Exploring The Web

Chapter 27

27.54 Confidence in Congress. The General Social Survey (GSS) is a sociological survey used to collect data on demographic characteristics and attitudes of residents of the United States. The survey is conducted face-to-face with an in-person interview by the National Opinion Research Center at the University of Chicago, of a randomly selected sample of adults (18+). SDA (Survey Documentation and Analysis) is a set of programs that allows you to analyze survey data and includes the GSS survey as part of its archive. Go to the website `sda.berkeley.edu/` and click on ARCHIVE. Unless there is a more recent file, open the 1972–2014 cumulative data file (without the quick tables option).

(a) In the first row of the right-hand portion of the window that opens, click on MEANS. You are to do an ANOVA that examines whether there is a difference in the mean age of the respondents for the three levels of confidence in the U.S. Congress. To do this, type in the dependent variable as "Age" and the row (treatment) variable as "Conlegis." For the selection filter, type in "Year(2014)," or the most recent year available. For weight, change it to "noweight." Finally, in Output options, the *only* boxes that should be checked are "Std dev," "N," and "ANOVA stats." Make sure checks are removed from the other boxes. Now click on RUN THE TABLE.

(b) How many respondents are included in the analysis? What are the three means and standard deviations? In the ANOVA table, explain how the degrees of freedom were obtained. What are the *F*- and *P*-values? Write a brief report explaining the relationship between the average respondent's age and their confidence in the Congress.

27.55 Confidence in Congress, continued. This exercise is a continuation of the previous web exercise. You are going to download the data set and reproduce the analysis, as well as provide some additional plots and multiple comparisons. First open the 1972–2014 cumulative data file following the instructions in the previous exercise.

(a) In the DOWNLOAD tab on the top of the page, click on DOWNLOAD CUSTOM SUBSET. For File Options, if you highlight the CSV bubble, the observations will be downloaded with commas between the variable values and variable names in the first row. This type of file can be imported into Excel or your statistical software package. For Select Cases, type in "year(2014)" or the year used in the previous exercise. For Select Variables, in the box for Specify individual variable names, enter "age" and "conlegis." Finally, for Create Files, first check to make sure you have entered everything correctly and then click on the Create Files bubble at the bottom of the window and then on Datafiles. You can now either open or save the data file to your computer.

(b) Import the data into your statistical software package. You first need to "clean" the data a little because there are observations for which either the "age" or "conlegis" variable is missing. For the "conlegis" variable, any value other than a 1, 2, or 3 is a missing value code. Delete these observations. For the "age" variable, the missing value codes are 0, 98, and 99. Eliminate any observations with these values for "age." You should now have the same number of observations as in the previous exercise.

(c) Draw comparative boxplots of the age distribution for the three values of "conlegis." Describe the shapes of the three distributions. What information can you obtain from the boxplots that was not included in the output for the previous exercise?

(d) Reproduce the one-way ANOVA table using your software. Your results should agree with the previous exercise.

(e) Find the Tukey simultaneous 95% confidence intervals for all pairwise differences among the three population means. Which pairs of means differ significantly at the overall 5% significance level?