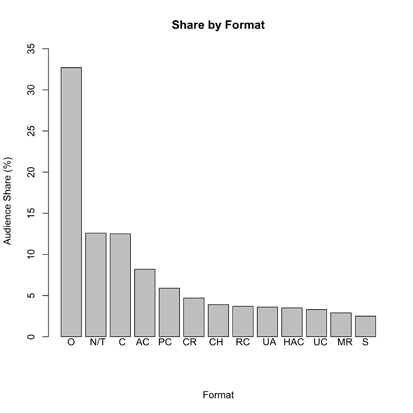
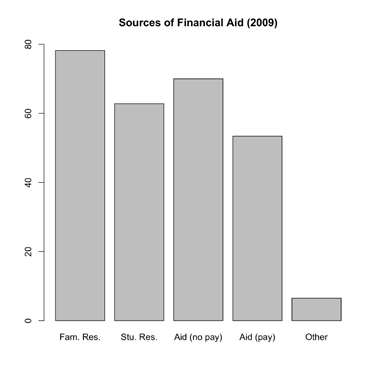
## Homework

1.3: (a) These shares sum to 67.3%. Hence, 100% – 67.3% = 32.7% of the radio audience listens to stations with other formats. (b)

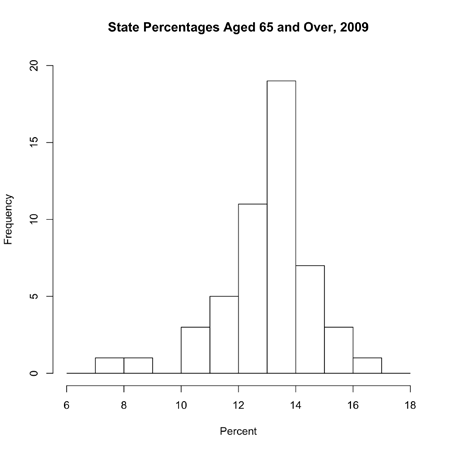


(c) A pie chart would be inappropriate based only on the data presented because the areas of the pie wedges would be relative to the total of the categories presented (67.3%). If you include a wedge for “other” that accounts for 32.7% of the total, a pie chart would be reasonable.

1.4: (a) Individuals fall into more than one of the categories. (b) A bar graph follows:



1.6: Two histograms with different bin widths. One should have bin widths of 1% as follows:



1.11: Here is a stemplot for health expenditure per capita (PPP). Data are rounded to units of hundreds. For example, Argentina’s “1322” becomes 13. Stems are thousands, and are split, as prescribed.

0 1 1 2 3

0 7 7 7 8 8 8 8 8

1 0 3

1 7

2 3

2 7 7 7 7 8

3 0 3 3 4 4

3 5 5 6 7 8 9

4 4

4 8

5

5

6

6

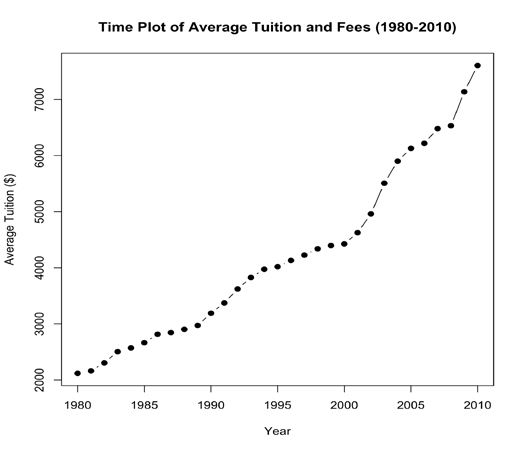
7 3

This distribution is somewhat right-skewed, with a single high outlier (United States). There are two clusters of countries. The center of this distribution is around 25 ($2500 spent per capita), ignoring the outlier. The distribution’s spread is from 1 ($100 spent per capita) to 73 ($7300 spent per capita).

**Rounding Values for a Stem Plot**

A stem plot is like a histogram but where you can still see your data values. Your stems represent the classes. The leaves are to be single digits to represent the number of data points in each class. Rounding is a way of creating the classes. If you have very large numbers like those in problem 1.11, rounding "1322" to the nearest hundred becomes 1300. You just plot the 13 and note that a stem of "1" with a leaf of "3" equals 1300. If you did not round, you would have a class/stem 132 with leaf 2. There is no other value that starts with "132" so that line will only have one value. Very few, if any, values start with the same 3 digits so you would have a class for every data point. You would not be able to see the shape of the distribution that way. If your values ranged from 0 to 100, you probably would not round because the ten's place would be the stem and the one's place the leaf. How you round will depend on what classes you want. Let me know if you have further questions.

1.12: (a) A time plot of Average tuition follows:



(b) Tuition has steadily climbed during the 30-year period, with sharpest absolute increases in the last 10 years. (c) It would be better to use percent increases, rather than dollar increases. A 10% increase in tuition in 1980 should correspond to a 10% increase in tuition in 2005, but the absolute dollar increases in these cases are very different.

1.32: (a) The distribution is slightly left-skewed. (b) The center is somewhere between 0% and 2.5%. (c) The smallest value is somewhere between –10% and –12.5%, and the largest value is between 12.5% and 15%. (d) There are about 130 negative returns, although your estimate could differ. This corresponds to about 42%.

1.33:

1. Are you male or female Histogram (c). There are two outcomes possible, and the difference in frequencies is likely to be smaller than the right-handed/left-handed difference in (2).

2. Are you right-handed or left-handed Histogram (b), since there are more right-handed people than left handed people, and the difference is likely larger than the sex difference in (1).

3. Heights Histogram (d). Height distribution is likely to be symmetric.

4. Time spent studying Histogram (a). The variable takes on more than one value, and time spent studying may well be a right-skewed distribution, with most students spending less time studying, and some students spending more time studying.