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

# Lecture - 26 Best Subset Regression, Multicollinearity

Hello and welcome everyone to this course on Quality Control and Improvement with MINITAB. We are in session 26, and I am Professor Indrajit Mukherjee from Shailesh J Mehta School of Management, IIT Bombay.

So, in last session, we are discussing about multiple regression and how to select the variables. You see there are confusions which arises when we have multiple x which is regressed with a y, single y. So, we are not sure which variable should be considered, which is not to be considered. So, some examples we have taken last time. So, let us try to see whether we can resolve that problem and dilemmas that we are facing.

(Refer Slide Time: 00:50)

Example						
The electric power consumed each	Y	$X_1$	X <sub>2</sub>	X <sub>3</sub>	X <sub>4</sub>	
month by a chemical plant is thought	240	25	24	91	100	
to be <b>related to the average ambient</b> temperature (x1), the number of	236	31	21	90	95	
days in the month (x2), the average	270	45	24	88	110	
product purity (x3), and the tons of	274	60	25	87	88	
product produced (x4).	301	65	25	91	94	
The past year's historical data are available and are presented in the	316	72	26	94	99	
given table. Obtain a least square	300	80	25	87	97	
regression model.	296	84	25	86	96	
	267	75	24	88	110	
	276	60	25	91	105	
	288	50	25	90	100	
	261	38	23	89	98	Data Causas Mantaaman D. C. (20)
~						Data Source: Montgomery, D. C. (200 Applied statistics and probal engineers. John Wiley & Son

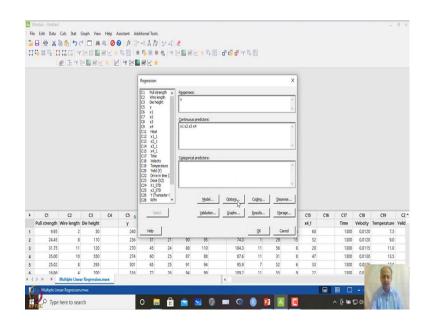
So, this was one of the problem that we are dealing with at the time point, that electrical power consumption is monitored over here. And this may be related to variable X1, X2, X3, and X4 and the details are given on the left hand side of your screen in this power point. And we want to see which is the best model that explains y with respect to given x, set of x like that.

# (Refer Slide Time: 01:20)

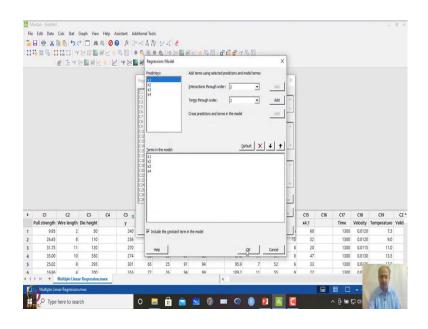
			Basic Statis Regression ANOVA DOE Control Ch Quality Too Reliability/? Predictive J Multivariab Time Series Tables Nonparam Equivalence	erts Is iurvival inalytic trics	3	Stabil Stabil Ortho Partia Binar Binar	I Line Plot Ission iftear Regres ity Study opgonal Regre I Least Squa y Fitted Line y Logistic Re ual Logistic Re inal Logistic	ion ssion res Plot gression egression	, , , , , , , , , , , , , , , , , , ,	Fit Regress Best Subse Predict Factorial P Contour P Surface Pb Overlaid C Response	sion Model. ets Plots Plot lot Contour Plot Optimizer	ta									
			Power and	Sample	Size	, Poiss	on Regressic	n	Ne		eet Ctrl+ eet Ctrl+ eet Ctrl+	Shift+N									
F	CI			Semple	Size C4	, Poiss CS p	on Regressic	n C7		New Proje	ect Ctrl+	Shift+N	C12	C13	C14	C15	C16	C17	C18	C19	Q
	Il strength V	C2 Wire length	C3 h Die hei	pht		CS g	C6 x1	C7 x2	Ne C8 x3	New Proje w Workshe C9 x4	ect Ctrl+ eet Ctrl+	rShift+N +N C11 g Heat	C12 x1_1	x2_1	C14 x3_1	x4_1	C16	Time	Velocity	Temperature	Yeild
	Il strength V 9.95	C2 Wire length	Power and C3 h Die hei 2	ght 50		C5 12 y 240	C6 x1 25	C7 x2 24	Ne C8 x3 91	New Proje w Workshe C9 x4 100	ect Ctrl+ eet Ctrl+	Cti z Heat 78.5		<b>x2_1</b> 26	x3_1 6	<b>x4_1</b> 60	C16	Time 1300	Velocity 0.0120	Temperature 7.5	Yeld
	Il strength V 9.95 24.45	C2 Wire length	Power and C3 h Die hei 2 8	<b>yht</b> 50		C5 g y 240 236	C6 x1 25 31	C7 x2 24 21	Ne C8 x3 91 90	New Proje w Workshe C9 x4 100 95	ect Ctrl+ eet Ctrl+	Cti 12 Heat 78.5 74.3	x1_1 7 1	x2_1 26 29	x3_1 6 15	x4_1 60 52	C16	Time 1300 1300	Velocity 0.0120 0.0120	Temperature 7.5 9.0	Yeild
Pul	Il strength V 9.95 24.45 31.75	C2 Wire length	Power and C3 Die hei 2 8 1	<b>yht</b> 50 110 120		C5 g y 240 236 270	C6 x1 25 31 45	C7 x2 24 21 24	Ne C8 x3 91 90 88	New Proje w Workshe w Vorkshe x4 100 95 110	ect Ctrl+ eet Ctrl+	C11 2 Heat 78.5 74.3 104.3	x1_1 7 1	x2_1 26 29 56	x3_1 6 15 8	x4_1 60 52 20	C16	Time 1300 1300 1300	Velocity 0.0120 0.0120 0.0115	Temperature 7.5 9.0 11.0	Yeik
	Il strength V 9.95 24.45 31.75 35.00	C2 Wire length	Power and C3 Die hei 2 8 1 0	ght 50 110 120		C5 g y 240 236 270 274	C6 x1 25 31 45 60	C7 x2 24 21 24 25	Ne C8 x3 91 90 88 87	New Proje w Workshe w 4 100 95 110 88	ect Ctrl+ eet Ctrl+	Ct1 2 Heat 78.5 74.3 104.3 87.6	x1,1 7 1 11	x2_1 26 29 56 31	x3_1 6 15 8 8	x4_1 60 52 20 47	C16	Time 1300 1300 1300 1300	Velocity 0.0120 0.0120 0.0115 0.0130	Temperature 7.5 9.0 11.0 13.5	Yeik
	Il strength V 9.95 24.45 31.75	C2 Wire length	Power and Die hei 2 8 1 0 8	<b>yht</b> 50 110 120		C5 g y 240 236 270	C6 x1 25 31 45	C7 x2 24 21 24	Ne C8 x3 91 90 88	New Proje w Workshe w Vorkshe x4 100 95 110	ect Ctrl+ eet Ctrl+	C11 2 Heat 78.5 74.3 104.3	x1_1 7 1	x2_1 26 29 56	x3_1 6 15 8	x4_1 60 52 20	C16	Time 1300 1300 1300	Velocity 0.0120 0.0120 0.0115 0.0130 0.0135	Temperature 7.5 9.0 11.0	Yelld

So, what we have done is that we have gone to MINITAB, and then what we have seen is that these are the variables this from C5 to C9, C5 is the y and C6 to C9 is the 4 variables. So, what we have done is that stat we have gone to regression, and we have gone to regression, and then we have used fit regression models.

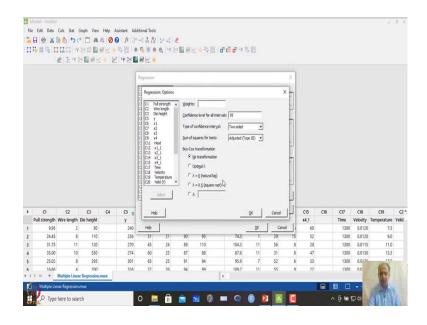
(Refer Slide Time: 01:40)



## (Refer Slide Time: 01:49)



(Refer Slide Time: 01:55)

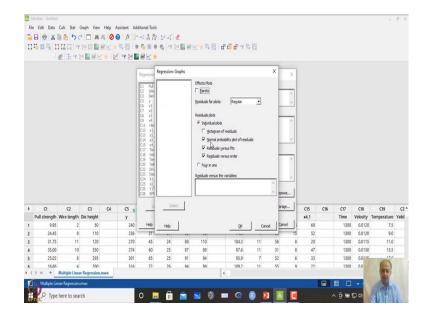


And then what we have done is that we have selected y and the set of x conditions that is X1, X2, X3, and X4. And in models what we have done is that all the variables we have considered include a constant term. And options we have not given any transformation over here.

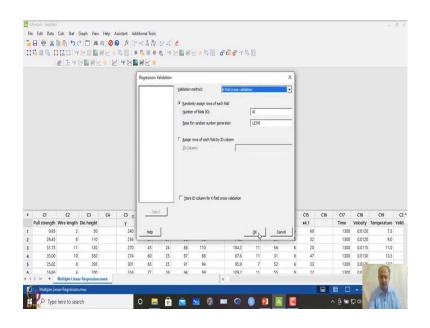
# (Refer Slide Time: 02:00)

Fil	tinitab - Untitled a Edit Data □ 👳 😹 1 □ III III III III III IIII □ III III II	Calc Stat		4 0 ( 8 k *	) fx   3 時間	Method: Repuise						0 ×
						[==]ctdat seets in our vended [==]ctdat isen in the retain restrict and a Apha to region: [0,15]						
	CI	C2	C3	C4	C5 g		C15	C16	C17	C18	C19	C2 *
	Pull strength				y		x4_1		Time		Temperature	
1	9.95	2	50		240		60		1300	0.0120	7.5	
2	24.45	8	110		236	Herarchy	52		1300	0.0120	9.0	
3	31.75	11	120		270		20		1300	0.0115	11.0	
4	35.00	10	550		274	G gisplay the table of model selection details	47		1300	0.0130	13.5	
5	25.02	8	295		301	Details about the method	33		1300	0.0175	17.0	
<b>К</b> Н (	_	4 Multiple Line	ar Regression	mwx	316	7 Digitive the graph of R-squared vs step	27		1300	0.0		1
		ear Regression here to sea				Hep (X Greet		· ·	ت ۱ هه ق	 12 40	SMA	1

# (Refer Slide Time: 02:07)

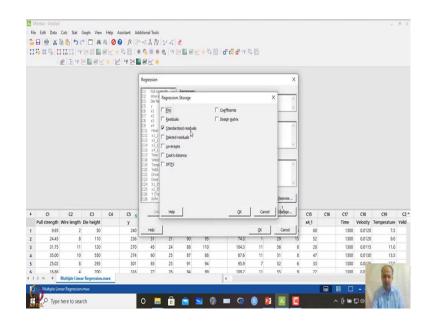


# (Refer Slide Time: 02:16)



And in this case first we have not done not given any let us say stepwise regression over here. And we may see graphs over here, normal probability plot, residual plots and order plots like that. So, this is possible. And validation tenfold cross-validation that we have discussed last time also we can put over here.

(Refer Slide Time: 02:21)



#### (Refer Slide Time: 02:25)

<b>E1.</b>	linitab - Untitled																			Ð
ю	Edit Data	Calc Stat Gra	ph View	Help /	ssistant Ada	itional Tool	í.													
	8 8 X	50 50	- M	10	a fx 3-	-14.8	21 1													
		I In Id Y							K + B	10	r d d	收回								
		@ EYN						C. 100 C.	L			- V US								
				E		e n.														
Re	gression Analys	is: y vers × ×																		
1	MULTIPLE LINEAR	R REGRESSION MW	x																	
Re	gression /	Analysis: y v	ersus >	(1, x2, :	(3, x4															
	Method																			
	Cross-validation	n 10-fold																		
1	Regression E	Equation																		
	y = -123+0	.757 x1 + 7.52 x2	+ 2.48 x3 -	0,481 x4				N												
								6												
1	Coefficients																			
		Loef SE Coef		P-Value	VIF															
		-123 157 .757 0.279	-0.78 2.71	0.459	2.32															
		7.52 4.01	1.87	0.030																
			G																	
	C1 Pull strength	C2 Wire length Di		C4	C5 g	C6 x1	C7 x2	C8 x3	C9 x4	C10	C11 g Heat	C12	C13 x2_1	C14 x3_1	C15 x4_1	C16	C17 Time	C18 Velocity	C19 Temperature	
	9.95	2	50		240	25	24	91	100		78.5	7	26	6	60		1300	0.0120	7.5	
	24,45	8	110		236	31	21	90	95		74.3	1	29	15	52		1300	0.0120	9.0	
					270	45	24	88	110		104.3	11	56	8	20					
	31.75	11	120		270						87.6	11	31	8			1300	0.0115	11.0	
		11	120		274	60	25	87	88		87.0	11	31	0	47		1300 1300	0.0115	11.0	
	31.75					60 65	25 25	87 91	88 94		95.9	7	52	6	33					
	31.75 35.00	10	550		274												1300	0.0130	13.5	
	31.75 35.00 25.02 16.86	10 8	550 295 200	mwx	274 301	65	25	91	94	4	95.9	7	52	6	33		1300 1300	0.0130	13.5	
	31.75 35.00 25.02 16.86 b н +	10 8 4 Multiple Linear R	550 295 200 legression.r	mwx	274 301	65	25	91	94	ł	95.9	7	52	6	33		1300 1300 1300	0.0130	13.5	
	31.75 35.00 25.02 16.86 b н +	10 8 4	550 295 200 legression.r	mwx	274 301 316	65	25 26	91 94	94		95.9	7	52	6	33		1300 1300 1300	0.0130	13.5	

### (Refer Slide Time: 02:32)

		0.1																		-	Ø
						Assistant Add															
						0 fx 3=															
16			3 4Y 3	101	sk *	時回:#	₽ <mark>0</mark> Ⅲ 4	F 选 LY	200	长*科	10	🐻 🗗 -Y	時間								
						- Y > S															
_	_	-					the late														
gress	ion Analy	ysis: y ve	rs * x																		
			ESSION MI																		
egre	ession	Analy	ysis: y	versus	x1, x2, )	x3, x4															
Tern	1	Coef	SE Coef	T-Value	P-Value	VIF															
Cons	tant	-123	157	-0.78	0.459																
x1		0.757	0.279	2.71	0.030	2.32															
x2		7.52	4.01	1.87	0.103	2.16															
x3		2.48	1.81	1.37	0.212	1.34															
×4		0.481	0.555	-0.87	0.415	1.01															
11.	7866 8	R-sq 5.20%	76.75%			old s 10-fol	R-sq 8.45%														
11. Ana	5 7866 8: Iysis of	R-sq 5.20% Varian	76.75%	52.4	8% 16,8	8035 4															
11. Ana	5 7866 8: lysis of re	R-sq 5.20% Varian	76.75%	52.4	8% 16.8 Walue D-	8035 4 Value	8.45%	9	68	<b>C9</b>	C10	C11 -	C12	C13	C14	CIS	C16	C17	Ct8	C19	
11. Ana	S 7866 8: lysis of re C1	R-sq 5.20% Varian DF A	76.75% ACCE ALISS &	52.4 di MS E- C3	8% 16,8	8035 4		C7 x2	C8 x3	C9 x4	C10	C11 5	C12	C13	C14 x3.1	C15 x4.1	C16	C17 Time	C18 Velocity	C19 Temperature	2
11. Ana	S 7866 8: lysis of re C1	R-sq 5.20% Varian DF A C Wire I	76.75% ACCE ALISS &	52.4	8% 16.8 Walue D-	8035 4 Value C5 g	8.45% C6			_	C10					C15 x4_1 60	C16			C19 Temperature 7.5	
11. Ana	S 7866 8: lysis of re C1 strength	R-sq 5.20% Varian DF A C Wire I	76.75% Ace Aliss A 2 length D	di MS F- C3 tie height	8% 16.8 Walue D-	8035 4 Value C5 g y	8.45% C6 x1	x2	x3	x4	C10	Heat		x2_1	x3_1	x4_1	C16	Time	Velocity	Temperature	5
11. Ana	S 7866 8: lysis of re C1 strength 9.93	R-sq 5.20% Varian DF A C Wire I S	76.75% Acce whi ss A 2 length D 2	di MS F- C3 tie height 50	8% 16.8 Walue D-	8035 4 Value C5 5 y 240	C6 x1 25	x2 24	x3 91	x4 100	C10	Heat 78.5		<b>x2_1</b> 26	x3_1 6	x4_1 60	C16	Time 1300	Velocity 0.0120	Temperature 7.5	5
11. Ana	S 7866 8: lysis of re C1 strength 9.92 24.43	R-sq 5.20% Varian DF A C Wire I 5 5	76.75% hce di ss A 2 kength D 2 8	52.4 rli MS F- C3 tie height 50 110	8% 16.8 Walue D-	8035 4 Value C5 g y 240 236	C6 x1 25 31	<b>x2</b> 24 21	x3 91 90	x4 100 95	C10	Heat 78.5 74.3	x1_1 7 1	x2_1 26 29	x3_1 6 15	x4_1 60 52	C16	Time 1300 1300	Velocity 0.0120 0.0120	Temperature 7.5 9.0	5
11. Ana	S 7866 8: lysis of re C1 strength 9.95 24.45 31.75	R-sq 5.20% DF A Wire I 5 5 5	76.75% ACE whi SS & A 22 kength D 2 8 11	di MS F- C3 ie height 50 110 120	8% 16.8 Walue D-	236 270	C6 x1 25 31 45	x2 24 21 24	x3 91 90 88	x4 100 95 110	C10	Heat 78.5 74.3 104.3	x1_1 7 1 11	x2_1 26 29 56	x3_1 6 15 8	x4_1 60 52 20	C16	Time 1300 1300 1300	Velocity 0.0120 0.0120 0.0115	Temperature 7.5 9.0 11.0	5
11. Ana	S 7866 8: lysis of re C1 strength 9.99 24.49 31.75 35.00	R-sq 5.20% Varian DF A C Wire I S S S S S	76.75% Acce Hiss A Rength D 2 8 11 10	<ul> <li>4i MS E-7</li> <li>C3</li> <li>tie height</li> <li>50</li> <li>110</li> <li>120</li> <li>550</li> </ul>	8% 16.8 Walue D-	Value C5 0 y 240 236 270 274	C6 x1 25 31 45 60	x2 24 21 24 25	x3 91 90 88 87	x4 100 95 110 88	C10	Heat 78.5 74.3 104.3 87.6	xt_1 7 1 11	x2_1 26 29 56 31	x3_1 6 15 8 8	x4_1 60 52 20 47	C16	Time 1300 1300 1300 1300	Velocity 0.0120 0.0120 0.0115 0.0130	Temperature 7.5 9.0 11.0 13.5	5
11. Ana	5 7866 8: lysis of ve C1 5trength 9.95 24.45 31.75 35.00 25.00	R-sq 5.20% DF A C Wire I S S S S S S	76.75% Ace Ali SS A 2 Rength D 2 8 11 10 8 4	<ul> <li>52.4</li> <li>Ali MS E-</li> <li>C3</li> <li>te height</li> <li>50</li> <li>110</li> <li>120</li> <li>550</li> <li>295</li> </ul>	8% 16.8 Makue P- C4	Value Value 240 236 270 274 301	C6 x1 25 31 45 60 65	x2 24 21 24 25 25	x3 91 90 88 87 91	x4 100 95 110 88 94	C10	Heat 78.5 74.3 104.3 87.6 95.9	xt_1 7 1 11 11 7	x2_1 26 29 56 31 52	x3_1 6 15 8 8 8 6	x4_1 60 52 20 47 33	C16	Time 1300 1300 1300 1300 1300	Velocity 0.0120 0.0120 0.0115 0.0130 0.0135	Temperature 7.5 9.0 11.0 13.5	5
11. Ana Sour	S 1/ysis of ve C1 strength 9.93 24.43 31.73 35.00 25.00 16.84 41 +	R-sq         S.20%           Varian         DE         A           DE         A         CC           n         Wire I         S           5         S         S           5         S         S           5         S         S           5         Multip         Multip	76.75% hce di SS A 2 kength D 2 8 11 10 8 4 4 ke Linear I	rli MS E-3 C3 ite height 50 110 120 550 295 200 Regression.	8% 16.8 Makue P- C4	Value Value 240 236 270 274 301	C6 x1 25 31 45 60 65	x2 24 21 24 25 25	x3 91 90 88 87 91	x4 100 95 110 88 94		Heat 78.5 74.3 104.3 87.6 95.9	xt_1 7 1 11 11 7	x2_1 26 29 56 31 52	x3_1 6 15 8 8 8 6	x4_1 60 52 20 47 33		Time 1300 1300 1300 1300 1300 1300	Velocity 0.0120 0.0120 0.0115 0.0130 0.0135 0.0	Temperature 7.5 9.0 11.0 13.5	5
11. Source	S 1/ysis of ve C1 strength 9.93 24.43 31.73 35.00 25.00 16.84 41 +	R-sq         S.20%           Varian         DE         A           DE         A         CC           n         Wire I         S           5         S         S           5         S         S           5         S         S           5         Multip         Multip	76.75% Ace Ali SS A 2 Rength D 2 8 11 10 8 4	rli MS E-3 C3 ite height 50 110 120 550 295 200 Regression.	8% 16.8 Makue P- C4	C5 2 y 240 236 270 274 301 316	C6 x1 25 31 45 60 65	x2 24 21 24 25 25	x3 91 90 88 87 91	x4 100 95 110 88 94 94		Heat 78.5 74.3 104.3 87.6 95.9	x1_1 7 11 11 11 7	x2_1 26 29 56 31 52	x3_1 6 15 8 8 8 6	x4_1 60 52 20 47 33		Time 1300 1300 1300 1300 1300 1300	Velocity 0.0120 0.0120 0.0115 0.0130 0.0135 0.0135	Temperature 7.5 9.0 11.0 13.5	5

And then storage I want to store that the standardized residual let us say and click ok. So, let us try to see what are the results that we are getting. So, MINITAB gives you automatically a regression equation over here, but what is surprising over here you see that although X1 is the only variable which is significant over here; that means, less than 0.05 and others are not significant terms over here.

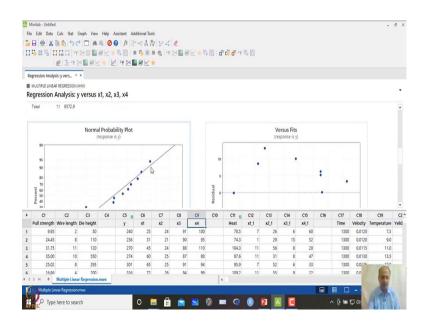
So, only X1 is showing significance, others are not showing significance over here. And also you see the R square value adjusted R square is about 76 and tenfold R square this

cross-validation is around 48. So, when there is a difference between these two, significant difference that exist over here there must be something going wrong over here. And maybe model fitting is not correct or we have over-fitted the model basically, ok.

(Refer Slide Time: 03:14)

	initab - Untitle	ed																			-	Ø
le	Edit Data	Ca	nic Stat	Graph V	iew Help	Ass	istant Add	itional Tools	5													
í	8 💩 X	D	6 5	0	H 14 6	00	fx a	七人名	244	1												
	6 m m										K N	S de	di Ly	的图								
							4Y >															
				_		-		er LL.														
	pression Anal																					
	NULTIPLE LINE																					
ke	gression				is x1, x	2, X3																
				•••																		
	Analysis of	Var	iance																			
	Source				C.Mahun	p-Va	due															
	Regression	4		Adj MS 1400.1	10.08		005															
	x1	1	1022.8	1022.8	7.36		.030															
	x2	1	488.4	488.4	3.52	0.	103															
	x3	1	261.6																			
			20110	261.6	1.88		212															
	xo4	1	104.3	104.3	1.88		212 415															
	Error	1 7	104.3 972.5																			
			104.3	104.3																		
	Error		104.3 972.5	104.3																		
	Error		104.3 972.5	104.3 138.9		0.	415								Ve	rsus Fits						٣
	Error Fotal C1	11	104.3 972.5 6572.9	104.3 138.9 Nor	0.75 mal Prol	0. babil	415 lity Plot C5 g	C6	67	C8	()	C10	C11 12	C12	C13	C14	C15	C16	C17	C18	C19	
	Error Fotal C1 Pull strengt	11 h W	104.3 972.5 6572.9 C2 ire length	104.3 138.9 Nor C3 Die heig	0.75 mal Prol C4 ht	0. babil	415 lity Plot CS g	xt	x2	x3	x4	C10	Heat	C12 x1_1	C13 x2_1	C14 x3_1	x4_1	C16	Time	Velocity	Temperature	Yello
	Error Fotal C1 Pull strengt 9.9	11 h W 5	104.3 972.5 6572.9 C2 ire length	104.3 138.9 Nor C3 Die heig	0.75 mal Prol C4 ht 50	0. babil	415 lity Plot C5 p y 240	x1 25	x2 24	x3 91	x4 100	C10	Heat 78.5		C13 x2_1 26	C14 x3_1 6	<b>x4_1</b> 60	C16	Time 1300	Velocity 0.0120	Temperature 7.5	Yeik
	C1 Pull strengt 9.9 24.4	11 h W 5 5	104.3 972.5 6572.9 C2 fire length 2 8	Nor C3 Die heig	0.75 mal Prol 64 ht 50 10	0. babil	415 lity Plot C5 12 y 240 236	x1 25 31	x2 24 21	x3 91 90	x4 100 95	C10	Heat 78.5 74.3	x1_1 7 1	C13 x2_1 26 29	C14 x3_1 6 15	x4_1 60 52	C16	Time 1300 1300	Velocity 0.0120 0.0120	Temperature 7.5 9.0	Yello
	C1 Pull strengt 9.9 24.4 31.7	11 h W 5 5 5	104.3 972.5 6572.9 C2 fre length 2 8 11	104.3 138.9 Nor C3 Die heig	0.75 mal Prol 64 ht 50 10 20	0. babil	415 lity Plot C5 12 y 240 236 270	x1 25 31 45	x2 24 21 24	x3 91 90 88	x4 100 95 110	C10	Heat 78.5 74.3 104.3	x1_1 7 1	C13 x2_1 26 29 56	C14 x3_1 6 15 8	x4_1 60 52 20	C16	Time 1300 1300 1300	Velocity 0.0120 0.0120 0.0115	Temperature 7.5 9.0 11.0	Yeik
	C1 Pull strengt 9.9 24.4	11 h W 5 5 5 5 0	104.3 972.5 6572.9 C2 fire length 2 8	104.3 138.9 Nor C3 Die heig	0.75 mal Prol 64 ht 50 10	0. babil	415 lity Plot C5 12 y 240 236	x1 25 31	x2 24 21	x3 91 90	x4 100 95	C10	Heat 78.5 74.3	x1_1 7 1	C13 x2_1 26 29	C14 x3_1 6 15	x4_1 60 52	C16	Time 1300 1300	Velocity 0.0120 0.0120	Temperature 7.5 9.0	Yello
	C1 Pull strengt 9.9 24.4 31.7 35.0	11 h W 5 5 5 5 5 5 5 5 5 5 5 5 5 5 5 5 5 5 5	104.3 972.5 6572.9 C2 ire length 2 8 11	104.3 138.9 Nor C3 Die heig 1 1 1 1 5 2	0.75 mal Prol 64 ht 50 10 20 50	0. babil	415 lity Plot C5 12 240 236 270 274	x1 25 31 45 60	x2 24 21 24 25	x3 91 90 88 87	x4 100 95 110 88	C10	Heat 78.5 74.3 104.3 87.6	x1_1 7 1 11 11	C13 x2_1 26 29 56 31	C14 x3_1 6 15 8 8	x4_1 60 52 20 47	C16	Time 1300 1300 1300 1300	Velocity 0.0120 0.0120 0.0115 0.0130	Temperature 7.5 9.0 11.0 13.5	
	C1 Pull strengt 9.9 24.4 31.7 35.0 25.0	11 h W 5 5 5 5 6	104.3 972.5 6572.9 C2 fire length 2 8 11 10 10 8	104.3 138.9 Nor C3 Die heig 1 1 1 1 5 2	0.75 mal Prol C4 ht 50 10 20 50 95 00	0. babil	415 lity Plot C5 p y 240 236 270 274 301	x1 25 31 45 60 65	x2 24 21 24 25 25	x3 91 90 88 87 91	x4 100 95 110 88 94	C10	Heat 78.5 74.3 104.3 87.6 95.9	x1_1 7 11 11 11 7	C13 x2_1 26 29 56 31 52	C14 x3_1 6 15 8 8 6	x4_1 60 52 20 47 33	C16	Time 1300 1300 1300 1300 1300	Velocity 0.0120 0.0120 0.0115 0.0130 0.0135	Temperature 7.5 9.0 11.0 13.5	Yello
	C1 Pull strengt 9.9 24.4 31.7 35.0 25.0 16.8 5 H +	11 h W 5 5 5 5 5 6 M	104.3 972.5 6572.9 C2 fire length 2 8 11 10 2 8 4 4 4 4 4 4 4 4 4 4 4 4 4 4 4 4 4 4	104.3 138.9 Nor C3 Die heig 1 1 1 1 5 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2	0.75 mal Prol C4 ht 50 10 20 50 95 00	0. babil	415 lity Plot C5 p y 240 236 270 274 301	x1 25 31 45 60 65	x2 24 21 24 25 25	x3 91 90 88 87 91	x4 100 95 110 88 94		Heat 78.5 74.3 104.3 87.6 95.9	x1_1 7 11 11 11 7	C13 x2_1 26 29 56 31 52	C14 x3_1 6 15 8 8 6	x4_1 60 52 20 47 33		Time 1300 1300 1300 1300 1300 1300	Velocity 0.0120 0.0120 0.0115 0.0130 0.0135	Temperature 7.5 9.0 11.0 13.5	Yeli
	Ct Pull strengt 9.9 24.4 31.7 35.0 25.0 16.8 b H + Wuttple 1	11 h W 5 5 5 5 5 6 Minear	104.3 972.5 6572.9 C2 fire length 2 8 11 10 10 8	104.3 138.9 Nor C3 Die heig 1 1 1 1 5 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2	0.75 mal Prol C4 ht 50 10 20 50 95 00	0. babil	415 lity Plot C5 2 y 240 236 270 274 301 316	x1 25 31 45 60 65	x2 24 21 24 25 25 25 26	x3 91 90 88 87 91 94	x4 100 95 110 88 94	1	Heat 78.5 74.3 104.3 87.6 95.9 109.2	xt_1 7 11 11 11 7	C13 x2_1 26 29 56 31 52	C14 x3_1 6 15 8 8 6	x4_1 60 52 20 47 33		Time 1300 1300 1300 1300 1300 1300	Velocity 0.0120 0.0120 0.0115 0.0130 0.0135 0.0135	Temperature 7.5 9.0 11.0 13.5	Yeil

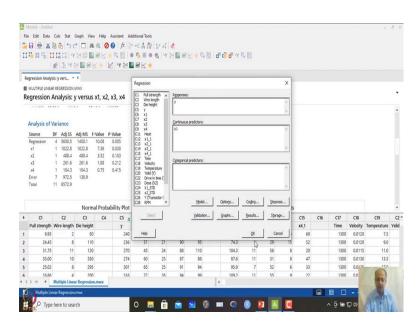
(Refer Slide Time: 03:18)



So, in this scenario when we are not sure which variable will go in which variable will go out. And analysis of variance also confirm that one of the variable X1 is significant others are not significant what we can do is that, we can remove all the variables all the

X2, X3, X4s like that and retain only X1 like that only retain X1 and then we can see. So what will happen? So, this is trial and error methods that I am showing over here without going into our stepwise regression like that.

(Refer Slide Time: 03:36)



(Refer Slide Time: 03:42)

Minitab - Unti						Regression: Stepwise ×						θ ×
					Assistant A	Hethod:						
۵. 👲 🖯	X 🗈	6 50	e 🗆 1	6 /4 (	<b>0 f</b> x	Dealer Income						
		1 13 4		Bk	*呜回!	Potential terms:						
	1	E Y	200	K×	12 Y 12	xi						
Regression Ar	alysis:	y vers *	×									
MULTIPLE LI												
Regressio	n Ar	nalysis:	y versus	5 x1, x	2, x3, x4							
			••• ••									
Analysis												
Source			Adj MS 1400.1		P-Value 0.005	E = Include term in every model E = Include term in the initial model						
Regression ×1	4	1022.8	1022.8	10.08	0.005	Alpha to egter: 0.15						
×2	1	438.4	488.4	3.52	0.103	Nena to egyer. Julia						
x3	1	261.6	261.6	1.88	0.212	Alpha to remove: 0.15						
x04	1	104.3	104.3	0.75	0.415							
Error	7	972.5	138.9									
Total	11	6572.9										
			Norn	nal Prol	bability Plot		s					÷
+ C1		C2	C3	C4			C15	C16	C17	C18	C19	C2
Pull streng	yth W	rire length	Die heigh	t	У		x4_1		Time	Velocity	Temperature	Yeild
1 9	.95	2	5	0	240		60		1300	0.0120	7.5	
2 24	.45	8	11	5	236	Herarchy	52		1300	0.0120	9.0	
3 31	.75	11	12	)	270		20		1300	0.0115	11.0	
4 35	.00	10	55	)	274	P Qrolay the table of model selection details	47		1300	0.0130	13.5	
5 25	.02	8	29	5	301	Details about the method v	33		1300	0.0175	170	
	86	4	710		316		1 77		1300	0.0		
ары +	M	ultiple Line	ar Regressie	m.mwx		P Digitary the graph of R-squared vs step						
Multipl	e Linea	r Regression	nmwx			Help OK Cancel		<b>.</b>			AN	
Ma.			cab.						<u>5</u> w	D 44	AL AL	
NPTEL	/pe he	ere to sea	ircn			0 📕 🗄 🕿 📽 🖗 💻 🧟 🙋 🛄 🚺			5 W)	E. 40	SVIMI	PA

# (Refer Slide Time: 03:47)

				4 0 8 K 1	Assistant Add	-:	1 2 4   北 4   4		<b>K</b> ∗ \$	10 d <b>a</b> 1	🖥 🗗 (Y )	<sup>4</sup>							-	θ ;
Reg	ession Analysi	is: y vers	×																	
II M	ULTIPLE LINEAR	REGRESSION	MWX																	
Reg	ression A	Analysis:	y versus	x1																•
M	ethod																			*
	oss-validation	10-fold																		
R	egression E	quation																		
у	= 219.4+	1.011 x1																		
						2														
						.0														
C	pefficients					-														
Te	rm Ci	oef SE Coef			VIF															
Te	rm Constant 21	9.4 14.3	15.38	0.000																
Te	rm Constant 21		15.38	0.000	VIF 1.00	19														v
Ti G X	rm Cr onstant 21 1.0 C1	9,4 14.3 011 0.238 C2	15.38 4.25 C3	0.000	1.00 C5 g	C6	67	C8	C9	C10	C11 g	C12	C13	C14	C15	C16	C17	C18	C19	
P	rm Cr onstant 21 1.0 C1 ull strength	9,4 14.3 011 0.238 C2 Wire length	C3 Die height	0.000	1.00 C5 g	C6 x1	x2	×3	×4	C10	Heat	C12 x1_1	x2_1	x3_1	x4_1	C16	Time	Velocity	Temperature	Yeild
Th G X <sup>1</sup>	rm Cr onstant 21 1.0 C1 ull strength 9.95	9.4 14.3 011 0.238 C2 Wire length 2	C3 Die height 50	0.000	1.00 C5 5 y 240	C6 x1 25	x2 24	x3 91	x4 100	C10	Heat 78.5		x2_1 26	x3_1 6	x4_1 60	C16	Time 1300	Velocity 0.0120	Temperature 7.5	Yeild
The Colored State	rm Co onstant 21 1.0 C1 ull strength 9.95 24.45	9.4 14.3 011 0.238 C2 Wire length 2 8	C3 Die height 50 110	0.000	1.00 C5 5 y 240 236	C6 x1 25 31	<b>x2</b> 24 21	x3 91 90	x4 100 95	C10	Heat 78.5 74.3	x1_1 7 1	x2_1 26 29	x3_1 6 15	<b>x4_1</b> 60 52	C16	Time 1300 1300	Velocity 0.0120 0.0120	Temperature 7.5 9.0	Yeild
Te G x <sup>1</sup>	rm Cr onstant 21 1.0 C1 ull strength 9.95 24.45 31.75	9,4 14.3 011 0.238 C2 Wire length 2 8 11	C3 Die height 50 110 120	0.000	1.00 C5 5 y 240 236 270	C6 x1 25 31 45	x2 24 21 24	x3 91 90 88	x4 100 95 110	C10	Heat 78.5 74.3 104.3	x1_1 7 1 11	x2_1 26 29 56	x3_1 6 15 8	x4_1 60 52 20	C16	Time 1300 1300 1300	Velocity 0.0120 0.0120 0.0115	Temperature 7.5 9.0 11.0	Yeild
Te Co x1 P P I S S S	rm Cr onstant 21 1.0 Cl ull strength 9.95 24.45 31.75 35.00	9,4 14.3 011 0.238 C2 Wire length 2 8 11 10	C3 Die height 50 110 120 550	0.000	1.00 C5 p y 240 236 270 274	C6 x1 25 31 45 60	x2 24 21 24 25	x3 91 90 88 87	x4 100 95 110 88	C10	Heat 78.5 74.3 104.3 87.6	x1_1 7 1 11 11	x2_1 26 29 56 31	x3_1 6 15 8 8	x4_1 60 52 20 47	C16	Time 1300 1300 1300 1300	Velocity 0.0120 0.0120 0.0115 0.0130	Temperature 7.5 9.0	
T4 G x1 F F F F F F F F F F F F F F F F F F	rm Cr onstant 21 1.0 C1 ull strength 9.95 24.45 31.75	9,4 14.3 011 0.238 C2 Wire length 2 8 11	3 15.38 4.25 Die height 50 110 120 550 295	0.000	1.00 C5 5 y 240 236 270	C6 x1 25 31 45	x2 24 21 24	x3 91 90 88	x4 100 95 110	C10	Heat 78.5 74.3 104.3	x1_1 7 1 11	x2_1 26 29 56	x3_1 6 15 8	x4_1 60 52 20	C16	Time 1300 1300 1300	Velocity 0.0120 0.0120 0.0115	Temperature 7.5 9.0 11.0 13.5	Yeild
Tr C/ x1 4 1 2 3 3 4 4 5 6	rm Cr onstant 21 1.0 Cl ull strength 9.95 24.45 31.75 35.00 25.02 16.86	9.4 14.3 211 0.238 C2 Wire length 2 8 111 10 8	15.38 4.25 C3 Die height 50 110 120 550 295 700	0.000 0.002 C4	1.00 CS 2 240 236 270 274 301	C6 x1 25 31 45 60 65	x2 24 21 24 25 25	x3 91 90 88 87 91	x4 100 95 110 88 94	C10	Heat 78.5 74.3 104.3 87.6 95.9	x1_1 7 11 11 11 7	x2_1 26 29 56 31 52	x3_1 6 15 8 8 8 6	x4_1 60 52 20 47 33	C16	Time 1300 1300 1300 1300 1300	Velocity 0.0120 0.0120 0.0115 0.0130 0.0135	Temperature 7.5 9.0 11.0 13.5	Yeild
Te G x1 b F F F F F F F F F F F F F F F F F F	rm Солькалт 21 1.0 С1 ull strength 9.95 24.45 31.75 35.00 25.02 16.86 H +	9.4 14.3 011 0.238 C2 Wire length 2 8 11 10 8 4 Multiple Lines	15.38 4.25 C3 Die height 50 110 120 550 295 200 ar Regression	0.000 0.002 C4	1.00 CS 2 240 236 270 274 301	C6 x1 25 31 45 60 65	x2 24 21 24 25 25	x3 91 90 88 87 91	x4 100 95 110 88 94		Heat 78.5 74.3 104.3 87.6 95.9	x1_1 7 11 11 11 7	x2_1 26 29 56 31 52	x3_1 6 15 8 8 8 6	x4_1 60 52 20 47 33		Time 1300 1300 1300 1300 1300 1300	Velocity 0.0120 0.0120 0.0115 0.0130 0.0135 0.0	Temperature 7.5 9.0 11.0 13.5	Yeld
Ti Ci X	rm <u>Ci</u> 21 1.0 <b>Ci</b> <b>ull strength</b> 9.95 24.45 31.75 35.00 25.02 16.86 H + H + <b>Multiple Lin</b>	9.4 14.3 211 0.238 C2 Wire length 2 8 11 10 8 4	C3 Die height 50 110 120 550 295 200 ar Regression	0.000 0.002 C4	1.00 CS 2 y 240 236 270 274 301 316	C6 x1 25 31 45 60 65	x2 24 21 24 25 25 25 26	x3 91 90 88 87 91 94	x4 100 95 110 88 94 94 99	1	Heat 78.5 74.3 104.3 87.6 95.9	x1_1 7 11 11 11 7	x2_1 26 29 56 31 52	x3_1 6 15 8 8 8 6	x4_1 60 52 20 47 33		Time 1300 1300 1300 1300 1300	Velocity 0.0120 0.0120 0.0115 0.0130 0.0135 0.0135	Temperature 7.5 9.0 11.0 13.5	Yeld

#### (Refer Slide Time: 03:49)

	Edit Data	Calc Stat (	Iranh View	Helo	Anistant Add	Dianal Tesh													-	Ø
		0000																		
1		I Los Los Y						× 8	长*教	0 o c	on ry i	時間								
		A LY	× 🛯 🕫 b	* 12	< 4Y 🖂 🗖	8K 1														
e	ession Analy	sis: y vers *	×																	
		R REGRESSION P	ewox.																	
		Analysis: y		1																
					Log															
		Toef SE Coef 19.4 14.3	15.38	P-Value 0.000	VIF															
					1.00															
		011 0,238																		
	and all former																			
1	lodel Sum	mary																		
		R-sq R-sq(ac				R-sq														
	15.2955 64	.41% 60.85	% 50.2	154 16	5254 5	0.14%														
	nalysis of '	Variance				0.1410														
-1	egression	Variance DF Adj SS 1 4233.41 1 4233.41	Adj MS 4233.41 4233.41		P-Value 0.002 0.002	0.1470														
- 1	egression	Variance DF Adj SS 1 4233.41	Adj MS 4233.41	F-Value 18.10	P-Value 0.002 0.002	C6	9	C8	09	C10	C11 m	C12	C13	C14	C15	C16	C17	C18	C19	
	egression cl ror C1	Variance DF Adj SS 1 4233.41 1 4233.41 10 2339.51	Adj MS 4233.41 4233.41 273.05 C3	F-Value 18.10 18.10	P-Value 0.002 0.002		C7 x2	C8 x3	C9 ×4	C10	C11 5	C12 x1_1	C13 x2_1	C14 x3_1	C15 x4_1	C16	C17 Time		C19 Temperature	2
	egression cl ror C1	Variance DF Adj SS 1 4233.41 1 4233.41 10 2330.51 C2 Wire length	Adj MS 4233.41 4233.41 273.05 C3	F-Value 18.10 18.10	P-Value 0.002 0.002 C5 g	C6				C10						C16				
	ct ct C1 ct ct ct ct ct ct ct ct ct	Variance DF Adj SS 1 4233.41 1 4233.41 10 233.61 C2 Wire length 2	Adj MS 4233.41 4233.41 733.05 C3 Die height	F-Value 18.10 18.10	P-Value 0.002 0.002 CS p y	C6 x1	x2	x3	x4	C10	Heat		x2_1	x3_1	x4_1	C16	Time	Velocity	Temperature	5
	ct C1 C1 Ull strength 9.95	Variance DF Adj SS 1 4233.41 1 4233.41 10 2340.51 C2 Wire length 2 8	Adj MS 4233,41 4233,41 233,41 233,05 C3 Die height 50	F-Value 18.10 18.10	P-Value 0.002 0.002 CS 5 Y 240	C6 x1 25	x2 24	x3 91	x4 100	C10	Heat 78.5		x2_1 26	x3_1 6	x4_1 60	C16	Time 1300	Velocity 0.0120	Temperature 7.5	5
	C1 C1 Ull strength 9.95 24.45	Variance DF Adj SS 1 4233,41 1 4233,41 10 2330,51 C2 Wire length 2 8 11	Adj MS 4233.41 4233.41 722.05 C3 Die height 50 110	F-Value 18.10 18.10	P-Value 0.002 0.002 CS p y 240 236	C6 x1 25 31	x2 24 21	x3 91 90	x4 100 95	C10	Heat 78.5 74.3	xU1 7 1	x2_1 26 29	x3_1 6 15	x4_1 60 52	C16	Time 1300 1300	Velocity 0.0120 0.0120	Temperature 7.5 9.0	5
	C1 C1 Ull strength 9.95 24.45 31.75	Variance DF Adj 55 1 423.41 1 423.41 1 23.41 1 23.41 C2 Wire length 2 8 11 10 10 10 10 10 10 10 10 10	Adj MS 4233.41 4233.41 733.05 C3 Die height 50 110 120	F-Value 18.10 18.10	P-Value 0.002 0.002 236 270	C6 x1 25 31 45	x2 24 21 24	x3 91 90 88	x4 100 95 110	C10	Heat 78.5 74.3 104.3	x1_1 7 1 11	x2_1 26 29 56	x3_1 6 15 8	x4_1 60 52 20	C16	Time 1300 1300 1300	Velocity 0.0120 0.0120 0.0115	Temperature 7.5 9.0 11.0	5
	egression (1 <b>C1</b> 9.95 24.45 31.75 35.00	Variance DF Adj 55 1 423,41 1 423,41 10 230,41 C2 Wire length 2 8 11 10 8	Adj MS 4233,41 4233,41 233,05 C3 Die height 50 110 120 550	F-Value 18.10 18.10	P-Value 0.002 0.002 y 240 236 270 274	C6 x1 25 31 45 60	x2 24 21 24 25	x3 91 90 88 87	x4 100 95 110 88	C10	Heat 78.5 74.3 104.3 87.6	x1_1 7 1 11 11	x2_1 26 29 56 31	x3_1 6 15 8 8	x4_1 60 52 20 47	C16	Time 1300 1300 1300 1300	Velocity 0.0120 0.0120 0.0115 0.0130	Temperature 7.5 9.0 11.0 13.5	5
-	egression (1 707 <b>C1</b> 24.45 31.75 35.00 25.02 16.86	Variance DF Adj 55 1 423,41 1 423,41 10 230,41 C2 Wire length 2 8 11 10 8	Adj MS 4233.41 4233.41 733.05 C3 Die height 50 110 120 550 295 200	F-Value 18.10 18.10 C4	P-Value 0.002 0.002 y 240 236 270 274 301	C6 x1 25 31 45 60 65	x2 24 21 24 25 25	x3 91 90 88 87 91	x4 100 95 110 88 94	C10	Heat 78.5 74.3 104.3 87.6 95.9	x1_1 7 11 11 11 7	x2_1 26 29 56 31 52	x3_1 6 15 8 8 8 6	x4_1 60 52 20 47 33	C16	Time 1300 1300 1300 1300 1300	Velocity 0.0120 0.0120 0.0115 0.0130 0.0135	Temperature 7.5 9.0 11.0 13.5	5
-	egression (1 <b>C1</b> <b>Ull strength</b> 9.95 24.45 31.75 35.00 25.02 16.86 ( ) +	Variance DF Adj 55 1 4233.41 1 4233.41 1 4233.41 1 233.45 C2 Wire length 2 8 111 10 8 111 10 8 4 Multiple Linea	Adj MS 4233,41 4233,41 233,05 C3 Die height 50 110 120 550 295 200 r Regression	F-Value 18.10 18.10 C4	P-Value 0.002 0.002 y 240 236 270 274 301	C6 x1 25 31 45 60 65	x2 24 21 24 25 25	x3 91 90 88 87 91	x4 100 95 110 88 94		Heat 78.5 74.3 104.3 87.6 95.9	x1_1 7 11 11 11 7	x2_1 26 29 56 31 52	x3_1 6 15 8 8 8 6	x4_1 60 52 20 47 33		Time 1300 1300 1300 1300 1300 1300	Velocity 0.0120 0.0120 0.0115 0.0130 0.0135 0.0	Temperature 7.5 9.0 11.0 13.5	5
	egression (1 <b>C1</b> <b>Ull strength</b> 9.95 24.45 31.75 35.00 25.02 16.86 ( ) +	Variance DF Adj SS 1 4233.41 1 4233.41 1 4233.41 1 0 230.61 Vire length 2 8 11 10 2 8 11 10 8 4	Adj MS 4233,41 4233,41 233,05 C3 Die height 50 110 120 550 295 200 r Regression	F-Value 18.10 18.10 C4	P-Value 0.002 0.002 240 236 270 274 301 316	C6 x1 25 31 45 60 65	x2 24 21 24 25 25	x3 91 90 88 87 91 91	x4 100 95 110 88 94 94		Heat 78.5 74.3 104.3 87.6 95.9 109.2	x1_1 7 11 11 11 7	x2_1 26 29 56 31 52	x3_1 6 15 8 8 8 6	x4_1 60 52 20 47 33		Time 1300 1300 1300 1300 1300 1300	Velocity 0.0120 0.0120 0.0115 0.0130 0.0135 0.0	Temperature 7.5 9.0 11.0 13.5	5

So, I can remove this X1 which are not significant over here and we could have regressed this one and a stepwise regression we have done none over here. So, in this case I can click ok and see what is the performance of that model. And what we are seeing is that, ok, X1 is significant that is shown over here regression coefficient and the equation is also given and, but the explained variability is very less R square adjusted is around 60.85 and tenfold cross-validation is about 50 percent.

Although, now, there is somewhat match between cross-validation and also R square adjusted somewhat close we can we can assume over here. But still I feel that there could be something more which can be done on this, ok.

		sis: y vers *	×															
	TIPLELINEA	R REGRESSION			-	Regression						×						
Rear		Analysis:		x1		C1 Pull stren	oth 🔺	Responses:										•
ie g.	content	· · ·	,		_	C2 Wreleng C3 Die heigh		У				^						
	S I	R-sq R-sq(a	dj) R-sq(p	red) 10-1	old S 10-fo	C5 y C6 x1						Y						1
15	5.2955 64.	.4196 60.8	5% 50.	24% 16	.5254	C7 x2 C8 x3		Continuous										
						C9 x4 C11 Heat		x1 x2 x3 x	4			^						
Ana	alysis of \	Variance				C12 x1_1 C13 x2_1												
Sou	irce	DF Adj SS		F-Value	P-Value	C14 x3_1 C15 x4_1						Y						
	ression	1 4233.41		18.10	0.002	C17 Time C18 Velocity	. 1	Categorical	predictors:									
x1 Erro		1 4233.41		18.10	0.002	C19 Tempera	àre 🛛					^						
	ck-of-Fit	9 2337.51		129.86	0.068	C20 Yeld (1) C22 Drive in	time (											
	ire Error	1 2.00				C23 Dose (X C24 X1_STD	9					v						
Tot	al	11 6572.92				C25 x2_STD C26 Y (frans C28 RPM	istor (		Model	Optiogs	Coding	Stepwise						
	1.61				<i></i>		1					1 40						
	CI	C2	G	C4	C5 👩	Select			Validation	Graphs	Results	Storage	C15	C16	C17	C18	C19	C
		Wire length			у								x4_1		Time		Temperature	
	ll strength				240	Help					Ök	Cancel	60		1300	0.0120	7.5	
Pul	9.95	2	50						95	74.3	1	29	15 52		1300	0.0120	9.0	1
Pul	9.95 24.45	2	110		236	31	21	90							10.00			
Pul	9.95 24.45 31.75	11	110		270	45	24	88	110	104.3	11	56	8 20		1300	0.0115	11.0	
Pul	9.95 24.45		110 120 550							104.3 87.6 95.9	11 11 7	56 31 52	8 20 8 47 6 33		1300 1300 1300	0.0115	11.0	

(Refer Slide Time: 04:12)

So, lack of fit test also shows that we are above 0.05, so in this case there is no sign of lack of fit. So, then and to avoid all this confusion what we can do is that directly we can use that stepwise regression over here. So, what you have to do is that fit regression model and go to instead of this variable you select all the variables that is X1, X2, X3, and X4 over here.

# (Refer Slide Time: 04:42)

Minitab - Untitler	1				Regression: Stepwise	×					8 ×
Regression Analy		×	x1	<ul> <li><b>∂ ∱ ↓</b> &lt;</li></ul>	Sethot: Toponie terre Datental terres: District determine 1 Bachdard eterston 2 Personal selecton with unlideton 3 Personal selecton with unlideton 3						•
\$ 15.2955 64				fold S 10-fo 6.5254							
Analysis of Source Regression x1 Error Lack-of-Fit Pure Error Total	Variance DF Adj St 1 4233.4' 1 4233.4' 10 2339.5' 9 2337.5' 1 2.00 11 6572.9' 4	4233.41 4233.41 233.95 259.72 2.00	F-Value 18.10 18.10 129.86	P-Value 0.002 0.002 0.068	E = bicket term in every nodel Alpha to regtor: 0.15 8/9/a to remove: 0.15	<ul> <li>Include term in the initial model</li> </ul>					
4 C1	C2	G	C4	C5 g			C15	C16 C17	C18	C19	C2 *
Pull strength	Wire length			У			x4_1	Time		Temperature	
1 9.95	2			240			60	1300	0.0120	7.5	
2 24.45	8			236	Herarchy		52	1300	0.0120	9.0	
3 31.75	11			270	Clisplay the table of model selection details		20	1300		11.0	
4 35.00	10							1300		13.5	
-	8 4 Multiple Line	200 ar Regression		301 316	Details about the method		33 22	1300		6	
	hear Regression here to sea				Heb O 🧮 🛱 💼 🖬 🍪 📟	QK Cancel		₩ 0 %	 D 40	sin A	2

And in stepwise options what you do is that right stepwise over here, and there will be alpha to enter alpha to remove this to you can keep, so this methodology works, this stepwise regression works like adding variables and removing variables like that simultaneously. It will work in adding and removing variables which is significant which is not significant like that.

So, that way it works and the theories can be seen in any books on stepwise regression, you have details about this stepwise regression. There are other methods of stepwise regression which is also forward selection and backward elimination these are the other two methods. But stepwise include forward selection and backward elimination both of them. So, we prefer to use stepwise regression over here. So, in this case I click ok.

### (Refer Slide Time: 05:29)

Tie fait Das Cole San Gooph View Hoje Austonet Bole 日 登 文 画 合 () () () () () () () () () () () () ()				
13 : 13 : 13 : 13 : 12 : 13 : 12 : 13 : 12 : 13 : 13				
13 : 13 : 13 : 13 : 12 : 13 : 12 : 13 : 12 : 13 : 13				
Nagression Analysis: y versus x1, x2, x3, x4           y = 13 + 6373 x1 + 8.88 x2           Conflicients           Term         Cold SC Cold T Yable Proble VIF           Center:         19				
w.u.thrutunew.recession.wink           Regression Analysis: y versus x1, x2, x3, x4           y = 19+0.371 x1+9.88 x2           Conflicients           Term         Coef SE Coef T Yake VVF           Constant: 30         81.9         0.55         0.83				
Coefficients         Coefficients           Term         Coef SE Coef         T-Value         P-Value         V/IF				
y = 3.9+0.573 x1+9.88 x2 Coefficients Term Coef 5E Coef 1*Value P-Value V/IF Contact 1.9 81.9 0.05 0.963				- 6
Coefficients Term Coef 5E Coef T-Value P-Value V/IF Contact 10 81.9 0.05 0.093				
Cert         SE Coef         T-Value         P-Value         V/S           Contraint         3.9         81.9         0.05         0.963				
Constant 3.9 81.9 0.05 0.963				
x1 0.573 0.250 2.29 0.048 1.77				
x2 9.88 3.72 2.66 0.026 1.77				
S         R-sq, R-sq, R-sq, Rolp, R-sq, Rolp, S         No-fold R-sq, R-sq, Rolp, R-sq, Rolp, R-sq, Rolp, R-sq, Rolp, R-sq, Rolp, Rol				
C1 C2 C3 C4 C5 g C6 C7 C8 C9 C10 C11 g C12 C13 C14 C15 C16	C17	C18	C19	
Pull strength Wire length Die height         y         x1         x2         x3         x4         Heat         x1_1         x2_1         x3_1         x4_1	Time	Velocity	Temperature	Ye
9.95 2 50 240 25 24 91 100 78.5 7 26 6 60	1300	0.0120	0 7.5	
24.45 8 110 236 31 21 90 95 74.3 1 29 15 52	1300			
31.75 11 120 270 45 24 88 110 104.3 11 56 8 20	1300			
35.00 10 550 274 60 25 87 88 87.6 11 31 8 47	1300			
25.02 8 295 301 65 25 91 94 95.9 7 52 6 33	1300		170	
16.86 4 200 316 72 26 94 99 109.2 11 55 9 22	1300	0.0		
4 b H + Multiple Linear Regression.mwx		1	100	
	ш П			
1 Multiple Linear Regression.mmx	- ē 99		A COL	

And then if you click ok over here what will happen is that you will you will find a scenario over here which is two variables are entered into the model that is X1. And you can see that X1 is also prominent over here 0.048.

And the other one is X2 is more prominent which is P value is 0.026 both are significant and less than 0.05 over here. And also R square adjusted has improved significantly from 60 to 75 over here and this tenfold cross-validation from 50 to 57 it has improved over here.

(Refer Slide Time: 05:58)

0	Edit Data				Assistant Add															
									11 S 11 M	50 : m .m	m									
ł								208	K×4	81 d' C	On the state	83								
		€ II Y		* 2	4 Y 🖂 🗖	BK #														
Re	gression Analysi	is: y vers ×	×																	
8	MULTIPLE UNEAP	REGRESSION	WWX																	
Re	gression A	Analysis:	versus :	x1. x2.	x3. x4															
	S F	R-sq R-sq(at	ij) R-sq(pr	ed) 10-f	old S 10-fold	d R-sq														
1	12.0724 80.0	04% 75.61	% 55.8	5% 15.	3323 5	57.08%														
	Analysis of V	ariance																		
			Adj MS F-	Value D.	-Value															
	Jource																			
1	Regression	2 5261.2	2630.6		0.001															
	Regression x1	2 5261.2 1 766.2	2630.6 766.2	18.05 5.26	0.001 0.048															
		1 766.2 1 1027.8		5.26																
	x1 x2 Error	1 766.2 1 1027.8 9 1311.7	766.2	5.26	0.048															
	x1 x2 Error	1 766.2 1 1027.8	766.2 1027.8	5.26	0.048															
	x1 x2 Error	1 766.2 1 1027.8 9 1311.7	766.2 1027.8	5.26	0.048		6													
	x1 x2 Error	1 766.2 1 1027.8 9 1311.7	766.2 1027.8 145.7	5.26 7.05	0.048								Ve	reue Cite						
	x1 x2 Error	1 766.2 1 1027.8 9 1311.7	766.2 1027.8 145.7	5.26 7.05	0.048 0.026 hility Plot	C6	6	C8	0	C10	C11 👦 C	12	C13	rene Fite C14	C15	C16	C17	C18	C19	
	x1 x2 Error Total	1 766.2 1 1027.8 9 1311.7 11 6572.9 C2	766.2 1027.8 145.7 Norma C3	5.26 7.05	0.048 0.026 hility Plot	_		C8 x3	C9 x4			12			C15 x4_1	C16	C17 Time		C19 Temperature	
	x1 x2 Error Total	1 766.2 1 1027.8 9 1311.7 11 6572.9 C2	766.2 1027.8 145.7 Norma C3	5.26 7.05	0.048 0.026 hility Plat	C6	<b>C7</b>						C13	C14		C16			Temperature	
	x1 x2 Error Total C1 Pull strength	1 766.2 1 1027.8 9 1311.7 11 6572.9 C2 Wire length	766.2 1027.8 145.7 Norma C3 Die height	5.26 7.05	0.048 0.026 hility Plot C5 g	C6 x1	C7 x2	x3	x4		leat x		C13 x2_1	C14 x3_1	x4_1	C16	Time	Velocity	Temperature 7.5	Ye
	x1 x2 Error Total C1 Pull strength 9.95	1 766.2 1 1027.8 9 1311.7 11 6572.9 C2 Wire length 2	766.2 1027.8 145.7 Norma C3 Die height 50	5.26 7.05	0.048 0.026 hility Plot C5 12 y 240	C6 x1 25	C7 x2 24	x3 91	x4 100		1eat x1 78.5		C13 x2_1 26	C14 x3_1 6	<b>x4_1</b> 60	C16	Time 1300	Velocity 0.0120	Temperature 7.5 9.0	Ye
	x1 x2 Error Total Pull strength 9.95 24.45	1 766.2 1 1027.8 9 1311.7 11 6572.9 <b>C2</b> Wire length 2 8	766.2 1027.8 145.7 Norma C3 Die height 50 110	5.26 7.05	0.048 0.026 hility Plat C5 p y 240 236	C6 x1 25 31	C7 x2 24 21	x3 91 90	x4 100 95		1eat x1 78.5 74.3	U 7 1	C13 x2_1 26 29	C14 x3_1 6 15	x4_1 60 52	C16	Time 1300 1300	Velocity 0.0120 0.0120	Temperature 7.5 9.0 11.0	Ye
	x1 x2 Error Total Pull strength 9.95 24.45 31.75	1 766.2 1 1027.8 9 1311.7 11 6572.9 C2 Wire length 2 8 11	766.2 1027.8 145.7 C3 Die height 50 110 120	5.26 7.05	0.048 0.026 hility Diot C5 12 y 240 236 270	C6 x1 25 31 45	C7 x2 24 21 24	x3 91 90 88	x4 100 95 110		feat x1 78.5 74.3 104.3	U 7 1 11	C13 x2_1 26 29 56	C14 x3_1 6 15 8	x4_1 60 52 20	C16	Time 1300 1300 1300	Velocity 0.0120 0.0120 0.0115	Temperature 7.5 9.0 11.0 13.5	Ye
	x1 x2 Error Total Pull strength 9,95 24,45 31,75 35,00	1 766.2 1 1027.8 9 1311.7 11 6572.9 C2 Wire length 2 8 11 10	766.2 1027.8 145.7 C3 Die height 50 110 120 550	5.26 7.05	0.048 0.026 hility Diat C5 12 y 240 236 270 274	C6 x1 25 31 45 60	C7 x2 24 21 24 25	x3 91 90 88 87	x4 100 95 110 88		leat x1 78.5 74.3 104.3 87.6	U 7 1 11 11	C13 x2_1 26 29 56 31	C14 x3_1 6 15 8 8	x4_1 60 52 20 47	C16	Time 1300 1300 1300 1300	Velocity 0.0120 0.0120 0.0115 0.0130 0.0135	Temperature 7.5 9.0 11.0 13.5	Ye
	x1 x2 Eror Pull strength 9.95 24.45 31.75 35.00 25.02 21.6.86	1 766.2 1 1027.8 9 1311.7 11 6572.9 Vire length 2 8 11 10 8	766.2 1027.8 145.7 C3 Die height 50 110 120 550 295 200	5.26 7.05	0.048 0.026 hility Plot C5 12 y 240 236 270 274 301	C6 x1 25 31 45 60 65	C7 x2 24 21 24 25 25	x3 91 90 88 87 91	x4 100 95 110 88 94		leat x1 78.5 74.3 104.3 87.6 95.9	U 7 1 11 11 11 7	C13 x2_1 26 29 56 31 52	C14 x3_1 6 15 8 8 6	x4_1 60 52 20 47 33	C16	Time 1300 1300 1300 1300 1300	Velocity 0.0120 0.0120 0.0115 0.0130 0.0135	Temperature 7.5 9.0 11.0 13.5	Ye
	x1 x2 Error Total Pull strength 9.95 24.45 33.50 25.02 16.86 b H +	1 766.2 1 1027.8 9 1311.7 11 6572.9 C2 Wire length 2 8 11 10 8 4	766.2 1027.8 145.7 C3 Die height 50 110 120 550 295 200 r Regression.	5.26 7.05	0.048 0.026 hility Plot C5 12 y 240 236 270 274 301	C6 x1 25 31 45 60 65	C7 x2 24 21 24 25 25	x3 91 90 88 87 91	x4 100 95 110 88 94	•	leat x1 78.5 74.3 104.3 87.6 95.9	U 7 1 11 11 11 7	C13 x2_1 26 29 56 31 52	C14 x3_1 6 15 8 8 6	x4_1 60 52 20 47 33		Time 1300 1300 1300 1300 1300	Velocity 0.0120 0.0120 0.0115 0.0130 0.0135 0.0	Temperature 7.5 9.0 11.0 13.5	

So, in this case, we can say that this may be the best model that we have considered over here and we should go we should go about go in implementing this one.

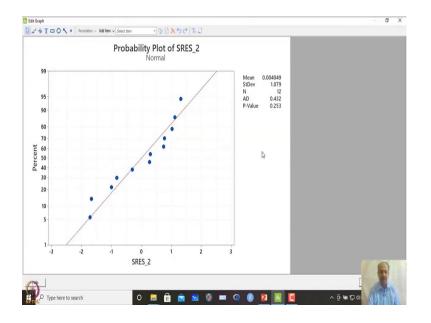
(Refer Slide Time: 06:14)

	tab - Untitled Edit Data Calc S	tat Graph View	Helo A	Arristant A	dditional Tool												-	8 >
		Basic Statistics	riep /		play Description													
		Regression ANOVA			re Descriptive phical Summ			长*药	10 C	<b></b>								
	ession Analysis: y ve	DOE Control Charts Quality Tools		· 1-S	ample Z ample t													
	ression Anal	Reliability/Survi Predictive Anal	ival	a state	ample t ved t													•
	<ul> <li>3.9 + 0.573 x1</li> <li>sefficients</li> </ul>	Multivariate Time Series Tables		, 남 2P	oportion oportions ample Poissor ample Poissor													*
-	rm Coef : instant 3.9 0.573	Nonparametric Equivalence Tes Power and Sam	ts	• 1 × 2-5	niance	1 Ndte												
×2	9.88	3.72 2.66	0.026	1./ 11 60	relation													
M	odel Summary	sa(adi) R-sa(a	red) 10-fc	at Co	mailty Test	De												
	S R-sq R	sq(adj) R·sq(p 75.61% 55.	red) 10-fo 85% 15.	0 <sup>2</sup> Co 1 No 01d - Ou 332		*0	50 <b>n</b>											
1	S R-sq R			α <sup>2</sup> Co ald No ald - Ou λ Go	mality Test Jier Test odness-of-Fit	*0	son C8	0	C10 (	11 😰 C12	C13	C14	C15	C16	C17	C18	C19	
1	S R-sq R 2.0724 80.04%	75.61% 55. C3	85% 15.	$a^2$ Co and $\lambda$ No and $\lambda$ Gov	mality Test Jier Test odness-of-Fit	Test for Poiss	_	C9 x4		111 😰 C12 Pat x1_1	C13 x2_1	C14 x3_1	C15 x4_1	C16	C17 Time		C19 Temperature	c
1 PL	S R-sq R 2.0724 80.04% C1 C2	75.61% 55. C3	85% 15.	0 <sup>2</sup> Ce 0 <sup>1</sup> No 332 - Ου λ Go C5 g	mality Test dier Test odness-of-Fit C6	Test for Poiss	C8			nat x1_1				C16				C. Yeil
1	S         R+sq         R           2.0724         80.04%            C1         C2         C2           ull strength         Wire let	75.61% 55. C3 ngth Die height	85% 15.	0 <sup>2</sup> Co 0 0 0 0 0 0 0 0 0 0 0 0 0	maity Test dier Test odness-of-Fit C6 x1	Test for Poiss C7 X2	C8 x3	x4		sat x1_1	x2_1	x3_1	x4_1	C16	Time	Velocity	Temperature	C. Yell
1 PL	S         R-sq         R           2.0724         80.04%         20           C1         C2         C2           ull strength         Wire let         9.95	75.61% 55.7 C3 ngth Die height 2 50	85% 15.	α <sup>2</sup> Co A No 332 A Go CS S y 240	maity Test dier Test odness-of-Fit C6 x1 25	Test for Poiss C7 x2 24	C8 x3 91	<b>x4</b>	н	rat x1_1 78.5	x2_1 7 26 1 29	x3_1 6	<b>x4_1</b> 60	C16	Time 1300	Velocity 0.0120 0.0120	Temperature 7.5	Ci Yeik
1 PL	S         R-sq         R           2.0724         80.04%         20.04%           C1         C2         C2           Jll strength         Wire let         9.95           24.45         24.45         24.45	75.61% 55.7 C3 ngth Die height 2 50 8 110	85% 15.	α <sup>2</sup> Co A No 332 λ Go CS 2 y 240 236	maity Test dier Test odness-of-Fit C6 x1 25 31	C7 x2 24 21	C8 x3 91 90	x4 100 95	н	rat x1_1 78.5 74.3	x2_1 7 26 1 29 1 56	x3_1 6 15	x4_1 60 52	C16	Time 1300 1300	Velocity 0.0120 0.0120 0.0115	Temperature 7.5 9.0	Ci Yeik
1 PL	S         R-sq         R           2.0724         80.04%         0.04%           C1         C2         Wire les           9.95         24.45         31.75	75.61% 55.7 ngth Die height 2 50 8 110 11 120	85% 15.	α <sup>2</sup> Co A No 332 → Ou λ Go CS s y 240 236 270	mailty Test lier Test cdness-of-Fit C6 x1 25 31 45	C7 K2 24 21 24	C8 x3 91 90 88	x4 100 95 110	н	eat x1_1 78.5 74.3 104.3 1 87.6 1	x2_1 7 26 1 29 1 56	x3_1 6 15 8	x4_1 60 52 20	C16	Time 1300 1300 1300	Velocity 0.0120 0.0120 0.0115 0.0130	Temperature 7.5 9.0 11.0	Yeild
1 PL 1 2 3 3 4 5 5	S         R:sq         R           2.0724         80.04%            C1         C2            Jll strength         Wire les         9.95           24.45         31.75            35.00         25.02            16.86	75.61% 55.7 C3 Die height 2 50 8 110 11 120 10 550	C4	α <sup>2</sup> Co Λ No 332 λ Go CS g y 240 236 270 274	mailty Test fier Test odness-of-Fit C6 x1 25 31 45 60	C7 X2 24 21 24 25	C8 x3 91 90 88 87	x4 100 95 110 88	н	eat x1_1 78.5 74.3 104.3 1 87.6 1	x2_1 7 26 1 29 1 56 1 31 7 52	x3_1 6 15 8 8	x4_1 60 52 20 47	C16	Time 1300 1300 1300 1300	Velocity 0.0120 0.0120 0.0115 0.0130 0.0130	Temperature 7.5 9.0 11.0 13.5	C2 Yeild
1	S         R:sq         R           2.0724         80.04%            C1         C2            Jll strength         Wire les         9.95           24.45         31.75            35.00         25.02            16.86	75.61% 55. C3 Die height 2 50 8 110 11 120 10 550 8 295 4 200 Linear Regression	C4	α <sup>2</sup> Co A No 332 λ Go CS 2 y 240 236 270 274 301	mailty Test fier Test odness-of-Fit C6 x1 25 31 45 60 65	C7 x2 24 21 24 25 25	C8 x3 91 90 88 87 91	x4 100 95 110 88 94	H	rat x1_1 78.5 74.3 104.3 1 87.6 1 95.9	x2_1 7 26 1 29 1 56 1 31 7 52	x3_1 6 15 8 8 6	x4_1 60 52 20 47 33		Time 1300 1300 1300 1300 1300	Velocity 0.0120 0.0120 0.0115 0.0130 0.0130 0.0135	Temperature 7.5 9.0 11.0 13.5	C2 Yeild

(Refer Slide Time: 06:17)

			201	sk *	時回!#	<b>™</b> 0 ₩ 1			K * * 8	<b>d<sup>m</sup> t<sup>rag</sup> d<sup>m</sup> - Y</b>	专图							
Ret	ression Analysis:		-	. <b>.</b>		B L												
	ULTIPLE LINEAR R																	
-	gression An			x1, x2, )	3, x4													
>	= 3.9 + 0.573	x1 + 9.88 x2					Normality Te	est				×						
0 ~ ~	Term Coe Constant 3.1 1 0.573 2 9.81 Aodel Summa	81.9 0.250 3.72	0.05 2.29 2.66	0.963 0.048 0.026 Md) 10-fc	ld S 10-fol	<b>i R-sq</b> 7.08%	Selet	ž	None     At Y values:     At data value     At data value     Tests for Normalit     Anderson-Dar     Ryan-Joiner     Kolmogorov-G  Tide:	r ing (Similar to Shapin	-Wik)							
-	S R+s	q R-sq(ad) 6 75.619		5% 15.3								- 1-						
-				5% 15.3 C4	C5	C6	Help	1		OK	Ca	ncel	C14	C15	C16 C17	C18	C19	-
-	12.0724 80.04	6 75.619 C2	55.8 C3		C5 g	C6 x1	Help x2	x3	x4	OK	Ca x1_1	x2_1	C14 x3_1	C15 x4_1	C16 C17 Time		C19 Temperature	
-	12.0724 80.04 C1	6 75.619 C2	55.8 C3					x3 91	x4 100		_					Velocity	Temperature	Y
-	C1 Will strength	C2 fire length	C3 Nie height		у	xt	x2			Heat	_	x2_1		x4_1	Time	Velocity 0.0120	Temperature 7.5	Y
-	C1 Pull strength 9.95	6 75.614 C2 ire length 1 2	C3 Ne height 50		y 240	<b>x1</b> 25	x2 24	91	100	Heat 78.5	_	<b>x2_1</b> 26	<b>x3_1</b> 6	x4_1 60	Time 130	Velocity 0 0.0120 0 0.0120	Temperature 7.5 9.0	Ye
	C1 Pull strength 9.95 24.45	6 75.619 C2 ire length 1 2 8	C3 Nie height 50 110		y 240 236	x1 25 31	x2 24 21	91 90	100 95	Heat 78.5 74.3	x1_1 7 1	x2_1 26 29	x3_1 6 15	<b>x4_1</b> 60 52	Time 130 130	Velocity 0 0.0120 0 0.0120 0 0.0120	Temperature 7.5 9.0 11.0	Ye
-	C1 V 9.95 24.45 31.75	6 75.619 C2 ire length 2 8 11	C3 Nie height 50 110 120		y 240 236 270	x1 25 31 45	x2 24 21 24	91 90 88	100 95 110	Heat 78.5 74.3 104.3	x1_1 7 1	x2_1 26 29 56	x3_1 6 15 8	x4_1 60 52 20	Time 130 130 130	Velocity 0 0.0120 0 0.0120 0 0.0115 0 0.0130	Temperature 7.5 9.0 11.0	
-	C1 W 9.95 24.45 31.75 35.00 25.02 16.86	6 75.619 C2 ire length 1 2 8 11 10	C3 Nie height 50 110 120 550 295 200	C4	y 240 236 270 274	x1 25 31 45 60	x2 24 21 24 25	91 90 88 87	100 95 110 88	Heat 78.5 74.3 104.3 87.6	x1_1 7 1 11 11	x2_1 26 29 56 31	x3_1 6 15 8 8	x4_1 60 52 20 47	Time 130 130 130 130	Velocity 0 0.0120 0 0.0120 0 0.0115 0 0.0130 0 0.0130	Temperature 7.5 9.0 11.0 13.5	Y

#### (Refer Slide Time: 06:22)



So, this may be this model that we have generated over here. Now, only thing is that residual whether it is normal distribution we can check by normality test over here. And the residual will be saved at the end values that we see over here residual 2 that is the final residual after we have run the models. And the residual also says that there is no abnormality in the normality assumptions that we are considering in the residuals over here.

0.253 is more than 0.05, so in this case residual also satisfies. So, there is no problem. So, if I consider X1 and X2 in the model with y. We are getting X1, X2 significant and also the model adequacy test are quite there is no significant deviation from that, ok.

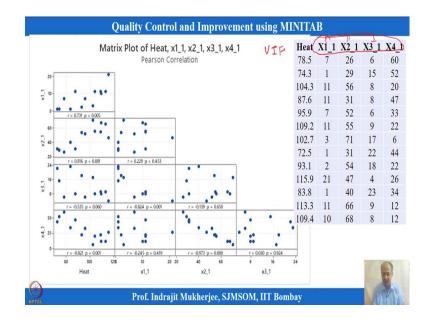
So, this is one of the scenario where we can use stepwise regression method. But for the second case if we consider this one that heat and X1, X2 again the confusion comes, again the confusion will arise because and we can see that one. So, what we will do is that I will take this second one where heat is the y characteristics and that has to be regressed with other variables over here.

# (Refer Slide Time: 07:14)

⊟ 5·0 ፼ -	Week 4 NPTEL Quality PPT 24 Dec (Revised) - PowerPoint (Proc	uct Activation Failed)	1 – 0 ×
File Home Insert Design Transitions			Sign in 🧏 Share
K Cut     Copy -     Stee     Ste	20         A* A*         ♦         Ξ + Ξ + Ξ         = Ξ         = Ξ         = Ξ         = Ξ <th></th> <th>ind leplace * elect *</th>		ind leplace * elect *
Clipboard rs Slides	Font 13 Paragraph	S Drawing S E	dting
billing -5	Quality Control and Impr	ovement using MINITAB	
	Coefficients	Analysis of Variance	
1 1 1997	Term Coef SE Coef T-Value P-Value VIF	Source DF Adj SS Adj MS F-Value P-Value	
	Constant -123 157 -0.78 0.459	Regression 4 5600.5 1400.1 10.08 0.005	
Advance of systems of a street	x1 0.757 0.279 2.71 0.030 2.32	x1 1 1022.8 1022.8 7.36 0.030	
(sets General Representation (BRIS))	x2 7.52 4.01 1.87 0.103 2.16	x2 1 488.4 488.4 3.52 0.103	
( <u>*</u> )	x3 2.48 1.81 1.37 0.212 1.34	x3 1 261.6 261.6 1.88 0.212	
Mari di Bella Changano Bella Changa Aliang Aliang Aliang Aliang Aliang Aliang Aliang Aliang Aliang	x4 -0.481 0.555 -0.87 0.415 1.01	x4 1 104.3 104.3 0.75 0.415 Erryer 7 972.5 138.9	
	Model Summary	Total 11 6572.9	
Biology and States	S R-sq R-sq(adj) R-sq(pred)	Regression Equation	
	11.7866 85.20% 76.75% 52.48%	y = -123 + 0.757 x1 + 7.52 x2 + 2.48 x3 - 0.481 :	x4
Landeline Will Alley 1			
Z°CA .	Prof. Indrajit Mukherjee	SJMSOM, IIT Bombay	
56 []8 English (India)		≜ Notes ♥ Comments 🔠 😚 🕸 🛡	
P Type here to search	0 📑 🔒 💼 😼 🚳 📼 🕥	😰 🛐 🔚 🧧 🔺 🙀	A la de la

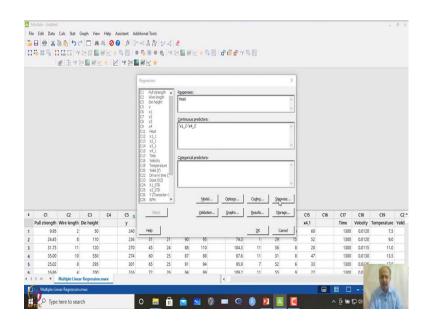
So, this is the first example that we have selected and these are the values that we have seen X1, X2, this is with X3.

(Refer Slide Time: 07:22)



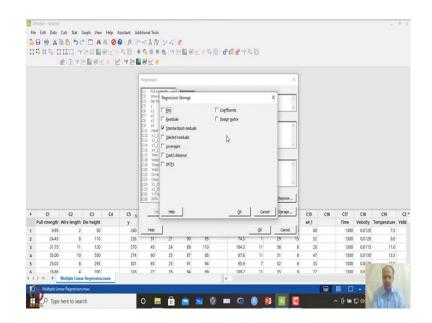
So this is now, this is the variables that we are talking about heat is with X1, X2, X3, and X4, 4 variables we are trying to regress.

# (Refer Slide Time: 07:40)



And then what we can do is that I can go to stat, and I can go to regression and then regression and fit regression models over here, and instead of this I give heat and then I give X1 to X4 variables I select this one. And I do not do stepwise regression at initially. So, I want to see what are the results.

(Refer Slide Time: 08:01)



#### (Refer Slide Time: 08:05)

		d																			8
File	Edit Data	Calc	Stat Gra	ph View	Help /	Assistant Add	itional Tool	5													
						0   fx ]*															
										U + M	50 .0	- anto	14. 59								
•	·0 = -0 1								CIP	ET X M	83 D.	<b>CO D</b> 1	-y 83								
		· @   .:	E 17 🖻	BR	. <b>.</b>	4 Y 🖂 🗖	BK 1														
Re	gression Analy	ysis: Hea	t * X																		
B	MULTIPLE LINE	AR REGRI	ESSION MV	x																	
Re	gression	Analy	ysis: He	eat vers	us x1_	1, x2_1, x3	1, x4 1														,
	•				-																
1	Method																				
	Cross-validati	on 10-	fold																		
h	Regression	Equat	ion																		
	Heat = 62	4+1.55	$1 \times 1 + 0$																		
				310 X2_1 *	0.102 x3	1 · 0.144 X4_1															
				510,82_1 *	0.102 x3	_1 • 0.144 X4_1	¢	8													
	Coefficient			510.42_1	0.102 x3	_1 • 0.144 X4_1		8													
1		s	SE Coef		P-Value	_1 - 0, 144 X4_1		8													
1	Coefficient	s						8													
1	Coefficient Term Constant x1_1	s 62.4 1.551	SE Coef 70.1 0.745	T-Value 0.89 2.08	P-Value 0.399 0.071	VIF 38.50		8													
1	Coefficient Term Constant ×1_1	S Coef 62.4	SE Coef 70.1	T-Value 0.89	P-Value 0.399 0.071	VIF		8													
	Coefficient Term Constant x1_1	s 62.4 1.551	SE Coef 70.1 0.745 0.724	T-Value 0.89 2.08	P-Value 0.399 0.071	VIF 38.50		6	C8	09	C10	CIT E	C12	C13	C14	C15	C16	C17	C18	C19	
	Coefficient Term Constant x1_1 x2_1	s 62.4 1.551 0.510	SE Coef 70.1 0.745 0.724	T-Value 0.89 2.08 0.70 C3	P-Value 0.399 0.071 0.501	VIF 38.50 254.42	C		C8 x3	C9 x4	C10	C11 12 Heat	C12	C13 x2_1	C14 x3_1	C15 x4_1	C16	C17 Time	Velocity	C19 Temperature	
	Coefficient Term Constant ×1_1 ×2_1 C1 Pull strength 9.93	5 62.4 1.551 0.510 C Wire I	SE Coef 70.1 0.745 0.724	T-Value 0.89 2.08 0.70 C3	P-Value 0.399 0.071 0.501	VIF 38.50 254.42 C5 g	C6	<i>a</i> -			C10	Heat 78.5				<b>x4_1</b> 60	C16		Velocity 0.0120		Yei
	Coefficient Term Constant x1_1 x2_1 C1 Pull strength 9.99 24.43	5 62.4 1.551 0.510 C Wire I 5 5	SE Coef 70.1 0.745 0.724 2 kength Di 2 8	T-Value 0.89 2.08 0.70 C3 ie height	P-Value 0.399 0.071 0.501	VIF 38.50 25442 <b>C5</b> 2 <b>y</b> 240 236	C6 x1 25 31	C7 = x2 24 21	<b>x3</b> 91 90	×4	C10	Heat 78.5 74.3		x2_1 26 29	x3_1 6 15	x4_1 60 52	C16	Time	Velocity 0.0120 0.0120	Temperature	Yei
	Coefficient Term Constant x1_1 x2_1 C1 Pull strength 9.99 24.42 31.75	5 62.4 1.551 0.510 Wire I 5 5	SE Coef 70.1 0.745 0.724 2 length Di 2 8 11	T-Value 0.89 2.08 0.70 C3 ie height 50 110 120	P-Value 0.399 0.071 0.501	VIF 38.50 254.42 <b>C5</b> <b>2</b> <b>2</b> 240 236 270	C6 x1 25 31 45	C7 x2 24 21 24	<b>x3</b> 91 90 88	x4 100 95 110	C10	Heat 78.5 74.3 104.3	xU1 7 1	x2_1 26 29 56	x3_1 6 15 8	x4_1 60 52 20	C16	Time 1300 1300 1300	Velocity 0.0120 0.0120 0.0115	Temperature 7.5 9.0 11.0	Yei
	Coefficient Term Constant x1_1 x2_1 C1 Pull strength 9.99 24.43 31.72 35.00	S Coef 62.4 1.551 0.510 Wire I S S S D	SE Coef 70.1 0.745 0.724 2 kength Di 2 8 11 10	T-Value 0.89 2.08 0.70 C3 ie height 50 110 120 550	P-Value 0.399 0.071 0.501	VIF 38.50 254.42 <b>C5</b> p <b>y</b> 240 236 270 274	C6 x1 25 31 45 60	C7 = x2 24 21 24 25	<b>×3</b> 91 90 88 87	x4 100 95 110 88	C10	Heat 78.5 74.3 104.3 87.6	xt_1 7 1 11 11	x2_1 26 29 56 31	x3_1 6 15 8 8	x4_1 60 52 20 47	C16	Time 1300 1300 1300 1300	Velocity 0.0120 0.0120 0.0115 0.0130	Temperature 7.5 9.0 11.0 13.5	Yei
	Coefficient Term Constant x1_1 x2_1 C1 Pull strength 9.90 24.40 31.72 35.00 25.00 25.00	S Coef 62.4 1.551 0.510 C Wire I 5 5 5 2	SE Coef 70.1 0.745 0.724 2 kength Di 2 8 11 10 8	T-Value 0.89 2.08 0.70 C3 ie height 50 110 120 550 295	P-Value 0.399 0.071 0.501	VIF 38.50 254.42 CS g y 240 236 270 274 301	C6 x1 25 31 45 60 65	C7 = x2 24 21 24 25 25	x3 91 90 88 87 91	x4 100 95 110 88 94	C10	Heat 78.5 74.3 104.3 87.6 95.9	x1_1 7 11 11 11 7	x2_1 26 29 56 31 52	x3_1 6 15 8 8 6	x4_1 60 52 20 47 33	C16	Time 1300 1300 1300 1300 1300	Velocity 0.0120 0.0120 0.0115 0.0130 0.0135	Temperature 7.5 9.0 11.0	Yei
	Coefficient Term Constant x1_1 x2_1 C1 Pull strength 9.92 24.42 31.72 35.00 25.00 25.00 16.84	S Coef 62.4 1.551 0.510 C 1 Wire I 5 5 5 5 5 5 5 5 5 5 5 5 5	SE Coef 70.1 0.745 0.724 2 kength Di 2 8 11 10 8 4	T-Value 0.89 2.08 0.70 C3 ie height 50 110 120 550 295 200	P-Value 0.399 0.071 0.501 C4	VIF 38.50 254.42 <b>C5</b> p <b>y</b> 240 236 270 274	C6 x1 25 31 45 60	C7 = x2 24 21 24 25	<b>×3</b> 91 90 88 87	x4 100 95 110 88		Heat 78.5 74.3 104.3 87.6	xt_1 7 1 11 11	x2_1 26 29 56 31	x3_1 6 15 8 8	x4_1 60 52 20 47	C16	Time 1300 1300 1300 1300	Velocity 0.0120 0.0120 0.0115 0.0130	Temperature 7.5 9.0 11.0 13.5	Yei
1 2 3 4 5	Coefficient Term Constant x1_1 x2_1 C1 Pull strength 9.90 24.40 31.72 35.00 25.00 25.00	S Coef 62.4 1.551 0.510 C 1 Wire I 5 5 5 5 5 5 5 5 5 5 5 5 5	SE Coef 70.1 0.745 0.724 2 kength Di 2 8 11 10 8 4	T-Value 0.89 2.08 0.70 C3 ie height 50 110 120 550 295	P-Value 0.399 0.071 0.501 C4	VIF 38.50 254.42 CS g y 240 236 270 274 301	C6 x1 25 31 45 60 65	C7 = x2 24 21 24 25 25	x3 91 90 88 87 91	x4 100 95 110 88 94	C10	Heat 78.5 74.3 104.3 87.6 95.9	x1_1 7 11 11 11 7	x2_1 26 29 56 31 52	x3_1 6 15 8 8 6	x4_1 60 52 20 47 33	C16	Time 1300 1300 1300 1300 1300	Velocity 0.0120 0.0120 0.0115 0.0130 0.0135	Temperature 7.5 9.0 11.0 13.5	Yei
	Coefficient Term Constant x1_1 x2_1 C1 Pull strength 9.92 24.42 31.72 35.00 25.00 25.00 16.84	S Coef 62.4 1.551 0.510 C 0 1 Wire l 5 5 5 5 2 2 5 5 5 5 5 5 5 5 5 5 5 5 5	SE Coef 70.1 0.745 0.724 2 kength Di 2 8 11 10 8 4 4 4 4 4 4 10 8	T-Value 0.89 2.08 0.70 C3 ie height 50 110 120 550 295 200 tegression.	P-Value 0.399 0.071 0.501 C4	VIF 38.50 254.42 CS g y 240 236 270 274 301	C6 x1 25 31 45 60 65	C7 = x2 24 21 24 25 25	x3 91 90 88 87 91	x4 100 95 110 88 94		Heat 78.5 74.3 104.3 87.6 95.9	x1_1 7 11 11 11 7	x2_1 26 29 56 31 52	x3_1 6 15 8 8 6	x4_1 60 52 20 47 33		Time 1300 1300 1300 1300 1300	Velocity 0.0120 0.0120 0.0115 0.0130 0.0135	Temperature 7.5 9.0 11.0 13.5	Yei
	Coefficient Term Constant x1_1 x2_1 CI Pull strength 9.90 24.40 31.70 35.00 25.00 16.8/ р. н. +  25.00 25.00 16.8/ р. н. +  26.00 26.00 26.00 27.00 26.00 27.00 26	S Coef 62.4 1.551 0.510 C C C C C S S S S Multip Inter Res	SE Coef 70.1 0.745 0.724 2 kength Di 2 8 11 10 8 4 4 4 4 4 4 10 8	T-Value 0.89 2.08 0.70 C3 ie height 50 110 120 550 295 200 tegression.	P-Value 0.399 0.071 0.501 C4	VIF 38.50 254.42 <b>C5</b> 9 <b>y</b> 240 236 270 274 301 316	C6 x1 25 31 45 60 65	C7 x2 24 21 24 25 25 25 26	x3 91 90 88 87 91 94	x4 100 95 110 88 94 94	1	Heat 78.5 74.3 104.3 87.6 95.9	xU1 7 11 11 11 7 11	x2_1 26 29 56 31 52	x3_1 6 15 8 8 8 6	x4_1 60 52 20 47 33		Time 1300 1300 1300 1300 1300 1300	Velocity 0.0120 0.0120 0.0115 0.0130 0.0115 0.01	Temperature 7.5 9.0 11.0 13.5	Ye

So, what will happen is that I have all the criteria. So, and I will store let us say the residuals over here and I click ok.

(Refer Slide Time: 08:08)

le Edit																				8
	Data Cal	: Stat Gr	ph View	Help /	Assistant Add	ditional Tools														
	XD	1 50		10 0	0 fx 3-	A. 1. 20	and a													
					時間 #				I S A M	50 : .m .	m	W NO								
.0								0.00	E × 4	SE : D. C	0.1.	99 BB								
	æ	Ξ чY ≥	BR	* 2	< 'Y 🖂 🗖	8K×														
erression	n Analysis: H	eat_ v >																		
	E LINEAR RE																			D
					1, x2_1, x3	5_1, X4_1														
Term		SE Coef			VIF															
Constar x1_1	1.551	70.1	0.89	0.399	38.50															
x1_1 x2_1	0.510		0.70	0.501	38.50															
	0.102	0.755	0.14	0.896	46.87															
x3.1																				
x4_1 Model	-0.144 Summar	y R-sq(adj)	-0.20 R-sq(pre 95.9-	ed) 10-f	282.51 old S 10-foli 0772 S	<mark>d R-sq</mark> 25.95%														
2.4460 Analys	-0.144 Summar S R-sq D1 98.24% sis of Vari	y R-sq(adj) 97.36% ance	R∙sq(pro 95.9	ed) 10-fr 4% 2.9	old S 10-foli 0772 S															
x4_1 Model	-0.144 Summar S R-sq D1 98.24% Sis of Vari	y R-sq(adj) 97.36% ance	R+sq(pre	ed) 10-fr 4% 2.9	old S 10-fold 0772 S P-Value	95.95%	9	6	<b>C9</b>	C10	Ct1 r	C12	C13	C14	C15	C16	C17	C18	C19	
x4_1 Model 2.4460 Analys Source C	-0.144 Summar S R-sq D1 98.24% Sis of Vari	y R-sq(adj 97.36% ance Arli ss C2	R-sq(pro 95.9- Arti MC F C3	ed) 10-fr 4% 2.9	old S 10-foli 0772 S		C7 x2	C8 x3	C9 x4	C10	C11 12 Heat	C12 x1_1	C13 x2_1	C14 x3_1	C15 x4_1	C16	C17 Time		C19 Temperature	
x4_1 Model 2.4460 Analys Source C	-0,144 Summar S R-sq D1 98.24% bis of Vari DF	y R-sq(adj 97.36% ance Arli ss C2	R-sq(pro 95.9- Arti MC F C3	ed) 10-fr 4% 2.9	old S 10-fold 0772 S P-Valme CS g	25.95% C6				C10						C16			Temperature	Yei
x4_1 Model 2.4460 Analys Source C	-0.144 Summan S R-sg D1 98.24% Sis of Vari DF 1 Trength Win	y R-sq(adj 97.36% ance Arli ss C2 e length D	R-sq(pro 95.94 Arti MS F C3 ie height	ed) 10-fr 4% 2.9	old S 10-foli 0772 S P-Valme CS s y	C6 x1	x2	<b>x</b> 3	x4	C10	Heat	x1_1	x2_1	x3_1	x4_1	C16	Time	Velocity	Temperature 7.5	Yei
x4_1 Model 2.4460 Analys Source C	-0,144 Summar S R-sg 21 98.24% Sis of Vari DF 1 rength Wir 9.95	y R-sq(ad) 97.36H ance Arli SS C2 e length D 2	R-sq(pro 95.94 Arli MS F C3 ie height 50	ed) 10-fr 4% 2.9	old S 10-fold 0772 S P-Maline CS 12 y 240	C6 x1 25	x2 24	<b>x</b> 3 91	x4 100	C10	Heat 78.5	x1_1	x2_1 26	<b>x3_1</b> 6	x4_1 60	C16	Time 1300	Velocity 0.0120	Temperature 7.5 9.0	Yei
x4_1 Model 2.4460 Analys Source C	-0,144 Summar S R-sg 1 98.24% sis of Vari nF 1 n 9.95 24.45	y R-sq(ad) 97.36H ance Adi SS C2 e length 2 8	R-sq(pre 95.9- Arti MC F C3 ie height 50 110	ed) 10-fr 4% 2.9	01d s 10-folo 0772 s P-Value C5 p y 240 236	C6 x1 25 31	x2 24 21	x3 91 90	x4 100 95	C10	Heat 78.5 74.3	x1_1 7 1	x2_1 26 29	x3_1 6 15	x4_1 60 52	C16	Time 1300 1300	Velocity 0.0120 0.0120	Temperature 7.5 9.0 11.0	Yel
x4_1 Model 2.4460 Analys Source C	-0,144 S R-sg D1 98,24% sis of Vari DF 1 1 1 24,45 31,75	y R-sq(adj) 97.36% ance Arli SS C2 C2 8 11	R-sq(pre 95.9- Ardi MC F C3 ie height 50 110 120	ed) 10-fr 4% 2.9	old s 10-fold 0772 s P-Value CS 10 y 240 236 270	C6 x1 25 31 45	x2 24 21 24	x3 91 90 88	x4 100 95 110	C10	Heat 78.5 74.3 104.3	x1_1 7 1 11	x2_1 26 29 56	x3_1 6 15 8	x4_1 60 52 20	C16	Time 1300 1300 1300	Velocity 0.0120 0.0120 0.0115	Temperature 7.5 9.0 11.0	
x4_1 Model 2.4460 Analys Source C	-0.144 Summar S R-sq D1 98.24% sis of Vari pF 1 9.95 24.45 31.75 35.00	y <u>R-sq(adj</u> 97.36# ance <u>Arli ss</u> C2 e length 2 8 11 10	R-sq(pre 95.94 Arli MS F C3 ie height 50 110 120 550	ed) 10-fr 4% 2.9	old <u>S</u> 10-foll 0772 <u>S</u> P-Value <u>CS</u> <u>p</u> 240 236 270 274	C6 x1 25 31 45 60	x2 24 21 24 25	x3 91 90 88 87	x4 100 95 110 88	C10	Heat 78.5 74.3 104.3 87.6	x1_1 7 1 11 11	x2_1 26 29 56 31	x3_1 6 15 8 8	x4_1 60 52 20 47	C16	Time 1300 1300 1300 1300	Velocity 0.0120 0.0120 0.0115 0.0130	Temperature 7.5 9.0 11.0 13.5	Yei

I want to see what model says? So, over here what you see is that X1, X2, P values, if you observe P values none of the P values is significant, but R square adjusted is highly I am getting a very good model fit over here, 97 over here, and the tenfold cross-validation is also very good over here. So, something wrong is happening because none of the variable is significant. But I am able to predict very high predictability over here what

we are seeing like that. So, this equations means we can say that let us adopt this one, but we will not do that because there is another issue coming over here which is known as variation inflation factor that is, a variation inflation factor over here.

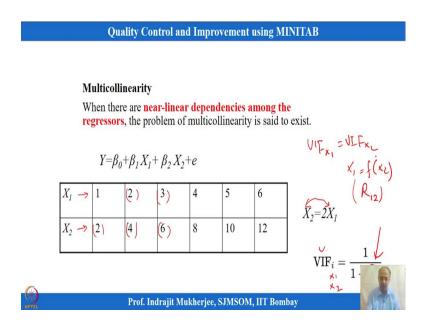
So, variation inflation factor indicates that whether there is a situation of multicollinearity in the data set. So, what is multicollinearity? We will try to explain. Multicollinearity in a sense it says that whenever the x are interrelated with each other let us say X1 with X2 or X2 with X3, and the correlation is very high that will influence the model and the model will not be correct and it will give you a bias judgment and the sign that you will get coefficient sign that you will get may interchange.

That means what it should be positive it is reflecting negative like that. So, that can happen over here. So, multicollinearity means there is a high significant relationship between the x variables over here. And this will be reflected by a index that is known as variation inflation factor, that is known as variation inflation factor.

(Refer Slide Time: 09:34)

Best Subset							
Regression	Heat	X1_	1 X2_	1 X3_1	1 X4_1	Response is Heat	
Following table	78.5	7	26	6	60		x
presents data	74.3	1	29	15	52		1234
concerning the	104.3	11	56	8	20		
concerning the heat evolved in	87.6	11	31	8	47	Vars R-Sq R-Sq (adj) R-Sq (pred) Ma	allows Cp S1111
heat evolved in	95.9	7	52	6	33	1 67.5 64.5 56.0	138.7 8.9639 X
calories per	109.2	11	55	9	22	1 66.6 63.6 55.7	142.5 9.0771 X
gram of cement	102.7	3	71	17	6	2 97.9 97.4 96.5	2.7 2.4063 X X
as a function of	72.5	1	31	22	44	2 97.2 96.7 95.5	5.5 2.7343 X X
four different	93.1	2	54	18	22	3 98.2 97.6 96.9	3.0 2.3087 X X X
chemical	115.9	21	47	4	26	3 98.2 97.6 96.7	3.0 2.3121 X X X
ingredients.	83.8	1	40	23	34	4 98.2 97.4 95.9	5.0 2.4460 X X X X
	113.3	11	66	9	12		
	109.4	10	68	8	12		
•		F	Prof. I	ndraii	t Mukh	Source: Montgomery, D. C., Pec Introduction to linear regression	

#### (Refer Slide Time: 09:36)



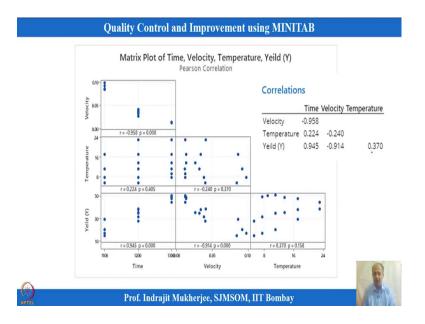
And this is what I mean to say over multicollinearity over here. So, you see the numbers of X1 is given over here for a particular and X2 is also simultaneously recorded. So, X1 is 1, X2 is 2 and when this is 2 this is 4 like this, this is 3 this is 6 like this. So, X2 has a functional relationship with X1 over here.

So that means, there is a high amount of correlation between the data set that I am having in X1 and the data set that I am having in X2. So, I can calculate the variation inflation factor between this data set. So, variation inflation factor for X1, I can calculate the similarly for X2, I can calculate like that only two variables over here.

So, then I calculate the Ri index that is the coefficient of determination over here, so Ri index can be calculated, where X1 let us say is regressed with a function of X2 like this. And then the R values are indicated, so that is with one and two like that. So, this values will be indicated and that will be reported over here we can put that value and I can get the variation inflation factor.

So, variation inflation factor for the first variable X1 will be same as variation inflation factor for because there is only two variables over here, say this one and this one like that; so, X1 and X2 if I am considering two variables over here.

# (Refer Slide Time: 10:51)



So, this variation inflation factor what would you see will lead to models; we will not get the best models out of this when multicollinearity exists, and then the beta coefficient estimation goes wrong, and in this case and the prediction will also go wrong.

So, if I consider that, if I ignore this multicollinearity what can happen is that my prediction model will show something different and actual scenario may be something else like that. So, in this case, I need to rectify this multicollinearity, there are different ways of rectifying the multicollinearity problem.

So, one of the one of the approach that takes care of this may be this what we are using as what we are using there are two methods over here. So, one of the method that is stepwise regression we have adopted like that and that may eliminate this multicollinearity problem that we are having. So, an another method is known as best subset regression based on which we can select variables which will go in and which will go out like that.

So, first is best substitute that we can talk about is stepwise regression like that. So, what we will do is that we will go to regression. So, when we have fitted this model that this variation inflation factor that you see over here. What; if I copy this one as image and we can paste it in excel, and let us try to see enlarge the image and try to see what is the results over here.

#### (Refer Slide Time: 12:31)

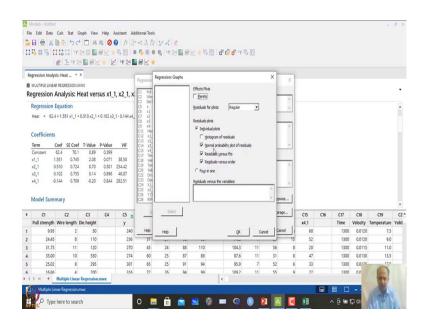
<b>3</b> & a	rmat Painter		u -   8 -	• A*	Data Review	b · ∰r Wrap ' I II II Merge	Text Gen & Center - 😭	eral - % →   % .		Format as Co Table - Styl	l Inse	t Delete Fo	tint ormat v	Clear *	A Z Sort & Fi Filter - Se iting	nd &	Share
Picture 1	• 1 3	< V	f <sub>x</sub>														
A	8	с	D	ε	F G	н	I J	К	L M	N	0	р	Q	R	s	т	U
-							0										
					Coeffic	ionte					-						
					Coeffic	ients											
					-	C		. Maluar	N to be								
-				-	Term	Coef	SE Coef T	-value I	-value	VIF							
					Constant	62.4	70.1	0.89	0.399								
				0							0						
					×1_1	1.551	0.745	2.08	0.071	38.50	-						
					x2_1	0.510	0.724	0.70	0.501	254.42							
					-		0 755			46.07							
					x3_1	0.102	0./55	0.14	0.896	46.87	_						
					x4_1	-0.144	0.709	-0.20	0.844	282.51							
					-												
				0-		1 1	0				-0						
-																	
															-	-	
																-	
( )	Sheet1	۲								•					-	N	
(da)_ \$7													Ħ	(D) E	] -	100	

So, variation inflation factor for each of these variables is indicative over here in the VIF, what we are seeing and we can we can just go to a simple and we can paste this one. And what we see over here is basically 38.5, X1 is having a variation inflation factor of 38.5, this is 254, this is 46, there is high amount of correlation that exists between the x variables over here. Which one is highly correlated which one we can see by the correlation matrix plot, and we will be able to know which is related with which one.

So, whenever this relationship strong relationship exists this variation inflation factor will be more than 5, will be more than 5 or 10 like that. So, generally statistician follows some rule, or thumb rules like that. If it is more than 5 we will take action and we want to eliminate multicollinearity problem in the regression equation.

So, that my prediction model becomes more accurate like that. So, anything more than 5 is we may consider a significant, we can take actions over there by using different methods and addressing the regression equations like that, developing the best regression equation like that.

So, here problem is that it is more than 10 or criteria 5 whatever you can select like that. And generally recommended is 10, but some statistician also suggest anything more than 5 is also a concern for me, so we should yeah we should try to remove that multicollinearity problem and then develop the regression equation like that, ok. So, this is a problem variation inflation factor. And in this case the options that we are having is that we will go for a stepwise equation. So, what we will do regression fit regression model over here. So, this is X1 to X4. So, in this case I will use stepwise regression and let us try to see how it works in this case.



(Refer Slide Time: 14:15)

(Refer Slide Time: 14:21)

						8K×														
gression An	alysis: He	rat v	×			Regression:	alidation						-	x						
MULTIPLE LIP				sus x1	1, x2_1, )			Yaldation	method:	K-fold o	ross-validat	ion		3						
Regressio Heat = 0 Coefficier	52.4 + 1.5		0.510 x2_1	+ 0.102 x	L1 - 0.144 x4	-		lyumbe gase f	nly assign rows ( r of folds (K): or random numb	er generator:		10								
Term	Coef	SE Coef	T-Value	P-Value	VIF			C Assign	rows of each fol	id by ID colum	`			-						
Constant	62.4	70.1		A 300	_															
	-807	/0,1	0.89	0.399							,									
x1_1	1.551	0.745	2.08	0.071	38.50						,									
x2_1	1.551 0.510	0.745	2.08	0.071	254.42						,			L						
	1.551	0.745	2.08 0.70 0.14	0.071	254.42 46.87															
x2_1 x3_1	1.551 0.510 0.102 -0.144	0.745 0.724 0.755 0.709	2.08 0.70 0.14	0.071 0.501 0.896	254.42 46.87			∏ §tore	ID column for K-	fold cross-val	dation									
x2_1 x3_1 x4_1 Model Su	1.551 0.510 0.102 -0.144	0.745 0.724 0.755 0.709	2.08 0.70 0.14 -0.20	0.071 0.501 0.896	254.42 46.87		3	∏ §tore	ID column for K-	-fold cross-val	dation			CI		C16	C17	C18	C19	
x2_1 x3_1 x4_1 Model Su C1 Pull streng	1.551 0.510 0.102 -0.144	0.745 0.724 0.755 0.709	2.08 0.70 0.14 -0.20 C3 Die height	0.071 0.501 0.896 0.844	254.42 46.87 282.51 C5 g	Sele	2	☐ Store	ID column for K-	-fold cross-val	dation			Ct: x4	1		Time	Velocity	Temperature	Ye
x2_1 x3_1 x4_1 Model Su C1 Pull streng	1.551 0.510 0.102 -0.144	0.745 0.724 0.755 0.709	2.08 0.70 0.14 -0.20 C3 Die height 50	0.071 0.501 0.896 0.844 C4	254.42 46.87 282.51 C5 g y 240		3	∏ §tore	ID column for K	-fold cross-vali	dation	QK.	Cancel					Velocity 0.0120	Temperature 7.5	Y
x2_1 x3_1 x4_1 Model Su C1 Pull streng 9, 24.	1.551 0.510 0.102 -0.144 mmary ath Wire 95 45	0.745 0.724 0.755 0.709 C2 2 length 2 8	2.08 0.70 0.14 -0.20 C3 Die height 50 110	0.071 0.501 0.896 0.844 C4	254.42 46.87 282.51 C5 g y 240 240 236	Help			20		. 1.0	13	**		60 52		Time 1300 1300	Velocity 0.0120 0.0120	Temperature 7.5 9.0	Y
x2_1 x3_1 x4_1 Model Su C1 Pull streng 9, 24, 31.	1.551 0.510 0.102 -0.144 mmary 95 45 75	0.745 0.724 0.755 0.709 C2 e length 2 8 11	2.08 0.70 0.14 -0.20 C3 Die height 50 110 120	0.071 0.501 0.896 0.844 C4	254.42 46.87 282.51 CS 5 y 240 236 270	Help 45	24	88	110		104.3	11	56		60 52 20		Time 1300 1300 1300	Velocity 0.0120 0.0120 0.0115	Temperature 7.5 9.0 11.0	• •
x2_1 x3_1 x4_1 Model Su C1 Pull streng 9, 24, 31, 35,	1.551 0.510 0.102 -0.144 mmary th Wire 95 45 75 00	0.745 0.724 0.755 0.709 C2 e length 2 8 11 10	2.08 0.70 0.14 -0.20 C3 Die height 50 110 120 550	0.071 0.501 0.896 0.844 C4	254.42 46.87 282.51 282.51 240 236 270 274	Help 45 60	24	88 87	110 88		104.3 87.6	11	56 31		60 52 20 47		Time 1300 1300 1300 1300	Velocity 0.0120 0.0120 0.0115 0.0130	Temperature 7.5 9.0 11.0 13.5	
x2_1 x3_1 x4_1 Model Su C1 Pull streng 9, 24, 31.	1.551 0.510 0.102 -0.144 mmary with Wire 95 45 75 00 02	0.745 0.724 0.755 0.709 C2 e length 2 8 11	2.08 0.70 0.14 -0.20 C3 Die height 50 110 120	0.071 0.501 0.896 0.844 C4	254.42 46.87 282.51 CS 5 y 240 236 270	Help 45	24	88	110		104.3	11	56		60 52 20		Time 1300 1300 1300	Velocity 0.0120 0.0120 0.0115	Temperature 7.5 9.0 11.0	

And then I will go for let me see storage of this residual is already there and I will click and graph. What we can do is that? We can see residual process with normal probability plot after doing stepwise regression, validation tenfold cross-validation we are doing over here.

# (Refer Slide Time: 14:24)

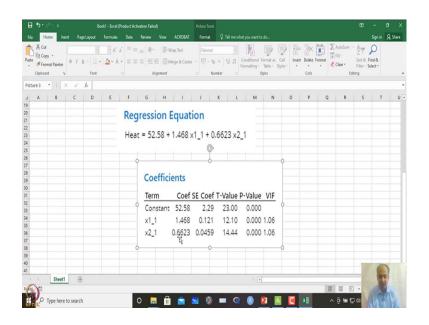
El.	nitab - Untitled																		-	Ø
ne.	Edit Data	Calc Stat	Graph View	Help A	ssistant Ada	fitional Tool	5													
1	3 @ XI	00 50	* 🗆 #	40	A I A	-148	24	2												
1		I I I V	ר	SK +	時回!#	Po 111 1	·尚 Y	×∎ø	K * 4	E d'	🐻 🗗 LY	時間								
		ELY																		
Per	ression Analysi		_																	
	ULTIPLE UNEAR																			
-	gression A			eue vi i	v2 1 v2	1														10
re	gression A	Analysis: I	reat ver	sus xi_i	, x2_1, x3	_1, X4_1														
1	Aethod																			
	ross-validation	10-fold																		
	1033-10100000	1 10-1010																		
			-					Q												
-	tepwise Sel	ection of 1	erms																	
¢	to enter = 0.1	15, a to remo	ve = 0.15																	
,	in a second s	-																		
	legression E																			
	legression E leat = 52.5		+ 0.6623 x2	_1																
1	leat = 52.5		+ 0.6623 x2	.1																
1	-		+ 0.6623 x2	.1																
-	leat = 52.5 Coefficients C1	8 + 1.468 ×1_1	G	_1 C4	C5 g	C6	67	C8	<b>C</b> 9	C10	C11 g	C12	C13	C14	C15	C16	C17	C18	C19	(
-	ieet = 52.5 Coefficients C1 Pull strength	8 + 1.468 ×1_1 C2 Wire length	C3 Die height		у	x1	x2	<b>x</b> 3	x4	C10	Heat	C12 x1_1	x2_1	x3_1	x4_1	C16	Time	Velocity	Temperature	Yei
-	coefficients Coefficients C1 Pull strength 9.95	C2 Wire length 2	C3 Die height 50		у 240	x1 25	x2 24	x3 91	<b>x4</b> 100	C10	Heat 78.5		x2_1 26	x3_1 6	x4_1 60	C16	Time 1300	Velocity 0.0120	Temperature 7.5	Yei
	coefficients Coefficients C1 Pull strength 9.95 24.45	C2 Wire length 2 8	C3 Die height 50 110		y 240 236	x1 25 31	<b>x2</b> 24 21	x3 91 90	x4 100 95	C10	Heat 78.5 74.3	x1_1 7 1	x2_1 26 29	x3_1 6 15	<b>x4_1</b> 60 52	C16	Time 1300 1300	Velocity 0.0120 0.0120	Temperature 7.5 9.0	Yei
	Coefficients Coefficients C1 Pull strength 9.95 24.45 31.75	C2 Wire length 2 8 11	C3 Die height 50 110 120		y 240 236 270	x1 25 31 45	x2 24 21 24	x3 91 90 88	x4 100 95 110	C10	Heat 78.5 74.3 104.3	xt_1 7 1 11	x2_1 26 29 56	x3_1 6 15 8	x4_1 60 52 20	C16	Time 1300 1300 1300	Velocity 0.0120 0.0120 0.0115	Temperature 7.5 9.0 11.0	Yei
-	Coefficients Coefficients C1 Pull strength 9.95 24.45 31.75 35.00	C2 Wire length 2 8 11 10	C3 Die height 50 110 120 550		y 240 236 270 274	x1 25 31 45 60	x2 24 21 24 25	x3 91 90 88 87	x4 100 95 110 88	C10	Heat 78.5 74.3 104.3 87.6	x1_1 7 1 11 11	x2_1 26 29 56 31	x3_1 6 15 8 8	x4_1 60 52 20 47	C16	Time 1300 1300 1300 1300	Velocity 0.0120 0.0120 0.0115 0.0130	Temperature 7.5 9.0 11.0 13.5	Yei
	coefficients C1 Pull strength 9.95 24.45 31.75 35.00 25.02	C2 Wire length 2 8 11 10 8	C3 Die height 50 110 120 550 295		y 240 236 270 274 301	x1 25 31 45 60 65	x2 24 21 24 25 25	x3 91 90 88 87 91	x4 100 95 110 88 94	C10	Heat 78.5 74.3 104.3 87.6 95.9	x1_1 7 11 11 11 7	x2_1 26 29 56 31 52	x3_1 6 15 8 8 8 6	x4_1 60 52 20 47 33	C16	Time 1300 1300 1300 1300 1300	Velocity 0.0120 0.0120 0.0115 0.0130 0.0135	Temperature 7.5 9.0 11.0 13.5	Yei
	coefficients C1 Pull strength 9.95 24.45 31.75 35.00 25.02 16.86	C2 Wire length 2 8 11 10 8 4	C3 Die height 50 110 120 550 295 200	C4	y 240 236 270 274	x1 25 31 45 60	x2 24 21 24 25	x3 91 90 88 87	x4 100 95 110 88		Heat 78.5 74.3 104.3 87.6	x1_1 7 1 11 11	x2_1 26 29 56 31	x3_1 6 15 8 8	x4_1 60 52 20 47	C16	Time 1300 1300 1300 1300	Velocity 0.0120 0.0120 0.0115 0.0130	Temperature 7.5 9.0 11.0 13.5	Yei
	C1 Pull strength 9.95 24.45 31.75 35.00 25.02 16.86 > H +	C2 Wire length 2 8 11 10 8 4 Multiple Lines	C3 Die height 50 110 120 550 295 700 r Regression	C4	y 240 236 270 274 301	x1 25 31 45 60 65	x2 24 21 24 25 25	x3 91 90 88 87 91	x4 100 95 110 88 94	C10	Heat 78.5 74.3 104.3 87.6 95.9	x1_1 7 11 11 11 7	x2_1 26 29 56 31 52	x3_1 6 15 8 8 8 6	x4_1 60 52 20 47 33		Time 1300 1300 1300 1300 1300 1300	Velocity 0.0120 0.0120 0.0115 0.0130 0.0135 0.0	Temperature 7.5 9.0 11.0 13.5	Yei
-	C1 Pull strength 9.95 24.45 31.75 35.00 25.02 16.86 > H +	C2 Wire length 2 8 11 10 8 4	C3 Die height 50 110 120 550 295 700 r Regression	C4	y 240 236 270 274 301	x1 25 31 45 60 65	x2 24 21 24 25 25	x3 91 90 88 87 91	x4 100 95 110 88 94		Heat 78.5 74.3 104.3 87.6 95.9	x1_1 7 11 11 11 7	x2_1 26 29 56 31 52	x3_1 6 15 8 8 8 6	x4_1 60 52 20 47 33		Time 1300 1300 1300 1300 1300 1300	Velocity 0.0120 0.0120 0.0115 0.0130 0.0135	Temperature 7.5 9.0 11.0 13.5	
	Coefficients Coefficients 2445 31.75 35.00 25.02 5.02 5.02 5.02 5.02 5.02 5.02	C2 Wire length 2 8 11 10 8 4 Multiple Lines	C3 Die height 50 110 120 550 295 200 r Regression mwx	C4	y 240 236 270 274 301 316	x1 25 31 45 60 65	x2 24 21 24 25 25 25 26	x3 91 90 88 87 91	x4 100 95 110 88 94 94		Heat 78.5 74.3 104.3 87.6 95.9 109.2	x1_1 7 11 11 11 7	x2_1 26 29 56 31 52	x3_1 6 15 8 8 8 6	x4_1 60 52 20 47 33		Time 1300 1300 1300 1300 1300 1300	Velocity 0.0120 0.0120 0.0115 0.0130 0.0135 0.0135	Temperature 7.5 9.0 11.0 13.5	Ye

# (Refer Slide Time: 14:27)

	itab - Untitled																			0
le	Edit Data Calc	Stat Gra	ph View	Help A	ssistant Add	itional Tools	5													
E	-	60		0	i fx 3º	-r. J. R.	1 st d													
									L . M	NO 1		M 50								
								CIP	ET # 24	ST : 0. 6	0 0. 1	-y 89								
		£чγ⊠	BR	× 2	- Y ≥ ■	BK 1														
legr	ession Analysis: He	at v x																		
M	ULTIPLE LINEAR REGI	RESSION.MW	x																	
lea	ression Anal	lysis: He	at vers	us x1 1	, x2 1, x3	1, x4 1														
_	epwise selection																			
a	o enter = 0.15, a	to remove	= 0.15																	
-																				
Re	egression Equa	tion		*																
Н	eat = 52.58 + 1.	468 m1_1 + 0	0.6623 x2_																	
_		0																		
Te	rm Coef		T-Value		VIF															
	rm Coef	SE Coef 2.29 0.121 0.0459	T-Value 23.00 12.10 14,44	0.000	VIF 1.06 1.06															
	rm Coef onstant 52.58 _1 1.468 _1 0.6623	2.29 0.121	23.00 12.10	0.000	1.06 1.06	C6	9	C8	<b>C9</b>	C10	C11 -	C12	C13	C14	C15	C16	C17	C18	C19	
Te Co x1 x2	rm Coef onstant 52.58 _1 1.468 _1 0.6623	2.29 0.121 0.0459 C2	23.00 12.10 14.44 C3	0.000 0.000 0.000	1.06 1.06	C6 x1	C7 x2	C8 x3	C9 x4	C10	C11 g Heat	C12 x1_1	C13 x2_1	C14 x3_1	C15 x4_1	C16	C17 Time		C19 Temperature	
Te Co x1 x2	rm Coef onstant 52,58 _1 1.468 _1 0.6623 C1	2.29 0.121 0.0459 C2	23.00 12.10 14.44 C3	0.000 0.000 0.000	1.06 1.06 C5 g					C10						C16				Ye
Te Co x1 x2	rm Coef onstant 52.58 _1 1.468 _1 0.6623 C1 ull strength Wire	2.29 0.121 0.0459 C2 length Die	23.00 12.10 14.44 C3 e height	0.000 0.000 0.000	1.06 1.06 CS g y	x1	x2	<b>x</b> 3	<b>x</b> 4	C10	Heat		x2_1	x3_1	x4_1	C16	Time	Velocity	Temperature	Ye
Te Co x1 x2	rm Coef onstant 52.58 _1 1.468 _1 0.6623 C1 ull strength Wire 9.95	2.29 0.121 0.0459 C2 length Dir 2	23.00 12.10 14.44 C3 c3 c3 c height 50	0.000 0.000 0.000	1.06 1.06 <b>C5</b> g y 240	x1 25	x2 24	x3 91	<b>x4</b> 100	C10	Heat 78.5		x2_1 26	x3_1 6	<b>x4_1</b> 60	C16	Time 1300	Velocity 0.0120	Temperature 7.5	Yei
Te Co x1 x2	rm Coef onstant 52,58 _1 1.468 _1 0.6623 C1 Vire 9,95 24,45	2.29 0.121 0.0459 C2 length Dia 2 8	23.00 12.10 14.44 C3 a height 50 110	0.000 0.000 0.000	1.06 1.06 <b>C5</b> g y 240 236	x1 25 31	<b>x2</b> 24 21	x3 91 90	x4 100 95	C10	Heat 78.5 74.3	x1_1 7 1	x2_1 26 29	x3_1 6 15	<b>x4_1</b> 60 52	C16	Time 1300 1300	Velocity 0.0120 0.0120	Temperature 7.5 9.0	
Te Co x1 x2	rm Coef onstant 52,58 _1 1.468 _1 0.6623 C1 Wire 9.95 24,45 31,75	2.29 0.121 0.0459 C2 8 8 11	23.00 12.10 14.44 <b>C3</b> a height 50 110 120	0.000 0.000 0.000	1.06 1.06 <b>C5</b> g y 240 236 270	x1 25 31 45	x2 24 21 24	x3 91 90 88	x4 100 95 110	C10	Heat 78.5 74.3 104.3	x1_1 7 1 11	x2_1 26 29 56	x3_1 6 15 8	x4_1 60 52 20	C16	Time 1300 1300 1300	Velocity 0.0120 0.0120 0.0115	Temperature 7.5 9.0 11.0	Yei
Te Cx x1 x2 P	rm Coef nstant 52.58 _1 1.468 _1 0.6623 C1 Wire 9.95 24.45 31.75 35.00 25.02 16.86	2.29 0.121 0.0459 cc length Dia 2 8 11 10 8 4	23.00 12.10 14.44 <b>C3</b> <b>a height</b> 50 110 120 550 295 700	0.000 0.000 0.000 C4	1.06 1.06 <b>C5 p</b> <b>y</b> 240 236 270 274	x1 25 31 45 60	x2 24 21 24 25	x3 91 90 88 87	x4 100 95 110 88	C10	Heat 78.5 74.3 104.3 87.6	x1_1 7 1 11 11	x2_1 26 29 56 31	x3_1 6 15 8 8	x4_1 60 52 20 47	C16	Time 1300 1300 1300 1300	Velocity 0.0120 0.0120 0.0115 0.0130	Temperature 7.5 9.0 11.0 13.5	Ye
Te Cc ×1 ×2	rm Coef nstant 52.58 _1 1.468 _1 0.6623 C1 Wire 9.95 24.45 31.75 35.00 25.02 16.86	2.29 0.121 0.0459 C2 8 8 11 10 8	23.00 12.10 14.44 <b>C3</b> <b>a height</b> 50 110 120 550 295 700	0.000 0.000 0.000 C4	1.06 1.06 <b>C5</b> 2 <b>y</b> 240 236 270 274 301	x1 25 31 45 60 65	x2 24 21 24 25 25 25	x3 91 90 88 87 91	x4 100 95 110 88 94	C10	Heat 78.5 74.3 104.3 87.6 95.9	x1_1 7 11 11 11 7	x2_1 26 29 56 31 52	x3_1 6 15 8 8 8 6	x4_1 60 52 20 47 33	C16	Time 1300 1300 1300 1300 1300	Velocity 0.0120 0.0120 0.0115 0.0130 0.0130	Temperature 7.5 9.0 11.0 13.5	Ye
Te Cc ×1 ×2	rm Coef nstant 52.58 _1 1.468 _1 0.6623 C1 Wire 9.95 24.45 31.75 35.00 25.02 16.86	2.29 0.121 0.0459 C2 8 9 11 10 8 4 11 10 8 4 4 10 10 10 10 10 10 10 10 10 10 10 10 10	23.00 12.10 14.44 <b>C3</b> <b>a height</b> 50 110 120 550 295 200 egression.	0.000 0.000 0.000 C4	1.06 1.06 <b>C5</b> 2 <b>y</b> 240 236 270 274 301	x1 25 31 45 60 65	x2 24 21 24 25 25 25	x3 91 90 88 87 91	x4 100 95 110 88 94		Heat 78.5 74.3 104.3 87.6 95.9	x1_1 7 11 11 11 7	x2_1 26 29 56 31 52	x3_1 6 15 8 8 8 6	x4_1 60 52 20 47 33		Time 1300 1300 1300 1300 1300 1300	Velocity 0.0120 0.0120 0.0115 0.0130 0.0130	Temperature 7.5 9.0 11.0 13.5	Yei
Te Co x1 x2	rm <u>Coef</u> nottant 52.58 _1 1.468 _1 0.6623 <b>C1 Wire</b> 9.95 2445 31.75 35.00 25.02 16.86 N + <u>Май</u>	2.29 0.121 0.0459 C2 8 11 10 8 4 11 10 8 4 11 10 8 4 12 10 10 8 4 11 10 10 8 4 11 10 10 10 10 10 10 10 10 10 10 10 10	23.00 12.10 14.44 c3 a height 50 110 120 550 295 200 295 200 295	0.000 0.000 0.000 C4	1.06 1.06 236 270 274 301 316	x1 25 31 45 60 65	x2 24 21 24 25 25 25 26	x3 91 90 88 87 91 94	x4 100 95 110 88 94 94		Heat 78.5 74.3 104.3 87.6 95.9	x1_1 7 11 11 11 7	x2_1 26 29 56 31 52	x3_1 6 15 8 8 8 6	x4_1 60 52 20 47 33		Time 1300 1300 1300 1300 1300 1300	Velocity 0.0120 0.0120 0.0115 0.0130 0.0135 0.0135	Temperature 7.5 9.0 11.0 13.5	Yei

And I click ok. What we observe is that only X1 and X2 is retained over here. So, you see the equation after doing stepwise regression what is coming out on this model.

# (Refer Slide Time: 14:38)

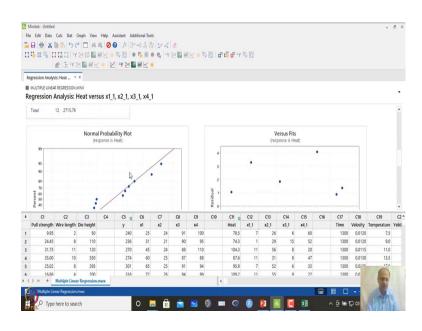


And we are getting a regression equation like this and heat is equal to 52.58 into 1.468. This is a coefficient plus coefficient what we have for X1 and X2. And the results also indicate that, now variation inflation factor if you look at this column what happened is that copy as picture and if we have taken this as the final equation, then the variation inflation vector is near to 1.

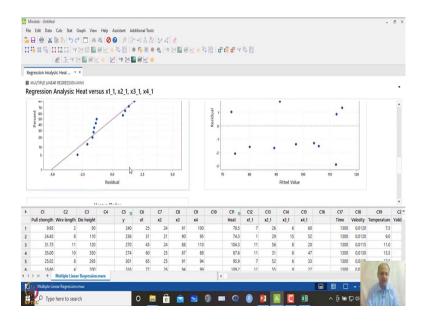
What you can see if we if I enhance this one that this is near to 1. When it is near to 1, that means, it is quite perfect and no multicollinearity problem does not exist now, and X1 and X2 are independent over here. We can assume independence in between the variables over here.

So, we have replaced this is X3 and X4 basically. We have just removed X3 and X4 over here. And the R square value and tenfold cross-validation more or less they are close to each other.

# (Refer Slide Time: 15:31)

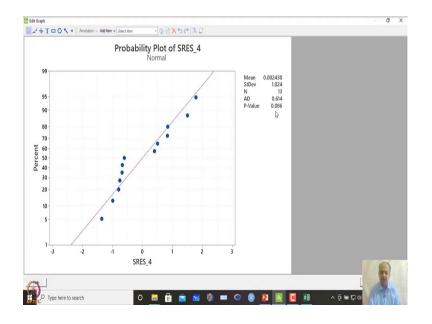


### (Refer Slide Time: 15:32)



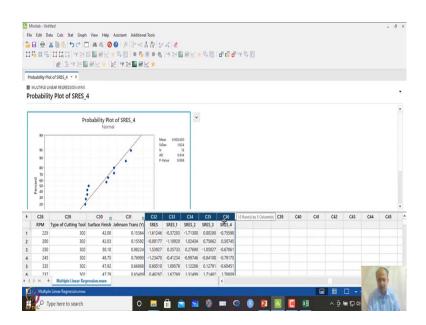
And ANOVA analysis also shows they are significant. And we can see that residual normal probability plot and we can also verify whether the final residual is normal or not, it is falling normal or not.

## (Refer Slide Time: 15:41)



So, what we observe over here is that, ok, this assumption is also validated. So, P is more than 0.05. So, this is there is no problem in the error or residuals like that. So, we can just remove for our benefit, so that later on we have only required information over here.

(Refer Slide Time: 15:54)

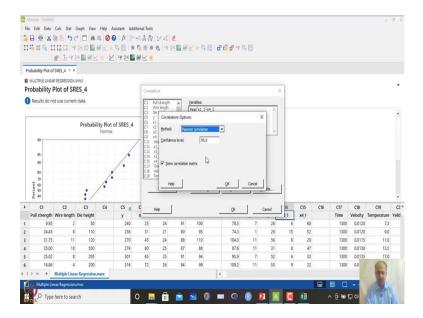


So, what we can do is that. So, this is verified over here. So, two variable goes in and two variable goes out over here. So, that means, this stepwise regression has taken care of this multicollinearity problem over here, somewhat we are fortunate that this is taken care of over here.

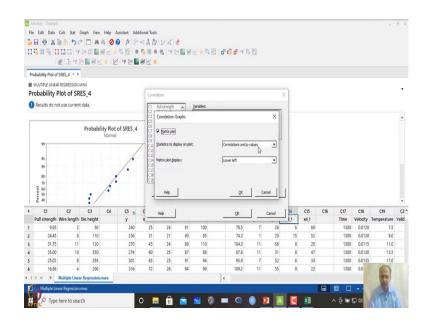
# (Refer Slide Time: 16:32)

20		500 13 44 2	nph View Help	• € 10 (#	二 古 計 称 雅 :	2   2 4   4 ● 株   Y ≥		1日本 4 日	d" d	<b>d</b> <sup>n</sup> - Y 23									
Probat	bility Plot of SRE	4 * X																	
-	ability Plot				Com	elation							×						
1 Re	isults do not us	e current d	ata.		C1 C2	Pull strength Wire length Die height	^	Variables:				^							
	99 95 90 80 70 60 50		Probability Plc		C1 C2 C3 C5 C6 C7 C8 C9 C11 C12 C13 C13 C14 C15 C17 C18 C19 C19 C19 C19 C19 C19 C19 C19 C10 C10 C10 C10 C10 C10 C10 C10 C10 C10	Y x1 x2 x3 x4 Heat x1_1 x1_1 x1_1 x1_1 x1_1 x1_1 x1_1 x1_	, ,	Optiong		graphs		v Besults							
		C2	G (4		-									C15	C16	C17	C18	C19	,
	I strength Win			CS g	1_	Help			_	<u>Ø</u> K		Cancel	31	x4_1	C10	Time		Temperature	
1	9.95	2	50	240	25	24	91	100		78.5	7	26	6	60		1300	0.0120	7.5	
2	24.45	8	110	236	31	21	90	95		74.3	1	29	15	52		1300	0.0120	9.0	
3	31.75	11	120	270	45	24	88	110		104.3	11	56	8	20		1300	0.0115	11.0	
4	35.00	10	550	274	60	25	87	88		87.6	11	31	8	47		1300	0.0130	13.5	
5	25.02	8	295	301	65	25	91	94		95.9	7	52	6	33		1300	0.0135	17.0	
6	16.86	4	200 Legression.mwx	316	72	26	94	99	4	109.2	11	55	9	22		1300	0.0		
2	Multiple Linear R		-			1		N ()			_		_	xI		▦ □	-		

# (Refer Slide Time: 16:37)



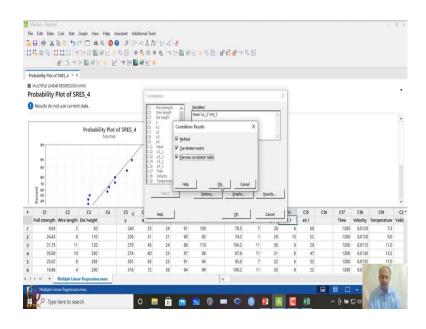
#### (Refer Slide Time: 16:43)



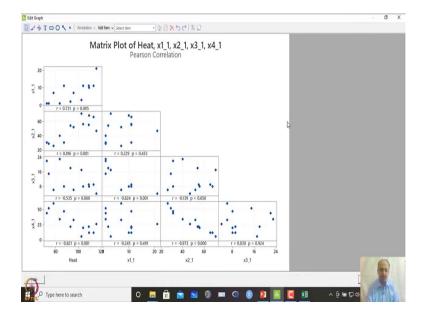
So, why; and let us try to see multicollinearity issues over here. So, what we can do is that basic stat over here and we go to correlation over here. So, correlations between what we can do is that we can see the correlation between heat and other variables over here from X1 to X4, and we can select this one I want to see the correlation matrix.

So, I go to options and I say Pearson correlation I want to say, I want to identify. And then in graphs what we can do is that I in the plot I want correlation with P values like that because I am interested in P values which is correlated with which one and which is significant like that. And we can see a pairwise correlation also.

#### (Refer Slide Time: 16:53)



(Refer Slide Time: 16:56)



And when I click ok what will happen you will get some graphs like this. So, this graph will indicate how what is the correlation between y variable and X1, X2 and inter relationship between X1 and X2, how is the correlation?

So, if I see the first column over here what you see heat is having showing a P value of 0.01 with X4 variables. So, heat is highly correlated with X4, heat is highly correlated with X3, not so much it is more than 0.05. So, we can say that this may not be significant

so much. But X2 i C p value is less than 0.05 X1, X1, X2 and X4 there are there is highly high amount of correlation that is existing.

Now, X1 if you see, X1 is highly if I see within the variables. So, X1 how it is related with X4, it does not have any correlation X4. But with X3 it is having a high level of correlation over here. So, X1 is related with X3; X1 and X3 are highly correlated over here. So, in this case, this is the observation.

Then, similarly X2 you see perfect relationship exist between X2 and X4. So, X2 and X4 are more or less perfect and the r coefficient is negative coefficient is 0.973. So, negative correlation exists between this variable X1, X2 and X4. So that means, these two variables are highly correlated. Similarly, X1 and X3 is highly correlated.

So, whenever high correlation exists and I want to do regression in that case what is required is that one of the variable has to go, one of the variable has to go out of this X1 and X3. We can think of an X2 and X4 basically has to go, one of the variables has to go. And the stepwise regression has correctly identified two variables instead of 4 and it has identified X1 and X2.

It has retained those two variables and remove the other two variables because there is a multicollinearity problem. So, some part of multicollinearity problem can be addressed by when I use stepwise regression. But that can always be verified by seeing the variation inflation factor and seeing the model adequacy and other checks. And finally, we select the models over here.

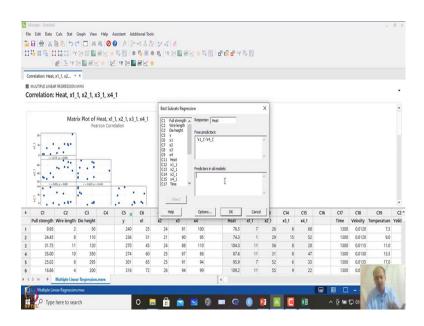
So, this correlation matrix helps you to understand that which variable is highly correlated to which one which variable is removed like that which variable is added by. So, this is one way of selecting the variables which is known as stepwise regression when there is a multicollinearity problem that is existing over here, ok.

So, this and there is another option which can be explored over here which is known as best subset regression which is another option. This stepwise regression what the limitation of this approach is that it will select the final variables X1 and X2 which is selected like that, but it will it has dropped X3 and X4.

But scenario can be that I want to explore what happens if I include X3 instead of X1; what happens if I include X4 in instead of X2 because those variables are easy to control maybe because I want a regression equation where variables can be easily controlled like that maybe X1 and X2 is too difficult, but regression equation by significance and best subset methodology we are getting X1 and X2.

But I want to see the complexity if I use different combination like that. If I use X3 X4 combination what is happening and like that; so what will happen. So, there is option which is known as best subset regression. So, if you go to regression, regression in MINITAB you will see an option of best subset over here.

(Refer Slide Time: 20:18)



So, if you click best subset regression for this then you identify which is the variable. So, I give heat as the response over here and then X1 to X4, these are the variables I want to include in the model which is free predicted over here. So, if you want some predicted to be always there in the model, so predicted in all models. So, in that case you can just write X1 to be there.

So, I am not identifying that one. So, I said every variable is free. So, show me all combinations and best combinations like that.

(Refer Slide Time: 20:48)

		Untitled																	-	8
File	e Edit	Data	Calc Stat	Graph View	Help Assis	tant Add	itional To	pols												
1	8	2 31	0 6 5	A 🗆 5	4 00	fx a	-2.4	12/4	12											
1	13 11	mit	I LI LI M		ek + 4	(四) #	<b>r</b> <sub>0</sub> #	<b>亲先</b>	YNER	K+A	10 d' co 1	PY英国								
					* 12															
				_			tor LL													
Be	est Subs	ets Regr	ression:	×																
₿	MULTIP	LE LINEA	R REGRESSION	MWX																
Be	est Su	bset	s Regress	ion: Heat	t versus x	1_1, x2	1, x3_	1, x4_1												
																				5
	Respo	nse is	Heat																	
							x x :	( X												
							1 2	3 4												
	Marca		0 Co (add)	D. Co. Inc. dt	Mallows Cp		di i													
	Vars 1	67.5	K-Sq (adj) 64.5	K-Sq (pred) 56.0		8.9639	K)	X												
	1	66.6	63.6	55.7		9.0771	х													
	2	97.9	97.4	96.5	2.7	2.4063	хх													
	2	97.2	96.7	95.5	2.7	2.4063 2.7343	x x x	х												
	2	97.2 98.2	95.7 97.6	95.5 96.9	2.7 5.5 3.0	2.4063 2.7343 2.3087	x x x x x	Х												
	2 3 3	97.2 98.2 98.2	96.7 97.6 97.6	95.5 96.9 96.7	2.7 5.5 3.0 3.0	2.4063 2.7343 2.3087 2.3121	x x x x x x x :	x												
	2 3 3	97.2 98.2	95.7 97.6	95.5 96.9	2.7 5.5 3.0 3.0	2.4063 2.7343 2.3087	x x x x x x x :	x												
	2 3 3 4	97.2 98.2 98.2	96.7 97.6 97.6	95.5 96.9 96.7	2.7 5.5 3.0 3.0	2.4063 2.7343 2.3087 2.3121 2.4460	x x x x x x x :	x	C8	C9	C10 C	11 m C12	C13	C14	C15	C16	C17	C18	C19	
	2 3 3 4	97.2 98.2 98.2 98.2	96.7 97.6 97.6 97.4	95.5 96.9 96.7 95.9	2.7 5.5 3.0 3.0 5.0	2.4063 2.7343 2.3087 2.3121 2.4460		x ( x	C8 x3	C9 x4		11 12 C12 at x1_1	C13 x2_1	C14 x3_1	C15 x4_1	C16	C17 Time		C19 Temperature	
	2 3 3 4	97.2 98.2 98.2 98.2	96.7 97.6 97.6 97.4 <b>C2</b>	95.5 96.9 96.7 95.9	2.7 5.5 3.0 3.0 5.0	2.4063 2.7343 2.3087 2.3121 2.4460	X X X X X X X X X X X X X	X ( X ( X ( X	x3							C16				
	2 3 3 4	97.2 98.2 98.2 98.2	96.7 97.6 97.4 97.4 <b>C2</b> Wire length	95.5 96.9 96.7 95.9 C3 Die height 50	2.7 5.5 3.0 3.0 5.0	2.4063 2.7343 2.3087 2.3121 2.4460 C5 g	X X X X X X 1 X X 1 X X 1 C6 x1	X (X (X (X (X) (X) (X) (X) (X) (X) (X) (	x3 91	x4		at x1_1	x2_1	x3_1	x4_1	C16	Time	Velocity	Temperature	Yei
2	2 3 3 4	97.2 98.2 98.2 98.2 98.2 98.2	96.7 97.6 97.6 97.4 <b>C2</b> Wire length 2	95.5 96.9 96.7 95.9 C3 Die height 50 110	2.7 5.5 3.0 3.0 5.0	2.4063 2.7343 2.3087 2.3121 2.4460 <b>C5</b> g <b>y</b> 240	X X X X X X X X X X X X X X X X X X X X	X C7 X2 5 2 5 2 5 2	x3 91 90 88	<b>x4</b> 100	H	at x1_1 78.5 7	x2_1 26 29	x3_1 6	<b>x4_1</b> 60	C16	Time 1300	Velocity 0.0120	Temperature 7.5	Yei
	2 3 3 4	97.2 98.2 98.2 98.2 98.2 1 rength 9.95 24.45	96.7 97.6 97.6 97.4 <b>C2</b> Wire length 2 8 11	95.5 96.9 96.7 95.9 <b>C3</b> <b>Die height</b> 50 1110 120	2.7 5.5 3.0 3.0 5.0	2.4063 2.7343 2.3087 2.3121 2.4460 <b>C5 p</b> <b>y</b> 240 236	X X X X X X X X X X X X X X X X X X X X	X C7 X2 5 2 5 2 5 2	x3 91 90 88	x4 100 95	H	at x1_1 78.5 7 74.3 1	x2_1 26 29	x3_1 6 15	x4_1 60 52	C16	Time 1300 1300	Velocity 0.0120 0.0120	Temperature 7.5 9.0 11.0	Yei
1	2 3 3 4	97.2 98.2 98.2 98.2 98.2 rength 9.95 24.45 31.75	96.7 97.6 97.6 97.4 <b>C2</b> Wire length 2 8 11	95.5 96.9 96.7 95.9 <b>C3</b> <b>Die height</b> 50 1110 120	2.7 5.5 3.0 3.0 5.0	2.4063 2.7343 2.3087 2.3121 2.4460 <b>C5 5</b> <b>y</b> 240 236 270	X X	X C7 X2 5 2 2 5 2 2 5 2 2 5 2 5 2 5 2 5 2 5 2 5 2 5 2 5 2 2 5 2 5 2 5 2 5 2 5 2 5 2 5 2 5 2 5 2 2 5 5 5 5 5 5 5 5 5 5 5 5 5	x3 91 90 8 88 5 87 5 91	x4 100 95 110	H	at x1_1 78.5 7 74.3 1 104.3 11	x2_1 26 29 56 31 52	x3_1 6 15 8	x4_1 60 52 20 47 33	C16	Time 1300 1300 1300	Velocity 0.0120 0.0120 0.0115	Temperature 7.5 9.0 11.0	Yei
1	2 3 3 4	97.2 98.2 98.2 98.2 98.2 1 rength 9.95 24.45 31.75 35.00	96.7 97.6 97.4 <b>C2</b> Wire length 2 8 111 10	95.5 96.9 96.7 95.9 C3 Die height 50 110 120 550	2.7 5.5 3.0 3.0 5.0	2.4063 2.7343 2.3087 2.3121 2.4460 <b>C5</b> g y 2.40 236 270 274	X X X X X X X X X X X X X 21 21 31 41 41 66	X C7 X2 5 2 2 5 2 2 5 2 2 5 2 5 2 5 2 5 2 5 2 5 2 5 2 5 2 2 5 2 5 2 5 2 5 2 5 2 5 2 5 2 5 2 5 2 2 5 5 5 5 5 5 5 5 5 5 5 5 5	x3 91 90 8 88 5 87 5 91	x4 100 95 110 88	H	at x1_1 78.5 7 74.3 1 104.3 11 87.6 11	x2_1 26 29 56 31 52	x3_1 6 15 8 8	x4_1 60 52 20 47	C16	Time 1300 1300 1300 1300	Velocity 0.0120 0.0120 0.0115 0.0130 0.0135	Temperature 7.5 9.0 11.0 13.5	Yei
	2 3 3 4	97.2 98.2 98.2 98.2 98.2 98.2 7 rength 9.95 24.45 31.75 35.00 25.02 16.86	96.7 97.6 97.6 97.4 Wire length 2 8 111 100 8 4	955 96.9 96.7 95.9 0ie height 50 110 120 550 295	2.7 5.5 3.0 5.0 C4	2.4063 2.7343 2.3087 2.3121 2.4460 <b>C5 p</b> <b>y</b> 240 236 270 274 301	X X	X C7 X2 5 2 2 5 2 2 5 2 2 5 2 5 2 5 2 5 2 5 2 5 2 5 2 5 2 2 5 2 5 2 5 2 5 2 5 2 5 2 5 2 5 2 5 2 2 5 5 5 5 5 5 5 5 5 5 5 5 5	x3 91 90 8 88 5 87 5 91	x4 100 95 110 88 94	H	at x1_1 78.5 7 74.3 1 104.3 11 87.6 11 95.9 7	x2_1 26 29 56 31 52	x3_1 6 15 8 8 6	x4_1 60 52 20 47 33	C16	Time 1300 1300 1300 1300 1300	Velocity 0.0120 0.0120 0.0115 0.0130 0.0135	Temperature 7.5 9.0 11.0 13.5	Yei
1 2 3 4 5 5	2 3 4 C Pull str	97.2 98.2 98.2 98.2 98.2 28.2 28.2 29.2 24.5 31.75 35.00 25.02 16.86 +	96.7 97.6 97.6 97.4 Wire length 2 8 11 10 8 4 4 Muttiple Line	95.5 96.9 96.7 95.9 Die height 50 110 120 550 295 200 ar Regression.	2.7 5.5 3.0 5.0 C4	2.4063 2.7343 2.3087 2.3121 2.4460 <b>C5 p</b> <b>y</b> 240 236 270 274 301	X X	X C7 X2 5 2 2 5 2 2 5 2 2 5 2 5 2 5 2 5 2 5 2 5 2 5 2 5 2 2 5 2 5 2 5 2 5 2 5 2 5 2 5 2 5 2 5 2 2 5 5 5 5 5 5 5 5 5 5 5 5 5	x3 91 90 8 88 5 87 5 91	x4 100 95 110 88 94	H	at x1_1 78.5 7 74.3 1 104.3 11 87.6 11 95.9 7	x2_1 26 29 56 31 52	x3_1 6 15 8 8 6	x4_1 60 52 20 47 33		Time 1300 1300 1300 1300 1300 1300	Velocity 0.0120 0.0110 0.0115 0.0130 0.0135 0.0	Temperature 7.5 9.0 11.0 13.5	Yei
1 2 3 4 5 6	2 3 4 C Pull str	97.2 98.2 98.2 98.2 98.2 98.2 98.2 98.2 16.2 31.75 31.75 35.00 25.02 16.86 +	96.7 97.6 97.6 97.4 Wire length 2 8 111 100 8 4	95.5 96.9 96.7 95.9 Die height 50 110 120 530 205 200 ar Regression.	2.7 5.5 3.0 5.0 C4	2.4063 2.7343 2.3087 2.3121 2.4460 <b>C5 s</b> <b>y</b> 240 236 270 274 301 316	X X X X X X X X X X X X X X X X X X X	X C7 X2 5 2 2 5 2 2 5 2 2 5 2 5 2 5 2 5 2 5 2 5 2 5 2 5 2 2 5 2 5 2 5 2 5 2 5 2 5 2 5 2 5 2 5 2 2 5 5 5 5 5 5 5 5 5 5 5 5 5	x3 4 91 90 4 88 5 87 5 91 5 94	x4 100 95 110 88 94 99	H	at x1_1 78.5 7 74.3 1 104.3 11 87.6 11 95.9 7 109.2 11	x2_1 26 29 56 31 52	x3_1 6 15 8 8 6	x4_1 60 52 20 47 33		Time 1300 1300 1300 1300 1300	Velocity 0.0120 0.0115 0.0130 0.0135 0.0	Temperature 7.5 9.0 11.0 13.5	Yei

So, I click ok over here, then you will get a information over here which I am copying and I will just paste this information over here. So, that it is easier to see also.

(Refer Slide Time: 21:03)

B					ook1 - Excel (P	oduct Act	ivation Failed)													
	Hon	ne Inse	nt Pag	pe Layout	Formulas		Review Vie	w ACROBAT	Format										Sign in 🖇	Q. Sha
-	X Cut				- A* A	==	= 8.	Wran Text	General		160		1 6	2 1		Σ	AutoSum -	Aw	ρ	
aste	Copy Copy	*								1.0.00	onditional	Termal.			rt Delete I	00-	50 ·	Sort & Fi		
*	💉 Forma	at Painter	BI	N - 1	2 · A ·		ㅋ 1 1	Merge & Center	- 19- %	5 % 43 F	ormatting	Table	- Sty	ei inse les • •	rt Delette i	• 🧶	Clear *	Filter - Se		
	Clipboard	- 6		Font			Alignme	nt	G NL	unber 5		Styles			Cells		Edi	ting		
Dicto	red v			6																
					1 81 1									1.22	1 2					
4	A	8	С	D	E	F	G	н	)	KL	М	3	N	0	P	Q	R	S	T	
ł					Ĭ			0					Ĭ							
					Resi	oons	e is Heat													
5											ХХ	хх								
5											12	34								
1																				
3							D. C ( Ib	0.0-1					-							
0								R-Sq (pred		ws Cp	511									
1					1	67.5	64.5	56.	0	138.7 8.96	39	Х								
2					Ĭ 1	66.6	63.6	55.	7	142.5 9.07	71 X		Ĭ.							
3						97.9	97.4													
									-	2.7 2.40			-							
					2	97.2	96.7	95.	5	5.5 2.73	43 X	Х	-							
,					3	98.2	97.6	96.	9	3.0 2.30	87 X X	Х	1							
3					2	98.2	97.6	96.	7	3.0 2.31	21 V V	v	-							
)														¢						
0					4	98.2	97.4	95.	9	5.0 2.44	60 X X	ХΧ		-						
1								0					_							
1 2 3 3 4 5 6 7 7 8 8 9 9 0 0 1 2 3 4 4 5 6 6 7 8 8 9 9 0 0 1 1 2 3 3 4 4 5 5 6 7 7 8 8 9 9 0 0 1 1 1 2 3 3 4 4 5 5 5 6 6 7 7 9 9 0 0 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1					-								~					1	0	
-		Sheet1	(+								1	41	_					-	1	
C	201	Pinesti	. 0													Ħ	(D) E		1	
10	100					_									_	888	000 CL		and the second	100

So, in this case, what you observe is that there is an indicator which is known as Mallow Cp over here which is known as Mallow Cp over here. So, this Mallow Cp is generally used to select the best model and combination. So, with one variable combination; so, over here what you see is that there are X1, X2, X3, X4 variables over here.

And there are with one variable that was selected the best model with one variable that is given over here. So, first two is giving you the best model. So, if I include X4 that is giving me a high value of R square adjusted over here and high value of R square also over here and this is the best model with one variable that is selected and that is the summary of that is given over here.

Similarly, with second best model that is with R square value 66.6 and I am getting a variable that should be that should be retained only X2. So, if you model with X4 what is happening; when model with X2 what is happening; best two models are given.

So, with single variable we have 4, so 4 models can be developed like that, 4 models can be developed like that. So, if you have k number of variables two to the power k is the total combinations of the permutation combination we can think of, all possible models, all possible models that can happen like that.

So, MINITAB only reports that best model with 1 variable, best model with 2 variable, best model with 3 variable. And finally, with all variables like that. So, we will ignore the last one because we want to reduce the number of variables and we will go by the lower combination which is giving you better options because I want to reduce the number of variables. Because I want to control less number of variables like that, and also I want to take care of multicollinearity that is an issue, ok.

So, in this case what we do is that there are 3 values that we see R square adjusted to make a compromise over here, R square adjusted, R square predicted and Mallow Cp. First we go by Mallow Cp in the indicator over here which indicates the variance is less. And in this case the error that is that is the sum of square error that we are committing over here that is that we are getting over here is much less as compared. So, how do you how do we compare that one?

There is an indicator that is known as Mallow Cp which says that the thumb rule over here is that the Mallow Cp value should be less than the number of variables considered for modeling plus 1. So, that is if you are going for let us say for first one over here, so this is number of variable considered is 1, X4 is considered over here in the model.

So, X4 plus 1 that is that is one variable that is 1 plus 1 is 2, and the Mallow Cp should be less than the 2 less than the value 2 over here and that is not the case over here 138.7

is the Mallow Cp indicator that we are getting over here. So, this criteria until and unless this criteria is fulfilled. So, this is not the best model that we should select.

Secondly, second one also you see that 142.5 which is very very higher than 2, that also can be eliminated over here. But in the third case what you see is that two variable model when X1 and X2 is considered Mallow Cp value is approximately 2.7. So, in these case this calculation of Mallow Cp any books will give you what is the calculation values of Mallow Cp considering, sum of square errors over here.

So, the formulation is given. So, I am not mentioning the formulation over here. That you can see in any standard textbook. So, what I am recommending is that this value you see two number of variables, so total number of 2.7 is less than 3.

So, this can be a possible combination. And also stepwise regression has also shown that this is the combination X1 and X2 is the best combination and that is the best model that stepwise regression has identified like that. Here also the suggestion is Mallow Cp is less than and very close to the number of variables plus 1. So, this value should be very close to number of variables plus 1. So, 2.7 is close to 3, so that means, this is one of the competitor models over here, ok.

And this is 5.5 it is more than, so this also goes, and this 3 is also less than 4 over here, so this can be one of the possibilities and this is another one possibilities over here. Now, you have to check that whether this model is with two variables is good with 3 variables what is happening with other 3 variables what is happening like that.

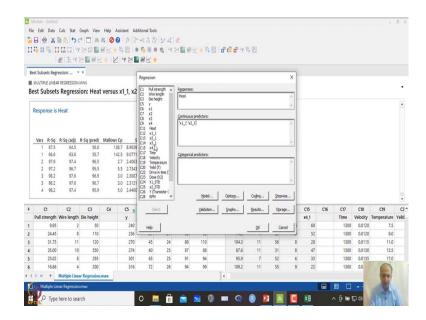
Because of correlation that exists between these two you will find that variation nucleation factor will be high, whenever I consider X1 and X3 together or X2 and X4 together like that there will be problems like that.

So, over here the closest model; that means based on Mallow Cp criteria also we are seeing that X1 and X2 is the best one. Mallow Cp is 2.7 R square predicted is 96 that is quite good enough, R square adjusted is 97.4. So, and this seems to be the closest one and we should select this one.

So, based on Mallow Cp criteria and based on our stepwise regression we are converging to the same models which can be suggested over here, which is the best model over here. So, X1, X2 variables regressed with y that is the best model over here. So, if you have if we have considered any other combination of that. So maybe we will not get the best models or maybe the assumptions will be violated, assumptions of normality and other assumptions that is there that can be violated head to schedule so, all these things.

So, whenever I have selected the best models there is always a requirement for checking the model adequacies. So, all the error terms and the assumptions of the errors are to be verified like that. So, in this case, what we see; and we see that X1 and X2 is the best selection over here, ok. We can check what is happening what is happening if I select X1, X2, and X4 over here? But X2 is highly related with X4. So, if you have selected that one variation inflation factor that problem will come.

(Refer Slide Time: 26:51)



# (Refer Slide Time: 26:58)

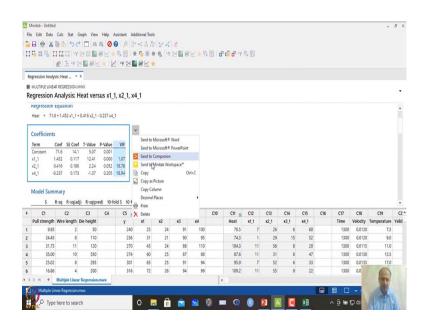
Best Sub	sets Rego	Calc Stat		Help Assis	fx   3	Bethod: None		<b>.</b>	×				. 8 ×
Resp	onse is	Heat											*
Vars	R-Sa	R-Sn (adi)	R-Sa (ored)	Mallows Cp		E = Include term in every model		I = Include term in the initial model					
1	67.5	64.5	56.0		8.963	Alpha to enter: 0.15							
1	66.6	63.6	55.7			Alpha to remove: 0.15							
2	97.9	97.4	96.5		2.406								
2	97.2	96.7	95.5		2.734								
3	98.2	97.6	96.9		2.308								
3	98.2 98.2	97.6 97.4	96.7 95.9		2.312 2.446								
	CI	C2	G	C4	C5 10				C15	C16 C	17 C18	C19	*
Pull s	trength	Wire length	Die height		у				x4_1	Tì	ne Velocit	Temperatu	re Yeld
1	9.95	2	50		240				60		1300 0.012	0 7	1.5
2	24,45	8	110		236	Herarchy			52		1300 0.012	0 9	0.0
3	31.75	11	120		270	Descriptor		₽.	20		1300 0.011	5 11	1.0
4	35.00	10	550		274	P Display the table of model selection details			47		1300 0.013	0 13	1.5
5	25.02	8	295		301	Details about the method	1		33		1300 0.013	5 17	0.1
6	16.86	4	200		316	locals about the means	-		22		1300 0.0	0	
4 D H		Multiple Line		umwx		Diplay the graph of R-squared vs step						-	16 Sa
		near Regression				Help		QK Cancel				AL	18
NPTEL	J Type	here to sea	rch			0 📃 🗄 🖻 🛂	0	💻 🥨 🕲 🔛	<u>I</u>	~ ĝ	<b>10</b> 💭 0	SAR	NA.

So, if I assume this one. So, in this case regression, if I assume X1, X2 and X4. So, in this case, if I remove this one X1, X2 and X4 and X2 and X4. We have seen highly correlated and I remove stepwise regression over here.

(Refer Slide Time: 27:01)

File																		-	e
26	Edit Data																		
	3 👲 🗶 🛛	10 50		14 0	0   fx   3*	子品品	12	2											
		In In Y	201	sk :	時回 #	心液 化	指 4	× 8	长*教	🗄 d' 🐻	ママシ語								
		C I Y		+ 14	- Y >	BK .													
Red	ression Analysi		_																
-	ULTIPLE LINEAP																		
Ke	gression A	Analysis: P	feat vers	us x1_	I, X2_1, X4	9													
N	Aethod																		
ç	cross-validation	10-fold																	
R	egression E	quation																	
	ieat = 71.6																		
- 5	ieat = /1.0	* 1/402 XI_1*	0.410 X2_1 -	0.237 X4	,														
C	oefficients									(	2								
7	ferm C	oef SE Coef	T-Value	P-Value	VIF														
		71.6 14.1	5.07	0.001															
		452 0.117		0.000	1.07														
-	-	416 0.186		0.052	18.78														
	C1 Pull strength	C2 Wire length	C3	C4	C5 👩	C6 x1	C7 x2	C8	C9	C10 1		C13	C14	C15	C16				
			Die height		N				vA		11 Z C12		v2.1	¥4.1		C17 Time	C18 Velocity	C19	
		-			y 240			x3	x4	н	nat x1_1	x2_1	x3_1 6	x4_1 60		Time	Velocity	Temperature	Ye
	9.95	2	50		240	25	24	91	100	н	nat x1_1 78.5	x2_1 7 26	6	60		Time 1300	Velocity 0.0120	Temperature 7.5	Ye
1	9.95 24.45	2	50 110		240 236		24 21	91 90	100 95	н	nat x1_1	x2_1 7 26 1 29	6 15	60 52		Time 1300 1300	Velocity 0.0120 0.0120	Temperature 7.5 9.0	Ye
	9.95	2	50		240	25 31	24	91	100	H	Nat x1_1 78.5 74.3	x2_1 7 26 1 29	6 15 8	60		Time 1300	Velocity 0.0120	Temperature 7.5 9.0 11.0	Ye
	9.95 24.45 31.75 35.00	2 8 11	50 110 120 550		240 236 270	25 31 45	24 21 24 25	91 90 88 87	100 95 110	H	nat x1_1 78.5 74.3 104.3	x2_1 7 26 1 29 11 56 11 31	6 15 8 8	60 52 20 47		Time 1300 1300 1300 1300 1300	Velocity 0.0120 0.0120 0.0115 0.0130	Temperature 7.5 9.0 11.0 13.5	Ye
	9.95 24.45 31.75 35.00 25.02	2 8 11	50 110 120 550 295		240 236 270 274 301	25 31 45 60 65	24 21 24 25 25	91 90 88 87 91	100 95 110 88 94	H	eat x1_1 78.5 74.3 104.3 87.6 95.9	x2_1 7 26 1 29 11 56 11 31 7 52	6 15 8 8 6	60 52 20 47 33		Time 1300 1300 1300 1300 1300	Velocity 0.0120 0.0120 0.0115 0.0130 0.0135	Temperature 7.5 9.0 11.0 13.5	Ye
	9.95 24.45 31.75 35.00 25.02 16.86	2 8 11 10 8 4	50 110 120 550 295 200		240 236 270 274	25 31 45 60	24 21 24 25	91 90 88 87	100 95 110 88		Rat x1_1 78.5 74.3 104.3 87.6	x2_1 7 26 1 29 11 56 11 31	6 15 8 8 8 6	60 52 20 47		Time 1300 1300 1300 1300 1300	Velocity 0.0120 0.0120 0.0115 0.0130 0.0135	Temperature 7.5 9.0 11.0 13.5	Ye
	9.95 24.45 31.75 35.00 25.02 16.86 > н +	2 8 11	50 110 120 550 295 200 Regression.	mwx	240 236 270 274 301	25 31 45 60 65	24 21 24 25 25	91 90 88 87 91	100 95 110 88 94	4	eat x1_1 78.5 74.3 104.3 87.6 95.9	x2_1 7 26 1 29 11 56 11 31 7 52	6 15 8 8 8 6	60 52 20 47 33		Time 1300 1300 1300 1300 1300	Velocity 0.0120 0.0115 0.0135 0.0135 0.0	Temperature 7.5 9.0 11.0 13.5	

#### (Refer Slide Time: 27:04)



And then, I do the calculation over here. What is being observed you see that the variation inflation factor that you observe over here is very high; variation inflation factor is very high over here.

(Refer Slide Time: 27:10)

	Home Inse	rt Page	Layout	Formulas Data Review Vi	ew ACROBAT	Format Q Tell me wh	it you want to do			Sign in 🔉 S
	opy  * ormat Painter		•   🖽 •	<ul> <li>▲、▲</li> <li>二 本</li> <li>▲</li> <li>▲</li> <li>本</li> <li></li> <li></li></ul>	🗒 Merge & Center 🕤	∰·%·}%#		Insert Delete Format		Find &
	* I >			an Augur		Number 1	Japas	con	Contray	
A	8	с	D	E F G	HII	JKL	MN	O P Q	R S	T
				Term	Coef	SE Coef	T-Value P	-Value	VIF	
			0	Constant	71.6	14.1	5.07	0.001		0
				x1_1	1.452	0.117	12.41	0.000	1.07	
				x2_1	0.416	0.186	2.24	0.052	18.78	
				x4_1	-0.237	0.173	-1.37	0.205	18.94	
						0				
										0
	Sheet1	۲					1	1		T

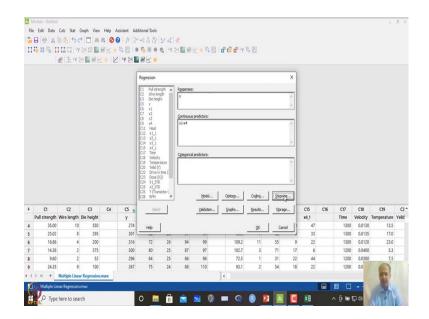
If I paste this one over here, and you will find that variation inflation factor is 18.78, X2 and X4, there is high correlation that exists between X2 and X4 that was prominent in correlation coefficient also. So, this whenever it is high this type of regression equation cannot be used that is the overall suggestion that we have, ok.

So, and in this case what happens is that y. So, let us take this example to finish off with this y equals 2. And we will continue with another examples and another situation in multi multiple regression when the error assumption fails in that case what we can do like that. And then we go into the design of experiment part of that.

So, why I am explaining this because when we develop a design of experiments regression equation we will we should be concerned about the variable interrelations between the variables, and how to select the best models out of many variables like that, how to eliminate variables like that, ok.

If this is the scenario y and X1 and X4, so we can eliminate this one and we go by this regression analysis, regression over here, fit regression models.

(Refer Slide Time: 28:20)



#### (Refer Slide Time: 28:30)

	Ainitab - Untitled					Regression: Stepwise	X						8 ×
•	e Edit Data	10 50		4 0 C	fx   3 時間   4	Method: Viane    Kore  Stopwe  Stopwe							
•	CI	62	G	C4	C5 g			C15	C16	C17	C18	C19	Q.4
	Pull strength				У		- 1	x4_1		Time		Temperature	
		10	550		274			47		1300	0.0130	13.5	Yeild
4	35.00												Yeild
5	25.02	8	295		301	Hierarchy	-	33		1300	0.0135	17.0	Yeild
5	25.02 16.86	4	200		316			22		1300	0.0120	17.0 23.0	Yeild
5 6 7	25.02 16.86 14.38	4	200 375		316 300	F gaplay the table of model election details		22 6		1300 1200	0.0120	17.0 23.0 5.3	Yeild
5	25.02 16.86	4	200		316			22		1300	0.0120	17.0 23.0	Yeld

(Refer Slide Time: 28:36)

		Calc Stat C																		
		0650									-									
-		I In In Y						× <b>•</b> 8	长大型	21 O C	Co LY 3	马間								
		₫ ± ч		1	LY 🖂	8K ¥														
eg	ession Analys	ils: y vers *	×																	
м	ULTIPLE UNEAR	R REGRESSION N	WX																	
eg	ression A	Analysis: y	versus	x1, x2, x	3, x4															
1	- 204.05	573 x1 + 9.88 x																		
y	· 3.9 / 0	7/3 X1 + 9.00 X																		
c	pefficients																			
		oef SE Coef	Tubba	D.Volue	VIF															
		3.9 81.9	0.05	0.963	VIP															
C																				
		573 0.250	2.29	0.048 1	.77															
x1 x2	i 0.2 2 9 Iodel Sumr	n.88 (mp.72) mary	2.66	0.026 1	.77	1 R-sa														
x1 x2 M	0.1 2 9 10del Sumr 5 F 12.0724 80.7 C1	0.88 ∰B.72 mary R·sq R·sq(ad 04% 75.61 C2	2.66 () R·sq(p % 55 C3	0.026 1 rred) 10-fol 85% 15.3 C4	.77 d s 10-fol 323 ! CS g	7.08% C6	67	C8	C9		C11 52	C12	C13	C14	C15		17	C18	C19	
×1 ×2 M	1 0.1 2 9 10del Sumr 5 F 12.0724 80.1 C1 ull strength	8.88 ∰3.72 mary R-sq R-sq(ad 04% 75.61 C2 Wire length	2.66 () R·sq(p 6 55 C3 Die heigh	0.026 1 rred) 10-fol 85% 15.3 C4	.77 d s 10-fol 323 5 y	C6 x1	x2	x3	<b>x</b> 4		Heat	x1_1	x2_1	C14 x3_1	x4_1		ne	Velocity	Temperature	
x1 x2 M	1 0.3 2 9 10del Sumr 5 F 12.0724 80.7 C1 ull strength 35.00	0.88 ∰3.72 mary R-sq R-sq(ad 0.4% 75.61 C2 Wire length 10	2.66 () <u>R-sq(p</u> % 55 C3 Die heigh 550	0.026 1 rred) 10-fol 85% 15.3 C4	.77 d s 10-fol 323 y y 274	C6 x1 60	x2 25	x3 87	<b>x4</b> 88		Heat 87.6	x1_1 11	x2_1 31	<b>x3_1</b> 8	<b>x4_1</b> 47		ne 1300	Velocity 0.0130	Temperature 13.5	
×1 ×2 M	0.0 2 9 100del Summ 5 1 12.0724 80.0 C1 uill strength 35.00 25.02	2.88 ₩3.72 mary R-sq R-sq(ad 0.4% 75.61 C2 Wire length 10 8	2.66 () R·sq(p 6 55 C3 Die heigh 550 295	0.026 1 rred) 10-fol 85% 15.3	CS 10-fol 323 2 y 274 301	C6 x1 60 65	x2 25 25	x3 87 91	x4 88 94		Heat 87.6 95.9	xU1 11 7	x2_1 31 52	<b>x3_1</b> 8 6	<b>x4_1</b> 47 33		ne 1300 1300	Velocity 0.0130 0.0135	Temperature 13.5 17.0	
×1 ×2 M	1 0.1 2 9 10del Sumr 5 F 12.0724 80.7 C1 ull strength 35.00 25.02 16.86	1.88 (₩3.72 mary R·sq R·sq(ad 04% 75.61 C2 Wire length 10 8 4	2.66 () R·sq(p 6 55 C3 Die heigh 550 299 200	0.026 1 rred) 10-fol 85% 15.3. C4	CS 10-fol 323 274 301 316	7.08% C6 x1 60 65 72	x2 25 25 26	x3 87 91 94	x4 88 94 99		Heat 87.6 95.9 109.2	x1_1 11 7 11	x2_1 31 52 55	<b>x3_1</b> 8 6 9	x4_1 47 33 22		me 1300 1300 1300	Velocity 0.0130 0.0135 0.0120	Temperature 13.5 17.0 23.0	
×1 ×2 M	0.0 2 9 100del Summ 5 1 12.0724 80.0 C1 uill strength 35.00 25.02	2.88 ₩3.72 mary R-sq R-sq(ad 0.4% 75.61 C2 Wire length 10 8	2.66 () R·sq(p 6 55 C3 Die heigh 550 295	0.026 1 rred) 10-fol 85% 15.3 C4	CS 10-fol 323 2 y 274 301	C6 x1 60 65	x2 25 25	x3 87 91	x4 88 94		Heat 87.6 95.9	xU1 11 7	x2_1 31 52	<b>x3_1</b> 8 6	<b>x4_1</b> 47 33	Ti	ne 1300 1300	Velocity 0.0130 0.0135	Temperature 13.5 17.0	
×1 ×2 M	1 0.2 2 9 100del Sumr 5 F 12.0724 80.1 12.0724 80.1 01 12.0724 80.1 01 12.0724 80.1 12.0724 80.1 12.0744 80.10744 80.10744	x.88         Imary           R×sq         R-sq(ad           04%         75.61           C2         Wire length           10         8           4         2	2.66 ) R·sq(p 6 55 C3 Die height 550 290 200 375	0.026 1 rred) 10-fol 85% 15.3	.77 d s 10-fol 323 y 274 301 316 300	C6 x1 60 65 72 80	x2 25 25 26 25	x3 87 91 94 87	x4 88 94 99 97		Heat 87.6 95.9 109.2 102.7	x1_1 11 7 11 3	x2_1 31 52 55 71	x3_1 8 6 9 17	x4_1 47 33 22 6	TI	me 1300 1300 1300 1300	Velocity 0.0130 0.0135 0.0120 0.0400	Temperature 13.5 17.0 23.0 5.3	
×1 ×2 M	1         0.1           2         9           Idodel Summ           S           12.0724         80.0           C1           uill strength           35.00         25.02           16.86         14.38           9.60         24.35	X88         UTB.72           mary         R-sq           R-sq         R-sq(ad)           0446         75.61           C2           Wire length           10         8           4         2           2         9	2.66 ) R·sq(p 6 55 C3 Die height 550 290 200 375 55 100	0.026 1 rred) 10-foi 85% 15.3	.77 d s 10-fol 323 y 274 301 316 300 296	C6 x1 60 65 72 80 84	x2 25 25 26 25 25 25	x3 87 91 94 87 86	x4 88 94 99 97 96		Heat 87.6 95.9 109.2 102.7 72.5	x1_1 11 7 11 3 1	x2_1 31 52 55 71 31	x3_1 8 6 9 17 22	x4_1 47 33 22 6 44	TI	me 1300 1300 1300 1200 1200	Velocity 0.0130 0.0135 0.0120 0.0400 0.0380	Temperature 13.5 17.0 23.0 5.3	
x1 x2 M	с 12.99 100del Sumr 5 И 12.0724 80.1 025.02 16.86 14.38 9.60 24.35 н +	x.88         (№3, 72)           mary         R×sq (ad 04%)           0.4%         75.61           C2         Wire length           10         8           4         2           2         2	2.66 () R-sq(p 6 55 C3 Die height 55 200 375 55 100 Regressio	0.026 1 rred) 10-foi 85% 15.3	.77 d s 10-fol 323 y 274 301 316 300 296	C6 x1 60 65 72 80 84	x2 25 25 26 25 25 25	x3 87 91 94 87 86	x4 88 94 99 97 96		Heat 87.6 95.9 109.2 102.7 72.5	x1_1 11 7 11 3 1	x2_1 31 52 55 71 31	x3_1 8 6 9 17 22	x4_1 47 33 22 6 44	TI	me 1300 1300 1300 1200 1200	Velocity 0.0130 0.0135 0.0120 0.0400 0.0380 0.0	Temperature 13.5 17.0 23.0 5.3	

So, in this case, what happens is that I selected y variables over here and then I select X1 to X4 and select this sorry this has to be coming over here continuous predictor X1 to X4 and I select this one and stepwise regression what we have done is that we use stepwise regression over here and use this one.

So, suggested model is X1 and X2. These are the variables. And variation inflation factor is less, so this can be the best model. Only thing is that R square adjusted is around 74,

75.61 and this is 57. So, there is some gap that we are observing over here 75 and 57 over here. So, whether we can improve this tenfold cross-validation over here.

(Refer Slide Time: 29:01)

File		Calc Stat		AL O	0 方 13 马图 14	-2 古 品 - 1 古 田 - 1	12 di 1 14 14 2		≤★脊間市	<b>o<sup>n</sup> too o</b> n +Y	4 E								8.3
Re	gression Analys	ils: y vers	×																
-	MULTIPLE LINEA																		
Re	gression /	Analysis:	y versus	x1, x2,	x3, x4														
		3.9 81.9		0.963			Best Sub	sets Regressi	on			×							
		573 0.250 0.88 3.72		0.048 0.026	1.77		C1 Pul C2 Wir C3 Die	strength A	Response: y										
	Model Sumi						C5 y	neight	Free predictors:			-							
1							C6 x1 C7 x2		X1_1-X1_1			^							
2	\$ 12.0724 80.	R-sq R-sq(a				d R-sq 57.08%	C3 x3 C9 x4 C11 2					~							
							IC12 81		Predictors in all m										
	Analysis of N	/ariance					C13 x2 C14 x3	1	Predictors in all m	2005:		A							
	Source	DF Adj SS	Adj MS F	Value P	Value		C15 x4 C17 Tr	ne v											
	Regression x1	2 5261.2 1 766.2	2630.6 766.2	18.05	0.001 0.048		s	elect	I			.v.							¥
ŧ.	CI	C2	3	C4	C5 g	C6	Help	1	Options	OK	Cancel	3	C14	C15	C16	C17	C18	C19	(2
	Pull strength	Wire length	Die height		У	x1	x2	x3	x4	Heat	x1_1	x2_1	x3_1	x4_1		Time	Velocity	Temperature	Yeld
4	35.00	10	550		274	60	25	87	88	87.6	11	31	8	47		1300	0.0130	13.5	
5	25.02	8	295		301	65	25	91	94	95.9	7	52	6	33		1300	0.0135	17.0	
6	16.86	4	200		316	72	26	94	99	109.2	11	55	9	22		1300	0.0120	23.0	
7	14.38	2	375		300	80	25	87	97	102.7	3	71	17	6		1200	0.0400	5.3	
8	9.60	2	52		296	84	25	86	96	72.5	1	31	22	44		1200	0.0380	7.5	
9	24.35	9	100		267	75	24	88	110	93.1	2	54	18	22		1200	0.0		
4	⊳н +	Multiple Line	r Regression	mwx						4									
Ø.	Multiple Lin	ear Regression	mwx												<b></b>	⊞ □		100	
V	Q Type	here to sea	celle			0		-	M 🔞 I	0				×B		^ @ @	CT 44	AL CA	
NP	rel vipe	nete to sea	null.						- V			e 18		×Π.		. ê. m.	E. W.	116	

So, again what we can do is that we can see this regression by best subset values. And the best subset regression we can do with y and X1 to X4 and try to see what the model recommendation like that.

(Refer Slide Time: 29:10)

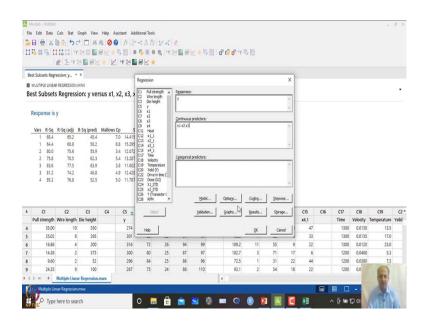
÷.		Intitled																		-	Ø
e	Edit	Data	Calc Stat	Graph View	Help Assist	ant Addit	ional Too	ls													
1		X	DA 5	4 0 5	1 00	fx Se.	24.2	240													
				Y 🖂 🖬 🔳						L + M	50		14.59								
'	0								CO	CT W M	83 · 0· 0	0.4	-9 BB								
			₽ ± ч	×∎@k		Y 🔁 🗖	SK1														
es	st Subset	ts Regr	ession: y	v x																	
		LINEA	R REGRESSION	A MARK																	
				sion: y ver	cue v1 v2	v2 v4															
e	st sui	DSet:	s Regres:	sion: y ver	Sus XI, XZ	, xɔ, x4															
P	Respon	ise is	У																		
	Mare	0.64	D. Car (add)	R-Sq (pred)	Mallaure Co.		x x 2 3														
-	vais 1	68.4	65.2	45.4		14.415	X	4													
	1	64.4	60.8	50.2		15.295 X															
	2	80.0	75.6	55.9	3.4	12.072 X	х														
		75.8	70.5	62.3		13.287 X															
		83.6	77.5	63.9		11.602 X															
		81.2 85.2	74.2	46.8		12.420 X															
		03.2																			
				52.5	5.0	11.787 X	XO	х													
				52.5	5.0	11.787 X	xO	X													
				52.5	5.0	11.787 X	хЮ	X													
	CI		C2	G		C5 g	C6	67	C8	C9	C10	C11 👦	C12	C13	C14	C15	C16	C17	C18	C19	
	Pull stre	ength	Wire length	C3 Die height		CS g y	C6 x1	C7 x2	х3	<b>x</b> 4		Heat	x1_1	x2_1	x3_1	x4_1	C16	Time	Velocity	Temperature	
	Pull stre	ength 35.00	Wire length	C3 Die height 550		C5 12 y 274	C6 x1 60	C7 x2 25	x3 87	x4 88		Heat 87.6	xU 11	<b>x2_1</b> 31	<b>x3_1</b> 8	<b>x4_1</b> 47	C16	Time 1300	Velocity 0.0130	Temperature 13.5	
	Pull stre	ength 35.00 25.02	Wire length	C3 Die height 0 550 3 295		C5 g y 274 301	C6 x1 60 65	C7 x2 25 25	x3 87 91	x4 88 94		Heat 87.6 95.9	xt_1 11 7	x2_1 31 52	<b>x3_1</b> 8 6	<b>x4_1</b> 47 33	C16	Time 1300 1300	Velocity 0.0130 0.0135	Temperature 13.5 17.0	
	Pull stre	ength 35.00 25.02 16.86	Wire length	C3 Die height 0 550 3 295 4 200		C5 12 y 274 301 316	C6 x1 60 65 72	C7 x2 25 25 26	x3 87 91 94	x4 88 94 99		Heat 87.6 95.9 109.2	xt_1 11 7 11	x2_1 31 52 55	x3_1 8 6 9	x4_1 47 33 22	C16	Time 1300 1300 1300	Velocity 0.0130 0.0135 0.0120	Temperature 13.5 17.0 23.0	
	Pull stre	ength 35.00 25.02 16.86 14.38	Wire length 11	C3 Die height 0 550 8 295 4 200 2 375		C5 g y 274 301 316 300	C6 x1 60 65 72 80	C7 x2 25 25 26 25	x3 87 91 94 87	x4 88 94 99 97		Heat 87.6 95.9 109.2 102.7	x1_1 11 7	x2_1 31 52 55 71	x3_1 8 6 9 17	x4_1 47 33 22 6	C16	Time 1300 1300 1300 1300 1200	Velocity 0.0130 0.0135 0.0120 0.0400	Temperature 13.5 17.0 23.0 5.3	
	Pull stre	ength 35.00 25.02 16.86 14.38 9.60	Wire length	C3 Die height 550 3 295 4 200 2 375 2 52		C5 g y 274 301 316 300 296	C6 x1 60 65 72 80 84	C7 x2 25 25 25 26 25 25 25	x3 87 91 94 87 86	x4 88 94 99 97 96		Heat 87.6 95.9 109.2 102.7 72.5	x1_1 11 7 11 3 1	x2_1 31 52 55 71 31	x3_1 8 6 9 17 22	x4_1 47 33 22 6 44	C16	Time 1300 1300 1300 1200 1200	Velocity 0.0130 0.0135 0.0120 0.0400 0.0380	Temperature 13.5 17.0 23.0	
	Pull stre	ength 35.00 25.02 16.86 14.38 9.60 24.35	Wire length	C3 Die height 550 2 200 2 375 2 52 9 100	C4	C5 g y 274 301 316 300	C6 x1 60 65 72 80	C7 x2 25 25 26 25	x3 87 91 94 87	x4 88 94 99 97		Heat 87.6 95.9 109.2 102.7	xt_1 11 7 11	x2_1 31 52 55 71	x3_1 8 6 9 17	x4_1 47 33 22 6	C16	Time 1300 1300 1300 1300 1200	Velocity 0.0130 0.0135 0.0120 0.0400	Temperature 13.5 17.0 23.0 5.3	
	Pull stre	ength 35.00 25.02 16.86 14.38 9.60 24.35	Wire length	C3 Die height 550 3 295 4 200 2 375 2 52	C4	C5 g y 274 301 316 300 296	C6 x1 60 65 72 80 84	C7 x2 25 25 25 26 25 25 25	x3 87 91 94 87 86	x4 88 94 99 97 96		Heat 87.6 95.9 109.2 102.7 72.5	x1_1 11 7 11 3 1	x2_1 31 52 55 71 31	x3_1 8 6 9 17 22	x4_1 47 33 22 6 44	C16	Time 1300 1300 1300 1200 1200	Velocity 0.0130 0.0135 0.0120 0.0400 0.0380	Temperature 13.5 17.0 23.0 5.3	
	Pull stre	ength 35.00 25.02 16.86 14.38 9.60 24.35 +	Wire length	C3 Die height 0 550 3 295 4 200 2 375 2 52 9 100 corr Regression.	C4	C5 g y 274 301 316 300 296	C6 x1 60 65 72 80 84	C7 x2 25 25 25 26 25 25 25	x3 87 91 94 87 86	x4 88 94 99 97 96		Heat 87.6 95.9 109.2 102.7 72.5	x1_1 11 7 11 3 1	x2_1 31 52 55 71 31	x3_1 8 6 9 17 22	x4_1 47 33 22 6 44		Time 1300 1300 1300 1200 1200 1200	Velocity 0.0130 0.0135 0.0120 0.0400 0.0380	Temperature 13.5 17.0 23.0 5.3	
	Pull stre	ength 35.00 25.02 16.86 14.38 9.60 24.35 + tiple Lin	Wire length	C3 Die height 5550 3 295 5 200 2 375 2 52 3 100 car Regression.	C4	CS g y 274 301 316 300 296 267	C6 x1 60 65 72 80 84 75	C7 x2 25 25 25 26 25 25 25	x3 87 91 94 87 86 88	x4 88 94 99 97 96		Heat 87.6 95.9 109.2 102.7 72.5	x1_1 11 7 11 3 1 2	x2_1 31 52 55 71 31	x3_1 8 6 9 17 22	x4_1 47 33 22 6 44		Time 1300 1300 1300 1200 1200 1200	Velocity 0.0130 0.0135 0.0120 0.0400 0.0380 0.0	Temperature 13.5 17.0 23.0 5.3	

So, in this case what you see is that X1 and X2 which is best subset is giving me a value of 3.4. So, this is about 3 over here this is about. So, this is more than 3 basically. So,

number of variables plus 1 and this is more than 3 over here. So, and this is 3.8 which is very close you see 1, 2, 3, 4.

So, X1, X2 and X3 variables if I consider and that is coming out to be very close. So, Mallow Cp, based on Mallow Cp index what we are seeing is that if I consider X1 and X2 and X3 variables over here that is giving me a Mallow Cp which is is approximately 3.8, which is very close to 4 and in that case 3 variables can be considered.

(Refer Slide Time: 29:52)



So, if I go over to 3 variables over here fit 3 variables; so, let us reduce this one, let us incorporate X1, X2 and X3. Stepwise says X1 and X2 only. So, we will remove stepwise over here and try to see what the model gives.

#### (Refer Slide Time: 30:03)

nie	Edit Data Ca	to the co	wh Man	-		No. 17 al														8 3
	8 8 × 6	6 50		4 0 ( 8 k *	)   fx   }= 時間   #	-2 古品 10 服用	12 4 1 # 14		长*节	n e d	d <sup>o</sup> uy 1	9, 四								
Re	gression Analysis:	y vers × >	2																	
8	MULTIPLE LINEAR R	EGRESSION M	NX																	
Re	gression Ar	nalysis: y	versus	x1, x2, >	:3															•
	Method																			*
	Cross-validation	10-fold																		
	0.000 101000001	10 1010																		
	Regression Eq	uation																		
	v = -162 + 0.74		1+121-2																	
	y = 1102 + 0.74	49 XI + 7709 X	2 + 6.04 3.0			G														
	Coofficiente					0														
	Coefficients					0														
	Term Coe	f SE Coef			VIF	0														
		2 148	T-Value -1.09 2.73	0.306	<u>VIF</u> 2.32	0														
	Term Coe Constant -16	2 148 9 0.275	-1.09	0.306		0														v
	Term         Coe           Constant         -16.           x1         0.741           x2         7.61           C1         C1	2 148 9 0.275 9 3.94 C2	-1.09 2.73 1.95 C3	0.306	2.32 2.16 C5 g	C6	67	C8	C9	C10	C11 😰	C12	C13	C14	C15	C16	C17	C18	C19	
	Term         Coe           Constant         -16:           x1         0.74!           x2         7.6!           C1         Pull strength	2 148 9 0.275 9 3.94 C2 fire length D	-1.09 2.73 1.95 C3 ie height	0.306 0.026 0.087	2.32 2.16 C5 g	C6 x1	x2	x3	<b>x</b> 4		Heat	x1_1	x2,1	x3_1	x4_1	C16	Time	Velocity	Temperature	Yeld
	Term         Coe           Constant         -16:           x1         0.74!           x2         7.6!           C1         Pull strength           W         35.00	2 148 9 0.275 9 3.94 C2 fire length D 10	-1.09 2.73 1.95 C3 ie height 550	0.306 0.026 0.087	2.32 2.16 C5 g y 274	C6 x1 60	x2 25	x3 87	<b>x4</b> 88		Heat 87.6	xU1 11	<b>x2_1</b> 31	<b>x3_1</b> 8	<b>x4_1</b> 47	C16	Time 1300	Velocity 0.0130	Temperature 13.5	Yeild
	Term         Coe           Constant         -16:           x1         0.74!           x2         7.6!           C1         Pull strength           W         35.00           25.02         25.02	2 148 9 0.275 9 3.94 C2 fire length D 10 8	-1.09 2.73 1.95 C3 ie height 550 295	0.306 0.026 0.087	2.32 2.16 <b>C5</b> g y 274 301	C6 x1 60 65	x2 25 25	x3 87 91	x4 88 94		Heat 87.6 95.9	x1_1 11 7	x2_1 31 52	<b>x3_1</b> 8 6	<b>x4_1</b> 47 33	C16	Time 1300 1300	Velocity 0.0130 0.0135	Temperature 13.5 17.0	Yeld
	Term         Coe           Constant         -16i           x1         0.74i           x2         7.6i           Pull strength         W           35.00         25.02           16.86	2 148 9 0.275 9 3.94 C2 10 8 4	-1.09 2.73 1.95 C3 ie height 550 295 200	0.306 0.026 0.087	2.32 2.16 C5 g y 274 301 316	C6 x1 60 65 72	x2 25 25 26	x3 87 91 94	x4 88 94 99		Heat 87.6 95.9 109.2	x[_1 11 7 11	x2_1 31 52 55	x3_1 8 6 9	x4_1 47 33 22	C16	Time 1300 1300 1300	Velocity 0.0130 0.0135 0.0120	Temperature 13.5 17.0 23.0	Yeild
	Term         Coe           Constant         -16:           x1         0.74!           x2         7.6!           C1         Pull strength           W         35.00           25.02         25.02	2 148 9 0.275 9 3.94 C2 fire length D 10 8	-1.09 2.73 1.95 C3 ie height 550 295	0.306 0.026 0.087	2.32 2.16 <b>C5</b> g y 274 301	C6 x1 60 65	x2 25 25	x3 87 91	x4 88 94		Heat 87.6 95.9	x1_1 11 7	x2_1 31 52	<b>x3_1</b> 8 6	<b>x4_1</b> 47 33	C16	Time 1300 1300	Velocity 0.0130 0.0135	Temperature 13.5 17.0	Yeild
	Term         Coe           Constant         -16i           x1         0.74i           x2         7.6i           Pull strength         W           35.00         25.02           16.86         14.38	2 148 9 0.275 9 3.94 C2 C2 10 10 8 4 2	-1.09 2.73 1.95 <b>C3</b> <b>ie height</b> 550 295 200 375	0.306 0.026 0.087	2.32 2.16 <b>C5</b> g y 274 301 316 300	C6 x1 60 65 72 80	x2 25 25 26 25	x3 87 91 94 87	x4 88 94 99 97		Heat 87.6 95.9 109.2 102.7	x[_1 11 7 11	x2_1 31 52 55 71	x3_1 8 6 9 17	x4_1 47 33 22 6	C16	Time 1300 1300 1300 1300 1200	Velocity 0.0130 0.0135 0.0120 0.0400 0.0380	Temperature 13.5 17.0 23.0 5.3	
	Term         Coe           Constant         -16:           x1         0.74!           x2         7.6!           Pull strength         W           35.00         25.02           16.86         14.38           9.60         24.35	2 148 9 0.275 9 3.94 C2 10 8 4 2 2 9	-1.09 2.73 1.95 <b>C3</b> <b>ie height</b> 550 295 200 375 52 100	0.306 0.026 0.087	2.32 2.16 <b>C5</b> g y 274 301 316 300 296	C6 x1 60 65 72 80 84	x2 25 25 26 25 25 25	x3 87 91 94 87 86	x4 88 94 99 97 96		Heat 87.6 95.9 109.2 102.7 72.5	xt_1 11 7 11 3 1	x2_1 31 52 55 71 31	x3_1 8 6 9 17 22	x4_1 47 33 22 6 44	C16	Time 1300 1300 1300 1200 1200	Velocity 0.0130 0.0135 0.0120 0.0400 0.0380	Temperature 13.5 17.0 23.0 5.3	
	Term         Coe           Constant         -16:           x1         0.74!           x2         7.6!           C1         W           35.00         25.02           16.86         14.38           9.60         24.35           ▷         H         ★	2 148 9 0.275 9 3.94 C2 10 10 8 4 2 2 2 9 wittiple Linear	-1.09 2.73 1.95 <b>C3</b> <b>ie height</b> 550 295 200 375 52 100 <b>Regression.</b>	0.306 0.026 0.087	2.32 2.16 <b>C5</b> g y 274 301 316 300 296	C6 x1 60 65 72 80 84	x2 25 25 26 25 25 25	x3 87 91 94 87 86	x4 88 94 99 97 96		Heat 87.6 95.9 109.2 102.7 72.5	xt_1 11 7 11 3 1	x2_1 31 52 55 71 31	x3_1 8 6 9 17 22	x4_1 47 33 22 6 44		Time 1300 1300 1300 1200 1200 1200	Velocity 0.0130 0.0135 0.0120 0.0400 0.0380 0.0	Temperature 13.5 17.0 23.0 5.3	Yei
1 5 7 3	Term         Coe           Constant         -16:           x1         0.74!           x2         7.6!           Pull strength         W           35.00         25.02           16.86         14.38           9.60         24.35	2 148 9 0.275 9 3.94 C2 10 10 8 4 2 2 2 9 wittiple Linear	-1.09 2.73 1.95 <b>C3</b> <b>ie height</b> 550 295 200 375 52 100 <b>Regression.</b>	0.306 0.026 0.087	2.32 2.16 <b>C5</b> g y 274 301 316 300 296	C6 x1 60 65 72 80 84	x2 25 25 26 25 25 25	x3 87 91 94 87 86	x4 88 94 99 97 96		Heat 87.6 95.9 109.2 102.7 72.5	xt_1 11 7 11 3 1	x2_1 31 52 55 71 31	x3_1 8 6 9 17 22	x4_1 47 33 22 6 44		Time 1300 1300 1300 1200 1200 1200	Velocity 0.0130 0.0135 0.0120 0.0400 0.0380	Temperature 13.5 17.0 23.0 5.3	Yell

(Refer Slide Time: 30:06)

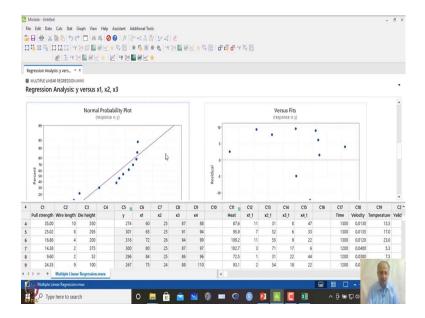
E						osistant Ad			1												
I.						時間 (#				K N	1	d d	英国								
						4Y >															
_			_				the lake														
	ression Anal																				
	ULTIPLE LINE																				
eg	gression	Analysi	5: y v	ersus >	(1, x2, )	(3															
Te	erm	Coef SE C	oef T	Value 1	P-Value	VIF															
C			148	-1.09	0.306																
x1			275	2.73		2.32															
X			.94	1.95	0.087	2.16															
	10del Sun S 11.6017 8	R-sq R-s	ą(adj) 7.47%	R·sq(pre 63.91		ild S 10-fol 8406 (	d R-sq 52.98%														
A	s	R-sq R-s 3.62% 7	7.47%	63.91	196 14.3	1406 (	d R-sq														
A	S 11.6017 8 nalysis of ource C1	R-sq R-s 3.62% 7 Variance DF Adj 5 C2	7.47% S Adj	63.91 MS F-\ C3	196 14.3	1406 (	d R-sq	C7	C8	C9	C10	C11 5	C12	C13	C14	C15	C16	C17	C18	C19	
A	S 11.6017 8 nalysis of ource C1 full strengtl	R-sq R-s 3.62% 7 Variance DF Adj 1 C2 Wire leng	7.47% S Adj	63.91 MS F-V C3 height	1% 14.3 Aalue P-	Value CS 12 Y	d R-sq 52.98% C6 x1	x2	x3	x4	C10	Heat	C12 x1_1	x2_1	C14 x3_1	x4_1	C16	Time	Velocity	Temperature	
A	s 11.6017 8 nalysis of ource C1 Yull strengtl 35.0	R-sq R-s 3.62% 7 Variance DF Adj ! C2 Wire leng	7.47% S Adj th Die 10	63.91 MS F-V C3 height 550	1% 14.3 Aalue P-	Value C5 12 y 274	C6 x1 60	x2 25	x3 87	x4 88	C10	Heat 87.6	xU 11	x2_1 31	x3_1 8	<b>x4_1</b> 47	C16	Time 1300	Velocity 0.0130	Temperature 13.5	5
A	S 11.6017 8 nalysis of ource C1 Vull strengti 35.0 25.0	R-sq R-s 3.62% 7 Variance DF Adj : C2 Wire leng	7.47% S Adj th Die	63.91 MS F-V C3 height 550 295	1% 14.3 Aalue P-	Value CS 10 y 274 301	C6 x1 60 65	x2 25 25	x3 87 91	x4 88 94	C10	Heat 87.6 95.9	x1_1 11 7	x2_1 31 52	<b>x3_1</b> 8 6	<b>x4_1</b> 47 33	C16	Time 1300 1300	Velocity 0.0130 0.0135	Temperature 13.5 17.0	5
A	5 11.6017 8 nalysis of ource C1 vull strengtl 35.0 25.0 16.8	R-sq R-s 3.62% 7 Variance DF Adj : C2 Wire leny 0 2 5	7.47% S Adj th Die 10	63.91 MS F-V C3 height 550 295 200	1% 14.3 Aalue P-	C5 10 Value C5 10 y 274 301 316	C6 x1 60 65 72	x2 25 25 26	x3 87 91 94	x4 88 94 99	C10	Heat 87.6 95.9 109.2	xt_1 11 7 11	x2_1 31 52 55	x3_1 8 6 9	x4_1 47 33 22	C16	Time 1300 1300 1300	Velocity 0.0130 0.0135 0.0120	Temperature 13.5 17.0 23.0	5
A	5 11.6017 8 nalysis of ource C1 1011 strengt1 35.00 25.00 16.80 14.30	R-sq         R-sq           3.62%         7           Variance         DF         Adj !           C2         1         Wire leng           0         2         5           8         3         3	7.47% S Adj th Die 10 8	63.91 MS F-V C3 height 550 295 200 375	1% 14.3 Aalue P-	Value CS y 274 301 316 300	C6 x1 60 65 72 80	x2 25 25 26 25	x3 87 91 94 87	x4 88 94 99 97	C10	Heat 87.6 95.9 109.2 102.7	x1_1 11 7 11 3	x2_1 31 52 55 71	x3_1 8 6 9 17	x4_1 47 33 22 6	C16	Time 1300 1300 1300 1300 1200	Velocity 0.0130 0.0135 0.0120 0.0400	Temperature 13.5 17.0 23.0 5.3	5
A	5 11.6017 8 nalysis of ource C1 Vull strengtl 35.00 25.00 16.80 14.33 9.60	R-sq         R-s           3.62%         7           Variance         DF         Adj !           C2         Wire leng           0         2           5         3           0         2	7.47% S Adj th Die 10 8	63.91 MS F-V C3 height 550 295 200 375 52	1% 14.3 Aalue P-	C5 10 Value C5 10 y 274 301 316	C6 x1 60 65 72 80 84	x2 25 25 26 25 25 25	x3 87 91 94 87 86	x4 88 94 99	C10	Heat 87.6 95.9 109.2 102.7 72.5	x1_1 11 7 11 3 1	x2_1 31 52 55 71 31	x3_1 8 6 9 17 22	x4_1 47 33 22 6 44	C16	Time 1300 1300 1300 1200 1200	Velocity 0.0130 0.0135 0.0120 0.0400 0.0380	Temperature 13.5 17.0 23.0 5.3	5
A	5 11.6017 8 nalysis of ource C1 1011 strengt1 35.00 25.00 16.80 14.30	R-sq         R-s           3.62%         7           Variance         DF         Adj !           C2         Wire leng           0         2           5         3           0         2	7.47% S Adj th Die 10 8	63.91 MS F-V C3 height 550 295 200 375	1% 14.3 Aalue P-	Value CS y 274 301 316 300	C6 x1 60 65 72 80	x2 25 25 26 25	x3 87 91 94 87	x4 88 94 99 97	C10	Heat 87.6 95.9 109.2 102.7	x1_1 11 7 11 3	x2_1 31 52 55 71	x3_1 8 6 9 17	x4_1 47 33 22 6	C16	Time 1300 1300 1300 1300 1200	Velocity 0.0130 0.0135 0.0120 0.0400	Temperature 13.5 17.0 23.0 5.3	5
A	5 11.6017 8 nalysis of ource C1 Vull strengtl 35.00 25.00 16.80 14.33 9.60	R-sq         R-s           3.62%         7           Variance         DF         Adj !           C2         Wire leng           V         5           3         5	7.47% S Adj 10 8 4 2 2 9	63.91 MS F-V C3 height 550 295 200 375 52 100	1% 14.2 Value P- C4	CS         p           Value         Value           Value         Value	C6 x1 60 65 72 80 84	x2 25 25 26 25 25 25	x3 87 91 94 87 86	x4 88 94 99 97 96	C10	Heat 87.6 95.9 109.2 102.7 72.5	x1_1 11 7 11 3 1	x2_1 31 52 55 71 31	x3_1 8 6 9 17 22	x4_1 47 33 22 6 44	C16	Time 1300 1300 1300 1200 1200	Velocity 0.0130 0.0135 0.0120 0.0400 0.0380	Temperature 13.5 17.0 23.0 5.3	5
A	S 11.6017 8 nalysis of ource C1 vull strength 35.00 25.00 16.80 14.30 9.64 24.30 24.30 24.30 24.30 24.30 24.30 24.30	R-sq         R-s           3.62%         7           Variance         DF           DF         Adj !           C2         Wire kenj           2         5           3         5	7.47% S Adj 10 8 4 2 2 9 9 mear Re	63.91 MS F-V C3 height 550 295 200 375 52 100 gression.	1% 14.2 Value P- C4	CS         p           Value         Value           Value         Value	C6 x1 60 65 72 80 84	x2 25 25 26 25 25 25	x3 87 91 94 87 86	x4 88 94 99 97 96		Heat 87.6 95.9 109.2 102.7 72.5	x1_1 11 7 11 3 1	x2_1 31 52 55 71 31	x3_1 8 6 9 17 22	x4_1 47 33 22 6 44		Time 1300 1300 1300 1200 1200 1200	Velocity 0.0130 0.0135 0.0120 0.0400 0.0380	Temperature 13.5 17.0 23.0 5.3	5

So, if you click ok over here what happens is that it gives a 3 model 3 variable models over here. So, over here what you see is that only X1 is coming prominent and others two are not coming prominent because the P value is more than 0.05 over here. Although the variation inflation factor is not significant over here, but R square predicted R square adjusted value somewhat improved and the tenfold cross-validation is also somewhat improved over here 62.98, ok.

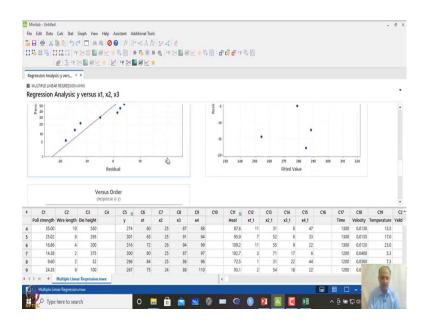
## (Refer Slide Time: 30:29)

Me	initab - Untitl	ed																		-	8
ile	Edit Data	Cal	lc Stat	Graph Vi	w Help	Assistant A	ditional Tool	5													
6	🗄 🧶 🐰		6 51	2	6.75	0 🕜 🛛 🍂 💡	-248	244	2												
11	10		113 4	NO	BK	***	H 76 H 1	# 兆 LY	×	K + N	10	🐻 🗗 LY	時間								
						V YM															
		-		_		-	the late														
	pression Ana																				
-	<b>ULTIPLE UN</b>																				1,
Reg	gression	An	alysis:	y versu	s x1, x	2, x3															
A	Analysis of	f Vari	iance																		
Se	Source			Adj MS		P-Value															
R	Regression		5496.1	1832.0	13.61	0.002															
			1001.1	1001.1	7.44	0.026															
,	x1 x2	1	512.2	512.2																	
,	x1 x2 x3		512.2 234.9	512.2 234.9	3.81	0.087															
2	×2	1						Þ													
) Er	x2 x3	1 1 8	234.9	234.9				R													
) Er	x2 x3 Error	1 1 8	234.9 1076.8	234.9				Q													
) Er	x2 x3 Error	1 1 8	234.9 1076.8	234.9 134.6	1.75	0.223		G						V	vreue Eite						
) Er	x2 x3 Error	1 1 8	234.9 1076.8	234.9 134.6	1.75	0.223 bability Plot		Q							ersus Fits						
a a Er Tr	x2 x3 Tror Tatel	1 1 8	234.9 1076.8 6572.9	234.9 134.6 Norr	1.75 nal Prol (respon	0.223 bability Plot ise is y)								(re:	iponse is y	)					
a a E E T	x2 x3 Error Fotal	1 1 8 11	234.9 1076.8 6572.9	234.9 134.6 Norr	1.75 nal Prol (respon	0.223 bability Plot use is y)		<b>C7</b>	CI	09	C10	C11 5	C12	(re: C13	c14	C15	C16	C17	C18	C19	
a Er Ti	x2 x3 Error Fotal C1 Pull strengt	1 1 8 11	234.9 1076.8 6572.9 C2 re length	234.9 134.6 Norr C3 Die heigt	1.75 nal Prol (respon	0.223 bability Plot isse is y)	x1	C7 x2	x3	<b>x</b> 4	C10	Heat	x1_1	(re: C13 x2_1	C14 x3_1	C15 x4_1	C16	Time	Velocity	Temperature	Yei
a Er Tr	x2 x3 Error Total  C1 Pull strengt 35.0	1 1 8 11	234.9 1076.8 6572.9 C2 re length 10	234,9 134,6 Norr C3 Die heigt 55	1.75 nal Prol (respon t 0	0.223 bability Plot ise is y) CS p y 274	x1 60	C7 x2 25	x3 87	<b>x4</b> 88	C10	Heat 87.6	xU 11	(re: C13 x2_1 31	C14 X3_1 8	C15 x4_1 47	C16	Time 1300	Velocity 0.0130	Temperature 13.5	Yei
Di D	x2 x3 fror fotal Pull strengt 35.0 25.0	1 1 8 11	234.9 1076.8 6572.9 C2 re length 10 8	234,9 134.6 Norr C3 Die heigh 55 29	1.75 nal Prol (respon t 0 5	0.223 bability Plot ise is y) CS is y 274 301	x1 60 65	C7 x2 25 25	x3 87 91	x4 88 94	C10	Heat 87.6 95.9	xt_1 11 7	(re: C13 x2_1 31 52	C14 x3_1 8 6	C15 x4_1 47 33	C16	Time 1300 1300	Velocity 0.0130 0.0135	Temperature 13.5 17.0	Yei
Diama di Antonio di An	x2 x3 Error Total  C1 Pull strengt 35.0	1 1 8 11 11	234.9 1076.8 6572.9 C2 re length 10	234,9 134.6 Norr C3 Die heigt 55	1.75 nal Prol (respon t 0 5 0	0.223 bability Plot ise is y) CS p y 274	x1 60	C7 x2 25	x3 87	<b>x4</b> 88	C10	Heat 87.6	xU 11	(re: C13 x2_1 31	C14 X3_1 8	C15 x4_1 47	C16	Time 1300	Velocity 0.0130	Temperature 13.5 17.0 23.0	Yei
i i i i	x2 x3 Error Fotal Pull strengt 35.0 25.0 16.8 14.3	1 1 8 11 11	234.9 1076.8 6572.9 C2 re length 10 8	234.9 134.6 Norr C3 Die heigh 55 29 20 37	1.75 nal Prol (respon t 0 5 5	0.223 bability Plot see is y)	x1 60 65 72	C7 x2 25 25 25 26 25	x3 87 91 94 87	x4 88 94 99 97	C10	Heat 87.6 95.9 109.2 102.7	xt_1 11 7 11	(res <b>C13</b> <b>x2_1</b> 31 52 55 71	C14 x3_1 8 6 9 17	C15 x4_1 47 33 22	C16	Time 1300 1300 1300 1300 1200	Velocity 0.0130 0.0135 0.0120 0.0400	Temperature 13.5 17.0 23.0 5.3	Yei
a Er Tr	x2 x3 fror fotal Pull strengt 35.0 25.0 16.8	1 1 8 11 1 1 1 1 1 1 1 1 1 1 1 1 1 1 8 11 1 1 8 11 1 1 8 11 1 1 8 11 1 1 8 11 1 1 8 11 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1	234.9 1076.8 6572.9 C2 re length 10 8	234.9 134.6 Norr C3 Die heigh 55 29 20	1.75 nal Prol (respon t 0 5 2	0.223 bability Plot CS g y 274 301 316	x1 60 65 72 80	C7 x2 25 25 26	x3 87 91 94	x4 88 94 99	C10	Heat 87.6 95.9 109.2	xt_1 11 7 11	(re: C13 x2_1 31 52 55	C14 x3_1 8 6 9	C15 x4_1 47 33 22 6	C16	Time 1300 1300 1300	Velocity 0.0130 0.0135 0.0120 0.0400 0.0380	Temperature 13.5 17.0 23.0	Yei

# (Refer Slide Time: 30:31)

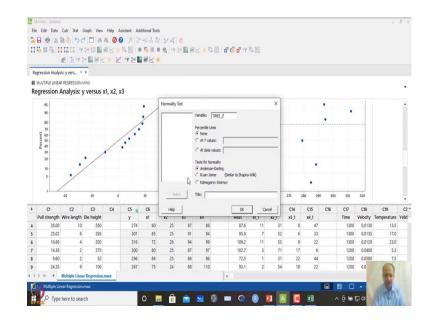


#### (Refer Slide Time: 30:31)

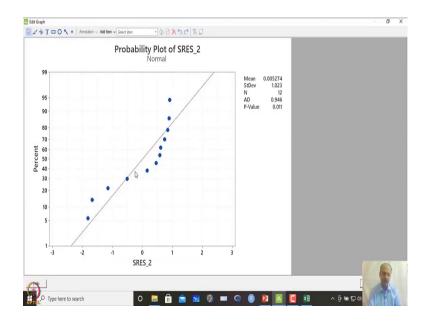


But the residuals that I have saved over here. So, this is the residual plot that you see. So, when I have used 3 3 variable models over here what happens is that; if I go to basic stat normality test what will happen is that. So if I go to the last variables and try to test this one what happens is that you see that there is a violation in the error distribution over here.

(Refer Slide Time: 30:42)

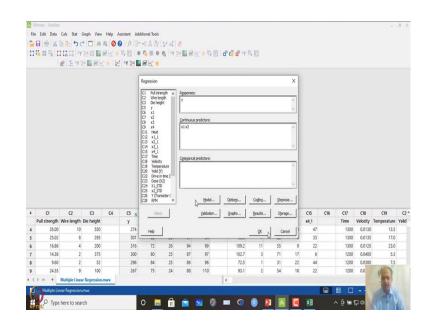


#### (Refer Slide Time: 30:45)



So, whenever I have added this one. So, if I have restricted this to 2 variable models like that. So I go to basic regression analysis regression fit regression model instead of X3, I go to X1 and X2 only which is suggested by stepwise regression.

(Refer Slide Time: 30:58)

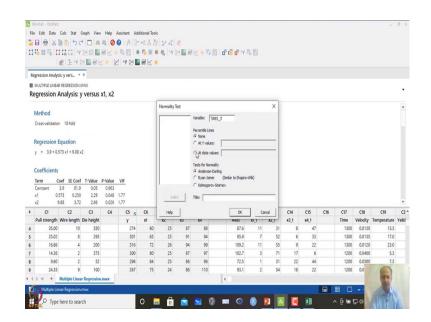


#### (Refer Slide Time: 31:02)

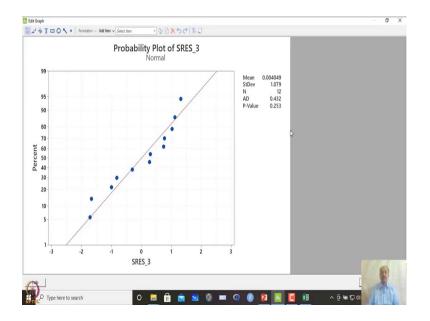
																			8
-	Edit Data	Calc Stat G	raph View	Help	Assistant Add	Stional Tools	5												
	8 8 X	06 50		11 0	0   fx ] ?=	-24.8	24	1											
1									K + M	1 d c d	Y 乾阳								
		@ EYP																	
0.		sis: y vers * 1	4			CP LL H													
	-																		
-		R REGRESSION M																	,
K	egression	Analysis: y	versus	x1, x2															
	the stand																		1
	Method																		
	Cross-validatio	n 10-fold																	
	Regression	Equation																	
	v = 3.9+0.	573 x1 + 9.88 x2																	
	,																		
	Coefficients																		
	Coefficients Term C	oef SE Coef	T-Value	P-Value	VIF														
	Coefficients Term C Constant	coef SE Coef 3.9 81.9	T-Value 0.05	0.963															
	Coefficients Term C Constant ×1 0.	oef SE Coef	T-Value	0.963	VIIF 1.77 1.77														
	Coefficients Term C Constant ×1 0.	Coef         SE Coef           3.9         81.9           573         0.250	T-Value 0.05 2.29	0.963	1.77 1.77	C6	67	CS	C9	C10 C11	72 C12	C13	C14	C15	C16	C17	C18	C19	
	Coefficients Term C Constant ×1 0, ×2 1 C1	Coef         SE Coef           3.9         81.9           573         0.250           9.88         3.72	T-Value 0.05 2.29 2.66 C3	0.963 0.048 0.026	1.77 1.77	C6 x1	C7 x2	C8 x3	C9 x4	C10 C11 Heat	Z C12 x1_1	C13 x2_1	C14 x3_1	C15 x4_1	C16	C17 Time		C19 Temperature	
	Coefficients Term C Constant ×1 0, ×2 1 C1	SE Coef         SE Coef           3.9         81.9           573         0.250           9.88         3.72           C2	T-Value 0.05 2.29 2.66 C3	0.963 0.048 0.026	1.77 1.77 C5 g						x1_1				C16				Yei
	Coefficients Term C Constant x1 0, x2 1 C1 Pull strength	SE Coef         SE Coef           3.9         81.9           573         0.250           9.88         3.72           C2         Wire length         E	T-Value 0.05 2.29 2.66 C3 Ne height	0.963 0.048 0.026	1.77 1.77 CS g y	x1	x2	x3	<b>x</b> 4	Heat	x1_1 6 11	x2_1	x3_1	x4_1	C16	Time	Velocity	Temperature	Yei
	Coefficients Term C Constant ×1 0, ×2 1 C1 Pull strength 35.00	SE Coef         SE Coef           3.9         81.9           573         0.250           9.88         3.72           C2         Wire length         D           10         10	T-Value 0.05 2.29 2.66 C3 Die height 550	0.963 0.048 0.026	1.77 1.77 C5 g y 274	x1 60	x2 25	x3 87	<b>x4</b> 88	Heat 8	x1_1 6 11 9 7	<b>x2_1</b> 31	<b>x3_1</b> 8	<b>x4_1</b> 47	C16	Time 1300	Velocity 0.0130	Temperature 13.5	Yei
	Coefficients Term C Constant ×1 0, ×2 1 C1 Pull strength 35,00 25,02	Coef         SE Coef           3.9         81.9           573         0.250           9.88         3.72           C2         Wire length         D           10         8	T-Value 0.05 2.29 2.66 C3 Die height 550 295	0.963 0.048 0.026	1.77 1.77 C5 g y 274 301	x1 60 65	x2 25 25	x3 87 91	x4 88 94	Heat 8 9	x1_1 6 11 9 7 2 11	x2_1 31 52	<b>x3_1</b> 8 6	<b>x4_1</b> 47 33	C16	Time 1300 1300	Velocity 0.0130 0.0135	Temperature 13.5 17.0	Yei
	Coefficients Term C Constant x1 0. x2 1 Pull strength 35.00 25.02 16.86	Coef         SE Coef           3.9         81.9           573         0.250           8.88         3.72           C2         Wire length         I           10         8           4         2	T-Value 0.05 2.29 2.66 C3 Die height 550 295 200	0.963 0.048 0.026	1.77 1.77 C5 g y 274 301 316	x1 60 65 72	x2 25 25 26	x3 87 91 94	x4 88 94 99	Heat 8 9 10	x1_1 6 11 9 7 2 11 7 3	x2_1 31 52 55 71 31	x3_1 8 6 9 17 22	x4_1 47 33 22 6 44	C16	Time 1300 1300 1300	Velocity 0.0130 0.0135 0.0120	Temperature 13.5 17.0 23.0	Yei
	Coefficients Term C Constant x1 0. x2 1 C1 Pull strength 35.00 25.02 16.86 14.38	Coef         SE Coef           3.9         81.9           573         0.250           8.88         3.72           C2         Wire length         I           10         8           4         2	T-Value 0.05 2.29 2.66 C3 Ne height 550 295 200 375	0.963 0.048 0.026	1.77 1.77 C5 g y 274 301 316 300	x1 60 65 72 80	x2 25 25 26 25	x3 87 91 94 87	x4 88 94 99 97	Heat 8 9 10 10	x1_1 6 11 9 7 2 11 7 3 5 1	x2_1 31 52 55 71 31	x3_1 8 6 9 17	x4_1 47 33 22 6	C16	Time 1300 1300 1300 1300 1200	Velocity 0.0130 0.0135 0.0120 0.0400	Temperature 13.5 17.0 23.0 5.3	Yei
	Coefficients Constant x1 0 x2 1 C1 Pull strength 35.00 25.02 16.86 14.38 9.60 24.35	Coef         SE Coef           3.9         81.9           573         0.250           9.88         3.72           C2         Wire length         D           10         8           4         2           2         2	T-Value 0.05 2.29 2.66 C3 Die height 550 295 200 375 52 100	0.963 0.048 0.026 C4	1.77 1.77 <b>C5</b> g y 274 301 316 300 296	x1 60 65 72 80 84	x2 25 25 26 25 25 25	x3 87 91 94 87 86	x4 88 94 99 97 96	Heat 8 9 10 10 7	x1_1 6 11 9 7 2 11 7 3 5 1	x2_1 31 52 55 71 31	x3_1 8 6 9 17 22	x4_1 47 33 22 6 44	C16	Time 1300 1300 1300 1200 1200	Velocity 0.0130 0.0135 0.0120 0.0400 0.0380	Temperature 13.5 17.0 23.0 5.3	Yei
1 5 7 3 9	Coefficients Term C Constant x1 0. x2 1 Pull strength 35.00 25.02 16.86 14.38 9.60 24.35 2 4.35	Code         SE Codef           3.9         81.9           573         0.250           8.88         3.72           C2           Wire length         D           10         8           4         2           2         2           9         Muttiple Linear	T-Value 0.05 2.29 2.66 C3 Die height 550 295 200 375 52 100 Regression	0.963 0.048 0.026 C4	1.77 1.77 <b>C5</b> g y 274 301 316 300 296	x1 60 65 72 80 84	x2 25 25 26 25 25 25	x3 87 91 94 87 86	x4 88 94 99 97 96	Heat 8 9 10 10 10 7 9	x1_1 6 11 9 7 2 11 7 3 5 1	x2_1 31 52 55 71 31	x3_1 8 6 9 17 22	x4_1 47 33 22 6 44		Time 1300 1300 1300 1200 1200 1200	Velocity 0.0130 0.0135 0.0120 0.0400 0.0380 0.0	Temperature 13.5 17.0 23.0 5.3	Yei
1 5 7 3 9	Coefficients Term C Constant ×1 0, ×2 1 C1 Pull strength 35,00 25,02 16,86 14,38 9,50 24,35 ▷ H + _ C4,45 ▷ H + _	Coef         SE Coef           3.9         81.9           573         0.250           8.88         3.72           C2         Vire length         C           10         8           4         2           2         2           9         9	T-Value 0.05 2.29 2.66 C3 Xie height 550 295 200 375 52 100 Regression	0.963 0.048 0.026 C4	1.77 1.77 <b>C5</b> 2 <b>y</b> 274 301 316 300 296 267	x1 60 65 72 80 84	x2 25 25 26 25 25 25 24	x3 87 91 94 87 86	x4 88 94 99 97 96 110	Heat 8 9 10 10 10 7 9	x1_1 6 11 9 7 2 11 7 3 5 1 1 2	x2_1 31 52 55 71 31	x3_1 8 6 9 17 22	x4_1 47 33 22 6 44		Time 1300 1300 1300 1200 1200 1200	Velocity 0.0130 0.0135 0.0120 0.0400 0.0380 0.0	Temperature 13.5 17.0 23.0 5.3	Yei

And I save the residual over here and I go the I do the normality test over here with the residuals residual 3 which is saved over here. And I do this and what I see is that the residuals are perfectly following normal distributions like that 0.253 like that.

(Refer Slide Time: 31:07)



#### (Refer Slide Time: 31:10)



So, we will always go by suggestions that is what statistician has suggested. So, we go by stepwise regression. We do not add unnecessary variables which are not nonsignificant terms, but whenever I am removing a non-significant terms please remember that we are losing some amount of information.

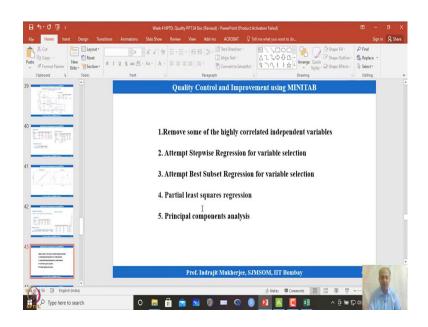
So, sometimes people can suggest why we should remove that one. So, there is a we can we can debate on that that which want to detain which one to remove like that. So, this is an art and this is not perfectly black and white scenarios like that in regression, at least multiple regression like that. But there are suggestions which can be incorporated like that. So, based on which we can select the variables..

So, one I have shown is that best subset regression is one of the option when we have different combinations of the variables and we can select one or two of them and then try to figure out which model is basically good.

Or we use stepwise regression and forget about everything of combinations like that, so whichever is the best will be the veiled and that model we will recommend like that, ok. But you should be careful about the model adequacy checks and all these things, ok.

Even if you have done stepwise regression also finally, you have to make a check of model adequacy over there, ok. So, that is the suggestions and there are other ways of dealing with multicollinearity coordinating which is more statistically sound like that.

### (Refer Slide Time: 32:37)



So, one is partially square regression and one is principle component analysis based regression like that, ok. So, these things can be adopted.

So, we will stop here and we will we will try to discuss about another example where the multiple regression fails like that and error assumption fails and in that case how we have to deal with that. That is not discussed. It is discussed in simple regression. So, we will start from here. And another example I have on this time velocity temperature and yield and selection of the variables over here also we will discuss.

And then, we will move into the core concept which is the improvement phase and that is design of experiment. We will try to emphasize now on design of experiments and how do we how do you do design of experiments what are the things. So, basic idea of design of experiments in our next session basically.

Thank you for listening.