PHOTOGRAPH BY G. KADER; ALL RIGHTS RESERVED

The least is made in San Frasico [*sic*] because it's farther away from us.

No, I think somebody else would have different result [*sic*] because I think they would have more of the coins made in Denver if their [*sic*] from the central U.S.

I think ones that have no symbol is [*sic*] 22%. The coins that are made in Denver are 26%.

Students gave several different responses to the fourth question, along with a variety of explanations. Many students selected Philadelphia, and some of the reasons given include "That's where most of the other ones were minted" and "It is the closest." One student thought these coins were minted in Denver because "That's in the middle." A few students selected San Francisco because "We had the least coins with an S printed on them."

Clearly, a meaningful interpretation of these data depended on where the N coins were actually minted. For example, if the N coins were minted in Denver, then this fact would lead to a different interpretation than if they were minted in San Francisco. Consequently, unless the mint location for the N coins could be determined, the investigation concerning the circulation of coins would remain inconclusive.

The difficulty encountered at this stage of the investigation was a common one in statistical problem solving. The data had yielded a significant number of unanticipated measurements, since almost one-fourth of the data was in a sense "no response." To continue the original investigation, the class had to ask new questions and obtain additional data to address these new questions.

To introduce the next phase of the investigation, the teacher posed this question, "If all N coins are minted at one of the other three mint locations, which one do you think it is?"

After giving students ample time to discuss the first question, she asked, "What other information is readily available from a coin?" The only other obvious measurement possibility was the year a coin was minted. It was not clear, however, that "*where* the coin was minted" should be related to "*when* the coin was minted." At this stage of their investigation, the class had come full circle and were ready to ask two new questions.

Second Round

Questions

- When were the coins in this collection minted?
- Are when and where the coins were minted related? Specifically, does when the coins were minted tell us anything about where the N coins were minted?

Data

The class divided into three groups—about one-half the class worked with the P coins, one-fourth with the D coins, and one-fourth with the N coins, while the teacher worked with the S coins. In each group students worked in pairs and devised strategies for organizing the data on the year each coin was minted.

Analysis

Before beginning the analysis, the class discussed and compared the measurement scales for the two variables of mint location and mint year. Most students agreed that the mint location was nominal data, that is, objects or people grouped into classes on the basis of a common characteristic. However, the mint year was ordinal data. Unlike nominal data, summary measures, such as the minimum, maximum, median, and so forth, could be determined. These statistical measures along with graphs would prove useful when comparing and contrasting two or more groups.

Each group of students combined their results and recorded the number, or frequency, of coins minted for each year. Then they recorded the frequency, cumulative frequency, and relative cumulative frequency on a summary sheet. Note that the cumulative frequency is simply the total number of coins with a mint year less than or equal to the specified year. The relative cumulative frequency is the ratio of the number of coins found for the cumulative frequency to the total number of coins in a sample expressed as percent. In **table 1,** for the

TABLE 1
Frequencies and Cumulative Frequencies for Four Mint Groups

Mint Location: Denver

Year	Frequency	Cumulative Frequency	Relative Cumulative Frequency
1968	1	1	0.1%
1969	1	2	1.6%
1970	9	11	8.6%
1971	7	18	14.1%
1972	7	25	19.5%
1973	7	32	25.0%
1974	3	35	27.3%
1975	4	39	30.5%
1976	5	44	34.4%
1977	7	51	39.8%
1978	6	57	44.5%
1979	6	63	49.2%
1980	4	67	52.3%
1981	2	69	53.9%
1982	4	73	57.0%
1983	7	80	62.5%
1984	5	85	66.4%
1985	10	95	74.2%
1986	5	100	78.1%
1987	6	106	82.8%
1988	5	111	86.7%
1989	3	114	89.1%
1990	2	116	90.6%
1991	3	119	93.0%
1992	3	122	95.3%
1993	2	124	96.9%
1994	2	126	98.4%
1995	2	128	100.0%

Mint Location: None

Year	Frequency	Cumulative Frequency	Relative Cumulative Frequency
1971	4	4	3.1%
1972	9	13	10.1%
1973	13	26	20.2%
1974	23	49	38.0%
1975	8	57	44.2%
1976	16	73	58.9%
1977	27	100	77.5%
1978	12	112	86.8%
1979	17	129	100.0%

Mint Location: Philadelphia

Year	Frequency	Cumulative Frequency	Relative Cumulative Frequency
1980	17	17	5.9%
1981	6	23	8.0%
1982	5	28	9.7%
1983	10	38	13.1%
1984	23	61	21.1%
1985	22	83	28.7%
1986	16	99	34.3%
1987	5	104	36.0%
1988	14	118	40.1%
1989	58	176	60.9%
1990	53	229	79.2%
1991	27	256	88.6%
1992	13	269	93.1%
1993	5	274	94.8%
1994	12	286	99.0%
1995	3	289	100.0%

Mint Location: San Francisco

Year	Frequency	Cumulative Frequency	Relative Cumulative Frequency
1968	1	1	25.0%
1969	2	3	75.0%
1970	1	4	100.0%

year 1983 for coins minted in Philadelphia, the cumulative frequency is 38 and the relative cumulative frequency is 13.1 percent ($38/289 \times 100$). The relative cumulative frequency is useful in finding the median of a large set of data. The results for the four different mint locations are summarized in **table 1.**

Each group determined the five-number summary (minimum, lower quartile, median, upper quartile, and maximum) for the dates on the Jefferson nickels. The teacher reminded them that the lower quartile had approximately 25 percent of the data values below it and 75 percent above it; the median, approximately 50 percent below and 50 percent above; and the upper quartile, approximately 75 percent below and 25 percent above.

For small data sets, the usual method for finding these statistics is to list the data in order and to find the median first by counting from both ends until the middle value is reached, when the number of