Assignment 2

58. The file P03_58.xlsx lists salaries of MLB players in the years 2007 to 2009. Each row corresponds to a particular player. As indicated by blank salaries, some players played in one of these years, some played in two of these years, and the rest played in all three years.

a. Create a new Yes/No variable, All 3 Years, that indicates which players played all three years.

b. Create two pivot tables and corresponding pivot charts. The first should show the count of players by position who played all three years. The second should show the average salary each year, by position, for all players who played all three years. (For each of these, put the All 3 Years variable in the Filters area.) Explain briefly what these two pivot tables indicate.

c. Consider the data set consisting of only the players who played all three years. Using this data set, create a table of correlations of the three salary variables. What do these correlations indicate about player salaries?

60. The file P02_12.xlsx includes data on the 50 top graduate programs in the United States, according to a U.S. News & World Report survey.

a. Create a table of correlations between all of the numerical variables. Discuss which variables are highly correlated with which others.

b. The Overall score is the score schools agonize about. Create a scatterplot and corresponding correlation of each of the other variables versus Overall, with Overall always on the Y axis. What do you learn from these scatterplots?