Index

arithmetic mean 18
average 18
Bar charts 9
bar graphs 9
bell curve 24
bimodal 18
bivariate data 3
Box and Whisker Plot 12
causality 4
cause and effect relationship 31
charts 7
coefficient of determination 32
coincidental 31
confounding 31
continuous variable 3
correlation 3
correlation coefficient 28
data 2
Data coding 34
Data collection errors 4
Data editing 34
dependent variable 3
descriptive statistics 2
discrete variable 3
dispersion 22
Empirical rule 25
extrapolation 32
frequency distribution 8
frequency polygon 10
graphs 7
histograms 9
independent variable 3
inferential statistics 2
Interval scales 34
IQR 22
Length of questionnaire 77
line of best fit 11
linear relationship 27, 31
lower quartile 22
mean 18
measure of dispersion 22
measure of spread 22
measures of central tendency 18
median 18
mode 18
multi-modal 18
Nominal scales 34
normal distribution curve 24
ordered data set 22
Ordinal scales 34
outlier 19
outlying data 22
Pareto chart 9
Pearson product moment product correlation coefficient 28
perfect negative linear relationship 28
perfect positive linear relationship 27
Pie charts 11
population 3
positive linear relationship 27
predictive statistics 2
qualitative 2
quantitative 2
Questionnaires 5
range 19
regression analysis 31
regression line 31
resistant to outliers 19
reverse cause and effect relationship ... 31
robust 19
Root Mean Squared Deviation 24
sample 3
scatter plot 11
scattergram 11
skewed 25
Standard deviation 24
statistical analysis 2
statistics 2
stem and leaf plot 8
stemplot 8
strong negative linear relationship 28
strong positive linear relationship 27
tables 7
univariate data 3
upper quartile 22
variability 22
variable 2
Variance 23
y hat 31

Training and Education Support
Industry Skills Unit
Meadowbank

Resource UEENEEE024B

Compile and produce an electrotechnology report

Version 1

Training Packages
Electrotechnology UEE07
ESI-TDR-UET06

Learner Resource
Product Code: UEENEEE024B/LLP/V1

TABLE OF CONTENTS

Introduction to engineering analysis and report writing 1
Introduction to statistics 2
Statistics and statistical analysis 2
Descriptive and inferential statistics 2
Data and variables 2
Populations and samples 3
Causation and correlation 3
Data collection errors 4
Design an effective questionnaire 5
Representing data visually 7
Presenting statistical data 7
Diagrams to display data 8
Frequency distribution 8
Stemplot 8
Bar charts 9
Histograms 9
Pareto charts 9
Frequency polygon 10
Pie charts 11
Scatter plot 11
Box Plot 12
Outliers in box plots 13
Review questions – statistical diagrams 14
Activity 1a: Draw statistical diagrams 14
Activity 1b: Draw statistical diagrams 14
Activity 2: Draw and interpret statistical diagram 15
Activity 3a: Interpret statistical diagrams 15
Activity 3b: Interpret statistical diagrams 16
Activity 3c: Interpret statistical diagrams 17
Describing univariate data numerically 18
Definition and calculation of mean 18
Definition and calculation of median 18
Definition and calculation of mode 18
Definition and calculation of range 19
The effect of outliers on measures of centre and spread 19
Review questions – describing univariate data numerically 20
Activity 4a: Describe data numerically 20
Activity 4b: Describe data numerically 20
Activity 5: Compare data sets 20
Activity 6: Determine effect of outliers 21
Activity 7: Choosing appropriate statistical measures 21
Analysing univariate data 22
Measures of spread, dispersion and variability 22
Definition and calculation of inter-quartile range 22
Definition and calculation of variance 23
Definition and calculation of standard deviation 24
Normal distribution curve and the empirical rule 24
Review questions – analysing univariate data 26
Activity 8 – Calculating IQR, S^2 and σ 26

BSI
ISO 9001
TAFE NSW
TRAINING AND EDUCATION SUPPORT

Analysing bivariate data 27
Linear relationships between variables 27
Definition and calculation of correlation in bivariate data 28
Correlation and causation 31
Definition and calculation of regression 31
Coefficient of determination 32
Review questions – analysing bivariate data 33
Activity 9 – Calculating r, y and r 33
Data analysis for reports 34
Quick review of simple data analysis 34
Revision activity 1 – interpreting data from a frequency distribution 34
Revision Activity 2a – Choosing data sources and collection methods 36
Revision Activity 2b – Choosing data sources and collection methods 36
Revision Activity 2c – Choosing data sources and collection methods 36
Revision Activity 3 – Analysing data 37
Revision Activity 4 – Drawing conclusions from data 37
Researching the report 38
The Report Brief 38
Targeting the Reader/s of the Report 40
Researching the Information 41
Writing the report 44
Writing in plain English 44
Tips for writing in Plain English 44
Examples of plain English (not) 48
Introduction to workplace documents 51
Using a word processor 53
Simple Business Reports 57
Using graphics 60
Writing consistently 64
Referencing 69
Preparing the document 72
Assessment task - Conduct an investigation and report on findings 74
Appendix 1 - Recommended websites 75
Appendix 2- question types 76
Appendix 3 – questionnaire checklist 77
Index 78

Introduction to engineering analysis and report writing

A simple search of the web will reveal a great number of resources related to report writing. Many focus on structure, language (grammar, punctuation and spelling) and presentation. Others provide advice about reports for specific purposes. For example: financial reports, business reports and sales reports. What is not easy to find is an example of an engineering report that shows how data and analysis are incorporated into the document and used to substantiate the findings and recommendations. Yet, this is exactly what engineers are required to do.

This resource is designed for an engineer who is required to present information, conclusions and/or recommendations to others. It is assumed that the reports will be underpinned by technical information and will require the collection and analysis of data as part of the research phase. As a result, this resource has two main foci. The first is statistical analysis. The second is on research and writing of engineering reports.

An engineering report is structured in much the same way as any other. It has an introduction, the body of the report, conclusion and/or recommendations and may also have appendices. Hence it is not the structure of the engineering report that makes it different. Rather, it is the content that separates this from other reports.

The task of an engineer is to draw conclusions and/or make recommendations. Both these require some research. That means collecting information, facts and data. This all needs to be sorted, understood, analysed and in some cases, used to predict future events.

Much of the information an engineer relies on is in the form of data; numbers that describe a phenomenon. For example, the data may be a series of ammeter and power factor meter readings over a period of time. Or the number and severity of noise spikes in a particular circuit. Or the number of times a particular component or category of components fails over a period of time.

These numbers are meaningless until we use statistics to create pictures out of them. Statistics are the result of applying mathematics to data sets so they describe or predict phenomenon.

In the following pages you will be introduced to the simpler statistical analyses and asked to apply those to given data and/or data from your own workplace. Then you will be introduced to research and the preparation of an engineering report. Ultimately, your demonstration of competence will be in the form of a completed engineering report based on your own research including some statistical analysis.

Introduction to statistics

Statistics and statistical analysis

Statistics is the name given to the mathematical science of collecting, organising, summarising, analysing, interpreting, explaining and presenting data. Statistics can be used to predict and forecast what may happen based on data. They can also be used to describe a problem.

Some people consider statistics to be a difficult branch of mathematics. However, like everything else, once you understand the purpose and processes of creating statistics, you will find them easier to work with.

Statistical analysis is the process of applying one or more mathematical models to data to determine trends, relationships or make predictions. A statistical model is a set of mathematical equations which describe the behaviour of the object of study in terms of one or more variables.

Descriptive and inferential statistics

When statistical methods are used to summarize or describe a collection of data it is called **descriptive statistics**. When patterns in data are used to draw inferences about the process or population being studied it is called **inferential** or **predictive statistics**.

For example, if we collected data about how long various parts of the distribution system have been in place without need for replacement cabling, the information would be descriptive of the current situation. However, using that information to make plans for progressive upgrades of cabling would require inferential statistics.

Data and variables

Data is the name given to information or facts collected through experience, observation or experiment. Data can also be thought of as the values of measurements used to describe a variable. They may consist of numbers, words, or images; particularly as measurements or observations of a set of variables.

The singular form of data is datum. This means that if you have one fact, it is datum. If you have many facts, they are data.

There are two main types of data: **qualitative** and **quantitative**. Generally, we call data 'quantitative' if it is in numerical form and 'qualitative' if it is not. It's important to note that qualitative data could be more than just words or text. Photographs, videos, sound recordings and so on, can also be considered qualitative data.

In statistics we talk a lot about variables. Variables are things that we measure, control, or manipulate in research. A **variable** is *any entity that can take on different values*. For example, *failure rate* can be considered a variable because failure rate can take different values for different components or for the same component under different circumstances. Similarly, *power consumption* can be considered a variable because how much power is consumed by a component or installation can be assigned a value.

It's important to understand the distinction between **independent** and **dependent** variables. This is particularly relevant when investigating cause-effect relationships. An *independent variable is the one that can be manipulated*. The *dependent variable is what is affected by the independent variable*. For example, if you are researching the effects of a new fertilizer on crops, the fertilizer is the independent variable and the growth of the crops is the dependent one.

Variables can also be described as either **discrete** or **continuous**. Measurements such as height, weight, length and time are **continuous.** They can be measured precisely and *can have any one of an infinite number of values*. Car might travel 3.87 km which could be measured precisely. On the other hand, discrete values have a limited number of values and there are "gaps" between the values. You cannot score 1.2 goals in a soccer game even though we could calculate that to be the average number you score per game during a season. In summary, discrete data are "counts" of things, while continuous data come from measurements.

Data can also be described as univariate or bivariate. **Univariate data** has only one variable. An example would be the number of blackouts each month over a twelve month period. The variable is blackouts.

Bivariate data has two variables and is used to describe and/or analyse the relationship between the two variables. An example would be the average temperature each day and the average power consumption of the household. The bivariate data (temperature and power) could be examined to see if there is a relationship between the two variables.

Populations and samples

The term population means the entire group items that are the subject of a study or investigation. For example, if you are trying to determine the number and cause of failures in distribution transformers in a particular geographic area, the population is the total number of distribution transformers in that area. The population excludes other types of transformers and those in other geographic areas.

A sample is a subset of the population. For example, if you were interested in the savings that may be gained from using a particular supplier you may choose to collect information about some of products they supply (a sample) rather than all the products available.

Causation and correlation

Often when we have gathered our data and displayed it in the form of a graph or table, we can see relationships between variables. However, it is very important that we can differentiate between a causal relationship and a coincidental relationship.

For example, if the frequency of transformer failures in a particular geographic area was plotted against the number of lightening strikes over a period of time, it may be possible to see a relationship. That doesn't necessarily mean that lightening caused the failures. It simply means there is a correlation (relationship) between them. The significance of the **correlation** is a measure of how strong the relationship is but is not an indicator of cause and effect.

Correlation is one of the most common and most useful statistics. A correlation is a single number that describes the degree of relationship between two variables (bivariate data). Correlation can describe a relationship regardless of whether that relationship is causal or coincidental.

Causality means that one variable is the direct result of another variable. To determine causality an experiment needs to be done where the assumed "cause" is deliberately applied to see the "effect". For example, to determine the causality between temperature and failure rates in circuit boards, a sample of circuit boards would need to be subjected to a range of temperatures and the failure rate recorded. The relationship between temperature and failure rate is called correlation.

Data collection errors

Data collection can be prone to errors. This is true regardless of whether the data is collected manually or through an electronic process.

A common error is to collect insufficient data. Even if working under time constraints it is important to collect enough data to make the subsequent analysis meaningful. Too much data will make analysis time consuming but is not likely to affect the result. Too little data can save time but can have a significant effect on the result.

How much is enough data? The answer depends on what you are studying and what type of analysis you are intending to do.

Another common error is to collect the wrong data. This is often caused by not spending enough time thinking about the problem or issue being studied and the possible symptoms or causes. Symptoms are indications of a problem but are not the cause.

The following are typical words or phrases that may be used to describe 'problems'. In fact, each of these is a symptom. The cause needs to be investigated.

> makes unusual noise, won't work, no power, machine down, broken tool, head froze up, contaminated, rough surface, porosity, shortage of parts, rattles, quality problem, worn out, line stopped, not to specification, labour problem, management problem, too much variation

Determining the cause can be as simple as applying the 5W2H questions.

Who?	...brought the problem to your attention
What?	...is happening or not happening
When?	...did the problem start
Where?	...is it occurring?
Why?	...is it happening – this is about the context of the problem
How?	...is it happening - in what mode or situation did the problem occur
How Many?	...times or errors or similar measures of the problem

If the preceding questions have been answered, another data collection error can be avoided. That is, collecting data at the wrong time. If an event only occurs at particular times or under particular circumstances, it is important to collect data at those times or under those circumstances.

Correlation is one of the most common and most useful statistics. A correlation is a single number that describes the degree of relationship between two variables (bivariate data). Correlation can describe a relationship regardless of whether that relationship is causal or coincidental.

Causality means that one variable is the direct result of another variable. To determine causality an experiment needs to be done where the assumed 'cause' is deliberately applied to see the 'effect'. For example, to determine the causality between temperature and failure rates in circuit boards, a sample of circuit boards would need to be subjected to a range of temperatures and the failure rate recorded. The relationship between temperature and failure rate is called correlation.

Data collection errors

Data collection can be prone to errors. This is true regardless of whether the data is collected manually or through an electronic process.

A common error is to collect insufficient data. Even if working under time constraints it is important to collect enough data to make the subsequent analysis meaningful. Too much data will make analysis time consuming but is not likely to affect the result. Too little data can save time but can have a significant effect on the result.

How much is enough data? The answer depends on what you are studying and what type of analysis you are intending to do.

Another common error is to collect the wrong data. This is often caused by not spending enough time thinking about the problem or issue being studied and the possible symptoms or causes. Symptoms are indications of a problem but are not the cause.

The following are typical words or phrases that may be used to describe 'problems'. In fact, each of these is a symptom. The cause needs to be investigated.

> makes unusual noise, won't work, no power, machine down, broken tool, head froze up, contaminated, rough surface, porosity, shortage of parts, rattles, quality problem, worn out, line stopped, not to specification, labour problem, management problem, too much variation

Determining the cause can be as simple as applying the 5W2H questions.

Who?	...brought the problem to your attention
What?	...is happening or not happening
When?	...did the problem start
Where?	...is it occurring?
Why?	...is it happening – this is about the context of the problem
How?	...is it happening - in what mode or situation did the problem occur
How Many?	...times or errors or similar measures of the problem

If the preceding questions have been answered, another data collection error can be avoided. That is, collecting data at the wrong time. If an event only occurs at particular times or under particular circumstances, it is important to collect data at those times or under those circumstances.

Design an effective questionnaire

Questionnaires are a very widely used tool to gather data. How effective they are in terms of gathering useful data depends on many things. However, four of the most important design considerations are (a) how easy it is to respond to, (b) how relevant the questions are to the goal of the research, (c) how well the questions are written, and (d) what type of questions are used.

In Appendix 1 are some sample questions including both open-ended and closed question types. Which you choose will make a difference to how you can display and analyse the data collected.

For example, if a multiple choice type question is used, the results can be tallied and displayed in a bar chart. Below is an example.

25 engineers are asked the following question.

Which of the following is your preferred supplier?

- ☐ Supplier A
- ☐ Supplier B
- ☐ Supplier C
- ☐ Supplier D

Their responses are shown below.

4	Supplier A
7	Supplier B
2	Supplier C
12	Supplier D

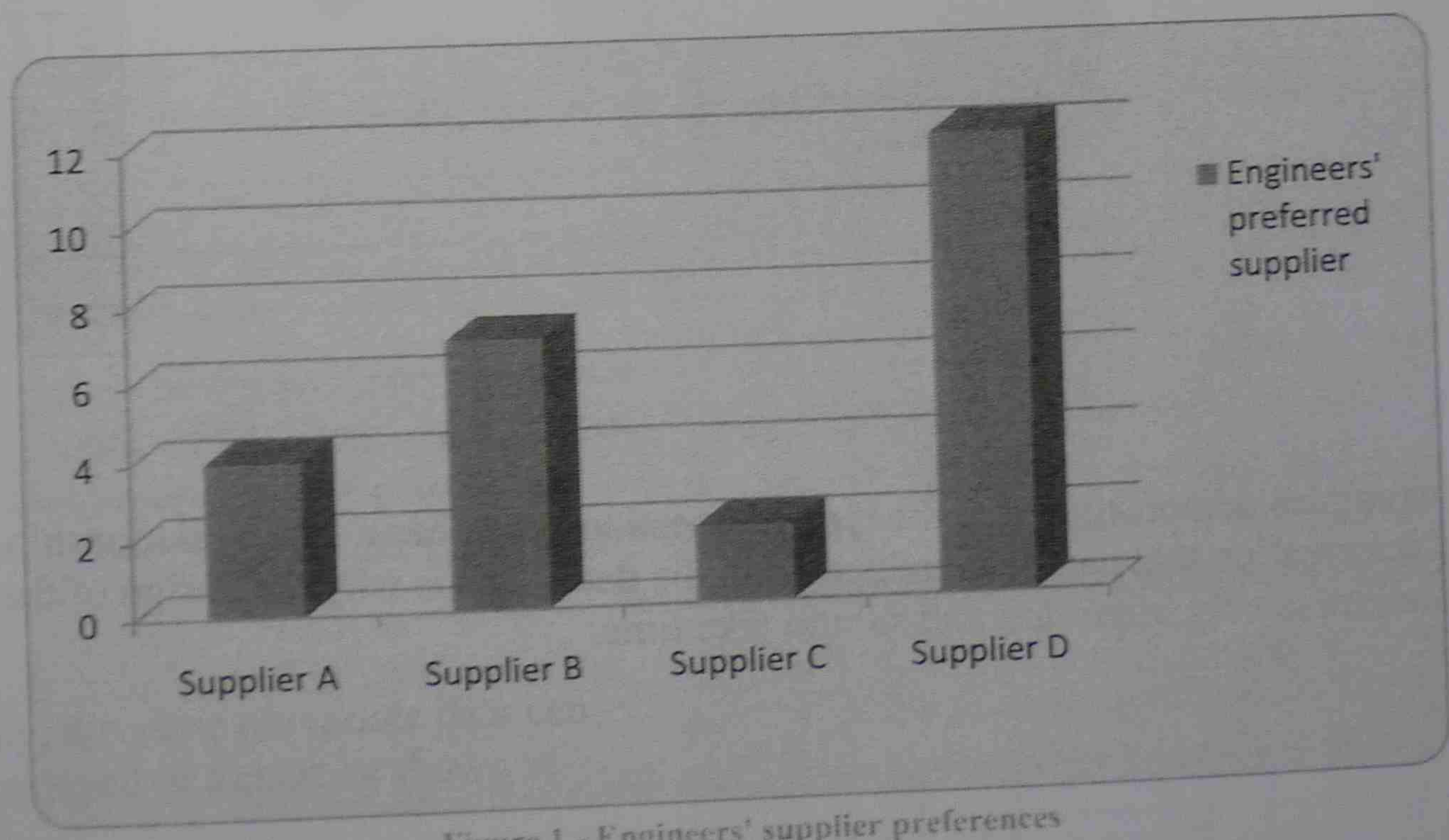


Figure 1 - Engineers' supplier preferences

If the same 25 engineers were asked to rank the suppliers from 1 to 4 more information could have been extracted from the data. For example, it would have also been possible to identify their ***least*** preferred supplier.

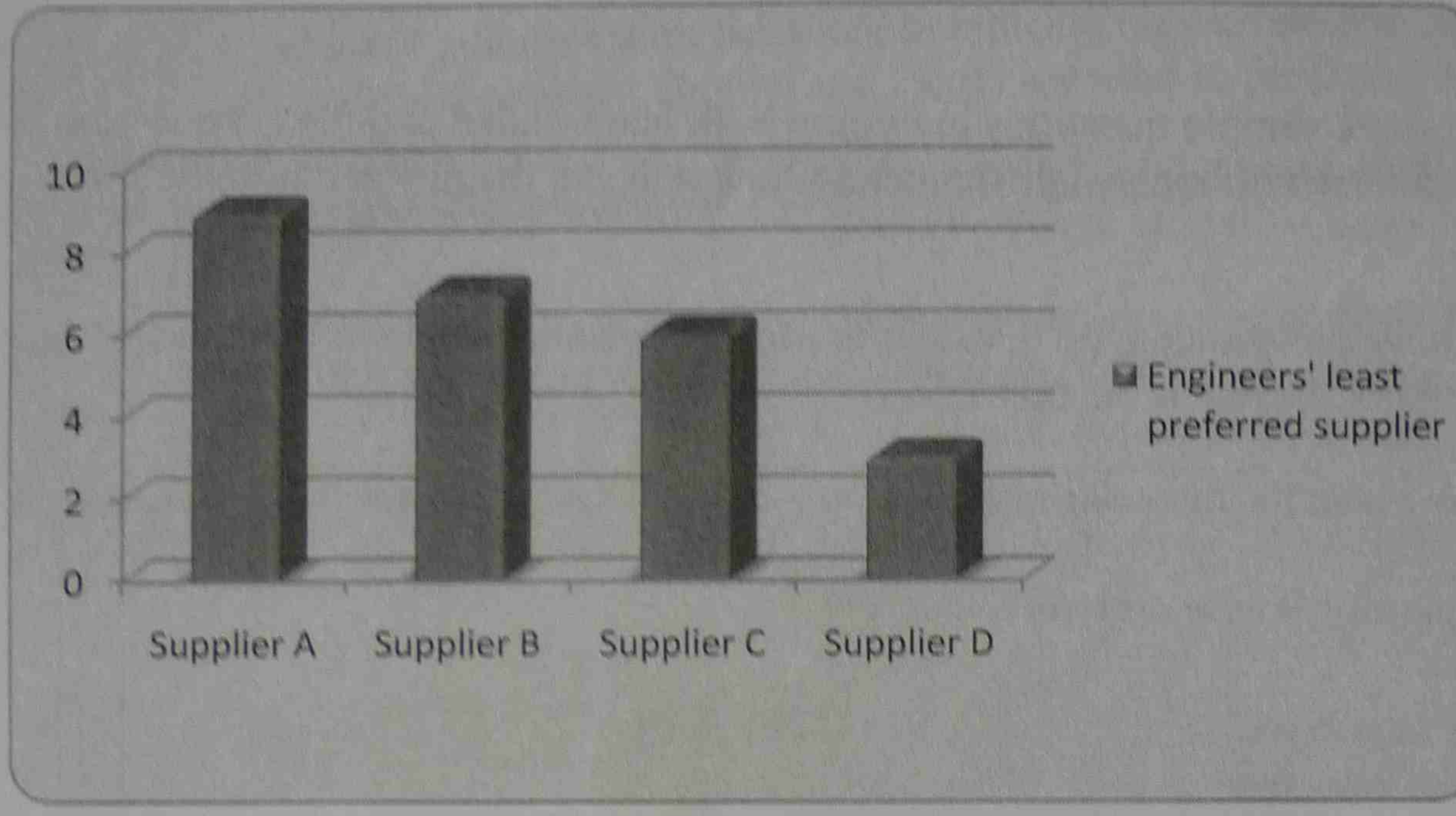


Figure 2 - Engineers' least preferred supplier

Rankings			
A	B	C	D
A	C	B	D
A	C	B	D
A	B	D	C
B	A	D	C
B	A	D	C
B	A	D	C
B	D	A	C
B	D	A	C
B	D	C	A
B	D	C	A
C	D	A	B
C	D	A	B
D	A	C	B
D	A	C	B
D	C	A	B
D	C	A	B
D	C	A	B
D	C	B	A
D	C	B	A
D	B	C	A
D	B	C	A
D	B	C	A
D	B	C	A
D	B	C	A

These examples demonstrate that the type of question asked has a direct relationship to the type of information that can be extracted from the data. This is why the design of the questionnaire is very important and should take time.

Representing data visually

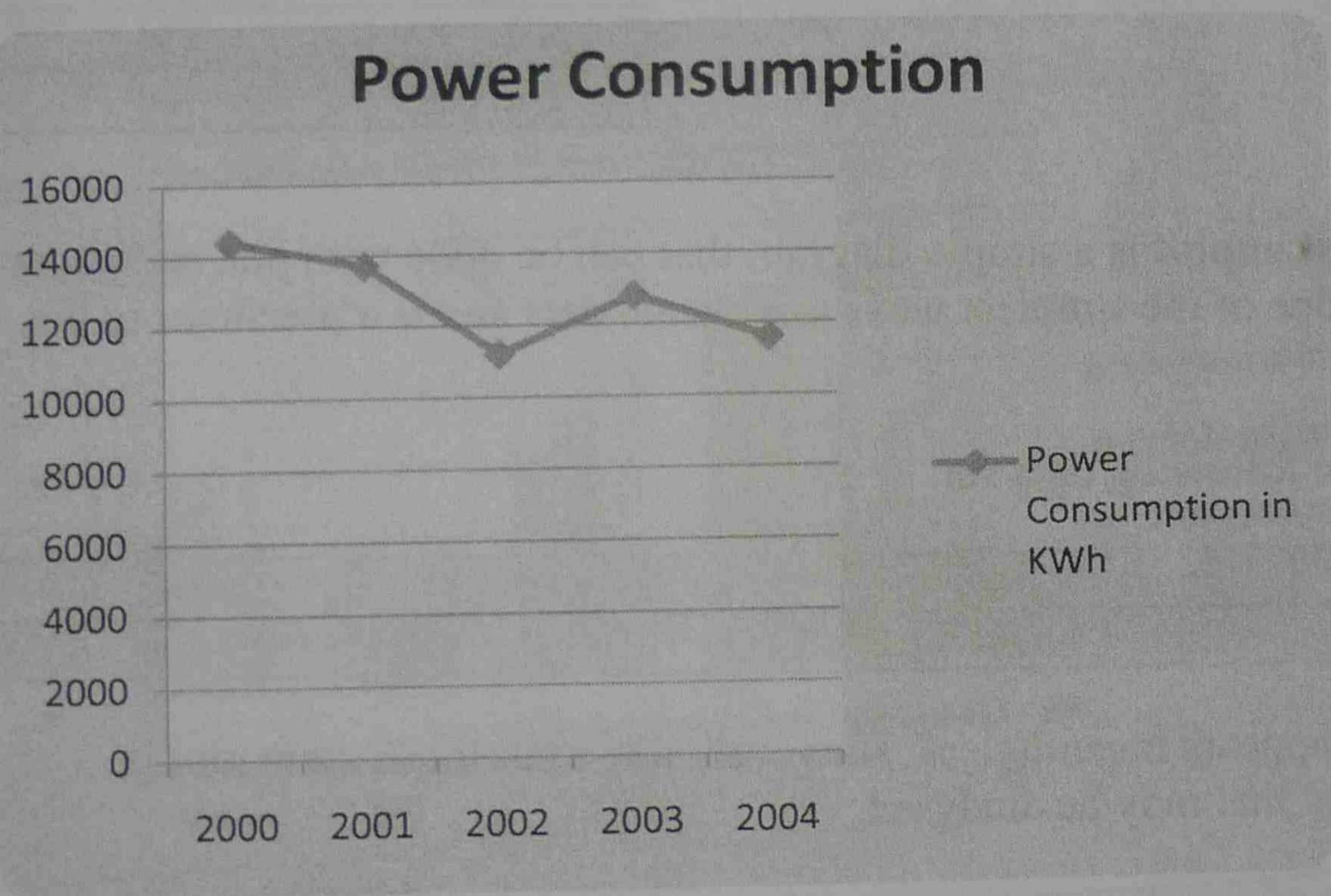
Presenting statistical data

The most common methods of presenting data are in tables, graphs or charts.

Tables are simply rows and columns with headings to indicate the variables or statistical analysis that has been applied to them. Graphs and charts are used to present data visually.

For example, the following table shows the annual power consumption of a household over a five year period.

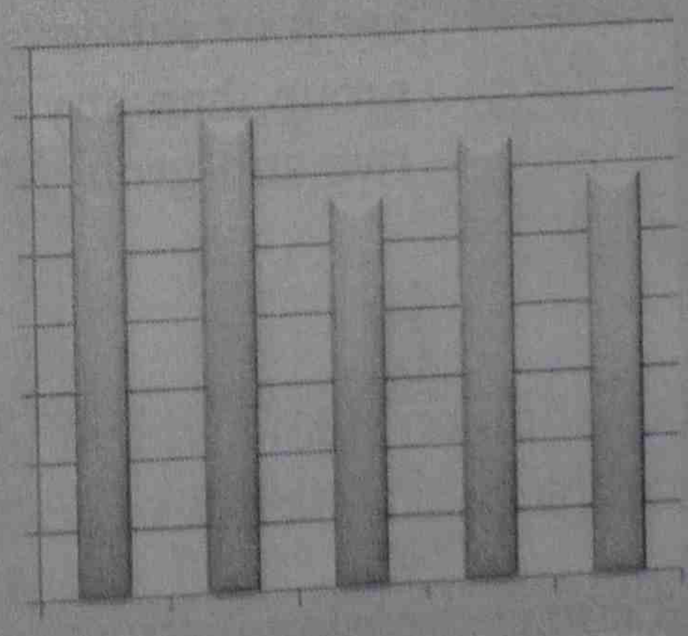
Year	2000	2001	2002	2003	2004
Power consumption in KWh	14400	14500	11200	9600	6300



The graph at left shows how power consumption (Y axis) in the household varied over time (X axis).

Figure 3 - Household power consumption over time

Finally, the same univariate data can be displayed in a chart as shown at right.



Diagrams to display data

Data can be displayed in a wide variety of graphical forms. The following diagrams types can all be created by hand or using the "insert chart" function in Microsoft Word® or Excel®.

Frequency distribution

In very simple terms, a **frequency distribution** is simply a count of data. This count is displayed in a way that makes sense to the observer. For example, a maintenance team may count the number of call outs each month of the year and record it in a frequency distribution such as the one at right.

Month	No. of call outs
January	175
February	209
March	186
April	155
May	102
June	134
July	178
August	172
September	118
October	130
November	156
December	169

Stemplot

A **stem and leaf plot** or **stemplot** is a simple diagram that can be done with pen and paper with no other tools. It is one of the simplest ways to indicate data and is a precursor to the histogram.

For example, consider the following data set.

22	45	28	36	34	29	24	37	31	49
52	58	38	26	54	34	39	39	28	41

In its raw form, this data appears meaningless. However, when put into a stem plot it begins to assume a 'shape' that may be analysed.

To create the stemplot the 1st number (ie most significant digit) is used as the stem and the 2nd number as the leaf as shown below.

2	2 4 6 8 8 9
3	1 3 4 4 6 7 9 9
4	1 5 9
5	2 4 8

From this very simple diagram we can immediately see that there are four sub-groups of data, there are more items in the 30-39 sub-group than any other and the majority of the data is in the lower two sub-groups (20-29 and 30-39).

For more information on stem and leaf plots go to
http://www.netmba.com/statistics/plot/stem/ (accessed in June 2009)

For simple but comprehensive instructions on the construction of stem and leaf plots go to
http://www.purplemath.com/modules/stemleaf.htm (accessed June 1009)

Bar charts

Bar charts or **bar graphs** are most commonly used to display univariate data. Each bar in the chart represents a variable. The height or length of the bar represents the count of that variable. The advantage of bar charts is that they are very simple to create and update with new data. Bar charts can be displayed horizontally or vertically.

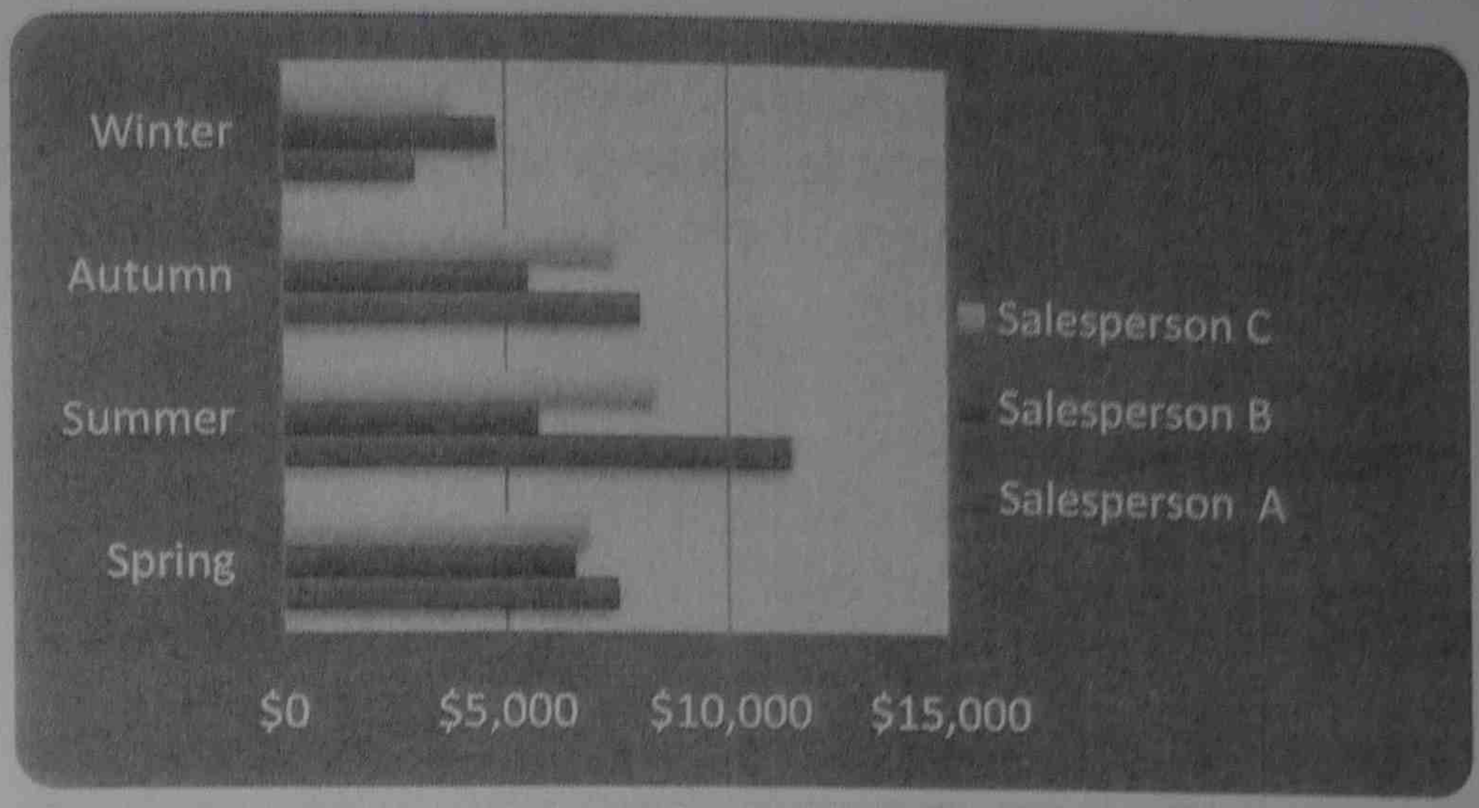


Figure 5 - Sample bar chart

Histograms

Histograms are similar to bar charts except each 'bin' (bar) represents a 'class' (subgroup) of data. For example, the histogram below shows the number of employees in age ranges (sub-groupings of ages) employed in different regions.

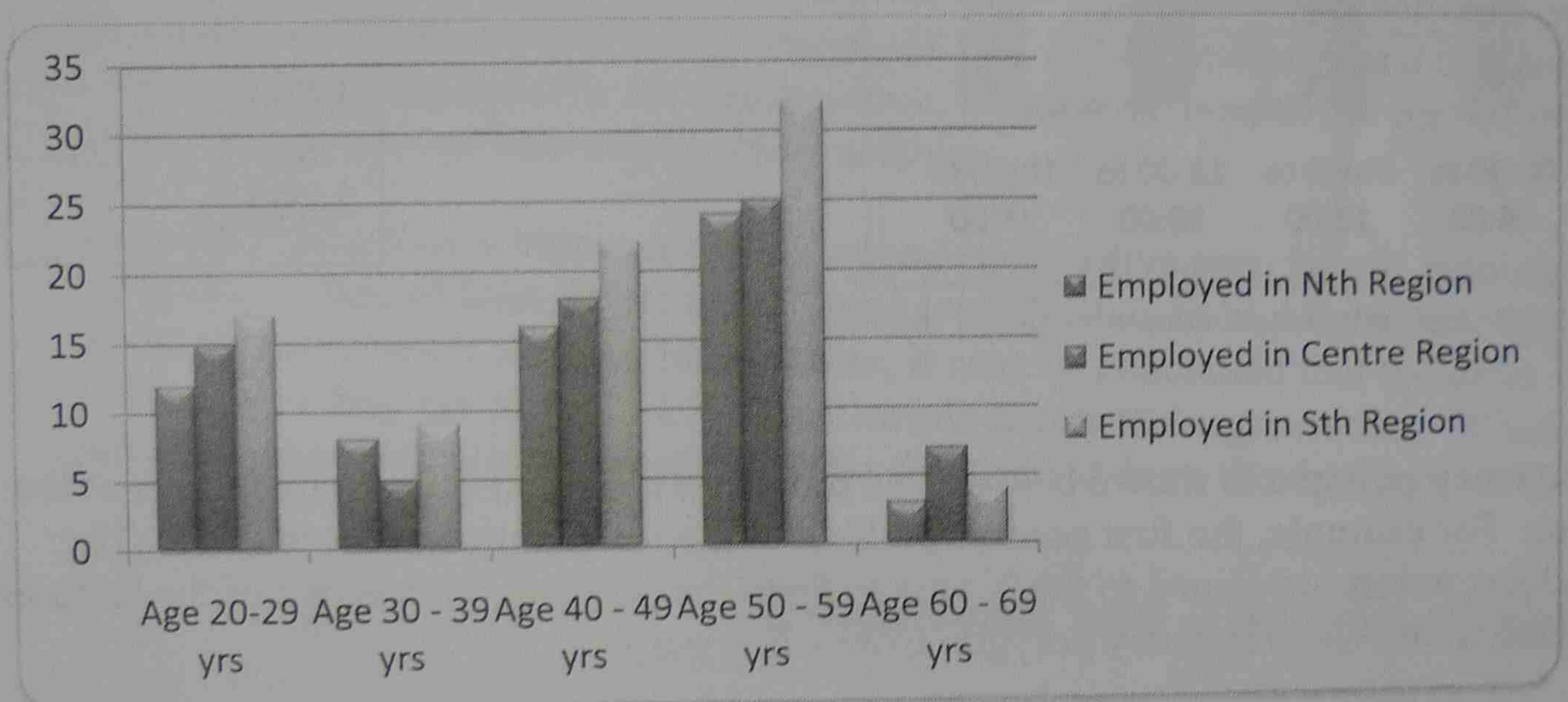


Figure 6 - Sample histogram

Pareto charts

A bar chart or histogram arranged from highest to lowest value is called a **Pareto chart**.

For more information on histograms go to http://www.netmba.com/statistics/histogram/ (accessed in June 2009)

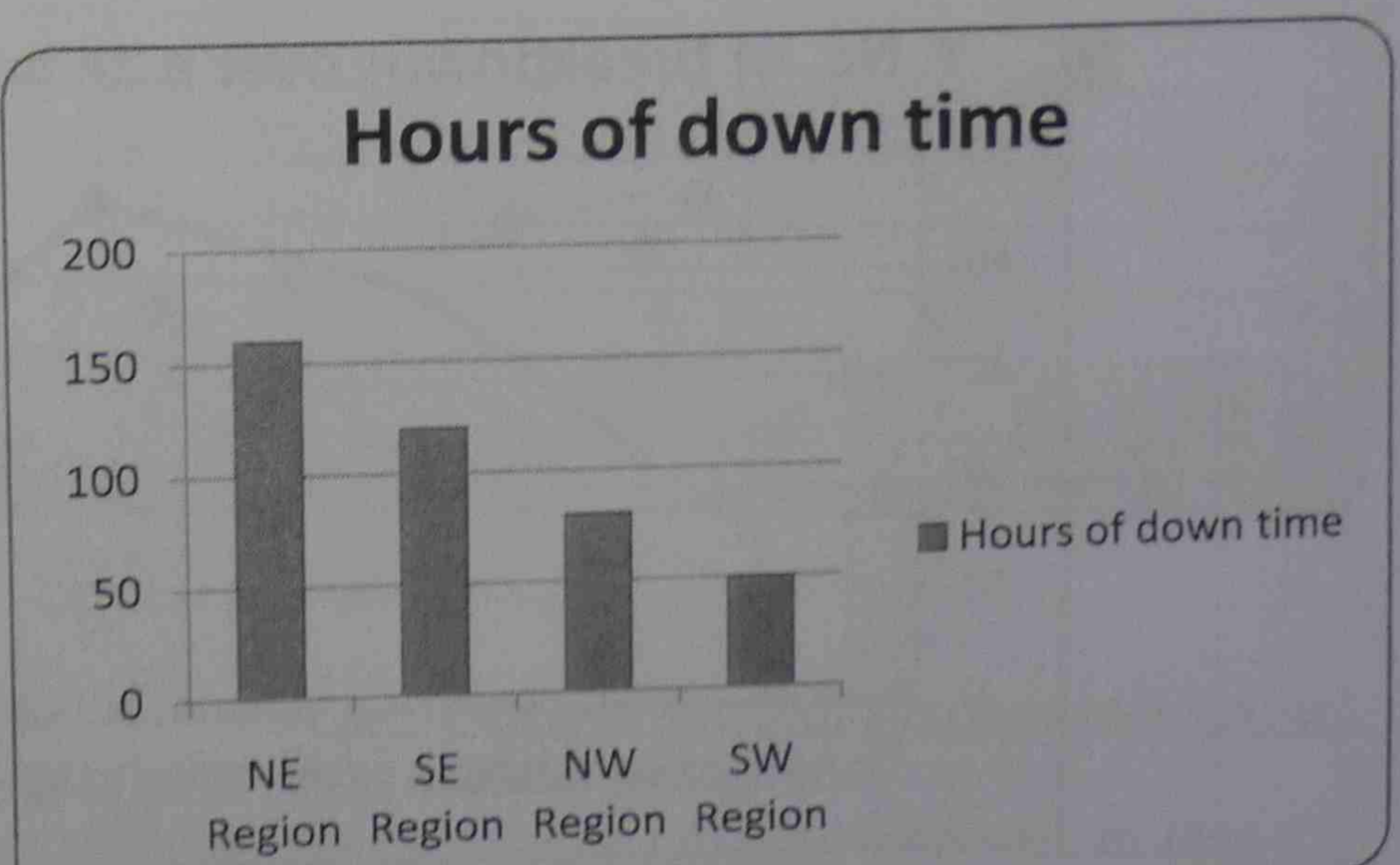


Figure 7 - Sample Pareto chart

Frequency polygon

A **frequency polygon** is a type of line graph. It can be created by joining the midpoints of the bars in a histogram or can be developed straight from the frequency distribution.

For example, at right is a frequency distribution showing the number of breakdowns in each of four, six hour periods in a day. The histogram for this data is shown below.

Time	No. of breakdowns
00:00 to 06:00	15
06:00 to 12:00	38
12:00 to 18:00	47
18:00 to 24:00	21

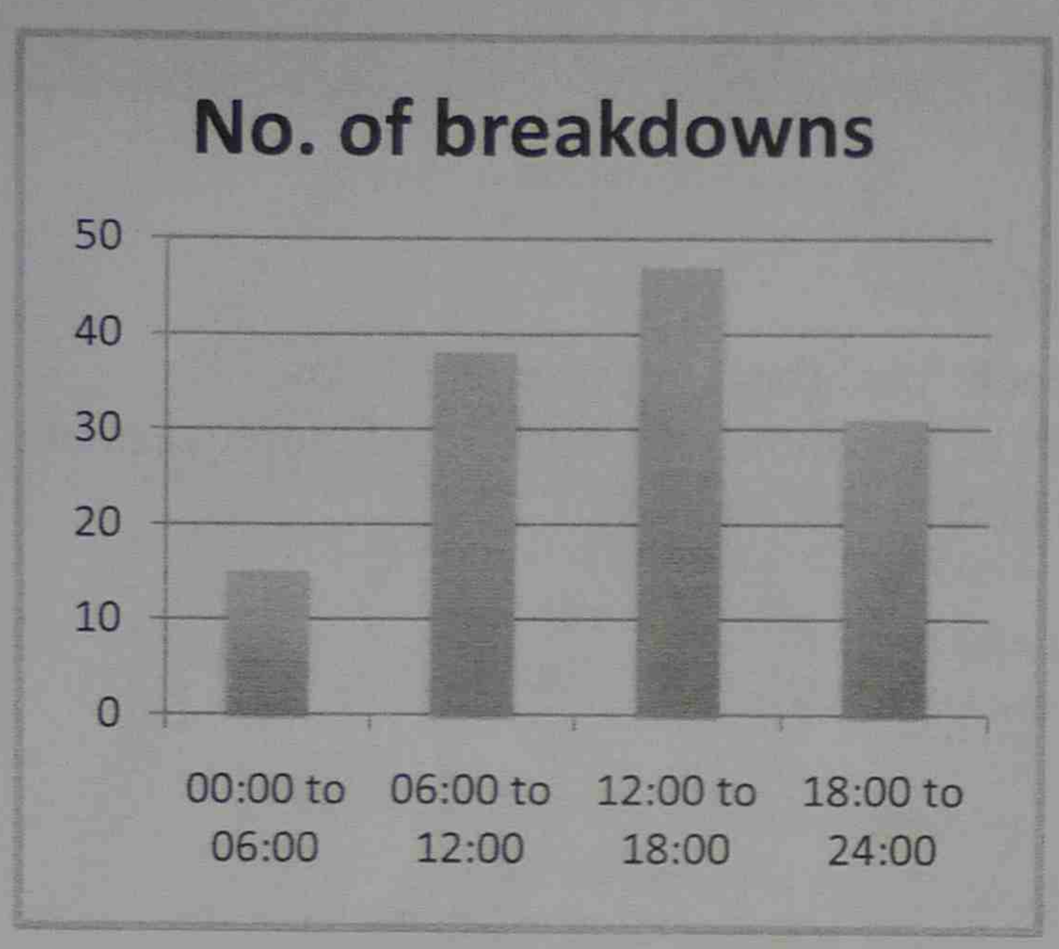


Figure 8 - Histogram based on sample frequency distribution

The frequency polygon is shown below. The points on the X axis are the midpoints of the subgroup. For example, the first point is at 03:00 – the midpoint of 00:00 to 06:00. The first and last points are joined to the X axis to form the polygon. This is one of the features that differentiates the frequency polygon from a line graph.

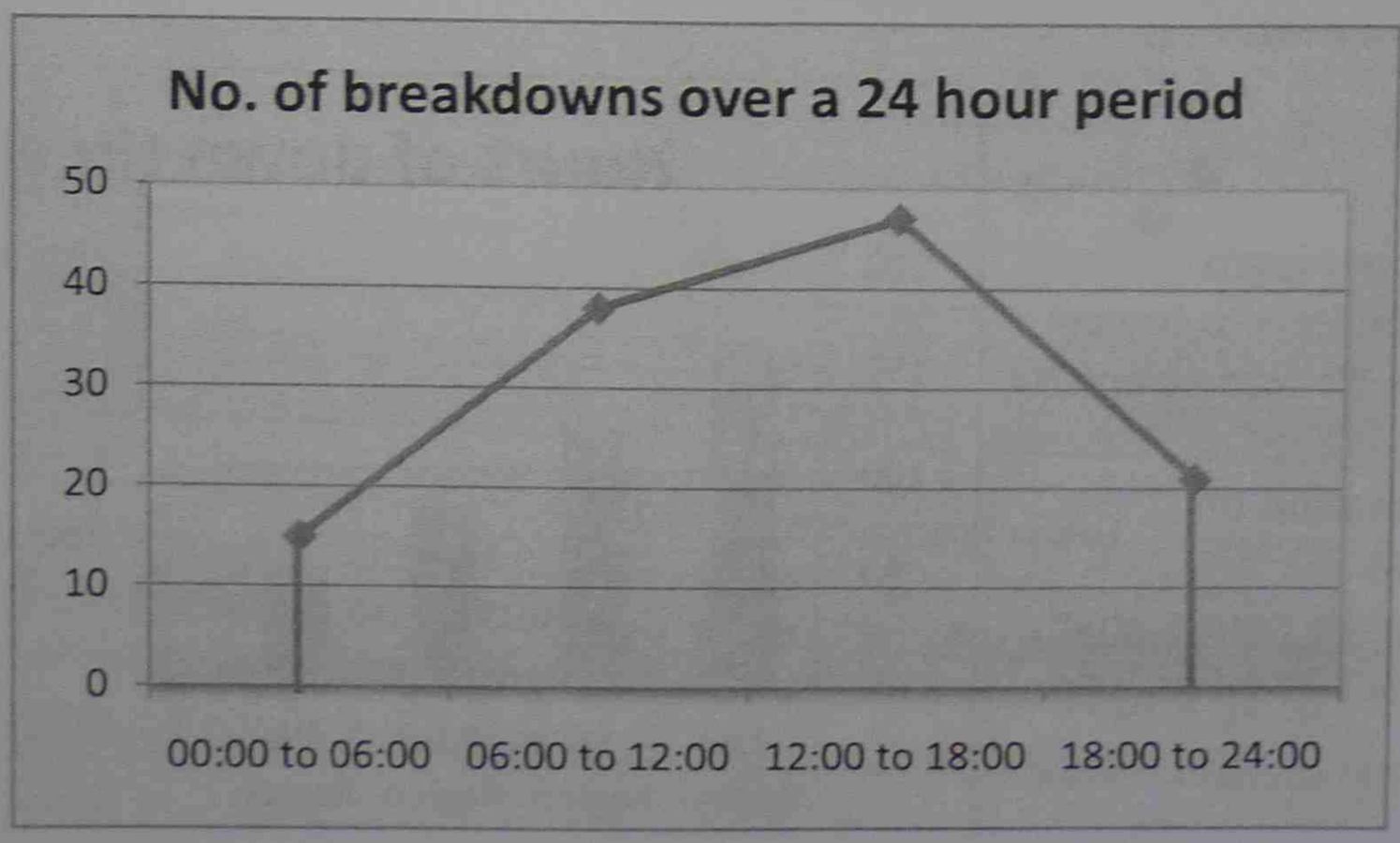


Figure 9 - Sample frequency polygon

For more information on frequency polygons go to http://www.math-mate.com/chapter16_10.shtml (accessed in June 2009)

Pie charts

Pie charts are very useful for showing relative values. In the example below, the total number of repairs completed in week 16 has been represented in terms of the proportion completed by each of four teams.

It is important to note that the 'pie' must always represent 100% of the data values.

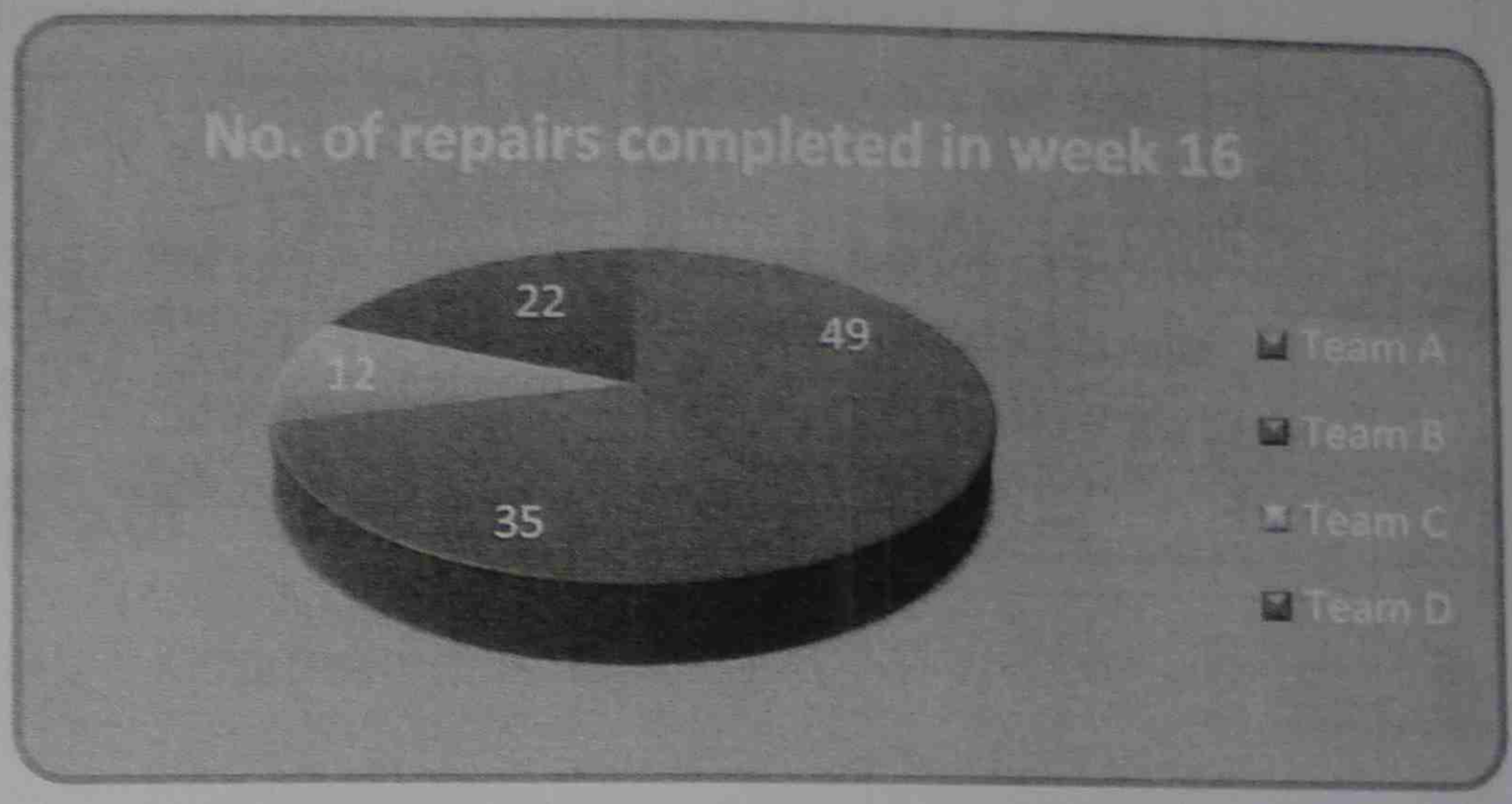


Figure 10 - sample pie chart

Scatter plot

A **scatter plot** or **scattergram** can be used to show the relationship between two variables. For example, the scatter plot below shows the age of 12 employees and the length of time they have been employed by the organisation. Being displayed in the scatter plot shows there is a relationship between the two variables.

The scatter plot shown below seems to indicate a 'positive correlation' between the two variables. A **line of best fit** has been drawn over the data to highlight the relationship. Based on how closely the data fits that line, it may be concluded that the older an employee is, the longer they are likely to have been employed by the organisation. Whether that is a causal or coincidental relationship can't be determined from the data provided.

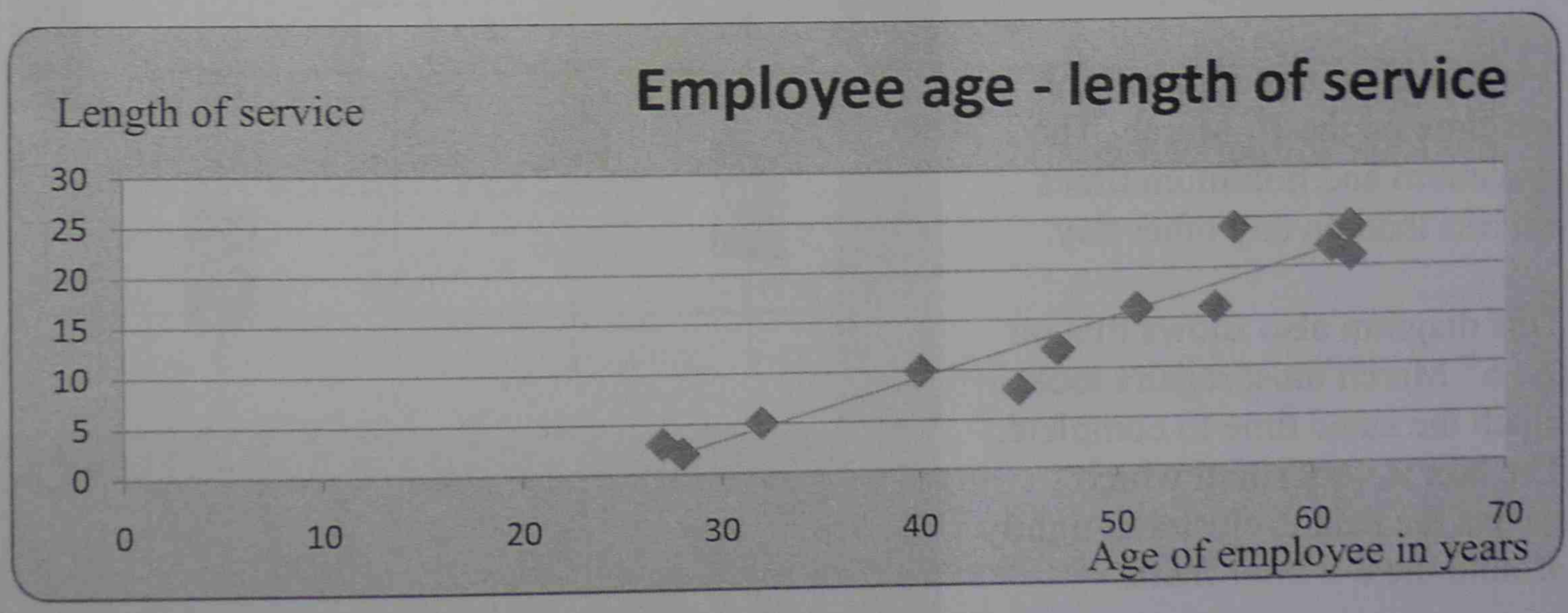
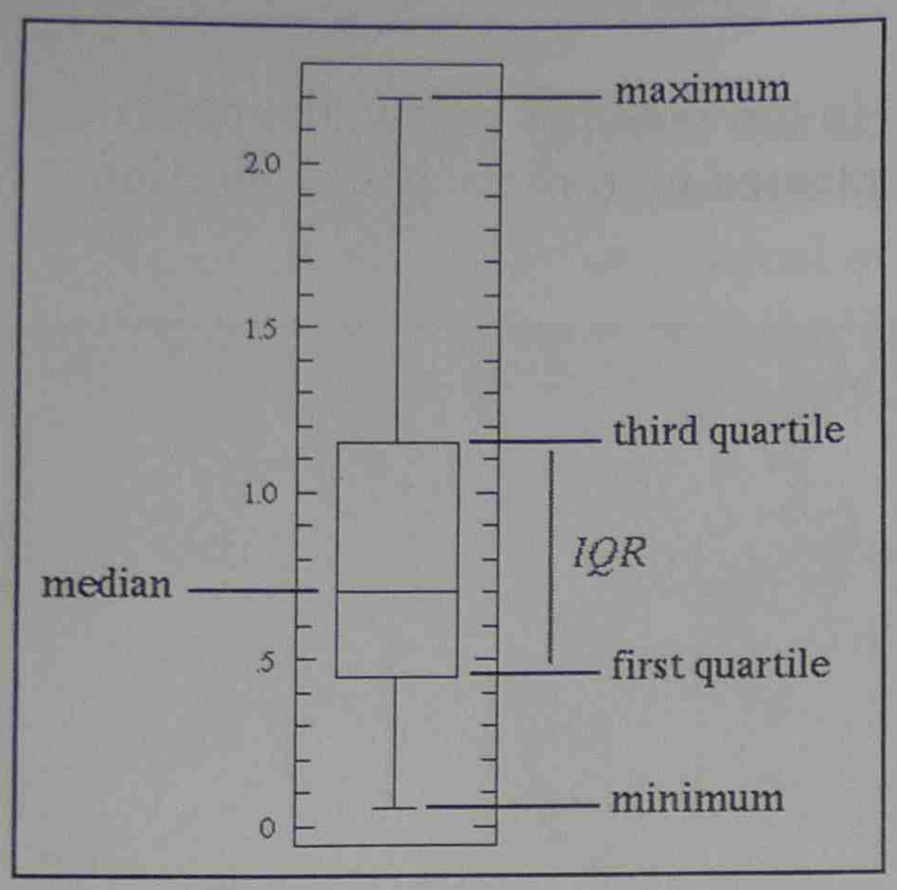


Figure 11 - Sample scatter plot

For more information on scatter plots go to http://www.netmba.com/statistics/plot/scatter/
For information on determining correlation from a scatter plot go to http://www.mste.uiuc.edu/courses/ci330ms/youtsey/scatterinfo.html (accessed in June 2009)



Box Plot

A **Box and Whisker Plot** is a way of looking at the whole range and distribution of numbers for a given variable.

The box contains the middle 50% of the data. If the median line in the box is not centred, the data is skewed. The ends of the vertical lines are the 'whiskers' and they represent the minimum and maximum values of the data set.

Note: ***IQR*** = inter quartile range

This box and whisker plot shows the time taken to repair equipment over a period of five days.

It is immediately clear that repairs took longer on the 8th March. The maximum repair time (12 hrs) and minimum repair time (2 hrs) were both longer than on any other day.

It is also clear that repairs took less time on the 7th March. The maximum and minimum times are less than on any other day.

This diagram also shows that on the 6th March most repairs took much the same time to complete. The box is very small which means the data is clustered tightly around the 6-7 hour mark.

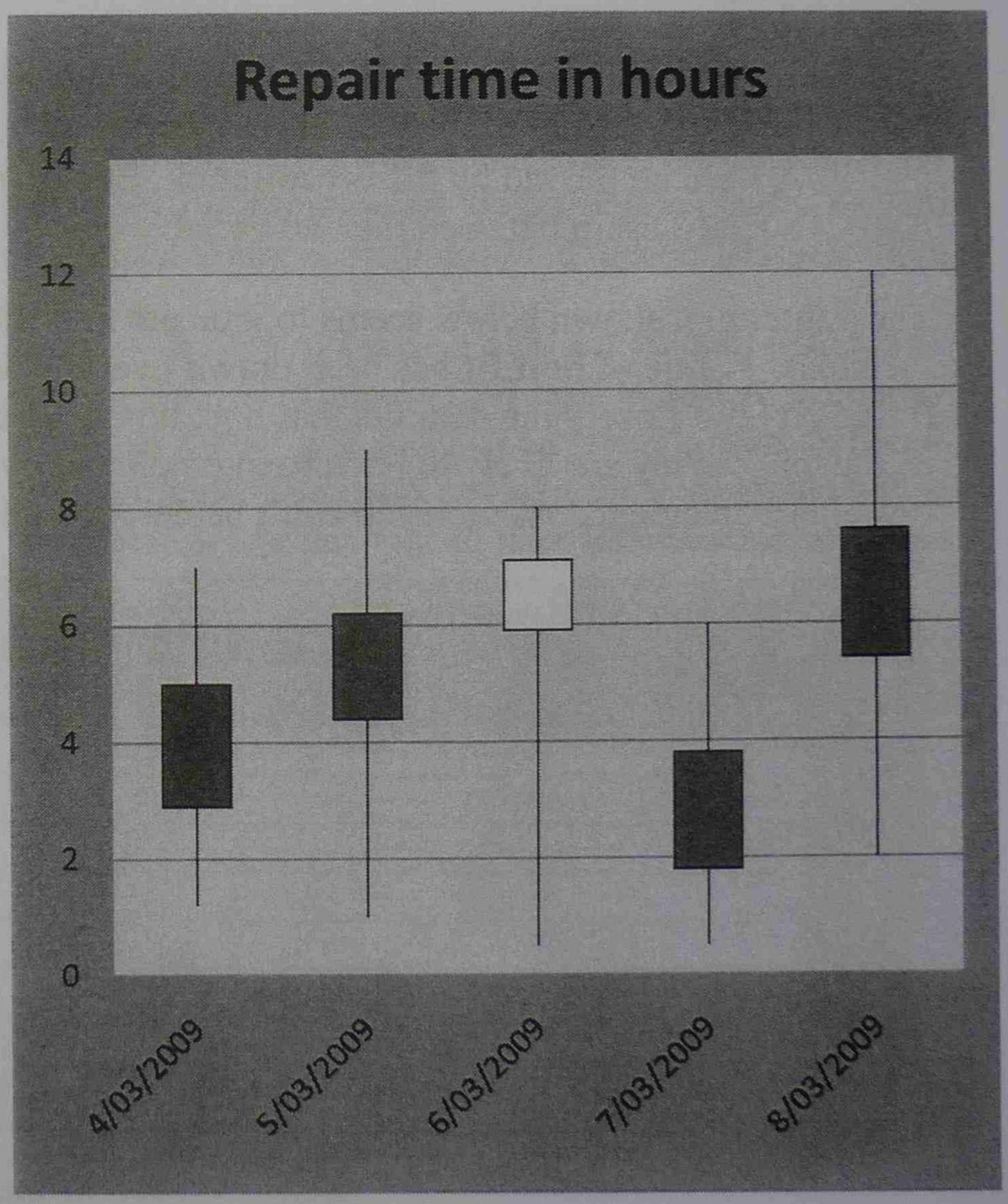


Figure 12 - Sample box plot

For more information on box and whisker plots go to http://www.netmba.com/statistics/plot/box/ (accessed in June 2009)

Outliers in box plots

A box plot assumes that all your data is clumped around the median value. In some data sets there are values that lie quite a long way from the median. These are called outliers and represent an unusual event or are the result of data collection error. The length of the box in the box plot is the "inter-quartile range". An outlier is defined as datum that lie more than 1.5 times the inter-quartile range away from the box. They are assumed to be erroneous data. If outliers are included in the box plot, the length of the whiskers will give a distorted view of the bulk of the data. For this reason, often the outliers are not included in the diagram.

For simple but comprehensive instructions on the construction of box plots go to http://www.purplemath.com/modules/boxwhisk.htm (accessed June 1009)

Review questions – statistical diagrams

Activity 1a: Draw statistical diagrams

Twenty rolls of cable were weighed to the nearest ½ kilogram. The results are shown in the table below.

64.5	65.0	64.0	68.5	62.0	65.5	71.5	70.0	69.0	67.5
63.5	66.5	68.0	72.0	69.5	66.0	66.0	68.5	64.5	65.0

Use the data to create each of the diagrams listed below. Your diagrams may be drawn by hand or using any technology or software available to you.

a. Frequency distribution
b. Stem and leaf plot
c. Bar chart
d. Histogram with class interval of 2 kg
e. Frequency polygon
f. Pie chart based on the class intervals in the histogram

Activity 1b: Draw statistical diagrams

The data at right was collected through a questionnaire distributed to ten Engineering Diploma students. It represents the number of hour's study they did in the five days prior to a final examination and the percentage scored in that event.

Hrs of study	% grade
3	80
5	90
2	75
6	80
7	90
1	50
2	65
7	85
1	40
7	100

Create a scatter plot based on the data provided and include a line of best fit.

Respond to the following.

a. Is there any indication of a relationship between the variables? ☐yes ☐no
b. What conclusion(s) would you draw from the data and/or shape and distribution of your scatter plot?

Activity 2: Draw and interpret statistical diagram

The following data represents current readings from an instrument panel over a period of three days (readings taken approximately every 2 hours).

4.3, 5.1, 3.9, 4.5, 4.4, 4.9, 5.0, 4.7, 4.1, 4.6, 4.4, 4.3, 4.8, 4.4, 4.2, 4.5, 4.4, 7.2

a. Calculate the median, range, 1st quartile, 3rd quartile and inter-quartile range
b. Identify any outliers
c. Draw the box plot for the data set including outliers
d. Draw the box plot excluding outliers
e. Describe the box plot in terms of skew

Activity 3a: Interpret statistical diagrams

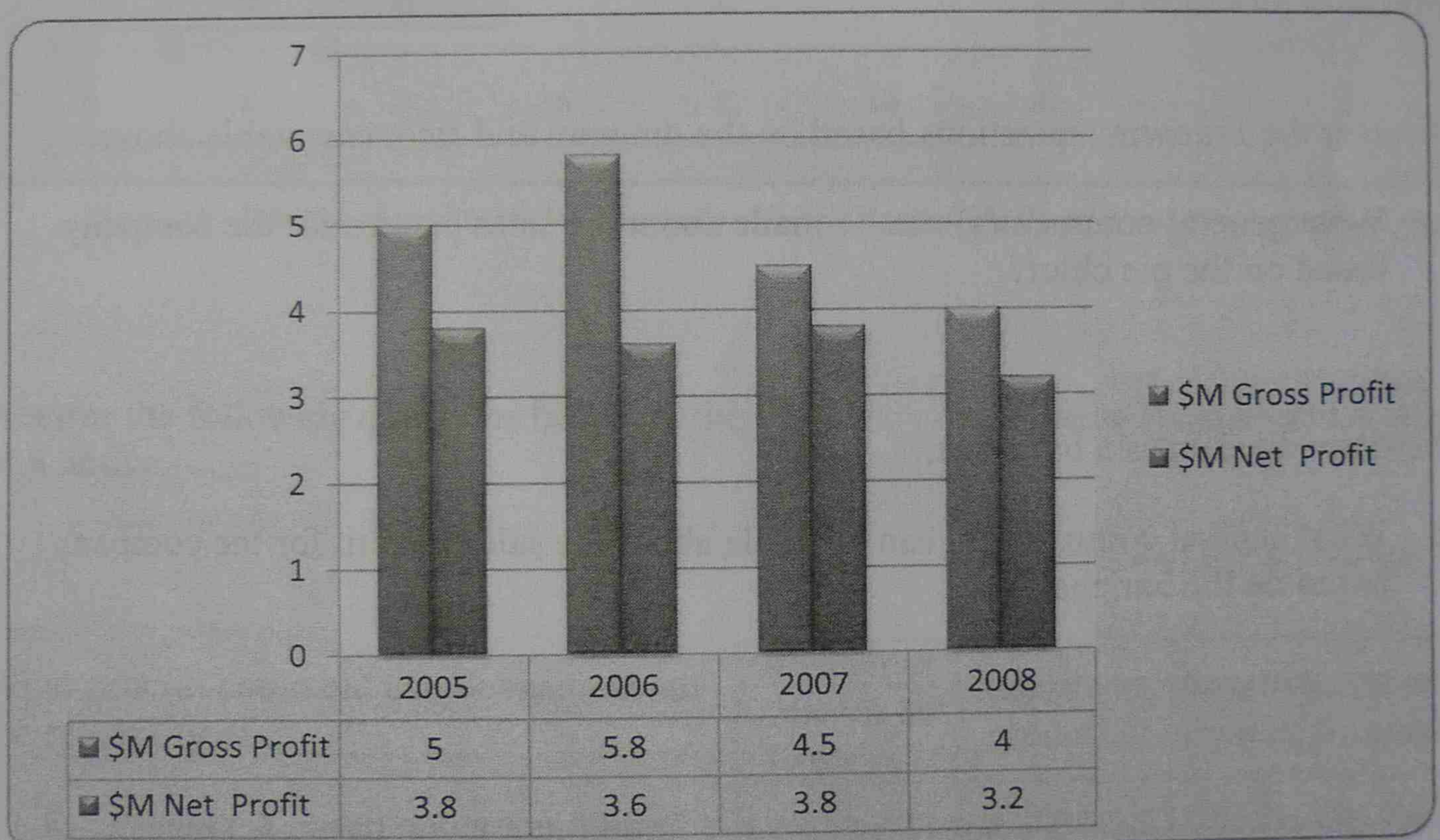


Figure 13 - Gross and Net profit for Engineering Solutions Pty Ltd

Answer the following questions using the information in the diagram above.

a. In what year did the company make the least net profit?
b. What comment can be made about the general trend of profit for the company?
c. Re-draw the diagram as a frequency polygon showing both variables.
d. Based on your frequency polygon, comment on the data for the years 2006 and 2007.

Activity 3b: Interpret statistical diagrams

Figure 14 - Percentage of sales in each quarter of 2008

	Sales in $M
1st Qtr	8.2
2nd Qtr	2.5
3rd Qtr	4.6
4th Qtr	1.2

Answer the following questions based on the diagram and frequency table above.

a. What general comment(s) can be made about the sales pattern for the company based on the pie chart?

Re-draw the data as a bar chart.

b. What general comment (s) can be made about the sales pattern for the company based on the bar chart?

Activity 3c: Interpret statistical diagrams

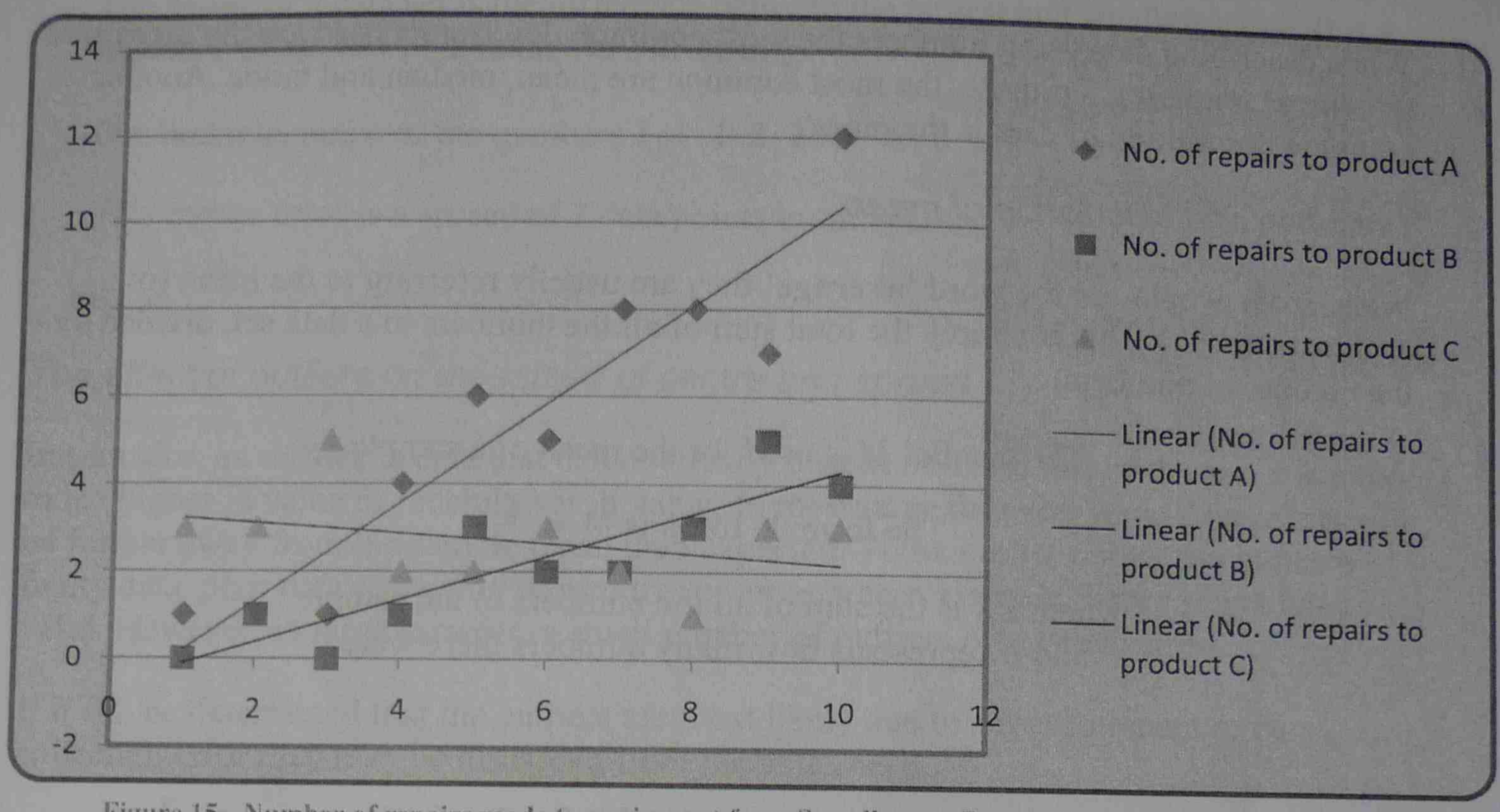


Figure 15 - Number of repairs made to equipment from Suppliers A, B and C based on months of service

Answer the following questions based on the frequency distribution (below) and scatter plot above.

What general comment can be made about the reliability of...

a. Product A based on its trend line?

b. Product B based on its trend line?

c. Product C based on its trend line

d. Product A compared to Product B?

e. Product C compared with Products A and B?

f. If you were asked to recommend one of the three products for use, what would your response be?

Number of months in service	Repairs to product A	Repairs to product B	Repairs to product C
1	1	0	3
2	1	1	3
3	1	0	5
4	4	1	2
5	6	3	2
6	5	2	3
7	8	2	2
8	8	3	1
9	7	5	3
10	12	4	3
Total	53	21	27

Describing univariate data numerically

When describing data using numbers the most common descriptors used are the **measures of central tendency**. Of those, the most common are mean, median and mode. Another common description of data is the range.

Definition and calculation of mean

When most people use the word '**average**' they are usually referring to the mean (or arithmetic mean). This is simply the total sum of all the numbers in a data set, divided by the number of numbers.

The symbol M is used for the mean of a sample.

The formula for m is $M = \frac{\Sigma X}{N}$

where ΣX is the sum of all the numbers in the sample
and N represents how many numbers there were.

For example, the mean of the numbers $1+1+1+2+2+3+6+8 = \frac{24}{8} = 3$

The mean is the central point where the sum of the data above that value equals the sum of the data below that value.

Definition and calculation of median

The **median** is the middle data point in a data set and is not related at all to the value of that data point. The median is found by locating the data point that is equidistant from either end of the data set. If there are an equal number of data points in the set, the median is the average of the two central points.

For example, the median of the numbers **1 1 1 2 2 3 6 8** is the number **2**.

The median is the point where there are an equal number of data points above and below.

Definition and calculation of mode

The **mode** is the most common data point in a data set. This is the value that occurs with greatest frequency.

For example, the mode of the numbers **1 1 1 2 2 3 6 8** is the number **1**. It occurs three times, which is more frequently than any other number in the set.

If two data points have the highest and equal frequency, the data is termed **bimodal** – that means it has two modes. If there are more than two modes, the data set is **multi-modal**.

Definition and calculation of range

The **range** of a data set is the difference between the largest and smallest value. It is a **measure of spread** or **measure of variability**.

For example, range of the numbers **1 1 1 2 2 3 6 8** is $8 - 1 = 7$.

This means there is a spread of 7 data points in the set.

The effect of outliers on measures of centre and spread

In statistics, an **outlier** is data that is numerically distant from the rest of the set. It is either much higher in value or much lower in value. In most large data sets some data points will be further away from the sample mean than expected. These outlier points can indicate faulty data, poor data collection procedures or areas where a certain theory might not be valid. However, in large samples, a small number of outliers is to be expected.

If it can be determined that the outliers are most likely due to a measurement or data collection error they may be discarded from the data set.

Outliers, being the most extreme observations, will include the sample maximum or sample minimum, or both. However, the maximum and minimum values in the data set are not outliers if they are not unusually far from other the rest of the data.

If a data set is analysed without any consideration of outliers, the results can be misleading. For example, if calculating the mean temperature of 10 objects in a kitchen and most are between 20 and 25 degrees Celsius, but the oven is at 150 °C, the median of the data may be 23 °C but the mean temperature may be between 35.5 and 40 °C. In this example the median is a better measure of central tendency than the mean. In this case, the outlier (temperature of the oven) is a data point that doesn't belong to the data set.

Some statistics are 'robust' or 'resistant to outliers'.

- The *median* is a robust measure of central tendency but the *mean* is not.
- The *median absolute deviation* and *interquartile range* are robust measures of dispersion but the *standard deviation* and *range* are not.

Review questions – describing univariate data numerically

Activity 4a: Describe data numerically

Twenty rolls of cable were weighed to the nearest ½ kilogram. The results are shown in the table below. Use the data to create a frequency distribution.

64.5	65.0	64.0	68.5	62.0	65.5	71.5	70.0	69.0	67.5
63.5	66.5	68.0	72.0	69.5	66.0	66.0	68.5	64.5	65.0

From the frequency distribution calculate the

a. Range b. Mean c. Median d. Mode

Activity 4b: Describe data numerically

A collection of 50 circuit breakers were tested for response times. The results are given below correct to the nearest 100^{th} of a second. Use the data to create a frequency distribution.

0.44	0.32	0.31	0.47	0.27	0.31	0.40	0.28	0..16	0.26
0.33	0.46	0.41	0.33	0.31	0.28	0.38	0.29	0.17	0.26
0.29	0.40	0.29	0.24	0.41	0.22	0.25	0.47	0.31	0.36
0.49	0.21	0.42	0.43	0.28	0.36	0.24	0.37	0.34	0.27
0.49	0.16	0.29	0.30	0.41	0.27	0.29	0.28	0.40	0.42

From the frequency distribution calculate the

a. Range b. Mean c. Median d. Mode

Activity 5: Compare data sets

45	48	46	49	49	42	40	42	35	43	47	46	41	48	43	45	42	49
45	48	43	48	49	41	40	45	40	44	46	46	47	48	43	45	44	49

The table above represents two data sets: green and blue.

a. For each data set, calculate the range, median, mode and mean
b. Draw comparison box plots to compare the data sets
c. What comments can be made about the two data sets based on your calculated and graphical measures of spread and central tendency?

Activity 6: Determine effect of outliers

Calculate the mean, median and range for the following sets of numbers.

Set A	6	8	11	5	2	9	7	8	Mean =	Median =	Range =
Set B	1	3	0.5	0.6	2	2.5	3.1	3	Mean =	Median =	Range =

a. How many of the data values were used to calculate the mean?

b. How many of the data values were used to calculate the range?

c. Which measure(s) (mean, median and/or mode) was/were influenced by outliers in the data set? Explain your answer.

d. Would the mode of a data set be influenced by outliers?

Activity 7: Choosing appropriate statistical measures

A repair crew keeps a record of call outs over a period of a month. The data is shown at right.

a. Calculate the mean, median, mode and range of the data set.

b. Which of the statistical measures would you use to estimate how many staff need to be on duty each day and why?

Day of month	Call outs
Monday	25
Tuesday	33
Wednesday	22
Thursday	34
Friday	40
Saturday	57
Sunday	48
Monday	36
Tuesday	35
Wednesday	27
Thursday	33
Friday	44
Saturday	53
Sunday	47
Monday	38
Tuesday	39
Wednesday	34
Thursday	34
Friday	45
Saturday	45
Sunday	60
Monday	44
Tuesday	33
Wednesday	35
Thursday	34
Friday	34
Saturday	45
Sunday	45

Analysing univariate data

Measures of spread, dispersion and variability

The previous section introduced the mean, median and mode – measures of central tendency. Also introduced was the range – a measure of spread or dispersion.

Measures of central tendency can be misleading on their own. For example, knowing that the average price of a collection of components is $390 is not meaningful if there is a small number of components that cost over $1000 and a larger number that cost around $50.

Range is one measure that helps provide better understanding of a data set. Knowing that the average cost is $390 ***and*** that the range is from $48 to $1200 is much more descriptive and meaningful.

In order to calculate measures of dispersion, it is necessary to understand some terminology.

If a data set is sequenced from lowest to highest value it is known as an **ordered data set**.

In an ordered data set the 25% of the data points at the lower end of the range are known as the **lower quartile** (LQ). The 25% of the data points at the upper end of the range are known as the **upper quartile** (UQ).

Definition and calculation of inter-quartile range

Like the range, the inter-quartile range is an easy measure of dispersion to calculate and is ***not*** very sensitive to outlying data (very small or very large values in the data set). Using an ordered data set, the inter-quartile range can be calculated by finding the difference between the lower and upper quartile values.

The steps to find the IQR are as follows.

1. Order the data set
2. Take the median of the data set
3. Find the medians of the upper and lower halves of the set
4. The inter-quartile range is the difference between these two secondary medians

In mathematical terms this is written as $IQR = UQ - LQ$

For example, consider the following scores achieved by Diploma students in a revision quiz.

3	3	3	3	4	5	6	6	6	6	7	7	8	8	8	8	9	9	9	10

The median of the set is 6.5. The UQ is 8.5 and the LQ is 4.5.
Therefore the IQR = 8.5 – 4.5 = 4
This means that the middle 50% of scores in the quiz span a four mark range.

Definition and calculation of variance

If all the data points in a set are clustered around the mean, there is little variability in the data. The more widely the data are scattered the larger the variance of the data set. Calculations of variance are sensitive to outlying data - very small or very large values in the data set.

Variance is the approximate average of the squared deviations from the mean of the data set. The formula looks challenging but in fact the calculation is quite simple if the steps are followed.

$$S^2 = \frac{\Sigma(\bar{x}-x^i)^2}{n-1}$$

1. Calculate the mean of the data set
2. Determine the variation of each data point from the mean
3. Square each deviation
4. Find the sum of the squares
5. Divide by the number of data points – 1 (n-1)

Consider data set used in the previous example.

3	3	3	3	4	5	6	6	6	6	7	7	8	8	8	8	9	9	9	10

The sum of the data points is 128 therefore the mean is 128 ÷ 20 = 6.4

Data point	Deviation from mean	Squared deviation
3	6.4 – 3 = 3.4	11.56
3	6.4 – 3 = 3.4	11.56
3	6.4 – 3 = 3.4	11.56
3	6.4 – 3 = 3.4	11.56
4	6.4 - 4 = 2.4	5.76
5	6.4 – 5 = 1.4	1.96
6	6.4 – 6 = 0.4	0.16
6	6.4 – 6 = 0.4	0.16
6	6.4 – 6 = 0.4	0.16
6	6.4 – 6 = 0.4	0.16

Data point	Deviation from mean	Squared deviation
7	6.4 – 7 = -1.4	1.96
7	6.4 – 7 = -1.4	1.96
8	6.4 – 8 = -2.4	5.76
8	6.4 – 8 = -2.4	5.76
8	6.4 – 8 = -2.4	5.76
8	6.4 – 8 = -2.4	5.76
9	6.4 – 9 = -3.4	11.56
9	6.4 – 9 = -3.4	11.56
9	6.4 – 9 = -3.4	11.56
10	6.4 – 10 = -4.4	19.36

The sum of the squares is 135.6

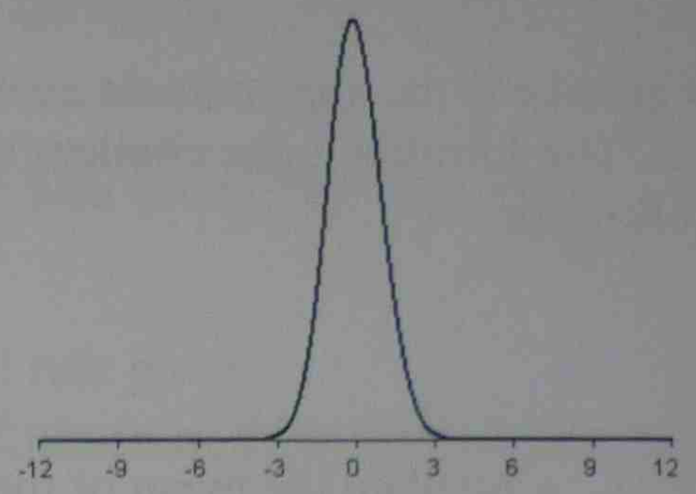
$$135.6 \div (n\text{-}1) = 125.6 \div (20 - 1) = 6.61$$

Therefore the variance of the data set is 6.61

Definition and calculation of standard deviation

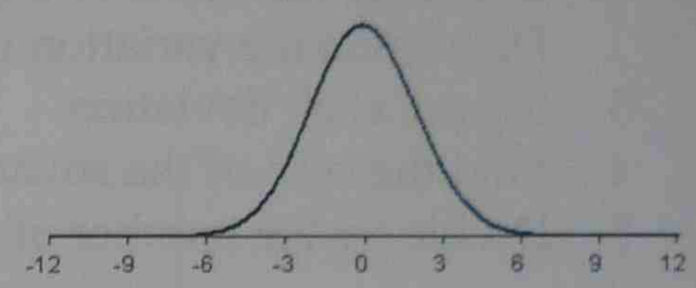
Standard deviation is the square root of the variance. Hence it is also known as the Root Mean Squared Deviation (RMSD). The symbol often used for standard deviation is σ (sigma). Calculations of standard deviation are sensitive to outlying data.

One way to think about standard deviation is that it measures how spread out the points in a data set are. If our data set follows a bell curve and had a standard deviation of 1 it would look like this. ➔

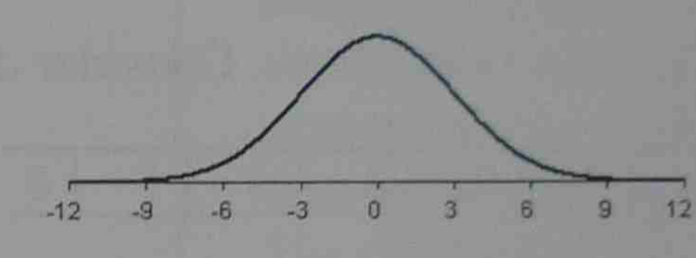


It is quite concentrated around the middle.

However, a data set with the same mean, median and range but a standard deviation of 2 it would look like this. ➔



Finally, a data set with the same mean, median and range but a standard deviation of 3 would look like this. ➔



Knowing the variance of a data set is just one step short of knowing the standard deviation. Standard deviation is the square root of the variance.

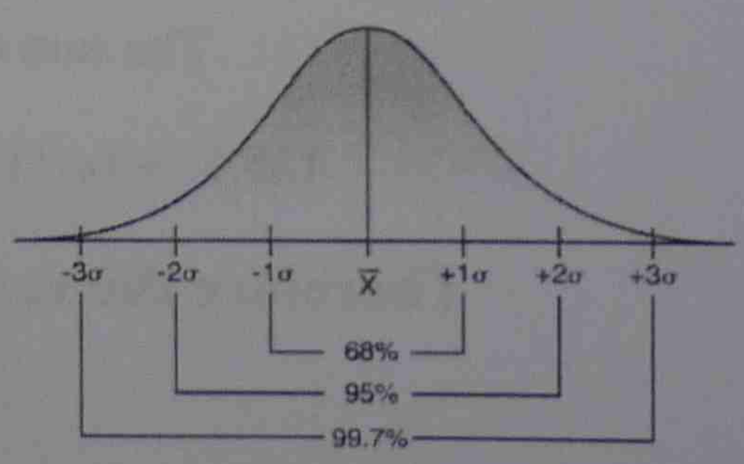
$$\sigma = \sqrt{S^2} = \sqrt{\frac{\sum(x^i - \bar{x})^2}{n-1}}$$

For our previous example $\sigma = \sqrt{6.61} = 2.57$

Normal distribution curve and the empirical rule

In statistics, the normal distribution curve or bell curve represents the shape of data that clusters around the mean of the data set. In theory, any variable that is made up of a large number of independent factors is likely to be distributed in a bell curve.

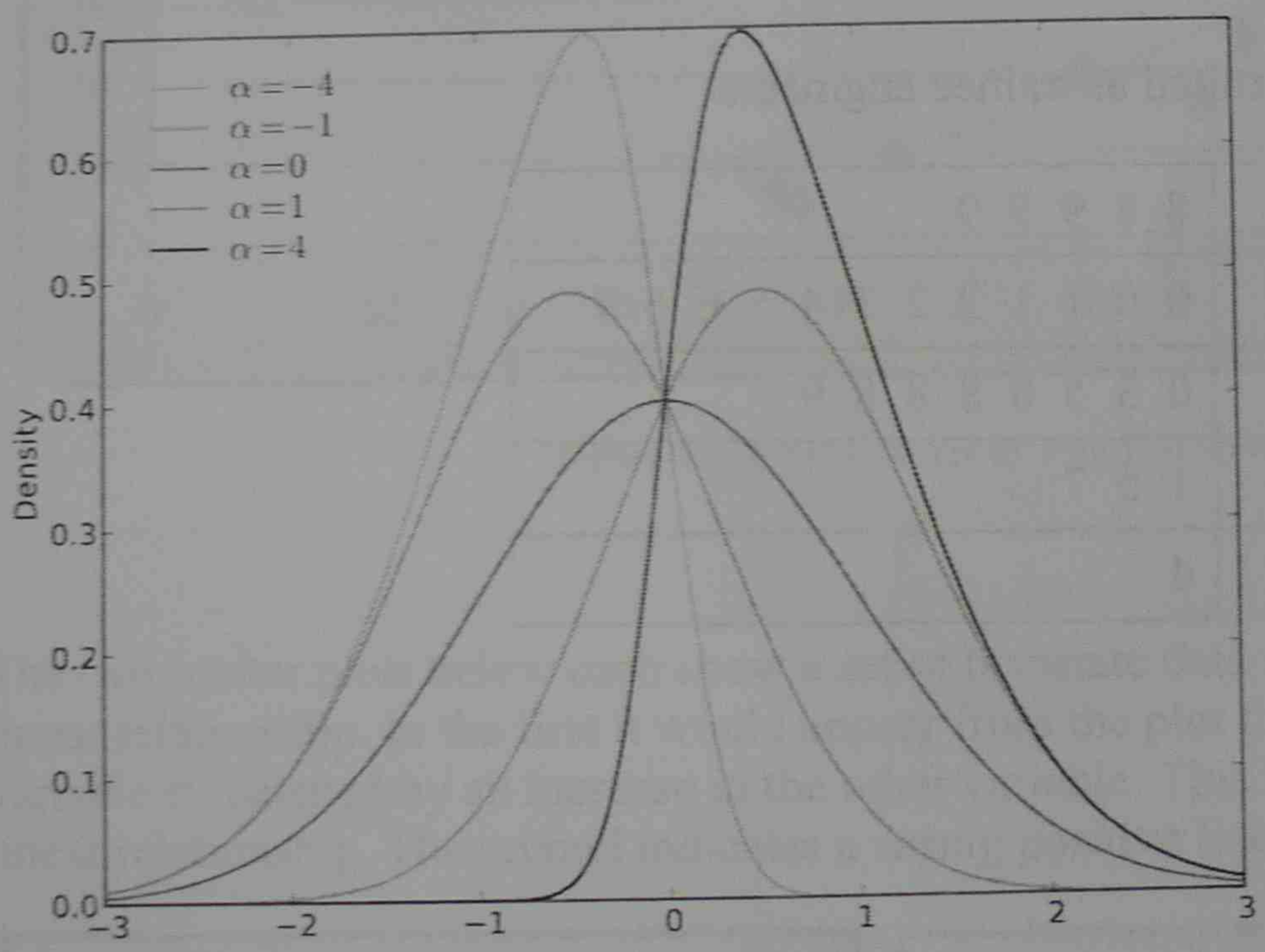
For example, if the weight of all persons living in a town was graphed the result will look like this. Some of the values will be at the extreme ends of the distribution and most will be clustered around the mean.



The Empirical rule states that approximately...

- 68% of values in a normally distributed data set will be clustered within one standard deviation of the mean $\mu + 1\sigma$
- 95% of values in a normally distributed data set will be within two standard deviations of the mean $\mu + 2\sigma$
- 99.7% of the values in a normally distributed data set will be within three standard deviations of the mean $\mu + 3\sigma$

To see an example calculation proving the Empirical rule go to http://www.nku.edu/~statistics/212_Justifying_the_Empirical_Rule.htm (accessed in June 2009)



I

If the data in a set have the shape of a normal distribution but are not centred on the mean, the data is said to be skewed. The diagram at left shows some examples of skewed bell curves.

Review questions – analysing univariate data

Activity 8 – Calculating IQR, S^2 and σ

For each of the following data sets, calculate interquartile range (IQR), variance (S^2) and standard deviation (σ).

Data set (a)

3	3	6	3	4	5	6	9	6	6	7	7	8	8	8	8	9	9	9	10

Data set (b)

Ages of employees when recruited as trainee engineers

1	8 8 9 9 9
2	0 0 1 1 2 2 3 4 5 6 6 7
3	0 3 5 8 8 8 8 9
4	1 5 7
5	4

Data set (c)

Average voltage measured across consumer mains of rural customers in a single 24 hour period

22	3 5 5 6 7 8 8 9 9
23	0 1 2 3 5 5 6 6 7 7 7 8 9 9 9 9
24	0 1 1 1 1 2 2 2 3 3 4 5 6 6 7

Analysing bivariate data

Linear relationships between variables

A linear 'relationship' between variables means that both change in a pattern consistent with each other. Generally we can see a relationship by using a scatter plot.

Consider the following example. This scatter plot would indicate that there is a positive linear relationship between the age of an employee and how long they have been employed by the organisation.

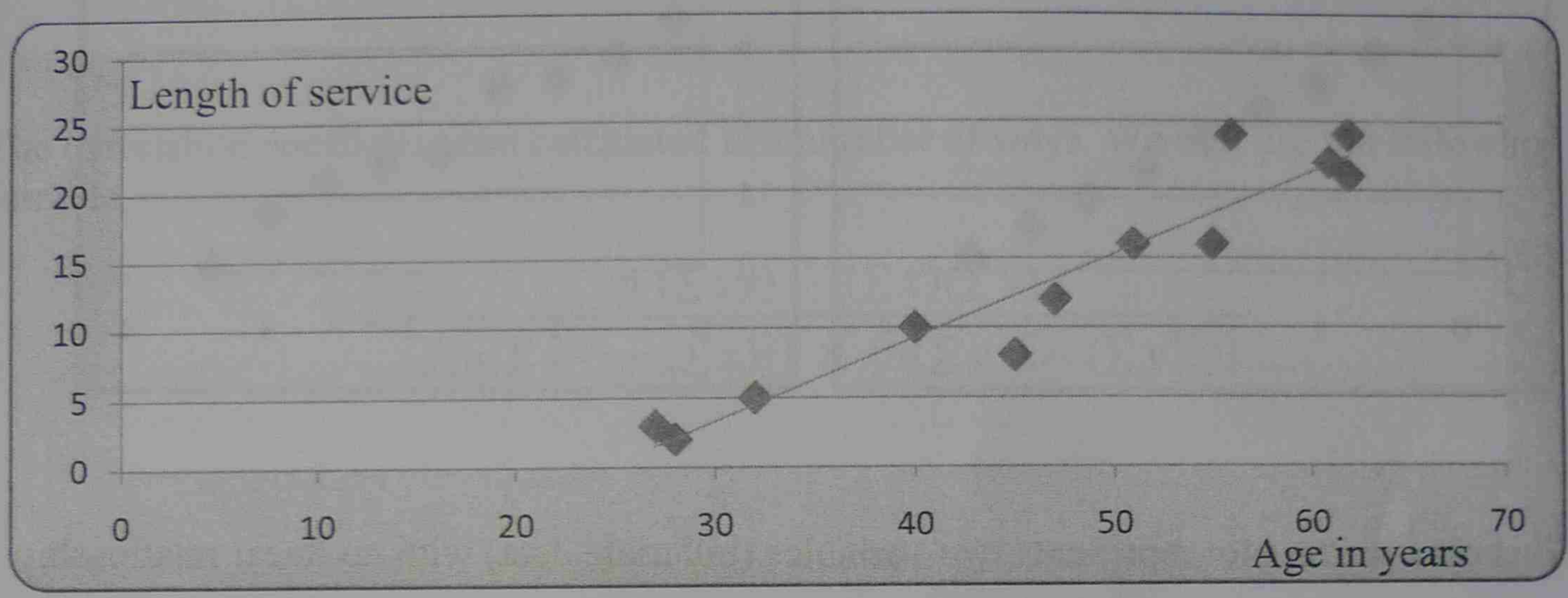
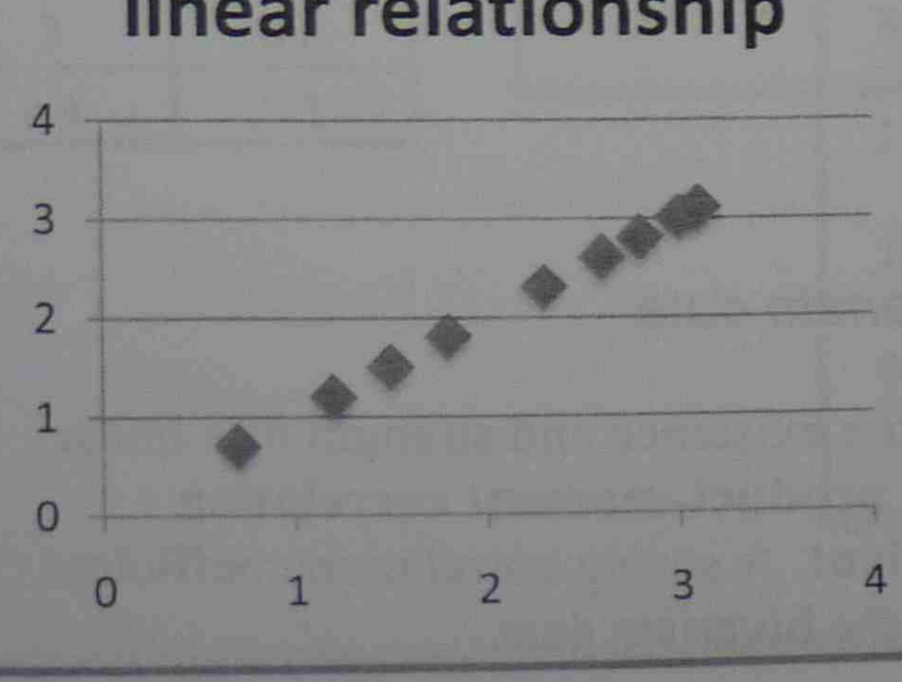


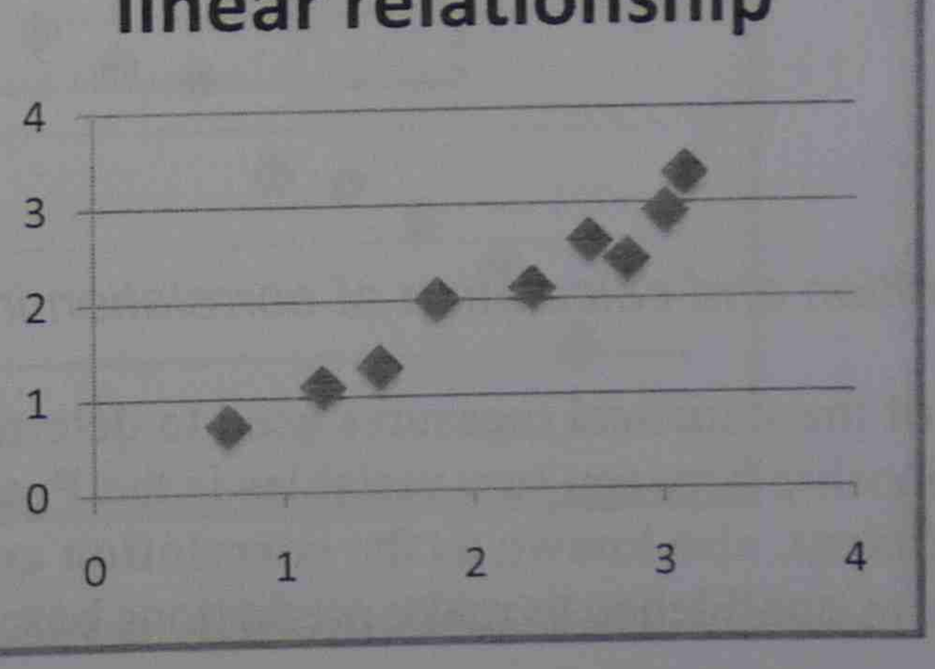
Figure 16 - Employee age by length of service

The two scatter plots below each show a set of bivariate data with a degree of positive linear relationship. In the first it would appear from the plot that any increase in one variable is matched by an increase in the other variable. This is known as a perfect positive linear relationship. The second indicates a strong positive linear relationship.

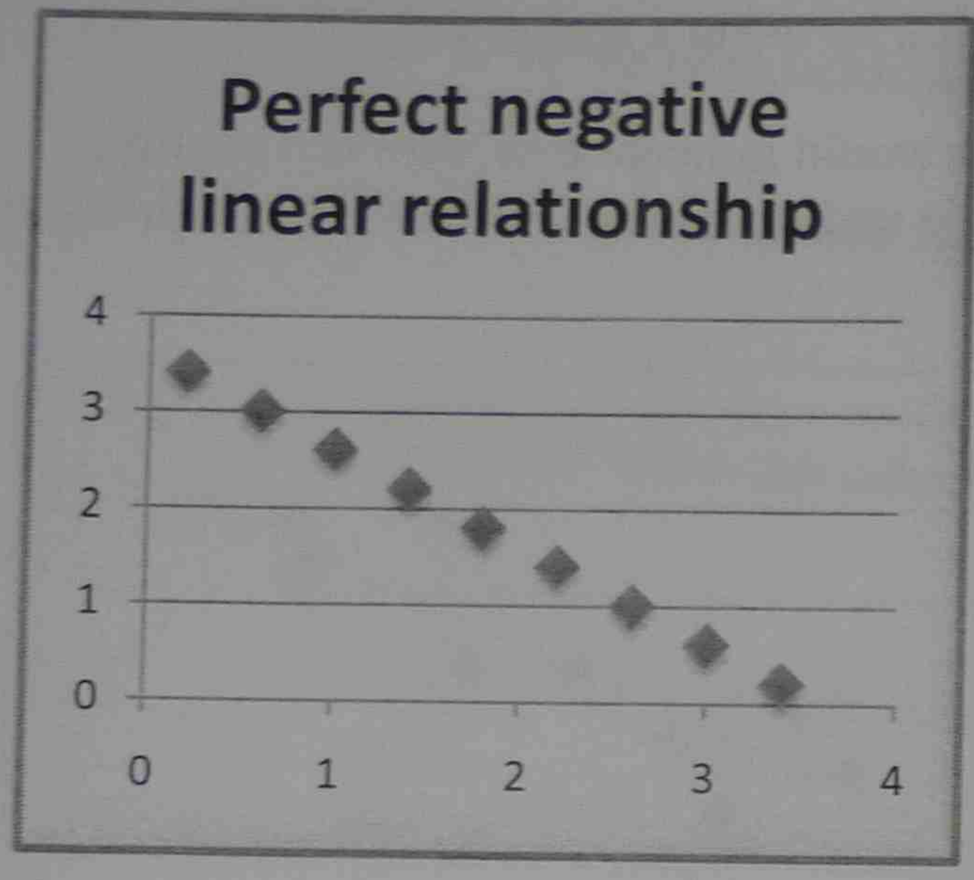
Perfect positive linear relationship

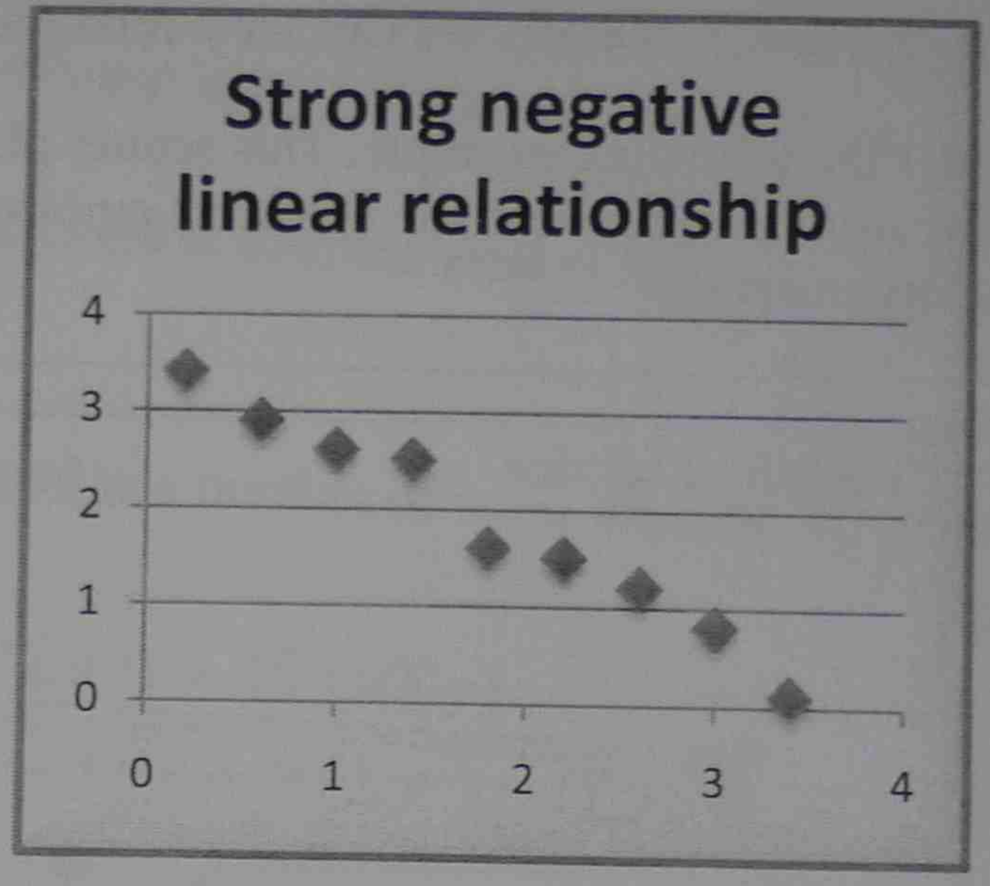


Strong positive linear relationship

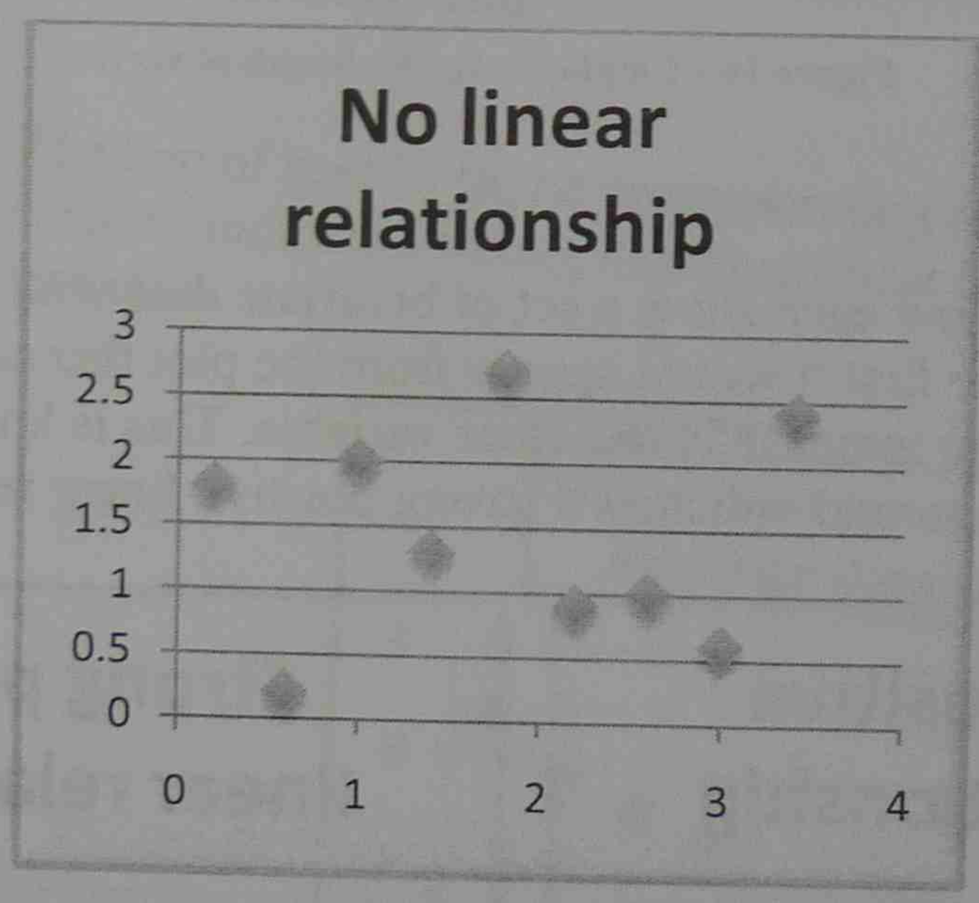


The two scatter plots below each show a set of bivariate data with a degree of negative linear relationship. In the first it would appear from the plot that any increase in one variable is matched by a decrease in the other variable. This is known as a perfect negative linear relationship. The second indicates a strong negative linear relationship.





This next scatter plot represents two variables (bivariate data) with no linear relationship.



Definition and calculation of correlation in bivariate data

One of the statistical measures used to determine the existence and strength of a linear relationship between two variables is the **Pearson product-moment correlation coefficient**, also known as the **correlation coefficient**. A strong correlation coefficient gives us confidence to make predictions based on the bivariate data.

NOTE: The correlation coefficient does not indicate whether the linear relationship is causal. It merely tells us the strength of the linear relationship.

Summary of correlation coefficient indications

- A correlation coefficient will always be a value between -1.0 and +1.0
- A correlation coefficient of +1.0 indicates a perfect positive linear relationship
- A correlation coefficient of -1.0 indicates a perfect negative linear relationship
- The closer the correlation coefficient is to 1.0, the stronger the positive linear relationship is
- The closer the correlation coefficient is to -1.0, the stronger the negative linear relationship is
- The closer the correlation coefficient is to 0, the weaker the linear relationship is

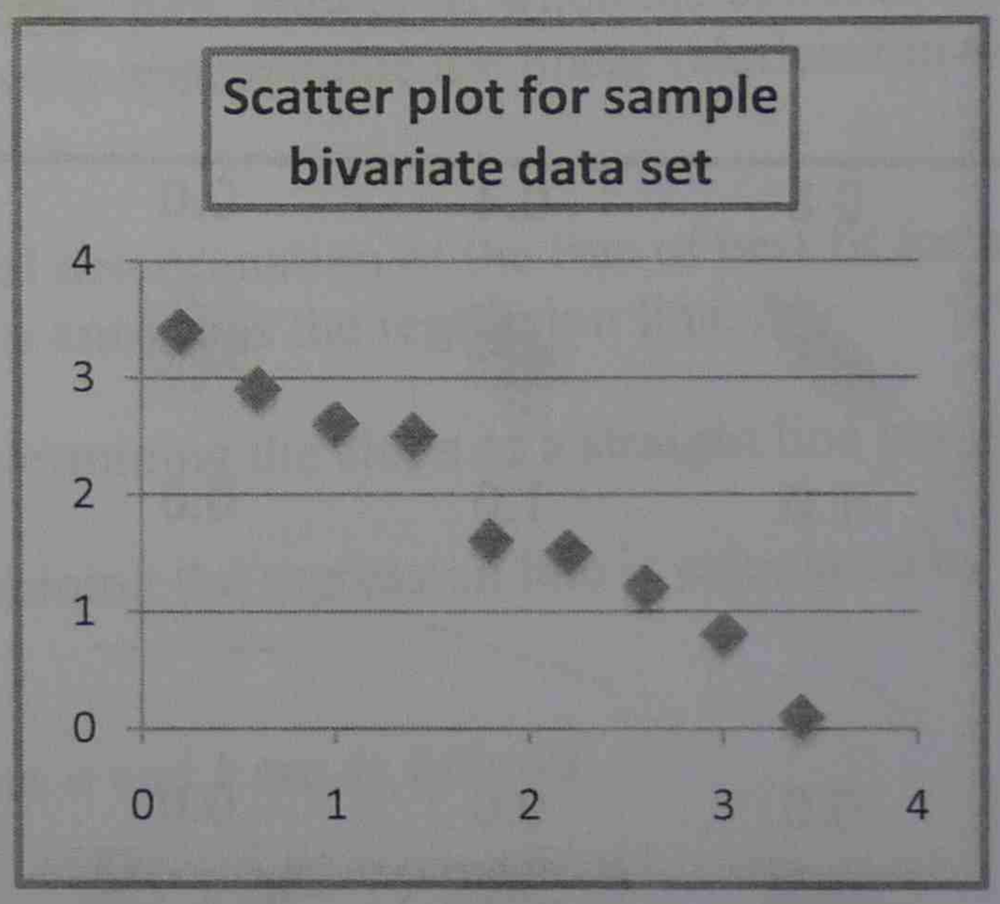
The correlation coefficient can calculated in a number of ways. We will use the following formula.

$$r = \frac{n\,(\sum xy) - (\sum x)(\sum y)}{\sqrt{[n(\sum x^2) - (\sum x)^2] \times [n(\sum y^2) - (\sum y)^2)]}}$$

Consider the following set of bivariate data.

x	y
0.2	3.4
0.6	2.9
1	2.6
1.4	2.5
1.8	1.6
22	1.5
2.6	1.2
3	0.8
3.4	0.1

The line of best fit in the corresponding scatter plot (below) would seem to indicate a strong negative linear relationship between the variables.



We can test the strength of the relationship by calculating the correlation coefficient.

The first step is to use the data to create a chart to help with the calculations.

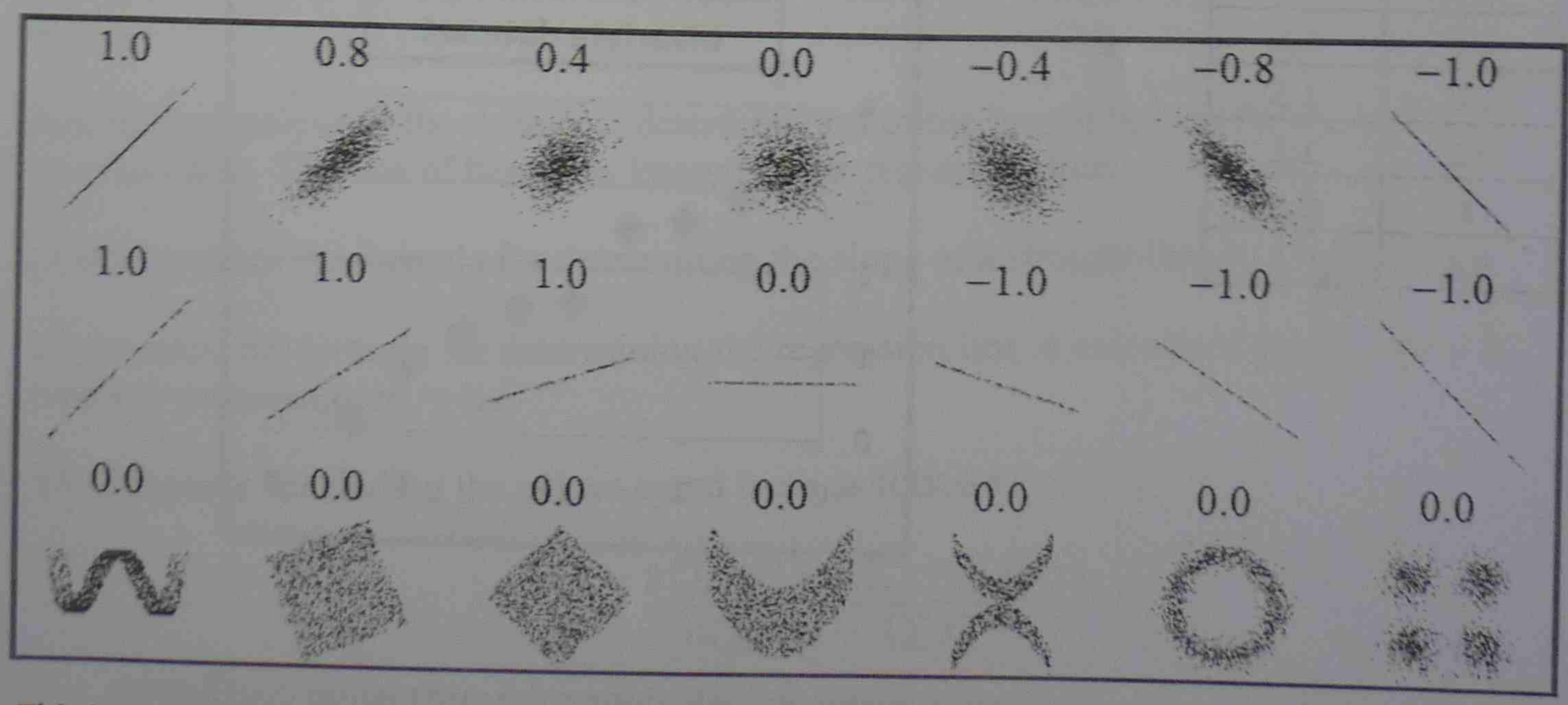
x	y	xy	X^2	Y^2
0.2	3.4	0.68	0.04	11.56
0.6	2.9	1.74	0.36	8.41
1	2.6	2.6	1	6.76
1.4	2.5	3.5	1.96	6.25
1.8	1.6	2.88	3.24	2.56
22	1.5	33	484	2.25
2.6	1.2	3.12	6.76	1.44
3	0.8	2.4	9	0.64
3.4	0.1	0.34	11.56	0.01
Σx	Σy	$\Sigma(xy)$	Σx^2	Σy^2
36	16.6	50.26	517.92	39.88

Applying the data to the formula gives the following equation and result.

$$r = \frac{9\,(50.26) - (36)(16.6)}{\sqrt{[9(517.92) - (36^2)] \times [9(39.88) - (16.6)^2)]}} = -0.955$$ [a strong negative linear relationship]

The table at right indicates a 'rule of thumb' for interpreting the correlation coefficient for a set of bivariate data. The images below show scatter plots that represent various values of correlation coefficient.

Correlation	Negative	Positive
Weak	−0.3 to −0.1	0.1 to 0.3
Medium	−0.5 to −0.3	0.3 to 0.5
Strong	−1.0 to −0.5	0.5 to 1



This is a public domain image from http://en.wikipedia.org/wiki/File:Correlation_examples.png

Correlation and causation

The correlation coefficient tells us about the linear relationship between an independent variable (x) and the dependent variable (y). However, it does not tell us the nature of that relationship.

When a linear relationship has been established a range of possibilities must be considered. These are as follows.

- The independent variable may cause the dependent variable – a cause and effect relationship. For example, an increase in temperature may cause an increase in product failure.
- There may be a reverse cause and effect relationship. For example, drinking alcohol may cause a decrease in driver reaction time.
- The relationship may be coincidental or due to chance. For example, if sales of multimeters had a strong correlation to the number of persons visiting the local museum, it must be assumed that the relationship is coincidental.
- Finally, the relationship may be due to 'confounding'. This means there are other variables in play and the relationship is actually due to the interrelationships of all those variables.

It is up to the engineer to interpret the data and the statistical analysis of the data. Assumptions should never be made. All conclusions must be based on sound reasoning.

Definition and calculation of regression

When the correlation coefficient is +1 or -1 there is a straight line that will pass through all data points when displayed on a scatter plot. However, when the correlation coefficient is strong but is not +1 or - 1, we can only approximate the linear relationship with a line of best fit.

Regression analysis is the statistical determination of the line of best fit for a set of bivariate data. The line of best fit is known as the regression line.

In mathematics the formula for determining the slope of a straight line is $y = mx+b$

In statistics, the formula for determining the regression line is calculated by $\hat{y} = ax + b$ where $\hat{y}$ is pronounced 'y hat'

The formulae for finding the values a and b are as follows.

$$a = \frac{n(\sum xy) - (\sum x)(\sum y)}{n(\sum x^2) - (\sum x)^2}$$

$$b = \frac{(\sum y)(\sum x^2) - (\sum x)(\sum xy)}{n(\sum x^2) - (\sum x)^2}$$

Consider the following problem. At a particular college, engineering students are asked to

be able to predict a student's grade in statistics on the basis of their math test. To do this they have the results from 5 randomly selected engineering students who have already completed both the math test and the statistics subject.

For this investigation the independent variable (x) is the maths test result. The dependent variable (y) is the grade gained in statistics.

The table below shows the data set and some preliminary calculations.

Student	x	y	xy	x^2	y^2
1	95	85	8075	9025	7225
2	85	95	8075	7225	9025
3	80	70	5600	6400	4900
4	70	65	4550	4900	4225
5	60	70	4200	3600	4900
	Σx	Σy	$\Sigma(xy)$	Σx^2	Σy^2
Sum	390	385	30500	31150	30275

Applying the data to the formula gives the following equation and result.

$$a = \frac{5(30500) - (390)(385)}{5(31150) - (390)^2} = 0.6438 \qquad b = \frac{(385)(31150) - (390)(30500)}{5(31150) - (390)^2} = 26.78$$

Using this information we can predict a new student's grade in statistics. For example, if a student scores 80% in the maths test we can calculate her grade in the statistics subject as follows.

$$\hat{y} = ax + b = (0.6438 \text{ x } 80) + 26.78 = 78.284\%$$

NOTE: It is important not to use the regression equation for any values outside the range of the values used to create the equation. Doing so is called **extrapolation** and can produce results that are not reliable.

Coefficient of determination

The coefficient of determination allows us to determine how confident we could be with any predictions made based on a given regression analysis. It measures how much of the variation in the dependent (y) variable can be explained by regression of the independent (x) value.

The symbol for the coefficient of determination is r^2. It is literally the square of the correlation coefficient and is always between 0 and 1. Note: $0 \leq r^2 \leq 1$

The close r^2 is to the value of 1, the more confidence we can have in the regression analysis. The closer r^2 is to 0, the less confidence we have in the regression analysis.

Review questions – analysing bivariate data

Activity 9 – Calculating r, ŷ and r

For each of the following data sets

a. Draw a scatter plot
b. Calculate the coefficient of correlation and comment on it's value
c. Calculate the coefficient of determination and comment on it's value
d. Complete the regression analysis to find the equation for the regression line
e. Determine $\hat{y}$ for the value highlighted in the data set table

Data set a

As part of an investigation into ways to cut costs, a senior management team asked a repair crew to drive their vehicle at various average speeds and calculate the fuel consumption. The following table shows the results.

Speed Km/hour	30	40	50	60	70	80	**85**	90	**95**	100	110
Fuel consumption Km/litre	16.58	16.58	15.31	13.6	11.48	10.2		9.35		8.54	7.9

Data set b

A group of eight technicians decided to increase their fitness by doing an exercise program during their lunch hour. They plotted their decrease in time taken to run a 5km course over the period of time each was involved in the program. The results are as follows.

No. of weeks in program	1	3	3	4	5	5	5	**6**	7	7	7
Reduction in 5km time in minutes	0.1	3.2	3.3	3.9	5.1	5.2	5.3		6.2	6.5	7

Data analysis for reports

Quick review of simple data analysis

Data analysis is converting data into a usable format then drawing out themes, patterns and other information.

When data is collected it can be in the form of nominal, ordinal or interval scales.

A nominal scale is where the response to a question is done by 'nominating' one of the available choices.

e.g. gender, occupation, geographic area, nationality

Ordinal scales are used for placing responses in 'order' or ranking.

e.g. first to last, highest to lowest, excellent-good-fair-poor-very poor

Interval scales allow for responses between two values at any point along pre-set intervals. They are sometimes called 'measurement data'.

e.g. temperature, weight, quantity, distance, % marks

Nominal and ordinal scales are considered 'descriptive' because statistical analysis can't be applied to the data itself. On the other hand, data collected on an interval scale can be statistically analysed.

Data analysis may occur after data collection is complete HOWEVER it must be planned before any data is collected. How data is collected and recorded has a significant impact on the analysis of that data.

Data coding is the application of alpha or numeric symbols to responses so they can be sorted, categorised and analyses by computer.

Data editing is checking the data collected for errors, omissions, illegibility or inconsistency. Irresolvable data may have to be excluded from the data analysis.

In terms of statistical analysis, we're particularly interested in measures of central tendency. These include: range, mean, median, mode, standard deviation and variance.

- Range is the difference between the extreme values (e.g. highest – lowest)
- Mean is the average of the data values.
- Median is the midpoint of the distribution (NOT the centre of the range)
- Mode is the most frequently occurring value in the distribution.

Revision activity 1 – interpreting data from a frequency distribution

Countryside Council members were concerned by the number of motor vehicle accidents occurring in the township. They decided they needed to do some research to help them come up with strategies to reduce the number of accidents.

The local Police Sergeant, Roads & Traffic Authority and Registrar of Births Deaths & Marriages were asked for information. This was summarised as follows.

Year	Motor Vehicle Accidents	Motor Vehicle Registrations	Population	Accidents per vehicle	Accidents per capita
1990	40	1274	4020		
1991	42	1328	4132		
1992	53	1360	4196		
1993	48	1412	5006		
1994	50	1475	5091		
1995	53	1522	5133		
1996	70	1607	5180		
1997	56	1641	5199		
1998	54	1656	5217		
1999	50	1697	5232		
Mean $\bar{x}$					

In preparation for the next Council meeting you are asked to do the following data analysis.

For each year

- ☐ the number of accidents per vehicle (calculate to 5 decimal places)
- ☐ the number of accidents per capita (calculate to 6 decimal places)

What year had the

- ☐ least number of accidents per vehicle
- ☐ least number of accidents per capita
- ☐ most number of accidents per vehicle
- ☐ most number of accidents per capita

The mean of

- ☐ accidents per year
- ☐ population per year
- ☐ accidents per capita
- ☐ registrations per year
- ☐ accidents per vehicle

How many years were above the mean of

- ☐ accidents per vehicle
- ☐ accidents per capita

What other information would you recommend the Council research and consider before developing any strategies to reduce the number of motor vehicle accidents?

Revision Activity 2a – Choosing data sources and collection methods

The local supply authority's call centre has recorded an increase in the time taken to resolve calls for assistance form their clients.

List at least 4 types of information the call centre manager will need in order to identify the problem.

Describe how this information would be collected.

Revision Activity 2b – Choosing data sources and collection methods

A electrical wholesaler is considering opening a retail outlet in a neighbouring town.

List at least 4 types of information the owner would need in order to make this decision.

Describe how this information would be collected.

Revision Activity 2c – Choosing data sources and collection methods

A retail electricity supplier has been using a variety of advertising mediums: radio, local paper and letter box drops. The manager wants to know which medium results in the most sales per $ spent.

List at least 3 types of information the manager would need to be able to identify the most effective medium.

Describe how this information would be collected.

Revision Activity 3 – Analysing data

A new car sales team has been given the following frequency distribution showing average vehicle sales per week over the past two months. Use the information in the table to answer the questions below.

Day	**Model A** $25,000 per unit	**Model B** $30,000 per unit	**Model C** $22,000 per unit
Monday	6	4	4
Tuesday	5	4	4
Wednesday	6	3	4
Thursday	6	4	3
Friday	5	3	3
Saturday	3	4	8
Sunday	4	3	9

Determine

- ☐ which model is most popular,
- ☐ which model results in the highest $$ sales

Display in the form of a graph, chart or similar, the following.

- How many of each model is sold as a % of total weekly sales.
- The $ value of each model sold as a % of total weekly sales.
- Daily fluctuations in the total number of vehicles sold each week.
- Daily fluctuations in the total $ value of sales each week.
- Daily fluctuations in sales of each model.
- Daily fluctuations in $ value of sales of each model

Revision Activity 4 – Drawing conclusions from data

Assume you are the sales manager viewing the information presented in your response to Revision Activity 3.

- Are there any conclusions you can draw from the information?
- Does the information prompt you to ask any further questions about sales?

Researching the report

The Report Brief

Planning a report is essential. A report is like a project - it has a specific purpose, a time deadline for completion and other people rely on it.

If you've been requested to produce a report your first step is to ask for a report brief. This is the outline of the…

- issue to be reported on,
- intended audience,
- expected length of the document,
- any other particulars to be included.
- purpose of the report,
- scope of included information,
- deadline for completion, and

If you're not given a report brief, you should develop one of your own and when the brief has been discussed and agreed to, have it signed off by the person requesting the report. Getting the brief agreed to and signed off is essential. It makes sure that everyone understands exactly what has been asked for.

The Issue

Reports centre round an issue. This is the reason for the report or the problem being investigated. Usually the issue is identified in the brief however, this isn't always the case.

For example, imagine the leader of a work team asks you to investigate the poor quality of report-writing her team produces. The team leader believes her team members make mistakes, don't think clearly and as a result, produce unusable reports. She says they need some training.

As you investigate you find that the problem isn't the team members -it's the team leader. She gives no feedback other than to point out their mistakes. She also makes unrealistic demands of them without providing appropriate support. So, the ***real*** issue is not the writing ability of the team members but the poor supervisory skills of the team leader.

This example shows that sometimes the person requesting a report doesn't really know what the issue is. Other times this is deliberate and indicates a ***hidden agenda*** on the part of the person requesting the report. Either way, it shows how important it is to try and find out the real issue.

Figuring out the real issue is like a doctor diagnosing a patient. The patient will describe one or more symptoms and the doctor uses these to identify the condition. Your job as a report writer is similar. The person requesting the report may not know the real issue.

Revision Activity 3 – Analysing data

A new car sales team has been given the following frequency distribution showing average vehicle sales per week over the past two months. Use the information in the table to answer the questions below.

Day	**Model A** $25,000 per unit	**Model B** $30,000 per unit	**Model C** $22,000 per unit
Monday	6	4	4
Tuesday	5	4	4
Wednesday	6	3	4
Thursday	6	4	3
Friday	5	3	3
Saturday	3	4	8
Sunday	4	3	9

Determine

- ☐ which model is most popular,
- ☐ which model results in the highest $$ sales

Display in the form of a graph, chart or similar, the following.

- How many of each model is sold as a % of total weekly sales.
- The $ value of each model sold as a % of total weekly sales.
- Daily fluctuations in the total number of vehicles sold each week.
- Daily fluctuations in the total $ value of sales each week.
- Daily fluctuations in sales of each model.
- Daily fluctuations in $ value of sales of each model

Revision Activity 4 – Drawing conclusions from data

Assume you are the sales manager viewing the information presented in your response to Revision Activity 3.

- Are there any conclusions you can draw from the information?
- Does the information prompt you to ask any further questions about sales?

Researching the report

The Report Brief

Planning a report is essential. A report is like a project - it has a specific purpose, a time deadline for completion and other people rely on it.

If you've been requested to produce a report your first step is to ask for a report brief. This is the outline of the…

- issue to be reported on,
- intended audience,
- expected length of the document,
- any other particulars to be included.
- purpose of the report,
- scope of included information,
- deadline for completion, and

If you're not given a report brief, you should develop one of your own and when the brief has been discussed and agreed to, have it signed off by the person requesting the report. Getting the brief agreed to and signed off is essential. It makes sure that everyone understands exactly what has been asked for.

The Issue

Reports centre round an issue. This is the reason for the report or the problem being investigated. Usually the issue is identified in the brief however, this isn't always the case.

For example, imagine the leader of a work team asks you to investigate the poor quality of report-writing her team produces. The team leader believes her team members make mistakes, don't think clearly and as a result, produce unusable reports. She says they need some training.

As you investigate you find that the problem isn't the team members -it's the team leader. She gives no feedback other than to point out their mistakes. She also makes unrealistic demands of them without providing appropriate support. So, the ***real*** issue is not the writing ability of the team members but the poor supervisory skills of the team leader.

This example shows that sometimes the person requesting a report doesn't really know what the issue is. Other times this is deliberate and indicates a ***hidden agenda*** on the part of the person requesting the report. Either way, it shows how important it is to try and find out the real issue.

Figuring out the real issue is like a doctor diagnosing a patient. The patient will describe one or more symptoms and the doctor uses these to identify the condition. Your job as a report writer is similar. The person requesting the report may not know the real issue.

They may only be able to tell you the symptoms. In these cases your first task will be to work but the underlying cause/s.

Scoping a report

Scoping a report is about figuring out what goes into it, and what doesn't. Without clear guidelines about the detail required, a report could end up being thousands of pages long! The scope helps you focus on what's relevant and important.

Here's an activity on scoping a report. Read the following scenario and respond to the question.

Imagine you work in a small but busy office that makes extensive use of computers. Your boss has decided to link all the computers together into a local area network (LAN). She asks you to examine the proposal and come up with most efficient type of LAN after investigating the market. Now you have to decide what information you will look for. You have to define the scope of your research.

Which of the following topics will you research? Why did you make that decision?

☐ how a LAN works	Yes	No
☐ what different systems are available	Yes	No
☐ cost of different systems	Yes	No
☐ names of contractors to wire up the network	Yes	No
☐ writing a manual for the chosen system	Yes	No
☐ how the LAN will benefit the office	Yes	No
☐ training staff to use the system	Yes	No
☐ is a LAN the best solution to the problem	Yes	No

Which topics did you decide to research?

☒ How a LAN works - This shouldn't be included. The fact that your boss is considering the installation of a LAN indicates that she already knows what one is and how it works.

☑ What different systems are available - This is important to include. Your boss can't make a choice without knowing what's available.

☑ Cost of different systems - This is another important inclusion. Cost will be one of the criteria used to choose a system.

☒ Names of contractors to wire up the network = No; far too early for this information. Your boss hasn't yet chosen a system so it's too early to be looking for installers.

☒ Writing a manual for the chosen system - No; again it's far too early for this. Until a system has been chosen it's not time to prepare a manual.

☒ How the LAN will benefit the office - This shouldn't be included either. Your boss has already made up her mind to install a LAN so she doesn't need to be told the benefits.

☒ Training staff to use the system - This would only be useful to research if one system was easier to learn than another and therefore the training of staff might become one of the criteria your boss uses to choose a LAN.

☒ Is a LAN the best solution to the problem - If your boss has already made up her mind there's really no point investigating this.

Targeting the Reader/s of the Report

Reports are written as though there is only one reader - the person who needs the information or will make decisions based on it. If a report is written for a group, that group is considered to be a single reader.

Questions you should ask yourself include the following.

- Who requested the report? Is it the same person who'll read it or use it?
- How much does my reader already know about the issue? Writing to suit the level of the reader's knowledge is important when writing a report. If your report is pitched too high, the reader won't be able to understand it. If it's pitched too low they may consider it unprofessional and dismiss the information in it.
- Why do they need the report? Most reports are requested because the reader has a problem to solve or a decision to make. Knowing how the report is going to be used will help you focus on including relevant information.
- What does the reader expect from the report? Most people reading a report expect it to be
    - easy to understand,
    - include all the information they need,
    - accurate and relevant,
    - structured logically, and
    - correct in terms of grammar, punctuation and spelling.
- What preferences do they have with regard to the way the report is written and/or presented? Some readers like things written in a particular way. They may like lots of heading or none at all. They may prefer the use of bullet points rather than long paragraphs. Some readers like the use of tables, graphs and other diagrams to indicate what the information is saying. Very often, managers prefer to have the detail left until the end of the report so they can refer to it if and when they want to. If you know the preferences of the reader you can tailor your report to suit them.

Researching the Information

In the previous section it was highlighted how important it is to focus on the real issue or problem. It's only when you know what the problem is that you can decide what sort of information you need and where you might find it.

One thing you need to do right from the beginning is to keep all your notes. You should keep notes on all your thoughts and conversations with people in a file where you can find them if your need them again for any reason. In some instances, these notes may be required for legal reasons. In other cases they might be essential to help you find where an error in the report was made or to relocate a particular source of information.

Your first source of information is yourself

It's very likely that you've been asked to prepare a report because you already know something about the issue. You have lots of information stored in your memory. The challenge is to recall those things that are relevant to the issue you're reporting on.

The first step is to make a note of everything you are conscious of knowing about the subject. This can be just a list of words or short phrases that represent larger thoughts or ideas. You may have included places or people you may be able to provide information.

Types of information

When you've noted everything you know about the issue it's time to look for other sources. Where you look will depend on the type of information you want. **Quantitative** information is objective, factual and often able to be measured or counted. For example, names, addresses, dates, numbers, statistics, engineering specifications.

Qualitative information is subjective, based on opinion or perspective. For example, feelings, values, preferences, interpretations.

Generally, information is categorised as either "quantitative" or "qualitative". Although there's a strong temptation to assume quantitative information is more credible, qualitative information can often tell you more.

For example, although the number of people who purchase a particular product (quantitative information) can tell you how popular it is, an opinion survey (qualitative information) can tell you WHY it's more popular.

Before you begin looking for information it's a good idea to decide what type of things you need to know. Make a note of the balance between qualitative and quantitative information you're going to look for. Depending on the type of report, you may need to shift the balance between the two types.

For example, if you're preparing a comparison of several brands of generators you'd include more quantitative information than qualitative. The reader of the report is more likely to be interested in the facts than in opinions about the machines.

Another way of thinking about information is whether or not it is "primary" or "secondary". **Primary** information is information that you've collected directly from the original source. For example, experiments you've conducted yourself, data you've collected.

Secondary information has come from someone else's work. For example, reports written by others.

Primary information is often considered more reliable because you've investigated it yourself. However, there is always the potential for your own biases and preferences to have influenced your investigations.

When relying on secondary information it's important to check that it's complete, can be verified by other sources and has evidence to support any claims or statements made.

Sources of information

Finding the information you want or need can be time consuming. Those who are familiar with the internet will prefer that as their source. Others will be more familiar with libraries, databases, registries or information sources inside their own organisation.

Your sources may include:

- Bibliographies
- Dictionaries
- Government documents
- Local government records
- National and international standards
- Opinion surveys
- People
- Company records
- Directories
- Journals
- Maps
- Newspapers
- Parliamentary papers
- Observations/experiments

Credibility of information

When you're gathering information for your report you need to consider the credibility of the source. There's an old saying – don't trust everything you read. This is a reminder that just because something is in print doesn't mean it's true or accurate.

The internet is becoming more and more relied on as a source of information. However, there's no way to know who published that information or whether it's already been plagiarised from another source. There's nothing wrong with using information that others have gathered as long as you acknowledge them as the original source.

Writing the report

Writing in plain English

Being able to communicate effectively in writing is an essential skill in the workplace. Unlike verbal communication, written words are long lasting and often read by other than the originally intended audience.

To ensure your document is read and understood you need to ensure that you

- write clearly,
- engage the reader, and
- use a visually engaging style.

The Plain English Campaign [http://www.plainenglish.co.uk/plainenglishguide.html] explain plain English by saying what it isn't.

- It's not 'cat sat on the mat or 'Peter and Jane' writing. Almost anything - from leaflets and letters to legal documents - can be written in plain English without being patronising or over-simplifying.
- It doesn't mean reducing the length or changing the meaning of your message. Most of the UK's biggest insurance companies produce policies that explain everything fully in plain English.
- It's not about banning new words, killing off long words or promoting completely perfect grammar. Nor is it about letting grammar slip.
- It is not an amateur's method of communication. Most forward-looking senior managers always write in plain English.
- And finally, it is not as easy as we would like to think.

Unfortunately, there's so much badly written English around that we tend to get used to it; we ignore it rather than complain or do something about it. However, as the Plain English Campaign point out, plain English has several advantages…

- "it is faster to write;
- it is faster to read; and
- you get your message across more often, more easily and in a friendlier way. "

Tips for writing in Plain English

Using 'first person'

Engaging your readers is about speaking to them directly and clearly. Writing in "first person" makes the reader feel as though you are talking directly to them. "You" gives the impression the document is intended for the reader in a way that "he", "she" or "they" cannot. Using "we" to refer to your own organisation makes the reader feel they are engaging in a conversation. Your organisation will have its own protocols about whether or

not to write in 1st or 3rd person and those should be followed. However, if none exist, remember that 1st person is a plainer form of English.

Organising the content

It's important to organise the information in the document in a structure that reflects the readers' interests. People read documents to find information or get answers. They want to know how to do something or what happens if they take a particular course of action. Hence, it makes sense to organise your document around the questions your audience is likely to be asking.

One way to do this is to use a question-and-answer format. This can be done by writing the subheadings as questions. By doing this the reader can scan the document and find the information they want. They may also see a question they hadn't thought of but need to know the answer to.

It helps to think through the questions the readers might ask and organise your document in that order. After each heading, answer the question immediately.

Use short sentences

Long complicated sentences are difficult to read. Take care to express only one idea in each sentence. Shorter sentences force you to break up the information into smaller, easier-to-process chunks. Compare the two versions of the same passage on the next page.

Long sentence structure…	Short sentence structure…
For good reasons, the Secretary may grant extensions of time in 30-day increments for filing of the lease and all required bonds, provided that additional extension requests are submitted and approved before the expiration of the original 30 days or the previously granted extension.	The Secretary may extend the time you have to file the lease and required bonds. Each extension will be for a 30–day period. To get an extension you must write to us giving the reasons that you need more time. We must receive your extension request in time to approve it before your current deadline or extension expires.

Use the present tense

Present tense avoids the confusion of compound verbs. Past tense or future tense should only be used when actually writing about something that has already happened or has not yet happened.

Subjects and objects

Keeping subjects and objects close to their verbs makes is easier for the reader to work out who is doing what to whom. A SUBJECT is a NOUN that is the FOCUS of the sentence.

The SUBJECT is the one who is the DO-ER of the VERB. The OBJECT is the thing/person the verb/action is done to. The **DOG** (subject) **ATE** (verb) the **BONE** (object). http://www.writershelper.org/objects.htm

If the subject and object are too far away from the verb, it becomes unclear what is going on. Consider the following sentence.

The equipment that had been installed at the Jacobsenn site about ten years ago when we used quite different connection and protection techniques is now causing damage to the entire network.

In this sentence the subject is the "equipment" and the object is the "entire network". However, they've been separated by explanatory information. Here is a way the sentence could be written more clearly.

The equipment at the Jacobsenn site is causing damage to the entire network. It was installed about ten years ago when we used quite different connection and protection techniques.

Modifiers

A 'modifier' explains or slightly alters a word. Some modifiers, especially simple modifiers (*only, just, nearly, barely)* have a bad habit of slipping into the wrong place in a sentence.

Modifiers need to be placed carefully in sentences. If they're in the wrong place the whole meaning of your sentence can alter. The following examples show how this can happen.

INCORRECT: He barely kicked that ball twenty yards.
[How can you 'barely' kick a ball?]
CORRECT: He kicked that ball barely twenty yards.

INCORRECT: The table was bought by a lady with sculpted legs.
[This is a very unusual way to describe someone's legs!]
CORRECT: The table with sculpted legs was bought by a lady.

Sentence sequence

In English we write our sentences so that the most important information is at the beginning. Here are some examples.

I went to the office yesterday.

[By placing "I" first, the author is emphasising who went to the office.]

Yesterday, I went to the office.

[By placing "yesterday" first, the author is emphasising when the trip to the office was made.]

Chris was with Leone when the invoice arrived.
This emphasises who was with whom at the time.
When the invoice arrived, Chris was with Leone.
This emphasises when the two people were together.

Each of the following sentences could be improved by using one of the hints outlined above. Rewrite each sentence to make it more easily understood.

1. If you take less than your entitled share of product for any month but you pay royalties on the full volume of your entitled share in accordance with the provisions of the contract, you will owe no additional royalty when you later take more than your entitled share to balance your account.
2. Individuals and organisations wishing to apply must file applications with the appropriate documentation to our office in a timely manner.
3. These sections describe types of equipment that would meet the criteria under clause 16:1B if they are installed according to the Australian regulations rather than the International guidelines.
4. Applicants who were in our employment at the time their injuries were sustained should have filed an incident report at that time. Failure to do so could have an effect on the degree to which the applicant can be covered by Workers Compensation.

It's very unlikely that you'll be the only person in your organisation responsible for preparing documents. Usually there are lots of people creating memos, work instructions, contracts, emails and other written communications. Some of these will be very well written and immediately clear to the reader. Others will be a mystery to everyone other than the author.

All organisations have an obligation to make sure information is available to the people who need it. This means everyone has responsibility for writing in plain English whenever possible.

You can encourage your colleagues to use plain English in many ways. The easiest method is by example - using plain English yourself. You can ask your colleagues for feedback on how clearly written your documents are to help bring the issue of plain English to their attention.

The **Plain English at Work** site is very helpful.
http://www.dest.gov.au/literacynet/plain_en/default.htm

Developing a "Plain English Guide" for your organisation is a very useful strategy. However, there are some already available. This one in particular gives a terrific overview

and some very helpful advice for writing all types of documents.
http://www.plainenglish.co.uk/plainenglishguide.html

At this web site you can find several free "plain English' guides for CVs, forms, letters, medical information, proofreading, reports, websites and wills.
http://www.plainenglish.co.uk/guides.html

Examples of plain English (not)

Before

High-quality learning environments are a necessary precondition for facilitation and enhancement of the ongoing learning process.

After

Children need good schools if they are to learn properly.

Before

If there are any points on which you require explanation or further particulars we shall be glad to furnish such additional details as may be required by telephone.

After

If you have any questions, please ring.

Before

It is important that you shall read the notes, advice and information detailed opposite then complete the form overleaf (all sections) prior to its immediate return to the Council by way of the envelope provided.

After

Please read the notes opposite before you fill in the form. Then send it back to us as soon as possible in the envelope provided.

Before

Your enquiry about the use of the entrance area at the library for the purpose of displaying posters and leaflets about Welfare and Supplementary Benefit rights, gives rise to the question of the provenance and authoritativeness of the material to be displayed. Posters and leaflets issued by the Central Office of Information, the Department of Health and Social Security and other authoritative bodies are usually displayed in libraries, but items of a disputatious or polemic kind, whilst not necessarily excluded, are considered individually.

After

Thank you for your letter asking permission to put up posters in the entrance area of the library. Before we can give you an answer we will need to see a copy of the posters to make sure they won't offend anyone.

Finally, if you thought those examples were bad, here's the longest sentence I've been able to find. It's a perfect example of just how complicated things can be if attention isn't paid to some of the simple suggestions made in this module. If you want to know how to write plain English, just do the exact opposite to everything shown in this example!

In the event that the Purchaser defaults in the payment of any instalment of purchase price, taxes, insurance, interest, or the annual charge described elsewhere herein, or shall default in the performance of any other obligations set forth in this Contract, the Seller may: at his option: (a) declare immediately due and payable the entire unpaid balance of purchase price, with accrued interest, taxes, and annual charge, and demand full payment thereof, and enforce conveyance of the land by termination of the contract or according to the terms hereof, in which case the Purchaser shall also be liable to the Seller for reasonable attorney's fees for services rendered by any attorney on behalf of the Seller, or (b) sell said land and premises or any part thereof at public auction, in such manner, at such time and place, upon such terms and conditions, and upon such public notice as the Seller may deem best for the interest of all concerned, consisting of advertisement in a newspaper of general circulation in the county or city in which the security property is located at least once a week for three (3) successive weeks or for such period as applicable law may require and, in case of default of any purchaser, to re-sell with such postponement of sale or resale and upon such public notice thereof as the Seller may determine, and upon compliance by the Purchaser with the terms of sale, and upon judicial approval as may be required by law, convey said land and premises in fee simple to and at the cost of the Purchaser, who shall not be liable to see to the application of the purchase money; and from the proceeds of the sale: FIRST, to pay all proper costs and charges, including but not limited to court costs, advertising expenses, auctioneer's allowance, the expenses, if any required to correct any irregularity in the title, premium for Seller's bond, auditor's fee, attorney's fee, and all other expenses of sale occurred in and about the protection and execution of this contract, and all moneys advanced for taxes, assessments, insurance, and with interest thereon as provided herein, and all taxes due upon said land and premises at time of sale, and to retain as compensation a commission of five percent (5%) on the amount of said sale or sales; SECOND, to pay the whole amount then remaining unpaid of the principal of said contract, and interest thereon to date of payment, whether the same shall be due or not, it being understood and agreed that upon such sale before maturity of the contract the balance thereof shall be immediately due and payable; THIRD, to pay liens of record against the security property according to their priority of lien and to the extent that funds remaining in the hands of the Seller are available; and LAST, to pay the remainder of said proceeds, if any, to the vendor, his heirs, personals representatives, successors or assigns upon the delivery and surrender to the vendee of possession of the land and premises, less costs and excess of obtaining possession.

Introduction to workplace documents

As a technician, engineer or tradesperson you'll use a variety of documents in your work. Some will be written by other people and some you'll write yourself. Some documents are simple - for example: forms where information is written into the spaces provided. Other documents are much more complex - for example: technical feasibility reports.

Documents have both a ***purpose*** and an intended ***audience***. Purpose is the intended outcome or result. For example, the purpose of a quotation is to inform the client of costs involved in a project. Audience is the intended or potential person/s who'll read the document. For example, the audience of a quotation is the client requesting the project.

Some documents have more than one purpose and more than one audience.

For example…

The purposes of a technical specification might be to ensure correct manufacture or to assist with repairs and/or maintenance. The audiences would include manufacturing staff, technicians and tradespersons.

The purposes of a tender document are to describe the capability of an organisation to complete a specified project, to outline the required budget, costs, time frames and other relevant information. The audience is the client financing the project or the project manager acting on their behalf.

For each of these documents, write what you think the purpose and audience/s might be

Document	Audience/s	Purpose
Letter of demand		
Standard Operating Procedure		
Technical Specification		
Bill of Quantities		

The purpose and audience of a document will determine its ***tone***. Tone is the emotion or mood created by the combination of language used and style of writing. For example: friendly, formal, angry, demanding, semi-formal, conciliatory, and neutral.

What words would you use to describe the tone of each of these types of documents?

Document	Tone
Letter to a friend	
Job application	
Letter responding to a client complaint	
Instructions for a new employee	

Formal language

Many workplace documents should be written in **formal** language. This is a style of writing very different than the casual tone of a letter to a friend or an email to a colleague. Formal language uses no slang or colloquialisms, does not contract words and has an impersonal tone.

Here is an example of a letter written in **informal** or **casual** language.

G'day Charlie. It was great to get your letter the other day. I'm glad to hear you're getting along OK in the new branch office.

Were you able to hire all the staff you needed or will you have to have another go at it? Don't forget I can come down and help out if you want.

Anyway, take care mate. See you later.

Sam Kronchite

Here is the same letter written in **formal** language.

Dear Charles,

Thank you for you letter of Wednesday 25th June. We were pleased to hear of your progress establishing the new branch office.

Please notify us of your progress with staff recruitment. I remind you that we're able to provide in-person support if necessary.

We look forward to your reply.

Yours truly, Samual Kronchite.

You'll notice that the formal example does not include…

- slang or colloquialisms,
- contractions, or
- subjective opinions.

Colloquialisms are slang terms or words only used in casual, everyday language.
Contractions are two words merged together such as it's (it is), who're (who are), can't (can not) or we're (we are).

Rewrite the following letter in a formal tone.

This guy, Ken Trent, dropped into the office last week sometime to see if we can do a job for him. He knows what he wants but needs someone to draw up the plans and get them checked out so the equipment and installation will meet all the regulations and stuff.

I reckon we should do the job even though it's not going to be much money for us. We did some work for him before, a big job, and I reckon we might get another big one out of him if we treat him right.

Can you let me know what your blokes are doing in the next week or so and if they can fit this in? I have to get back to Ken by the end of this week with some figures so be quick. OK?

Using a word processor

Before you can begin to create workplace documents you need basic skills in using your computer. The most frequently used software used to create documents is called a "word processor". Typical programs are Microsoft Word and Lotus WordPro. Each of these may be installed separately on your computer or as part of an "Office Suite" - a collection of programs (word processor, spreadsheet, database, etc.).

If you're a competent user of the word processing program used in your workplace, you can skip this section and proceed to the subheading

"Simple Business Reports".

Your first challenge is to reproduce the following document using the word processing program used in your workplace. Print it and ask your workplace learning mentor/tutor to check it for accuracy.

2 July 1999

Mr Craig Lyons

Manager, Occupational Health & Safety

North Coast Workwise

99 Main Street

Port Macquarie NSW 2444

Dear Mr Lyons

I'd like to formally thank you for the attention to detail given in your work with our organization. The recommendations from your final report have been well received by our Board of Directors.

In particular, they were impressed by the data provided on our accident and incident reporting. Clearly it's an area of our OHS procedures that needs to be addressed.

If your recommendations are accepted and the Board allocates the necessary resources to act on them, would you be available to assist in the capacity of Project Coordinator? Obviously we would need to discuss the detail of such an arrangement before entering into a contract. This initial inquiry is to ascertain your availability and willingness to undertake the role.

Please consider this suggestion and make contact with me in the next week or so. I am available by phone or email.

Yours sincerely

Delina Smyth-Jones B.A. M.Eng.

Senior Resource and Quality Analyst

Coastal Constructions Pty Ltd

Your next challenge is to duplicate the following text. This task is specifically designed to highlight whether or not you are familiar with common editing functions in a word processor.

This is a bullet list

- line one
- line two
- line three

This is a number list

1. item one
2. item two
3. item three
4. item four

> This paragraph has been positioned two "tabs" from the margin. Note that all the lines in the paragraph have been moved - not just the first line.

Now we're going to explore the **bold function** and the underline function as well as the use of *italics* within text.

This text has been changed to a font called Arial.

This text has been centred to the middle of the page.

Can you change the size of the font to 16 point?

Finally, print this document and show it to your workplace learning mentor/tutor.

If you've successfully completed these two challenges, you're ready to go ahead with the rest of this module.

If you had any difficulty with one or both of these challenges, ask your workplace learning mentor/tutor to recommend how you might improve your work processing skills before you go any further with this module.

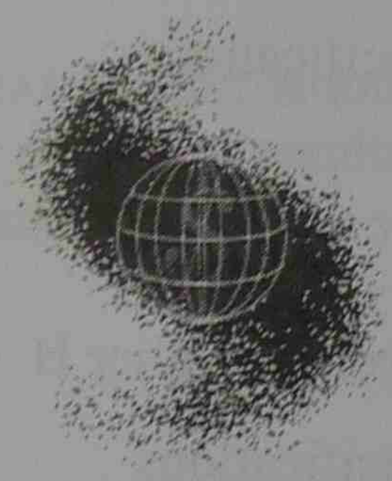
Simple Business Reports

Many organisations have their own "template" for preparing reports. However, the following format is typical and acceptable for most situations. A template is the outline used for a specific type of document. For example: letters, business cards, pamphlets, reports, invoices and memos. The template specifies the font; size of text used for headings and subheadings; indents for paragraphs, bullets and lists; page and margin settings; and other structural elements of the document.

Title Page

Your report should have a professional title page. Using the business logo or symbol adds credibility and promotes the organisation.

Sample Title Page



A Comparison Of

Programmable Logic Controllers

from Supplier A and Supplier B

for Kwan Ho

Director, Design Engineering

by

Frederick Bathstone

Electronics Engineer

August 2002

Executive Summary

A long report often includes a brief summary so that the reader can decide whether to read the whole document. The main point(s) of each section of the report should be included in this summary.

Table of Contents

The section headings of the report should be listed, using a consistent numbering system. Page numbers for each section should also be included.

(SAMPLE) Table of Contents

	Executive Summary	i
1.0	Introduction	1
2.0	The intended purpose for Programmable Logic Controllers (PLCs)	1
3.0	Characteristics of PLCs from Supplier A	2
	3.1 Selected technical specifications	2
	3.2 Special features	4
	3.3 Price schedule	5
4.0	Characteristics of PLCs from Supplier B	6
	4.1 Selected technical specifications	6
	4.2 Special features	7
	4.3 Price schedule	8
5.0	Comparison chart	9
6.0	Conclusion and recommendation	11
	Appendix A - Full technical specifications	12

Introduction

The introduction needs to include the purpose and scope of the report, background information, and an outline of the main points of the report.

Body of the Report

Headings are used to divide the report into a logically organised presentation of the results of your investigation and research. The numbering system and headings should correspond to your table of contents.

Conclusion

You need to draw conclusions based on the evidence you have presented in your report.

Recommendations

Sometimes you'll be asked to make recommendations for further action. These should be specific, concrete actions based on the conclusions you have reached. If a recommendation can't be supported by the information and conclusions in the report - it should not be included.

Bibliography

List the sources of information you have used to compile the report whether you have referred to them in the text or not.

Appendices

You may need to attach (or append) extra information to your report. For example: technical data, specification sheets or a table of results. These appendices should be clearly labeled (for example 'Appendix A') and included in your bibliography.

More information on report writing can be found at
http://www.wisc.edu/writing/Handbook/ScienceReport.html

Using graphics

Some workplace documents include information that can be displayed in the form of graphs and charts. For example, the following graph may be used to indicate the dollar value of contracts entered into by engineering teams in the three regions of the organisation.

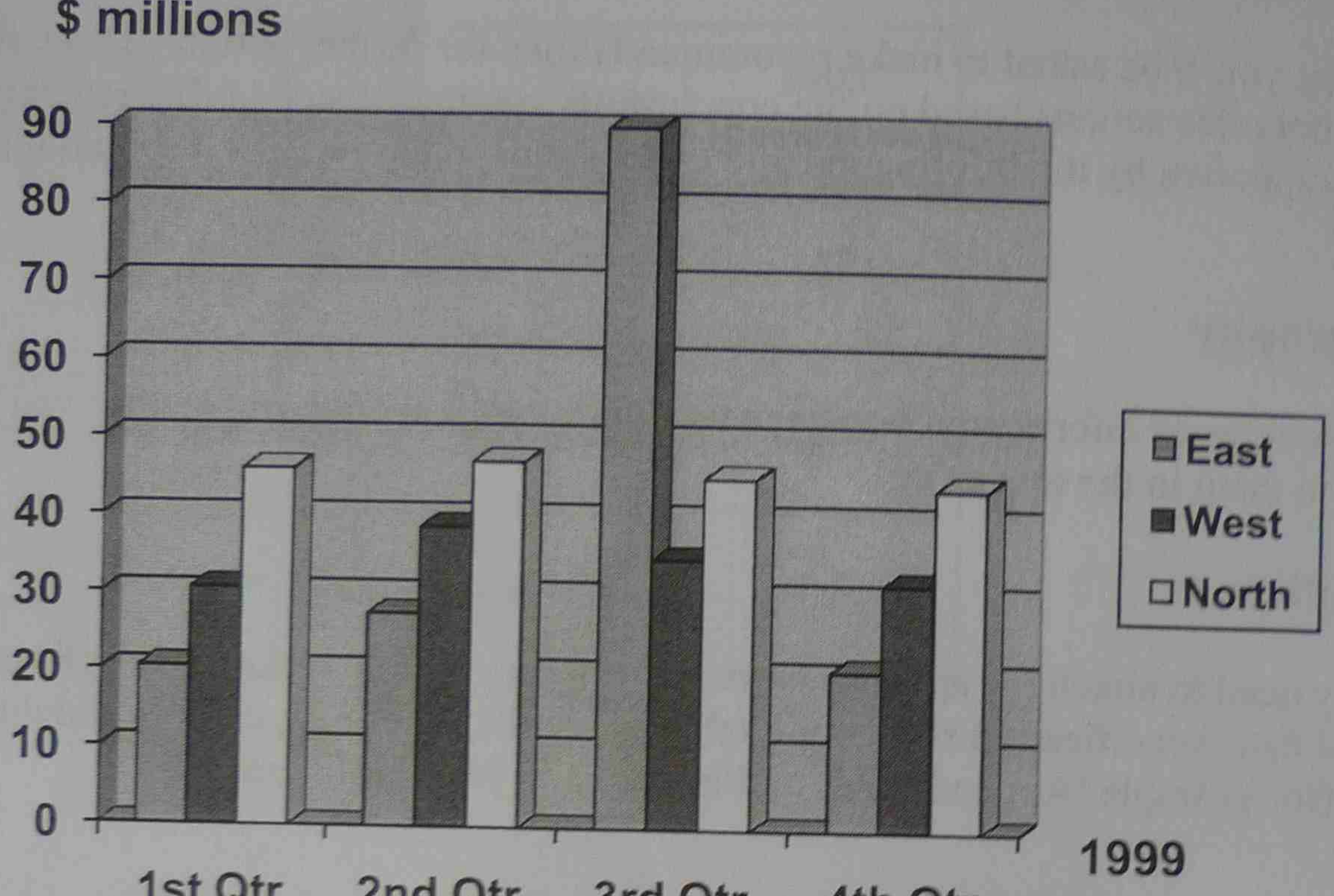


Figure 1 - Value of contracts entered into by engineering teams in 1999

In this case, the use of a graph makes it immediately obvious that…

- the Northern Engineering Team consistently enter into contracts totalling more than the other two teams.
- generally, the Eastern Engineering Team enter into contracts totalling less than either of the other two teams.
- in the 3rd quarter of 1999, the Eastern Engineering Team entered into contracts totalling a significant amount of money.

This example demonstrates that graphs or charts can display information so that meaning can be gained in a single glance. However, it's important that if graphs or charts are used, the raw information (the data) is available to the reader in an appendix or an easily accessible reference document.

Draw a rough sketch of the type of graph/chart you would use to display the following information in a report.

Number of days lost due to accident/injury	1999	2000	2001	2002	2003 (prediction)
Office staff	253	198	215	170	150
On-site sub-contracted staff	375	352	301	220	250
On-site technical staff	147	123	118	89	80

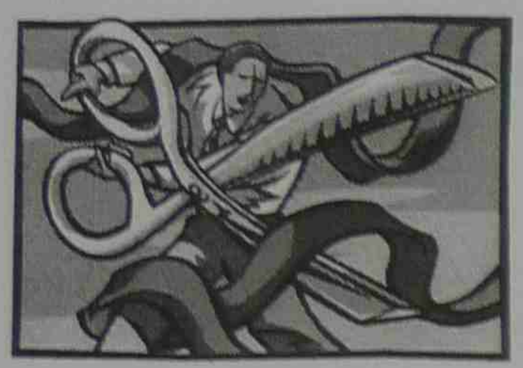
1999	Proportion of income
Bell Engineering Group	32%
Acorn Engineering Group	35%
Halicon Engineering Group	22%
Phonics Engineering Group	11%

Graphs and charts are not the only type of graphics used in workplace documents. Small, stylised pictures (icons) can be used to indicate emotions or concepts.

For example…

…this icon may indicate a bright idea.

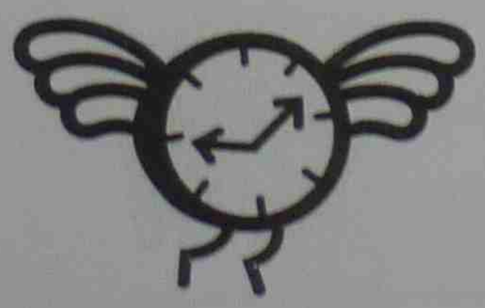
…this icon may  indicate time pressures.

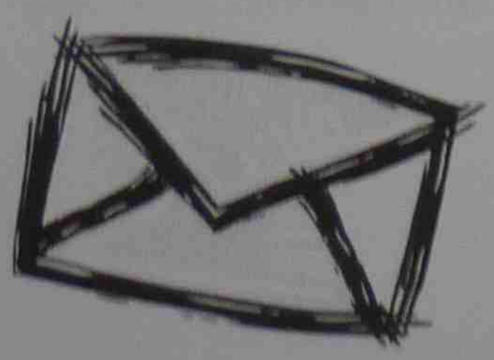


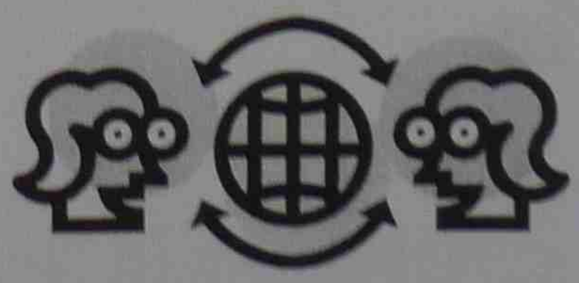
…and this one might indicate frustration with 'red tape' in an organisation.

Icons are not appropriate for formal documents such as reports, tender proposals, letters or academic documents. However, they can be useful in instructions, procedures, learning resources and even internal memos.

Next to each of the following icons, indicate what you think they might be used to represent.







Draw or describe an icon that could be used to indicate the following messages/concepts.

Make a note of this.
This is a group task.
This is important.
Pass this information on.

Writing consistently

Some documents are difficult to read, not because of the content, but because of the way they're written. To make a document easy to read you should take care with…

- the language used,
- the tone of the document,
- spelling, grammar and punctuation, and
- the structure of the document.

Language

One aspect of analysing your target audience is determining whether or not they're experts, technicians, executives or non-specialists. The amount of technical 'jargon' you use should vary to suit the audience. Obviously you can use more technical language if the primary audience is made up of experts. On the other hand, you need to limit the use of technical language if the audience is intended to be non-specialists.

Tone

Another important aspect of a document is the tone. Some documents need to be quite formal. Others are more effective if they're kept friendly and casual. This was discussed in a previous section of this workbook.

Spelling, Grammar and Punctuation

Spelling, grammar and punctuation are an integral part of writing whether you're a professional writer or just want to write a note to a colleague. Together they help get your point across effectively and focus attention on **what** you have to say instead of how you choose to say it. The following are absolute essentials to good writing that everyone should master.

Agreement – Agreement in a sentence refers to all of the parts of the sentence working well with each other. It's important that the subject of a sentence and the verb associated with it are both singular or both plural.

> For example, you wouldn't say "John have two pieces of toast and I has three."
> You would instead say, "John has two pieces of toast and I have three."

For more information on 'agreement' go to

http://www.shared-visions.com/explore/english/verbagr.html

Tense – Tense refers to time. What time is it in your sentence? Whatever time it is, it should remain consistent throughout your whole piece of writing. If it was last week you are talking about, stay there. There are three tenses in writing, past tense, present tense and future tense.

> Here is an example of writing with mixed tenses: "Julie wondered how she is going to finish in time, but Joe will help her."

This sentence contains all three tenses. "wondered" = past tense, "is" = present tense and "will" = future tense. Pick a tense and stick to it! The sentence could read "Julie wonders how she will finish in time, but Joe will help."

Find more information on "tense" at
http://www.shared-visions.com/explore/english/tense.html

Point of View – The point of view refers to whoever is telling the story or "speaking." When you write a letter you are writing in "first person" which includes I, me, my, we and our. Second person writing occurs when we talk about you and yours and third person includes he, she, they and theirs. In third person writing, the author does not put him/herself into the story.

Using the first and second person results in an informal, conversational tone in the document. The reader feels like you're talking directly to them.

The third person results in a formal, objective and impersonal effect. This can add credibility to the document. Academic and business documents are usually written in third person.

Find more on "person" at
http://www.shared-visions.com/explore/english/person.html

Spelling – Correct spelling is essential. Without it your document, and you, have very little credibility in the mind of the reader. Spell checkers on computer programs are poor substitutes for knowing how to spell and can leave behind more errors than you realize. There are many different forms of words and your spell checker doesn't know which form you wanted to use.

For example, "When Mark washed they're care, he forgot too putt on the wacks."

A spell check program wouldn't realise this sentence should read, "When Mark washed their car, he forgot to put on the wax."

For some great spelling tips go to
http://writing.englishclub.com/spelling.htm

Run-On Sentences – A run-on sentence is one that is just too long! Not only is it too long, it is incorrect. Usually, a run-on sentence can be made into two or more sentences by using better punctuation and style.

An example of a run-on sentence might be: "We walked over to the commissary to get something to eat but it was closed so we didn't know what to do so we kept walking until we saw a restaurant and decided to go in and get something to eat but Andrew didn't want to eat there so we kept going for another mile."

This sentence could have gone on for another mile too! Break the sentence into smaller, more coherent parts.

Find out more at
http://www.wisc.edu/writing/Handbook/CommonErrors_Sprawl.html

Punctuation – It is very important to know your punctuation; even if you never plan on using a semicolon for the rest of your life. The most important thing to learn is where to put your commas, a common mistake among writers. Avoid using commas after conjunctions like "but" and "and." Commas are used to separate parts of sentences that stand alone, such as those that are parenthetical.

> For example "There were too many flowers, not that I minded, but they took up most of the room."

Usage – If you are going to use a word, you really ought to know how to use it. Some writers think big words look impressive but actually the reverse is true if the word is used incorrectly. Words don't have to be big to be misused, consider its vs. it's.

Capitalisation – Words at the beginning of sentences aren't the only ones worthy of capital letters. Always capitalise proper names such as people and places. Titles of all kinds deserve capital letters and so do acronyms.

Information on the correct use of capitals can be found at
http://writing.englishclub.com/caps.htm or
http://www.shared-visions.com/explore/english/capital.html

Sentence Fragments – A sentence fragment is an incomplete sentence that does not include both noun and verb. An example of a sentence fragment might be, "Really dumb." Make sure your sentences reflect a complete thought unless you are writing dialogue.

More on sentence fragments at
http://www.wisc.edu/writing/Handbook/CommonErrors_Frag.html

Wasted Words – A big no-no. Sometimes we throw in words just to round out our sentences, or we over-describe something, like, "The really complex spreadsheet program was on the E drive." If the purpose of the sentence is to say where the spreadsheet was, you don't need to point out that it was really complex. Economise your words and you'll reduce the chance of grammatical errors.

Inflated words are also a problem in report writing. Find out about them at
http://www.wisc.edu/writing/Handbook/CCS_inflated.html

For more information, try these sites

http://www.learnenglish.de/grammarpage.htm

http://www.wisc.edu/writing/Handbook/CommonErrors.html

http://www.wisc.edu/writing/Handbook/GramPunct.html

http://www.wisc.edu/writing/Handbook/Style.html

The following passage is full of errors in punctuation, spelling and grammar. Your challenge is to rewrite it without any errors, using your organisation's preferred word processing program.

Print your final document and discuss it with your workplace learning mentor/tutor.

Message from the Executive Director

I am pleased to present the organisations strategic plan for the next five years. The plan had been developed after wide consultating with staff, the Board of Director's, major client's and local community representatives.

We're getting feedback that our previous plan needed more input from people other than the Board of Directors, that is why we used a consultative process.

I am confidant that because of the consultative process followed, you will fined the strategic plan a true reflection of we're we're wanting to head as a organisation in the future. Our vision of being the leader in engineering services in the region continued to drives our planning. Our new goals reflect our determination to support the development of the local region, respond to our customer needs and keep a positive working environment for all our staff.

Vision

To be the leader in engineering in the mid north coast region.

Roll

To advance the sustainability of the region economically and environmentally by providing engineering services which meet our clients needs and are environmentally sensitive.

Values

We…

…**value people** and embrace diversity.

…**respond** with versatility **to our customers** needs

…are **ethical**, **innovative** and committed to quality.

…**cooperate** with other organisations for **mutual benefit**.

…practice **continuous improvement** in every aspect of our work.

…support the principals of **sustainable development** and **environmental sensitivity**.

Referencing

If you quote or paraphrase information from another source you need to "reference" it. Referencing means showing where the information that you quoted or paraphrased originally came from. It's OK to use someone else's words, ideas or information in your work but you must show that they are not your own by indicating their source.

One of the reasons we reference is to distinguish between our ideas and someone else's. If you don't acknowledge the work of others you may be accused of *plagiarism*.

Plagiarism occurs when you present someone else's ideas, thoughts, information, etc, as if they're your own. It's plagiarism if you use someone else's ideas and deliberately fail to acknowledge the source. Its also plagiarism if you attempt to paraphrase or rewrite ideas in your own words but your ideas resemble the words or the original source too closely, or you don't use indented quotations or "quotation marks" to indicate which words are yours and which ones are the words of others.

Quoting the work of other authors/writers/sources

A quotation is where you put down the exact words of the original source. A quotation may be as short as one word but, if that word is significant, it must be put in quotation marks and referenced.

Quotations should include the exact words of your source inside quotation marks. For example:

> "Everything we do is an experience of a kind." (Kenny 1996: 45).

If you look on page 45 in the book written by Kenny, published in 1996 you will find the words ***Everything we do is an experience of a kind*** with no alterations or omissions.

Sometimes a quotation is too long. When this is the case you can leave out a word or words by using an ellipsis. Here's an example.

ORIGINAL

> "In many academic circles in America, literary translation is still considered a secondary activity, mechanical rather than creative, neither worthy of serious critical attention nor of general interest to the public" (Gentzler 1993: 34).

SHORTENED

> "In many academic circles in America, literary translation is still considered a secondary activity . . . neither worthy of serious critical attention nor of general interest to the public" (Gentzler 1993: 34).

Notice that where words have been removed from the quotation, three dots are inserted. These three dots are known as an ellipsis. No matter how many words you take out, you only use three dots - except if the quoted sentence comes to an end, in which case you use three plus a full stop, i.e. four.

Sometimes it helps to make small alterations/clarifications to a quotation. In the following example the writer clarifies who "The engineers" are, and changes capital T to small to fit the writer's sentence structure.

ORIGINAL

Smithson (1996) argues that, "The engineers never knew what good technical specifications were." (p.245).

ALTERED

Smithson argues that "[t]he engineers [on the Clarke Brothers worksite] never knew what good technical specifications were." (1996: 245).

Where possible it's better to include quotations in your own sentence. In the following example the writer includes a quotation from Shilton in his/her own sentence. Notice how the quotation is altered to fit grammatically into the sentence. The writer has carefully chosen the point at which to start.

ORIGINAL

"Everything we do adds to our experience." (Shilton 1994: 32).

INCLUDED IN WRITER'S SENTENCE

It can be argued that "[e]verything we do adds to our experience." (Shilton 1994: 32).

If a quotation is included in your paragraph, it must be indicated with quotation marks (" . . ."). However, if you are using a long quotation it should be set in an indented block. Here's an example.

> In Scandinavia, Canada, Tasmania, and the Snowy Mountains in mainland Australia, hydroelectric power is generated by running large volumes of water through low speed turbines. The water is collected at as high a level as possible in reservoirs. A power station is then built at a lower level. (Jenneson 1996:27)

Generally, quotations should be kept short and kept to a minimum (i.e. only use quotations when the words themselves are important). It advisable to avoid using a large number of quotations as they mean you are letting your sources present ideas instead of you presenting your own.

Read through the text below and highlight the sections that are quotations. Ask your workplace mentor/tutor to check your work.

Many years in the future it's likely that man will communicate via very different means that those used today. Robinson (2000:24) suggests that "at an early age children will receive implants in their ears to allow direct radio communication to their ear from a remote source." Although this sounds like science fiction, it may be closer to reality than we first think.

Both Robinson and Gregson predict a continued decrease in the size of communication devices. Gregson (1998:317) in particular, points out that the only reason for the comparatively large size of mobile phones is that "the device must be large enough to accommodate a keypad that adult fingers can operate." Although we have the technology to produce a telephone as small as a matchbox, it would be impractical from the consumer's point of view.

"Electronic devices of various types… would be more secure if placed under the skin of the owner rather than collected in a wallet ready for a pickpocket to acquire". (Robinson 2000:42) However, what Robinson fails to point out is that there are issues of maintaining such devices. At no point does he describe the possible complications associated with installation, removal and re-installation of subdermal communication devices. Gregson (1998) on the other hand admits that there will be resistance on the part of some consumers to the idea of electronic apparatus residing inside the body, even if they are only placed under the upper layers of the skin.

As a matter of interest, studies by Eriksann (2002) involving subdermal implantation of electronic devices in various mammals, indicates that "there is a very low rate of complication with the most severe reaction recorded being a case of Staphylococcus aureus in one of the equine subjects." In that particular case Eriksann was not able to eliminate other possible sources of the infection hence his ultimate finding was that "subdermal implants are unlikely to present any significant risk to the host".

Paraphrasing

Paraphrasing means putting another author's ideas or information into your own words. In the following example the writer has no need to use Tenn's exact words as it is his information (not his words) that are important. Paraphrasing also allows the writer to compare the author (Tenns) with another (Keffer).

ORIGINAL

"This has led to the belief that the Macintyre method of measuring power factor is more accurate than calculation." (Tenns, 1991:54).

PARAPHRASED

In contradiction to Keffer's claim, Tenns argues that calculation is not as accurate as the Macintyre method of measuring power factor (1991).

Remember when paraphrasing you are using someone else's ideas or information so you must reference the source. However, unlike quotations we don't enclose paraphrasing with quotation marks or indentation.

One of the most commonly used referencing systems is APA (Australian Psychological Association). Details on this system can be found using these URLs.

http://campus.cpit.ac.nz/learning/documents/apastyle2003.pdf

http://auckland.massey.ac.nz/dept/ss/documents/notes_on_referencing.pdf

Preparing the document

Documents need to be planned before they can be written. This is a relatively simple process. The steps involved are as follows.

Step	Task	Explanation
1	Decide the purpose of the document	Refer to Section 1 of this appendix. The purpose is the intended outcome or result. For example, the purpose of a quotation is to inform the client of costs involved in a project.
2	Determine who the audience/s is/are	Refer to Section 1 of this appendix. Audience is the intended or potential person/s who'll read the document. For example, the audience of a quotation is the client requesting the project.
3	Analyse the audience	Analysing the audience is necessary to adapt your writing to their needs, interests and background knowledge.
4	Plan the document	This is where you create the 'skeleton' of the document. Creating a list of bullet points or subheadings is useful. Under each you can make notes of the type of information required or references to check.
5	Draft the document	This first draft can be done by hand or word processor depending on your level of computing skill. The draft is a 'working document'. This means it can be changed and rearranged.
6	Check for accuracy of information	This is one of the most important steps. Its essential that all information in a workplace document is checked and the source noted in the bibliography.

Step	Task	Explanation
7	Format the draft	Formatting is where you use a word processing package to put in headings, subheadings, lists, page numbering and other 'structural' items.
8	Check for consistency in language, tone, grammar, and document structure	Refer to Section 6 of this resource booklet. It's very difficult to read reports that aren't consistent. Not only that, spelling, grammar and punctuation errors make the document appear unprofessional.
9	Edit and reformat	This is the time to make your corrections, improve the formatting and get it looking professional.
10	If possible, have another person check the document for accuracy and consistency	Time and time again it's been shown that having a second or third person look at a document will pick up errors. The author has seen it so often they don't see the mistakes.
11	Final edit and print	After fixing up the last mistakes, print the document out in its final form.

Assessment task - Conduct an investigation and report on findings

To demonstrate competence you are required to research, analyse and interpret engineering data then use your findings as the basis of an engineering report. The report and the processes you use to prepare it should address the following criteria.

- ☐ Use the internet to obtain relevant data
- ☐ Prepare a report brief that investigates an engineering problem, issue or phenomena
- ☐ Design, conduct and report on an experiment to investigate a relationship between two variables
    - o collect and analyse experimental results expressed in two variables, using technology as required
    - o investigate a practical problem using correlation and regression
- ☐ Describe the statistical method and design chosen to meet the aim of the investigation
- ☐ Carry out the statistical analysis and report results
- ☐ Comment critically on choices of model and analyses resulting from them
- ☐ Evaluate and interpret the results of the investigation
- ☐ Discuss the investigation with reference to real world applications
- ☐ Describe the chronology of the investigation and reflect on the statistical process in journal
- ☐ Present a well formatted report with a clearly stated aim, method and conclusions

Appendix 1 - Recommended websites

The following websites cover material related to this competency. The urls were valid as at June 2009.

- Engineering Technical reports – Colorado State University
- http://writing.colostate.edu/guides/documents/ce-trpt/ - this is an excellent resource complete with a sample report including the comments from the assessor who graded it.
- Writing technical reports – Monash University
- http://www.monash.edu.au/lls/llonline/writing/engineering/technical-report/index.xml
- Guide to technical report writing – University of Sussex
- http://www.sussex.ac.uk/engineering/1-3-11-2.html
- Report Writing Style Guide – University of South Australia
- http://www.unisa.edu.au/ltu/students/study/specific/report-engineering.pdf
- Sample report – Midwest Electrical
- http://www.midwestelectrical.com/_resources/pdfs/mec_sample_engineering_report.pdf

- Research Methods Knowledge Base http://www.socialresearchmethods.net/kb/index.php
- STEPS statistical glossary http://www.stats.gla.ac.uk/steps/glossary/alphabet.html
- OECD glossary of statistical terms http://stats.oecd.org/glossary/
- Introduction to descriptive statistics http://www.mste.uiuc.edu/hill/dstat/dstat.html
- (Electronic Version): StatSoft, Inc. (2007). Electronic Statistics Textbook. Tulsa, OK: StatSoft. WEB: http://www.statsoft.com/textbook/stathome.html
- SISA Simple Interactive Statistical Analysis http://www.quantitativeskills.com/sisa/
- Survey and questionnaire design – free online tutorial http://www.statpac.com/surveys/index.htm#TOC

Appendix 2- question types

Just about any item in a questionnaire can be classified as open ended or closed ended.

Here are some examples of four common types of closed questions.

Name	Description	Example
Dichotomous	Answer can be one of two choices	Has the back-up power supply been fully tested? ☐Y ☐N
Multiple choice	Answer can be one of three or more choices	What is your employment status? ☐ casual ☐ part time ☐ contract ☐ permanent
Closed numeric	Respondent can only choose from the number categories provided	How many times have you serviced the Pittown generator? ☐ never ☐ 5 or less ☐ less than 20 ☐ 20 or more
Likert scale	Asks the respondent to show amount of agreement or disagreement	The product technical information was readily accessible Strongly Disagree / Disagree / Neither agree or disagree / Agree / Strongly Agree
Semantic differential	The respondent selects a point that represents the direction and intensity of their response on a scale between two opposite words.	The product technical information was... Useless ___ ___ ___ ___ ___ Useful Boring ___ ___ ___ ___ ___ Interesting Inaccessible ___ ___ ___ ___ ___ Accessible Jargonistic ___ ___ ___ ___ ___ Plain English
Rank order	Respondent is asked to rank given options	Rank the following suppliers from most preferred (1) to least preferred (5) Motorolla ____ Isnox ____ Siemens ____ Challenge ____ Castils ____

Appendix 3 – questionnaire checklist

- ☐ Is there a cover letter? Include clear and concise instructions on how to complete the questionnaire, how to return it for analysis, and a due date.
- ☐ Do the questions directly relate to the goal of the study? There is no point asking a question if you are not going to use the responses.
- ☐ Is the questionnaire brief, simple and in plain English?
- ☐ Have you structured the questions in a way that you can analyse the responses?
- ☐ Are the first couple of questions interesting and/or non-threatening? Many people look at the first few questions then decide whether or not to complete the questionnaire.
- ☐ Are the most important questions in the first half of the questionnaire? If so, it is less of a problem if it comes back only partially completed.
- ☐ Is there room for comments? This allows the respondent to add information that your closed questions may not capture.
- ☐ Are there changes in question format every now and then? This helps avoid the respondent getting into a habit of answering in a particular way without reading the question properly.
- ☐ Is the questionnaire convenient? If you have posted it out, include a self-addressed postage-paid envelope.