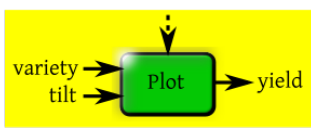
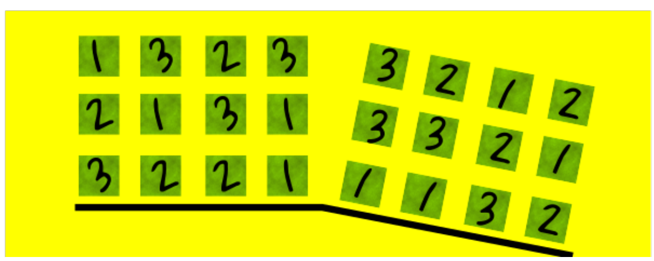
Two-Factor Model: Interaction Chart

Srikar babu Pilli

August 2022

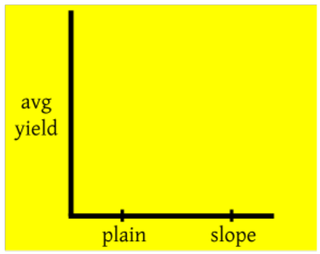
Let us try to analyze a data set from our fictitious example of three varieties,two tilts and 24 plots.We have this setup,which we have already seen:

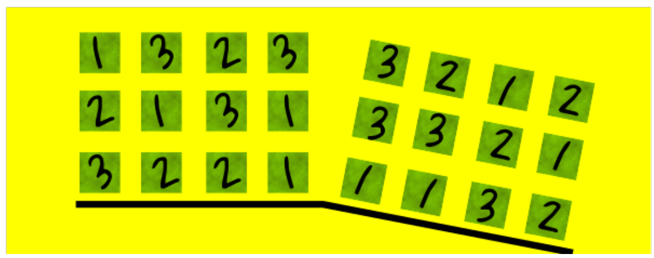




We have got 12 plots in plain land and the other 12 in a sloping land.We have got three varieties of say,paddy and two tilts namely plain and slope and we have measured the yields from each of these 24 plots,so we have got 24 numbers.Now,we have to start analyzing this data.

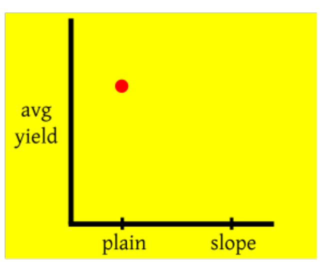
We expect that if we control the tilt,and we sow the same variety of paddy i.e. if we take 2 plots which are in same tilt and use the same variety of paddy,in that case we will expect that their yield will be more or less the same,the only variation will be because of the random error.So,if we take all the four plots(with same tilt,of course)in which a same variety of paddy is sown,then the average of these 4 yields will give us an idea about the effect of e on yield of that particular variety and the tilt combination(remember,there are 12 plots on plain land and the sloping land each and hence we will have 4 plots where variety 1 is sown and 4 plots where variety 2 is sown and 4 plots where variety 3 is sown on both the plain and the sloping lands respectively).So,we can first find out all those averages and it will help us if we use this simple graphical device:

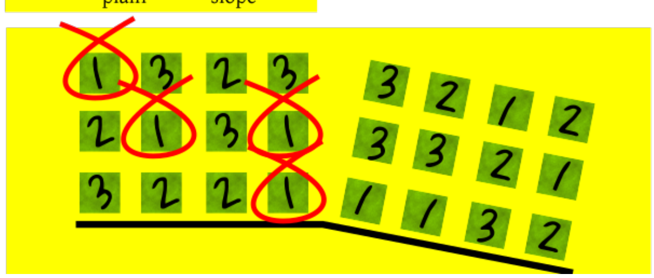




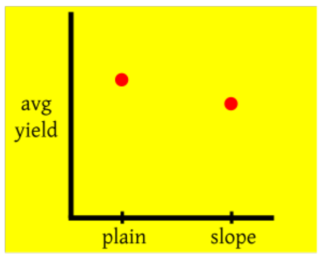
In this graphical device,the X-axis is not a continuous axis,we are showing the two tilts(plain and slope).On the Y-axis,we are showing the average yields for each combination of variety and tilt.We shall plot the average yield as a single point.

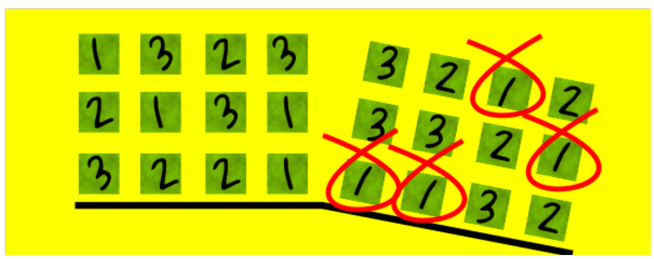
So,we shall start with all the plots in the plain land where we have sown variety one.We will take the yields from all these four plots,average them and plot it at as shown:



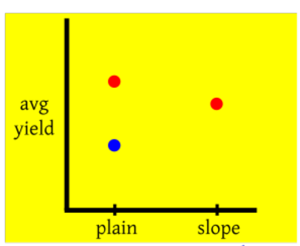


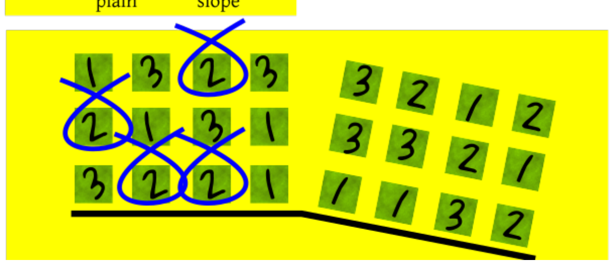
So the height from the X-axis gives the average yield and we have plotted it against plain,to show that it's the average yield from the plots on the plain land and we used the red colour as we are working with plots where we have sown the variety one i.e. red colour here indicates that the average yields are from plots where variety one is sown.Similarly,we shall use blue for variety two and green for variety three respectively.Now,we shall do the same for all the plots in the sloping land where we have sown variety one,so we get another red point as we have sown variety one and we shall plot it against slope as shown:



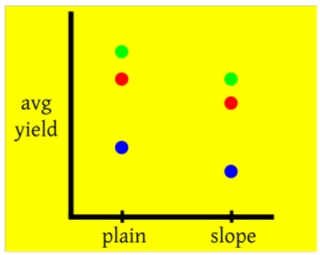


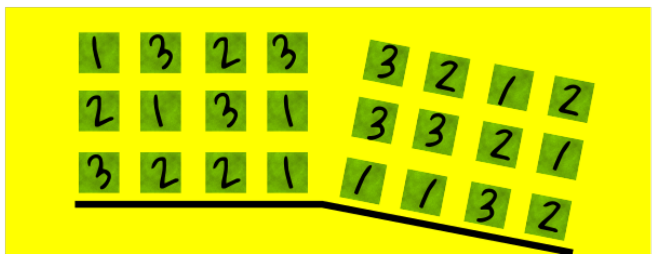
The height of a point from X-axis gives the average yield of that particular tilt and variety combination.Now,suppose we do the same with all the plots in the plain land where we have sown variety two,so we get a blue point now,plotted against plain as shown(blue stands for variety two):



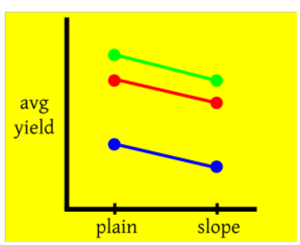


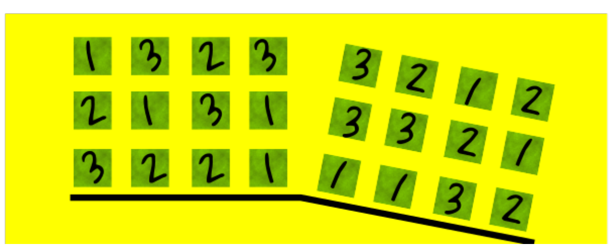
Similarly,if we take all the plots on the sloping land,where we have sown variety two,then we will get another blue point,plotted against slope.We will do the same for all the plots,where we have sown variety three, on the plain and sloping lands respectively and plot them in green against plain and slope respectively as shown(green stands for variety three):





So,we have 6 combinations of variety-tilt combination and we have got 6 dots accordingly(the red dots correspond to variety one and the blue dots correspond to variety two and the green dots correspond to variety three).This is based on the fictitious data,the positions of the points can be completely different based on what type of data we have.Now to just make it visually appealing,we shall join the points of same colour with the lines as shown:





These lines do not mean anything because this X-axis is just a discrete axis,just for visual appeal,we have joined them with the lines and we call this red line as the variety one profile,and the blue line as the variety two profile and the green line as the variety three profile.Now the first thing that attracts our attention for this particular fictitious data set is that all the profiles are parallel.This is interesting,there is no guarantee why they should be parallel,but in this case they are parallel.

Now once we observe this interesting thing,we can say that there is a certain relationship,say for example,if we have the values for say,variety 1,then by adding(in this case,it is rather,by subtracting) a certain constant to this,we can get the values for variety 2 and,that constant does not depend on the slope.The constant that we need to add(or subtract) to go from red point to blue point on the plain land is same as the constant that we need to add(or subtract) to go from red point to blue point on the sloping land.So,the constant does not depend on the tilt,when the things are parallel.We have to capture this thing mathematically.This is not necessarily always the case,but in this case,where the profiles are parallel,we would like to capture this parallel pattern using some mathematical notations.