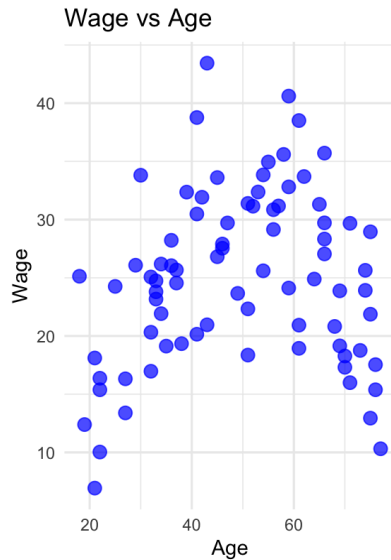# Assignment More Regression

Effect age has on wages:

a) **Plot Wage against Age and evaluate whether a linear or quadratic model would better capture the relationship.**

Wage vs Age

By examining the scatter plot above we can conclude that the relationship between wage and age is nonlinear . Given this, a quadratic model may be more appropriate, as it can capture the curvilinear relationship between age and wage."

- To better visualize this, I evaluated both a linear regression model and a quadratic model. By plotting both, we can observe that the quadratic model's curve fits the relationship between age and wage more accurately.

Wage vs Age with Quadratic Model

Wage vs Age with Linear Model

b) Estimate a multiple regression model of Wage using Age and Education as independent (X) variables; assume a standard linear relationship between Wage and Age

Coefficients:

	Estimate Std.	Error	t value	Pr(>|t|)
(Intercept)	2.63808	2.36649	1.115	0.268
Age	0.04717	0.03062	1.541	0.127
Educ	1.44101	0.13123	10.981	<2e-16

This model shows a positive correlation between wages and both age and education. The connection between education and wages is especially strong, with a p-value of "<2e-16," which is far below the 0.001 significance level. Age doesn't seem to have as big an impact and we conclude that from its p-value "0.127" being greater than the 0.05 level.

c) Estimate another multiple regression model of Wage using Age and Education as independent (X) variables; this time fit Age using a quadratic relationship. Verify your choice from part a. by comparing the distribution of residuals and the goodness of fit between the models in parts b and c.

Compared to the standard relationship when we fit the quadratic model it provides a better fit. We conclude by taking a look at the residual standard error. The quadratic model has a lower standard error meaning it has smaller prediction errors. The higher adjusted R-Squared tells us this model does a better job of predicting wages based on the factors of age and education. It shows that these factors actually matter and aren't just random.

	RESIDUAL STANDARD ERROR	MULTIPLE R-SQUARED	ADJUSTED R-SQUARED
LINEAR	4.678	0.6187	0.6088
QUADRATIC	3.123	0.8323	0.8257

d) Use the appropriate model to predict hourly wages for someone with 16 years of education and age equal 30, 50, or 70.

Age	30	50	70
Predicted Wage	$25.85	$31.54	$26.56

From this we can confirm the quadratic relationship where the downward curve illustrates wage increasing with age and eventually decreasing at some point.

e) According to the model, at what age will someone with 16 years of education attain the highest wages?

According to this model at age around 51 is when someone with 16 years of education will attain the highest wage.

Rental Market in Ann Arbor:

a) Plot Rent against each of the three predictor variables and evaluate whether the relationship is best captured by a line or a curve. Identify variables that may benefit from a log-transformation.

Based on these scatter plots, the bedroom and bathroom variables would likely benefit from a log transformation, while the square footage variable appears to have a more linear relationship.
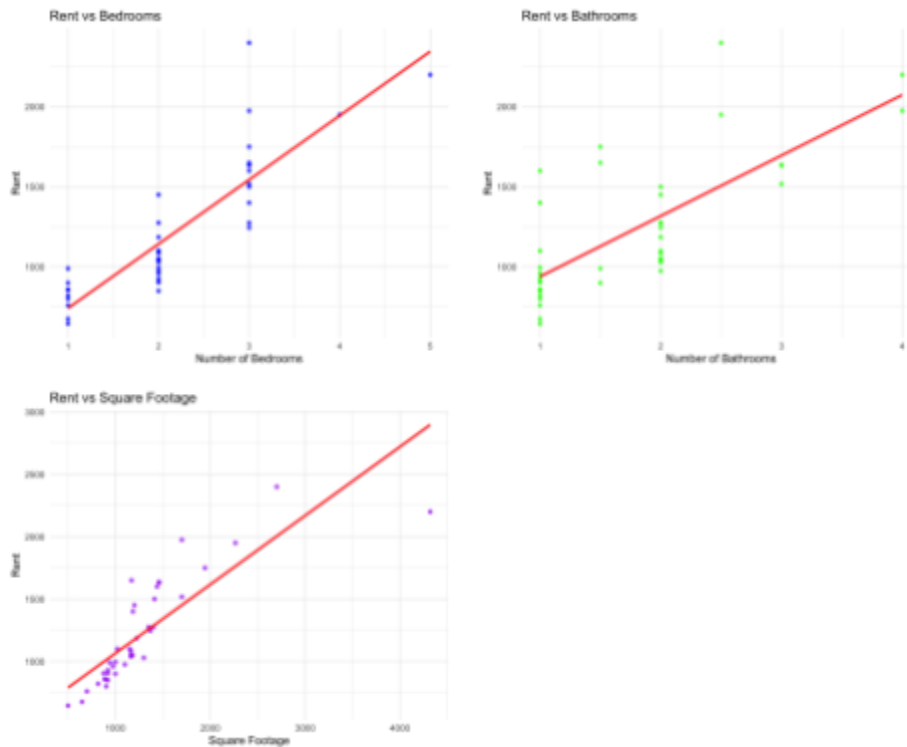
b) Estimate a multiple regression model (with any appropriate log-transformations) to predict rent for a 1,600-square-foot rental with 3 bedrooms and 2 bathrooms.

I used log transformation for both the bedroom and bathroom variable and estimated the rent to be $1504 for a 1,600 square foot rental with 3 bedrooms and 2 bathrooms. This seems to be fairly accurate because if we go back to the dataset I found a 1412 square foot rental with 3 bedrooms and 2 bathrooms to be $1500 in rent.