

15.5: Multiple Regression

Simple linear regression as presented here is only a stepping stone towards an entire field of research and application. Regression is an incredibly flexible and powerful tool, and the extensions and variations on it are far beyond the scope of this chapter (indeed, even entire books struggle to accommodate all possible applications of the simple principles laid out here).

Multiple Regression

The next step in regression is to study multiple regression, which uses multiple X variables as predictors for a single Y variable at the same time. In other words, we can use regression to “predict” one score/variable from many other scores/variables, as well as show which of the multiple variables contribute to the score on the target variable. This shows us, statistically, which variables are most related to the changes of the variable that we’re trying to predict.

The general formula is pretty simple:

$$\hat{Y} = a + (b \times X_1) + (b \times X_2) + (b \times X_\infty)$$

But the math of multiple regression is very complex but the logic is the same: we are trying to use variables that are statistically significantly related to our outcome to explain the variance we observe in that outcome.

We can keep adding as many predictor variables as we have data for! Imagine that you’d like to know what all contributes to getting an Associate degree in 2 years.

Note

If my target variable was graduation rates, what variables could affect it?

The first variables that come to mind are how many hours a student works, how many kids in the student's household, and how many units the student passes each semester. Did you think of others? Statistically, you throw all of the variables that you can think of into the regression to try to predict graduation. Then, you use the regression equation to figure out which variables actually contribute to graduation, and which don't. You can see how this "simple" statistical can be a really powerful tool to help college administrators make decisions. The tool can be used in any industry that has a modeling sample (a group of people who can provide all of the variables).

And More!

Other forms of regression include curvilinear models that can explain curves in the data rather than the straight lines used here, as well as moderation models that change the relation between two variables based on levels of a third. The possibilities are truly endless and offer a lifetime of discovery.

Before we wrap-up regression analyses, here are some practice exercises to see if you have the concepts down.

This page titled [15.5: Multiple Regression](#) is shared under a [CC BY-NC-SA 4.0](#) license and was authored, remixed, and/or curated by [Michelle Oja](#).

- [13.6: Multiple Regression and Other Extensions](#) by [Foster et al.](#) is licensed [CC BY-NC-SA 4.0](#). Original source: <https://irl.umsl.edu/oer/4>.