3.7 CHAPTER SUMMARY

This chapter introduced ***descriptive statistics***—ways of *describing* data to summarize key characteristics of the data.

3.7.1 Overview of Tables and Graphs

- A ***frequency distribution*** summarizes the counts for possible values or intervals of values. A ***relative frequency*** distribution reports this information using percentages or proportions.
- A ***bar graph*** uses bars over possible values to portray a frequency distribution for a categorical variable. For a quantitative variable, a similar graphic is called a ***histogram***. It shows whether the distribution is approximately bell shaped, U shaped, skewed to the right (longer tail pointing to the right), or whatever.
- The ***stem-and-leaf plot*** is an alternative portrayal of data for a quantitative variable. It groups together observations having the same leading digit (stem), and shows also their final digit (leaf). For small samples, it displays the individual observations.
- The ***box plot*** portrays the quartiles, the extreme values, and any outliers. The box plot and the stem-and-leaf plot also can provide back-to-back comparisons of two groups.

Stem-and-leaf plots and box plots, simple as they are, are relatively recent innovations in statistics. They were introduced by the great statistician John Tukey (see Tukey 1977), who also introduced the terminology "software." See Cleveland (1994) and Tufte (2001) for other innovative ways to present data graphically.

3.7.2 Overview of Measures of Center

Measures of center describe the center of the data, in terms of a typical observation.

- The ***mean*** is the sum of the observations divided by the sample size. It is the center of gravity of the data.
- The ***median*** divides the ordered data set into two parts of equal numbers of observations, half below and half above that point.

- The lower quarter of the observations fall below the ***lower quartile***, and the upper quarter fall above the ***upper quartile***. These are the 25th and 75th ***percentiles***. The median is the 50th percentile. The quartiles and median split the data into four equal parts. They are less affected than the mean by outliers or extreme skew.
- The ***mode*** is the most commonly occurring value. It is valid for any type of data, though usually used with categorical data or discrete variables taking relatively few values.

3.7.3 Overview of Measures of Variability

Measures of variability describe the spread of the data.

- The ***range*** is the difference between the largest and smallest observations. The ***interquartile range*** is the range of the middle half of the data between the upper and lower quartiles. It is less affected by outliers.
- The ***variance*** averages the squared deviations about the mean. Its square root, the ***standard deviation***, is easier to interpret, describing a typical distance from the mean.
- The ***Empirical Rule*** states that for a bell-shaped distribution, about 68% of the observations fall within one standard deviation of the mean, about 95% fall within two standard deviations, and nearly all, if not all, fall within three standard deviations.

Table 3.9 summarizes the measures of center and variability. A ***statistic*** summarizes a sample. A ***parameter*** summarizes a population. ***Statistical inference*** uses statistics to make predictions about parameters.

TABLE 3.9: Summary of Measures of Center and Variability

Measure	Definition	Interpretation
Center		
Mean	$\bar{y} = \Sigma y_i/n$	Center of gravity
Median	Middle observation of ordered sample	50th percentile, splits sample into two equal parts
Mode	Most frequently occurring value	Most likely outcome, valid for all types of data
Variability		
Standard deviation	$s = \sqrt{\Sigma(y_i - \bar{y})^2/(n - 1)}$	Empirical Rule: If bell shaped, 68%, 95% within $s, 2s$ of $\bar{y}$
Range	Difference between largest and smallest observation	Greater with more variability
Interquartile range	Difference between upper quartile (75th percentile) and lower quartile (25th percentile)	Encompasses middle half of data

3.7.4 Overview of Bivariate Descriptive Statistics

Bivariate statistics are used to analyze data on two variables together.

- Many studies analyze how the outcome on a ***response variable*** depends on the value of an explanatory variable.
- For categorical variables, a ***contingency table*** shows the number of observations at the combinations of possible outcomes for the two variables.
- For quantitative variables, a ***scatterplot*** graphs the observations, showing a point for each observation. The response variable is plotted on the y-axis and the explanatory variable is plotted on the x-axis.
- For quantitative variables, the ***correlation*** describes the strength of straight-line association. It falls between -1 and $+1$ and indicates whether the response variable tends to increase (positive correlation) or decrease (negative correlation) as the explanatory variable increases.
- A ***regression analysis*** provides a straight-line formula for predicting the value of the response variable using the explanatory variable. We study correlation and regression in detail in Chapter 9.

GETTING STARTED

The first step in using performance tasks is to find or create some useful tasks. On the right and in the following section are a few examples of the kind of task or question that might be used for classroom work or for assessment.

Page 16 lists criteria for a good performance task. Page 17 describes a process for task development, originated by the Connecticut State Department of Education.

A quick source of tasks is the textbook. Consider how students can demonstrate their understanding of a concept—through manipulatives, class presentations, physical construction of mathematical models, journals, bulletin boards, or even music and drama.

The concluding question of a textbook word problem can be changed to elicit further explanation. Instead of "How many were left?," the directions can become "Explain how you would find out how many were left," or "Explain how you might check your answer."

Other good sources of tasks are some of the many existing books of problems listed in catalogs by teaching supply outlets. Open the problems, if necessary, by changing the question, by introducing additional information or extra questions, or by deleting some necessary information. See page 31 of this book for a list of possible questions.

Look also at the many ideas in the *Arithmetic Teacher* or the *Mathematics Teacher.* Almost every section of the *Curriculum and Evaluation Standards* (NCTM 1989) is full of good problems. Look for those that are best solved or completed using a means other than paper and pencil.

It is vital that students learn to set up their own problem-solving strategies, decide how to organize data, identify their own mistakes, and demonstrate their own thinking as much as possible. It may be difficult at the start, but if we persist, they will amaze us with their ideas.

SAMPLE TASKS

Fractions

Ask third-grade students who are learning about fractions to show you with manipulatives how they would divide different items, such as 5 candy bars, 10 pencils, or 11 comic books, among 4 students.

Place Value

Have students explain how they would teach a younger sibling to understand the meaning of tens and ones in place value.

Long Division

Instead of a traditional division test with twenty items, give each group of students a different problem. Ask the groups to make posters for the class that explain the methods they tried in solving their problem.

Pi—Organizing and Displaying Data

Ask a group of students to find and demonstrate the value of pi by measuring diameters and circumferences, expressing the ratio, and finding decimal equivalents on a calculator. Allow the students to choose a way to explain and display their findings.

Functions, Data, and Computer Use

Give students the task of finding the function that "best fits" certain data by using computer software and a printout of their results.

LOGISTICS

We should not feel that we must observe or evaluate every idea or every student on the same day.

Start by working on or assessing a single simple but important idea. For example, for several days we might want to check only to see whether students can paraphrase the problem they are working on, teaching in our usual way the rest of the time. The next skill to observe might be whether students can formulate a plan.

We can make notes as we happen to see other qualities, making sure that a variety of assessment techniques are used and that the goals of the *Curriculum and Evaluation Standards* (NCTM 1989) are well represented. Included in this section are some checklists showing objectives that can be assessed through performance tasks.

Some teachers carry a clipboard with self-stick labels (with student names already entered if you have that luxury!) so that notes can be entered easily into individual student records.

Another system is to carry a class list and make a note of anything we notice about a student, looking for the unusual rather than the expected. We especially take note of those students for whom nothing is written down—we may want to take a special look at how they are doing. We might write a summary paragraph about each student once a week or every two weeks. It's a good idea to be alert for things we think the student herself should know about, as well as our own information.

"It is very easy to think that every piece of work that the students carry out should be assessed. This dramatically increases the teacher's workload and the student's stress level, and it does not necessarily produce a more effective assessment of students than can be achieved by carrying out a more selective assessment programme."

(New Zealand Department of Education, 1988, Section 1.6)

ANOTHER SAMPLE TASK

Data Collection

How many bicycles are there within two miles of this school? Your group's task is to make a plan for investigating this question and to prepare an oral report, with overheads or other displays, for the class. Your planning report is due in three days. Please keep a daily log of your work. Final reports will be due two weeks from today.

(Or instead of bicycles, substitute automobiles, animals, chimneys, trees, skateboards, or have students choose their own survey topic.)

Assessment Ideas

- Are students able to make a workable plan for useful data collecting?
- Do they have a hypothesis?
- Do they effectively use statistical ideas, such as matrix sampling or surveys?
- Have they used outside resources, such as the library, computers, or the telephone?
- Can they justify their choice of techniques?
- Have they organized their information in a reasonable manner?
- Do they compare their results with their hypothesis?
- Can they communicate their ideas and their results to the class orally and with visual aids?
- Has every member of the group contributed?
- Do they go beyond the immediate problem and pose new questions?

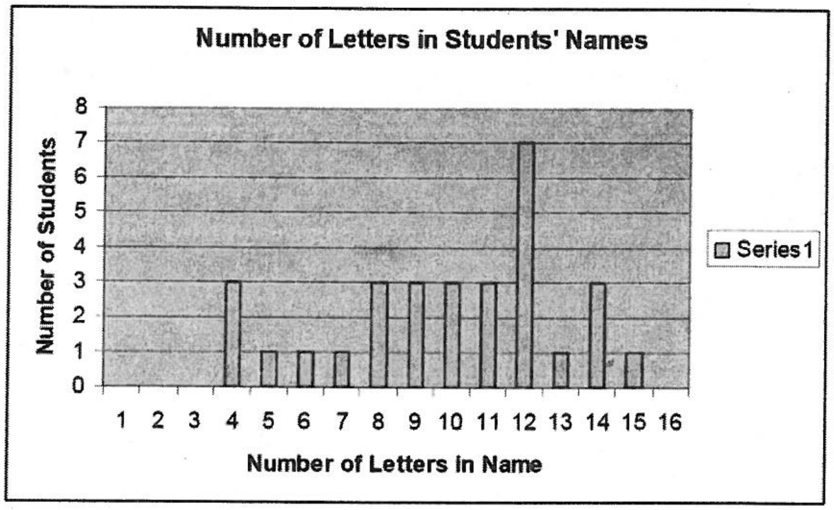
Levels of Abilities to Read Graphs

There are varying levels of abilities to read graphs. As a teacher, you will want to help your students be able to interpret graphs at both of these levels.

Levels:

1. An elementary level which focuses on extracting individual data points from the graph
2. An overall level that involves looking at the data points on the graph in their totality, usually comparing trends or seeing groups. (In this case, the students are looking at the data points as a whole.)

Suppose students were provided with a bar graph showing the number of students with a certain number of letters in their name. The graph might look something like the following:



Determine which of the Levels of Abilities to Read Graphs are illustrated by the following questions:

a. If a new student joined our class, how many letters would you predict that student would have in her name?
b. How many students have 12 letters in their names?
c. How many students have more than 12 letters in their names?
d. Looking at the data about name lengths from several different classes of students, what kinds of patterns do you observe in name lengths across the classes?

References:

Friel, S. N., & Bright, G. W. (1996, April). *Building a theory of graphicacy: How do students read graphs?* Paper presented at the Annual Meeting of the American Educational Research Association, New York.

Curcio, F. R. (1987). Comprehension of mathematical relationships expressed in graphs. *Journal for Research in Mathematics Education, 18*(5), 382-393.

Wainer, H. (1992). Understanding graphs and tables. *Educational Researcher, 21*(1), 14-23.