

53	8
54	1 2 6
55	0 4 5 5
56	1 1 2 3 5 8 8 9
57	0 2 3 5 5 8 8 8 9 9
58	8 8
59	1 4

Fig. 3 Stem-and-leaf plot for thirty head circumferences

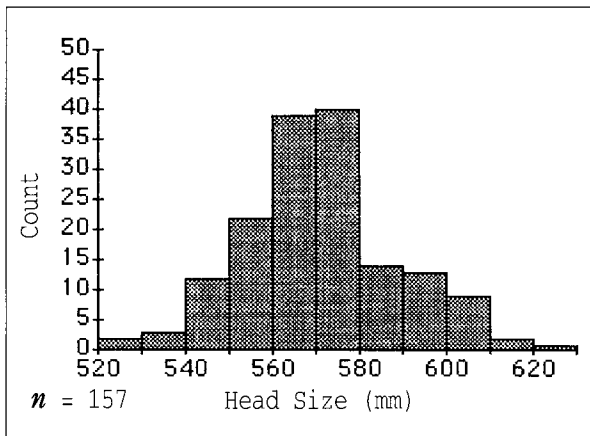


Fig. 4 Histogram produced by Statistics Workshop of 157 head circumferences

number of measurements. The most common stems are 56 and 57, with 8 and 10 leaves, respectively. So, just as with the line plot, we observe that eighteen of the thirty circumferences are between 560 mm and 580 mm.

Expanded study

Making a decision about how many of each-size hat should be ordered would require a more in-depth study. Certainly, data should be collected from more people. A more extensive study might also include information on such additional characteristics as sex, age, and so on. Measurements of head circumference were collected from several other classes of college students, along with those of the instructors, and combined with the original data set for a total of 157 measurements.

Let's continue our computer analysis of the larger data set using Statistics Workshop. The *frequency histogram* produced by Statistics Workshop is shown in **figure 4**. We notice that the interval lengths are the same as those used in the stem-and-leaf plot previously shown for the smaller data set. The range of values for the larger data set is somewhat wider than the range for the smaller data set, with all values between 520 and 630. Also, like the previous analysis, this graph suggests a pattern in the head circumferences for the entire group that is similar to that of the smaller group—the values in the middle of

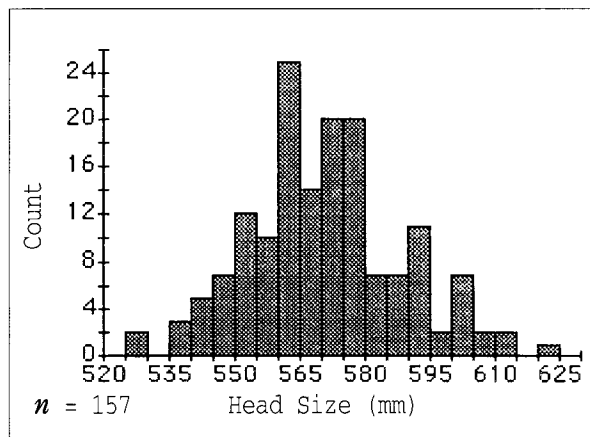


Fig. 5 Histogram with interval of length 5 of 157 head circumferences

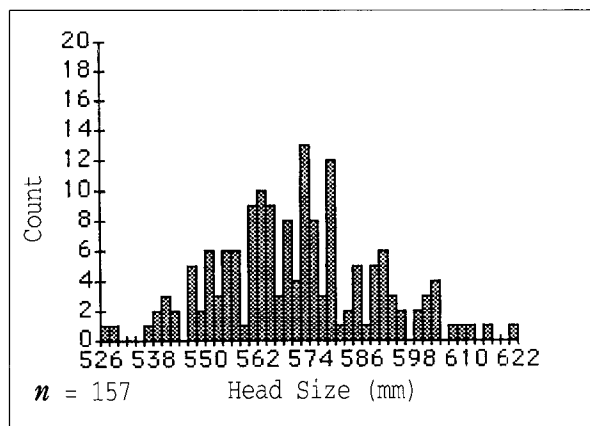


Fig. 6 Histogram with interval of length 2 of 157 head circumferences

the data are more frequent, whereas those in the ends are less frequent.

A primary reason for using computer software in statistical problem solving is to allow us more easily to address "What if?" questions. For example, what if hat sizes changed for each 5-mm increase in head circumference instead of for each 10-mm increase? For each 2-mm increase? The frequency histograms produced by Statistics Workshop using intervals of length 5 and length 2 are shown in **figures 5 and 6**. **Figures 5 and 6** furnish more detail about the specific data values but less information about the overall pattern in the distribution. If an interval of length 1 was used, then the histogram would be equivalent to the line plot. Note in **figure 6** that Statistics Workshop forced a change in the beginning and ending points in the histogram.

What happens if the interval length is increased? The frequency histograms using intervals of length 20 and length 50 are shown in **figures 7 and 8**. Notice that **figures 7 and 8** supply much less information about the specific data values as well as less information about the

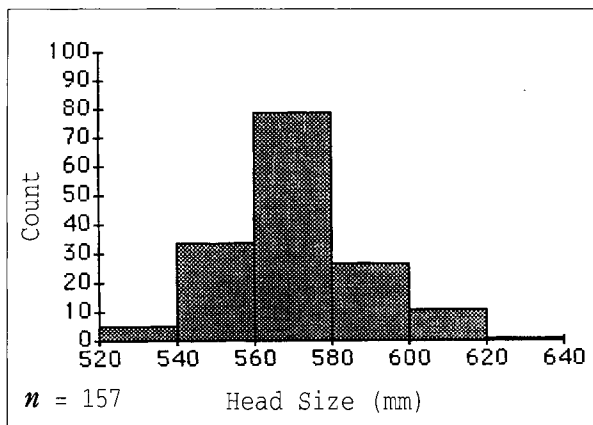


Fig. 7 Histogram with interval of length 20 of 157 head circumferences

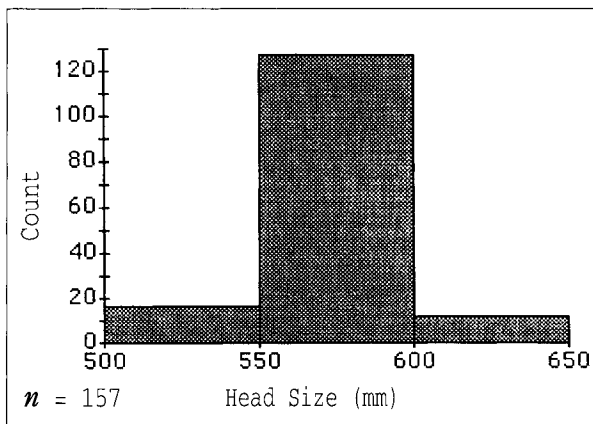


Fig. 8 Histogram with interval of length 50 of 157 head circumferences

overall pattern in the distribution. If an interval of length 100 was used beginning at 525 and ending at 625, then the histogram would simply be one rectangle and its count, or frequency, would be 157. Also note that the beginning and ending points in the histogram in **figure 8** are different from those used in **figure 4**.

In the study of head sizes as they relate to hat sizes, the most appropriate interval length is 10 mm, which conveniently corresponds to the most appropriate histogram. However, studying the effects of different-length intervals on the shape of a histogram is an important component in statistical education. Computer software allows these concepts to be explored and developed without the tedious work required in finding the frequency tables and associated histograms by hand.

The reader may have noticed that all histograms produced by Statistics Workshop are frequency histograms. Actually, Statistics Workshop allows for relative-frequency histograms; however, the vertical axis is always scaled from 0 to 100 percent in incre-

TABLE 2

INTERVAL	FREQUENCY	RELATIVE FREQUENCY (PERCENT)
520-529	2	1.3
530-539	3	1.9
540-549	12	7.6
550-559	22	14.0
560-569	39	24.8
570-579	40	25.5
580-589	14	8.9
590-599	13	8.3
600-609	9	5.7
610-619	2	1.3
620-629	1	0.6
Total:	157	99.9

TABLE 3

HAT SIZE	FITS HEAD CIRCUMFERENCE (MM)	NUMBER TO ORDER
520	< 520	0
530	520-529	13
540	530-539	19
550	540-549	76
560	550-559	140
570	560-569	249*
580	570-579	255
590	580-589	89
600	590-599	83
610	600-609	57
620	610-619	13
630	620-629	6

**Arbitrarily raised this total by 1 to get to 1000 hats*

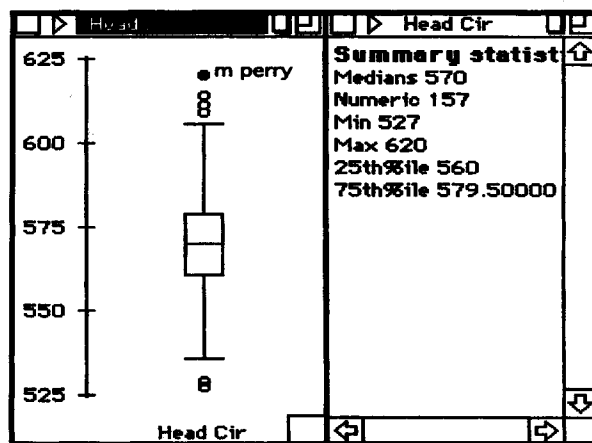


Fig. 9 Box plot and five-number summary for head circumferences