

Auto Pricing

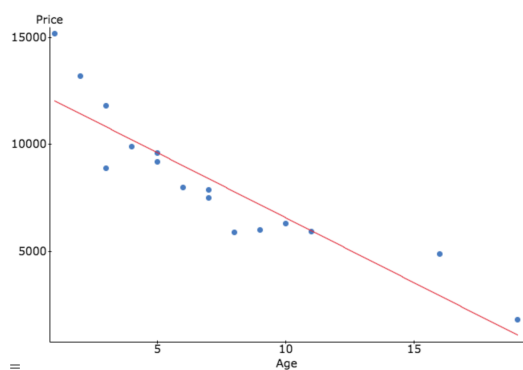
Search craigslist for a popular car type (make and model, e.g. Honda Civic). Randomly sample 20 vehicles of this type from those listed (you may want to refine your search to a particular sub-model (e.g. S or LX) and type of seller (individual or dealer)). Record the asking price, the mileage, and the age (in years since 2020) for each vehicle.

- Find a linear model for the cost as a function of the mileage. Remark on the rate of depreciation (including units).
- Use your result to (a) to predict the cost of a car of the same make and model with 100,000 miles on it.
- Find a linear model for the cost as a function of the age. Remark on the rate of depreciation (including units).
- Use your result to (c) to predict the cost of a car of the same make and model that is 10 years old.
- Comment on which model seems to be a better predictor of cost (the one tied to age or the one based on mileage) – explain. You should include the r^2 value in your commentary and explain to your reader what r^2 tells you about the model you chose to represent the data. Consider other variables that might influence price and discuss them.
- Using the model you chose in part (e), list the price–intercept and interpret its meaning in this context. Explain whether this value reflects reality and if not, give both a statistical explanation for why the model would fail and a real-world explanation for why the model would fail.
- You may have already considered that the best model for predicting price might include both age and mileage. Under the regression sub-menu there is an option for multiple-linear. Choose this and select the appropriate variables. Record the resulting equation and use it to predict the price of a vehicle that is ten years old and has 100,000 miles. Compare this with the results in part (b) and part (d).

★ You will want to use a spreadsheet like StatCrunch or Google Sheets (or MS Excel). Generate a chart from your data for each of the equations you're required to produce. Your charts should show a trend line, the equation of the line, and the r^2 value.

★ Your final paper should be typeset (e.g. Google Doc or MS Word or OpenOffice) and should address the questions and instructions above. Your paper should include your original data (in tabular form) and an introductory paragraph explaining the nature of your study. Write as though you were writing to someone who knows the same level of statistics as you do, but is unfamiliar with the assignment. Remember to include your data (in a table), the type of car, and where you collected your data. Include three scatterplots (showing regression lines) - one for the data used in answering (a), the second for the data used in answering (c), and the third for the multivariable regression. Your paper should also include the tables showing regression results for each of the scatterplots (equation, r and r^2 values).

Example of scatter plot showing regression line and table of regression results:



Simple linear regression results:

Dependent Variable: Price
 Independent Variable: Age
 Price = 12662.962 - 608.25342 Age
 Sample size: 16
 R (correlation coefficient) = -0.90596611
 R-sq = 0.82077459
 Estimate of error standard deviation: 1451.292