

Chapter 2

2.54 Home run leaders. The three top players on the career home run list are Barry Bonds, Hank Aaron, and Babe Ruth. You can find their home run statistics by going to the website www.baseball-reference.com and then clicking on the PLAYERS tab at the top of the page. Construct three side-by-side boxplots comparing the yearly home run production of Barry Bonds, Hank Aaron, and Babe Ruth. Describe any differences that you observe. It is worth noting that in his first four seasons, Babe Ruth was primarily a pitcher. If these four seasons are ignored, how does Babe Ruth compare with Barry Bonds and Hank Aaron?

2.55 Global health statistics The global health observatory data repository maintains an extensive database containing a variety of indicators of global health at apps.who.int/gho/data/node.home. We are going to provide the details for using this database to download life expectancy data by country as an example, although you are free to choose any category and variable of interest to you. Here are the details:

After going to the data repository website, select BY CATEGORY on the left side of the page, and then click on WORLD HEALTH STATISTICS on the right side. In the window that opens, click on MORTALITY AND GLOBLE HEALTH ESTIMATES, and then click on LIFE EXPECTANCY. A data table will open that contains several variables related to life expectancy. You can reduce the number of variables reported by clicking on FILTER TABLE at the top left-hand corner of the table. For example, you can limit the table to life expectancy at birth for females in 2013. Be sure to click on APPLY to make the changes. Once you have the data table of interest, download the data as a CSV file using the download option on the top-right side of the table. The CSV file can be opened in Excel and imported into your statistical software. If you were interested in comparing life expectancies of males and females throughout the world in 2013, then you would have included both sexes when filtering the table. Selection of other indicators of global health would proceed in a similar manner.

Give a full description of the distribution of the global health indicator selected. Your description of the distribution should include appropriate graphical and numerical summaries and a brief report describing the main features of the distribution. If you included both sexes in the data table, then your report should include a comparison of the distributions of life expectancies of males and females.