

Student workbook

MSL924003

Process and Interpret Data

Semester 2 2019

**TAFE NSW would like to pay our respect and acknowledge Aboriginal and Torres Strait Islander Peoples as the Traditional Custodians of the Land, Rivers and Sea. We acknowledge and pay our respect to the Elders, both past and present of all Nations.**

Version: v 1.0

Date created: 07/05/2018

Date modified: 27/02/2019

For queries please contact:

Innovative Manufacturing, Robotics and Science SkillsPoint

© TAFE NSW 2019  
RTO Provider Number 90003 | CRICOS Provider Code: 00591E

This resource can be found in the TAFE NSW Learning Bank.

The content in this document is copyright © TAFE NSW 2019, and should not be reproduced without the permission of the TAFE NSW. Information contained in this document is correct at time of printing: 16 May 2019. For current information please refer to our website or your teacher as appropriate.

Contents

[Introduction – Working in a laboratory 7](#_Toc8891346)

[Working in a lab 7](#_Toc8891347)

[What is a laboratory? 8](#_Toc8891348)

[What is your role in the lab? 9](#_Toc8891349)

[Following workplace procedures 9](#_Toc8891350)

[Process and interpret data workflow 11](#_Toc8891351)

[Analytical problems 11](#_Toc8891352)

[Analytical procedures 11](#_Toc8891353)

[Topic 1 – Metrology 13](#_Toc8891354)

[What is metrology? 14](#_Toc8891355)

[The international (metric, SI) system 14](#_Toc8891356)

[How the metric system works in Australia 14](#_Toc8891357)

[Regulatory metrology in Australia 15](#_Toc8891358)

[Base units 17](#_Toc8891359)

[Derived units 18](#_Toc8891360)

[Uncertainty in measurement 19](#_Toc8891361)

[Error in analytical measurement 20](#_Toc8891362)

[Accuracy and precision 21](#_Toc8891363)

[Relative and absolute descriptions of error 21](#_Toc8891364)

[Calibration 23](#_Toc8891365)

[Traceability 23](#_Toc8891366)

[Topic 2 - Data 27](#_Toc8891367)

[Data sources 27](#_Toc8891368)

[Storing data 29](#_Toc8891369)

[Retrieving data 30](#_Toc8891370)

[Checking data 30](#_Toc8891371)

[Rectifying errors 31](#_Toc8891372)

[Topic 3 - Calculating scientific quantities 36](#_Toc8891373)

[Calculating scientific quantities 37](#_Toc8891374)

[Calculation skills 37](#_Toc8891375)

[Estimations 37](#_Toc8891376)

[Significant figures 38](#_Toc8891377)

[Scientific notation 39](#_Toc8891378)

[Unit conversions 41](#_Toc8891379)

[Using equations and formulae 52](#_Toc8891380)

[Formula(e) or equation? 52](#_Toc8891381)

[Transposition 52](#_Toc8891382)

[Substitution 54](#_Toc8891383)

[Solving 55](#_Toc8891384)

[Expressing results 58](#_Toc8891385)

[Calculations with measurements 59](#_Toc8891386)

[Percentage calculations 59](#_Toc8891387)

[Geometric calculations 62](#_Toc8891388)

[Example calculations - pathology 64](#_Toc8891389)

[Example calculations - chemistry 65](#_Toc8891390)

[Example calculations - food 67](#_Toc8891391)

[Topic 4 – Descriptive statistics 74](#_Toc8891392)

[What is statistics? 75](#_Toc8891393)

[Samples and populations 76](#_Toc8891394)

[The descriptive statistical process 77](#_Toc8891395)

[Step 1 - Getting the data 77](#_Toc8891396)

[Step 2 – Tally and graph the data 78](#_Toc8891397)

[What is ‘a distribution of data’? 79](#_Toc8891398)

[The ‘Normal’ distribution 79](#_Toc8891399)

[Non-normal distributions 80](#_Toc8891400)

[Standard deviation 81](#_Toc8891401)

[What is ‘central tendency’ of data 83](#_Toc8891402)

[Topic 5 - Presenting data 88](#_Toc8891403)

[Tabulating data 88](#_Toc8891404)

[Visualising data 88](#_Toc8891405)

[What is a *plot* or a *graph*, or even a *chart*? 89](#_Toc8891406)

[General graphing principles 89](#_Toc8891407)

[Column and bar charts 92](#_Toc8891408)

[Line graphs 92](#_Toc8891409)

[XY scatter graphs 93](#_Toc8891410)

[Other types of charts 93](#_Toc8891411)

[The art of being ‘honest’! 95](#_Toc8891412)

[Semi-quantitative observations 96](#_Toc8891413)

[Reporting data 97](#_Toc8891414)

[Short forms 98](#_Toc8891415)

[LIMS 98](#_Toc8891416)

[Issuing Analytical Reports or Certificates of Analysis 99](#_Toc8891417)

[Record keeping and confidentiality 101](#_Toc8891418)

[Topic 6 - Interpreting data 105](#_Toc8891419)

[What is interpretation? 106](#_Toc8891420)

[What type of data was collected? 106](#_Toc8891421)

[Is the analysis qualitative or quantitative? 107](#_Toc8891422)

[Trends, patterns & noise 107](#_Toc8891423)

[Trends 107](#_Toc8891424)

[Patterns 108](#_Toc8891425)

[Noise (no trend or pattern) 108](#_Toc8891426)

[Proportionality 109](#_Toc8891427)

[How to read a graph 109](#_Toc8891428)

[Appendix A – Metric (SI) tables (reference only) 117](#_Toc8891429)

[Appendix B - Tables of Relationships (SI – Imperial) 118](#_Toc8891430)

[Appendix C - Answers to practice questions 122](#_Toc8891431)

[Appendix D – Answers to Self-Check Questions 131](#_Toc8891432)

[Bibliography 138](#_Toc8891433)

[References 138](#_Toc8891434)

[Image attribution labels 138](#_Toc8891435)

Introduction

Working in a

laboratory

# Introduction – Working in a laboratory

**Topic glossary**

*Several new terms will be introduced in this topic. These terms are* ***bold*** *in the body text. Please take the time to review the meaning of these terms with your facilitator before you begin this section.*

|  |  |
| --- | --- |
| Term (abbreviation) | Definition |
| Laboratory (lab, labs) |  |
| Analysis/analytical |  |
| Proficiency |  |
| Quality |  |
| Risk |  |

## Working in a lab

This unit of competency applies to laboratory assistants, field/laboratory technicians and instrument operators in all industry sectors. It covers the ability to retrieve data, evaluate formulae and perform scientific calculations, present and interpret information in tables and graphs and keep accurate records. You will learn how to work with scientific quantities and deal with data in the laboratory workplace.



Figure 0.1 – Examples of different types of laboratory.

## What is a laboratory?

Traditionally, we can define a **laboratory** (abbreviated to *lab*, or *labs*) as any building that has the equipment and staff to perform scientific analysis and tests. Typically, this will result in the **analysis** of some material, be it chemical, food, biological material or any other matter. In modern times though, the term ‘laboratory’ has taken on a wider role to include *computer labs*, *data labs* and many other areas outside of traditional analysis.

Let’s consider the what, when, where, why and how of laboratories. To give you a bit more understanding, in terms of the role that labs play in business, laboratories can take the form of:

* **Analytical** laboratories who analyse materials of any nature
* **Quality** control labs which look after the quality of the processes within a lab
* **Proficiency** laboratories check how well the labs are doing as a whole

|  |  |
| --- | --- |
| Practice activity icon | Practice activity 0.1 |

*Perform an internet search of the different types of laboratory. Complete the table below identifying the type of lab and what they analyse. A soil laboratory has been included as an example.*

|  |  |
| --- | --- |
| Type of laboratory | What do they analyse? |
| E.g. Soil laboratory | E.g. Physical and chemical soil properties |
|  |  |
|  |  |
|  |  |
|  |  |
|  |  |
|  |  |

## What is your role in the lab?

To ensure safety and quality, laboratory staff need to follow workplace procedures. Well written procedures and staff who follow them carefully ensure a safe, scientifically credible outcome. Without this, the health of staff may be put at **risk** and the results will lack the required quality.

|  |  |
| --- | --- |
| Practice activity icon | Practice activity 0.2 |

*List some examples of potential consequences of not following workplace procedures in a laboratory.*

|  |  |
| --- | --- |
| What do you think could happen if… | Potential consequence |
| You don’t use glassware properly? |  |
| You don’t use electricity properly? |  |
| You don’t use chemicals properly? |  |
| You don’t use biological material properly? |  |
| You don’t use radioactive material properly? |  |
| You don’t use mechanical equipment properly? |  |
| You don’t handle data correctly? |  |

## Following workplace procedures

When we do work in a lab, the work is usually *standard* work, meaning that it is the type of work that is so routinely performed that the organisation has written a work instruction on how to do it.

One common problem is what these instructions are called, as different industries (and even different companies in the same industry) can use any name they want when designing these instructions. These days there are usually two different types of instruction; *general instructions* and *risk managed instructions*. The only difference between the two is that the process of *hazard identification, risk assessment and control* has been performed on a general instruction to create a risk managed instruction.

Examples of the names used for *general instructions* include:

* Standard Operating Procedures (SOP, one of the oldest terms)
* Written Work Procedure
* Work Instructions (or variants of)

Examples of the names for *risk managed* *instructions* include:

* Safe Work Method Statements (used in the construction and related industries)
* Job Safety/Environment Analysis (JSA/JSEA)
* Task Hazard Analysis (THA)

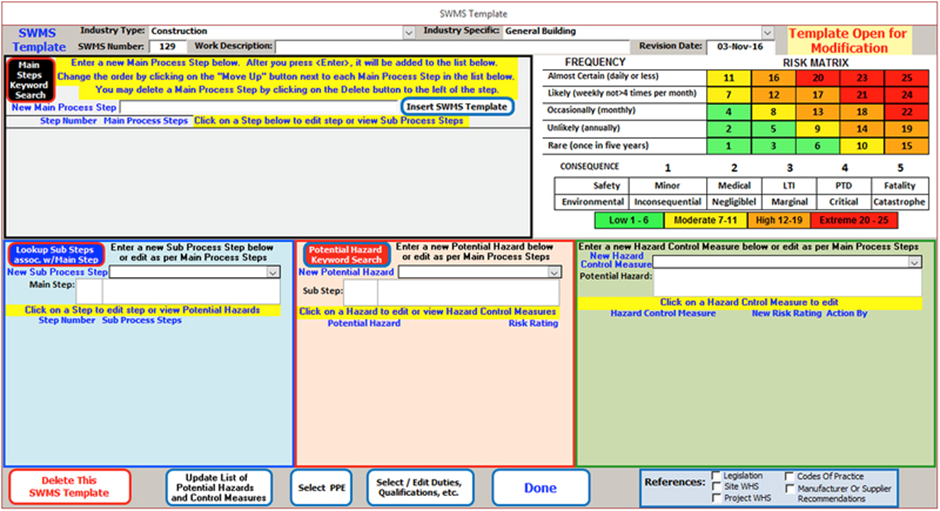


Figure 0.2 – Example of a Safe Work Method Statement.

Examples of the types of work that procedures would cover include:

**Administration related**

* Induction and orientationprocedures
* Customer survey procedures
* Staff training procedures

**Sample related**

* Sample receipt and condition
* Sample collection procedures
* Sample transportation procedures

**Data related**

* Verification of data quality
* Rectifying errors in data
* Managing data (data entry, storage, retrieval, communication and security)

**Analysis related**

* Analytical procedures
* Results calculation procedures
* Statistical analysis procedures
* Results validation procedures

**Maintenance related**

* Equipment installation procedures
* Equipment decommission procedures
* Equipment repair procedures
* Decontaminations procedures

## Process and interpret data workflow

Although every laboratory is different, there are workflows that generally apply. The workflow presented below is scalable to large companies and small, and can be used internally or externally relative to the organisation.

### Analytical problems

The most important aspect of any analysis is to ensure that it will provide useful and reliable data on the material being analysed. The laboratory technician must often communicate with other scientists and non-scientists to establish things like:

* the amount and quality of the information required
* the time-scale for the work to be completed
* Budget.

It is essential for the laboratory technician to have an appreciation of the objectives of the analysis and an understanding of the capabilities of the various analytical techniques at their disposal so they can select or develop the most appropriate and cost-effective method.

### Analytical procedures

The stages or steps in an overall analytical procedure can be summarized as follows:

1. **Definition of the work**

Analytical information and level of accuracy required.

1. **Choice of technique and method**

Selection of the best technique for the required analysis.

1. **Sampling**

Selection of a small sample of the material to be analysed.

1. **Sample pre-treatment or conditioning**

Conversion of the sample into a form suitable for detecting or measuring.

1. **Qualitative/quantitative analysis**

Tests on the sample under specified and controlled conditions.

1. **Preparation of report or certificate of analysis**

Summary of the analytical procedure, the results and their statistical assessment.

The six topics in this workbook will help you understand the basis of measurement, managing data to minimise and rectify errors, calculating scientific quantities and presenting and interpreting data.

Topic 1

Metrology

# Topic 1 – Metrology

**Topic glossary**

*Several new terms will be introduced in this topic. These terms are* ***bold*** *in the body text. Please take the time to review these terms with your facilitator before you begin this section.*

| Term (abbreviation) | Definition |
| --- | --- |
| Measurement |  |
| Measurand |  |
| Value |  |
| Unit |  |
| Quantity |  |
| Reference |  |
| Number |  |
| Metric |  |
| Decimal |  |
| Derived unit |  |
| Uncertainty |  |
| Error |  |
| Reliability |  |
| Systematic |  |
| Random |  |
| Accuracy |  |
| Precision |  |
| Calibration |  |
| Traceability |  |

## What is metrology?

Metrology is the science of **measurement**. Measurement is the assignment of a number to a characteristic of an object or event, which can be compared with other objects or events. Consider that you travelled at a rate of 60 kilometres per hour (60 km/h). What are the *numbers* actually representing and what is a *kilometre* or an *hour*? This is what this topic is about.

The quantity that you are measuring is referred to as the **measurand.** The **value** *of a quantity* is generally expressed as a *number and a unit* (such as 60 km/h). The **unit** is simply the **quantity** which is used as a **reference** (a constant ‘thing’ we can refer to, such as the speed of light), and the **number** represents how many units of the *quantity* we have.

For a particular quantity, many different units may be used. For example, the speed  of a particle may be expressed in the form  = 25 m/s = 90 km/h, where metre per second and kilometre per hour are alternative units for expressing the same value of the quantity speed. This means that you need to convert between units, a skill you learn later.

### The international (metric, SI) system

The system of units and references used in Australia (and the majority of the rest of the world) is called the **metric** system. All this really means is that we use a system of numbers (and units that is based on the number 10). The international body in charge of this is the *International Bureau of Weights and Measures* (in French, *Bureau Internationale des Poids et Mesures*, or BIPM). The task of the BIPM is to ensure worldwide unification of measurements by ensuring:

* fundamental standards and scales for the measurement of the principal physical quantities and maintain any physical standards (such as the kilogram)
* carry out comparisons of national and international standards to correct differences
* ensure the world is using these techniques properly
* carry out and coordinate measurements of the fundamental physical constants relevant to these activities.

### How the metric system works in Australia

BIPM are the international body which feed their equivalents in each country, which, in Australia is the *National Measurement Institute (NMI)*, who also runs the *National Association of Testing Authorities* (*NATA),* which administer the certification of labs that require formal recognition of their laboratory capabilities. Ultimately, the metric system will come down to the people working in the lab (who are required to use the metric system in their work).

Two other contributors to the metric universe are the *Organisation for Economic Cooperative Development* (*OECD*) and the *International Standards Organisation (ISO).* The OECD, aside from looking after trade ties between countries (the numerical base of trade is metric), implement Good Laboratory Practice (GLP). The ISO create industry endorsed Standards that are built around the metric system. The figure below demonstrates these relationships.

Figure 1.1 – A flowchart of quality from international to laboratory level. 

The diagram shows a heirachy with OECD, ISO and BIPM at the top, the second level is NATA and NMI and the third level is the laboratory indicating frameworks which influence laboratory work. 

Figure 1.1 – A flowchart of quality from international to laboratory level.

### Regulatory metrology in Australia

As mentioned above, the National Measurement Institute (NMI) is the arm of the Australian government (through the Department of Industry, Innovation and Science) that are responsible for the implementation of the metric system in Australia.

The principal piece of legislation/regulation that exists is the *National Measurement Act 1960* (NM Act) and the *National Measurement Regulations 1999* (both available from legislation.gov.au). From Section 7B(1) of the NM Act, the NMI issues guidelines, the latest of which is the *National Measurement Guidelines 2016*.

Before you ‘write this off’ as another pointless government document, let’s have a look at what it legally defines. The Act defines the legal units of measurement in Australia, including:

* The SI base (and derived) units of measurement
* The non-SI units for legal use with SI units
* Additional legal units used in Australia (such as inches)

Other details for the correct use of units from the legislation and regulation include;

**Defining how to combine units of measurement**

This is about using the metric prefix system to correctly express unit notation (how we write it out). If we measured a length as being 1.2 x 10-9 metres, we can express this as 1.2 nanometres (nm) but we cannot state this as 1.2 milli-micrometres (mm – wrong!). Single use of the prefix only! Also, the measurement unit of tonne can only have prefixes that are decimal multiples. Never prefix a fraction of a tonne –kilotonne = 🗸, millitonne = 🗴.

**Defining which units must not have a prefix at all**

There is a list of 16 units that cannot be prefixed at all.

|  |  |  |  |
| --- | --- | --- | --- |
| Unit | | | |
| Kilogram | Decibel | Ounce | Second of arc |
| Minute | Inch | Foot/minute | Radian |
| Hour | Foot | Degree (Arc) | Steradian |
| Day | Hectare | Minute | Degree (Celsius) |

Table 1.1 – Legally defined units which must never be prefixed.

Millihour is ‘illegal’ by this definition.

Other key criteria that come from this legislation includes definitions on working with **decimal** *multiple* or *submultiple* units. Consider the following expressions:

* 1.02 kilometres is a *decimal multiple* (greater than 1, expressed with a positive power when written in standard form, 1.02 x 103 m).
* 102 nanometres is a *decimal sub-multiple* (greater than 0, but less than 1, expressed with a negative power when written in standard form as metres 1.02 x 10-7 m).

**Formal notation in metrology**

One potentially confusing part of working with calculations is that the same thing can be expressed in different ways. Consider the extract from Table 1.4 below.

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Name | Symbol | Quantity | Equivalents | SI base unit equivalents |
| Hertz | Hz | Frequency | 1/s | s-1 |

Table 1.2 – Extract of Table 1.4 below for discussion (International Bureau of Weights and Measures, 2006).

In the table extract above, the headings are very informative.

Name: This tells us the name of the unit (Hertz)

Symbol: The name is abbreviated to this symbol (Hz)

Quantity: The quantity is the physical phenomenon being ‘defined’ by the unit (Frequency)

Equivalents: This shows you how else the unit can be expressed as an equivalent (s-1)

SI Base Unit:This is the ‘formal’ notation

Compare the last two columns in Table 1.2. What do you notice? There are no division signs (/) used in the last column. Why? Because there are three ways that we can express the same thing and you need to be able to ‘read’ each of these expressions.

Consider at some point you have travelled at a rate of 14 metres per second. This can be written in three ways:

Or,

Or,

These are all the same thing. The first two are exactly the same apart from the ‘angle’ of the oblique stroke (division sign), and the third way is reciprocal (that is, ‘1 over…’), as division is the same as multiplying one number by the reciprocal of another number!

The final expression here (product expression) is the preferred (but not compulsory) scientific expression. Whichever one you use, be sure you can ‘read’ and understand all three.

### Base units

The term base units refers to the seven primary units from which all other units are derived. The table below lists the properties of the seven base units.

|  |  |  |
| --- | --- | --- |
| Dimension | Unit name | Unit symbol |
| Length | Metre | m |
| Time | Second | s |
| Amount of substance | Mole | mol |
| Electric current | Ampere | A |
| Temperature | Kelvin | K |
| Luminous intensity | Candela | cd |
| Mass | Kilogram | kg |

Table 1.3 – Metric (SI) base units.

Image shows the 7 base units as symbols plus some basic inter-relationships between the units. 


Figure 1.2 – Image showing the 7 base units as symbols plus some basic inter-relationships between the units.

### Derived units

When any of the base units are combined with each other they form what is called a **derived unit** (meaning the new unit was derived from multiple base units).

**Example**

You have been given a box and you want to know its *volume*. You are told that volume is equal to the product of the boxes length, width and height, all of which you measure and find each dimension equal to 1 metre. So, you have performed three individual measurements, all using the base unit of metre (for length). From this we can multiply various measures together to end up with new derived units:

* Multiply the width and length to find the area in metres squared (m2)
* Multiply the area by the height to form volume (m3)

The units of metres squared (m2) and cubed (m3) are examples of derived units, as is below.

| Name | Symbol | Quantity | Equivalents | SI base unit equivalents |
| --- | --- | --- | --- | --- |
| Hertz | Hz | Frequency | 1/s | s−1 |
| Radian | rad | Angle | m/m | 1 |
| Steradian | sr | Solid angle | m2/m2 | 1 |
| Newton | N | Force, weight | kg⋅m/s2 | kg⋅m⋅s−2 |
| Pascal | Pa | Pressure, stress | N/m2 | kg⋅m−1⋅s−2 |
| Joule | J | Energy, work, heat | N⋅m  C⋅V  W⋅s | kg⋅m2⋅s−2 |
| Watt | W | Power, radiant flux | J/s  V⋅A | kg⋅m2⋅s−3 |
| Coulomb | C | Electric charge or quantity of electricity | s⋅A  F⋅V | s⋅A |
| Volt | V | Voltage, electrical potential difference, electromotive force | W/A  J/C | kg⋅m2⋅s−3⋅A−1 |
| Farad | F | Electrical capacitance | C/V  s/Ω | kg−1⋅m−2⋅s4⋅A2 |
| Ohm | Ω | Electrical resistance, impedance, reactance | 1/S  V/A | kg⋅m2⋅s−3⋅A−2 |
| Siemens | S | Electrical conductance | 1/Ω  A/V | kg−1⋅m−2⋅s3⋅A2 |
| Weber | Wb | Magnetic flux | J/A  T⋅m2 | kg⋅m2⋅s−2⋅A−1 |
| Tesla | T | Magnetic induction, magnetic flux density | V⋅s/m2  Wb/m2  N/(A⋅m) | kg⋅s−2⋅A−1 |
| Henry | H | Electrical inductance | V⋅s/A  Ω⋅s  Wb/A | kg⋅m2⋅s−2⋅A−2 |
| Degree Celsius | °C | Temperature relative to 273.15 K | K | K |
| Lumen | lm | Luminous flux | cd⋅sr | cd |
| Lux | lx | Illuminance | lm/m2 | m−2⋅cd |
| Becquerel | Bq | Radioactivity (decays per unit time) | 1/s | s−1 |
| Gray | Gy | Absorbed dose (of ionizing radiation) | J/kg | m2⋅s−2 |
| Sievert | Sv | Equivalent dose (of ionizing radiation) | J/kg | m2⋅s−2 |
| Katal | kat | Catalytic activity | mol/s | s−1⋅mol |

Table 1.4 – The 22 ‘named’ derived units of the metric system.

To summarise, when we take measurements we end up with numbers and units to express the physical meaning of the measured value. The units we use are called metric, for which there are seven base units, and when we combine any two or more of these base units in our routine calculations, we end up with derived units.

## Uncertainty in measurement

Because every measurement is just an *estimate*, we need to have a better understanding of the facts about taking a measurement.

When we perform a measurement, the result is a combination of many factors and not just the reading of an output signal or value of a ruler, for example – there is far more to it than that, and as for more complex measurements, we rarely know all the factors that contribute to the measurement, which ultimately leads to questions about the measurement.

The result of this is called **uncertainty**, and it is something that needs to be calculated and reported so that the end user of the results can trust the measurement and understand the limitations of the value. There will always be slight differences in the measurements if readings are repeated a *number of times under the same conditions*.

In addition, errors may go undetected if the true or accepted value is not known for comparison purposes. Errors must be controlled and managed so that valid analytical measurements can be made and reported. The reliability of such data must be demonstrated so that an end-user can have an acceptable degree of confidence in the results of an analysis.

### Error in analytical measurement

To improve our confidence, scientists have developed ways to estimate the uncertainty in the measured value(s) and how they differ from the true or correct answer(s). Laboratory scientists use a series of terms when describing uncertainty and errors.

The amount by which the true value has been missed is called the **error**. Some errors are brought about by the limitations of the equipment. Analysis of these limitations can be used to calculate the size of the errors. Other errors are linked to limitations in the skill or vigilance of the operator or shortcomings in the method. These errors are much more difficult to identify and compensate for.

Repeated performance of a procedure which achieves good results every time means it is reliable and its errors are under control; this is termed **reliability**. This is used to describe the quality of a technician, a measurement, a technique or a method. Reliability is also known as *repeatability*.

All scientific work is subject to errors and the skill of a competent worker is to understand where errors may come from and how to minimise or compensate for them. Errors normally can only arise from defects in:

* the operator
* the equipment
* the method being followed

*Operator errors* are those errors for which the operator is responsible and can be caused by physical handicaps (e.g. colour blindness), bias, prejudice or poor attitude toward quality and accuracy. Examples of operator errors include:

* mechanical loss or gain of materials (such as during weighing, filtering or liquid transfer)
* failure to obey essential conditions (such as temperatures, times or conditions)
* incorrect performance of required techniques (such as dilution for example)

*Equipment errors* are due to defects in the analyst’s tools and equipment or the effects of environmental factors upon them. Common equipment errors are:

* random defects (such as a balance or oven not operating as expected)
* systematic defects producing results which appear acceptable but are actually in error
* equipment which is not appropriate for the task
* reaction of reagents and samples with glassware and other containers
* use of reagents containing impurities

*Method errors* are due to errors in the procedure or the technique. Common method errors are:

* inaccuracy (the method used is not adequate)
* lack of validity (the method cannot cope with limitations of your sample)
* calculations and processing of data which have been incorrectly carried out
* incorrect sampling or sample preparation
* failure of the necessary reactions to go to completion
* occurrence of side reactions and by-products

Another system uses the terms **systematic** (or determinate) errors and **random** (or indeterminate) errors to classify the errors which affect an experimental result.

*Systematic* errors are those that have a definite value which can be measured and accounted for, that is, they are errors that are possible to avoid, minimise or compensate for.

*Random* errors result from the person, the equipment or the method operating outside its limitations.

These errors may or may not be positively identified and so will not always have a definite measurable value.

### Accuracy and precision

These two terms are related in the sense that they both describe where the value of a measurement sits in relationship to the *reference* or *true value* (that does exist, but we can never know because every measurement is an estimate).

The term **accuracy** refers to the closeness of agreement between a measured quantity value and a true quantity value, whereas the term **precision** refers to a closeness of agreement between indications or measured quantity values obtained by replicate measurements on the same or similar objects under specified conditions. Consider the image below.

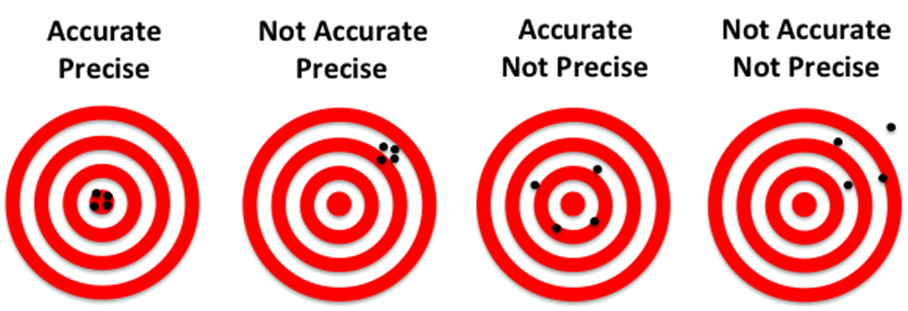


Figure 1.3 – The dartboard analogy for accuracy and precision.

### Relative and absolute descriptions of error

The **absolute error** (or absolute accuracy) is the difference between the observed valueand the true value.

The **relative error** (or relative accuracy) is the absolute error expressed as a percentageof the accepted value. The sign of the error may be positive or negative, indicating thatthe result is high or low respectively.

The **absolute precision** is half of the range of the measurements.

The **relative precision** is the absolute precision expressed as a percentage of the mean ofthe measurements.

| Symbol | Interpretation | Formula |
| --- | --- | --- |
| X | A measured or observed value |  |
| R | The range from biggest to smallest of all replicates for this mesurement |  |
| µ | The average of all replicates for this measurement |  |
| Xtrue | The true or correct value |  |
| Eabs | The absolute error or accuracy | or |
| Erel | The relative error or accuracy |  |
| Pabs | The absolute precision |  |
| Prel | The relative precision |  |

Table 1.5 – Symbols, interpretations and formulas used in error calculations.

To illustrate these symbols and formulae, assume we know the *true value* for vitamin C analysis to be 24.31 mg/L. The data so far has calculated a mean of 24.28 mg/L for a range of figures from 24.18 up to 24.39 mg/L. What are our error calculations for this?

**Absolute error** = 24.31 mg.L-1 – 24.28 mg.L-1

= +0.03 mg.L-1

**Relative error** = (+0.03 mg.L-1 / 24.31 mg.L-1) × 100

= +0.1%

**Range** = 24.39 mg.L-1 – 24.18 mg.L-1

= 0.21 mg.L-1

**Absolute precision** = 0.11 mg.L-1

**Relative precision** = (0.11 mg.L-1 / 24.28 mg.L-1) × 100

= ±0.41 %

How do you interpret these numbers?

An error of 0.1% is superb and the analyst should be proud. Many tasks have much higher error levels and generally errors of 1–2% are quite acceptable.

The precision figure of 0.4% is worse than the error figure and could mean that the operator has fluked a really lucky average figure. The measurement process is giving a scatter of values much wider apart than the final gap to the true answer.

Often precision is the only guide to the quality of work and a precision of 0.4% is still excellent in most forms of testing.

### Calibration

How do you know if the measurement device you are using is telling the truth? Instruments/devices used for measurement need to be checked to find out how accurate they are and adjusted as necessary. This is called **calibration**. The process of calibration is considered an operation that, under specified conditions, as a first step, establishes a relationship between the quantity values with measurement uncertainties provided by measurement standards and corresponding indications with associated measurement uncertainties. In a second step, calibration uses this information to establish a relationship for obtaining a measurement result from an indication.

That is, you measure quantities that are known to see what results you get and use the relationship between the results and the known quantities to find out how accurately your instrument or process is measuring.

The result of a calibration can be given as:

* correction value(s) at discrete measurement point(s)
* calibration equation (curve)
* re-calculated constant(s)
* statement that the error is in given specifications

A calibration result always includes the uncertainty and the probability level should always be given – discussed later.

### Traceability

Metrological **traceability** is the property of a measurement result whereby the result can be related to a reference through a documented unbroken chain of calibrations, each contributing to the measurement uncertainty. It is critical that data is recorded in such a way that it clearly links to the relevant samples, sampling points, production batches or processing records.

Having traceability in lab processes enables a company, for example, to identify the exact batch of suspect raw materials that were used to make a product-lot that caused customer complaints.

In this example below, you can see how properly recorded data connects the wheat moisture test report to the actual batch of wheat.

There are two flow charts in this diagram sitting above a table containing laboratory results for wheat. The flow chart on the left has four components: Wheat silo 10 flows into Sampling which flows into Sample 15 Silo 10 15/2/01, 1530hrs, ID:Sampler Name and this flows into the last component Testing. 
 The flow chart on the right shows that text in boxes and text next to each box. To follow, the text next to each box is in brackets after the text in the boxes. The flow chart starts with Definition of the unit (metre convention) which flows into Reaslisation of the unit: National measurement standard (National Metrology Institute/Laboratory). These two components have a box around them indicating they are part of the SI System of Units. 
These two components then flow into Secondary measurement standard (National Metrology Institute/Laboratory), then Referene measurement standard (accredited laboratory), then Working measurement standard (industrial laboratory) and finally Traceable measurement (end user).

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Wheat Central Testing Laboratory** | | | | |
| Moisture content (AS 1289.2.1.6) | | | Date | **15/2/01** |
| Project No. | BR-1234 | Location: | Rockbank | |
| Sample No. | **15** | Tested by: | **J.Smith** | |
| Description: **Silo 10** | | | | |
| Determination No. | | **1** | **2** | **3** |
| Moisture content (%) | | **9.3** |  |  |

Figure 1.4 – Example of a ‘traceability ladder’.

|  |  |
| --- | --- |
| Practice activity icon | Practice activity 1.1 |

With your teacher’s assistance, find one analytical procedure from the campus or workplace laboratory and discuss the traceability of the test back to the top national level (NMI).

|  |  |
| --- | --- |
| Self-check questions icon | Self-check questions – topic 1 |

*Check your understanding of the basic topics in this section by answering the questions below (or other questions provided by your teacher).*

1. A measurement is a combination of two parts. What are the two parts that make a measurement?
2. Generally speaking, what is a unit of measure? Briefly explain how we reference a constant.
3. What is the difference between a base unit and a derived unit of measure?
4. Can anyone trust results that come from uncalibrated equipment? Explain.
5. Name one international body that contributes to metrology. What does this organisation do?
6. What do you think would be one consequence of having poor traceability?
7. A series of repeated measurements were taken. The results were very close to each other, but not quite near the true value. Would these measurements be precise or accurate? Why?

Topic 2

Data

# Topic 2 - Data

#### Topic glossary

*Several new terms will be introduced in this topic. These terms are* ***bold*** *in the body text. Please take the time to review these terms with your facilitator before you begin this section.*

|  |  |
| --- | --- |
| Term (acronym) | Definition |
| Data |  |
| Sample receipt |  |
| Calculations |  |
| Quality checks |  |
| Laboratory Information Management System (LIMS) |  |
| Audit |  |
| Quality assurance (QA) |  |
| Quality control (QC) |  |

Now that we have a better understanding of what a laboratory is and does, as well as understanding the values (that is, numbers and units) associated with laboratory analysis, we can discuss the flow of data that will accompany the work done to get it.

## Data sources

The first question we can ask is *where can* **data** *come from*?

This question is best answered by exploring what a laboratory does from beginning to end. Generally speaking (as each lab is different), the workflow in a lab will consist of the following stages:

* Customer enquiry
* **Sample receipt**
* Sample pre-treatment
* Sample testing and analysis
* **Calculations**/determinations
* **Quality checks**
* Reporting of results

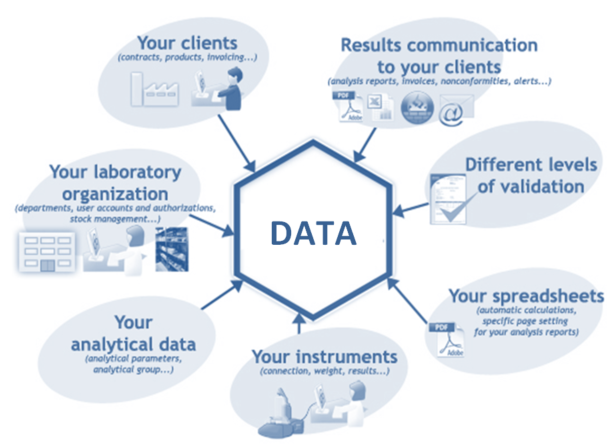


Figure 2.1 – Example of a generalised laboratory data workflow.

|  |  |
| --- | --- |
| Practice activity icon | Practice activity 2.1 |

What type of data could be generated from each stage of data workflow in your workplace or college? Once the data is collected, how could that data be processed? Two examples have been provided for clarity. Complete the remaining rows.

| For this type of work… | What type of data is it? | How could the data be processed? |
| --- | --- | --- |
| Customer service | Names, locations, addresses, type of work | Analysed for sales trends |
| Sample receipt |  |  |
| Sample preparation |  |  |
| Analysis |  |  |
| Calculation processes |  |  |
| Statistical treatment |  |  |
| Quality | Compliance data | Compared to other ‘benchmark’ data |

## Storing data

The previous section explored the sources of data. Once data has been generated, the data needs to be stored appropriately. The ‘path’ data takes depends on how it was collected in the first place – physically or electronically.

If the data is generated electronically, then it is likely that the data will be stored electronically. If the data was physically recorded (that is, written on paper for example), then the physical paper will need to be stored in an equally physical location. Data recorded physically is typically transcribed into electronic data for easier handling by computers or **Laboratory Information Management System (LIMS)**, which leads to a workflow termed ‘data entry’.

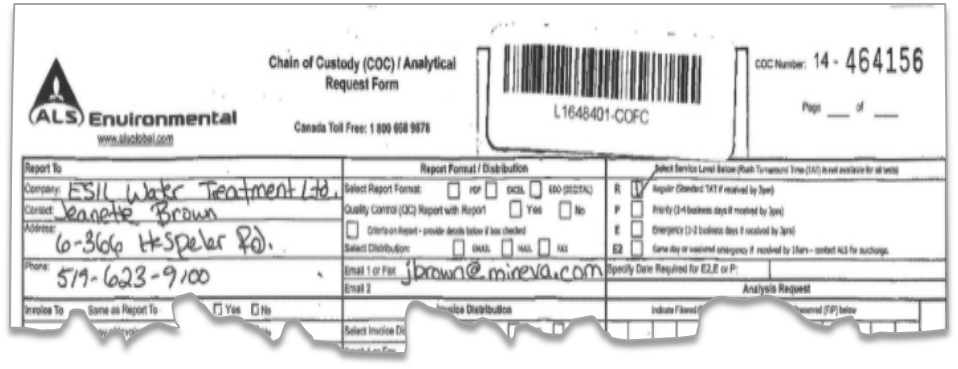


Figure 2.2 – Example of routine paperwork used by laboratories. In this case a document called a Chain of Custody form which is used for legal reasons to ensure the integrity of the sample is kept during transportation.

Physical data can be stored in:

* Filing cabinets

Electronic data can be stored in:

* Computer databases
* Spreadsheets
* LIMS (online or cloud databases)

The storage of data is a standard workflow process. A sample is worthless unless its origin and/or history are recorded. The origin of a sample is the place and time where it was taken, such as a particular work area or batch.

The history of a sample is a record of how the sample was taken, how it was transported and stored and what has been done to test it. Samples must be carefully labelled to avoid confusion and to make sure that the sample can always be linked to the place and time where it was taken.

All data obtained must be coded. This means that the data must be identified in some way so that it is linked to its origin and history. For example, if you have obtained data from testing wine bottles, when you record that data you must identify the crates or batch that each sample was obtained from.

The same applies to any equipment that are calibrated (like thermometers). These are coded so that accurate calibration records can be kept. Careful coding will give you an **audit** trail. An audit trail is a way that you can trace any sample or result back to its source.

This is particularly important for two reasons:

1. In a quality system, you must be able to prove the reliability of your test results by being able to trace back from the result to find out exactly how tests were done and how the samples were collected.
2. If there is a problem with a result, it may be necessary to take some form of corrective action. For example:
   1. If there was a food poisoning outbreak caused by salami, it would be necessary to know exactly which batch of salami caused the problem.
   2. If a customer questioned the reliability of the results of your testing, you must be able to show exactly how the results were obtained, what equipment was used and whether it was in proper working order at the time.

## Retrieving data

Data is stored for many reasons. If data is stored for a reason, then it makes sense that it may need to be retrieved for similar reasons that it was originally stored for. The data retrieval will be based on how the data is stored, so if you physically store, you will physically retrieve the data, and the same for electronic data.

Most companies have developed forms for recording and storing data as part of a Quality Assurance (QA) program. In a QA system, forms are usually coded by number or date and filed in order so they can easily be found again. The forms are usually filed in a QA filing system maintained by the plant’s QA manager. This is done so that the results can easily be retrieved when needed.

In addition to hard copy forms, much data these days is stored electronically on spreadsheets or databases to allow speedy retrieval and processing of information.

## Checking data

At some point during the workflow, the data that is stored needs to be checked to see if it is fit for purpose. This can be performed at any stage, and often at multiple stages, which makes sense.

But how do we check data? How do you check anything? Usually by comparing it to a known value, or acceptable range of values. Believe it or not, you do this all the time when you shop – the cashier will ask for money from you and you compare the stated value against the ticketed price to check the cost. This is exactly what we do in the workplace.

Checking the quality of data is part of a much bigger process called **Quality Assurance** (**QA**), as part of a work function called **Quality Control** (**QC**, although the two terms are combined and commonly called **QA/QC**). Ignoring the bigger QA/QC part at the moment, and focusing just on data checking, the theory states that there are in fact six processes (core dimensions) associated with the checking of data alone, as illustrated in the figure below.



Figure 2.3 – Checking the quality of data using the 6 data quality dimensions.

So, data is only useful if it is correct and presented in the appropriate format. The quality of data relies on a number of factors that include:

* correct recording of data onto primary documents such as test sheets and QC sheets
* correct spelling of names and addresses
* correct numerical data such as date of birth and driver’s licence number
* checking that the check digits are correct
* correct transcription of this data from the primary document to a secondary document or to a computer spreadsheet or database
* checking for transcription errors
* checking that all coding that links this data to a particular customer, patient or process is correct; for example, a pathology sample will be linked by at least two pieces of information such as full name of patient and date of birth
* ensuring that the sample is suitable for testing or processing
* ensuring that the correct tests were performed on the sample
* checking that QC data that may confirm the quality of the data is within limits
* checking that all sign offs and double checks have been completed

A common expression in the computer industry is ‘garbage in, garbage out!’ The computer cannot fix poor data. It will just process that data to produce ‘garbage’ results. Adopt the habit of always double checking data.

## Rectifying errors

Only by acknowledging errors and investigating them can we begin to understand the causes of problems. When you know what the causes of the problems are, you can put work procedures in place to ensure that, in the future, the quality system works more effectively.

There are a number of golden rules for handling errors in data that must be followed for an enterprise to function properly. They are not difficult rules to understand or apply and can be summed up as:

1. Have processes in place to identify the error. A process of systematic checking should discover the mistakes before they are passed onto a customer.
2. Acknowledge the mistake openly rather than cover it up. Every laboratory will make mistakes sooner or later.
3. Mistakes and errors must be promptly dealt with. This may mean amending a calculation or repeating a test.
4. The new (correct) data or result should be added to the record, and clearly linked to the original data that was found to be incorrect.
   1. The incorrect data should not be erased, but it must be made absolutely clear that it is no longer valid and a brief reason should be given for the change.
   2. Neatly crossing out a faulty result, rather than hiding it with correcting fluid allows you to keep a record of the problems that occur with testing and data processing.
5. The person making the correction must initial/sign and date the correction.
6. Reports should be issued free from erasures if possible. You don’t need to advertise your mistakes to a customer if you pick them up and can correct them before the report is issued.
7. If results have already been issued to a customer before a mistake was discovered, the customer must be promptly informed of the error and a new statement of results issued that is clearly annotated ‘Amended Result’.
8. Learn from the error and amend processes to prevent this happening again.

There are a number of reasons why data errors in the laboratory are amended in such a way as to leave both the incorrect and correct result in the records. These are:

* It will be clear if someone tries to alter the record, at a later date, for dishonest purposes.
* Correct data that is incorrectly thought to be incorrect will not be erased and lost forever!
* The causes of the problems may be properly investigated in order to prevent the mistakes from being repeated.

|  |  |
| --- | --- |
| Self-check questions icon | Self-check questions – topic 2 |

*Check your understanding of the basic topics in this section by answering the questions below (or other questions provided by your teacher).*

1. Why is it so important to check data quality? Give at least three reasons why.
2. Discuss with your teacher and provide two examples where poor data could cause problems in an analysis.
3. A 20mL volumetric flask is filled to the mark and weighed. This is repeated five times. Assume that the temperature is the same at each weighing. Which of the following results might be a transcription or other type of error: 19.968g, 19.994g, 19.399g, 20.013g, 20.002g. Discuss the reasons for the variation and sources of error with your teacher.
4. Why do you think the process of rectifying errors is a formal process?
5. A technician has written down the following results in the laboratory and then typed them into a table. Have any errors been made? What is the source of the error?

|  |
| --- |
| Notebook  3/5/18 Readings: SL2018-01 8.562mg/L SL2018-02 9.362mg/L  4/5/18 Reading: SL2018-03 4.695mg/L |

|  |  |  |
| --- | --- | --- |
| Date | SL Number | Reading (mg/L) |
| 03/05/2018 | SL2018-01 | 8.662 |
| 04/05/2018 | SL2018-02 | 9.362 |
| 04/05/2018 | SL2018-03 | 8.695 |

1. Identify 4 sources of data from your workplace or campus laboratory. Indicate whether the data is in written or digital form.
2. What does LIMS stand for? What does it do for a laboratory?

Topic 3

Calculating

Scientific

quantities

# Topic 3 - Calculating scientific quantities

#### Topic glossary

*Several new terms will be introduced in this topic. These terms are* ***bold*** *in the body text. Please take the time to review these terms with your facilitator before you begin this section.*

| Term | Definition |
| --- | --- |
| Measurement |  |
| Estimate |  |
| Ratio |  |
| Significant figures |  |
| Standard form |  |
| Scientific notation |  |
| Unit conversion |  |
| Prefix |  |
| Order of magnitude |  |
| Fractional |  |
| Decimal |  |
| Algebra |  |
| Equation |  |
| Formula |  |
| Transposition |  |
| Confidence level |  |

## Calculating scientific quantities

Remember, a number such as 0.01 is just a number, which to us in the real world is meaningless (the number of what?). To give meaning to numbers, scientists use *units* (or *dimensions*). These provide a meaningful reference to which we can compare the physical world. Think of a typical school ruler used to measure length; they are typically in the *unit* of *centimetres*. This topic will teach you how to perform calculations using numbers and their quantity values.

## Calculation skills

One definition of a calculation is:

“A mathematical determination of the amount or number of something.” (Oxford Dictionaries, 2019)

You would definitely be familiar with some simple daily calculations. Consider the routine calculations you do every day such as calculating how much money you have left over from shopping, or how much fuel you need to get somewhere, or how long an activity will take in terms of time. These are all calculations, even though you may perform them in your head, and even though they are only estimated.

### Estimations

Every **measurement** is an **estimate** because we will never know the true value (at least in terms of how many decimal places we can report). So, every calculation you perform with those measurements is also an estimate. So is every other determination you perform with the data or information you collect. In fact, it doesn’t matter how careful you are, or how complex or detailed the calculation is, *it will always be an estimate*.

Why? Because most physical phenomena are not discrete values, but rather continuous values, which means that we round off (or chop) the values to express how precise we are. Estimating can involve rounding the individual numbers in a calculation to the nearest 0.1, 1, 5, 10 or 100, as appropriate. In this way the calculation is very easy and often able to be done in your head.

For this topic, you also need to remember that calculations come in two varieties, with or without units. Calculations that are truly unitless (in other words, playing with just the numbers) are actually quite rare. We generally don’t calculate something without using their physical meaning (we are analysing the physical world after all!). Note that in the lab, almost every number will have a unit. The number for *pi* (~3.14159) is an example of a unitless number and so is every **ratio** as the units will cancel out.

#### Example for review

Routine preparation of 11 litres (L) of an acid solution requires using 97 millilitres (mL) of the concentrated acid. One day you need to make 6 L of the solution instead. How much of the compound is needed? A quick estimation: 6 L is about ½ of the usual volume and 97 mL is approximately 100 mL, therefore you need around 50 mL. The accurate calculation (6/11 x 97 mL) shows you should use 52.9 mL. Because the calculation agreed with the estimate you have confidence that the answer was correct.

|  |  |
| --- | --- |
| Practice activity icon | Practice activity 3.1 |

Estimate the answer for each of the following calculations and compare your answer to the accurate answer from your calculator:

|  |  |  |
| --- | --- | --- |
| 1. (9.30 x 10.4)/1.30 | Estimate: | Calculated answer: |
| 1. (0.12 x 5.9)/96 | Estimate: | Calculated answer: |
| 1. (0.0022 x 8.9)/4.5 | Estimate: | Calculated answer: |
| 1. (√26 x 4.7)/92 | Estimate: | Calculated answer: |

### Significant figures

We have already seen that all measurements are not exact. Often if calculations are performed (such as in working out uncertainties) it is tempting to quote answers to an excessive number of decimal places.

* How many decimal places should an answer be quoted to?

To answer this question we must learn a little bit about significant figures.

* Which numbers are significant?

The number of significant figures is not the same as the number of decimal places. Below are some guidelines for determining the number of **significant figures** in a measurement:

**All non-zero digits are significant:**

For example, 356.4 has 4 significant figures.

**All zeros between significant digits are significant:**

For example, 30.045 has 5 significant figures.

**All zeros to the right of a decimal point are significant:**

For example, 2.00 has 3 significant figures.

**Zeros to the left in measurements less than one are not significant:**

For example, 0.0024 has 2 significant figures.

**Zeros trailing in a whole number are not significant:**

For example, 200 is considered to have only ONE significant figure while 25,000 has two.

The figures below all have three significant figures.

|  |  |  |  |
| --- | --- | --- | --- |
| 24.0 | 0.385 | 0.0490 | 907 |
| 10.0 | 0.00621 | 2.00 x 106 | 9070 |

|  |  |
| --- | --- |
| Practice activity icon | Practice activity 3.2 |

How many significant figures are there in the following?

a) 183

b) 1024

c) 0.04

d) 15.301

e) 91.00

### Scientific notation

Scientists often use **Standard Form** (also known as **Scientific Notation** – used interchangeably) to write a number because it is easier to express numbers that are very large or very small. Scientific Notation expresses any number as:

(a number between 1 and 9.999999...) x 10power

Firstly, it is necessary to look at the meaning of negative powers of 10.

Which means for example

The following table is helpful:

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| **Number** | 100 | 10 | 1 | 0.1 | 0.01 | 0.001 |
| **10power** | 102 | 101 | 100 | 10-1 | 10-2 | 10-3 etc. |

To express any number in Standard Form, move the decimal point in the original number to ‘make’ a number between 1 and 9.99999 and continue as in the following examples.

#### Examples for review

1. Express 456 000 in Standard Form.

*456 000 = 4.56 x 100 000 = 4.56 x 105*

1. Express 0.003 456 in Standard Form.

*0.003 456 = 3.456 x 0.001 = 3.456 x 10-3*

Note that the power of 10 is always equal to the number of places the decimal point has moved – positive if the decimal point moves to the left, negative if it moves to the right.

Always look at your answer to see if it makes sense.

For example, multiplying 4.56 by 104 would make a larger number, which is what we started out with. Likewise, multiplying 3.456 by 10-3 would make a smaller number, again as in the original.

1. Express 5.980 x 105 as a decimal number.

5.980 x 105 = 598 000 (just move the decimal point 5 places to the right)

Notice that the number in Standard Form clearly shows that this number has 4 significant figures – the last zero does count!

1. Express 3.456 x 10-4 as a decimal number.

3.456 x 10-4 = 0.0003456 (just move the decimal point 4 places to the left).

|  |  |
| --- | --- |
| Practice activity icon | Practice activity 3.3 |

Express the following in Standard Form (Scientific Notation):

|  |  |
| --- | --- |
|  |  |
| 1. 270000 | Answer: |
| 1. 430 | Answer: |
| 1. 0.00008 | Answer: |
| 1. 0.5 | Answer: |
| 1. 803000000 | Answer: |

Express as decimal numbers:

|  |  |
| --- | --- |
|  |  |
| 1. 3 x 104 | Answer: |
| 1. 2 x 10-3 | Answer: |
| 1. 2.6 x 10-1 | Answer: |
| 1. 8.95 x 10-5 | Answer: |
| 1. 7.32 x 104 | Answer: |

### Unit conversions

Converting a unit of measure from one unit expression to a different expression is referred to formally as *dimensional analysis*, though you will typically hear this referred to as **unit conversions**.

Why do we need to do this? Because you *need* to. Especially when you are *working* with *standard* calculations where the units are already defined, but you need to *report* in a different unit, then you will need to undertake a unit conversion.

Imagine you have a friend who lives in America, where they use British imperial units, and you are describing how fast the cars were driving in a race you both watched. Your friend was describing the fastest car as travelling at 212 miles per hour (mph), but what does that mean? You ask your friend what 212 mph is in kilometres per hour (km/h), but unfortunately your friend cannot help, as they only use Imperial units. You need to convert *from* Imperial units *to* Metric units. There are many science related stories where incorrect unit conversions have caused problems (see below).

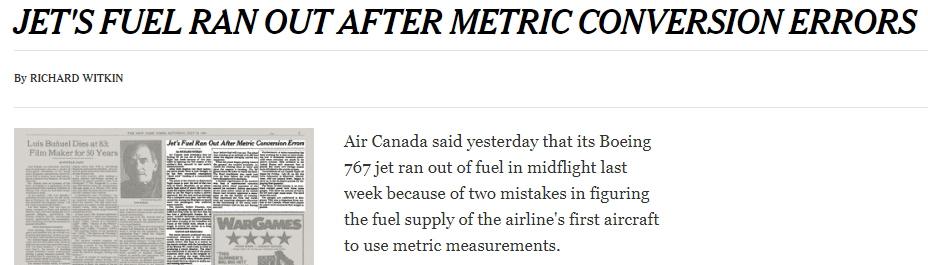


Figure 3.1 – Story from the New York Times from 30/07/1983.

#### Metric unit conversions

You would recall from Topic 1 that we discussed the metric SI system of units. A defining component of that system is the use of decimals. The table below should help clarify what we mean.

| Prefix Name | Number of Base Units (Magnitude) | Notation (Scientific) | Symbol |
| --- | --- | --- | --- |
| yotta | 1 000 000 000 000 000 000 000 000 | 1024 | Y |
| zetta | 1 000 000 000 000 000 000 000 | 1021 | Z |
| exa | 1 000 000 000 000 000 000 | 1018 | E |
| peta | 1 000 000 000 000 000 | 1015 | P |
| tera | 1 000 000 000 000 | 1012 | T |
| giga | 1 000 000 000 | 109 | G |
| mega | 1 000 000 | 106 | M |
| kilo | 1 000 | 103 | k |
| hecto | 1 00 | 102 | h |
| deca | 1 0 | 101 | da |
| Base unit | 1 | Any Base Unit |  |
| deci | 0.1 | 10-1 | d |
| centi | 0.01 | 10-2 | c |
| milli | 0.001 | 10-3 | m |
| micro | 0.000 001 | 10-6 | u |
| nano | 0.000 000 001 | 10-9 | n |
| pico | 0.000 000 000 001 | 10-12 | p |
| femto | 0.000 000 000 000 001 | 10-15 | f |
| atto | 0.000 000 000 000 000 001 | 10-18 | a |
| zepto | 0.000 000 000 000 000 000 001 | 10-21 | z |
| yocto | 0.000 000 000 000 000 000 000 001 | 10-24 | y |

Table 3.1 – The metric pre-fixes and their magnitudes from yocto to yotta.

The metric system uses **prefixes**, which are names describing the change in the number of the zero’s used. Each change (up or down) is termed an **order of magnitude**, but you may notice that the numbers in Table 3.1 change by three zeros, or three orders of magnitude. This is because we use the numbers to describe the magnitude of the prefix in-between three orders:

[**0.001** grams= **1** milligram] [**0.010** grams= **10** milligrams] [**0.100** grams= **100** milligrams]

So *one* prefix covers *three orders* of magnitude! This means that we don’t have to remember the prefix name for every single order – only every third. *And really, we just move the decimal place!*

#### How do the prefixes work?

The first skill is learning to use the prefixes themselves. When we perform a metric unit conversion, we simply combine the *prefix* with the *base* *unit*. The unit of *millimetre* is simply the base unit *metre* with the prefix of *milli*. It’s that simple!

|  |  |
| --- | --- |
| Practice activity icon | Practice activity 3.4 |

*With your teacher’s assistance, complete the following examples of working with metric prefixes (the first one is done for you).*

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Prefix | Base unit | Combined unit | Prefix | Base unit | Combined unit |
| nano | metre | nanometre | kilo | gram |  |
| milli | second |  | micro | ampere |  |
| milli | mole |  | kilo | metre |  |
| micro | ampere |  | peta | candela |  |

#### How does the conversion work?

The second step is performing the actual unit conversion.

These conversions are reversible, so there is two directions that you need to consider. To aid your understanding, consider the change in the magnitude. This can only have two outcomes:

* When changing from a larger unit to a smaller unit, the *number* will get bigger.

**10** *grams* = **10 000** *milligrams*

* When changing from a smaller unit to a larger unit, the number will get smaller.

**10** *grams* = **0.1** kilograms

How can this be? Haven’t we changed the quantity of material? Think about it. No, all we are doing is changing the *number*, not the *quantity* of the *physical attribute*. So why do we do this at all? Well, sometimes it is as simple as the new expression is easier to read (similar effect to scientific notation), other times, it is a legal requirement to express things in a particular unit. Consider the following the following.

***If you have 10 grams of salt, you have 10 000 milligrams of salt, or 0.010 kilograms of salt***

It is always the same amount of substance; *that does not change* (in fact, if it has, you may have done something wrong).

So, how do we perform metric unit conversions? In these notes we will use a technique called *dimensional analysis* (explained later in this Topic) because it is an almost foolproof method to perform conversions, but your teacher may prefer to show you another method.

Table 3.1 explains to us the magnitude of the prefixes relative to the base unit of 1, but how do we actually use this information? Take the *milli* prefix for example, as a fraction, it is ‘1 1000th of a base unit’, or written, we write 1/1000, and as a decimal we write 0.001. It is these last two that are the ‘problem’, because although they look very different, they are the same thing expressed differently!

‘1 1000th’ is the same as saying or **0.001**

There are two expressions we can use: the **fractional**  or the **decimal** 0.001. At this point, it does not matter which expression we use due to how we calculate them, so let’s solve for both expressions. How do we use our choice? Consider the relationship table below, you can clearly see that irrespective of what relationship we use (fractional or decimal) we end up with the same answer.

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| 10 grams | = | 10 g x | 1000 mg | = | 10 000 mg |
| 1 g |
|  |  |  |  |
| 10 g x | 1 mg | = | 10 000 mg |
| 0.001 g |

#### Unit conversions using dimensional analysis

In science and engineering, generally we find unit conversions get a bit wild and crazy the further we get into hard science, and we need a universal technique to solve more complex unit conversions. The ‘proper’ term for unit conversions is *dimensional analysis.*

The term is literal…we use this to analyse *dimensions* (where *dimensions* means what you are converting from and to, and *analysis* refers to the relationships between the *dimensions*). We speak in terms of laboratory related calculations, but dimensional analysis is far bigger than that and can be used to convert anything to anything – assuming that you know the relationship between the units!

We will keep this very simple here, but just to show you what happens at the coal face of hard science, consider the example of a complex dimensional analysis (unit conversion) below.

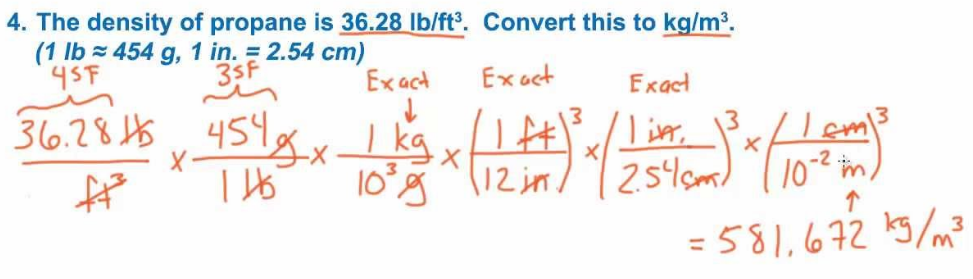


Figure 3.2 – Example of a complex unit conversion. You will not necessarily need to calculate something like this, but then again…

So how does dimensional analysis work? We will use a relationship table to practice with, and this technique will serve you well and carry you through into any industry, so learn it well!

In relationship tables, vertical lines (|) are multiplication, horizontal lines (  ) are division.

**Step 1: Write the conversion unit out the front (left of the = sign)**

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| 15 kg | = |  |  |  |  |
|  |  |  |  |

**Step 2: Rewrite the conversion unit in the top left cell of the table**

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| 15 kg | = | 15 kg |  |  |  |
|  |  |  |  |

**Step 3: Write the first relationship, being converted ‘to’ into the second column**

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| 15 kg | = | 15 kg | 1000 g |  |  |
|  | 1 kg |  |  |

**Step 4: Cancel out the starting units (not the numbers!)**

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| 15 kg | = | 15 kg | 1000 g |  |  |
|  | 1 kg |  |  |

**Step 5: Separately multiply through the top and the bottom**

= 15 x 1 000

1

**Step 6: Solve to finalise the conversion**

= 15 x 1 000

= 15 000 g

If we wanted to perform a more complex (multistep) conversion, we simply keep adding relationships into the table (before we solve) and then work it through, in this case to micrograms.

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| 15 kg | = | 15 kg | 1000 g | 1000 mg | 1000 g |
|  | 1 kg | 1 g | 1 mg |

= 15 x 1000 x 1000 x 1000

1 x 1 x 1

= 15 000 000 000 μg (or 1.5 x 1010 g)

**A note about the internet**

Access to the internet makes this process very easy to get results, but in no way helps you learn how to perform these calculations. Learn this technique first, then you can Google “convert…’

**Example**

In this example the relationship between the *existing* unit (kg) and the *desired* unit (g) is 1 to 1000 (there are 1000 g in one kg). The example follows all of the simple rules; *the units are swapped between the top and bottom, we cancel out and we solve*. Note that we were converting from a ‘large’ unit (kg) to a ‘small’ unit (g), which means we convert from a ‘small’ number (15) to a ‘large’ number (15000).

1. Convert 15 kg to grams.

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| 15 ~~kg~~ | 1000g |  |  |  |
|  | 1 ~~kg~~ |  |  |  |

15 kg =

= 15x1000

1

= 15 x 1000

= 15 000 g

In this example, we can see that it is simply the ‘reverse’ of the problem above, which is due to the fact that we are going from a ‘small’ *unit* to a ‘large’ *unit*, which means we must go from a ‘large’ *value* to a ‘small’ *value*.

1. Convert 1236 grams to kilograms.

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| 1236 ~~g~~ | 1 kg |  |  |  |
|  | 1000 ~~g~~ |  |  |  |

1236 g =

= 1236 x 1

1000

= 1236 / 1000

= 1.236 kg

Finally, in this example we have included scientific notation, which as you can see, is simply calculated through the conversion without converting it to a whole number.

Convert 3.6 x 10-3 grams to nanograms.

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| 3.6x10-3 ~~g~~ | 1000 ~~mg~~ | 1000 ~~μg~~ | 1000 ng |  |
|  | 1 ~~g~~ | 1 ~~mg~~ | 1 ~~μg~~ |  |

3.6x10-3 g =

= (3.6x10-3) x 1000 x 1000 x 1000

1

= 3.6 x 106

= 3 600 000 ng

|  |  |
| --- | --- |
| Practice activity icon | Practice activity 3.5 |

With your teacher’s assistance, complete the following unit conversions.

1. How many ***milligrams*** is 3.674 ***kg***?

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  |  |  |  |  |
|  | ~~kg~~ | ~~g~~ |  |  |

3.674 kg =

=

=

=

1. Convert 9.7 grams to milligrams.

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  |  |  |  |  |
|  | ~~g~~ |  |  |  |

9.7 g =

=

=

=

1. Convert 35856 picograms to kilograms (express in both normal and sc. not.).

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
|  |  |  |  |  |  |
|  | ~~pg~~ | ~~ng~~ | ~~µg~~ | ~~mg~~ | ~~g~~ |

35856 pg =

=

=

=

1. Convert 1.12 ***carats*** (c) ***milligrams*** (note 1 carat = 200mg).

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  |  |  |  |  |
|  | ~~c~~ |  |  |  |

1.12 carats =

=

=

=

|  |  |
| --- | --- |
| Practice activity icon | Practice activity 3.6 |

*With your teacher’s assistance, complete the following metric unit conversion problems.*

| **No.** | **Convert** | **To** | **Answer** |
| --- | --- | --- | --- |
| 1 | 789 nm | m |  |
| 2 | 0.848 g | µg |  |
| 3 | 4.25 x 10-4 kL | mL |  |
| 4 | 270 mm2 | m2 |  |
| 5 | 22 yards | m |  |
| 6 | 62.3 kg | pounds |  |
| 7 | 8456 kL/hr | L/s |  |
| 8 | 82 µg/g | mg/kg |  |
| 9 | 23.45 g/100 mL | mg/L |  |

#### Unit Conversions using the factor method

This is the preferred method if you have access to tables of conversion factors such as the one below.

|  |
| --- |
| **Length**  1 m = 39.37 in 1 ft = 0.305 m |
| **Volume** |
| 1 litre = 1  10-3 m3 = 1  103 cm3 1 cm3 = 1 cc = 1 mL |
| 1 ml = 1  10-6 m3 1 gallon = 4.544 litres |
| **Mass** |
| 1 metric tonne = 1  103 kg |
| **Density** |
| 1 kg/m3 = 1  10-3 g/cm3 |
| **Pressure** |
| 1 Pa = 1 N/m2 = 1.45  10-4 psi 1 torr = 133.3 Pa |
| 1 atm = 1.013  105 Pa 1 lb/ft2 = 47.880259 (6dec places)Pa 1 mbar = 100Pa |
| **Energy** |
| 1 J = 9.49  10-4 Btu 1 cal = 4.186 J 1 dietary cal = 4.186kg |
| **Power** |
| 1 horsepower = 745.7 W |
| **Magnetic Flux Density** |
| 1 T = 1  104 Gauss (G) 1 T = 1  109 Gamma |
| **Illumination** |
| 1 lux = 9.290  10-2 foot-candle |
| **Luminance** |
| 1 candela/metre2 = 3.142  10-4 lambert (or lumen/cm2) |

Table 3.2 – Some common conversion factors to use in unit conversions. Refer to Appendix A for a more comprehensive list of factors.

So, in this method, because we have tables of values that state the numerical factor for a defined relationship, we simply take our quantity value and multiply it by the conversion factor.

**Step 1** – Find the prefix and its relationship (conversion factor) from Table 3.2 (or similar)

**Step 2** – Multiply the problem value by the conversion factor.

**Example**

Convert 0.74 metres to inches. We find the relationship in the table to find the conversion factor (in this case 1 m = 39.37 inches). Now, in reality, you simply multiply them together, but the example below shows the whole expression and shows you how the metre units cancel out.

|  |  |
| --- | --- |
| Practice activity icon | Practice activity 3.7 |

*Refer to table 3.2 above for conversion factors and perform the following conversions.*

a) Convert 1.45 litres to m3 Answer:

b) Convert 4500 lb/ft2 to Pascals Answer:

c) Convert 3.7 Horsepower (hp) to Watts Answer:

d) Convert 20 mL to m3 Answer:

e) Convert 100 000 Pascals to atmospheres Answer:

## Using equations and formulae

When it comes time to perform laboratory calculations, you are entering the field of mathematics called **algebra**. In its simplest form, algebra is mathematics with symbols, where the use of the symbols (and numbers) follows a strict set of rules. The way our calculations work is simple:

1. We attach meaning to the symbols (for example, *m = 6*)
2. We arrange the symbols into a functional formula (for example, *y = mx + b*)
3. We then substitute the symbols for numbers (for example, y = 0.345 x 20 + 0.0012)
4. We then solve the equation to find the unknown value (y = 6.9)

The good news is that we only use a limited amount of calculations, and they are usually very routine, not to mention that in the workplace you are typically given all the formulas because no employer wants to pay you to waste time figuring out calculations which increases the risk of you performing the calculation incorrectly.

### Formula(e) or equation?

This question is old and the answer is not always clear, but generally, the following definition will work for now.

An **equation** describes an equality, such as A+6=13.

Because this is an equation, defining that both sides are equal, we can re-arrange to solve for the unknown term in this equation.

A **formula** is a special type of equation (or other mathematical expression not involving equalities) that shows the relationships between different variable terms. So, we can view a formula as a general form of an equation, and can view an equation as the working version of a general formula. How does that make sense? Consider the following formula for the dilution equation (a mass balance equation):

*C1V1=C2V2*

We can describe this as a formula as it tells us everything we need to know to calculate a mass balance problem. The issue is that it is not in a useful form, so we rearrange (transcribe) the formula so that it is only solving for 1 term (the one we don’t know the answer to), and then we substitute and solve.

### Transposition

Sometimes when we are working with an equation, we find that the variable out the front of the equation is not the one we wish to solve for. When this happens, we need to rearrange the equation so that we can solve for the term we want to use through a process known as **transposition**.

We more commonly refer to this process as *rearranging the equation*, or *finding the subject of the formula*, or similar terms depending on your schooling.

This problem is commonly found with the dilution equation which is expressed as:

Where (ignoring units):

C1 = The starting concentration

V1 = The transfer volume

C2 = The final concentration

V2 = The final volume

In the current form, *C1V1=C2V2*, we have the *formula* (which generalises how dilution works) from which we need to find a specific *equation* to solve for a variable, and to do that we will *transpose* the formula following some basic rules which will allow us to rearrange the equation and finally solve what we are after.

#### The rules of transposition

**Rule 1:**

You can add, subtract, multiply and divide by anything, as long as you **do the same thing to both sides of the equals sign**. In an equation, doing the same operation to both sides keeps the meaning of the equation from changing.

Let's use the equation for a line to illustrate an example of how to use Rule #1. The general equation for a line is:

[equation for a line](https://serc.carleton.edu/details/images/11093.html)

If we wish to solve for b in this equation, we must subtract mx from both sides.

[first step in solving line equation for b](https://serc.carleton.edu/details/images/11094.html)

If we perform the math on each side (that is, subtract mx from mx on the right), we end up with an equation that looks like this:

[equation for a line solved for b](https://serc.carleton.edu/details/images/11095.html)

This equation can also be written *b = y - mx*, if you prefer to have the solved variable on the left.

**Rule 2:**

To move or cancel a quantity or variable on one side of the equation, perform the *opposite* operation with it on both sides of the equation. For example if you had *g-1=w* and wanted to isolate *g*, add 1 to both sides (*g-1+1 = w+1*). Simplify (because *(-1+1)=0*) and end up with *g = w+1*.

|  |  |
| --- | --- |
| Practice activity icon | Practice activity 3.8 |

*With your teacher’s assistance, use the techniques presented above to transpose the following formula.*

Use the techniques presented above to transpose the following formula:

Transpose this equation to make Q, P and C the subject (three separate transpositions):

Q =

P =

C =

### Substitution

Once we have identified which equations we need to work with, if the equation is in the correct form, then we can use it as it is presented. The process of replacing the letters with numbers in an equation or formula is referred to as substitution as we are literally replacing (substituting) the letters for numbers, and then solving.

|  |  |
| --- | --- |
| Practice activity icon | Practice activity 3.9 |

*With your teacher’s assistance, use the techniques presented above to substitute the following values into the formula below:*

Re-write the formula below (but don’t solve it) by substituting the values into this formula when:

Q = 55

P = 2.5

C = 75

### Solving

The process of performing the actual calculation is referred to as *solving the equation* and is the step in our calculation process that gives us the first part of our overall result. But, like everything in science, there are rules, and the rules to solving an equation must follow the *Order of Operations.*

#### Order of operations

The **Order of Operations** states the following rules which must be followed when solving equations. This is commonly referred to BODMAS (or BEDMAS).

1. Solve from left to right (where possible)
2. Solve any **B**rackets first
3. Solve **O**perators/**E**xponents
4. Solve for **D**ivision and **M**ultiplication
5. Solve for **A**ddition and **S**ubtraction

**Example**

NOTE: This example uses *logarithms*. **You do not need to know about these at the moment, we have included this example as it uses all the steps in the process**. Consider the following equation that we are solving for ‘x’:

Where:

Pi = 0.255

Po = 0.00002

Log10 = A mathematical operator

To answer this we simply substitute the given data into the formula and solve in accordance with the rules:

1. **Solve from left to right** (where possible)

This only applies when there are multiple instances of the same operator.

1. **Solve brackets first**

x = 10 + log10**(0.255 / 0.00002)**2

(= 12750)

1. **Solve operators/exponents second**

x = 10 + log10**(12750)2** (here, log10(12750)2 reads as log10 **of** 127502)

(=162562500)

1. **Solve multiplication and division**

x = 10 + **log10(162562500)** (here, Log10 of 316306225 is solved)

(=8.2)

1. **Solve addition and subtraction last, hopefully to find the correct answer!**

x = 10 + 8.2

**x = 18.5**

#### Example

Five packets of Cheese Things costs $5.50. How much did each one cost? We can calculate this using the formula:

C = NP

We know the quantities C (cost) and N (number) and want to find P (price), but P is not the subject of the formula, C = NP. The formula can be rearranged so that P is the subject. The rearranged formula can then be used to calculate P. In this example the transposed formula becomes:

P = C/N.

P = $5.50/5 = $1.10

The steps for doing this are given below.

**Step 1**

* Identify what steps happen to the quantity you want to be the subject in the formula to give the current subject.
* If we knew the value of P, what happens to it to find C?
* P is multiplied by N to give C.

**Step 2**

* To make this quantity the subject, start with what was the subject and reverse the steps. In mathematics when you reverse steps:
* x becomes ÷
* ÷ becomes x
* + becomes –
* - becomes +
* square becomes square root and square root becomes square.
* C was the subject.

The reverse of multiplying by N is to divide by N (as in C/N)

This will give you P.

P = C/N

**Example 2:**

Make **t** the subject of the formula

**Step 1**

What steps happen to t to calculate C? Starting with t

1. First 100 is added t+ 100
2. next this is multiplied by N and P NP(t+ 100)
3. next this is divided by 100
4. finally *D* is added

This successfully gives us C.

**Step 2**

The reverse of these steps is:

1. Starting with C
2. first subtract D C - D
3. next multiply by 100 (C - D ) 100
4. next divide by N and P
5. Finally subtract 100

This successfully gives us *t*.

## Expressing results

Once all the work has been done in terms of measurements and calculations, you find yourself in a position to report your result, but what do you report, and how do you do it correctly?

Imagine you have been busy in the lab performing some sort of analysis and you now have a result for Biological Oxygen Demand (BOD) to report.

**BOD = 360.4 mg/L**

Great work…but this is technically speaking a lie, as what you are stating appears to be stated as a fact and that is not true – *because all measurements are only estimates and need the error associated with that measurement stated clearly*.

When you express the error associated with a measurement, it must;

* be an expression of absolute or relative error (i.e. value or percentage)
* appear after the result and be used with a ‘plus or minus’ sign (±)
* be reported to *one significant figure* when the most significant digit in the measurement is 5 or more; otherwise to a *maximum of two significant figures*
* The measurement is quoted to the same number of decimal places as the uncertainty.

**Result = 360.40 mg/L ± 0.05 mg/L**

**Example:**

The absorbance of a milk sample was found to be 0.564 ± 0.023. Round this to the appropriate number of significant figures. The error should be quoted to one significant figure. This becomes 0.02. We then have 0.564 ± 0.02.

The measurement should be quoted to the same number of decimal places as the uncertainty. We then have:

**Abs = 0.56 ± 0.02**

|  |  |
| --- | --- |
| Practice activity icon | Practice activity 3.10 |

Write the following measurements and their associated uncertainty to the appropriate number of significant figures.

1. 123 ± 12 kPa
2. 14 ± 2.3 ms-1
3. 4.765 ± 0.36 g/mL

## Calculations with measurements

You learnt from the Introduction that there are many different types of laboratories in Australia. Each industry has different analytical requirements, and as such we need to provide a wide range of calculation types to support each of the specific types of calculations you may encounter in your course.

### Percentage calculations

You would be familiar with some percentages, especially with pie charts. All of the pie chart is equal to 100%, half of a pie is 50% and so on. The key point is that percentage calculations make quantities relative to 100, so it is *normalising* to a standard value (100).

|  |  |  |
| --- | --- | --- |
|  |  |  |
| Percentage | ÷ 100  x 100 | Fraction  Decimal |

These types of calculations are used in all industries, yet you do not need to look far to find examples in a laboratory or in real life (like a beer whose label says 4.7% v/v). If we view a percentage calculation as having four terms:

Then, excluding the 100, there are three terms that could be found, which leads to three transpositions to work with:

|  | % Question | Example |
| --- | --- | --- |
| 1 | Find A% of B | 20% of 250g  = 50 g |
| 2 | What % of X is Y? | What % is 35g of 140g?  = 25% |
| 3 | C% is an amount of D. What is the whole amount? | 18% is an amount of 45g. What is the whole amount?  = 250g |

|  |  |
| --- | --- |
| Practice activity icon | Practice activity 3.11 |

1. Convert to the simplest fraction:
   1. 85 %
   2. 8⅓ %
   3. 1.2 %

Convert to decimals:

1. 24 %
2. 0.5 %
3. 112⅔ %
4. Convert to percentages:
   1. 2/5
   2. 0.08
   3. 3.9
5. 35% of an amount is 2.8 kg. What is the whole amount?
6. A mixture contains 84% of a certain compound. How many grams of the mixture are required to obtain 5.6 g of the compound?

### Geometric calculations

These types of calculations involve the base unit of length, except where named units such as the Litre (L) are used. The Litre is a special unit of decimetre cubed (dm3). Other example calculations include:

* areas (m2)
* volumes (mL, L, m3)

#### Volume calculations and conversions

|  |  |
| --- | --- |
|  |  |
| 1000mm  1000mm  1000mm | 1 m3 = 1 000 mm x 1 000 mm x 1000 mmm  = 1 000 000 000 mm3  = 109 mm3  How many cubic millimetres in one cubic metre? |

X 1000

Number of cubic metres

\

Number of cubic millimetres

**Rule: for a cubic measures, cube the usual conversion.**

For example:

How many micrometres (µm3) are in 3.17 cubic kilometres (km3)?

|  |  |
| --- | --- |
|  |  |
| Mega (M)  kilo (k)  Multiply by 109  unit  milli (m)  micro (µ) | To convert km to µm, multiply by 109  To convert km3 to µm3, multiply by (109)3 = 1027  3.17 km3 = 1.35 x 1027 |

The trick here is not being able to use the simple formula to calculate these quantities, but more in the ability to convert between different units (especially with volumes).

Use these formulas to calculate:

**Length** = Just a measurement

**Area** = Length squared, *l* x *l* or *l*2

**Volume** = Length cubed, *l* x *l* x *l*, or *l3*

So, now we have the unit problem to solve. With reference to the figure above, and Appendix A, we find that both area and volume have a wide range of unit expressions, so being able to convert between these units becomes a highly important skill. Examples of area units include m2, ft2, acre and hectare. Examples of volume units include m3, ft3, litre (dm3), and mL.

**Example**

Let’s say we wanted to convert 3657 litres to m3. There are 1000 L in 1 m3, so:

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| 3657 L | = | 3657 ~~L~~ | 1 m3 |  |  |
|  | 1000 ~~L~~ |  |  |

= 3657 / 1000

= 3.657 m3

|  |  |
| --- | --- |
| Practice activity icon | Practice activity 3.12 |

*With your teacher’s assistance, provide answers to the following questions.*

1. Give the volume of a prism 0.1mm x 250 µm x 300 µm in mm.

There is an image of a rectangular prism showing the following dimensions.
Height: 0.1 millimetres
Width: 250 micrometres
Length: 300 micrometres

1. Perform the following conversions:
   1. 2.3 m3 to mm3
   2. 1.6 µm3 to m3

### Example calculations - pathology

The following calculations will typically be part of workplace procedures or automatically calculated within the LIMS. You do not need to remember these equations, but they are useful for practicing the transposition and substitution techniques discussed in this topic.

Mean cell (or corpuscular) volume (MCV) is used to measure the average volume of red blood cells.

*MCV (fL) =*

fL = femtolitres = 10-15L

Mean cell haemoglobin (MCH) measures the average weight of haemoglobin in the RBC.

pg = picograms

Mean cell haemoglobin concentration (MCHC) measures the average concentration of haemoglobin (Hb) in the RBC volume.

g/dl = grams per decilitre

Serial dilutions are important in both chemistry and in pathology, particularly in serology and microbiology. It is important to understand the decrease in concentration caused by each step in the dilution process.

The International Normalised Ration (INR) is used to monitor coagulation.

Note, as a ration the INR has no units.

### Example calculations - chemistry

Usually the workplace procedure for a test or preparation notes for making a solution will tell you exactly how much solute to weigh out to get the right concentration and volume of solution needed for the task. However, in some situations you will need to work out the required mass of a solute on your own to obtain a solution with a particular concentration. This is not too hard if you follow a few basic rules.

There are two measurements that you will have to make:

* amount of solute
* volume of solution so you can add the right amount of solvent

The amount of solute to be measured depends on the volume of solution that you want to make and the concentration that you require.

Concentrations can be specified in several ways.

Percent volume – volume (% v/v) involves measuring the volume of a liquid solute and adding sufficient solvent to reach a particular final volume for the solution. It is calculated by using the following:

Where the same volume unit (either mL or L) is required for the solute and the solution. When you know the volume of solution and %v/v concentration required, the volume of solute to measure is given by:

The calculation for making 200 mL of 10% v/v alcohol solution is:

Therefore, to get 200 mL of 10% v/v alcohol solution, you must measure 20 mL of alcohol and make up the volume of the solution to 200 mL with pure water.

Percent weight – volume (% w/v) involves measuring the weight (mass) of a solid or liquid solute and adding sufficient solvent to reach a particular final volume for the solution. The concentration is calculated by using the following rule:

Where the mass unit for the solute is grams (g) and the volume unit for the solution is millilitres (mL). When you know the volume of solution and %w/v concentration required, the mass of solute to measure is given by:

The calculation for making 200 mL of 0.1% w/v peptone solution is:

So, to get 200 mL of 0.1% w/v peptone solution, you must measure 0.2 g of peptone and make up the volume of the solution to 200 mL with pure water.

Molarity concentration is given by:

Some texts use C = Concentration and the unit for Molarity Concentration is: mol/L or M. If the concentration is a molarity (M) the mass of solute is measured in grams and this is converted to moles.

Moles is calculated by:

A mole is a measure of the number of particles of chemical. Each chemical will have a different formula mass. The formula mass for a chemical is on the label of the bottle.

Volume is measured in litres (L) or millilitres (mL) but must be converted to litres, for the calculation of the concentration, by:

When you know the volume of solution and molarity concentration required, the mass of solute to measure is given by:

The calculation for making 200.0 mL of 0.100M silver nitrate solution (formula mass on label of bottle = 169.9) is:

Therefore, to get 200.0 mL of 0.100 M silver nitrate solution, you must measure and dissolve 3.40 g of silver nitrate and make up the volume of the solution to 200.0 mL with pure water.

**Density** is calculated using the density (**) formula, **=mass/volume. Consider a substance with a mass (m) of 18.78 g and volume (v) of 3.06 cm3.

d =

=

= 6.137 254 9

The result is given to 3 significant figures (the least number of sig. figures in the data)

d = 6.14 g/cm3

### Example calculations - food

Specific gravity (or relative density) of a liquid is defined as the ratio of the liquid’s density (normally at 200C) relative to the density of water at 40C and does not have any dimensions.

Specific gravity (SG) is commonly used in the food industry to calculate either:

* the unknown volume of a given weight of liquid, or
* the unknown weight of a given volume of liquid.

**Example**

If 1 L of cordial weighs 1.178 kg, and 1 L of water weighs 1.000 kg, then what is the SG of this cordial?

SG = Mass (kg) / Volume (L)

=1.178 / 1.000

=1.178 (note: SG has no units)

If 2,000L of this cordial (SG = 1.178) was stored in a tank, what would it weigh?

Rearranging the above formula:

Mass (kg) = SG x Volume (L)

= 1.178 x 2,000

= 2356 kg

What is the volume of 2000 kg of this cordial?

Rearranging the above formula:

Volume = mass (kg) / SG

= 2000 / 1.178

= 1698L

Pearson’s square: can be used to calculate the ratios of ingredients required in a formulation. The ingredients may vary in composition such as fat, protein, water or sugar content. The food manufacturer needs to deliver a product with a standard composition from the raw materials.

Pearson’s square is set out in this way:

1. Determine the **critical component** in the mix on which the calculations will be based.
2. Put the values for the critical component in the ingredients at the left hand corners of the square. (A, B in the square below.)
3. Put the desired value of the critical component in the centre of the square, E.
4. The differences are taken across the diagonals of the square and these are placed at the right hand corners of the square. (C, D)

|  |  |  |  |
| --- | --- | --- | --- |
| **A** |  | **C** | C = |B-E| |
|  | **E** |  |  |
| **B** |  | **D** | D = |A-E| |

Note that the | symbol around a calculation, such as |A-E| means absolute value. That is, whether the end result of the subtraction of E from A is positive or negative you write it as positive.

**Example 1:**

A manufacturer wants to prepare 1000kL of a fruit drink which contains 10% sugar. They need to blend fruit juice which contains 8% sugar with a sugar syrup which contains 40% sugar. How much sugar syrup and fruit juice needs to be blended to give 1000kL of a fruit drink containing 10% sugar?

Following the steps for Pearson’s square, set out above:

1. The mix will contain critical component of 10% sugar.

2. Put 8 at A for the fruit juice and put 40 at B for the sugar syrup.

3. Put 10 at E as this is the critical component.

4. Calculate C as the absolute value of B – E = 30 and calculate D as the absolute value of A–E=2

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Fruit juice | **8** |  | **C** | C = |40 - 10| =30 parts fruit juice |
|  |  | **10** |  |  |
| Sugar syrup | **40** |  | **D** | D = |8-10| = 2 parts sugar syrup |

This means that in 1000kL of fruit drink there will 30 parts of fruit juice and 2 parts sugar syrup.

One part equals:

Then 30 parts fruit juice equals:

And 2 parts of sugar syrup equals :

To make 1000kL of a drink containing 10% sugar the manufacturer needs to use:

1000kL = 62.5 kL of sugar syrup + 937.5 kL of fruit juice

**Example 2:**

In this example you are given the quantities of the ingredients rather than the quantity of the final product.

How much cream at 35% fat must be added to 1000kg of milk containing 3.0% fat in order to standardise the milk to 3.3% fat.

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Milk | **3.0** |  | **C** | C = |35-3.3| =31.7 parts milk |
|  |  | **3.3** |  |  |
| Cream | **35** |  | **D** | D = |3.0-3.3| = 0.3 parts cream |

The amount of milk used is 1000kg which is 31.7 parts of the total.

If 31.7 parts equals 1000kg then one part equals:

The amount of cream needed equals:

|  |  |
| --- | --- |
| Self-check questions icon | Self-check questions – topic 3 |

*Check your understanding of the basic topics in this section by answering the questions below (or other questions provided by your teacher).*

1. You determine that there are 240 microorganisms on a pour plate of 10:3 dilution. How many microorganisms were the original sample?
2. What is the mean cell volume if haemocrit is 37% and RBC is 4.2 x 1012 /L?
3. Calculate the concentration in g/L of a solution if 0.25 g of a substance is dissolved in a final volume of 0.4 L.
4. Find the mass of dye present in 5 mL of a solution whose concentration is 70 mg/dL.
5. There are 24 g of solute in 300 ml of solution.
   1. What is the % concentration?
   2. Is this concentration %w/w or %w/v?
6. A stock solution of dye in water has its concentration given as 300 g per litre.
   1. What mass of dye is contained in one litre of solution?
   2. Find the mass of dye in one decilitre of solution.
   3. Find the mass of dye in 5 ml of solution.
7. How much 40% v/v sulphuric acid (H2S04) would be required to make 1 000 ml of 0.5% v/v H2S04?
8. Which one of the following four solutions of a salt in water is the most concentrated?
   1. 11 g of salt dissolved in 10 mL of water
   2. 9 g of salt dissolved in 10 mL of water
   3. 7 g of salt dissolved in 5 mL of water
   4. 5 g of salt dissolved in 5 mL of water
9. A manufacturer wants to prepare 200kg of a biscuit dough which contains 24% sugar by weight. The operator already has two other standard mixes which contain 20% sugar (batch 1) and 30% sugar (batch 2). How much of each batch should be combined to make 200kg of biscuit dough containing 24% sugar by weight?
10. How much cream at 28% fat ?ust be added to 1000kg of milk containing 2.0% fat to give a fat content of 3.5%.

Topic 4

Descriptive

statistics

# Topic 4 – Descriptive statistics

#### Topic glossary

*Several new terms will be introduced in this topic. These terms are* ***bold*** *in the body text. Please take the time to review these terms with your facilitator before you begin this section.*

| Term | Definition |
| --- | --- |
| Statistics |  |
| Descriptive statistics |  |
| Inferential statistics |  |
| Population |  |
| Sample |  |
| Census |  |
| Representative |  |
| Tally chart |  |
| Bin |  |
| Frequency histogram |  |
| Distribution |  |
| Normal distribution |  |
| Standard deviation |  |
| Central tendency |  |
| Mean |  |
| Median |  |
| Mode |  |

## What is statistics?

One definition of statistics is:

“The practice or science of collecting and analysing numerical data in large quantities, especially for the purpose of inferring proportions in a whole from those in a representative sample.” (Oxford dictionaries, 2019)

**Statistics** is a collection of techniques which makes sense of numbers - numbers which are the results of observations, tests or measurements. These numbers are also known as data. Figure 4.1 shows the range of processes that are part of the statistical treatment of data.

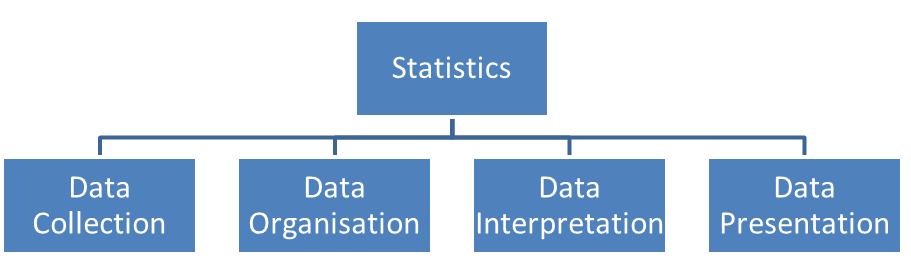


Figure 4.1 – Descriptive statistics tree.

When we have a table of raw data, it is difficult to ‘see’ what the data is actually saying, it can be like looking at static on a television screen – no discernible picture, just ‘noise’. Essentially, statistics enables the viewing of data in different ways, and allows us to either describe a data set using **descriptive statistics** techniques, or allows us to infer that something is true or false using **inferential statistics** techniques. We shall only describe the descriptive statistics here.

|  |  |
| --- | --- |
| Practice activity icon | Practice activity 4.1 |

Below are some statements that quote statistics (not real ones!). Identify whether the statistics are *descriptive* or *inferential*.

|  |  |
| --- | --- |
| Statistic | Descriptive or Inferential? |
| The average December temperature in Sydney has increased by 1°C in the last 50 years. |  |
| It is expected that the average December temperature in Sydney will increase by another 1°C in 25 years. |  |
| 25% of people surveyed at a shopping centre indicated that they were aware of increasing temperatures in Sydney. |  |
| That same survey therefore tells us that 75% of Sydney-siders are ignorant of the changing climatic conditions in their city. |  |

## Samples and populations

The distinction between descriptive and inferential statistics leads onto another pair of terms that are perhaps the most significant you will meet in this section: *population* and *sample*.

A **population** is the entire set of individual items about which observations are made, tests performed on and data recorded. This is very rare due to cost, time and inconvenience.

A **sample** is a subset or proportion of the population.

However, in most cases the measurements will be made on a proportion of the population, not on the entire set, i.e. a sample. If however a population is used to collect data and make determinations about a system then that data collection method is referred to as a **census**.

|  |  |
| --- | --- |
| Practice activity icon | Practice activity 4.2 |

Identify the sample and the population in the following scenarios.

|  |  |
| --- | --- |
| Statistic | Sample and population? |
| A bottle of water is taken from a dam to be tested. | Sample  Population |
| The frog population of a large wetland is checked by looking at two separate hectares. | Sample  Population |
| The levels of lead in dust ‘fallout’ around a smelter are assessed by testing a selection of properties. | Sample  Population |
| Environmental opinions are surveyed in a shopping centre in a study to determine the level of awareness in society. | Sample  Population |

Descriptive statistics refer to the sample from which the data was collected, while inferential statistics make the assumption (one that is not always true) that the results from the sample can be applied to the population.

The relationship between sample and population is critical in making predictions and judgements based on statistics, and as such, the sample must be **representative** of the population, and by that we mean that the characteristics of the sample are the same as those of the population. If not, the assumption above breaks down, and any decisions made are likely to be incorrect.

*Be warned!* You may encounter occasions where there is little visible distinction between “what is a sample?” and “what is a population?” To overcome this dilemma, data people usually define the *boundary* of the population so that the *sample* can be positioned against the relative scale of the population. Compare the sampling of a lake to chocolate bars on a food factory production line – how do you define a population of rapidly moving chocolates?

## The descriptive statistical process

### Step 1 - Getting the data

Before we discuss this massively important concept, consider a food factory that makes jelly beans and all that happens in that process to make all the beans the same. Let’s see how similar these beans really are using descriptive statistics!

A B

Figure 4.2 A) – Jelly beans, and B) Example of an analytical balance used to weigh jelly beans.

To create our descriptive statistics, we need to collect data on our jelly beans. Consider what *property* of these beans *could* be measured. For example, for each *individual* bean we could measure any number of aspects to obtain data such as the length, width or weight of the bean when we want to look at the *quality* of individual beans, or we could do a comparative assessment by looking at how similar beans are *different*. Either way…we’re getting data!

|  |  |
| --- | --- |
| Practice activity icon | Practice activity 4.3 |

1. With your teacher’s assistance, *complete* the following table by writing down other aspects or properties of beans that we could measure (some examples already included).
2. *Discuss* how much data we could collect from experiments such as these.

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Weight |  |  | Hardness |  |
|  |  | Width |  |  |
|  | Density |  |  | Length |

Almost every laboratory has an *analytical balance* for determining the weights of materials, so let’s work with the *weights* of the beans. The following table contains the weights (in the unit of *grams*) for 100 beans from one bag bought from a local shop.

|  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| 0.96 | 1.11 | 0.88 | 1.02 | 0.96 | 1.11 | 0.95 | 1.15 | 0.89 | 1.03 |
| 0.90 | 0.93 | 0.94 | 1.02 | 1.11 | 0.92 | 1.10 | 0.97 | 0.93 | 1.07 |
| 1.12 | 0.84 | 1.01 | 0.94 | 1.04 | 1.06 | 0.94 | 1.05 | 0.97 | 0.92 |
| 1.03 | 0.94 | 1.09 | 0.96 | 0.97 | 1.14 | 1.17 | 1.19 | 0.86 | 1.19 |
| 0.93 | 0.94 | 1.08 | 0.94 | 0.98 | 0.90 | 1.06 | 0.91 | 0.97 | 1.00 |
| 1.08 | 1.12 | 1.11 | 0.85 | 1.01 | 1.19 | 0.93 | 0.88 | 1.01 | 1.02 |
| 0.99 | 0.99 | 0.89 | 0.83 | 1.15 | 1.16 | 0.86 | 0.98 | 1.06 | 0.94 |
| 0.89 | 0.88 | 1.07 | 1.00 | 1.08 | 0.95 | 0.99 | 0.90 | 1.14 | 1.02 |

Table 4.1 - Raw data from weighing 100 jelly beans on an analytical balance (see Figure 4.2).

As you can see, simply looking at the table above tells us almost nothing about the data – it is after all just a random collection of values, but can you see any trends or patterns in the data? Probably not. So what can we *do* to this data to make it more sensible?

### Step 2 – Tally and graph the data

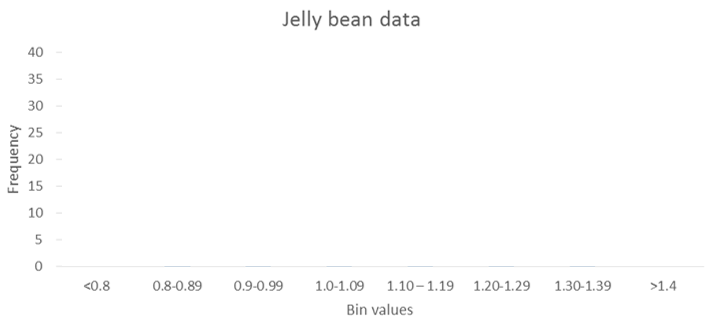
The first action is to do a **tally chart** (a simple scoring system) to see how many numbers fall within some pre-defined range (called a **bin**), and then tally (score) each bin that a value falls into. From a statistical viewpoint, we can determine the frequency of the occurrences by creating bins that contain a range of data, as in the table below, and then recording how many instances of each number fall within each bin.

|  |  |
| --- | --- |
| Practice activity icon | Practice activity 4.4 |

1. With your teacher’s assistance, *complete* the following table by tallying which values from Table 4.1 above fit into the range of values (bins) below.

|  |  |  |
| --- | --- | --- |
| **Bin** | **Tally**  **(score)** | **Frequency**  **(tally sum)** |
| **Example>>>** | ~~||||~~ ||| | 8 |
| **<0.8** |  |  |
| **0.8-0.89** |  |  |
| **0.9-0.99** |  |  |
| **1.0-1.09** |  |  |
| **1.10 – 1.19** |  |  |
| **1.20-1.29** |  |  |
| **1.30-1.39** |  |  |
| **>1.4** |  |  |

1. With your teacher’s assistance, *graph* the bin data above to see if the data *behaves* in a particular way.



Now we can start to see if the data *does* behave in a particular way. You should notice that our graph has two axes; *bins* and *frequencies*, and that there is a *pattern to the graph* which tells us a considerable amount of information.

In summary, that is the basic process on how we achieve descriptive statistics, but there is a lot more we can get from this data. If we kept collecting measurements, say a million data points to create the graph (called a **Frequency histogram**), we find that the shape of the graph is very symmetrical and shows a range of data across the bins whose frequency tends to gather around a central point.

The range of data across the *x axis* is called a *distribution*, and the frequency values on the y *axis* build up to a central point of the curve and is referred to as the data’s *central tendency*.

## What is ‘a distribution of data’?

With reference to the graph above, we see that the data in the bins is distributed over a wide range. This is referred to as the **distribution** of data, and is very important to science as it allows us to:

* Understand the general behaviour of data
* Calculate the probability of what might happen with the next measurement
* Decide whether something is being made well or badly

### The ‘Normal’ distribution

The concept of a distribution of data is incredibly important. ‘Most’ measurements reveal similar patterns, with a relatively narrow distribution and a strong central tendency, and you can easily prove this by measuring something and graphing the data. This pattern of data behaviour is so common in fact that they have given the general form a specific name; the normal (standard) distribution.

You will hear or read about whether data *follows the normal distribution*. This is because the **normal distribution** is actually a mathematical model of the general pattern. The normal distribution exhibits unique properties that makes it such a valuable tool, including:

* The mean, mode and median are equal in value and at the centre of the curve.
* The area under the curve is considered equal to 1 (that is, 100%).
* The frequencies ‘fall away’ from the centre in calculated equidistant fractions called **standard deviations,** given the symbol of **s** (sample) and **** (population).

This can clearly be seen in Figure 4.3. The mean, mode and median are all 0 (centre value of the mathematical model). The standard deviations are expressed as whole values, typically from 1-3 (although theoretically there is an infinite number of standard deviations, we only ever really use up to 6).

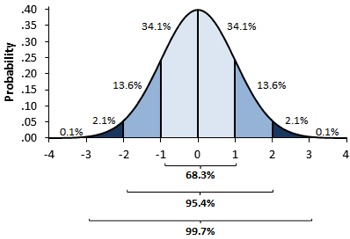


Figure 4.3 - The mathematical model of the Normal Distribution showing standard deviations broken down into ‘halves’ with the full percentages displayed underneath.

### Non-normal distributions

Given the variety of data that exists, there are many other distributions of data that occur. Consider the distributions below.

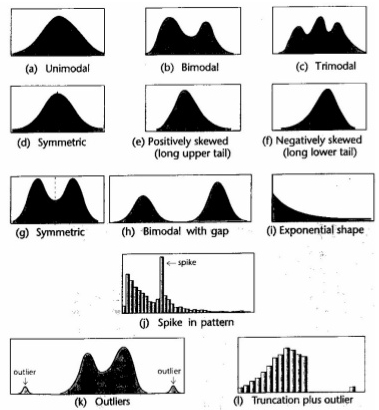


Figure 4.4 – Examples of non-normal distributions.

If there is any deviation in the pattern of the normal distribution, then we describe this as being non-normal. These can arise for many reasons, but typically associate with:

* The data just behaves that way (such as the distribution of the wavelengths of light).
* Overlapping distributions due to sampling from more than one population.

One example is missing from Figure 4.4, a rectangular shape! Consider the distribution of data from a lottery. Lotteries work by ensuring that each number being drawn has the same chance of being drawn as every other number. If this was not the case, you could predict which numbers are more likely to be drawn, and place all your money on the most frequently occurring numbers.

### Standard deviation

Standard deviation is a measure used to express the variation within a data set - similar to the size of the range, the difference between the maximum and minimum value. Earlier in this topic we referred to normal distributions where standard deviations are used to express the shape of the distribution. For any group of numbers a standard deviation can be calculated. If the standard deviation is large then it means that the data is quite variable or spread over more values than if the standard deviation is small. The calculation for standard deviation can be completed:

1. By hand using a calculator to assist
2. Using a statistics function on your calculator
3. Using a spreadsheet (such as Excel) function

Your teacher can show you how to use the functions on your calculator and on a spreadsheet.

The following formula shows how to calculate standard deviation by hand. It might look complex but what it is really doing is asking you to find out the mean/average value of the data set and then find out the difference between each one of the values in the data set and the mean. So if the values are all very close together the difference between each value and the mean will be small. If the values are very spread out then the difference between each value and the mean will be large. The formula involves squaring the differences, adding them and taking the square root. This is just so that any negative differences don’t cancel out any positive differences – we’re interested in how big the difference of each point is from the mean, not whether it is bigger or smaller.

The formula for standard deviation depends on whether the data is being considered as a population or a sample representing a larger population. If the data is being considered as a population we divide by the number of data points (N). If it is a sample we divide by one less than the number of data points (n-1).

Population standard deviation:

Where:

σ = population standard deviation

N = the number of data points in the population

refers to each value

µ = the mean value of the data set (the population)

means that for every value, do what is in the brackets and add them up.

Sample standard deviation:

Where:

= sample standard deviation

n = the number of data points in the sample

refers to each value

= the mean value of the data set (the sample)

means that for every value, do what is in the brackets and add them up.

For example:

The following data is an entire population representing the ages of students in a small class. Calculate the standard deviation.

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  |  |  |  |  |
| 18 | 34 | 23 | 56 | 25 |

**Step 1:** Calculate the population mean (µ), this is further explained in the section to follow:

µ=(18+34+23+56+25) / 5

= 31.2

**Step 2:** Subtract the mean from each data value.

|  |  |  |
| --- | --- | --- |
|  |  |  |
| 18 | - 31.2 = | -13.2 |
| 34 | - 31.2 = | 2.8 |
| 23 | - 31.2 = | -8.2 |
| 56 | - 31.2 = | 24.8 |
| 25 | - 31.2 = | -6.2 |

**Step 3:** Square the resulting amount and add up the squares.

|  |  |  |
| --- | --- | --- |
|  |  |  |
| -13.2 x | -13.2 = | 174.24 |
| 2.8 x | 2.8 = | 7.84 |
| -8.2 x | -8.2 = | 67.24 |
| 24.8 x | 24.8 = | 615.04 |
| -6.2 x | -6.2 = | 38.44 |
|  |  | 902.8 |

**Step 4:** Divide by the number of values.

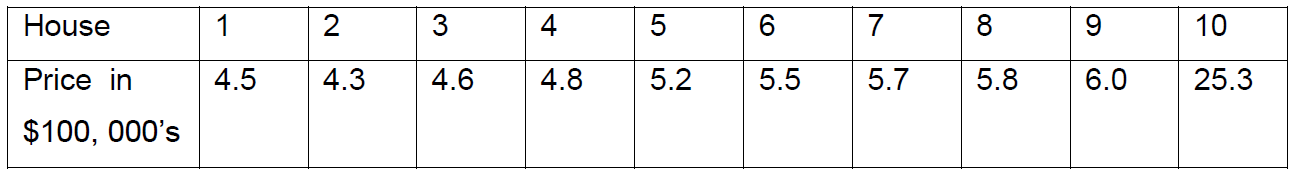
902.8 / 5 = 180.56

**Step 5:** Take the square root of the resulting value to find the population standard variation.

= 13.43

### What is ‘central tendency’ of data

A measure of **central tendency** is a single value that tries to give us information about a set of data by identifying the central position within that set of data. It is the highest point on the normal distribution. The **mean***,* given the symbols of *x̅* (sample) ** (population) is the most commonly used measure of central tendency. It is often used to give us information about a population from a sample set of observed values. The major limitation in using the mean is that its value may be distorted by outlying results. For example, if we look at the house prices in a suburb from the table below.

 We find the mean of the ten house prices is $717,000 (the sum of all values divided by the number of values) but we find that most houses are actually less than or equal to $600,000. Therefore the very high price of one house distorts the picture of house prices to seem higher.

We calculate the mean using the formula:

Where:

* The sum of values is literally the adding together of all the values you have.
* Number of values is literally how many values you are summing up

In sentence form, this states

“Divide the sum of all the numbers by the number of values summed”

The **median** is either the *middle value* or the mean of the two middle values in an even number of organised data points (sorted from smallest to largest, for example). In real data (like the jelly beans), the median will not likely be equal to the mean.

In the above example, the median would be the mean of the two middle values, the fifth and sixth values, that is, ($520,000 + $550,000)/2 = $535,000 which is closer to the majority of prices in the suburb. This ignores the skewing effect of the very expensive house (which is why house prices in the media are usually described using the median).

We calculate the median by sorting the data from highest to lowest and finding the middle value, or the average of the two middle values.

The **mode** is the value that occurs with the *greatest frequency* (the most popular value). It is often used for categorical data. The mode has applications in manufacturing. For example, it is important to manufacture more of the most popular shoe sizes because manufacturing different shoe sizes in equal numbers would cause a shortage of some shoes and an oversupply of others.

We calculate the mode by sorting the frequency from highest to lowest. The mode will be the highest value (or bin range).

|  |  |
| --- | --- |
| Self-check questions icon | Self-check questions – topic 4 |

*Check your understanding of the basic topics in this section by answering the questions below (or other questions provided by your teacher).*

1. How do a sample and a population differ?
2. Under what circumstances can we use the word census?
3. Why is a representative sample so important to obtain?
4. What is the difference between descriptive and inferential statistics?
5. Briefly describe the process of creating a frequency histogram.
6. What is the difference between the distribution of a data set and the central tendency of the data?
7. What is the normal distribution and why is it so fundamental in statistics?
8. Calculate the mean, median, mode and sample standard deviation of the following data set.

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| 5 | 40 | 177 | 40 | 11 | 29 | 88 | 49 | 117 | 64 | 53 | 29 | 57 | 69 | 3 | 30 | 9 | 66 | 58 | 46 |

**Mean:**

**Median:**

**Mode:**

**Standard deviation:**

Topic 5

Presenting data

# Topic 5 - Presenting data

#### Topic glossary

*Several new terms will be introduced in this topic. These terms are* ***bold*** *in the body text. Please take the time to review these terms with your facilitator before you begin this section.*

|  |  |
| --- | --- |
| Term | Definition |
| Axes |  |
| X Axis |  |
| Y Axis |  |
| Line graph |  |
| Scatter plot |  |
| Gridlines |  |
| Legend |  |

## Tabulating data

What is a table? For our purposes, it is simply a box of rows and columns. Traditionally, rows go ‘left-to-right’ and columns go ‘up-and-down’. This is true for most tables, except the Periodic Table which uses the term ‘Periods’ for rows and ‘Groups’ for columns. The process of creating a table is referred to as *tabulating* data, and when we display data in a table, we describe the view as being *tabular*.

Tabulating data is done for any number of reasons, one of the most common is to simply sort the data. If you have lots of data from different sources, then you might collect key data from each source and place it in a table. Or, you can use the table as a tool to actually do work on the data using a functional table such as the frequency tally example mentioned below.

## Visualising data

The problem with data (especially when in a table) is that it is not very easy to see what the data is trying to say. To overcome this problem, we frequently use visualization tools to display the data in the form of pictures (or info-graphics), the most common of which are the graphical representations called a plot, graph or chart of the data. Graphs are pictorial versions of numerical data. They are intended to tell any ‘story’ the data wants to say in a visual way.

### What is a *plot* or a *graph*, or even a *chart*?

The literature is vague on this topic and we often use these terms interchangeably (and incorrectly) without too much trouble, but one *way* to differentiate between the three terms is to assume that one is for *points*, one is for *lines* and one is for *shapes or symbols*; such that:

* A *plot* is created when points are placed on a coordinate system (x-y-z)
* A *graph* is created when the data points display a line from an event or formula
* A *chart* is created when shapes are used to display data

These definitions are open to debate and there are always exceptions to the rule, and to that end, ask yourself where a photograph would fit into this definition?

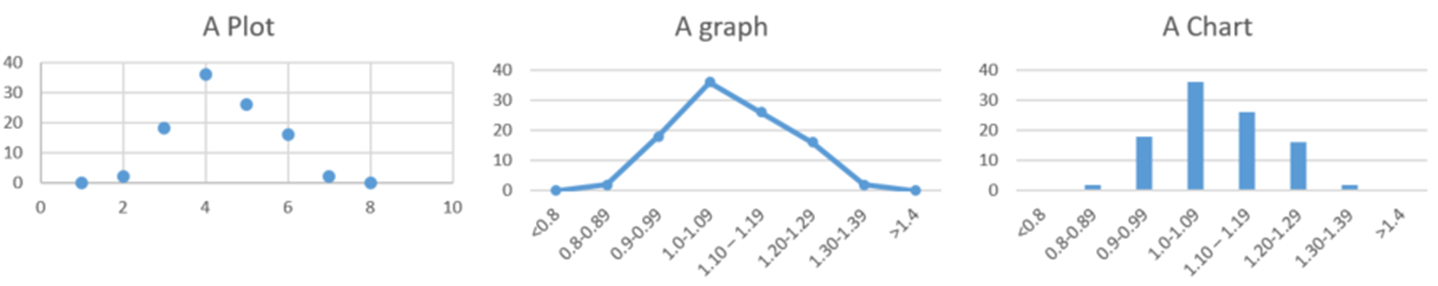


Figure 5.1 – Examples of plots, graphs and charts.

### General graphing principles

You should apply the following general rules when constructing a graph, manually or by computer:

* **axes** labelled clearly (**X axis** and **Y axis**)
* informative title
* axis scale is the same all the way along – if 0-10 covers 1 cm, then so should 90-100 etc.
* the axes scales must be shown
* where multiple series are plotted on the same graph, make it clear which data belongs to which line/column by the use of legends
* don’t over-complicate the graph by the use of too many colours/**gridlines**/3-D (see below)
* the measured variable should be plotted on the vertical axis

Excel’s Chart function has so many options, it becomes a mixed blessing. Some of the appearance options tend to swamp the graph with unnecessary “bells and whistles”. Do not accept the default settings without thinking “can I make this look better?”

#### Using 2D vs 3D columns

There is no doubt that “3-dimensional” columns look better than plain rectangles. However, they can make it difficult to judge where the top of the column is, relative to the vertical axis. Gridlines (see below) can help in this regard.

However, 3D pie charts are to be avoided like the plague. This will be explained later in this chapter.

#### Gridlines

As noted above, they help with 3D columns, but are less needed with 2D columns, and are totally distracting with line-based graphs.

#### Legend

A **legend** is crucial where there are more than one data series, but totally useless when there is only one item (a calibration graph for example).

When the legend is necessary, it is made useless if the labels are simply Series 1, Series 2 etc. Make sure you fix this using the Series/Name in the Chart wizard.

#### Background

Excel defaults to a graph background of mid-grey. On the printed page, this becomes very distracting, and should be changed to no background (or white).

However, on a PowerPoint slide a background is probably going to be needed if the background of the slide is relatively dark. White may be too much contrast (the difference between two colours), so it will be a matter of trial and error.

#### Colours & patterns

If you have to change colours of features in a chart, make sure they don’t end up similar brightness, because black & white printing will mask the difference. Also, avoid patterns – they are too hard on the eyes.

#### The general parts of a graph

Consider the figure below which identifies the major parts of a graph to help you understand the detailed descriptions that follow. It certainly doesn’t show every part of every graph, but the basics are there. Note, in Excel the area where the chart sits is called the chart area, and the area on that where the data is displayed is termed the plot area.

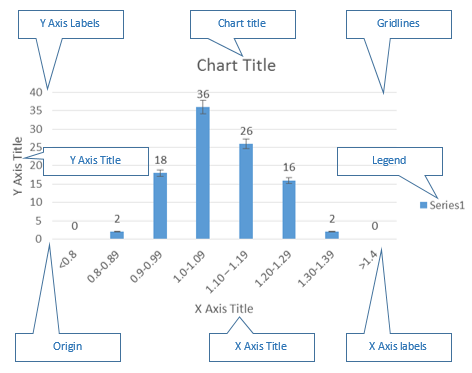


Figure 5.2 – The major parts of a graph (or chart, or plot).

Further to this, there are accepted terms that come with most of these chart features, so consider this chart below (which mainly applies to scatter graphs, which will become particularly important in later units of study).

|  |  |
| --- | --- |
| Practice activity icon | Practice activity 5.1 |

Enter the term *misleading graphs* into [Wikipedia](http://www.wikipedia.org). Choose an appropriate graph from the selection displayed throughout the page to find a graph of particular interest to you (or use any bad graph).

|  |  |
| --- | --- |
| Does the graph have…? | Yes or no? |
| Title |  |
| The units of the quantity of the x-axis |  |
| Quantity represented on the y-axis |  |
| The scale used on the y-axis |  |
| Quantity represented on the x-axis |  |
| The scale used on the x-axis |  |
| The units of the quantity of the y-axis |  |
| Describe the shape of the plot |  |
| Any other problems with the graph? Describe in the space below. | |

### Column and bar charts

This includes bar, cylinder, cone and pyramid charts offered by Excel. Column type graphs plot frequency (or relative frequency) on the vertical (with the exception of bar graphs, where it is the horizontal) axis, and the category value on the horizontal axis.

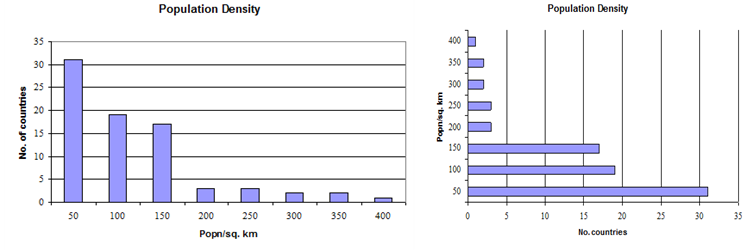
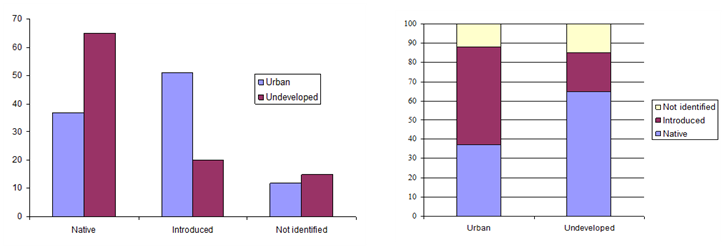


Figure 5.3 – Example of a Column chart (left) and a Bar chart (right).

Where you have multiple data sets with the same categories (such as Table 1.3), they work better when plotted together rather than separately, just as the combined table does. There are two ways of approaching this, depending on the data:

* multiple column/bars per category (Figure 5.4 left side)
* stacked columns (Figure 5.4 right side) - only appropriate where the data makes sense when it is added together

Figure 5.4 – Examples of column graphs with data displayed in different ways.

### Line graphs

It is easy to mistake line and scatter graphs in Excel, and imagine they are the same thing. In reality, line charts are best used for data that can be assumed to be continuous, for example weather data, as shown in Figure 5.5. A scatter graph might also have a line which shows the relationship between the data but the data points might fall or not fall on that line. In a line graph each data point is joined by the line and shows that it is reasonable to expect that the line represents the data values that lie between actual, measured data points.

The only point of using a **line graph** is to put a line in, so don’t leave it as dots. Since it has to be a “join-the-dots” line, it isn’t always necessary to plot both the line and the data points. It is sometimes necessary to check that the line is thick enough (in Excel) to be easily seen and this might mean you need to change the default.

Figure 5.5 – Example of a line graph.

### XY scatter graphs

A line is actually an equation: *y = mx + b*. When you want to draw a graph with a line in it (as opposed to a curve or a ‘line’ through raw data), this should your first port of call. To make a sensible line-based graph, both axes need a number associated with them. In other words, two measurements should have been made about that particular item, for example *light absorbed versus concentration*. If there is only one measurement, it is not possible.

The **scatter** **plot** shows the correlation between the two variables, that is, whether one changes in a consistent way when the other changes. This doesn’t have to be a perfect straight line, like in calibration graphs.

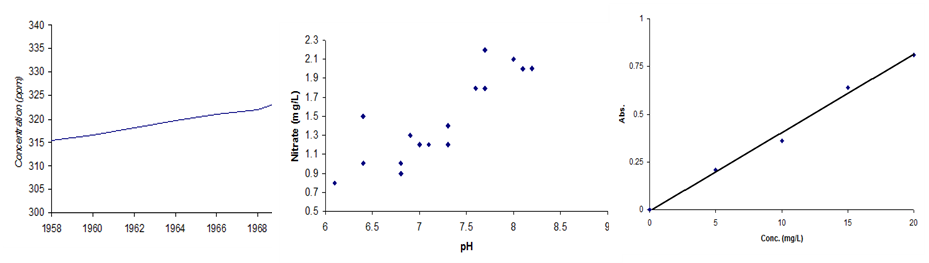


Figure 5.6 – Three different examples of scatter plots.

### Other types of charts

#### The Statistical Process Control (SPC) chart

You will recall from the descriptive statistics section of this manual that a data set has distinctive, definable properties associated with the distribution of data, such as a mean (average) and standard deviation. Outside of the descriptive statistical properties that we collect from data, we can display that statistical data in a unique way called Control Charts.

These charts are very data intensive and it is a complex calculation process to get the correct data for them, in fact there are several national and international Standards (such as AS/NZS ISO 9001:2016 and others) that apply just to these types of charts. In short though, instead of graphing the data as a plot of frequency versus occurrence and viewing the distribution (mean and standard deviation), we plot the raw data versus time (or event or similar) as well as statistical data such as means or standard deviations (but this time they display as lines from left to right). See the figures below.

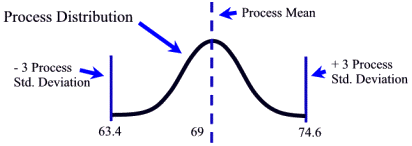


Figure 5.7 – The normal distribution viewed as a line chart with the mean (dotted line) and Standard Deviations (solid vertical lines).

When we apply this to a control chart, the mean is used as a ‘central measure’ for the raw data to be compared to and the Standard Deviations (1, 2 or 3) are displayed on the chart to show various critical limits for the process to follow. The resulting chart looks like the figure below.

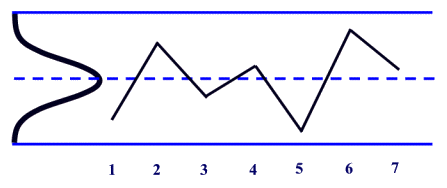
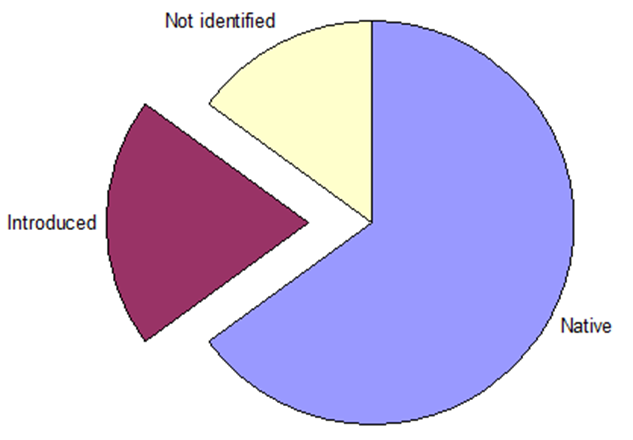


Figure 5.8 – The basic concept of a Statistical Process Control (SPC) chart. Mean and SD are as Figure 5.7 above. The zig-zag pattern is the raw data from the process being monitored. The x-axis in this case is event based (not time).

#### Pie & doughnut graphs

Pie charts can also be used for category data, where the categories are represented by segments of a circle. The size of the segments is proportional to relative frequency (or proportion) of each category. A pie chart, therefore, displays information in a similar way to each of the stacked columns in Figure 5.4.

Only data collected from the same population should be grouped into a pie chart, for example weights of the different types of recycled materials collected. If related measurements come from different populations – for example masses of paper in recycling from different suburbs – then a pie chart is incorrect.

Figure 5.9 – The old pie chart (this one exploded).

### The art of being ‘honest’!

Consider this fact – you can use these plots, graphs and charts to tell lies! It is true, and it happens all the time, all over the world in newspapers and online. Sometimes it is done by accident, and sometimes it is on purpose, and it is easy to do. This section will teach you:

* How to spot a bad graph
* How to read a bad graph properly
* How to ensure ‘honesty’ when you use visual methods

The basic intention of a graph is to pictorially display data in a sensible and meaningful way. However, it is true to say that there are many graphs in the public domain that are constructed so that they fail to show the true meaning of the data. There are two basic reasons why this happens; poor design or to be intentionally deceiving.

#### Poor design

Apart from the various aspects described in the previous section, poor graph design can come about through:

* plotting the wrong data
* carelessness
* over-complication
* duplication of information

#### Intentionally deceiving

An English prime minister of the late nineteenth century said that there were three kinds of lies: *lies*, *damned lies* and *statistics*. Graphs are powerful story-tellers, particularly for people in a hurry. It is quite possible for the truth to be hidden in the detail of a graph, and for the picture to distort this truth. There are various ways that this can be done:

* hiding elements of the graph
* distorting elements of the graph
* overemphasising one element of the graph
* making unfair comparisons

## Semi-quantitative observations

So far we have focussed on measurements that are numerically based. In the next topic you will explore quantitative and qualitative ways of reporting information, that is, numerical as well as descriptive.

Sometimes we present data in a way that is semi-quantitative, using symbols or number to describe qualitative observations.

For example: you might create a scale of 1 to 4 with 1 being the lightest and 4 being the heaviest, to describe the amount of dots on each of the boxes below.

|  |  |  |  |
| --- | --- | --- | --- |
|  |  |  |  |
| Box A | Box B | Box C | Box D |

In this example we could say that in terms of lightest (1) to darkest (4) we have:

Box A = 2

Box B = 4

Box C = 1

Box D = 3

For example, soil scientists use a semi-quantitative method for determining the field assessment of stone content in soil, where the assessor can use the chart below to estimate the numerical value of the percentage of stone in the soil.

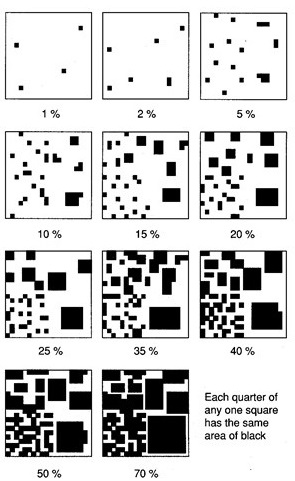


Figure 5.10 – Field assessment of stone content.

## Reporting data

The reporting of data is the final objective in our flow of data. The most common type of reporting is simply that of results reporting. This can be performed physically, or electronically depending upon the client requirements.

There are many ways that data can be reported. Consider the workflow from the Introductory section and the types of data recorded. This data can be reported:

* In simple tables
* In a spreadsheet of calculations
* With plots, graphs or charts (or other visual methods)
* As a presentation
* As a full report

### 

### Short forms

The most common form of reporting data is to simply fill in standard workplace forms, such as test results sheets or any similar document, such as the example below.

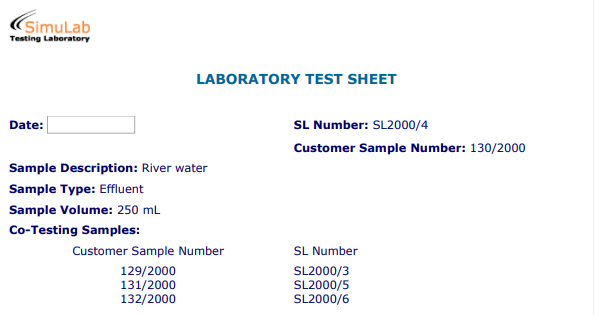


Figure 5.11 – Example of short style ‘reports’ that lab technicians use on a routine basis.

### LIMS

The Laboratory Information Management System, or LIMS, is a computer based system for connecting instruments with spreadsheets and databases for the purposes of ensuring quality and ultimately to manage the data from the analytical services the lab provides.

In a large lab, most technicians will use the LIMS for almost everything including sourcing methods for tests, entering results and generating reports for either internal use within the lab, or even external reports for clients. As this is a computer based application, it is difficult to ’see’ what a LIMS does, so the figure below will hopefully show you how they work.

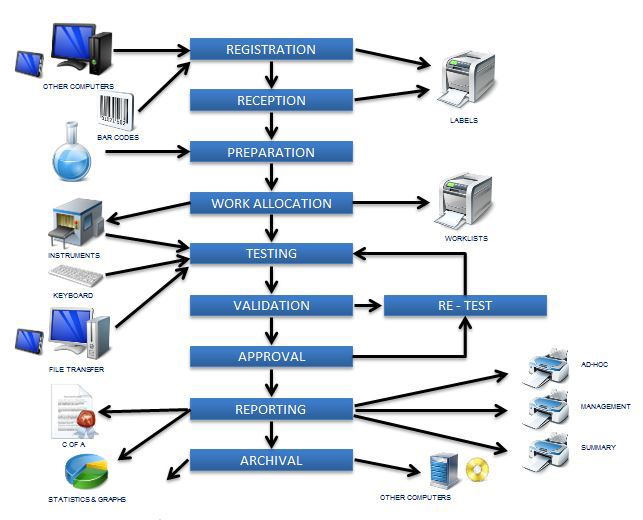


Figure 5.12 – The role that LIMS can play in a laboratory. Data storage, treatment and reporting can be all done by this system.

### Issuing Analytical Reports or Certificates of Analysis

For many laboratories, their business depends on the quality of the service they provide. To this end, labs who have been deemed competent by NATA can issue Certificates of Analysis which is a document that certifies the quality of the analysis. An example of a Certificate of Analysis is provided to follow.



**Simulab Testing Laboratory**

**Unit 8/345 Jones Crescent, Smithville**

**Tel: (607) 354 9604 Fax: (607) 354 9554**

**Email: info@simulab.com.au**

**CERTIFICATE OF ANALYSIS**

|  |  |
| --- | --- |
| SIMULAB REF | KLNR32185 |
| **REPORT DATE** | **2015-01-19** |
| **COMPANY** | **ABC Sweets and Confectionary**  **696 Smith Street,**  **Jonesville** |
| **CONTACT PERSON** | **Anne Jones**  **Tel: 0421 222 111**  **Email: anne@abcsweets.com.au** |

**:**

**:**

**:**

**:**

**DATE SAMPLE RECEIVED 2015-10-29**

**SAMPLE MARKING CRISPY SWEET CRISPS**

**SAMPLE DESCRIPTION ONE (1) SAMPLE**

**SIMULAB SAMPLE ID KLNR32185**

**ANALYSIS RESULTS**

**(As per sample)**

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| No. | Test Parameters | Result | Units | Method Reference |
| **1.** | **ARSENIC (As)** | **ND ( < 0.5 )** | **mg/kg** | **QW1-QF/17-42 based on AOAC 999.11, 986.15 and APHA3120. APHA 3112 and APHA 3125** |
| **2.** | **CADMIUM (Cd)** | **ND ( < 0.1 )** | **mg/kg** | **QW1-QF/17-41 based on AOAC 999.11, 986.15 and APHA3120. APHA 3112 and APHA 3125** |
| **3.** | **IRON (Fe)** | **0.2** | **mg/100g** | **QW1-QF/17-41 based on AOAC 999.11, 986.15 and APHA3120. APHA 3112 and APHA 3125** |
| **4.** | **MERCURY (Hg)** | **ND ( <0.05 )** | **mg/kg** | **QW1-QF/17-40 based on AOAC 999.11, 986.15 and APHA3120. APHA 3112 and APHA 3125** |

**(ND – Not Detected)**

**\*Test is performed at Accredited Laboratory Simulab.**

**\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_**

**Sean Ying**

**Lab Manager – Food Division**

**BSc (Hons), PHD (Biotechnology)**

**MJMM No: AD11 IKM No: L/2032/7080/85**

Figure 6.13 – Example of a Certificate of Analysis

### Record keeping and confidentiality

Strict laws govern privacy in Australia and organisations need to have clear policies and procedures in place to protect private information. Government, professional associations, employers and funding bodies require accurate and appropriate records to be kept of any data generated by the lab. Record keeping is an important aspect of being professional and accountable for the services that practitioners provide to clients. Accurate and up-to-date records support quality service delivery and assist in making accurate interpretations when reporting risks to clients or to other parties, and to ensure clients can receive continuity of service.

|  |  |
| --- | --- |
| Self-check questions icon | Self-check questions – topic 5 |

*Check your understanding of the basic topics in this section by answering the questions below (or other questions provided by your teacher).*

1. What is the benefit to visualising data?
2. What is the difference between a plot and a chart?
3. Identify and list three general graphing principles
4. Why should a graph be labelled and formatted correctly?
5. What is the key difference between a line graph and an xy scatter graph?
6. How is a statistical process control chart (SPC) related to the concepts of distribution and central tendency of data?
7. How do we identify a bad graph?
8. Why is confidentiality such an important part of data treatment?
9. Work with your teacher to draw:
   * a line graph
   * a calibration graph
   * a column chart
   * a histogram

Using data provided by your teacher. Draw each graph by hand and also using the graphing function on a spreadsheet.

Topic 6

Interpreting data

# Topic 6 - Interpreting data

#### Topic glossary

*Several new terms will be introduced in this topic. These terms are* ***bold*** *in the body text. Please take the time to review these terms with your facilitator before you begin this section.*

|  |  |
| --- | --- |
| Term | Definition |
| Interpretation |  |
| Category variable |  |
| Ordinal variable |  |
| Nominal variable |  |
| Numerical variable |  |
| Discrete variables |  |
| Continuous data |  |
| Truncate |  |
| Qualitative data |  |
| Quantitative data |  |
| Trend |  |
| Pattern |  |
| Noise |  |
| Correlation |  |

## What is interpretation?

The term **interpretation** means the process of attaching meaning to data. This is because numbers (or raw data) do not speak for themselves.

When we wish to interpret data, we need to position that data with reference to the following questions:

* Why was the analysis done?
* What type of data was collected?
* Is the assessment to be qualitative or quantitative?
* Is the interpretation to be based on values to compare to a reference?
* Is the interpretation to be based off graphical presentations?

This topic will answer these questions and train you in how best to interpret the data.

### What type of data was collected?

It is almost impossible to truly interpret what data means if you do not understand the original purpose for which the data was collected. The type of questions we are answering here include:

* What is being analysed?
* Why is it being analysed?
* Was the test accurate and precise enough?

Before we dive into the interpretation of data, let’s have a quick look at the types of data that can be collected. Consider the following figure which illustrates the major categories of data.

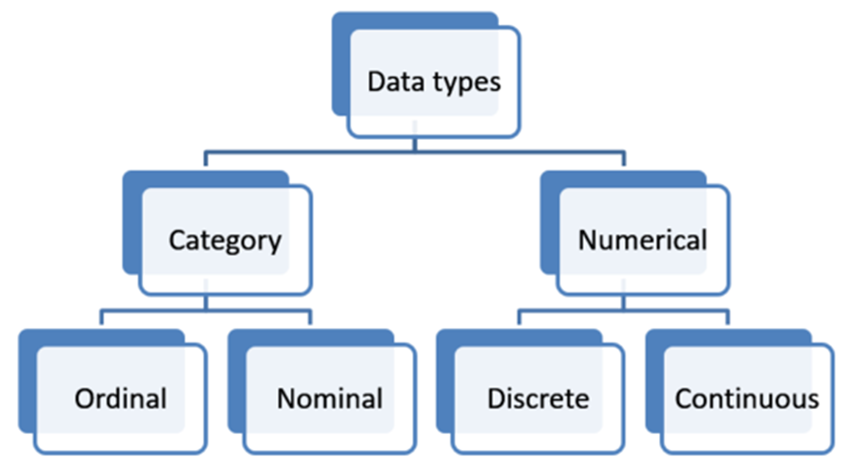


Figure 6.1 – The ‘data tree’.

So what does each of these terms mean?

**Category** **variables** are where the results of measurement are not numerical and will be a particular class, ranking, option or type of whatever is being measured.

**Ordinal** **variable** involves an ordering of the possible classes – for example, first/second/third/etc. or agree strongly/agree slightly/disagree slightly/disagree strongly. Again, the end result will be a number of each response type.

**Nominal** **variables** usually involve the data having a name, for example, yes/no, male/female, car/truck/bicycle/motorcycle/4WD etc.

When taking a sample of a population using a category variable, the end result will be a number of occurrences of each category.

**Numerical variables** involve the measurement producing a number instead of a categorical response. This is the most common type of data found in a laboratory environment (but categorical data does exist).

**Discrete** **variables** are numbers that are whole. Classic examples include statements such as “you can’t have half a cat”. For clarity…if you did have halves of cats, then your analysis would not be on whole cats, rather parts of cats, which are still discrete measures.

**Continuous** **data** is like the value for pi…it just goes on for infinity. For example, the number of children in a family is a discrete variable, because there can’t be any in-between values, but their age is a continuous one. One problem students may have with lab data is that we **truncate** it (that is, we limit the decimal places), and by doing so, sometimes ‘hide’ the continuous nature of values.

When we know the type of data we are working with, we can categorise the type of analysis as either qualitative or quantitative to further aid our interpretation.

### Is the analysis qualitative or quantitative?

#### Qualitative

**Qualitative** **data** is a data concerned with descriptions, which can be observed but cannot be computed easily. It is typically associated with categorical data such as nominal ‘yes/no’ or ordinal ‘1st, 2nd, 3rd…’ data sets.

#### Quantitative

**Quantitative** **data** is the one that focuses on numbers and mathematical calculations and can be calculated and computed. The main purpose of quantitative analysis is to quantify (accurately measure) the data and assess it statistically if required. It is typically associated with numerical data, but continuous data more so than discrete data (which can sometimes be analysed qualitatively as well).

## Trends, patterns & noise

### Trends

The term **trend** refers to whether the data exhibits a general direction which is developing or changing. Consider the figure below where we see the smooth blue line exhibiting an upwards trend over time. In this case, we could easily predict that the next temperature reading will be warmer again.

### Patterns

A **pattern** is described as a repeating sequence of change in direction. With reference to the figure below, the red zig zag line (temperature) reveals a distinct pattern of temperature cycling up and down. In this instance, that pattern correlates well with changing from winter and summer and back again.

Also, the pattern itself exhibits a trend…upwards!

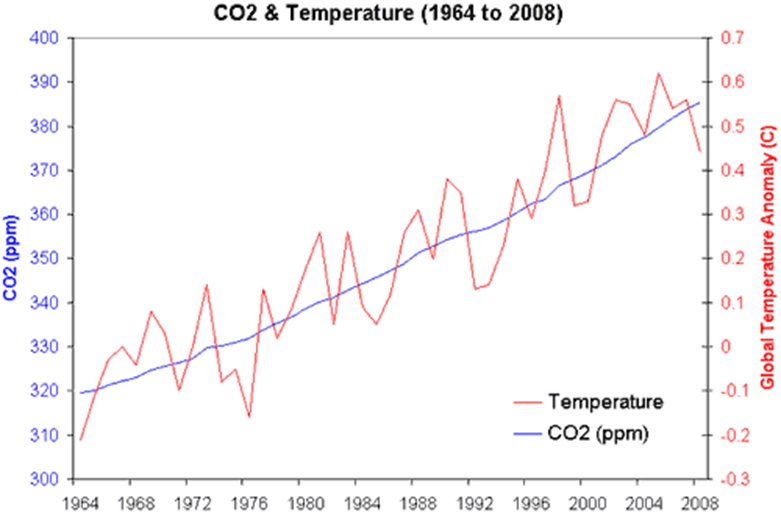


Figure 6.2 – Example of a graph with trends and patterns. The smooth blue line is a trend, and the red fluctuating line exhibits both a trend (generally upward) and a pattern of repeating ups and downs. Also, is this a good graph?

### Noise (no trend or pattern)

The term **noise** when applied to data or signals simply means that there are no obvious visible trends or patterns, and as such, there is no meaningful information to be taken from the graph.

When noise is generated, it usually means that there is literally no trend or pattern, or it could be that the interpretive technique you are using is wrong (that is, the wrong graph has been used). This is easiest to see using scatter graphs as in the figure below.



Figure 6.3 – We clearly see what is meant by *noise* when we compare *negative* and *positive* correlations with a *‘no correlation’* graph.

### Proportionality

So, see the graphs above? We need to describe how they *behave*, and we do that using the term *proportionality* (meaning ‘in proportion to’, where the ‘to’ refers to the opposite axis). If the data trend is positive (trends upwards, left to right), we describe the behaviour as being directly proportional. If the opposite occurs (trends downwards, down to right), then we describe it as inversely proportional. Positive simply means that the values are increasing.

## How to read a graph

All graphs tell a story, so we literally interpret graphs by reading them, but before we do that, we will usually find the trendline of the graph, which is a mathematical operation that ‘averages’ out the path the data is taking. Trendlines can take two common ‘shapes’:

* straight lines (just called ‘lines’)
* curved

Consider the following figures which explain the different types of curved lines that we can encounter in a scientific laboratory.

#### Parabolic graph

* The curve is symmetrical about an axis. The x and y-values are not directly proportional.
* The equation has the form of y = ax2 + bx + c

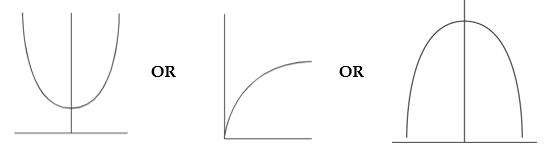


Figure 6.4 – Examples of parabolic curved lines.

#### Hyperbolic graph

* The curve never actually meets either axis (asymptotic).
* The equation has the form of ax2 + by2 + c = 0 (c > 0)

The hyperbolic graph looks like:

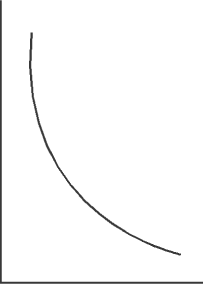


Figure 6.5 – Example of hyperbolic curved lines.

#### Interpreting linear graphs

The following graph can be used to set the required speed of the laboratory centrifuge.

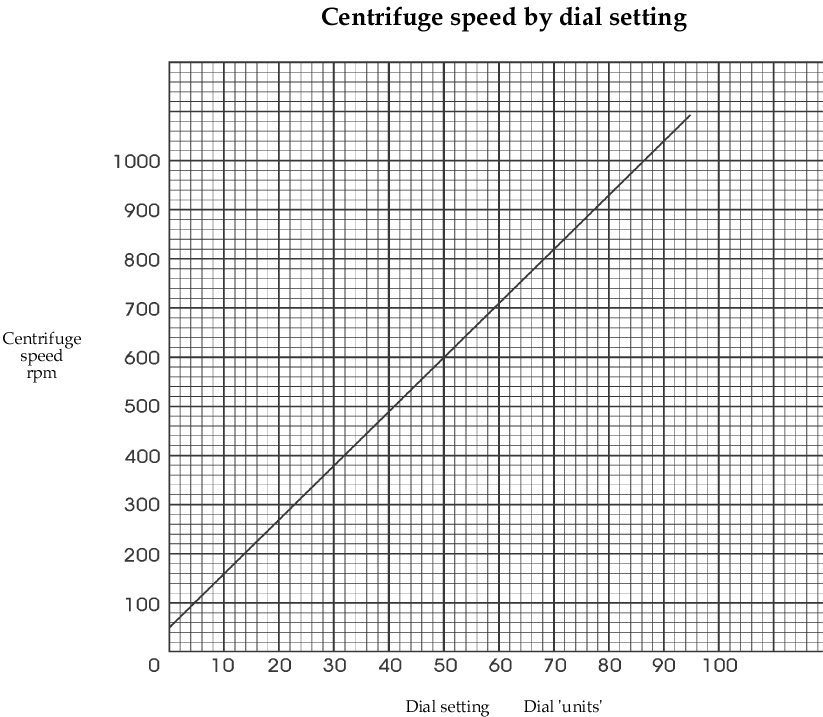


Figure 6.6 – A graph of centrifuge speed vs dial setting.

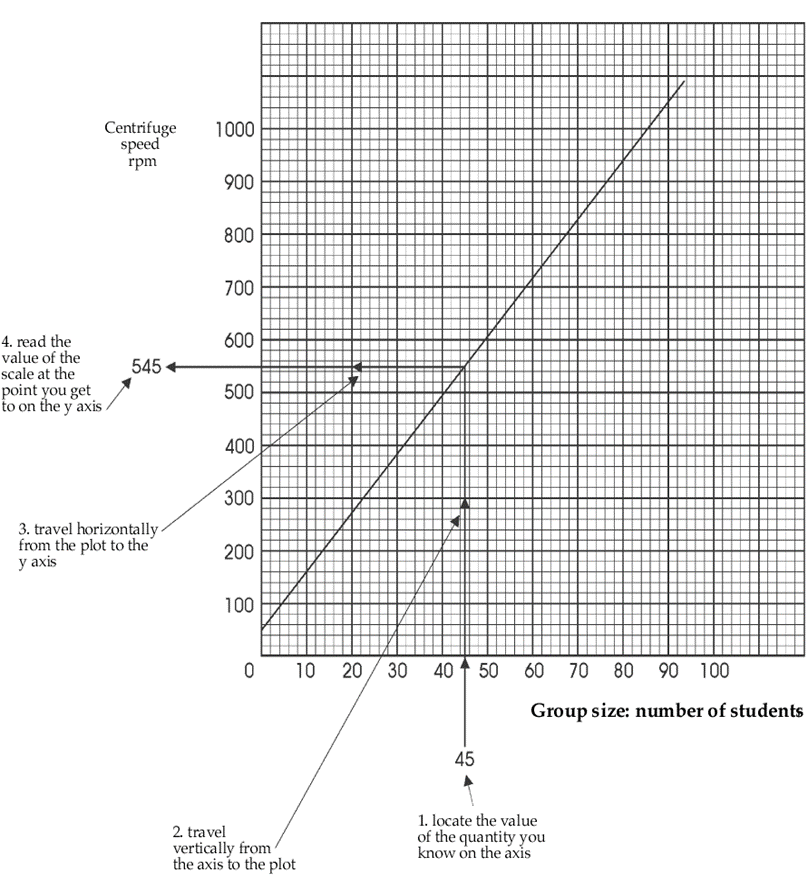
Figure 6.6 above displays a graph of data from a piece of laboratory equipment called a centrifuge. If the dial setting is 45, you can read the graph to find out how fast the centrifuge is spinning. But how would we read the graph to obtain this information? The steps in reading a graph are:

1. Check if the quantity you know is represented by the *x axis* or the *y axis.*

|  |  |
| --- | --- |
| **If it is on the x-axis then:** | **If it is on the y-axis then:** |
| Locate the value of the quantity you know on the x-axis | Locate the value of the quantity you know on the y-axis |
| Travel vertically from the axis to the plot | Travel horizontally from the axis to the plot |
| Travel horizontally from the plot to the y-axis | Travel vertically from the plot to the x-axis |
| Read the value of the scale of the point you get to on the y axis | Read the value of the scale of the point you get to on the x axis |
| This is the reading | This is the reading |

Table 6.1 – A procedure for reading a graph, chart or plot.

So, in our example we know the required dial setting for the centrifuge setting is 45. This is represented on the x-axis. Following the other steps we find:



Centrifuge speed

Figure 6.7 – Applying the instructions to the graph.

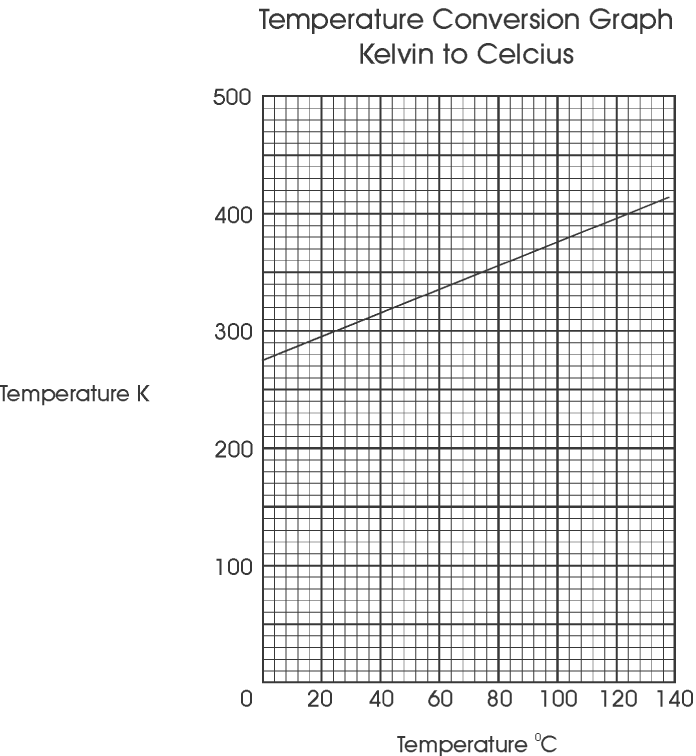
In this case, we interpret the graph using the forecast equation *y = mx + b*, so we read from the *x axis* to find the value of the *y axis*. But we could easily do it the other way around using the calibration equation, *x = (y – b) / m* and find an *x axis* value by reading from the *y axis* (in other words, the reverse direction in Figure 6.7 above.

One last comment about interpreting graphs (actually, reading them); and that is, what does an expression such as *Absorbance versus Concentration* actually mean? It literally means *y versus x*. In other words, it relates the quantity value to the name of the measure (a physical meaning) and states which axis the values belong to.

In the example of *Absorbance versus Concentration*, the first word or description (absorbance) belongs to the *y (vertical) axis* and the second word or description (concentration) belongs to the *x (horizontal) axis*.

|  |  |
| --- | --- |
| Practice activity icon | Practice activity 6.1 |

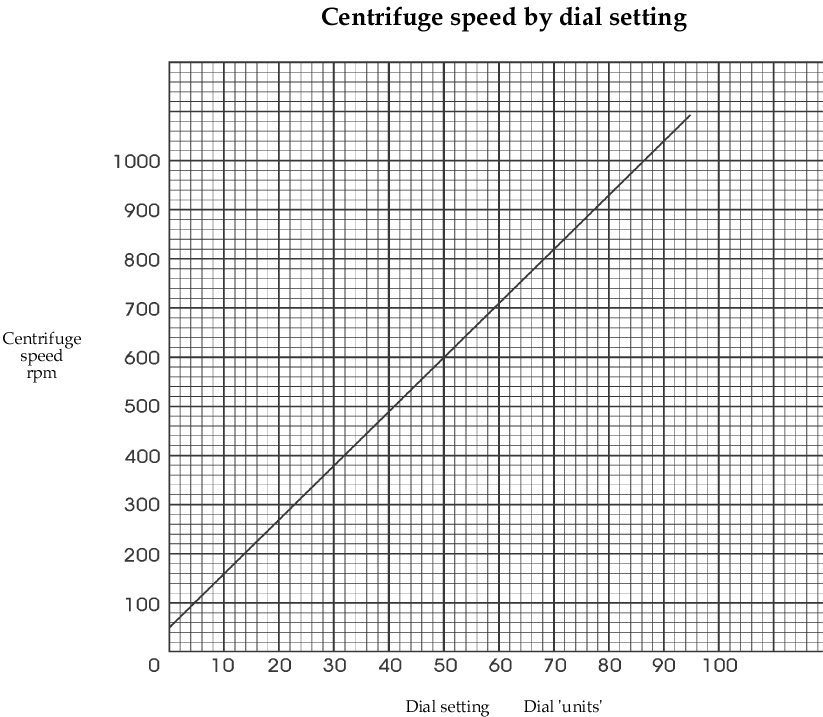
Using the techniques above on how to read a graph, read the values indicated on the following graphs and answer the questions that follow.

****

1. What is the temperature in kelvin when it is 20 °C?

**Your answer:**

1. What is the temperature in degrees Celsius when it is 350 K?

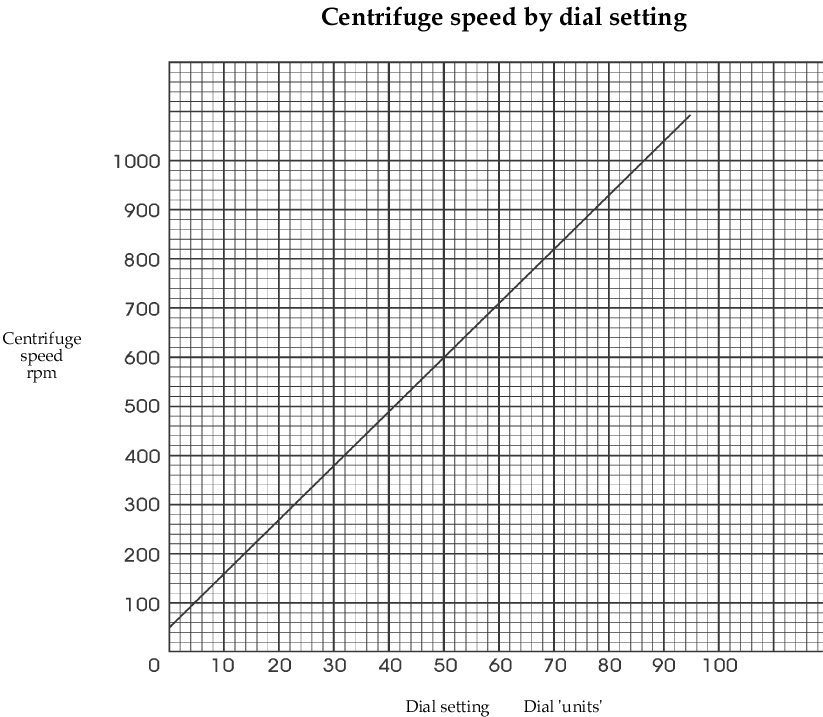
**Your answer:**

1. Find the centrifuge speed when the dial setting is 70.

**Your answer:**

1. What will be the dial setting if a speed of 226 rpm is required?

**Your answer:**

****

Use the graph above to find the speed of the centrifuge when:

1. the dial setting is 30

**Your answer:**

1. the dial setting is 80

**Your answer:**

1. Use the formula S = 11D + 50 to calculate the centrifuge speed

Where:

* S is the centrifuge speed, and
* D is the dial setting.

When:

* 1. the dial setting is 30

**Your answer:**

* 1. the dial setting is 80

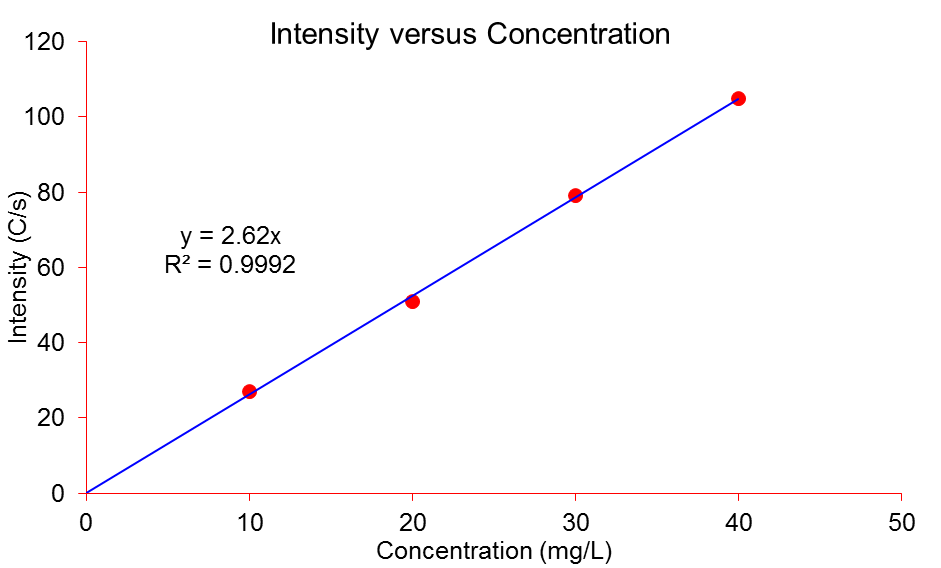
**Your answer:**

NOTE: The values for centrifuge speed were the same using the graph and using the formula! Graphs and formulae can be used to find the same information. We can find the formula that gives the same information as a graph. This can be done for most plot shapes.

|  |  |
| --- | --- |
| Self-check questions icon | Self-check questions – topic 6 |

*Check your understanding of the basic topics in this section by answering the questions below (or other questions provided by your teacher).*

1. Why do we need to interpret data? Why would an incorrect interpretation be bad?
2. Would an opinion poll on preferred political party be considered qualitative or quantitative data? Why?
3. With reference to Figure 6.2, would this data be positively or negatively proportional?
4. What is the problem with the data when a graph turns out to be noise? What could we do to fix it?
5. With your teacher’s assistance, interpret the graph and determine the approximate values when:
   1. The value of the concentration when the intensity is 80 c/s.
   2. The intensity when the concentration is 25 mg/L.



# Appendix A – Metric (SI) tables (reference only)

#### Table of Metric Prefixes

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Prefix Name | Number of Base Units (Magnitude) | Scientific Notation | Symbol | Expression |
| yotta | 1 000 000 000 000 000 000 000 000 | 1024 | Y | Decimal multiple |
| zetta | 1 000 000 000 000 000 000 000 | 1021 | Z |
| exa | 1 000 000 000 000 000 000 | 1018 | E |
| peta | 1 000 000 000 000 000 | 1015 | P |
| tera | 1 000 000 000 000 | 1012 | T |
| giga | 1 000 000 000 | 109 | G |
| mega | 1 000 000 | 106 | M |
| kilo | 1 000 | 103 | k |
| hecto | 1 00 | 102 | h |
| deca | 1 0 | 101 | da |
| Base unit | 1 | Any Base Unit |  |  |
| deci | 0.1 | 10-1 | d | Decimal sub-multiple (fractions) |
| centi | 0.01 | 10-2 | c |
| milli | 0.001 | 10-3 | m |
| micro | 0.000 001 | 10-6 | u |
| nano | 0.000 000 001 | 10-9 | n |
| pico | 0.000 000 000 001 | 10-12 | p |
| femto | 0.000 000 000 000 001 | 10-15 | f |
| atto | 0.000 000 000 000 000 001 | 10-18 | a |
| zepto | 0.000 000 000 000 000 000 001 | 10-21 | z |
| yocto | 0.000 000 000 000 000 000 000 001 | 10-24 | y |

# Appendix B - Tables of Relationships (SI – Imperial)

| Length units | |
| --- | --- |
| Angstrom (Å) | 0.1 nanometre (exactly)  0.000 1 micrometre (exactly)  0.000 000 1 millimetre (exactly)  0.000 000 004 inch |
| 1 centimetre (cm) | 0.393 7 inch |
| 1 chain (ch) | 66 feet (exactly)  20.1168 metres |
| 1 fathom | 6 feet (exactly)  1.828 8 metres |
| 1 furlong (fur) | 10 chains (surveyors) (exactly)  660 feet (exactly)  1/8 U.S. statute mile (exactly)  201.168 metres |
| [1 hand] | 4 inches |
| 1 inch (in) | 2.54 centimetres (exactly) |
| 1 kilometre (km) | 0.621 mile |
| 1 metre (m) | 39.37 inches  1.094 yards |
| 1 mile (mi) (U.S. statute) | 5280 survey feet (exactly)  1.609 kilometres |
| 1 mile (mi) (international) | 5280 international feet (exactly) |
| 1 mile (mi) (int. nautical) | 1.852 kilometres (exactly)  1.151 survey miles |
| 1 Point (typography) | 1/72 inch (approximately)  0.351 millimetre |
| 1 rod (rd), pole, or perch | 16 1/2 feet (exactly)  5.029 2 metres |
| 1 yard (yd) | 0.914 4 metre (exactly) |

| Area units | |
| --- | --- |
| 1 acre | 43 560 square feet (exactly)  0.405 hectare |
| 1 hectare | 2.471 acre |
| [1 square (building)] | 100 square feet |
| 1 square centimetre (cm2) | 0.155 square inch |
| 1 square decimetre (dm2) | 15.500 square inches |
| 1 square foot (ft2) | 929.030 square centimetres |
| 1 square inch (in2) | 6.451 6 square centimetres (exactly) |
| 1 square kilometre (km2) | 247.104 acre  0.386 square mile |
| 1 square metre (m2) | 1.196 square yards  10.764 square feet |
| 1 square mile (mi2) | 258.999 hectares |
| 1 square millimetre (mm2) | 0.002 square inch |
| 1 square rod (rd2), sq pole, or sq perch | 25.293 square metres |
| 1 square yard (yd2) | 0.836 square metre |

| Volume units | |
| --- | --- |
| 1 cubic centimetre (cm3) | 0.061 cubic inch |
| 1 cubic decimetre (dm3) | 61.024 cubic inches |
| 1 cubic foot (ft3) | 7.481 gallons  28.316 cubic decimetres |
| 1 cubic inch (in3) | 0.554 fluid ounce  4.433 fluid drams  16.387 cubic centimetres |
| 1 cubic metre (m3) | 1.308 cubic yards |
| 1 cubic yard (yd3) | 0.765 cubic metre |
| 1 cup, measuring | 8 fluid ounces (exactly)  237 millilitres  1/2 liquid pint (exactly) |
| 1 gallon (gal) (U.S.) | 3.785 litres  0.833 British gallon  128 U.S. fluid ounces (exactly) |
| [1 gallon (gal) (British Imperial)] | 277.42 cubic inches  1.201 U.S. gallons  4.546 litres  160 British fluid ounces (exactly) |
| 1 pint (pt), liquid | 28.875 cubic inches exactly  0.473 litre |
| 1 tablespoon, measuring | 3 teaspoons (exactly)  15 millilitres |
| 1 teaspoon, measuring | 1/3 tablespoon (exactly)  5 millilitres |

| Mass units | |
| --- | --- |
| 1 carat (c) | 200 milligrams (exactly) |
| 1 gram (g) | 15.432 grains  0.035 ounce, avoirdupois |
| 1 kilogram (kg) | 2.205 pounds |
| 1 milligram (mg) | 0.015 grain |
| 1 ounce, avoirdupois (oz avdp) | 437.5 grains (exactly)  0.911 troy or apothecaries ounce  28.350 grams |
| 1 point | * 1. carat   2 milligrams |
| 1 pound, avoirdupois (lb avdp) | 7000 grains (exactly)  1.215 troy or apothecaries pounds  453.592 37 grams (exactly) |
| 1 tonne, gross or long | 2240 pounds (exactly)  1.12 net tonnes (exactly)  1.016 metric tonnes |
| 1 tonne, metric (t) | 2204.623 pounds  0.984 gross tonne  1.102 net tons |

| Time units | |
| --- | --- |
| 1 year (yr) | 12 months  365 days |
| 1 day | 12 hours  720 minutes |
| 1 minute | 60 seconds  1/60 hours |

By action of the 12th General Conference on Weights and Measures (1964) the litre is a special name for the cubic decimetre.

# Appendix C - Answers to practice questions

**Practice activity 0.2**

|  |  |
| --- | --- |
| What do you think could happen if… | Potential consequence |
| You don’t use glassware properly? | Breakage resulting in cuts  Inaccurate results |
| You don’t use electricity properly? | Electrocution, death |
| You don’t use chemicals properly? | Chemical burns, poisoning from inhalation, ingestion or skin contact, death  Inaccurate results |
| You don’t use biological material properly? | Infection, death |
| You don’t use radioactive material properly? | Radiation poisoning, burns, death |
| You don’t use mechanical equipment properly? | Crushing of fingers or limbs  Inaccurate results |
| You don’t handle data correctly? | Inaccurate results |

**Practice activity 2.1**

What type of data could be generated from each stage of data workflow in your workplace or college? Once the data is collected, how could that data be processed? Two examples have been provided for clarity. Complete the remaining rows.

| For this type of work… | What type of data is it? | How could the data be processed? |
| --- | --- | --- |
| Customer service | Names, locations, addresses, type of work | Analysed for sales trends |
| Sample receipt | Sample code(s) | Stored on database for historical reference and traceability |
| Sample preparation | Processes used, reagents used, grinding times, particle size etc. | Comparative techniques to reference methods or standards. Error analysis and uncertainty calculations |
| Analysis | Numerical data for calculations | Determination of results or calibration checks etc. |
| Calculation processes | Numerical data for calculations | Determination of results or calibration checks etc. |
| Statistical treatment | Statistical data determined from raw data | Usually visually and numerically to help with understanding the behaviour, accuracy and precision of test processes and results |
| Quality | Compliance data | Compared to other ‘benchmark’ data |

**Practice activity 3.1**

Estimate the answer for each of the following calculations and compare your answer to the accurate answer from your calculator:

|  |  |  |
| --- | --- | --- |
| 1. (9.30 x 10.4)/1.30 | Estimate: 70 | Calculated answer: 74.4 |
| 1. (0.12 x 5.9)/96 | Estimate: 0.006 | Calculated answer: 0.0074 |
| 1. (0.0022 x 8.9)/4.5 | Estimate: 0.0036 | Calculated answer: 0.0044 |
| 1. (√26 x 4.7)/92 | Estimate: 0.25 | Calculated answer: 0.26 |

**Practice activity 3.2**

How many significant figures are there in the following?

a) 183 3

b) 1024 4

c) 0.04 1

d) 15.301 5

e) 91.00 4

**Practice activity 3.3**

Express the following in Standard Form (Scientific Notation):

|  |  |
| --- | --- |
|  |  |
| 1. 270000 | Answer: 2.7 x 105 |
| 1. 430 | Answer: 4.3 x102 |
| 1. 0.00008 | Answer: 8.0 x 10-5 |
| 1. 0.5 | Answer: 5.0 x 10-1 |
| 1. 803000000 | Answer: 8.03 x 108 |

Express as decimal numbers:

|  |  |
| --- | --- |
|  |  |
| 1. 3 x 104 | Answer: 30000 |
| 1. 2 x 10-3 | Answer: 0.002 |
| 1. 2.6 x 10-1 | Answer: 0.26 |
| 1. 8.95 x 10-5 | Answer: 0.0000895 |
| 1. 7.32 x 104 | Answer: 732000 |

**Practice activity 3.4**

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Prefix | Base unit | Combined unit | Prefix | Base unit | Combined unit |
| nano | metre | nanometre | kilo | gram | *kilogram* |
| milli | second | *millisecond* | micro | ampere | *microampere* |
| milli | mole | *millimole* | kilo | metre | *kilometre* |
| micro | ampere | *microampere* | peta | candela | *petacandela* |

**Practice activity 3.5**

With your teacher’s assistance, complete the following unit conversions.

1. How many ***milligrams*** is 3.674 ***kg***?

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| 3.674 ~~kg~~ | ~~1000 g~~ | 1000 mg |  |  |
|  | ~~1 kg~~ | ~~1 g~~ |  |  |

3.674 kg =

= 3.674 x 1000 x 1000

*= 3674000 mg*

1. Convert 9.7 grams to milligrams.

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| 9.7 ~~g~~ | 1000 mg |  |  |  |
|  | ~~1 g~~ |  |  |  |

9.7 g =

= 9.7 x 1000

*= 9700 mg*

1. Convert 35856 picograms to kilograms (express in both normal and sc. not.).

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| 35856pg | ~~1 ng~~ | ~~1 µg~~ | ~~1 mg~~ | ~~1 g~~ | 1 kg |
|  | ~~1000 pg~~ | ~~1000 ng~~ | ~~1000 µg~~ | ~~1000 mg~~ | ~~1000 g~~ |

35856 pg = 35856 / 1000 x 1000 x 1000 x 1000 x 1000

*= 3.5856x10-11 kg*

1. Convert 1.12 ***carats*** (c) ***milligrams*** (note 1 carat = 200mg).

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| 1.12 ~~carats~~ | 200 mg |  |  |  |
|  | ~~1 c~~ |  |  |  |

1.12 carats = 1.12 x 200

*= 0.224 mg*

**Practice activity 3.6**

| **No.** | **Convert** | **To** | **Answer** |
| --- | --- | --- | --- |
| 1 | 789 nm | m | 7.89 x 10-7 or 0.000000789 |
| 2 | 0.848 g | µg | 8.48 x 10-6 or 0.000000848 |
| 3 | 4.25 x 10-4 kL | mL | 425 |
| 4 | 270 mm2 | m2 | 2.7 x10-4 or 0.00027 |
| 5 | 22 yards | m | 20.1168 |
| 6 | 62.3 kg | pounds | 137 |
| 7 | 8456 kL/hr | L/s | 2349 |
| 8 | 82 µg/g | mg/kg | 82 |
| 9 | 23.45 g/100 mL | mg/L | 234500 |

**Practice activity 3.7**

a) Convert 1.45 litres to m3 Answer: 1.45 x 10-3 m3

b) Convert 4500 lb/ft2 to Pascals Answer: 215 461 Pa

c) Convert 3.7 Horsepower (hp) to Watts Answer: 2760 Watts

d) Convert 20 mL to m3 Answer: 2.0 x 10-5

e) Convert 100 000 Pascals to atmospheres Answer: 0.977 atm

**Practice activity 3.8**

Use the techniques presented above to transpose the following formula:

Transpose this equation to make Q, P and C the subject (three separate transpositions):

**Practice activity 3.9**

**Practice activity 3.10**

Write the following measurements and their associated uncertainty to the appropriate number of significant figures.

1. 123 ± 12 kPa

123 ± 10kPa

1. 14 ± 2.3 ms-1

14.0 ± 2.0 ms-1

1. 4.765 ± 0.36 g/mL

4.8 ± 0.4 g/mL

**Practice activity 3.11**

1. Convert to the simplest fraction:
   1. 85 %
   2. 8⅓ %
   3. 1.2 %

Convert to decimals:

1. 24 %

0.24

1. 0.5 %

0.005

1. 112⅔ %

1.1267 (to 4 decimal places)

1. Convert to percentages:
   1. 2/5

40%

* 1. 0.08

8%

* 1. 3.9

350%

1. 35% of an amount is 2.8 kg. What is the whole amount?

8kg

1. A mixture contains 84% of a certain compound. How many grams of the mixture are required to obtain 5.6 g of the compound?

6.667g

**Practice activity 4.1**

Below are some statements that quote statistics (not real ones!). Identify whether the statistics are *descriptive* or *inferential*.

|  |  |
| --- | --- |
| Statistic | Descriptive or Inferential? |
| The average December temperature in Sydney has increased by 1°C in the last 50 years. | Descriptive |
| It is expected that the average December temperature in Sydney will increase by another 1°C in 25 years. | Inferential |
| 25% of people surveyed at a shopping centre indicated that they were aware of increasing temperatures in Sydney. | Descriptive |
| That same survey therefore tells us that 75% of Sydney-siders are ignorant of the changing climatic conditions in their city. | Inferential |

**Practice activity 4.2**

Identify the sample and the population in the following scenarios.

| Statistic | Sample and population? |
| --- | --- |
| A bottle of water is taken from a dam to be tested. | Sample  Population |
| The frog population of a large wetland is checked by looking at two separate hectares. | Sample  Population |
| The levels of lead in dust ‘fallout’ around a smelter are assessed by testing a selection of properties. | Sample  Population |
| Environmental opinions are surveyed in a shopping centre in a study to determine the level of awareness in society. | Sample  Population |

**Practice activity 4.3** discuss with your teacher

**Practice activity 4.4** discuss with your teacher

**Practice activity 5.1** discuss the results with your teacher

**Practice activity 6.1**

Question 1: 295 K

Question 2: 75oC

Question 3: 820 rpm

Question 4: Dial setting 16

Question 5: 380 rpm

Question 6: 930 rpm

Question 7: a) 380 rpm b) 930 rpm

# Appendix D – Answers to Self-Check Questions

**Topic 1**

1. A measurement is a combination of two parts. What are the two parts that make a measurement?

*A number and a unit, for example, 60 kilometre per hour*

1. Generally speaking, what is a unit of measure? Briefly explain how we reference a constant.

*A unit of measure is a reference to something that does not change, a constant. Then the number indicates how many of the unit, or constant, you have.*

1. What is the difference between a base unit and a derived unit of measure?

*There are seven base units. Derived units are formed by combining derived units.*

1. Can anyone trust results that come from uncalibrated equipment? Explain.

*No. If a measuring device is not calibrated then you have no way of knowing the level of uncertainty and you cannot provide a statement of error for the results obtained.*

1. Name one international body that contributes to metrology. What does this organisation do?

*The International Bureau of Weights and Measures (in French, Bureau Internationale des Poids et Mesures, or BIPM) ensures worldwide unification of measurements*

1. What do you think would be one consequence of having poor traceability?

*It would not be possible to know which results related to which samples.*

1. A series of repeated measurements were taken. The results were very close to each other, but not quite near the true value. Would these measurements be precise or accurate? Why?

*These results would be precise but not accurate. They are precise because they are close to each other but they are not accurate because they are not the true value.*

**Topic 2**

1. Why is it so important to check data quality? Give at least three reasons why.

* *To ensure the data is within specification*
* *To ensure customer satisfaction*
* *To ensure compliance*
* *To ensure correct interpretations are made*
* *Any answer that reflects quality would be acceptable.*

1. *Discuss this question with your teacher.*
2. A 20mL volumetric flask is filled to the mark and weighed. This is repeated five times. Assume that the temperature is the same at each weighing. Which of the following results might be a transcription or other type of error: 19.968g, 19.994g, 19.399g, 20.013g, 20.002g. Discuss the reasons for the variation and sources of error with your teacher.

*19.399g because it is significantly different from the other measurements.*

1. Why do you think the process of rectifying errors is a formal process?

*So that any mistakes are learned from, so that systematic errors can be identified and rectified and so that there is a traceable way of ensuring that customers receive the correct results.*

1. A technician has written down the following results in the laboratory and then typed them into a table. Have any errors been made? What is the source of the error?

|  |
| --- |
| Notebook  **3/5/18 Readings: SL2018-01 8.562mg/L SL2018-02 9.362mg/L**  **4/5/18 Reading: SL2018-03 4.695mg/L** |

|  |  |  |
| --- | --- | --- |
| Date | SL Number | Reading (mg/L) |
| 03/05/2018 | SL2018-01 | 8.662 |
| 04/05/2018 | SL2018-02 | 9.362 |
| 04/05/2018 | SL2018-03 | 8.695 |

*Yes, errors have been made in entering the data.*

*SL2018-01 has an incorrect reading, the reading should b 8.562mg/L.*

*SL2018-02 has the wrong date, the date should be 03/05/2018.*

*SL2018-03 has an incorrect reading, the reading should be 4.695mg/L.*

*The source of these errors is the technician.*

1. Identify 4 sources of data from your workplace or campus laboratory. Indicate whether the data is in written or digital form. *Discuss this answer with your teacher in relation to your particular workplace or laboratory.*
2. What does LIMS stand for? What does it do for a laboratory?

*LIMS stands for Laboratory Information Management System.*

*A LIMS manages information for all parts of the laboratory workflow such as: customer information, organisational structure, policies and procedures, storing results, quality control of results, communication of results to clients, information on instrument maintenance and calibration, to help ensure the overall quality of analytical data that the lab generates for its clients.*

**Topic 3**

1. You determine that there are 240 microorganisms on a pour plate of 10:3 dilution. How many microorganisms were the original sample?

*800*

1. What is the mean cell volume if haemocrit is 42% and RBC is 4.2 x 10-12?

*88fL*

1. Calculate the concentration in g/L of a solution if 0.25 g of a substance is dissolved in a final volume of 0.4 L.

*0.625g/L*

1. Find the mass of dye present in 5 mL of a solution whose concentration is 70 mg/dL.

*3.5mg*

1. There are 24 g of solute in 300 ml of solution.
   1. What is the % concentration?

*8%*

* 1. Is this concentration %w/w or %w/v?

*%w/v concentration*

1. A stock solution of dye in water has its concentration given as 300 g per litre.
   1. What mass of dye is contained in one litre of solution?

*300g*

* 1. Find the mass of dye in one decilitre of solution.

*30g*

* 1. Find the mass of dye in 5 ml of solution.

*1.5g*

1. How much 40% v/v sulphuric acid (H2S04) would be required to make 1 000 ml of 0.5% v/v H2S04?

*V1 = C2V2/C1*

*= 0.5 \* 1000 / 40*

*= 12.5 mL*

1. Which one of the following four solutions of a salt in water is the most concentrated?

*c is the most concentrated*

1. A manufacturer wants to prepare 200kg of a biscuit dough which contains 24% sugar by weight. The operator already has two other standard mixes which contain 20% sugar (batch 1) and 30% sugar (batch 2). How much of each batch should be combined to make 200kg of biscuit dough containing 24% sugar by weight?

*120kg of batch 1 and 80kg of batch 2.*

1. How much cream at 28% fat must be added to 1000kg of milk containing 2.0% fat to give a fat content of 3.5%.

*61.2kg of cream with 28% fat needs to be added to the milk.*

**Topic 4**

1. How do a sample and a population differ?

*A population is the entire set of individual items about which observations are made, tests performed on and data recorded. This is very rare due to cost, time and inconvenience.*

*A sample is a subset or proportion of the population.*

1. Under what circumstances can we use the word census?

*A census is the process of collecting data from an entire population.*

1. Why is a representative sample so important to obtain?

*If you want to understand and make judgements about an entire population then you need to ensure that the sample you are testing is going to give you the same results as if you tested the whole population. If it does then the sample is representative of the population. If the sample is not representative then you cannot apply the sample results to the population.*

1. What is the difference between descriptive and inferential statistics?

*Descriptive statistics are ways of describing aspects of the data set that has been sampled or measured. Inferential statistics uses the information from the sample to make judgements about the broader population.*

1. Briefly describe the process of creating a frequency histogram.

*Look at the data and decide how to organise it into groupings (bins) or individual values. Create a tally chart that counts how many data points fall into each grouping, that is, the frequency of each. Create a graph with the values for the groupings (bins) or individual values along the x-axis and the frequency on the y-axis. Ensure the histogram has a title and labelled axes with a suitable scale.*

1. What is the difference between the distribution of a data set and the central tendency of the data?

*The distribution describes the spread of the data, how spread out it is and if there are any patterns. Central tendency is a measure of the centre point that the data is spread around. In a normal distribution this is the highest point.*

1. What is the normal distribution and why is it so fundamental in statistics?

*The normal distribution is a pattern of data in which the central tendencies (mean, mode, median) are all equal and the frequencies fall away from the central value in a standard way, measured in standard deviations. It is important because many natural phenomena follow a standard deviation and also because it is mathematically easy to work with.*

1. Calculate the mean, median, mode and sample standard deviation of the following sample data set.

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| 5 | 40 | 177 | 40 | 11 | 29 | 88 | 49 | 117 | 64 | 53 | 29 | 57 | 69 | 3 | 30 | 9 | 66 | 58 | 46 |

*Mean: 52*

*Median: 47.5*

*Mode: 40*

*Sample standard deviation: 40.9*

**Topic 5**

1. What is the benefit to visualising data?

*It is hard to describe data or to see patterns and trends when data is in numerical form, especially if there are a lot of data points. When the data is put into graphical or other format it can be more easily understood.*

1. What is the difference between a plot and a chart?

*A plot is created when points are placed on a coordinate system.*

*A chart is create when shapes are used to display data.*

1. Identify and list three general graphing principles

* *axes labelled clearly (X axis and Y axis)*
* *informative title*
* *axis scale is the same all the way along – if 0-10 covers 1 cm, then so should 90-100 etc.*

1. Why should a graph be labelled and formatted correctly?

*Unless a graph has appropriate labels, scale and accuracy then it is not a true representation of the data. Incorrect graphs can be used to manipulate data to give meanings that are not there.*

1. What is the key difference between a line graph and an xy scatter graph?

*An xy scatter graph shows the relationship (or absence of a relationship) between two variables, that is, whether when one changes the other changes, a line can be used to represent the relationship but it could be that only some or no data actually falls on the line. A line graph is plotted data points joined together by a line or a curve. It implies that the data is continuous between the points.*

1. How is a statistical process control chart (SPC) related to the concepts of distribution and central tendency of data?

*A statistical process control chart uses previous data from the process to track how the process is currently going. It shows the mean and also one standard deviation so that when you plot the current data you can see if it is going to follow the acceptable range or if the process is moving into an unacceptable range.*

1. How do we identify a bad graph?

*A bad graph might not have a title, labelled axes or correct data or might not have used appropriate scales.*

1. Why is confidentiality such an important part of data treatment?

*When dealing with any type of personal data there are laws that require confidentiality. Laboratories have systems and policies in place to ensure that personal data is protected and all staff need to understand and adhere to these policies.*

1. *Work with your teacher on this question.*

**Topic 6**

1. Why do we need to interpret data? Why would an incorrect interpretation be bad?

*Interpreting data gives it meaning, without interpretation you won’t know what the data is really telling you. If the data is incorrectly interpreted then decisions might be made on the wrong basis.*

1. Would an opinion poll on preferred political party be considered qualitative or quantitative data? Why?

*Poll data is quantitative because it is based on numerical measurement.*

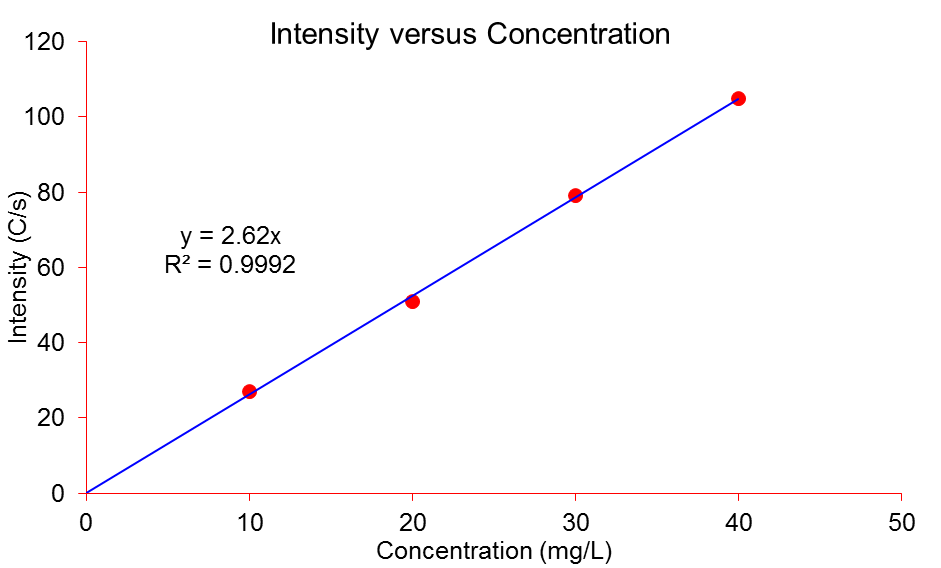
1. With reference to *Figure 6.2 – CO2 and Temperature (1964 to 2008)*, would this data be positively or negatively proportional?

*This data is positively proportional as temperature increases as CO2 increases.*

1. What is the problem with the data when a graph turns out to be noise? What could we do to fix it?

If no obvious correlation exists, we would change the axis data types as we have clearly graphed data that makes no sense.

1. With your teacher’s assistance, interpret the graph and determine the approximate values when:
   1. The value of the concentration when the intensity is 80 c/s.
   2. The intensity when the concentration is 25 mg/L.



* 1. *Approximately 30 mg/L*
  2. *Approximately 65 c/s*

# Bibliography

DAMA UK . (2013). *The six primary dimensions for data quality assessment: Defining data quality dimensions.* UK: DAMA UK.

Heinonen, M. (2013). *Introduction to metrology - presentation.*

*International vocabulary of metrology — Basic and general concepts and associated terms (VIM), 3rd ed.,.* (2008).

National Measurement Act 1960.

National Measurement Guidelines 2016.

National Measurment Regulation 1999.

Organisation Intergouvernementale de la Convention du Mètre. (2006). *The International System of Units (SI).* Paris: BIPM.

PACFA. (2014). *Guidelines for client records.* Sydney: PACFA.

# References

International Bureau of Weights and Measures (2006), The International System of Units (SI) (PDF) (8th ed.), [ISBN](https://en.wikipedia.org/wiki/International_Standard_Book_Number) 92-822-2213-6.

Oxford Dictionaries (2019). *calculation | Definition of calculation in English by Oxford Dictionaries*. [online] Available at: https://en.oxforddictionaries.com/definition/calculation [Accessed 2 May 2019].

Oxford Dictionaries | English. (2019). *statistics | Definition of statistics in English by Oxford Dictionaries*. [online] Available at: https://en.oxforddictionaries.com/definition/statistics [Accessed 2 May 2019].

## Image attribution labels

| Image | Attribution |
| --- | --- |
| Title page image | [Person writing in notebook](https://www.pexels.com/photo/person-writing-on-notebook-669615/) by [Lukas](https://www.pexels.com/@goumbik) under [CC0 1.0](https://creativecommons.org/publicdomain/zero/1.0/deed.en) |
| Figure 0.1 – Examples of different types of laboratory | Copied Under s113P, The University of Utah, <https://www.che.utah.edu/projects_lab>, accessed 27 February 2019 |
| Figure 0.2 – Example of a Safe Work Method Statement | Copied Under s113P, unknown, JSEasy <https://www.jseasy.com.au/new-swms-template>, accessed 21 March 2019 |
| Figure 1.1 – A flowchart of quality from international to laboratory level | © TAFE NSW 2018 |
| Figure 1.2 – Image showing the 7 base units as symbols plus some basic inter-relationships between the units. | Copied under s113P, [Dono](https://en.wikipedia.org/wiki/User:Dono), <https://commons.wikimedia.org/wiki/File:SI_base_unit.svg>, accessed 27 February 2019 |
| Figure 1.4 – The dartboard analogy for accuracy and precision | Copied Under s113P, unknown, <https://bigideasgreatscience.wordpress.com/denk-en-werkwijzen/data-analyseren-en-interpreteren/>, accessed 27 February 2019 |
| Figure 1.5 – Example of a ‘traceability ladder’. | Copied under s113P, Manufacturing Skills Australia, *Process and interpret data Assessment Guide for Trainers and Assessors*, Manufacturing Skills Australia, North Sydney, December 2007 |
| Figure 2.1 – Example of a generalised laboratory data workflow. | Copied under s113P, unknown, <https://area-info.net/global-laboratory-information-management-system-lims-market-2017-promium-perkinelmer-genologics/> accessed 21 March 2019 |
| Figure 2.2 – Example of routine paperwork used by laboratories. | © TAFE NSW 2018 |
| Figure 2.3 – Checking the quality of data using the 6 data quality dimensions | Copied under s113P, DAMA UK, <https://www.whitepapers.em360tech.com/wp-content/files_mf/1407250286DAMAUKDQDimensionsWhitePaperR37.pdf>, accessed 21 March 2019 |
| Figure 3.1 – Story from the New York Times from 30/07/1983. | Copied under s113P, NY Times, <https://www.nytimes.com/1983/07/30/us/jet-s-fuel-ran-out-after-metric-conversion-errors.html>, accessed 21 March 2019 |
| Figure 3.2 – Example of a complex unit conversion. You will not necessarily need to calculate something like this, but then again… | © TAFE NSW 2018 |
| Figure 4.1 – Descriptive statistics tree. | © TAFE NSW 2018 |
| Figure 4.2 A) Jelly beans | Copied under s113P, Gracefruit <https://www.gracefruit.com/item/466/Gracefruit/Jelly-Beans-Fragrance-Oil.html>, accessed 28 March 2019 |
| Figure 4.2 B) Example of an analytical balance to weight jelly beans | Copied under s113P, Sartorius, <https://www.coleparmer.ca/i/sartorius-entris64i-1sus-analytical-balance-60-x-0-1mg-internal-calibration/1195219>, accessed 28 March 2019 |
| Figure 4.3 - The mathematical model of the Normal Distribution showing standard deviations broken down into ‘halves’ with the full percentages displayed underneath. | Copied under s113P, Charley Kyd, <http://www.exceluser.com/formulas/statsnormal.htm>, accessed 28 March 2019 |
| Figure 4.4 – Examples of non-normal distributions | Copied under s113P, Chance Encounters by C.J. Wild and G.A.F. Seber, © John Wiley & Sons, 2000. Figure 2.3.10 |
| Figure 5.1 – Examples of plots, graphs and charts | © TAFE NSW 2018 |
| Figure 5.2 – The major parts of a graph (or chart, or plot) | © TAFE NSW 2018 |
| Figure 5.3 – Example of a Column chart (left) and a Bar chart (right) | © TAFE NSW 2018 |
| Figure 5.4 – Examples of column graphs with data displayed in different ways. | © TAFE NSW 2018 |
| Figure 5.5 – Example of a line graph | © TAFE NSW 2019 |
| Figure 5.6 – Three different examples of scatter plots | © TAFE NSW 2018 |
| Figure 5.7 – The normal distribution viewed as a line chart with the mean (dotted line) and Standard Deviations (solid vertical lines). | Copied under s113P, unknown, <https://www.moresteam.com/toolbox/statistical-process-control-spc.cfm>, accessed 28 March 2019 |
| Figure 5.8 – The basic concept of a Statistical Process Control (SPC) chart. | Copied under s113P, unknown, <https://www.moresteam.com/toolbox/statistical-process-control-spc.cfm>, accessed 28 March 2019 |
| Figure 5.9 – The old pie chart | © TAFE NSW 2018 |
| Figure 5.10 – Field assessment of stone content | Copied under 2113P, D.L. Rowell, *Soil Science: Methods and applications*, Longman Scientifici & Technical, Essex UK, 1994, ISBN 0-592-087848 |
| Figure 5.11 – Example of short style ‘reports’ that lab technicians use on a routine basis | © TAFE NSW 2018 |
| Figure 5.12 – The role that LIMS can play in a laboratory. Data storage, treatment and reporting can be all done by this system | Copied under s113P, unknown, <https://www.autoscribeinformatics.com/resources/blog/what-is-a-lims>, accessed 29 March 2019 |
| Figure 5.13 – Example of a Certificate of Analysis | © TAFE NSW 2019 |
| Figure 6.1 – The ‘data tree’. | © TAFE NSW 2018 |
| Figure 6.2 – Example of a graph with trends and patterns. | Copied under s113P, unknown, <https://static.skepticalscience.com/images/co2_temp_1964_2008.gif>, accessed 1 April 2019 |
| Figure 6.3 – We clearly see what is meant by noise when we compare negative and positive correlations with a ‘no correlation’ graph. | Copied under s113P, unknown, <https://thesalience.wordpress.com/psychology-101/research-methods-in-psychology/correlations/> accessed 1 April 2019 |
| Figure 6.4 – Examples of parabolic curved lines | © TAFE NSW 2018 |
| Figure 6.5 – Example of hyperbolic curved lines. | © TAFE NSW 2018 |
| Figure 6.6 – A graph of centrifuge speed vs dial setting. | Copied under s113P, unknown, <https://scs.illinois.edu/resources/facilities/high-throughput-screening-facility/equipment/centrifuge> accessed 1 April 2019 |
| Figure 6.7 – Applying the instructions to the graph. | © TAFE NSW 2018 |