

Study Guide-Unit 4B- Statistics

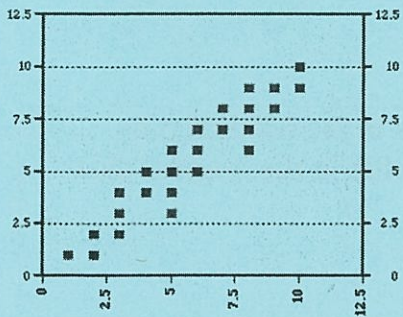
1.	Correlation	A relationship between two events.
2.	Causation	One set of data is causing the affect that is occurring in the other set of data
3.	Line of Best Fit	Line that runs through the majority of the data points (Best trend line, Regression Line)
4.	Interpolation	Looking for values that fall inside (within) the given data
5.	Extrapolation	Looking for values that fall outside the given data. The further away from the plotted values you go, the less reliable your prediction is.
6.	Residuals	Residual = Actual Value – Predicted Value (“R = A – P”)

CORRELATION COEFFICIENT:

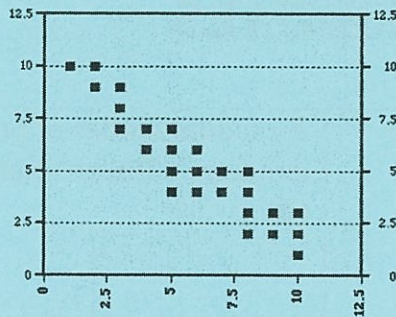
Correlation coefficient is the “r-value” that measures how close the data values are to the line of best fit.
 $(-1 \leq r \leq 1)$

To find the correlation coefficient in the calculator, you must turn your Stat Diagnostics ON!
 Turn the diagnostics on: MODE → Stat Diagnostics On → Enter 2 times

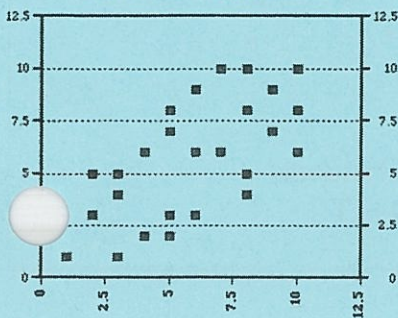
High Positive Correlation



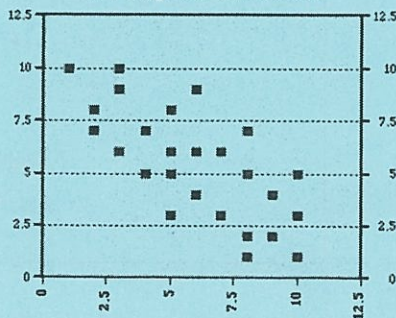
High Negative Correlation



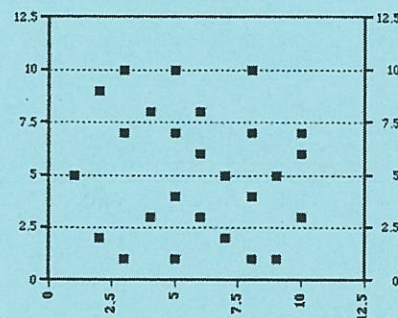
Low Positive Correlation



Low Negative Correlation



No Correlation



Closer to 1 = positive/strong correlation
 Closer to -1 = negative/strong correlation
 Close to 0 = no correlation

CALCULATOR STEPS FOR LINEAR REGRESSION

Step 1- Hit the “STAT” button on your calculator.

Step 2- Click “EDIT”

Step 3- Under L₁, list all your x values

Step 4- Hit the right arrow

Step 5- Under L₂, list all your corresponding y- values

Step 6- Click STAT, then right arrow so that it’s highlighted over “CALC”

Step 7- Click “4” for “Lin Reg”

Step 8- Make sure your screen looks like this:

LinReg(ax+b)

Xlist:L1

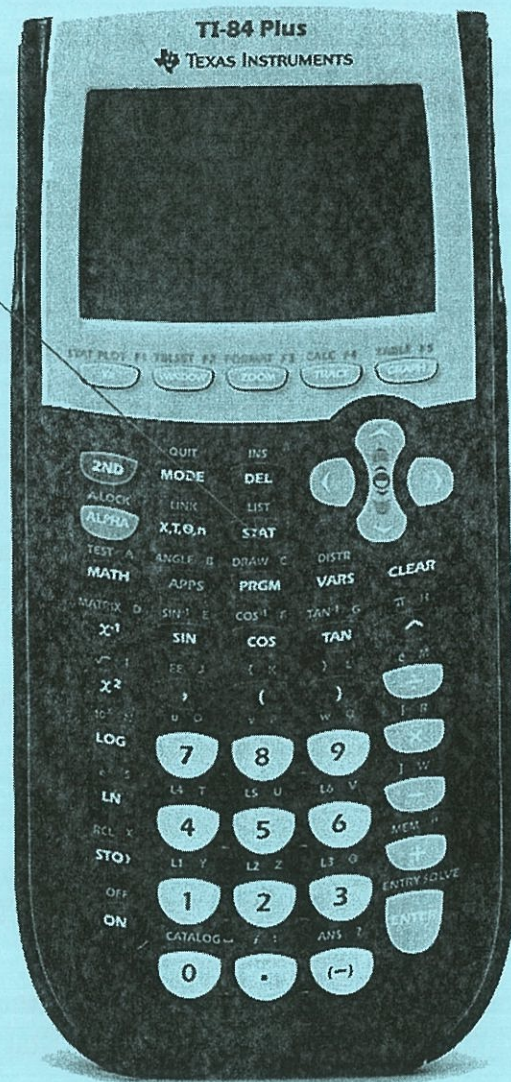
Ylist:L2

FreqList:

Store RegEQ:

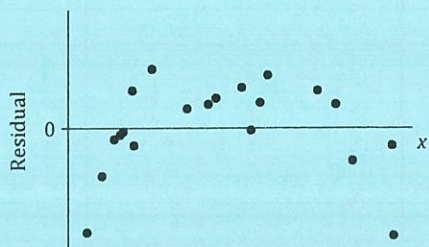
Calculate

Step 9- Hit enter on “Calculate” to get you’re a and b values for $y = ax + b$.



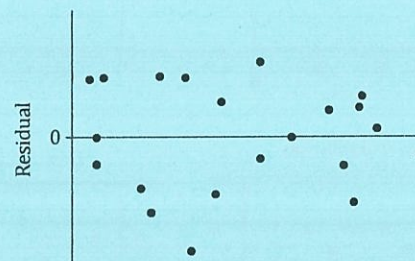
RESIDUALS

A curve or pattern in the residual plot indicates a **curved (non-linear)** relationship in the original data set.



Non-linear (Curved)
NOT A GOOD FIT!!

A random scatter of points in the residual plot indicates a **linear** relationship in the original data set.



Linear (Random)
GOOD FIT!!