

Stat 150 – Day 28
More Regression Examples

Example 1: Speaking and Intelligence

Recall the data that you analyzed earlier in the Minitab worksheet `gesell.mtw`. The goal of this study is to predict a child's aptitude score on the Gesell test, based on the age (in months) at which the child first speaks.

a) Produce a scatterplot with the regression line superimposed (`Stat > Regression > Fitted Line Plot`). Record the equation of the line, using good statistical notation.

b) Report and interpret the value of R^2 .

c) Report and interpret the value of the slope coefficient. (Remember to include a probabilistic aspect to your interpretation.)

d) Now remove from the analysis the child who took the longest to speak. Report the new regression equation, slope coefficient, and R^2 value. Comment on how much the removal of just one child affected the regression line.

e) Now remove the child who took the second longest to speak. Answer question d) again.

f) What term do we use to describe these two observations, whose removal had a substantial impact on the regression line?

Example 2: House Prices

Recall again the data in the Minitab worksheet `HousePrices150.mtw`, which lists prices and sizes (in square feet) for a random sample of houses that sold in the year 2006 in Arroyo Grande, California

- a) What is the natural choice for a response variable with these data?

- b) What are the potential explanatory variables with these data?

- c) For each of the three explanatory variables, fit the regression line for predicting house price based on that predictor. Record the value of R^2 for each. Which explanatory variable is most strongly correlated with price? Which is least strongly correlated?

- d) Record the regression equation and value of R^2 based on the best single predictor of price.

- e) Now add the second best predictor to the regression model for predicting price. Report the multiple regression equation and value of R^2 . Does it appear that adding this second predictor provides a large improvement in the ability to predict price? Explain.

- f) Now add the third predictor to the regression model for predicting price. Report the multiple regression equation based on all three predictors, and also record value of R^2 . Does it appear that adding this third predictor provides a large improvement in the ability to predict price? Explain.