

An introduction to Linear Regression with R

AUA.R community

Linear Regression – Overview I

- Linear regression is used to predict the value of an outcome variable Y based on one or more input predictor variables X .

$$Y = \beta_1 + \beta_2 X + \epsilon$$

- We have data and we estimate unknown β that best fit the data. Then we use this formula to estimate the value of the response Y , when only the predictors (X s) values are known.

Linear Regression – Example I

A REAL EXAMPLE

The case study "[SAT and College GPA](#)" contains high school and university grades for 105 computer science majors at a local state school. We now consider how we could predict a student's university GPA if we knew his or her high school GPA.

Figure 3 shows a scatter plot of University GPA as a function of High School GPA. You can see from the figure that there is a strong positive relationship. The correlation is 0.78. The regression equation is

$$\text{University GPA}' = (0.675) (\text{High School GPA}) + 1.097$$

Therefore, a student with a high school GPA of 3 would be predicted to have a university GPA of

$$\text{University GPA}' = (0.675) (3) + 1.097 = 3.12.$$

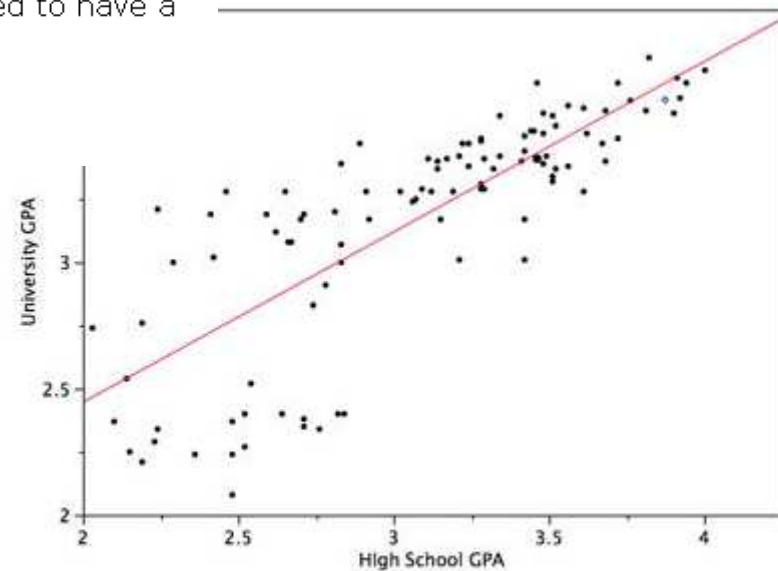
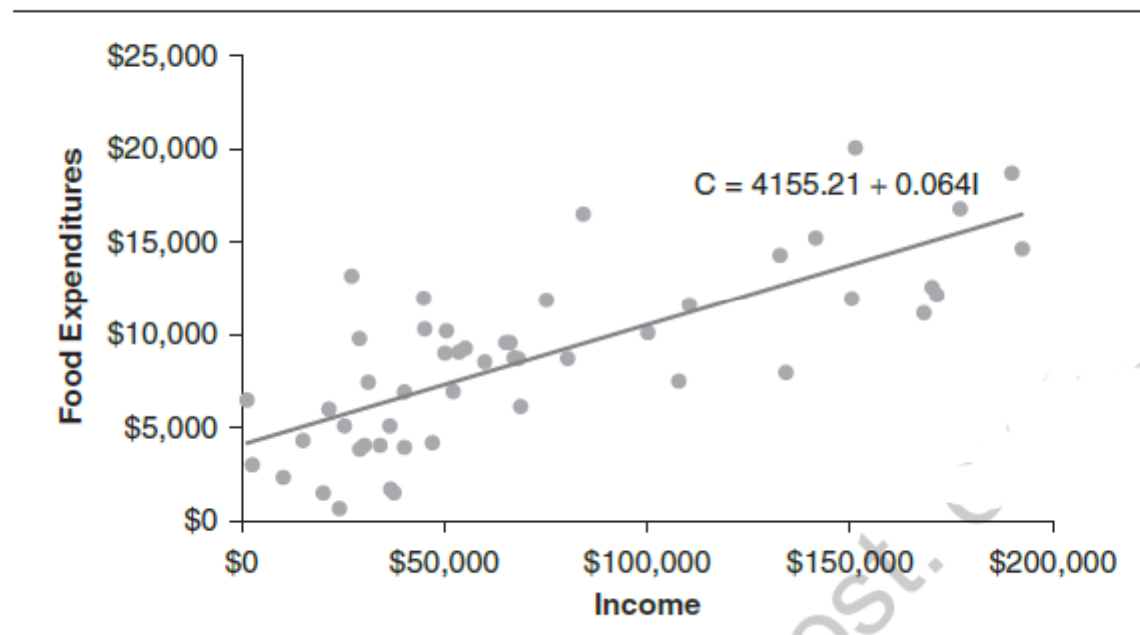


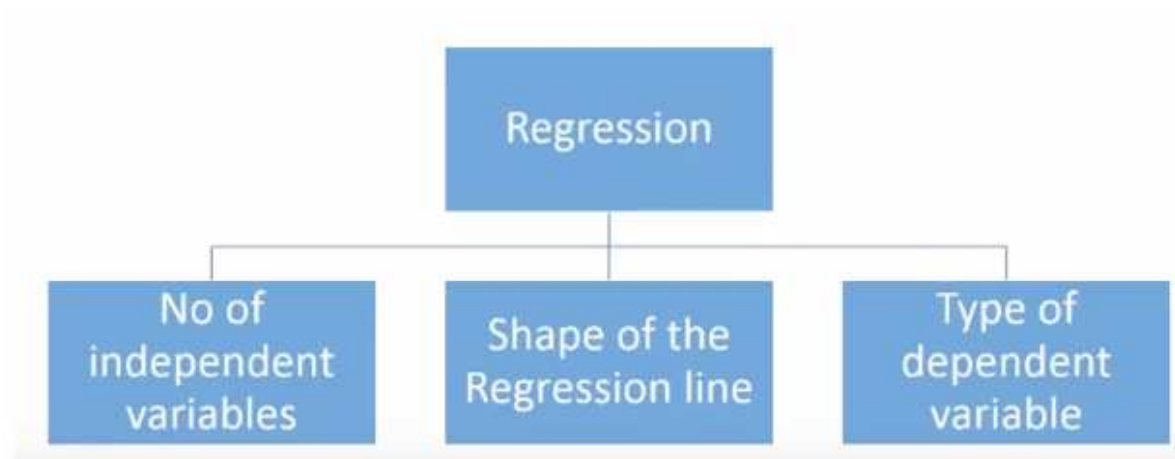
Figure 3. University GPA as a function of High School GPA.

Linear Regression – Example II

Figure 1.3 “Best Fitting” Regression Line



Linear Regression – Types




<https://www.analyticsvidhya.com/blog/2015/08/comprehensive-guide-regression/>

Simple Linear Regression with R

- Go to `linear_regression_code.R`

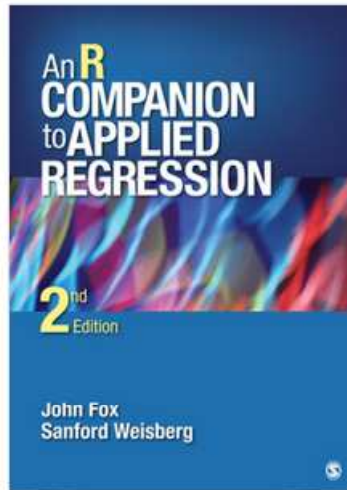
Multiple Linear Regression with R

Are a person's brain size and body size predictive of his or her intelligence?

Interested in answering the above research question, some researchers (Willerman, *et al.*, 1991) collected the following data ([iqsize.txt](#) ) on a sample of $n = 38$ college students:

- Response (y): Performance IQ scores (**PIQ**) from the revised Wechsler Adult Intelligence Scale. This variable served as the investigator's measure of the individual's intelligence.
- Potential predictor (x_1): Brain size based on the count obtained from **MRI** scans (given as count/10,000).
- Potential predictor (x_2): **Height** in inches.
- Potential predictor (x_3): **Weight** in pounds.

Regression resources I



[🔍 Preview now](#)

[⬇️ Save Preview](#)

An R Companion to Applied Regression

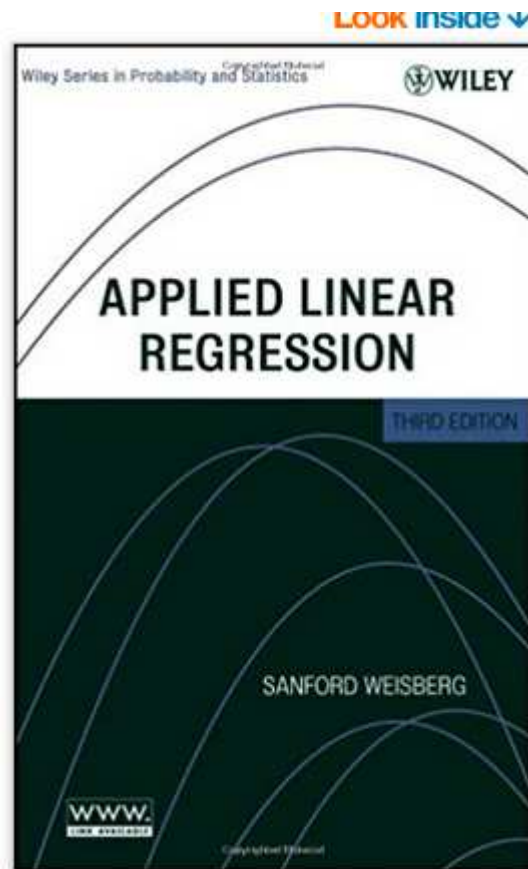
by [Dr. John Fox](#), [Professor Harvey Sanford Weisberg](#)

Synopsis

This is a broad introduction to the R statistical computing environment in the context of applied regression analysis. It is a thoroughly updated edition of John Fox's bestselling text *An R and S-Plus Companion to Applied Regression* (SAGE, 2002). The **Second Edition** is intended as a companion to any course on modern applied regression analysis. The authors provide a step-by-step guide to using the high-quality free statistical software R, an emphasis on integrating statistical computing in R with the practice of data analysis, coverage of generalized linear models, enhanced coverage of R graphics and programming, and substantial web-based support materials.

- *car* package

Regression resources II



- *alr3* package

Regression resources II

- <http://r-statistics.co/Linear-Regression.html>
- <http://tutorials.iq.harvard.edu/R/Rstatistics/Rstatistics.html>