Chapter 19 - Linear Patterns

37. Diamond Rings This data table contains the listed prices and weights of the diamonds in 48 rings offered for sale in *The Singapore Times.* The prices are in Singapore dollars, with the weights in carats.

- (a) Scatterplot the listed prices of the rings on the weights of the rings. Does the trend in the average price seem linear?
- (b) Estimate the linear equation using least squares. Interpret the fitted intercept and slope. Be sure to include their units. Note if either estimate represents a large extrapolation and is consequently not reliable.
- (c) Interpret the summary values r^2 and s_e associated with the fitted equation. Attach units to these summary statistics as appropriate.
- (d) What is the estimated difference in price (on average) between diamond rings with diamonds that weigh 0.25 and 0.35 carat?

41. Seattle Homes This data table contains the listed prices (in thousands of dollars) and the number of square feet for 28 homes listed by a realtor in the Seattle area.

- (a) Create a scatterplot for the price of the home on the number of square feet. Does the trend in the average price seem linear?
- (b) Estimate the linear equation using least squares. Interpret the fitted intercept and slope. Be sure to include their units. Note if either estimate represents a large extrapolation and is consequently not reliable.
- (c) Interpret the summary values r^2 and s_e associated with the fitted equation. Attach units to these summary statistics as appropriate.
- (d) If a homeowner adds an extra room with 500 square feet to her home, can we use this model to estimate the increase in the value of the home?
- (e) A home with 2,690 square feet lists for $625,000. What is the residual for this case? Is it a good deal?
- (f) Do the residuals from this regression show patterns? Does it make sense to interpret s_e as the standard deviation of the errors of the fit? Use the plot of the residuals on the predictor to help decide.

47. Promotion These data describe spending by a major pharmaceutical company for promoting a cholesterol-lowering drug. The data cover 39 consecutive weeks and isolate the area around Boston. The variables in this collection are shares. Marketing research often uses the notion of voice to describe the level of promotion for a product. In place of the absolute spending for advertising, *voice* is the share of a type of advertising devoted to a specific product. Voice puts this spending in context; $10 million might seem like a lot for advertising unless everyone else is spending $200 million.

The column *Market Share* is sales of this product divided by total sales for such drugs in the Boston area. The column *Detail Voice* is the ratio of detailing for this drug to the amount of detailing for all cholesterol-lowering drugs in Boston. Detailing counts the number of promotional visits made by representatives of a pharmaceutical company to doctors' offices.

- (a) Do timeplots of Market Share and Detail Voice suggest an association between these series? Does either series show simple variation?
- (b) Create a scatterplot for Market Share on Detail Voice. Are the variables associated? Does a line summarize any association?
- (c) Estimate the least squares linear equation for the regression of Market Share on Detail Voice. Interpret the intercept and slope. Be sure to include the units for each. Note if either estimate represents a large extrapolation and is consequently not reliable.
- (d) Interpret r^2 and s_e associated with the fitted equation. Attach units to these summary statistics as appropriate.
- (e) According to this equation, how does the average share for this product change if the detail voice rises from 0.04 to 0.14 (4% to 14%)?
- (f) Plot the residuals from the regression fit in part (c) on the sizes of the files. Does this plot suggest that the residuals possess simple variation?