

Chapter 7 Examples

1. Consider each of the two following variables, and imagine and describe what the scatter plot might look like.

- Drug dosage and degree of pain relief

Answer: The association is likely to be strong, positive and curved. Assuming, of course, that the drug is an effective pain reliever, as the dosage increases, the degree of pain relief will increase. Eventually, the association is likely to level off, until no further pain relief is possible, since the pain will be gone.

- Calories consumed and weight loss

Answer: The association is likely to be moderate, negative and linear. As fewer calories are consumed, more weight is likely to be lost. The association will not be strong; since some people lose weight easier than others, and there are other variables involved like overall health, exercise and beginning weight.

- Hours of sleep and score on a test

Answer: The association is likely weak, positive and possibly linear. Generally, a well rested person is expected to score higher on a test. The relationship is weak, since there are other variables involved. Maybe a person got less sleep because they were up studying.

- Shoe size and grade point average

Answer: There is no association between shoe size and GPA. The scatter plot is likely to be randomly scattered.

- Time for a mile run and age

Answer: The association between time for a mile run and age is likely to be moderate and curved, with no dominant direction. The very young will likely have high run times. Run times are likely to be the lowest for people in their late teens or early twenties. Older people are likely to have high run times.

- Age of car and cost of repairs

Answer: The association between age of car and cost of repairs is positive, moderate, and linear. As cars get older, they usually require more repairs.

2. Straighten a scatter plot.

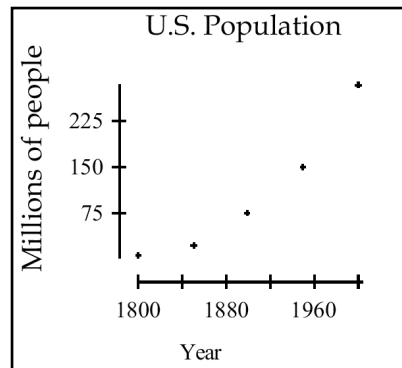
Try logarithmic and square root re-expressions of US population data.

Solution:

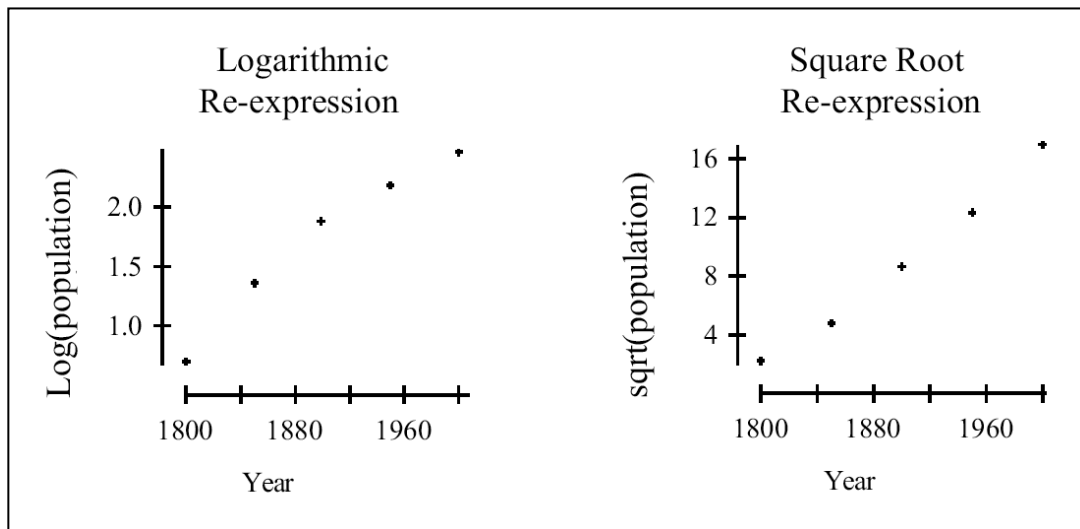
Think

The variables are *year* and *U.S. population*, in millions of people. Both variables are quantitative. The association between *year* and *population* is strong, positive, and curved. Population has been increasing over the last 200 years. Furthermore, the rate of population growth has been increasing. The U.S. population has been growing faster in more recent years. We will attempt to straighten the scatter plot using a logarithmic re-expression and a square root re-expression.

Year	Population (millions)
1800	5
1850	23
1900	76
1950	151
2000	285



Show



Tell

The scatter plot of $\log(\text{population})$ and *year* shows a strong, positive association, but it is still curved. This is not a good re-expression. The scatter plot of square root of *population* and *year* shows a strong, positive, more linear association. The square root re-expression straightens the scatter plot well.