

Residual Plot Activity

Name: _____

AP Statistics

Period: _____ Date: _____

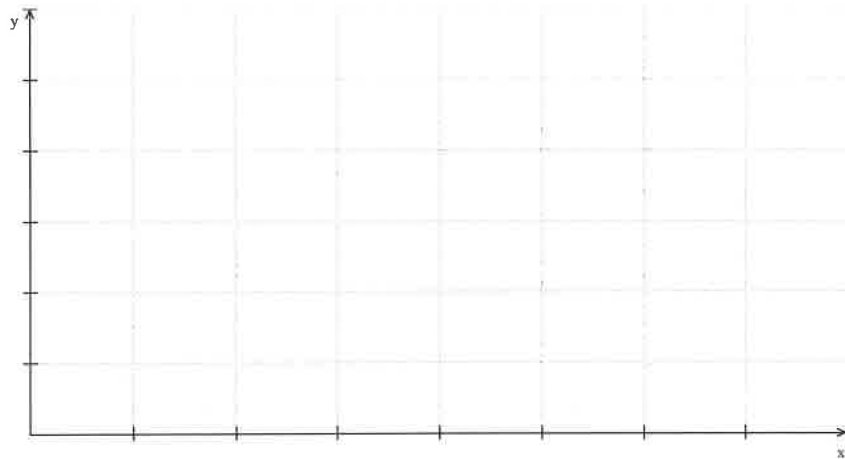
A.M.D.G.

Formulas:

$$\text{residual} = y - \hat{y}$$
 (the residual value is simply the actual y - value minus the predicted y - value)

Plot the scatterplot for death rate by heart attack for each year of the data below on the graph (label and scale your axes appropriately):

Year:	1950	1960	1970	1975	1980
Death rate:	307.6	286.2	253.6	217.8	202.0



Does the data appear linear? Does the trend appear to be positive or negative?

The regression line for this data is $\widehat{\text{death rate}} = 7386.87 - 3.627(\text{year})$ with $r^2 = 0.963$

According to the coefficient of determination, what percentage of the death rate can be explained by the year?

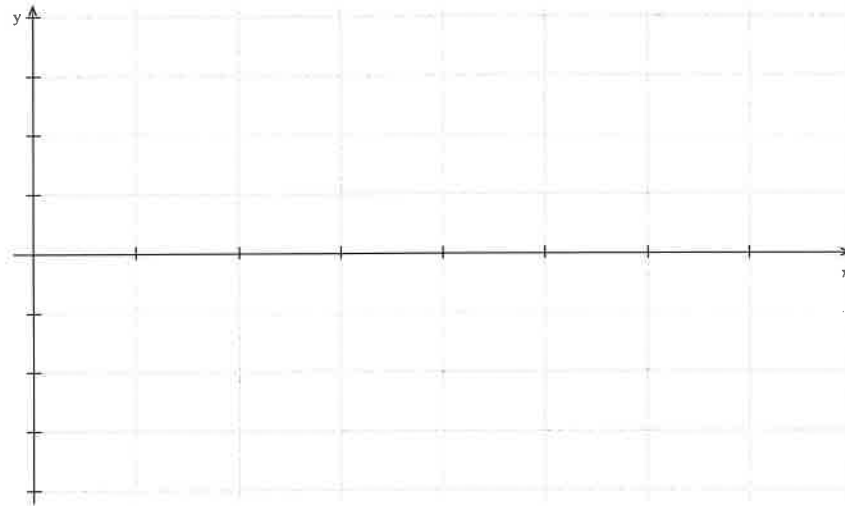
What is the value of r for this regression? Interpret the value of r for this situation.

Using the predictor line, calculate the residual for each y value.

$$\widehat{\text{death rate}} = 7386.87 - 3.627(\text{year})$$

Year:	1950	1960	1970	1975	1980
Death rate:	307.6	286.2	253.6	217.8	202.0
Residual:					

Graph the value of each residual against the year on the axis system below. Make sure you label the axes.



Does there appear to be a pattern in the residuals? What does this indicate about the trend being linear or not?

Use the regression line to predict the death by heart rate in 2023. Do you think that this prediction is appropriate? Why or why not?

A chemistry student collected the following data on a decomposition of crystal violet experiment:

Time (seconds)	10	20	30	40	50	60	70	80	90
Absorbance	0.90	0.78	0.70	0.62	0.57	0.51	0.45	0.41	0.37

After performing a linear regression, he gets the regression equation and r value below:

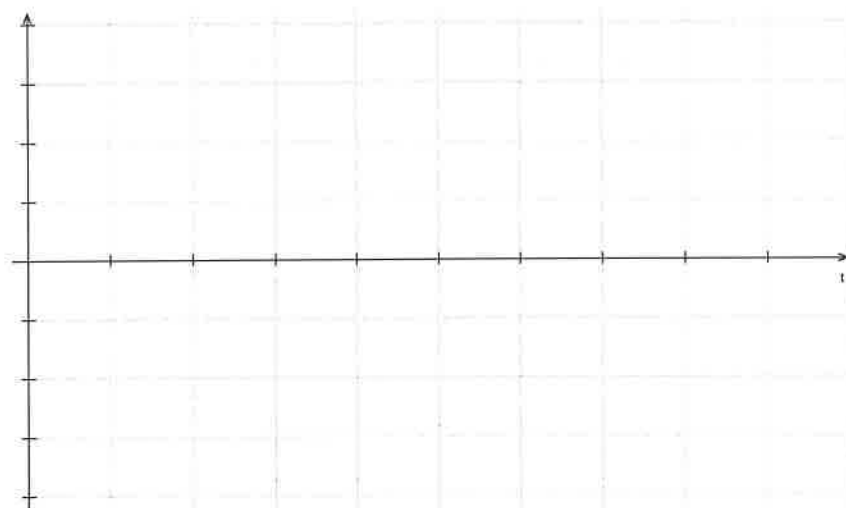
$$\widehat{\text{Absorbance}} = 0.91 - 0.0064(\text{time}) \quad r = -0.987$$

From this, he asserts that the trend **must be** linear, and that his predictor line should work quite well for extrapolating values. Is his assertion correct? Explain why or why not.

Use the regression line to calculate the residuals for each value, and create a residual plot below:

$$\widehat{\text{Absorbance}} = 0.91 - 0.0064(\text{time})$$

Time (seconds)	10	20	30	40	50	60	70	80	90
Absorbance	0.90	0.78	0.70	0.62	0.57	0.51	0.45	0.41	0.37



Would this residual plot be evidence that the trend is actually linear or not? Explain briefly.