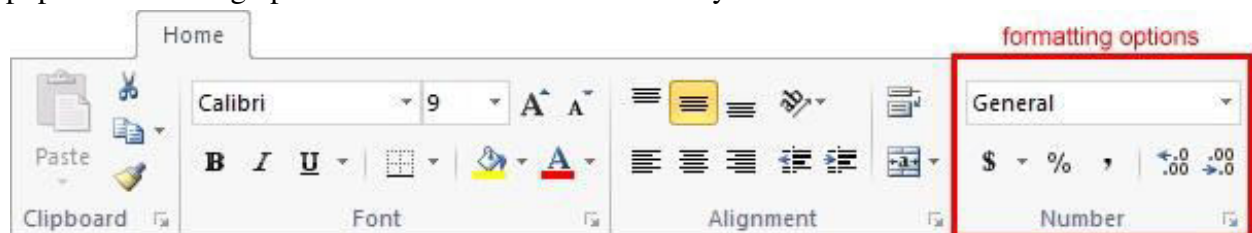
- The time 12:00 hrs
- The date and time 01-Jan-2020 12:00
- The currency value \$0.50
- The percentage value 50%
- The simple number 0.5

The spreadsheet below shows different ways of formatting Excel cells containing numeric values.

	A	B	C
1		Formatted Value	Underlying Value
2	A number formatted as a percentage:	59%	0.59
3	A number formatted as a currency:	\$54.27	54.27
4	A number formatted as a date:	01 August 2019	39661
5	A number formatted as a time:	12:27 PM	0.51875
6	A number formatted as a fraction:	1 3/4	1.75

How to Change Formatting in Excel

To format data in Excel cells, you first need to select the cells to be formatted. Some of the more popular formatting options can then be accessed directly from the Home tab of the ribbon

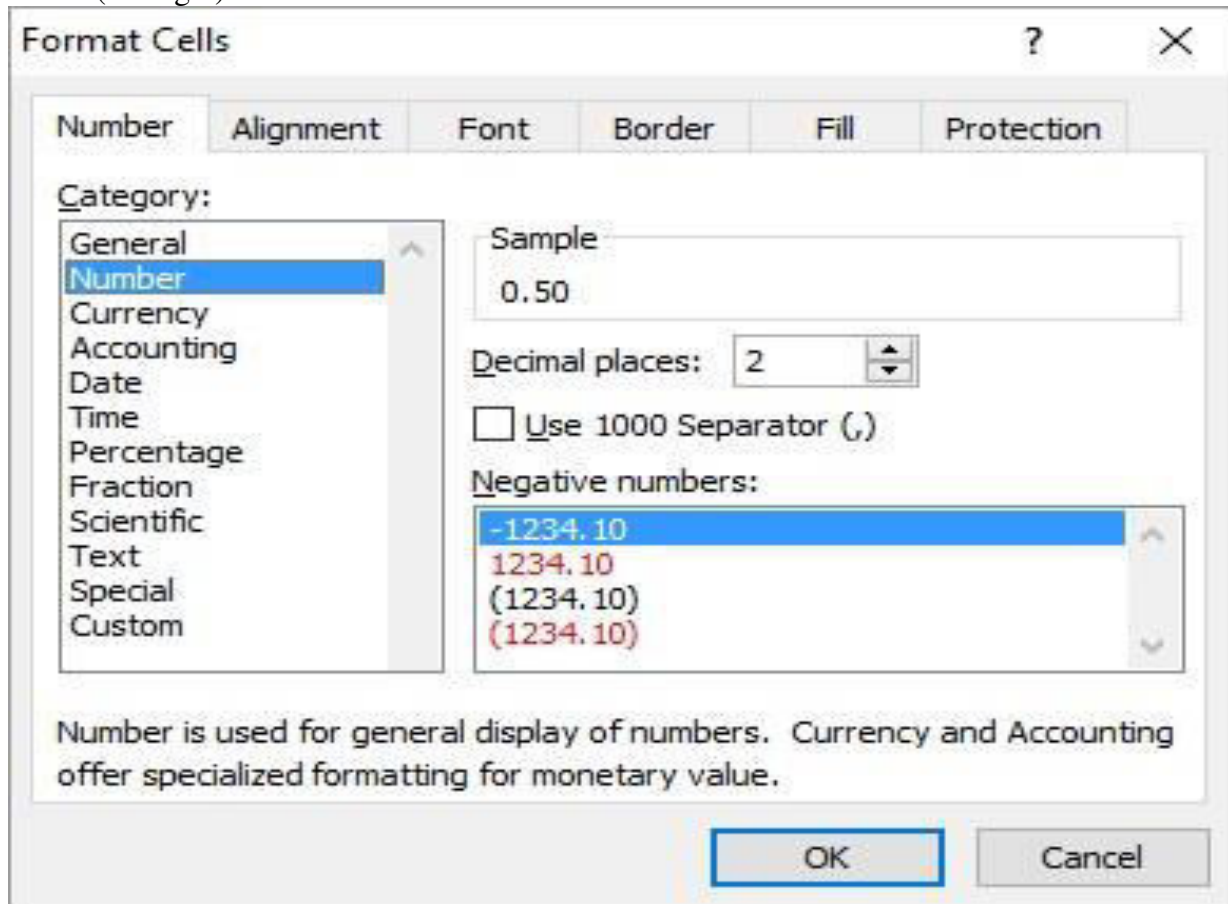


However, for the full menu of formatting options, you will need to use the 'Format Cells' dialog box, which can be accessed by any of the following methods:

- Right-click on the selected cell or range and select the **Format Cells ...** option from the right-click menu
- or
- Press Ctrl + 1 (i.e. Select the Ctrl key and while holding this down, select the "1" (one) key)



- Click on the dialog box launcher in the Number grouping within the **Home** tab of the Excel ribbon (see right).



The 'Format Cells' dialog box is shown on the left.

Make sure that the Number tab at the top of the dialog box is selected then, from the Category options (on the left of the dialog box), select the Excel formatting type that you

want to apply to your cell (e.g. Number, Percentage, Date, etc).

When you select a category, you may be presented with further options on the right hand side of the dialog box, which you can use to specify your formatting style.

Once you have specified the formatting details that you require, click OK to close the dialog box.

Note that the listed formats will only work with numbers, so if you apply a date format to a cell containing text, the appearance of the cell will remain unchanged.

Excel Date Format

Excel provides several standard date formats that you can apply to cells containing date values. These can be accessed in different ways (see below).

However, if none of the standard date formats are satisfactory, Excel also allows you to define your own custom date format.

In recent versions of Excel (Excel 2007 and later) a selection of standard formats, including some date formats, are available via the drop-down box in the 'Number' grouping on the **Home** tab of the ribbon.

To apply a standard date format to the cell(s) of your spreadsheet, simply select the cell(s) to be formatted and click on the format cells drop-down box. You will then be presented with several formats which, if selected, will be applied to the current selected cell(s).

The image on the right shows two Excel date format options that are available in the drop-down menu when a PC is set up for US times.

The standard date formats in your version of Excel may be different, depending on the settings on your computer.

- **Excel Custom Number Formatting**
- **Merge Cells**
- **Wrap Text In Excel**
- **Conditional Formatting**

Excel Functions

Excel Functions written in layman's terms and in a language we all understand. No computer speaks just plain English and practical examples of Excel functions used in formulas. Before we dive into the examples I want to cover some terminology that I'll be using.

Anatomy of an Excel Function

Excel functions comprise of the function name and arguments, as you can see in the example below:

SORTBY Sort a range or arrays based on criteria.

UNIQUE Extract a unique or distinct list from a range or array.

Database Functions

DSUM Sum a range based on criteria. Alternative to SUMIFS.

Date and Time Functions

DATE Build a date with separate year, month and day values.

DATEDIF Find the difference between two dates in days, months or years.

DATEVALUE Convert dates stored as text to a date Excel can use in formulas etc.

DAY Returns the day of the month between 1 and 31 from a date or text.

DAYS Returns the number of whole days between two date serial numbers. It ignores time portions where included.

DAYS360 Returns the number of days between two date serial numbers based on a 360-day year, or twelve 30-day months.

EDATE Returns the same date of the month, n months before or after the specified date.

EOMONTH	Returns the last day of the month, n months before or after the specified date.
HOUR	Returns the hour integer ranging from 0 to 24, from a time.
ISOWEEKNUM	Returns the week number of a date serial number. The first Monday of the year marks the start of week 1.
MINUTE	Returns the minute, ranging from 0 to 59, from a time.
MONTH	Returns the month number, between 1 and 12, from a date.
NETWORKDAYS	Returns the number of whole working days between two date serial numbers, excluding weekends.
NETWORKDAYS.INTL	Returns the number of whole working days between two date serial numbers, excluding weekend days of your choice.
NOW	Returns the current date and time from your computer clock.
SECOND	Returns the seconds, ranging from 0 to 59, from a time.
TIME	Build a time value by entering separate hour, minute and second values.
TIMEVALUE	Convert times stored as text to a time Excel can

use in formulas etc.

TODAY	Returns the current date serial number from your computer clock.
WEEKDAY	Returns the day number of the week from a date serial number. The default ranges from 1 for Sunday, through to 7 for Saturday.
WEEKNUM	Returns an integer representing the week number (from 1 to 54) of the year from a date serial number.
WORKDAY	Returns the day number of the week from a date serial number. The default return type ranges from 1 for Sunday to 7 for Saturday.
WORKDAY.INTL	Returns a date serial number that is the specified number of working days before or after the start date.
YEAR	Returns the year as an integer between 1900 and 9999, from a date.
YEARFRAC	Calculates the fraction of the year represented by the number of whole days between two dates.

Engineering Functions

CONVERT	Converts inches to centimetres, pounds to grams etc.
---------	--

Financial Functions

ACCRINT	Returns the accrued interest that gets paid periodically.
EFFECT	Returns the effective annual interest rate, from the nominal annual interest rate (APR).
FV	Calculate the future value of an investment.
NPER	Calculates the number of periods to reach an investment goal.
PMT	Returns the regular and constant repayments for a loan or mortgage required to reduce the balance to zero, or an amount you specify.

Information Functions

CELL	Returns information about the formatting, contents or location of a cell.
N	Converts a value to a number, but it's clever uses is where it shines.
SHEET	Returns the sheet number of a cell reference.
SHEETS	Returns the count of sheets in a given reference.

Logical Functions

IF	Test for a condition to be met and return a result if true, or false.
----	---

IFS	Allows for multiple logical tests without the need for nesting. New in Excel 2016.
IFERROR	Hide or handle formula errors.
SWITCH	Looks up a value in a list of values, and returns the result corresponding to the first matching value. New in Excel 2016.
XOR	Performs an 'exclusive or' test on an array, range or one or more expressions. New in Excel 2013.

Lookup and Reference Functions

ADDRESS	Return a text value of a cell address.
CHOOSE	Return a value from a list based on the position specified.
FILTER	Filter cells based on criteria.
GETPIVOTDATA	Get values from a PivotTable.
HLOOKUP	Look up a value in the top row of a table.
HYPERLINK	Insert links to internal or external locations.
INDEX	lookup a range of cells and return a single value, an array of values, a reference to a cell or range of cells.
INDIRECT	Convert text into a cell reference.

MATCH	Returns the relative position of a value. The lookup range can take the shape of a row or column.
OFFSET	Return a range of cells offset from a starting cell.
SORT	Sort cells or arrays in ascending or descending order.
SORTBY	Sort cells or arrays based on criteria.
UNIQUE	Extract a unique or distinct list from a range or array.
VLOOKUP	Look up a value in the first column of a table.
XLOOKUP	Like VLOOKUP but much better!
Math Functions	
AGGREGATE	Returns an aggregate (sum, average, min, max, count etc.) of a list or database.
MOD	Returns the remainder after a number is divided by a divisor.
MROUND	Rounds numbers to the nearest multiple.
RANDARRAY	Returns an array of random numbers between 0 and 1.
SEQUENCE	Returns list of sequential numbers that increment as specified.

SIGN	Returns the sign of a number.
SUBTOTAL	Choose the aggregation method e.g. SUM, AVERAGE, COUNT etc. and ignore or include filtered rows.
SUMPRODUCT	Much more than just multiplying arrays and returning the sum of the results.

Statistical Functions

FORECAST	Forecast future values using linear regression.
FORECAST.ETS	Forecast future values using Exponential Triple Smoothing and machine learning. New in Excel 2016
FORECAST.ETS.CONFINT	Returns a confidence interval for forecasted values. New in Excel 2016
FORECAST.LINEAR	Forecast future values using linear regression. New in Excel 2016, replacing the FORECAST function.
RANK	Return a rank for a dataset. Includes RANK.EQ and RANK.AVG

Text Functions

CHAR	Returns a character based on its Windows-1252 character number.
CLEAN	Removes non-printing characters from text.
CONCATENATE	Joins text from separate cells together.

SUBSTITUTE	Replaces new text for old text in a text string.
T	Checks whether a value is text, and returns the text if it is, otherwise returns a blank.
TEXT	Converts numbers to text in the format you specify.
TEXTJOIN	Joins text together and can ignore empty cells. New in Excel 2019.
TRIM	Remove the excess spaces from text, except for single spaces between words.

Excel Formula

A **formula** is an expression telling the computer what mathematical operation to perform upon a specific value. When referring to computer software, formulas are most often used in spreadsheet programs, such as Microsoft Excel.

Formulas in Excel are used to perform calculations or other actions on data entered into the formula and/or stored in program files.

They can range from basic mathematical operations, such as addition and subtraction, to complex engineering and statistical calculations.

Formulas are great for working out “what if” scenarios that compare calculations based on changing data. Once the formula is entered, you need only change the amounts to be calculated. You don’t have to keep entering “plus this” or “minus that” as you do with a regular calculator.

Formulas Start With the Equal Sign

Formulas begin with an equal (=) sign and, for the most part, are entered into the worksheet cell(s) where you want the results or answer to appear.

For example, if the formula =5 + 4 - 6 is entered into cell A1, the value 3 appears in that location. When cell A1 is selected, the formula displays in the formula bar above the worksheet.

Formula Breakdown

A formula can also contain any or all of the following:

- Values
- Constants
- Cell references
- Functions
- Operators

Values

Values in formulas are not just restricted to numbers but can also include:

- Dates
- Text: Words surrounded by quotation marks (" ")
- Boolean values: TRUE or FALSE only

Formula Constants

A constant is a value that does not change and is not calculated. Although constants can be well-known ones such as Pi (Π), the ratio of a circle's circumference to its diameter, they can also be any value, such as a tax rate or a specific date, which changes infrequently.

Cell References in Formulas

Cell references, such as A1 or H34, indicate the location of data in a worksheet. Rather than enter data directly into a formula, it is better to enter the data into worksheet cells and then enter the cell references to the location of the data into the formula.

The advantages of this are that:

- If you later change your data, the formula automatically updates to show the new result.
- In certain instances, using cell references makes it possible to copy formulas from one location to another in a worksheet.

To simplify entering multiple contiguous cell references into a formula, they can be entered as a range that just indicates the start and end points. For example, the references A1, A2, A3 can be written as the range A1:A3.

To simplify things even further, frequently used ranges can be given a name that can be entered into formulas.

Functions: Built-in Formulas

Excel also contains a number of built-in formulas called functions.

Functions make it easier to carry out:

- Commonly performed tasks, such as adding up columns or rows of numbers with the SUM function.
- Long or complex operations, such as finding specific information with the VLOOKUP function.

Formula Operators

An arithmetic or mathematical operator is the symbol or sign that represents an arithmetic operation in an Excel formula.

Operators specify the type of calculation being carried out by the formula.

Types of operators

The different types of calculation operators that can be used in formulas include:

- Arithmetic: Used for basic math, such as addition and subtraction.
- Comparison
- Text concatenation

Arithmetic Operators

Some of the arithmetic operators, such as the ones for addition and subtraction, are the same as those used in hand-written formulas. Arithmetic operators for multiplication, division, and exponents are different.

The arithmetic operators are:

- Subtraction: Minus sign (-)
- Addition: Plus sign (+)
- Division: Forward slash (/)
- Multiplication: Asterisk (*)
- Exponentiation: Caret (^)

If more than one operator is used in a formula, there is a specific order of operations that Excel follows in deciding which operation occurs first.

Comparison Operators

A comparison operator carries out a comparison between two values in the formula and the result of that comparison is either TRUE or FALSE.

There are six comparison operators:

- Equals (=)
- Less than (<)
- Less than or equal to (< =)
- Greater than (>)
- Greater than or equal to (> =)
- Not equal to (< >)

The AND and OR functions are examples of formulas that use comparison operators.

Concatenation Operator

Concatenation means to join things together and the concatenation operator is the ampersand (&). It can be used for joining multiple ranges of data in a formula.

An example of this would be

```
{=INDEX (D6:F11, MATCH (D3 & E3, D6:D11 & E6:E11, 0), 3)}
```

where the concatenation operator is used to combine multiple data ranges in a lookup formula using Excel's INDEX and MATCH functions.

Seven Basic Excel Formulas For Your Workflow

- SUM. The **SUM function**. The **function** will sum up cells that are supplied as multiple arguments. ...
- AVERAGE. The **AVERAGE function**. ...
- COUNT. The **COUNT function**. ...
- COUNTA. Like the **COUNT function**, COUNTA. ...
- IF. The **IF function**. ...
- TRIM. The **TRIM function**. ...
- MAX & MIN. The **MAX**.

Define chart and its features

A chart is a powerful tool that allows you to visually display data in a variety of different chart formats such as Bar, Column, Pie, Line, Area, Doughnut, Scatter, Surface, or Radar charts. With Excel, it is easy to create a chart. Here are some of the types of charts that you

can create in Excel.

Microsoft Office's Excel application allows users to store, model and manipulate data sets. Excel spreadsheets organize this data into worksheets, each with a number of rows and columns. Each row or column has one or more cells, with each cell holding a single data value. Excel provides a range of automated functions for working with these stored data values, including the chart function. Once data is stored in an Excel spreadsheet, anyone with access to that spreadsheet can manipulate the data in order to analyze and communicate its meaning. The charting function can be a key element in these processes.

Advantages of Charts

Excel charts allow spreadsheet administrators to create visualizations of data sets. By highlighting a set of data within an Excel spreadsheet and feeding it into the charting tool, users can create various types of charts in which the data is displayed in a graphical way. This can aid understanding of a data set, as well as communication of it, with Excel charts suitable for use within management or corporate presentations.

A chart can create a clearer picture of a set of data values than a table with rows of numbers in it, allowing managers to incorporate this understanding into analysis and future planning. Benefits of charts and graphs also come into play in presentations, where they can be used to quickly illustrate trends in data for others to see.

It's often easier to spot trends and patterns in graphs and charts than by simply looking through a long table of data, especially for big datasets, so one of the common uses of chart tools is in simply coming to grips with what a new dataset actually means.

Automation of Chart Creation

The Excel application automates the process of generating charts from existing data sets. If a spreadsheet already contains updated data, the chart function can transform this data into a chart with a minimal amount of user input. The Recommended Charts tool in Excel is a key part of this process. Using the chart function, spreadsheet administrators can generate a chart in a few clicks, choosing a chart type, as well as options such as labels, axes and titles.

This is often a big advantage over drawing a chart by hand with an illustration tool or even using a more complex visualization tool, many of which require programming or configuration.

Customization of Charts in Excel

The chart function in Excel enables users to strike a balance between automation and customization. Although the program can generate a chart quickly and easily, the user is also able to assert control over many chart details if necessary. As well as customizing the data presented within a chart, users can finely tweak the chart options to any detail. Charts in Excel can also be altered after they have been initially generated, by selecting a chart and choosing from the available options to amend the chart type, data or other details.

Integration into Existing Sheets

If a business or other organization is using spreadsheet data managed within Excel, using the chart function within Excel aids integration of the data. For example, when an Excel spreadsheet generates a chart from the data within a worksheet, the chart will automatically update whenever the data itself is edited. This allows business managers and administrators to manage their data and visualizations within a single application, with the results updating instantly. Excel spreadsheets can also be easily shared between users and computers.

This can be one of the advantages of Excel tables and charts over other data visualization systems where charts and datasets are effectively stored and managed separately.

Working with graphics in excel

Definition: - In Microsoft Excel, a chart is often called a graph. ... A chart is a powerful tool that allows you to visually display data in a variety of different chart formats such as Bar, Column, Pie, Line, Area, Doughnut, Scatter, Surface, or Radar charts.

Excel provides several features that make inserting graphics into your worksheet a breeze. However, once your graphic is in place, you may wonder how to format it. The following articles explore the many ways you can edit graphics in Excel.

Adding a Drop Shadow to a Text Box

One way to make your text boxes "stand off" the page is to add a drop shadow to them. This tip shows just how easy it is to add this formatting touch.

Adding Auto Shapes

The graphics features of Excel allow you to add a number of predefined AutoShapes to a workbook. If you want to add shapes to the AutoShapes feature, however, you are out of luck. The shapes are apparently hard-coded into Excel, and cannot be modified. There are workarounds, however, and they are covered here.

Adding Data Labels to Your Chart

Adding labels to a chart can make the information presented in the chart more understandable. Excel allows you to add different types of data labels to your charts, as discussed in this tip.

Adding Drop Shadows

Want your shapes to really "pop" off the page? Add a drop shadow to them, as described in this tip.

Adding Text to an AutoShape

You can add text to all sorts of drawing shapes, not just text boxes. Here's how easy it is.

Adjusting Picture Appearance

Excel provides the Picture toolbar to help you modify any images in your worksheet. This tip explains how to use the toolbar to adjust the appearance of a graphic.

Assigning Macros to Graphics

The graphics you place in a worksheet can do more than just look pretty. You can also assign macros to a graphic, which are triggered when the graphic is later clicked.

Capturing a Screen

A picture is worth a thousand words, but getting the picture—particularly a screen shot—into a workbook may seem daunting. Here's a couple of easy ways to get the picture you need.

Changing an AutoShape

Need to change a shape you previously added to your worksheet? It's easy to do using the graphics tools provided by Excel. Here's how.

Changing Chart Types

Want to change an existing bar chart to a different type of chart, such as a line chart or a column chart? It's easy to do when you follow the steps in this tip.

Changing How Arrows Look

If you use Excel's graphic capabilities to insert a line or an arrow into a worksheet, you can change how that arrow looks. For instance, you can change the type of arrowhead used and the ends of the line on which those arrowheads appear.

Changing Line Color in a Drawing Object

Don't like the color of the lines that Excel chose for your drawing object? It's easy to choose your own colors, as pointed out in this tip.

Changing the Axis Scale

When creating a chart, you may want to adjust the default scaling that Excel applies to an axis. This is relatively easy to do by following the steps outlined in this tip.

Changing the Pattern Used in a Data Series

When you create a chart, Excel attempts to assign colors to your data series that it thinks will work best for you. If you want to change those colors to something else—including a variety of patterns—you can easily do so.

Changing the Size of a Drawing Object

Add a drawing object to your worksheet, and at some point you may want to change that object's size. You can easily change sizes by using the technique described in this tip.

Changing the Size of a Graphic

Adding a graphic to a worksheet is easy. Getting that graphic to just the right size may take a little bit of trial and error. Here's how to adjust the size easily.

Colorizing Charts

Need to change the color of different parts of your chart? It's easy to do when you apply the technique described in this tip.

Controlling Display of the Scroll Bars

The scroll bars can be very handy in navigating your worksheet, but did you know you can turn them off or on at will? Here's how.

Controlling Display of the Status Bar

Turning off the status bar through the options menu.

Creating a Drawing Object

Creating simple drawing objects is easy in Excel. All you need to do is use the tools made available on the Drawing toolbar.

Creating a Log/Log Chart

If you need to create a chart that uses logarithmic values on both axes, it can be confusing how to get what you want. This tip explains which of the chart types in Excel is best suited for the type of chart you need.

Creating an AutoShape

Ever wanted to create a simple drawing in your worksheet? Excel has made this simple. This tip explains how Excel uses AutoShapes.

Creating an Organization Chart

The graphics capabilities of Excel are flexible enough that you can use the program to create organization charts. Here's how quick overview of how easy creating such a chart can be.

Creating Custom Chart Formats

Excel allows you to create custom chart formats that go beyond the standard formats provided in the program. You can then apply the custom chart format to future chart-creating endeavors. Here's how to create and save your custom chart formats.

Coming Topics of this Unit- Descriptive Statistics and steps involved in calculation of descriptive statistics in MS Excel, Mean, Median, Mode, Range, Standard deviations, Skewness, Kurtosis.

Descriptive Statistics

If we simply present our raw data, it would be hard to visualize what the data is showing, especially if there is a lot of it. In order to reduce the complexity of data and to make them comparable, we can resort to descriptive statistics. Descriptive statistics, therefore, enables us to present the data in a more meaningful way, which allows simpler interpretation of the data.

Descriptive statistics consists of methods for organizing and summarizing information. Descriptive statistics are numbers that are used to summarize and describe data. The word "data" refers to the information that has been collected from an experiment, a survey or a historical record, etc. Descriptive statistics includes the construction of graphs, charts & tables and the calculation of various descriptive measures such as averages, measures of dispersion and measures of variation.

MEASURES OF CENTRAL TENDENCY

The most important objective of statistical analysis is to obtain one single value representing the entire data. Such a value is called average or central value. The following definitions may be given to explain its meaning more clearly.

According to Simpson and Kafka, "A measure of central tendency is a typical value around which other figure congregate."

According to Clark, "Average is an attempt to find one single figure to describe whole of figures."

An average is a single value which represents the whole set of figures and all other individual items concentrate around it. In other words, an average is single value within the range of the data that is used to represent all of the values in the series. Such an average is somewhere within range of the data, it is therefore called measure of central tendency.

There are three measures of central tendency - Arithmetic Mean, Median and Mode.

Arithmetic Mean: Arithmetic mean is the most popular and widely used measure of central tendency, representing the entire data by a single value. Generally, it is termed as an 'average' to a layman and for a statistician it means "arithmetic mean' or 'mean'. Arithmetic mean is defined as the value which is obtained by adding all the items of a series and dividing this by total by the number of items. The mean is computed as:

$$\bar{X} = \frac{\sum X}{N}$$

where \bar{X} =(X-bar) is the symbol for the mean.

\sum (The Greek letter sigma) is the symbol for summation.

X is the symbol for the values.

N is the symbol for the total number of values.

Example: The marks of 9 students in economics are given 21, 16, 18, 13, 15, 19, 17, 20, and 23. Compute the mean from the data.

Solution: The mean would be $=\frac{21+16+18+13+15+19+17+20+23}{9}=18$

The mean marks of these students will be 18.

Median: Median is a positional average and important measure of central tendency. Median is defined as the middle value of the series when arranged either in ascending or descending order. It is a value which divides the arranged series into two equal parts in such a way that the number of observations smaller than median is equal to the number greater than it. Median is thus, a positional average. The following definitions may be given to explain its meaning more clearly.

According to Cornor, "The median is that value of the variable which divides the group into two equal parts, one part of comprising all values greater and the other values less than the median."

According to SeCrist, "Median of the series is the value of the item actual or estimated when a series is arranged in order of magnitude which divides the distribution into two equal parts."

Median is always determined by first arranging the series in an ascending or descending manner. Median is denoted by the symbol 'M'. The formula used for calculating median is:

$$\text{Median} = \left(\frac{n+1}{2}\right)^{\text{th}} \text{ item}$$

Example: The marks of 9 students in economies are given 122, 116, 118, 113, 115, 119, 117, 120, 123. Compute the median from the data.

Solution: By arranging the data in ascending order of magnitude, we obtain: 113, 115, 116, 117, 118, 119, 120, 122, 123. The median will be the value of the 5th observation arranged as above. Therefore, the value of median equals 118. This means 50 percent of the students are having marks below 118 and remaining 50 percent are having marks above 118.

Mode: The word Mode is derived from a French word la mode which means fashion or the most popular phenomenon. Thus, mode is defined as the value which occurs most frequently in a series. In other words, it is a value which has the greatest frequency in a distribution. Mode is denoted by the symbol 'Z'. For example, the mode of the series;

19, 20, 22, 22, 22, 22, 25, 26, 26

would be 22, since this value occurs most frequently (4 times) than any other values.

According to Kenny and Keeping. "The value of the variable which occurs most frequently in a distribution is called the mode"

Mode is a concept which is widely used in business, eg, a shoe store owner would be naturally interested in knowing the size of shoe that the majority of the customers ask for.

4.2 MEASURE OF DISPERSION

Measures of central tendency do not indicate how terms are spread out on either side of the centre. Therefore, there is a need to study the spread of a distribution of a variable and the methods which provide that are called measures of dispersion. The study of dispersion helps in taking better decisions. Measures of dispersion help us to study the variability of the items, ie... the extent to which the items vary from each other and from the central value. This is because small dispersion indicates high uniformity of the items, whereas, large variability denotes less uniformity. The following definitions may be given to explain the meaning dispersion more clearly

According to Bowley, "Dispersion is a measure of the variations of items."

According to Spiegel, "The degree to which numerical data tend to spread about average is called variation or dispersion of data."

METHODS OF MEASURING DISPERSION

Following are the main methods of measuring dispersion:

1. Range
2. Interquartile Range and Quartile Deviation
3. Mean Deviation
4. Standard Deviation
5. Co-efficient of Variation

(In this chapter we will cover only range and standard deviation.) **Range**: This is simplest measure of dispersion and is defined as the difference between the largest and smallest values in the series. Its formula is:

$$\text{Range} = L - S$$

where: R Range, L- Largest value in the series, S-Smallest value in the series

Range

Range shows the mathematical distance between the lowest and highest values in the data set. Range measures the variability of the data set. A wide range indicates greater variability in the data, or perhaps a single outlier far from the rest of the data. Outliers may skew, or shift, the mean value enough to impact data analysis.

In the sample group, the lowest value is 20 and the highest value is 36.

To calculate range, subtract the lowest value from the highest value. Since

$$36 - 20 = 16$$

the range equals 16.

In the sample set, the high data value of 36 exceeds the previous value, 25, by 11. This value seems extreme, given the other values in the set. The value of 36 might be an outlier data point.

Standard Deviation: The concept of standard deviation was introduced by a great statistician **Karl Pearson** in 1893. It is the most important and widely used measure of dispersion. Standard deviation is also called as root mean square deviation. Standard deviation is defined as the square root of arithmetic mean of the squares of the deviation of the values taken from the mean.

In statistics, the **standard deviation (S.D)**, also represented by the Greek letter sigma, σ for the population standard deviation or **s** for the sample standard deviation is a measure that is used to quantify the amount of variation or dispersion of a set of data values. A standard deviation close to 0 indicates that the data points tend to be very close to the mean (also called the expected value) of the set, while a high standard deviation indicates that the data points are spread out over a wider range of values.

The formula for the standard deviation is:

$$\sigma = \sqrt{\frac{\sum d^2}{N}}$$

Where: σ (Greek letter little sigma) is the standard deviation.

d^2 is the square of value's deviation from the mean.

N is the number of values.

A large standard deviation indicates that the data points can spread far from the mean and a small standard deviation indicates that they are clustered closely around the mean.

SKEWNESS

Skewness refers to lack of symmetry in a frequency distribution means a tendency of 'twist' or 'turn'. Skewness denotes the degree of departure of a distribution from symmetry and reveals the direction of scatterness of the items. It gives us an idea about the shape of the frequency curve. When a distribution is not symmetrical, it is called a skewed distribution. Skewness tells us about the asymmetry of the frequency distribution. The following definitions may be given to explain the meaning of skewness more clearly.

According to M.R.Speigal, "Skewness is the degree of asymmetry or departure from symmetry of a distribution."

According to Croxten and Cowden, "When a series is not symmetrical, it is said to be asymmetrical or skewed"

There are three types of distribution, viz. symmetrical distribution, positively skewed distribution and negative skewed distribution.

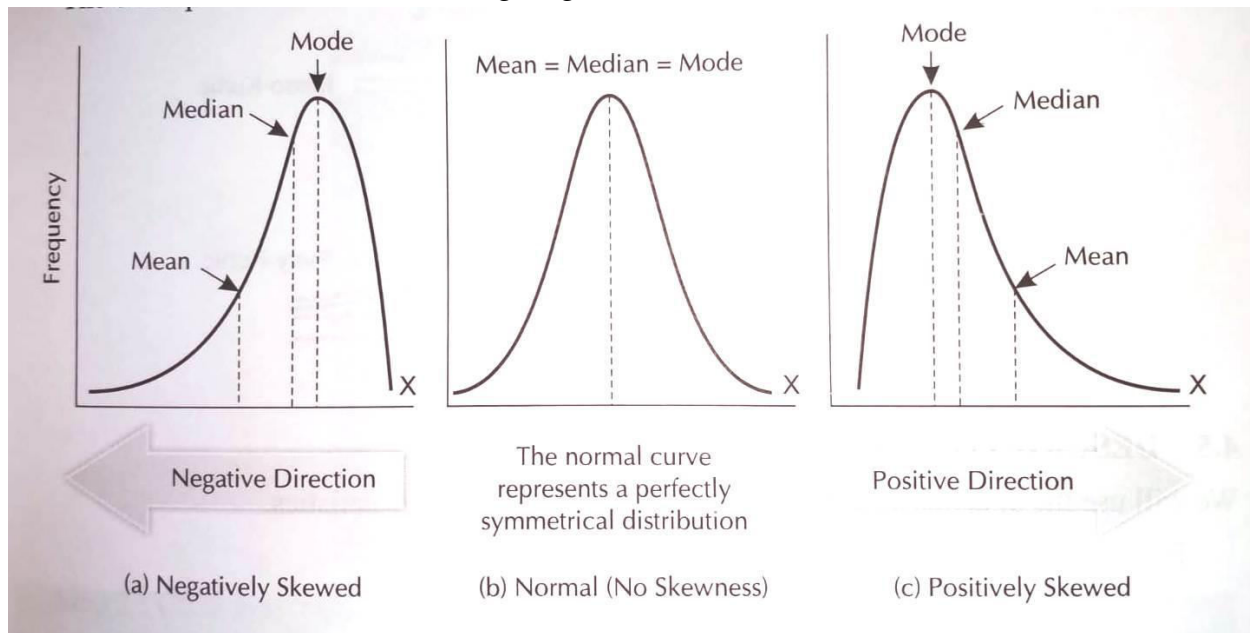
Normal or Symmetrical Distribution: In a symmetrical distribution or symmetrical curve, skewness is not present. The values of mean, median and mode coincide, i.e.,

$\bar{X} = M = Z$. The spread of the frequencies is the same on the both sides of the central point or the curve.

Positively Skewed Distribution: A distribution is called positively skewed when the value of the mean is greater than median and median is greater than mode, i.e., $\bar{X} > M > Z$ and the longer tail of the frequency curve of distribution lies to the right of the central point.

Negatively Skewed Distribution: A distribution is called negatively skewed when the value of the mean is less than median and median is less than mode. i.e. $\bar{X} < M < Z$ and the longer tail of the

frequency curve of the distribution lies to the left of the central point. The concept of skewness will be made clearer from the following diagram.



KURTOSIS

Like average, dispersion and skewness, **kurtosis** is the fourth characteristic of a frequency distribution which give us an idea about the shape of the frequency distribution. Kurtosis is a Greek word meaning bulkiness. In statistics, it refers to the degree of flatness or peakedness of a frequency curve. The degree of kurtosis (or peakedness) of a distribution is measured relative to the peakedness of the normal curve.

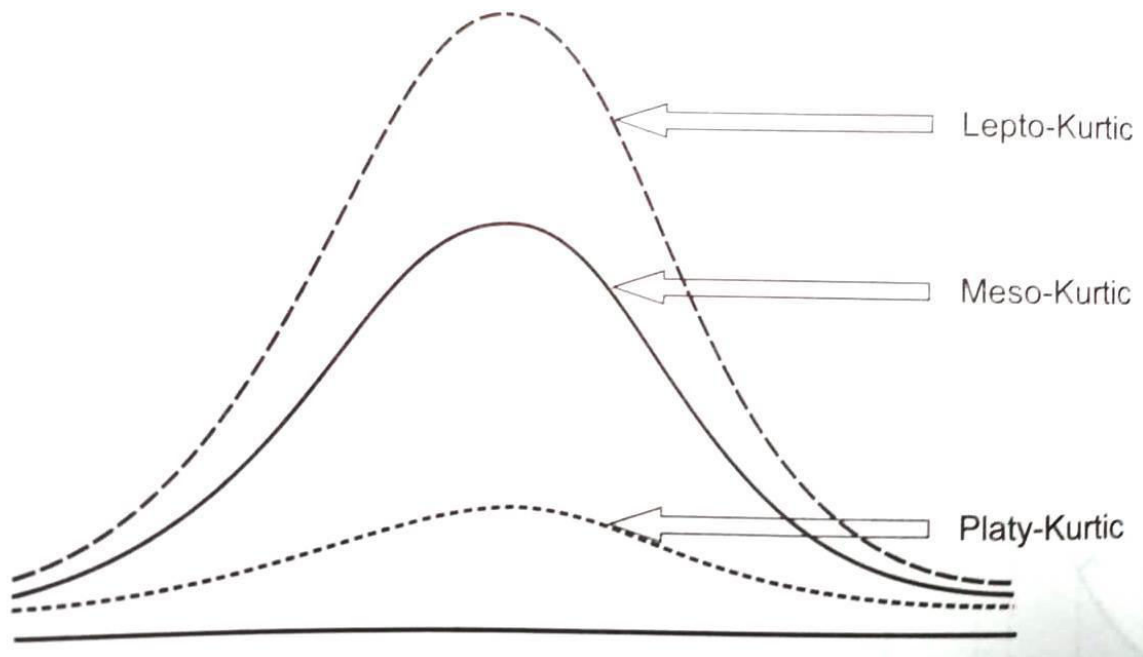
According to Croxten and Cowden, “A measure of kurtosis indicates the degree to which a frequency distribution is peaked or flat-topped.”

Thus, a measure of kurtosis tells us the extent to which a distribution is more peaked or flat-topped than the normal curve.

There are three types of kurtosis in a distribution;

1. **Lepto-kurtic** : if a curve is having more peak than the normal curve, it is called leptokurtic. In such a curve, there is too much concentration of the items near the centre.
2. **Meso-kurtic** : A curve having normal peak to the normal curve itself is called mesokurtic. In such a curve, there is equal distribution of items around the central value.
3. **Platy-kurtic**: A curve having a low peak(or flat topped) than the normal curve is called platykurtic. In such a curve, there is less concentration of items near the centre.

The figure below illustrates three different types of kurtosis.



Types of Kurtosis

DESCRIPTIVE STATISTICS USING MS EXCEL

We will use the analysis ToolPak add-in to generate descriptive statistics.

First, make sure that the “**Analysis ToolPak**” is installed.

Steps to Follow

Under “Data” tab is the option “Data Analysis” present?

If Yes- Analysis ToolPak is installed. Then go further.

If No – Install Analysis ToolPak.

Installing Analysis ToolPak

To load the Analysis ToolPak add-in, executes the following steps.

1. Click on the green **File** tab in excel 2010 replaces the **Office Button** (for file menu)in previous versions of Excel.
2. Click on **options**.
3. Under Add-ins, select **Analysis ToolPak** and click on the **Go** button.
4. Check Analysis ToolPak and click on OK.
5. Under the Data tab, we can now click on **Data Analysis**.

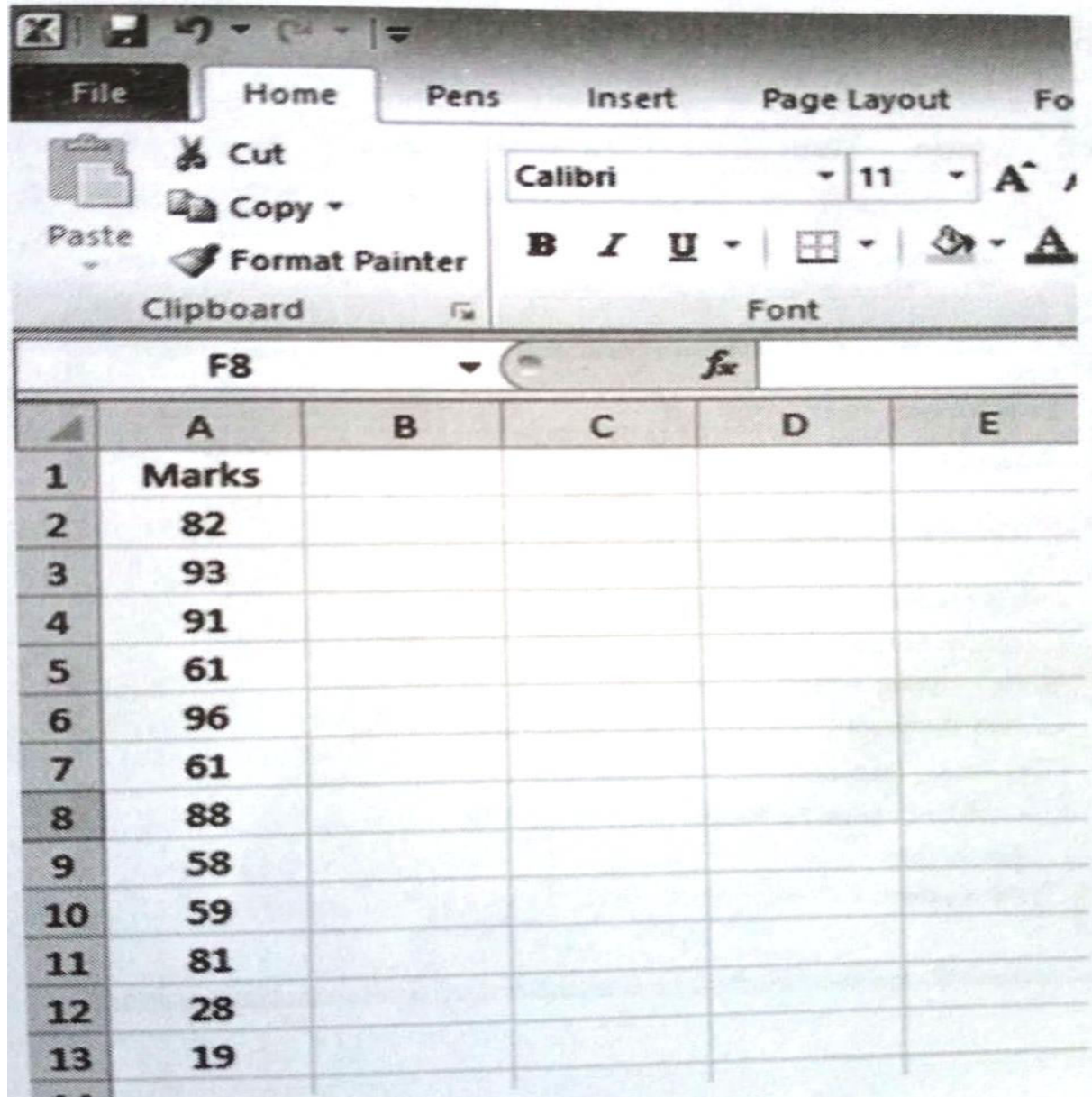
Illustration 1: The marks Obtained by some students in the statistics examination are:

82	93	91	61	96	61
88	58	59	81	28	19

Using MS Excel, generate descriptive statistics viz. mean, median, mode, range, standard deviation, skewness and kurtosis.

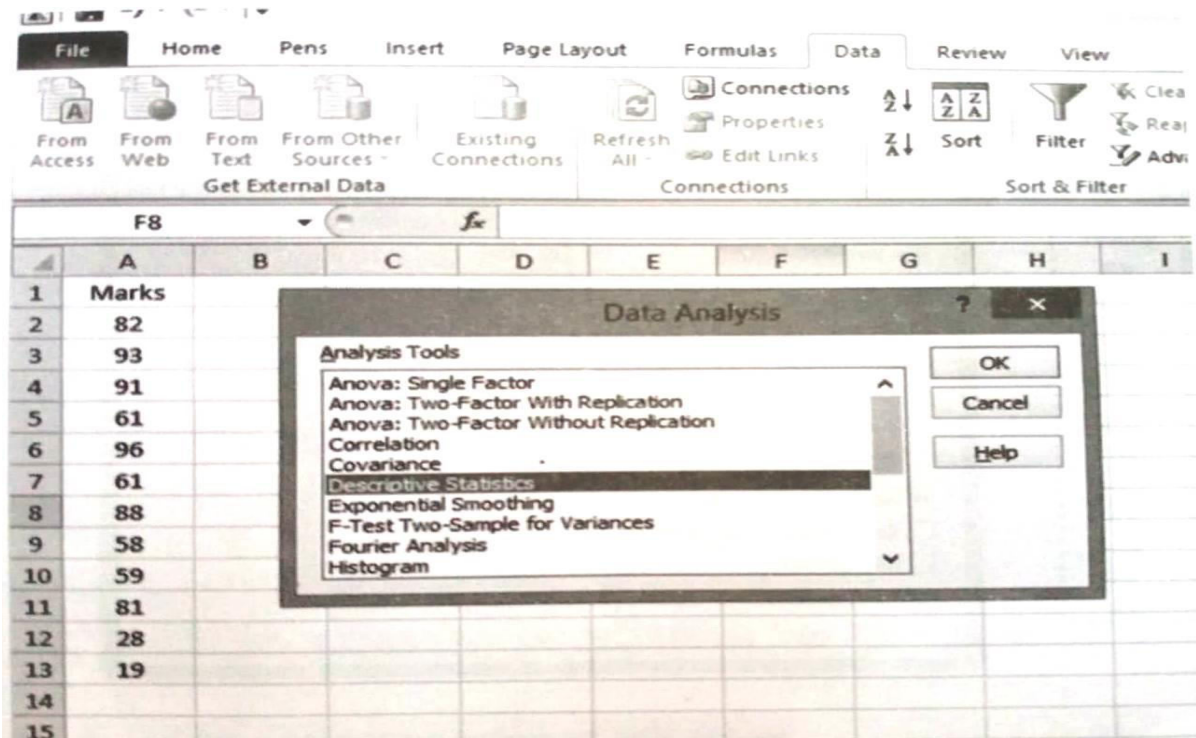
Excel Solution: Steps to calculate descriptive statistics using MS Excel.

Step 1: Type data into Excel, in a single column. For example, we have twelve items in our data set; type them in cells A2 through A13.

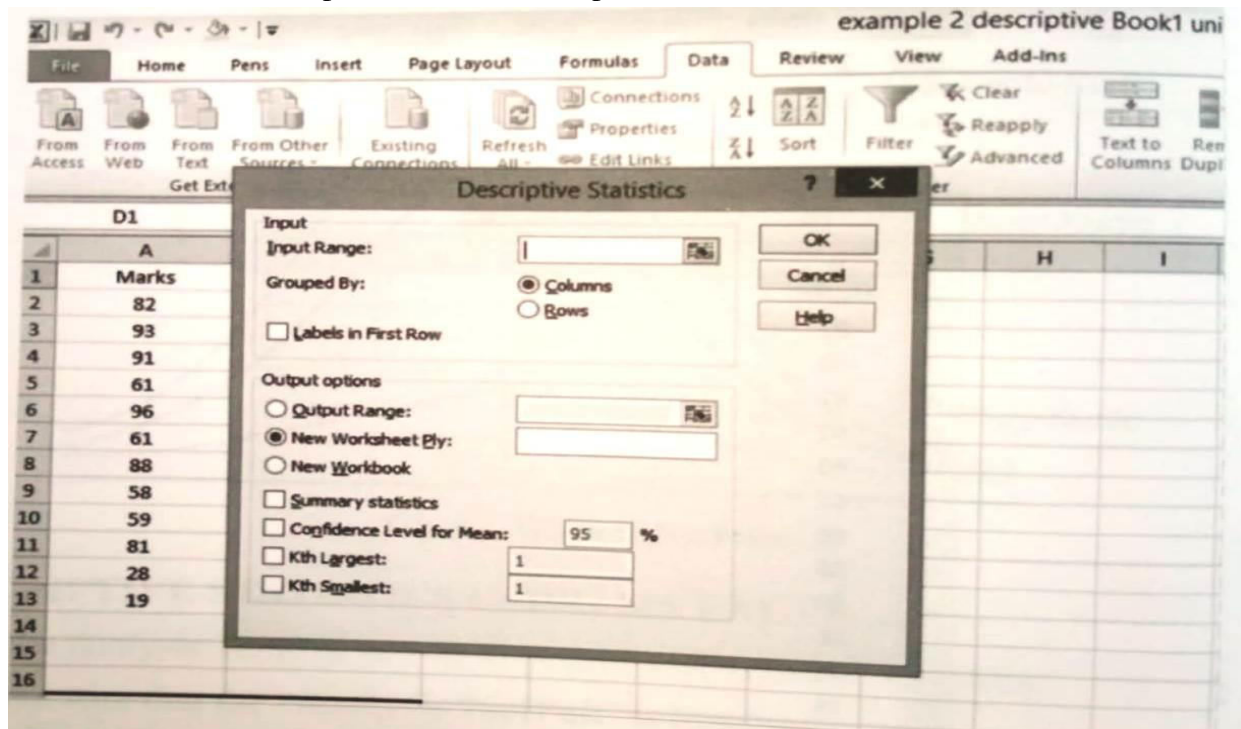


Step 2: Click the “Data” tab and then click **“Data Analysis”**

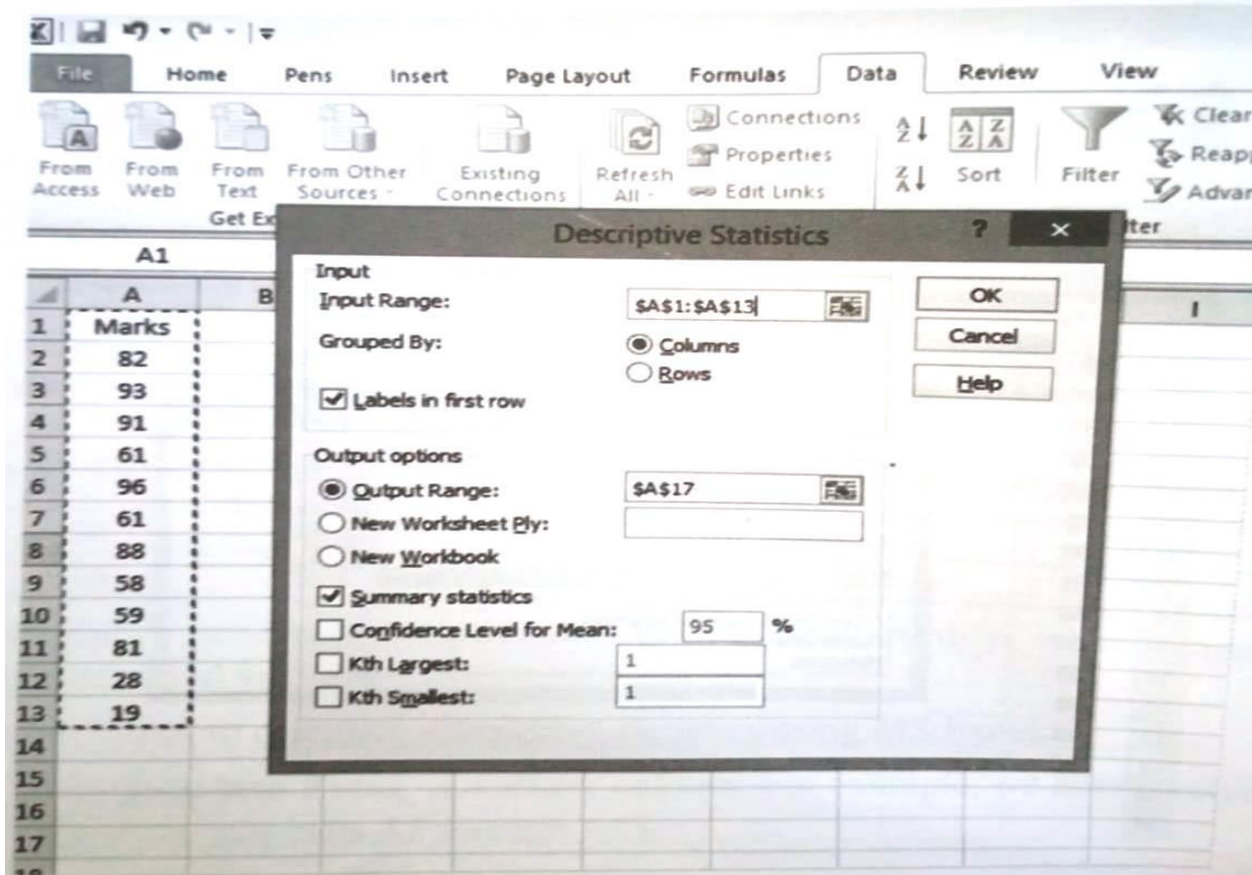
Step 3: Select “Descriptive Statistics” and click on OK in the pop-up Data Analysis window.



A window named Descriptive Statistics will open.



Step 4: Type an input range into the “Input Range” text box. For this example, “A1:A13”.



Step 5: Check the box “Labels in first row”, as we have titled the column in row 1, otherwise leave the box unchecked.

Step 6: Type a cell location into the “Output Range” box. For example, type “A17”, Make sure that two adjacent columns do not have data in them.

Step 7: Click the “Summary Statistics” check box and then click on “OK” to display descriptive statistics.

A list of descriptive statistics will return in the column which we have selected as the Output Range.

The screenshot shows the Microsoft Excel interface with the Data tab selected. The ribbon includes options for 'From Access', 'From Web', 'From Text', 'From Other Sources', 'Existing Connections', 'Refresh All', 'Connections', 'Properties', and 'Edit Links'. The active cell is A17, containing the text 'Marks'. Below this, a table of descriptive statistics is displayed in the range A17:E31.

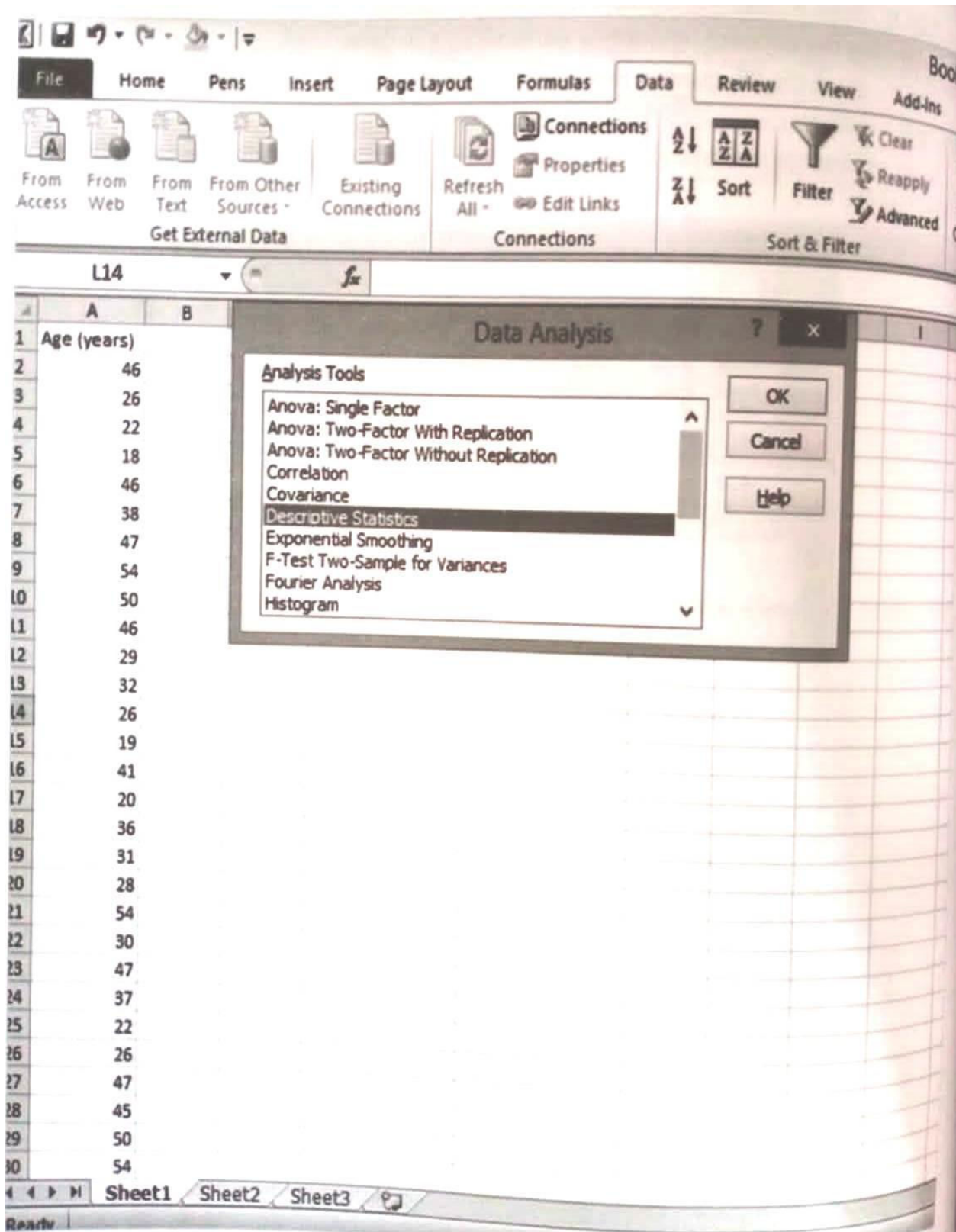
	A	B	C	D	E
15					
16					
17	<i>Marks</i>				
18					
19	Mean		68.08333333		
20	Standard Error		7.273299378		
21	Median		71		
22	Mode		61		
23	Standard Deviation		25.19544812		
24	Sample Variance		634.8106061		
25	Kurtosis		-0.162602451		
26	Skewness		-0.839784032		
27	Range		77		
28	Minimum		19		
29	Maximum		96		
30	Sum		817		
31	Count		12		
32					
33					

Illustration 2: The age of 29 consumers are given below:

46,26,22,18,46,38,47,54,50,46,29,32,26,19,41,20,36,31,28,54,30,47,37,22,26,47,45,50,54
calculate the descriptive statistics using MS Excel.

Excel Solution:

By following the steps as stated in **illustration no. 1.**



We will get the results as shown here:

The screenshot shows an Excel spreadsheet with a data table in column A and its descriptive statistics in columns D and E. The data table lists ages from 19 to 54. The descriptive statistics include Mean, Standard Error, Median, Mode, Standard Deviation, Sample Variance, Kurtosis, Skewness, Range, Minimum, Maximum, Sum, and Count.

Age (years)	Mean
46	36.79310345
26	2.183333258
22	37
18	46
46	11.75760942
38	138.2413793
47	-1.411343835
54	-0.070168748
50	36
46	18
29	54
32	1067
26	29
19	
41	
20	
36	
31	
28	
54	
30	
47	
37	
22	
26	
47	
45	
50	
54	

LIMITATIONS OF DESCRIPTIVE STATISTICS

Descriptive statistics are limited in use so much that they only allow us to make summations about the people or objects that we have actually measured. We cannot use the data that we have collected to generalize to other people or objects (i.e., using data from a sample to infer the properties/parameters of a population). For example, if we tested a drug to beat cancer and it worked in our patients, we cannot claim that it would work in other cancer patients only relying on descriptive statistics, but inferential statistics would give us this opportunity.

UNIT 3RD

Hypothesis Testing: Chi-square test, t-Test and Analysis of Variance (ANOVA) with the help of MS Excel. An introduction to SPSS: coding, graphs, pivot tables, missing values etc. Application of Chi-square test, t-Test and Analysis of Variance (ANOVA) using SPSS.

Hypothesis Testing: - Hypothesis testing is a form of statistical inference that uses data from a sample to draw conclusions about a population parameter or a population probability distribution. A hypothesis refers to an assumption that is made find out solution of a problem. The procedure to utilize statistic to check the probability that hypothesis is true or not called hypothesis testing.

What is chi-square t test and ANOVA?

Chi-square test is used on contingency tables and more appropriate when the variable you want to test across different groups is categorical. It compares observed with expected counts. Both t test and ANOVA are used to compare continuous variables across groups.

Chi-Square Test in Excel

The chi-square test is a non-parametric test that compares two or more variables from randomly selected data. It helps find the relationship between two or more variables. In Excel, we calculate the chi-square p-value. Since Excel does not have an inbuilt function, mathematical formulas are used to perform the chi-square test.

Chi Square Test in Excel

Chi-Square Points= (Observed-Expected)^2/Expected			
Service	Salary		
	Low	Medium	High
Excellent	0.055576923	1.265014	1.256622889
Good	1.734215686	1.652353	4.868871353
Poor	2.925217391	0.515958	4.384400848
CHI-SQUARE	18.65823041		
Critical Value of Chi-square =	9.487729037		

$$\chi^2 = \sum \frac{(O_i - E_i)^2}{E_i}$$

There are two types of chi-square tests which are listed as follows:

1. Chi-square goodness of fit test
2. Chi-square test for independence

Chi-Square Goodness of Fit Test

The goodness of fit test helps determine whether the sample data matches the population or not. In other words, it shows how well the sample data fits a set of observations.

The symbol of the chi-square test is “x²” (i.e., “x” raised to the power 2). “x²” is the summation of the (observed count–expected count)²/expected count.

The formula of the chi-square goodness of fit test is given as follows:

$$x^2 = \sum_{i=1}^k (O_i - E_i)^2 / E_i +$$

Where,

- “x²” is the chi-square statistic
- “O_i” is the observed frequency
- “E_i” is the expected frequency
- “i” is the “ith” position in the contingency table
- “k” is the category
- Degrees of freedom (df)=k-1

The Uses of the Goodness of Fit Test

The test is used in the following situations:

- To evaluate the creditworthiness of borrowers based on their age group and history of debt obligations
- To establish a relation between the performance of sales representatives and the training received by them
- To compare the returns of a single stock with the returns on stocks of the whole sector
- To evaluate the impact of a television campaign on a category of viewers

Chi-Square Test for Independence

It helps determine whether the variables are independent of one another or not. Two random variables are called independent if the probability distribution of one variable is not affected by the other.

The formula of the chi-square test for independence is given as follows:

$$x^2 = \sum_{i=1}^r \sum_{j=1}^c \frac{(O_{ij} - E_{ij})^2}{E_{ij}}$$

Where,

- “x²” is the chi-square statistic
- “O_{ij}” is the observed frequency in the ith row and jth column
- “E_{ij}” is the expected frequency in the ith row and jth column
- “r” is the number of rows
- “c” is the number of columns
- Degrees of freedom (df)=(r-1)(c-1)

The formula for calculating the expected frequency in the ith row and jth column is given as follows:

$$E_{ij} = \frac{\text{row } i \text{ total} * \text{col } j \text{ total}}{\text{grand total}}$$

The Uses of the Chi-Square Test for Independence

The test is used in the following situations:

- There are two categorical variables and the relation between them is to be determined.
- There are cross-tabulations (two-way tables) and the relationship between various categorical variables needs to be studied.
- There are non-quantifiable variables. For example, one needs to determine the reason behind variations of health plans across age groups.

The Characteristics of the Chi-Square Test

The features of the chi-square test are listed as follows:

- It evaluates whether observed frequencies vary significantly from the expected frequencies under a given set of assumptions.
- It determines how well an assumed distribution fits the data.
- It uses the contingency tables (or crosstabs) to summarize the relationship between various categorical variables.
- It supports nominal level measurements.

Note: In the simplest form, the chi-square distribution is the square of the standard normal distribution.

How to Perform the Chi-Square Test in Excel? (With Example)

Chi Square Test Excel Template - Microsoft Excel

Home Insert Page Layout Formulas Data Review View

Clipboard Font Alignment Number Styles Cells

B18 =CHITEST(B4:D6,B14:D16)

Observed Frequencies					
Service	Salary			Total	
	Low	Medium	High		
Excellent	9	10	7	26	
Good	11	9	31	51	
Poor	12	8	3	23	
Total	32	27	41	100	
N	100				

Chi-Square Points=(Observed-Expected) ² /Expected				
Service	Salary			
	Low	Medium	High	
Excellent	0.055576923	1.265014	1.256622889	
Good	1.734215686	1.652353	4.868871353	
Poor	2.925217391	0.515958	4.384400848	
CHI-SQUARE	18.65823041			
Critical Value of Chi-square =	9.487729037			

Expected Frequencies(Variables Perfectly Independent)				
Service	Salary			
	Low	Medium	High	
Excellent	8.32	7.02	10.66	
Good	16.32	13.77	20.91	
Poor	7.36	6.21	9.43	

Chi-Test (P)Value = 0.00091723

A restaurant manager wants to find the relationship between quality of service and the salary of customers waiting to be served.

She organizes the task in the following way:

- A random sample of 100 customers is considered.
- Every customer is asked to rate the service of the restaurant as “excellent,” “good,” and “poor.”

She constructs the following hypothesis:

- Null hypothesis (H₀)–The quality of service is not dependent on the salary of customers waiting to be served.
- Alternative hypothesis (H₁)–The quality of service is dependent on the salary of customers waiting to be served.

The manager divides the customers into three categories based on their salaries–“low,” “medium,” and “high.” The level of significance (α) is 0.05.

The findings are presented as nine data points shown in the following table.

	A	B	C	D	E
1	Observed Frequencies				
2	Service	Salary			Total
3		Low	Medium	High	
4	Excellent	9	10	7	
5	Good	11	9	31	
6	Poor	12	8	3	
7	Total	32	27	41	
8					

Let us calculate the sum of all the rows and columns. We apply the following SUM formula to add the numbers of the fourth row.

"=SUM(B4:D4)"

Observed Frequencies				
Service	Salary			Total
	Low	Medium	High	
Excellent	9	10	7	=SUM(B4:D4)
Good	11	9	31	
Poor	12	8	3	
Total	32	27	41	

Press the "Enter" key and the sum appears in cell E4. The output is 26. Similarly, we apply the SUM formula to the remaining rows and columns. There are 27 respondents with medium salary and 51 respondents who rated the service quality as "good."

Observed Frequencies				
Service	Salary			Total
	Low	Medium	High	
Excellent	9	10	7	26
Good	11	9	31	51
Poor	12	8	3	23
Total	32	27	41	100

We apply the formula " $(r-1)(c-1)$ " to calculate the degrees of freedom (df).

$$df = (3-1)(3-1) = 2 \times 2 = 4$$

We apply the following formula to calculate the expected frequency for column B and row 4.

$$"=(B7 * E4 / B9)"$$

The calculation is shown in the following image.

Observed Frequencies				
Service	Salary			Total
	Low	Medium	High	
Excellent	9	10	7	26
Good	11	9	31	51
Poor	12	8	3	23
Total	32	27	41	100

N	100
---	-----

Expected Frequencies(Variables Perfectly Independent)			
Service	Salary		
	Low	Medium	High
Excellent	=B7 * E4 / B9		
Good			
Poor			

The expected number of customers who have “low” salary but rated the restaurant service as “excellent” is 8.32.

In the following calculations, E11 is the expected frequency of the first row and the first column. E12 is the expected frequency of the first row and the second column.

- $E11=(26*32)/100=8.32$, $E12=7.02$, $E13=10.66$
- $E21=16.32$, $E22=13.77$, $E23=20.91$
- $E31=7.36$, $E32=6.21$, $E33=9.43$

Similarly, we calculate the expected frequencies for the entire table, as shown in the succeeding image.

	A	B	C	D	E
11	Expected Frequencies(Variables Perfectly Independent)				
12	Service	Salary			
13		Low	Medium	High	
14	Excellent	8.32	7.02	10.66	
15	Good	16.32	13.77	20.91	
16	Poor	7.36	6.21	9.43	
17					

Let us calculate the chi-square data points by using the following formula.

Chi-square points= $(\text{observed}-\text{expected})^2/\text{expected}$

We apply the formula “ $=(B4-B14)^2/B14$ ” to calculate the first chi-square point.

		Chi-Square Points= (Observed-Expected) ² /Expected		
		Salary		
Service		Low	Medium	High
Excellent	$=(B4-B14)^2/B14$			
Good				
Poor				

We copy and paste the formula to the remaining cells. This is done to fill values in the entire table, as shown in the following image.

Excel formula bar: $= (B4 - B14)^2 / B14$

Chi-Square Points= (Observed-Expected) ² /Expected			
Service	Salary		
	Low	Medium	High
Excellent	0.055577	1.265014	1.256622889
Good	1.734216	1.652353	4.868871353
Poor	2.925217	0.515958	4.384400848

Let us calculate the chi-square calculated value by adding all the values given in the succeeding table.

Excel formula bar: $= \text{SUM}(I5:K7)$

Chi-Square Points= (Observed-Expected) ² /Expected			
Service	Salary		
	Low	Medium	High
Excellent	0.055576923	1.265014	1.256622889
Good	1.734215686	1.652353	4.868871353
Poor	2.925217391	0.515958	4.384400848

CHI-SQUARE $= \text{SUM}(I5:K7)$

The chi-square calculated value is 18.65823.

Excel formula bar: $= \text{SUM}(I5:K7)$

Chi-Square Points= (Observed-Expected) ² /Expected			
Service	Salary		
	Low	Medium	High
Excellent	0.055576923	1.265014	1.256622889
Good	1.734215686	1.652353	4.868871353
Poor	2.925217391	0.515958	4.384400848

CHI-SQUARE **18.65823**

To calculate the critical value, we use either the chi-square critical value table or the CHISQ formula. The formula "CHISQ.INV.RT" contains two parameters—the probability and the degrees of freedom.

The probability is 0.05, which is a significant value. The df is equal to 4.

=CHISQ.INV.RT(0.05,4)				
	H	I	J	K
2	Chi-Square Points= (Observed-Expected)^2/Expected			
3	Service	Salary		
4		Low	Medium	High
5	Excellent	0.055576923	1.265014	1.256622889
6	Good	1.734215686	1.652353	4.868871353
7	Poor	2.925217391	0.515958	4.384400848
8				
9	CHI-SQUARE	18.65823041		
10				
11	Critical Value of Chi-square =	=CHISQ.INV.RT(0.05,4)		
12				

The chi-square critical value is 9.487729037.

=CHISQ.INV.RT(0.05,4)				
	H	I	J	K
2	Chi-Square Points= (Observed-Expected)^2/Expected			
3	Service	Salary		
4		Low	Medium	High
5	Excellent	0.055576923	1.265014	1.256622889
6	Good	1.734215686	1.652353	4.868871353
7	Poor	2.925217391	0.515958	4.384400848
8				
9	CHI-SQUARE	18.65823041		
10				
11	Critical Value of Chi-square =	9.487729037		
12				

Let us find the chi-square p-value with the help of the following formula.

"=CHITEST (actual_range,expected_range)"

We apply the formula "=CHITEST(B4:D6,B14:D16)."

Observed Frequencies				
Service	Salary			Total
	Low	Medium	High	
Excellent	9	10	7	26
Good	11	9	31	51
Poor	12	8	3	23
Total	32	27	41	100
N		100		

Expected Frequencies(Variables Perfectly Independent)				
Service	Salary			Total
	Low	Medium	High	
Excellent	8.32	7.02	10.66	
Good	16.32	13.77	20.91	
Poor	7.36	6.21	9.43	

Chi-Test (P)Value = =CHITEST(B4:D6,B14:D16)

The chi-square p-value is= 0.00091723.

Observed Frequencies				
Service	Salary			Total
	Low	Medium	High	
Excellent	9	10	7	26
Good	11	9	31	51
Poor	12	8	3	23
Total	32	27	41	100
N		100		

Expected Frequencies(Variables Perfectly Independent)				
Service	Salary			Total
	Low	Medium	High	
Excellent	8.32	7.02	10.66	
Good	16.32	13.77	20.91	
Poor	7.36	6.21	9.43	

Chi-Test (P)Value = 0.00091723

The chi-square calculated value is significant when equal to or more than the chi-square critical value (tabulated value). The null hypothesis (H₀) is

rejected if the chi-square calculated value is greater than the chi-square critical value.

Here χ^2 (calculated) $>$ χ^2 (tabulated) or $18.65 > 9.48$. Hence, we reject the null hypothesis and accept the alternative hypothesis.

The p-value can also determine whether the null hypothesis must be accepted or rejected. For this, the p-value is compared with alpha (α) in the following way:

- If $p\text{-value} \leq \alpha$, the null hypothesis is rejected.
- If $p\text{-value} > \alpha$, the null hypothesis is accepted.

In this example, $p\text{-value} < \alpha$ or $0.0009172 < 0.05$. So, we reject H_0 and accept H_1 .

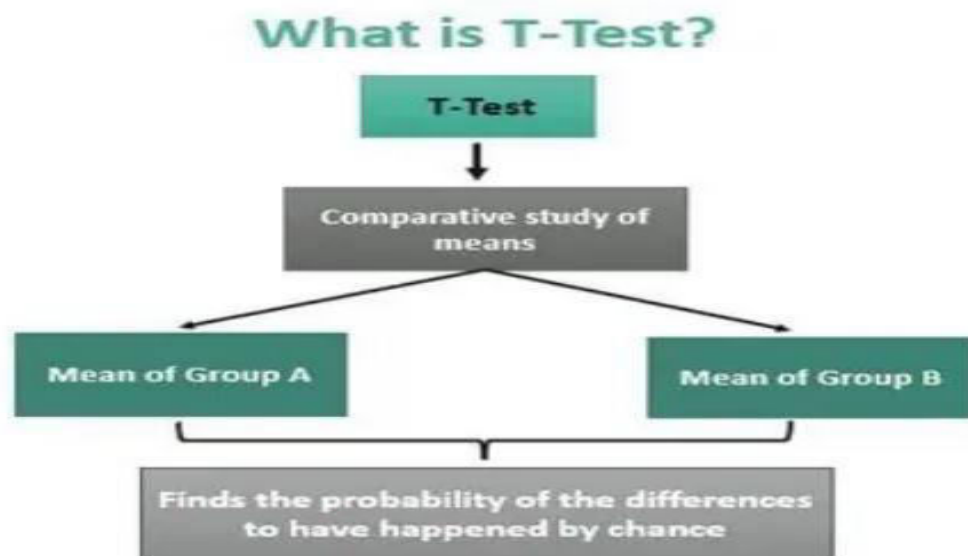
We conclude that the quality of service is dependent on the salary of customers waiting to be served.

T.TEST

A t-test is a statistical test that is used to compare the means of two groups. It is often used in hypothesis testing to determine whether a process or treatment actually has an effect on the population of interest, or whether two groups are different from one another.

T-Test Meaning

A T-test is the final statistical measure for determining differences between two means that may or may not be related. The testing uses randomly selected samples from the two categories or groups. It is a statistical method in which samples are chosen randomly, and there is no perfect normal distribution.



- A T-test is a statistical method of comparing the means or proportions of two samples gathered from either the same group or different categories.
- It is aimed at hypothesis testing, which is used to test a hypothesis pertaining to a given population.
- It is the difference between population means and a hypothesized value.
- One-sample, two-sample, paired, equal, and unequal variance are the types of T-tests users can use for mean comparisons.

T-Test Explained

A T-test studies a set of data gathered from two similar or different groups to determine the probability of the difference in the result than what is usually obtained. The accuracy of the test depends on various factors, including the distribution patterns used and the variants influencing the collected samples. Depending on the parameters, the test is conducted, and a T-value is obtained as the statistical inference of the probability of the usual resultant being driven by chance.

For example, if one wishes to figure out if the mean of the length of petals of a flower belonging to two different species is the same, a T-test can be done. The user can select petals randomly from two other species of that flower and come to a standard conclusion. The final **T-test interpretation** could be obtained in either of the two ways:

- A null hypothesis signifies that the difference between the means is zero and where both the means are shown as equal.
- An alternate hypothesis implies the difference between the means is different from zero. This hypothesis rejects the null hypothesis, indicating that the data set is quite accurate and not by chance.

This T-test, however, is only valid and should be done when the mean or average of only two categories or groups needs to be compared. As soon as the number of comparisons to be made is more than two, conducting this is not recommended.

Assumptions

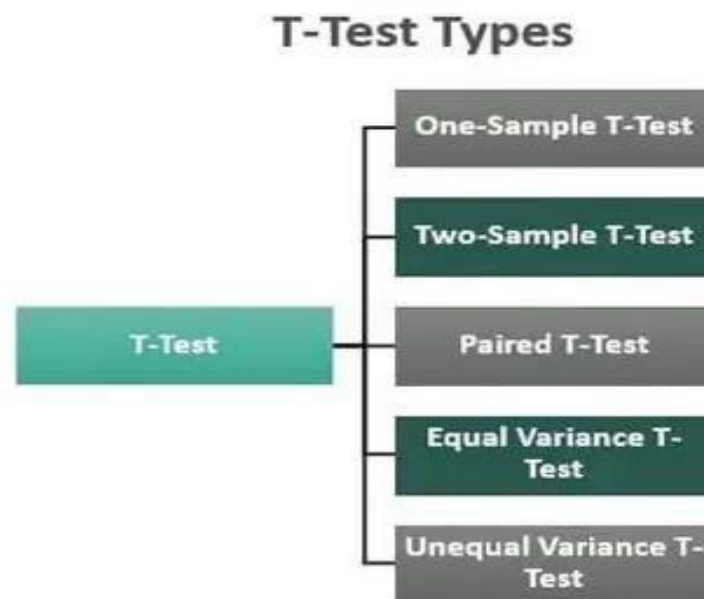
The test runs on a set of assumptions, which are as follows:

- The measurement scale used for such hypothesis testing follows a set of continuous or ordinal patterns. The accounted parameters and variants influencing the samples and surrounding the groups are based on the standard consideration.

- The tests are completely based on random sampling. As no individuality is maintained in the samples, the reliability is often questioned.
- When the data is plotted with respect to the T-test distribution, it should follow a normal distribution and bring about a bell-curved graph.
- For a clearer bell curve, the sample size needs to be bigger.
- The variance should be such that the standard deviations of the samples are almost equal.

Types

Some of the widely used **T-test types** are as follows:



1 – One-Sample T-Test

While performing this test, the mean or average of one group is compared against the set average, which is either the theoretical value or means of the population. For example, a teacher wishes to figure out the average height of the students of class 5 and compare the same against a set value of more than 45 kgs.

The teacher first randomly selects a group of students and records individual weights to achieve this. Next, she finds out the mean weight for that group and checks if it meets the standard set value of 45+. The formula used to obtain one-sample t-test results is:

$$t = \frac{m - \mu}{s/\sqrt{n}}$$

Where,

- T = t-statistic
- m = mean of the group
- = theoretical mean value of the population
- s = standard deviation of the group
- n = sample size

2 – Independent Two-Sample T-Test

This is the test conducted when samples from two different groups, species, or populations are studied and compared. It is also known as an independent T-test. For example, if a teacher wants to compare the height of male students and female students in class 5, she would use the independent two-sample test.

The **T-test formula** used to calculate this is:

$$t = \frac{m_A - m_B}{\frac{\sqrt{s^2}}{\sqrt{n_A}} + \frac{\sqrt{s^2}}{\sqrt{n_B}}}$$

Where,

- mA – mB = means of samples from two different groups or populations
- nA – nB = respective sample sizes
- s2 = standard deviation or common variance of two samples

3 – Paired Sample T-Test

This hypothesis testing is conducted when two groups belong to the same population or group. The groups are studied either at two different times or under two varied conditions. The formula used to obtain the t-value is:

$$t = \frac{m}{s/\sqrt{n}}$$

Where,

- T = t-statistic
- m = mean of the group
- = theoretical mean value of the population
- s = standard deviation of the group
- n = sample size

4 – Equal Variance T-Test

This test is conducted when the sample size in each group or population is the same or the variance of the two data sets is similar. It is also referred to as pooled T-test. The formula applied here is as follows:

$$T - value = \frac{mean1 - mean2}{(n1 - 1) \times var1^2 + (n2 - 1) \times var2^2 \sqrt{\frac{1}{n1} + \frac{1}{n2}}}$$

Where,

- Mean1 and mean2 = average value of each set of samples
- var1 and var2 = variance of each set of samples
- n1 and n2 = number of records in each set

5 – Unequal Variance T-Test

The unequal variance testing is used when the variance and the number of samples in each group are different. It is often referred to as Welch's test, and the formula is:

$$T - value = \frac{mean1 - mean2}{\sqrt{\left(\frac{var1}{n1} + \frac{var2}{n2}\right)}}$$

Where,

- mean1 and mean2 = Average value of each set of samples
- var1 and var2 = Variance of each set of samples
- n1 and n2 = number of records in each set

Example With Calculation

Let us consider the scores for each subject in the examination held in two phases.

Subject	Marks Phase 1	Marks Phase 2
<i>Maths</i>	45	62
<i>Physics</i>	45	55
<i>Chemistry</i>	45	55
<i>Biology</i>	50	65
<i>History</i>	55	68
<i>Geography</i>	80	70

Step 1: Subtract the marks scored in both the phases

	A	B	C	D	E
	Subject	Mark Phase 1	Mark Phase 2	Phase 1 - Phase 2	
1	Maths	45	62	-17	
2	Physics	45	55	-10	
3	Chemistry	45	55	-10	
4	Biology	50	65	-15	
5	History	55	68	-13	
6	Geography	80	70	10	
7	ADD			-55	

Step 2: Add up all the differences, i.e., -55

Step 3: Square up the differences

	A	B	C	D	E
1	Subject	Mark Phase 1	Mark Phase 2	Phase 1 - Phase 2	(Phase 1 - Phase 2)^2
2	Maths	45	62	-17	289
3	Physics	45	55	-10	100
4	Chemistry	45	55	-10	100
5	Biology	50	65	-15	225
6	History	55	68	-13	169
7	Geography	80	70	10	100
8	ADD			-55	983

Step 4: Add up all the squares of difference, i.e., 983

Step 5: Usage of formula to calculate the T value

$$T = \{(\sum D)/N\} / \sqrt{\{\sum D^2 - (\sum D)^2/N\}/(N-1)} - N$$

- = $-9.16 / \sqrt{\{983 - (-55)^2/6\}} / (6-1) * 6$
- = $-9.16 / \sqrt{15.96}$
- = $-9.16 / 3.99$

T Value = -2.29

Now, get the degrees of freedom. To obtain this, subtract 1 from the sample size (6 - 1 = 5). The next thing is to find out the p-value, which, if smaller in value, supports the null hypothesis result. For example, if the p-value is something around 0.9, i.e., 90%, it indicates that the T-value obtained has the probability of being a random observation. On the other hand, if the p-value is around 0.025, i.e., 2.5%, the result or t-value obtained is significant.

What is ANOVA in MS Excel?

ANOVA (Analysis of Variance) in Excel is the single and two-factor method used to perform the null hypothesis test, which says if the test will be PASSED for Null Hypothesis if all the population values are exactly equal to each other.

What is ANOVA?

While ANOVA has many varieties, the essential purpose of this family of analyses is to determine whether factors have an association with an outcome variable.

Factors are the variables that you will use to categorize your outcome variable into groups. For example, if you want to know whether tapes from three different suppliers have the same peel strength, the suppliers are your factor. All the strength measurements for the same supplier's tape form a group of measurements.

ANOVA is an inferential statistical analysis

Inferential analysis is the formal way of saying that we want to look at a sample of measurements and make an educated guess about what all of the possible measurements might be like if we could take them.

Let's return to the tape example. If you could tape 1 million boxes from a batch of tape, those million might represent the entire population that we want to know about. But if we taped those million boxes and measured the peel strength, we would have used up all of the tape. Instead, we'll measure the strength from a sample of taped boxes and use those measurements to guess what the numbers would look like if we taped a million boxes.

What does ANOVA do?

An important point is that we won't expect all the measurements in a group to be the same.

Consider the tape example again. The differences in strength measurements from the same supplier's tape give us within-group variation. Another important point is that we won't expect the average strength of our sample to be the same as the average strength if we taped a million boxes. This variation between the sample average and the overall average we'll call bias.

Because of within-group variation and bias, comparisons among groups become harder. We'll know that our sample average is not the same as the real average, there's no easy way to know when our guess is too high or too low. If we guess too high for one group and too low for another group, we might easily reach an incorrect conclusion, such as predicting that the supplier with the strongest tape on average has the weakest tape.

ANOVA gives us mathematical sets of rules that hold certain given assumptions, to decide when we can have confidence that the real average of one group is different from the real average of one or more other groups. ANOVA sets up these rules by asking how sure we are that the means are the same, a concept that we refer to as the null hypothesis. Remember that the null hypothesis is a useful concept for helping us make comparisons, even though we already know that for real group averages to all be the same would be a remarkable coincidence.

Most of the time, a key result of an ANOVA analysis is a p-value. The p-value has meaning only with respect to the null hypothesis of the ANOVA analysis. For one-way ANOVA, the null hypothesis is that the means for each level of your factor are the same.

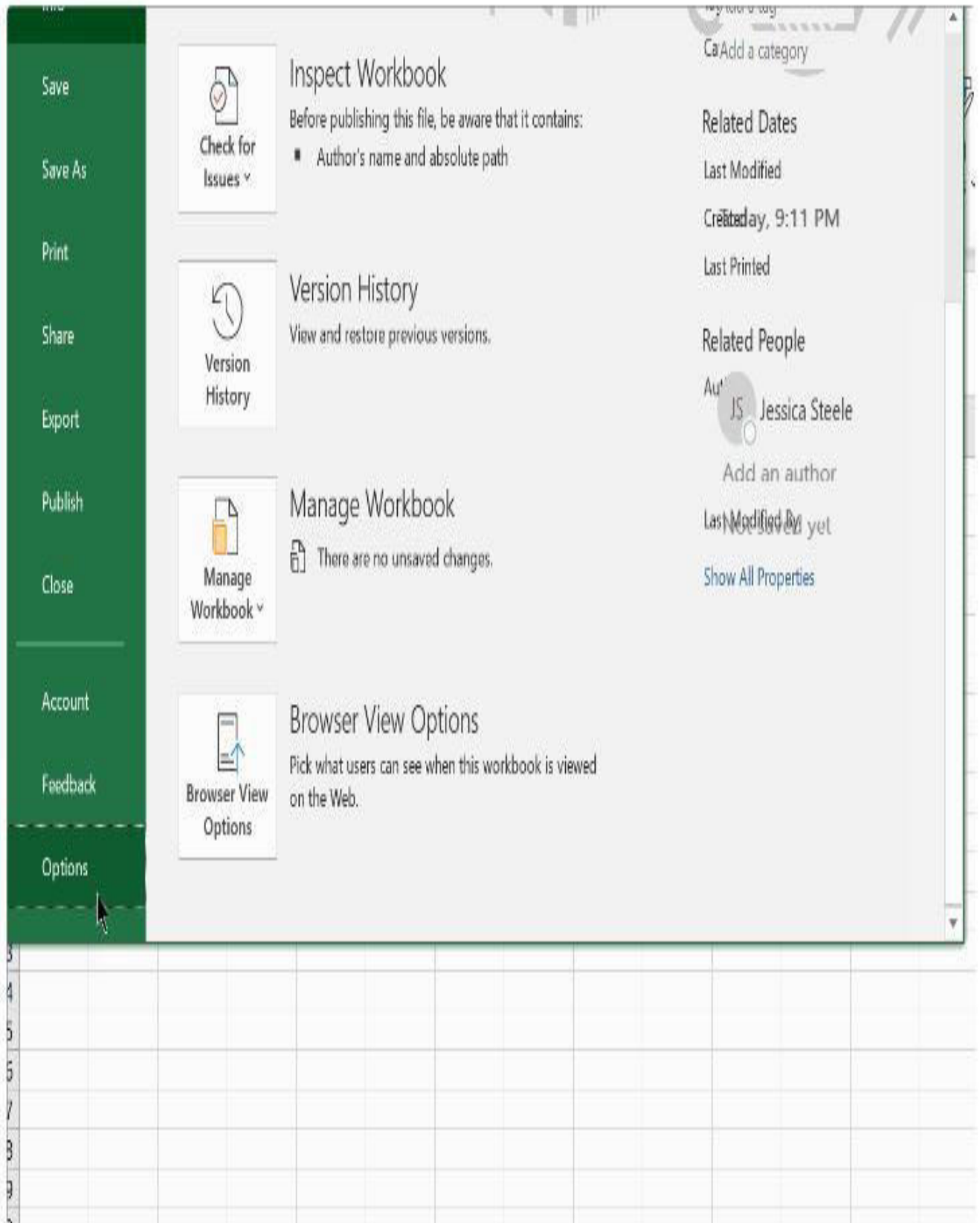
A rough interpretation would be that the p-value reflects how much confidence you can have that the null hypothesis is a reasonable model. Small p-values make you think that the null hypothesis is not a reasonable model. Large p-values might lead you to act like the null hypothesis is true, even though you know that it's not really true, just a reasonable model.

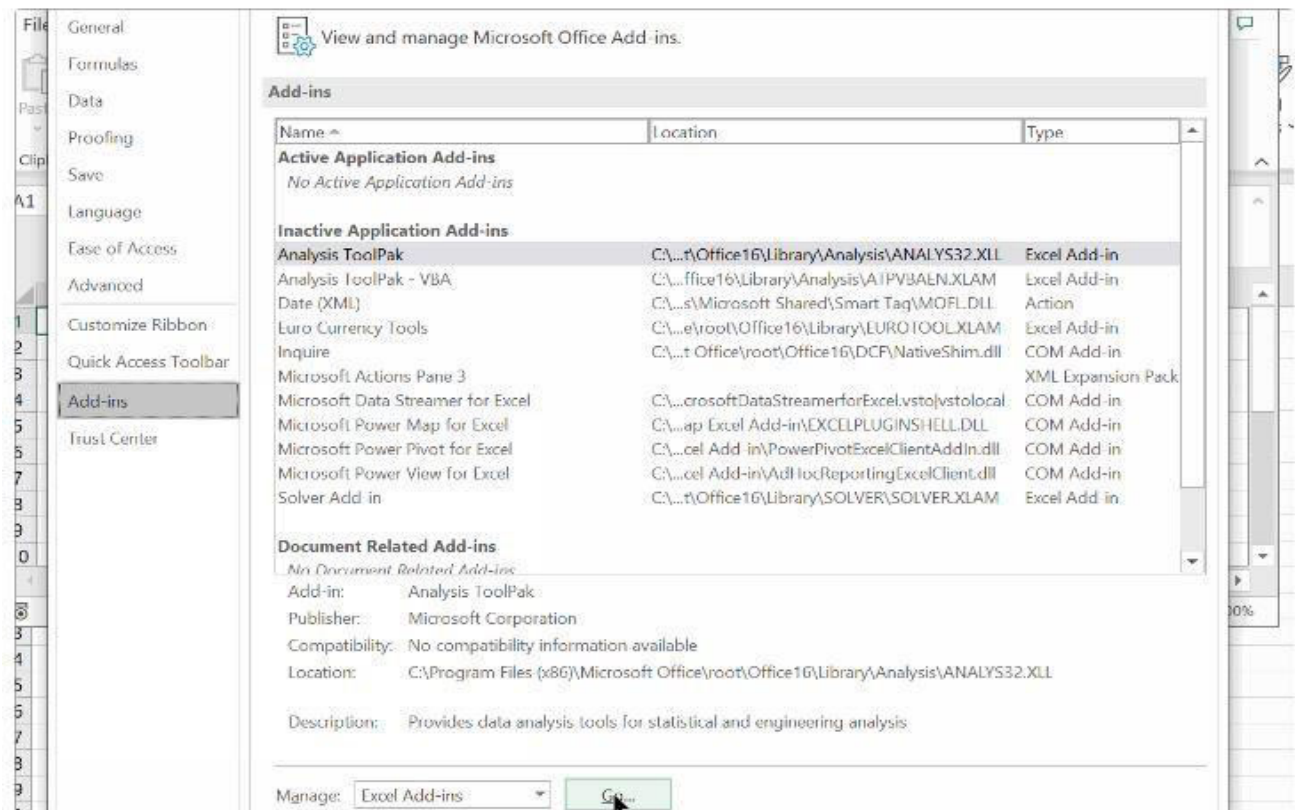
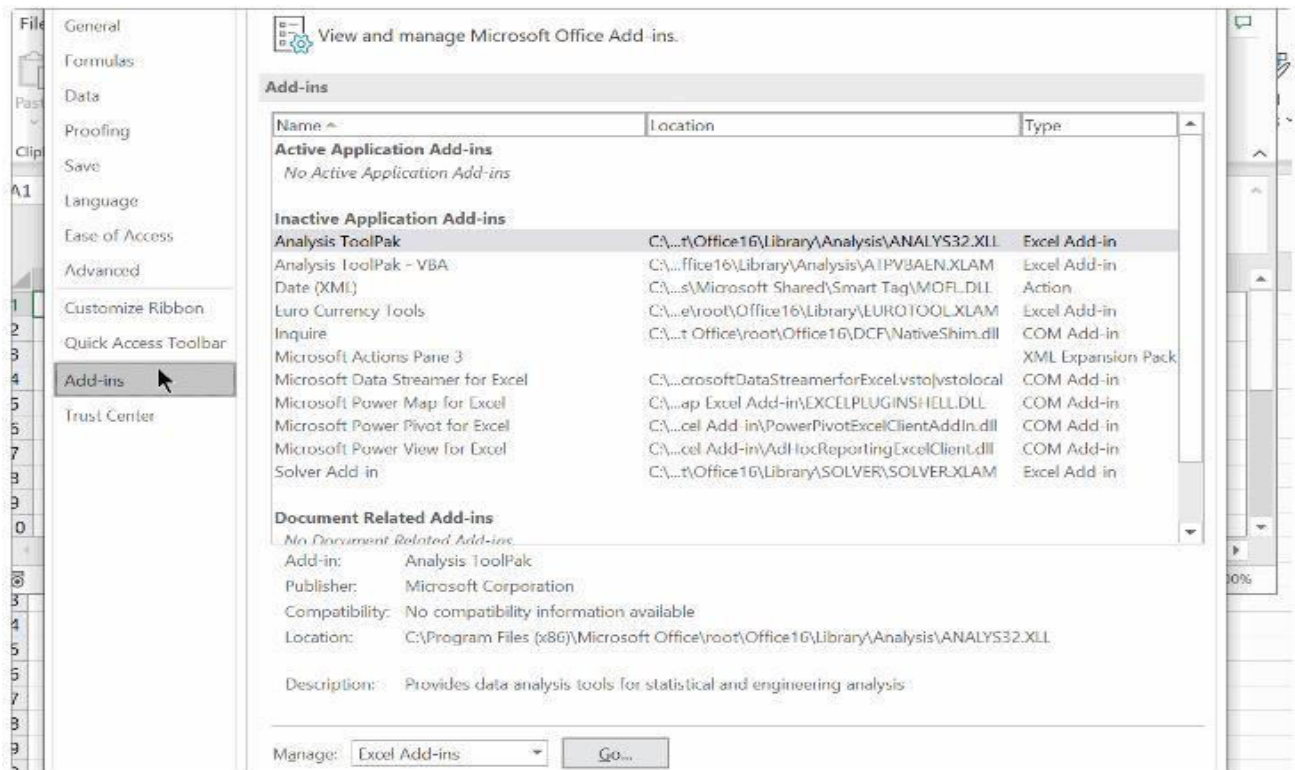
The Data Analysis Toolpak in Excel

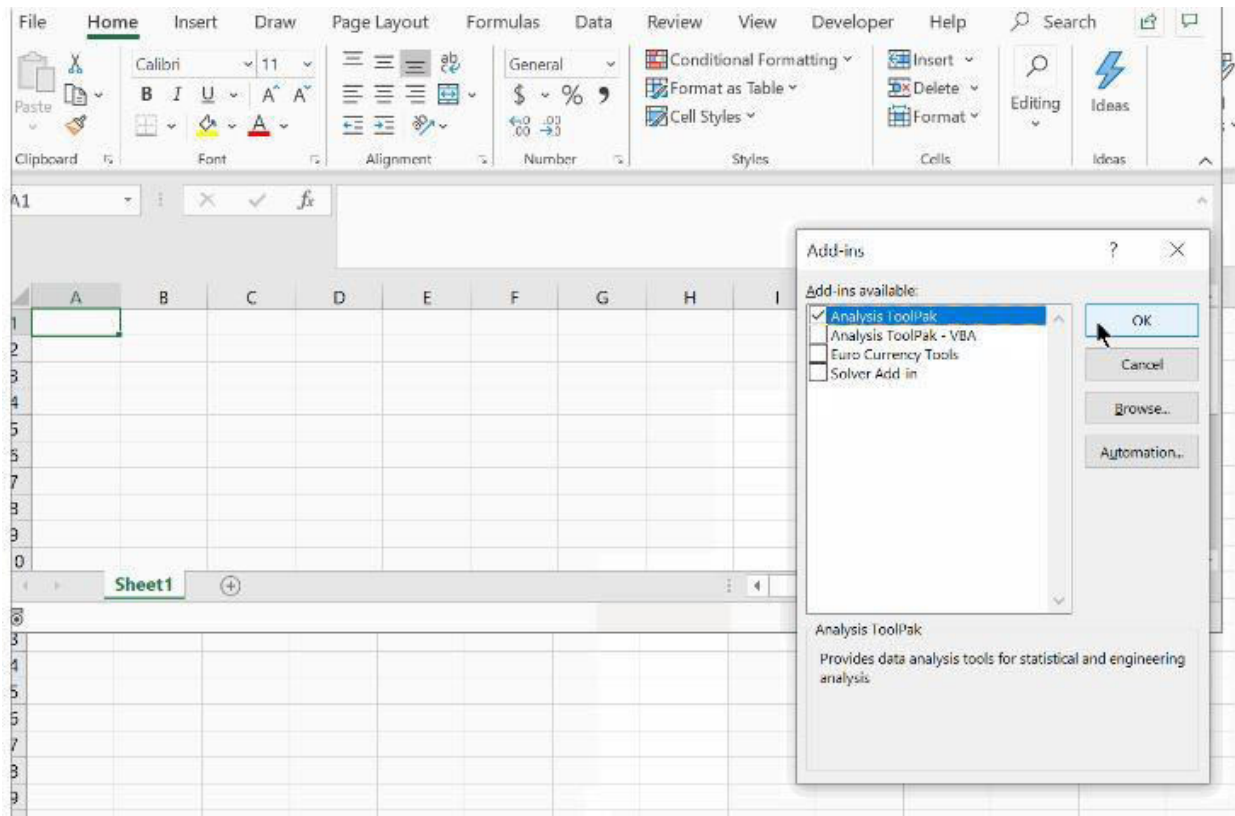
If you're analyzing data in Excel, then it's natural to make use of the tools that Microsoft provides for you. One of the less obvious features in Excel is the Data Analysis Toolpak. The Toolpak is an Excel add-in from Microsoft that's included with Excel, but isn't turned on.

Here's how to turn it on in the Microsoft Windows operating system.

1. Choose **File**, then **Options**
2. In the **Excel Options Window**, choose **Add-ins**
3. Next to **Manage**, select **Excel Add-ins** and click **Go**
4. In the **Add-ins** window, select **Analysis ToolPak** and click **OK**







A new button on your **Data** ribbon will appear.



Example of one-way ANOVA in Excel's Data Analysis Toolpak

While it can sometimes seem like a simple analysis will have the fewest applications, it's easy to find practical ways to use one-way ANOVA. Essentially, you can use it anytime you have only one set of groups to compare.

Let's keep going with our tape example. You've invested in an automatic taping machine that applies heat to tape to create strong bonds. You've decided that you're going to measure the strengths of tape samples from different suppliers yourself so that you can see whether there's any practical difference in the strengths of the bonds using your machine and your boxes.

Data arrangement for one-way ANOVA in Excel

If you've been using Excel for a long time, you've gotten used to the idea that the spreadsheet is cell-based. That is, there's very little difference between putting numbers in the spreadsheet in rows or in columns.

Data in rows:

	A	B	C	D	E	F	G	H	I	J
1	Tape supplier 1	9.5	10.5	9.6	10.1	9.4	9.6	8.9	10.0	9.3
2	Tape supplier 2	9.8	9.8	9.6	9.1	9.5	9.2	9.4	9.9	10.0
3	Tape supplier 3	10.0	9.6	9.9	9.9	9.1	9.6	10.1	10.3	10.0

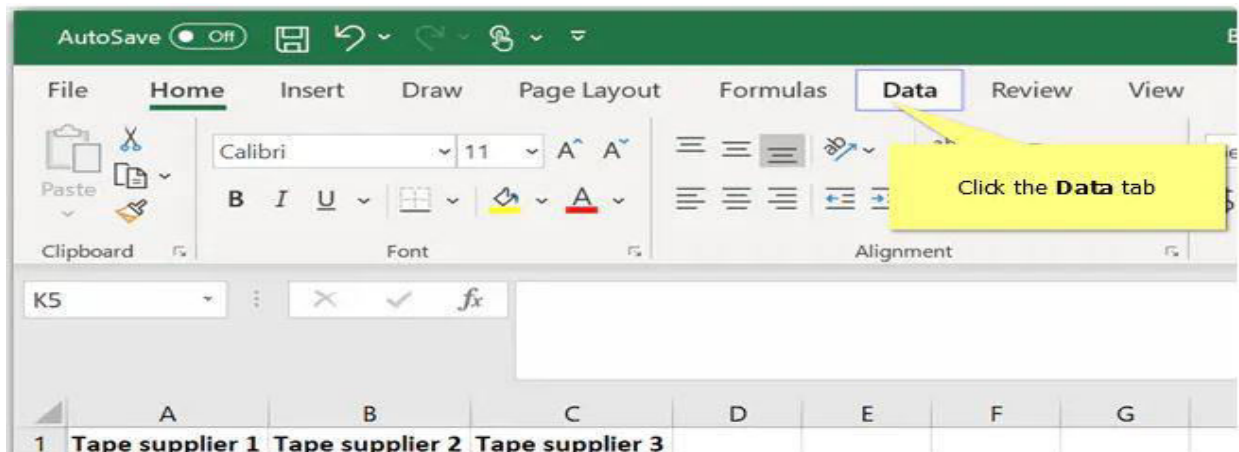
Data in column

	A	B	C
1	Tape supplier 1	Tape supplier 2	Tape supplier 3
2	9.5	9.8	10.0
3	10.5	9.8	9.6
4	9.6	9.6	9.9
5	10.1	9.1	9.9
6	9.4	9.5	9.1
7	9.6	9.2	9.6
8	8.9	9.4	10.1
9	10.0	9.9	10.3

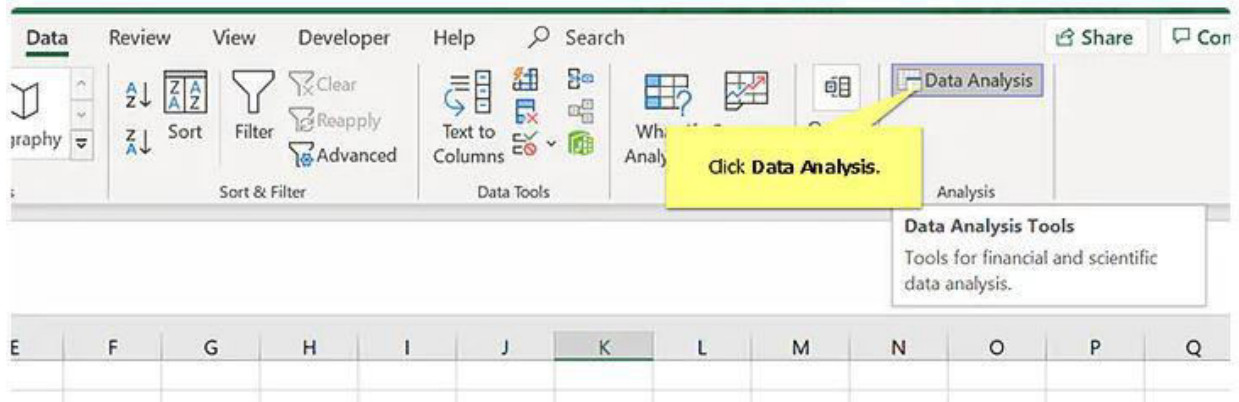
How to use one-way ANOVA in Excel

With the Data Analysis Toolpak installed and your data in columns, you can perform the following steps in Excel to get the results of the one-way ANOVA analysis.

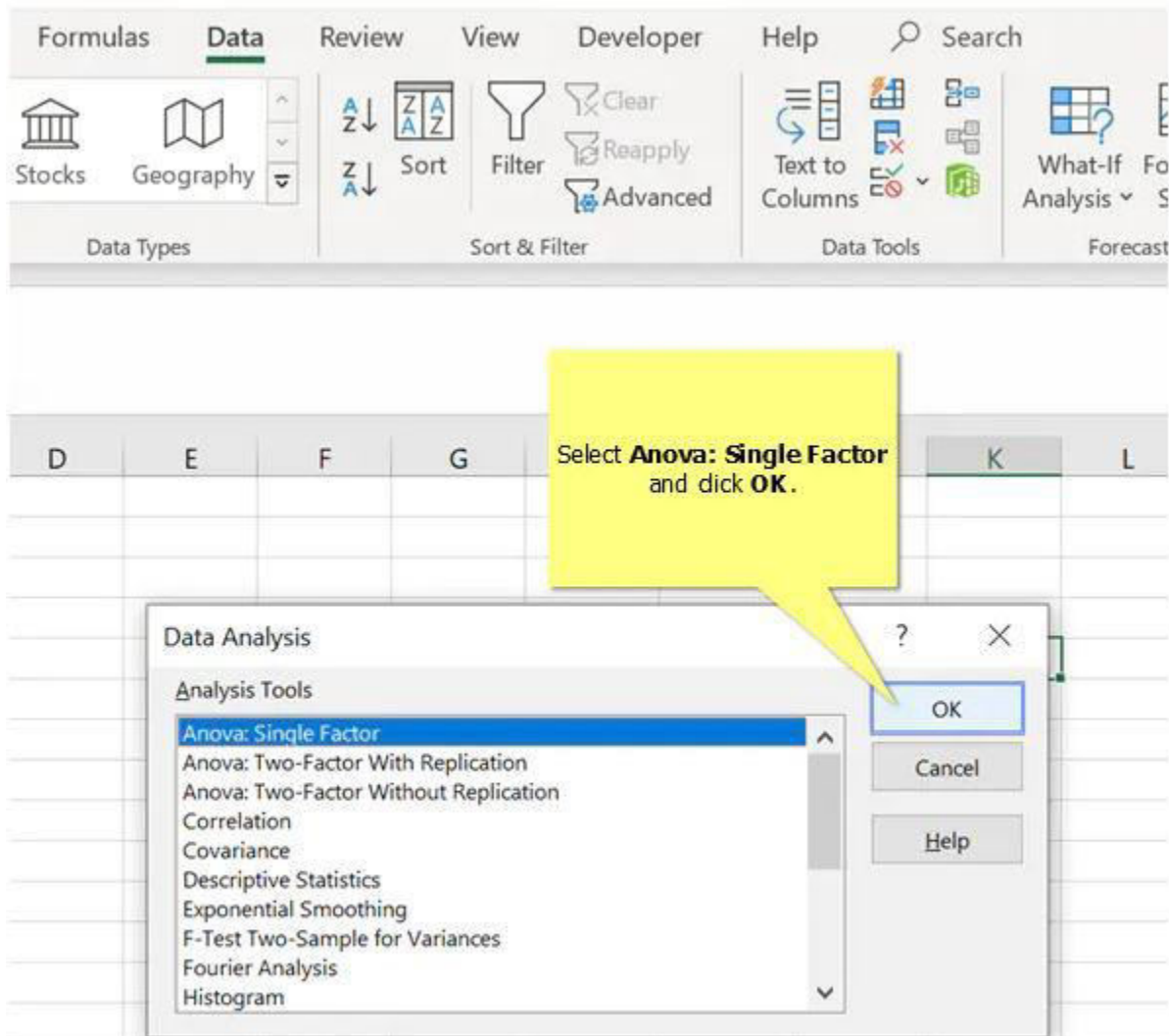
1. Click the **Data** tab



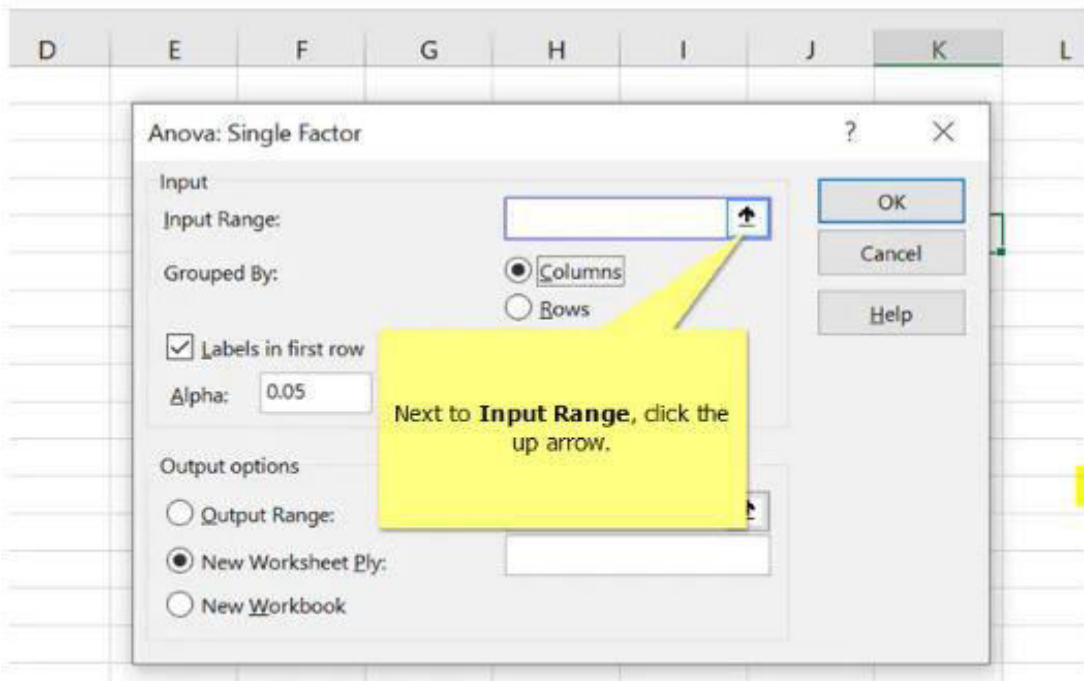
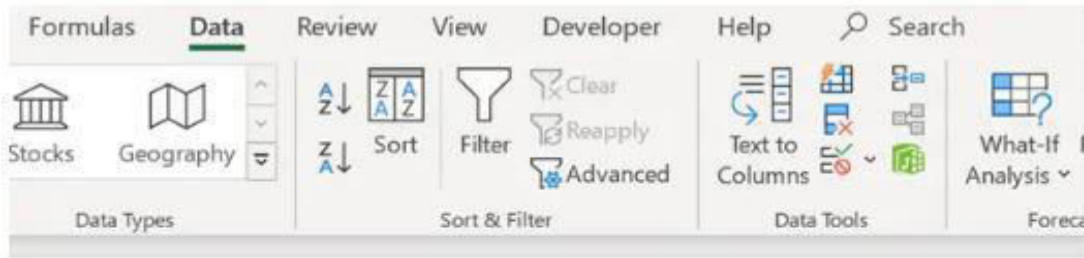
2. Click Data Analysis



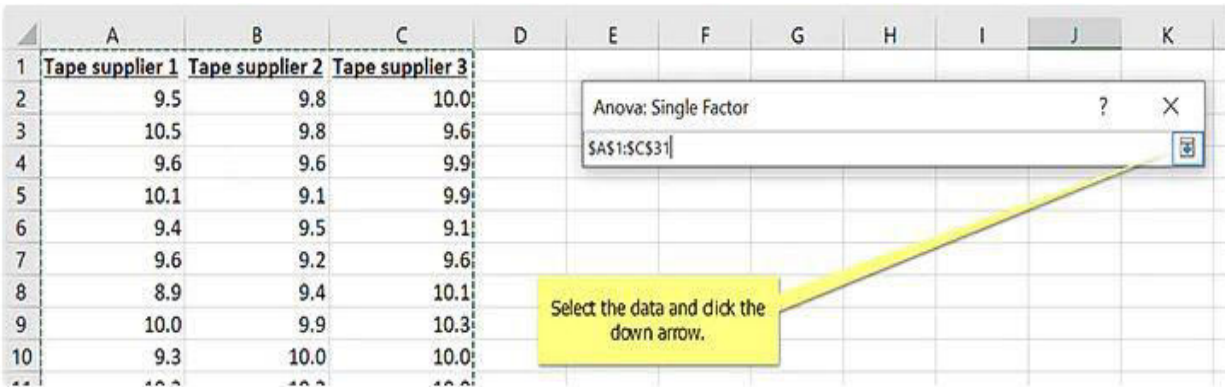
3. Select Anova: Single Factor and click OK



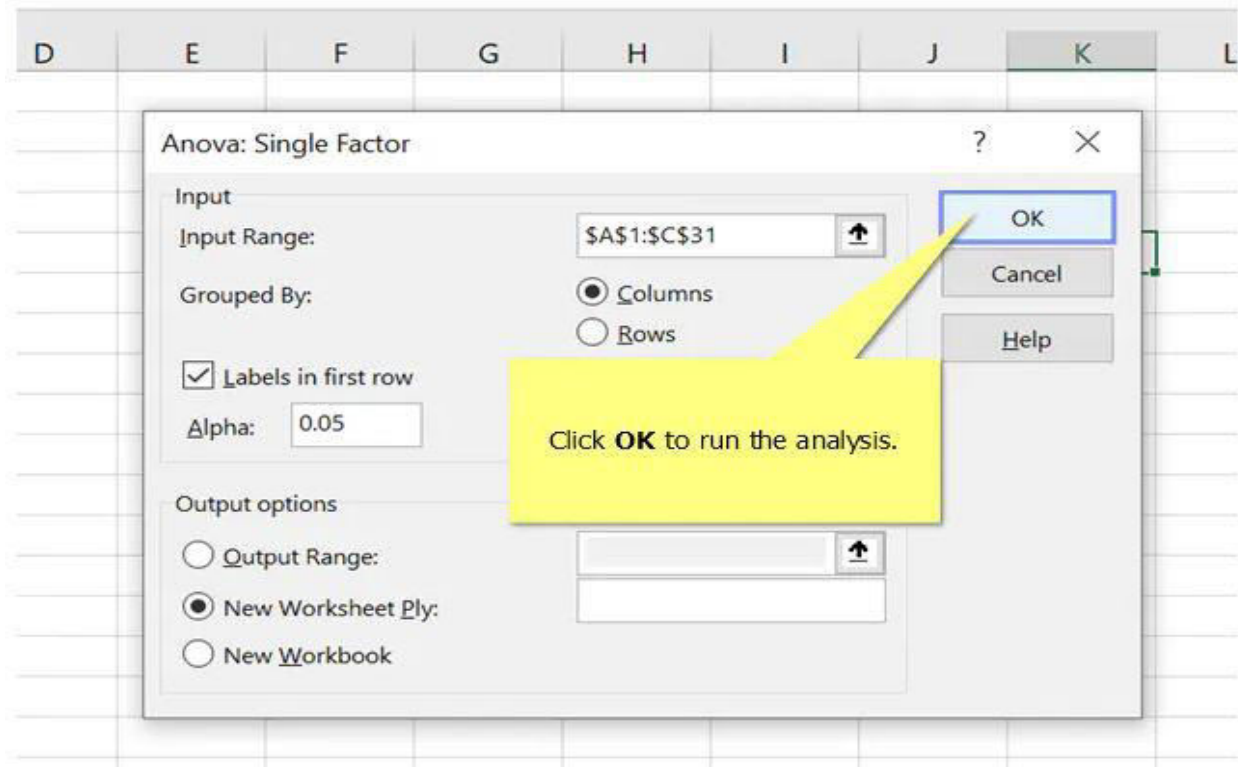
4. Next to Input Range click the up arrow



5. Select the data and click the down arrow



6. Click OK to run the analysis



Results for one-way ANOVA in Excel: Summary statistics

The results will look like this

	A	B	C	D	E	F	G
1	Anova: Single Factor						
2							
3	SUMMARY						
4	<i>Groups</i>	<i>Count</i>	<i>Sum</i>	<i>Average</i>	<i>Variance</i>		
5	Tape supplier 1	30	293.1303	9.771009	0.265686		
6	Tape supplier 2	30	290.2275	9.674249	0.125547		
7	Tape supplier 3	30	295.4848	9.849493	0.126999		
8							
9							
10	ANOVA						
11	<i>Source of Variation</i>	<i>SS</i>	<i>df</i>	<i>MS</i>	<i>F</i>	<i>P-value</i>	<i>F crit</i>
12	Between Groups	0.462329	2	0.231165	1.33819	0.267665	3.101296
13	Within Groups	15.02874	87	0.172744			
14							
15	Total	15.49107	89				
16							

First, let's take a minute to look at the summary statistics of each group. In particular, the averages, in ascending order, are about 9.67, 9.77, and 9.84. That is, each of the tapes holds almost 10 kg before breaking. The difference between the largest mean and the smallest mean is about 0.17 kg. If kilograms aren't very familiar to you, you can think of the tape with the lowest average being strong enough to hold about 60 apples and the tape with the highest average being strong enough to hold about 62 apples.

That should be enough for us to start to think about what we expect about the null hypothesis for the ANOVA. If you think that the means are similar, then you'll expect to see a larger p-value for the hypothesis test.

Results for one-way ANOVA in Excel: Hypothesis tests

Remember that small p-values give us low confidence in the null hypothesis. The value of about 0.27 is higher than the level where people traditionally agree that there is strong evidence against the null hypothesis. While most people learn 0.05 as a traditional cutoff, that value is mutable depending on the consequences of making an error either by deciding to act as if the means are the same or by acting like the means are not all the same.

Even so, 0.27 is such a large p-value that a lot of uncertainty remains about whether any of the averages are different. By extension, there's a lot of uncertainty about whether any one average is larger than another.

If those 2 apples worth of strength are so much that you would make a different decision about the tape suppliers because of that difference, then you'll need more data.

On the other hand, if those 2 apples don't sound like a big deal, this is a good place to decide that you can choose the supplier with other criteria. For example, you might consider price or your confidence that the supplier can fill your orders on time.

What if you have more factors?

Let's suppose that you're considering not only the tape supplier, but also choosing among some different boxes.

You know that the roughness and absorbency of the box might affect how strong the tape holds to it. Instead of doing the test only on the factor of tape supplier, you want to make sure that you have the right tape for the right box.

One approach could be to do a one-way ANOVA where you use more than one factor to define the groups. For example, one of the groups might be the first

tape supplier on the first box type. Another group might be the second tape supplier on a second box type.

The disadvantage of this approach is that it doesn't let you distinguish the effect of different factors. If the one-way ANOVA said that there was a difference between those two groups, then you still wouldn't know how much of the difference was from the change in tape, the change in box, or a change that depended on both simultaneously.

An analysis to get this type of information when you have two factors is two-way ANOVA.

Data Arrangement for Two-Way ANOVA in Excel

Excel can be flexible with your data arrangement for one-way ANOVA, but is strict about the data arrangement when you do a two-way ANOVA with replication through the Data Analysis Toolpak. Data for one factor need to be in different columns.

Data for the second factor need to be in consecutive rows.

For Excel to work, you'll need to have the same number of measurements for all of your groups.

You don't necessarily have to provide the factor label for the rows, but it's good practice, especially if you might want to graph your data in Excel later.

This data arrangement, called a two-way table, would look like this:

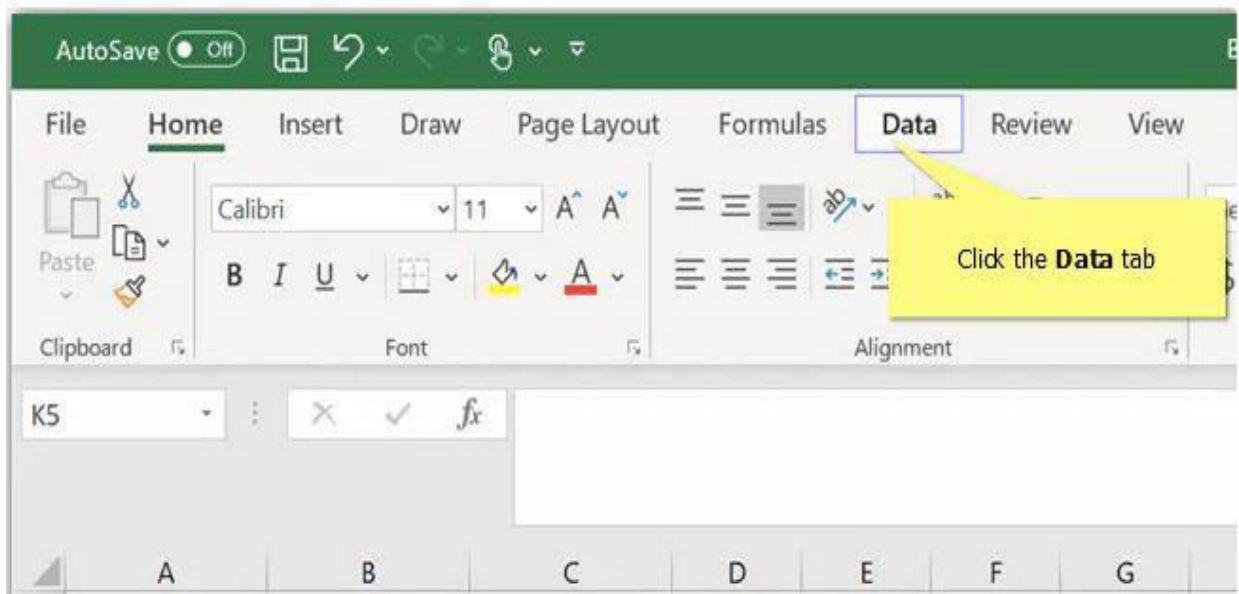
	A	B	C	D
1	Box Type	Tape supplier 1	Tape supplier 2	Tape supplier 3
2	Type 1	9.5	9.8	10.0
3	Type 1	10.5	9.8	9.6
4	Type 1	9.6	9.6	9.9
5	Type 1	10.1	9.1	9.9
6	Type 1	9.4	9.5	9.1
7	Type 1	9.6	9.2	9.6
8	Type 1	8.9	9.4	10.1
9	Type 1	10.0	9.9	10.3
10	Type 1	9.3	10.0	10.0
11	Type 1	10.2	10.2	10.0
12	Type 1	9.3	9.2	9.7
13	Type 1	9.6	9.4	10.0
14	Type 1	11.6	9.4	10.7
15	Type 1	9.7	10.0	9.9
16	Type 2	8.5	8.8	10.3
17	Type 2	9.0	8.6	9.5
18	Type 2	8.6	8.8	9.8
19	Type 2	8.9	8.8	10.1
20	Type 2	8.3	8.5	9.5
21	Type 2	9.0	8.6	9.6
22	Type 2	9.1	9.5	9.7
23	Type 2	8.8	8.8	10.2
24	Type 2	8.2	9.0	9.7
25	Type 2	8.4	7.9	9.3
26	Type 2	8.3	8.2	9.9
27	Type 2	9.4	9.1	9.4
28	Type 2	8.6	8.9	10.3
29	Type 2	8.6	8.9	9.4
30	Type 2	9.2	8.7	10.0
31	Type 2	9.0	8.3	9.8

How to use two-way ANOVA in Excel

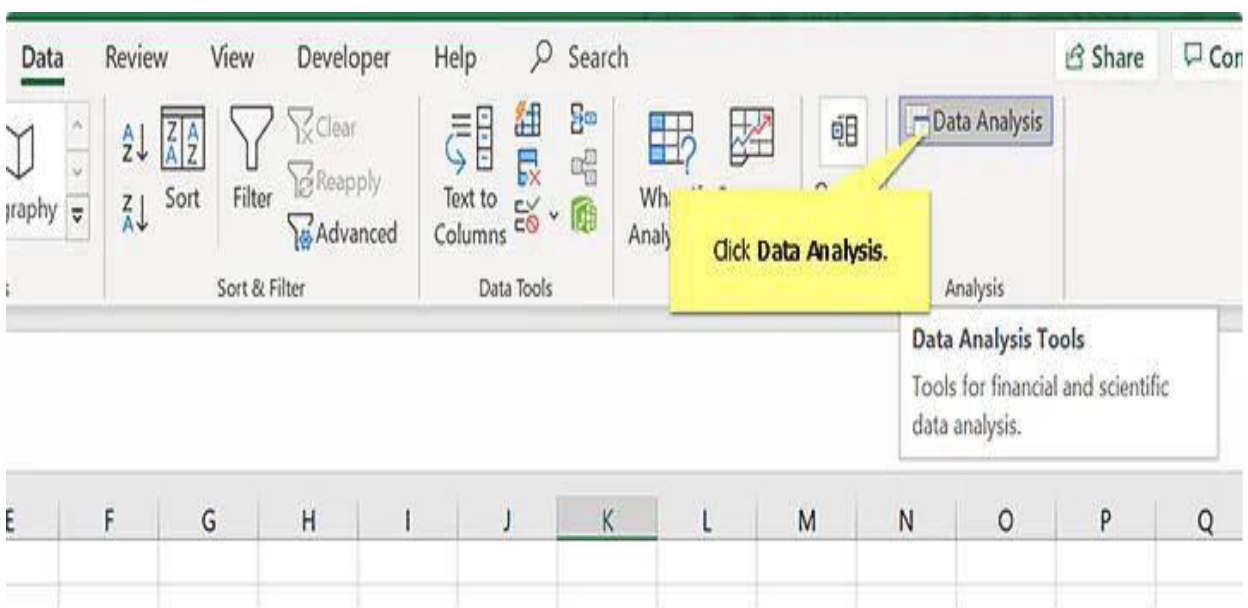
With the Data Analysis Toolpak installed and your data in columns, you can perform the following steps in Excel to get the results of the two-way ANOVA analysis. You'll begin as you did for one-way ANOVA.

Follow along with the two-way ANOVA steps

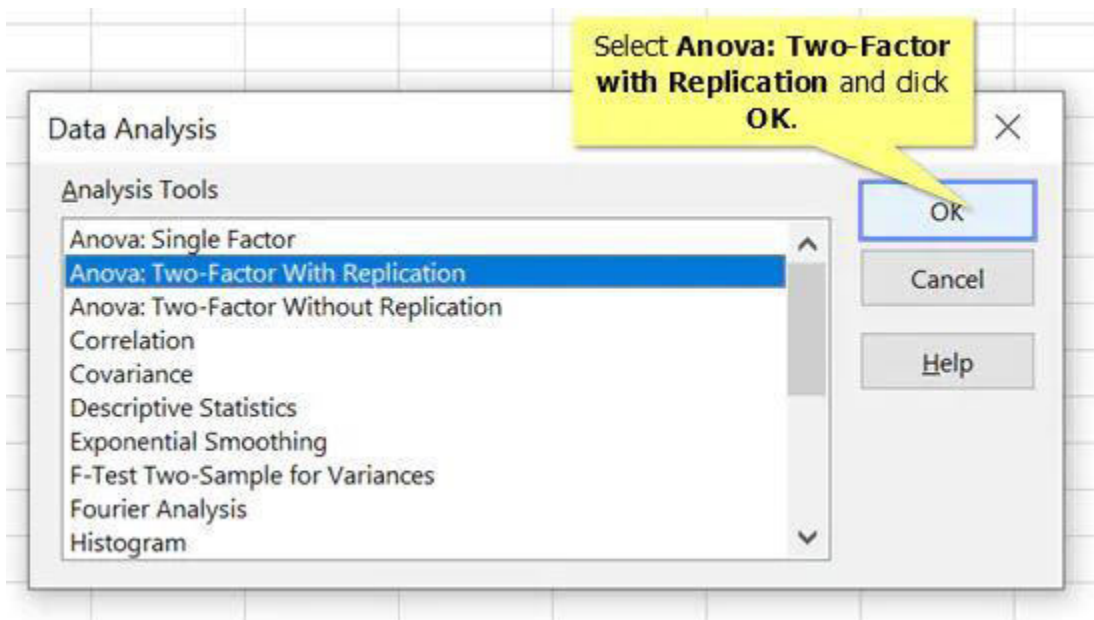
1. Click the Data tab



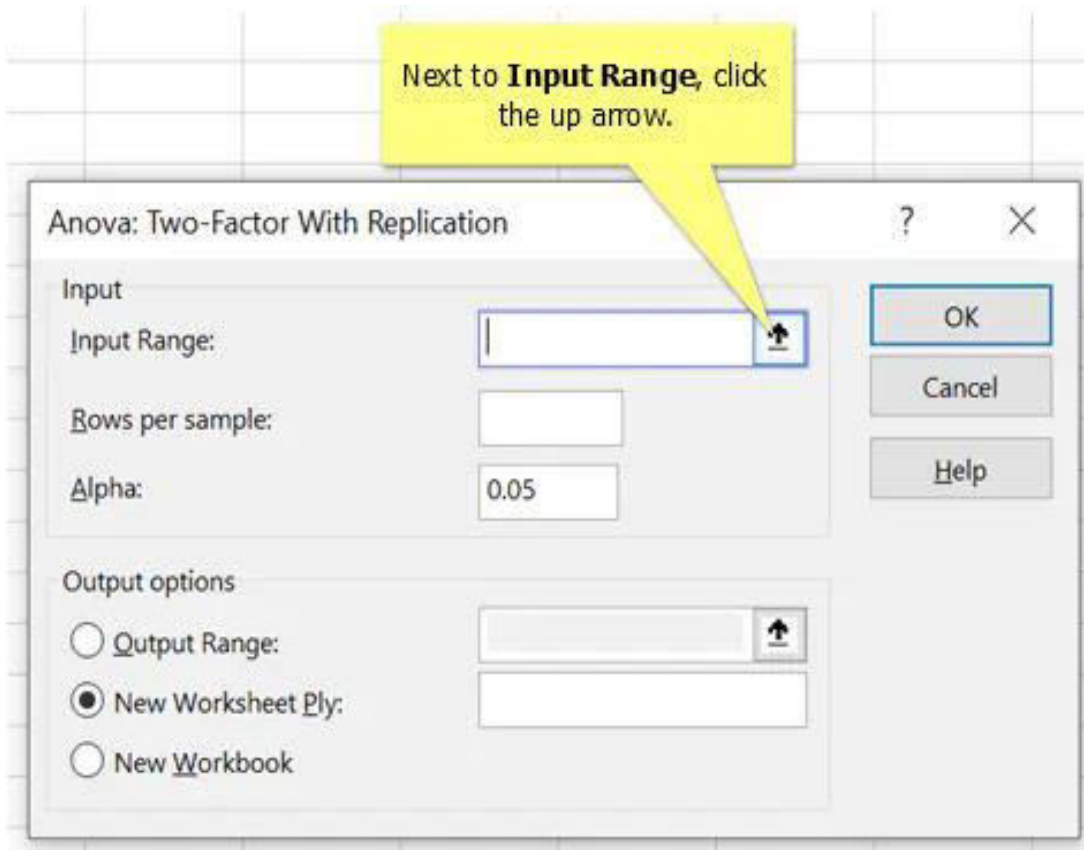
2. Click Data Analysis



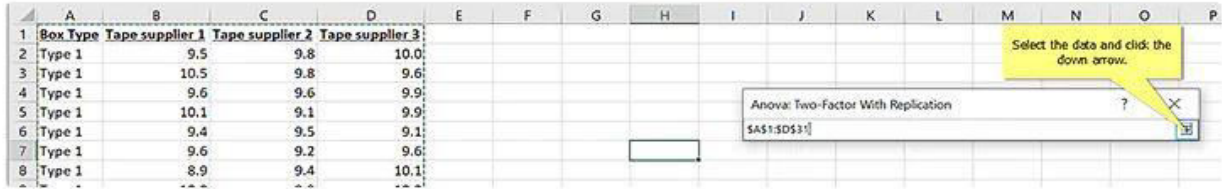
3. Select Anova: Two Factor with Replication and click OK



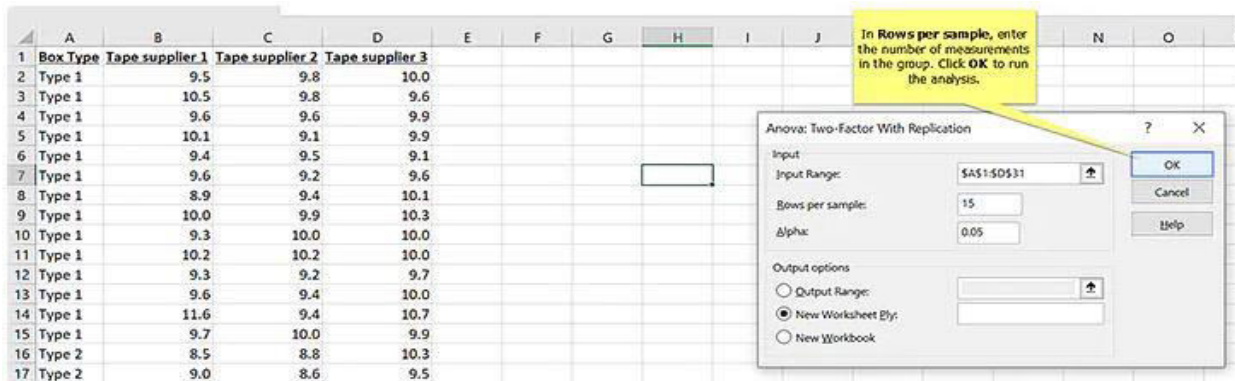
4. Next to Input Range, click the up arrow



5. Select the data and click the down arrow



6. In Rows per sample, enter the number of measurements in the group, then click OK to run



In this data, you can see that rows 2 to 15 have the measurements for the first box type. Those rows have 15 data points. Since the groups all have to have the same amount of data for the analysis to work in Excel, we know that the second box type must also have 15 rows.

Results for two-way ANOVA in Excel: Summary statistics

As with one-way ANOVA, your results will come in two parts. The first part will be summary statistics about your groups. I've added the highlighting.

	A	B	C	D	E
1	Anova: Two-Factor With Replication				
2					
3	SUMMARY	Tape supplier 1	Tape supplier 2	Tape supplier 3	Total
4	<i>Type 1</i>				
5	Count	15	15	15	45
6	Sum	145.7547785	143.6043962	149.1854396	438.5446
7	Average	9.716985235	9.573626414	9.945695974	9.745436
8	Variance	0.533654834	0.144809627	0.144865911	0.28598
9					
10	<i>Type 2</i>				
11	Count	15	15	15	45
12	Sum	131.3754857	130.6230684	146.2993576	408.2979
13	Average	8.758365711	8.708204557	9.753290507	9.073287
14	Variance	0.129259258	0.150741122	0.098371318	0.357277
15					
16	<i>Total</i>				
17	Count	30	30	30	
18	Sum	277.1302642	274.2274646	295.4847972	
19	Average	9.237675473	9.140915486	9.84949324	
20	Variance	0.557687335	0.336374928	0.126998972	

The blue highlighting shows the overall averages for the two different box types in the data. The difference is about 0.67 kilograms. The gray highlighting shows the averages for the 3 different tape suppliers. The averages for tape supplier 3 is closest to 10, while the averages for tape suppliers 1 and 2 are closer to 9.

The averages for the individual groups have gold highlighting. If the tapes from the different suppliers all work the same on both types of boxes, then the averages for the individual groups should follow the same patterns: The average for box type 1 should be higher and the average for tape supplier 3 should be higher.

The group averages show a different pattern than the overall averages for the two factors. Tape supplier 3's average is higher than the other two because there is a larger difference between the suppliers for the second box type.

This comparison of the averages should prepare us for what to expect about the null hypothesis for two-way ANOVA that the factors do not affect the response variable.

Results for two-way ANOVA in Excel: Hypothesis tests

For our one-way ANOVA analysis, the p-value was relatively large. That value led us to conclude that we couldn't be certain whether there was any difference between the tape suppliers.

For the two-way ANOVA, our largest p-value is about 0.002. That is much smaller than the traditional cutoff value for statistical significance of 0.05.

ANOVA						
Source of Variation	SS	df	MS	F	P-value	F crit
Sample	10.16514471	1	10.16514471	50.75373	3.32E-10	3.954568
Columns	8.857659099	2	4.428829549	22.11278	1.93E-08	3.105157
Interaction	2.621802115	2	1.310901057	6.545222	0.002282	3.105157
Within	16.82382898	84	0.200283678			
Total	38.46843491	89				

Because the p-value for the interaction is small, we cannot make a simple statement that one supplier or box type leads to a higher peel strength.

The hypothesis test confirms what we might have expected from the examination of the averages: The effect of the different tapes depends on the box type. (We could equivalently say that the effect of the different box types depends on the tape.)

From the default results in Excel, you can conclude that not all of the groups have the same peel strength. To make a more precise statement about the

relationships among the groups, you should proceed to a multiple comparisons analysis.

While the examination of the averages for the two-way ANOVA analysis suggests that the choice of tape matters only if you're going to use the second box type, you'll want to consider your decisions carefully from both a statistical and a practical perspective.

If you need the tape to have peel strength of only 5 kilograms, then the peel strengths are probably all adequate. If a difference of 0.4 kilograms might lead you to choose one group over another, then more analysis of the data is in order.

You can learn more techniques for analyzing ANOVA data in this course on hypothesis testing. If you've already acquired the basics, then you're ready to proceed to more advanced considerations in this design of experiments course. Either way, the knowledge that you gain will help you prepare to ensure that your projects exceed your expectations.

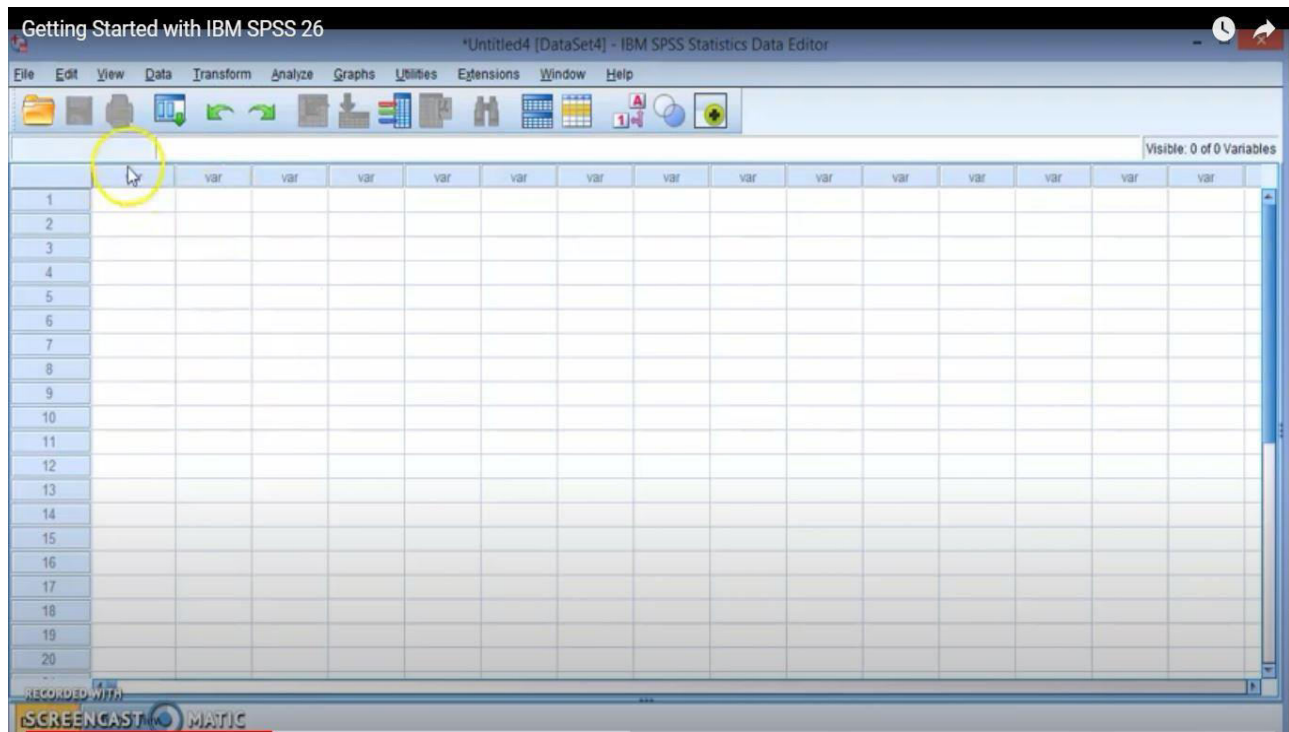
SPSS: An Introduction

Statistical Package for Social Sciences (SPSS) is one of the most popular software packages to perform statistical analysis on survey data. Its first version was released in 1968 and since then, it has come a long way. It is used by researchers in educational institutions, research organizations, government, marketing firms etc. SPSS can be used for:

1. Processing Questionnaires
2. Reporting in the form of tables and graphs
3. Analyzing-Means, Chi-Square, Regression and much more

SPSS comes into picture after data has been collected from the primary sources of data. Questionnaire is one of the popular method among all the methods of primary data collection.

- **Getting Started:** When SPSS is initially installed, the SPSS program is created in the Programs menu. To start with SPSS, click on the IBM SPSS Statistics 20 icon. A dialog box will be opened in front of SPSS grid listing several options to choose from.



The following options will appear in the dialog box:

- Run the tutorial
- Type in data
- Run an existing enquiry
- Create new query using Database Wizard

Select the option **Type in data** and click on OK. By default, the Data Editor view is initially selected.

- **SPSS Data Editor:** SPSS has 2 windows:

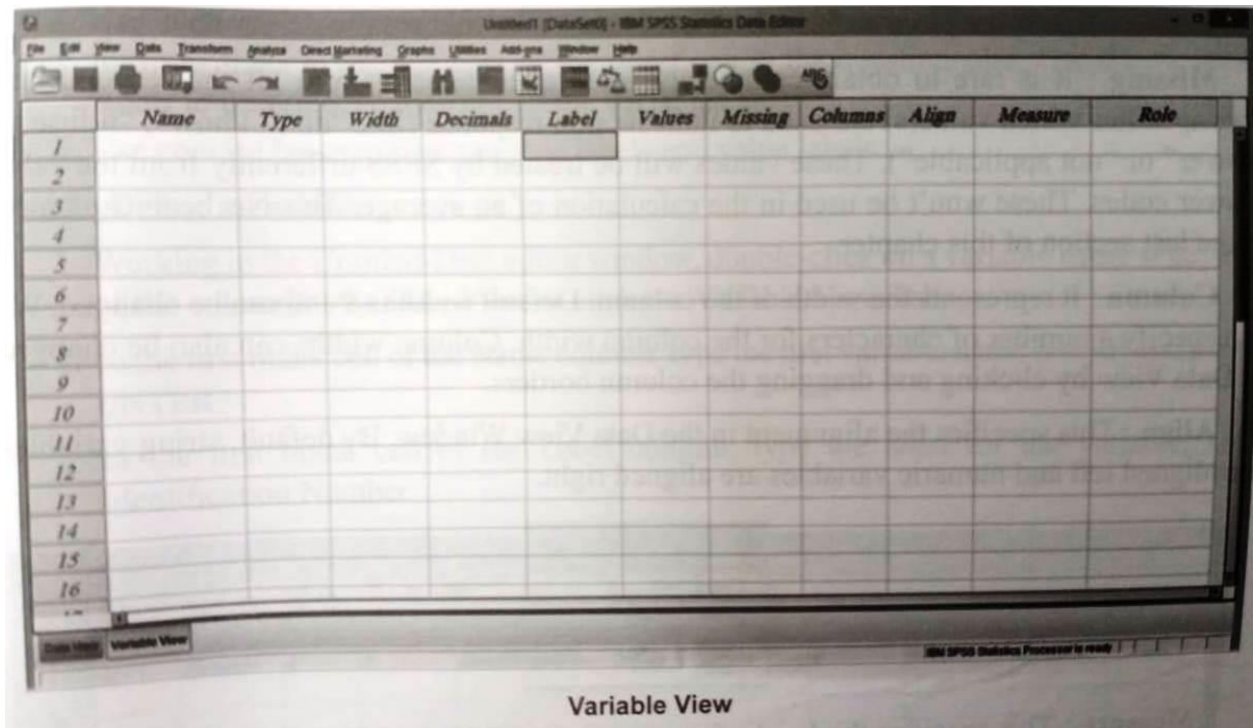
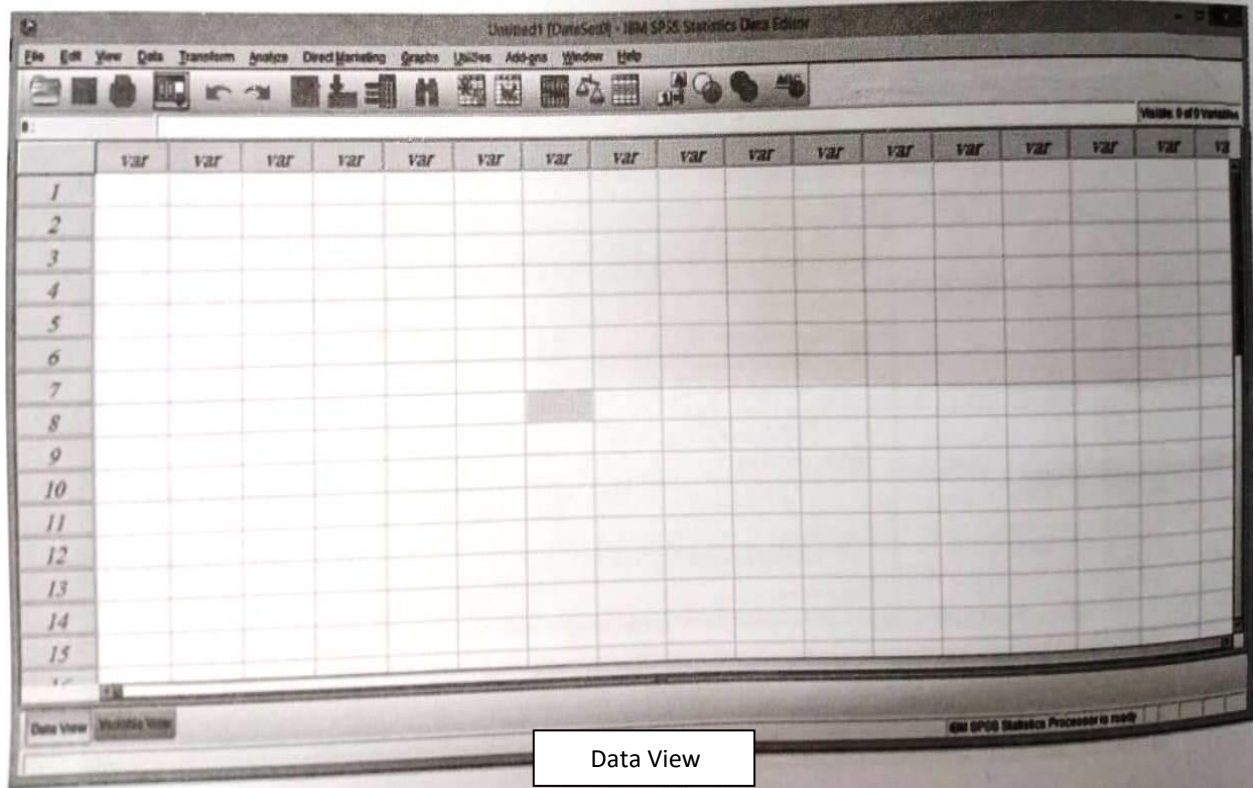
1. Data Editor

2. Viewer or draft viewer which displays the output files

The data editor is a versatile spread sheet-like system for defining, entering, editing and displaying data. The data editor is like a spread sheet in which cases are represented in rows and columns. Data Editor has two views:

Data View: It contains the actual data.

Variable View: It is used to define variables that will store the data, such as variable names, type, values, labels, formats and missing value indicators.



DEFINING VARIABLES

The first step is to open the "Variable View" window of the data editor and define variables. Suppose, we have developed a questionnaire that asks questions relating to an individual's confidence in learning statistics. The variables we have measured include: gender, age and agreement level on some statements. We have assigned an identification number to each case/respondent.

Name: Each variable can be named. No space or special characters are allowed. Just keep it simple to one word name or use under score (_) between the words. e.g. Marital status is not allowed, write it either Maritalstatus or Marital_status.

Type: By default, SPSS assumes that all new variables are numeric with two decimal places. However, it is possible to select other variable types (such as date, currency, string etc.). Data made of words is called string data.

Width: If we have a string variable type (i.e., data is entered in text instead of numerical value), we may adjust the number of text characters by changing the width value. By default, SPSS sets the length of string variables to eight characters.

Decimals: By default, SPSS predetermines two decimal places for all numeric variables. However, it is possible to adjust the decimal places according to our preference.

Label: Here, we will clearly describe the meaning of the variable. Try to be concise. SPSS uses this label to refer to the variable in all its output.

Values: This is the place to describe the answer codes that we use.

Missing: It is rare to obtain completely filled questionnaire from all the respondents. This specifies which value(s) are used to indicate non response (like "don't know", "refuse to answer" or "not applicable"). These values will be treated by SPSS differently from the valid answer codes. These won't be used in the calculation of an average. This has been explained in the last section of this chapter.

Column: It represents the width of the column. Default width is 8 and can be changed. We can specify a number of characters for the column width. Column widths can also be changed in Data View by clicking and dragging the column borders.

Align: This specifies the alignment in the Data View Window. By default, string variables are aligned left and numeric variables are aligned right.

Align
Right
Left

Measure: This specifies the level of measurement of the variable. SPSS takes values at nominal or ordinal or scale level.

Nominal: The values (the answer options) are a list of categories. For example, the variable Gender is coded as 1=male, 2=female and 9=unknown (the missing value). Note that an average of gender has no meaning.

Ordinal: A variable is ordinal if the list of answer options has a logical ordering. Ordering t-shirt sizes alphabetically (L, M, S, XL, XS, XXL) makes no sense. We all know that the logical ordering should be XS, S, M, L, XL, XXL. Another well known example is options used in Likert Scale: 1 = Strongly disagree, 2 = disagree, 3= Neither agree nor disagree, 4 = Agree, 5 Strongly agree. The higher the value, stronger will be the agreement. If you like, you can reverse the order, but scrambling it makes no sense.

Scale: A variable where the answers are actual numbers and where arithmetic mean makes sense. Examples are age, price, income etc.

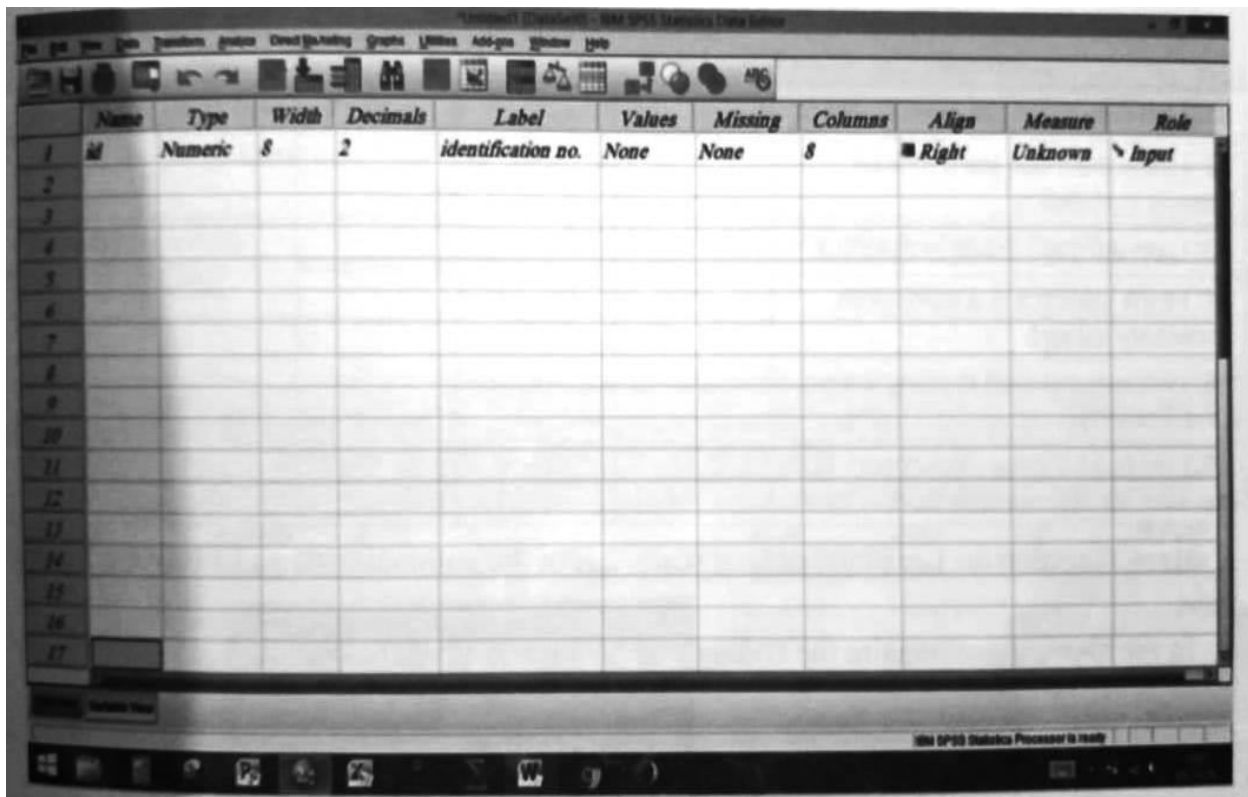
Role: This feature is new to SPSS 20. Here, we will indicate the role which a variable plays in our data analysis. For example, in regression analysis a variable can be used as cause (input) or as effect (target).

Basic Guidelines

1. Give every completely filled questionnaire a number or other unique identification number (ID No.). Create a variable in SPSS to contain this ID No.
2. Use numeric variables whenever possible.
3. Use such variable name for each question that will clearly define about it in a short form, as links to the numbers of the questions in the questionnaire. Use the variable labels to describe the meaning of the variables.
4. Use concise but complete variable labels and value labels.

- **To Define a Variable:**

1. Working in the Untitled-Data editor window, double-click on a variable name at the top of the column in the Data View or click the Variable View Tab.
2. In the first blank cell of the Name column, type the first variable name (i.e. id) and press ENTER.
3. In the first blank cell of the Label column, type the label for the variable, i.e., Identification Number.



CODING

Coding is the mechanism in which the researcher provides some codes to data in the form of symbols, characters and numbers. In other words, it involves assigning numbers or other symbols to answers so that the responses can be grouped into a limited number of categories. In coding, categories are partitions of a data set of a given variable (for example, if the variable is gender, the partitions are male and female).

Coding helps the researcher in interpretation of data and deriving accurate results. When the data is generated with the help of a questionnaire; it can be

coded either at the time of framing the questionnaire (pre-coded data) or after collecting the data (post-coded data).

Let us try to understand the concept of coding with the help of an example. Given below is a questionnaire that aims to measure student's personal confidence in learning statistics in a university.

Students' Personal Confidence in Learning Statistics					
(I, a research scholar, request your cooperation in filling up the following questionnaire. Your inputs are greatly valued.)					
Q1)	Age				
	Below 20 Years	<input type="text"/>	20-25 Years	<input type="text"/>	Above 25 Years <input type="text"/>
Q2)	Gender				
	Male	<input type="text"/>	Female	<input type="text"/>	
Q3) Show your opinion regarding these statements :					
(SA-strongly Agree, A-Agree, N-Neutral, DA-Disagree, SDA-Strongly Disagree)					
Statements	SA	A	N	DA	SDA
S1: I am sure that I can learn statistics.					
S2: I think that I can learn difficult lessons in statistics					
S3: I can get good grade in statistics.					
S4: I think statistics is a worthwhile, necessary subject.					
S5: I will use statistics in many ways as a researcher.					
S6: I will need a good understanding of statistics for my research work.					

Here, Question no. 1 is of the ordinal scale, 2 is of the nominal scale and 3 is of the interval scale.

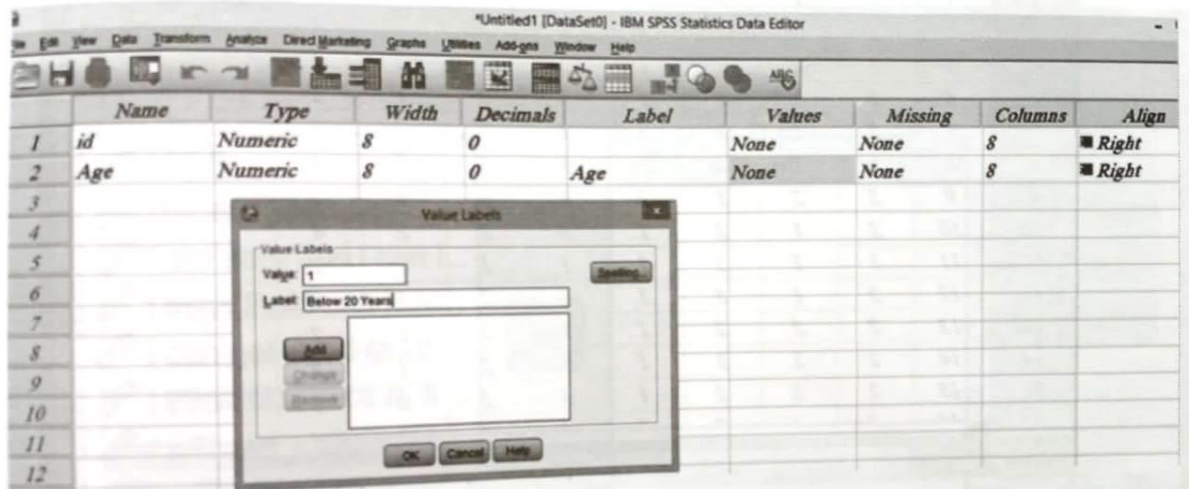
In the above questionnaire the coding will be done as given below:

Variable Name	Value labels	Values/coding
Age	Below 20 years	1
	20-25 years	2
	Above 25 years	3
Gender	Male	1
	Female	2
Statements	SA	5
	A	4
	N	3
	DA	2
	SDA	1

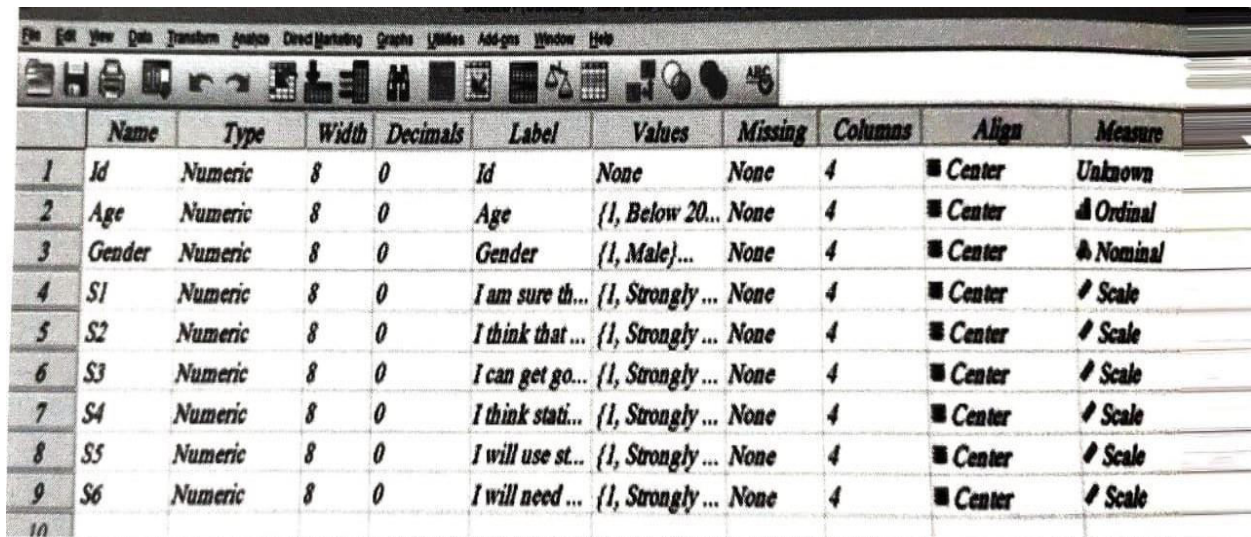
We can easily understand the procedure for defining values labels for age in SPSS with the help of table 10.1.

Table 10.1: Assigning Value Labels in SPSS

S. No.	SPSS Command/Action	Comments
1.	In variable view, in the column headed Values click on the small shaded square next to None .	This shows the Value Labels dialog box so that labels can be assigned.
2.	In the Value box: Type 1	
3.	In the Label box: Type Below 20 years	When we start typing Add suddenly darkens, whereas, it was previously faint.
4.	Click on Add	This pastes the information into the adjacent area so that 1="Below 20 years". The cursor will automatically jump to the Value box.
5.	In the Value box: Type 2	
6.	In the Value Label box: type 20-25 years	
7.	Click on Add.	2 = "20-25 years" will be added to the list
8.	In the Value box: Type 3	
9.	In the Value Label box: Type Above 25 years	
10.	Click on Add .	3 = "Above 25 years" will be added to the list
11.	Click on OK	



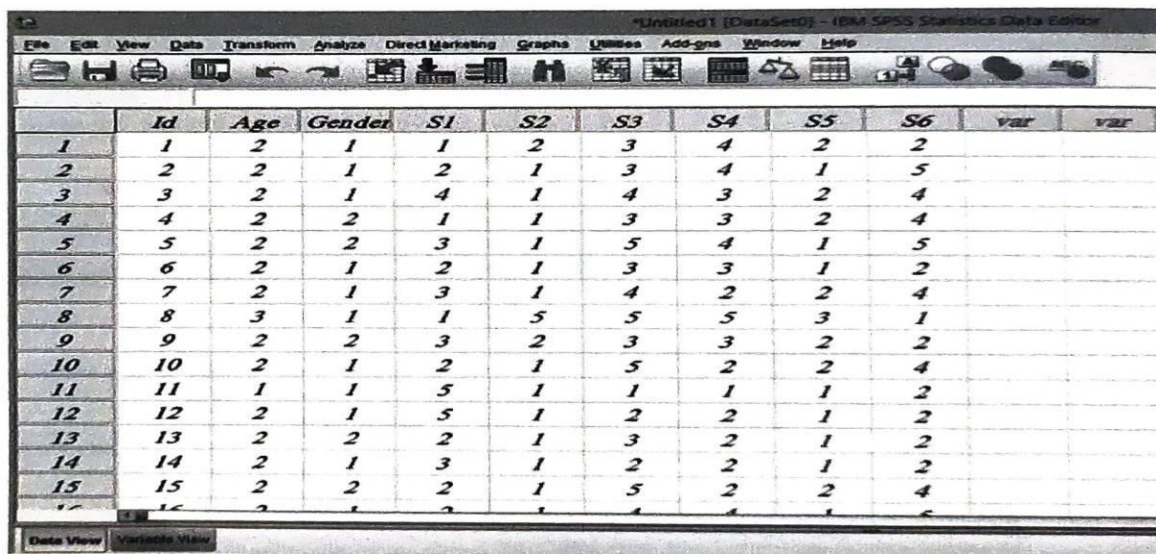
Repeat the above procedure for other variables also. Once you have finished defining all variables you can check your work by clicking variable view. In above example, variable view will look like this:



	Name	Type	Width	Decimals	Label	Values	Missing	Columns	Align	Measure
1	Id	Numeric	8	0	Id	None	None	4	Center	Unknown
2	Age	Numeric	8	0	Age	{1, Below 20...	None	4	Center	Ordinal
3	Gender	Numeric	8	0	Gender	{1, Male}...	None	4	Center	Nominal
4	S1	Numeric	8	0	I am sure th...	{1, Strongly ...	None	4	Center	Scale
5	S2	Numeric	8	0	I think that ...	{1, Strongly ...	None	4	Center	Scale
6	S3	Numeric	8	0	I can get go...	{1, Strongly ...	None	4	Center	Scale
7	S4	Numeric	8	0	I think stati...	{1, Strongly ...	None	4	Center	Scale
8	S5	Numeric	8	0	I will use st...	{1, Strongly ...	None	4	Center	Scale
9	S6	Numeric	8	0	I will need ...	{1, Strongly ...	None	4	Center	Scale

An overview of Variable View

In the data view, we start entering the data one by one of all the respondents. The complete data set in data view will look like as given below:



	Id	Age	Gender	S1	S2	S3	S4	S5	S6	var	var
1	1	2	1	1	2	3	4	2	2		
2	2	2	1	2	1	3	4	1	5		
3	3	2	1	4	1	4	3	2	4		
4	4	2	2	1	1	3	3	2	4		
5	5	2	2	3	1	5	4	1	5		
6	6	2	1	2	1	3	3	1	2		
7	7	2	1	3	1	4	2	2	4		
8	8	3	1	1	5	5	5	3	1		
9	9	2	2	3	2	3	3	2	2		
10	10	2	1	2	1	5	2	2	4		
11	11	1	1	5	1	1	1	1	2		
12	12	2	1	5	1	2	2	1	2		
13	13	2	2	2	1	3	2	1	2		
14	14	2	1	3	1	2	2	1	2		
15	15	2	2	2	1	5	2	2	4		

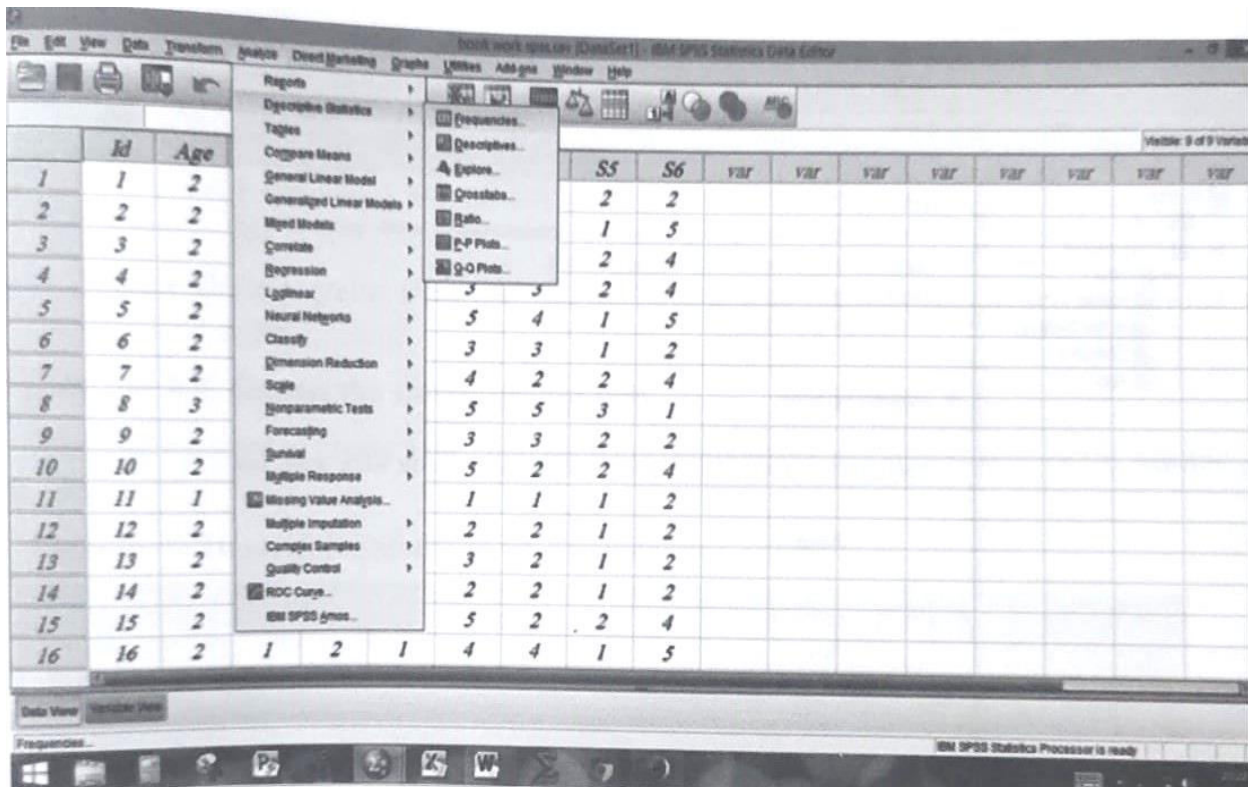
An overview of Data View

FREQUENCY DISTRIBUTION

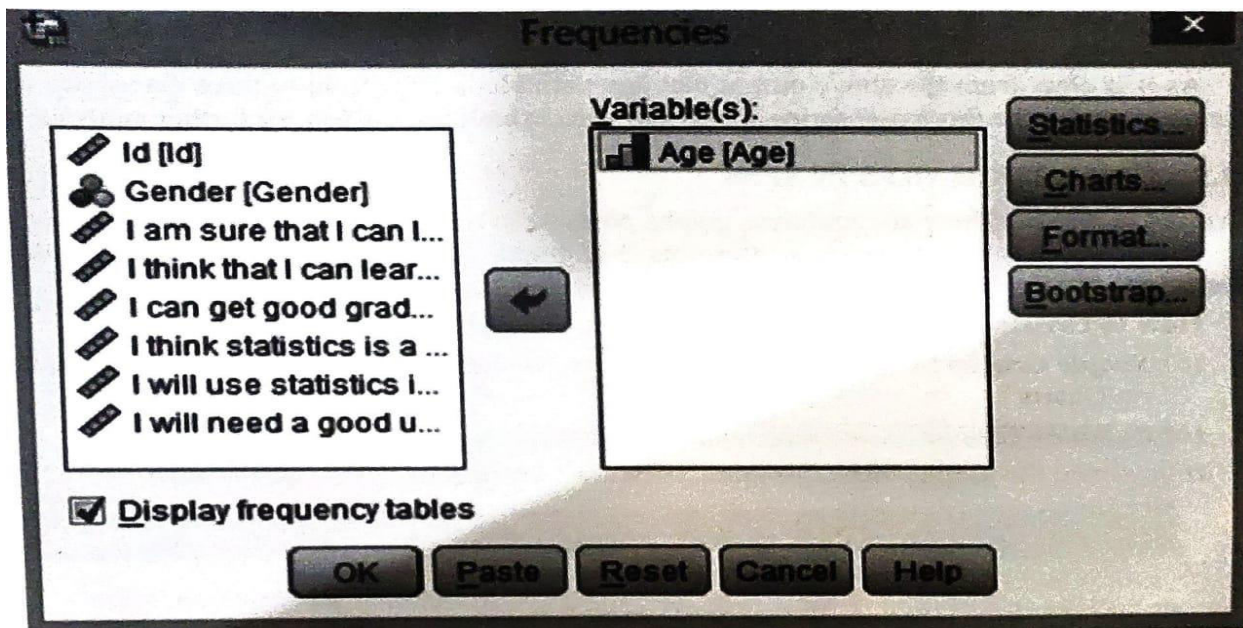
After applying coding, we can easily obtain the frequency table of above variables corresponding to their codes. Like if we want to have a look at the frequency table of age variable,

Go to:

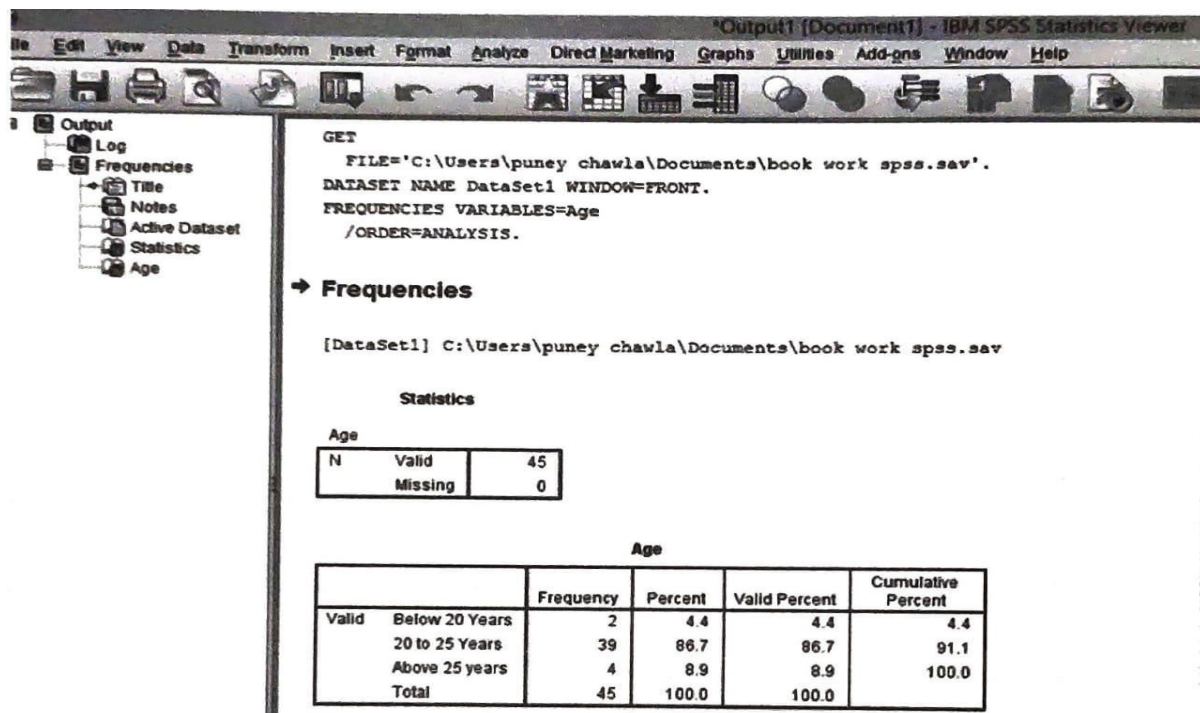
Analyze => (Right arrow) Descriptive Statistic => Frequencies



A dialogue box named Frequencies will appear. Select age and click on the button in entre to move this variable into Variable(s): box. Click on OK.



We will get the results in the form of table in output window.



As it is clear from the above output that age variable is divided in to three categories and frequency of all the three categories are shown here, which we can use for further analysis.

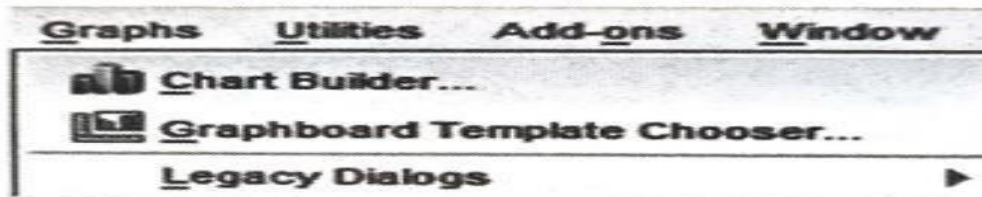
GRAPHS/CHARTS IN SPSS

Graphs in SPSS: There are countless graphs possible in SPSS. High-resolution, full-colour pie charts, bar charts, histograms, scatterplots, 3-D graphics and more are included as standard features of SPSS. Charts and graphs words are used interchangeably in SPSS.

How to Create Graph in SPSS: There are various options to create graphs in SPSS:

- a) **Simple Graphs:** Use chart option in Analyse => Descriptive Statistics => Frequencies=> Charts
- b) **Complex Graphs:** Use graphs in menu bar.

If you open the graphs menu, you find three main entries

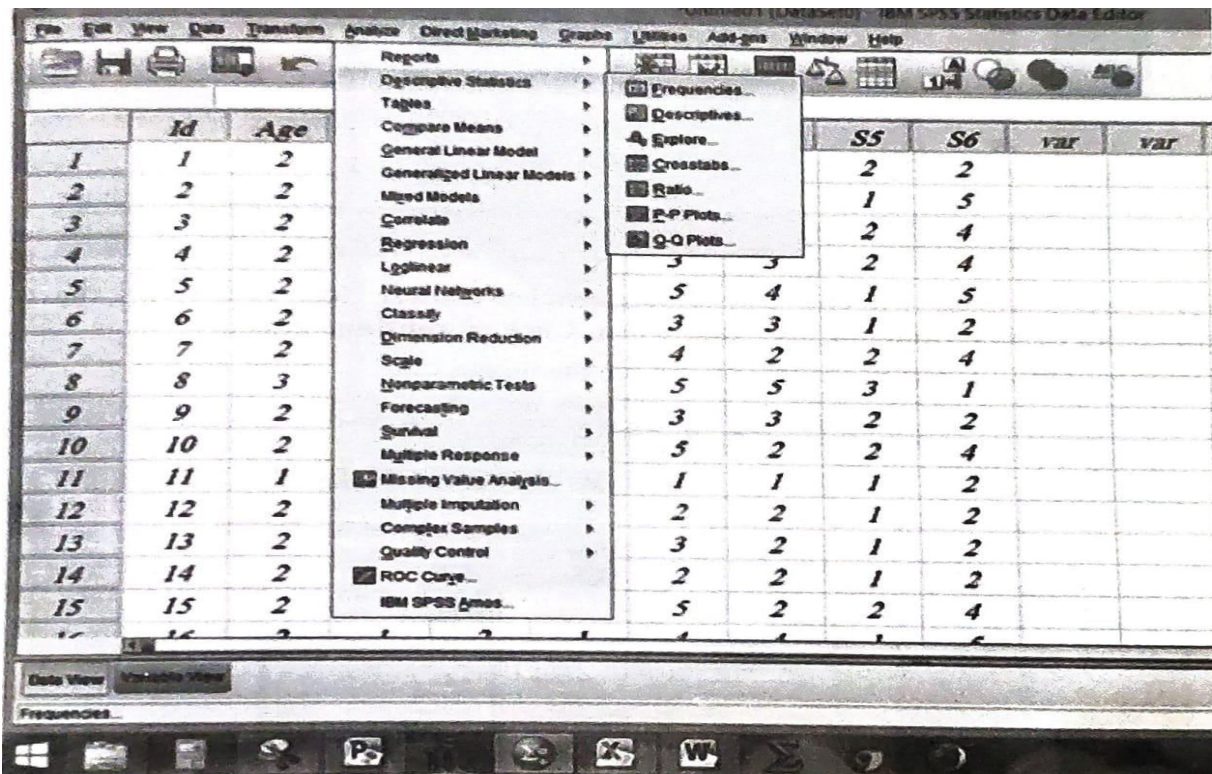


1. In **Chart Builder**, we drag elements to a preview screen to create our graph. Here, it is important that the level of measurement of our variables is set correctly.
2. Via **Graphboard Template Chooser**, we create graphs based on predefined templates. Often, these templates are specially designed to expand on the normal options. For beginners this is less relevant.
3. **Legacy Dialogs** stems from the "old SPSS" and uses dialog boxes to determine what you want to do.

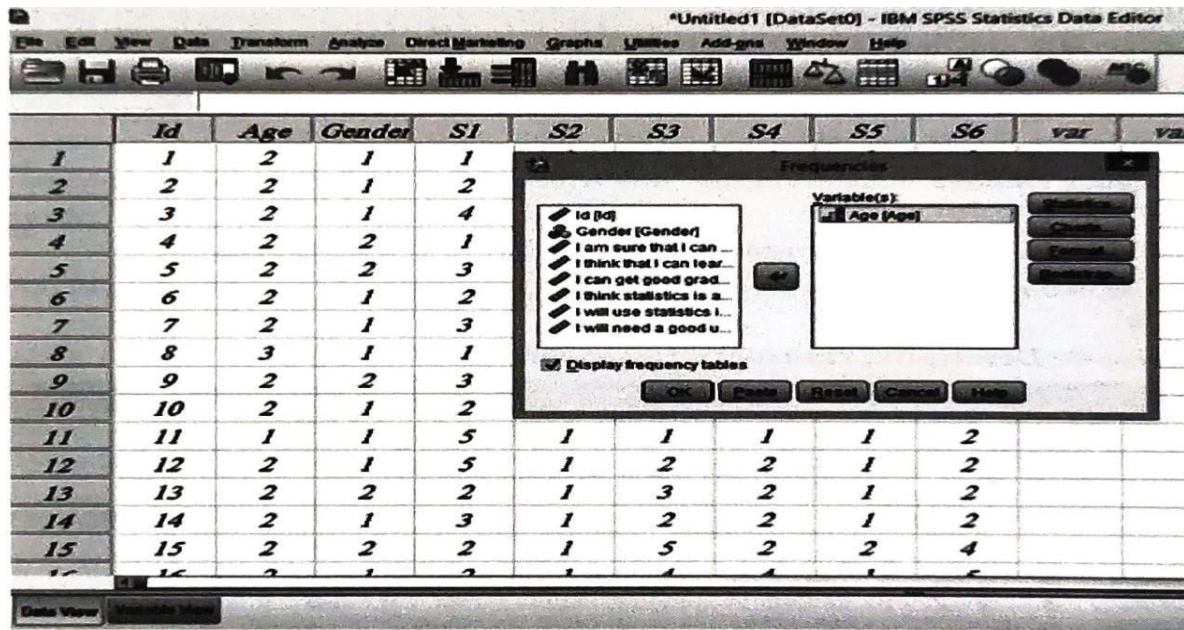
Here, we will discuss the simple graphs only.

Suppose, we want to create a pie chart showing respondents' age distribution of our sample data. Go to:

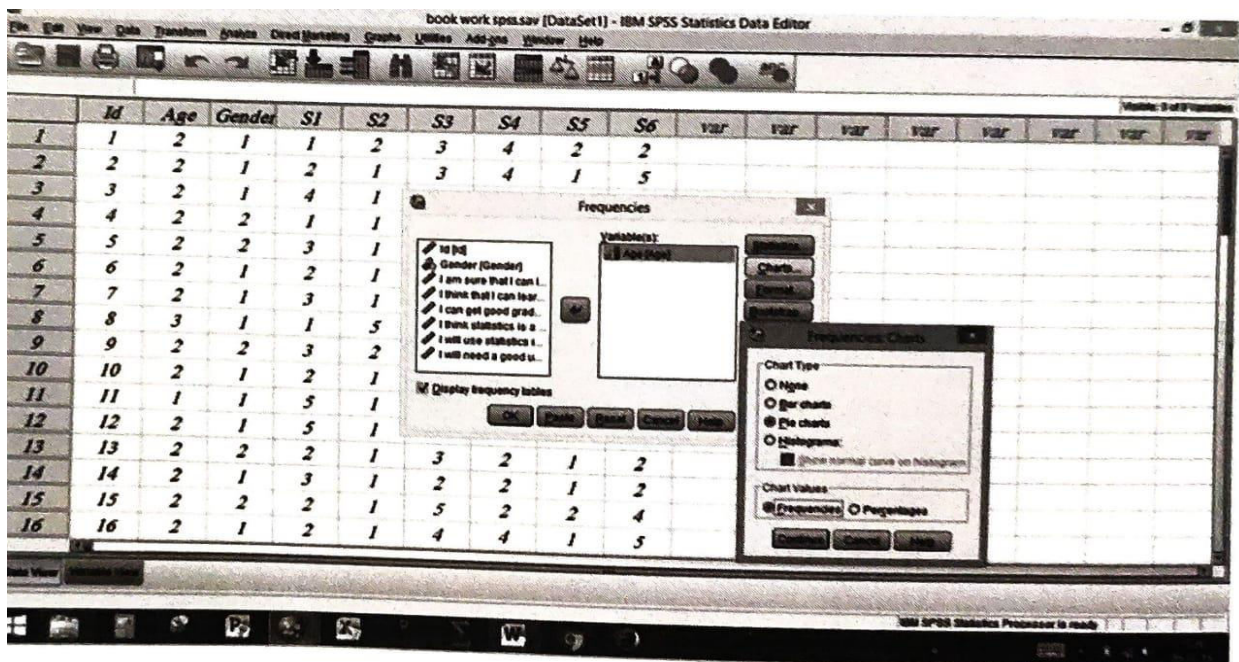
Analyse => Descriptive statistics => frequencies.



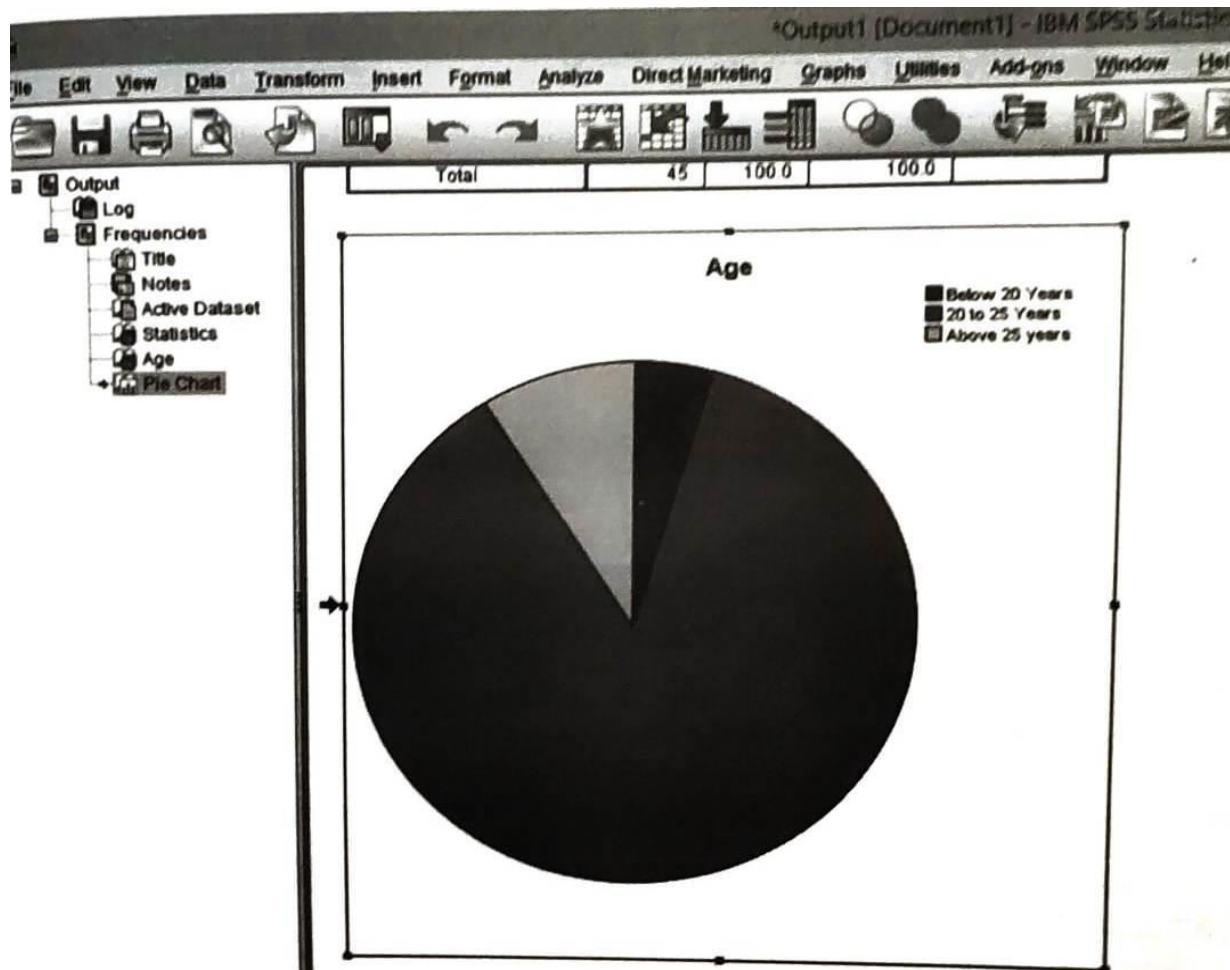
A window named Frequencies will open. After this, move the variable of interest (age) to the variable(s): box, click on Charts button at the right of the window.



Now, a new window will open named Frequencies: Charts. Here, click on Pie charts and Frequencies radio button in Chart values box. Click on continue button. Click on OK in Frequencies window to see the output.



This is the output window we will see. Now, we have a pie chart that graphically shows the distribution of the survey represents' age.



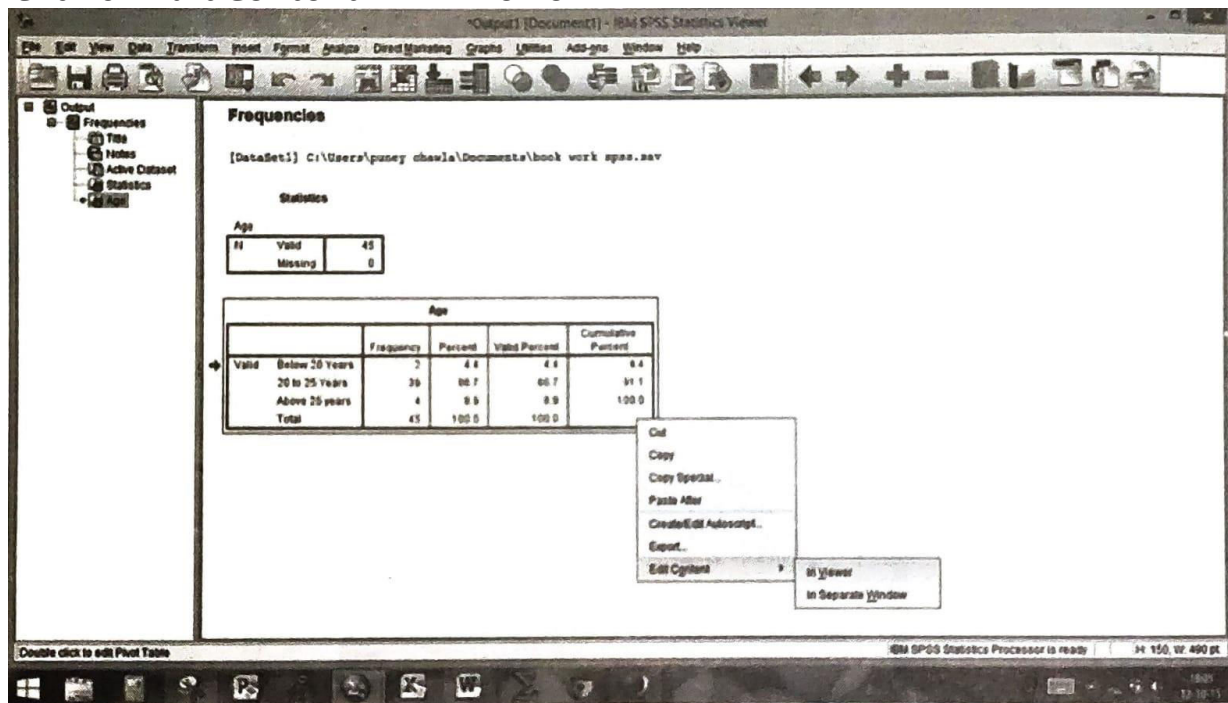
PIVOT TABLE IN SPSS

A pivot table is a table capable of dynamically displaying and rearranging multiple dimensions. There are three display areas of a pivot table: Rows, Columns, and Layers. The dimensions of any table can be easily moved from row to column to layer. SPSS adds a Title, Caption, and Footnotes to the general idea of a pivot table.

Pivot Table Editor: Output displayed in pivot tables can be modified in different ways. Using the editor, it is possible to edit text, rearrange rows, columns and layers, add colours, create multidimensional tables and selectively hide and display results. To create a pivot table, Go to:

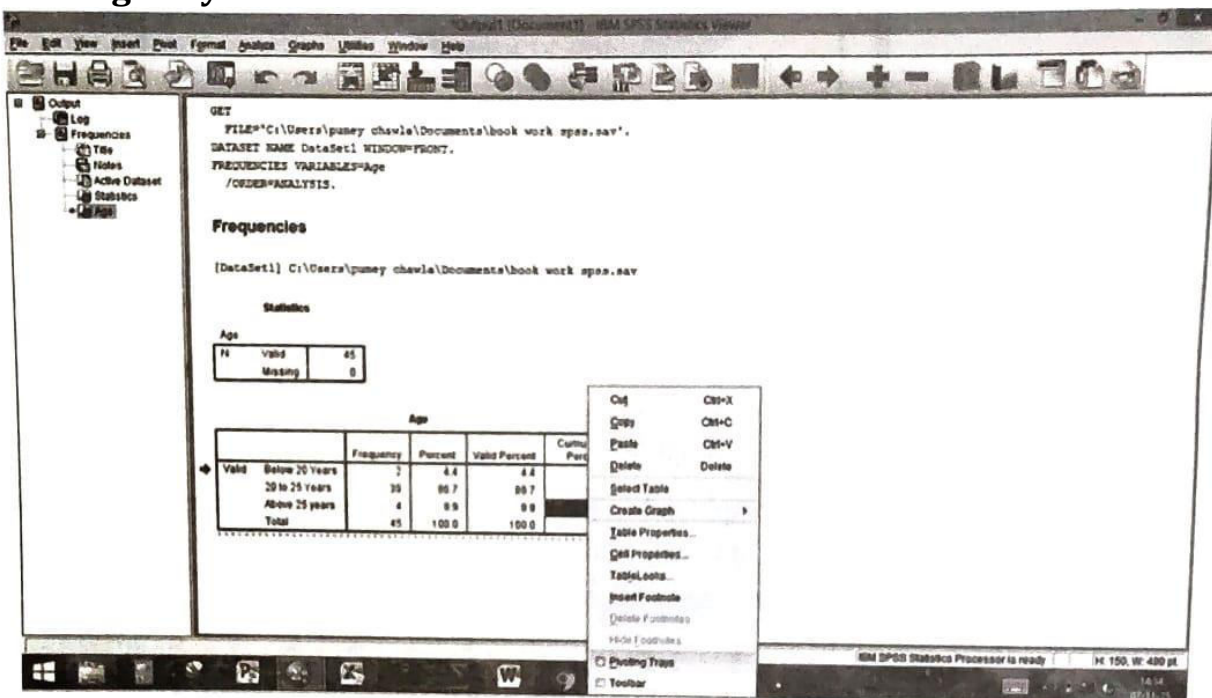
Open the **output window** containing a table for which you want to create a pivot table.

Right Click on the Table and a list of options will appear.
Click on **Edit Content => In Viewer.**



As a result of it, our table gets highlighted.

Right click on the highlighted table. A list of options will appear and check **Pivoting Trays** box.



A small window named **Pivoting Trays** will open at the right hand side of the main output window. We may also predefine the number of rows we would like for our table to display by selecting the set rows to display feature.

The screenshot shows the IBM SPSS Statistics Viewer interface. The main window displays the output for the 'Frequencies' command. The command is: `GET FILE='C:\Users\puney chawla\Documents\book work spss.sav'. DATASET NAME DataSet1 WINDOW=FRONT. FREQUENCIES VARIABLES=Age /ORDER=ANALYSIS.`

The 'Frequencies' window shows the following table for the variable 'Age':

N		Valid	Missing
		45	0

The main table shows the distribution of 'Age' values:

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	Below 20 Years	2	4.4	4.4	4.4
	20 to 25 years	39	86.7	86.7	91.1
	Above 25 years	4	8.9	8.9	
	Total	45	100.0	100.0	

The 'Pivoting Trays' dialog box is open, showing options for Layer, Row, and Column. The 'Statistics' section is also visible.

Double-click at any point inside the table we want to edit and then click the right mouse button to select Pivoting Trays and Toolbar. This will reveal the Pivot Table, formatting toolbar and pivot trays and allow editing of tables to begin.

MISSING VALUES IN SPSS

Missing data are the information from a participant or case that is not available for one or more variables of interest. Missing values can arise for many reasons. For any survey question there is always the possibility that someone simply did not answer the question, or else wrote with illegible handwriting. For whatever reason, when we do not have a useful data for an individual case for a specific variable we need to enter a missing value into the relevant cell, indicating that a valid response was not provided in that instance and therefore, should not be included in any analysis of that variable.

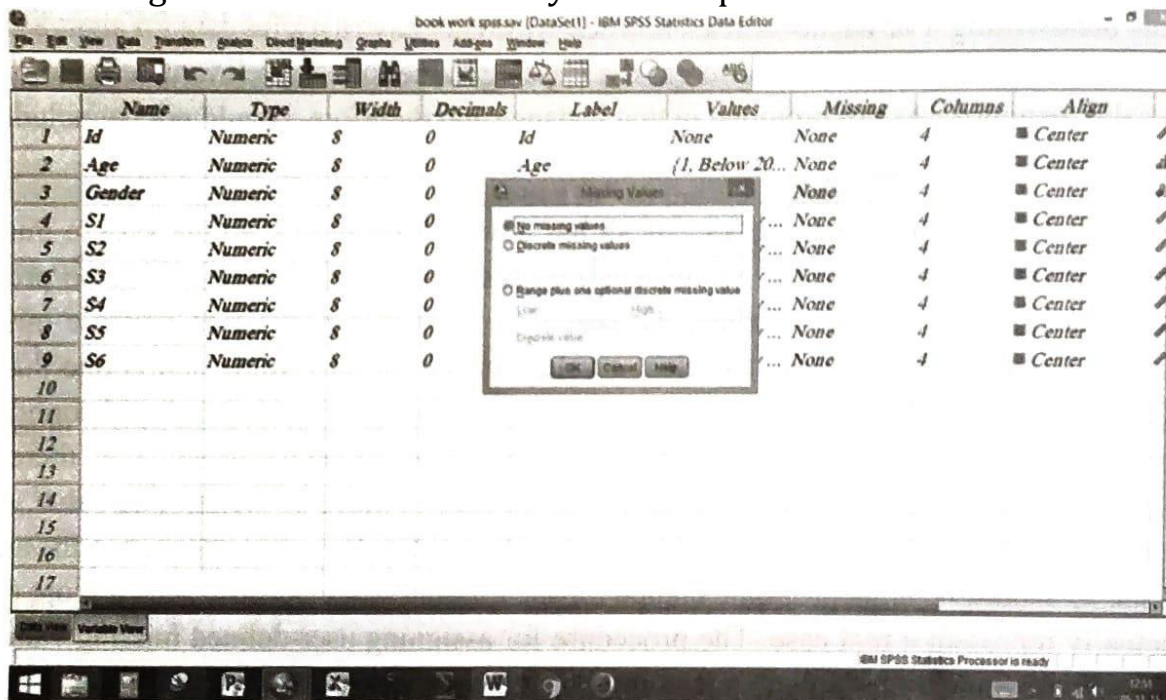
Dealing Missing Values in SPSS: SPSS has a default setting called the system missing value that appears as a dot (.) in a cell where no response is entered. We can also provide user-defined missing values, whereby, we specify a particular number to indicate missing responses. We need to be careful to select a value for the missing value that is one that the variable cannot

possibly take. In the previous example of measuring the students' confidence in learning statistics, if some of the students have not ticked the gender option, then there will be missing data in gender. We can choose 9 to be entered in place of the missing value, since it is impossible for the variable 'Gender' to take on this value. If, we are measuring the age of students, 99 would be an appropriate choice for the missing value because such a score could not actually represent a real case. The procedure for assigning user-defined missing values to the "Gender of student" variable in our example is given below:

Table 10.2: Assigning User-defined Missing Values in SPSS

S. No.	SPSS command/action	Comments
1.	In variable view, in the column headed Missing, click on the small shaded square next to None.	This brings up the Missing Values: dialogue box. The radio button next to No missing values is selected indicates that no user-defined missing values is the by default setting.
2.	Click on Discrete missing values radio button.	
3.	In the first box: Type 9	
4.	Click on OK	

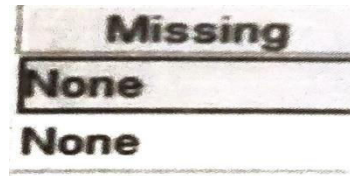
Having defined the missing value when we set up the data file, we can then enter 9 (nine) while entering the data for missing value, if we encounter a case for where gender box is left blank by some respondents.



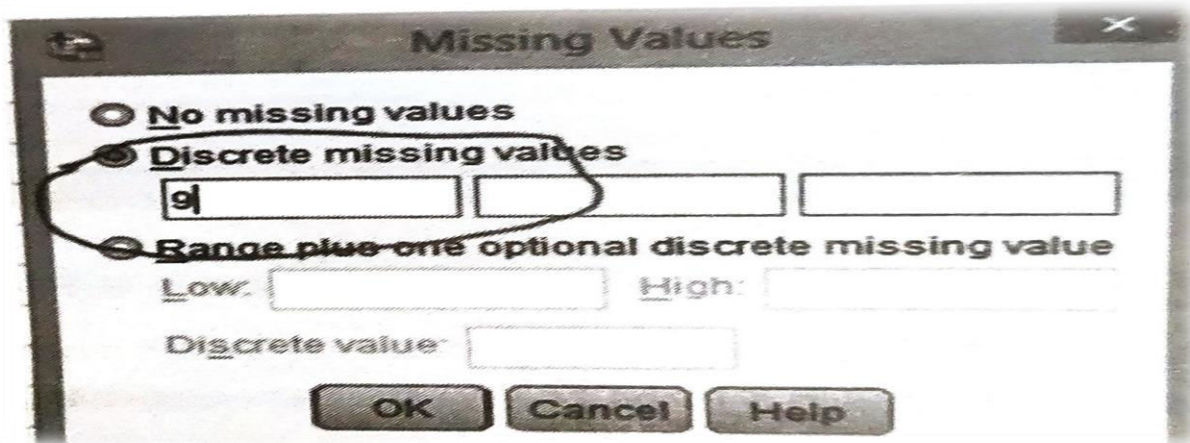
How to Handle Missing Data in SPSS

Before entering the data:

1. If you want that SPSS doesn't include missing data in any cell, you must follow important step. Go to the Variable View.
2. The seventh column from the left is called 'Missing'. Click on the first cell under this column, and click on the blue box that appears in the cell.



3. Select Discrete Missing Values and enter in the box the number that you choose for missing data say 9. Click on OK.



4. After this in variable view we have the value 9 as missing as shown below.

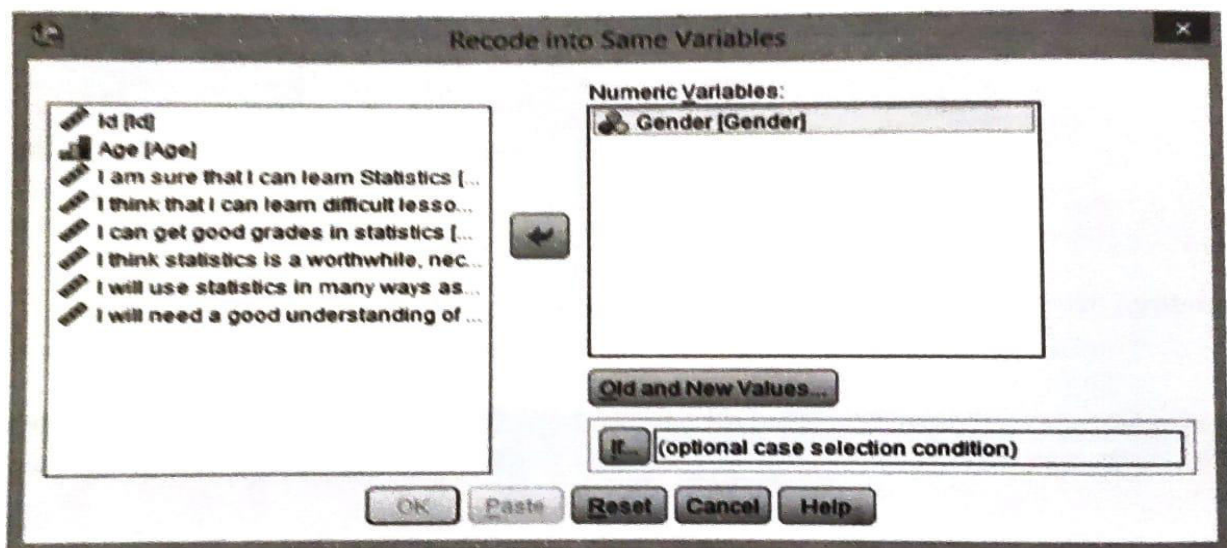
	Name	Type	Width	Decimals	Label	Values	Missing	Columns
1	Id	Numeric	8	0	Id	None	None	9
2	Age	Numeric	8	0	Age	{1. Below 2	None	7
3	Gender	Numeric	8	0	Gender	{1. Male}	9	8

Handling Missing Data: After entering the data:

1. You have entered all the data. After that you found some missing values. Like in this example ID 5 and 10 have missing values in gender.

	ID	Age	Gender	S1	S2	S3	S4	S5	S6	var	var
1	1	2	1	1	2	3	4	2	2		
2	2	2	1	2	1	3	4	1	5		
3	3	2	1	4	1	4	3	2	4		
4	4	2	2	1	1	3	3	2	4		
5	5	2		3	1	5	4	1	5		
6	6	2	1	2	1	3	3	1	2		
7	7	2	1	3	1	4	2	2	4		
8	8	3	1	1	5	5	5	3	1		
9	9	2	2	3	2	3	3	2	2		
10	10	2		2	1	5	2	2	4		
11	11	1	1	5	1	1	1	1	2		
12	12	2	1	5	1	2	2	1	2		
13	13	2	2	2	1	3	2	1	2		
14	14	2	1	3	1	2	2	1	2		
15	15	2	2	2	1	5	2	2	4		
16	16	2	1	2	1	4	4	1	5		
17	17	2	2	2	3	3	3	4	1		
18	18	3	1	2	2	4	2	2	3		
19	19	2	1	2	1	4	4	1	5		
20	20	3	1	1	2	2	3	3	2		
21	21	3	1	5	1	3	2	2	2		
22	22	2	1	5	1	4	1	1	5		
23	23	2	2	2	2	4	2	2	2		

2. To get SPSS to fill in all the empty cells, go to **Transform => Record into same variables.**
3. Move variable (Gender) into the right hand box and click on **Old and New Values.**



4. On the left, select **System-or user missing** and on the right enter a number that will not otherwise occur in your data set (e.g. 9) in the 'New Value' box. Click on Add, and then Continue. Click on Ok in **Recode into same variables window**.

The image shows a dialog box titled "Recode into Same Variables: Old and New Values". It is divided into two main sections: "Old Value" and "New Value".

Old Value:

- Value:
- System-missing
- System- or user-missing
- Range:
- Range, LOWEST through value:
- Range, value through HIGHEST:
- All other values

New Value:

- Value: [9]
- System-missing

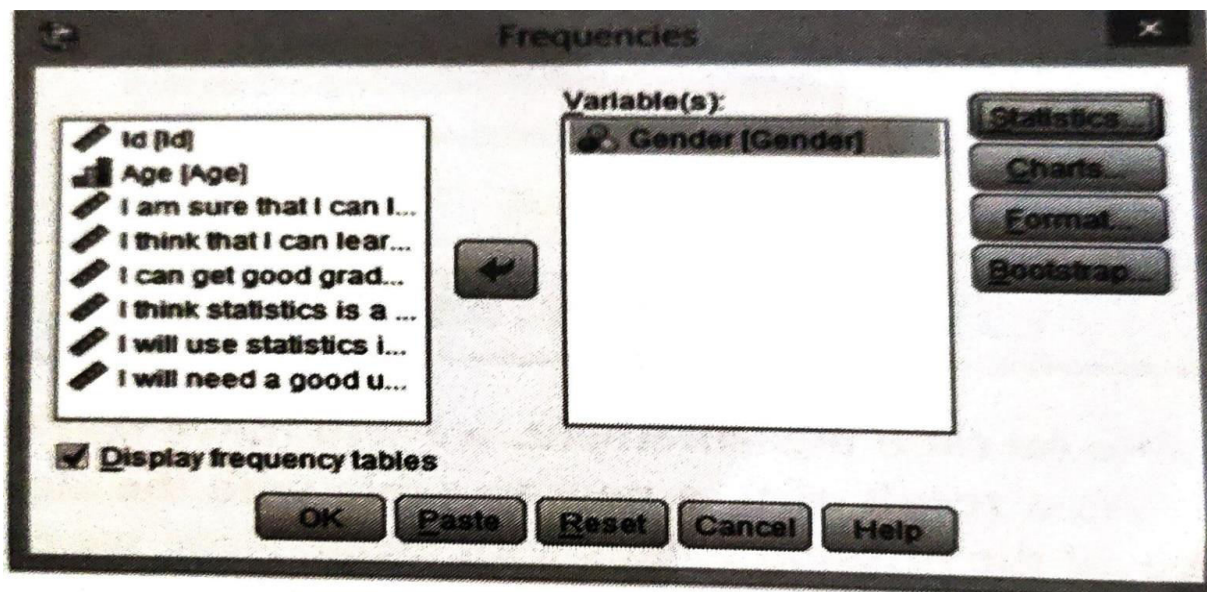
Old -> New:

Buttons: Add, Change, Remove, Continue, Cancel, Help

5. All blank cells will now be replaced with the value (9) you entered in the previous step.

	Id	Age	Gender	S1	S2	S3	S4	S5	S6	var
1	1	2	1	1	2	3	4	2	2	
2	2	2	1	2	1	3	4	1	5	
3	3	2	1	4	1	4	3	2	4	
4	4	2	2	1	1	3	3	2	4	
5	5	2	9	3	1	5	4	1	5	
6	6	2	9	2	1	3	3	1	2	
7	7	2	1	3	1	4	2	2	4	
8	8	3	1	1	5	5	5	3	1	
9	9	2	2	3	2	3	3	2	2	
10	10	2	9	2	1	5	2	2	4	
11	11	1	1	5	1	1	1	1	2	
12	12	2	1	5	1	2	2	1	2	
13	13	2	2	2	1	3	2	1	2	
14	14	2	1	3	1	2	2	1	2	
15	15	2	2	2	1	5	2	2	4	
16	16	2	1	2	1	4	4	1	5	
17	17	2	2	2	3	3	3	4	1	
18	18	3	1	2	2	4	2	2	3	
19	19	2	1	2	1	4	4	1	5	
20	20	3	1	1	2	2	3	3	2	
21	21	3	1	5	1	3	2	2	2	
22	22	2	1	5	1	4	1	1	5	
23	23	2	2	2	2	4	2	2	2	

6. After this we can view the frequency table to see that how many cases have missing values. For this you need to go to **Analyze => Descriptive Statistics => Frequencies**. A dialog box named frequencies will appear. Move the Variable (Gender) to variable box.



Click on OK.

You will receive the output as given below from where you can check how many times 9 (i.e. missing frequencies) are there.

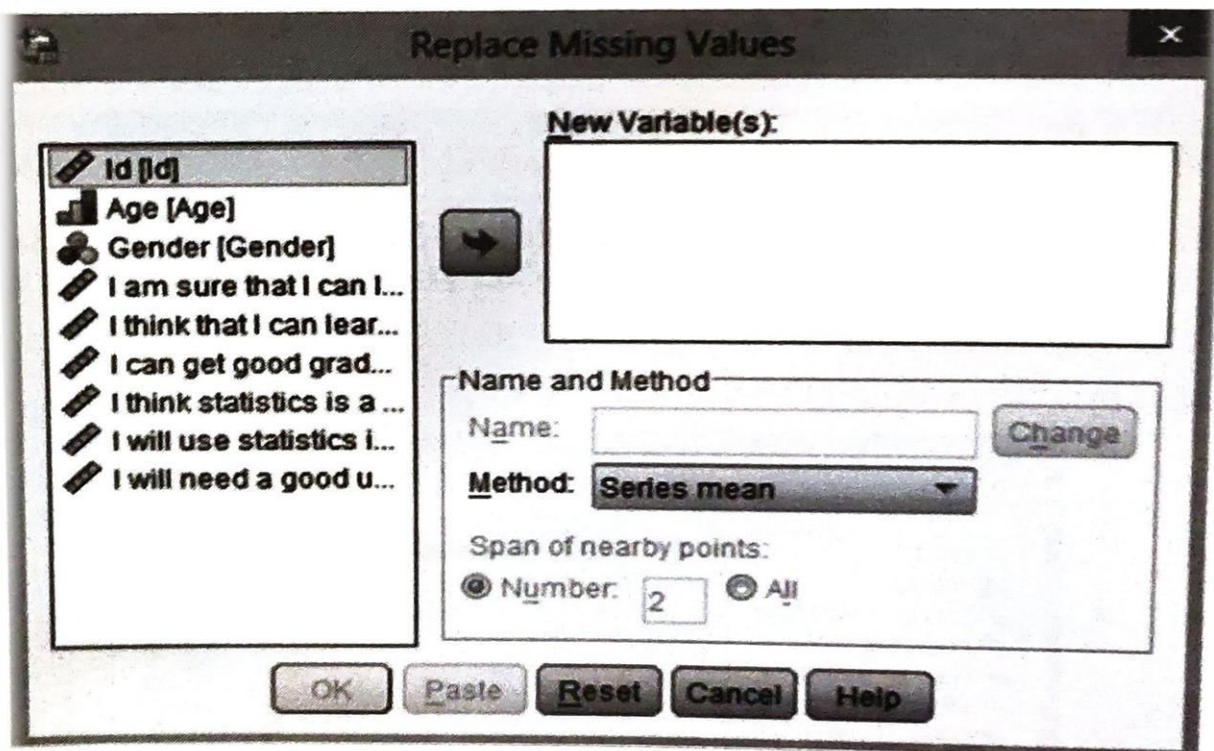
Statistics

Gender

N	Valid	45
	Missing	0

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	Male	30	66.7	66.7	66.7
	Female	13	28.9	28.9	95.6
	9	2	4.4	4.4	100.0
Total		45	100.0	100.0	

Replacing Missing Values with Mean: We can replace the missing values mean. This is very much popular in the variables having ratio measurement. For this go to **Transform => Replace Missing Values**. The following screen appears.



If we have any variable on ratio measurement, we will move that variable in new variable box and the click on ok. As a result of it, a new variable with the almost same name will be created and it will replace all the missing values with mean. Now, we can use this new variable for further analysis having no missing values.

Applications of Chi-Square Test

Introduction:

A Chi-square test is any statistical hypothesis test where the sampling distribution of the test statistics is a chi-squared distribution when the null hypothesis is true. Chi-square is a very versatile test used both as a non-parametric and a parametric measure. Chi-square test is an approximate test for large values of n.

Conditions:

For the validity of chi-square test of 'goodness of fit' between theory and experiment, the following conditions must be satisfied,

- i) The sample observation should be independent.
- ii) Constraints on the cell frequencies, if any, should be linear
- iii) N, the total frequency should be reasonably large, say, greater than 50.
- iv) No theoretical cell frequency should be less than 5.

Applications of Chi-square Distribution:

Chi-square distribution has a large number of applications in Statistics, some of which are enumerated below:

- i) To test if the hypothetical value of the population variance is $\sigma^2 = \sigma_0^2$
- ii) To test the 'goodness of fit'.
- iii) To test the independence of attributes.
- iv) To test the homogeneity of independent estimates of the population variance.

v) To combine various probabilities obtained from independent experiments to give a single test of significance.

vi) To test the homogeneity of independent estimates of the population correlation coefficient.

Example: 1 (Sports Analytics)

Chi-square can be used in Sports analytics as an example of validating a belief or hypothesis using analytics. It can also be used in a business analytics or business statistics course as an application of the Chi-square goodness of fit

Techniques and concepts:

Hypothesis testing, Contingency table, Chi square goodness of fit.

TENNIS:

Tennis is played by millions of recreational players and is also a popular worldwide spectator sport. The four Slam tournaments (also referred to as the Majors) are especially popular: the Australian Open(19905) played on hard courts, the French Open (1891) played on red clay courts, Wimbledon(1877) played on grass courts, and the US Open(1881) also played on hard courts.

TYPE OF COURT SURFACE:

A variety of surfaces can be used to create a tennis court, each with its own characteristics which affect the playing style of the game. There are four main types of courts depending on the materials used for the court surface: clay courts, hard courts, grass courts and carpet courts.

Hard Court:

Hard court made up the most common tennis-court surfaces. Theses courts played faster than clay courts but slower than grass courts. Hard courts also offered a more predictable bounce. Due to the fast nature of these courts, players commonly tried to keep the length of their rallies to a minimum.

Grass court:

Grass courts were the fastest of all court surfaces. Due to the slippery surface of the grass, the ball was inclined to bounce low and skid across the surface.

The fast nature of the surface encouraged players to put a quick end to each point.

Clay Court:

Clay courts constituted the slowest courts because their surface absorbed much of the bounce and reduced the forward motion of the ball. Due to the slow nature of the surface, rallies were lengthy and required great athleticism on the part of the player. These courts favoured defensive style players who engaged in long rallies from the baseline.

Chi square Analysis:

In general, we presume that winning or losing any tennis match is completely based on the coach, the practice and fitness of the player. But logically the type of court they play has significant impact in winning or losing the match. To understand that, Chi square test is performed between the players and different surfaces. The results clearly tell us the type of court makes the difference.

Wins and Losses on different court surfaces

	Grass		Hard		Clay	
	Win	Loss	Win	Loss	Win	Loss
Federer	94	5	427	61	185	37
Nadal	50	12	254	81	283	26
Djokovic	43	12	288	61	141	42
Murray	45	11	243	63	76	33

Based on above data, describe a hypothesis test to determine whether each player's performance depends on the type of surface.

Null Hypothesis (H₀) : The type of court surface does not make a differences in Roger federer performance.

Alternative Hypothesis (H₁) : The type of court surface does make a differences in Roger federer performance.

Chi-Square Statistics:

Player	P-Value	Conclusion
Federer	0.0152	Reject the Null
Nadal	0.000	Reject the Null
Djokovic	0.2904	Fail to reject Null
Murray	0.0984	Fail to reject Null

Application of t-test

The t test is type of inferential statistics used to determine whether there is a significant difference between the means of two groups (compare mean between 2 samples). In simple words it can be put as t test compares "signal" to "noise" ratio.

lower the noise, greater is the t value and vice versa.

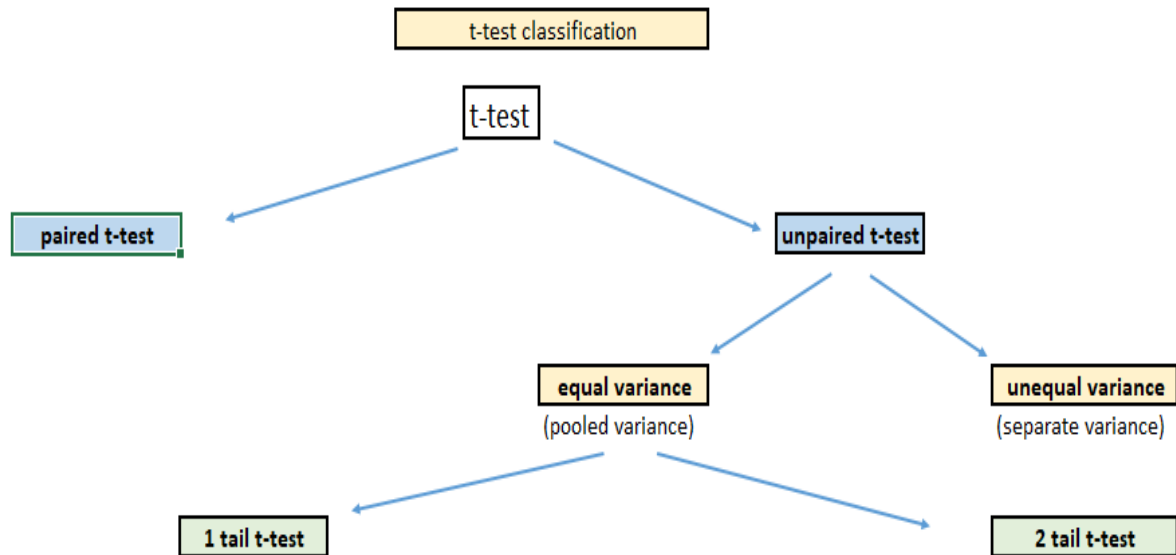
$$t = \frac{\bar{x}_2 - \bar{x}_1}{S_{x_1x_2} \sqrt{\frac{1}{n_1} + \frac{1}{n_2}}}$$

← Signal (pointing to $\bar{x}_2 - \bar{x}_1$)
← noise (pointing to $S_{x_1x_2} \sqrt{\frac{1}{n_1} + \frac{1}{n_2}}$)

Assumptions of t test:-

1. that the data is normally distributed (bell curve)
2. that the two samples come from distributions that may differ in their mean value, but not in the standard deviation.
3. only a single variable is being compare (univariate).
4. variable are continuous variables.
5. that the observations are independent of each other.

Classification of t -test:-



paired t-test	unpaired t-test
used when we have samples "before & after taken from same subject", e.g:- data taken before a particular drug was administered, and after the drug was administered	the procedure compares the averages of two samples that were selected independently of each other, e.g:- height measurement taken from different groups

equal variance	unequal variance
Assumes that the variance between 2 groups is same.	Assumes that the variance between 2 groups is not the same. This is more conservative as compared to equal variance.

1 tail t-test	2 tail t-test
Used when we know that mean of group is higher than that of other group.	Used when we do not know that mean of group is higher than that of other group. This is more conservative as compared to 1 tail t-test.

Application of t-test:-

1) If we question regarding yield of field A & field B, which has a higher yield ?

The most obvious answer is to compare mean yield between the both and determine the same, but it may be the case that because of sampling the difference appears, but actually it may not be, how do we confirm this?, the answer is t-test.

1) Calculate the t value (signal \ noise ratio), if it is >1 means we have more signal than noise.

2) Formulate the null hypothesis, and test the null hypothesis

Ho (NULL hypothesis) = there is no statistical difference between mean of 2 samples (any difference we find is simply because of chance)

H1 (Alternate Hypothesis) = there is inequality (difference exists between mean of 2 samples)

3) perform t statistic on data, for p value of 0.05, and degrees of freedom ([size of sample size 1 + size of sample size 2] -2), if the t value in step 1 is greater than critical we reject the null hypothesis and conclude there is significant difference between mean's of 2 samples.

We can extend the same logic for comparing heights of individuals' between 2 samples, or for comparing the lengths between leave size of 2 plants (on in sunlight & other in shade).

Disadvantages of T Test:-

1) The main disadvantage of t-test is it does not give accurate results on large datasets ($n > 30$ samples). If you have larger sample sizes go for Z-test.

2) The sample sizes used for comparing the means has to be approximately the same.

Analysis of Variance

When we have only two samples we can use the t-test to compare the means of the samples but it might become unreliable in case of more than two samples. The analysis of variance, popularly known as the ANOVA, can be used in cases where there are more than two groups.

ANALYSIS OF VARIANCE

Analysis of Variance (ANOVA) is a statistical method used to test differences between two or more means. The technique of analysis of variance was first originated by R. A. Fisher in 1920s which was later on improved in systematic manner by some other eminent academicians. It may seem odd that the technique is called "Analysis of Variance" rather than "Analysis of Means." But, the name is appropriate because inferences about means are made by analysing variance. The following definitions may be given to explain its meaning more clearly.

According to Yule and Kendall, "The analysis of variance is essentially a procedure for testing the difference between different groups of data of homogeneity." According to R.A. Fisher, "Analysis of variance is the separation of the variance ascribe to one group of causes from the variance ascribe to other group."

This method is often used in scientific or medical or agriculture experiments when treatments, processes, materials or products are being compared. For example, make a list of 10 food household items purchased regularly by your family. Obtain the current prices of the items from three different shops; preferably a small "corner" shop, a small supermarket and a large supermarket or hyper market. Compare different shop prices. The variations arise because of many factors is known as total variation or total variance. In analysis of variance, the degree of variance between two or more data as well as the factors contributing towards the variance is studied.

WHY NOT MULTIPLE t-TESTS?

T-test can only be used to test differences between two means. When there are more than two means, it is possible to compare each mean with another mean using many t-tests. Every time you conduct a t-test, there is a chance that you will make a Type 1 error. This error is usually 5%. By running two t-tests on the same data, we will have increased our chance of "making a mistake" to 10%. The formula for determining the new error rate for multiple t-tests is not as simple as multiplying 5% by the number of tests. However, if we are only making a few multiple comparisons, the results are very similar if we do. As such, by conducting 1-tests three times, error would be 15% (actually, 14.3%) and so on. These are intolerable errors. An ANOVA controls for these errors so that the Type 1 error remains at 5% and we can be more assured that any significant result we find is not just due to chance.

Assumptions: The analysis of variance technique is based on the following assumptions:

1. Each sample is drawn randomly from a normal population and the sample statistic tends to reflect the characteristics of the population.
2. The populations from which the samples are drawn have identical means, i.e.

$$\mu_1 = \mu_2 = \mu_3 = \dots = \mu_k$$

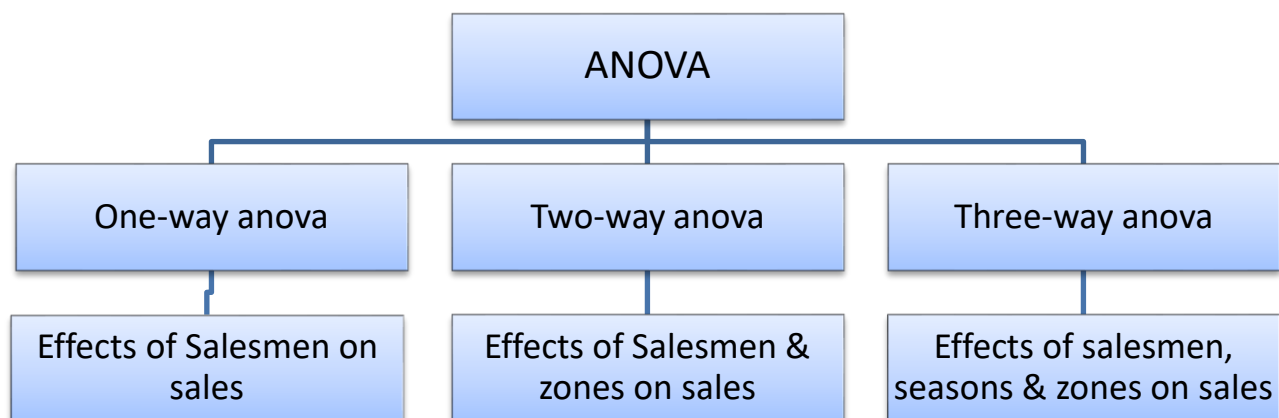
3. The populations from which the samples are drawn have identical variances, ie.

$$\sigma_1^2 = \sigma_2^2 = \sigma_3^2 = \dots = \sigma_k^2$$

In case we are not in a position to make these assumptions in a particular problem, the analysis of variance should not be used. In such cases, we should consider using a "Non parametric (distribution-free) technique."

TECHNIQUES OF ANOVA

The technique of analysis of variance can be divided into three parts.



A. ONE-WAY ANOVA

When data is classified on the basis of one attribute is called "one-way anova", e.g., to test the level of efficiency of various salesmen on the basis of their average sales or to understand whether exam performance differs based on anxiety levels amongst students, by dividing students into three independent groups (e.g., low, medium and high-stressed students). The one-way analysis of variance (ANOVA) is used to determine whether there are any significant differences between the means of two or more independent (unrelated) groups (although we tend to only see it is used when there are a minimum of three, rather than two groups).

What does this test do? The one-way ANOVA compares the means between the groups we are interested in and determines whether any of these means are significantly different from each other. Specifically, it tests the null hypothesis:

$$H_0: \mu_1 = \mu_2 = \mu_3 = \dots = \mu_k$$

Where μ = group mean and k = number of groups.

If, however, the one-way ANOVA returns a significant result, we accept the alternative hypothesis (H_a), which is that there are atleast 2 group means that are significantly different from each other. At this point, it is important to realize that the ANOVA is an omnibus test statistic and cannot tell us which specific groups were significantly different from each other. To determine which specific groups differed from each other, we need to use a post hoc test.

In the case of one-way classification, the analysis of variance can be computed as follow: Set up the hypotheses:

Null hypothesis: $H_0: \mu_1 = \mu_2 = \mu_3 = \dots = \mu_k$ (ie. the arithmetic mean of the population from which samples have been taken, is the same and there is no difference.)

Alternative hypothesis: $H_a: \mu_1 \neq \mu_2 \neq \mu_3 \neq \dots = \mu_k$

(a) Variance between samples

(1) The mean of all samples is calculated separately and denoted by $\bar{X}_1, \bar{X}_2, \bar{X}_3$

(ii) The grand mean ($\bar{\bar{X}}$) of all sample means is calculated as follows:

$$\bar{\bar{X}} = \frac{\bar{X}_1 + \bar{X}_2 + \bar{X}_3 + \dots + \bar{X}_k}{k}$$

Where k No. of samples

(iii) The deviations of samples mean from the grand mean are calculated and squared. The square of the deviation is multiplied by the number of items in the sample and the total of the product is known as Sum of squares between Samples (SSB). The following formula can be used:

$$SSB = n_1(\bar{X}_1 - \bar{\bar{X}})^2 + n_2(\bar{X}_2 - \bar{\bar{X}})^2 + n_3(\bar{X}_3 - \bar{\bar{X}})^2 + \dots + n_k(\bar{X}_k - \bar{\bar{X}})^2$$

(iv) The SSB is divided by degree of freedom and variance between samples is obtained. Following formula can be used to find degree of freedom:

$$v_1 \text{ or } df_1 = k-1$$

Where k No. of samples

(b) Variance Within Samples.

(1) Mean of each sample is calculated i.e. $\bar{X}_1, \bar{X}_2, \bar{X}_3$

(ii) Deviation of each unit of sample from respective mean of sample is calculated. These deviations are squared and summed. It can be done as follows:

$$SSW = \Sigma(X_1 - \bar{X}_1)^2 + \Sigma(X_2 - \bar{X}_2)^2 + \Sigma(X_3 - \bar{X}_3)^2 + \dots + \Sigma(X_k - \bar{X}_k)^2$$

(iii) The sum of squares within sample is divided by degree of freedom and Mean Sum of Squares within Sample (**MSW**) is obtained. The degree of freedom is calculated as follows:

$$v_2 \text{ or } df_2 = N - k$$

Where N = Total number of items in all the samples

k = No. of samples

n = No. of items in each sample

(c) Analysis of Variance Table: After following the above steps, the sources of variance, degree of freedom etc. are presented in the form of table as follows:

Analysis of Variance Table or ANOVA Table

	Sources of Variance	Degree of Freedom	Sum of Squares	Variance	Variance Ratio or F
(i)	Between Samples	k - 1	$\Sigma [n(\bar{X} - \bar{\bar{X}})^2]$ (SSB)	$SSB \div (k - 1) = (MSB)$	$\frac{MSB}{MSW}$
(ii)	Within Samples	N - k	$\Sigma (X - \bar{X}_k)^2$ (SSW)	$SSW \div (N - k) = (MSW)$	
(iii)	Total	N - 1	$\Sigma (X - \bar{\bar{X}})^2$ (SST)	$SST \div (N - 1)$	

(d) **Calculation of Variance Ratio or F Statistic** : To calculate the variance ratio or F statistic, we use the following formula :

$$F = \frac{\text{Variance between Samples}}{\text{Variance within Samples}} \text{ or } \frac{\text{Greater variance}}{\text{Smaller variance}}$$

(e) **Levels of Significance** : The table value of F is obtained according to different levels of significance but practically 5% or 1% levels of significance are used.

(f) **Table/Critical Value** : The table value of F is obtained from the table according to level of significance and degree of freedom. The degree of freedom of greater variance is known as ν_1 and that of smaller variance as ν_2 .

(g) **Interpretation of F** : The computed value of F is compared with the table value. If the calculated value of F is more than the table value, the null hypothesis is rejected and difference is regarded significant and otherwise null hypothesis is accepted.

Illustration: 1

To test the significance of the variations of the retail prices of a commodity in three Metro cities: Mumbai, Kolkata and Delhi, four shops were chosen at random from each city and prices observed in rupees were as follow:

Shops	Prices (in ₹)		
	Mumbai	Kolkata	Delhi
1	5	4	4
2	6	4	3
3	2	2	2
4	7	6	3

Using Excel, check that do the data indicate the prices in the three cities are significantly differ at 5% level of significance?

Excel Solution : Here no. of samples (k) = 3

Size of sample, $n_1 = 4, n_2 = 4, n_3 = 4$

Total no. of items = $N = n_1 + n_2 + n_3 = 12$

Level of Significance = 5% = 0.05 (*i.e.* Alpha)

Let us take the hypotheses that :

Null hypothesis : $H_0 : \mu_1 = \mu_2 = \mu_3$ (*i.e.*, the mean prices in the three cities are the same.)

Alternate hypothesis : $H_a : \mu_1 \neq \mu_2 \neq \mu_3$ (*i.e.*, the mean prices in the three cities are not the same.)

Calculate Mean of each Sample and Grand Mean.

Mumbai	Kolkata	Delhi
X_1	X_2	X_3
5	4	4
6	4	3
2	2	2
7	6	3
$\Sigma X_1 = 20$	$\Sigma X_2 = 16$	$\Sigma X_3 = 12$
$\bar{X}_1 = \frac{20}{4} = 5$	$\bar{X}_2 = \frac{16}{4} = 4$	$\bar{X}_3 = \frac{12}{4} = 3$

Grand Mean or = $\bar{\bar{X}} = \frac{5+4+3}{3} = 4$

Variance Between Samples

Sum of squares of the deviations between samples (SSB)

$$SSB = n_1 (\bar{X}_1 - \bar{\bar{X}})^2 + n_2 (\bar{X}_2 - \bar{\bar{X}})^2 + n_3 (\bar{X}_3 - \bar{\bar{X}})^2$$

$$SSB = 4(5 - 4)^2 + 4(4 - 4)^2 + 4(3 - 4)^2 = 8$$

$$\text{Degree of freedom, } v_1 = k - 1 = 4 - 1 = 3$$

$$\text{Variance between samples (MSB)} = \frac{SSB}{k - 1} = \frac{8}{2} = 4$$

Variance within samples

Mumbai			Kolkata			Delhi		
X_1	$(X_1 - \bar{X}_1)$	$(X_1 - \bar{X}_1)^2$	X_2	$(X_2 - \bar{X}_2)$	$(X_2 - \bar{X}_2)^2$	X_3	$(X_3 - \bar{X}_3)$	$(X_3 - \bar{X}_3)^2$
5	0	0	4	0	0	4	1	1
6	1	1	4	0	0	3	0	0
2	-3	9	2	-2	4	2	-1	1
7	2	4	6	2	4	3	0	0
ΣX_1 = 20	$\Sigma(X_1 - \bar{X}_1)$ = 0	$\Sigma(X_1 - \bar{X}_1)^2$ = 14	ΣX_2 = 16	$\Sigma(X_2 - \bar{X}_2)$ = 0	$\Sigma(X_2 - \bar{X}_2)^2$ = 8	ΣX_3 = 12	$\Sigma(X_3 - \bar{X}_3)$ = 0	$\Sigma(X_3 - \bar{X}_3)^2$ = 2

$$\bar{X}_1 = 5$$

$$\bar{X}_2 = 4$$

$$\bar{X}_3 = 3$$

$$\bar{X} = \frac{5 + 4 + 3}{3} = 4$$

Sum of squares of the deviations within samples (SSW)

$$SSW = \Sigma(X_1 - \bar{X}_1)^2 + \Sigma(X_2 - \bar{X}_2)^2 + \Sigma(X_3 - \bar{X}_3)^2 = 14 + 8 + 2 = 24$$

$$\text{Degree of freedom, } v_2 = N - k = 12 - 3 = 9$$

$$\text{Variance with samples (MSW)} = \frac{SSW}{N - k} = \frac{24}{9} = 2.67$$

ANOVA TABLE

Source of Variance	Sum of Square	Degree of Freedom	Mean Sum of Square	F Ratio
Between City	SSB = 8	$k - 1 = 2$	$\frac{SSB}{k - 1} = \frac{8}{2} = 4$	$F = \frac{4}{2.67} = 1.498$
Within City	SSW = 24	$N - k = 12 - 3 = 9$	$\frac{SSW}{N - k} = \frac{24}{9} = 2.67$	
Total Variance	TSS = 32	$N - 1 = 12 - 1 = 11$		

Table value of F at 5% level (for $v_1 = 2, v_2 = 9$) = 4.26

Interpretation : The computed value of F (1.498) is less than the table value of F (4.26) at 5% level of significance, hence the null hypothesis is accepted and concluded that the difference is not significant. It means the mean prices in the three cities are not significantly different.

• ANOVA using MS Excel:

One-way ANOVA using Excel : We will use the analysis ToolPak add-in to conduct a one-way ANOVA.

Steps to Follow

Under "Data", tab is the option "Data Analysis" present?

If Yes – Analysis ToolPak is installed. Then go further.

If No – Install Analysis ToolPak.

Installing Analysis ToolPak

To load the Analysis ToolPak add-in, execute the following steps.

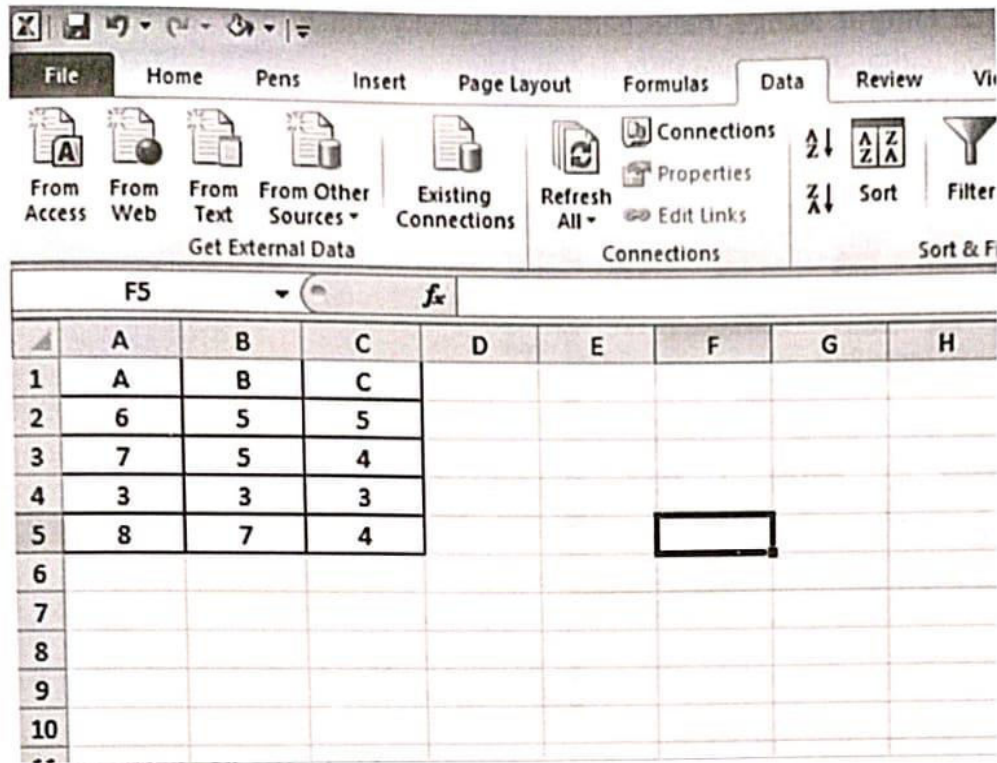
1. Click on the green **File** tab. The File tab in Excel 2010 replaces the **Office Button** (or File Menu) in previous versions of Excel.
2. Click on **Options**.
3. Under Add-ins, select **Analysis ToolPak** and click on the **Go** button.
4. Check Analysis ToolPak and click on OK.
5. Under the Data tab, we can now click on **Data Analysis**.

Excel Solution:

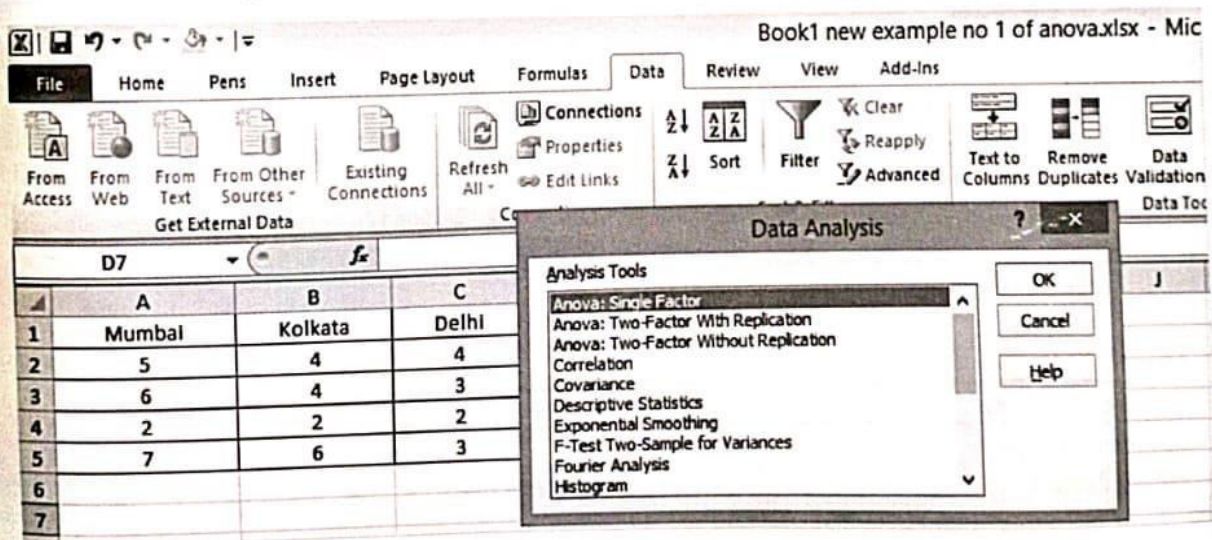
Open a worksheet and enter the data as shown below.

The screenshot shows the Microsoft Excel 2010 interface. The ribbon is set to the 'Data' tab, which includes options like 'Sort & Filter', 'Connections', and 'Refresh All'. Below the ribbon, a worksheet is visible with the following data:

	A	B	C	D	E	F	G
1	Mumbai	Kolkata	Delhi				
2	5	4	4				
3	6	4	3				
4	2	2	2				
5	7	6	3				
6							



Under the **Data** tab, click on **Data Analysis**. A window named Data Analysis will open. Select **Anova : Single Factor** and click on **OK**.



A window named **Anova : Single Factor** will open.

In **Input Range** box, select the range *i.e.*, A1 : C5.

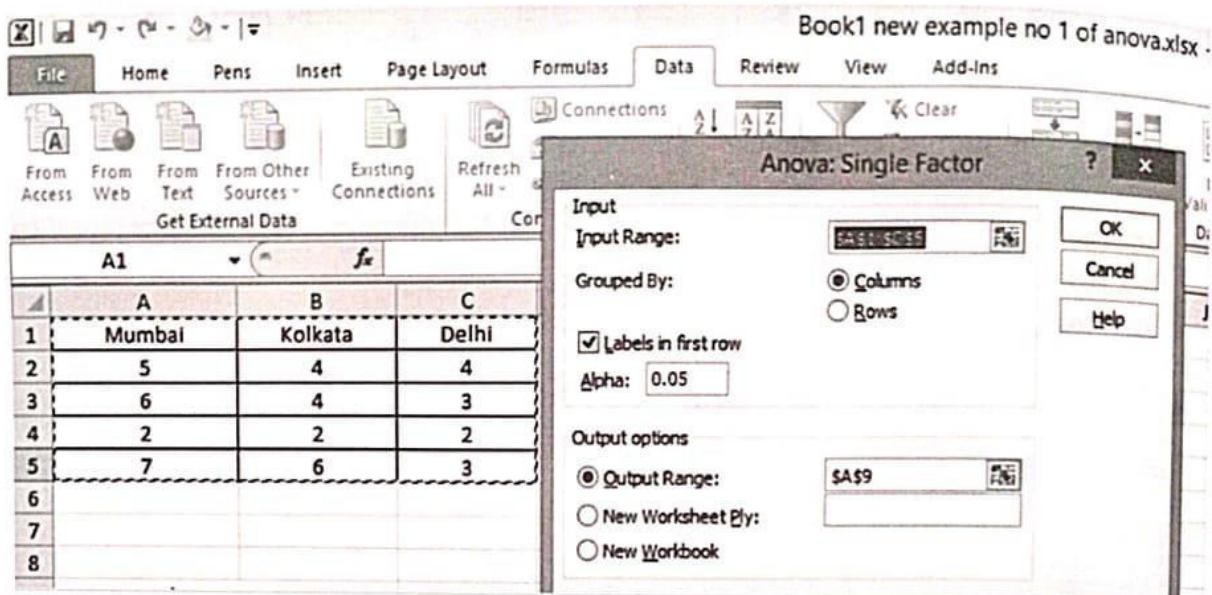
In **Grouped by** option, click on columns radio button, because our data are in columns.

Check the box, **labels in first row**.

In **Alpha** box type 0.05. (*i.e.*, Level of significance)

Click on **Output Range**, radio button. Select any cell in box(*i.e.*, A9) where we want results.

Click on **OK**.



We will get the **results** as shown below in the form of ANOVA Table.

Groups	Count	Sum	Average	Variance
Mumbai	4	20	5	4.666667
Kolkata	4	16	4	2.666667
Delhi	4	12	3	0.666667

Source of Variation	SS	df	MS	F	P-value	F crit
Between Groups	8	2	4	1.5	0.274016	4.256495
Within Groups	24	9	2.666667			
Total	32	11				

Interpretation : The computed value of F (1.5) is less than the table value of F or F crit (critical) (4.256495) at 5% level of significance, and according to *p*-value approach, *p*-value is also more than 0.05, hence the null hypothesis is accepted and concluded that the mean prices in the three metro cities are not significantly different.

Illustration : 2

Four varieties A, B, C and D of wheat are shown in five plots each and the following yields per acre were obtained:

Plots	Varieties			
	A	B	C	D
1	32	34	31	29
2	34	33	34	26
3	33	36	35	30
4	35	37	32	28
5	37	35	36	29

Using MS Excel, test the significance of difference between the yields of the four varieties at 5% level of significance.

Excel Solution : Let us take the hypotheses.

Null Hypothesis: $H_0 : \mu_1 = \mu_2 = \mu_3 = \mu_4$ (The yields of four varieties are same)

Alternate Hypothesis: $H_a : \mu_1 \neq \mu_2 \neq \mu_3 \neq \mu_4$ (i.e. atleast one of the means is different)

Level of significance = 5% = 0.05 (i.e. Alpha)

Open a worksheet and enter the data as shown below and follow the steps as stated in **Illustration No. 1:**

The screenshot shows the Microsoft Excel interface with the following data table:

	A	B	C	D	E
1	Plots	A	B	C	D
2	1	32	34	31	29
3	2	34	33	34	26
4	3	33	36	35	30
5	4	35	37	32	28
6	5	37	35	36	29

The ANOVA: Single Factor dialog box is open with the following settings:

- Input Range: \$B\$1:\$E\$6
- Grouped By: Columns
- Labels in first row
- Alpha: 0.05
- Output options: Output Range: \$A\$9

Results are in the form of ANOVA Table.

Groups	Count	Sum	Average	Variance
A	5	171	34.2	3.7
B	5	175	35	2.5
C	5	168	33.6	4.3
D	5	142	28.4	2.3

Source of Variation	SS	df	MS	F	P-value	F crit
Between Groups	134	3	44.66667	13.95833333	9.89E-05	3.238872
Within Groups	51.2	16	3.2			
Total	185.2	19				

Interpretation : The calculated value of F (13.9583333) is less than table value of F or F crit (critical) (13.95833) and according to *p*-value approach, *p*-value is also less than 0.05; hence difference is significant and null hypothesis is rejected. It means that yields of four varieties are not same.

Illustration : 3

Units produced by three different plants of a company are as under:

Plant	Units manufactured								
	A	84	60	40	47	34			
B	67	92	95	40	98	60	59	108	86
C	46	93	100						

Using MS Excel, test that is there any significant difference in the production of three plants at 5% level of significance?

Excel Solution: Set up the hypothesis :

Null hypothesis : $H_0 : \mu_1 = \mu_2 = \mu_3$ (i.e. there is no difference between the production of three plants.)

Alternative hypothesis : $H_a : \mu_1 \neq \mu_2 \neq \mu_3$

Level of Significance = 5% = 0.05 (i.e. Alpha)

Open a worksheet and follow the steps as stated in previous Illustrations.

But, in **Grouped by**, click on **Rows**, radio button. (Because our data are in rows)

The screenshot shows the Excel interface with the 'Data' tab selected. The ribbon includes 'Sort & Filter' and 'Connections'. The data table is as follows:

	A	B	C	D	E	F	G	H	I	J	K
1	A	84	60	40	47	34					
2	B	67	92	95	40	98	60	59	108	86	
3	C	46	93	100							

The 'Anova: Single Factor' dialog box is open, showing the following settings:

- Input Range:
- Grouped By: Columns, Rows
- Labels in first column
- Alpha:
- Output options: Output Range: , New Worksheet Ply:, New Workbook

Results are in the form of ANOVA Table.

The screenshot shows an Excel spreadsheet with the following data:

Groups	Count	Sum	Average	Variance
A	5	265	53	394
B	9	705	78.33333	512.25
C	3	239	79.66667	862.3333

Source of Variation	SS	df	MS	F	P-value	F crit
Between Groups	2329.098	2	1164.549	2.203598	0.147225	3.738892
Within Groups	7398.667	14	528.4762			
Total	9727.765	16				

Interpretation : The computed value of F (2.203598) is less than critical value of F (3.738892) and according to P -value approach, p -value is more than 0.05, therefore, the null hypothesis is accepted and concluded that the difference is not significant. It means productions of all the three plants are equal.

B. TWO-WAY ANOVA

One way anova compares three or more levels of one factor. But some experiments involve two factors. When it is believed that two independent factors might have an effect on the response variable of interest, it is possible to design the test so that an analysis of variance can be used to test for the effects of the two factors simultaneously. Such a test is called two-way anova. Let us suppose that the Human Resources Department of a company desires to know if occupational stress varies according to age and gender. The variable of interest is, therefore occupational stress as measured by a scale.

There are two factors being studied - age and gender.

Further suppose that the employees have been classified into three groups or levels:

(i) Age less than 40 Years (ii) 40 to 55 Years (iii) above 55 Years

In addition employees have been labelled into gender classification (levels):

(i) Male (ii) female

In **two way ANOVA**, the analysis of variance is studied in following three parts:

- (i) Sum of Squares Between Columns (SSC)
- (ii) Sum of Squares Between Rows (SSR)
- (iii) Residual Variation (SSE)

Total Sum of Squares = Sum of Squares between Columns + Sum of Squares between
Rows + Residual Variation

or
$$TSS = SSC + SSR + SSE$$

Following steps should be followed to calculate Variance Ratio (F) in the case of Two Way ANOVA.

Coding method can be used to simplify the calculations.

- (i) Calculation of Correction Factor

$$\text{Correction Factor (c.f.)} = \frac{T^2}{N}$$

- (ii) **Total Sum of Squares (TSS)**: It is obtained by subtracting correction factor from the total of squared values of the sample, *i.e.*

$$TSS = \Sigma X_1^2 + \Sigma X_2^2 + \Sigma X_3^2 + \dots - \frac{T^2}{N}$$

- (iii) **Sum of Squares Between Columns (SSC)** : The total of each column is squared and divided by the number of items in respective columns. The correction factor is subtracted from the total of thus arrived values and SSC is obtained:

$$SSC = \Sigma \left\{ \frac{\Sigma(X_c)^2}{n_c} \right\} - \frac{T^2}{N}$$

Where ΣX_c^2 = Total of squared Values in each Column

n_c = Number of items in each column

- (iv) **Sum of Squares between Rows (SSR)** : The total of the sample values in each row is squared and divided by the number of items in the respective row. From the total of the values thus arrived correction factor is deducted and remaining is known as sum of squares between rows or SSR.

$$SSR = \Sigma \left[\left(\frac{\Sigma X_R}{n_R} \right)^2 \right] - \frac{T^2}{N}$$

Where,

ΣX_R = Sum of the squared value of each row

n_R = No. of items in each row.

- (v) **Sum of Squares of the Residual (SSE)** : The sum of the squares of the residual is obtained by deducting the sum of squares between the columns and the sum of squares between the rows from the total sum of squares:

$$SSE = TSS - (SSC + SSR)$$

- (vi) **Number of Degree of Freedom** : It is calculated as follows:

No. of degree of freedom between columns = $(c - 1)$

No. of degree of freedom between rows = $(r - 1)$

No. of degree of freedom for residual = $(c - 1)(r - 1)$

Total No. of degree of freedom = $N - 1$ or $cr - 1$

Where r refers to numbers of rows

c refers to number of columns

N refers to total of items in the samples

- (vii) **ANOVA Table** : In a two way anova the analysis of variance table is prepared in the following form:

Source of Variance	Sum of Squares	Degrees of Freedom	Mean Sum of Squares (MSS)	F Ratio
Between columns	SSC	$c - 1$	$SSC \div (c - 1) = MSC$	$F = \frac{MSC}{MSE}$
Between Rows	SSR	$r - 1$	$SSR \div (r - 1) = MSR$	$F = \frac{MSR}{MSE}$
Residuals	SSE	$(c - 1)(r - 1)$	$SSE \div (c - 1) \times (r - 1) = MSE$	
Total	TSS	$N - 1$ or $cr - 1$		

Interpretation : The calculated value of F is compared with the table value and if the calculated or computed value of F is greater than the table value at a specific level of significance, the null hypothesis is rejected and concluded that the difference is significant otherwise *vice-versa*.

Illustration : 4

The following table gives the average monthly sale of four salesmen in the three different types of zones:

Zone	Salesmen				Total
	A	B	C	D	
X	5	4	4	7	20
Y	7	8	5	4	24
Z	9	6	6	7	28
Total	21	18	15	18	72

Using Excel, make an analysis of variance table from the above facts at 5% level of significance and give your comments.

Excel Solution:

Let us make the Hypothesis that:

- (i) The mean sales for three different zones are the same.
- (ii) Four salesmen do not differ with respect to their mean sales.

Level of Significance = 5% = 0.05 (i.e., Alpha)

Applying the two way Anova without replication using excel.

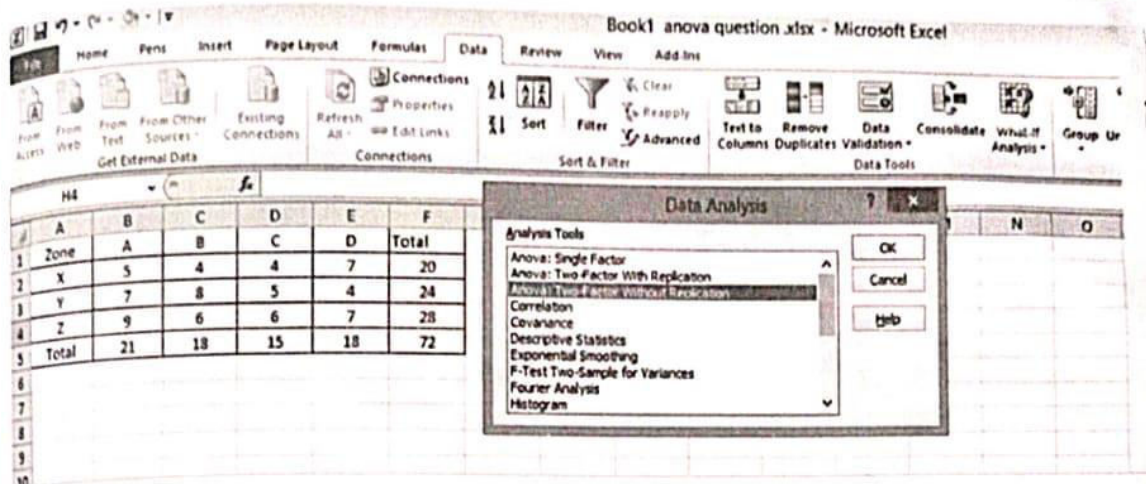
Open a worksheet and enter the data as shown below.

Under **Data** tab, click on **Data Analysis**.

A window named **Data Analysis** will open.

In Data Analysis window, select **Anova: Two Factor Without Replication**.

Click on **OK**.

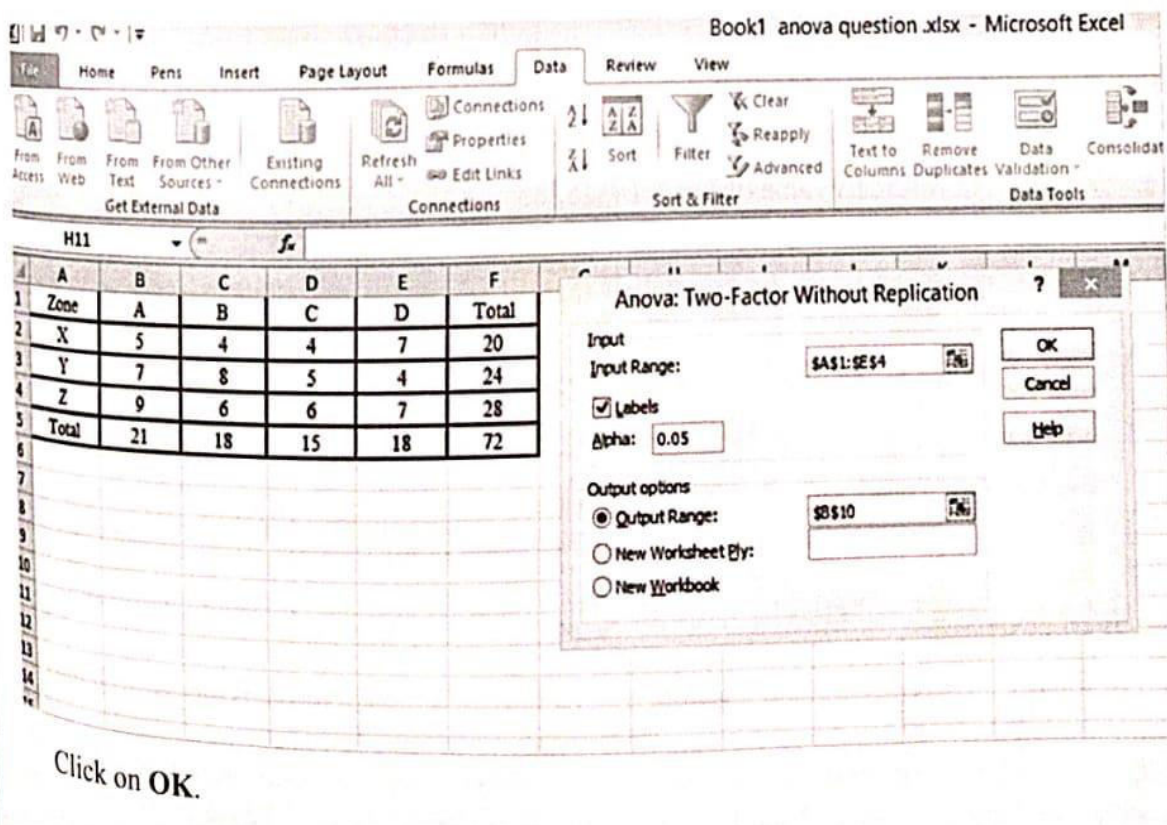


Anova : Two-Factor Without Replication window will open. In **Input Range** box, select range *i.e.* A1 : E4,

Check the **labels** box.

In **Alpha** box, type 0.05 (*i.e.* level of significance)

In **Output options**, click on **Output Range** radio button and select any cell (*e.g.* B10) where we want results.



Click on **OK**.

BOOK1: anova question .xlsx - MICROSOFT EXCEL

File Home Insert Page Layout Formulas Data Review View

Clipboard Font Alignment Number Styles

Calibri 11 A A

Wrap Text

General

Conditional Formatting as Table Styles

Format Painter Merge & Center

J18

	A	B	C	D	E	F	G	H	I
8									
9									
10		Anova: Two-Factor Without Replication							
11									
12		SUMMARY	Count	Sum	Average	Variance			
13		X	4	20	5	2			
14		Y	4	24	6	3.333333333			
15		Z	4	28	7	2			
16									
17		A	3	21	7	4			
18		B	3	18	6	4			
19		C	3	15	5	1			
20		D	3	18	6	3			
21									
22									
23		ANOVA							
24		Source of Variation	SS	df	MS	F	P-value	F crit	
25		Rows	8	2	4	1.5	0.296296296	5.14325285	
26		Columns	6	3	2	0.75	0.56094764	4.757062663	
27		Error	16	6	2.666666667				
28									
29		Total	30	11					
30									
31									

Interpretation:

(a) **For Zones** : The calculated value of $F = 1.5$

Table/crit (critical) value of $F = 5.14325285$

Since, the calculated value of F is less than the table value of F , the null hypothesis is accepted and it can be concluded that all the zones are similar so far as sales is concerned.

(b) **For Salesmen** : The calculated value of $F = 0.75$

Table/crit value of $F = 4.757062663$

Since, the calculated value of F is less than the table value of F , the null hypothesis is accepted and it can be concluded that there is no difference in the sales of four salesmen.

Illustration : 5

A hospital doctor wished to compare the effectiveness of 4 brands of painkiller A, B, C and D. She arranged that when patients on a surgical ward requested painkillers they would be asked if their pain was mild, severe or very severe. The first patient who said mild would be given brand A, the second who said mild would be given brand B, the third brand C and the fourth

brand D. Painkillers would be allocated in the same way to the first four patients who said their pain was severe and to the first four patients who said their pain was very severe. The patients were then asked to record the time, in minutes, for which the painkillers were effective. The following data were collected.

Brand	A	B	C	D
Mild	165	214	173	155
Severe	193	292	142	211
Very severe	107	110	193	212

Using Excel carry out a two factor analysis of variance and test at the 5% significance level for differences between brands and between symptoms.

Excel Solution:

Let us take the hypothesis that there is no significant difference between the brand and symptoms of pain.

The above data are classified according to criteria (i) brand and (ii) symptoms of pain.

Level of Significance (i.e. Alpha) = 5% = 0.05

Applying the **Anova two-way without replication** using Excel.

Open a worksheet in Excel and enter the data as shown below. Under **Data** tab, click on **Data Analysis**.

A window named **Data Analysis** will open.

In Data Analysis window select **Anova: Two Factor without Replication**.

Click on **OK**.

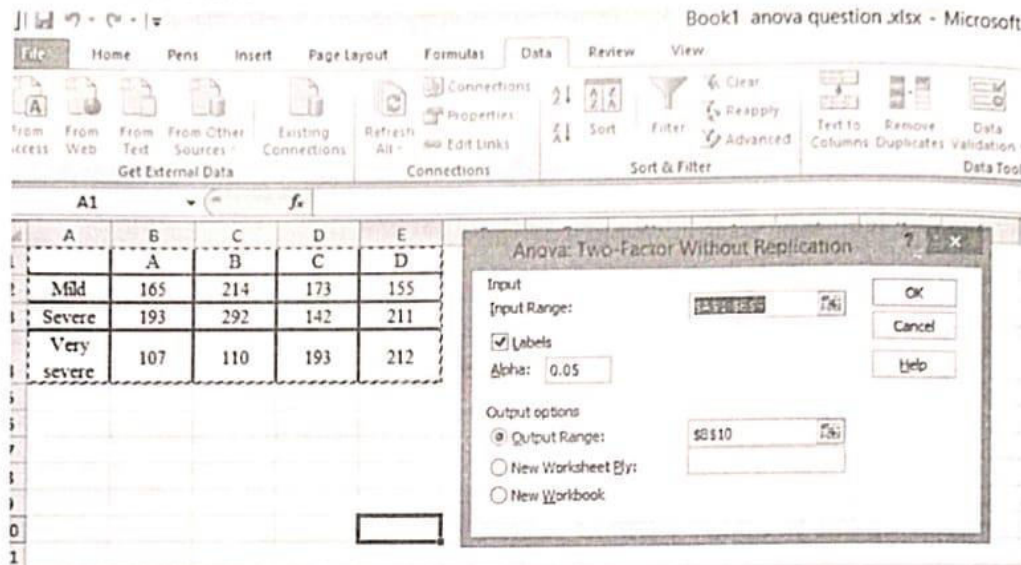
The screenshot shows the Microsoft Excel interface. The 'Data' tab is active, and the 'Data Analysis' tool is open. The 'Data Analysis' window is titled 'Data Analysis' and contains a list of analysis tools. The tool 'Anova: Two-Factor Without Replication' is selected. The background shows a worksheet with the data from the table above entered into cells A1:D4.

A window named **Anova : Two factor without Replication** will open.

In **Input Range** box, select the data range that is A1 : E4.

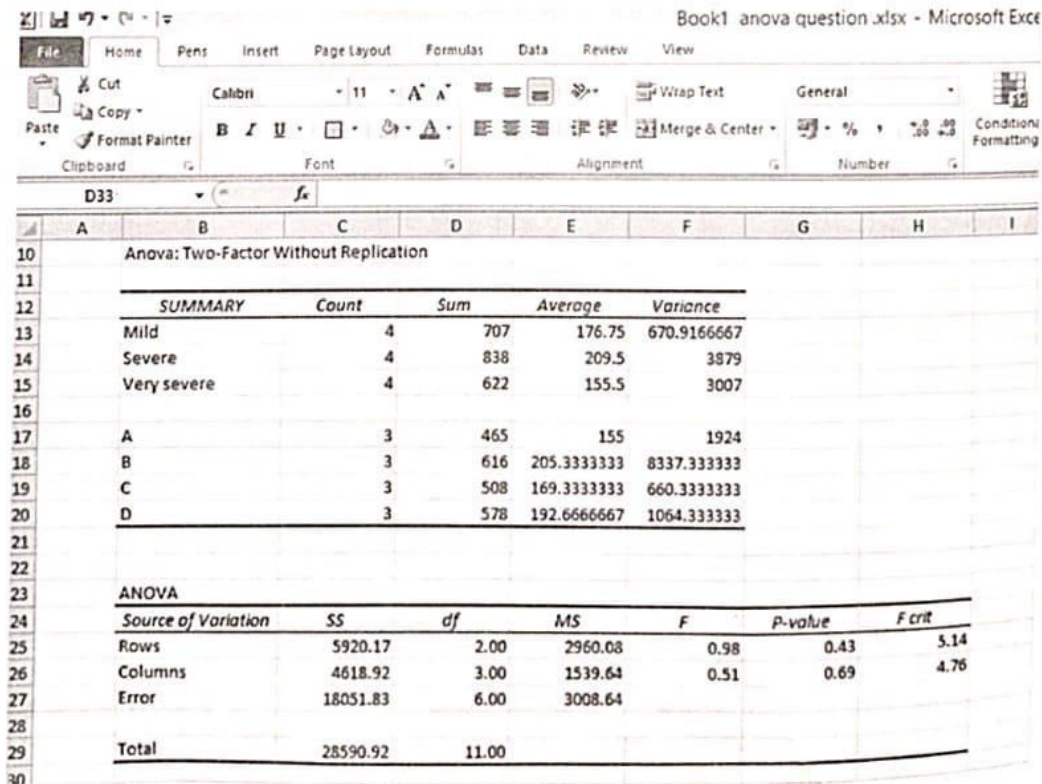
Check the **labels** box.

In **Output options**, click on **Output Range**, radio button and select the cell (e.g. B10) where we want the results.



Click on **OK**.

Results are as shown below:



Interpretation:

(a) **For Zones** : The calculated value of $F = 0.98$

Table/crit (critical) value of $F = 5.14$

Since the calculated value of F is less than the table value of F , the null hypothesis is accepted and it can be concluded that there is no significant difference in the brands as far as the pain symptoms are concerned.

(b) **For Salesmen** : The calculated Value of $F = 0.51$

Table/crit (critical) value of $F = 4.76$

Since the calculated value of F is less than the table value of F , the null hypothesis is accepted and it can be concluded that the brands and the pain symptoms are alike so far as effect is concerned.