Six Sigma
Prof. Jitesh J. Thakkar
Department of Industrial and System Engineering
Indian Institute of Technology, Kharagpur

Lecture - 40
Randomized Block Design: Minitab Application

Hello friends, I welcome you to our ongoing journey on Six Sigma. And we are step by step moving ahead in our DMAIC cycle. We have talked about define phase, measure phase, analyze phase and now we are in the improve phase. As a part of this we have started with introduction to design of experiment. And then subsequently we had seen in the previous lecture Randomized complete Block Design. So, now I would like to give you a minute of exposure how we can do the randomized block design in minitab and this will say over time from the manual computation and will help us to focus more on interpretation of the results.

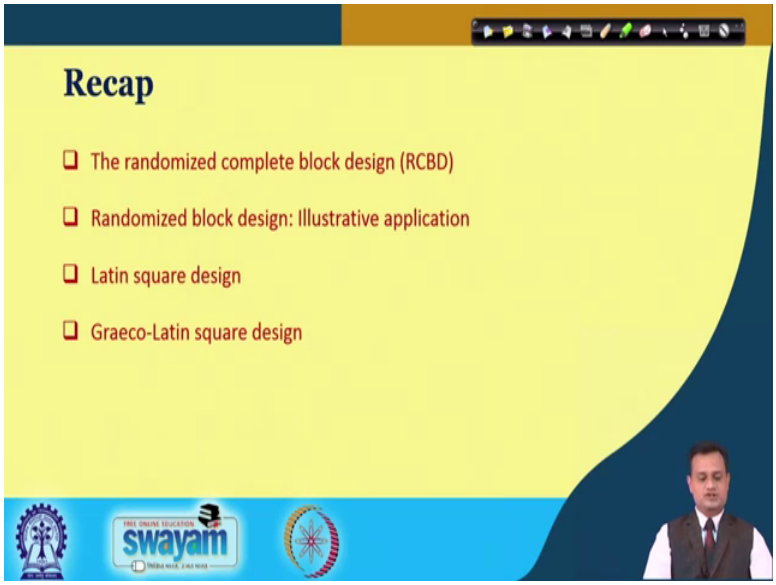
(Refer Slide Time: 01:10)

So, I would like to share one very beautiful quote that experiment, fail, learn, repeat. So, not necessary that experiments they will succeed, they fail, but they provide lot of insights. We can learn from that, we can look at the factors, we can look at the process characterization and we can change our setting factor level and then we can repeat the experiment. So, in case of my improved phase of DMAIC, I would like to see that the

maximum issues can be addressed at the design stage through a powerful technique like design of experiment.

And these improvements once are ensured at the design stage of product or process manufacturing or service, then subsequently I will have the better control to establish on my processes and the better quality a-grade products can be produced

(Refer Slide Time: 02:13)



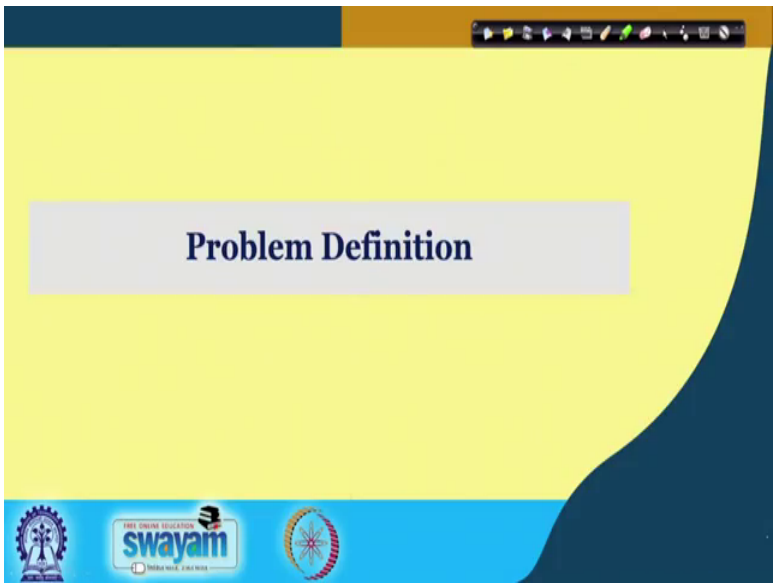
So, this is what we have discussed in the last lecture. The randomized complete block design concept, I need to block the nuisance factor in order to have better inferences, good quality inferences from my design of experiment. So, we had seen the illustrative application, we have also seen that not necessary that you have only one nuisance factor, you might be having two nuisance factor, in that case you would like to go for the Latin square design.

If you have three nuisance factor, you may superimpose two Latin square design which will emerge as Graeco-Latin square design, and you can consider the effect of or consideration of three nuisance factor in analyzing your design.

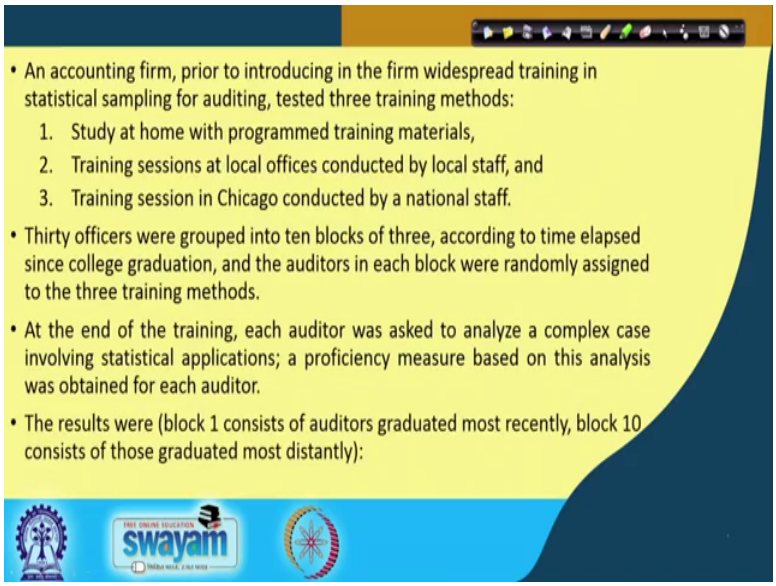
(Refer Slide Time: 03:04)

So, now let us see minitab application for randomized complete block design and the problem definition is like this.

(Refer Slide Time: 03:06)



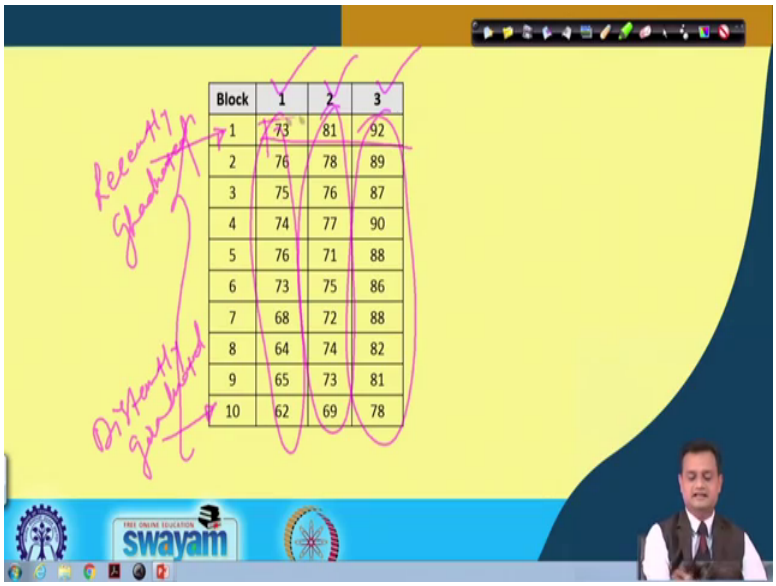
(Refer Slide Time: 03:08)



There is an accounting firm and prior to introducing in the form widespread training in statistical sampling for auditing, tested three training methods. So, basically they are interested to test three training methods. Number 1 study at home with program training material you are given the training material study at home. Number 2 training sessions at local offices conducted by the staff one to one training. Number 3 training session in some Chicago conducted by national staff. So, these are the three different methods of training they want to evaluate.

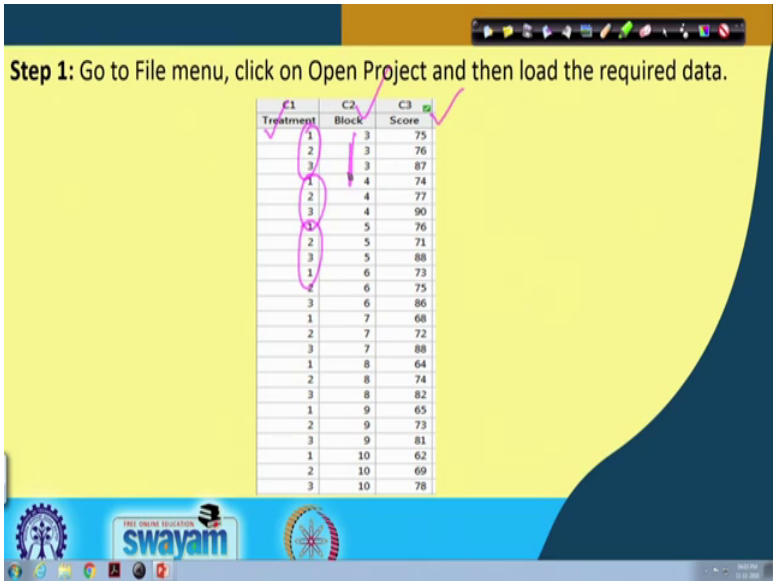
Now, what they have done thirty officers were grouped into say ten blocks, each one will consist three and according to time elapsed since college graduation, the auditors in each block were randomly assigned to the three training methods. So, now there are ten blocks and when I say block number 1, then it consists of auditors graduated most recently. And block number 10, consists of those graduated most distantly. So, somebody has passed out may be in 2017 or 18, I will consider them as part of my block 1, within this block there would be 3 auditors to be tested and these 3 auditor will be assigned to three different methods of training randomly.

(Refer Slide Time: 04:50)



So, this is exactly what I did. And you can see here that this is my block 1 to 10. And when I say block 1 means recently graduated and these are distantly graduated, I have three methods of training 1, 2 and 3. And I am finding the scores of a particular training method how effective it is maybe by conducting that test or examination and these are these scores. Now, within this as I mentioned a particular block this auditors are assigned randomly and these are the scores of those auditors.

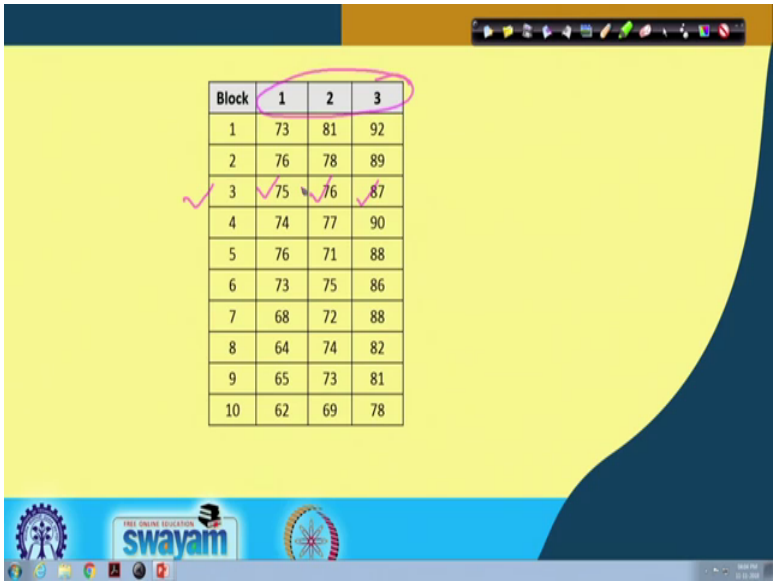
(Refer Slide Time: 05:50)



Now, once you have understood this that you have 30 auditors to be evaluated. Now, you open the step one in minitab go to file menu, then click on the open project, new project, load the required data. So, the data is very simple to interpret. You have treatment 1, 2, 3, 1, 2, 3, 1, 2, 3, it is put, then you have the block and then you have the score.

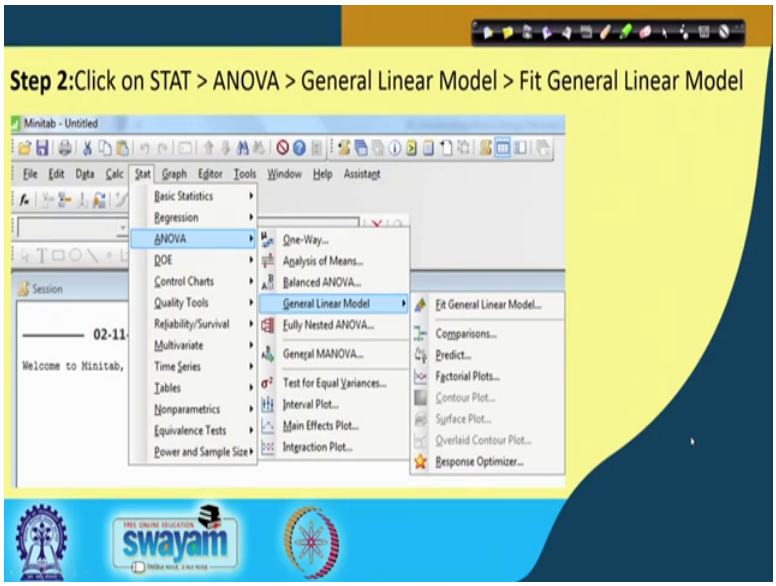
So, let us try to see that treatment means three different training methods. So, I am repeating 1, 2, 3, 1, 2, 3, 1, 2, 3, in order to put my data in the desired format. Now, let us say I have just considered here starting from block three and this is 75, 76, 87.

(Refer Slide Time: 06:51)



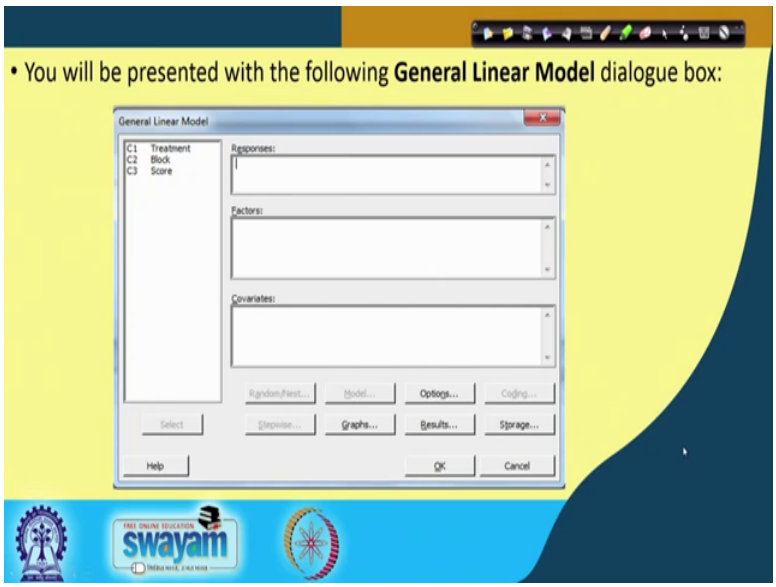
Now, if you just go back and see, then how it looks like say just see this, this is 75, 76 and 87. So, this is 1, 2, 3 training method 1, training method 2, training method 3. And this is my block 3 and score specific to this is 75, 76, 87.

(Refer Slide Time: 07:13)



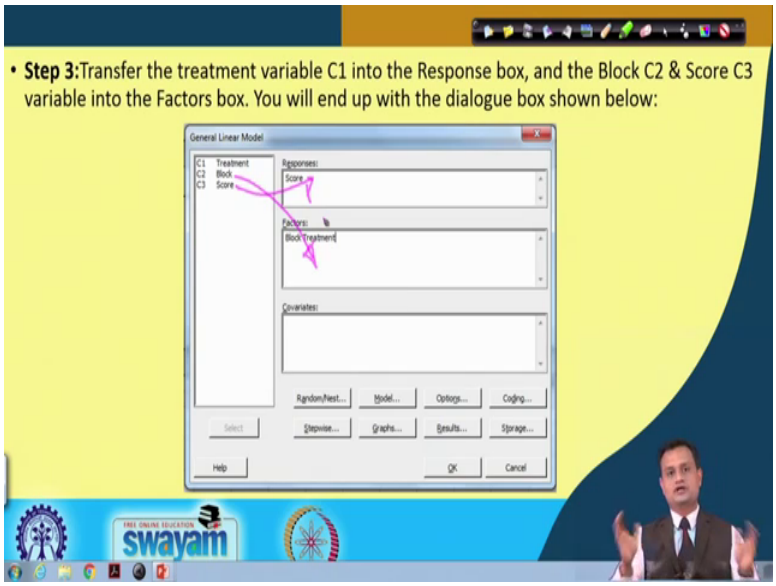
Now, once I put to my data into this sheet, then I will go to step two click on statistics, then go to ANOVA, then go to general linear model, then fit general linear model.

(Refer Slide Time: 07:30)



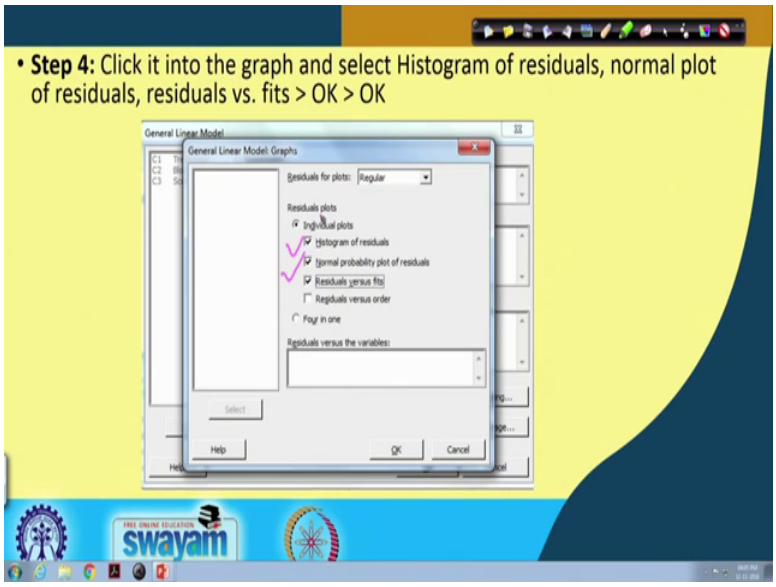
And then you advance further. So, you will have this window and in this window you have to define the response you have to define the factor and you have covariates or simply you be focused on these two. And you have C 1, C 2 and C 3 here. So, C 1, C 2, C 3 basically pertains to your column in the minitab worksheet.

(Refer Slide Time: 07:56)



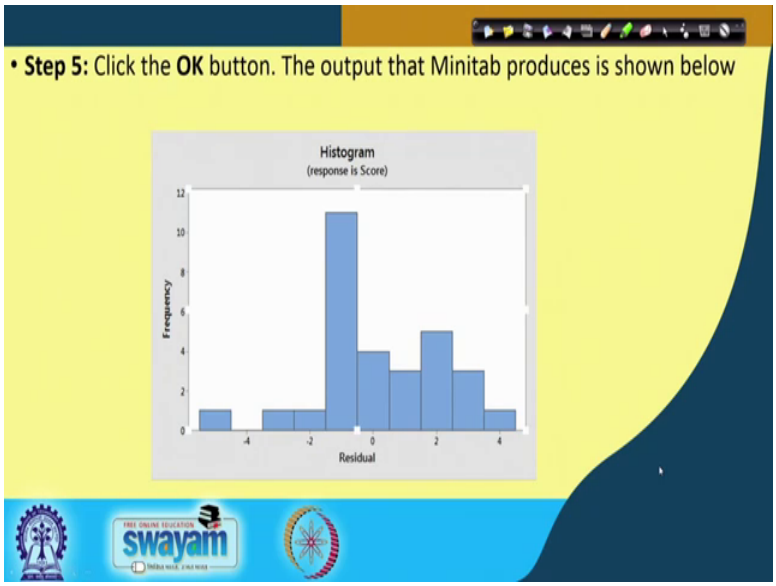
And you have put here or dragged here this score has the response variable, then block has the factor. So, here I have the treatment block and score. So, the factor is the block treatment and I am just trying to dragging it.

(Refer Slide Time: 08:16)



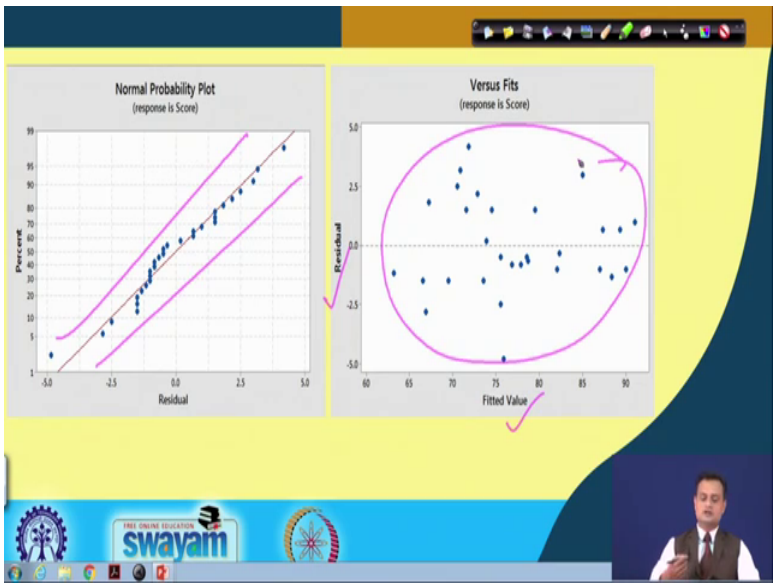
So, now with this data you go to residual for plots that is regular and you click on histogram normal probability, residual versus fits.

(Refer Slide Time: 08:26)



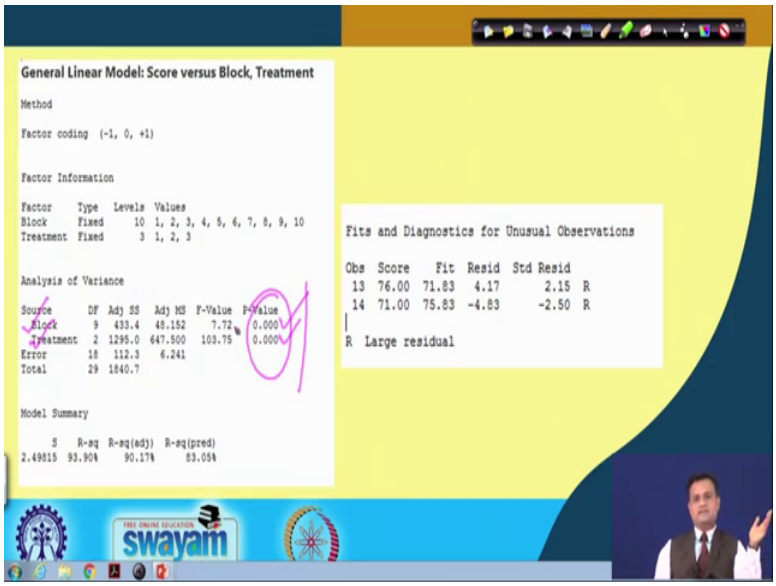
And then you will have histogram that will just give you a brief idea that to what extent your data distribution is normal or not.

(Refer Slide Time: 08:40)



And then you better check the normal probability plot so more or less your line is passing through the data points and you can say that your assumption about the normality is valid. You also check the fitted value versus residual. And you can see that more or less, it is randomly scattered and no peculiar pattern or behaviour is observed. So, I can go ahead with my assumption of independence.

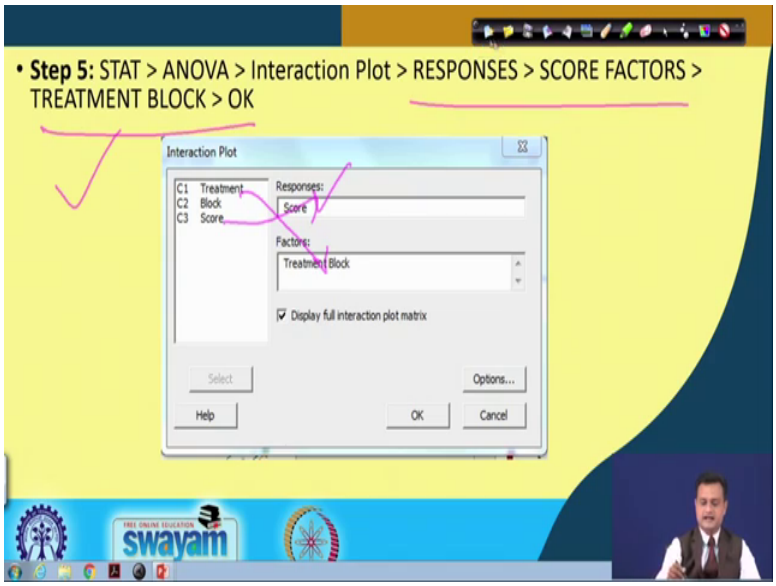
(Refer Slide Time: 09:13)



And then you will have this ANOVA results. So, what you can see here, you just refer the P-value specific to block, refer the treatment value specific to this. Now, here your block means what you have 10 blocks and block means they one block, block 1 means recently graduated auditors and block 10 means distantly may be many years back 10 years, 8 years back graduated auditors that is your block. You have 10 blocks and treatment. There are 3 treatments, number 1 testing a training method, number 2 training method and number 3 training method.

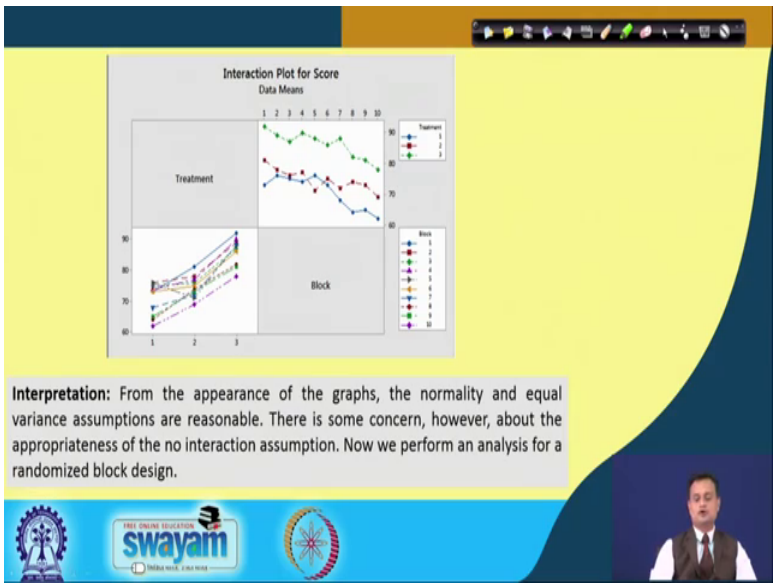
Now, what you can see here that both the P-values are 0, and even if you check it for alpha is equal to 0.05 level of significance, then they both fall in the rejection region, it means you say that null hypothesis is rejected. It means there is a significant block effect, it means those who have graduated distantly and those who have graduated recently, yes, there is a difference in their score and another is the treatment. So, there is a significant difference in the training method, 1, 2 and 3, so you cannot say that all the training methods are equally effective.

(Refer Slide Time: 10:41)



So, this is how you can interpret the results and as I told minitab helps us to focus more on the interpretation, and you can analyze the data very easily without having much difficulty, complexity in manual calculation. So, then step 5, you go to stat ANOVA interaction plot because we have not yet checked the interaction response score factor treatment block ok. So, here you drag the score, this is your C 3 and then you have the treatment and block that is your factors. So, once you have done this, then you press the ok and go ahead in your minitab window.

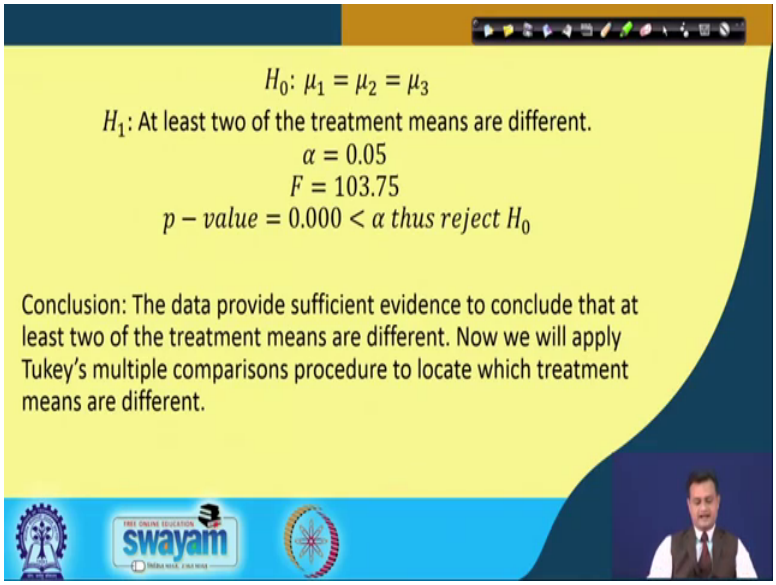
(Refer Slide Time: 11:19)



So, you will have interaction plot for this course and what you can see here that from the appearance of the graphs, the normality and the equal variance assumptions are reasonable. There is some concern, however, about the appropriateness of the number and no interaction assumption. So, here you say are not very much confident to say that there is no interaction there could be an interaction, interaction between what interaction between the graduation year recently graduated, distantly graduated and the training method.

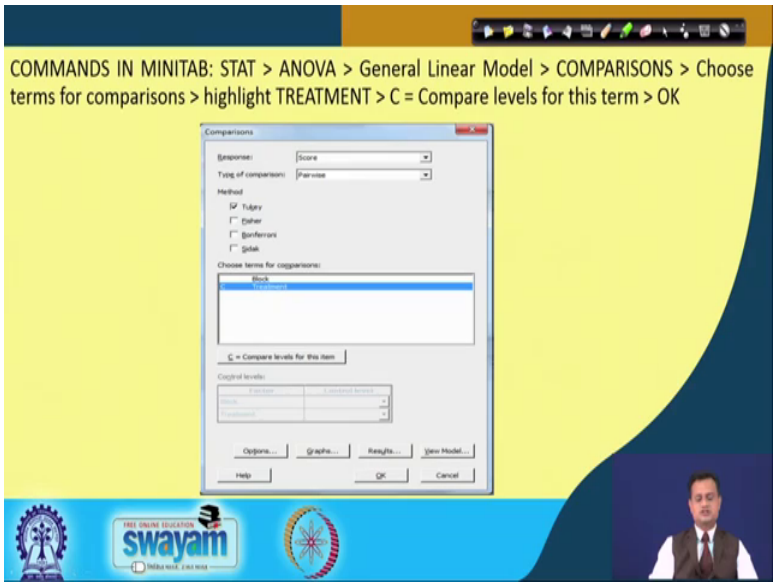
So, many a times say young people they are more comfortable with the distance learning, but old age people them they need one to one interaction and this can happen. So, there could be some interaction effect possible.

(Refer Slide Time: 12:18)



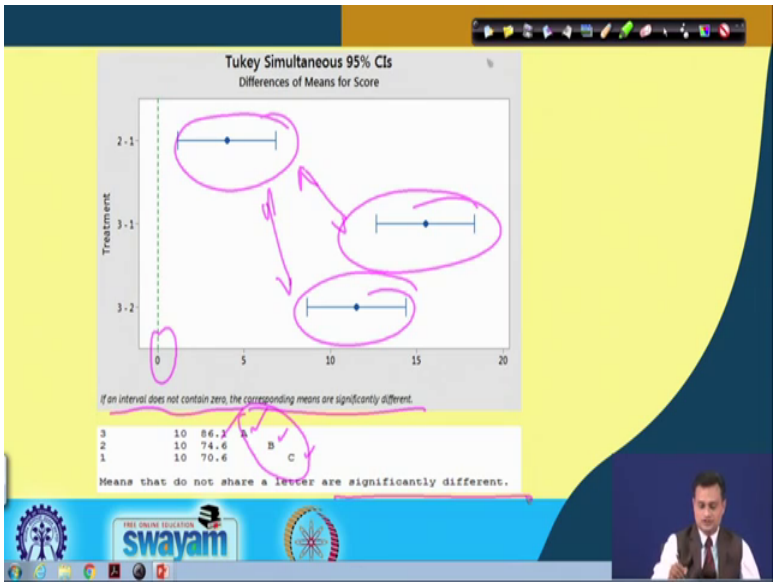
So, this is what exactly we can figure out. So, you have the two treatment means are different if I want to check.

(Refer Slide Time: 12:25)



Then you can just operate with the Tukey test. So, go to minitab stat I have written here ANOVA general linear comparison and then treatment and compare, so, press ok. So, you will have this Tukey test to click, so that will basically compare the pair of means.

(Refer Slide Time: 12:48)

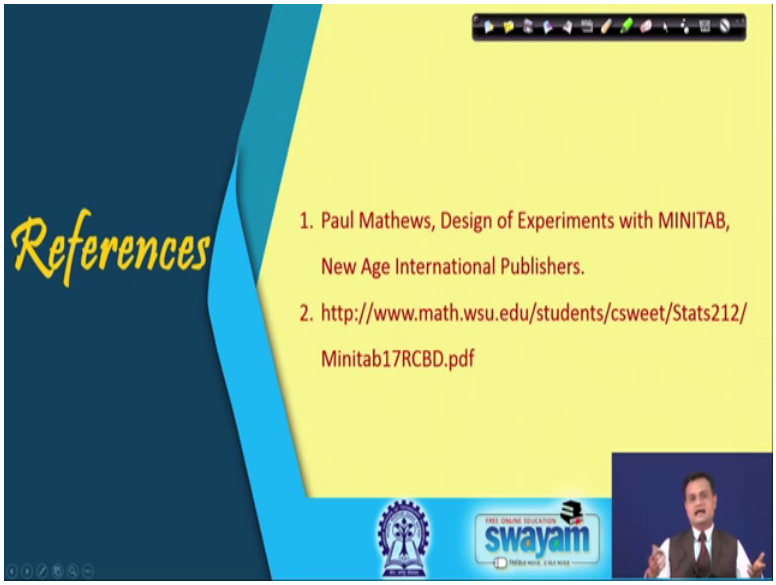


And what you will get that if there is significant difference in the mean score of three different methods, here you have three different training methods, then typically this interval does not contain 0. So, 0 is here and you can see that no interval typically contains 0 and the corresponding means are significantly different. So, we can see that

this difference lies between these and these and these and these. And you can also check it as per minitab say understanding that means, that do not share a letter. So, here all these three are different A, B, C are significantly different.

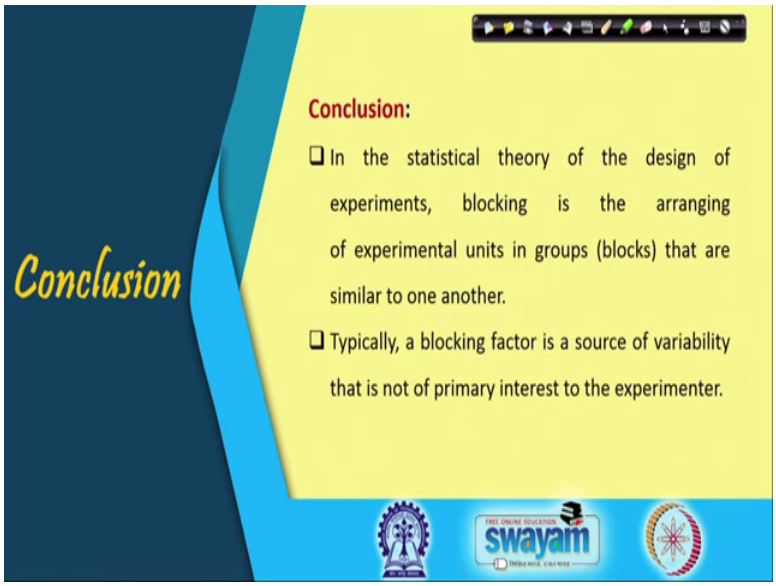
So, my conclusion here is that yes all the three training methods are different in terms of their effectiveness in imparting the knowledge and this is reflected in the score of the candidates. So, this is what we can do in say minitab and that would really help you to conduct the analysis with ease and comfortably say analyze and interpret the results.

(Refer Slide Time: 13:55)



You can use these references. You can also refer this book called Mathew, Design of Experiment with MINITAB, New Age International Publisher. And I have also indicated the link.

(Refer Slide Time: 14:10)



So, conclusion, in this statistical theory, the design of experiment blocking is the arranging of experimental unit in groups or blocks that are similar to one another within that block they are homogeneous. And typically blocking factor is a source of variability that is not of primary interest to experimenter. So, something in which I am not interested I must try to block it, so that I do not have that particular effect coming into my results, and my results should not get distorted.

So, this is what I wanted to teach you as a part of minitab exposer. And I hope you have appreciated just take a hypothetical dataset and try to conduct this analysis follow the steps indicated and see that whether you can really interpret the results and conduct the RCBD in minitab or not.

So, thank you very much, be with me, enjoy.