edX **Microsoft:** DAT210x Programming with Python for Data Science

⊙ **Outline** > 1. The Big Picture > Lecture: The Possibilities > Regression

Regression

⊓ Bookmark this page

Regression (Supervised)

The goal of regression is to predict a continuous-valued feature associated with many samples. *Continuous-valued* meaning small changes in the input result in small changes of the output.

Imagine hiking from Portland, OR to Seattle, WA. As time progresses, your distance from Portland increases and your distance to Seattle decreases. Even though you stop for meals and to rest, these distance values transition smoothly. Throughout the course of your hiking, you never magically *teleport* a large distance. Instead, you smoothly and incrementally make your way step-by-step.

With regression, a mathematical relationship is modeled for your samples so that as you gently alter an input feature, an output feature responds by being altered as well. For example, your time journeying and your distance to your destination.
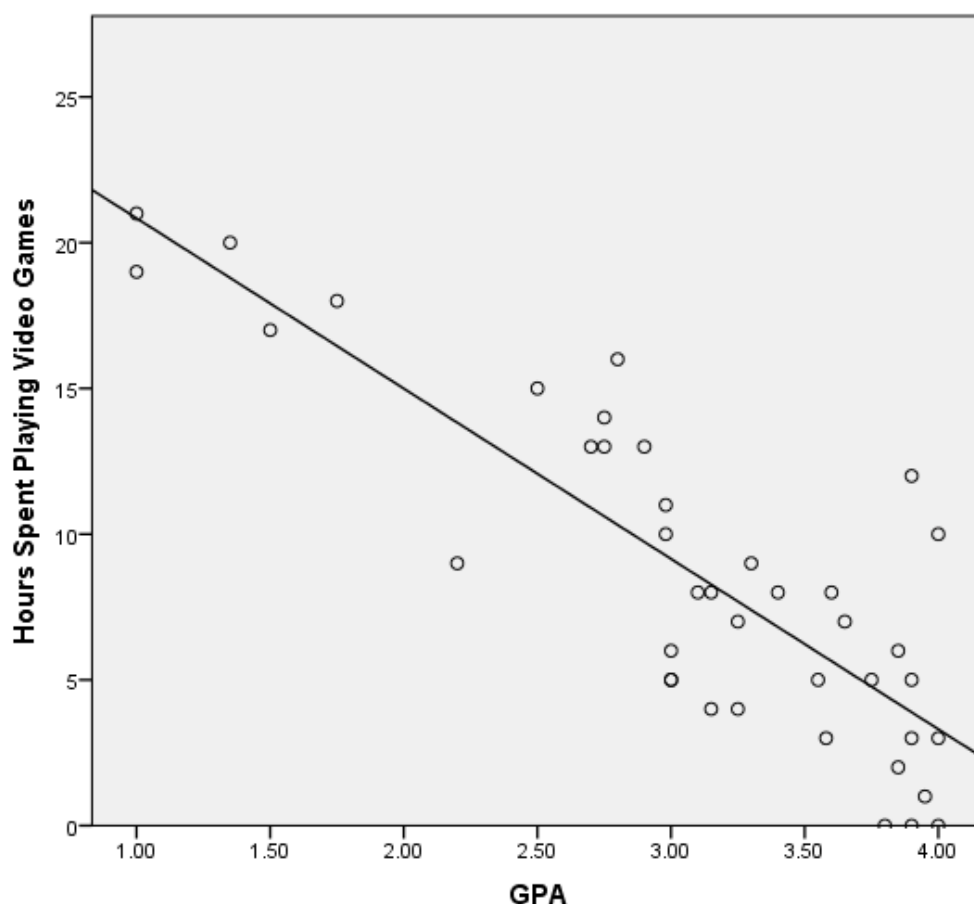
More Examples

- Calculate an equation to predict the size of a house given its price; or the price of a house given its size.

- Explore if a correlation exists between the hours a student spends studying, spends watching TV, and their final exam score.

- Estimate how many power plants should be constructed in the next 50 years, based upon the historical energy consumption per household.

- Figure out how many days a person has left to live based on the severity of their symptoms.

Regression falls into the realm of supervised learning because you have to provide the computer with labeled samples. It then attempts to fit an equation to the samples' features.