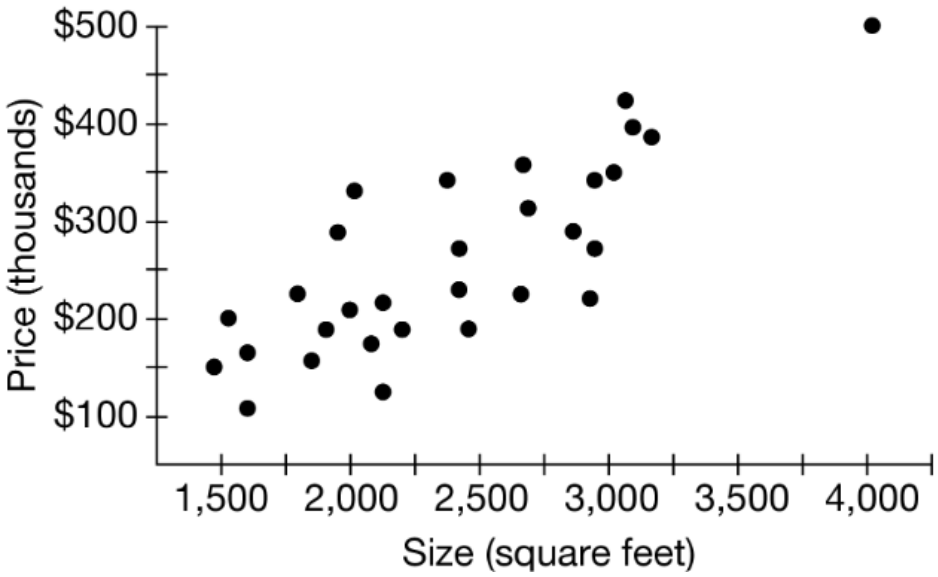
2023 AP Daily: Practice Sessions

AP Statistics

Session 7 – FRQ (Part B, Question 6: Investigative Task)



The above scatterplot shows the size, in square feet, and the selling price, in thousands of dollars, for a sample of 30 houses for sale in a certain area.

a. Describe the association shown in the scatterplot.

b. In the region, houses are considered large if they are greater than 2,500 square feet and expensive if the selling price is greater than $300,000. The following two-way table summarizes the houses in the sample.

   i. Use the information in the table to construct a graphical display of the data.

	Large	Not Large	Total
Expensive	8	2	10
Not Expensive	4	16	20
Total	12	18	30

Source: AP Topic Question; Taken from: AP Classroom

ii. Assume there is no association between size (large, not large) and price (expensive, not expensive). Use the given totals to complete the following table with the expected number of houses for each classification if there was no association.

	Large	**Not Large**	**Total**
Expensive			10
Not Expensive			20
Total	12	18	30

For associations displayed in the scatterplot, the strength of linear association is measured by the correlation coefficient. For the scatterplot of houses, *r* = 0.82.

For associations that are summarized in two-way tables, the strength of association is measured by the chi-square statistic. The formula for the chi-square statistic is:

$$\chi^2 = \sum \frac{(observed - expected)^2}{expected}$$

where expected is the count assuming no association and observed is the count shown by the data. Greater values of χ^2 indicate stronger association. For the table of counts in part (b), χ^2 = 10.

c. Suppose the selling price for the most expensive house in the sample is decreased from $489,000 to $325,000.

   i. What effect would the decrease have on the value of r? Explain your reasoning.

   ii. What effect would the decrease have on the value of χ^2? Explain your reasoning.

d. Based on your answer to part (c), explain one benefit and one drawback of using χ^2 rather than *r* to measure the strength of an association.

Source: AP Topic Question; Taken from: AP Classroom