AP STAT Worksheet: LSRL, Correlation, & Predictions Name _____

Ch 3.2: Least Squares Regression Lines

1. Karl keeps his savings under his mattress. He began with $500 from his mother and add $100 each year. His total savings *y* after *x* years are given by the equation *pred*(savings) = 500 + 100(years)

a) After 20 years, how much will Karl have under his mattress?

b) If Karl had added $200 instead of $100 each year to his initial $500, what is the equation that describes his savings after *x* years.

2. The following computer output is the result of plotting hours studied for an exam against the score obtained on the exam.

The regression equation is

Predictor	Coef	St Dev	*t* ratio	*P*
Constant	59.026	2.863	20.62	.000
Hours	6.767	1.092	6.20	.000

s = 6.135 R-sq = 74.7% R-sq(adj) = 72.8%

a) Write an equation (in context) for exam score after hours studying.

b) Interpret the slope in context of the problem.

c) Interpret the y-intercept in context of the problem.

3. The following calculations are from a data set plotting school grade point average (GPA), **y**, against IQ test score , **x**, for 78 seventh-grade students.

$\bar{x}$ = 108.9 s_x = 13.17 $\bar{y}$ = 7.447 s_y = 2.10 r = 0.6337

a) Find the equation of the least squares line for predicting GPA from IQ.

b) What percent of the observed variation in these students' GPAs can be explained by the linear relationship between GPA and IQ?

c) One student has an IQ of 103 but a very low GPA of 0.53. What is the predicted GPA for a student with IQ = 103? What is the residual for this particular student and what does the residual say about this student?

4. Keeping water supplies clean requires regular measurement of levels of pollutants. The measurements are indirect - a typical analysis involves forming a dye by a chemical reaction with the dissolved pollutant, then passing light through the solution and measuring its "absorbence." To calibrate such measurements, the laboratory measures known standard solutions and uses regression to relate absorbence to pollute concentration. This is usually done every day. Here is one series of data on the absorbence for different levels of nitrates. Nitrates are measured in milligrams per liter of water.

Nitrates	50	50	100	200	400	800	1200	1600	2000	2000
Absorbence	7.0	7.5	12.8	24.0	47.0	93.0	138.0	183.0	230.0	226.0

a) Chemical theory says that these data should lie on a straight line. If the correlation is not at least 0.997, something went wrong and the calibration procedure is repeated. Find the correlation and decide if the calibration must be done again.

b) What is the equation of the LSRL for predicting absorbence from concentration? If the lab analyzed a specimen with 500 mg of nitrates per liter, what do you expect the absorbence to be?

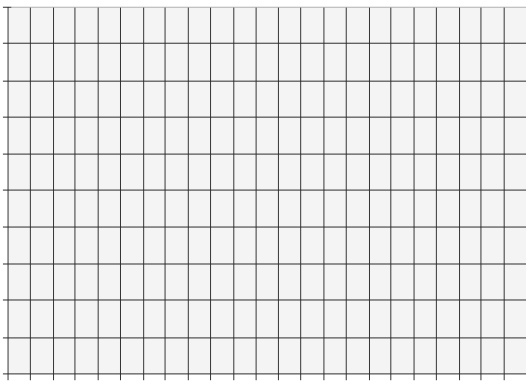
c) What type of prediction is this?

5. The number of people living on American farms has steadily declined during this century. Here are the data on the farm population (millions of persons) from 1935 to 1980.

Year	1935	1940	1945	1950	1955	1960	1965	1970	1975	1980
Population	32.1	30.5	24.4	23.0	19.1	15.6	12.4	9.7	8.9	7.2

a) Make a scatterplot. State and graph the least-squares regression line (in context) of farming on year.

LSRL: _____



b) According to the regression line, how much did the farm population decline, on the average, each year during this period?

c) What is the *proportional reduction in error* (also called the coefficient of determination)? Interpret this value in context of the problem.

d) Predict the number of people living on farms in 1990.
This type of prediction is called _____.
Is this result reasonable? Explain.

e) What conclusions can you make concerning making predictions obtained from values of x which stray far from your data?