

SOS POLITICAL SCIENCE AND PUBLIC ADMINISTRATION

MBA HRD 206

SUBJECT NAME: QUANTITATIVE TECHNIQUES FOR MANAGERS

TOPIC NAME:

What is regression analysis and what does it mean to perform a regression?

Regression analysis is a reliable method of identifying which variables have impact on a topic of interest. The process of performing a regression allows you to confidently determine which factors matter most, which factors can be ignored, and how these factors influence each other.

In order to understand regression analysis fully, it's essential to comprehend the following terms:

- **Dependent Variable:** This is the main factor that you're trying to understand or predict.
- **Independent Variables:** These are the factors that you hypothesize have an impact on your dependent variable.

In our application training example above, attendees' satisfaction with the event is our dependent variable. The topics covered, length of sessions, food provided, and the cost of a ticket are our independent variables.

How does regression analysis work?

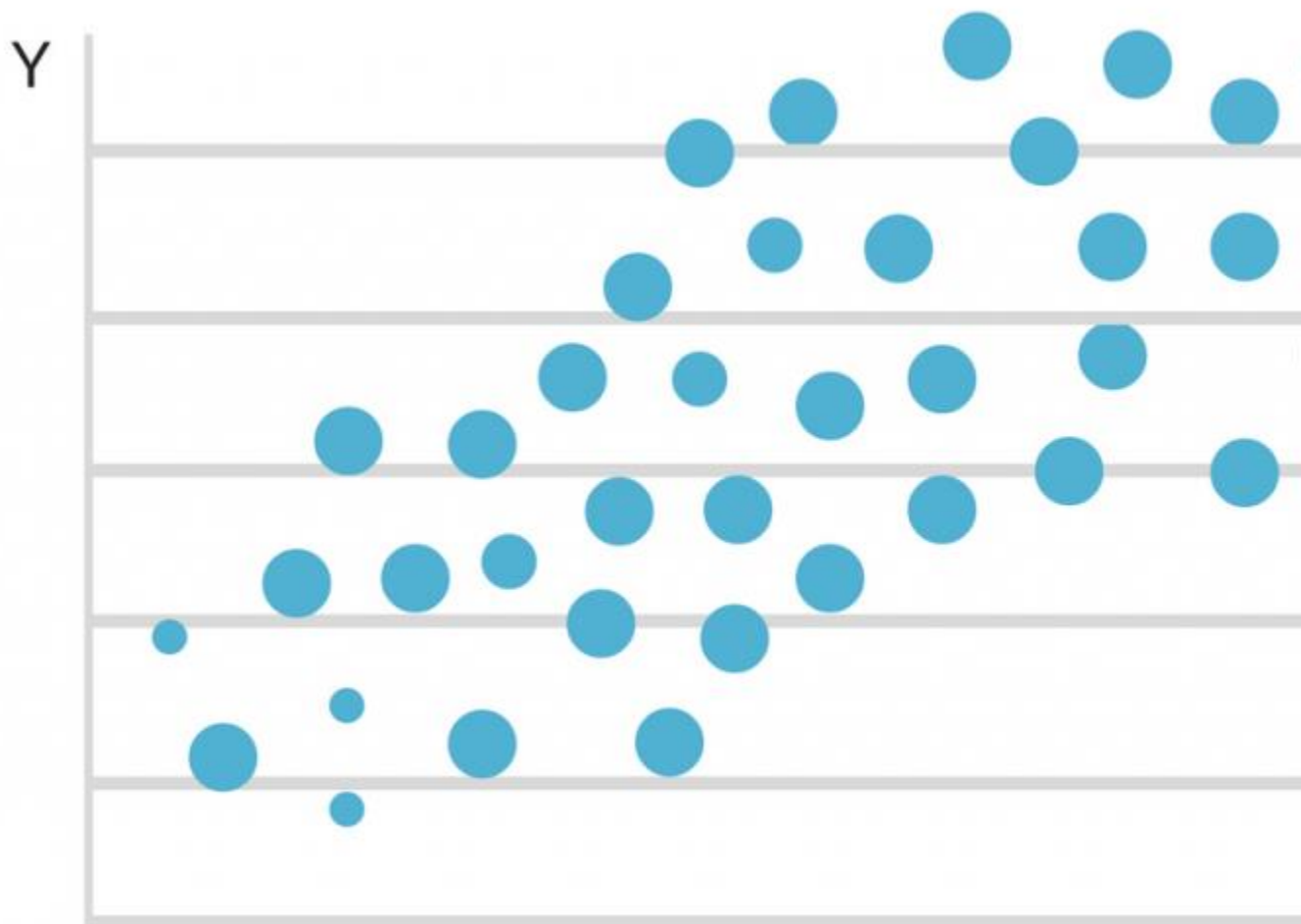
In order to conduct a regression analysis, you'll need to define a dependent variable that you hypothesize is being influenced by one or several independent variables.

You'll then need to establish a comprehensive dataset to work with. Administering surveys to your audiences of interest is a terrific way to establish this dataset. Your survey should include questions addressing all of the independent variables that you are interested in.

Let's continue using our application training example. In this case, we'd want to measure the historical levels of satisfaction with the events from the past three years or so (or however long you deem statistically significant), as well as any information possible in regards to the independent variables.

Perhaps we're particularly curious about how the price of a ticket to the event has impacted levels of satisfaction.

To begin investigating whether or not there is a relationship between these two variables, we would begin by plotting these data points on a chart, which would look like the following theoretical example.



(Plotting your data is the first step in figuring out if there is a relationship between your independent and dependent variables)

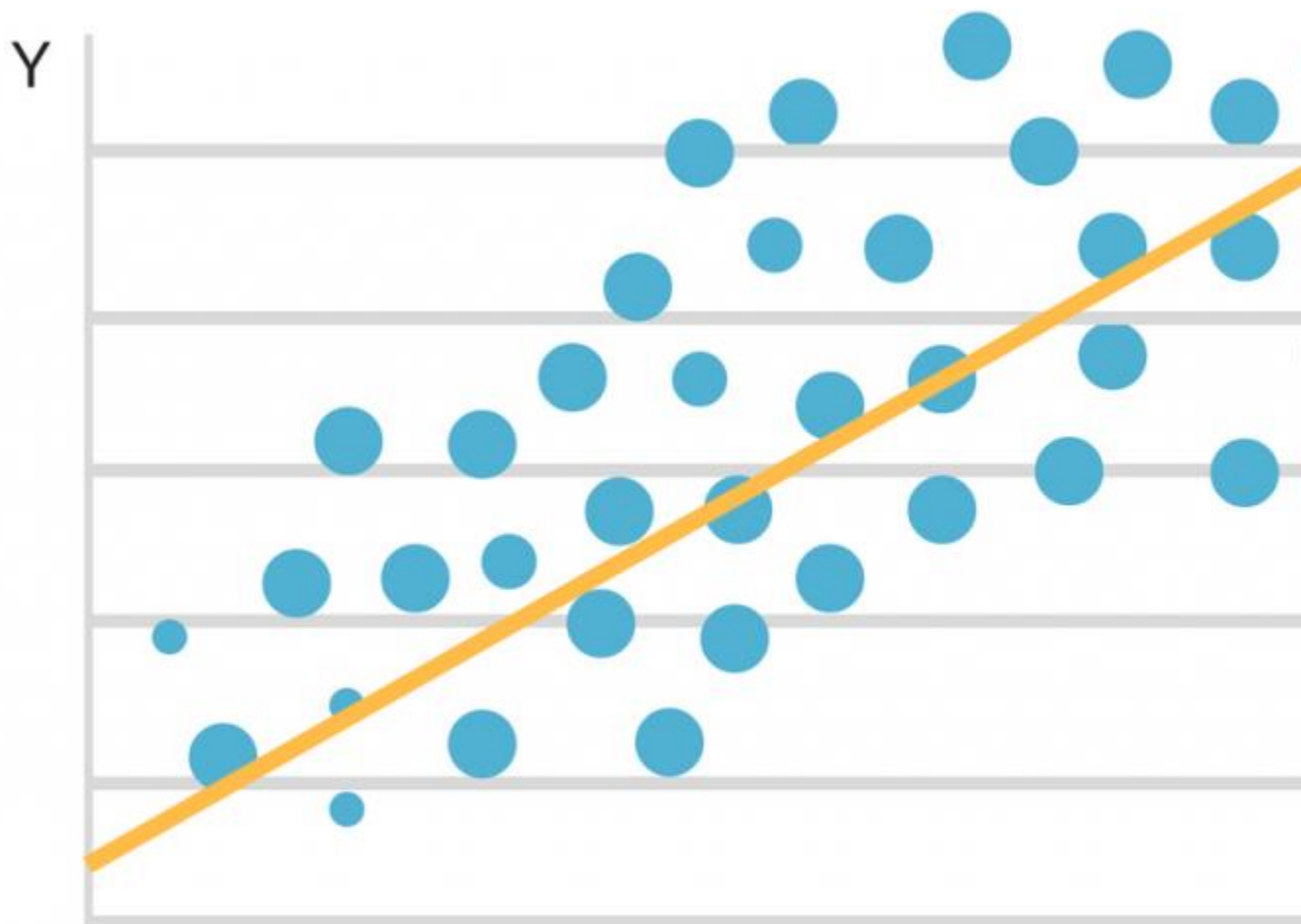
Our dependent variable (in this case, the level of event satisfaction) should be plotted on the y-axis, while our independent variable (the price of the event ticket) should be plotted on the x-axis.

Once your data is plotted, you may begin to see correlations. If the theoretical chart above did indeed represent the impact of ticket prices on event satisfaction, then we'd be able to confidently say that the higher the ticket price, the higher the levels of event satisfaction.

But how can we tell the degree to which ticket price affects event satisfaction?

To begin answering this question, draw a line through the middle of all of the data points on the chart. This line is referred to as your regression line, and it can be precisely calculated using a standard statistics program like Excel.

We'll use a theoretical chart once more to depict what a regression line should look like.



The regression line represents the relationship between your independent variable and your dependent variable.

Excel will even provide a formula for the slope of the line, which adds further context to the relationship between your independent and dependent variables.

The formula for a regression line might look something like $Y = 100 + 7X + \text{error term}$.

This tells you that if there is no “X”, then $Y = 100$. If X is our increase in ticket price, this informs us that if there is no increase in ticket price, event satisfaction will still increase by 100 points.

You’ll notice that the slope formula calculated by Excel includes an error term. Regression lines always consider an error term because in reality, independent variables are never precisely perfect predictors of dependent variables. This makes sense while looking at the impact of ticket prices on event satisfaction — there are clearly other variables that are contributing to event satisfaction outside of price.

Your regression line is simply an estimate based on the data available to you. So, the larger your error term, the less definitively certain your regression line is.

Why should your organization use regression analysis?

Regression analysis is helpful statistical method that can be leveraged across an organization to determine the degree to which particular independent variables are influencing dependent variables.

The possible scenarios for conducting regression analysis to yield valuable, actionable business insights are endless.

The next time someone in your business is proposing a hypothesis that states that one factor, whether you can control that factor or not, is impacting a portion of the business, suggest performing a regression analysis to determine just how confident you should be in that hypothesis! This will allow you to make more informed business decisions, allocate resources more efficiently, and ultimately boost your bottom line.