suppose the entering freshman at a certain college have mean combined SAT scores of 1222, with a standard deviation of 123. in the first semester these students attained a mean GPA of 2.66 with a standard deviation of 0.56. a scatterplot showed the association to be reasonably linear, and the correlation between SAT and GPA was 0.47

a) write the equation of the regression line

b) explain what the y-intercept of the regression line indicates

c) interpret the slope of the regression line

d) predict the GPA of a freshman who scored a combined 1400

e) based upon these statistics how effective do you think SAT scored would be in predicting academic success during the first semester of the freshman year at this college? explain

f) as a student would you rather have a positive or negative residual in this context? Explain

Solution

a) Let's use SAT as x and GPA as y.

$$covariance = rs_xs_y = 0.47 \cdot 123 \cdot 0.56 = 20.81.$$

The equation is

$$y = B_o + B_1x$$

$$B_1 = \frac{cov}{var(x)} = \frac{cov}{s_x^2} = \frac{20.81}{123 \cdot 123} = 0.0021.$$

$$B_o = mean_y - mean_x = 2.66 - 1833 \cdot 0.0021 = -1.189$$

So the equation is

$$y = -1.189 + 0.0021x$$

b) GPA can't be negative. So it is meaningless value. Just consider your regression line only in the area with y>0.

c) For each additional SAT score, the model predicts an increase of 0.0021 in GPA score.

d) predict the GPA of a freshman who scored a combined 1400

$$y(1400) = -1.189 + 0.0021 \cdot 1400 = 1.751.$$

e) $r^2 = 0.47^2 = 0.2209.$

SAT scores is not an effective predictor of college GPA, There are only 22.09% of variability in GPA can be explained by the model.

f) Positive, it means GPA is higher than model predicts.