The Nature of Regression Analysis

The term regression was introduced by Francis Galton. He found that there was a tendency for tall parents to have tall children and for short parents to have short children. The average height of children born of parents of a given height tended to move or “regress” toward the average height in the population as a whole.

In the modern view our concern is not with explanation but rather with finding out how the average height of sons changes, given the father’s height. In other words, our concern is with prediction of the average height of sons knowing the height of their fathers.

Corresponding to any given height of a father is a range or distribution of the heights of the sons. Despite the variability of the height of sons for a given value of a father’s height, the average (mean) height of sons generally increases as the height of the father increases. The line connecting these averages (means) is known as the regression line. The similar way to approach the heights of boys measured at fixed ages.