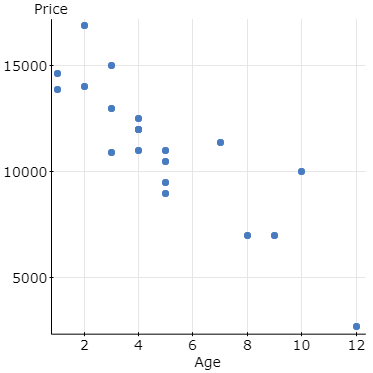
1. Classified ads at www.auto123.com (on July 22, 2013) offered used Toyota Corollas for sale in central Ontario. Listed in the data set UsedCars\_2013 (on Moodle) are ages and advertised prices for some Corolla CEs with automatic transmission.
   1. Make a scatterplot for these data.



* 1. Describe the association between age and price of a used Corolla.

The association is moderate, negative and linear

* 1. Do you think the linear model is appropriate?

Yes, the linear model is appropriate for this distribution.

* 1. Computer software says that R2 = 74.8%. What is the correlation between age and price?

-0.86502557

* 1. Explain the meaning of R2 in this context.

R2 in this context means that the linear model is good for representing this data, but it is not perfect.

* 1. Why doesn’t this model explain 100% of the variability in the price of a used Corolla?

Because there is variation from the linear model.

* 1. Find the equation of the regression line.

y = -935.869x + 15728.996

* 1. Explain the meaning of the slope of the line

For every year that you have the Corolla, the value of the car will decrease by $935.87

* 1. Explain the meaning of the y-intercept of the line.

The price for a brand-new Toyota Corolla should cost you approximately $15,729.

* 1. If you want to sell a 6-year-old Corolla, what price seems appropriate?

$10113.78

* 1. You have a chance to buy one of two cars. They are about the same age and appear to be in equally good condition. Would you rather buy the one with a positive residual or a negative residual? Explain.

I would rather buy a car with a positive residual, because that would mean that the car is priced under where it “should be” according to that linear function of age and price.

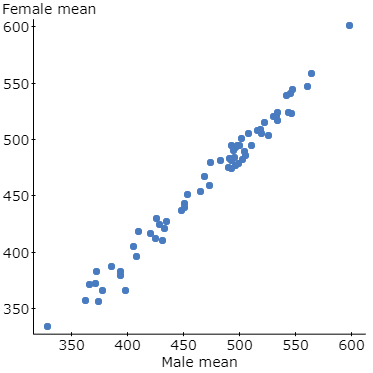
* 1. You see a “For Sale” sign on a 5-year-old Corolla stating the asking price as $9000. What is the residual?

2049.65

* 1. Would this regression model be useful in establishing a fair price for a 17-year-old car? Explain.

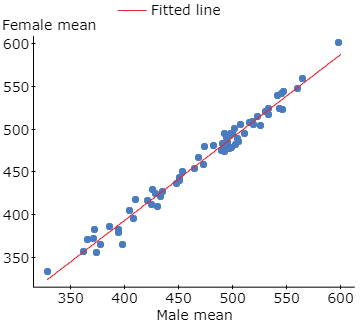
No, this would not be appropriate for pricing a 17-year-old car. Following this model, a 17-year-old car would cost you $-180.777, which would mean you’re getting given both a car and $181, which is not appropriate. This model fits the values given, but in reality, the function would need to have some sort of asymptote and wouldn’t continue decreasing infinitely.

1. In the data set MathAndGEnder\_2009 are mean PISA (Programme for International Student Assessment) math scores for samples of 15-year-old male and female students from a number of randomly selected schools in each various OECD and other industrialized nations (40 000-10000 per country).
   1. Plot female mean score versus male mean score. Describe the relationship.



This relationship is strong positive and linear

* 1. Fit a straight line for predicting female means from male means, and find and interpret the R2 value.



R2 is 0.9799155 – meaning that the relationship is very strong and positive.

* 1. Plot residuals versus male means or versus predicted (fitted) values. What do you learn?
  2. In which country or countries do the female students perform unusually poorly in relation to the male students, according to the regression model? Give the residual(s).

In a lot of the South American countries, mostly Columbia, the males preform better than the females according to the regression model.