

Scatter Plots and Line of Best Fit

Name _____ Class Period _____

The **best fitting line or curve** is the line that lies as close as possible to all the data points.

Regression is a method used to find the equation of the best fitting line or curve.

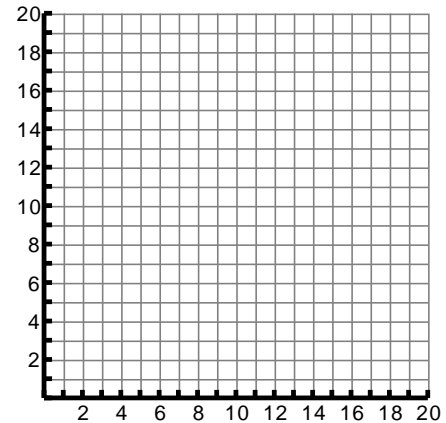
Extrapolation – the use of the regression curve to make predictions outside the domain of values of the independent variable.

Interpolation – Interpolation is used to make predictions within the domain of values of the independent variable.

Line of Best Fit by Hand:

1) The environment club is interested in the relationship between the number of canned beverages sold in the cafeteria and the number of cans that are recycled. The data they collected are listed in this chart.

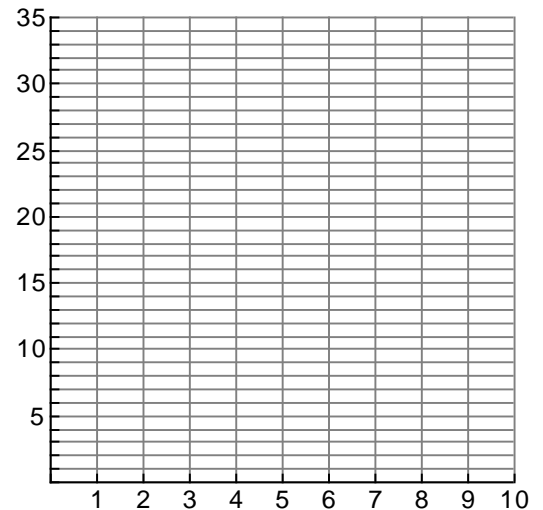
Beverage Can Recycling								
Number of Canned Beverages Sold	18	15	19	8	10	13	9	14
Number of Cans Recycled	8	6	10	6	3	7	5	4



- Plot the points to make a scatter plot.
- Use a straightedge to approximate the line of best fit by hand.
- Find an equation of the line of best fit for the data.

2. Mike is riding his bike home from his grandmother's house. In the table below, x represents the number of hours Mike has been biking and y represents the number of miles Mike is away from home. Make a scatter plot for this data on the grid below.

Hours (x)	1	2	3	4	5	6	7	8
Miles (y)	35	29	26	20	16	9	6	0



- Describe the association between the data points on the scatter plot.
- Use a straightedge to approximate the line of best fit.
- Find an equation of the line of best fit for the data.
- What does the slope represent in the context of the problem? What does the y-intercept represent in the context of the problem?

- Could you use your equation to predict how far Mike would be after 10 hours? Use mathematics to justify your answer.

Line of Best Fit using the calculator

3) Use the table below to answer the questions about the population p (in millions) in Florida.

Year, t	2002	2003	2004	2005
Population (millions)	16.4	17.0	17.4	17.8

a) Find the best-fitting line for the data and the correlation coefficient.

b) Using this model, what will be the population in 2020?

4) Use the table below to answer the questions about the U.S. residential carbon dioxide emissions from 1993 to 2002. Emissions are measured in million metric tons.

Year, t	1993	1994	1995	1996	1997	1998	1999	2000	2001	2002
Emissions	1027.6	1020.9	1026.5	1086.1	1077.5	1083.3	1107.1	1170.4	1163.3	1193.9

a) Find the best-fitting line for the data and the correlation coefficient.

b) Using this model, how many residential tons were emitted in 1990? In 2010?

5) Use the table below to answer the questions about the operating costs in thousands of a small business from 2000 to 2007.

Year, t	2000	2001	2002	2003	2004	2005	2006	2007
Operating Costs	2.3	2.6	3.1	3.3	4.0	5.2	5.9	7.0

a) Find the best-fitting line for the data and the correlation coefficient.

b) Using this model, what will be the operating costs in 2015?