1.82 All three studies are experiments since the scientists actively control the treatment (tears or salt solution).

1.88 Yes, this study provides evidence that louder music causes people to drink more beer, because the explanatory variable (volume of music) was randomly determined by the researchers and an association was found.

1.92 (a) We randomly divide the participants into two groups of 25 each. Half will be given fluoxetine and half will get a placebo.

(b) The placebo pills will look exactly like the fluoxetine pills and will be taken the same way, but they will not have any active ingredients.

(c) The patients won't know who is getting which type of pill (the fluoxetine or the placebo) and the people treating the patients and administering the questionnaire won't know who is in which group.

1.94 (a) Randomly assign 25 people to carbo-load and 25 people to not carbo-load and then measure each person's athletic performance the following day.

(b) We would have each person carbo-load and not carbo-load, on different days (preferably different weeks). The order would be randomly determined, so some people would carbo-load first and other people would carbo-load second. In both cases athletic performance would be measured the following day and we would look at the difference in performance for each person between the two treatments.

(c) The matched pairs experiment is probably better because we are able to compare the different effects for the same person. It is more precise comparing one person's athletic performance under two different treatments, rather than different people's athletic performance under two different treatments.

2.74 Remember that a standard deviation is an approximate measure of the average distance of the data from the mean. Be sure to pay close attention to the scale on the horizontal axis for each histogram.

(a) V

(b) III

(c) IV

(d) I

(e) VI

(f) II

2.76 Remember that the five number summary divides the data (and hence the area in the histogram) into fourths.

(a) This shows a distribution pretty evenly spread out across the numbers 1 through 9, so this five number summary matches histogram W.

(b) This shows a distribution that is more closely bunched in the center, since 50% of the data is between 4 and 6. This five number summary matches histogram X.

(c) Since the top 50% of the data is between 7 and 9, this data is left skewed and matches histogram Y.

(d) Since both the minimum and the first quartile are 1, there is at least 25% of the data at 1, so this five number summary matches histogram Z.

2.80 The mean appears to be about 68. Since the data is relatively bell-shaped, we can estimate the standard deviation using the 95% rule. Since there are 100 dots in the dotplot, we want to find the boundaries with 2 or 3 dots more extreme on either side. This gives boundaries from 59 to 76, which is 8 or 9 units above and below the mean. We estimate the standard deviation to be about 4.5.

2.82 We see that the minimum value is 58 and the maximum is 77. We can count the dots to find the value at the 25^{th}-percentile, the 50^{th}-percentile, and the 75^{th}-percentile to find the quartiles and the median. We see that $Q_1 = 65$, the median is at 68, and $Q_3 = 70$. The five number summary is (58, 65, 68, 70, 77).

2.86 This data appears to be quite symmetric about the median of 36.3.

2.88 We have

$$Z\text{-score} = \frac{\text{Value} - \text{Mean}}{\text{Standard deviation}} = \frac{88 - 96}{10} = -0.8.$$

This value is 0.80 standard deviations below the mean, which is likely to be relatively near the center of the distribution.

2.94 The 95% rule says that 95% of the data should be within two standard deviations of the mean, so the interval is:

$$\begin{array}{rcl} \text{Mean} & \pm & 2 \cdot \text{StDev} \\ 1500 & \pm & 2 \cdot (300) \\ 1500 & \pm & 600 \\ 900 & \text{to} & 2100. \end{array}$$

We expect 95% of the data to be between 900 and 2100.

2.98 (a) We use technology to see that the mean is 56.10 and the standard deviation is 7.50.

(b) We see that 4 of the 10 values are larger than the mean. None of these four are in the early five years and all 4 of them are in the later five years. People seem to be getting better at eating hot dogs!

2.132 (a) Action movies appear to have the largest budgets, while horror and drama movies appear to have the smallest budgets.

(b) Action movies have by far the biggest spread in the budgets, with dramas appearing to have the smallest spread.

(c) Yes, there definitely appears to be an association between genre and budgets, with action movies having substantially larger budgets than the other three types.

2.144 (a) The explanatory variable is whether the traffic lights are on a fixed or flexible system. This variable is categorical. The response variable is the delay time, in seconds, which is quantitative.

(b) Using technology we find the mean and standard deviation for each sample:
Timed: $\overline{x}_T = 105$ seconds and $s_T = 14.1$ seconds
Flexible: $\overline{x}_F = 44$ seconds and $s_F = 3.4$ seconds
This shows that the mean delay time is much less, 61 seconds or more than a full minute, with the flexible light system. We also see that the variability is much smaller with the flexible system.

(c) For the differences we have $\overline{x}_D = 61$ seconds and $s_D = 15.2$ seconds.

(d) The boxplot is shown. We see that there are 3 large outliers. Since this is a boxplot of the differences, this means there were three simulation runs where the flexible system *really* improved the time.