

How much disk space does your music use? Here are the files sizes (in megabytes) for 18 randomly selected files on Tim's mp3 player:

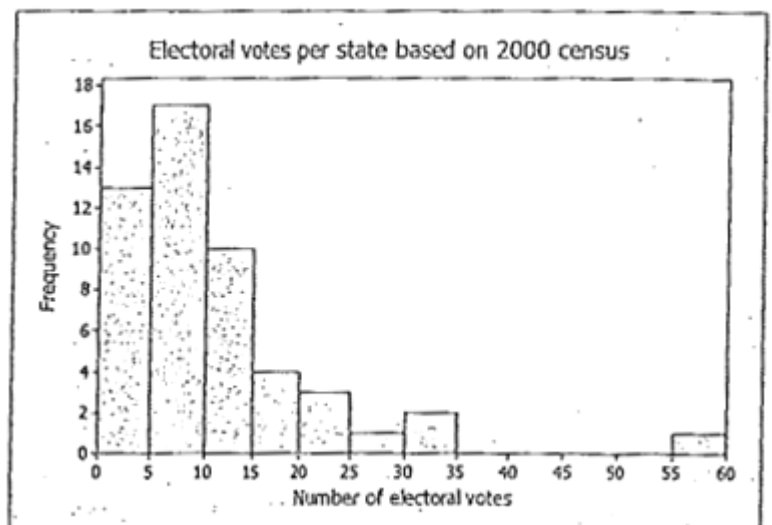
1.1 1.3 1.3 1.6 1.9 1.9 2.1 2.2 2.4
2.5 2.7 3.0 4.4 4.7 5.0 5.6 6.2 7.5

#1. Make a dotplot of these data.

#2. Find the mean, standard deviation, and 5-number summary (min, Q1, median, Q3, max).

#3. Describe the overall pattern of the distribution and any possible outliers.

#4. The histogram shows the distribution of electoral votes for the 50 United States and the District of Columbia. Describe the shape, center, and spread of the distribution.



Of the 50 species of oaks in the United States, 28 grow on the Atlantic coast and 11 grow in California. We are interested in the distribution of acorn volumes among oak species. Here are back-to-back stemplots on the volumes of acorns (in cubic centimeters) for these 39 oak species:

Volume of Acorns (cubic centimeters)

<u>Atlantic coast</u>		<u>California</u>
9 9 8 6 4 3	0	4
8 8 8 6 4 2 1 1 1 1 1	1	0 6
5 0	2	0 6
6 6 4 0	3	
8	4	1
	5	5 9
	6	0
	7	1
1	8	
1	9	
5	10	
	11	
	12	
	13	
	14	
	15	
	16	
	17	1

Key: 2|6 = 2.6 cm³

#5. Use the stemplots to compare the distribution of acorn sizes between Atlantic Coast and California oak species.

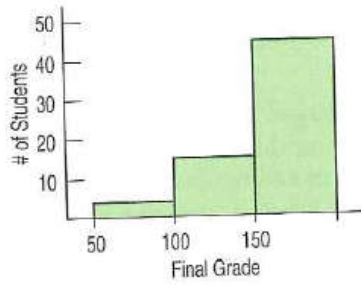
3. **Thinking about shape.** Would you expect distributions of these variables to be uniform, unimodal, or bimodal? Symmetric or skewed? Explain why.
- The number of speeding tickets each student in the senior class of a college has ever had.
 - Players' scores (number of strokes) at the U.S. Open golf tournament in a given year.
 - Weights of female babies born in a particular hospital over the course of a year.
 - The length of the average hair on the heads of students in a large class.

12. **The Great One.** During his 20 seasons in the NHL, Wayne Gretzky scored 50% more points than anyone who ever played professional hockey. He accomplished this amazing feat while playing in 280 fewer games than Gordie Howe, the previous record holder. Here are the number of games Gretzky played during each season:

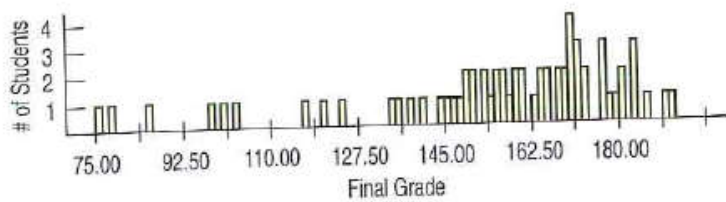
79, 80, 80, 80, 74, 80, 80, 79, 64, 78, 73, 78, 74, 45, 81, 48,
80, 82, 82, 70

- Create a stem-and-leaf display for these data using split stems.
- Describe the shape of the distribution.
- Describe the center and spread of this distribution.
- What unusual feature do you see? What might explain this?

25. **Final grades.** A professor (of something other than Statistics!) distributed the following histogram to show the distribution of grades on his 200-point final exam. Comment on the display.



27. **Final grades revisited.** After receiving many complaints about his final grade histogram from students currently taking a Statistics course, the professor from Exercise 25 distributed the following revised histogram.



- Comment on this display.
- Describe the distribution of grades.

38. **Drunk driving.** Accidents involving drunk drivers account for about 40% of all deaths on the nation's highways. The table tracks the number of alcohol-related fatalities for 20 years.

Year	Deaths (thousands)	Year	Deaths (thousands)
1982	25.2	1992	17.9
1983	23.6	1993	17.5
1984	23.8	1994	16.6
1985	22.7	1995	17.2
1986	24.0	1996	17.2
1987	23.6	1997	16.5
1988	23.6	1998	16.0
1989	22.4	1999	16.0
1990	22.0	2000	16.7
1991	19.9	2001	16.7

- Create a stem-and-leaf display or a histogram of these data.
- Create a timeplot.
- Using features apparent in the stem-and-leaf display (or histogram) and the timeplot, write a few sentences about deaths caused by drunk driving.

42. **Rainmakers.** The table lists the amount of rainfall (in acre-feet) from 26 clouds seeded with silver iodide.

2745	200
1697	198
1656	129
978	119
703	118
489	115
430	92
334	40
302	32
274	31
274	17
255	7
242	4

- Why is "acre-feet" a good way to measure the amount of precipitation produced by cloud seeding?
- Plot these data, and describe the distribution.
- Create a re-expression of these data that produces a more advantageous distribution. (suggestion: try taking the \log_{10} of the rainfall values)
- Explain what your re-expressed scale means.