

1. Wetlands offer a diversity of benefits. They provide habitat for wildlife, spawning grounds for U.S. commercial fish, and renewable timber resources. In the last 200 years, the United States has lost more than half its wetlands. *Environmental Almanac* gives the percentage of wetlands lost in each state in the last 200 years. For the lower 48 states, the percentage loss of wetlands per state is as follows:

46	37	36	42	81	20	73	59	35	50	87	52
24	27	38	56	39	74	56	31	27	91	46	9
54	52	30	33	28	35	35	23	90	72	85	42
59	50	49	48	38	60	46	87	50	89	49	67

- Make a stemplot, boxplot, and histogram.
- Describe the distribution.
- Find the numerical information: mean, standard deviation, five number summary, outliers, etc.
- Describe the relationship between the numerical information and the graphical information in context of this problem.

2. I have a data set consisting of 33 whole number observations. Its five number summary is: (16, 20, 22, 30, 46).

- What is the range of the data?
- How many observations are strictly less than 22?
- Is it possible that there are no observations equal to 22? Explain.
- What is the IQR?
- Show the test for outliers. Are there any outliers?

3. The average playing time of compact discs in a large collection is 35 minutes with a standard deviation of 5 minutes. Assume a normal distribution.

- Draw the distribution for this problem.
- What values are at one standard deviation above and below the mean?
- Assuming a normal distribution, at least what percentage of the times are between 25 and 45 minutes?
- What value is at the 84th percentile? The 50th percentile?
- What can be said about the percentage of times that are either less than 20 minutes or greater than 50 minutes?

4. At a ski area in Vermont, the daytime high temperature is normally distributed during January, with a mean of 22 degrees and a standard deviation of 10 degrees (U.S. Department of Commerce, *Environmental Data Services*). You are planning to ski there this January. What is the probability that you will encounter daytime highs of:

- 42 degrees or higher?
- 15 degrees or lower?
- between 29 and 40 degrees?

5. Eleanor scores 680 on the mathematics part of the SAT. The distribution of SAT scores is normal with a mean of 500 and standard deviation 100. Gerald takes the ACT mathematics test and scores 27. ACT scores are normally distributed with mean 18 and standard deviation 6. Assuming both tests measure the same kind of ability, who has the higher score? Explain.

6. According to *The New York Times* (April 2, 1993), the average monthly rate for basic television cable service has increased as follows:

Year: 1986	1987	1988	1989	1990	1991	1992
Cost: \$11	13.20	13.90	15.20	16.80	18.00	20.00

- Draw a scatterplot for the data.
- Find the equation of the least squares regression line.
- Slope = _____. Interpret the slope in context.
- y-intercept = _____. Interpret the y-intercept in context.
- $r =$ _____. Interpret the correlation coefficient in context.
- $r^2 =$ _____. Interpret the coefficient of determination in context.
- Predict the average monthly rate in 2007.
- Predict in what year the rate will reach \$50.
- Do you feel confident in the prediction you made in part g? Explain.
- Find the residuals for each year and make a residual plot. Use the plot to determine whether the LSRL is a good model to predict the cost of basic cable service.

For the next two questions (7 & 8), state whether the relationship between the two variables involves causation, common response, or confounding. Identify the lurking variables. Draw a diagram of the relationship and write a brief description.

7. People who do well tend to feel good about themselves. Perhaps helping people feel good about themselves will help them do better in school and life. Raising self-esteem became for a time a goal in many schools. California even created a state commission to advance the cause. Is there an association between high self-esteem and good performance other than “Self-esteem causes better work in school?”

8. A study of elementary school children, ages 6 to 11, finds a high positive correlation between shoe size and score on a test of reading comprehension. Does shoe size cause higher scores?

9. Ball players have been signing ever larger contracts. The highest salaries (in millions of dollars per season) for some notable players are given in the following table.

	Year	Salary (million \$)
Ryan	1980	1
† Foster	1982	2.04
†uckett	1990	3
†anseco	1990	4.7
Clemens	1991	5.3
†iffey, Jr.	1996	8.5
Belle	1997	11
Martinez	1998	12.5
†iazza	1999	12.5
†ughn	1999	13.3
†rown	1999	15
†elgado	2001	17
†odriguez	2001	25.2

- Draw a scatterplot. Is it linear?
- The LSRL after taking the logarithms of Salary is: $\text{Log}(\text{Salary}) = -113.47 + 0.057(\text{Year})$. Convert that to an exponential equation.
- Predict a superstar salary in 2006.
- Alfonso Soriano just signed a new contract for an average of \$16.5 million. Does that match your prediction from part (c)? Explain.

10. In a large city school system with 20 elementary schools, the school board is considering the adoption of a new policy that would require elementary students to pass a test in order to be promoted to the next grade. The PTA wants to find out whether parents agree with this plan. Listed below are some of the ideas proposed for gathering data. For each, indicate what kind of sampling strategy is involved and what (if any) biases might result.

- Put an ad in the newspaper asking people to log their opinions on the PTA web site.
- Randomly select one of the elementary schools and contact every parent by phone.
- Send a survey home with every student, and ask parents to fill out and return the next day.
- Randomly select 20 parents from each elementary school. Send them a survey, and follow up with a phone call if they do not return the survey within a week.

What’s the design? Read each (11 & 12) and identify whether it was an observational study or an experiment.

If observational, identify:

- the subjects studied, and how they were selected.
- the conclusion of the study

If an experiment, identify:

- the subjects studied
- the factor(s) in the experiment, and the number of levels for each.
- the number of treatments
- the design (completely randomized, blocked, or matched pairs)
- whether it was blind (or double-blind)
- The conclusion of the experiment

11. After menopause many women take supplemental estrogen. There is some concern that if these women also drink alcohol, their estrogen level will rise too high. Twelve volunteers who were receiving supplemental estrogen were randomly divided into two groups, as were other volunteers not on estrogen. In each case, one group drank an alcoholic beverage, the other a non-alcoholic beverage. An hour later everyone's estrogen level was checked. Only those on supplemental estrogen who drank alcohol showed a marked increase.

12. In 2001 a report in the Journal of the American Cancer Institute indicated that women who work nights have a 60% greater risk of developing breast cancer. Researchers based these findings on the work histories of 763 women with breast cancer and 741 women without the disease.

13. Hoping to learn how to control crop damage by a certain species of beetle, a researcher plans to test two different pesticides in small plots of corn. A few days after application of the chemicals, the researcher will check the number of beetle larvae found on each plant. The researcher wants to know if either pesticide works, and whether there is a significant difference in effectiveness between them. Design an appropriate experiment.

14. A pharmaceutical company wishes to test a new medication it thinks will reduce cholesterol. A group of 20 volunteers is formed and each has his or her cholesterol measured. After 6 months the volunteers' cholesterol is measured again and any change from the beginning of the study recorded. The researcher believes that regular exercise may influence the change in the cholesterol level. Create a randomized block design that takes account of subjects who exercise regularly.

15. A manufacturer of tires wants to conduct an experiment with a new type of tread design. This tread is supposed to last longer in hot climates than the existing tread design. The tire company wants to conduct this experiment in Phoenix, Arizona during the months of April through September. One hundred taxicabs will be used in this experiment, and only the two wheels will have the new tires. Design an experiment to determine if the new type of tire will last longer than the existing tire.