

**5.16 True / False: paired.** Determine if the following statements are true or false. If false, explain.

- (a) In a paired analysis we first take the difference of each pair of observations, and then we do inference on these differences.
- (b) Two data sets of different sizes cannot be analyzed as paired data.
- (c) Consider two data sets that form paired data. Each observation in one data set has a natural correspondence with exactly one observation from the other data set.
- (d) Consider two data sets that form paired data. In the analysis, each observation in one data set is subtracted from the average of the other data set's observations.

**5.17 Paired or not, Part I?** In each of the following scenarios, determine if the data are paired.

- (a) Compare pre- (beginning of semester) and post-test (end of semester) scores of students.
- (b) Assess gender-related salary gap by comparing salaries of randomly sampled men and women.
- (c) Compare artery thicknesses at the beginning of a study and after 2 years of taking Vitamin E for the same group of patients.
- (d) Assess effectiveness of a diet regimen by comparing the before and after weights of subjects.

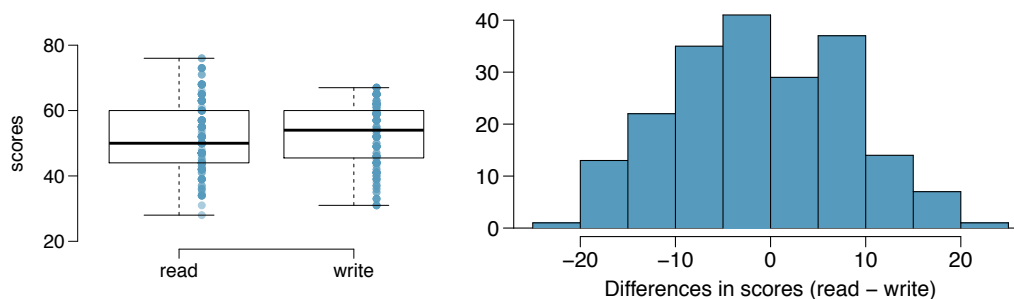
**5.18 Paired or not, Part II?** In each of the following scenarios, determine if the data are paired.

- (a) We would like to know if Intel's stock and Southwest Airlines' stock have similar rates of return. To find out, we take a random sample of 50 days, and record Intel's and Southwest's stock on those same days.
- (b) We randomly sample 50 items from Target stores and note the price for each. Then we visit Walmart and collect the price for each of those same 50 items.
- (c) A school board would like to determine whether there is a difference in average SAT scores for students at one high school versus another high school in the district. To check, they take a simple random sample of 100 students from each high school.

**5.19 Global warming, Part I.** Is there strong evidence of global warming? Let's consider a small scale example, comparing how temperatures have changed in the US from 1968 to 2008. The daily high temperature reading on January 1 was collected in 1968 and 2008 for 51 randomly selected locations in the continental US. Then the difference between the two readings (temperature in 2008 - temperature in 1968) was calculated for each of the 51 different locations. The average of these 51 values was 1.1 degrees with a standard deviation of 4.9 degrees. We are interested in determining whether these data provide strong evidence of temperature warming in the continental US.

- (a) Is there a relationship between the observations collected in 1968 and 2008? Or are the observations in the two groups independent? Explain.
- (b) Write hypotheses for this research in symbols and in words.
- (c) Check the conditions required to complete this test.
- (d) Calculate the test statistic and find the p-value.
- (e) What do you conclude? Interpret your conclusion in context.
- (f) What type of error might we have made? Explain in context what the error means.
- (g) Based on the results of this hypothesis test, would you expect a confidence interval for the average difference between the temperature measurements from 1968 and 2008 to include 0? Explain your reasoning.

**5.20 High School and Beyond, Part I.** The National Center of Education Statistics conducted a survey of high school seniors, collecting test data on reading, writing, and several other subjects. Here we examine a simple random sample of 200 students from this survey. Side-by-side box plots of reading and writing scores as well as a histogram of the differences in scores are shown below.



- Is there a clear difference in the average reading and writing scores?
- Are the reading and writing scores of each student independent of each other?
- Create hypotheses appropriate for the following research question: is there an evident difference in the average scores of students in the reading and writing exam?
- Check the conditions required to complete this test.
- The average observed difference in scores is  $\bar{x}_{read-write} = -0.545$ , and the standard deviation of the differences is 8.887 points. Do these data provide convincing evidence of a difference between the average scores on the two exams?
- What type of error might we have made? Explain what the error means in the context of the application.
- Based on the results of this hypothesis test, would you expect a confidence interval for the average difference between the reading and writing scores to include 0? Explain your reasoning.

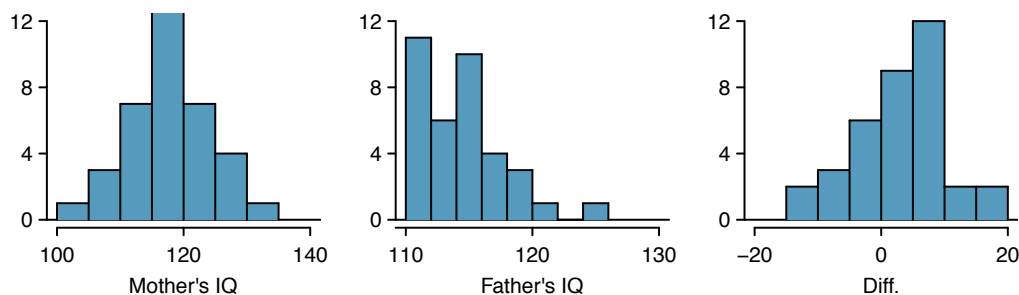
**5.21 Global warming, Part II.** We considered the differences between the temperature readings in January 1 of 1968 and 2008 at 51 locations in the continental US in Exercise 5.19. The mean and standard deviation of the reported differences are 1.1 degrees and 4.9 degrees.

- Calculate a 90% confidence interval for the average difference between the temperature measurements between 1968 and 2008.
- Interpret this interval in context.
- Does the confidence interval provide convincing evidence that the temperature was higher in 2008 than in 1968 in the continental US? Explain.

**5.22 High school and beyond, Part II.** We considered the differences between the reading and writing scores of a random sample of 200 students who took the High School and Beyond Survey in Exercise 5.20. The mean and standard deviation of the differences are  $\bar{x}_{read-write} = -0.545$  and 8.887 points.

- Calculate a 95% confidence interval for the average difference between the reading and writing scores of all students.
- Interpret this interval in context.
- Does the confidence interval provide convincing evidence that there is a real difference in the average scores? Explain.

**5.23 Gifted children.** Researchers collected a simple random sample of 36 children who had been identified as gifted in a large city. The following histograms show the distributions of the IQ scores of mothers and fathers of these children. Also provided are some sample statistics.<sup>37</sup>



	Mother	Father	Diff.
Mean	118.2	114.8	3.4
SD	6.5	3.5	7.5
n	36	36	36

- Are the IQs of mothers and the IQs of fathers in this data set related? Explain.
- Conduct a hypothesis test to evaluate if the scores are equal on average. Make sure to clearly state your hypotheses, check the relevant conditions, and state your conclusion in the context of the data.

**5.24 Sample size and pairing.** Determine if the following statement is true or false, and if false, explain your reasoning: If comparing means of two groups with equal sample sizes, always use a paired test.

<sup>37</sup>F.A. Graybill and H.K. Iyer. *Regression Analysis: Concepts and Applications*. Duxbury Press, 1994, pp. 511–516.