

MAKING REGRESSION ANALYSIS EASY USING A CASIO SCIENTIFIC CALCULATOR

Astrid Scheiber
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.

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.

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.

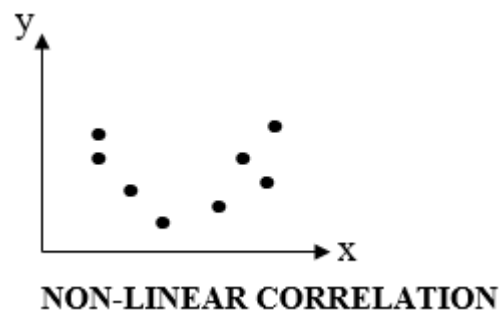
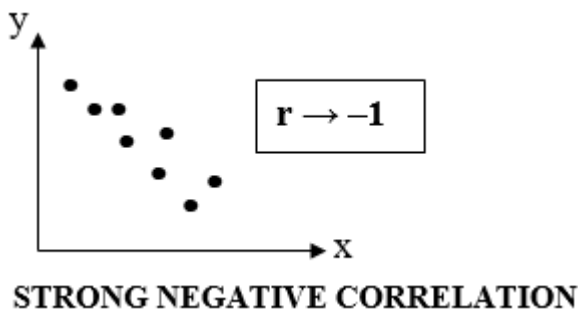
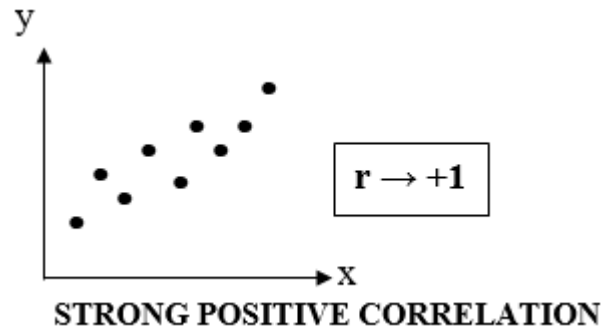
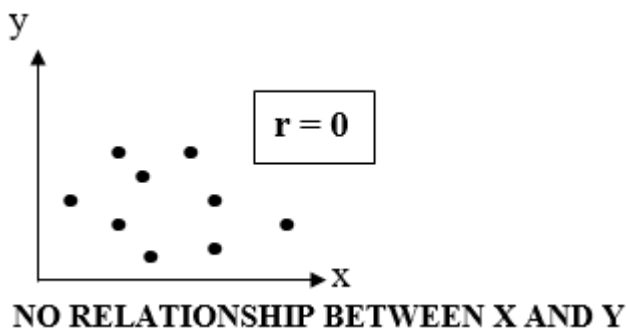
Linear Regression predicts a relationship between a dependent variable (y) and an independent variable (x), where the relationship approaches that of a straight line.

$$y = A + Bx$$

Correlation Co-efficient (r) is the measure of the strength of the relationship between the variables. It tells us the strength and direction of the correlation.

$$-1 \leq r \leq 1$$

Scatter Plots showing Correlation:



- 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
OR as one quantity decreases, the other one decreases
- **Negative correlation:** As one quantity increases, the other one decreases
OR as one quantity decreases, the other one increases

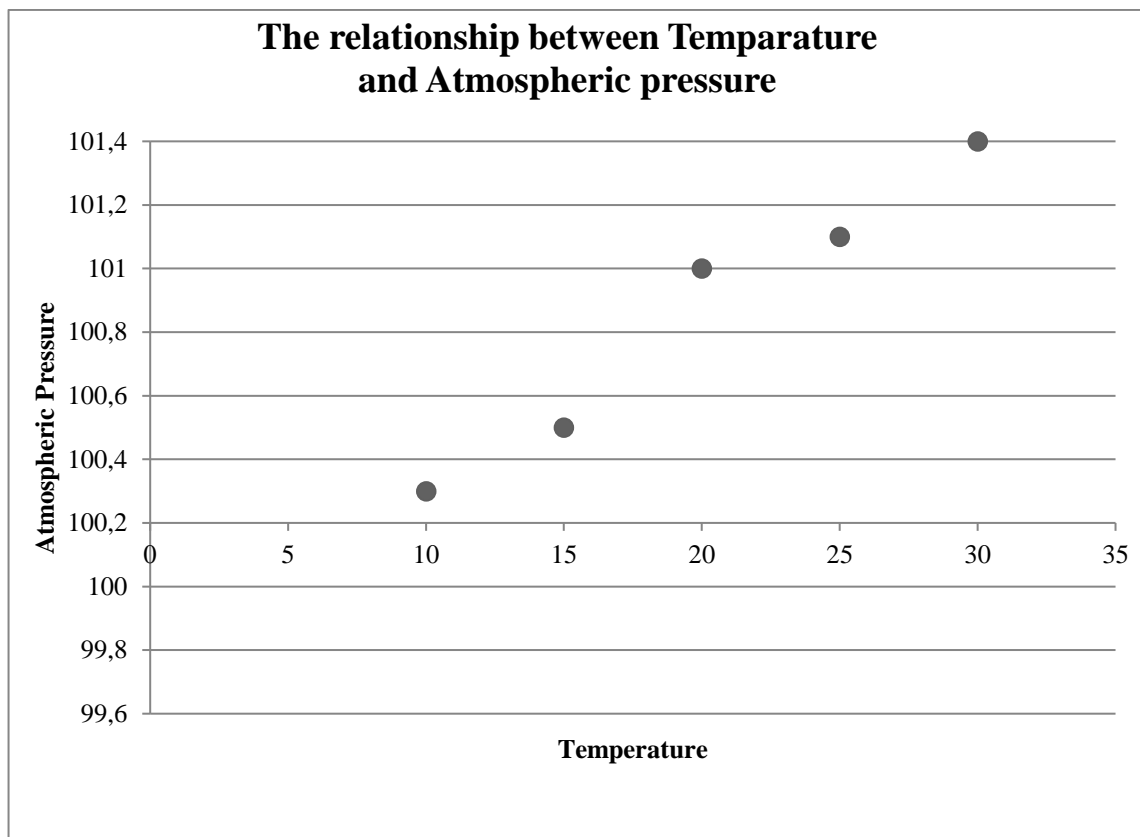
Consider the following table:

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**

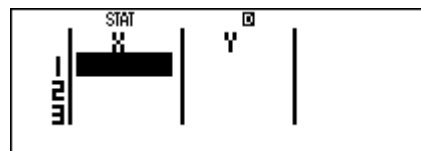
Using the CASIO Scientific Calculator:

MODE 2: STATISTICS

1: 1-VAR	2: A+BX
3: ax^2+bx+c	4: $\ln x$
5: e^x	6: $A \cdot B^x$
7: $A \cdot x^B$	8: $1/x$

1. Single variable / Data handling
2. **Linear regression**
3. Quadratic regression
4. Logarithmic regression
5. Exponential regression
6. AB exponential regression
7. Power regression
8. Inverse regression

Enter STAT MODE for Bivariate data 2



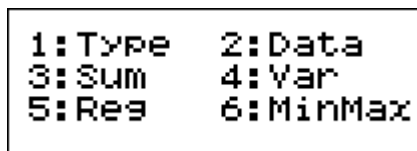
Enter the data into the table: Input x-values, use the arrows to move the cursor to the y-column & then input y-values

1 0 = 1 5 = 2 0 = 2 5 = 3 0 = ▼ ▶ 1 0 0 . 3 = 1
0 0 . 5 = 1 0 1 = 1 0 1 . 1 = 1 0 1 . 4 =



Clear the screen - ready for the **Regression sub menu**

AC SHIFT 1



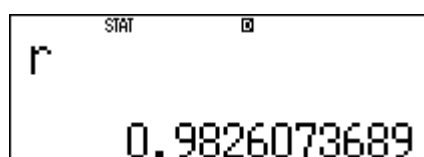
5

Breakdown of **Regression sub menu**

Key	Menu Item	Explanation
5: Reg	<div style="border: 1px solid black; padding: 5px; display: inline-block;"> 1:A 2:B 3:r 4:x 5:y </div>	1. Regression co-efficient of A 2. Regression co-efficient of B 3. Correlation co-efficient r 4. Estimated value of x 5. Estimated value of y

1. CALCULATE THE CORRELATION CO-EFFICIENT

3 =



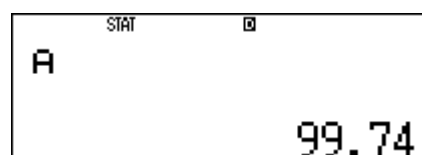
There is a **strong, positive** linear correlation between temperature and atmospheric pressure.

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

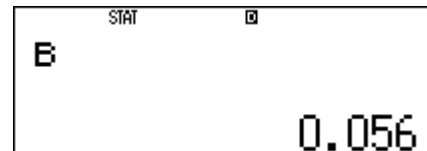
$y = A + Bx$ where **A** is the **y-intercept** and **B** is the **gradient/slope**

2. CALCULATE THE VALUES OF A & B

• Calculate A AC SHIFT 1 5 1 =

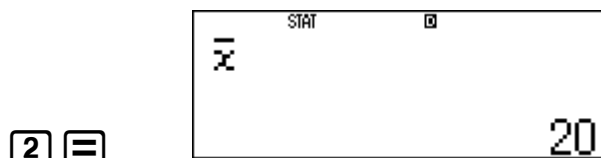
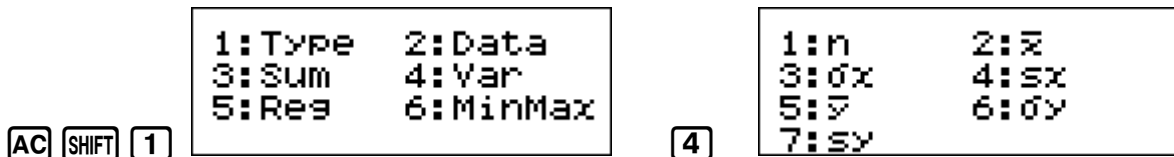


- Calculate B $\boxed{\text{AC}} \boxed{\text{SHIFT}} \boxed{1} \boxed{5} \boxed{2} \boxed{=}$

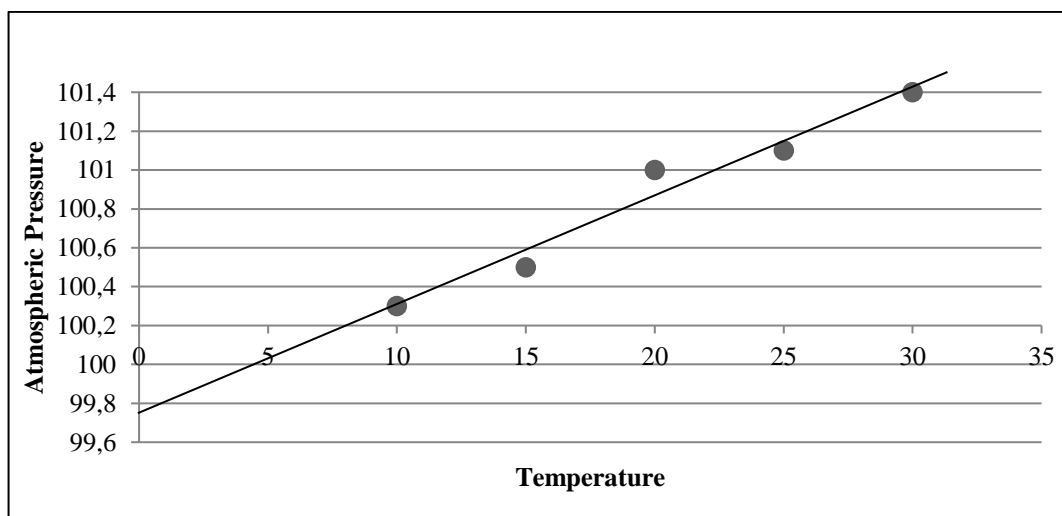


So the line of best fit is: $y = 0,056x + 99,74$

3. FIND A SECOND POINT TO PLOT THE LINE OF BEST FIT



Once you plot the linear equation you can then make projections.

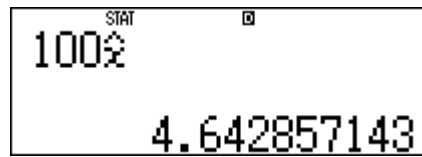


HOW TO MAKE PROJECTIONS ON THE CALCULATOR

- RULE:**
- Step 1: Input what is given
 - Step 2: Regression sub menu select which variable is required

A. What is the approximate temperature if the atmospheric pressure is 100 kPa.

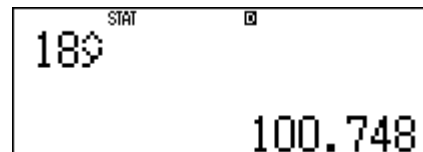
AC 1 0 0 SHIFT 1 5 4 =



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

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

AC 1 8 SHIFT 1 5 5 =



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

LIKE US ON

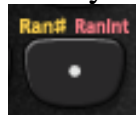


CASIO CALCULATORS SOUTH AFRICA

VISIT OUR WEBSITE FOR
RESOURCES

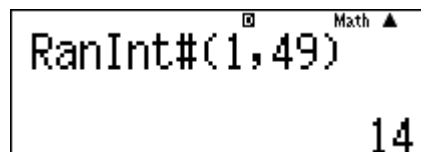
www.casiocalcs.co.za

RANDOM INTEGERS – The simplest way to choose a random sample of numbers



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

ALPHA . 1 SHIFT) 4 9) =



NOTE

- The calculator will give a different string of numbers each time
- Integers are repeated
- This function can be used to flip a coin (1,2)
- This function can be used to throw a die (1,6)

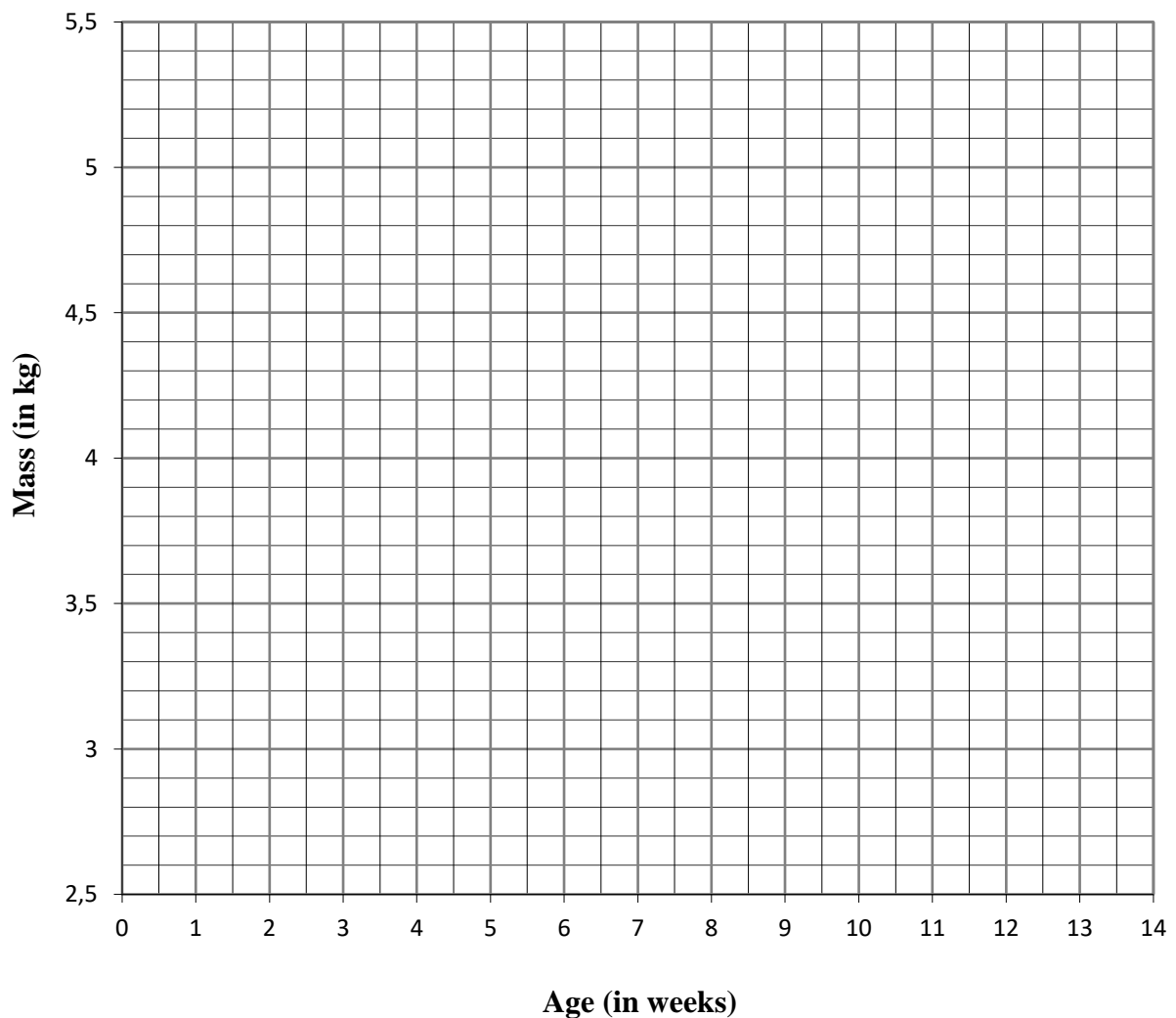
LINEAR REGRESSION ACTIVITY

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

Age (in weeks)	4	5	12	10	6	1	9	13	2	7
Mass (in kg)	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.

THE RELATIONSHIP BETWEEN THE MASS OF A BABY AND ITS AGE



b) What pattern do the points form?

c) Find the correlation coefficient (rounded off to 3 decimal places)

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>Key Sequence:</p> <p>SHIFT MODE 6</p> <p>Now select decimal places</p> <p>3</p>	<p>Fix 0~9?</p>
--	---	-----------------

d) How would you describe the relationship between the data?

e) Find the equation of the line of best fit.

f) Plot the line of best fit on the scatter plot.

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

g) Read the estimated value for the mass of Mrs van der Merwe's baby off the graph and note it down

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

i) Are the 2 values similar?

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>Keys Sequence:</p> <p>SHIFT MODE 8</p> <p>Select</p> <p>2</p>	<p>Norm 1~2?</p>
--	---	------------------

Norm 1 is the **default setting** and gives answers in scientific notation.

e.g. $1 \div 50\,000 = 2 \times 10^{-5}$

Norm 2 is **generally preferred** as answers are only expressed in scientific notation when they are too big to fit on the screen.

e.g. $1 \div 50\,000 = 0.00002$

**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.