

Review of Key Concepts

2.1 Data Analysis With Graphs

Refer to the Key Concepts on page 100.

- The following data show monthly sales of houses by a real-estate agency.

6	5	7	6	8	3	5	4	6
7	5	9	5	6	6	7		

- Construct an ungrouped frequency table for this distribution.
 - Create a frequency diagram.
 - Create a cumulative-frequency diagram.
- A veterinary study recorded the masses in grams of 25 kittens at birth.

240	300	275	350	280	260	320
295	340	305	280	265	300	275
315	285	320	325	275	270	290
245	235	305	265			

- Organize these data into groups.
 - Create a frequency table and histogram.
 - Create a frequency polygon.
 - Create a relative-frequency diagram.
- A class of data-management students listed their favourite board games.

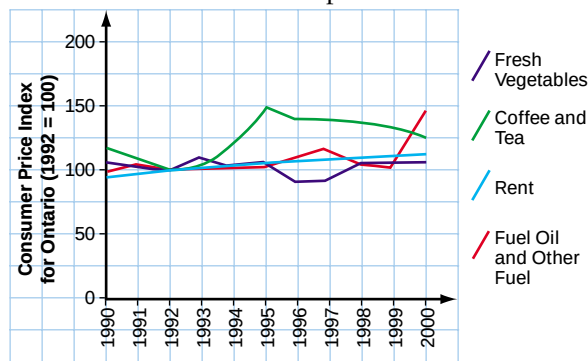
Game	Frequency
Pictionary®	10
Chess	5
Trivial Pursuit®	8
MONOPOLY®	3
Balderdash®	6
Other	4

- What type of data does this table show? Explain your reasoning.
- Graph these data using an appropriate format.
- Explain why you chose the type of graph you did.

2.2 Indices

Refer to the Key Concepts on page 109.

The following graph shows four categories from the basket of goods and services used to calculate the consumer price index.



- What is this type of graph called?
 - Which of the four categories had the greatest increase during the period shown?
 - Why do all four graphs intersect at 1992?
 - Which category was
 - the most volatile?
 - the least volatile?
 - Suggest reasons for this difference in volatility.
- If a tin of coffee cost \$5.99 in 1992, what would you expect it to cost in
 - 1995?
 - 1990?
 - What rent would a typical tenant pay in 2000 for an apartment that had a rent of \$550 per month in 1990?
 - What might you expect to pay for broccoli in 2000, if the average price you paid in 1996 was \$1.49 a bunch?

2.3 Sampling Techniques

Refer to the Key Concepts on page 116.

6.
 - a) Explain the difference between a stratified sample and a systematic sample.
 - b) Describe a situation where a convenience sample would be an appropriate technique.
 - c) What are the advantages and disadvantages of a voluntary-response sample?
7. Suppose you are conducting a survey that you would like to be as representative as possible of the entire student body at your school. However, you have time to visit only six classes and to process data from a total of 30 students.
 - a) What sampling technique would you use?
 - b) Describe how you would select the students for your sample.
8. Drawing names from a hat and using a random-number generator are two ways to obtain a simple random sample. Describe two other ways of selecting a random sample.

2.4 Bias in Surveys

Refer to the Key Concepts on page 122.

9. Identify the type of bias in each of the following situations and state whether the bias is due to the sampling technique or the method of data collection.
 - a) A survey asks a group of children whether or not they should be allowed unlimited amounts of junk food.
 - b) A teacher asks students to raise their hands if they have ever told a harmless lie.

- c) A budding musician plays a new song for family members and friends to see if it is good enough to record professionally.
 - d) Every fourth person entering a public library is asked: “Do you think Carol Shields should receive the Giller prize for her brilliant and critically acclaimed new novel?”
10. For each situation in question 9, suggest how the statistical process could be changed to remove the bias.

2.5 Measures of Central Tendency

Refer to the Key Concepts on page 133.

11.
 - a) Determine the mean, median, and mode for the data in question 1.
 - b) Which measure of central tendency best describes these data? Explain your reasoning.
12.
 - a) Use your grouped data from question 2 to estimate the mean and median masses for the kittens.
 - b) Determine the actual mean and median masses from the raw data.
 - c) Explain any differences between your answers to parts a) and b).
13.
 - a) For what type of “average” will the following statement always be true? “There are as many people with below-average ages as there are with above-average ages.”
 - b) Is this statement likely to be true for either of the other measures of central tendency discussed in this chapter? Why or why not?

14. Angela is applying to a university engineering program that weights an applicant's eight best grade-12 marks as shown in the following table.

Subjects	Weighting
Calculus, chemistry, geometry and discrete mathematics, physics	3
Computer science, data management, English	2
Other	1

Angela's grade-12 final marks are listed below.

Subject	Mark	Subject	Mark
Calculus	95	Computer science	84
English	89	Chemistry	90
Geometry and discrete mathematics	94	Mathematics of data management	87
Physical education	80	Physics	92

- a) Calculate Angela's weighted average.
 b) Calculate Angela's unweighted average.
 c) Explain why the engineering program would use this weighting system.
15. Describe three situations where the mode would be the most appropriate measure of central tendency.

2.6 Measures of Spread

Refer to the Key Concepts on page 147.

16. a) Determine the standard deviation, the interquartile range, and the semi-interquartile range for the data in question 1.
 b) Create a box-and-whisker plot for these data.
 c) Are there any outliers in the data? Justify your answer.

17. a) Explain why you cannot calculate the semi-interquartile range if you know only the difference between either Q_3 and the median or median and Q_1 .
 b) Explain how you could determine the semi-interquartile range if you did know both of the differences in part a).

18. a) For the data in question 2, determine
 i) the first and third quartiles
 ii) the 10th, 25th, 75th, and 90th percentiles
 b) Would you expect any of the values in part a) to be equal? Why or why not?

19. The scores on a precision-driving test for prospective drivers at a transit company have a mean of 100 and a standard deviation of 15.

- a) Determine the z -score for each of the following raw scores.
 i) 85 ii) 135 iii) 100 iv) 62
 b) Determine the raw score corresponding to each of the following z -scores.
 i) 1 ii) -2 iii) 1.5 iv) -1.2

20. Dr. Simba's fourth-year class in animal biology has only 12 students. Their scores on the midterm examination are shown below.

50	71	65	54	84	69	82
67	52	52	86	85		

- a) Calculate the mean and median for these data. Compare these two statistics.
 b) Calculate the standard deviation and the semi-interquartile range. Compare these statistics and comment on what you notice.
 c) Which measure of spread is most suitable for describing this data set? Explain why.