General Maths Semester 2 Summary

Matrix Summary

A matrix is a rectangular array of numbers arranged in rows and columns. The numbers in a matrix are called the **elements** of the matrix. The **order** of a matrix is the number of rows and columns in the matrix.

Example 1

 $A = \begin{bmatrix} 3 & 5\\ 0 & 4\\ -5 & 6 \end{bmatrix}$ is a **3 by 2 or 3** × **2** matrix as it has 3 rows and 2 columns. Matrices are often

denoted by capital letters.

Types of Matrices

 $C = \begin{bmatrix} 4 \\ 9 \\ 2 \end{bmatrix}$ is a 3 by 1 **column** matrix and $D = \begin{bmatrix} 8 & -3 & 0 & 15 \end{bmatrix}$ is a 1 by 4 **row** matrix.

A null or zero matrix has all elements zero.

 $O = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}$ is a 2 by 3 **zero** matrix.

Note that the null matrix is the identity matrix for addition and is often denoted by 0.

 $F = \begin{bmatrix} 4 & -3 & 0 \\ 5 & -5 & 6 \\ 0 & 4 & 1 \end{bmatrix}$ is a 3 by 3 square matrix. A square matrix has the same number of rows and

columns.

A diagonal matrix is a square matrix that has non-zero elements in the leading diagonal only.

 $A = \begin{bmatrix} 3 & 0 & 0 \\ 0 & -7 & 0 \\ 0 & 0 & 4 \end{bmatrix}$ is a diagonal matrix.

Elements of a Matrix

The elements of a matrix are often denoted by small letters with subscripts. For example:

$$A = \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix}$$

Where $a_{11} = element in row 1, column 1, a_{12} = element in row 1, column 2 etc$

Two matrices are equal when they have the same order and corresponding elements are equal.

If
$$A = \begin{bmatrix} 3 & 9 & -4 \\ 0 & -5 & -2 \\ 1 & 0 & -2 \end{bmatrix}$$
 and $B = \begin{bmatrix} 3 & 9 & -4 \\ 0 & -5 & -2 \\ 1 & 0 & -2 \end{bmatrix}$ then $A = B$

Addition and Subtraction of Matrices

When matrices have the same order they can be added and subtracted by simply adding or subtracting corresponding elements.

Example 2

$$A = \begin{bmatrix} -4 & 0 \\ 3 & 8 \\ -5 & 7 \end{bmatrix} \text{ and } B = \begin{bmatrix} -3 & 7 \\ 8 & -3 \\ 4 & -5 \end{bmatrix}, \qquad \text{Calculate} \qquad \text{i} \ A - B \qquad \text{ii} \ 3A - 2B$$

i)
$$A - B = \begin{bmatrix} -4 & 0 \\ 3 & 8 \\ -5 & 7 \end{bmatrix} - \begin{bmatrix} -3 & 7 \\ 8 & -3 \\ 4 & -5 \end{bmatrix} = \begin{bmatrix} -1 & -7 \\ -5 & 11 \\ -9 & 12 \end{bmatrix}$$

ii)
$$3A - 2B$$

$$3A = 3 \begin{bmatrix} -4 & 0 \\ 3 & 8 \\ -5 & 7 \end{bmatrix} = \begin{bmatrix} -12 & 0 \\ 9 & 24 \\ -15 & 21 \end{bmatrix}$$

$$2B = 2 \begin{bmatrix} -3 & 7 \\ 8 & -3 \\ 4 & -5 \end{bmatrix} = \begin{bmatrix} -6 & 14 \\ 16 & -6 \\ 8 & -10 \end{bmatrix}$$

$$So \qquad 3A - 2B = \begin{bmatrix} -12 & 0 \\ 9 & 24 \\ -15 & 21 \end{bmatrix} - \begin{bmatrix} -6 & 14 \\ 16 & -6 \\ 8 & -10 \end{bmatrix} = \begin{bmatrix} -6 & -14 \\ -7 & 30 \\ -23 & 31 \end{bmatrix}$$

Note that you can easily do the above calculations on the calculator.

Multiplying Matrices

Two matrices can be multiplied if the number of columns in the first matrix is the same as the number of rows in the second matrix.

If matrix A is of order $m \times n$ and matrix B is of order $n \times p$ then the product $A \times B$ exists and will be of order $m \times p$.

Example 3.

$$A = \begin{bmatrix} 3 & 2 \\ -6 & 0 \\ -4 & 4 \end{bmatrix} \text{ and } B = \begin{bmatrix} -3 & 2 \\ 0 & -6 \end{bmatrix} \text{, find } A \times B \text{ if it exists.}$$

Since A is a 3×2 matrix and B is a 2×2 matrix the product exists and will be a 3×2 matrix.



Matrix will be 3 x 2

$$A \times B = \begin{bmatrix} 3 & 2 \\ -6 & 0 \\ -4 & 4 \end{bmatrix} \times \begin{bmatrix} -3 & 2 \\ 0 & -6 \end{bmatrix} = \begin{bmatrix} 3 \times -3 + 2 \times 0 & 3 \times 2 + 2 \times -6 \\ -6 \times -3 + 0 \times 0 & -6 \times 2 + 0 \times -6 \\ -4 \times -3 + 4 \times 0 & -4 \times 2 + 4 \times -6 \end{bmatrix} = \begin{bmatrix} -9 & -6 \\ 18 & -12 \\ 12 & -32 \end{bmatrix}$$
$$A \times B = \begin{bmatrix} -9 & -6 \\ 18 & -12 \\ 12 & -32 \end{bmatrix} \text{ which is a } 3 \times 2 \text{ matrix.}$$

Note that the matrix product $B \times A$ does **not** exist since.

2 × 2 and 3 × 2 unequal so product does not exist In general $A \times B \neq B \times A$, so the order of matrix multiplication is important!

Matrices and Equations

The Identity Matrix for Multiplication

The identity matrix for multiplication is a square matrix in which all the elements are zero except those in the leading diagonal which are 1.

Examples of identity matrices:

$$I = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$$
 or $I = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$

The identity matrix is always denoted by I

When any matrix is multiplied by the identity matrix that matrix will remain unchanged. It is like the number 1 in normal multiplication. In normal multiplication 1 is the identity for multiplication of all numbers as any number multiplied by 1 is not changed.

For example: $1 \times 21 = 21 = 21 \times 1$ Notice that multiplying 21 by 1 leaves 21 unchanged.

If
$$A = \begin{bmatrix} 1 & 4 \\ 3 & 5 \end{bmatrix}$$
 and $I = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$ then $AI = \begin{bmatrix} 1 & 4 \\ 3 & 5 \end{bmatrix} \times \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} 1 & 4 \\ 3 & 5 \end{bmatrix}$
and $IA = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \times \begin{bmatrix} 1 & 4 \\ 3 & 5 \end{bmatrix} = \begin{bmatrix} 1 & 4 \\ 3 & 5 \end{bmatrix}$

Note that AI = IA. This is an exception to the rule $AB \neq BA$

The Inverse of a Matrix

When you multiply 4 by $\frac{1}{4}$ you will get the answer 1. ie. $4 \times \frac{1}{4} = 1 = \frac{1}{4} \times 4$. We say that $\frac{1}{4}$ is the multiplicative inverse of 4. Also 4 is the multiplicative inverse of $\frac{1}{4}$.

Consider the following product: $\begin{bmatrix} 2 & 3 \\ 3 & 5 \end{bmatrix} \times \begin{bmatrix} 5 & -3 \\ -3 & 2 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$

The result is the identity matrix *I*. We say that $\begin{bmatrix} 5 & -3 \\ -3 & 2 \end{bmatrix}$ is the multiplicative inverse of $\begin{bmatrix} 2 & 3 \\ 3 & 5 \end{bmatrix}$ and $\begin{bmatrix} 2 & 3 \\ 3 & 5 \end{bmatrix}$ is the multiplicative inverse of $\begin{bmatrix} 5 & -3 \\ -3 & 2 \end{bmatrix}$

Also $\begin{bmatrix} 5 & -3 \\ -3 & 2 \end{bmatrix} \times \begin{bmatrix} 2 & 3 \\ 3 & 5 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$

In matrices we use the symbol A^{-1} to denote the multiplicative inverse of A.

So if
$$A = \begin{bmatrix} 2 & 3 \\ 3 & 5 \end{bmatrix}$$
 then its multiplicative inverse is given by:
$$A^{-1} = \begin{bmatrix} 5 & -3 \\ -3 & 2 \end{bmatrix}$$

and

$$\mathbf{A}\mathbf{A}^{-1} = \mathbf{I} = \mathbf{A}^{-1}\mathbf{A}$$

Only square matrices have multiplicative inverses.

So in general for any square matrix A:

If $AA^{-1} = I = A^{-1}A$ then A^{-1} is called the multiplicative inverse of A

Finding the Multiplicative Inverse of a 2 by 2 matrix.

In order to find the inverse of a 2 by 2 matrix, you need to find the determinant of the given matrix first.

The Determinant of a 2 by 2 Matrix

If $A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$ then the determinant of matrix A, denoted by |A| or det(A), is given by:

|A| = ad - bc. Note that if the determinant of a matrix is zero then that matrix is called a **singular** matrix.

Example 1: Find the determinant of $A = \begin{bmatrix} 3 & -5 \\ 2 & -4 \end{bmatrix}$

 $det(A) = 3 \times (-4) - (-5) \times 2 = -12 - (-10) = -12 + 10 = -2$

Example 2: Find the determinant of $B = \begin{bmatrix} 4 & 8 \\ 3 & 6 \end{bmatrix}$

 $det(B) = 4 \times 6 - 8 \times 3 = 24 - 24 = 0$ Matrix B is **singular** because its determinant is **zero**.

The multiplicative inverse of any 2 by 2 matrix can be found using the following steps, provided the inverse exists.

If
$$A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$$

a. Find the determinant of the matrix.

$$\det(\mathsf{A}) = = ad - bc$$

b. Multiply the matrix by $\frac{1}{\det(A)} = \frac{1}{ad-bc}$

 $\frac{1}{ad-bc} \begin{bmatrix} a & b \\ c & d \end{bmatrix}$

c. Swap the elements in the main diagonal and change the sign of the elements in the other diagonal. The inverse matrix $A^{-1} = \frac{1}{ad-bc} \begin{bmatrix} d & -b \\ -c & a \end{bmatrix}$

Note that the inverse of a matrix can be easily found using the calculator. Simply raise the matrix to the power of -1.

Example 3: If $A = \begin{bmatrix} 3 & 5 \\ 2 & 4 \end{bmatrix}$, find the inverse, A^{-1}

First find the determinant of A. d

 $\det(A) = 3 \times 4 - 5 \times 2 = 2$

$$A^{-1} = \frac{1}{2} \begin{bmatrix} 4 & -5 \\ -2 & 3 \end{bmatrix}$$

Check that $A \times A^{-1} = I$

$$\begin{bmatrix} 3 & 5 \\ 2 & 4 \end{bmatrix} \times \frac{1}{2} \begin{bmatrix} 4 & -5 \\ -2 & 3 \end{bmatrix} = \frac{1}{2} \begin{bmatrix} 3 & 5 \\ 2 & 4 \end{bmatrix} \times \begin{bmatrix} 4 & -5 \\ -2 & 3 \end{bmatrix} = \frac{1}{2} \begin{bmatrix} 2 & 0 \\ 0 & 2 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} = I$$

Note that if the matrix is singular (the determinant is zero) then the inverse of the matrix will **not** exist. This is because $\frac{1}{\det(A)} = \frac{1}{0}$ which is not defined. You **cannot** divide by zero.

Solving Matrix Equations

Inverses enable us to easily solve matrix equations. We use the following technique:

a. If AX = B, where A, X and B are matrices. We can find the matrix X by doing the following:

Pre multiply both sides of the equation by A^{-1}

$$A^{-1}AX = A^{-1}B$$

 $IX = A^{-1}B$, since $A^{-1}A = I$

 $X = A^{-1}B$ since IX = X

In general if AX = B then $X = A^{-1}B$

b. If XA = B, where A, X and B are matrices. We can find the matrix X by doing the following:

Post multiply both sides of the equation by A^{-1}

$$XAA^{-1} = BA^{-1}$$

 $XI = BA^{-1}$, since $AA^{-1} = I$

 $X = BA^{-1}$ since XI = X

In general if XA = B then $X = BA^{-1}$

Example 4

If
$$A = \begin{bmatrix} 5 & 5 \\ 0 & 6 \end{bmatrix}$$
 and $B = \begin{bmatrix} 6 & -3 \\ 3 & 3 \end{bmatrix}$ find X if
a. $AX = B$ and b. $XA = B$
a. For $AX = B$
 $X = A^{-1}B$
So $X = \begin{bmatrix} 5 & 5 \\ 0 & 6 \end{bmatrix}^{-1} \times \begin{bmatrix} 6 & -3 \\ 3 & 3 \end{bmatrix}$

The quickest way is to use the calculator which will give the following answers, depending on your settings:

$$X = \begin{bmatrix} \frac{7}{10} & -\frac{11}{10} \\ \frac{1}{2} & \frac{1}{2} \end{bmatrix} or \begin{bmatrix} 0.7 & -1.1 \\ 0.5 & 0.5 \end{bmatrix}$$
 The first answer can be simplified to $\frac{1}{10} \begin{bmatrix} 7 & -11 \\ 5 & 5 \end{bmatrix}$

b. For XA = B

$$X = BA^{-2}$$

 $X = \begin{bmatrix} 6 & -3 \\ 3 & 3 \end{bmatrix} \times \begin{bmatrix} 5 & 5 \\ 0 & 6 \end{bmatrix}^{-1}$

The quickest way is to use the calculator which will give the following answers, depending on your settings:

$$X = \begin{bmatrix} \frac{6}{5} & -\frac{3}{2} \\ \frac{3}{5} & 0 \end{bmatrix} \text{ or } \begin{bmatrix} 1.2 & -1.5 \\ 0.6 & 0 \end{bmatrix}$$
 The first answer can be simplified to $\frac{1}{10} \begin{bmatrix} 12 & -15 \\ 6 & 0 \end{bmatrix}$

Notice that the answers to part a and part b are different. The order of multiplication is important.

Application of Matrices to Simultaneous Equations

We can use a similar approach to solve a system of simultaneous equations with two unknowns.

Example 5

Solve the following simultaneous equations using matrix methods:

2x + 3y = 135x + 2y = 16

We can write the system of equations as a matrix equation AX = B as shown below. $\begin{bmatrix} 2 & 3 \\ 5 & 2 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} 13 \\ 16 \end{bmatrix}$

Notice that A is the matrix of the coefficients, X is a column matrix of the pronumerals x and y and B is a column matrix of the values on the right hand side of the equations.

We can now use $X = A^{-1}B$ to solve the simultaneous equations.

ie. $\begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} 2 & 3 \\ 5 & 2 \end{bmatrix}^{-1} \begin{bmatrix} 13 \\ 16 \end{bmatrix}$ Since $\begin{bmatrix} 2 & 3 \\ 5 & 2 \end{bmatrix}^{-1} = \frac{1}{4-15} \begin{bmatrix} 2 & -3 \\ -5 & 2 \end{bmatrix} = \frac{1}{-11} \begin{bmatrix} 2 & -3 \\ -5 & 2 \end{bmatrix}$ $\begin{bmatrix} x \\ y \end{bmatrix} = \frac{1}{-11} \begin{bmatrix} 2 & -3 \\ -5 & 2 \end{bmatrix} \times \begin{bmatrix} 13 \\ 16 \end{bmatrix}$ $\begin{bmatrix} x \\ y \end{bmatrix} = \frac{1}{-11} \begin{bmatrix} -22 \\ -33 \end{bmatrix} = \begin{bmatrix} 2 \\ 3 \end{bmatrix}$ So x = 2 and y = 3.

Or you can use the calculator to solve for *x* and *y*.

Simply enter $\begin{bmatrix} 2 & 3 \\ 5 & 2 \end{bmatrix}^{-1} \times \begin{bmatrix} 13 \\ 16 \end{bmatrix}$ to give $\begin{bmatrix} 2 \\ 3 \end{bmatrix}$ So x = 2 and y = 3

It is worth checking your answers by substituting the values in the original equations.

Statistics

Univariate Data

Measures of Central Tendency

There are 3 common ways of measuring the typical or representative value or centre of a set of data. These averages are called the mean, median and mode.

The Mode

The mode is the value or values that occur most often. You can have more than one mode.

The Median

The median is the middle value when all the observations or outcomes are arranged in order. When there are 2 middle values, find the mean of the 2 middle values. The median is the $\frac{(n+1)}{2}$ th position.

The Mean

The mean (often called the arithmetic average) is the sum of all the observed values divided by the total number of observations. It is written as \bar{x} , and calculated using the formula:

$$\bar{x} = rac{Total \ of \ observed \ values}{Total \ number \ of \ observations}$$

Note that the mean is affected by extreme values or outliers.

Example 1:

Twenty VCE students were asked in a survey to record how much money, in dollars, they earned in a particular week. The data is shown below.

- 0, 0, 25, 0, 260, 15, 30, 10, 0, 13, 0, 0, 0, 20, 30, 80, 20, 79, 5, 40
- a Calculate the median and mean for this sample.
- b Which is the more reasonable measure of centre?

Sorted data: 0, 0, 0, 0, 0, 0, 0, 5, 10, 13, 15, 20, 20, 25, 30, 30, 40, 79, 80, 260 a Since *n* = 20 (even), the median is the mean of the middle two values.

Notice that the median is the $\frac{(n+1)}{2}$ th or $\frac{(20+1)}{2}$ th $= 10\frac{1}{2}$ th position in the list.

Median =
$$\frac{13+15}{2} = 14$$

The mean $\bar{x} = \frac{(0+0+25+0+260+15+30+10+0+13+0+0+0+20+30+80+20+79+5+40)}{20}$

$$\bar{x} = \frac{627}{20} = 31.35$$

b The large value (260) distorts the mean, and the median does not reflect the fact that 7 out of the 20 students earned \$0. In fact, there is no clear centre in this data!

Example 2

The number of siblings for each of 21 students in a Year 12 class was tallied and tabulated as shown. Note that the **frequency** of an observation indicates the number of times the observation occurs. In this example 3 students had 0 siblings, 6 students had 1 sibling etc

Number of Siblings	Frequency
0	3
1	6
2	8
3	1
4	2
5	1

a. Find the modal number of siblings.

b. Find the median number of siblings.

c. Find the mean number of siblings.

a. The modal number of siblings for the class is 2.

b. The median number of siblings is in the $\frac{(n+1)}{2}$ th position = $\frac{(21+1)}{2}$ th position = 11th position

Reading down the frequency column, the median occurs in the 3rd row and will be 2.

Alternatively you could write out the values in order of size and find the median.

0, 0, 0, 1, 1, 1, 1, 1, 1, 2, (2), 2, 2, 2, 2, 2, 2, 3, 4, 4, 5

c. To find the mean number of siblings, you could total the list above and divide by 21, but it is easier to find the mean by extending the table.

Number of Siblings (x)	Frequency (<i>f</i>)	Number of Siblings x Frequency (xf)
0	3	0
1	6	6
2	8	16
3	1	3
4	2	8
5	1	5
Totals	21	38

Mean number of siblings $\bar{x} = \frac{38}{21} = 1.81$

Note: it is easier to enter the data into Lists and Spreadsheets on your calculator to find the mean.

See later

Example 3

The table opposite records the birth weights of 45 babies (to the nearest gram) born in a public hospital in Melbourne in a particular week.

 Estimate the mean weight for these babies (to the nearest gram).

Weight range (g)	Frequency
1500-<2000	2
2000-<2500	4
2500-<3000	4
3000-<3500	18
3500-<4000	16
4000-<4500	1

a It is not possible to add up the exact weights of the babies. However, we can estimate the weight of babies within each class by assuming the weights are all halfway (the midpoint) between the endpoints of interval. For example, we can estimate that the two babies that 1500 + 2000

weighed between 1500 and 2000 g each weighed $\frac{1500 + 2000}{2} = 1750$ g. This number

is used as the typical value (the x) for the purpose of calculating the weighted mean.

Weight range (g)	Frequency (f)	Midpoint (<i>x</i>)	Frequency \times Midpoint (<i>xf</i>)
1500 - < 2000	2	1750	3500
2000 - < 2500	- < 2500 4		9000
2500 - < 3000	2500 - < 3000 4		11000
3000 - < 3500 18		3250	58500
3500 - < 4000	16	3750	60000
4000 - < 4500	1	4250	4250
Totals	45		146250

Mean weight of baby is $\bar{x} = \frac{146250}{45} = 3250g$ (nearest gram)

Note this is an estimate of the mean weight of baby as the actual weights, within each given class is not known.

Note: it is easier to enter the data into Lists and Spreadsheets on your calculator to find the mean.

Cumulative Frequency Graphs

Often it is useful to draw a cumulative frequency curve when data is grouped into classes. The table below shows the results of 40 students in a maths test. We can complete the cumulative frequency column. This column is like a running total of the scores. 3 students obtained a mark of less than 40, 8 students obtained a mark of less than 50, 15 students obtained a mark of less than 60, and so on.

Note that 40 students obtained a mark of less than 100.

Score	Frequency	Cumulative Frequency
30 - < 40	3	3
40 - < 50	5	8
50 - < 60	7	15
60 - < 70	11	26
70 - < 80	8	34
80 - < 90	5	39
90 - < 100	1	40

We can now plot the cumulative frequency graph of the scores on the grid below. The cumulative frequency is always the y-axis and the observations (scores) the x-axis. The first 3 points have been plotted for you. Notice that the graph starts at (30, 0). This indicates that zero students obtained a mark below 30. We then plot the point (40, 3) to indicate that 3 students obtained a mark below 40. The next point is (50, 8) to indicate that 8 students obtained a mark below 50. The next point will be (60, 15) to indicate that 15 student obtained a mark below 60. The other points (70, 26), (80, 34), (90, 39) and (100, 40) are then plotted likewise.

Complete the cumulative frequency curve by plotting the remaining points and joining up the points with a pencil. You can either join the points with a ruler or with a curve. The curve should look like an elongated S.



Mark

Using the cumulative frequency graph above, we can answer questions similar to those that follow:

a. How many students scored 64 marks or less?	
---	--

b. How many students scored 46 marks or less?

- c. If the pass mark was 58%, how many students failed the test? _____
- d. How many students passed the test? ______

The cumulative frequency curve allows us to find a good estimate for the median mark, lower quartile, upper quartile and thus the interquartile range.

e. Find the median score. To do this, find the score corresponding to 20 students on the cumulative frequency axis. 20 students are 50% of the class total.

f. Find the lower quartile score. To do this, find the score corresponding to 10 students on the cumulative frequency axis. 10 students are 25% of the class total.

g. Find the upper quartile score. To do this, find the score corresponding to 30 students on the cumulative frequency axis. 30 students are 75% of the class total.

h. Hence, find the interquartile range. _____

Using the Calculator

This next example will be done using the calculator.

The following table shows the number of daily errors that a machine makes over a 50 day period. Find

- a. The mean number of errors made.
- b. The standard deviation.
- c. The median.
- d. The range
- e. The interquartile range.
- f. Represent the data in a box plot.

Number of Errors per Day	Frequency
0	9
1	18
2	13
3	6
4	3
5	1

- 1. Open Lists and Spreadsheet View.
- 2. Enter the data labelling, each column as shown.

1.1	1	RAD APPRX REAL				
A er	rrors	[■] freq	С	D		
•						
1	0.	9.			G	
2	1.	18.				
3	2.	13.				
4	3.	6.				
5	4.	3.				

 Click Menu – Statistics – Statistics Calculations – One Variable Statistics. Accept 1 list for the first small pop up box that appears. You should then see the One Variable Statistics pop up box below:

One–Variable Statist	ICS
X1 List:	'errors 🗢
Frequency List:	'freq 🗸 🗸
Category List:	▼
Include Categories:	▼
1st Result Column:	c[]
	OK Cancel

In the X1 List box, press the Click button, and select **errors** from the drop down menu. Press TAB to go to the next box. Press the Click button and select **freq** from the drop down menu. TAB to OK and click the OK button.

4. You will now see statistical data in column C and D.

1.1	RAD APPRX REAL					
[■] freq		C	D	E		
•			=OneVar('			
1.	9.	Title	One-Var			
2.	18.	x	1.58			
3.	13.	Σx	79.			
4.	6.	Σx²	197.			
5.	3.	SX := Sn	1.2137			

5. Scroll through the data and write down the appropriate answers.

	Aerrors	Bfreg	С	D	E	a. The mean number of errors
•				=OneVar('errors,'	ı.	$\bar{x} = 1.58$
1	0.	9.	Title	One–Variable S		= b. The standard deviation $sx =$
2	1.	18.	x	1.58		- 1.2137 c The median - MedianX - 1
3	2.	13.	Σx	79.		-d The range = MaxX - MinX = 5 - 0 =
4	З.	6.	Σx²	197.		_ 5
5	4.	З.	SX := Sn-1X	1.2137		-e. The lower quartile Q1 = Q1X = 1
6	5.	1.	σx := σnx	1.2015		and the upper quartile Q3 = Q3X =
7			n	50.		2. The IQR = 2 – 1 = 1
8			MinX	0.		
9			QıX	1.		_ Notice that n gives the total of the
10			MedianX	1.		_ frequency column. Ie the total number
11			Q3X	2.		of days 50.
12			MaxX	5.		_
13			SSX := Σ(x	. 72.18		_

A box plot can now be drawn based on the 5-number summary.



The median and lower quartile are equal in this question.

Measures of Spread of Data

To determine how spread out or dispersed data is we can determine either

- The range (not a good measure as affected by extreme values) The range is the difference between the maximum and minimum value.
- The IQR (Inter Quartile Range) The IQR is the difference between the upper and lower quartile.
- The Standard Deviation (has to be found using the calculator) The standard deviation is represented by sx on the calculator.

The larger the measure of spread, the more spread out or dispersed the data is.

Stem and Leaf Plots

Example 1

The data below shows the weights in kg of 20 possums arranged in order of size:

0.7	0.9	1.1	1.4	1.5	1.6
1.7	1.7	1.8	1.8	1.9	2.0
2.1	2.1	2.2	2.3	2.3	2.5
3.0	3.2				

We can represent this data in a stem and leaf plot as shown below: Key: 0|7 = 0.7kg

Stem	Le	af							
0	7	9							
1	1	4	5	6	7	7	8	8	9
2	0	1	1	2	3	3	5		
3	0	2							

In a stem and leaf plot the numbers are **arranged in order of size**. The key is given as 0|7 kg means stem 0 and leaf 7 which represents 0.7 kg. You should always include a key in the stem and leaf plot.

When preparing a stem and leaf plot keep the number s in neat vertical columns because a neat plot will show the distribution of the scores. It is like a sideways bar chart or histogram.

The interquartile range can be found from the stem and leaf plot.

- 1. Find the **median** weight. The median weight Q2 is the $\frac{(20+1)}{2}$ th score. ie the 10.5th score. The median lies between the 10th and 11th scores. Count through the data to find the position of the median. It can be seen from the plot that the median lies between 1.8 and 1.9. The median weight is $\frac{(1.8+1.9)}{2} = 1.85$ kg.
- 2. The lower quartile Q1 will be the $\frac{(10+1)}{2}$ th score in the lower half. ie the 5.5th score in the lower half. Count through the data to find the position of the lower quartile. Q1 = $\frac{(1.5+1.6)}{2}$ = 1.55 kg.
- 3. The upper quartile Q3 will be the 5.5th score in the upper half of the plot. Count through the data to find the position of the upper quartile. Q3 = $\frac{(2.2 + 2.3)}{2}$ = 2.25 kg
- 4. The interquartile range = Q3 Q1 = 2.25 1.55 = 0.7 kg
- 5. See diagram below:



Example 2

Find the interquartile range of the data presented in the following stem and leaf plot.



The median is the $\frac{(30+1)}{2}$ th score. ie the 15.5th score which lies between 179 and 179. So the median is 179.

The lower quartile Q1 will be the $\frac{(15+1)}{2}$ th score in the lower half. ie the 8th score in the lower half. Q1 = 168.

The upper quartile Q2 will be the 8th score in the upper half of the data. ie 188.

The interquartile range = Q3 - Q1 = 188 - 168 = 20.

See the diagram above.

Using CAS.

You could check your answers by entering the data into your CAS calculator to determine the median, lower and upper quartiles.

Box Plots

Five-number summary

A five number summary is a list consisting of the lowest score (Xmin), lower quartile (Q1), median (Q2), upper quartile (Q3) and the greatest score (Xmax) of a set of data.

A five number summary gives information about the spread or variability of a set of data.

Box Plots

A box plot is a graph of the 5-number summary. It is a powerful way of showing the spread of data. A box plot consists of a central divided box with attached "whiskers". The box spans the interquartile range. The median is marked by a vertical line inside the box. The whiskers indicate the range of scores. **Box plots are always drawn to scale** and a scale is often attached.



Interpreting a Boxplot

A boxplot divides the data into four sections. 25% of the scores lie between the lowest score and the lower quartile, 25% between the lower quartile and the median, 25% between the median and the upper quartile and 25% between the upper quartile and the greatest score.



Extreme Values

Extreme values often make the whiskers appear longer than they should and hence give the appearance that the data is spread over a much greater range than they really are. If an extreme value occurs in a set of data it can be denoted by a small cross on the boxplot. The whisker is then shortened to the next largest or smallest score.

In the figures below, a symmetric distribution is represented in the histogram and in the boxplot.

The characteristics of this boxplot are that the whiskers are about the same length and the median is located about halfway along the box.



The figures below show a negatively skewed distribution. In such a distribution, the data peak to the right on the histogram and trail off to the left.

In corresponding fashion on the boxplot, the bunching of the data to the right means that the left-hand whisker is longer and the right-hand whisker is shorter; that is, the lower 25% of data are sparse and spread out whereas the top 25% of data are bunched up.

The median occurs further towards the right end of the box.





Negatively skewed histogram



In the figures below, we have a positively skewed distribution. In such a distribution, the data peak to the left on the histogram and trail off to the right.

In corresponding fashion on the boxplot, the bunching of the data to the left means that the left-hand whisker is shorter and the right-hand whisker is longer; that is, the upper 25% of data are sparse and spread out whereas the lower 25% of data are bunched up.

The median occurs further towards the left end of the box.



Comparing Two Sets of Data

Back to Back Stem and Leaf Plots

Two sets of data can be compared using back to back stem and leaf plots. The data below shows the life time of a sample of 40 batteries in hours of each of two brands when fitted into a child's toy.

Some of the toys are fitted with an ordinary battery and some with Brand X. Which battery is best?

				(Ordi	nary	Bra	nd	Stem	Bra	and	Х						
							Le	eaf		Lea	əf							
				8	6	2	0	0	6	9								
	9	9	9	8	8	6	4	0	7	3	5							
8	8	7	5	3	1	1	1	0	8	2	4	8						
9	6	6	4	2	2	2	0	0	9	0	1	4	5	5	9			
		8	7	5	3	1	1	1	10	0	0	2	5	8	8	9	9	
							4	2	11	0	0	1	1	3	3	6	7	9
									12	1	4	6	6	6	7	8	8	
									13	3	5							
									14	6								

The spread of each set of data can be seen graphically from the stem and leaf plot. It can be seen that although brand X showed a little more variability than the ordinary brand the batteries generally lasted longer.

Parallel Box Plots

The above data can also be compared by using parallel boxplots. The boxplots share a common scale. Quantitative comparisons can be made between the sets of data.

The 5-Number Summaries of both types of batteries are given below. You can work them out from the stem plots or by using your calculator.

Brand X

Xmin	Lower Quartile Q1	Median Q2	Upper Quartile Q3	Xmax
69	95	109.5	122.5	146

Ordinary Brand

Xmin	Lower Quartile Q1	Median Q2	Upper Quartile Q3	Xmax
60	78.5	87.5	97.5	114

The following parallel boxplots can be drawn to compare the data.

Time in Hours



From the box plots it can be seen that:

- Brand X showed more variability in its performance than the ordinary brand. Brand X range = 77, ordinary brand range = 54. Brand X interquartile range = 27.5 and ordinary brand interquartile range =19.0
- 2. The longest lifetime recorded was that of a Brand X battery of 146 hours
- 3. The shortest lifetime recorded was that of an ordinary battery of 60 hours.
- 4. Brand X battery median lifetime (109.5 hours) was better than that of an ordinary battery (87.5 hours)
- Over one quarter of Brand X batteries were better performers than the best ordinary brand battery (that is, had longer lifetimes than the longest of the ordinary brand batteries' lifetimes)

Bivariate Data

Scatter Plots

We often want to know if there is a relationship between two numerical variables. A scatterplot, which gives a visual display of the relationship between two variables, provides a good starting point.

Below is a gallery of scatterplots showing the various patterns we look for.



Example

The scatterplot at right shows the number of hours people spend at work each week and the number of hours people get to spend on recreational activities during the week.

Decide whether or not a relationship exists between the variables and, if it does, comment on whether it is positive or negative; weak, moderate or strong; and whether or not it has a linear form.



There is a moderate, negative linear relationship between the two variables.

Pearson's Correlation Coefficient (r)

A more precise tool for measuring correlation between two variables is Pearson's product-moment correlation coefficient. This coefficient is used to measure the strength of *linear relationships* between variables.

The symbol for Pearson's product–moment correlation coefficient is r. The value of r ranges from -1 to 1; that is, $-1 \le r \le 1$.

Following is a gallery of scatterplots with the corresponding value of r for each.



The two extreme values of r (1 and -1) are shown in the first two diagrams respectively. From these diagrams we can see that a value of r = 1 or -1 means that there is perfect linear association between the variables.

The value of the Pearson's product–moment correlation coefficient indicates the strength of the linear relationship between two variables. The diagram below gives a rough guide to the strength of the correlation based on the value of r.



The coefficient of determination (r^2)

The coefficient of determination is given by r^2 . It is very easy to calculate — we merely square Pearson's product-moment correlation coefficient (*r*). The value of the coefficient of determination ranges from 0 to 1; that is, $0 \le r^2 \le 1$.

The coefficient of determination is useful when we have two variables which have a linear relationship. It tells us the percentage of variation in one variable which can be explained by the variation in the other variable.

Example.

A set of data giving the number of police traffic patrols on duty and the number of fatalities for the region was recorded and a correlation coefficient of r = -0.8 was found. Calculate the coefficient of determination and interpret its value.

THINK

Calculate the coefficient of determination by squaring the given value of *r*.

2 Interpret your result.

WRITE

Coefficient of determination = r^2 = $(-0.8)^2$ = 0.64

We can conclude from this that 64% of the variation in the number of fatalities can be explained by the variation in the number of police traffic patrols on duty. This means that the number of police traffic patrols on duty is a major factor in predicting the number of fatalities.

Linear Regression

If 2 variables have a moderate or strong association (positive or negative), we can find the equation of the line of best fit of the data and make predictions. The general process of fitting curves to data is called **regression** and the fitted line is called a regression line.

Lines of Best Fit

By Eye

The simplest method is to plot the data on a scatter graph, and place a line over the plot by eye which seems to best represent the pattern in the data values. This method will often give the approximate location of the regression line.

The Least Squares Regression Method

The Least Squares Regression method finds the line that minimizes the total of the squares formed by the points and the line. We normally use the calculator to generate the equation of the least squares regression line. It is given in the form y = mx + b on the calculator.

Example:

The manager of a small ski resort has a problem. He wishes to be able to predict the number of skiers using his resort each weekend in advance so that he can organise additional resort staffing and catering if needed. He knows that good deep snow will attract skiers in big numbers but scant covering is unlikely to attract a crowd. To investigate the situation further he collects the following data over twelve consecutive weekends at his resort.

$Depth \ of \ snow \ (m)$	Number of skiers
0.5	120
0.8	250
2.1	500
3.6	780
1.4	300
1.5	280
1.8	410
2.7	320
3.2	640
2.4	540
2.6	530
1.7	200

You would expect the number of skiers to depend on the depth of snow. The **independent** variable is the depth of snow and **dependent** variable is the number of skiers.

Create a scatterplot of the data. This can be done on the calculator.

1. In Lists and Spreadsheet view, enter the data in the table.

1	1.1 🕨	*Un	saved 🗢		
	A depth	Bskiers	C	D	
٠					
1	0.5	120.			
2	0.8	250.			
3	2.1	500.			
4	3.6	780.			
5	1.4	300.			
6	15	200			
	B13				

2. Hit the Home button and go to Data and Statistics view.

4	1.1	1.2	•	*Unsaved 🗢		
	Са	ption:	depth			
ble						
varia			015			
add 7	0	1 9		●1.7 ●3.6		
< to	0	1.0	-	0.8		
Click		• 21	00.5	2	•27	
	0	2.4		●3.2	0	
			OI.4	4 ck to add variab	10	
	N		Cu	un to aud valiab	ite	

3. Tab to the horizontal axis and select the independent variable **depth** and tab to the vertical axis and select the dependent variable **skiers**. The scatterplot will form.



4. It can be seen that there is a linear, positive strong correlation between the depth of snow and the number of skiers. There is evidence to suggest that as the depth of the snow increases the number of skiers increases.

5. Next find r, the coefficient of correlation and the coefficient of determination r^2 .

Hit Ctrl and Left Arrow to return to Lists and Spreadsheet View. Hit Menu, Statistics, Stat Calculations, Linear Regression (mx + b). Hit the Click button and select depth from the drop down list for X List. Hit tab and select skiers from the drop down list for the Y List.

Linear Regression (n	nx+b)
× List:	'depth
Y List:	'skiers
Save RegEqn to:	f1
Frequency List:	1
Category List:) ×
Include Categories:	
	OK Cancel

◀ 1.1	1.2	*Un	saved 🗢	4	×
A d	epth	skiers	С	D	
•				=LinRegM	
-	0.8	250.	Keg⊵qn	m*x+b	
3	2.1	500.	m	186.418	
4	3.6	780.	b	28.3373	
5	1.4	300.	r²	0.781492	
6	1.5	280.	r	0.88402	
7	1.8	410.	Resid	{-1.54620	V
D1	="Lines	ar Regressi	on (mx+b))"	•

There is no need to enter data into the other boxes. Tab to OK and hit Enter.

The coefficient of correlation r = 0.88402

This indicates that there is a strong, positive correlation between the depth of snow and the number of skiers.

The **coefficient of determination**, r^2 is 0.781492

We can say that 78% of the variation in the number of skiers can be explained by the variation in the depth of snow.

The data also gives us the line of best fit, the least squares regression equation.

y = 186.418x + 28.3373

We can write this more clearly as

Number of skiers = $186.418 \times \text{depth of snow in } m + 28.3373$

6. The equation of the least squares regression line can also be determined in Data and Statistics view. Hit Ctrl + right arrow to return to your scatterplot. Hit Menu, Analyse, Regression, Show Linear (mx + b)



The least squares regression equation is y = 186.418x + 28.3373

Interpreting the Gradient and y-intercept

The **gradient** 186.418 indicates that for every 1 metre increase in depth of snow the number of skiers increases by 186.

The **y-intercept** 28.3373 indicates that if the depth of snow is 0, there would be 28 skiers attending the resort.

Using the Least Squares Regression Equation to make Predictions

Suppose we want to estimate the number of skiers when the depth of snow is 3m. Using

Number of skiers = $186.418 \times \text{depth of snow} + 28.3373$

Number of skiers = $186.418 \times 3 + 28.3373 = 587.5913$

That is 588 skiers

This result is reliable because we have **interpolated**. 3m lies within the bounds of the depth of snow given in the table. That is it is between 0.5 and 3.6m.

Suppose we want to estimate the number of skiers when the depth of snow is 4m.

Number of skiers = $186.418 \times 4 + 28.3373 = 774$

That is 774 skiers.

The result is unreliable because we have **extrapolated**. 4m lies outside the bounds of the depth of snow given in the table. It is outside the range of 0.5 to 3.6m.

Financial Maths

Simple Interest

When you borrow money or invest money over a certain period of time you would expect to pay interest or receive interest.

Simple Interest or Flat Rate of Interest is paid only on the original amount borrowed or invested.

Simple Interest
$$I = \frac{PrT}{100}$$

I = interest (\$),

P = principal (\$), the sum of money borrowed or invested

r = rate of interest (% per annum) or flat rate of interest (% per annum)

T = term of interest (years)

Note that you may be given I and will have to determine P, r, or T. You can either use Solve on your calculator or rearrange the formula as shown below:

$$\mathbf{P} = \frac{100I}{rT} \qquad \mathbf{r} = \frac{100I}{PT} \qquad \mathbf{T} = \frac{100I}{Pr}$$

The total amount owing or invested after the period of time is given by: A = P + I

A = total amount at the end of the term (\$), P = principal (\$) and I = simple interest (\$).

When calculating simple interest the interest earned is the **same** for each time period.

Simple interest is represented by a **straight line** graph with a positive gradient.



Hire Purchase

Under a hire purchase agreement, the principal and interest of the loan taken must be repaid with a certain number of instalments. Often a deposit is paid and interest is calculated on the outstanding balance or principal. Hire purchase uses the flat rate of interest or simple interest, given by the formula

$$I = \frac{PrT}{100}$$

- or interest = instalment amount x number of instalments principal
- or interest = total amount paid purchase price
- Also total amount paid = purchase price + interest
- or total amount paid = deposit + instalment amount × number of instalments.

Effective Rate of Interest

The flat rate of interest, r %, is misleading as it is charged on the total amount owed over the whole period of repayment even though the principal owing is reduced by each instalment.

The **Effective Rate** of Interest or **Real Rate** of Interest is a better measure of the rate of interest paid as it is based on the average amount owing. It is given by:

$$r_e = \frac{2n \times r}{n+1}$$

 r_e = real or effective rate of interest (% pa)

r= flat rate of interest (% pa)

n = total number of instalments to be made.

$r_e pprox 2r$ The effective rate of interest is around twice the flat rate of interest.

Interest Earned on Bank Balances

Minimum Monthly Balance

Banks often base the interest earned on bank account balances on the minimum monthly balance. The interest is calculated using the simple interest formula for 1 month, making the **assumption** that

each month has an equal number of days. Take T as $\frac{1}{12}$ year in the formula: $I = \frac{PrT}{100}$.

Daily Balance

Interest is sometimes based on a daily balance. The daily balance method will always produce more interest than the minimum monthly balance method.

Compound Interest

When using compound interest, the interest is calculated on the interest as well as the principal over each time period. The interest gains interest as well. The result is that the balance of the account increases at regular intervals and so does the interest earned.

 $A = P(1 + \frac{r}{100})^n \text{ or}$ $A = PR^n, where R = \left(1 + \frac{r}{100}\right), \text{ R is called the growth factor or rate of appreciation.}$

A = amount at the end of n compounding periods (\$)

P = principal (\$)

r = rate of interest per period (%)

n = number of compounding periods.

Remember: r is the rate of interest per period, not per annum and n is the number of compounding periods, not years.

Compound Interest is represented by an exponential graph, **not** a straight line.

The compound interest formula can be used for compound growth.

When solving compound interest questions, you may be given the amount and have to calculate one of the other variables. You can use **Solve** on the calculator. Also you can use **Finance Solver** where **Pmt** is set to zero.



Comparing Simple and Compound Interest

If the same principal \$1000 is invested for 20 years, interest earned at the end of each year, then the growth rate of simple interest and compound interest can easily be compared by plotting the graphs on the same axes. The straight line is simple interest and the curve is compound growth. Clearly it is wise to ensure that your investment is compounded!



Depreciation

Depreciation means an item reduces in value from the time that it is bought. The original price of an item is called the **purchase price** of the item. The decrease in value of an item is called the **depreciation**. The value of an item at any given time is called the **book value**.

Book Value = purchase price – depreciation

Scrap value is the value that an item is said to have at the end of its useful life.

There are 3 ways to calculate depreciation:

Straight Line Depreciation or Flat Rate Depreciation or Constant Depreciation Method

The item or asset loses the **same amount** each year. It is calculated by the simple interest method.

Book Value = purchase price – depreciation

$$\mathsf{B.V} = \mathsf{P} - \frac{PrT}{100}$$

or B. V = P - dT, where d is the annual depreciation in \$

B.V = book value of asset or item (\$)

P = purchase price (\$)

r = depreciation rate (% per annum)

T = period of depreciation (years)

Note: the formula for calculating the depreciation is similar to the formula for simple interest.

The graph of the book value over time is a **straight** line with negative gradient.



Reducing Balance Depreciation

Compared with straight line depreciation in which the item's value decreases by the same amount each time period, reducing balance depreciation decreases by a different amount each year, as the book value decreases by a percentage of the previous book value.

The formula for the book value of an item that depreciates under the reducing balance method is:

$$B.V = P(1 - \frac{r}{100})^n = P \times R^n$$
, where $R = 1 - \frac{r}{100}$ is called the rate of decay.

B.V = book value of item (\$)

P = purchase price (\$)

- r = annual rate of depreciation (% per annum)
- n = number of depreciation periods (years)

Notice that the formula for calculating the book value is similar to the formula for calculating compound interest.

The graph of reducing balance depreciation is an exponential curve. As time increases the book value decreases by different amounts.



Unit Cost Depreciation.

Unit cost depreciation applies to machinery that performs a repetitive job. The method relates depreciation to use rather than to time. The rate of depreciation is determined from a description of unit cost.

Book value = purchase price – unit cost x amount of use

REMEMBER

- 1. Unit cost depreciation is based on how much an item is used.
- 2. Current book value (\$) = previous book value (\$) amount of depreciation (\$)
- 3. Amount of depreciation (\$) = amount of use \times rate of depreciation (\$ per use)
- 4. Rate of depreciation (\$ per use) = $\frac{\text{amount of depreciation ($)}}{\text{mount of depreciation ($)}}$

amount of use