

Lesson 12: Relationships Between Two Numerical Variables

Classwork

A scatter plot is an informative way to display numerical data with two variables. In your previous work in Grade 8, you saw how to construct and interpret scatter plots. Recall that if the two numerical variables are denoted by x and y , the scatter plot of the data is a plot of the (x, y) data pairs.

Example 1: Looking for Patterns in a Scatter Plot

The National Climate Data Center collects data on weather conditions at various locations. They classify each day as clear, partly cloudy, or cloudy. Using data taken over a number of years, they provide data on the following variables.

x = elevation above sea level (in feet)

y = mean number of clear days per year

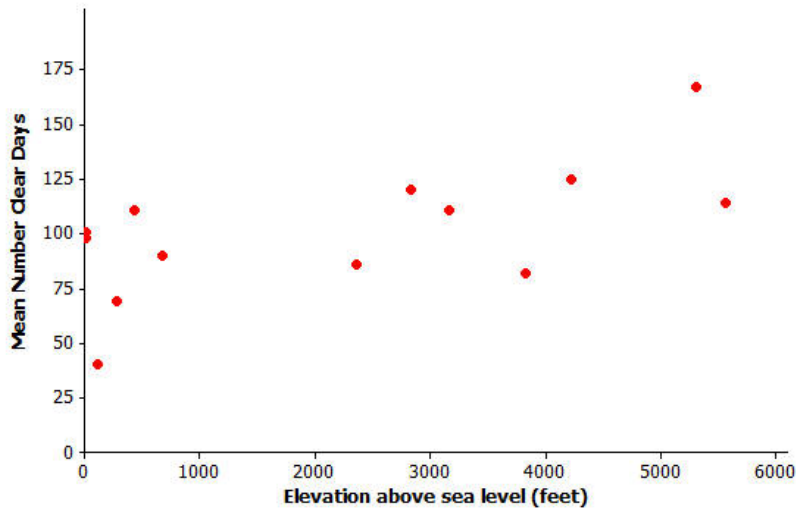
w = mean number of partly cloudy days per year

z = mean number of cloudy days per year

The table below shows data for 14 U.S. cities.

City	x = Elevation Above Sea Level (ft.)	y = Mean Number of Clear Days per Year	w = Mean Number of Partly Cloudy Days per Year	z = Mean Number of Cloudy Days per Year
Albany, NY	275	69	111	185
Albuquerque, NM	5,311	167	111	87
Anchorage, AK	114	40	60	265
Boise, ID	2,838	120	90	155
Boston, MA	15	98	103	164
Helena, MT	3,828	82	104	179
Lander, WY	5,557	114	122	129
Milwaukee, WI	672	90	100	175
New Orleans, LA	4	101	118	146
Raleigh, NC	434	111	106	149
Rapid City, SD	3,162	111	115	139
Salt Lake City, UT	4,221	125	101	139
Spokane, WA	2,356	86	88	191
Tampa, FL	19	101	143	121

Here is a scatter plot of the data on elevation and mean number of clear days.



Data Source: <http://www.ncdc.noaa.gov/oa/climate/online/ccd/cldy.html>

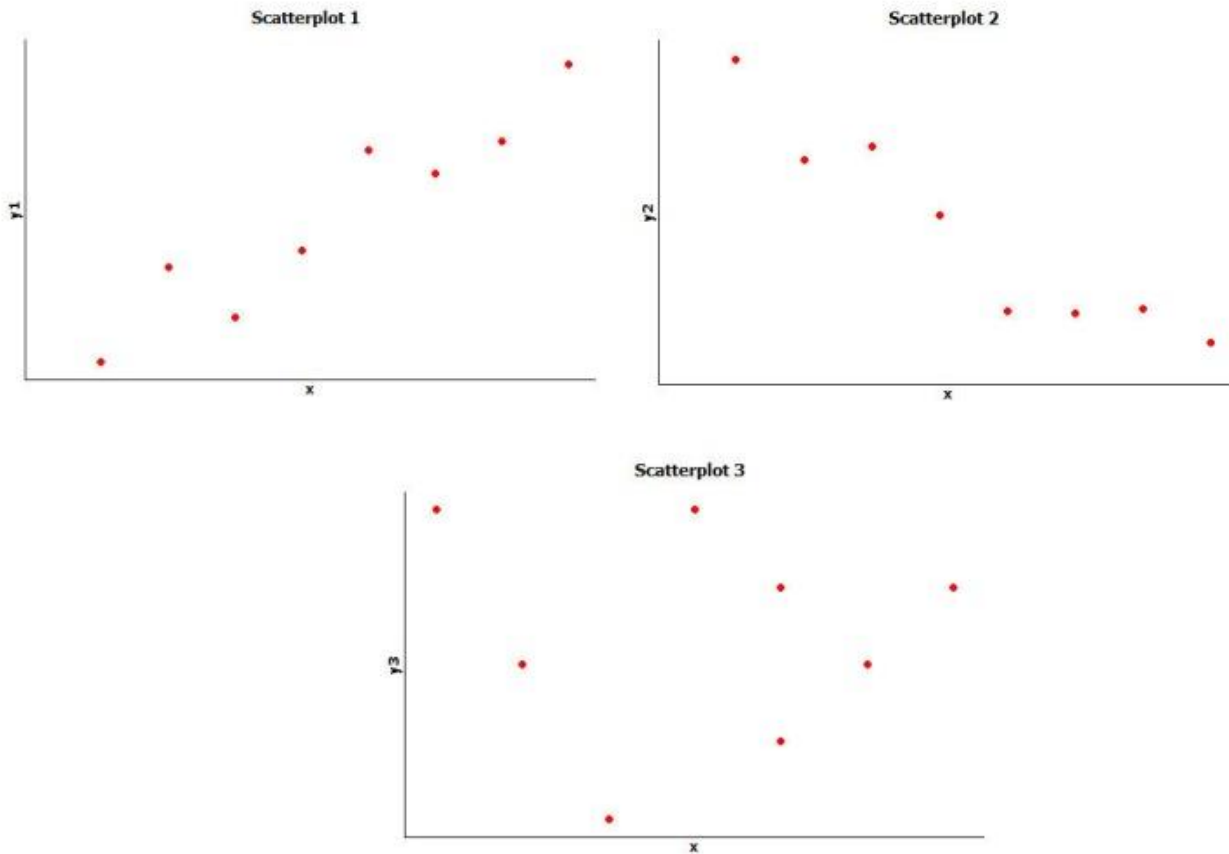
Exercises 1–3

1. Do you see a pattern in the scatter plot, or does it look like the data points are scattered?
2. How would you describe the relationship between elevation and mean number of clear days for these 14 cities? That is, does the mean number of clear days tend to increase as elevation increases, or does the mean number of clear days tend to decrease as elevation increases?
3. Do you think that a straight line would be a good way to describe the relationship between the mean number of clear days and elevation? Why do you think this?

Exercises 4–7: Thinking about Linear Relationships

Below are three scatter plots. Each one represents a data set with eight observations.

The scales on the x - and y -axes have been left off these plots on purpose so you will have to think carefully about the relationships.



- If one of these scatter plots represents the relationship between height and weight for eight adults, which scatter plot do you think it is and why?

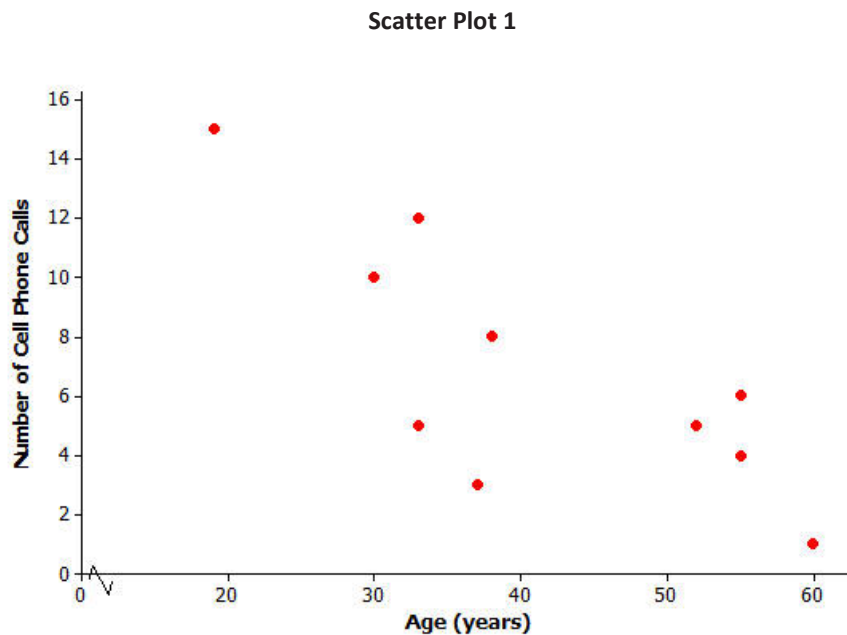
- If one of these scatter plots represents the relationship between height and SAT math score for eight high-school seniors, which scatter plot do you think it is and why?

- If one of these scatter plots represents the relationship between the weight of a car and fuel efficiency for eight cars, which scatter plot do you think it is and why?
- Which of these three scatter plots does *not* appear to represent a linear relationship? Explain the reasoning behind your choice.

Exercises 8–13: Not Every Relationship Is Linear

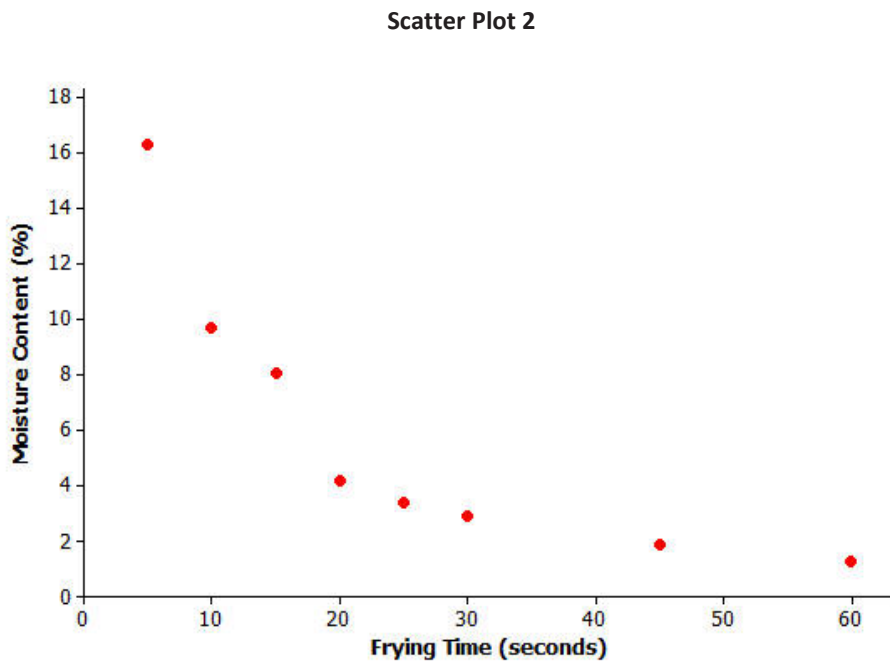
When a straight line provides a reasonable summary of the relationship between two numerical variables, we say that the two variables are *linearly related* or that there is a *linear relationship* between the two variables.

Take a look at the scatter plots below and answer the questions that follow.



- Is there a relationship between number of cell phone calls and age, or does it look like the data points are scattered?

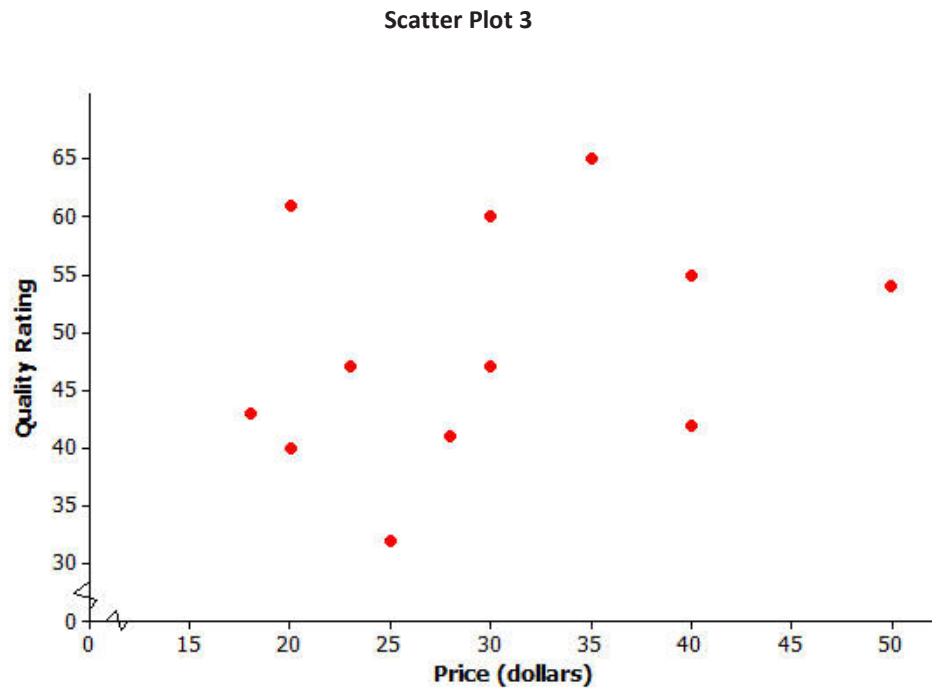
9. If there is a relationship between number of cell phone calls and age, does the relationship appear to be linear?



Data Source: R.G. Moreira, J. Palau, V.E. Sweat, and X. Sun, "Thermal and Physical Properties of Tortilla Chips as a Function of Frying Time," *Journal of Food Processing and Preservation*, 19 (1995): 175.

10. Is there a relationship between moisture content and frying time, or do the data points look scattered?

11. If there is a relationship between moisture content and frying time, does the relationship look linear?



Data Source: www.consumerreports.org/health

12. Scatter plot 3 shows data for the prices of bike helmets and the quality ratings of the helmets (based on a scale that estimates helmet quality). Is there a relationship between quality rating and price, or are the data points scattered?
13. If there is a relationship between quality rating and price for bike helmets, does the relationship appear to be linear?

Lesson Summary

- A scatter plot can be used to investigate whether or not there is a relationship between two numerical variables.
- A relationship between two numerical variables can be described as a linear or nonlinear relationship.

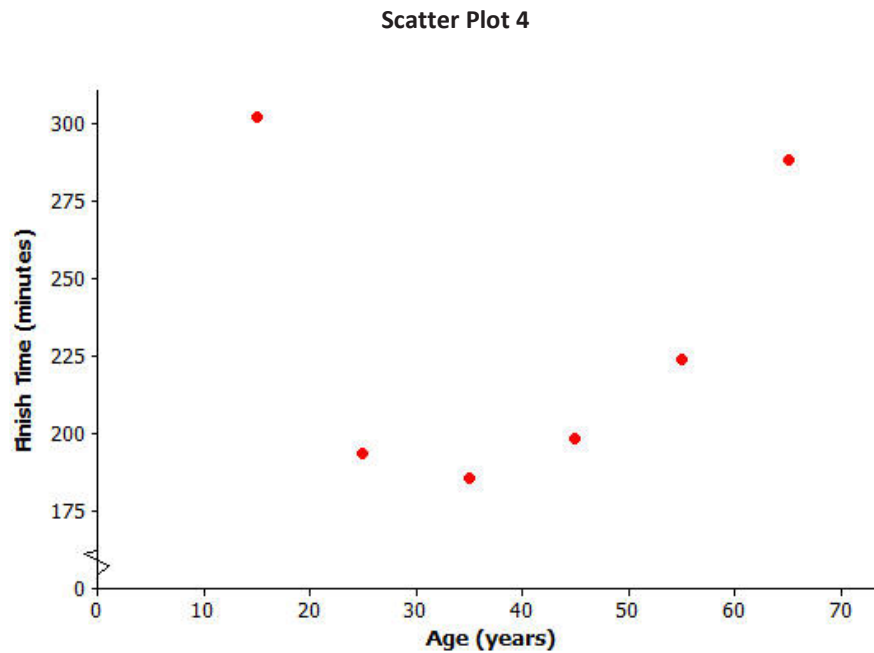
Problem Set

1. Construct a scatter plot that displays the data for x = elevation above sea level (in feet) and w = mean number of partly cloudy days per year.

City	x = Elevation Above Sea Level (ft.)	y = Mean Number of Clear Days per Year	w = Mean Number of Partly Cloudy Days per Year	z = Mean Number of Cloudy Days per Year
Albany, NY	275	69	111	185
Albuquerque, NM	5,311	167	111	87
Anchorage, AK	114	40	60	265
Boise, ID	2,838	120	90	155
Boston, MA	15	98	103	164
Helena, MT	3,828	82	104	179
Lander, WY	5,557	114	122	129
Milwaukee, WI	672	90	100	175
New Orleans, LA	4	101	118	146
Raleigh, NC	434	111	106	149
Rapid City, SD	3,162	111	115	139
Salt Lake City, UT	4,221	125	101	139
Spokane, WA	2,356	86	88	191
Tampa, FL	19	101	143	121

2. Based on the scatter plot you constructed in Question 1, is there a relationship between elevation and the mean number of partly cloudy days per year? If so, how would you describe the relationship? Explain your reasoning.

Consider the following scatter plot for Questions 3 and 4.

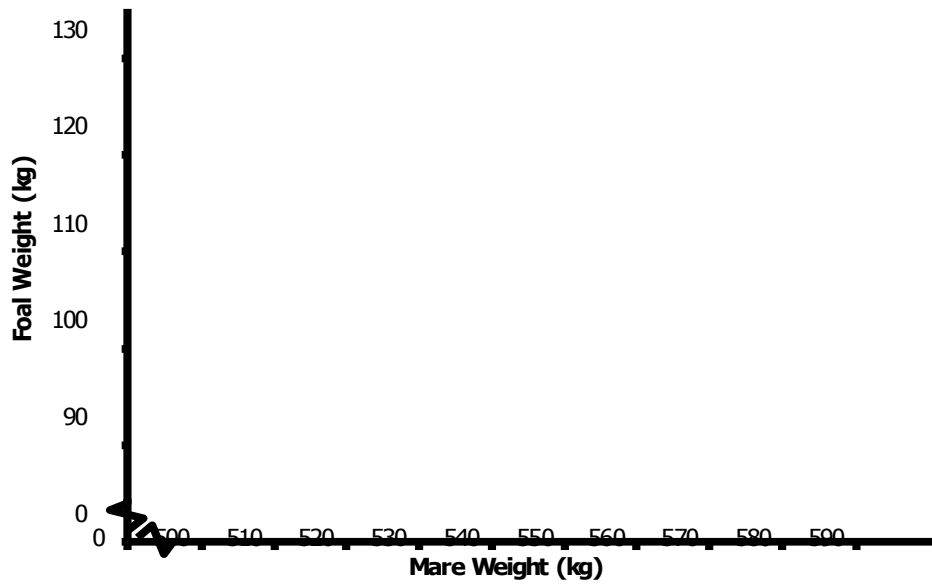


Data Source: Sample of six women who ran the 2003 NYC marathon

3. Is there a relationship between finish time and age, or are the data points scattered?
4. Do you think there is a relationship between finish time and age? If so, does it look linear?

Consider the following scatter plot for Questions 5 and 6.

Scatter Plot 5



Data Source: Elissa Z. Cameron, Kevin J. Stafford, Wayne L. Linklater, and Clare J. Veltman, "Suckling behaviour does not measure milk intake in horses, equus caballus," *Animal Behaviour*, 57 (1999): 673.

5. A mare is a female horse and a foal is a baby horse. Is there a relationship between a foal's birth weight and a mare's weight, or are the data points scattered?
6. If there is a relationship between baby birth weight and mother's age, does the relationship look linear?