Line of Best Fit & Mean Fit Line

Definition:

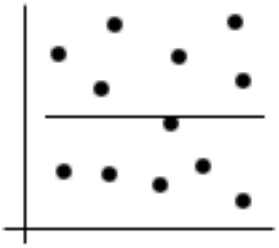
THE LINE OF BEST FIT is a straight line through a scatter plot of data that indicates the pattern and design of the scatter plot. The line should have roughly an equal number of points on each side and should spread out the points, rather than concentrating them at one end.

Note:

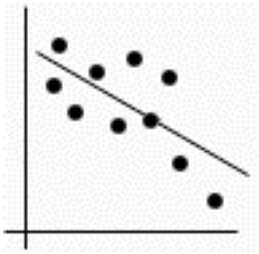
The line of best fit shows a trend in the data and can help you make predictions about the data.

Examples:

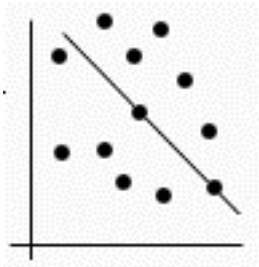
Recall, CORRELATION is the apparent relationship between the variables in a set of data. Here are some examples of different scatter plots and the line of best fit. Note that these scatter plots are for demonstration purposes only. Your graphs should be properly labelled, including a title.



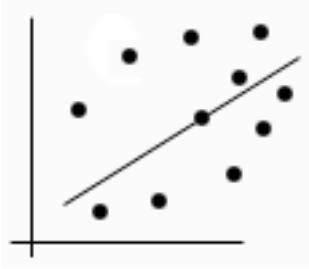
No Correlation



Strong Negative Correlation



Weak Negative Correlation



Weak Positive Correlation

Predictions:

The line of best fit can be used to make predictions for values not actually recorded or plotted. There are two methods in which this can be done:

INTERPOLATE: To calculate or estimate a value between known values.

EXTRAPOLATE: To calculate or estimate a value by following a pattern and going beyond values already known.

Example 1:

Here is an example of extrapolation. In an experiment on stretching, masses were added to a spring and the following measurements were obtained.

Weight added (g)	20	40	60	80	100	120	140
Length of Spring (cm)	14	21	28	35	41	46	50

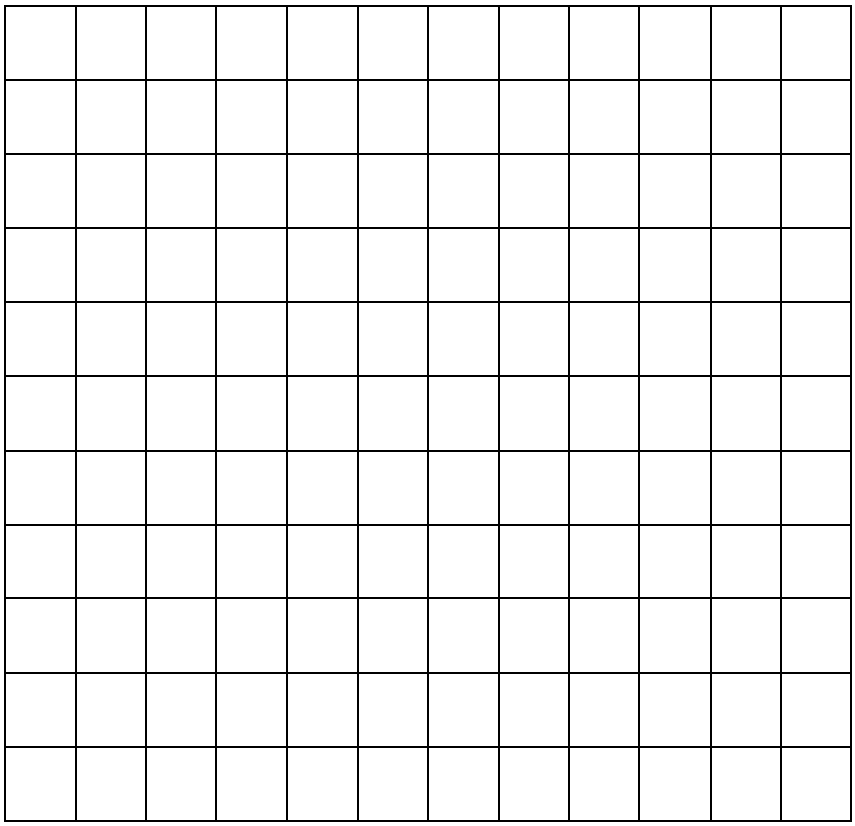
It can be noticed that, after the first few masses, the length of the spring increased by 1cm less for each 20g added. This pattern suggests, for a 160g mass, the length of the spring may be extrapolated to be 50 + 3 = 53cm. However, there can be no certainty about a value calculated by extrapolation.

Example 2:

Students from a previous Mathematics course conducted an experiment to determine if marks scored on a test were related to the number of hours of television watched. This data was collected for ten students.

Student	1	2	3	4	5	6	7	8	9	10
Mark (%)	82	64	84	70	74	76	85	73	94	90
Hours of TV	2	4	0	3	2	2	1	3	1	2

i. Create a scatter plot for this data. Does the data show a trend?



ii. How would you describe the relationship between hours of TV watched and test score?

iii. Draw a straight line that passes close to the data points. How many points lie:
- ❖ on your line?
- ❖ above your line?
- ❖ below your line?

iv. Compare your line with the students on either side of you. Are all three lines the same? Explain.

v. Using the line that you drew, predict the test score of a student who watched 2.5 hours of TV.

Mean Fit Line

Introduction:

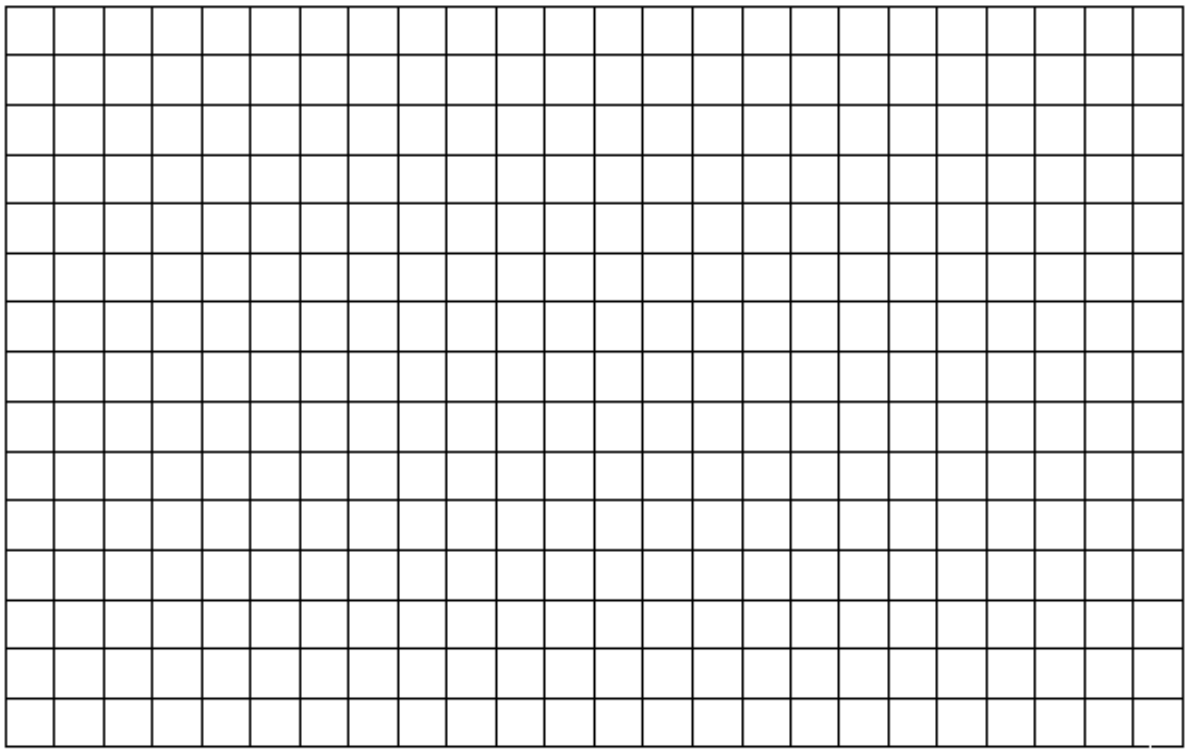
The Mean Fit Line gives a better model for a set of data, compared to a line drawn freely. To draw the MEAN FIT LINE, find the mean coordinates of a set of points by calculating the mean of the first set of data and then the second set of data. Plot this point and then draw a line through it so that about the same number of points are on each side of the line. Drawing a Mean Fit Line allows everyone to make similar predictions about the same set of data.

Example 3:

A hockey team is interested in the relationship between the number of shots on goal they make and the number of goals they score. Plot the data and construct the mean fit line. Use it to predict the number of goals they will score if they take 35 shots. Make a prediction for 45 shots.

Shots	11	20	22	24	28	32	32	40
Goals	1	6	2	7	6	4	8	11

Step 1: Plot the data by thinking about which variable is the dependent (y-axis) variable, and which variable is the independent (x-axis) variable. Be sure that you include a title, and label the axis.



Step 2: Calculate the mean number of shots, and the mean number of goals.

mean shots $= \mu_{shots} = \frac{\sum x_{shots}}{N_{shots}}$

mean goals $= \mu_{goals} = \frac{\sum x_{goals}}{N_{goals}}$

Step 3: Plot the point (,) and draw the Line of Best Fit so that it passes through this point. Be sure that the line through the point has the same number of points on each side of it.

i. Use interpolation and the graph (- - -) to predict the number of goals if 35 shots are taken.

ii. Use extrapolation and the graph (-- -- --) to predict the number of goals if 45 shots are taken.