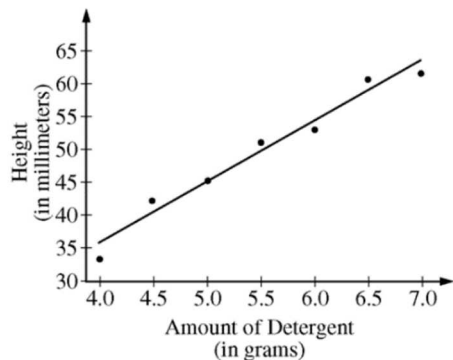


### AP Stats Ch3 and 4 Review and Free Response Questions

#### Q1: 2006 FRQ

A manufacturer of dish detergent believes the height of soapsuds in the dishpan depends on the amount of detergent used. A study of the suds' heights for a new dish detergent was conducted. Seven pans of water were prepared. All pans were of the same size and type and contained the same amount of water. The temperature of the water was the same for each pan. An amount of dish detergent was assigned at random to each pan, and that amount of detergent was added to the pan. Then the water in the dishpan was agitated for a set amount of time, and the height of the resulting suds was measured.

A plot of the data and the computer output from fitting a least squares regression line to the data are shown below.



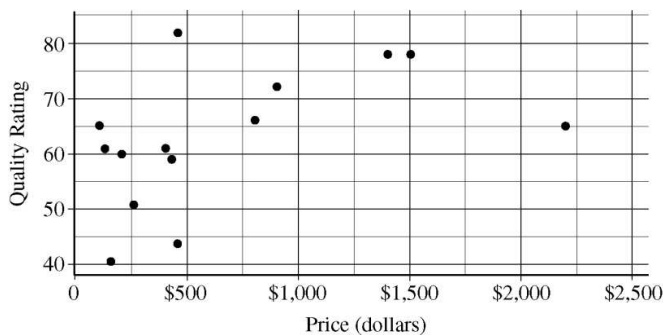
Predictor	Coef	SE Coef	T	P
Constant	-2.679	4.222	-0.63	0.554
Amount	9.5000	0.7553	12.58	0.000

$S = 1.99821$     $R\text{-}Sq = 96.9\%$     $R\text{-}Sq(\text{adj}) = 96.3\%$

- Write the equation of the fitted regression line. Define any variables used in this equation.
- Note that  $s = 1.99821$  in the computer output. Interpret this value in the context of this study.
- Identify and interpret the standard error of the slope.

#### Q2 :2012 FRQ

The scatterplot below displays the price in dollars and quality rating for 14 different sewing machines.



- Describe the nature of the association between price and quality rating for the sewing machines.
- One of the 14 sewing machines substantially affects the appropriateness of using a linear regression model to predict quality rating based on price. Report the approximate price and quality rating of that machine and explain your choice.
- Chris is interested in buying one of the 14 sewing machines. He will consider buying only those machines for which there is no other machine that has both higher quality and lower price. On the scatterplot reproduced below, circle all data points corresponding to machines that Chris will consider buying.



Q3: 2011 FRQ

Windmills generate electricity by transferring energy from wind to a turbine. A study was conducted to examine the relationship between wind velocity in miles per hour (mph) and electricity production in amperes for one particular windmill. For the windmill, measurements were taken on twenty-five randomly selected days, and the computer output for the regression analysis for predicting electricity production based on wind velocity is given below. The regression model assumptions were checked and determined to be reasonable over the interval of wind speeds represented in the data, which were from 10 miles per hour to 40 miles per hour.

Predictor	Coef	SE Coef	T	P
Constant	0.137	0.126	1.09	0.289
Wind velocity	0.240	0.019	12.63	0.000

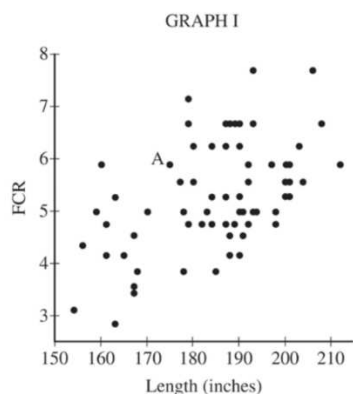
$S = 0.237$	$R\text{-Sq} = 0.873$	$R\text{-Sq (adj)} = 0.868$
-------------	-----------------------	-----------------------------

- Use the computer output above to determine the equation of the least squares regression line. Identify all variables used in the equation.
- How much more electricity would the windmill be expected to produce on a day when the wind velocity is 25 mph than on a day when the wind velocity is 15 mph? Show how you arrived at your answer.
- What proportion of the variation in electricity production is explained by its linear relationship with wind velocity?

Q: 2014FRQ

Jamal is researching the characteristics of a car that might be useful in predicting the fuel consumption rate (FCR); that is, the number of gallons of gasoline that the car requires to travel 100 miles under conditions of typical city driving. The length of a car is one explanatory variable that can be used to predict FCR. Graph I is a scatterplot showing the lengths of 66 cars plotted with the corresponding FCR. One point on the graph is labeled A.

Jamal examined the scatterplot and determined that a linear model would be a reasonable way to express the relationship between FCR and length. A computer output from a linear regression is shown below.



Linear Fit  

$$\text{FCR} = -1.595789 + 0.0372614 * \text{Length}$$

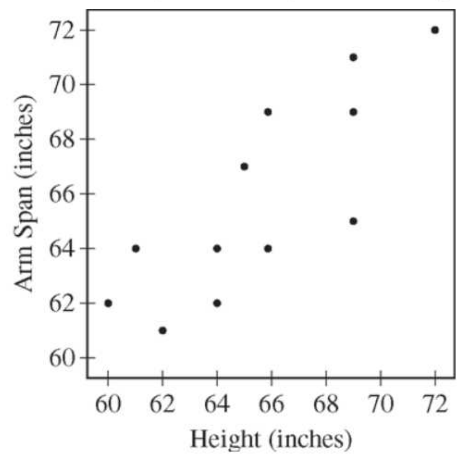
Summary of Fit  
 RSquare 0.250401  
 Root Mean Square Error 0.902382  
 Observations 66

- The point on the graph labeled A represents one car of length 175 inches and an FCR of 5.88. Calculate and interpret the residual for the car relative to the least squares regression line.



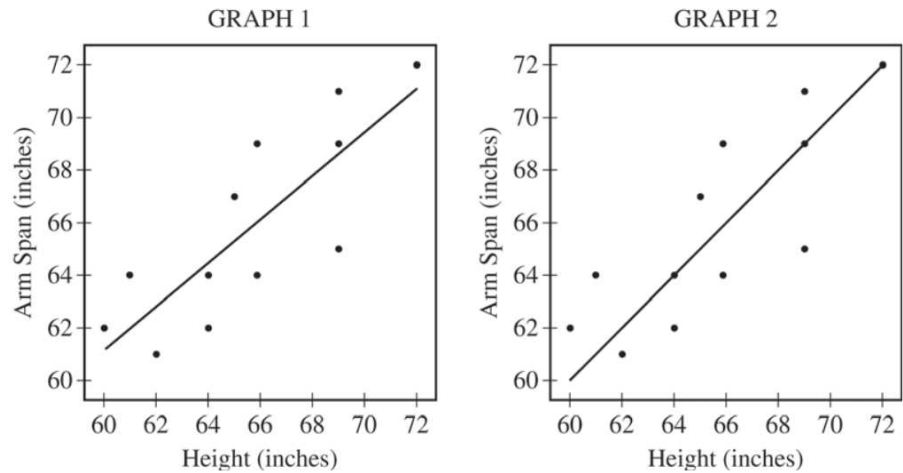
Q: 2015 FRQ

A student measured the heights and the arm spans, rounded to the nearest inch, of each person in a random sample of 12 seniors at a high school. A scatterplot of arm span versus height for the 12 seniors is shown.



(a) Based on the scatterplot, describe the relationship between arm span and height for the sample of 12 seniors.

Let  $x$  represent height, in inches, and let  $y$  represent arm span, in inches. Two scatterplots of the same data are shown below. Graph 1 shows the data with the least squares regression line  $\hat{y} = 11.74 + 0.8247x$ , and graph 2 shows the data with the line  $y = x$ .



(b) The criteria described in the table below can be used to classify people into one of three body shape categories: square, tall rectangle, or short rectangle.

Square	Tall Rectangle	Short Rectangle
Arm span is equal to height.	Arm span is less than height.	Arm span is greater than height.

(i) For which graph, 1 or 2, is the line helpful in classifying a student's body shape as square, tall rectangle, or short rectangle? Explain.

(ii) Complete the table of classifications for the 12 seniors.

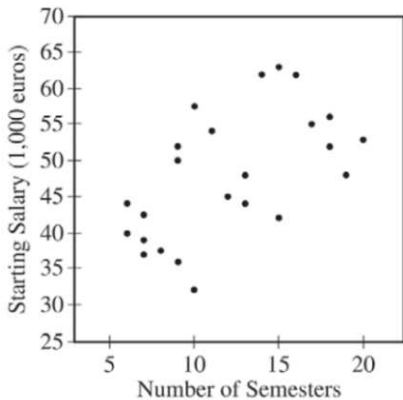
Classification	Square	Tall Rectangle	Short Rectangle
Frequency			

(c) Using the best model for prediction, calculate the predicted arm span for a senior with height 61 inches.



Q: 2016 FRQ

A newspaper in Germany reported that the more semesters needed to complete an academic program at the university, the greater the starting salary in the first year of a job. The report was based on a study that used a random sample of 24 people who had recently completed an academic program. Information was collected on the number of semesters each person in the sample needed to complete the program and the starting salary, in thousands of euros, for the first year of a job. The data are shown in the scatterplot below.



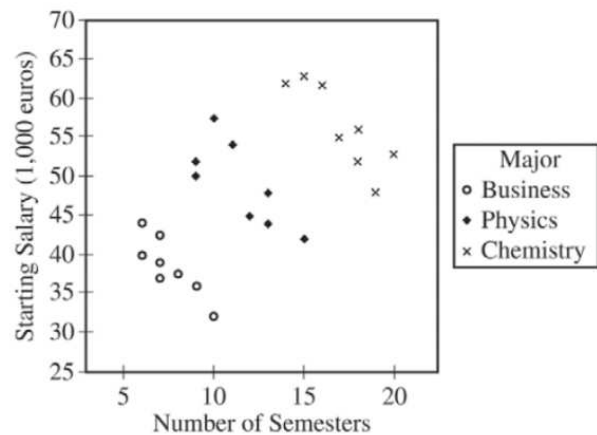
The table below shows computer output from a linear regression analysis on the data.

Predictor	Coef	SE Coef	T	P
Constant	34.018	4.455	7.64	0.000
Semesters	1.1594	0.3482	3.33	0.003

S = 7.37702      R-Sq = 33.5%      R-Sq(adj) = 30.5%

- (a) Does the scatterplot support the newspaper report about number of semesters and starting salary? Justify your answer.
- (b) Identify the slope of the least-squares regression line, and interpret the slope in context.

An independent researcher received the data from the newspaper and conducted a new analysis by separating the data into three groups based on the major of each person. A revised scatterplot identifying the major of each person is shown below.



- (c) Based on the people in the sample, describe the association between starting salary and number of semesters for the business majors.
- (d) Based on the people in the sample, compare the median starting salaries for the three majors.
- (e) Based on the analysis conducted by the independent researcher, how could the newspaper report be modified to give a better description of the relationship between the number of semesters and the starting salary for the people in the sample?