The only age group showing a clear lead for In is the young – but only 52% of them said they are certain to vote on June 23. Voters aged over 35 back Brexit, the poll showed.

The poll showed an overall four point lead for the Brexit camp but undecided voters were said to lean toward staying in.

---

4.5 The Use and Misuse of Statistics

Assessing the Validity of Conclusions

Work with a partner.

Refer to Case Study: Women in the Workforce.

- Do you think the conclusion is valid or invalid?
  - Include answers to these questions in your response.
  - Are the data reliable?
  - Is there a possible bias in the person analysing the data?
  - Is the sample size reasonable?
  - Is the correlation strong?
  - Is there any evidence to support a cause-and-effect relationship?
  - Does the graph represent the data appropriately?
Case Study
Women in the Workforce

The members of a high school debating club are preparing for a debate.

Debate position:
It would be better for families if women stayed at home rather than joined the workforce.

Preparation:
- Analyse data about women in the workforce and single-parent families in Canada.
- Use data collected by Statistics Canada.

Data:

<table>
<thead>
<tr>
<th>Year</th>
<th>Number of women in workforce (thousands)</th>
<th>Number of single-parent families</th>
</tr>
</thead>
<tbody>
<tr>
<td>2000</td>
<td>6790.4</td>
<td>1317 760</td>
</tr>
<tr>
<td>2001</td>
<td>6910.3</td>
<td>1406 390</td>
</tr>
<tr>
<td>2002</td>
<td>7126.0</td>
<td>1404 250</td>
</tr>
<tr>
<td>2003</td>
<td>7324.2</td>
<td>1451 150</td>
</tr>
<tr>
<td>2004</td>
<td>7466.4</td>
<td>1444 150</td>
</tr>
</tbody>
</table>

Conclusion:
These data support our position.
There is a strong positive correlation between the two variables.

As the number of women in the workforce increases, the number of single-parent families increases. If we want to reduce the number of single-parent families in Canada, women should stay at home and not go out to work.

Example 1
Assessing Graphs

The graphs in each pair show the same data. Choose the graph that displays the data more accurately. Justify your choice.

a) Canada’s population by age according to the 2001 census

i) Ages of Canadians, 2001 Census
ii) Ages of Canadians, 2001 Census

b) A company’s profits over a 5-year period

i) Company Profits
ii) Company Profits
Example 2 Assessing How Data Were Collected and Graphed

Four Grade 9 students collected data on school lunch preferences.

They concluded:

We asked students to tell us their favourite lunch meals and displayed the results in this bar graph. We conclude that the school cafeteria should serve more pizza since it is clearly the favourite lunch of students.

Is this conclusion valid?

Solution

- **Was the sample size appropriate?**
  By adding the frequencies, you can see that 50 students were surveyed. Depending on the size of the school, this may not be enough data.

- **Was the sample representative?**
  You cannot tell from the information given. Perhaps the researchers surveyed only their friends or only Grade 9 students. You need more information about the sampling technique to judge this.

- **Was the survey question biased?**
  It appears that students were simply asked their favourite lunch meal. This is an unbiased question since it does not try to influence the answer.

- **How was the survey conducted?**
  It appears the survey was conducted orally. This could bias the results because some students might be self-conscious about their eating habits.

- **Is the graph constructed accurately?**
  In general, a three-dimensional bar graph tends to distort the relative quantities being displayed. Starting the vertical axis at 4 also distorts the relative quantities.
Example 3

Assessing Assumptions about Cause and Effect

A group of Grade 12 students performed a linear regression on data they collected from Statistics Canada about the number of seniors and the number of weapons crimes in Canada.

The $r$-value is a measure of how strong a correlation is. The closer to 1 or -1, the stronger the correlation.

They concluded:

There is a strong positive correlation between the two variables. As the number of seniors increases, weapons charges increase. Therefore, criminals in Canada are becoming bolder because of our ageing population.

Is this conclusion valid?

To assess the validity of the conclusion, ask yourself these questions.

■ Was there bias in the data collection?
The students gathered the data from Statistics Canada, which is a reliable source of data. However, only 5 years of data were included.

■ Is the graph constructed accurately?
Yes. Because the numbers are large, it would be impractical to start the scale on the vertical axis at zero.

■ Is the correlation strong?
Yes. The points are close to the regression line.

■ Does the analysis support a cause-and-effect relationship?
Not necessarily. Both variables may be increasing because the population of Canada is increasing.

hw p. 229 1-5, 8,11,12