CHAPTER 1

Introduction

In a business context, we frequently collect information with a variety of different objectives in mind. Some examples are:

- To understand our customers better – which products do they prefer to purchase?
- To make decisions about the size of our business – should we expand our sales team?
- To analyse financial information – by how much did our market share increase last year?
- To plan for the future – which location would be suitable for our new manufacturing facility?

However, business-related information, also referred to as **data**, has no practical use unless it can be properly processed and understood. We need statistics to help us achieve this understanding: statistics is a mathematical science that involves the collection, analysis, interpretation and presentation of data.

KEY TERMS

data
data set
descriptive statistics
discrete
continuous
inferential statistics
observation
qualitative data
quantitative data
variable

Collection

In collecting information, we need to ensure that the methods used are efficient in terms of both time and expense. If we choose to focus on a small selection of people or objects then it is important that this group is representative of the entire set of people or objects that we could investigate.

Chapter 2 describes some data collection techniques that can be used when people are involved in an investigation. It also provides a comparison of sampling methods to help us choose people to be included in our group of interest, and it explains how bias can occur when we ask people for information.

Analysis

Throughout the next six chapters, we develop a wide variety of processes that allow us to analyse our information more effectively using three main forms: tabular, graphical and numerical. As data are often collected in large quantities, it is necessary to condense and summarise the detail so that we can begin to understand its meaning. Collectively, these processes are called **descriptive statistics**.

Chapters 3 and 4 provide examples to illustrate the use of a range of tables, charts and graphs for summarising many different types of information. In Chapters 5 and 6 we focus on numerical methods, describing a 'typical value' and a measure of the variation within our data. When we are especially interested in analysing potential relationships between two data sets, correlation and regression techniques are used, as described in Chapters 7 and 8.

Interpretation

In the final two chapters of this book, we start to introduce some of the theory and techniques known as **inferential statistics**. Here we attempt to reach conclusions about the entire set of people or objects that we could investigate based on a smaller, representative group.

Covering the basics of probability, Chapter 9 lays the foundation of learning that is needed to understand inferential statistical methods. It explains techniques that can be applied in situations where there is uncertainty and allows us to assign a numerical value to the likelihood of something occurring. Chapter 10 extends our knowledge, providing instructions in using statistical tables to estimate numerical values that describe our collected data.

Presentation

When we have investigated our information thoroughly and discovered new facts, we often need to present these ideas to other people; this might involve writing a report, developing a website or giving a verbal presentation. It is important that we justify our choice of techniques and fully explain the way in which our conclusions have been reached.

Throughout this book, each chapter provides advice on how to present your information. For example, Chapter 4 lists some straightforward guidelines to ensure that graphs and charts are always well presented. Chapters 5 and 6 explain when it is appropriate to use each numerical measure, and Chapter 7 describes the limitations that should be considered before reaching your final conclusions.

Terminology

Before we can apply any of the techniques involved in descriptive and inferential statistics, we need to understand the meaning of some basic terminology.

Variable

A **variable** is a characteristic or an attribute that can have different values. When you are collecting information, each person or object might provide a different value for each variable.

Observation

The value of a variable for a specific person or object in the group that is of interest is known as an **observation**. When we have collected many observations for a range of variables, we describe this collection as a **data set**.

Quantitative and qualitative

Every variable, and therefore each observation, can be classified as quantitative or qualitative depending on the nature of the information. **Quantitative data** are information that can be counted or measured on a numerical scale. We say that quantitative data can be expressed numerically. **Qualitative data** represent a characteristic or an attribute that cannot be described using a numerical value. We use words or a single letter for qualitative data.

Discrete and continuous

For quantitative variables only, we can make a further distinction between discrete and continuous data. A **discrete** variable can assume only specific numerical values, whereas if a variable can assume any numerical value within a specific range, then it is known as **continuous**. Discrete data can be counted but continuous data are usually measured.

The table below demonstrates how to classify information using this terminology.

Variable	Examples of observations	Quantitative or qualitative	Discrete or continuous
number of employees working for a company	2350; 31; 175,000	quantitative	discrete
reason for customers making their purchase choice	brand; colour; functionality	qualitative	n/a
weight of a cereal packet in a production facility	500.5 g; 499.2 g; 502.3 g	quantitative	continuous

Whenever you collect data for an investigation, they should be classified immediately before the application of any statistical techniques. Deciding whether your data are quantitative or qualitative is an important, but often overlooked, first step: the nature of your information will determine which techniques are appropriate for use, and any conclusions reached at the end of your investigation will depend on correct classification.

Using Excel

Although Microsoft Excel is a general-purpose spreadsheet package, it provides a range of features that we can use to analyse and present our data.

In particular, Excel allows us to:

- construct formulae for performing calculations;
- use built-in functions;
- create graphs and charts to display data.

In this section, we will explain how to construct basic formulae and introduce some simple built-in mathematical functions. The use of statistical functions for measures of central tendency and dispersion, correlation and regression, and for graphical displays will be described at the end of each appropriate chapter.

The screen shots and instructions are associated with the use of MS Excel 2010. Earlier and more recent versions of the software package may have different menu options and may refer to different function names in comparison to the examples provided.

CONSTRUCTING BASIC FORMULAE

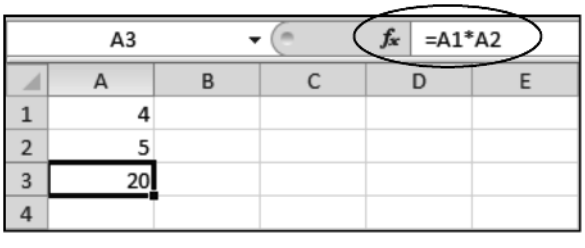
To construct a basic formula in Excel, click in the cell in which you want the result to appear and then start your formula with an equals sign (=); this indicates that the subsequent typed information should be regarded by the software package as a formula rather than ordinary text.

Once you have typed an equals sign, you will need to use an appropriate combination of numbers, cell references and operators to produce the calculation in which you are interested. Cell references can be typed, or alternatively you can click in the appropriate cell. When your formula is complete, press **<enter>** to see the result of the calculation.

The following tables show screenshots of the most common operators being used.

Operator	Description	Example: formula	Example: result
*	multiplication	A 1 =3*4 2	A B 1 12 2
		A 1 4 2 5 3 =A1*A2	A B 1 4 2 5 3 20

To view a completed formula, click in the cell which contains the result and the formula will be displayed in the **formula bar** at the top of the spreadsheet, as follows:

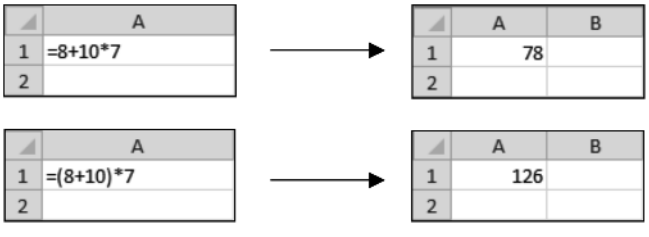


Operator	Description	Example: formula	Example: result
/	division	A 1 =18/6 2 A 1 50 2 10 3 =A1/A2	A B 1 3 2 A B 1 50 2 10 3 5

Operator	Description	Example: formula	Example: result
+	addition	A 1 =9+14 2 A 1 17 2 8 3 =A1+A2	A B 1 23 2 A B 1 17 2 8 3 25

Operator	Description	Example: formula	Example: result
–	subtraction	A 1 =19-3 2 A 1 4 2 15 3 =A1-A2	A B 1 16 2 A B 1 4 2 15 3 -11

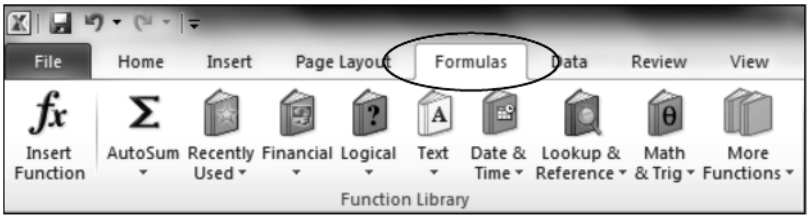
When you are constructing a formula, it is important to consider the order of precedence for the operators that you are using. Excel applies the usual order of precedence so that multiplication and division are performed before addition and subtraction. You will need to use brackets if a change to the order of precedence is required.



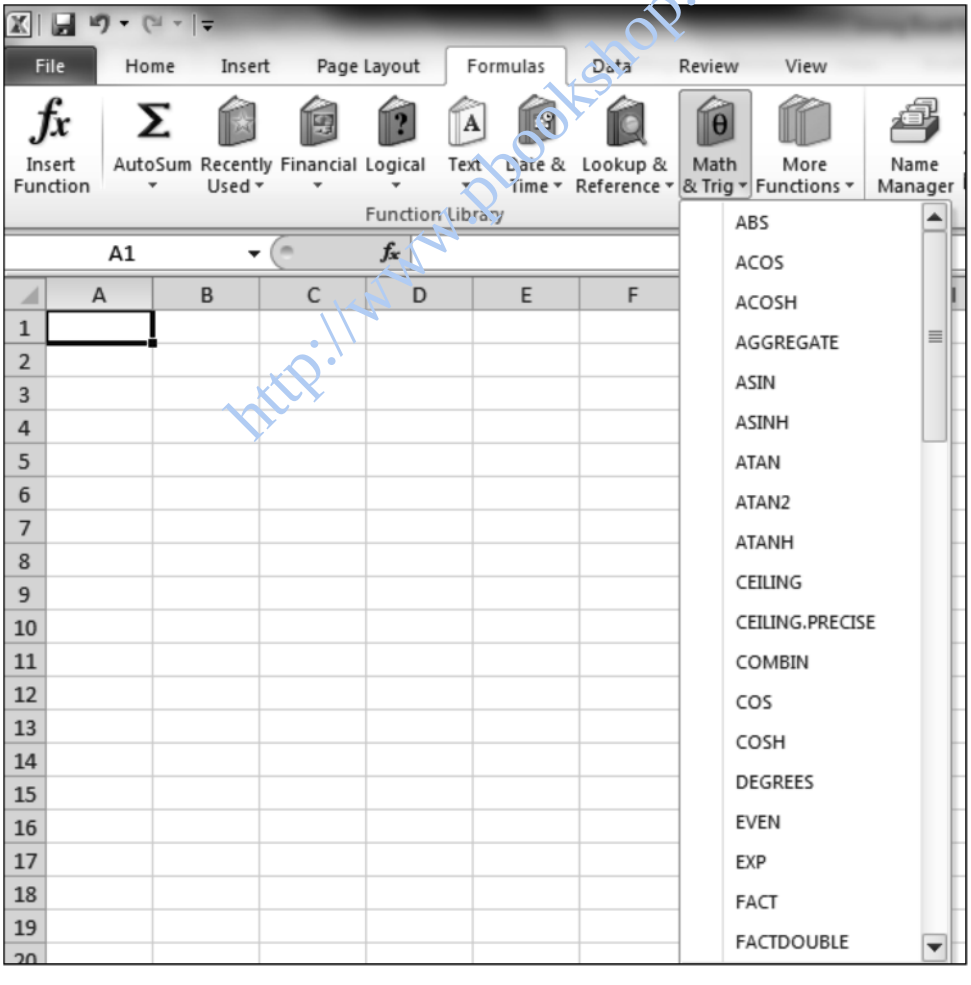
USING BUILT-IN FUNCTIONS

Excel provides a variety of built-in mathematical functions which can be used in addition to creating your own formulae using operators such as multiplication, division, addition and subtraction.

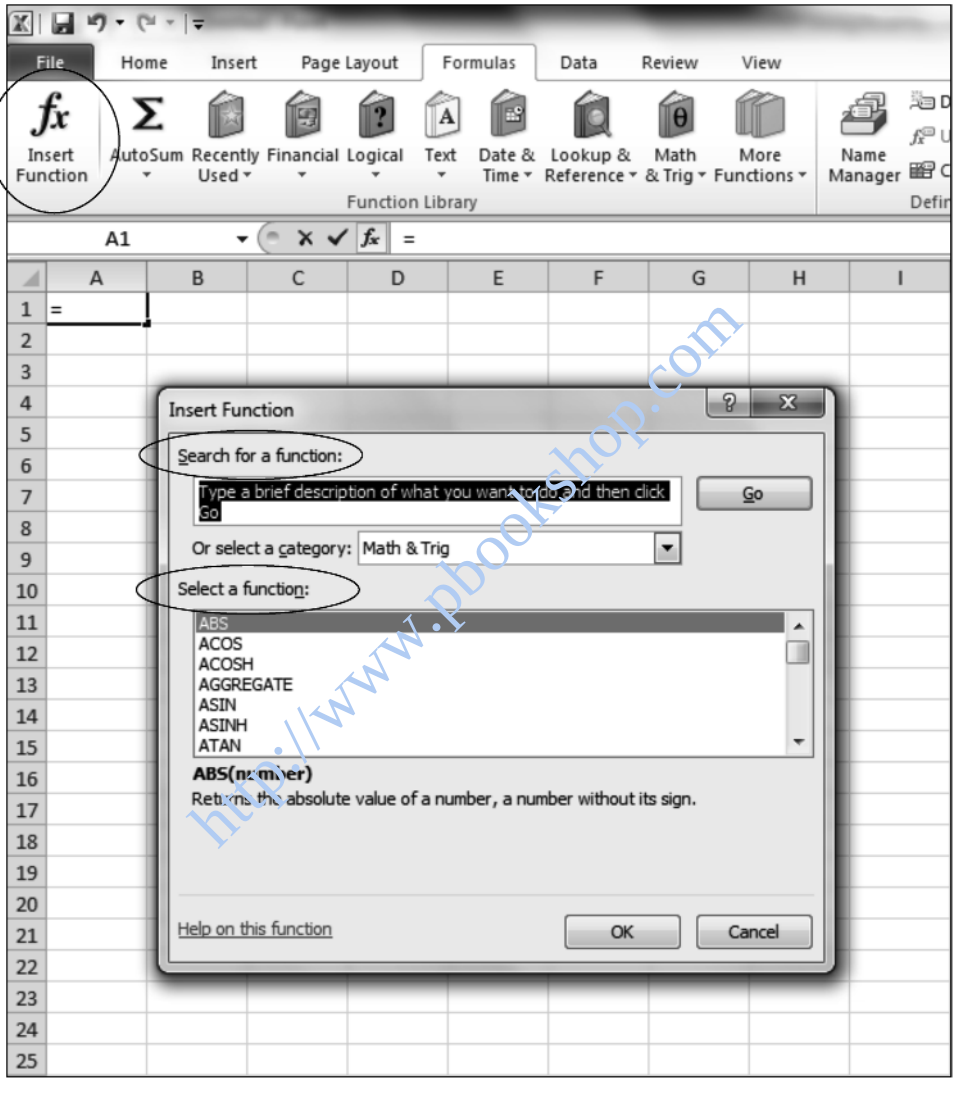
Using the **Formulas** tab on the ribbon, you can access built-in functions in two different ways, as described below.



- Click the cell in which you wish the result to appear and then click on **Math & Trig**. Scroll down the list until you find the function required.



- Alternatively, click in the cell in which you wish the result to appear and then click on **Insert Function**. Select the **Math & Trig** category. Either type some search text in the **Search for a function** box or use the **Select a function** box to find the function you require and click **OK**.

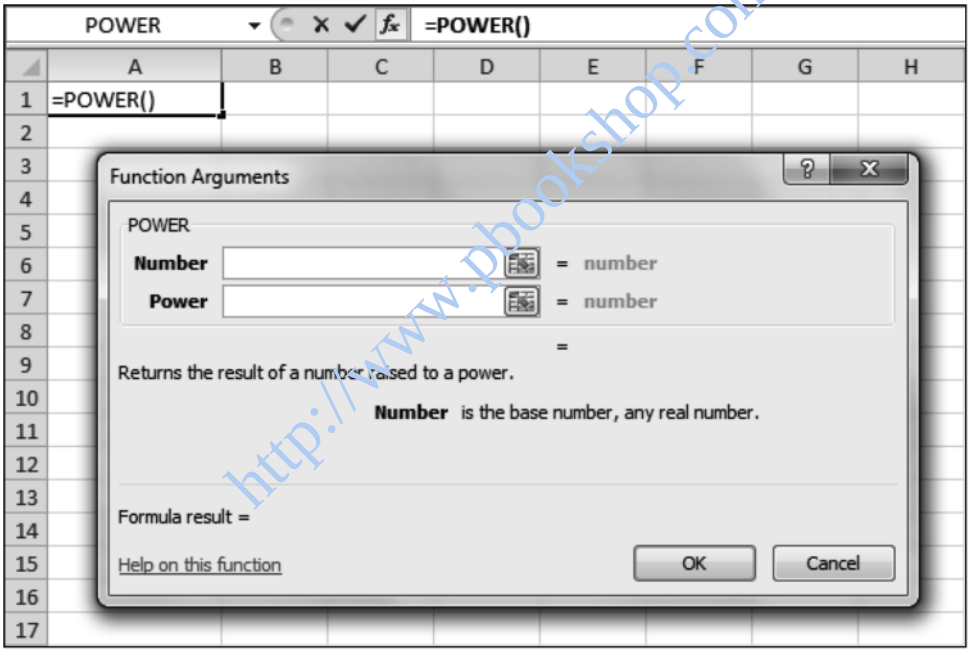


When you have selected the required function from the list, the **Function Arguments** dialog box will be displayed. This provides basic information, including a description of the value that the function will return and details of the arguments required for its operation.

Arguments in bold type must be entered; other arguments are optional and may contain assumed values. Once you have entered a combination of numbers and/or cell references as arguments, the dialog box will show you the result of the function. Finally, press **<enter>** or click **OK** to see the result of the calculation in the spreadsheet cell.

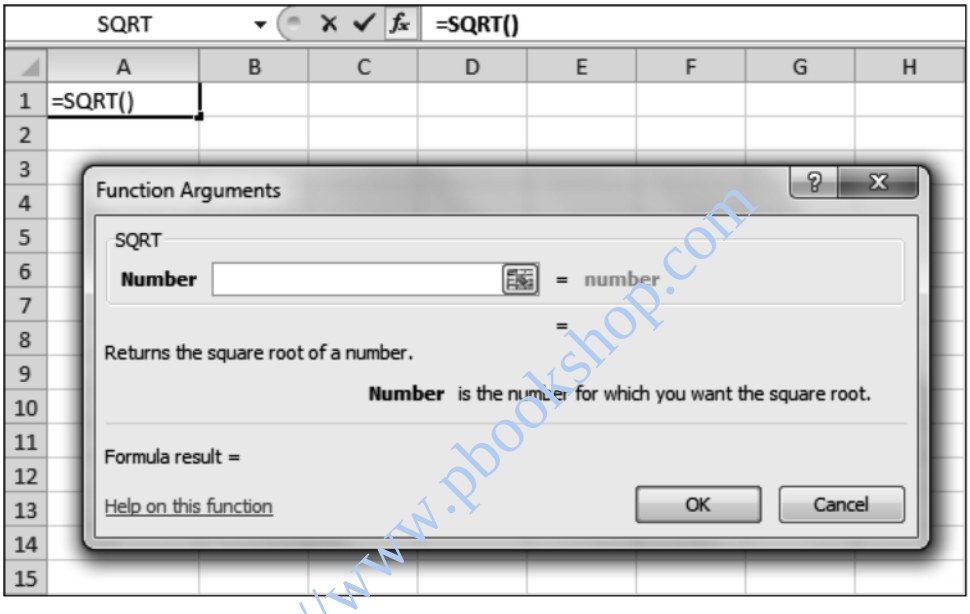
When constructing statistical formulae that include built-in functions, it is most likely that you will use **POWER**, **SQRT** and **SUM**. A description, screenshots of the **Function Arguments** dialog box and examples of use are shown below for each function each of these functions.

Function	Description	Syntax	Arguments
POWER	calculates the result of a number raised to a power	POWER(number, power)	**Number**: the base number which will be raised to the power **Power**: the exponent to which the base number will be raised



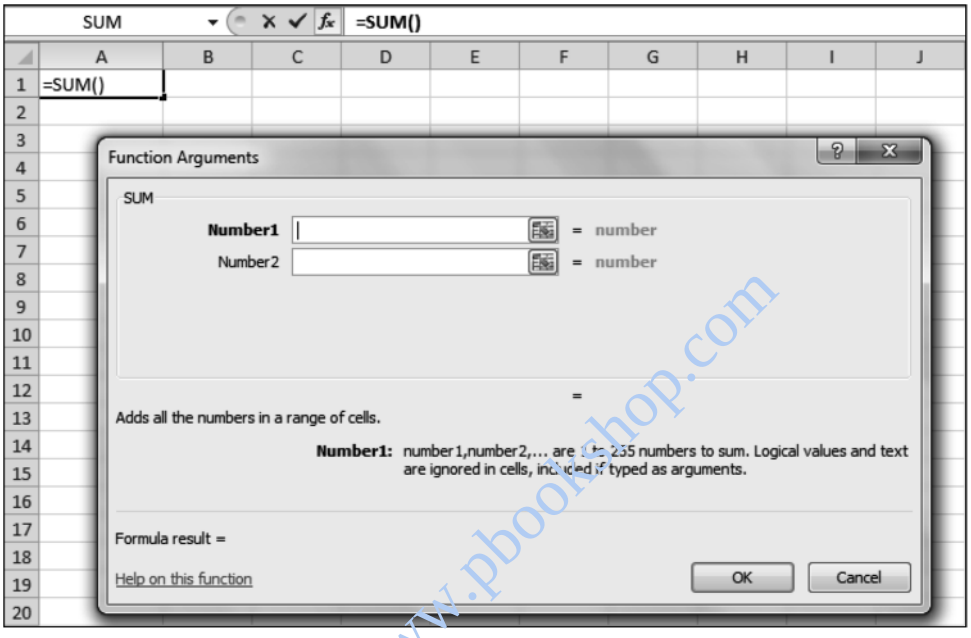
Example: function	Example: result
A 1 =POWER(5,3) 2	A B 1 125 2
A 1 10 2 2 3 =POWER(A1,A2)	A B 1 10 2 2 3 100

Function	Description	Syntax	Arguments
SQRT	calculates the square root of a number; an error will be returned **#NUM!** if a negative number is used	SQRT(number)	**Number**: the number for which the square root will be calculated



Example: function	Example: result
A 1 =SQRT(49) 2	A B 1 7 2
A 1 144 2 =SQRT(A1) 3	A B 1 144 2 12 3
A 1 -64 2 =SQRT(A1) 3	A B 1 -64 2 #NUM! 3

Function	Description	Syntax
SUM	calculates the total of all the numbers used as arguments	SUM(number1, number2 . . .)

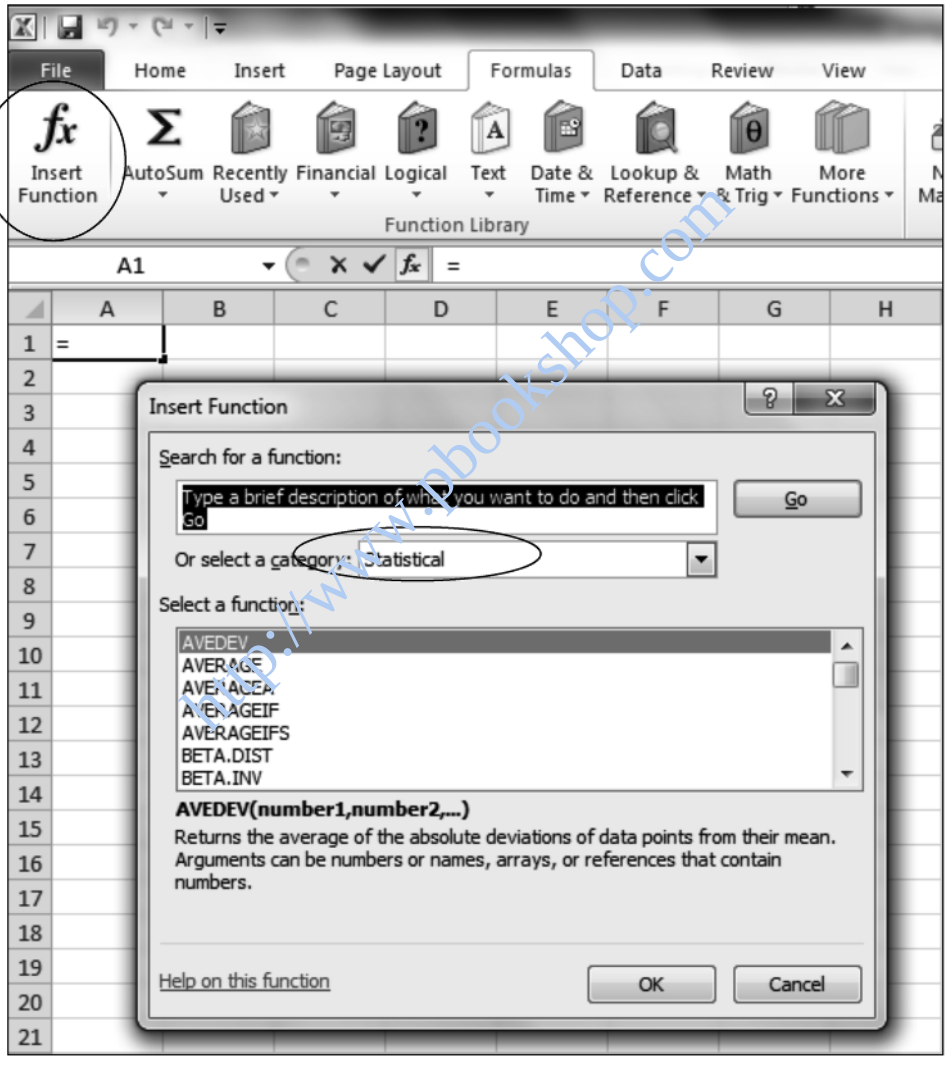


Example: function	Example: result
A 1 =SUM(19,24) 2	A B 1 43 2
A 1 8 2 17 3 3 4 27 5 =SUM(A1:A4)	A B 1 8 2 17 3 3 4 27 5 55

STATISTICAL FUNCTIONS

You can also use the **Formulas** tab to access a selection of statistical functions. As before, there are two methods:

- Click the cell in which you wish the result to appear and then click on **Insert Function**. Select the **Statistical** category. Either type some search text in the **Search for a function** box or use the **Select a function** box to find the function you require, and click **OK**.



- Alternatively, click the cell in which you wish the result to appear and then click on **More Functions**, and select **Statistical** from the drop-down menu. Scroll down the list until you find the function required.