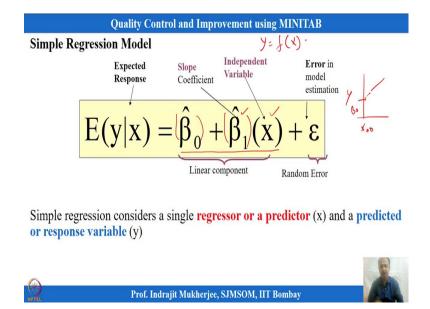
Quality Control and Improvement with MINITAB Prof. Indrajit Mukherjee Shailesh J. Mehta School of Management Indian Institute of Technology, Bombay

Lecture - 24 Linear Regressions

Hello everyone and welcome to our session 24 on Quality Control and Improvement with MINITAB. So, I am Professor Indrajit Mukherjee from Shailesh J. Mehta School of Management IIT Bombay. So, in previous session what we are doing is that we are discussing on Basics of Regression and which is an important tool to identify variables which can be considered for further experimentation.

(Refer Slide Time: 00:46)



And so the basic models that regression uses is proposed by Gauss and what we have and this is the basic fundamental model that we are using over here to understand relationship between y and x. And here we are modeling expected value of y with respect to x over here. So, there are 2 coefficients that is similar in line with what we know about basic line equations that is y = mx + c, c is known as intercept of the models and m is known as slope of the models like that ok.

So, this is a simple linear regression. So, I am assuming one single x I am having and E(y|x) is the expectation of y for a given x this is the conditional values that is mean

values that we can expect ok. So, at different conditions of x what we will have? We have different values of y and the condition if I can change the condition and again reset that one to that same condition even the output will be defined.

So, that was observed in when we are talking about analysis of variance and so expected value of mean value is basically modeled over here with respect to x. So, this is the linear component that is y = mx + c, this is the component where m is we can think of a slope over here.

So, for 1 unity 1 unit change in x what is the change in y expected value of y that is basically β_1 over here and if you just extrapolate this regression equation if I have developed something y is β_0 and if you can extend the line what is the intercept? That is this is the β_0 intercept over here. So, this is this will be the β_0 component over here, the value of y expected value of y when x is equals to 0 basically.

So, this is the but generally in regression we do not; we do not extrapolate and we do not extend that one, but this is the intercept concept that we have over here. So, physical interpretation is not possible for β_0 , but β_1 has a physical interpretation like that for every unity increase in x. What is the expected value change in y basically change in y? So, that is the interpretation.

So, β_0 and β_1 are the two important parameters which needs to be estimated from this model. So, if I can get the values then I can write the function over here and if I have the value of β_0 and β_1 .

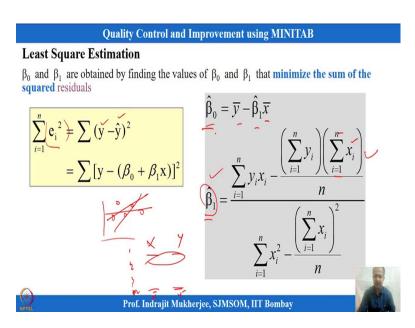
(Refer Slide Time: 02:57)



So, how β_0 and β_1 is estimated that is important for us and so estimation is important over here. So, estimation of unknown parameters that is β_0 and β_1 I need to estimate this one. So, I am writing $\hat{\beta}_0$ hat and $\hat{\beta}_1$ hat and whenever we have estimated that one then we have to do some model adequacy checks, that is also necessary like in nano work we have done.

Similar kind of tests are required in this regression analysis, I told that regression is an extension of analysis of variance basically. So, that is also important model adequacy checks like that ok.

(Refer Slide Time: 03:25)



So, this is the mathematical equation. So, I have a so what it will do is that many many lines can be constructed. So, it will plot the lines and innumerable line out of the innumerable lines. What it will do is that where the error is minimized basically, where the predicted value with the actual value so this will give you an error.

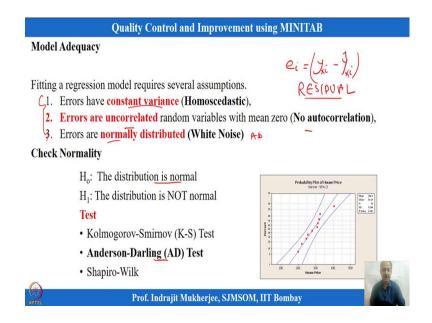
So, error minimization basically how can I minimize the error and that will give me the best field line, so there will be possibilities over here. So, I can place a line over here I can place a line over here like this. So, I can place innumerable lines like that and whichever gives me the minimum error over here that will be the best fit line like that and MINITAB does it automatically for you ok.

So, derivatives of this error square over here will give you the mathematical relationship of β_0 and β_1 and if we can equate it with 0 and that is the normal equation we consider and when we do that what we get is that β_0 estimation and β_1 estimation. So, I have a set of x condition I have set of y conditions. So, I have 1, 2 observations like this n number of observations, so every pair of observations that I get over here.

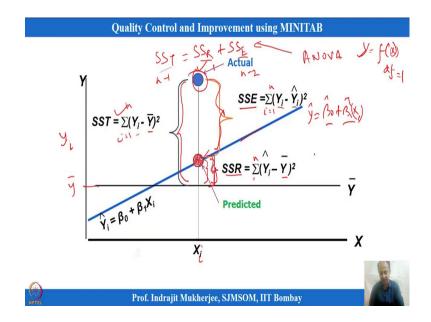
So, specifically I can calculate what is y average of that values and x average of the data set like that. So, this value will be used and β_1 estimation is given over here. So, this is a complex 1, but this is not difficult because I know what are the values of xi, i varies from 1 to n n number of observations that we are considering over here.

So, all these values can be calculated and β_1 can be calculated and when β_1 is estimated β_0 can also be estimated. Now MINITAB does it automatically this is based on normal equation solution of normal equations over here and this is known as least square estimation this is known as least square estimation and these are unbiased estimates basically these are unbiased estimate, this is statistician what they have suggested these are unbiased estimation.

(Refer Slide Time: 05:13)



(Refer Slide Time: 05:14)



And so we can adopt that one, so in this case what we will do is that this estimation MINITAB will automatically give it for you and then there is an ANOVA analysis concept over here in MINITAB also like earlier Anova analysis over here. So, this is let us say y is on this side and x is on this side xi values over here yi values on this direction over here single x and single y I am just representing one point over here.

So, one actual value is located over here and based on a line equation which I can assume for a given β_0 and β_1 . So, this can be one of the line equation over here β_0 plus β_1 for a given estimation let us say and we have developed some equation and this is the line equation. So, for a given value of x that is actually xi over here, what we can do is that we can also get a predicted value over here we can get a predicted value over because this is the line equation.

So, whenever I put xi in this equation I will get a predicted value which is nothing but what we are seeing over here. So, but the actual value is over here and this is the predicted value over here ok. So, this point from the overall average how much it differs.

So, this is yi variables over here. So, this will have some average which is the y average or over average although all the y values over here. So, what is the total deviation over here? This is the total deviation that we are seeing over here.

So, this is for one observation I am saying this is the total deviation from the mean values over here and out of this how much is explained by regression equation, this is known as this is the part that is explained by regression equation. But this part of the variability is unexplained by the regression equation over here. So, the total variability from the mean overall mean over here is known as SST this is known as SST which is represented over here.

And the part of explained variability that we are seeing over here is known as SS regression, basically SST is equals to SS regression over here and the remaining which is unexplained over here that variability which we are seeing is SSE over here. Now this is for one variable one observation over here there can be n number of observations. So, this summation equals to 1 to n over here summation equals to 1 to n over here.

So, like this summation 1 to n like that so SST plus SSR plus SSE over here. So, same concept of regression again SS sum of square regression sum of square error over here

and because we are predicting y is a function of x over here, x is a single variable and the degree of freedom for a single predictor will be equals to 1 basically, so that is considered is Anova analysis.

And if I have n number of observation SST degree of freedom will be n minus 1 and the regression degree of freedom will be 1 and error degree of freedom will be n minus 2 like that. So, and that is the interpretation n minus 1 minus 1 will be n minus 2 basically ok. So, the same concept is used over here only thing is that Xi is continuous variable and it can take any values not that specific values of predefined level 1 level 2 like that that is not the case here X is continuous Y is continuous also.

So, every values I can calculate what is the SS total over here, that is the deviation from the overall average that is taken over here considered this one and this is actual values over here and this is the predicted one that will give me SS error. And the regression will be predicted minus average from the average how much is explained basically what we are getting over here.

So, this gives an idea that analysis of variance can also be adopted over here and this is used for regression analysis for adequacy model adequacy checks like that ok. And that we will see in MINITAB the interpretation of that and here also model adequacy checks are required. So, in this case constancy of variance that we have homoscedastic has to be checked. Errors are uncorrelated that means, Durbin-Watson statistics is used and normal distribution assumptions we have Anderson darling test over here to prove that one ok.

Anderson darling test means the distribution is normal, so Anderson darling test we can adapt over here. We can store the residual and residual over here is nothing but so error over here.

So, error is known as residual and that is nothing but actual values minus predicted values like that for a given observation x i condition like that ok. So, for a given x i what is the prediction over here? So, this is the actual x i values over here, so this gives you the error which is known as residual which is known as a residual ok.

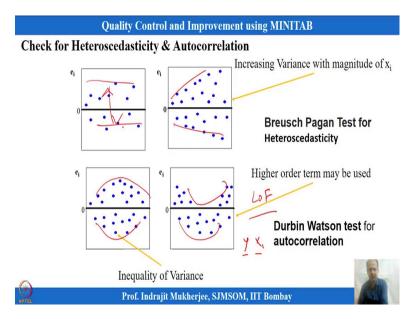
And MINITAB in MINITAB you can save the residuals and do all kind of analysis. So, sometimes residuals are stored sometimes standard residuals are stored. So, preferably we use standard standardized residual like that. So, that takes care of certain other

aspects like that. So, what we will do is that we will we will just analyze the residual and try to see model adequacy check.

So, this is important over here this is the assumptions has to be satisfied to use the regression model which can be, if model assumptions are satisfactory immediately we can say it can be generalized and irrespective. If you give me the within the range of x if you give me any values I can predict. Even if I do not have the observation in earlier in the historical data set any new observation within that boundary conditions of x I can predict what will be the expected value of y basically ok.

So, prediction is possible of y for a given value of x ok, only thing I have to understand that it should be within the boundaries where the regression model is developed. So, I cannot extrapolate basically regression equations. So, that consideration is always there ok.

(Refer Slide Time: 10:24)



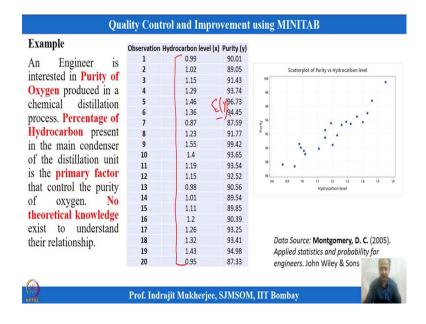
So, heteroscedasticity we have already discussed that if this is the funnel shape in that case heteroscedastic behavior, if this is the scenario then it says that linear model may not be sufficient you may have to incorporate like second order terms or something like that.

So, this is also lack of fit this will be reflected in lack of fit which is known as lack of fit integration. So, this if you have multiple observations like xi over here and yi

observations over here. And if you have x repetition observation for a given x there are at a given condition of x I have multiple observations like that, then only I can calculate this lack of fits like that.

So, that will give me nonlinearity whether it is non-linearity is there or not this is also non constancy of variance over here and this is more or less what is expected. So, there is no as such deviations that is happening, so that is residual versus fit we can plot that one and see that one ok.

(Refer Slide Time: 11:19)



So, if that is the scenario if this is the scenario again that conversion or transformation has to be used on y basically and then regress transform y with the x variables like that ok. So, here we are taking one example to illustrate regression in MINITAB. So, for this what we will do is that this is an example an engineer is interested this is taken from one of the mode is applied statistics and probability for engineers, this is the data set that that is given.

So, purity of data purity of oxygen over here is basically y variable and percentage of hydrocarbon is considered as one of the factor this is the historical data not based on experimentation statistical experimentation. But just historical data which says and I want to; I want to check whether the hydrocarbon levels, when I change whether it is expected value of purity is changing and whether I can develop a generalized equation prediction equation and within this range that is given over here.

Whether I can predict the expected value of y for a whether it is possible or not by developing a regression equation which can be generalized like that, if all model adequacy checks are ok then in that case we can do that ok.

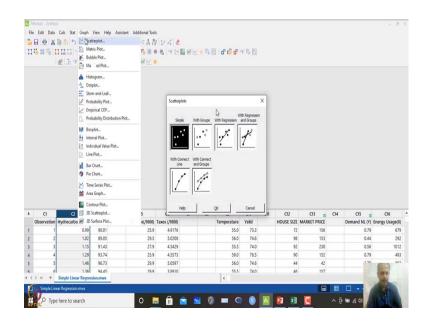
So and why we are doing this because why we are doing this because we do not have this we do not have this theoretical functions that can be used to model purity with hydrocarbon levels basically. And so that is one of the; one of the constraints that we have, that is why we are doing empirical modeling over here ok.

(Refer Slide Time: 12:57)

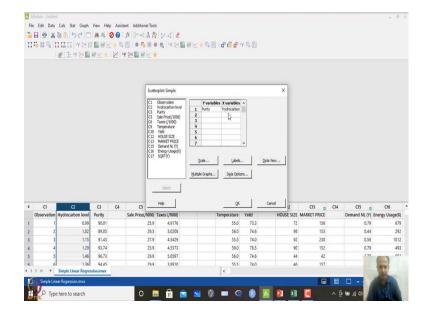
	Calc Stat Graph	istics	> EX Disp > X Stor > #₽ Gra	play Descriptiv re Descriptive ! phical Summa	ve Statistics Statistics	5	(★教習	c ⁿ c <mark>0</mark> d ⁿ	Y 鸟 器						- 6
	Control C Quality To Reliability Predictive Multivaria Time Seni Tables Nonparan Equivalen	iols /Survival : Analytics ate es metrics	小 1-S 売 2-S 小 Pain 量 2 Pr 量 1 Pr 量 2 Pr 副 1-S と 2 S 1 S 2 S 1 S 2 S 1 S 2 S 1 S 2 S 1 S 2 S 1 S 2 S 1 Pr 1 S 2 S 1 S 2 S 1 Pr 1 S 2 S 1 S 2 S 1 Pr 1 S 2 S 1 S 2 S 2 S 2 S 2 S 2 S 2 S 2 S 2				Open lew Project	ab ctri+0 ctri+shift+N							
			A Nor Out	variance rmality Test tlier Test odness-of-Fit 1	Test for Poisson	N	Worksheet	Ctrl+N							
e a	a	G	A Nor Out	rmality Test tlier Test	Test for Poisson C6	C7	Worksheet C8	Ctrl+N C9	C10	C11	C12	C13	C14	C15 g	C16
	Hydrocarbon level	Purity	λ Nor Out λ Good C4	rmality Test tiler Test odness-of-Fit 1 CS Price(/1000)	C6 Taxes (/1000)			C9 Temperature	Yeid	C11	HOUSE SIZE	MARKET PRICE		Demand NL (Y)	Energy Usage()
Observation 1	Hydrocarbon level 0.99	Purity 90.01	λ Nor Out λ Good C4	rmality Test tlier Test odness-of-Fit 1 CS Price(/1000) 25.9	C6 Taxes (/1000) 4.9176			C9 Temperature 55.0	Yeild 73.3	C11	HOUSE SIZE	MARKET PRICE		Demand NL (Y) 0.79	Energy Usage() 67
Observation 1	Hydrocarbon level 0.99 1.02	Purity 90.01 89.05	λ Nor Out λ Good C4	rmality Test tlier Test odness-of-Fit 1 CS Price(/1000) 25.9 29.5	C6 Taxes (/1000) 4.9176 5.0208			C9 Temperature 55.0 56.0	Yeild 73.3 74.6	C11	HOUSE SIZE 72 98	MARKET PRICE		Demand NL (Y) 0.79 0.44	Energy Usage() 67 29
	Hydrocarbon level 0.99 1.02 1.15	Purity 90.01 89.05 91.43	λ Nor Out λ Good C4	cs Price(/1000) 25.9 29.5 27.9	C6 Taxes (/1000) 4.9176 5.0208 4.5429			C9 Temperature 55.0 56.0 55.5	Yeild 73.3 74.6 74.0	C11	HOUSE SIZE 72 98 92	MARKET PRICE 156 153 230		Demand NL (Y) 0.79 0,44 0.56	Energy Usage() 67 29 101
Observation 1 2 3	Hydrocarbon level 0.99 1.02 1.15 1.29	Purity 90.01 89.05 91.43 93.74	λ Nor Out λ Good C4	mailty Test tiler Test CS Price(/1000) 25.9 29.5 27.9 25.9	C6 Taxes (/1000) 4.9176 5.0208 4.5429 4.5573			C9 Temperature 55.0 56.0 55.5 59.0	Yeild 73.3 74.6 74.0 78.5	C11	HOUSE SIZE 72 98 92 90	MARKET PRICE 156 153 230 152		Demand NL (Y) 0.79 0.44 0.56 0.79	Energy Usage() 67 29 101 49
Observation 1 2 3 4 5	Hydrocarbon level 0.99 1.02 1.15 1.29 1.46	Purity 90.01 89.05 91.43 93.74 96.73	λ Nor Out λ Good C4	rmality Test titer Test odness-of-Fit T CS Price(/1000) 25.9 29.5 25.9 25.9 25.9 25.9 25.9	C6 Taxes (/1000) 4.9176 5.0208 4.5429 4.5573 5.0597			C9 Temperature 55.0 56.0 55.5 59.0 56.0	Yeild 73.3 74.6 74.0 78.5 78.5 74.6	C11	HOUSE SIZE 72 98 92 90 44	MARKET PRICE 156 153 230 152 42		Demand NL (Y) 0.79 0,44 0.56	Energy Usage() 67 29 101
Observation 1 2 3 4 5 6	Hydrocarbon level 0.99 1.02 1.15 1.29 1.46 1.36	Purity 90.01 89.05 91.43 93.74 96.73 94.45	λ Nor Out λ Good C4	mailty Test tiler Test CS Price(/1000) 25.9 29.5 27.9 25.9	C6 Taxes (/1000) 4.9176 5.0208 4.5429 4.5573			C9 Temperature 55.0 55.5 59.0 56.0 55.5	Yeild 73.3 74.6 74.0 78.5	C11	HOUSE SIZE 72 98 92 90	MARKET PRICE 156 153 230 152		Demand NL (Y) 0.79 0.44 0.56 0.79	Energy Usage() 67 29 101 49
Оbservation 1 2 3 4 5 6 4 5 6 6 4 5 6 6	Hydrocarbon level 0.99 1.02 1.15 1.29 1.46	Purity 90.01 89.05 91.43 93.74 96.73 94.45	λ Nor Out λ Good C4	rmality Test titer Test odness-of-Fit T CS Price(/1000) 25.9 29.5 25.9 25.9 25.9 25.9 25.9	C6 Taxes (/1000) 4.9176 5.0208 4.5429 4.5573 5.0597			C9 Temperature 55.0 56.0 55.5 59.0 56.0	Yeild 73.3 74.6 74.0 78.5 78.5 74.6	C11	HOUSE SIZE 72 98 92 90 44	MARKET PRICE 156 153 230 152 42		Demand NL (Y) 0.79 0.44 0.56 0.79	Energy Usage() 67 29 101 49

So, how do we do that that is important for us? So, I will just go to that interface of MINITAB and this is the C1, C2 observation Ct observation that is given over here. I have just taken the data 1 is purity data 1 is hydrocarbon data over here and in this case what I have to do is that I have to go to stat and regression analysis. So, let us just try to see whether scatter plot what does caterpillar shows over here.

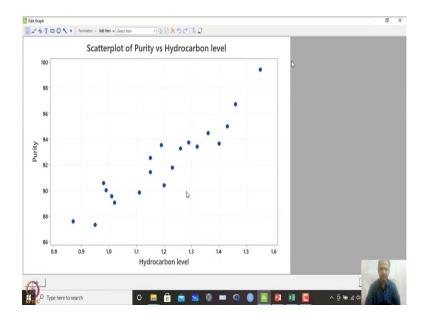
(Refer Slide Time: 13:19)



(Refer Slide Time: 13:23)



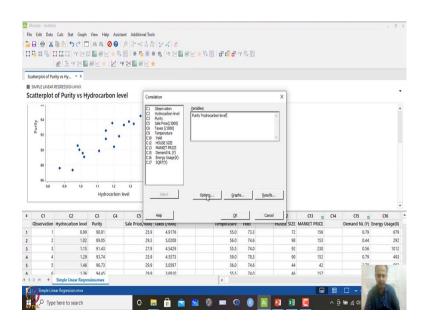
(Refer Slide Time: 13:30)



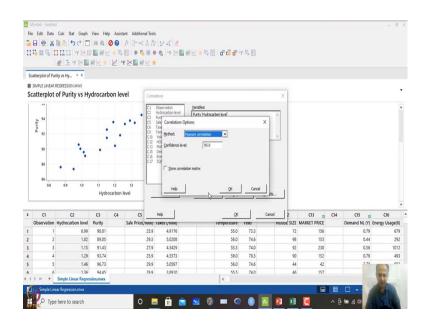
So, whether we can we see an see that linear relationship or not. So, you can go to graph and scatter plot. So, in this case what we can do is simple scatter plot over here and y variable is purity over here and x variable is hydrocarbon.

And if you click ok over here what will happen is that it will give you some this is the graph that we that we can see over here and it shows that there is a linear relationship that exists over here and also we can check the correlation. So, this is positive correlation what we can see is that hydrocarbon increases purity also increases.

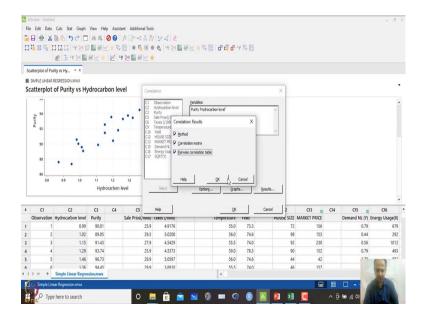
(Refer Slide Time: 13:50)



(Refer Slide Time: 13:56)

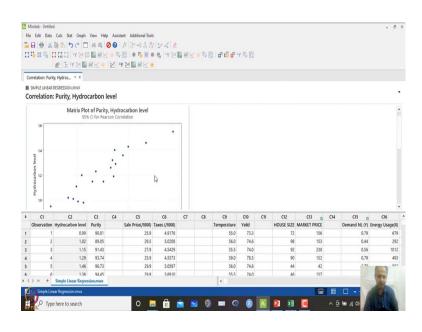


(Refer Slide Time: 14:02)



So, we can just check the correlation over here. So, basic statistics we can just check correlations and correlations I want to see purity with hydrocarbon levels over here and in options we can just do this and results we can just interpret this one and I click ok.

(Refer Slide Time: 14:04)



(Refer Slide Time: 14:07)

	Calc Stat Graph	View Hele	Arristant	Additional Tool											- (
	B € 5 € □														
							M. S								
·• · · · ·						BR	* 548	0 0 C 0	Y 49 83						
	1 ± ч ≥ ∎	BK*	12 VY	S B S K 1	r										
orrelation: Pur	ity, Hydroc × ×														
SIMPLE UNEAR	REGRESSION MWX														
orrelation	Purity, Hydro	carbon I	evel												
Correlation ty	pe Pearson														
	ws used 20														
o painaite F	learson correlation														
P. Postana a															
Correlation	10														
conciation															
	Purity														
Hydrocarbon	level 0.937														
Hydrocarbon	level 0.937														
	earson Correlation	IS			6										
Pairwise Pe Sample 1	earson Correlation Sample 2	N Correlat													
Pairwise Pe Sample 1	earson Correlation	N Correlat	on 95% C		2										
Pairwise Pe Sample 1	Earson Correlation Sample 2 1 level Purity 2	N Correlat			2										
Pairwise Po Sample 1 Hydrocarbon C1	earson Correlation Sample 2 I level Purity 2 C2	Correlat COCC	0.844 C4	(.0.975) 0.00 CS	C6	C7	C8	C9	C10	C11	C12	C13 g	C14	C15 g	
Pairwise Po Sample 1 Hydrocarbon C1	Sample 2 1 Ievel Purity 2 C2 Hydrocarbon level	Correlat C 0.5 C3 Purity	0.844 C4	C5 Sale Price(/1000)	C6 Taxes (/1000)	C7	C8	Temperature	Yeild	C11	HOUSE SIZE	MARKET PRICE	C14	Demand NL (Y)	Energy Usage
Pairwise Pe Sample 1 Hydrocarbon C1 Observatior	Sample 2 1 Ievel Purity 2 C2 Hydrocarbon Ievel 0.99	C3 Purity 90.01	0.844 C4	C5 Sale Price(/1000) 25.9	C6 Taxes (/1000) 4.9176	C7	C8	Temperature 55.0	Yeild 73.3	C11	HOUSE SIZE	MARKET PRICE 156	C14	Demand NL (Y) 0.79	Energy Usage
Pairwise Pe Sample 1 Hydrocarbon C1 Observation	C2 C	C3 Purity 90.01 89.05	0.844 C4	C5 C5 Sale Price(/1000) 25.9 29.5	C6 Taxes (/1000) 4.9176 5.0208	C7	C8	Temperature 55.0 56.0	Yeild 73.3 74.6	C11	HOUSE SIZE 72 98	MARKET PRICE 156 153	C14	Demand NL (Y) 0.79 0.44	Energy Usage
Pairwise Po Sample 1 Hydrocarbon C1 Observation	earson Correlation Sample 2 I level Purity 2 Hydrocarbon level 0.99 1.02 1.15	C3 C3 Purity 90.01 89.05 91.43	0.844 C4	C5 Sale Price(/1000) 25.9 29.5 27.9	C6 Taxes (/1000) 4.9176 5.0208 4.5429	C7	C8	Temperature 55.0 56.0 55.5	Yeild 73.3 74.6 74.0	C11	HOUSE SIZE 72 98 92	MARKET PRICE 156 153 230	C14	Demand NL (Y) 0.79 0.44 0.56	Energy Usage
Pairwise Pro Sample 1 Hydrocarbon C1 Observation	carson Correlation Sample 2 level Purity 2 Hydrocarbon level 0.99 1.02 1.15	C3 Purity 90.01 89.05 91.43 93.74	0.844 C4	C5 Sale Price(/1000) 25.9 29.5 27.9 25.9	C6 Taxes (/1000) 4.9176 5.0208 4.5429 4.5573	C7	C8	Temperature 55.0 56.0 55.5 59.0	Yeild 73.3 74.6 74.0 78.5	C11	HOUSE SIZE 72 98 92 90	MARKET PRICE 156 153 230 152	C14	Demand NL (Y) 0.79 0.44 0.56 0.79	Energy Usage
Pairwise Po Sample 1 Hydrocarbon C1 Observation	earson Correlation Sample 2 1 level Punty 2 C2 Hydrocarbon level 0.99 1.02 1.15 1.29 1.46	C3 Purity 90.01 89.05 91.43 93.74 96.73	0.844 C4	0.0975) 0.000 C5 Sale Price(/1000) 25.9 29.5 27.9 25.9 25.9 25.9	C6 Taxes (/1000) 4.9176 5.0208 4.5429 4.5573 5.0597	C7	CS	Temperature 55.0 56.0 55.5 59.0 56.0	Yeild 73.3 74.6 74.0 78.5 74.6	C11	HOUSE SIZE 72 98 92 90 44	MARKET PRICE 156 153 230 152 42	C14	Demand NL (Y) 0.79 0.44 0.56	Energy Usage
Pairwise Pr Sample 1 Hydrocarbon C1 Observation	Correlation Sample 2 I level Purity 2 C2 Purity 2 Hydrocarbon level 0.99 1.02 1.02 1.15 1.29 1.46 1.36 1.36 1.36 1.36	C3 Purity 90.01 89.05 91.43 93.74 96.73 94.45	0.844 C4	C5 Sale Price(/1000) 25.9 29.5 27.9 25.9	C6 Taxes (/1000) 4.9176 5.0208 4.5429 4.5573	C7	C8	Temperature 55.0 56.0 55.5 59.0 56.0 55.5	Yeild 73.3 74.6 74.0 78.5	C11	HOUSE SIZE 72 98 92 90	MARKET PRICE 156 153 230 152	C14	Demand NL (Y) 0.79 0.44 0.56 0.79	Energy Usage
Pairwise Pr Sample 1 Hydrocarbon C1 Observation	earson Correlation Sample 2 1 level Punty 2 C2 Hydrocarbon level 0.99 1.02 1.15 1.29 1.46	C3 Purity 90.01 89.05 91.43 93.74 96.73 94.45	0.844 C4	0.0975) 0.000 C5 Sale Price(/1000) 25.9 29.5 27.9 25.9 25.9 25.9	C6 Taxes (/1000) 4.9176 5.0208 4.5429 4.5573 5.0597	C7	C8	Temperature 55.0 56.0 55.5 59.0 56.0	Yeild 73.3 74.6 74.0 78.5 74.6	C11	HOUSE SIZE 72 98 92 90 44	MARKET PRICE 156 153 230 152 42 157	C14	Demand NL (Y) 0.79 0.44 0.56 0.79	Energy Usage
Pairwise Pr Sample 1 Hydrocarbon Observation 1 2 2 2 4 4 5 4 5 1 5 1 1 1 2 2 2 4 4 5 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1	Correlation Sample 2 I level Purity 2 C2 Purity 2 Hydrocarbon level 0.99 1.02 1.02 1.15 1.29 1.46 1.36 1.36 1.36 1.36	C3 Purity 90.01 89.05 91.43 93.74 96.73 94.45	0.844 C4	0.0975) 0.000 C5 Sale Price(/1000) 25.9 29.5 27.9 25.9 25.9 25.9	C6 Taxes (/1000) 4.9176 5.0208 4.5429 4.5573 5.0597	C7	C8	Temperature 55.0 56.0 55.5 59.0 56.0 55.5	Yeild 73.3 74.6 74.0 78.5 74.6	C11	HOUSE SIZE 72 98 92 90 44	MARKET PRICE 156 153 230 152 42	C14	Demand NL (Y) 0.79 0.44 0.56 0.79	Energy Usage 6 2 10 4
Pairwise Pi Sample 1 Hydrocarbon Observation 1 2 2 3 4 4 5 5 1 5 1 8 1 1 1 1 2 2 3 1 2 2 3 1 1 1 1 1 1 1 1 1	C2 Purity 2 Purity 2 Purity 2 Purity 2 Phydrocarbon level 0.99 1.02 1.15 1.29 1.46 Simple Linear Regree	C3 Purity 90.01 89.05 91.43 93.74 96.73 94.45	0.844 C4	0.0975) 0.000 C5 Sale Price(/1000) 25.9 29.5 27.9 25.9 25.9 25.9	C6 Taxes (/1000) 4.9176 5.0208 4.5429 4.5573 5.0597 3.8910	<i>с</i> 7		Temperature 55.0 56.0 55.5 59.0 56.0 55.5 4	Yeild 73.3 74.6 74.0 78.5 74.6 74.0	C11	HOUSE SIZE 72 98 92 90 44	MARKET PRICE 156 153 230 152 42 157		Demand NL (Y) 0.79 0.44 0.56 0.79 2.70	Energy Usage

And what we see is that correlation p-values that we are seeing over here is near to 0 and less than 0.05. So, that indicates that this and purity is basically highly correlated and correlation value what you can see is that 0.937 and that is more than 0.7 and that is significant over here 0.93 is very good ok.

So, a p-value indicates that there is a statistical significance over here and the p-value is quite p-value shows significance over here; that means, they are highly correlated and linearly correlated basically. So, Pearson correlation is used over here.

(Refer Slide Time: 14:42)

110		Regression	n	•	Fitted Line Plot Regression	N	• 2	Fit Regression		YS						
		DOF) K		sion 🗟		Best Subsets.								
	elation: Purity	,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,		,	Stability Study			Predict								
-	relation:	Reliability		, 🛛	Orthogonal Regr	ession		Factorial Plot								
		Predictive		• 2	Partial Least Squ	ires	8									
	umber of row			14	Binary Fitted Line	Plot	K									
P	o: pairwise Peu		5	1	Binary Logistic R		• *	Response Op	timizer							
_		Nonparan		, la												
Co	orrelations	Equivalen														
Hy	vdrocarbon le	Power and evel 0.937	d Sample Size		Poisson Regressi CI for p P-Valu		•									
Hy Pai Sai	vdrocarbon le nirwise Pea Imple 1 vdrocarbon le	Power ani rvel 0.937 arson Correlation Sample 2 M rvel Purity 2	d Sample Size S 4 Correlati 0 0.5	on 95% 137 (0.84	CI for p P-Valu 4, 0.975) 0.00	2	•									
Hy Pai Sar	irwise Pea imple 1 drocarbon le	Power an Power an arson Correlation Sample 2 M tvel Purity 2 C2	d Sample Size	on 95% 137 (0.84 C4	CI for p P-Valu 4, 0.975) 0.00 CS	6 C6	,	C8	C9	C10	C11	C12	CI3 g	C14	C15 2	
Hy Pai Sar	irwise Pea imple 1 drocarbon le	C2 Hydrocarbon level	S Correlati	on 95% 137 (0.84 C4	CI for p P-Valu 4, 0.975) 0.00 CS Sale Price(/1000)	C6 Taxes (/1000)	, c7	C8	Temperature	Yeild	C11	HOUSE SIZE	MARKET PRICE	C14	Demand NL (Y)	age
Hy Pai Sat	irwise Pea imple 1 drocarbon le	Power ani Power ani sample 2 t tevel Purity 2 C2 Hydrocarbon level 0.99	s Correlati Correlati Correlati Correlati Correlati Correlati Correlati Correlati Correlati Correlati	on 95% 137 (0.84 C4	CI for p P-Valu 4, 0.975) 0.00 CS Sale Price(/1000) 25.9	C6 Taxes (/1000) 4.9176	, C7	C8	Temperature 55.0	Yeild 73.3	C11	HOUSE SIZE	MARKET PRICE 156	C14	Demand NL (Y) 0.79	age
Hy Pai Sat	irwise Pea imple 1 drocarbon le	C2 Hydrocarbon level	S Correlati	on 95% 137 (0.84 C4	CI for p P-Valu 4, 0.975) 0.00 CS Sale Price(/1000)	C6 Taxes (/1000)	, C7	CS	Temperature	Yeild	C11	HOUSE SIZE	MARKET PRICE	C14	Demand NL (Y)	age (
Hy Pai Sat	vdrocarbon le invise Pea imple 1 vdrocarbon le C1 bservation 1 2	equivalence of the second seco	S C3 C3 C3 C3 Purity 90.01 89.05	on 95% 137 (0.84 C4	CI for p P-Valu 4, 0.975) 0.00 C5 Sale Price(/1000) 25.9 29.5	C6 Taxes (/1000) 4.9176 5.0208	, C7	C8	Temperature 55.0 56.0	Yeild 73.3 74.6	C11	HOUSE SIZE 72 98	MARKET PRICE 156 153	C14	Demand NL (Y) 0.79 0.44	age (2 10
Hy Pai Sai	vdrocarbon le invise Pea imple 1 vdrocarbon le C1 bservation 1 2	Cover ani vet 0.937 vet 0.937 vet 0.937 vet 0.937 vet 0.99 C2 Kydrocarbon kevel 0.99 1.02 1.15	4 Correlati 0 0.5 C3 Purity 90.01 89.05 91.43	on 95% 137 (0.84 C4	CI for p P-Valu 4, 0.975) 0.00 C5 Sale Price(/1000) 25.9 29.5 27.9	C6 Taxes (/1000) 4.9176 5.0208 4.5429	,	C8	Temperature 55.0 56.0 55.5	Yeild 73.3 74.6 74.0	Cti	HOUSE SIZE 72 98 92	MARKET PRICE 156 153 230	C14	Demand NL (Y) 0.79 0.44 0.56	

(Refer Slide Time: 14:48)

Å	Ainitab - Untitles	1															14	ē x
Fil	e Edit Data	Calc Stat Graph	View He	Ip Assist	tant Ad	ditional Tools												
•	BOX	0 500	144	00	fx 3	-2 L 20	14											
									+ 私田		松阳							
		€ EY>■									- V U.S							
7		y, Hydroc × ×											_					
					-	Regression						>	<					
_		REGRESSION MWX	carbon	loval		C1 Observa	tion	Responses:										٠
C		Purity, Hydro	carbon	level			bon level	Purity				2						
	Number of ros					C5 Sale Prio	e(/1000)											
						C6 Taxes (/ C9 Tempera		Continuous pr	adatase				-					
	p: pairwise Pe	arson correlation				C10 Yeld C12 HOUSE!	SIZE	Hydrocarbox				7	1					
		~				C13 MARKET C15 Demand												
	Correlation					C16 Energy C17 SQRT(Y)	(i)age(i)											
		Purity				en squitt						_	-					
	Hydrocarbon I	evel 0.937						Categorical pr	edictors:				-					
						I		ľ										
	Pairwise Pe	arson Correlation	15			I	- 1											
	Sample 1		N Correla															
	Hydrocarbon I	evel Purity 2	0 0	.937 (0.8	344, 0.973				Model	Options	Coding	Stepwise	1					
					_		-		3									٣
+	CI	C2	C3	C4		Select			Validation	Graphs	Besults	Storage		C13 🛛	C14	C15 👩		
	Observation	Hydrocarbon level	Purity		Sale P							1	E MAR	ET PRICE		Demand NL (Y)		
1	1	0.99	90.01			Help					<u>QK</u>	Cancel	1	156		0.79		679
2	2	1.02	89.05		1	29.5	5.020			56.0	74.6		98	153		0.44		292
3	3	1.15	91.43			27.9	4.542			55.5	74.0		92	230		0.56		012
4	4	1.29	93.74			25.9	4.557			59.0	78.5		90	152		0.79		493
5	5	1.46	96.73			29.9	5.059			56.0	74.6		44	42		2.70	(503
6	D H +	1.36 Simple Linear Regree	94.45			79.9	3,891	0		55.5	74.0		46	157			100	
			ssion.nwx			_	_	_	_	4	_	_	_			-	1 de la	5
1	Simple Line	ear Regression.mixe														 /	and the second	
ì	O Type	here to search				0	-	Ê 1	4 🚳			· 😰			^ ĝ	10 1 40		
NI	TEL					_	-		-				-			and the second se		

So, if this is so then I can develop the regression equation. So, let us develop the regression equation. So, what I do is that I go to stat I go to regression and I go to fit regression models over here, then it will ask what is the response? I will say purity is the response, what is the continuous predictor I have hydrocarbon there is no categorical predictor over here that is also possible to incorporate.

So, we will not go into that complexity now ok. So, if purity is the response and hydrocarbon is the only x variable that is continuous and then all these options there are many more options over here what you can see.

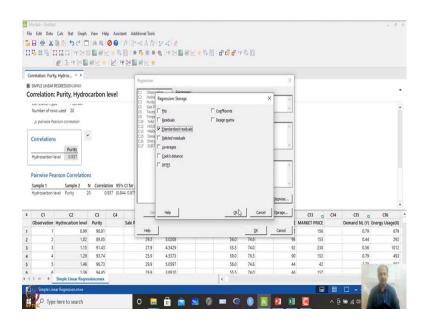
	1 III. I	I LI LI Y XO		< + 1h.	24 : 4	1 10	# # # · V >> ==		no no Ly J	6. 59							
		€ EY≥∎						- EX C)		× 0.4	×	1					
Correlati	ion: Purit	ty, Hydroc × ×					Predictors:	Add terms using select	ted predictors and m	odel terms:							
SIMPLI	E LINEAR I	REGRESSION MWX				Reg	Hydrocarbon level	Interactions through o	eder: 1		Add	×					
orrel	lation:	Purity, Hydro	carbon	level		C1				_							
-		-				382222		Terms through order:	1	•	Add	^					
Numb	ber of rov	ws used 20				00		Cross predictors and t	erms in the model		Add	~					
p:pi	ainise Pe	sarson correlation				C10	_										
2001						C12 C13						^					
Colle	elations	s				C15 C16				1							
Corre	elations	s Purity				C15 C16 C17	Terms in the model:		De	auit 🗙	4 +	~					
Hydro	ocarbon k	Purity level 0.937					Ierms in the model: Hydrocarbon level		<u>De</u>	aut X	<u>+ +</u>	> (
Hydro Pairv Samp	ocarbon le wise Peu ole 1	Purity evel 0.937 arson Correlation Sample 2	15 N Correla	tion 951 1937 (0.8				L9	<u>. 0</u>	aut X	<u>+</u> +	L					
Hydro Pairv Samp Hydro	ocarbon k wise Peu ole 1 ocarbon k	Purity ievel 0.937 arson Correlation Sample 2 ievel Purity 2	15 N Correla		44, 0.97			4	<u>عا</u>	at X	<u>+</u> +	< >	C13 12	CI4	CIS		C16
Hydro Pairv Samp Hydro	ocarbon k wise Peu ole 1 ocarbon k	Purity evel 0.937 arson Correlation Sample 2 ievel Purity 2 C2 Hydrocarbon level	15 N Correla 0 0 C3 Purity	.937 (0.8			Hydrocarbon levef		<u>عا</u>	iault X	4 +	< >	MARKET PRICE	CI4	Demand NL (Y	Energy	Usag
Hydro Pairv Samp Hydro	ocarbon k wise Peu ole 1 ocarbon k	Purity evel 0.937 arson Correlation Sample 2 ievel Purity 2 C2 Hydrocarbon level 0.99	15 N Correla 10 0 C3 Purity 90.01	.937 (0.8	44, 0.97				<u>.</u>	iault X	4 +	< >	MARKET PRICE 156	C14	Demand NL (Y 0.75	Energy	Usage
Hydro Pairv Samp Hydro	ocarbon k wise Peu ole 1 ocarbon k	Purity evel 0.937 arson Correlation Sample 2 evel Purity 2 C2 Hydrocarbon level 0.99 1.02	15 N Correla 10 0 C3 Purity 90.01 89.05	.937 (0.8	44, 0.97		Hydrocarbon levef		<u>D</u>	ault X	<u>+</u> +	< >	MARKET PRICE 156 153	C14	Demand NL (Y 0.79 0.44	Energy	Usag
Hydro Pairv Samp Hydro	ocarbon k wise Peu ole 1 ocarbon k	Purity level 0.937 arson Correlation Sample 2 level Purity 2 C2 Hydrocarbon level 0.99 1.02 1.15	15 N Correla 10 0 C3 Purity 90.01 89.05 91.43	.937 (0.8	44, 0.97		Hydrocarbon level ✓ Include the constant ter Help			auit X	Cancel		MARKET PRICE 156 153 230	C14	Demand NL (Y 0.79 0.44 0.56	Energy	Usag 1
Hydro Pairv Samp Hydro	ocarbon k wise Peu ole 1 ocarbon k	Purity level 0.937 arson Correlation Sample 2 evel Purity 2 Hydrocarbon level 0.99 1.02 1.15 1.29	15 N Correla 10 0 C3 Purity 90.01 89.05 91.43 93.74	.937 (0.8	44, 0.97		Hydrocarbon level Fundade the gonstant ter Heb		vin	OK .		~ ;; ; ; ; ; ; ; ; ; ; ; ; ; ; ; ; ; ;	MARKET PRICE 156 153 230 152	C14	Demand NL (Y 0.79 0.44 0.56	Energy	Usag
Hydro Pairv Samp Hydro	ocarbon k wise Peu ole 1 ocarbon k	Purity level 0.937 arson Correlation Sample 2 level Purity 2 C2 Hydrocarbon level 0.99 1.02 1.15	15 N Correla 10 0 C3 Purity 90.01 89.05 91.43	.937 (0.8	44, 0.97		Hydrocarbon level ✓ Include the constant ter Help			QK .			MARKET PRICE 156 153 230 152 42	C14	Demand NL (Y 0.79 0.44 0.56	Energy	Usag 1

(Refer Slide Time: 15:09)

So, only thing what we will consider include the constant term in the model. So, this should be clicked over here, because statistician says that when I include the intercept when I include the constant term which is the intercept over here, generally the model performance is quite good and this is seen by many research like that.

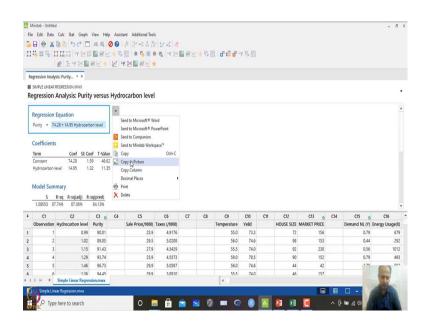
So, we will not omit this β_0 estimation over here we will keep that one. So, this is the only thing that you have to; you have to remember, these options coding stepwise that that is not required at this stage validation graphs results and only thing what you can do is that you can store the residuals.

(Refer Slide Time: 15:43)

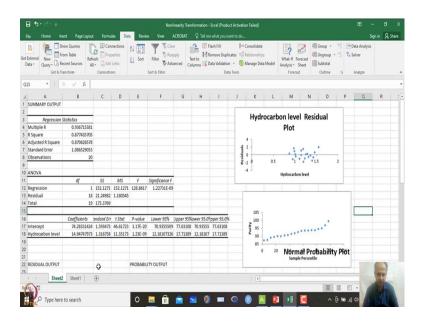


Now, I can store the residuals of standardized residual let us store, standardized residual which is generally recommended and this is nothing but residuals divided by standard deviation of residuals. So, that is known as standardized residual and we want to save that one and when I save this one and I click ok.

(Refer Slide Time: 16:09)



(Refer Slide Time: 16:16)



(Refer Slide Time: 16:20)

File Ho	Shu	ow Queries om Table cent Sources	ayout Formulas D Connections Refresh D All - D Edit Links Connections	nta <u>Review</u> ⊉↓ <u>X A</u> X↓ Sort F Sort	1 6	ceapply	👔 🕼 Flash Fil	Duplicates D	Consolidate Relationship	25	What-If F Analysis * Foreca	orecast Sheet	Group · Ungroup Subtotal Outline		🔁 Data Analysis	ign in A	Caller
Picture 2	• 1 2	< v 1															
A	8	c	DE	G	н	1.1	J K	L	м	N	0	р	Q	R	S	T	L
D 0 1 2 8 8 8			gression rity = 74.28 Coefficier	+ 14.95	5 Hyd	drocarl				0							
1		0	Term	0	oef s	SE Coef	T-Value	P-Value	VIF	0							
1			Constant	7.	4.28	\$1.59	46.62	0.000									
-																	
			Hydrocarbo	n level 14	4.95	1.32	11.35	0.000	1.00								
		0				0				-0					-	-	
	Sheet2	Sheet3	Sheet1 🕀	0					1				III	0 6 %			6

What will happen is that I will get the regression equation. When I get the regression equation, so this is the equation that you are seeing. So, I can copy this one copy as picture and I can just paste this one. Let us say I want to see the results. So over here what I have done is that this is the first result, the equation that MINITAB has generated is purity is equals to this is the intercept $\beta_0 = 74.28$ this is estimated based on the formulation that we have shown.

And 14.95 is the β_1 that is the slope and 1 unit increase of hydrocarbon level, how much will be the average increase in the purity that is given by 14.95 like that. So, β_0 estimation β_1 estimation over here and then if you go to the second results these are the this is the second results that you will get and which we can paste over here and see this one.

And here the coefficient is given, so when I see constant that is the β_0 estimation coefficient is 74.28 which is reflected over here and hydrocarbon is β_1 intercept is 14.95 that is given over here ok. Standard error of this is also given and the corresponding T values and P-values are given P less than 0.05 will indicate that β_1 is significant over here and this is statistically significant.

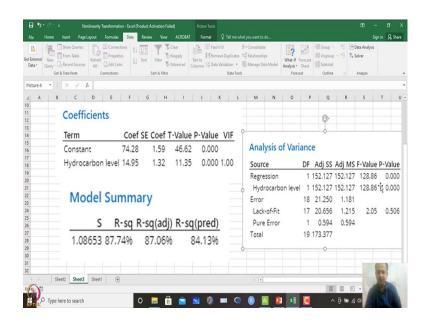
So, it is basically saying that there is a slope and we can consider this hydrocarbon level as one of the variables which is explaining the variability of purity over here ok. Constant is also β_0 is significant statistically, so we should retain this one β_0 ok. So, that is also signifies signified over here.

(Refer Slide Time: 17:36)

Fil	linitab - Untitle																		- 8
."	Edit Data	Calc	Stat	Graph	View Help	Assistan	nt Additional Tool	5											
2	6 8 X	0 6	5	0	14 14	00	fx 3=== 1 1	1240											
1							8 + 1 + 1		Rk	+ 12,0		Y 收回							
							MERK												
			_	_															
R	gression Analy	rsis: Puri	y '	×															
8	SIMPLE UNEAR	REGRESS	ONM	WX															
R	gression	Analy	sis:	Purity	versus	Hydro	carbon level												
	Model Sum	mary				*													
	s	R-sq	-sq(a	idj) R·s	q(pred)														
	1.08653 8	7.74%	\$7.0	1696	84.13%														
		_	-			C	5												
	Analysis of	Varian	e				a.												
	Source		DF	Adj SS	Adj MS	F-Value	P-Value												
	Regression			152.127	152.127	128.86	0.000												
				152.127	152.127	120.00	0.000												
	Hydrocarbor	level	1	152.127	152.127	128.85	0.000												
	Hydrocarbor Error	level	1 18	152.127 21.250	152.127 1.181	128.85	0.000												
	Hydrocarbor Error Lack-of-Fit	level	1 18 17	152.127 21.250 20.656	152.127 1.181 1.215														
	Hydrocarbor Error Lack-of-Fit Pure Error	level	1 18 17 1	152.127 21.250 20.656 0.594	152.127 1.181	128.85	0.000												
	Hydrocarbor Error Lack-of-Fit	level	1 18 17 1	152.127 21.250 20.656	152.127 1.181 1.215	128.85	0.000												
	Hydrocarbor Error Lack-of-Fit Pure Error	level	1 18 17 1	152.127 21.250 20.656 0.594	152.127 1.181 1.215	128.85	0.000	C6	C7	C8	C9	C10	C11	C12	C13 12	C14	C15 2	C16	6
	Hydrocarbor Error Lack-of-Fit Pure Error Total		1 18 17 1 19 C2	152.127 21.250 20.656 0.594 173.377	152.127 1.181 1.215 0.594	128.86 2.05 C4	0.000		67	C8	C9 Temperature	C10 Yeild	C11		C13 S MARKET PRICE	C14	C15 Z Demand NL (Y) E		
	Hydrocarbor Error Lack-of-Fit Pure Error Total		1 18 17 1 19 C2	152.127 21.250 20.656 0.594 173.377	152.127 1.181 1.215 0.594	128.86 2.05 C4	0.000 0.506 CS		C7	C8		Yeild 73.3	C11		MARKET PRICE	C14	Demand NL (Y) E 0.79		sage(
	Hydrocarbor Error Lack-of-Fit Pure Error Total		1 18 17 1 19 C2	152.127 21.250 20.656 0.594 173.377 n level	152.127 1.181 1.215 0.594 C3 g Purity 90.01 89.05	128.86 2.05 C4	0.000 0.506 CS Sale Price(/1000) 25.9 29.5	Taxes (/1000) 4.9176 5.0208	C7	C8	Temperature 55.0 56.0	Yeild 73.3 74.6	CII	HOUSE SIZE	MARKET PRICE 156 153	C14	Demand NL (Y) 8 0.79 0.44		sage(6) 2!
	Hydrocarbor Error Lack-of-Fit Pure Error Total		1 18 17 1 19 C2	152.127 21.250 20.656 0.594 173.377 n level 0.99	152.127 1.181 1.215 0.594 C3 2 Purity 90.01 89.05 91.43	128.86 2.05 C4	0.000 0.506 Sale Price(/1000) 25.9 29.5 27.9	Taxes (/1000) 4.9176 5.0208 4.5429	C7	C8	Temperature 55.0	Veild 73.3 74.6 74.0	C11	HOUSE SIZE 72 98 92	MARKET PRICE 156 153 230	C14	Demand NL (Y) 8 0.79 0.44 0.56		sage(6: 2! 10
	Hydrocarbor Error Lack-of-Fit Pure Error Total		1 18 17 1 19 C2	152.127 21.250 20.656 0.594 173.377 n level 0.99 1.02 1.15 1.29	152.127 1.181 1.215 0.594 C3 2 Purity 90.01 89.05 91.43 93.74	128.86 2.05 C4	0.000 0.506 Sale Price(/1000) 25.9 29.5 27.9 25.9	Taxes (/1000) 4.9176 5.0208 4.5429 4.5573	C7	C8	Temperature 55.0 56.0 55.5 59.0	Yeild 73.3 74.6 74.0 78.5	C11	HOUSE SIZE 72 98 92 90	MARKET PRICE 156 153 230 152	C14	Demand NL (Y) 8 0.79 0.44 0.56 0.79		sage(6) 2! 10 4!
	Hydrocarbor Error Lack-of-Fit Pure Error Total		1 18 17 1 19 C2	152.127 21.250 20.656 0.594 173.377 n level 0.99 1.02 1.15 1.29 1.46	152.127 1.181 1.215 0.594 C3 2 Purity 90.01 89.05 91.43 93.74 96.73	128.86 2.05 C4	0.000 0.506 Sale Price(/1000) 25.9 29.5 27.9 25.9 25.9 25.9 25.9 25.9	Taxes (/1000) 4.9176 5.0208 4.5429 4.5573 5.0597	C7	C8	Temperature 55.0 56.0 55.5 59.0 56.0	Yeild 73.3 74.6 74.0 78.5 74.6	C11	HOUSE SIZE 72 98 92 90 44	MARKET PRICE 156 153 230 152 42	C14	Demand NL (Y) 8 0.79 0.44 0.56		sage(6) 2! 10 4!
	Hydrocarbor Error Lack-of-Fit Pure Error Total Observation 1 2 3 4 4 5 6	Hydro	1 18 17 1 19 C2 arbo	152.127 21.250 20.656 0.594 173.377 n level 0.99 1.02 1.15 1.29 1.46 1.36	152.127 1.181 1.215 0.594 C3 g Purity 90.01 89.05 91.43 93.74 96.73 94.45	128.86 2.05 C4	0.000 0.506 Sale Price(/1000) 25.9 29.5 27.9 25.9	Taxes (/1000) 4.9176 5.0208 4.5429 4.5573 5.0597	C7	C8	Temperature 55.0 56.0 55.5 59.0 56.0 56.0 55.5	Yeild 73.3 74.6 74.0 78.5	C11	HOUSE SIZE 72 98 92 90	MARKET PRICE 156 153 230 152	C14	Demand NL (Y) 8 0.79 0.44 0.56 0.79		sage 6 2 10 4
	Hydrocarbor Error Lack-of-Fit Pure Error Total	Hydro	1 18 17 1 19 C2 arbo	152.127 21.250 20.656 0.594 173.377 n level 0.99 1.02 1.15 1.29 1.46	152.127 1.181 1.215 0.594 C3 g Purity 90.01 89.05 91.43 93.74 96.73 94.45	128.86 2.05 C4	0.000 0.506 Sale Price(/1000) 25.9 29.5 27.9 25.9 25.9 25.9 25.9 25.9	Taxes (/1000) 4.9176 5.0208 4.5429 4.5573 5.0597	C7	C8	Temperature 55.0 56.0 55.5 59.0 56.0	Yeild 73.3 74.6 74.0 78.5 74.6	CII	HOUSE SIZE 72 98 92 90 44	MARKET PRICE 156 153 230 152 42 157		Demand NL (Y) E 0.79 0.44 0.56 0.79 2.70		
	Hydrocarbor Error Lack-of-Fit Pure Error Total Observation 1 2 3 4 4 5 6	Hydro	1 18 17 1 19 C2 arbo	152.127 21.250 20.656 0.594 173.377 n level 0.99 1.02 1.15 1.29 1.46 1.36 r Regress	152.127 1.181 1.215 0.594 C3 g Purity 90.01 89.05 91.43 93.74 96.73 94.45	128.86 2.05 C4	0.000 0.506 Sale Price(/1000) 25.9 29.5 27.9 25.9 25.9 25.9 25.9 25.9	Taxes (/1000) 4.9176 5.0208 4.5429 4.5573 5.0597	C7	C8	Temperature 55.0 56.0 55.5 59.0 56.0 56.0 55.5	Yeild 73.3 74.6 74.0 78.5 74.6	C11	HOUSE SIZE 72 98 92 90 44	MARKET PRICE 156 153 230 152 42		Demand NL (Y) 8 0.79 0.44 0.56 0.79		sage(6) 2! 10 4!

And then model summary you will find over here model summary. So, I can just copy this one also copy as picture and we can press this one ok.

(Refer Slide Time: 17:48)



So, we can place this one over here, so in this case model adequacy. So, this is one of the; one of the measures that like R values that we have told correlation coefficient is nothing but the correlation coefficient that you see over here R square which is known as coefficient of determination, which is same in case there is one variable over here. So, it will come out to be same 87.74. So, it is converted into percentage.

So, I can convert into between minus 1 and plus 1, so 0.877 you can think of this is more than 0.7 ok and this is calculated by another formulas over here that is known as SS regression. How much of the variability by SS total basically which Anova analysis will tell you and the Anova tables will summarize that one. So, SS regression but SS total will tell me how much of the variability of y is explained by basically this hydrocarbon level variable over here.

So, it is around 87 percent like that so you can think of ok. So, that is quite good one of the variable is explaining so much of variability of the y that you observe. So, when I change the x it is influencing the expected value of y basically that is the interpretation that we can make out of this ok.

(Refer Slide Time: 18:55)

	itab - Untitled																- 6
	Edit Data C	alc Stat	Graph	View Help	Assistan	nt Additional Tool	ls .										
E		16		4. 46	00	fx 3= -2 1 8	here allo										
									- N. 5		W 16 150						
•						图 # 70 # 1		BR	× 591	0 0° C 0° .	Y 44 E						
	4	EITA		SK *	Y Y	NESK!	K.										
err	ession Analysis:	Purity	× x														
	VPLE LINEAR REG																
					Under	and an Invel											
eg	ression Ar	nalysis:	Purity	versus	Hydro	carbon leve											
A	nalysis of Va	riance				~											
se	ource	DF	Adj SS	Adj MS	E-Value	P-Value											
	egression	1	152.127		128.86	0.000											
+	Hydrocarbon lev	vel 1	152.127	152.127	188.86	0.000											
Er	ror	18	21.250		0												
	ack-of-Fit	17	20.656		2.05	0.506											
F	Pure Error	1	0.594														
To	lase	19	173.377														
-	Obs Purity 9 99.420		Kesid St 1.968	2.07 R	- C4	G	C6	C7	C8	C9	C10	CII	C12	C13 m	C14	C15 g	C16
0	bservation Hy		n local	Purity		Sale Price(/1000)		C/	Co	Temperature	Yeild	CII		C13 MARKET PRICE	C14	C15 Demand NL (Y)	
-	1	yuiocaroo	0.99	90.01		25.9	4.9176			55.0	73.3		72	156		0.79	chergy usage
	2		1.02	89.05		29.5	5.0208			56.0	74.6		98	153		0.44	2
	3		1.15	91.43		27.9	4.5429			55.5	74.0		92	230		0.56	10
	4		1.29	93,74		25.9	4.5573			59.0	78.5		90	152		0.79	4
	5		1.46	96.73		29.9	5.0597			56.0	74.6		44	42		.3.70	
	6		1.36	94.45		29.9				55.5	74.0		46	157			
							20410			4	74.0						10000
•	H + SI	imple Lines															100 100
+		imple Linea		HOLINWX						•	_	-	_			-	Aller .
-	H + Si Simple Linear			ROUTINAX						•		_		6			

So, then what are the other results that we are getting over here this is the analysis of variance that you see over here. So, this I can copy again and I can paste it over here. So, I can paste it here and just try to see what is the interpretation of this ok. So, over here what you see is that regression this is the regression equation that is developed, it is showing that P-value is less than 0.05.

That means, this regression is quite this equation is quite significant and in that case we can adopt this equation and hydrocarbon level is that is basically the variable x that we are considering P-value is significant over here. And we will also find a lack of fit testing over here, that means whether there is any nonlinearity in the model, that is we have to adopt and go to higher order equations like that, that will be given by lack of fit test and a formula is given in any books like that.

So, if you have multiple observation at a given level of x then lack of fit can be calculated and lack of fit over here is calculated as 0.5 which is not more than less than 0.05; that means, there is no lack of fit as such. So, linear model is quite sufficient to explain the variability and that is adequate over here. So, I do not need to go to higher order terms over here.

So, this is lack of fit what we can get ok. So, this regression equation one way is seeing these coefficients over here β_1 is significant β_0 is significant like that and overall if you want to see that whether the regression equation is making sense.

So, this value of regression this P-value we have to we have to see and generally it will agree. So, both of the coefficient is significant then only regression will be significant like that, so that is the interpretation that we can make ok.

Minitab - Untitle File Edit Data Calc Stat * IIY X BOK * ZYMERKY Analysis: Purity ... ' SIMPLE LINEAR R Regression Analysis: Purity versus Hydrocarbon level
 Hug as
 Hug ms
 F value
 F value

 152.127
 152.127
 128.86
 0.000

 152.127
 152.127
 128.86
 0.000
 152.127 152.127 1.181 1.215 0.594 152.127 152.127 21.250 20.656 2.05 0.506 0.594 its and Diagr ostics for Unusual Observation
 Obs
 Purity
 Fit
 Resid
 Std Resid

 9
 99.420
 97.452
 1.968
 2.07 R
 Send to Comp Send to Minitab Works C3 8 Purity 90.01 89.05 91.43 C2 C4 C1! Sale 🖗 0.99 1.02 X Delete 73.3 74.6 29.5 5.0208 4.5429 55.5 27.9 25.9 29.9 29.9 74.0 78.5 74.6 74.0 4.5573 5.0597 3.8910 1.29 93.74 59.0 56.0 55.5 96.73 0 🥫 -1 . 0 R ere to search

(Refer Slide Time: 20:32)

(Refer Slide Time: 20:42)

(acc)	Show Queries 2 Connections of 716 The Clear	ا Duplicates ۹	Consolidate Relationships Manage Data Model Anaysis - Forece	Sheet Subtota	- 👘 🔚 Data Analysis - 🕂 2, Solver	ign in 🔑 Shaw
icture 5 *	4 ∨ ×					
A	B C D E F G H I J K	L	M N O	P Q	R S	T
			Regression	1 152.127	152.127 128.8	5 0.000
			Hydrocarbon level	1 152.127	152.127 128.8	5 0.000
	Model Summary		Error	18 21.250		
	Woder Summary					
	-		Lack-of-Fit	17 20.656		5 0.506
	S R-sq R-sq(adj) R-sq(pred)		Pure Error	1 0.594	0.594	
			Total	19 173.377		
	1.08653 87.74% _@ 87.06% 84.13%					
	-					
	Fits and Diagnostics for Unusual Observations					
	Obs Purity Fit Resid Std Resid 9 99.420 97.452 1.968 2.07 R					
	,					
	R Large residual					
	0 0 0					
	iheet2 Sheet3 Sheet1 (+)		1.			0

Then some unusual observations over here when standard residual is more than. So if you can see this one copy as picture; so we can press this one. So, when we have certain observation which is beyond two that is standardized residual which indicates that this is the unusual observations like that. So, then we have to see whether to include that one or exclude that one. So, whether it is outlier like that whether we need to eliminate that one, regression also this type of information is also useful when we do regression analysis. So, that is why standardized residual is used to identify any outlier observations something we can do that ok.

So, and we have to be careful in dealing with outliers, so there are many ways of dealing with outliers ok. So, what we can understand based on this simple example over here is that every condition is satisfactory.

(Refer Slide Time: 21:22)

E1.	Ainitab - Untitled	ł																	-	8
n#	e Edit Data	Calc Sta	Graph	View H	elp Assistar	nt Additional Tools														
•	8 8 X	DA C	000	A AL	00	fx 2	244													
						🗄 i 🕸 🍢 🏨 🕸			4 XL 5	a transfer	JULY 2	. 59								
1						NERK		and the LL												
	1	C I.	T IC I	195.7	1 1 2 1 M	CODE X														
Re	gression Analys	sis: Purity	ν х																	
	SIMPLE LINEAR P	REGRESSION	MWX																	
Re	earession	Analysi	: Pur	ty versi	is Hydro	carbon level														
					singuio	early off for er														
	1.08653 87	R-SQ R-SI	(adj) 7.06%	84.13%																1
	1000000 07	11410 0	10017	04.137																
	Analysis of	variance																		
	Source	D			IS F-Value															
	Regression Hydrocarbon	level 1				0.000														
	Hydrocarbon	level 18				0.000														
	Lack-of-Fit	15				0.506														
	Pure Error	1	0.5	94 0.59	4															
	Total	19	173.3	77																
						1.00														
	Fits and Dia	gnostics	for Un	usual Ob	servations	×														
-						·	C17	40	C19	C20	(21	C22	C23	C24	(25	C26	C27	C28	(29	,
	C12	C13		C14	C15	c16	C17	station of the second s	C19	C20	C21	C22	C23	C24	C25	C26	C27	C28	C29	,
	C12 HOUSE SIZE	C13	RICE	C14	C15 emand NL (V	c16 () Energy Usage()	SQRT(Y)	SRES	C19	C20	C21	C22	C23	C24	C25	C26	C27	C28	C29	,
	C12 HOUSE SIZE 90	C13	RICE 152	C14	C15 emand NL (V 0.7	C16 () Energy Usage() (9 49	5 SQRT(Y) 3 0.88882	SRES 0.16585	C19	C20	C21	C22	C23	C24	C25	C26	Q7	C28	C29	
1	C12 HOUSE SIZE	C13	RICE	C14	C15 emand NL (V	C16 C Energy Usage() 9 49 0 58	 SQRT(Y) 0.88882 1.64317 	SRES	C19	C20	C21	C22	C23	C24	C25	C26	C27	C28	C29	
	C12 HOUSE SIZE 90 44	C13	RICE 152 42	C14	C15 emand NL (N 0.7 2.7	C16 C16 CEnergy Usage(X 9 4 0 58 4 115	 SQRT(Y) 0.88882 1.64317 1.90788 	SRES 0.16585 0.62315	C19	C20	C21	C22	C23	C24	C25	C26	C27	C28	C29	
1	C12 HOUSE SIZE 90 44 46	C13	RICE 152 42 157	C14	C15 emand NL (V 0.7 2.7 3.6	C16 C16 D Energy Usage(X 9 49 0 58 4 115 3 99	 SQRT(Y) 0.88882 1.64317 1.90788 2.17486 	SRES 0.16585 0.62315 -0.15615	C19	C20	(21	C22	C23	C24	C25	C26	Q7	C28	C29	
	C12 HOUSE SIZE 90 44 46 90	C13	RICE 152 42 157 113	C14	C15 emand NL (N 0.7 2.7 3.6 4.7	 C16 Energy Usage(X) For S8 S8 S115 S990 C15 C16 C16	 SQRT(Y) 0.88882 1.64317 1.90788 2.17486 3.08221 	SRES 0.16585 0.62315 -0.15615 0.31234	C19	C20	(21	C22	C23	C24	C25	C26	(27	C28	C29	
	C12 HOUSE SIZE 90 44 46 90 150 94	C13	RICE 152 42 157 113 573 202	C14 D4	C15 emand NL (N 0.7 2.7 3.6 4.7 9.5 5.3	 C16 Energy Usage(X) For S8 S8 S115 S990 C15 C16 C16	 SQRT(Y) 0.88882 1.64317 1.90788 2.17486 3.08221 	SRES 0.16585 0.62315 -0.15615 0.31234 -0.84939	C19	C20	C21	C22	C23	C24	25	C26	C27	(28	C29	
1 5 7 3	С12 HOUSE SIZE 90 44 46 90 150 91 5 н +	C13 MARKET P	152 152 42 157 113 573 202 car Regi	C14 D4	C15 emand NL (N 0.7 2.7 3.6 4.7 9.5 5.3	 C16 Energy Usage(X) For S8 S8 S115 S990 C15 C16 C16	 SQRT(Y) 0.88882 1.64317 1.90788 2.17486 3.08221 	SRES 0.16585 0.62315 -0.15615 0.31234 -0.84939	C19		C21	C22	C23	C24		C26		C28	C29	c

(Refer Slide Time: 21:31)

File	initab - Untitled Edit Data Calc	Stat Gr	inh View	Helo Assistant	Additional Tools														8
-			Statistics																
		Regr	ession		Fitted Line Plot	1	a	2 + Y. F		III LA IA	50								
ч.		ANO	VA		Regression		11	Fit Regression	Model	an 1 2	0.5								
	<u> </u>	DOE		. 17.	Nonlinear Regressio	h	1	Best Subsets.	1.83										
Reg	gression Analysis: Pu	e Cont	rol Charts		Stability Study			Predict											
	SIMPLE LINEAR REGRE	Qual	ty Tools	,				Factorial Plot											
-	gression Ana	0.5.	bility/Surviva	🕅	Orthogonal Regress	00	- 6	Contour Plot											
Ne		Pred	ctive Analytic	cs 🔸 🖄	Partial Least Squares														
	S K-sq	Mult	variate		Binary Fitted Line Pl														
	1.08653 87.74%	Time	Series	, Ľ				Response Op											
		Table	\$	10	Binary Logistic Regn Ordinal Logistic Reg			response op	unitedia										
A	Analysis of Varia	Non	arametrics)															
5	Source	Equi	alence Tests	, litts	Nominal Logistic Re	gression													
R	Regression	Pow	er and Sample	e Size +	Poisson Regression														
	Pure Error Total		0.594 0. 3.377	594															
-	its and Diagnos	itics for l	Jnusual O	bservations	*														
F		ctics for U	Jnusual O	bservations C15 g	~ C16	C17	C18	C19	C20	C21	C22	C23	C24	C25	C26	C27	C28	C29	
F		C13 😰	C14	C15 g		C17 SQRT(Y)	C18	C19	C20	C21	C22	C23	C24	C25	C26	C27	C28	C29	
F	C12	C13 😰	C14	C15 g	C16 Energy Usage(X)		C18	C19	C20	C21	C22	C23	C24	C25	C26	C27	C28	C29	
F	C12 HOUSE SIZE MAR	C13	C14	C15 g Demand NL (Y)	C16 Energy Usage(X) 493	SQRT(Y)	C18	C19	C20	C21	C22	C23	C24	C25	C26	C27	C28	C29	
F	C12 HOUSE SIZE MARI 90	C13 S KET PRICE	C14	C15 E Demand NL (Y) 0.79	C16 Energy Usage(X) 493 582	SQRT(Y) 0.88882	C18	C19	C20	C21	C22	C23	C24	C25	C26	C27	C28	C29	
F	C12 HOUSE SIZE MARI 90 44	C13 (KET PRICE 152 42	C14	C15 Demand NL (V) 0.79 2.70	C16 Energy Usage(X) 493 582 1156	SQRT(Y) 0.88882 1.64317	C18	C19	C20	C21	C22	C23	C24	C25	C26	C27	C28	C29	(
F	C12 HOUSE SIZE MARI 90 44 46	C13 KET PRICE 152 42 157	C14	C15 2 Demand NL (Y) 0.79 2.70 3.64	C16 Energy Usage(X) 493 582 1156 997	SQRT(Y) 0.88882 1.64317 1.90788	C18	C19	C20	Q1	C22	C23	C24	C25	C26	C27	C28	C29	(
F	C12 HOUSE SIZE MARI 90 44 46 90	C13 52 KET PRICE 152 42 157 113	C14	C15 Demand NL (Y) 0.79 2.70 3.64 4.73	C16 Energy Usage(X) 493 582 1156 997 2189	SQRT(Y) 0.88882 1.64317 1.90788 2.17486	C18	C19	C20	Q1	C22	C23	C24	C25	C26	Q7	C28	C29	
F	C12 HOUSE SIZE 90 44 46 90 150 94	C13 52 KET PRICE 152 42 157 113 573 202	C14	C15 2 Demand NL (Y) 0.79 2.70 3.64 4.73 9.50 5.34	C16 Energy Usage(X) 493 582 1156 997 2189	SQRT(Y) 0.88882 1.64317 1.90788 2.17486 3.08221	C18	C19	C20	Q1	C22	C23	C24	C25	C26	27	C28	C29	
F	C12 HOUSE SIZE 90 44 46 90 150 94	C13 152 KET PRICE 152 42 157 113 573 202 4e Linear Re	C14	C15 2 Demand NL (Y) 0.79 2.70 3.64 4.73 9.50 5.34	C16 Energy Usage(X) 493 582 1156 997 2189	SQRT(Y) 0.88882 1.64317 1.90788 2.17486 3.08221	C18	C19		Q1	C22	C23	C24		C26	Q1		C29	

(Refer Slide Time: 21:32)

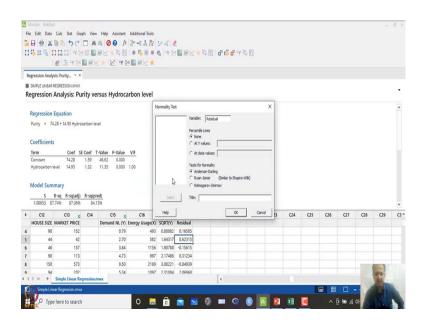
Bulk LinkAn RedetSourcewick Keymics Keymics Keymics Source OF Adj SS Adj MS F-Value P-Value Source Month RedetSourcewick Compared pendetres Source OF Adj SS Adj MS F-Value P-Value Source Month RedetSourcewick Compared pendetres Source OF Adj SS Adj MS F-Value P-Value Compared pendetres Compared pendetres Production bend Compared pendetres Compared pendetres Compared pendetres Compared pendetres Production bend Compared pendetres Compared pendetres Compared pendetres Compared pendetres Production bend Compared pendetres Compared pendetres Compared pendetres Compared pendetres Compared pendetres Productin bend Compared pe	Minitab - Untitled	1 Addieud Tech			- 8
13 III III IIII IIIIIIIIIIIIIIIIIIIIIII					
Bit Mit Link King * * Bit Mit Link King King * * Bit Mit Link King King King King King King King King					
Regression Analysis: Purity versus Hydrocant Regression Analysis: Purity versus Hydrocant 1.0653 87.268 87.068 Alaysis of Variance Source of Adj SS Adj SS Fully and the purity versus Hydrocant Improvide the provide the				4 3	
Bulk LinkAn RedetSourcewick Keymics Keymics Keymics Source OF Adj SS Adj MS F-Value P-Value Source Month RedetSourcewick Compared pendetres Source OF Adj SS Adj MS F-Value P-Value Source Month RedetSourcewick Compared pendetres Source OF Adj SS Adj MS F-Value P-Value Compared pendetres Compared pendetres Production bend Compared pendetres Compared pendetres Compared pendetres Compared pendetres Production bend Compared pendetres Compared pendetres Compared pendetres Compared pendetres Production bend Compared pendetres Compared pendetres Compared pendetres Compared pendetres Compared pendetres Productin bend Compared pe	· @ II Y ⊠ B K ★ IØ Y	×∎øk*			
B MARL MARK REGISSION MARK Regression Analysis: Purity versus Hydrocan 5 Year Mark Mark Regression 10653 87.268 87.068 84.138 Analysis of Variance Improve the first or the first	Regression Analysis: Purity Y X				
s n + n K	SIMPLE LINEAR RECRESSION MWY	Regression		×	
5 K44 K-92000 K-920000 K-920000000 K-920000000000000000		C1 Observation	Responses:		
S 5.45 R V R 1000 R V R 10000 R V R 1000 R V R 10000 R V R 100000 R V R 100000 R V R 1	• • • •		Purity	×	
Source OF Adj SS		C5 Sale Price(/1000)		w	
Analysis of Variance Updatase Updatase<	1.08053 87.74% 87.06% 84.13%	C9 Temperature	Contra ou a made tour		
Analysis of Variance Cite Model 7, 100 Cite Model 7, 100 Cite Model 7, 100 Source Of Adj SS Adj MS FVAlue Proj. Eith Model 7, 100 Eith Model 7, 100 Eith Model 7, 100 Mydissicationed 1 152,127 152,127 122,86 Cite Model 7, 100 Eith Model 7, 100 Mydissicationed 1 152,127 122,20 Lith Model 7, 100 Eith Model 7, 100 Eith Model 7, 100 Parterior 1 0.594 0.594 0.594 0.594 0.594 Parterior 1 0.594 0.594 0.594 0.594 0.594 0.594 MOUSE SZZ MMAKT PROCC Demand NL (7) no Model 1.514 0.594				<u></u>	
Source 0F Ags Ags Ags Free Bit Disposition I I Hydrocarbol Held 1 52,127 152,177 164 164 157 152,177	Analysis of Variance	C13 MARKET PRICE			
Nagratume 1	Source DF Adj SS Adj MS F-Value	P-Val C16 Energy Usage(X)	т		
Error 18 21.20 1.181 Lossefic 7000 2005 1.125 2000 10000	Regression 1 152.127 152.127 128.86	0.0 C17 SQRT(Y)	LL	<u></u>	
Label 45, 17, 2005, 1215, 203 0. Pure fror 1, 0354, 0394 Teal 19, 173,377 Fits and Diagnostics for Unusual Observations		0.4	Cgtegorical predictors:		
Burger 1 0.954 0.954 Teal 19 173.377 Fits and Diagnostics for Unual Observations				~	
Total 19 173.377 Fits and Diagnostics for Unusual Observations		0.5			
Bidd:::::::::::::::::::::::::::::::::::					
PHS and Disglostics for Unusual Observations Select yukken gentu gentu gentu gentu gentu HOUSS SZZ MAKKET PROC CAS CS CS CAS CS CAS	10001 19 173.377		L		
CLC CLC <thclc< th=""> <thclc< th=""> <thclc< th=""></thclc<></thclc<></thclc<>	Fits and Diagnostics for Unusual Observation		Model Optiogs Co	drg Stepwise	
HOUSE SUZE MARKET PRICE 00 152 0.77 e/db 00 Const 44 42 2.70 552 154517 54 51 72 2.44 155 15708 55 73 9.50 2.116 15708 155 573 9.50 2.116 3.0221 1 14 * Safe Data Represionance 4 P H * Safe Data Represionance 4 const 1 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2	C12 C13 c14 C15	Select	<u>Yalidation</u> <u>Graphs</u> <u>Be</u>	sults Storage 4 C25 C26 C27 C28	C29 C3
44 42 2.70 382 1454317 46 137 3.64 1156 150708 90 138 4.73 999 2.1766 150 573 9.50 2.189 30021 1 H + Single Toker Regression.mux 4	HOUSE SIZE MARKET PRICE Demand NL (*) Ene			
1 44 137 3.44 1194 10000 10 113 4.73 997 2.1746 101 101 5.73 6.90 2100 1021 101 1	4 90 152 0.7	9 Help		QK Cancel	
1 90 113 4.73 99 2.1746 1 150 573 5.95 2.19 3.0221 4 2 × + Single Liear Regression.mix 4 p × + Single Liear Regression.mix 4 0 mg/d Liear Regression.mix 4 0 mg/d Liear Regression.mix	5 44 42 2.7	0 582 1.643	17		
1 150 573 9.50 2189 3.00221 1 1 4 Single Toor Regression.mus 4 I 4 Single Toor Regression.mus 4 Confide Lever Regression.mus	6 46 157 3.6	4 1156 1.907	88		
S4 2027 S.S4 1097 2.310M 4 1 P.H. + Simple Linear Regression.mms 6 If Simple Linear Regression.mms If Simple Linear Regression.mms	7 90 113 4.7	3 997 2.174	86		
4 >> + Simple Lieur Regression.mm	8 150 573 9.5	0 2189 3.083	21		-
🕼 Simple Linear Regression.max	a 94 202 5.3	4 1097 2.310	R4		
	4 b H + Simple Linear Regression.mwx		4		H
	Simple Linear Repression may				A PARA
🖓 🖓 Type here to search 💦 🔗 🧰 🛜 🔁 🔞 💷 🚱 🔞 🚺 🛐 🙀 A 🖗 👘 🖉 🖓					
	P Type here to search	0 🥫 🖥	💼 🕺 🔞 🚥 😨 📵) 🔟 😰 👪 🥃 🔹 ^ 🖗 🖗 🕼 🚺	A NUMBER

(Refer Slide Time: 21:34)

Minitab - Untitled															0.0
File Edit Data Calc Stat Graph View Help Assi	stant Additio	nal Tools													
🔓 日 👳 & 🐚 🏠 ちぐ 🗂 株共 🖉 🚱	fx 3= -1	1.22	24. 2												
	(四) # #	兼兼		BK .	400 00	0" LY 1	中国								
Regression Analysis: Purity Y X	-														
SIMPLE LINEAR REGRESSION MWX	Re	gression							×						
Regression Analysis: Purity versus Hyd	rocarb	Obser	vation Ree					-	_						٠
	- G		Regression: Stor	age				×	^						
5 K-sq K-sq(adj) K-sq(pred) 1.08653 87.74% 87.06% 84.13%	C5 06	Taxes	E Bts		□ Cogfficier	nts			~						^
	C9 C1	0 Yeld	E Residuals		C Design m	atrix			_						
Analysis of Variance	Ci	2 HOUS	Standardized	residuals					^						
and a second and a second	C1	5 Dena	C Deleted residu	als											
Source DF Adj SS Adj MS F-Valu Regression 1 152.127 152.127 128.3			Leverages												
Hydrocarbon level 1 152.127 152.127 128.			Cosk's distance												
Error 18 21.250 1.181			IT DEITS	Do				-	~						
Lack-of-Fit 17 20.656 1.215 2.	05 0.5														
Pure Error 1 0.594 0.594 Total 19 173.377															
Total 19 173.377								-	<u> </u>						
Fits and Diagnostics for Unusual Observation	ins	_						Ştepwise.	<u> </u>						v
+ C12 C13 g C14 C15		Sele	Help		_	QK	Cancel	Storage	4	C25	C26	C27	C28	C29	C3 ⁴
HOUSE SIZE MARKET PRICE Demand N	L (Y) Ene														
4 90 152	0.79	Help					QK	Cancel							
	2.70	583	2 1.64317			-			_						
	3.64	115													
	4.73	99													
	9.50	218											1.000		
	5.34	109	7 2.31084		1.1									and a	
H 4 D H + Simple Linear Regression.mwx					4		_							AN A	
Simple Linear Regression.mwx														1.0	
P Type here to search	0		1 🔒 🚺		(A) ==	0	R	5	XI			5 m @	40		
NPTEL						11			100	-				1.1	

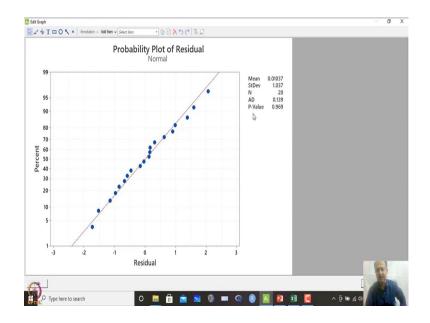
Now, one more thing what we have to do is that we have to save the residuals. So, let me just delete this one, I do not know whether this is we have saved or not. So, we I go to residual regression over here fit regression like that this is the one storage whether we have standardized residual were saved over here.

(Refer Slide Time: 21:42)

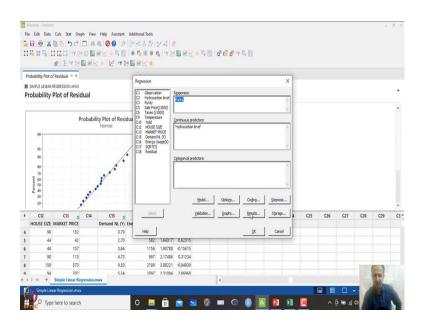


So, this is ok and if I save this one the last one is standardized residually this can be considered as residual over here and there are 3 checks that has to be under we have to undergo 3 test over here which is the one of the test is normality assumptions like that, so one test what we can do is that we can see the basic statistics we can go to normality test over here and we can see the residual whether it is normal or not.

(Refer Slide Time: 22:02)

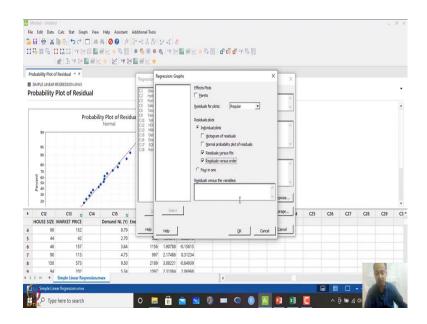


So, in this case what we observe is that P-value is more than 0.05. So, there is no problem in the assumptions of normality over here there is another graph which can be seen over here. So, when I am; when I am doing this regression.



(Refer Slide Time: 22:15)

(Refer Slide Time: 22:17)



There are other possible graphs that we can draw. So, one is residual versus fit, so this will indicate heteroscedasticity is there or whether there is any autocorrelation that exists this is the second one with respect to order.

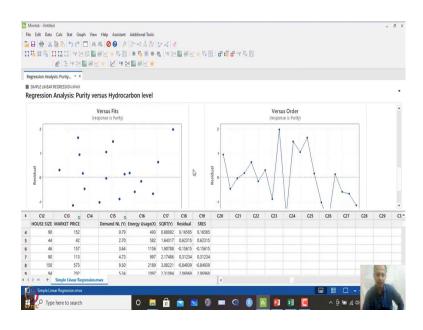
(Refer Slide Time: 22:28)

111	Ainitab - Untitle	d																-	Ð
	e Edit Data	Calc Stat Gr	ph View	Help Assistant	Additional Tools														
	8 8 %	06 50	0 4	1 001	8-24.8	24	E.												
1	16 II 18 1	I LI LI Y D	:0	●比★我国	* * * *	売 47 >		* 4 图	d" ("	OR LY 2	图								
				< * 12 4															
D 4		ysis: Purity × ×																	
		REGRESSION MWX																	
-					anh an Ioual														
<(egression	Analysis: Pt	inty ve	ersus Hydroc	arbon level														
	Regression	Equation																	
	Purity = 74	4.28 + 14.95 Hydro	carbon le	vel															
	Coefficient	s																	
	Term	Coef S	E Coef	T-Value P-Value	VIF														
	Constant	74.28	1.59	46.62 0.000															
	Hydrocarbon	level 14.95	1.32	11.35 0.000	1.00			2											
								rê.											
	Model Sum	imary																	
	s	R-sq R-sq(adj)	R-sq(p	red)															
	1.08653 87	7.74% 87.06%	84,	13%															
-	C12	C13 👩	C14	C15 8		C17	C18	C19	C20	C21	C22	C23	C24	C25	C26	C27	C28	C29	
				Demand NL (V)	Energy Usage(X)	SQRT(Y)	Residual	SRES											
	HOUSE SIZE	MARKET PRICE																	
	90	152		0.79	493	0.88882	0.16585	0.16585											
	90 44	152 42		0.79	493 582	1.64317	0.62315	0.62315											
	90 44 46	152 42 157		0.79 2.70 3.64	493 582 1156	1.64317 1.90788	0.62315	0.62315											
	90 44 46 90	152 42 157 113		0.79 2.70 3.64 4.73	493 582 1156 997	1.64317 1.90788 2.17486	0.62315 -0.15615 0.31234	0.62315 -0.15615 0.31234											
	90 44 46 90 150	152 42 157 113 573		0.79 2.70 3.64 4.73 9.50	493 582 1156 997 2189	1.64317 1.90788 2.17486 3.08221	0.62315 -0.15615 0.31234 -0.84939	0.62315 -0.15615 0.31234 -0.84939											
	90 44 46 90 150 94	152 42 157 113 573 202		0.79 2.70 3.64 4.73 9.50 5.34	493 582 1156 997 2189	1.64317 1.90788 2.17486	0.62315 -0.15615 0.31234	0.62315 -0.15615 0.31234	1.										
	90 44 46 90 150 94 5 н +	152 42 157 113 573 202 Simple Linear Re	-	0.79 2.70 3.64 4.73 9.50 5.34	493 582 1156 997 2189	1.64317 1.90788 2.17486 3.08221	0.62315 -0.15615 0.31234 -0.84939	0.62315 -0.15615 0.31234 -0.84939	1									P	
	90 44 46 90 150 94 5 н +	152 42 157 113 573 202	-	0.79 2.70 3.64 4.73 9.50 5.34	493 582 1156 997 2189	1.64317 1.90788 2.17486 3.08221	0.62315 -0.15615 0.31234 -0.84939	0.62315 -0.15615 0.31234 -0.84939	L.										

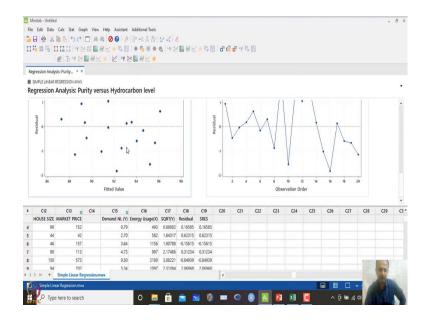
(Refer Slide Time: 22:29)

Reg		E		1681	K * 16	14 LY 2	× Bek*														
Reg																					
	pression Analysis:	Purity	* X																		
	SIMPLE LINEAR REG			ty ve	ersus P	Hydroca	arbon level														
A	Analysis of Var	iance																			
	Source	DF	Adj	A 22	MIMS F	F-Value	P-Value														
	Regression	1	152.1			128.86	0.000														
	Hydrocarbon leve	el 1	152.	27 15	52.127	128.85	0.000														
E	Error	18	21.3	50	1.181																
	Lack-of-Fit	17	20.4		1.215	2.05	0.506														
	Lack-of-Fit Pure Error Total	1	20.4	94	1.215 0.594	2.05	0.506														
	Pure Error	1	20.4	94		2.05	0.506		1	à											
T	Pure Error	1 19	20.0 0.5 173.3	94 77	0.594		0.506		I	rds -											
Fi	Pure Error Total	1 19 ostics	20.0 0.9 173.3	94 77 usual	0.594 Observ		0.506		I	à											
Fi	Pure Error Total	1 19 ostics Fit	20.4 0.5 173.3 for Ur Resid	94 77 usual Std Re	0.594 Observ		0.506		I	à											
Fi	Pure Error Total Fits and Diagn Obs Purity	1 19 ostics Fit	20.4 0.5 173.3 for Ur Resid	94 77 usual Std Re	0.594 Observesid		0.506		I	à											
Fi	Pure Error Total Fits and Diagn Obs Purity 9 99.420 S C12	1 19 ostics Fit 97.452 C13	20.0 0.5 173.3 for Ur Resid 1.968	94 77 usual Std Re	0.594 Observ esid 2.07 R	vations C15 g	C16	C17	C18	C19	C20	(21	C22	C23	C24	C25	C26	C27	C28	C29	
Fi	Pure Error Total Fits and Diagn Obs Purity 9 99.420 S C12 HOUSE SIZE MA	1 19 ostics Fit 97.452 C13	20.4 0.5 173.5 for Ur Resid 1.968 sz RICE	94 77 USUA Std Re	0.594 Observ esid 2.07 R	cts s nd NL (Y)	C16 Energy Usage(X	SQRT(Y)	C18 Residual	C19 SRES	C20	21	C22	C23	C24	C25	626	C27	C28	C29	
Fi	Pure Error Total Fits and Diagn Obs Purity 9 99.420 S C12 HOUSE SIZE MA 90	1 19 ostics Fit 97.452 C13	20.0 0.5 173.3 for Ur Resid 1.968 RICE 152	94 77 USUA Std Re	0.594 Observ esid 2.07 R	C15 g nd NL (Y) 0.79	C16 Energy Usage(X 49)	SQRT(Y) 0.88882	C18 Residual 0.16585	C19 SRES 0.16585	C20	(21	C22	C23	C24	C25	C26	(27	C28	C29	
Fi	Pure Error Total Fits and Diagn Obs Purity 9 99.420 1 C12 HOUSE SIZE MA 90 44	1 19 ostics Fit 97.452 C13 ARKET PI	20.4 0.3 173.3 for Ur Resid 1.968 RICE 152 42	94 77 USUA Std Re	0.594 Observ esid 2.07 R	C15 2 nd NL (Y) 0.79 2.70	C16 Energy Usage(X 49) 58;	SQRT(Y) 0.88882 1.64317	C18 Residual 0.16585 0.62315	C19 SRES 0.16585 0.62315	C20	(21	C22	C23	C24	C25	C26	Q1	C28	C29	
Fi	Pure Error Total Fits and Diagn Obs Purity 9 99.420 5 C12 HOUSE SIZE MA 90 44 46	1 19 oostics Fit 97.452 C13 NRKET PI	20.4 0.5 173.3 for Ur Resid 1.968 s RICE 152 42 157	94 77 USUA Std Re	0.594 Observ esid 2.07 R	C15 p nd NL (Y) 0.79 2.70 3.64	C16 Energy Usage(X 49) 582 1156	SQRT(Y) 0.88882 1.64317 1.90788	C18 Residual 0.16585 0.62315 -0.15615	C19 SRES 0.16585 0.62315 -0.15615	C20	C21	C22	C23	C24	C25	C26	(27	C28	C29	
Fi	Pure Error Total Fits and Diagn Obs Purity 9 99.420 S C12 HOUSE SIZE MAA 90 44 46 90	1 19 ostics Fit 97.452 C13 ARKET PI	20.4 0.173.3 for Ur Resid 1.968 x RICE 152 42 157 113	94 77 USUA Std Re	0.594 Observ esid 2.07 R	C15 2 nd NL (Y) 0.79 2.70 3.64 4.73	C16 Energy Usage(X 49) 58: 115/ 99:	SQRT(Y) 0.88882 1.64317 1.90788 2.17486	C18 Residual 0.16585 0.62315 -0.15615 0.31234	C19 SRES 0.16585 0.62315 -0.15615 0.31234	C20	C 21	C22	C23	C24	C25	C26	C27	C28	C29	
Fi	Pure Error Total Fits and Diagn Obs Purity 9 99:420 1 C12 HOUSE SIZE MA 400 44 46 90 150	1 19 Oostics 1 97.452 C13 NRKET PI	20.4 0.: 173.3 for Ur Resid 1.968 x RICE 152 42 157 113 573	94 77 USUA Std Re	0.594 Observ esid 2.07 R	C15 g nd NL (Y) 2.70 3.64 4.73 9.50	C16 Energy Usage(X 49; 588; 115 99; 218;	SQRT(Y) 0.88882 1.64317 1.90788 2.17486 3.08221	C18 Residual 0.16585 0.62315 -0.15615 0.31234 -0.84939	C19 SRES 0.16585 0.62315 -0.15615 0.31234 -0.84939	C20	21	C22	C23	C24	C25	C26	Q1	C28	C29	
Fi	Pure Error Total Fits and Diagn Obs Purity 9 99,420 5 C12 HOUSE SIZE MA 90 44 46 90 150 90 44	1 19 Oostics 1 97.452 C13 NRKET PI	20.4 0.: 173.3 for Ur Resid 1.968 n RICE 152 42 157 113 573 202	94 77 Std Ro 2 C14	0.594 Observ esid 2.07 R Opeman	C15 2 nd NL (Y) 0.79 2.70 3.64 4.73	C16 Energy Usage(X 49; 588; 115 99; 218;	SQRT(Y) 0.88882 1.64317 1.90788 2.17486 3.08221	C18 Residual 0.16585 0.62315 -0.15615 0.31234	C19 SRES 0.16585 0.62315 -0.15615 0.31234	C20	Q1	C22	C23	C24	C25	C26	Q1	C28	C29	

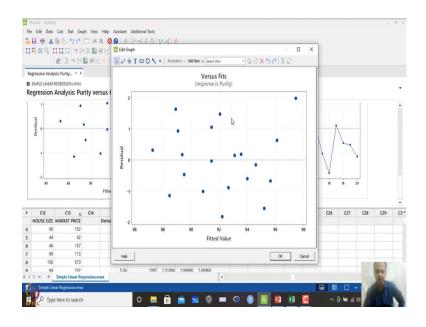
(Refer Slide Time: 22:29)



(Refer Slide Time: 22:30)

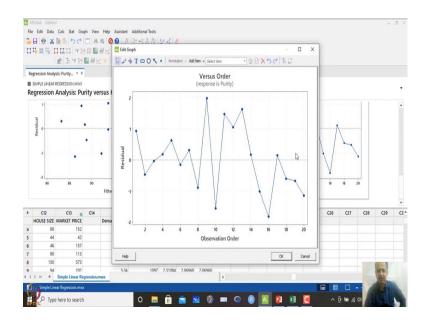


(Refer Slide Time: 22:31)



So, when you draw these two graphs also what you will observe is that this is one of the graph which will more or less you see residuals with fitted value is more or less random over here on the 0 line. So, there is no heteroscedasticity as such, but we can prove that by Breusch pagan test like that. So, we can take the residuals and we can test that one and that is available in R what I have told earlier also. So, this can be done.

(Refer Slide Time: 22:49)

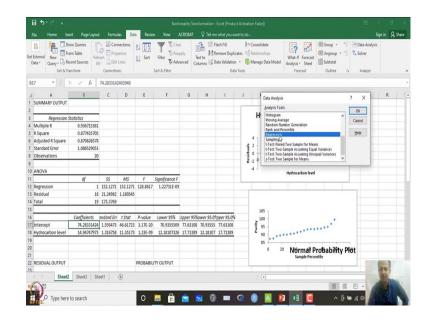


And if there is any pattern that we observe in this data set like that, but it is seems to be random. So, in this case autocorrelation may also be come out to be negligible over here.

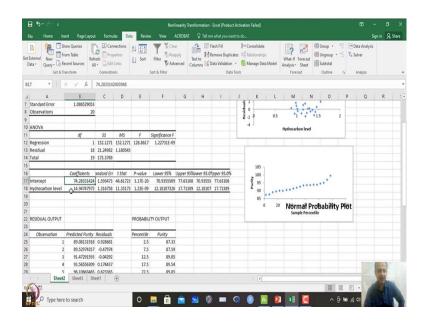
So, this is also not. So, all the assumptions we can check, but what we are seeing is that at least preliminary assumptions are satisfactory and this if you go to this book you will find that all the assumptions are true in this case and so we can generalize the equation.

That means, this equation that we have written over here can be generalized and the values that you are getting 74.28 and 14.95 of the β_0 and β_1 estimation that is also possible in excel. So, if you go to excel over here and you run the regression equation and which I have done earlier over here save this one and what happens is that the values that you see constant values of intercept that you see over here by using excel simple excel.

(Refer Slide Time: 23:37)



(Refer Slide Time: 23:45)



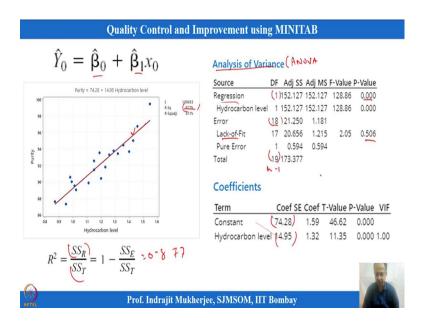
So, what I do data and I have data analysis tools and in this case you have regression analysis tools and based on that you can do the regression analysis and you can get the coefficients and corresponding P-values are also indicated over here regression is significant or not.

So, this is more than three place of decimal over here what you can see 10 to the power minus 9 and it is showing it will give you the values that is a this is not 0 what we can see over here and all the MINITAB reports this has 0. So, this MINITAB does not report more than three plus or decimal over here ok. But excel reports beyond three place of decimal over here so this is true and the values are exact what you see intercepts is 74.28 and MINITAB also has given 74.28.

So, if you go to the original value 74.28 74.28 14.95 is the β_1 estimation that you see and excel also shows the β_1 estimation 14.95 approximately that is also over here what you see 14.95. So, regression is also possible you can verify that one if you want to see what is the exact P-values like that and MINITAB is not giving you that because it does not go beyond 3 place of decimal.

So, in that case what you can do is that you can do it in excel and see that one or you can transfer this to R and do the analysis and you will get the exact P-values what is what comes out from the analysis like that ok. So generalized equation that we can use is 74.28 14.95 and this is hydrocarbon level that we can take ok.

(Refer Slide Time: 25:11)

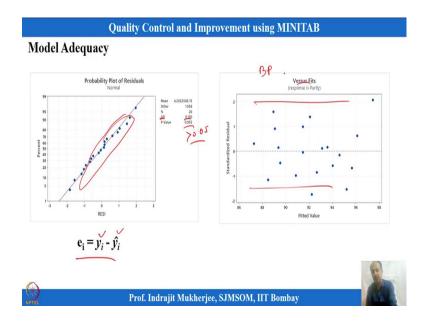


So, this is one of the example that we have seen and the data set is this is the data set and these are the graphs and equations over here. So, this is the model fit that you can see, this is the line equation that is denoted over here R square is around 0.87 7 like that. This β_0 and β_1 is estimated over here.

So, this is β_0 and this is β_1 estimation over here and this regression is found to be significant and lack of fit there is no lack of fit. So, in this case linear equation is sufficient and R square value is SS regression by SST which is approximately equals to 0.877.

What you are seeing over here 0.877 this is the value over here ok. So, this is the Anova analysis that you see analysis of variance which is the Anova analysis that you see and regression degree of freedom is 1, because 1 variable is there ok and error degree of freedom is basically n minus 2. So 20 observations are there so 20 minus 2 is18 and n minus 1 observation is over here, so this is 19 over here.

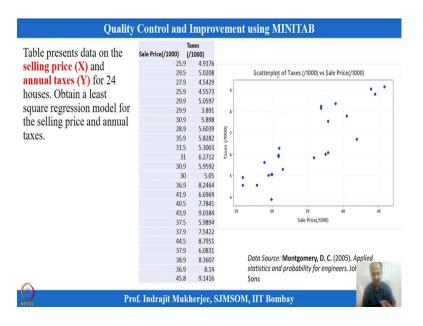
(Refer Slide Time: 26:01)



So, this is the basic interpretation of this and we can generate the errors because regression equation we have got. So, actual value and predicted value will give me residuals and residuals are probability plots are done over here Anderson darling test is done over here and P-value is greater than 0.05. So, that indicates that this is more or less normal.

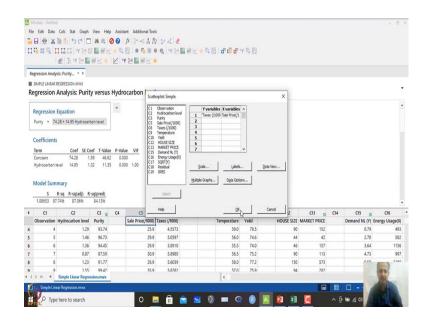
And the data set is on the line most of the points are on the line and also versus fit what do we see? There is no as such trend over here and more or less seems to be random. And also we can do the Breusch-Pagan test to confirm that one and other autocorrelation tests like Durbin Watson statistics can also be done.

(Refer Slide Time: 26:38)



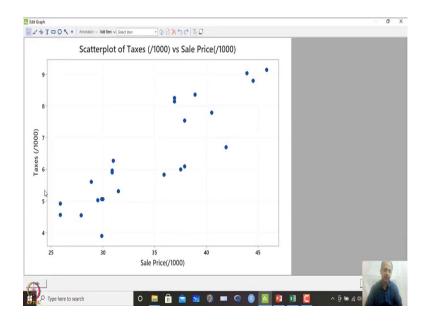
This is another example of selling price and annual taxes I will just repeat the analysis over here, so that one more examples we are taking over here. So, let me go to that examples and try to figure out how we have done this one. So, just repetition one more examples where we want to see that whether annual tax is related with the selling price of the house like that ok.

(Refer Slide Time: 27:10)



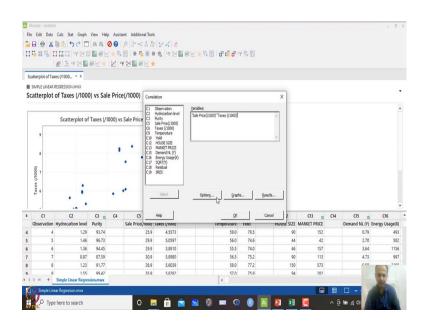
And we want to obtain the estimation relationship like that. So, what we do is that first step what we have to do is that this is the data set is given over here, this is the second one c 5 and c 6 and this data is taken from again Montgomery's applied statistics and probability and I want to see the relationship of tax with sales price and tax will be y and sales price will be x over here. So, this is dependent on the sales price. So, what we will do is that we will we can plot the scatter diagram and try to see whether linear relationship is we can think of.

(Refer Slide Time: 27:37)



So, this y will be tax and x will be sales price and we can draw the graphical scatter plot over here and what you see in the scatter plot is more or less again positive relationship that exists over here and I can confirm what is the correlation coefficient between these 2.

(Refer Slide Time: 27:48)



(Refer Slide Time: 27:52)

M	initab - Uni	titled															- 8	
File	Edit D	ata C	alc Stat Graph	View Help	p Assist	tant Additional Tool	5											
•	9	X 🗈	6 50 0	14 14	00	fx 2=== 1 #	2260											
1	1 II II		L L Y > 6	Bek	+ 12	图 · # 档册 #	···	86	* 24	d o c o	Y 构图							
						YMERK												
			ce(/10 * ×															
-			SRESSION MWX	00) T		200												,
.0	rrelatio	on: S	ale Price(/10	00), Tax	es (/10	000)												
																		1
						0), Taxes (/1000)											
			959	CI for Pear	rson Cor	relation												
	- 6							٦										
							•											
	8				1													
						•												
	Taxes (/1000)						3											
	5																	
	e e					•												
																		ļ
1	CI		C2	C3 n	C4	CS	C6	C7	C8	C9	C10	C11	C12	C13 m	C14	C15 m	C16	i
	Observati	ion Hy	ydrocarbon level	Purity		Sale Price(/1000)	Taxes (/1000)			Temperature	Yeild		HOUSE SIZE	MARKET PRICE		Demand NL (Y)	Energy Usage()	j
		4	1.29	93.74		25.9	4.5573			59.0	78.5		90	152		0.79	49	3
		5	1.46	96.73		29.9	5.0597			56.0	74.6		44	42		2.70	58	2
		6	1.36	94.45		29.9	3.8910			55.5	74,0		46	157		3.64	115	
		7	0.87	87.59		30.9	5.8980			56.5	75.2		90	113		4.73	99	
		8	1.23	91.77		28.9	5.6039			58.0	77.2		150			0.50		*
		9	1.55 imple Linear Regres	99,42		35.9	5.8282			57.0	75.9		94	202			N.	
	2 1 1	_		sion.mwx					_	4	_	_				-	distant.	
μ	Simple	Linear	Regression.mwx														Contraction of	
Ś	01	vpe h	ere to search			0	i 🔒 I	-	31 🕅			die			~ ē	i 90 @ 40		į

(Refer Slide Time: 27:52)

El.		d														-	Ð
		Calc Stat Graph															
		B 6 50 0															
11	i 🗆 🖷 🛛	III V >6	Bek	* 聪	🗄 i 🕸 🎁 🚻 🕯	青いと	Bek	* 4	i d" ti d" i	Y马图							
		@ IYN	sk*	12 4	(MESK)	t.											
Cor	relation: Sale	Price(/10 × ×															
8 9	IMPLE UNFAR	REGRESSION MWX															
-		: Sale Price(/10	00), Taxe	es (/10	00)												
1	Method																
	Correlation typ																
1	Number of ro	ws used 24															
	p: pairwise Pe	terson correlation															
(orrelation																
1	onclation																
		Cale				N											
		Sale Price(/1000)				3											
	Faxes (/1000)					Q											
	Faxes (/1000)	Price(/1000)				3											
		Price(/1000)	15			Q											
F	Pairwise Pe C1	Price(/1000) 0.876 Parson Correlation C2	C3 👩	C4	G	C6	C7	C8	C9	C10	C11	C12	C13	5 C14	C15 g		
F	Pairwise Pe C1	Price(/1000) 0.876 arson Correlation C2 Hydrocarbon level	C3 Z Purity	C4	Sale Price(/1000)	C6 Taxes (/1000)	C7	C8	Temperature	Yeild	C11	HOUSE SIZE	MARKET PRIC	E	Demand NL (Y)	Energy Usage	
F	Dairwise Pe C1 Observation 4	Price(/1000) 0.876 arson Correlation C2 Hydrocarbon level 1.29	C3 g Purity 93.74	C4	Sale Price(/1000) 25.9	C6 Taxes (/1000) 4.5573	C7	C8	Temperature 59.0	Yeild 78.5	C11	HOUSE SIZE	MARKET PRIC	2E 52	Demand NL (Y) 0.79	Energy Usage	493
F	CI CI Observation 4 5	Price(/1000) 0.876 arson Correlation C2 Hydrocarbon level 1.29 1.46	C3 2 Purity 93.74 96.73	C4	Sale Price(/1000) 25.9 29.9	C6 Taxes (/1000) 4.5573 5.0597	67	C8	Temperature 59.0 56.0	Yeild 78.5 74.6	C11	HOUSE SIZE 90 44	MARKET PRIC	2E 52 12	Demand NL (Y) 0.79 2.70	Energy Usage	493 582
F	C1 C1 Observation 4 5 6	Price(/1000) 0.876 arson Correlation C2 Hydrocarbon level 1.29 1.46 1.36	C3 2 Purity 93.74 96.73 94.45	C4	Sale Price(/1000) 25.9 29.9 29.9	C6 Taxes (/1000) 4.5573 5.0597 3.8910	C7	C8	Temperature 59.0 56.0 55.5	Yeild 78.5 74.6 74.0	C11	HOUSE SIZE 90 44 46	MARKET PRIC	2E 52 52 57	Demand NL (Y) 0.79 2.70 3.64	Energy Usage	493 582 156
F	C1 C1 Observation 4 5 6 7	Price(/1000) 0.876 arson Correlation C2 Hydrocarbon level 1.29 1.46 1.36 0.87	C3 2 Purity 93.74 96.73 94.45 87.59	C4	Sale Price(/1000) 25.9 29.9 29.9 30.9	C6 Taxes (/1000) 4.5573 5.0597 3.8910 5.8980	C7	C8	Temperature 59.0 56.0 55.5 56.5	Yelid 78.5 74.6 74.0 75.2	C11	HOUSE SIZE 90 44 46 90	MARKET PRIC	2E 32 32 57 13	Demand NL (Y) 0.79 2.70 3.64 4.73	Energy Usage	493 582 156 997
	CI CI Observation 4 5 6 7 8	Price(/1000) 0.876 arson Correlation C2 Hydrocarbon level 1.29 1.46 1.36 0.87 1.23	C3 g Purity 93.74 96.73 94.45 87.59 91.77	C4	Sale Price(/1000) 25.9 29.9 29.9 30.9 28.9	C6 Taxes (/1000) 4.5573 5.0597 3.8910 5.8980 5.6039	C7	C8	Temperature 59.0 56.0 55.5 56.5 56.5 58.0	Yeild 78.5 74.6 74.0 75.2 77.2	C11	HOUSE SIZE 90 44 46 90 150	MARKET PRIC 15 15 15	2E 52 57 13 73	Demand NL (Y) 0.79 2.70 3.64	Energy Usage	493 582 156
	C1 C1 Observation 4 5 6 7	Price(/1000) 0.876 arson Correlation C2 Hydrocarbon level 1.29 1.46 1.36 0.877 1.23 1.55	C3 g Purity 93.74 96.73 94.45 87.59 91.77 99.42	C4	Sale Price(/1000) 25.9 29.9 29.9 30.9	C6 Taxes (/1000) 4.5573 5.0597 3.8910 5.8980	C7	C8	Temperature 59.0 56.0 55.5 56.5 58.0 57.0	Yelid 78.5 74.6 74.0 75.2	C11	HOUSE SIZE 90 44 46 90	MARKET PRIC	2E 52 57 13 73	Demand NL (Y) 0.79 2.70 3.64 4.73	Energy Usage	493 582 156 997
	Ct Observation 4 5 6 7 7 8 8 9 H + 2	Price(/1000) 0.876 arson Correlation C2 Hydrocarbon level 1.29 1.46 0.87 1.23 1.45 Simple Linear Regree	C3 g Purity 93.74 96.73 94.45 87.59 91.77 99.42	C4	Sale Price(/1000) 25.9 29.9 29.9 30.9 28.9	C6 Taxes (/1000) 4.5573 5.0597 3.8910 5.8980 5.6039	C7	C8	Temperature 59.0 56.0 55.5 56.5 56.5 58.0	Yeild 78.5 74.6 74.0 75.2 77.2	C11	HOUSE SIZE 90 44 46 90 150	MARKET PRIC 15 11 11 11 55 20	2E 32 32 37 33 37 3 37 3 73	Demand NL (Y) 0.79 2.70 3.64 4.73 0.50	Energy Usage	493 583 154 993
	Ct Observation 4 5 6 7 7 8 8 9 H + 2	Price(/1000) 0.876 arson Correlation C2 Hydrocarbon level 1.29 1.46 1.36 0.877 1.23 1.55	C3 g Purity 93.74 96.73 94.45 87.59 91.77 99.42	C4	Sale Price(/1000) 25.9 29.9 29.9 30.9 28.9	C6 Taxes (/1000) 4.5573 5.0597 3.8910 5.8980 5.6039	C7	C8	Temperature 59.0 56.0 55.5 56.5 58.0 57.0	Yeild 78.5 74.6 74.0 75.2 77.2	C11	HOUSE SIZE 90 44 46 90 150	MARKET PRIC 15 11 11 11 55 20	2E 52 57 13 73	Demand NL (Y) 0.79 2.70 3.64 4.73 0.50	Energy Usage	493 582 156 997

(Refer Slide Time: 27:54)

In El Die Cell Sitz Gung New Help Ausser <																- 6
Sample 1 Sample 2 Cl 4 Cl 7		Calc Stat Graph	View Help	Assistant	Additional Tool	8										
•••••••••••••••••••••••••••••	8 👲 🕹	B 6 50 C	1 4 4 1	001	2-248	244 2										
ametalians Sale Price(/1000), Taxes (/1000) Tarest Unitable REGESSION MARK Tarest Unitable REGESSION MARK Tarest Unitable Regession correlations Sale <u>Price(/2000)</u> Tarest Unitable Regession Correlations Sample 1 Sample 2 A Correlation 95% Clife p Public Tarest Unitable Regession Correlations CC Cl cl cl cl cl p Cl Cl Cl p Cl Cl Cl p Cl Cl p Cl	16 II Th		Bek	* 時間		· 尚 · Y >>	BK	* 构图	d" c" d" .	Y英国						
Conclusions Safe Particle Model RESERSON_LINKS Conclusions Safe particle Reservation Conclusions Safe Particle Reservation Safe Conclusions Safe Safe Particle Reservation Safe Conclusions Safe																
Sumple Linkuk AldedSonumed and Linkuk AldedSonumed Terrer (17000) Sale Price(/1000), Taxes (/1000) Terrer (17000) Sale Price(/1000) Sale <u>Price(/1000)</u> 0.076 Paring Person Concelations Sample 1 Sample 2 A Concelations Sample 1 A Concelations Sample 1 A Concelations 95% Chic p Private Terrer (1700) 24 0.02% (1703 0.04%) 0.00% Concelations Sample 1 A Concelations 95% Chic p Private Terrer (1700) 24 0.02% (1703 0.04%) 0.00% Concelations 2 Concelations Sample 1 A Concelations 95% Chic p Private Terrer (1700) 24 0.02% (1703 0.04%) 0.00% Concelations 2 Concelations Sample 1 A Concelations 95% Chic p Private Terrer (1700) 24 0.02% (1703 0.04%) 0.00% Concelations 2 Concelations Sample 1 A Concelations 2 Concelations Sample 1 A Concelations S	orrelation: Sal	e Price(/10 × ×														
Sale Price(/1000), Taxes (/1000) Twenter Name Area # premer Prace Area Sale <u>Prece//0000</u> Sale <u>Prece//0000</u> Sale <u>Prece//0000</u> Sample 1 Sample 1 Sample 1 Sample 1 Sample 2 Sample 1 Sample 2 Sample 3 Sample 4 Sample 3 Sample 4 Sample 4 Sample 5 Sample 1 Sample 3 Sample 4 Sample 4 Sample 5 Sample 5 Sample 5 Sample 7																
Terreter Farson contention Sub- preverse farson contentions Sub- preverse farson contentin- stations Sub- preverse fars			00). Tax	es (/100	0)											
prevent Person Correlations Test (1000) 2.07 Pairwise Person Correlations Singhel X = Market Person Correlations <t< td=""><td></td><td></td><td></td><td></td><td>-,</td><td></td><td></td><td></td><td></td><td></td><td></td><td></td><td></td><td></td><td></td><td></td></t<>					-,											
Correlations Same Texes(1000) 0.076 Pairwise Person Correlations Sample 1 Sample 1 Sample 2 A Correlation 595 (1 for p P-toble Texes(1000) 24 0.076 (1000) Observation hydrocation level Nurly Sale Picot/0000 (Taxes (1000) CO C1 C12 C3 C4 C5 C16 Observation hydrocation level Nurly Sale Picot/0000 (Taxes (1000) Temperature Yeld HOUSS SZE MARKET PRCE Demand Ru (Y) Energy Use 4 129 9.374 22.99 5.097 5.00 726 01 132 0.79 6 13.6 64.45 22.99 3.690 5.557 2.40 133 4.47 8 12.19 17.7 22.65 5.080 5.630																
Safe Tares (1000) Safe Safe Prior (1000) Correlation Safe Prior (1000) Correlation C	p: pairwise F	rearson correlation														
Safe Tares (1000) Safe Safe Prior (1000) Correlation Safe Prior (1000) Correlation C																
Price/0000 Taxes (1/000) Data Pairwise Person Correlations Sample 1 Ample 2 N Correlation 95% Chr p Publice Privile Cl C2 C4 C5 C6 C7 C0 C0 C1 C12 C13 C4 C5 C16 Observation Hydrocation level Putry Sale Price/1000 24 0.876 (0.731, 0.845) 0.000 Observation Hydrocation level Putry C3 C6 C7 C0 C00 C11 C12 C13 C4 C5 C16 4 129 9.374 22.99 5.097 5.00 77.5 90 132 0.79 6 13.6 94.45 22.99 3.690 55.57 7.40 46 157 3.64 7 0.87 5.097 5.00 77.2 190 133 4.47 8 12.19 17.7 22.89 3.690 54.5 75.2 100 133 4.47 1	Correlation	ns														
Tares (11000) 0.276 Pairwise Pearson Correlations Sample 2 N. Carrelation 95% Cf for p. Publice Simple 1 Sample 2 N. Carrelation 95% Cf for p. Publice C1 C2 C1 g. C4 C5 g. C6 C1 C1 C2 C1 g. C4 C5 g. C6 C6 C1 C0 C1 C2 C1 g. C4 C5 g. C6 C6 C1 C10 g. C4 C5 g. C14 C5 g. C16 C10 C1 C12 C1 g. C4 C5 g. C16 C10 C1 C12 g. C14 C15 g. C16 C11 C12 C11 C14 C5 g. C16 C11 C12 C11 C14 C15 g. C16 C11 C12 C11 C12 C11 C14 C15 C16 C11 C12 C11 C14 C15 C16 C11 C12 C11																
Circuit Strangle 1 Sample 2 M. Correlations 55% Chir p. P-tuble Transf/1000 Sample 2 M. Correlation 55% Chir p. P-tuble Sale Proc/1000 Cold CSI CCI CSI p. Chi CCI CSI p. Chi Chi CSI p. Chi CSI p. Chi	-															
Sample 1 Sample 2 N. Cerrelation 95% Clf or p. P-Value Terres (7000) Same Proof/1000 2.4 Outpoint Produce 0.000 C1 C2 C1 c4 C5 c C6 C7 C8 C9 C0 C1 C1 C4 C5 c C6 Observation Hydrocarbon level Purity Sale Price/0000 Temperature Yeld HOUSE SIZE MARKET PRCC Demand NL (Y) Energy Us 4 1.29 93.44 22.9 4.5373 59.0 74.5 90 15.2 0.79 5 1.46 94.77 22.99 3.5097 55.0 74.6 44 42.2 2.70 6 1.36 94.55 2.99 3.900 55.5 74.0 46 157 3.44 7 0.07 2.89 5.6039 55.00 72 100 57.3 1.41 8 1.23 9.17 2.8.9 5.6039 55.00 7.7 150 57.3 1 H<	Taxes (/1000)	0.876														
C1 C2 C3 C4 C5 C6 C7 C8 C9 C0 C1 C2 C1 g C4 C5 g C6 Observation Hydrocarbon level Punty Sale Price/Y000 Temperature Veld HOUSE SIZE MARKET PROEL Demand NL Y0 Energy Us 4 1.29 93.74 25.9 4.5373 59.0 78.5 60 132 0.079 5 1.46 94.73 25.9 5.0597 50.0 74.6 44 4.22 2.70 6 1.36 94.45 2.29 3.390 55.5 7A.0 464 137 3.44 7 0.07 27.9 3.90 55.55 7A.0 464 137 3.44 8 1.23 91.77 2.83 5.6039 55.0 7A.0 461 137 3.44 9 1.33 91.77 2.83 5.6039 50.0 77.2 150 57.3 8 <																
Observation Hydrocarbon level Puthy Sale Price/0000 Temperature Veid HOUSE SZE MARKET PRICE Demand NL (?) Energy Use 4 1.28 9.374 23.9 4.573 50.0 71.5 60 152 0.79 5 1.46 98.73 23.9 5.5957 50.0 74.6 44 42.2 2.70 6 1.36 94.45 2.29 3.3910 55.5 74.0 464 157 3.64 7 0.37 23.9 5.8900 55.5 75.2 90 11.3 3.44 8 1.23 91.77 2.8.9 5.6039 55.0 77.2 150 57.3 6 1.32 94.7 2.8.9 5.6039 50.0 77.2 150 57.3 8 1.23 94.7 2.8.9 5.6039 50.0 77.2 150 57.3 9 1.3.4 5.17/7 7.7 44 2.7 7.7	Taxes (/1000,	Sale Price(/1000)	24	.876 (0.73	1, 0.945) 0.0											
4 129 93,74 259 4,5573 590 765 90 152 0,79 5 1,46 96,73 259 5,5697 5,60 74,6 44 42 2,70 6 1,36 94,45 239 3,8910 555 74,0 44 127 3,64 7 0,07 87,59 30,0 5,669 55,5 75,2 90 113 4,75 8 1,23 91,77 28,9 5,6699 55,0 72,2 150 573 0 9 1,55 9,42 35,9 5,609 55,0 72,0 130 4,75 9 1,55 9,60,0 75,0 72,0 150 573 0 14 14 14 2,90 14						6										
5 1.46 64.73 259 5.097 560 74.6 44 42 2.70 6 1.38 64.45 239 3.2910 53.5 74.0 46 157 3.64 7 0.07 87.99 3.09 5.556 75.2 90 113 4.77 6 1.22 9.1.77 25.8 5.6039 54.0 77.2 150 57.3 90 1 1.55 94.07 3.54 5.002 57.2 90 113 4.77 6 1.52 94.77 2.58 5.6039 57.0 90 153 4.0 1.5 94.77 2.58 5.6039 57.0 90 150 57.3 90 1.8 1.56 94.77 150 57.3 90 150 57.3 90 150 57.3 90 150 150 160 160 160 160 160 160 160 170 1	CI	(2	C3 👩	C4	CS	C6	C7	C8			C11			C14		
6 13.6 94.45 25.9 3.090 55.5 7.6.0 46 15.7 3.6.4 7 0.87 65.9 30.0 5.6500 96.5 75.2 90 11.3 4.7.7 8 12.2 91.71 22.6 5.009 55.0 77.2 150 57.3 150 5.7.7 150 57.3 150<	C1 Observation	C2 Hydrocarbon level	C3 g Purity	C4	C5 Sale Price(/1000)	C6 Taxes (/1000)	C7	C8	Temperature	Yeild	C11	HOUSE SIZE	MARKET PRICE	C14	Demand NL (Y)	Energy Usage
7 0.87 8759 30.9 5.8900 545 752 90 113 4.73 8 123 91.77 28.9 5.6039 58.0 77.2 150 57.3 0 9 154.5 94.0 95.0 57.0 75.0 94.0 17.0 19.0 17.0 19.0 17.0 19.0 17.0 19.0 10.0	C1 Observation	C2 Hydrocarbon level 1 1.29	C3 purity 93.74	C4	C5 Sale Price(/1000) 25.9	C6 Taxes (/1000) 4.5573	C7	C8	Temperature 59.0	Yeild 78.5	C11	HOUSE SIZE	MARKET PRICE 152	C14	Demand NL (Y) 0.79	Energy Usage
8 1.23 91.77 25.9 56.00 58.0 77.2 150 57.3 b # + Single Liner Regression.mx 4 4 20 4 20 4 20 4 20 4 20 4 20 4 20 4 20 4 20 4 20 4 20 4 20 4 20 4 20 4 20 4 20 4 20 4 4 20 4	C1 Observation	C2 Hydrocarbon level 1.29 5 1.46	C3 2 Purity 93.74 96.73	C4	C5 Sale Price(/1000) 25.9 29.9	C6 Taxes (/1000) 4.5573 5.0597	C7	C8	Temperature 59.0 56.0	Yeild 78.5 74.6	C11	HOUSE SIZE 90 44	MARKET PRICE 152 42	C14	Demand NL (Y) 0.79 2.70	Energy Usagel 4 5
	C1 Observation 4 5	C2 Hydrocarbon level 4 1.29 5 1.46 5 1.36	C3 p Purity 93.74 96.73 94.45	C4	C5 Sale Price(/1000) 25.9 29.9 29.9	C6 Taxes (/1000) 4.5573 5.0597 3.8910	C7	C8	Temperature 59.0 56.0 55.5	Yeild 78.5 74.6 74.0	C11	HOUSE SIZE 90 44 46	MARKET PRICE 152 42 157	C14	Demand NL (Y) 0.79 2.70 3.64	Energy Usage 4 5 11
1 > + + Simple Linear Regression.max	C1 Observation 2 2 2 2 2 2 2 2 2 2	C2 Hydrocarbon level 5 1.29 5 1.46 5 1.36 7 0.87	C3 p Purity 93.74 96.73 94.45 87.59	C4	C5 Sale Price(/1000) 25.9 29.9 29.9 30.9	C6 Taxes (/1000) 4.5573 5.0597 3.8910 5.8980	C7	C8	Temperature 59.0 56.0 55.5 56.5	Yeild 78.5 74.6 74.0 75.2	C11	HOUSE SIZE 90 44 46 90	MARKET PRICE 152 42 157 113	C14	Demand NL (Y) 0.79 2.70 3.64 4.73	Energy Usagel 4 5 11 9
Smpliner Reyesion.mx 🖩 🖩 🗆	C1 Observation 4 5 6 7 7 8	C2 Hydrocarbon level 4 1.29 5 1.46 5 1.36 7 0.87 8 1.23	C3 2 Purity 93.74 96.73 94.45 87.59 91.77	C4	C5 Sale Price(/1000) 25.9 29.9 29.9 30.9 28.9	C6 Taxes (/1000) 4.5573 5.0597 3.8910 5.8980 5.6039	C7	C8	Temperature 59.0 56.0 55.5 56.5 56.5 58.0	Yeild 78.5 74.6 74.0 75.2 77.2	C11	HOUSE SIZE 90 44 46 90 150	MARKET PRICE 152 42 157 113 573	C14	Demand NL (Y) 0.79 2.70 3.64 4.73	Energy Usagel 4 5 11 9
	C1 Observation 4 5 6 7 8 8 8	C2 Hydrocarbon level 4 1.29 5 1.46 5 1.36 7 0.87 7 1.23 3 1.23 9 1.55	C3 2 Purity 93.74 96.73 94.45 87.59 91.77 90.42	C4	C5 Sale Price(/1000) 25.9 29.9 29.9 30.9 28.9	C6 Taxes (/1000) 4.5573 5.0597 3.8910 5.8980 5.6039	C7	C8	Temperature 59.0 56.0 55.5 56.5 58.0 58.0 57.0	Yeild 78.5 74.6 74.0 75.2 77.2	C11	HOUSE SIZE 90 44 46 90 150	MARKET PRICE 152 42 157 113 573	CI4	Demand NL (Y) 0.79 2.70 3.64 4.73	Energy Usage(4) 5) 11: 9)
	Cl Observation 4 5 6 7 8 8 8 8 8 8 8 8 8 8 8 8 8 8 8 8 8 8	C2 Hydrocarbon level 4 1.29 5 1.46 5 1.36 7 0.87 3 1.23 9 1.55 Simple Linear Regre	C3 2 Purity 93.74 96.73 94.45 87.59 91.77 90.42	C4	C5 Sale Price(/1000) 25.9 29.9 29.9 30.9 28.9	C6 Taxes (/1000) 4.5573 5.0597 3.8910 5.8980 5.6039	C7	C8	Temperature 59.0 56.0 55.5 56.5 58.0 58.0 57.0	Yeild 78.5 74.6 74.0 75.2 77.2	C11	HOUSE SIZE 90 44 46 90 150	MARKET PRICE 152 42 157 113 573 207		Demand NL (Y) 0.79 2.70 3.64 4.73 0.50	Energy Usage(4) 5) 11: 9)
📲 💭 Type here to search 🛛 🖸 🧮 📩 👘 🔂 🚱 💷 💿 🔞 📠 🛐 🖬 📕 🖌 🎯 🖦 🖉 🕸	Cl Observation 4 5 6 7 8 8 8 8 8 8 8 8 8 8 8 8 8 8 8 8 8 8	C2 Hydrocarbon level 4 1.29 5 1.46 5 1.36 7 0.87 3 1.23 9 1.55 Simple Linear Regre	C3 2 Purity 93.74 96.73 94.45 87.59 91.77 90.42	C4	C5 Sale Price(/1000) 25.9 29.9 29.9 30.9 28.9	C6 Taxes (/1000) 4.5573 5.0597 3.8910 5.8980 5.6039	C7		Temperature 59.0 56.0 55.5 56.5 58.0 57.0 4	Yeild 78.5 74.6 74.0 75.2 77.2	C11	HOUSE SIZE 90 44 46 90 150	MARKET PRICE 152 42 157 113 573 207		Demand NL (Y) 0.79 2.70 3.64 4.73 0.50	Energy Usage(49 51 11

And I can go to correlation and I can just check sales price and taxes over here and I click ok and what I see is that P-value is significant over here and 0.876 is the more or less approximation that we are getting over here. Correlation coefficient this is positive; that means, there is a positive relationship between sales and tax that is relationship that we have.

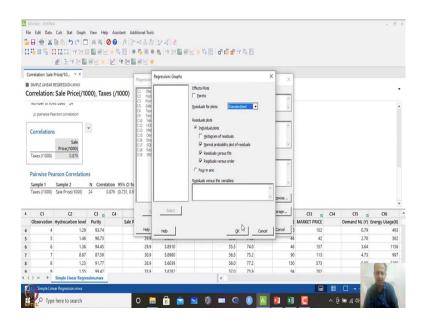
(Refer Slide Time: 28:16)

	1 2 1 5	© 6 5 २ 0 1 द्व 4 ⊠ 6 @ 1 £ 4 ⊠ ■	. sk	* 4	1 + I	6 ₩ #		: • • • • • • • • • • • • • • • • • • •	1 c^a c<mark>a</mark> c^a · y	马图							
C		Price(/10 * X		2	-												
8	IMPLE UNEAR	REGRESSION MWX			Re	gression						×					
C	rrelation	Sale Price(/10	00), Tax	es (/100	0)	1 Observ	ation arbon level	Responses:				_					•
	vomuer or ros	10 UDEU 24			C	5 Sale Pri	ce(/1000)	"Taxes (/ 1000)"				~					
	p: pairwise Pe	arson correlation			000	Temper		Continuous predictors:									
Г			¥		C	12 HOUSE		'Sale Price(/1000)				^					
	orrelation	5			C	15 Deman		I									
		Sale Price(/1000)			C	17 SQRTO	Usage(X) Y)					¥					
	Faxes (/1000)	0.876			C	18 Residu 19 SRES	al	Categorical predictors:									
												~					
-	Pairwise Pe	arson Correlation	15														
	Pairwise Pea Sample 1	arson Correlation Sample 2	N Correl	ation 959	6 CI fo							~					
	Sample 1	Sample 2	N Correl	ation 959 0.876 (0.7				Model.	. Optiogs	Coging	. Stepwise	~					
	Sample 1	Sample 2	N Correl			Selec	t	Model		Coging Besuits.			C13 12	C14	CIS	5 C16	
+	Sample 1 Faxes (/1000) C1	Sample 2 Sale Price(/1000)	N Correli 24 0	0.876 (0.7 C4		Selec	t						C13 g MARKET PRICE	C14	C15 Demand NL (Y		
+	Sample 1 Faxes (/1000) C1	Sample 2 Sale Price(/1000) C2	N Correll 24 0 C3 g Purity 93.74	0.876 (0.7 C4	31, 0.9	Help	: :			Besults. QK				C14) Energy Us	iage(X) 493
+	Sample 1 Faxes (/1000) C1	Sample 2 Sale Price/1000) C2 Hydrocarbon level 1.29 1.46	N Correl 24 0 Purity 93.74 96.73	0.876 (0.7 C4	31, 0.9	Help 29.9	5.059	<u>Y</u> aldation	<u>Graphs</u> 56.0	Besuits. QK 74.6	. Storage.		MARKET PRICE 152 42	C14	Demand NL (Y 0.7 2.7)	9 0	493 582
4 4 5 6	Sample 1 Faxes (/1000) C1	Sample 2 Sale Price(/1000) C2 Hydrocarbon level 1.29 1.46 1.36	N Correli 24 0 Purity 93.74 96.73 94.45	0.876 (0.7 C4	31, 0.9	Help 29.9 29.9	5.059		<u>Graphs</u> 56.0 55.5	<u>B</u> esuits. <u>QK</u> 74.6 74.0	. Storage.		MARKET PRICE 152 42 157	C14	Demand NL (Y 0.7 2.7 3.6) Energy Us 9 0 4	age(X) 493 582 1156
4 4 5 6 7	Sample 1 Faxes (/1000) C1	Sample 2 Sale Price/1000) C2 Hydrocarbon level 1.29 1.46 1.36 0.87	N Correl. 24 C3 2 Purity 93.74 96.73 94.45 87.59	0.876 (0.7 C4	31, 0.9	Help 29.9 29.9 30.9	5.059 3.891 5.898	 	graphs 56.0 55.5 56.5	Besuits. QK 74.6 74.0 75.2	. Storage.		MARKET PRICE 152 42 157 113	C14	Demand NL (Y 0.7 2.7 3.6 4.7	 i) Energy Us. 9 0 4 3 	493 582 1156 997
4 5 6 7 8	Sample 1 Faxes (/1000) C1	Sample 2 Sale Price/1000) C2 Hydrocarbon level 1.29 1.46 1.36 0.87 1.23	N Correll 24 0 Purity 93.74 96.73 94.45 87.59 91.77	0.876 (0.7 C4	31, 0.9	Help 29.9 29.9 30.9 28.9	5.059 3.891 5.898 5.603		<u>Graphs</u> 56.0 55.5 56.5 58.0	<u>R</u> esults. <u>QK</u> 74.6 74.0 75.2 77.2	. Storage.		MARKET PRICE 152 42 157 113 573	C14	Demand NL (Y 0.7 2.7 3.6	 i) Energy Us. 9 0 4 3 	age(X) 493 582 1156
4 4 5 6 7	5ample 1 Faxes (/1000) C1 Observation 4 5 6 7 8 9	Sample 2 Sale Price/1000) C2 Hydrocarbon level 1.29 1.46 1.36 0.87	N Correll 24 93.74 96.73 94.45 87.59 91.77 99.42	0.876 (0.7 C4	31, 0.9	Help 29.9 29.9 30.9	5.059 3.891 5.898		graphs 56.0 55.5 56.5	Besuits. QK 74.6 74.0 75.2	. Storage.		MARKET PRICE 152 42 157 113 573	C14	Demand NL (Y 0.7 2.7 3.6 4.7	 i) Energy Us. 9 0 4 3 	493 582 1156 997
4 5 6 7 8	5ample 1 Гахев (/1000) Observation 4 5 6 7 8 8 9 9 н +	Sample 2 Sale Price/1000) C2 Hydrocarbon level 1.29 1.46 1.36 0.087 1.23 1.55	N Correll 24 93.74 96.73 94.45 87.59 91.77 99.42	0.876 (0.7 C4	31, 0.9	Help 29.9 29.9 30.9 28.9	5.059 3.891 5.898 5.603		graphs 56.0 55.5 56.5 58.0 57.0	<u>R</u> esults. <u>QK</u> 74.6 74.0 75.2 77.2	. Storage.		MARKET PRICE 152 42 157 113 573	C14	Demand NL (Y 0.7 2.7 3.6 4.7	 i) Energy Us. 9 0 4 3 	493 582 1156 997

(Refer Slide Time: 28:25)

6 5	linitab - Untitle	d																× 8	×
		Calc Stat Graph																	
2	🖯 👲 🕺	00 500	14.44	🖉 🕜 🤅 fx	3-24.89	2 2 2													
	16 II II II	1 1 1 1 Y 1 Y	9886	(★呜囵	i # 🖧 🏦 #	告 ₩≥	BK *	時間 6	🐻 🗗 -Y	英国									
		<u>@</u> ± Υ ≥∎	BK*	12 Y	×∎øk*														
C	rrelation: Sale	Price(/10 * ×			-														
Ħ	SIMPLE UNEAR	REGRESSION MWX			Regression							×							
C	orrelation	: Sale Price(/1	000), Tax	es (/1000)) C1 Obser C2 Hydro	Regression: Stor					×	_							'
	Number of ru	WS USED 24			CS Sale P		eye				^	Ô						6	į.
	p: pairwise P	earson correlation			C6 Taxes	E E		Cogff											
Ē					C10 Yeld	E Besiduals		☐ Design	Шаріх			-							
	Correlation	IS	*		C13 MARK	14 Sraupa orseo													
		Sale			C16 Energ	C Deleted residu	als												
		Price(/1000)			C18 Resid	□ Leveragests													
	Taxes (/1000)	0.876				Cook's distance	e					_							
						□ DRIES						^							
		earson Correlatio																	
	Sample 1	Sample 2		ation 95%								v							
	Taxes (/1000)	Sale Price(/1000)	24	0.876 (0.731	, 0.9							stepwise							
_		1				Help			OK	1 0	incel	Storage		_			_		1
+	CI	C2	C3 g	C4		nep			- 20		110	alte alte	C13	2	C14		8	C16	1
4	Observation	Hydrocarbon level 1.29		S	ale P Helo	1				QK.	1	Cancel	MARKET PI	152		Demand NL (Y 0.7		gy Usage(X) 493	
5	4	1.46			29.9	5.0597			56.0	74.6		Carko 4		42		2.7		582	
6	6				29.9	3.8910			55.5	74.0		46		157		3.6		1156	
7	7				30.9	5.8980			56.5	75.2		90		113		4.7		997	
8	8				28.9	5.6039			58.0	77.2		150		573		0.6		2100	
9	9	1.55	99,42		35.9	5.8282			57.0	75.9		9.		202				-	
14 1	рн +	Simple Linear Regr	ession.mwx					4										2	
1	Simple Lin	ear Regression.mwx													=		1	44	
Y	O Two	e here to search			0		- 1	()	0		de				^ ñ	900 A 40		1	
NP	TEL	e mere to search							100			- e-			5			1	

(Refer Slide Time: 28:31)



Now, everything is fine so I want to predict the equation. So, I will go to regression and regression over here fit regression models over here and instead of this purity we will change this one to tax and instead of continuous variable I will sales purity I will use over here.

And I can store the residual again over here standardized residual and I click that will be saved at the end and then I will click on the graph. So, I can see residual plots I can also see normal plot of the residual and residual for plot may be standardized residual I want to check and then I will do ok.

(Refer Slide Time: 28:41)

Ter lin Dur Cir Sur Gap Vier Hoj Aussen Addeou Dub 다 응 한 가 가 가 가 가 가 가 가 가 가 가 가 가 가 가 가 가 가
113 目前 [1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1
15 目前 日日日 11 11 12 11 12 12 12 12 12 12 12 12 12
Image:
Impression Analysis: Taxes (1/000) versus Sale Price(/1000) Regression Equation Trem: Color Stoce 1 Stoce 1 T-Value Pride Impression State Price(/1000) Coefficients Impression State Price(/1000) Coefficients Impression State Price(/1000) Impression State Price(/1000) Coefficients Impression State Price(/1000) Coefficients Impression State Price(/1000) State Price(/1000) Coefficients Impression State Price(/1000)
Sub-Lunck REGRESSON.MAX Regression Analysis: Taxes (/1000) versus Sale Price(/1000) Regression Equation Term Col Sice Price(/1000) Coefficients Term Col Sice Price(/1000) Coefficients
Regression Analysis: Taxes (/1000) versus Sale Price//1000) Regression Equation Trees //1000 * -1.554 + 0.208 Sale Price//1000 Coefficients Quert Coefficients Quert Coefficients
Regression Equation Tame: Cole 5 St Coll 7 Value P-Value Ver Control 0: 02208 Six Prior(1000) Coefficients December Coll 3: 0208 Six Prior(1000) Model Summary S ** R*sq. Fsq. Baglight R: spg.pred) 0.707002 78,7719 Col cl n CC Cl 0 n CC
Tares (1/100) • 1/244 + 0.2208 Safe Proce/(1/000) Term Coeff Scient T Value Produce Vie Commune Term Coeff Scient T Value Produce Vie Commune 3/6 P-S2 0.000 1.00 Model Summary Cl C2 C1
C1 C2 C3 c4 C5 C6 c7 C8 C9 C10 C11 C12 C13 c14 C55 c16 Observation hydrocarbon level Purity Sale Prior/M00 Temperature Veld HOUSE SZE MARKET PRCE Demand M:(1) Energy Us 4 1.29 90.74 25.59 4.5573 59.0 78.5 60 152 0.79 5 1.64 56.79 25.99 5.5597 54.00 78.5 44 4.22 2.79
4 1.29 93.74 25.9 4.5373 59.0 78.5 90 152 0.79 5 1.46 96.73 25.9 5.0597 56.0 74.6 44 42 2.70
5 1.46 96.73 29.9 5.0597 56.0 74.6 44 42 2.70
6 1.36 94.45 29.9 3.8910 55.5 74.0 46 157 3.64
7 0.87 87.59 30.9 5.8980 56.5 75.2 90 113 4.73
8 1.23 91.77 28.9 5.6039 58.0 77.2 150 573 0.50
9 1.55 99.42 35.9 5.8282 57.0 75.9 94 202
4 D H + Simple Linear Regression.mwx

(Refer Slide Time: 28:44)

1	X	00500	1 4 4 1	00	fx 3= -= 1 3	12100										
							av	- ¥1.5		N 14. 59						
• •							00	× 94.0	0 to 0.	1 24 83						
		£ Υ⊠∎		No. of	EDBK!											
leg	ression Analy	sis: Taxes Y X														
51	IMPLE LINEAR	REGRESSION MWX														
ec	aression	Analysis: Taxe	s (/1000)	versus	Sale Price(/	1000)										
2	oemcient					,										
				a tab	100											
	erm onstant	Coef SE Co -1.584 0.9														
	ale Price(/10)				1.00											
_	1odel Sum S 0.780238 7	R-sq R-sq(adj)	R-sq(pred) 73.20%													
ASR	s	R-sq R-sq(adj) I 76.73% 75.68% Variance DF Adj SS I 44.170	73.20%	-Value P 72.56	- <u>Value</u> 0.000											
A	S 0.780238 malysis of ource legression	R-sq R-sq(adj) I 76.73% 75.68% Variance DF Adj SS I 44.170	73.20% Adj MS F 44.1702	72.56	0.000	C6 12	C7	C8	09	C10	C11	C12	C13 g	C14	C15	c16
R	S 0.780238 malysis of ource legression C1	R-sq R-sq(adj) 75.68% Variance DF Adj SS 1 44.170	73.20% Adj MS F 44.1702	72.56 C4	0.000		C7	C8	C9 Temperature	C10 Yeild	C11		C13 m MARKET PRICE	C14	Demand NL (Y)	Energy Usag
A	S 0.780238 analysis of ource degression C1 Dbservation 4	R-sq R-sq(adj) 1 DF Adj SS 1 44,170 C2 Hydrocarbon level 1.29	73.20% Adj MS F 44.1702 C3 2 Purity 93.74	72.56 C4	0.000 C5 Sale Price(/1000) 25.9	Taxes (/1000) 4.5573	C7	C8		Yeild 78.5	C11		MARKET PRICE 152	C14	Demand NL (Y) 0.79	Energy Usag
A	S 0.780238 analysis of ource legression C1 Dbservation	R-sq R-sq(adj) 1 75.68% 75.68% Variance DF Adj SS 1 44,170 C2 Hydrocarbon level 1.29 1.46 1.46	73.20% Adj MS F 44.1702 C3 2 Purity 93.74 96.73	72.56 C4	0.000 C5 Sale Price(/1000)	Taxes (/1000) 4.5573 5.0597	C7	C8	Temperature 59.0 56.0	Yeild 78.5 74.6	C11	HOUSE SIZE 90 44	MARKET PRICE 152 42	C14	Demand NL (Y) 0.79 2.70	Energy Usag
ASR	S 0.780238 inalysis of ource kegression C1 Dbservation 4 5 6	R-sq R-sq(adj) 1 16.73% 75.68% 75.68% DF Adj SS 1 44.170 C2 Hydrocarbon level 1.29 1.46 1.36 1.36 1.36 1.36	73.20% Adj MS F 44.1702 C3 2 Purity 93.74 96.73 94.45	72.56 C4	C5 Sale Price(/1000) 25.9 29.9 29.9	Taxes (/1000) 4.5573 5.0597 3.8910	C7	C8	Temperature 59.0 56.0 55.5	Yeild 78.5 74.6 74.0	C11	HOUSE SIZE 90 44 46	MARKET PRICE 152 42 157	C14	Demand NL (Y) 0.79 2.70 3.64	Energy Usag
ASR	S 0.780238 inalysis of ource degression C1 Dbservation 4 5	R-sq R-sq(adj) 1 75.68% 75.68% Variance DF Adj SS 0 1 45.170 C2 Hydrocarbon level 1.29 1.46 1.36 0.87	73.20% Adj MS F 44.1702 C3 2 Purity 93.74 96.73	72.56 C4	C5 Sale Price(/1000) 25.9 29.9	Taxes (/1000) 4.5573 5.0597 3.8910 5.8980	C7	C8	Temperature 59.0 56.0	Yeild 78.5 74.6 74.0 75.2	C11	HOUSE SIZE 90 44	MARKET PRICE 152 42	C14	Demand NL (Y) 0.79 2.70	Energy Usag
ASR	S 0.780238 inalysis of ource kegression C1 Dbservation 4 5 6	R-sq R-sq(adj) 1 16.73% 75.68% 75.68% DF Adj SS 1 44.170 C2 Hydrocarbon level 1.29 1.46 1.36 1.36 1.36 1.36	73.20% Adj MS F 44.1702 C3 2 Purity 93.74 96.73 94.45	72.56 C4	C5 Sale Price(/1000) 25.9 29.9 29.9	Taxes (/1000) 4.5573 5.0597 3.8910	C7	C8	Temperature 59.0 56.0 55.5	Yeild 78.5 74.6 74.0	C11	HOUSE SIZE 90 44 46	MARKET PRICE 152 42 157	C14	Demand NL (Y) 0.79 2.70 3.64	Energy Usag
Ass	S 0.780238 malysis of ource degression C1 Dbservation 4 5 6 7 7 8 9	R-sq R-sq(adj) 1 75.68% 75.68% DF Adj SS C2 Hydrocarbon level 1.29 1.46 1.36 0.87 1.45 1.36 1.36 1.35	73.20% Adj MS F 44.1702 Purity 93.74 96.73 94.45 87.59 91.77 90.42	72.56 C4	0.000 C5 Sale Price(/1000) 25.9 29.9 29.9 30.9	Taxes (/1000) 4.5573 5.0597 3.8910 5.8980	C7	C8	Temperature 59.0 55.0 55.5 56.5 58.0 57.0	Yeild 78.5 74.6 74.0 75.2	C11	HOUSE SIZE 90 44 46 90	MARKET PRICE 152 42 157 113	C14	Demand NL (Y) 0.79 2.70 3.64 4.73	Energy Usag
ASR	S 0.780238 inalysis of ource legression Ct Dbservation 4 5 6 6 7 8	R-sq R-sq(adj) 1 75.68% 75.68% Variance DF Adj SS 1 44.170 C2 Hydrocarbon level 1.29 1.46 0.87 1.23 1.23	73.20% Adj MS F 44.1702 Purity 93.74 96.73 94.45 87.59 91.77 90.42	72.56 C4	0.000 CS Sale Price(/1000) 25.9 29.9 29.9 30.9 28.9	Taxes (/1000) 4.5573 5.0597 3.8910 5.8980 5.6039	C7	C8	Temperature 59.0 56.0 55.5 56.5 56.5 58.0	Yeild 78.5 74.6 74.0 75.2 77.2	C11	HOUSE SIZE 90 44 46 90 150	MARKET PRICE 152 42 157 113 573	C14	Demand NL (Y) 0.79 2.70 3.64 4.73	Energy Usag

And then what will happen is that it will give me the equations, then it will say whether beta is constant is significant or not. So, what we see is that constant is not significant, but I told that researcher suggest that we should keep the constant. So, we will keep that one and sales price is significant over here; that means β_1 is significant. So, we will retain this one, so there is a positive relationship that we are seeing.

So, coefficient is positive over here ok and constant is negative that constant cannot be interpreted in regression. So, in that case physical interpretation is not possible. So, and

the R square value is 0.7673 what we have seen calculation also previous calculations like that.

(Refer Slide Time: 29:19)

-		d															- 6
	e Edit Data	Calc 1	itat Graph	View He	lp Assis	tant Additional Tool	5										
						fx 3= -2 3											
						1 + to # +		au	+ M. 8		v 16. 59						
1						YMESKY		0.0		S . D to D	1 -9 03						
		<u> </u>	1 K .	100 11. 7	12	T C B B L I											
R	egression Analy	sis: Taxes	к., * X														
Ħ	SIMPLE UNEAR	REGRESSI	ON.MWX														
R	egression	Analy	sis: Taxe	s (/1000) versi	us Sale Price(/	1000)										
	Model Sum	mary															
	s			R-sq(pred)													
	0.780238		75.68%	73.20%													
		B															
	Analysis of	Variand	e														
	Source	D	F Adj SS	AdiMS	c.Vahuo	D.Ualua											
					r vanue	P-value											
	Regression		1 44,170	44.1702	72.56	0.000											
	Sale Price(/1)	100)	1 44.170	44.1702 44.1702													
	Sale Price(/1) Error	100) 2	1 44.170 2 13.393	44.1702 44.1702 0.6088	72.56	0.000											
	Sale Price(/1)	100) 2 1	1 44.170	44.1702 44.1702	72.56 72.56	0.000											
	Sale Price(/1) Error Lack-of-Fit	2000) 2 1	1 44.170 2 13.393 7 11.573	44.1702 44.1702 0.6088 0.6808	72.56 72.56	0.000											
	Sale Price(/1) Error Lack-of-Fit Pure Error Total	100) 2 1 2	1 44.170 2 13.393 7 11.573 5 1.820	44.1702 44.1702 0.6088 0.6808 0.3640	72.56 72.56 1.87	0.000 0.000 0.253	C6 -	9	C8	C9	C10	C11	C12	C13	C14	C15	C16
	Sale Price(/1) Error Lack-of-Fit Pure Error	100) 2 1 2	1 44.170 2 13.393 7 11.573 5 1.820 3 57.563 C2	44.1702 44.1702 0.6088 0.6808 0.3640	72.56 72.56 1.87	0.000	C6 📁 Taxes (/1000)	67	C8	C9 Temperature	C10 Yeild	C11	C12 HOUSE SIZE	C13 g MARKET PRICE	C14	C15 Demand NL (V)	
	Sale Price(/1) Error Lack-of-Fit Pure Error Total	100) 2 1 2	1 44.170 2 13.393 7 11.573 5 1.820 3 57.563 C2	44.1702 44.1702 0.6088 0.6808 0.3640 C3 2 Purity	72.56 72.56 1.87	0.000 0.000 0.253 CS		C7	C8			C11			C14		Energy Usage(
	Sale Price(/1) Error Lack-of-Fit Pure Error Total C1 Observation	100) 2 1 2	1 44.170 2 13.393 7 11.573 5 1.820 3 57.563 C2 arbon level	44.1702 44.1702 0.6088 0.6808 0.3640 C3 g Purity 93.74	72.56 72.56 1.87	0.000 0.000 0.253 CS Sale Price(/1000)	Taxes (/1000)	C7	C8	Temperature	Yeild	C11	HOUSE SIZE	MARKET PRICE	C14	Demand NL (Y)	Energy Usage(4
	Sale Price(/1) Error Lack-of-Fit Pure Error Total C1 Observation 4	100) 2 1 2	1 44.170 2 13.393 7 11.573 5 1.820 3 57.563 C2 arbon level 1.29	44.1702 44.1702 0.6088 0.6808 0.3640 C3 2 Purity 93.74 96.73	72.56 72.56 1.87	0.000 0.000 0.253 CS Sale Price(/1000) 25.9	Taxes (/1000) 4.5573	C7	C8	Temperature 59.0	Yeild 78.5	C11	HOUSE SIZE	MARKET PRICE	C14	Demand NL (V) 0.79	Energy Usage(4) 51
	Sale Price(/1) Error Lack-of-Fit Pure Error Total C1 Observation 4 5	100) 2 1 2	1 44.170 2 13.393 7 11.573 5 1.820 3 57.563 C2 1.29 1.46 1.36 0.87	44.1702 44.1702 0.6088 0.6808 0.3640 C3 2 Purity 93.74 96.73 94.45 87.59	72.56 72.56 1.87	0.000 0.000 0.253 Sale Price(/1000) 25.9 29.9	Taxes (/1000) 4.5573 5.0597 3.8910 5.8980	C7	C8	Temperature 59.0 56.0	Yeild 78.5 74.6 74.0 75.2	C11	HOUSE SIZE 90 44	MARKET PRICE 152 42 157 113	C14	Demand NL (Y) 0.79 2.70 3.64 4.73	Energy Usage(4) 5) 11: 9)
	Sale Price(71) Error Lack-of-Fit Pure Error Total Observation 4 5 6 7 7 8	100) 2 1 2	1 44.170 2 13.393 7 11.573 5 1.820 3 57.563 C2 1.29 1.46 1.36 0.87 1.23	44.1702 44.1702 0.6088 0.6808 0.3640 C3 Purity 93.74 96.73 94.45 87.59 91.77	72.56 72.56 1.87	0.000 0.000 0.253 Sale Price(/1000) 25.9 29.9 29.9 29.9 30.9 28.9	Taxes (/1000) 4.5573 5.0597 3.8910 5.8980 5.6039	C7	C8	Temperature 59.0 56.0 55.5 56.5 56.5 58.0	Yeild 78.5 74.6 74.0 75.2 77.2	C11	HOUSE SIZE 90 44 46 90 150	MARKET PRICE 152 42 157 113 573	C14	Demand NL (Y) 0.79 2.70 3.64	Energy Usage(4) 5) 11: 9)
	Sale Price(/11 Error Lack-of-Fit Pure Error Total Cl Observation 4 5 6 6 7 7 8 9 9	200) 2 1 2 Hydroc	1 44.170 2 13.393 7 11.573 5 1.820 3 57.563 C2 arbon level 1.29 1.46 1.36 0.87 1.23 1.55	44.1702 44.1702 0.6088 0.6808 0.3640 Purity 93.74 96.73 94.45 87.59 91.77 90.42	72.56 72.56 1.87	0.000 0.000 0.253 Sale Price/7000) 25.9 29.9 29.9 30.9	Taxes (/1000) 4.5573 5.0597 3.8910 5.8980	C7	C8	Temperature 59.0 56.0 55.5 56.5 58.0 57.0	Yeild 78.5 74.6 74.0 75.2	C11	HOUSE SIZE 90 44 46 90	MARKET PRICE 152 42 157 113	C14	Demand NL (Y) 0.79 2.70 3.64 4.73	Energy Usage(4) 5) 11: 9)
	Sale Price(/1) Error Lackof-Fit Pure Error Total C1 Observation 4 5 6 7 8 9 1: H +	2 1 2 Hydroc Simple	1 44,170 2 13,393 7 11,573 5 1,820 3 57,563 C2 arbon level 1,29 1,46 1,36 0,87 1,23 1,55 Linear Regr	44.1702 44.1702 0.6088 0.6808 0.3640 C3 Purity 93.74 96.73 94.45 87.59 91.77	72.56 72.56 1.87	0.000 0.000 0.253 Sale Price(/1000) 25.9 29.9 29.9 29.9 30.9 28.9	Taxes (/1000) 4.5573 5.0597 3.8910 5.8980 5.6039	C7	C8	Temperature 59.0 56.0 55.5 56.5 56.5 58.0	Yeild 78.5 74.6 74.0 75.2 77.2	C11	HOUSE SIZE 90 44 46 90 150	MARKET PRICE 152 42 157 113 573 202		Demand NL (Y) 0.79 2.70 3.64 4.73	Energy Usage(4) 5) 11: 9)
	Sale Price(/11 Error Lack-of-Fit Pure Error Total Cl Observation 4 5 6 6 7 7 8 9 9	2 1 2 Hydroc Simple	1 44,170 2 13,393 7 11,573 5 1,820 3 57,563 C2 arbon level 1,29 1,46 1,36 0,87 1,23 1,55 Linear Regr	44.1702 44.1702 0.6088 0.6808 0.3640 Purity 93.74 96.73 94.45 87.59 91.77 90.42	72.56 72.56 1.87	0.000 0.000 0.253 Sale Price(/1000) 25.9 29.9 29.9 29.9 30.9 28.9	Taxes (/1000) 4.5573 5.0597 3.8910 5.8980 5.6039	C7	C8	Temperature 59.0 56.0 55.5 56.5 58.0 57.0	Yeild 78.5 74.6 74.0 75.2 77.2	C11	HOUSE SIZE 90 44 46 90 150	MARKET PRICE 152 42 157 113 573		Demand NL (Y) 0.79 2.70 3.64 4.73	Energy Usage(4) 5) 11: 9)

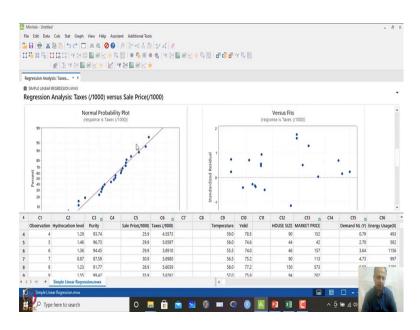
So, it will be close to R square that is calculated for manual analysis regression by its just a total and it is significant and there is no lack of fit; that means, 0.25 that you are seeing over here. So, approximately what we can see is that copy if I copy this image and I paste it over here in another sheet.

(Refer Slide Time: 29:35)

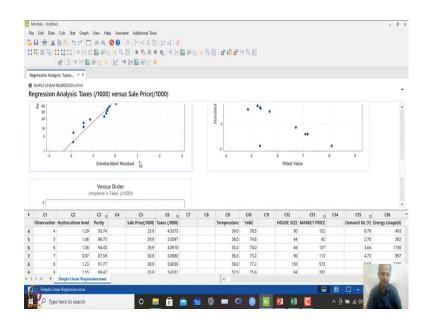
File Et Extern Data *	Show Querie		erties Links	Review	Filter Advanced	Format	Flash Fill Remove I	Uplicates C	t you want to i ¹⁰ Consolidate Relationship Manage Da	5	What-If F Analysis * Foreca	orecast Sheet		p - 9 -	🔁 Data Analysi		A Shar
Picture			ns	20	rt & Filter			Uasa roois			Poreci	154	Outin	e 14	Anarjze		
A	A .	B	c	D	E	F	G	н	i i i	,	к	L	M	N	0	P	C
	Demand NL (Y)	Energy Usage(X)			Hydrocarbon level												
-	0.79	679	0.89		0.99	90.01	0					-				-0	
	0.44	292	0.66		1.02	89.05											
	0.56	1012	0.75		1.15	91.43		Anal	ysis of	Vari	ance						
	0.79	493	0.89		1.29	93.74			/								
	2.7	582	1.64		1.46	96.73		Sourc	ce	[DF Adi S	S Adi	MS F	-Value	P-Value		
	3.64	1156	1.91		1.36	94.45											
	4.73	997	2.17		0.87	87.59		Regre	ssion		1 44.1/	0 44.1	702	72.56	0.000		
	9.5	2189	3.08		1.23	91.77	-	Sale	Price(/1)	(000	1 44.17	0 44.1	702	72.56	0.000	6	
0	5.34	1097	2.31		1.55	99.42	Ĭ	Error			22 13.39		088			Ĭ	
1	6.25	2078	2.50		1.4	93.65		Error			22 13.39	3 0.0	088				
2	5.84	1818	2.42		1.19	93.54		Lack	-of-Fit	1	17 11.57	3 0.6	808	1.87	0.253		
	5.21	1700	2.28		1.15	92.52		Dure	Error		5 1.82	0 0 2	640				
	3.25	747	1.80		0.98	90.56			Entor				040				
5	4.43	2030	2.10		1.01	89.54		Total			23 57.56	3					
6	3.16	1643	1.78		1.11	89.85											
1	0.5	414	0.71		1.2	90.39	0									-0	
3	0.17	354	0.41		1.26	93.25											
	1.88	1276	1.37		1.32	93.41											
)	0.77	745	0.88		1.43	94.98											
	1.39	435	1.18		0.95	87.33										63	
2	0.56	540	0.75													and the	
	Sheet2 Shee	t3 Sheet1 (÷						E							6	-
(ne)	27												E	8 (0)	m -		

So, what we are seeing over here is that analysis. So, in this case what we are seeing is that regression is significant black of it is not there. So, in this case model can be generalized.

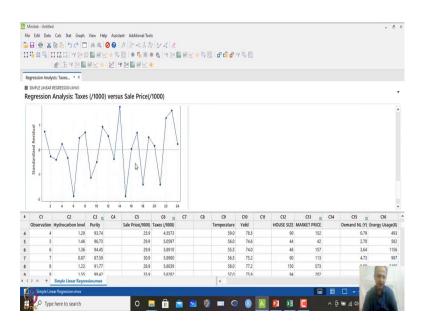
(Refer Slide Time: 29:47)



(Refer Slide Time: 29:49)

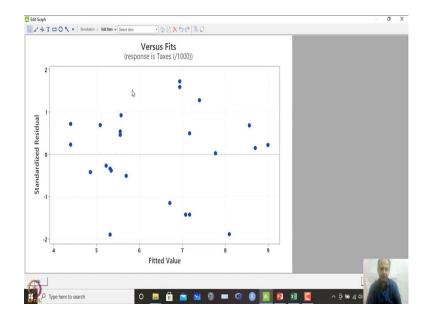


(Refer Slide Time: 29:51)

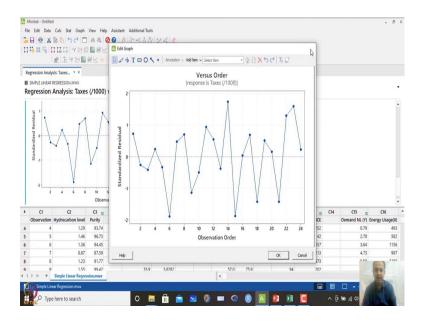


So, normal probability part more or less is seems to be ok, but we can do Anderson darling test and there is as such no pattern in the residual versus fit diagram equation over here.

(Refer Slide Time: 29:54)

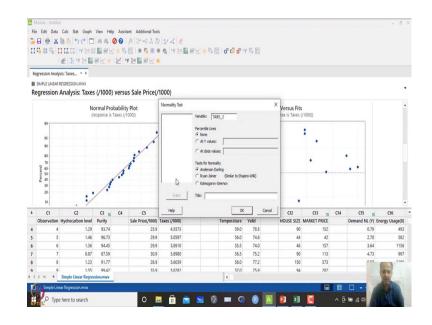


(Refer Slide Time: 30:04)

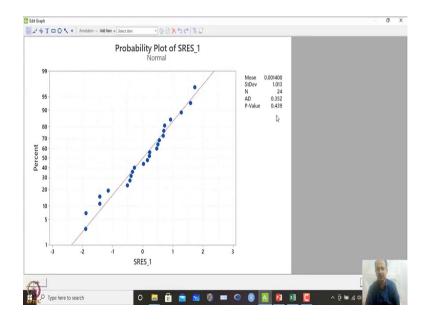


So, heteroscedasticity may not be a problem over here and also the autocorrelation does not seems to be significant, because there is no trend as such that we are seeing and maybe random observations. But we can do Durbin Watsons we can do that test Durbin-Watson test and or Breusch-Pagan test BP test we can do over here to confirm that one ok and we have stored the residuals.

(Refer Slide Time: 30:23)



(Refer Slide Time: 30:26)



So, we can see at least the normality test and when I go to the last residual over here and do this test what will be what we see is that the p-value is not significant. So, in this case we can assume that the residual is normal. So, what we do is that we test the residual over here and if the condition is not satisfactory then in that case we go for transformation.

So, next case what we will see is that ok errors are non normal. So, in that scenario what is to be done and how the regression models has to be developed. So, that we can see in another examples in our next session like that and some more complexities we can see for now we will go to multiple regression.

So, some time we will spend in multiple regression, because in design of experiment what has what happens is that there is not single x that we are dealing with design of experiments. There can be multiple x single x is the one way analysis of variance that we are dealt with and that is the simplest of condition but that is not reality basically.

So, there will be multiple x that will influence the y and we need to develop the mathematical model and we need to develop the regression equation which we have to optimize basically finally ok. So that case we will discuss about that, but at present we will stop over here and we will continue in our next session on simple regression complexities when error is not normal in that case how do I deal with that.

Like analysis of variance and we will also go to multiple regression and see the complexities what, in modeling what we face basically. So we will continue with that.

Thank you.