Business Analytics for Management Decision
Prof. Rudra P Pradhan
Vinod Gupta School of Management
Indian Institute of Technology, Kharagpur

Lecture – 29
Predictive Analytics (Non linear Regression Modelling)

Hello everybody, this is Rudra Pradhan here; welcome you all to BMD lecture series and that too today's discussion is on Predictive Analytics; that is part 1. And in the last couple of lectures, we have discussed about the regression modeling; that too the simple one and the multiple one.

So; that means, typically right now we get to know how you will do the kind of predictions by using the regression technique? So, in the regression kind of structure; so you must have actually the one dependent variables and at least one independent variable and if that is the case that is called as simple regression modeling.

And in the kind of complex environment, so you may have the situation like one dependent variable with multiple independent variables. And again more complexity, more dependent variable with more independent variables; usually that is the particular structure called as a multivariate regression modelling.

And for example, like structural equation modeling, simultaneous equation system; so, these are the examples of multivariate case. But this is again the complex kind of modelling for prediction, but right now we are dealing with the problems in the context of multiple regression modeling; and where we have one dependent variable with the many independent variables.

In the previous class, we have already discussed about the particular case with respect to two independent variables and one dependent variables and we have already gone through the kind of process. So, whether it is you one independent variable or two independent variable or three independent variable; so typically, so you have to connect the kind of structures and then you will get the regression output. And on the basis of regression output, you will do the predictions and then you can go for the kind of management requirement or management decision.

But so, connecting these variables that too dependent with you all independent variables is very easy and having the data and having the software and you can quickly get the kind of regression outputs. But in many instances; so when we will get the regression output; so, these outputs may not be actually reliable. So, it may have some kind of problem which in fact, already highlighted these specific issues like; if the variables are not statistically significant or your R square is not very supportive, then in this kind of situation your regression output or the kind of model; which you have built for the prediction and the kind of management decisions may not actually work.

So, then what will you do? So, you have to re estimate and restructure the process till you get the particular model, where you can do the better predictions and then it can go for better management decision. So, technically it is a kind of continuous process; continuous search process, till you get the best model as per the business requirement that too predict the dependent variable and then go for the kind of management decision.

So, now in this lectures; I will highlight some of the problems and then I will connect with one particular problem that is what it is called as the structure of the non-linear regression modeling. Till now, whatever we have discussed that is with respect to simple one and the multiple one. So, we have not discussed about the non-linearity kind of regression modelling.

But here, we will be discussing the concept called as non-linear regression modeling. So; that means, in the last couple of lectures we have discussed this model that for the kind of predictive analytics, various models dealing with linear relationship among the dependent variable and independent variable. But in a real life scenario, you will find plenty of problems are there; they are actually connected or associated in a kind of non-linear format so; that means, they are not actually linearly related.

So, if they are linearly related then it is very easy to deal with the problem and do the kind of predictions. But if that is not the case then it will give you some kind of problems and this problem itself will get a kind of situation, where the R square value may not be supportive or if R square supportive; then the significance of the parameters may not be supportive.

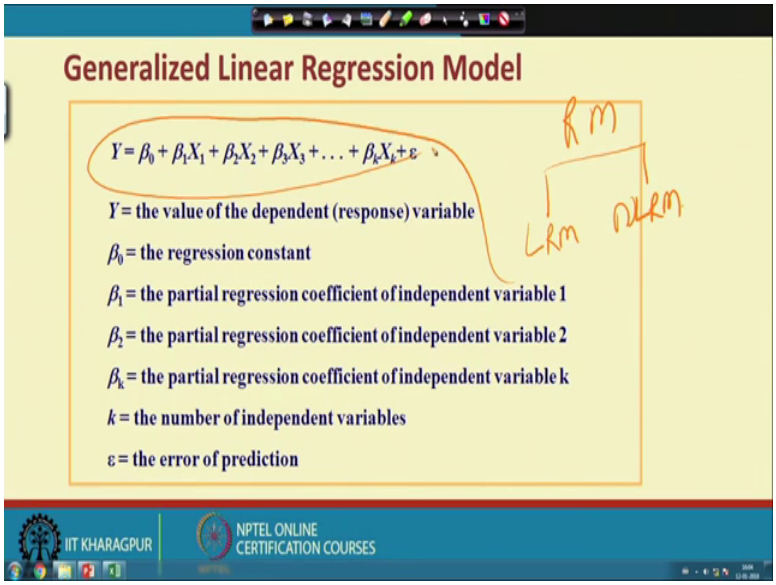
So, what is actually required here; so, you have to go for the kind of visualization process with respect to the data and the kind of problem. And check the right

functionality or functional form and once mark the actual functional form, then you can develop the model accordingly. And once your function of the problem or the building of the model is perfect, then the estimation process will not be having some kind of problem.

So; that means the fast problems may be with respect to the kind of functionality between dependent variable and independent variables. So, you have to fix the models as per they are actually the kind of the actual linkage; whether the linear linkage or the non-linear linkage. So, if it is following non-linear in case; then you have to follow that one only; then that will give you the better predicted models and then the management decision will be very effective.

So, with this basic introduction about this particular process; so, we will go for the kind of discussion.

(Refer Slide Time: 06:15)



So, generalized regression modelling is like this and then what will you do here in this particular; this is actually the multiple regression modeling, and in this case the typical structure is that there is a linear relationship among the dependent variable and independent variable.

So; that means, your regression modeling; there are many different ways you can actually classify the regression modeling. Because this is a very interesting or very

reliable tool and it will not actually directly help to predict the kind of business problems and then to go for management decision; it will also give some kind of solutions indirectly by helping other models as well.

So, that is how this particular technique is very useful technique so, far as predictive analytics typical is concerned. And it is also give you the kind of root for prescriptive analytics. However the requirement of predictive analytics, it depends upon the descriptive analytics and the inferential analytics.

So, once they are well connected; then you will get the kind of predictive analytic structures and that will give you the route to the kind of prescriptive analytics. So, once you know all these details then automatically you can get to know the kind of linkage and that time you can understand that what is the beauty of this particular predictive analytics; through which you can predict the business requirement and go for the kind of management decision.

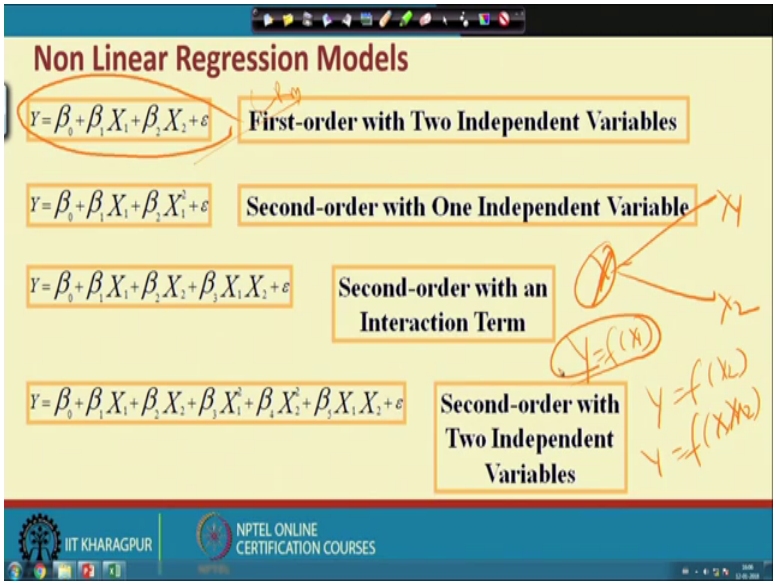
So, by default regression modelling is two types; so the linear regression modelling and non-linear regression modelling. So, in the linear regression modeling; so, there is a kind of the expression is called as linear combination of all these variables. And in the non-linear regression modeling, it may be any form; it may be it means depends upon the kind of functional form.

So, it may be in inverse function; it may be quadratic function, so it depends actually. So, what will you do? So, you will go for the data visualization and then pick up the right functional for; so, once you pick up the right functional form; then the model will be estimated on the basis of that functional form only.

So, then it will give you the perfect fit models through which; there is a high chance that you will get all these parameters; statistically significant and then the goodness fit of the model only be also very perfect or very accurate; so that you can do the better prediction and to come to a better management decision. So, likewise so, we are targeting here the non-linear regression modeling, so we get to know what is actually the concept of non-linear regression modelling and in fact, there are so, many other issues are there, but this is a typical structure through which you have to be very careful. Because most of the business problems may; have some kind of connection about the non-linear relationship

between dependent variable and independent variable. So, having this background; so, let us say go through the kind of structure.

(Refer Slide Time: 09:26)



So, here; so far as a non-linear regression modelling is concerned; so, I am just giving you some kind of snapshots here. So this is actually simple regression modeling and this is with respect to dependent variable and two independent variable and that too first order, which you can represent first order. And that too it is a called as linear regression modeling, so it is actually having linear relationship between Y and X 1 and X 2.

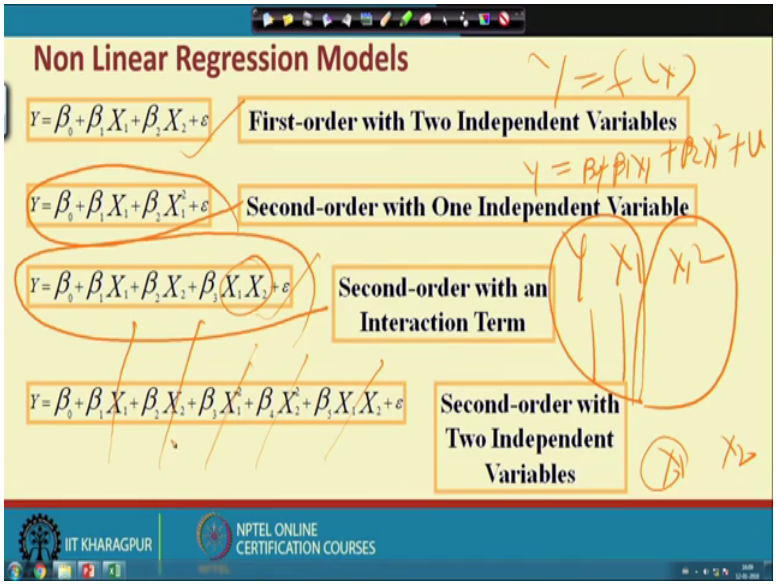
If that is not the case so; that means, in our basket; we have actually Y and then it will be connected with the X 1 and X 2. So; that means, typically you may have a two different models or three different models. So, first models can be Y equal to function of X 1's and Y can be equal to function of X 2 and then Y can be function of X 1 and X 2; so, if it is F equal to function of X 1.

So, that can be actually can be cleared further kind of multiple framework by doing the kind of study the interactive effect, kind of integration with the functional for so; that means, with the particular problems. So, you can create the particular structure through which you can do the right modeling; then go for the right prediction.

So; that means, in a real life scenario; so, you will have actually such kind of environment; so here this is actually the simple one. So, this is called as a linear

regression modelling and then with respect to X 1. So, what we will have actually; it is called as quadratic equations so; that means, we have actually Y; X 1 and X 1 square.

(Refer Slide Time: 11:20)



So; that means, typically we start with the game Y and X 1; then we connect with the X 1 square. So, as a result; so, initially it will be Y equal to function of X 1 then finally, the model will be Y equal to beta 0 plus beta 1 plus beta 1; X 1 plus beta 2; X 1 square plus error terms.

So, this is how the kind of simple model can be represented and with respect to two independent and two independent variable; the non-linearity representation can be like this also. So, this X 1 and X 2 is called as interactive effect; sometimes the interactive effect can also studied to predict the Y. For instance, while predicting sales with respect to price and price or advertising.

So, the link between price and advertising should be connected with the particular sales. So; that means, how much advertising expenditure and how much particular price will be actually required. So, that if the right combination is not there; then the sales prediction subject to price and advertising may be getting affected.

So, as a result; so, this may be an issue, so what will you do? So you have to find out the kind of linkage and then go for the kind of estimations through which you can do the predictions. And again with respect to three variables, so this is X 1 variable, X 2

variables and then you will have a X 3 variables and against you will find a couple of interactive terms like X 1, X 2; X 1, X 3 and X 2, X 3 like this

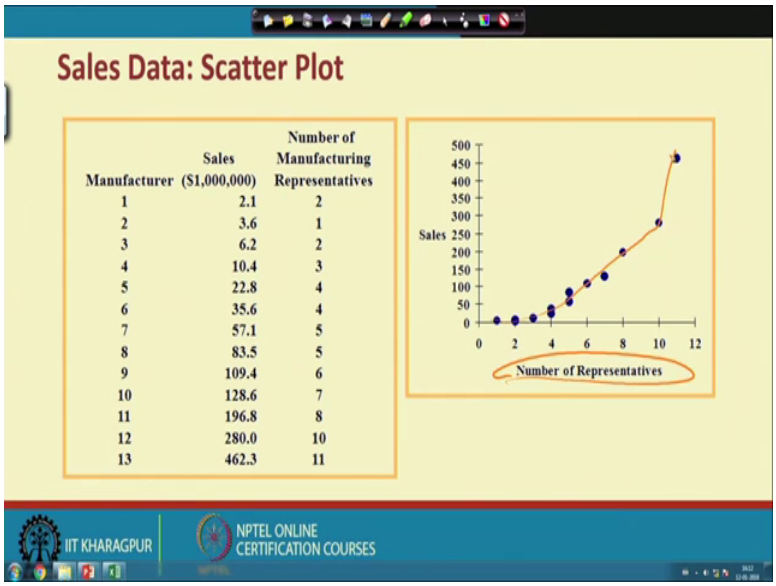
So, again having actually X 1 and X 2; so, here the model can be X 1 square; that is called as dot product and this is again dot product with respect to X 2 and this is what actually called as a cross product. So; that means, actually; so, we have X 1 and X 2; so, X 1 with degree 1, X 1 with degree 2; then X 2 with degree 1 X 2 with degree 2 and then the interactive effect. So, X 1 into X 2; so all the kind of possibilities we are trying and then you are checking whether they can contribute something towards this particular prediction.

So; that means, these are all actually various different modes through which actually you have to actually go for the search to find out the best fit model through which you do the predictions and do the kind of for castings and then come with right management decisions. So, these are the various forms of non-linear regression model and in fact, there are many other non-linear regression models are there.

So, like you know dummy modelling case, you like probably it legit which we will discuss in the later stage. But this is one point particular structure through which non-linearity can be connected and then the prediction can be done that too help the kind of business process by giving the right management decision.

With this basic structures; so, we will move.

(Refer Slide Time: 14:56)



So, how is this particular problem; so, let us say this is example and here this we have again sales data and that too, we are connecting; we are going to predict the sales subject to number of market representatives.

So, the hypotheses or theoretical knowledge argument is that; higher the kind of means; more and more is democratic representatives, then more will be the kind of sales. Because they will go for kind of more and more promotions; so, as a result; so, there is a high chance that sales of this particular organization can increase.

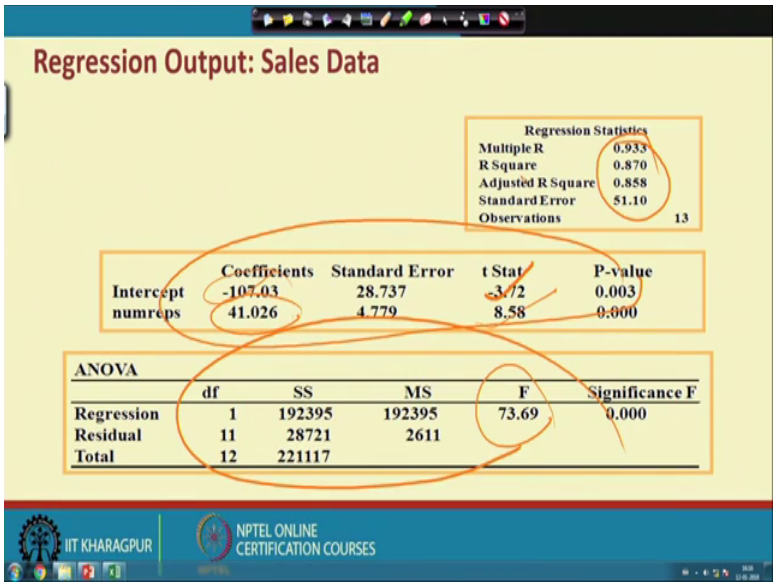
So, now knowing all these things; so, you have to check actually how these kind of connection. So, now whether these sales can be connected with the marketing representatives; in a kind of linear format or non-linear format. So, first what will you do? So, you go for data visualization. So, this is how the number of market representatives and then sales.

So, now you refine; so this will be showing if you plot or all these points and join all these points. So, it gives you a kind of linear structure; it creates a kind of non-linearity. So, as a result you have to find out the right choice and the right equation through which you can actually do the kind of predictions. For instance, actually sometimes as the software by default will show you the accurate functional; through which you will do the prediction.

The classic example is the origin software and where if you plot peak; means if you go for graphical visualization then it will be by default will show you the kind of the functional form. It will give you different kind of structure and they it will fix where the accuracy will be very high.

So; that means, typically the suggestion is that you have to go for data visualization process first and go for the right choice of the functionality; that means, functional relationship between dependent and independent variable. And then do the kind of estimations and get the estimated model; so that you will go for right predictions and then right management decision.

(Refer Slide Time: 17:19)



So, I having this kind of structure so, again with the help of this data so; that means, technically; so, we have actually this is your spreadsheet and this is what the data visualization. And then with the connection so you can actually estimate again you can go to the excel spreadsheet and estimate the particular output.

So, regression output will be coming in the first end like this. So, this is actually ANOVA and followed by; so, the kind of model outputs that is actually intercept and the market representative that is X 1 coefficients beta 1; so, it is coming actually positive.

So; that means, as per our theoretical expectation; so, the market representatives, the number really go up then by default. So, the kind of sales of that particular in

organization will go up so; that means, there is a positive association between these two and a sample data also validating the kind of theoretical expectations. So, now after knowing the theoretical knowledge, going through the data visualization process, fixing up the kind of right functional forms.

So once you estimate the particular process; then you will get the estimated output and before you go for the prediction again. So, you have to check that the model outputs are very accurate or very reliable so that the prediction should not be any defective and management decisions cannot be again and defective.

So, as a result; so you will go through the particular process and unit check. So, having the kind of link; so, in this model; so, your intercept is statistically significant which is not actually a highly irrelevant, but what is the relevant is the variable impact; which is coming statistically significant positive and statistical significant; that means, we are in the right track to do these predictions.

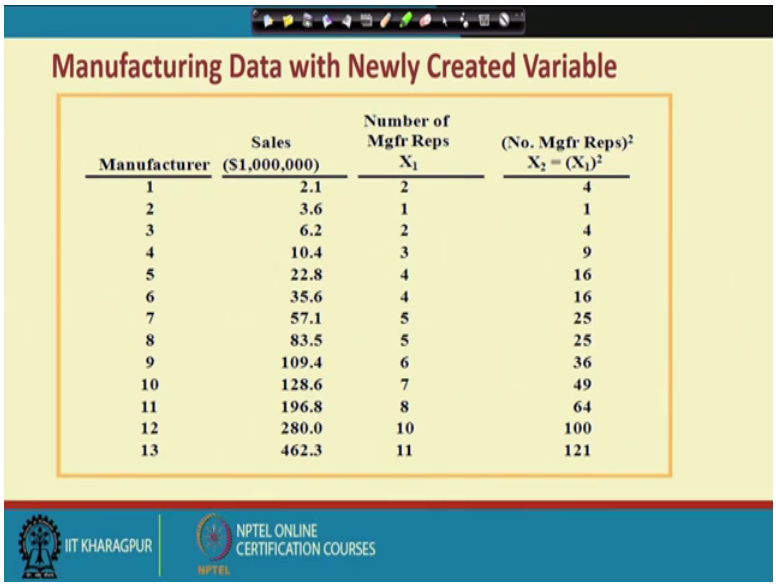
And the F statistic is also coming statistically significant and since F is coming high, by default we are expecting that R square and adjusted R square will be very high and that is also reflected here. So, adjusted R square also coming actually fair enough to predict the kind of environment.

So; that means, typically the overall fitness of the models reflected by R square adjusted R square and F R at the right side means in positive sides. And then the intercept or the parameters corresponding to the independent variable are also statistically significant and they are also giving the positive signal.

So; that means, since the overall structure is ok; so, you can go the kind of prediction. Sometimes there are two issues having wrong choice; so, model may be perfect, model may not be perfect, but we know what is actually best; that you go through all these screening and then you will come with right kind of modelling so, that the prediction for the future requirements should not be issue or it should not be any kind of defective.

Otherwise it may be a issue and your management decision can now be reflected negatively. So, with this basic prediction; so you can go for structuring, so the same structure here.

(Refer Slide Time: 20:41)



Manufacturer	Sales ($1,000,000)	Number of Mgfr Reps X_1	(No. Mgfr Reps)^2 $X_2 = (X_1)^2$
1	2.1	2	4
2	3.6	1	1
3	6.2	2	4
4	10.4	3	9
5	22.8	4	16
6	35.6	4	16
7	57.1	5	25
8	83.5	5	25
9	109.4	6	36
10	128.6	7	49
11	196.8	8	64
12	280.0	10	100
13	462.3	11	121

So, now what will you do? The same problem; so, this is actually linear estimates and where we are just putting; Y; this is sales and then X is nothing, but marketing representative. And with having say linear relationship, then we are getting perfect kind of situation through which you can predict these sales with respect to marketing representative.

And then when will it change the functional form; so, means we like to actually what we will called as in the kind of predictive analytics, the validation to justify that which model is perfectly ok; so, that is called as the structure called as a diagnostic check or sometimes we call as the vastness check. So, it will give you some kind of accuracy so that you can actually predict the kind of business requirement and then go for their right management decision. The thing is that; if any error is there in between, so this may actually reflect your predictions.

So, you must be very careful how you will do the kind of processing so; that means, at a particular point of time you must have a first hand output and that is the kind of route through which you have to go through other checks and other alternatives. It is actually all together; it is a decision making process and the whole idea about the decision making process is not with respect to only the kind of problem formulation.

So, it is with respect to also model estimation also; so you may get plenty of estimated models; that means, we are creating a kind of various alternatives for a particular

business prediction. And out of several alternatives, we have to pick up which one is the best for the particular requirement the business requirement. And once you pick up the right one then; obviously, the prediction will be the right and then the management decision will be also perfect.

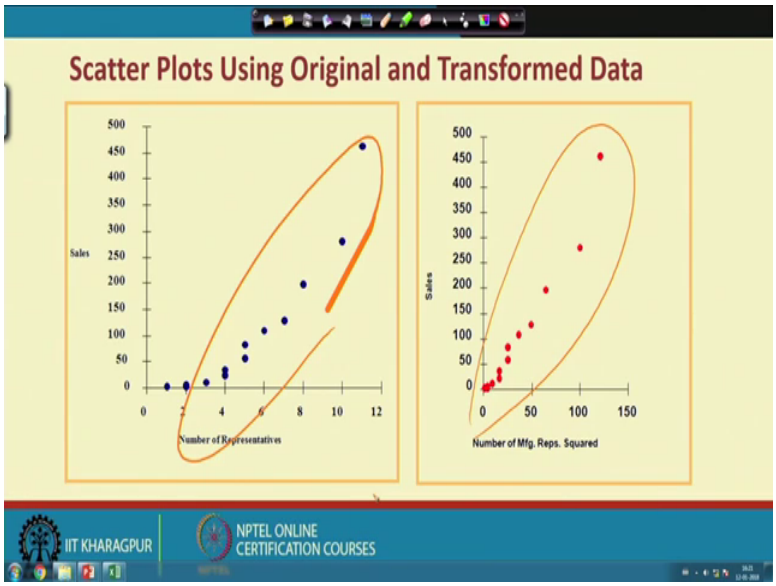
So, that is how we are actually in the process of checking that is what we called as going through robustness check or the kind of diagnostic check through, which it will give you the kind of justification or the kind of consistency that you are in a right track and the your estimated model is the estimated model is perfectly ok for the kind of prediction or the kind of business requirement.

So, now; so far just we are changing the functional form; initially we have a linear relationship, now putting non-linear relationship and that too we are trying to explore this relationship in a quadratic format; because your data visualization showing some kind of non-linearity here. And that is how we are trying, but we usually if you do not actually go for graphical plotting; so, you may not actually know the kind of functional form.

So, having actually big set of data, more number of variables; usually it is not so, easy actually. For two variables and having small data points or some considerable data points you can actually plot and get to know the particular structure. Otherwise you what happens; even if two variables with the huge data. So, you will find the non-linearity can be actually up ups and down. So, in that case actually sometimes it is very complex to put a kind of right structure through which you do the estimation.

But anyway; so, what will you do actually? We need to check through graphical visualization, then some kind of quantitative exploration that is what we are doing the kind of diagnostic check or robustness check. Because if your functional form is wrong or there is something or other obstacles or the other issues; then by default the model outcome will not be perfect to go for the prediction and the kind of inner cross checks right. So, that is why we will find the kind of structure through which you have to see which one is the best fit model and then go for the kind of predictions.

(Refer Slide Time: 24:51)

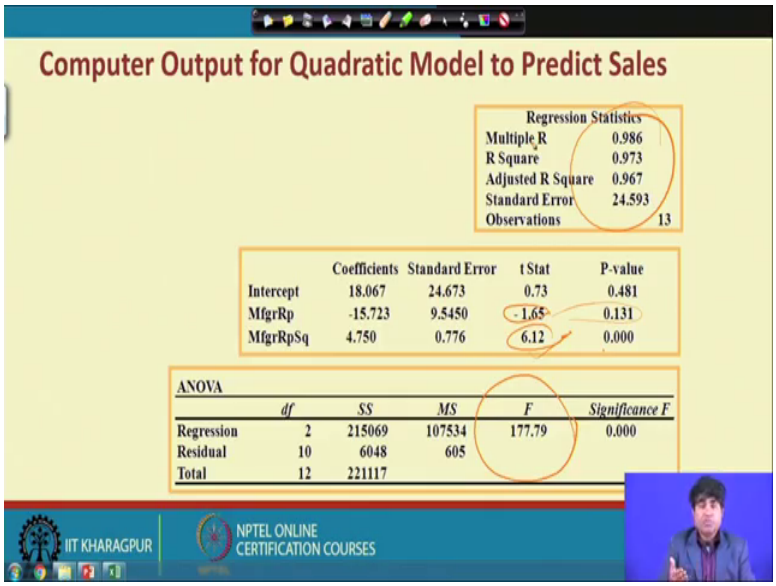


So, this is actually kind of what will you do? Sometimes, we can actually if you are finding some kind of non-linearity, so you can go for data transformations; then that will be transferred into some non-linearity, linearity. Then you will go for the kind of semi linear modeling; so, sometimes in the data there may be outlier. Outlier means a particular data point which is a highly distance from other data point.

So, again through data transformations; so, we can normalize the particular process. Because regression modelling require consistent in our data requirement so that we can do the perfect kind of structure, through which you do the prediction. So, this is the actual kind of plotting and this is actually go for transport data. So, you can go for various transformation, log transformation, exponential transformation, inverse transformation, first difference transformation; so, many transformations are there.

You can standardize the particular variables; then accordingly you can do the kind of output through which you can do the predictions.

(Refer Slide Time: 26:06)

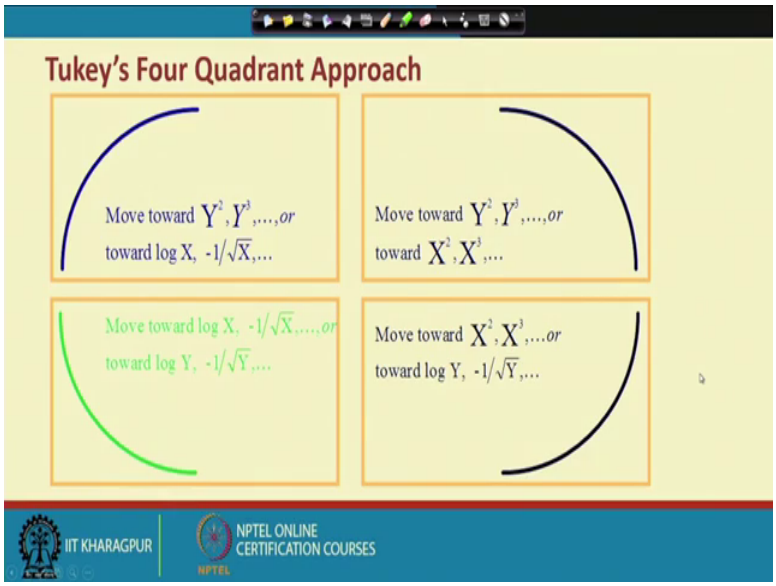


So, corresponding to this structure; so we can now estimate the process. Again after the transformation or the kind of connecting within non-linearity the quadratic form; so we are getting some kind of different output here. So, in this case this is actually X 1 output and this is X 1 square output; where this is coming actually positive and statistically significant, but this is coming negative and not coming statistically significant.

And here; in fact; F is very high and your goodness of fit is also very high; compared to the previous situation where we are integrating simply in a linear format. So; that means, sometimes if you pick up the right functional form, then the model validation will be perfect so that we must go for data visualization first before we use the predictive analytics, for any kind of prediction and any kind of management decision.

So, now; so having this output; so, this is actually fair enough to go for the predictions.

(Refer Slide Time: 27:14)

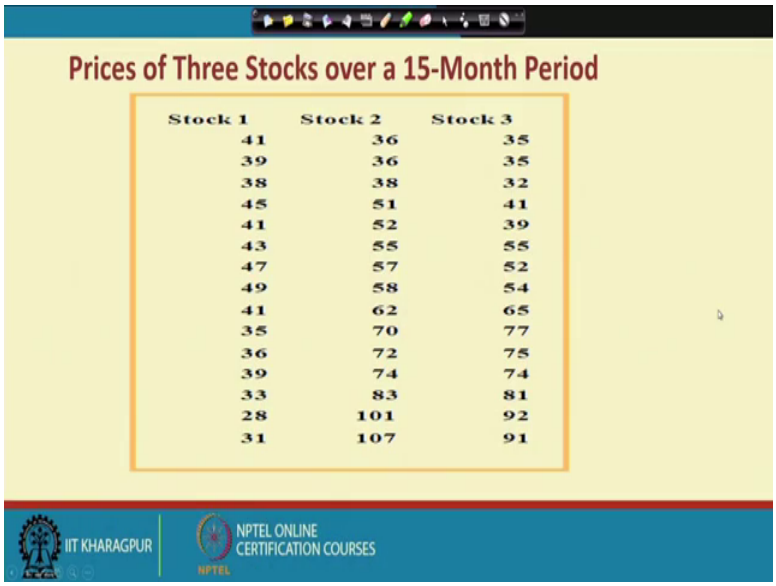


And so now; so, means there are many transformation which you can do. So, which I have already highlighted some extent, so you can go for log transformations, exponent transformation, standardized transformation; so these are kind of various formats through which you will go for the data transformation.

So, actually it is called as a data structuring; so it is not that you have to (Refer Time: 27:37) the data and putting a kind of proper structure, but after having the data, putting in the kind of a spreadsheet so, you need lots of standardizations as per the model requirement or the kind of business requirement. Because predictive analytics tools in some kind of techniques; so, there is a need of some accuracy of the data to predict the kind or to estimate the process and to go for the kind of prediction.

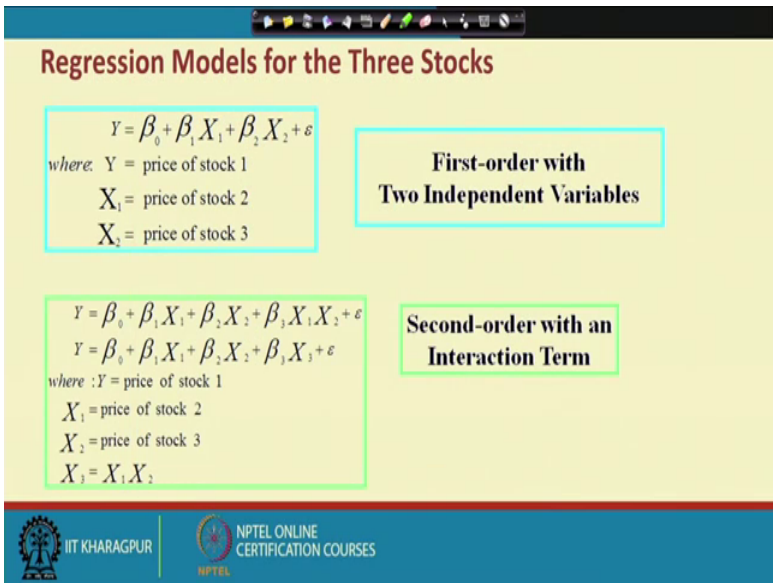
So data transformation is a kind of structure, we have to use frequently so that you can adjust the kind of error and go for the best bit of the models and then go for the best prediction in a requirement as per the business requirement.

(Refer Slide Time: 28:18)



So, this is another kind of model and here; so three variables and stock 1, stock 2, stock 3 and then we can actually study the interactive effect.

(Refer Slide Time: 28:30)



So, this is how the kind of structure; so, generally when there are three stocks. So, particular stocks behavior depends upon other stocks behavior and the intersection between these two or the kind of interactive kind of structure between other two kind of stocks. So, as a result; so this can be the kind of model non-linear model through which you can predict a price of stock 1; with respect to price of stock 2 and stock 3. So,

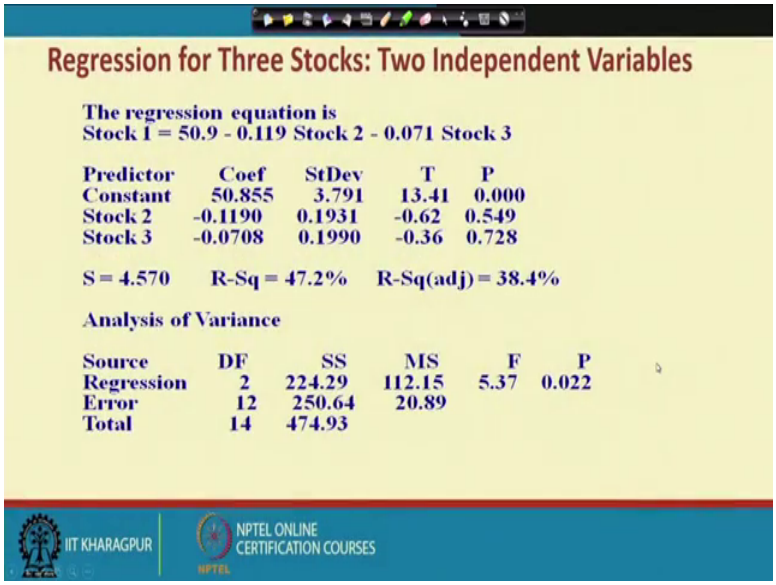
typically; so, Y is the kind of stock 1 structure and then X 1 is the stock 2, stock 3 and this is actually stock against stock 3.

So, this is what actually interactive effect; that is with respect to X 1 and X 2. So; that means, actually the kind of interactive, the kind of functional form; it exclusively depends upon the data visualization and some kind of theoretical linkage.

In this particular problem, if you have no financial understanding then you may not in a position to integrate all these things. So, the kind of restructuring of the data and the kind of transformations and the kind of functional form; exclusively depends upon the kind of theoretical knowledge and the kind of data visualization. So, once you are clear about all these things then you may in a right position to pick up the right model and then go for the right estimations. And; obviously, you can get the right estimated model for the kind of prediction.

So, now knowing this particular structure; you can actually move further and we can check; how is this particular structure with respect to this data?

(Refer Slide Time: 30:10)

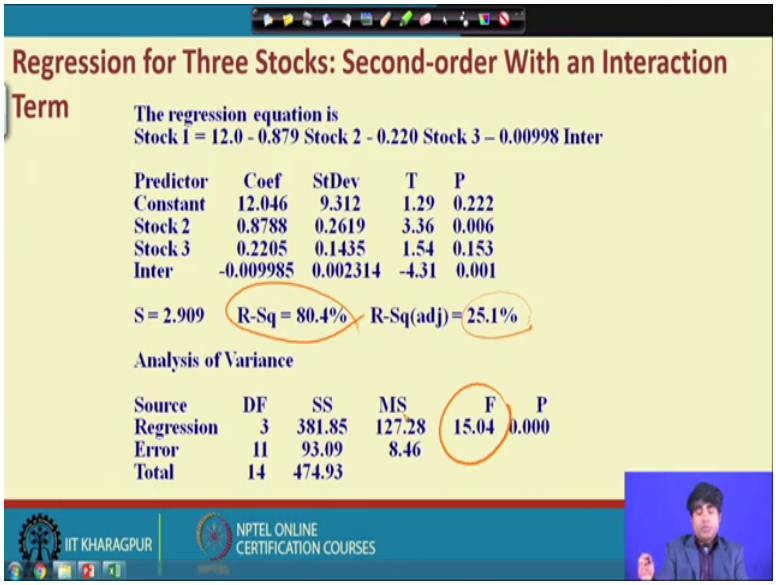


So, this is how the entrepreneur structure; this simple one and here actually we are dealing with predicting stock 1; with respect to stock 2 and stock 3 and this is the first end output and again changing the find up functional form; so, you will get second end output.

In this case, your F is coming very low, but it is statistically coming significant and R square is not very poor, but fairly, but still we need to see the whether there can be further improvement about the R square increase or the model validation increase. In fact, actually the T statistic is not actually statistically significant here. So, that is what actually the kind of signal negative signal; so, this is what the negative signal.

So, that is why this model cannot be used for predictions and the kind of future many requirement; so, again, so what will you do? You will look for alternatives.

(Refer Slide Time: 31:12)



So, the alternative structure is like this; in this case we are allowing interactive effect and that is actually with respect to the previous modelling framework. So, here we are allowing interactivity between X 1 and X 2 and that will be now treated as a third variable and as a result; so, the kind of an outcome is now more accurate to validate this.

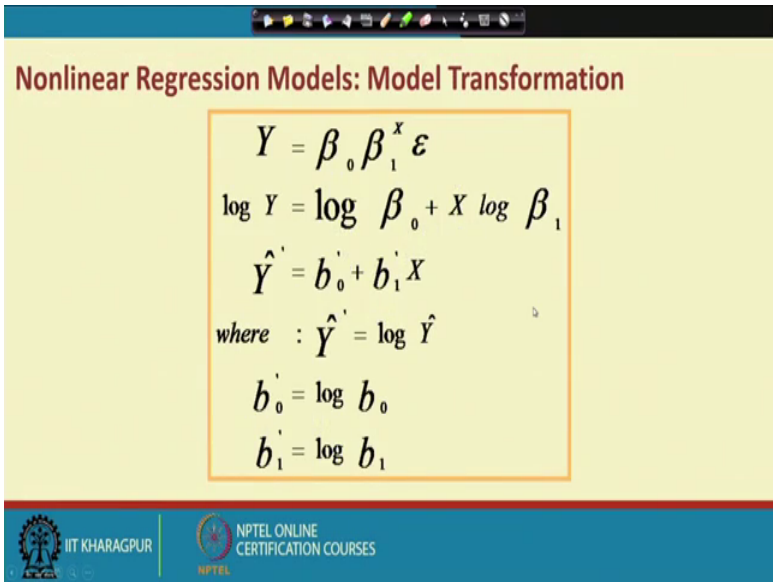
So, here in this case; so, after the right functional form, so F is improving and in fact, little bit R square is also improving; little bit question about registered R square, but it is still fairly to go for the prediction. So, the most important thing is that now one variable is coming statistically significant and also interactive effect is also coming significant.

So; that means, we try to find out all such possibilities or explorations then if actually the interactive effect is in order there by default; the statistical output will give you some kind of insignificant result. And later stage again you can throw this particular variables

and again look for the best alternative. So; that means, it is a continuous search process till you get the best fit models as per the business requirement.

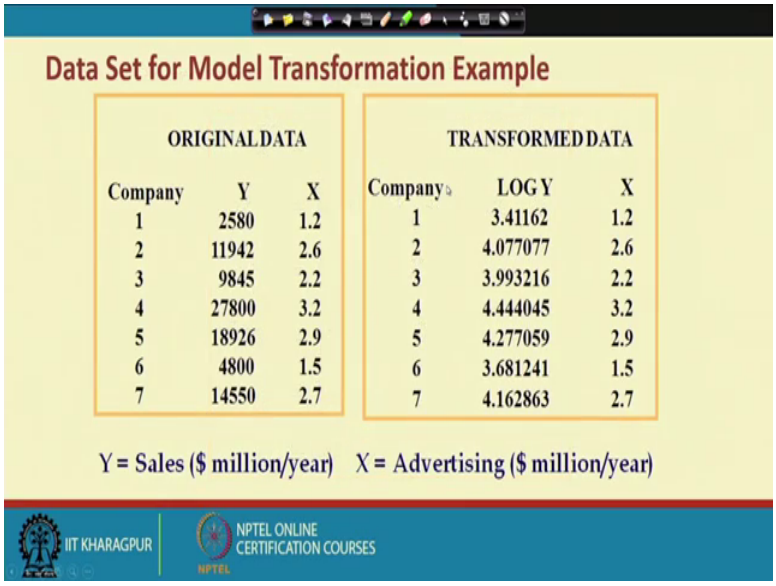
And once you find out the right choice or right model, then your management predictions or your management decision; will be very accurate and that will be actually having more accuracy with more kind of needful. So, with this actually typical structures; so, we will see the kind of structure; so, these are all various different functional part.

(Refer Slide Time: 32:50)



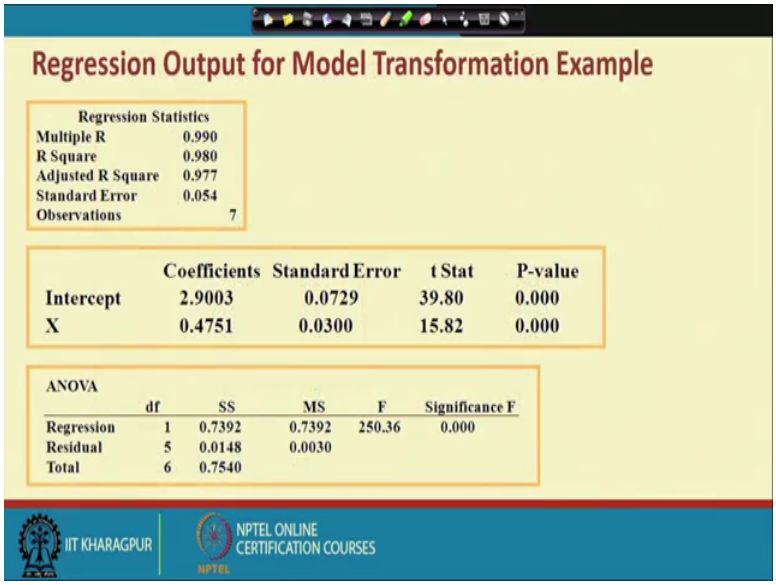
So, in this case here log transformation is allowed to do the kind of structuring.

(Refer Slide Time: 32:59)



And the same problems can be investigated; so, this is how the model and going for transformation the way you will do the model transformation like this. So, here you will find there is a kind of transformation, the same way variables can be adjusted. So, in the later stage; we are doing the adjustment and then we check how is the kind of estimation; this is the original data and this is what the transport data.

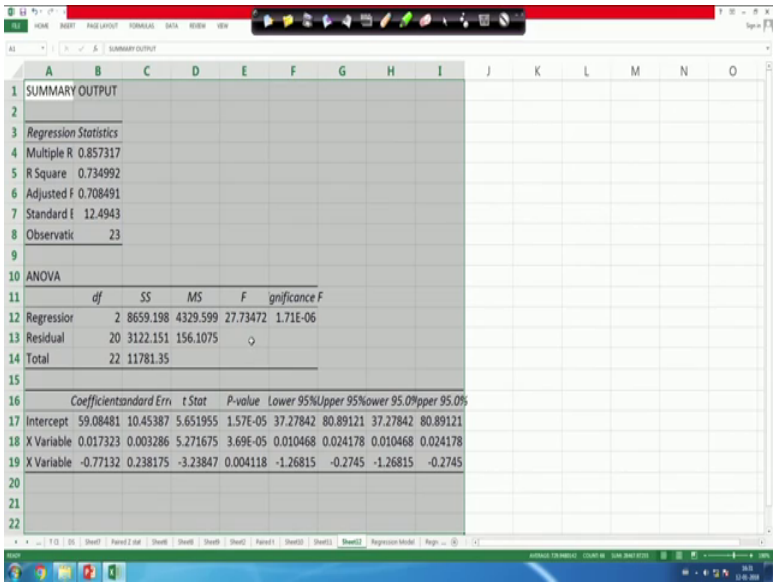
(Refer Slide Time: 33:25)



So, now again; so, you will check the kind of you output with the transport data. So, it is coming fairly actually very to get the kind of estimations; so, this is what the kind of structure and accordingly you can go for various problems through which you can do the transformation. By the way, some of the software directly connect with a functional format give you the output. And if not then in the excel spreadsheet, you can do some kind of manual adjustment and then do the kind of structuring.

So, let us see here is the kind of examples; I will give you here is the kind of examples; so that you can understand. So, what will you do here? So, I will take this problem and what we have already discussed?

(Refer Slide Time: 34:18)

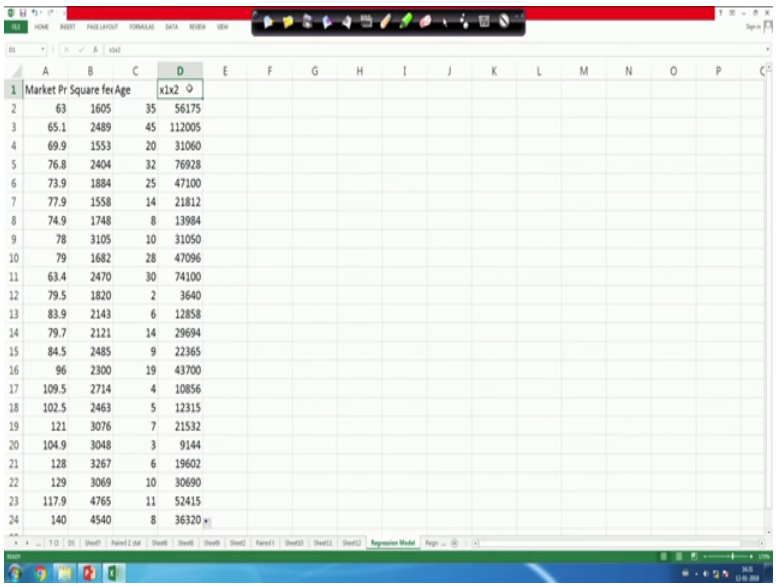


	A	B	C	D	E	F
1	SUMMARY OUTPUT					
2						
3	*Regression Statistics*					
4	Multiple R	0.857317				
5	R Square	0.734992				
6	Adjusted F	0.708491				
7	Standard E	12.4943				
8	Observatic	23				
9						
10	ANOVA					
11		*df*	*SS*	*MS*	*F*	*gnificance F*
12	Regressior	2	8659.198	4329.599	27.73472	1.71E-06
13	Residual	20	3122.151	156.1075		
14	Total	22	11781.35			
15						
16		*Coefficients*	*andard Erri*	*t Stat*	*P-value*	*Lower 95%*
17	Intercept	59.08481	10.45387	5.651955	1.57E-05	37.27842
18	X Variable	0.017323	0.003286	5.271675	3.69E-05	0.010468
19	X Variable	-0.77132	0.238175	-3.23847	0.004118	-1.26815
20						
21						
22						

	G	H	I	J	K	L	M	N	O
1									
2									
3									
4									
5									
6									
7									
8									
9									
10									
11									
12									
13									
14									
15									
16	*Upper 95%*	*ower 95.0%*	*pper 95.0%*						
17	80.89121	37.27842	80.89121						
18	0.024178	0.010468	0.024178						
19	-0.2745	-1.26815	-0.2745						
20									
21									
22									

So, this is the problem which we have discussed earlier and again you go to the data analysis package.

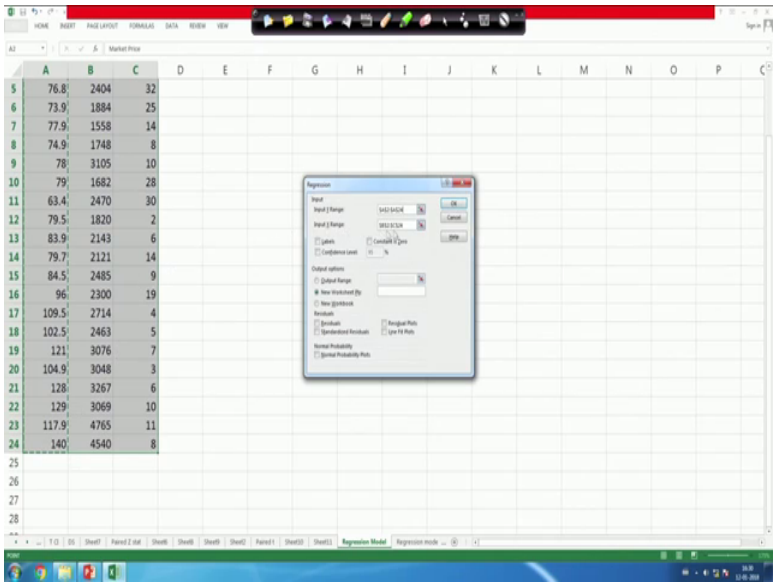
(Refer Slide Time: 34:24)



	A	B	C	D
1	Market Pr	Square fee	Age	x1x2
2	63	1605	35	56175
3	65.1	2489	45	112005
4	69.9	1553	20	31060
5	76.8	2404	32	76928
6	73.9	1884	25	47100
7	77.9	1558	14	21812
8	74.9	1748	8	13984
9	78	3105	10	31050
10	79	1682	28	47096
11	63.4	2470	30	74100
12	79.5	1820	2	3640
13	83.9	2143	6	12858
14	79.7	2121	14	29694
15	84.5	2485	9	22365
16	96	2300	19	43700
17	109.5	2714	4	10856
18	102.5	2463	5	12315
19	121	3076	7	21532
20	104.9	3048	3	9144
21	128	3267	6	19602
22	129	3069	10	30690
23	117.9	4765	11	52415
24	140	4540	8	36320

And for dada analysis; choose the regressions and ok here you can actually replace the particular structures and you can choose the right variable.

(Refer Slide Time: 34:42)



And then again, you can actually do the kind of linda dot structures; then again you reflect with the two variables and this is what actually some; so, now what will you do? So, we allow all the variables to the sita system and since uniform structure is there, just you put the ok. So, this is what actually the model output with respect to one dependent variable with the two independent variables. So, now after doing the kind of adjustment; so, you can check the various functional form and do the adjustment.

For instance, you go to the data view; so, this is what data view. So, you need actually let us say this is Y and this is X 1 and X 2; so, you need actually X 1 and X 2. So, let us say this is X 1 and X 2; so, what will you do? So, you just create another data set, allowing these two interaction; so, with respect to this one and this one.

So, then you create this particular structure; so now, you allow this particular in a series. So, this is now interactive effect, so now what will you do? So, this is again it is treated as third variable; so now, you can actually connect Y with the X 1, X 2 and X 1 and X 2. So, like that so; that means, what will you do? Any kind of structuring, you can do in the excel sheet and then you allow the kind of software to give you the final output.

So, what will you do? You structure the data, then you feed the models and then allow the software to give you the output. And then check these outputs; whether it is fairly ok to go for the predictions by checking the significance of the parameter, the goodness of

fit of the model and the kind of validation of the model. If all are fairly ok and then you can go for the predictions and then that will give you write management decision.

In fact, even if the goodness fit of the model is perfectly and something like all parameters are statistically significant, so there is no harm that to check the basic alternatives. So, checking the basic alternatives; so always give you some kind of strength or confidence or the kind of consistency to use the model for the kind of prediction and the kind of management decision.

So; that means, actually if this particular technique and the predictive analytics techniques is such a kind of beautiful structure, through which; by default you will find plenty of alternatives. And out of all these alternatives, you have to find out the right ones. So; that means, it is kind of question of searching for opportunity cost.

So, next best alternative; the Honda is here next best alternatives. So, every time you will actually check; whether there is any kind of model improvement. So, there is corresponding to the original estimated output, so you change this particular structure; even if you can increase the sample size, decrease the sample size; add one variable, remove one variables, add some kind of; I mean change the functional form, change the kind of structuring.

Then every time you will get another estimated output and again you have to compare with the original estimated output with the kind of change output. If the original output is fair enough or good enough to the change situation, then you can fix the original one; you reject this one. Or the change situation is more effective, more efficient, more accurate then you can pick up that ones and you can just remove this one.

So, this is how the continuous search process till you get the right models for the right prediction. And as a results, you can go for some kind of better management decision so; that means, technically in this lecture; we have addressed some problems relating to regression modeling. Because it gives some kind of exposure that; it is very interesting to find out some alternatives, that it is not one kind of utput through which you will do the prediction, which may be very high risk.

So, now you have series of alternatives by different structuring, restructuring. And then with various alternatives, you are picking up a right one; as a result so you may have a

very less risks for some kind of business predictions and then the kind of management decision. And likewise you will have a several kind of diagnostic and robustness checks; to find out the best fit or best alternatives and that we will discuss in the next lectures; with this we will stop here.

And thank you very much; have a nice time.