

## Stat 145 Homework Solutions: Chapter 1

### Problem 1.1

- a) The individuals are motor vehicles.
- b) The variables given are make and model (categorical), vehicle type (categorical), transmission type (categorical), number of cylinders (quantitative), city MPG (quantitative), and highway MPG (quantitative).

### Problem 1.3

- a) The sum of the given percentages is 88.6%, so  $100\% - 88.6\% = 11.4\%$  are some other color.
- b) Yes, it would be correct to make a pie chart if the "Other" category were included.

### Problem 1.5

Frequency table:

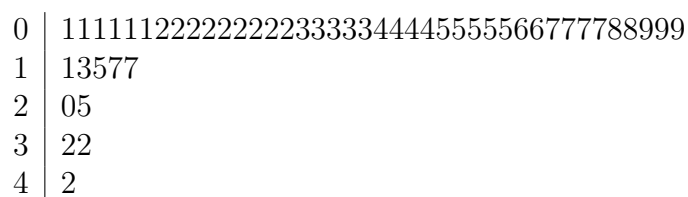
<u>Class</u>	<u>Count</u>
10 - 14	1
15 - 19	2
20 - 24	6
25 - 29	9
30 - 34	3
35 - 39	0
40 - 44	0
45 - 49	0
50 - 54	0
55 - 59	1

Problem 1.7

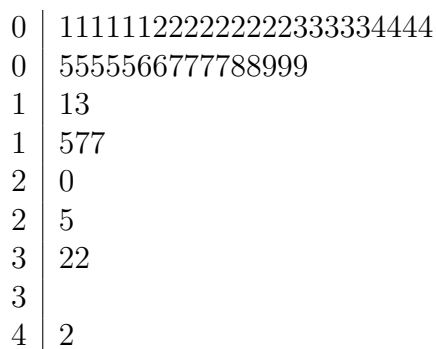
The center is between \$26,000 and \$28,000 and the spread is from \$8,000 to \$38,000. For public schools, the center is between \$10,000 and \$12,000 and the spread is from \$8,000 to \$16,000. For private schools, the center is between \$28,000 and \$30,000 and the spread is from \$20,000 to \$38,000.

Problem 1.9

Stemplot #1:



Stemplot #2:



Most students should prefer Stemplot #2 above or the stemplot presented in Figure 1.6 in the textbook. Stemplot #1 is too compact, which causes the extremely high states to not stand out as much as they should.

### Problem 1.11

The sum of the given percentages is 99%, so  $100\% - 99\% = 1\%$  are some other color.

The most significant difference is that 78% of the Japanese cars were gray or white, while only 39% of North American cars were.

### Problem 1.13

- a) The individuals are baseball players.
- b) The data set contains four other variables: team (categorical), position (categorical), age (quantitative), and salary (quantitative).
- c) Age is in years and salary is in dollars.

### Problem 1.17

- a) The shape is roughly symmetric with an outlier on the low end.
- b) The center is between 0% and 2%.
- c) The smallest return was between -24% and -22% (or between -16% and -14% ignoring the outlier) and the largest return was between 16% and 18%.
- d) Roughly 40% had returns less than 0.

### Problem 1.21

Histogram (a) is the length of words in Shakespeare's plays, as one would expect the distribution to be skewed to the right (many short words, fewer long words). Histogram (b) is the heights of students in the same class, as one would expect the distribution to be roughly symmetric. Histogram (c) is the handedness of students in the class, as one would expect right-handers to outnumber left-handers by a significant amount. Histogram (d) is the gender of the students in a large college course, as one would expect these counts to be somewhat close.

Problem 1.25

2		25
3		45
4		1166679
5		449
6		0

The distribution is roughly symmetric. In a typical year, Ruth hit about 46 home runs (the center of the distribution). His 1927 season is not an outlier.

Problem 1.26

McGwire		Ruth
99		0
		1
92		25
99322		45
92		1166679
82		449
5		0
0		7

McGwire has two very low values (the 1993 and 1994 seasons) and his distribution is skewed to the right with more spread than Ruth's.