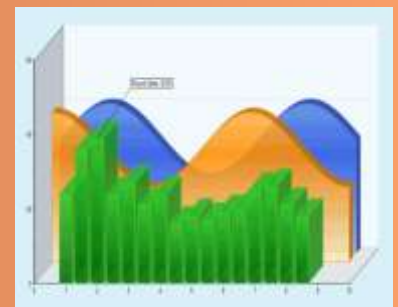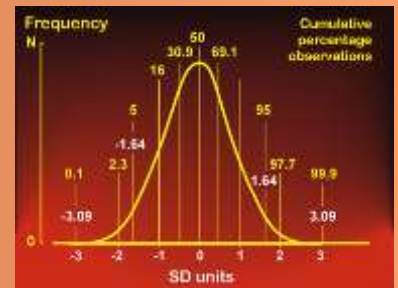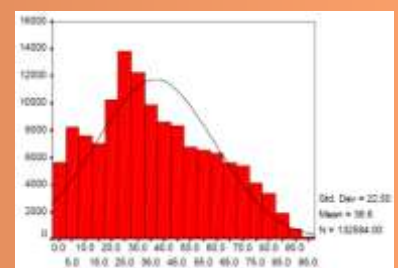# MEM12025A

Use graphical techniques and perform simple statistical computations

First Published January 2013

Edition 1 – January 2013

Conditions of Use:

Unit Resource Manual

Manufacturing Skills Australia Courses

This Student's Manual has been developed by BlackLine Design for use in the Manufacturing Skills Australia Courses.

All rights reserved. No part of this publication may be printed or transmitted in any form by any means without the explicit permission of the writer.

Statutory copyright restrictions apply to this material in digital and hard copy.

Copyright Ⓒ BlackLine Design 2013

Feedback:

Your feedback is essential for improving the quality of these manuals.

This resource has not been technically edited. Please advise the appropriate industry specialist of any changes, additions, deletions or anything else you believe would improve the quality of this Student Workbook.   Don't assume that someone else will do it.   Your comments can be made by photocopying the relevant pages and including your comments or suggestions.

Forward your comments to:

> BlackLine Design
>
> blakline@bigpond.net.au
>
> Sydney, NSW   2000

Aims of the Competency Unit:

This unit covers interpreting and constructing graphs and charts from given or determined data, and performing basic statistical calculations.

Graphs and charts may be applied to information from various work contexts, quality processes, production and market trends and other engineering applications. A range of devices may be used to assist with calculations.

Unit Hours:
18 Hours

Prerequisites:
MEM12024A Perform computations

Elements and Performance Criteria

1.	Read and construct graphs from given or determined data	1.1	Complex information is extracted from graphical representation.
		1.2	Data is analysed with respect to emerging trends.
		1.3	Graphs are constructed as required from data and drawn with respect to scale and accepted method.
		1.4	Significant features of graphical representation are understood such as limit lines, gradients (straight line graphs), intercepts, maximum and minimum values.
		1.5	A wide variety of graphs are constructed as required including histograms, control charts, straight line graphs and parabolic graphs.
2.	Perform basic statistical calculations	2.1	Mean, median and mode are calculated from given data.
		2.2	Standard deviation is calculated.
		2.3	Application of standard deviation and limits to process improvement techniques is understood.

Required Skills and Knowledge

Required skills include:

- obtaining required information by interpreting data presented in graphical form
- determining the trend(s) indicated by the data presented in graphical form
- constructing graphs to scale
- labelling the axes appropriately
- selecting scales appropriate to the purpose for which the graph is intended
- constructing histograms, control charts, straight line and parabolic graphs
- determining for a given set of data the mean, median and mode
- determining for a given set of data the standard deviation
- reading, interpreting and following information on written job instructions, specifications, standard operating procedures, charts, lists, drawings and other applicable reference documents
- planning and sequencing operations
- checking and clarifying task related information
- checking for conformance to specifications
- undertaking numerical operations, geometry and calculations/formulae within the scope of this unit

Required knowledge includes:

- characteristics of straight line, parabolic and hyperbolic curves
- procedures for determining the slope/rate of change of a curve
- the trend(s) indicated by changes in gradient of a graph
- procedures for drawing the line of best fit for the coordinates plotted
- standard form of equations relating to straight lines and parabolic curves
- gradient, intercepts, maximum and minimum values and limit lines for straight line and parabolic curves
- function of control charts
- the meaning of the terms mean, median and mode
- the meaning of the term standard deviation
- the significance of 1, 2 and 3 sigma limits
- safe work practices and procedures

Lesson Program:

The topics can be divided into the following program.

Topic	Skill Practice Exercise
Topic 1 – Statistical Data:	MEM-SP-0101 to MEM--SP-010
Topic 2 – Types of Charts:	MEM-SP-201 to MEM-SP-0210
Topic 3 – Histograms & Box Plots:	MEM-SP-301 to MEM-SP-0306
Topic 4 – Control Charts:	MEM-SP-401 to MEM-SP-0403
Topic 5 – Parabolic Charts:	MEM-SP-501 to MEM-SP-0502

Contents:

Topic 1 – Statistical Data:

Required Skills:

- Calculate various statistical data.
- .
- .

Required Knowledge:

-
-

Statistics:

Statistics is the study of the collection, organization, analysis, interpretation and presentation of data; it deals with all aspects of data, including the planning of data collection in terms of the design of surveys and experiments.

Some consider statistics a mathematical body of science that pertains to the collection, analysis, interpretation or explanation, and presentation of data, while others consider it a branch of mathematics concerned with collecting and interpreting data. Because of its empirical roots and its focus on applications, statistics is usually considered a distinct mathematical science rather than a branch of mathematics. Much of statistics is non-mathematical: ensuring that data collection is undertaken in a way that produces valid conclusions; coding and archiving data so that information is retained and made useful for international comparisons of official statistics; reporting of results and summarised data (tables and graphs) in ways comprehensible to those who must use them; implementing procedures that ensure the privacy of census information.

Statistics itself also provides tools for prediction and forecasting the use of data and statistical models. Statistics is applicable to a wide variety of disciplines, including natural and social sciences, government, and business. Statistical methods can summarize or describe a collection of data which is called descriptive statistics and is particularly useful in communicating the results of experiments and research. In addition, data patterns may be modelled in a way that accounts for randomness and uncertainty in the observations. These models can be used to draw inferences about the process or population under study, a practice called inferential statistics. Inference is a vital element of scientific advance, since it provides a way to draw conclusions from data that are subject to random variation. To prove the propositions being investigated further, the conclusions are tested as well, as part of the scientific method. Descriptive statistics and analysis of the new data tend to provide more information as to the truth of the proposition.

Data Types:

Four major data types whose more detailed descriptions can be found:

Quantitative:

Quantitative data (metric or continuous) is often referred to as the measurable data. This type of data allows statisticians to perform various arithmetic operations, such as addition and multiplication, to find parameters of a population like mean or variance. The observations represent counts or measurements, and thus all values are numerical. Each observation represents a characteristic of the individuals in a population or a sample.

Example:

A set containing the profits (or losses) of a number franchises operating under the name of a multinational company contains quantitative data. A possible data set for this companies are: Mother Flowers Sydney $260,000, Mother Flowers Melbourne $430,000, Mother Flowers Brisbane -$35,000, Mother Flowers Perth $132,000, Mother Flowers Adelaide $65,000.

Discrete:

Statistically speaking, discrete data result from either a finite or a countable infinity of possible options for the values present in a given discrete data set. The values of this data type can constitute a sequence of isolated or separated points on the real number line. Each observation of this data type can therefore take on a value from a discrete list of options.

The discrete data type usually represents a count of something; some examples of this type include the number of cars per crossing, a buildings height, the number of times an employee makes an error during a day, a number of defective lights on a production line, and a number of tosses of a coin before a head appears (which process could be infinite in length).

Discrete data can be divided into three types of data, Discrete Numerical Data, Discrete Ordinal Data and Discrete Qualative (Nominal) Data.

Discrete Numerical Data:

Discrete numerical data is discrete data that consists of numerical measurements or counts. A set that consists of discrete numerical data contains numbers. It is all quantitative data, which allows us to find population or sample parameters like mean, variance, and others. Discrete numerical data sets do not always consist of whole-numbers (or integers), but they may also take on the values of fractions and decimals.

Example:

A set containing the heights of shrubs in a building's landscaping, rounded to the nearest metre, represents such data type. It is very important to remember here to round the numbers up. If we do not, and we accept measurements such as 6.28 m, 6.42 m, 6.84 m, we will be dealing with continuous data type, described below. With discrete data type, there is a countable number of observations involved. For example, a set containing possible shrub heights will consist of integers starting at 0 in and ending at perhaps 7 m, unless there are shrubs over 7 metres tall. Integers are countable and that is what makes this set discrete.

Discrete Ordinal Data:

Discrete ordinal data is discrete data that may be arranged in some order or succession, but differences between values either cannot be determined or are meaningless. There are only relative comparisons made about the differences between the ordinal levels.

Example:

Consider the following statement: "In a group of twenty workers, five are "excellent," ten are "good," and five "need improvement." Although there are obvious differences between each category (excellent, good, and need improvement), and we can arrange them in order of worse to excellent or vice versa, there is not much more we can do to compare them. We do not know how much better is "excellent" from "good" or "good" from "need improvement." In this case, we could have also used numbers instead of words, ex. 1 for best, 2 for good, and 3 for need improvement, and the data type would still be ordinal. The numbers still lack any computational significance.

Discrete Qualitative (Nominal) Data:

Discrete qualitative (nominal) data is discrete data that cannot be arranged in any order. It can be represented by numbers, letters, words, and other forms of notation or symbolism, but there are no ranking differences to be determined. Each category or group will certainly be different from the others, but it will be equally significant. This data type will only consist of names, labels, or categories.

Example:
Gender, political parties, or religions are just some of many qualitative sets that exist in our modern society. Take, for example, these statements: "In a group of twenty workers, there are ten women and ten men," or "In a group of twenty workers, there are ten Liberals, eight Labour, and 2 Greens." The categories such as women and men, or republicans, democrats, and independents, can be talked about, described, and even criticized, but not officially ranked. There are no accepted schemes to put these categories in any meaningful order.

Continuous:
Continuous quantitative data result from infinitely many possible values that the observations in a set can take on. The term "infinitely," however, does not refer to the "countable" term we have seen with discrete data types. Continuous data types involve the uncountable or non-denumerable kind of infinity, which is frequently referred to as the number of points on a number line (or an interval on a number line). In other words, the observations of this data type can be associated with points on a number line, where any observation can take on any real-number value within a certain range or interval.

Example:
Temperature readings are one example of such data set.   Each reading can take on any real number value on a thermometer. If we agree that during a particular day the temperatures between 10:00 am and 6:00 pm will be somewhere between 0°C and 38°C, the truth is that these temperatures could take on any value in that range. For example, consider the following possible temperature readings given in degrees Celsius: 28.333..., 15.324, 30.23..., 12 or 8.

Another example will be a different approach to the Farmland Dairies ultra-pasteurized whole milk bottle example used with a description of the discrete numerical data. If, instead of measuring the number of bottles in different stores, we measure the amount of milk in each one half gallon bottle in different stores, those values could, for instance, be 0.498 gallon, or 0.5025 gallon, or any value in between. The observed values will be represented by real-line values, and there is an uncountable number of possibilities for that to occur.

Categorical:
Categorical data, also called qualitative or nominal, result from placing individuals into groups or categories.   The values of a categorical variable are labels for the categories. Both ordinal and qualitative categorical data types have been previously described.

Engineering Statistics:
Statistics are used to help to analyse and understand the performance and trends in various areas of work and may include financial trends, ideas concerned with the population, or mechanisms involved with manufacturing. The information may be presented visually with easy to understand graphics, graphs and charts. Typical charts used are the Pie Chart (Figure 1.1) and Bar Chart (Figure 1.2) which display the percentages on material types used by a company.
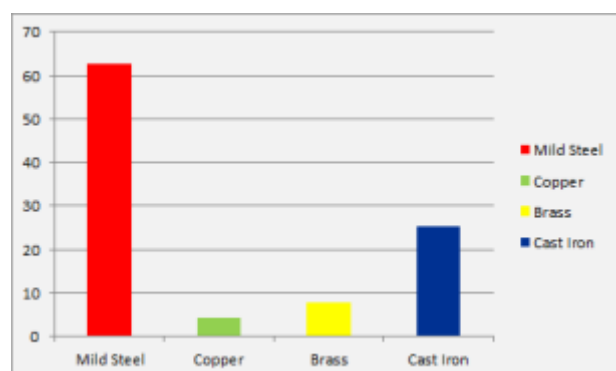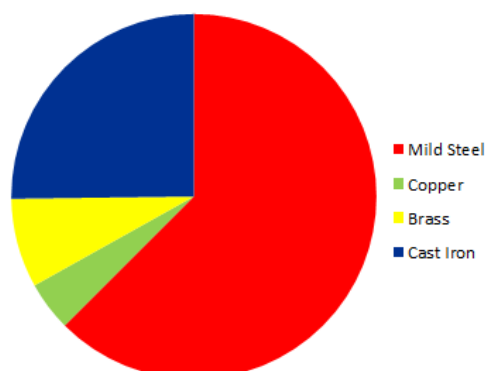

Figure 1.2

Figure 1.1

Raw Data:

Raw data in statistics is the details given by an investigator or collector from sources; it is the first hand information which has undergone no mathematical or statistical treatment.

Example:
A set of statistics was compiled for the lengths of timber measured to 0.01 m and used for a specific project and resulted in the following table. *This is the raw data*.

Sample	1	2	3	4	5	6	7	8	9
Length	1.45	1.56	1.37	1.44	1.32	1.42	1.55	1.29	1.37
Sample	10	11	12	13	14	15	16	17	18
Length	1.49	1.47	1.34	1.56	1.28	1.35	1.62	1.46	1.38

Ranked Data:

To rank a set of objects is to arrange them in order with respect to some characteristic. The set of objects could, for example, be a group of men and the characteristic could be height. When the men are arranged in order of height, the tallest is assigned the rank 1, the next tallest the rank 2, and so on. In this case the characteristic is a measurable one, and the ranking is merely a transformation of variables.

There is usually a distortion in such a transformation; thus, consider four men whose heights are 1.880 m; 1.829 m; 1.803 m; and 1.702 m respectively. The differences in height between consecutive men are 51 mm, 26 mm and 101 mm.

It is seen that in the case of a measurable characteristic rank is a rather rough way of assigning a numerical value to the degree in which the characteristic is possessed. However, there are certain advantages in using ranks. One of these is that the numbers involved in statistical computations and analyses are usually simpler. Another is that sometimes a set of numerical data will be dominated by one or two large items, whereas if the items are ranked the undue influence of these items are eliminated.

It is frequently possible to rank objects according to some characteristic which is difficult or even impossible to measure. Individuals can be ranked according to intelligence or personality, manufactured articles can be ranked according to beauty of design, and aircraft can be ranked according to performance or efficiency. Some of the characteristics just mentioned are too vague to allow of measurement, yet they do permit ranking.

Example:
Continuing with the example, the lengths have been ranked in order starting with the longest and ending with the shortest; the ranking can also be shortest to longest. *This is the ranked data*.

Sample	1	2	3	4	5	6	7	8	9
Length	1.28	1.29	1.32	1.34	1.34	1.35	1.37	1.38	1.42
Sample	10	11	12	13	14	15	16	17	18
Length	1.44	1.45	1.46	1.47	1.49	1.55	1.56	1.56	1.62

Data presented in this form is called Discrete Data because values can be jumped in steps, 0.01 m in the example.

Class or Bands:

If the exact heights were used as measurements, it is highly unlikely that two lengths of timber were exactly the same, so the values can be rounded off; in the example the 1.34 m lengths could have been measured as 1.342 m and 1.345 m.   The handling of large numbers of samples can result with huge lists of data which can be simplified by creating bands or classes within which the rounded measurements fall. The more the values are rounded off, the more likely of obtaining several similar values in a given class. The number of lengths of timber in the example is called the frequency.   Determining the number of lengths that fall in each class is then added to a table which may contain 5, 7, 9, 11,13, or 15 etc. classes.

Example:
Continuing with the example, it has been determined to use 5 classes. The result is a Frequency Distribution Table.

Height	1.2 – 1.3	1.3 – 1.4	1.4 – 1.5	1.5 – 1.6	1.6 – 1.7
Midpoint or Mark	1.25	1.35	1.45	1.55	1.65
Frequency	2	5	7	3	1

Graphs or Charts:

One goal of statistics is to present data in a meaningful way. It's one thing to see a list of data on a page; it's another to understand the trends and details of the data. Many times data sets involve millions (if not billions) of data values and maybe too many to print out in an article or report. One effective tool in the statistician's toolbox is to depict data by the use of a graph.

The adage states "a picture is worth a thousand words". The same thing could be said about a graph. Good graphs convey information quickly and easily to the user. Graphs highlight salient features of the data and they can show relationships that are not obvious from studying a list of numbers. Graphs can also provide a convenient way to compare different sets of data.

List of Common Graphs in Statistics

Different situations call for different types of graphs, and it helps to have a good knowledge of what graphs are available. Many times the type of data determines what graph is appropriate to use. Qualitative data, quantitative data and paired data each use different types of graphs. Seven of the most common graphs in statistics are the Bar Graph, Pie Chart, Histogram, Stem and Left Plot, Dot Plot, Scatter Plot and Time Series Graph.

Pareto Diagram or Bar Graph:
A bar graph contains a bar for each category of a set of qualitative data. The bars are arranged in order of frequency, so that more important categories are emphasized.

Pie Chart or Circle Graph:
A pie chart displays qualitative data in the form of a pie. Each slice of pie represents a different category.

Histogram:
A histogram in another kind of graph that uses bars in its display. This type of graph is used with quantitative data. Ranges of values, called classes, are listed at the bottom, and the classes with greater frequencies have taller bars.

Stem and Left Plot:
A stem and left plot breaks each value of a quantitative data set into two pieces, a stem, typically for the highest place value, and a leaf for the other place values. It provides a way to list all data values in a compact form.

Dot plot:
A dot plot is a hybrid between a histogram and a stem and leaf plot. Each quantitative data value becomes a dot or point that is placed above the appropriate class values.

Scatterplots:
A scatterplot displays data that is paired by using a horizontal axis (the x axis), and a vertical axis (the y axis). The statistical tools of correlation and regression are then used to show trends on the scatterplot.

Time-Series Graphs:
A time-series graph displays data at different points in time, so it is another kind of graph to be used for certain kinds of paired data. The horizontal axis shows the time and the vertical axis is for the data values. These kinds of graphs can be used to show trends as time progresses.

Be Creative:
What if none of the above graphs work for your data? Don't worry, although the above is a listing of some of the most popular graphs, it is not exhaustive as there are more specialized graphs that may work.

At times it might be considered that sometimes situations call for graphs that haven't been invented yet; there once was a time when no one used bar graphs, because they didn't exist. Now bar graphs are programmed into Excel and other spreadsheet programs, and many companies rely heavily upon them. If confronted with data that needs to be displayed, don't be afraid to use your imagination. Perhaps you'll think up a new way to help visualize data, and engineers, audiences, and students of the future will get to do homework problems based on your graph.

Mean or Average:
The Mean is one of the most commonly used statistics and is easy to calculate using a calculator, or computer. The Mean or Average is calculated by ***adding up all the values*** in a set of data, and then ***divide that sum by the number of values*** in the dataset.

Example:
Continuing with the example, the height is represented by the variable "*x*" and the frequency by "*f*".

Sample	1	2	3	4	5	6	7	8	9
Length	1.28	1.29	1.32	1.34	1.34	1.35	1.37	1.38	1.42
Sample	10	11	12	13	14	15	16	17	18
Length	1.44	1.45	1.46	1.47	1.49	1.55	1.56	1.56	1.62
Total =	25.72	Number of Values =		18				

The mean is denoted by $\overline{x}$; therefore = 25.72/18 = 1.429 m.

When large values and/or number of values are given in the dataset, a computer may be more advisable.

Using Excel:
When using Excel, the syntax for entering the formula to calculate the mean or average is:

=AVERAGE(CELL_REFERENCE:CELL_REFERENCE)

The equal sign (=) MUST be entered into the cell or the formula will revert to a label.

Example:
The average must be determined for the following numbers, 58, 67, 34, 19, 27, 59, 38, 89, 79, 91. The data can be entered in one or more columns as shown in

	A
1	58
2	67
3	34
4	19
5	27
6	59
7	38
8	89
9	78
10	91
11	=AVERAGE(A1:A10)
12	56

	A	B
1	58	59
2	67	38
3	34	89
4	19	78
5	27	91
6	=AVERAGE(A1:B5)	
7	56	

	A	B	C
1	58	19	38
2	67	27	89
3	34	59	78
4			91
5	=AVERAGE(A1:C4)		
6		56	

Figure 1.3 *Figure 1.4* *Figure 1.5*

Figure 1.3, Figure 1.4 and Figure 1.5 are extracts from an Excel worksheet. In Figure 1.3, the numbers are in the one column with the formula being used displayed in Cell A11; note the cell references A1 to A10 are indicated between the brackets within the formula and separated by a colon (:). The number 56 in Cell A12 is the number that would normally appear in Cell A11 if the Show Formula function was turned OFF.

In Figure 1.4, the data is evenly distributed in two columns (A & B) with the Average equation being shown in Cell A6 and includes cells ranging from A1 to B5. The number 56 in Cell A7 is the number that would normally appear in Cell A11 if the Show Formula function was turned OFF.

In Figure 1.5 the data is distributed over three columns (A, B & C) with the Average equation shown in Cell A5 and includes cells ranging from A1 to C4. Note there are three numbers in Columns A & B and four in Column C; although the formula includes the values in Cells A4 & B4 are included in the calculation, as there is no data in those cells they are not included in the values or number of values.

Median:

The Median is the middle number in a sorted list of numbers. To determine the median value in a sequence of numbers, the numbers must first be arranged in value order from lowest to highest. If there is an odd amount of numbers, the median value is the number that is in the middle, with the same amount of numbers below and above. If there is an even amount of numbers in the list, the middle pair must be determined, added together and divided by two to find the median value. The median can be used to determine an approximate average.

To find the median value in a list with an odd amount of numbers, one would find the number that is in the middle with an equal amount of numbers on either side of the median. To find the median, first arrange the numbers in order from lowest to highest and then arranged in numerical order.

Example:
In the example being used, the lengths of timber in a project were 1.45, 1.56, 1.37, 1.44, 1.32, 1.42, 1.55, 1.29, 1.37, 1.49, 1.47, 1.34, 1.56, 1.28, 1.35, 1.62, 1.46 & 1.38.

Arrange the numbers in numerical order:

1.28	1.29	1.32	1.34	1.34	1.35	1.37	1.38	1.42
1.44	1.45	1.46	1.47	1.49	1.55	1.56	1.56	1.62

As there are an even number of lengths, the middle two lengths are 1.42 and 1.44; therefore the median is determined as (1.42 + 1.44)/2 = 1.43

If the last length was not included in the sample lengths then the middle number would be 1.42 with eight values before and after.

Compare the medians to the average and they are similar, 1.42 & 1.43 to 1.429.

Using Excel:
When using Excel, the syntax for entering the formula to calculate the median is:

=MEDIAN(CELL_REFERENCE:CELL_REFERENCE)

Mode:

The mode of a set of data values is the value(s) that occurs most often.

The mode has applications in printing. For example, it is important to print more of the most popular books; because printing different books in equal numbers would cause a shortage of some books and an oversupply of others.

Likewise, the mode has applications in manufacturing. For example, it is important to manufacture more of the most popular fastenings; because manufacturing different fastenings in equal numbers would cause a shortage of some fasteners and an oversupply of others.

NB:
- It is possible for a set of data values to have more than one mode.
- If there are two data values that occur most frequently, we say that the set of data values is bimodal.
- If there is no data value or data values that occur most frequently, we say that the set of data values has no mode

Example:
Continuing with the example and using the table created for Classes and Bands, it can be seen that the 1.4 – 1.5 band contains a count of 7 therefore the 1.45 would become the mean.

Height	1.2 – 1.3	1.3 – 1.4	1.4 – 1.5	1.5 – 1.6	1.6 – 1.7
Midpoint or Mark	1.25	1.35	1.45	1.55	1.65
Frequency	2	5	7	3	1

Using Excel:
When using Excel, the syntax for entering the formula to calculate the mode is:

=MODE(CELL_REFERENCE:CELL_REFERENCE)

If there are no duplicates there is no mode so the error message #NUM is displayed.

Ogive & Quartiles:

There are several ways to measure the centre of a set of data. The mean, median, mode and midrange all have their advantages and limitations in expressing the middle of the data. Of all of these ways to find the average, the median is the most resistant to outliers; it marks the middle of the data in the sense that half of the data is less than the median.

There's no reason to stop at finding just the middle; to continue this process the median of the bottom half of our data can be calculated. One half of 50% is 25%, therefore half of a half, or one quarter, of the data would be below this and since a quarter of the original set is being dealt with, this median of the bottom half of the data is called the first quartile, and is denoted by Q1.

The Third Quartile

There is no reason why the bottom half of the data is looked at, instead the top half could have been looked at and performed the same steps as above. The median of this half, which is denoted by Q3 also splits the data set into quarters; however, this number denotes the top one quarter of the data. Thus three quarters of the data is below our number Q3. This is why the third quartile is called the Q3 (and explains the 3 in the notation.

Using Excel:

When using Excel, the syntax for entering the formula to calculate the mode is:

=QUARTILE(CELL_REFERENCE:CELL_REFERENCE,QUART)

Quart numbers refer to:
- 0   Minimum value.
- 1   First quartile (25%)
- 2   Median value (50%)
- 3   Third quartile (75%)
- 4   Maximum value (100%)

Tutorial 1-1:

The following results were recorded following a series of acoustic tests emanating from chipping machines (in Decibels):

100, 98, 121, 107, 105, 98, 102, 105, 104, 107, 104, 98, 99, 100, 101, 104, 99, 105, 107, 110, 104, 103, 106, 99, 104, 98, 107, 100, 111, 108, 102, 106, 99, 98, 101, 102.

Using the Excel program, determine the average, median, mode and quartiles of the acoustic results.

Solution:
1. Enter the Excel program.
2. As there are 36 results, they can be entered using 1 to 6 columns. *It is probably preferable to use 1 column until conversant with the program*.
3. In Cell A1 type **100** and then press the **Enter** key.
4. In Cell A2 type **98** and then press the **Enter** key.
5. Continue to Cell A36 entering the results as shown in each cell.
6. In Cell A37 type **Average** and then press the **Enter** key.
7. In Cell A38 type **Median** and then press the **Enter** key.
8. In Cell A39 type **Mode** and then press the **Enter** key.
9. In Cell A40 type **Quartile 1** and then press the **Enter** key.
10. In Cell A41 type **Quartile 2** and then press the **Enter** key.
11. In Cell A42 type **Quartile 3** and then press the **Enter** key.
12. In Cell A43 type **Quartile 4** and then press the **Enter** key.

Obtain the Average.
13. In Cell B37 type **=AVERAGE(A1:A36)** and then press the **Enter** key.
The number 103.89 is displayed.

Obtain the Median.
14. Activate Cell B38; click on the **Formula** tab in the Ribbon menu and then select **Insert Function**.
15. In the Insert Function dialog box expand the "Or select a category" box and select **Statistical**.
16. Scroll down the function options and click **Median** then click **OK**.
17. Click on Cell A36 and holding down the button drag the cursor to Cell A1 and then release the button.
18. Click **OK** and the result 104 is displayed.

Obtain the Mode.
19. Activate Cell B39 and then select **Insert Function** from the Ribbon menu.

20. Expand the function options and select **MODE.MULT** then click **OK**.
21. Click on Cell A1 and holding down the button drag the cursor to Cell A36 and then release the button.
22. Click **OK** and the result 98 is displayed.

Obtain the Quartile 1.
23. Activate Cell A40 and then select **Insert Function** from the Ribbon menu.
24. Scroll down the function options and click **QUARTILE.INC** then click **OK**.
25. Click on Cell A36 and holding down the button drag the cursor to Cell A1 and then release the button.
26. In the Quart box, type **1** and then click **OK** and the result 99.75 is displayed.

Obtain the Quartile 2.
27. Activate Cell A40 and then select **Insert Function** from the Ribbon menu.
28. In Cell B40 type **=QUARTILE(A1:A36,2)** and then press the **Enter** key.
29. The result 99.75 is displayed.

Obtain the Quartile 3.
30. Activate Cell A40 and then select **Insert Function** from the Ribbon menu.
31. In Cell B40 type **=QUARTILE(A1:A36,3)** and then press the **Enter** key.
32. The result 106.25 is displayed.

Obtain the Quartile 4.
33. Activate Cell A41 and then select **Insert Function** from the Ribbon menu.
34. Scroll down the function options and click **QUARTILE.INC** then click **OK**.
35. Click on Cell A36 and holding down the button drag the cursor to Cell A1 and then release the button.
36. In the Quart box, type **4** and then click **OK** and the result 121 is displayed.

Answers:

Average	= 103.89
Median	= 104
Mode	= 98
Quartile 1	= 99.75
Quartile 2	= 104
Quartile 3	= 106.25
Quartile 4	= 121

Standard Deviation:

Standard deviation is an important calculation for math and sciences, particularly for lab reports. Standard deviation usually is denoted by the lowercase Greek letter σ.

Standard deviation is the average or mean of all the averages for multiple sets of data. Scientists and statisticians use the standard deviation to determine how closely sets of data are to the mean of all the sets. Standard deviation is an easy calculation to perform. Many calculators have a standard deviation function, but can be the calculation can be performed by hand.

The two methods for calculating Standard Deviation are the Population Standard Deviation and the Sample Standard Deviation. If data is collected from all members of a population or set, the Population Standard Deviation is applied; if the data represents a sample of a larger population, the Sample Standard Deviation formula is applied. The equations/calculations are nearly the same, except the variance is divided by the number of data points (N) for the population standard deviation, but is divided by the number of data points minus one (N-1, degrees of freedom) for the sample standard deviation.

In general, if data that represents a larger set is being analysed, Sample Standard Deviation is used, however, if data from every member of a set is collected, Population Standard Deviation is applied.

Examples are:
- Population Standard Deviation - Analysing results of slump tests.
- Population Standard Deviation - Analysing age of respondents on a national census.
- Sample Standard Deviation - Analysing the effect of caffeine on reaction time on machinery operators aged 18-25.
- Sample Standard Deviation - Analysing the amount of contaminants in a reactor's heavy water or deuterium oxide.

Calculate the Sample Standard Deviation:
The following method can be applied to calculate the Sample Standard Deviation:
1. Calculate the mean or average of each data set. To do this, add up all the numbers in a data set and divide by the total number of pieces of data. For example, if you have found numbers in a data set, divide the sum by 4. This is the mean of the data set.
2. Subtract the deviance of each piece of data by subtracting the mean from each number. Note that the variance for each piece of data may be a positive or negative number.
3. Square each of the deviations.
4. Add up all of the squared deviations.
5. Divide this number by one less than the number of items in the data set. For example, if you had 4 numbers, divide by 3.
6. Calculate the square root of the resulting value. This is the sample standard deviation.

Tutorial 1-2:
The lengths of 20 machine screws are measured before being secured in an assembly, the following lengths were recorded:

9, 2, 5, 4, 12, 7, 8, 11, 9, 3, 7, 4, 12, 5, 4, 10, 9, 6, 9 & 4.

Calculate the sample standard deviation of the length of the machine screws.

Solution:
1. Calculate the mean of the data. Add up all the numbers and divide by the total number of data points.

    = (9+2+5+4+12+7+8+11+9+3+7+4+12+5+4+10+9+6+9+4)÷20

    = 7

2. Subtract the mean from each data point (*or the other way around, if preferred as the number will be squared, so it does not matter if it is positive or negative*).

$(9 - 7)^2 = (2)^2 = 4$	$(7 - 7)^2 = (0)^2 = 0$
$(2 - 7)^2 = (-5)2 = 25$	$(4 - 7)^2 = (-3)^2 = 9$
$(5 - 7)^2 = (-2)2 = 4$	$(12 - 7)^2 = (5)^2 = 25$
$(4 - 7)^2 = (-3)^2 = 9$	$(5 - 7)^2 = (-2)^2 = 4$
$(12 - 7)^2 = (5)^2 = 25$	$(4 - 7)^2 = (-3)^2 = 9$
$(7 - 7)^2 = (0)^2 = 0$	$(10 - 7)^2 = (3)^2 = 9$
$(8 - 7)^2 = (1)^2 = 1$	$(9 - 7)^2 = (2)^2 = 4$
$(11 - 7)^2 = (4)^2 = 16$	$(6 - 7)^2 = (1)^2 = 1$
$(9 - 7)^2 = (2)^2 = 4$	$(9 - 7)^2 = (2)^2 = 4$
$(3 - 7)^2 = (4)^2 = 16$	$(4 - 7)^2 = (-3)^2 = 9$

3. In Cell A1 type **100** and then press the **Enter** key.

In Cell A2

Skill Practice Exercises:

Skill Practice Exercise MEM-SP-0101

A company manufactures Ø25mm steel bars cut into 1.8 metre lengths. The diameter of each bar is measured at the middle of the length for the purpose of quality control. The results for the bars are 19.9, 19.8, 20.1, 19.9, 19.7, 20.1, 20.0, 19.6, 19.7, 20.1, 20.2, 20.0, 19.9, 19.8, 20.1, 20.0, 19.7, 19.6, 19.9 and 20.2.

Determine the Median, Mode, upper, lower and mid quartiles.

Skill Practice Exercise MEM-SP-0102

The following measurements were recorded for the drying time, in hours, of a certain brand of latex paint.

3.4	2.5	4.8	2.9	3.6
2.8	3.3	5.6	3.7	2.8
4.4	4.0	5.2	3.0	4.8

Calculate the Median, Mode, upper, lower and mid quartiles for this data

Skill Practice Exercise MEM-SP-0103

According to the journal Chemical Engineering, an important property of a fibre is its water absorbency. A random sample of 20 pieces of cotton fibre was taken and the absorbency on each piece was measured.
The following are the absorbency values:

18.71	21.41	20.72	21.81	19.29	22.43
20.17	23.71	19.44	20.50	18.92	20.33
23.00	22.85	19.25	21.77	22.11	19.77
18.04	21.12				

Calculate the Median, Mode, upper, lower and mid quartiles for this data

Skill Practice Exercise MEM-SP-0104

A certain polymer is used for evacuation systems for aircraft. It is important that the polymer be resistant to the aging process. Twenty specimens of the polymer were used in an experiment. Ten were assigned randomly to be exposed to an accelerated batch aging process that involved exposure to high temperatures for 10 days. Measurements of tensile strength of the specimens were made, and the following data were recorded on tensile strength in kPa:

No aging:	227	222	218	217	225
	218	216	229	228	221
Aging:	219	214	215	211	209
	218	203	204	201	205

Calculate the Median, Mode, upper, lower and mid quartiles for both conditions of aging.