Vertical Progression:

6th Grade	**ELG 6.10 Summarize and describe distributions.** o **6.SP.B.4** Display numerical data in plots on a number line, including dot plots, histograms, and box plots. o **6.SP.B.5** Summarize numerical data sets in relation to their context, such as by: o **6.SP.B.5c** Giving quantitative measures of center (median and/or mean) and variability (interquartile range and/or mean absolute deviation), as well as describing any overall pattern and any striking deviations from the overall pattern with reference to the context in which the data were gathered.
7th Grade	**ELG 7.7 Analyze proportional relationships and use them to solve real-world and mathematical problems.** o **7.RP.A.2** Recognize and represent proportional relationships between quantities. o **7.RP.A.2.c** Represent proportional relationships by equations o **7.RP.A.2.d** Explain what a point (x, y) on the graph of a proportional relationship means in terms of the situation, with special attention to the points (0, 0) and (1, r) where r is the unit rate. **ELG 7.8 Draw informal comparative inferences about two populations.** o **7.SP.B.4** Use measures of center and measures of variability for numerical data from random samples to draw informal comparative inferences about two populations.
8th Grade	**ELG 8.10 Investigate patterns of association in bivariate data.** o **8.SP.A.1** Construct and interpret scatter plots for bivariate measurement data to investigate patterns of association between two quantities. Describe patterns such as clustering, outliers, positive or negative association, linear association, and nonlinear association. o **8.SP.A.2** Know that straight lines are widely used to model relationships between two quantitative variables. For scatter plots that suggest a linear association, informally fit a straight line, and informally assess the model fit by judging the closeness of the data points to the line. o **8.SP.A.3** Use the equation of a linear model to solve problems in the context of bivariate measurement data, interpreting the slope and intercept. *For example, in a linear model for a biology experiment, interpret a slope of 1.5 cm/hr as meaning that an additional hour of sunlight each day is associated with an additional 1.5 cm in mature plant height.* o **8.SP.A.4** Understand that patterns of association can also be seen in bivariate categorical data by displaying frequencies and relative frequencies in a two-way table. Construct and interpret a two-way table summarizing data on two categorical variables collected from the same subjects. Use relative frequencies calculated for rows or columns to describe possible association between the two variables. *For example, collect data from students in your class on whether or not they have a curfew on school nights and whether or not they have assigned chores at home. Is there evidence that those who have a curfew also tend to have chores?*
Algebra 1	**ELG.MA.HS.S.2 Summarize, represent, and interpret data on two categorical and quantitative variables.** o **S-ID.B.5** Summarize categorical data for two categories in two-way frequency tables. Interpret relative frequencies in the context of the data (including joint, marginal, and conditional relative frequencies). Recognize possible associations and trends in the data. o **S-ID.B.6** Represent data on two quantitative variables on a scatter plot, and describe how the variables are related. o **S-ID.B.6a** Fit a function to the data; use functions fitted to data to solve problems in the context of the data. Use given functions or choose a function suggested by the context. Emphasize linear, quadratic, and exponential models. o **S-ID.B.6b** Informally assess the fit of a function by plotting and analyzing residuals. o **S-ID.B.6c** Fit a linear function for a scatter plot that suggests a linear association.

	ELG.MA.HS.S.3 Interpret linear models o **S-ID.C.7** Interpret the slope (rate of change) and the intercept (constant term) of a linear model in the context of the data. o **S-ID.C.8** Compute (using technology) and interpret the correlation coefficient of a linear fit.

Students will demonstrate command of the ELG by:

- Constructing scatter plots and describing patterns such as clustering, outliers, positive or negative association, linear association, and nonlinear association.
- Constructing and interpreting a two-way table summarizing data on two categorical variables collected from the same subjects.
- Informally fitting a straight line to data, and informally assessing the model fit by judging the closeness of the data points to the line.
- Using the equation of a linear model to solve real-world and mathematical problems.

Vocabulary:

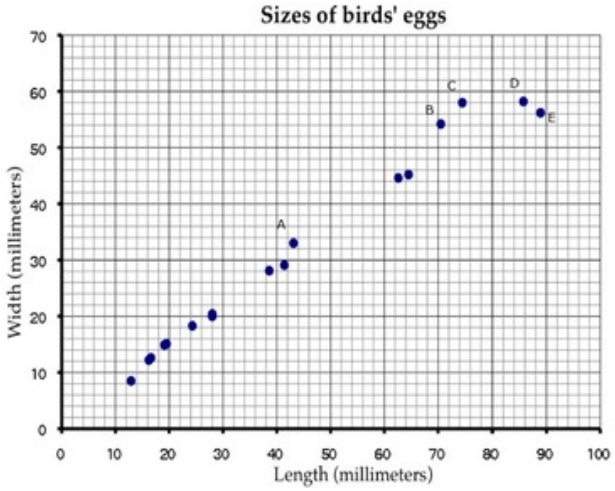
- bivariate data
- categorical data
- clustering
- frequency
- intercepts
- linear association
- negative association
- non-linear association
- outliers
- positive association
- scatter plot
- slope
- two-way table
- relative frequency

Sample Instructional/Assessment Tasks:

1) **Standard(s): 8.SP.A.1**

Source: https://www.illustrativemathematics.org/content-standards/8/SP/A/1/tasks/41

Item Prompt:
This scatter diagram shows the lengths and widths of the eggs of some American birds.



a. A biologist measured a sample of one hundred Mallard duck eggs and found they had an average length of 57.8 millimeters and average width of 41.6 millimeters. Use an X to mark a point that represents this on the scatter diagram.

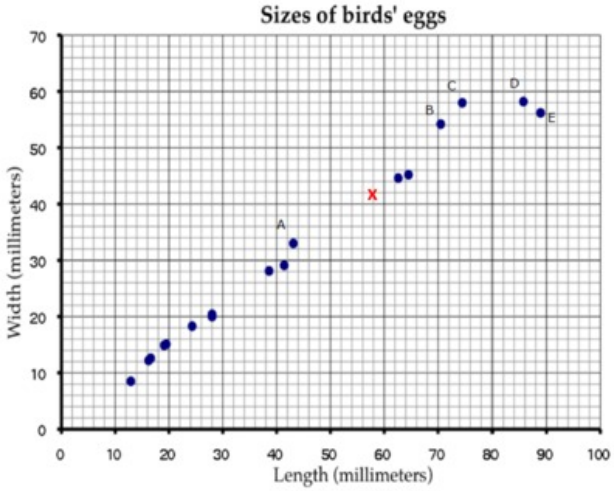
b. What does the graph show about the relationship between the lengths of birds' eggs and their widths?

c. Another sample of eggs from similar birds has an average length of 35 millimeters. If these bird eggs follow the trend in the scatter plot, about what width would you expect these eggs to have, on average?

d. Describe the differences in shape of the two eggs corresponding to the data points marked C and D in the plot.

e. Which of the eggs A, B, C, D, and E has the greatest ratio of length to width? Explain how you decided.

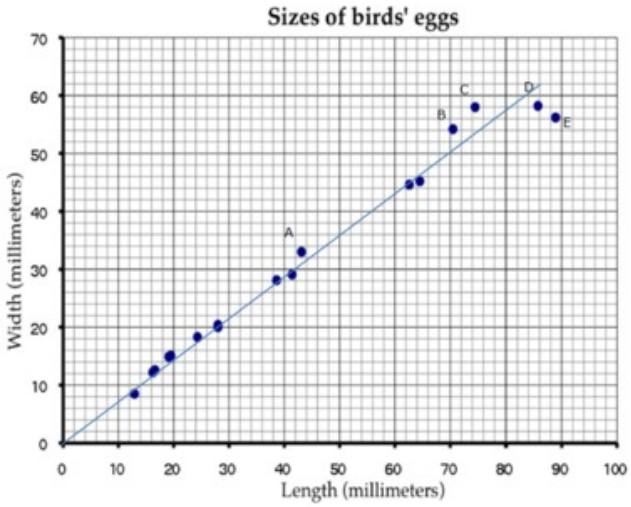
Solution:



a.

b. There seems to be a positive linear relationship between the length and width of the eggs.

c. The line below appears to fit the data fairly well:



Since it passes through (0,0) and (50,36), its slope is $\frac{36}{50} = 0.72$, so the equation of the line is y=0.72x

If x=35, then our line would predict that y=0.72·35=25.2. So we would expect the width of these eggs to be, on average, about 25 mm. Answers using different lines can vary up to 1 mm in either direction.

d. Without reading off precise numerical values from the plot, we can see that eggs C and D have very nearly the same width, but egg D is about 12 millimeters longer than egg C.

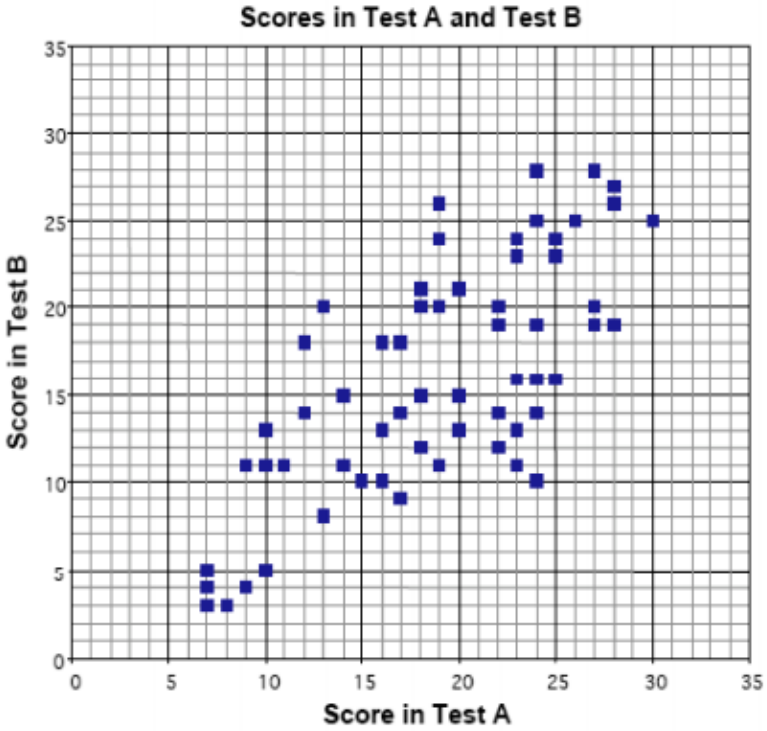
e. First we note that egg E certainly has a higher length-to-width ratio than C or D, since it is both longer and narrower. Similarly, E has a higher ratio than B because it is significantly longer, and only a tad wider. It is harder to visually identify the difference between A and E, we compute their respective length-to-width ratios numerically, which turn out to be approximately 1.3 for A and 1.6 for E. So E has the greatest ratio of length to width.

2) Standard(s): 8.SP.1

Source: Engage NYhttp://www.insidemathematics.org/assets/common-core-math-tasks/scatter%20diagram.pdf

Item Prompt:

A group of 66 students took two tests; Test A and Test B.
In the scatter diagram, each square represents one student and shows the scores that student got in the two tests.



1. The mean score for Test A was 19 and the mean score for Test B was 16.
   Plot a point to show this on the scatter diagram.

2. Draw a line of best fit on the scatter diagram.
   How can a line of best fit be used?

   _____

   _____

3. Here are five statements about the scores shown on the scatter diagram.

   If a statement is true check (√) it.

   If it is not true, write a correct statement.

Statement	Check (√) or write correct statement
The scatter diagram shows positive correlation between the scores on Test A and the scores on Test B.	
The lowest score on Test A is lower than the lowest score for Test B.	
The range of scores on Test B is 25.	
The student with the highest score on Test A also has the highest score on Test B.	
The biggest difference between a student's scores on the two tests is 5.	

Correct Answer:

1. Point correctly plotted

2. Lines will vary, but should be a good fit. Answers questions with reasoning such as, "A line of fit can be used to estimate a student's score in one test if you know his/her score in the other.

3. Table: Correctly identifies true statements and rewrites false statements with statements such as

   a. Correct
   b. No. The lowest score on Test A is greater than the lowest score on Test B.
   c. Correct
   d. No. The student with the highest score on Test A does not have the highest score on Test B.
   e. No. The biggest difference is more than 5.