MAP4C1

Unit 4: Statistical Literacy

Learning Goals: We are learning to…

1. Use statistical measures such as measures of central tendency and measures of spread as well as statistical reasoning to interpret data
2. Assess the validity of survey results based on bias and the sampling techniques used
3. Differentiate between the different types of sampling and identify the pros and cons of each
4. Assess the use and misuse of data in the media
5. Assess data based on graphs, method of collection and looking at cause and effect
6. Interpret and use price indices and price index to solve problems

Day	Topic	Expectations, Learning Goals	Practice/Homework
1	4.1 Interpreting Statistics	LG 1 D2.1	p.201 #1, 3, 4, 8-11, 14
2	4.2 Surveys and Questionnaires	LG 2, 3 D1.2, 2.4	p.214 #1-3, 5, 6, 8-11, 14
3	4.3 The Use and Misuse of Statistics	LG 4, 5 D2.3	p.229 #1-5, 7-10
4	4.4 Understanding Indices	LG 6 D2.2	p.237 #1-5, 9, 10, 12
5	**Two-Variable Data Assignment** **Due:** _____	D1.3, 1.4, 1.8	
6	Review		See review package
7	**Unit 4 Quest** **Date:**_____		

Subject to change based on school activities and class needs

4.1 Interpreting Statistics

Learning Goals: I am learning to…

- ☐ Determine the measures of central tendency and spread.
- ☐ Identify quartile values in a set of data and know what this means and represents
- ☐ Interpret data from its given form and explain/make conclusions from it

Recall: The measures of central tendency are → _____, _____ and _____.

	All values in a set of data added up and divided by the number of values in the data set
	The value that lies in the **middle** of a sorted data set
	The value that occurs the most in the set of data
	The highest value **subtract** the lowest value in the set of data

Example 1: The 14 students in a math class all measured their heights to the nearest centimetre. The results are shown below.

160 178 180 168 157 164 179 153 182 176 165 175 167

a) Determine the mean, median, mode and range for this set of data.

b) What percent of the class is shorter than each measure of central tendency?

c) Ryan is taller than 65% of the class. How many students are shorter than Ryan? What is Ryan's height?

Measure of Spread

- **Standard Deviation →** Measures how _____ the data is centered around the _____.
- **Percentiles →** Tells us what percentage of the data is _____ _____ a particular data value.
    - **Example:** _____ of the data is than or equal to the _____ percentile.
- **Quartiles →** Divide a set of ordered data into four equal parts.
    - The **2nd Quartile (Q_2),** is the median of the _____ set of data. It cuts the data set in half.
    - The **1st Quartile (Q_1),** is the median of the _____ _____ of the set of data, below Q_2. It divides the lower half of the set of data in half, so it is the same as the _____ percentile.
    - The **3rd Quartile (Q_3),** is the median of the _____ _____ of the set of data, above Q_2. It divides the upper half of the set of data in half, so it is the same as the _____ percentile.

Example 2: Below are the hourly rates in dollars for 17 high school students with part-time jobs.

11.50 10.50 8.00 8.25 9.00 9.15 9.75 7.50 8.00
12.50 13.00 11.25 10.75 9.50 9.25 9.45 7.75

a) What are the quartiles for this set of data?

b) Damien's pay is in the 85th percentile for this group of data. What does the percentile mean? What is Damien's hourly pay rate?

Data Reliability – Comparing Data Sources

Given the topic and possible data sources below, decide which source will provide more accurate data.

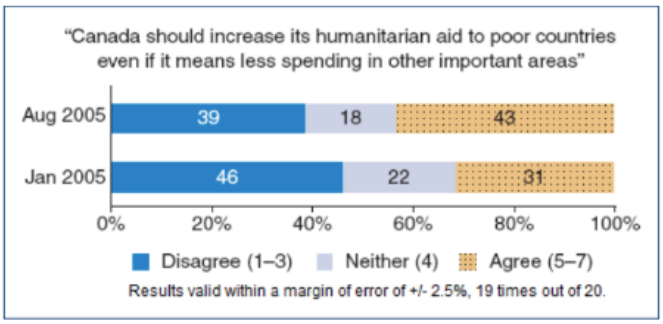
Research Topic	Data Source #1	Data Source #2
a) The benefits of adverse effects of drinking milk	A pamphlet from an animal rights group that opposes dairy farming	Canada's food guide produced by Health Canada
b) The effects of logging on the population of a bird species	A pamphlet from a wildlife protection organization	A forestry company advertisement
c) Possible complications of the flu shot	A Ministry of Health website	A website run by an organization again vaccinations

Explanations

a)
b)
c)

Example: Interpreting Poll Results

Results of a poll conducted by EKOS in 2005 are shown.



a) What question were people asked?

b) How did their favourable responses compare in January and August?

c) A line below the graph states that "the results are valid within a margin of error plus or minus 2.5% points, 19 times out of 20." What does this mean?

4.1 Interpreting Statistics Homework

1. Determine the quartiles for each data set.

a) 10, 8, 12, 15, 9, 9, 11, 12, 12, 8, 14, 11

b) 170, 162, 150, 165, 180, 165, 154, 163, 168, 164, 172

3. The Grade 9 students in a high school participated in a national standardized math test. The principal reported on the school Web site that three of the students placed in the upper quartile. Which sentence best describes the meaning of this statement?

i) Three students received a mark of at least 75%.

ii) Three students did better than 75% of all those who wrote the test.

4. In each case, a research topic and two sources of information are described. Decide which data source is more likely to provide reliable data. Justify your answers.

	Topic	Source 1	Source 2
a)	The sound quality of a particular stereo	An advertisement in a magazine	A review in a consumer magazine
b)	The possible side effects of a medication	A health information Web site run by a hospital	A blog written by someone who has taken the medication
c)	Job prospects in a particular field	A PDF document written by the Ontario Ministry of Training, Colleges, and Universities	A brochure you receive advertising mail-order courses in this field

8. Use the growth chart for girls aged 10 to 20.

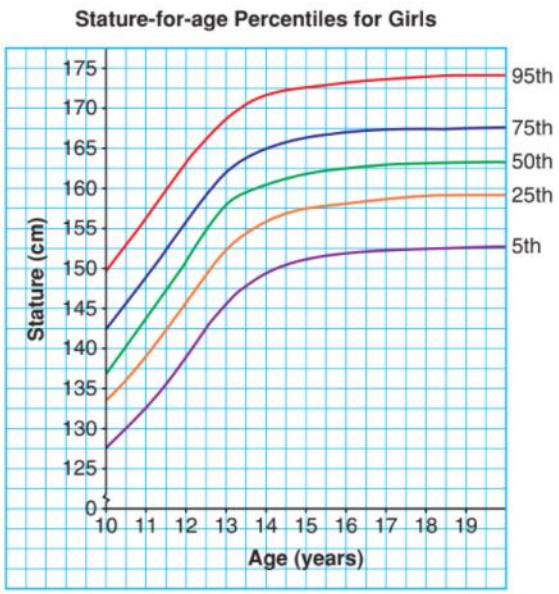
a) Determine the percentile ranking for each girl.

i) Tameika is 14 years old and 165 cm tall.

ii) Audra is 16 years old and 152 cm tall.

iii) Sabrina is 19 years old and 174 cm tall.

b) Asayo is 17 years old and 165 cm tall. Between what 2 quartiles is her percentile ranking?



9. Here are the exam marks for a class of 20 math students.

35	72	74	84	90	60	93	48	70	68
75	63	65	75	82	65	54	77	64	59

a) Determine the mean, median, and mode.
Which measure of central tendency best represents the data?
b) What are the quartiles for this data set?
c) Vince's mark is in the 37th percentile for this group.
Explain what the percentile means. What is Vince's mark?

10. In November 2007, Ipsos Reid conducted a poll of 1314 randomly selected Ontarians. The report stated:

> …eight in ten (80%) Ontarians support "legislation that would ban smoking in cars and other private vehicles where a child or adolescent under 16 years of age is present". Moreover, a majority of non-smokers (86%) and smokers (66%) would support this legislation…

a) What question were people asked?
b) How did the responses of smokers and non-smokers compare?
c) The poll results are considered accurate to within ±2.7 percentage points, 19 times out of 20. Explain what this statement means.

14. Transportation engineers are considering changing the speed limit on a rural road. Every day for 1 week, a technician records speeds of vehicles using the road in kilometres per hour. Here is a representative sample of data:

76	74	78	75	69	68	87	90	73	70
68	72	85	78	72	70	75	75	76	65

a) What are the quartiles for this data set?
b) Determine the 85th percentile speed for these data.
c) The current speed limit on this road is 70 km/h. Based on these data, would you recommend changing it? What other factors should be considered?

Answers:

8. **a)** **i)** 75th percentile
ii) 5th percentile
iii) 95th percentile
b) Between 50th and 75th percentile
9. **a)** Mean: 68.65; median: 69; modes: 65, 75
b) 1st quartile: 61.5; 3rd quartile: 76
c) 37% of the people that wrote the test received a mark below Vince's. Vince's mark is 65.
10. **a)** "Do you support the legislation that would ban smoking in cars and other private vehicles where a child or an adolescent under 16 of age is present?"/"Do you smoke?"
c) Poll results are not always accurate, but 19 times out of 20 (95%), the results are within 2.7% of the true public opinion.
11. **a)** Are you likely to/certain to/unlikely to/certain not to avoid toys made in China because of concerns about health or safety risks?
b) 55%
d) "Majority of Canadians Likely to Avoid Toys Made in China"
14. **a)** 1st quartile: 70, 2nd quartile: 74.5, 3rd quartile: 77
b) 85 km/h

4.2 Surveys and Questionnaires

Learning Goals: I am learning to…
- ☐ Identify and explain the different types of sampling techniques
- ☐ Identify and explain the different types of bias in surveys and questionnaires

Representative Sample: A sample that is typical of the _____. If the sample is not representative, it is _____ and the survey results are invalid.

Sample Size: In a survey, the sample size can affect results.
- If the sample is too _____, the survey results may not be _____.
- If the sample is too _____, the survey may _____ too much to complete and may to too difficult to conduct fairly.

Part A: Sampling Techniques

Random Techniques: Where each of the population has an _____ chance of being selected.

1. _____ _____ Sampling → Participants are picked randomly.
2. _____ Sampling → The population is grouped, and a few individuals are picked from each group.
3. _____ Sampling → The population is organized into groups and one group is chosen
4. _____ Sampling → Every n^{th} individual is selected.

Non-Random Techniques: Will not necessarily provide a _____.

1. _____ Sampling → Individuals who are easy to sample are chosen
2. _____ Sampling → The person who is doing the sampling uses their judgement to create a representative sample.
3. _____ Sampling → Participants volunteer.

Example 1: A town has a population of 20,000 people. The town council conducts a vote at a public meeting about constructing a new ice-hockey rink.
- 50 people attend the meeting
- 40 of the people at the meeting are in favour of the new hockey rink
- The council decides to build the hockey rink since the majority of the people support the idea

a) What percent of the people at the meeting voted for the rink?

b) What percent of people in the town attended the meeting?

c) Is this sample representative? **Justify** your answer.

Part B: Bias in Surveys

Bias: Occurs when the results of a _____ do not reflect the _____ population.

Biased Questions: These type of questions _____ people's choices or use _____ that could influence people to answer in a certain way. For results to be valid in a survey, questions must be _____.

Types of Bias:

- **Leading Questions** → Contain wording or information to _____ a specific response
- **Loaded Question** → Suggest a socially desirable answer or are emotionally charged.
- **Response Bias** → When people intentionally _____ or give _____ information.
- **Non-Response Bias** → When people fail to _____ one or more questions.
- **Sampling Bias** → When you have a _____ sample.

Example 2: People walking by in the mall were asked "We harm the planet when we use pesticides on our lawns. Should the government ban all residential pesticide use?"

a) Will the survey results be valid? **Justify** your answer.

b) How could this survey be improved?

Example 3: About 4000 people visited a large sports equipment store during its annual sale. The store surveyed 100 customers after they paid for their purchases. An employee recorded their answers.


1. Good sports equipment can greatly improve performance. How much do you spend on equipment each year?
___ $200 or less ___ $200–$400
___ $400–$600 ___ $600–$800
___ $800–$1000 ___ More than $1000

2. How much do you earn per year?
___ Less than $10 000 ___ $10 000–$20 000
___ $20 000–$40 000 ___ $40 000–$60 000
___ $60 000–$80 000 ___ More than $80 000

a) Is this sample size large enough?

b) Is this sample representative?

c) Are the survey questions unbiased?

d) Was the collection method appropriate?

4.1 Surveys and Questionnaires Homework

1. Three schools each survey 300 students about whether they want a longer lunch. What percent of students in each survey want a longer lunch?

	School	Number of students who want a longer lunch
a)	1	60
b)	2	270
c)	3	175

2. For each population, determine how many people should be surveyed to include 10% of the population.

a) 350 people b) 930 people c) 1180 people d) 10 360 people

3. The student council at a school surveys 50 students. What percent of each population is this? Choose one population and explain whether you think it is a large enough sample.

a) 450 students b) 750 students c) 1200 students

5. Identify whether each survey question is biased or unbiased.

a)

Old gasoline powered lawn mowers pollute more than cars. People should be forced to replace them with more efficient mowers.

❑ Agree ❑ Disagree

b)

We will offer yoga classes one weeknight each week. Which night would you prefer?

❑ Thursday
❑ Monday
❑ Wednesday
❑ Tuesday
❑ Friday

c)

Should owners of hybrid vehicles be given an energy efficiency rebate from the government?

❑ Yes ❑ No

d)

Speed kills! Speed limits on our highways should be reduced to 90 km/h.

❑ Agree ❑ Disagree

6. An Internet survey asks people's opinions about a new software package. Which question is unbiased? How is the other question biased?

i)

This software is used by some of the biggest names in business.
If you have tried this software, what did you think about it?

___ Excellent ___ Good ___ Fair ___ Poor

ii)

Have you tried this software? ___ Yes ___ No
If you have tried this software, what did you think about it?

___ Excellent ___ Good ___ Fair ___ Poor

8. For each part of question 5 that involved a biased question, explain how the question is biased and suggest how it could be improved.

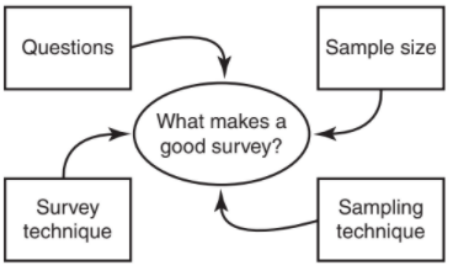
9. An urban music radio station asks its listeners to e-mail or text an answer to this question:
"Do you think students in our city should wear school uniforms?"
95% of respondents say "No". The radio station announces that city schools should not introduce school uniforms since 95% of city residents are against the idea. Is the sample representative? If not, how could it be improved?

10. A newspaper columnist wants to find out what people think of a proposed by-law that would limit the height of fences they can build in their yards. He writes this survey question in his weekly column.
Will the survey results be valid?
Justify your answer.
If you feel the survey is not valid, how could it be improved?

> Once again the government is trying to control us. This time they are interfering with our backyards.
> Do you agree with the proposed law to limit the height of a fence residents can put up in their yards to 2.44 m?
> ❑ No ❑ Yes

11. Describe how the town council in *Example 1* could conduct a valid survey to collect people's opinions about the arena.

14. **Literacy in Math** Create a concept map describing a good survey. Copy and complete this map, adding explanations of each component as well as any additional features you feel are missing, or create your own map.



Answers:

1. **a)** 20%
b) 90%
c) About 58%

2. **a)** 35 people
b) 93 people
c) 118 people
d) 1036 people

3. **a)** About 11%
b) About 7%
c) About 4%

5. **a)** Biased
b) Unbiased
c) Unbiased
d) Biased

6. **i)** Biased
ii) Unbiased

4.3 The Use and Misuse of Statistics

Learning Goals: I am learning to…
- ☐ Assess data presented in graphical form
- ☐ Assess data based on how it was collected
- ☐ Assess data based on cause and effect assumptions

A **valid conclusion** is one that is supported by _____ data that has been _____ appropriately.

When you read a conclusion, someone has made based on statistics, you must decide whether the conclusion is valid or not. To do this, ask yourself:

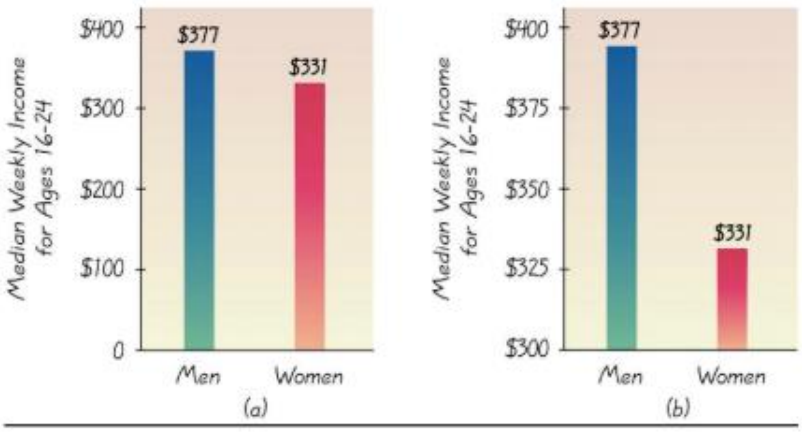
- Is there any _____ in the data collection, in the way the
  - Sample was _____
  - Questions were _____ or _____
  - Survey was _____?
- If the data involved _____, were they accurate?
- Are any _____ drawn accurate or do they _____ the viewer?

Part A: Assessing Graphs

The graphs in each pair show the SAME DATA. Choose the graph that displays the data MORE accurately. Justify your choice.

1. The following graphs both compare men and women's weekly income.

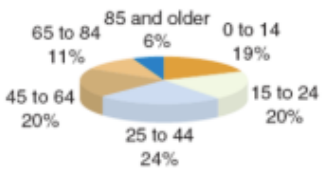
Justification:



2. The following pie charts show Canada's population by age, according to the 2001 census.

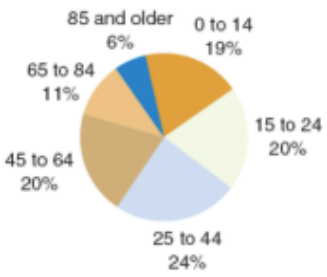
Justification:

i) Ages of Canadians, 2001 Census



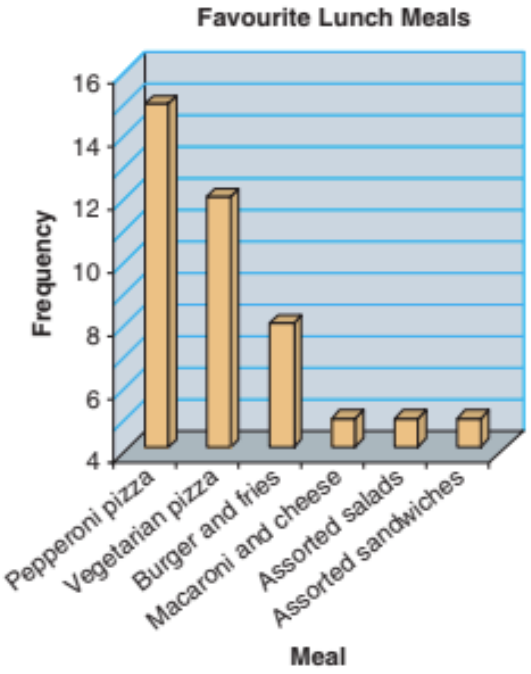
OR

ii) Ages of Canadians, 2001 Census



Part B: Assessing How Data was Collected and Graphed

Four Grade 9 students collected data on school lunch preferences. Is their conclusion valid?



Their conclusion:

> 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.

In order to decide if a conclusion is valid, ask yourself the following questions.

- Was the sample size appropriate?
- Was the sample representative?
- Was the survey question biased?
- How was the survey conducted?
- Is the graph constructed accurately?

1. Was the sample size appropriate?

2. Was the sample representative?

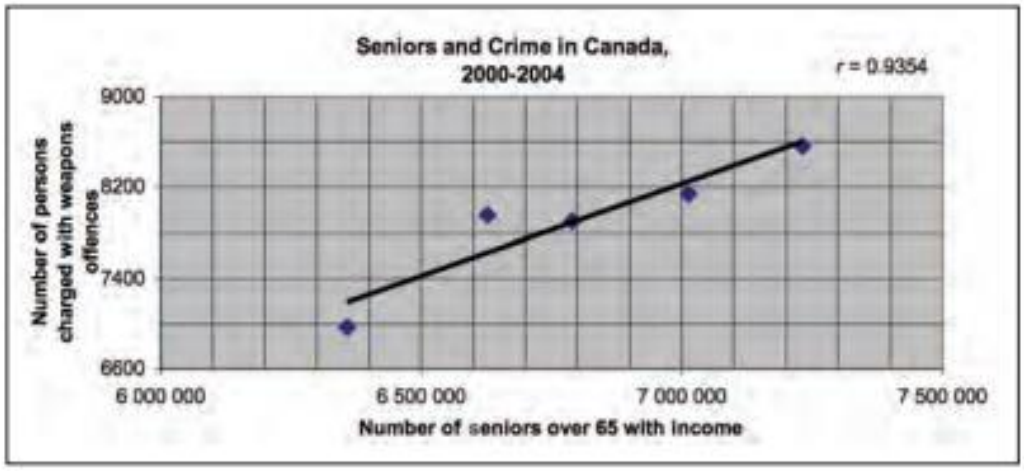
3. Was the survey question biased?

4. How was the survey conducted?

5. Is the graph constructed accurately?

Part C: 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. Is their conclusion valid?



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.

Conclusion:

To assess the validity of the conclusion, ask yourself the following questions:

- Was there bias in the data collection?
- Is the graph constructed accurately?
- Is the correlation strong?
- Does the analysis support a cause-and-effect relationship?

1. Was there bias in the data collection?

2. Is the graph constructed accurately?

3. Is the correlation strong?

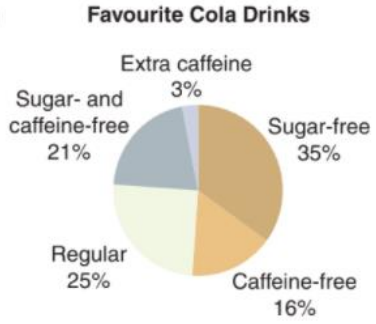
4. Does the analysis support a cause-and-effect relationship?

4.3 The Use and Misuse of Statistics Homework

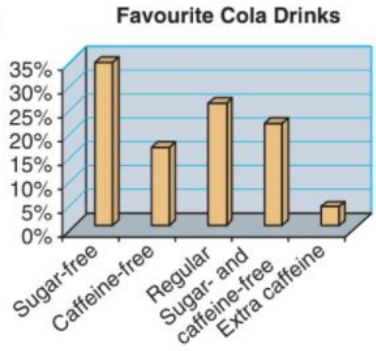
1. The graphs in each pair show the same data.
Choose the graph that displays the data more accurately.

a) Favourite cola drinks of 95 shoppers in a city mall

i)

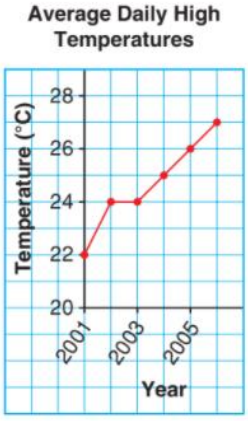


ii)

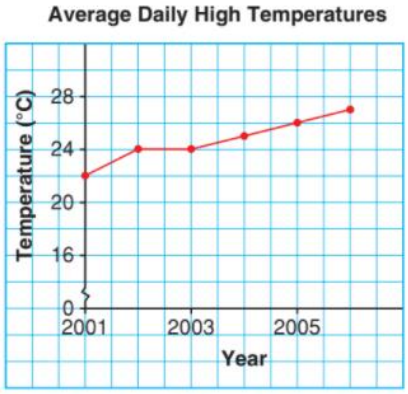


b) Temperature change over time

i)



ii)



2. For each part in question 1, describe both graphs. What features misrepresent the data in the graph that represents the data *less* accurately?

3. For each survey, who do you think would be less biased in collecting data?

a) A survey on recycling rates for plastic bottles

i) A bottled water manufacturer **ii)** A town's public works department

b) A survey on people's opinions about health care

i) A college student doing a project **ii)** A group of doctors

4. A soft-drink company wants to test consumers' reaction to a new soft drink. Which group should collect the data? Why?

i) The sales and marketing department of the company

5. Decide whether you would expect there to be a correlation between each pair of variables.

a) The numbers of students and teachers in a school

b) The number of years a person has worked for a company and the number of vacation days he or she receives each year

c) The population of a town and the amount of precipitation the town receives each year

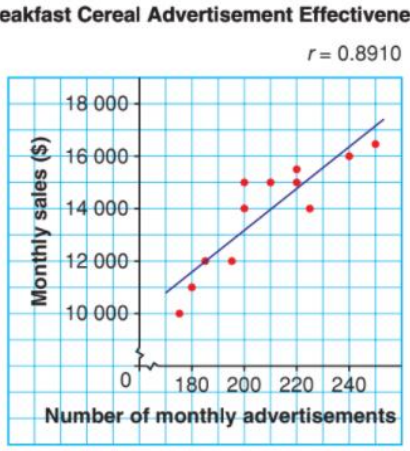
d) A person's height and her or his mark in mathematics

7. **Literacy in Math** What additional information would you need before deciding whether each statistical analysis is valid?

a) The host of a TV infomercial demonstrates a cleaning product. Then a man in a laboratory coat says, "Studies have shown that this product eliminates more bacteria from household surfaces than the leading brands."

b) You research athletic shoes on the Internet before you purchase a new pair. On one site, a pop-up advertisement displays results from an online survey in a bar graph. The graph shows that people prefer shoes made by Robur to those made by several other brands.

8. Which of the following statements best describes the information in the scatter plot? Justify your choice.



i) There is no correlation between the number of advertisements shown per month and the monthly cereal sales.

ii) There is a strong positive correlation between the number of advertisements shown per month and monthly cereal sales.

iii) As the number of advertisements shown per month increases, cereal sales increase.

9. A reporter from a TV news show asks 5 people on the street this question: "In light of the many recent home invasions, do you think police are doing all they can to keep us safe?"
Four of those interviewed say the police are not keeping us safe.
On the news that evening, the reporter announces, "4 out of 5 citizens are worried about personal safety," and then shows the interviews. What is wrong with this statistical analysis?

Answers:

1. **a)** Part i
b) Part ii

2. **a)** The data on the bar graph is difficult to read accurately.
b) The vertical scale makes changes in temperature look more dramatic than they really are.

3. **a)** Part ii
b) Part i

4. Part ii; an outside agency will have a more objective (unbiased) opinion. It will also have the expertise required in statistical surveys.

5. **a)** Yes **b)** Yes
c) No **d)** No

7. **a)** Additional information about the person's credentials.
b) Additional information about the party who conducted the study, so that one can decide about its bias.

8. Parts ii and iii

9. Misleading report; the sample size is too small and the survey question is biased.

4.4 Understanding Indices

Learning Goals: I am learning to…
- ☐ Interpret data to determine the price indices and price index
- ☐ Interpret, analyze and explain the meaning of data that has been presented

Price Indices help citizens, businesses and industries follow and _____ in prices.

A **price index** describes the price of an item compared to a _____ measured at a particular time or in a particular place.

Statistics Canada tracks price changes using several different indices. The most important is the **CPI,** known as the _____ _____ _____.

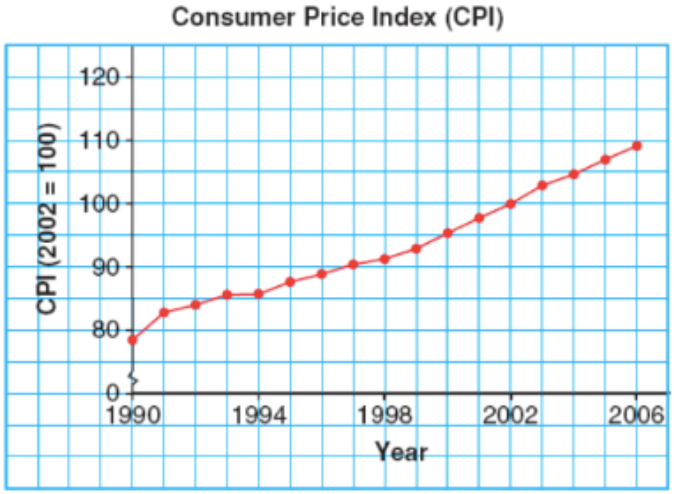
To determine the CPI, Statistics Canada collects thousands of price quotations from across the country for a basket of about 600 popular consumer goods and services, from French fries and bus fares to tuition and Internet services.

Example 1: Use the CPI graph on the right to answer the following questions.

a) What is the base year for the CPI?

b) In what year was the cost of the basket of goods about 90% of the base cost?

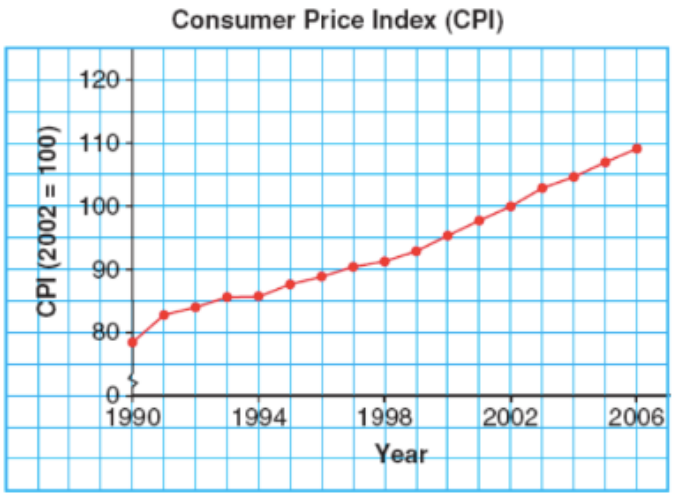
c) What was the CPI in 1990? What does this mean?



d) Describe the change in the CPI from 1990 to 1991. What do you notice about the line segment representing this period?

e) Describe the overall trend in the CPI and its significance.

Example 2: Use the same graph as example 1.



a) Calculate the average annual rate of inflation from 1990 to 2006.

b) Use your answer from part a) to predict the CPI for 2010. Justify your prediction.

Example 3: The 2006 UBS Prices and Earning report includes a comparison of clothing prices in 71 cities. The base price is the price in New York.

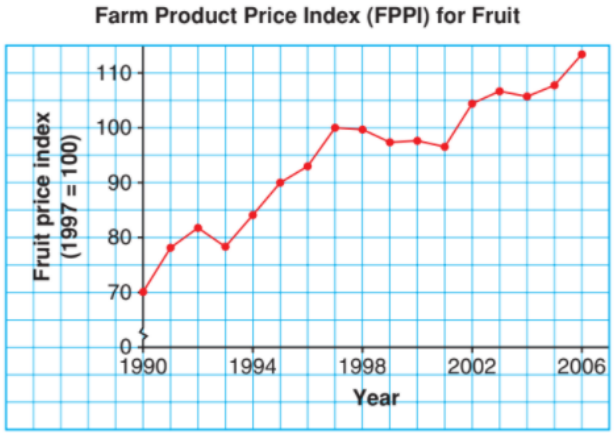
a) Which cities in this table have index values less than 100? What does this tell you?

b) How do clothing prices in Zurich and Toronto compare to clothing prices in New York?

City	Clothing Price Index (New York = 100)
Zurich	115.6
Oslo	114.4
Dublin	97.5
New York	100.0
Toronto	73.8
Tokyo	148.1
Rome	87.5
Hong Kong	75.0
Delhi	43.8

4.4 Understanding Indices Homework

1. a) What is this price index measuring?
b) What is the base year for the index?
c) Estimate the index value for each year.
i) 1994 ii) 2002

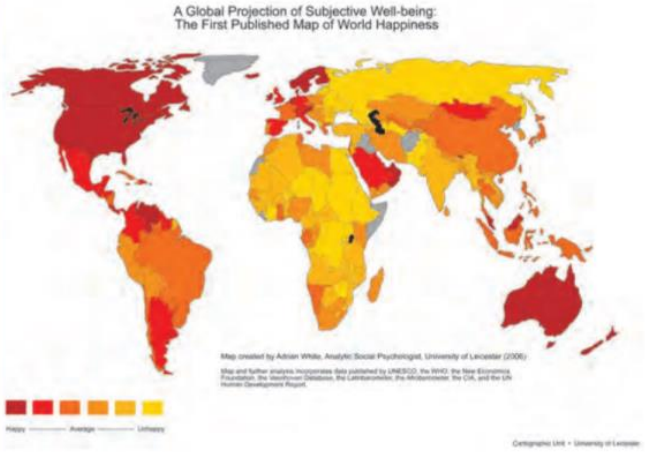


2. For each price, calculate the percent price increase from a base value of $124. Round each answer to the nearest percent.
a) $186 b) $155 c) $248 d) $131

3. For each price, calculate the percent price decrease from a base value of $124. Round each answer to the nearest percent.
a) $92 b) $62 c) $115 d) $25

4. Order these top 10 happiest countries from most to least happy.

Country	SWB Index
Austria	260
The Bahamas	257
Bhutan	253
Brunei	253
Canada	253
Denmark	273
Finland	257
Iceland	260
Sweden	257
Switzerland	273



5. Use the Consumer Price Index graph in *Example 1* to answer these questions.
a) What was the CPI in January 1996? What does this value mean?
b) What was the CPI in January 2001? What does this value mean?
c) Describe the change in the CPI from January 1996 to January 2001.
d) Calculate the average annual inflation rate from January 1996 to January 2001.

9. a) What is the base year for this index? Explain how you know.

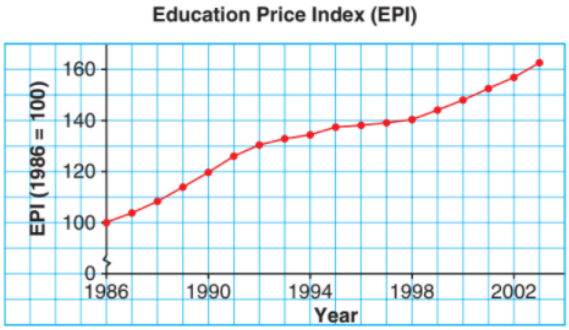
b) Estimate the EPI for each year.

i) 1992 ii) 1998 iii) 2003

c) By what percent did spending on education rise during each time period?

i) Base year to 1992 ii) 1992 to 1998 iii) 1998 to 2003

d) Compare your answers to part c. Which period had the greatest increase? Which period had the least increase? How does this relate to the line segments on the graph? Explain your thinking.



10. a) Calculate the overall change in the EPI from 1986 to 2003. What was the average rate of change per year for this 17-year period?

b) Predict the EPI for 2010 if the rate of change you determined in part a continues. Explain your method.

12. The 2006 UBS *Prices and Earnings* report compares the cost of a basket of food in 71 cities. The base cost is the cost in New York.
Data for 10 cities is given.

City	Food Price Index (New York = 100)
Zurich	115.6
Oslo	112.1
Dublin	86.6
New York	100.0
Copenhagen	99.5
Toronto	80.8
Tokyo	130.3
Rome	87.8
Hong Kong	86.6
Delhi	35.1

a) Which cities have index values greater than 100? What does this tell you about food prices in these cities?

b) How do food prices in Oslo and Delhi compare to food prices in New York?

c) Name a pair of cities that have similar food prices. Justify your answer.

d) Write a question someone could answer using these data. Answer the question.

Answers:

1. a) The change in the price of fruit from 1990 to 2006
b) 1997
c) i) About $84
ii) About $104

2. a) 50%
b) 25%
c) 100%
d) 6%

3. a) 26%
b) 50%
c) 7%
d) 80%

4. Denmark and Switzerland, Austria and Iceland, The Bahamas, Finland, and Sweden, Bhutan, Brunei, and Canada

5. a) 88.9%
b) 97.8%
c) CPI increased by 8.9% between January 1996 and January 2001.
d) 1.8%

9. a) 1986
b) i) 130 ii) 140 iii) 164
c) i) 30% ii) 7% iii) 17%
d) 1986 to 1992 had the greatest increase.
1992 to 1998 had the least increase.

10. a) The overall change in EPI: about 62%
The average annual change: about 3.6%
b) About 187.2%

12. a) Zurich, Oslo, Tokyo
It is more expensive to buy food in these cities than in New York.
b) It is about 12.1% more expensive to buy food at Oslo than New York. It will cost about 64.9% less to purchase food in Delhi than in New York.