Implementing Multiple Regression Models in Python

Vitthal Srinivasan
CO-FOUNDER, LOONYCORN

www.loonycorn.com

Overview

Implement multiple regression in Python

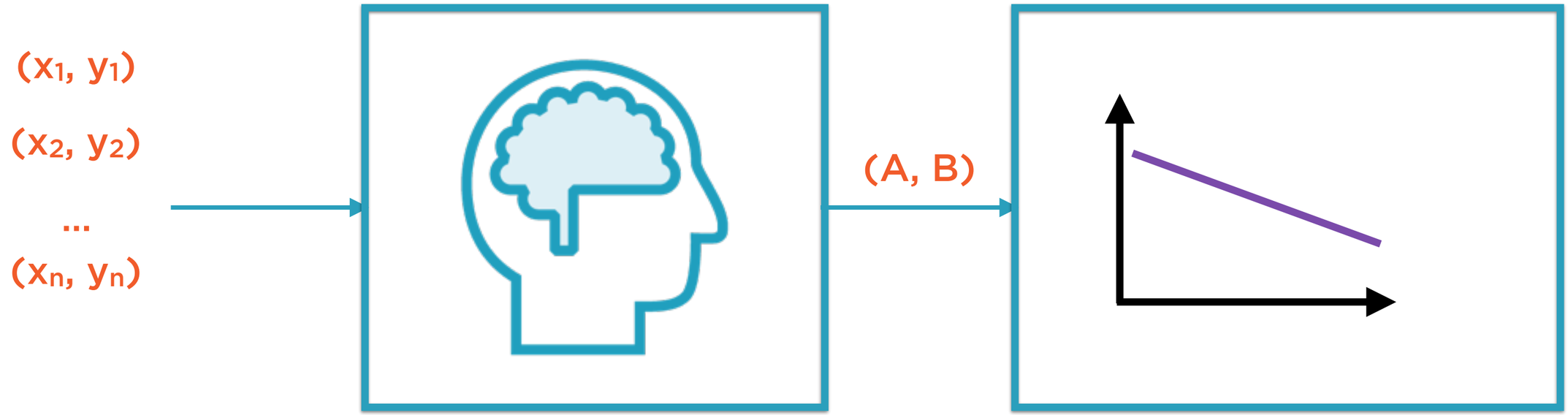
Interpret results of a multiple regression

Carry out multiple regression in Python to include categorical variables

Demo

Implement multiple regression in Python

ML-based Regression Model



Corpus

Regression Algorithm

Regression Line: y = A + Bx

ML-based Regression Model

Corpus

NumPy Linear Regression

Regression Line: y = A + Bx

Demo

Perform regression with categorical variables in Python

Regression Without Intercept

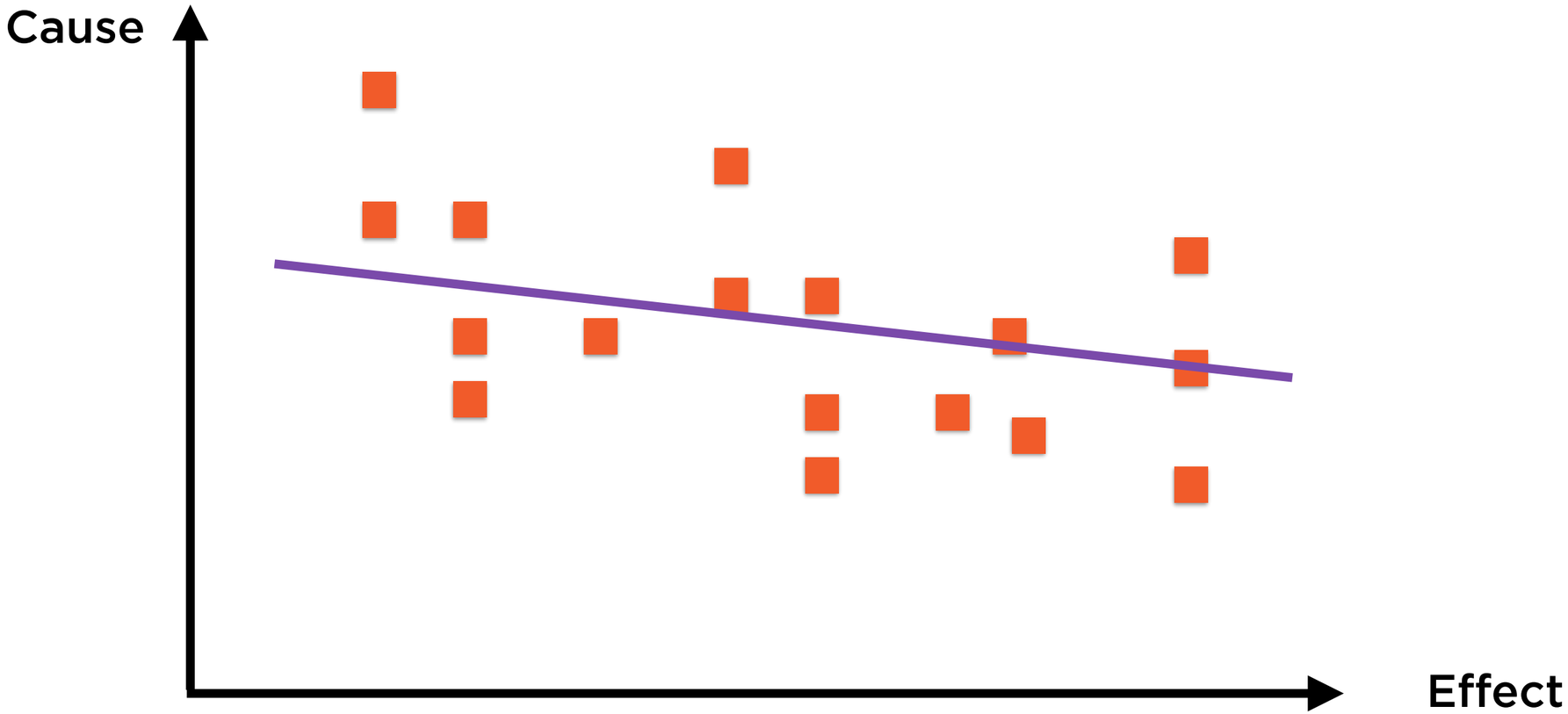
Regression R^2 can go negative

Excel, Python and R all adjust R^2 formula in this case

Excel and R usually agree

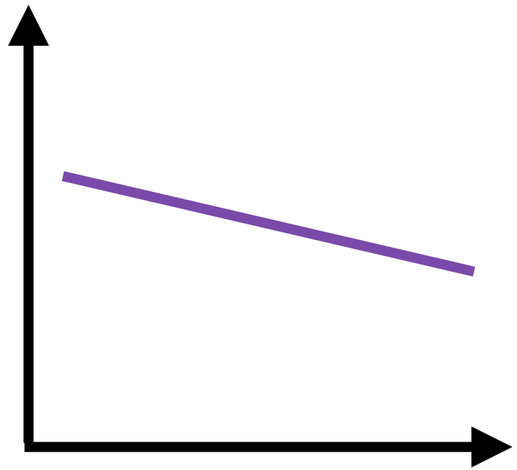
Python statsmodel R^2 sometimes differs

Data in Two Dimensions



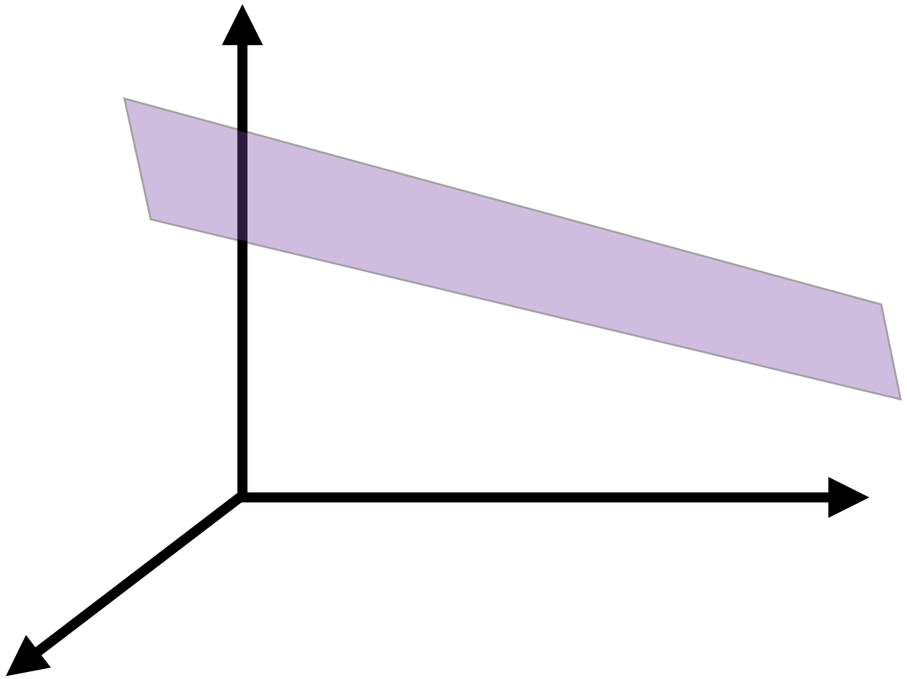
Finding the “best” such straight line is called Linear Regression

Simple and Multiple Regression



Simple Regression

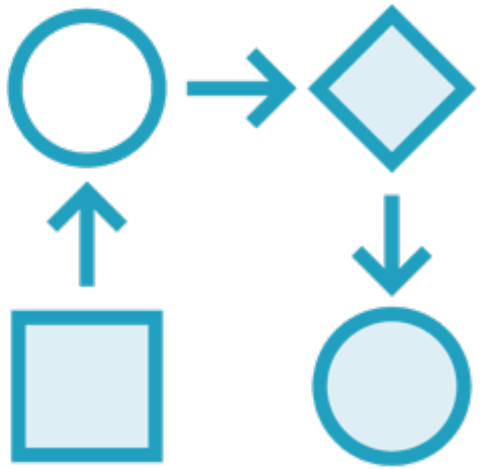
Data in 2 dimensions



Multiple Regression

Data in > 2 dimensions

Two Common Applications of Regression



Explaining Variance

How much variation in one data series is caused by another?

Making Predictions

How much does a move in one series impact another?

Regression Is a Great Tool

Powerful

Perfectly suited to two common use-cases

Versatile

Easily extended to non-linear relationships

Deep

The first “crossover hit” from Machine Learning

ML-based Predictor

Corpus

Regression Algorithm

ML-based Predictor

Regression Line:

$$y = \alpha + \beta x$$

Summary

- Implemented multiple regression in Python
- Interpreted results of a multiple regression
- Carried out multiple regression in Python to include categorical variables