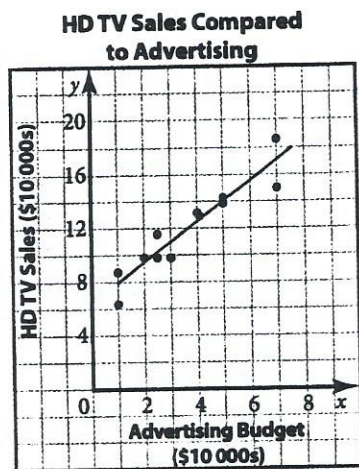


Analysis and Conclusions

When studying trends in data, it is important to consider the reasonableness of conclusions drawn. Sometimes there can appear to be a relationship between two variables but in reality, there are other factors at play which affect the conclusions that can be drawn.

In a **Cause and Effect** relationship, a change in one variable **causes** a change in the other variable. For example, buying more items at a fixed price results in a predictable increase in the amount you pay for those items.

Example 1 The sales manager for an electronics manufacturer analysed the sales of high definition televisions relative to the company's television advertising budgets in 12 markets across Canada. The scatter plot and line of best fit are shown.



$$y = 1.57x + 6.34; r = 0.94$$

- Describe the strength and type of correlation.
- Do you think we can conclude that this is a cause and effect relationship? What other factors might influence this relationship?
- Is it safe to **extrapolate** to estimate the television sales if a market has an advertising budget of \$350 000? Explain.

Outliers vs. Influential Points

Outliers and influential points both affect the line of best fit in some way.

An **outlier** is a point that does not follow the pattern of the other points in the data set.

An **influential point** is a point that lies a significant distance from the other points and greatly affects the slope and y-intercept of the line of best fit.

Errors in analysis can occur for a variety of reasons, including:

- Not enough data
- Using linear regression for non-linear data (ie. finding the line of best fit for data that is not linear!)
- Applying linear regression when the correlation is weak
- Reversing the cause and effect relationship (saying x causes y when in fact y causes x)
- Extrapolating outside of the reasonable range of data
- Not removing outliers and influential points prior to analysis
- Assuming cause and effect when it might actually just be a correlation or even an accidental relationship

Example Describe the possible error in regression or analysis for each situation.

