

# MAKING REGRESSION ANALYSIS EASY USING A CASIO SCIENTIFIC CALCULATOR

Astrid Scheiber  
astrids@jamesralph.com  
JAMES RALPH / CASIO

*Adequate knowledge of calculator skills makes the teaching of Statistics to Grade 12 learners easier and enables the educator to assist their learners more efficiently. This workshop will guide you through Linear Regression Analysis, including finding relationships between variables, the line of best fit and making projections, using the CASIO Scientific calculator.*

*Equipment required: CASIO FX-82ZA PLUS Scientific Calculator.*

**Target audience:** Further Education & Training – Mathematics

**Duration:** 2 hours

**Maximum number of participants:** 30

**Motivation:** As of 2014, the Grade 12 Statistics syllabus involves the learners' making use of available technology to: **[12.10.1 (b)] calculate the linear regression line which best fits a given set of bivariate numerical data** and **[12.10.1 (c)] calculate the correlation co-efficient of a set of bivariate numerical data**. As stated by the current Maths CAPS document. This workshop serves to increase educators' understanding of the CASIO Scientific calculator. In turn, it will foster self confidence and a positive attitude towards Statistics, enhancing both the educators' and learners' understanding of the topic.

**Content:** This workshop will cover: Identifying the relationship between bivariate numerical data, inputting bivariate data into the CASIO scientific calculator, calculating the correlation co-efficient, finding the equation of the regression line, calculating projected values - Interpolation & Extrapolation, using TABLE MODE to find the co-ordinates of the line of best fit & selecting random samples.

Introduction	5 mins
Identifying the relationship between bivariate numerical data	10 mins
Inputting bivariate data into the CASIO scientific calculator	10 mins
Calculating the correlation co-efficient	10 mins
Finding the equation of the regression line	15 mins
Calculating projected values - Interpolation & Extrapolation	15 mins
Using TABLE MODE to find the co-ordinates of the line of best fit	15 mins
Selecting random samples	10 mins
Discussion & worked example	30 mins

## WORKSHEET:

When we investigate statistical information, we often find there are connections between sets of data. When working with **two variables** ( $x ; y$ ) this is considered working with bivariate data.

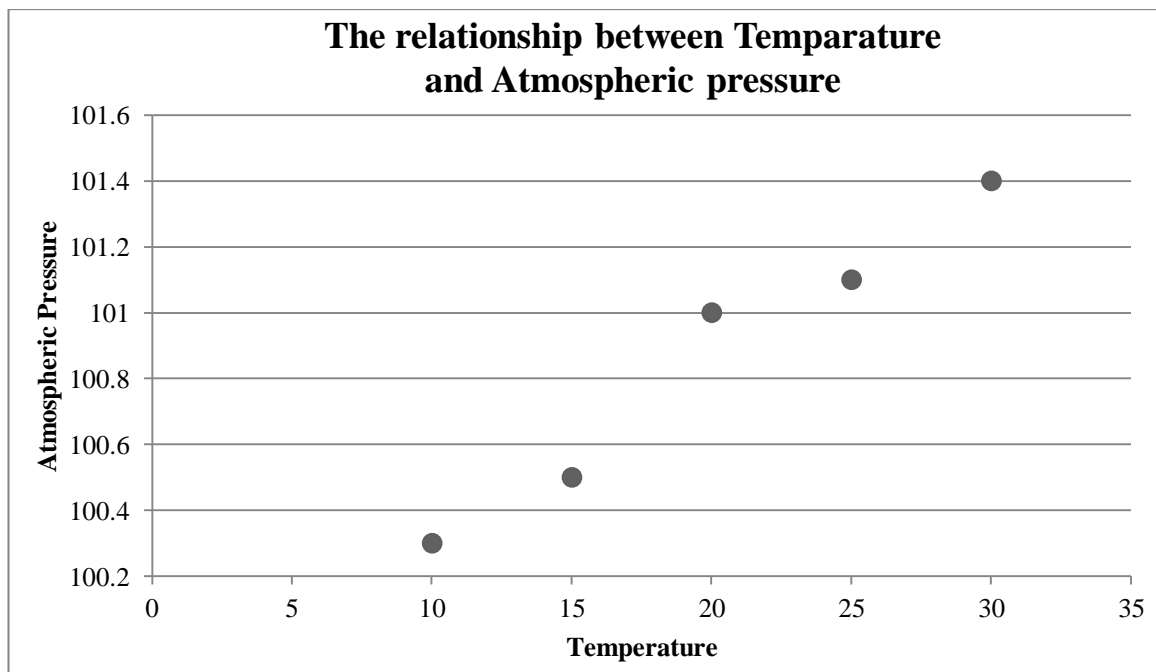
Consider the following table:

Temperature (°C)	Atmospheric pressure (kPa)
10	100,3
15	100,5
20	101,0
25	101,1
30	101,4

### Pressure is dependent on Temperature

Hence: temperature is the  $x$  variable and pressure the  $y$  variable.

We can represent this data by means of a **SCATTER PLOT**. If the data shows a pattern, we can identify whether it forms a linear function or some other type of function such as a Quadratic or Exponential function.



It is clear from the graph that the points tend to form a pattern which resembles a **straight line**

## LINEAR REGRESSION

**predicts** a relationship between a dependent variable ( $y$ ) & an independent variable ( $x$ ), where the relationship approaches that of a straight line.

$$y = A + Bx$$

Using the CASIO Scientific Calculator:

[MODE] [2: STAT]

Key	Menu Item	Explanation
1.	1-VAR	Single variable / Data handling
2.	A + BX	Linear regression
3.	$\_ + CX^2$	Quadratic regression
4.	ln X	Logarithmic regression
5.	$e ^ X$	Exponential regression
6.	$A . B ^ X$	AB exponential regression
7.	$A . X ^ B$	Power regression
8.	1/X	Inverse regression

### LINEAR REGRESSION [2: A+BX]

Enter the data into the double variable table

Input the data into the cell where the cursor is located

Use [=] to enter the data items

Input  $x$ -values first and then  $y$ -values

Use the [REPLAY] arrows to move the cursor to the  $y$ -column

	$x$	$y$
1	10 [=]	100,3 [=]
2	15 [=]	100,5 [=]
3	20 [=]	101 [=]
4	25 [=]	101,1 [=]
5	30 [=]	101,4 [=]

Clear the screen - ready for the STAT sub menu

[AC]

[SHIFT] [1] (STAT)

## STAT Linear Regression sub menu

Key	Menu Item	Explanation
5: Reg	1. A	Regression co-efficient of A
	2. B	Regression co-efficient of B
	3. r	Correlation co-efficient r
	4. $\hat{x}$	Estimated value of x
	5. $\hat{y}$	Estimated value of y

**Correlation** is an indication of the relationship between the two variables.

**Correlation co-efficient (r)** tells us the strength and direction of the correlation.

**r** lies between -1 and +1 (  $-1 \leq r \leq 1$  )

- If r is close to 0, then there is a **weak linear** relationship
- If r is close to -1 or +1, there is a **strong linear** relationship

The sign of r indicates whether the data has a positive or negative correlation (sloping line of best fit)

- Positive correlation  
As one quantity increases, the other one increases  
As one quantity decreases, the other one decreases
- Negative correlation  
As one quantity increases, the other one decreases  
As one quantity decreases, the other one increases

### CALCULATE THE CORRELATION CO-EFFICIENT

[SHIFT] [1] [5: Reg] [3: r] [=]

**r** = .....

**r** is very close to ....., hence there is a .....  
**correlation** between temperature and atmospheric pressure.

Once it is determined that  $-0,7 < r < -1$  OR  $0,7 < r < 1$  we can calculate the **linear regression line** OR the **line of best fit**, which helps us to predict future values.

$$y = A + Bx$$

where **A** is the **y-intercept** and **B** is the **gradient/slope**

**CALCULATE THE VALUE OF A**

[SHIFT] [1] [5: Reg] [1: A] [=]

$$A = \dots\dots\dots$$

**CALCULATE THE VALUE OF B**

[SHIFT] [1] [5: Reg] [2: B] [=]

$$B = \dots\dots\dots$$

So the equation of the line of best fit is

$$y = \dots\dots\dots + \dots\dots\dots x$$

### PROJECTIONS

#### CALCULATOR RULE:

**Step 1:** Input what is given

**Step 2:** Select which variable is required

**Interpolation:** the value predicted lies **within** the domain and range of the data set.

- *What is the approximate atmospheric pressure when the temperature is 18°C.*

18 [SHIFT] [1] [5: Reg] [5:  $\hat{y}$ ] [=]

$$\hat{y} = \dots\dots\dots$$

The pressure is ..... kPa when the temperature is 18 °C

**Extrapolation:** the value predicted lies **outside** the domain and range of the data set.

- *What is the approximate temperature if the atmospheric pressure is 100 kPa?*

**100 [SHIFT] [1] [5: Reg] [4:  $\hat{x}$ ] [=]**

$\hat{x}$  = .....

The temperature is ..... °C when the pressure is 100 kPa

### TABLE MODE

Using TABLE MODE you can find the **co-ordinates** to plot the **line of best fit**.

**[MODE] [3: TABLE]**

Enter the equation of the line of best fit

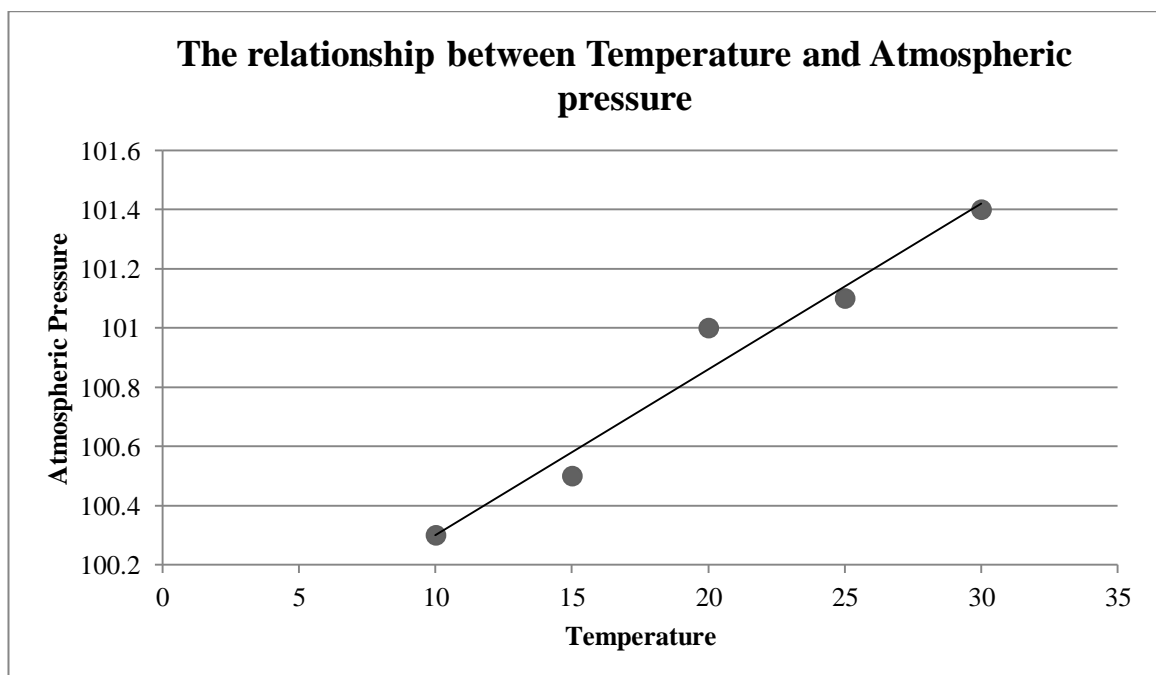
Input a START  $x$ -value of 10

Input a END  $x$ -value of 30

Input STEPS (INTERVALS) of 5

**The co-ordinates to plot are:**

( 10 ; 100.3 ) ( 15 ; 100.58 ) ( 20 ; 100.86 ) ( 25 ; 101.14 ) ( 30 ; 101.42 )

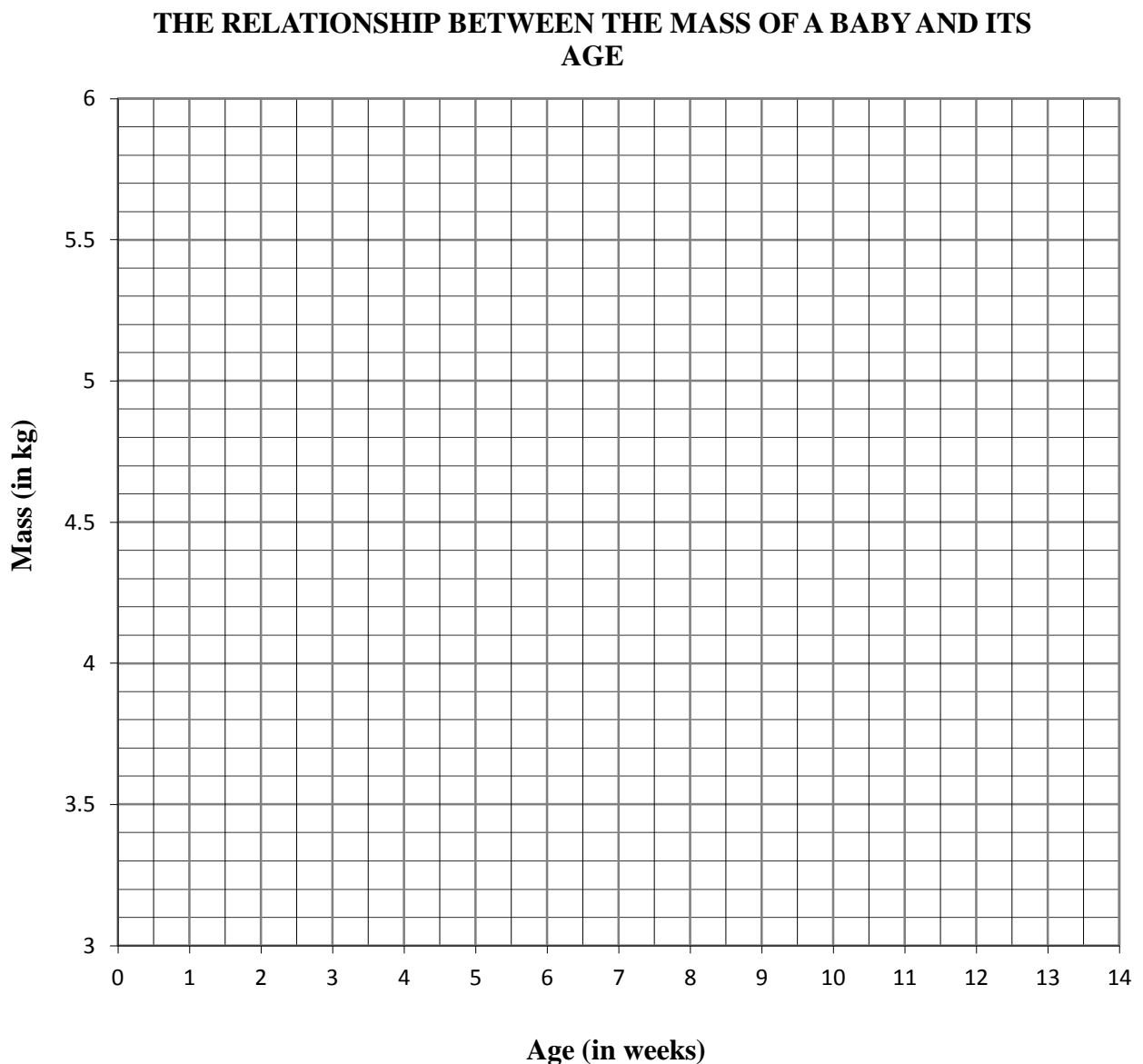


## **LINEAR REGRESSION ACTIVITY**

The following table shows the age and mass of ten babies:

<b>Age (in weeks)</b>	4	5	12	10	6	1	9	13	2	7
<b>Mass (in kg)</b>	3,3	3,8	5,5	5	4	3	4,2	5	3,5	4,7

a) Use the information to draw a scatter plot.



b) What pattern do the points form?

- c) Using the calculator: find the correlation coefficient – efficient (rounded off to 3 decimal places)

$r$  lies between 0,7 and 1, hence there is a:  
 (select) WEAK / STRONG and a POSITIVE / NEGATIVE  
 correlation between the data.

### How to set your calculator to round off to 3 decimal places

<pre>1:MthIO 2:LineIO 3:Deg   4:Rad 5:Gra   6:Fix 7:Sci   8:Norm</pre>	<p><b>Key Sequence:</b></p> <p><b>SHIFT MODE 6</b></p> <p>Now select decimal places</p> <p><b>3</b></p>	<pre>Fix 0~9?</pre>
--	---	---------------------

- d) Using the calculator: find the equation of the line of best fit.

Mrs van der Merwe's baby is 10,5 weeks old.

- e) Using the calculator: find the estimated mass of her baby.

- f) Using TABLE MODE plot the line of best fit on the given axis.

Is the estimated value for the mass of Mrs van der Merwe's baby similar when read off the graph?

### How to clear your calculator from rounding off

<pre>1:MthIO 2:LineIO 3:Deg   4:Rad 5:Gra   6:Fix 7:Sci   8:Norm</pre>	<p><b>Keys Sequence:</b></p> <p><b>SHIFT MODE 8</b></p> <p>Select</p> <p><b>2</b></p>	<pre>Norm 1~2?</pre>
<p>Norm 1 is the <b>default setting</b> and gives answers in scientific notation.          e.g. <math>1 \div 50\,000 = 2 \times 10^{-5}</math></p> <p>Norm 2 is <b>generally preferred</b> as answers are only expressed in scientific notation when they are too big to fit on the screen.          e.g. <math>1 \div 50\,000 = 0.00002</math></p>		



## RANDOM INTEGERS

The simplest way to choose a random sample of numbers is to let the calculator do it for you.

Select a random sample of 6 numbers, between 1 and 49, to play the lotto:

**[ALPHA] [.] (RanInt) 1 [SHIFT] [D] (,) 49 [D] [=]**

### **\*NOTE\***

- every time you use one of these key sequences, you will get a different string of numbers
- Integers are repeated
- This key sequence can be used to flip a coin (1,2)
- This key sequence can be used to throw a die (1,6)

## **HINTS TO MAKE THE TEACHING OF THE CALCULATOR EASIER**

- If you are introducing calculator work to a new class it is easier if all learners have the same calculator.
- Always keep the instruction booklet that you receive with your calculator so that you can refer to it whenever you are not sure how to do a calculation.
- There are 3 modes on the CASIO Scientific Calculator fx-82ZA PLUS, always make sure that your calculator is in the right mode before you begin.

### **[MODE]**

1. Computational – normal scientific calculations
2. Statistics – data handling & regression
3. Table – graph work & functions

**Calculators play a vital role in the classroom:  
not by *substituting* Mathematics, but by *supplementing* our subject.  
It's conventional Mathematics by new methods.**

### **REFERENCES:**

RADMASTE Centre, ACE – Data Handling and Probability FET (2010)  
University of the Witwatersrand, SA.

MARC ANCILLOTTI, Data Handling – Scatter plots of bivariate data.