Predictors for Success in College

When a college accepts a new student, it would like to have some positive indication that the student will be successful in his or her studies. College admissions deans consider SAT scores, standard achievement tests, rank in class, difficulty of high school courses, high school grades, and extracurricular activities. In a study of characteristics that make good predictors of success in college, it was found that class rank and scores on standard achievement tests are better predictors than SAT scores. A multiple regression equation with college grade-point average predicted by class rank and achievement test score was not improved by including another variable for SAT score. This particular study suggests that SAT scores should not be included among the admissions criteria, but supporters argue that SAT scores are useful for comparing students from different geographic locations and high school backgrounds.

Table 10-5 Select Key Results from Data Set 2 in Appendix B

Predictor Variables	Adjusted R^2	*P*-Value	
Age	0.1772	0.004	← **Not best:** Adjusted R^2 is far less than 0.7014 for Foot Length
Foot Length	**0.7014**	**0.000**	← **Best:** High adjusted R^2 and lowest *P*-value.
Shoe Print Length	0.6520	0.000	← **Not best:** Adjusted R^2 is less than 0.7014 for Foot Length
Foot Length/Shoe Print Length	0.7484	0.000	← **Not best:** The adjusted R^2 value is not very much higher than 0.7014 for the single variable of Foot Length.
Age/Foot Length/ Shoe Print Length/ Shoe Size	0.7585	0.000	← **Not best:** There are other cases using fewer variables with adjusted R^2 that is not too much smaller.

the objective of using evidence to estimate the height of a suspect, we use *critical thinking* as follows.

1. Delete the variable of age, because criminals rarely leave evidence identifying their ages.
2. Delete the variable of shoe size, because it is really a rounded form of foot length.
3. For the remaining variables of foot length and shoe print length, use only foot length because its adjusted R^2 value of 0.7014 is greater than 0.6520 for shoe print length, and it is not very much less than the adjusted R^2 value of 0.7484 for both foot length and shoe print length. In this case, it is better to use one predictor variable instead of two.
4. Although it appears that the use of the single variable of foot length is best, we also note that criminals usually wear shoes, so shoe print lengths are more likely to be found than foot lengths.

Interpretation

Although blind use of regression methods suggest that when estimating the height of a subject, it is best to use all four predictor variables of age, foot length, shoe print length, and shoe size, other practical considerations suggest that it is best to use the single predictor variable of foot length, so the best regression equation appears to be this: Height = 64.1 + 4.29 (Foot Length). However, given that criminals usually wear shoes, it is best to use the single predictor variable of shoe print length, so the best regression equation appears to be this: Height = 80.9 + 3.22 (Shoe Print Length). The *P*-value of 0.000 suggests that the regression equation yields a good model for estimating height.

Because the results of this example are based on sample data from only 40 subjects, estimates of heights will not be very accurate. As is usually the case, better results could be obtained by using larger samples.

Tests of Regression Coefficients The preceding guidelines for finding the best multiple regression equation are based on the adjusted R^2 and the *P*-value, but we could also conduct individual hypothesis tests based on values of the regression coefficients. Consider the regression coefficient of β_1. A test of the null hypothesis