6.5 Multiple Linear Regression - Variance Inflation Factor - Part 2

Monday, 07 November 2022 18:57

Summary	 An example where explanatory variables on their own are significant (SLR), but when taken together become insignificant (MLR). Signs and Remedies of Collinearity 						
	Interpreting Multiple Regression						
	Example: Estimating the price of a house						
	Response variable: Price of the house (INR).						
	 Three explanatory variables: Size of the house (in square foot), number of bedrooms and the number of parking lots provided. 						
	Apartment data is given:						
	Area (Sq ft) # of bedrooms Parking lot Price						
	9.5 2 2 5.68 10 3 2 8.9						
	8.7 2 1 7.6						
	10 3 3 10						
	11.45 2 2 8 20 2 3 9.8						
	9 2 2 8.1						
	8.34 2 2 7.1 11 3 2 9.1						
	13 3 1 5						
	14.5 4 3 12 16 3 3 11.5						
	16 3 3 11.5 8.19 1 2 6.4						
	7.9 1 2 7						
	8.6 2 3 8 10 2 2 7.9						
	11.45 3 1 9						
	12 3 3 9.35						
	15 4 3 10.3 11 3 2 14						
	 We're trying to predict the price of an apartment. Price given in the data is in 10 Lakhs. Area is given in 100 sq ft. 						
	Here's the correlation coefficient matrix:						
	Area (Sq ft) # of bedrooms Parking lot Price						
	Area (Sq ft) 1						
	# of bedrooms 0.503295389 1						
	Parking lot 0.434453051 0.274318858 1						
	Price 0.467910037 0.605720798 0.501392503 1						
	Price seems to be affecting by all of the three factors.						
	SLR on Area-Price						
	Response Variable: Price Explanatory Variable: Area						
	SUMMARY OUTPUT						
	Regression Statistics						
	Multiple R 0.467910037						
	R Square 0.218939803 Adjusted R Square 0.17554757 Standard Error 1.972955837 Observations 20 ANOVA df SS MS F Significance F						
	Regression 1 19.64026979 19.64026979 5.045598878 0.037477004						
I	Residual 18 70.06598521 3.892554734						

	Coefficients S		t Stat	P-value	Lower 95%	Upper 95%
Intercept	5.040157858		2.958396796 0 2.246241055 0			8.619454166
Area (Sq ft)	0.327646336	0.143604261	2.240241033 0	.037477004	0.021196854	0.034093619
• Significance	F values tells	us that the	regression	is significa	ant ($p-value$	e < 0.05).
		SLR on N	lumber of	Bedroo	m-Price	
Response Varial		CD . 1				
Explanatory Vai	riable: Numi	er of Bearo	oms			
SUMMARY OUTPUT						
Regression St						
Multiple R	0.605720798					
R Square Adjusted R Square	0.366897685					
Standard Error	1.776282599					
Observations	20					
ANOVA						_
	df	SS	MS	F	Significance I	<u> </u>
Regression	1		32.91301731	10.43142345	0.00464791	1
Residual	18		3.155179872			
Total	19	89.706255				_
	Coefficients	Standard Error	t Stat	P-value	Lower 95%	Unner 95%
Intercept	4.758615385	1.294091339	3.677186642	0.00172406		<i>Upper 95%</i> 9 7.477400401
# of bedrooms	1.591153846		3.229771424			9 2.626177613
kesponse variai	ble: Price					
Explanatory Vai		ng lot				
		ng lot				
Explanatory Vai	riable: Parkii	ng lot				
Explanatory Var SUMMARY OUTPUT Regression St	riable: Parkii	ng lot				
Explanatory Var SUMMARY OUTPUT Regression St Multiple R R Square	riable: Parkin atistics 0.501392503 0.251394442	ng lot				
Explanatory Var SUMMARY OUTPUT Regression St Multiple R R Square Adjusted R Square	atistics 0.501392503 0.251394442 0.209805244	ng lot				
Explanatory Var SUMMARY OUTPUT Regression St Multiple R R Square Adjusted R Square Standard Error	atistics 0.501392503 0.251394442 0.209805244 1.931530785	ng lot				
Explanatory Var SUMMARY OUTPUT Regression St Multiple R R Square Adjusted R Square	atistics 0.501392503 0.251394442 0.209805244	ng lot				
Explanatory Variable SUMMARY OUTPUT Regression St. Multiple R R Square Adjusted R Square Standard Error	atistics 0.501392503 0.251394442 0.209805244 1.931530785	ng lot				
Explanatory Var SUMMARY OUTPUT Regression St Multiple R R Square Adjusted R Square Standard Error Observations ANOVA	atistics 0.501392503 0.251394442 0.209805244 1.931530785 20	SS	MS		ignificance F	
Explanatory Variable Summary Output Regression St. Multiple R R Square Adjusted R Square Standard Error Observations ANOVA Regression	tiable: Parking attistics 0.501392503 0.251394442 0.209805244 1.931530785 20 df 1	SS 22.55165391	22.55165391	F S	ignificance F 0.024307596	
Explanatory Variable Summary Output Regression St. Multiple R R Square Adjusted R Square Standard Error Observations ANOVA Regression Residual	tiable: Parking attistics 0.501392503 0.251394442 0.209805244 1.931530785 20 df 1 18	SS 22.55165391 67.15460109	22.55165391			
Explanatory Variable Summary Output Regression St. Multiple R R Square Adjusted R Square Standard Error Observations ANOVA Regression	tiable: Parking attistics 0.501392503 0.251394442 0.209805244 1.931530785 20 df 1	SS 22.55165391	22.55165391			
Explanatory Variable Summary Output Regression St. Multiple R R Square Adjusted R Square Standard Error Observations ANOVA Regression Residual	tiable: Parking attistics 0.501392503 0.251394442 0.209805244 1.931530785 20 df 1 18	SS 22.55165391 67.15460109 89.706255	22.55165391 (3.730811171		0.024307596	Upper 95% ı
Explanatory Variable Explanatory Variable Explanatory Variable Expression St. Multiple R R Square Adjusted R Square Standard Error Observations ANOVA Regression Residual Total	distics 0.501392503 0.251394442 0.209805244 1.931530785 20 df 1 18 19	SS 22.55165391 67.15460109 89.706255	22.55165391 (3.730811171	5.04470526 P-value	0.024307596	
Explanatory Variable Explanatory Variable Explanatory Variable Expression St. Multiple R R Square Adjusted R Square Standard Error Observations ANOVA Regression Residual Total	distics 0.501392503 0.251394442 0.209805244 1.931530785 20 df 1 18 19 Coefficients Si	SS 22.55165391 67.15460109 89.706255 candard Error 1.466039445	22.55165391 (3.730811171 t Stat	P-value 002003183	0.024307596 Lower 95%	8.372099798
Explanatory Variable Explanatory Variable Explanatory Variable Expression St. Multiple R R Square Adjusted R Square Standard Error Observations ANOVA Regression Residual Total Intercept	distics 0.501392503 0.251394442 0.209805244 1.931530785 20 df 1 18 19 Coefficients Si 5.292065217 1.565652174	SS 22.55165391 67.15460109 89.706255 candard Error 1.466039445 0.636806842	22.55165391 (3.730811171	P-value 002003183 024307596	0.024307596 Lower 95% 2.212030636 0.227770645	8.372099798 2.903533703
Explanatory Var SUMMARY OUTPUT Regression St Multiple R R Square Adjusted R Square Standard Error Observations ANOVA Regression Residual Total Intercept Parking lot • Significance	distics 0.501392503 0.251394442 0.209805244 1.931530785 20 df 1 18 19 Coefficients St 5.292065217 1.565652174 F values tells	\$\$ 22.55165391 67.15460109 89.706255 andard Error 1.466039445 0.636806842	22.55165391 3.730811171 t Stat 3.609770008 0. 2.458598231 0. regression	P-value 002003183 024307596 is significa	0.024307596 Lower 95% 2.212030636 0.227770645 ant (p - value	8.372099798 2.903533703 e < 0.05).
Explanatory Var SUMMARY OUTPUT Regression St Multiple R R Square Adjusted R Square Standard Error Observations ANOVA Regression Residual Total Intercept Parking lot • Significance	distics 0.501392503 0.251394442 0.209805244 1.931530785 20 df 1 18 19 Coefficients St 5.292065217 1.565652174 F values tells	\$\$ 22.55165391 67.15460109 89.706255 andard Error 1.466039445 0.636806842	22.55165391 3.730811171 t Stat 8.609770008 0. 2.458598231 0. regression i	P-value 002003183 024307596 is significa	0.024307596 Lower 95% 2.212030636 0.227770645 ant (p - value	8.372099798 2.903533703 e < 0.05).
Explanatory Van SUMMARY OUTPUT Regression St Multiple R R Square Adjusted R Square Standard Error Observations ANOVA Regression Residual Total Intercept Parking lot Significance So, by themse	distics 0.501392503 0.251394442 0.209805244 1.931530785 20 df 1 18 19 Coefficients St 5.292065217 1.565652174 F values tells	\$\$ 22.55165391 67.15460109 89.706255 andard Error 1.466039445 0.636806842	22.55165391 3.730811171 t Stat 3.609770008 0. 2.458598231 0. regression	P-value 002003183 024307596 is significa	0.024307596 Lower 95% 2.212030636 0.227770645 ant (p - value	8.372099798 2.903533703 e < 0.05).
Explanatory Van SUMMARY OUTPUT Regression St Multiple R R Square Adjusted R Square Standard Error Observations ANOVA Regression Residual Total Intercept Parking lot Significance Response Varial	distics 0.501392503 0.251394442 0.209805244 1.931530785 20 df 1 18 19 Coefficients St 5.292065217 1.565652174 F values tells elves, each experies the company of the com	\$\$ 22.55165391 67.15460109 89.706255 20.636806842 8 us that the Explanatory V	22.55165391 3.730811171 t Stat 3.609770008 0. 2.458598231 0. regression i	P-value 002003183 024307596 is significa	0.024307596 Lower 95% 2.212030636 0.227770645 ant (p - value) o predict the	8.372099798 2.903533703 e < 0.05).
Explanatory Var SUMMARY OUTPUT Regression St Multiple R R Square Adjusted R Square Standard Error Observations ANOVA Regression Residual Total Intercept Parking lot Significance Response Varial	distics 0.501392503 0.251394442 0.209805244 1.931530785 20 df 1 18 19 Coefficients St 5.292065217 1.565652174 F values tells elves, each experies the company of the com	\$\$ 22.55165391 67.15460109 89.706255 20.636806842 8 us that the Explanatory V	22.55165391 3.730811171 t Stat 3.609770008 0. 2.458598231 0. regression i	P-value 002003183 024307596 is significa	0.024307596 Lower 95% 2.212030636 0.227770645 ant (p - value) o predict the	8.372099798 2.903533703 e < 0.05).
Explanatory Var SUMMARY OUTPUT Regression St Multiple R R Square Adjusted R Square Standard Error Observations ANOVA Regression Residual Total Intercept Parking lot Significance Response Varial	distics 0.501392503 0.251394442 0.209805244 1.931530785 20 df 1 18 19 Coefficients St 5.292065217 1.565652174 F values tells elves, each exception of the control of the	\$\$ 22.55165391 67.15460109 89.706255 20.636806842 8 us that the Explanatory V	22.55165391 3.730811171 t Stat 3.609770008 0. 2.458598231 0. regression i	P-value 002003183 024307596 is significa	0.024307596 Lower 95% 2.212030636 0.227770645 ant (p - value) o predict the	8.372099798 2.903533703 e < 0.05).
Explanatory Van SUMMARY OUTPUT Regression St Multiple R R Square Adjusted R Square Standard Error Observations ANOVA Regression Residual Total Intercept Parking lot Significance So, by themse Response Varial Explanatory Van SUMMARY OUTPUT	distics 0.501392503 0.251394442 0.209805244 1.931530785 20 df 1 18 19 Coefficients St 5.292065217 1.565652174 F values tells elves, each experience riable: Area,	\$\$ 22.55165391 67.15460109 89.706255 20.636806842 8 us that the Explanatory V	22.55165391 3.730811171 t Stat 3.609770008 0. 2.458598231 0. regression i	P-value 002003183 024307596 is significa	0.024307596 Lower 95% 2.212030636 0.227770645 ant (p - value) o predict the	8.372099798 2.903533703 e < 0.05).
Explanatory Varial Explanatory V	distics 0.501392503 0.251394442 0.209805244 1.931530785 20 df 1 18 19 Coefficients St 5.292065217 1.565652174 F values tells elves, each experiable: Area,	\$\$ 22.55165391 67.15460109 89.706255 andard Error 1.466039445 0.636806842 s us that the splanatory v Number of l	22.55165391 3.730811171 t Stat 3.609770008 0. 2.458598231 0. regression i	P-value 002003183 024307596 is significa	0.024307596 Lower 95% 2.212030636 0.227770645 ant (p - value) o predict the	8.372099798 2.903533703 e < 0.05).
Explanatory Van SUMMARY OUTPUT Regression St Multiple R R Square Adjusted R Square Standard Error Observations ANOVA Regression Residual Total Intercept Parking lot Significance So, by themse Response Varial Explanatory Van SUMMARY OUTPUT	distics 0.501392503 0.251394442 0.209805244 1.931530785 20 df 1 18 19 Coefficients St 5.292065217 1.565652174 F values tells elves, each experience riable: Area,	\$\$ 22.55165391 67.15460109 89.706255 andard Error 1.466039445 0.636806842 s us that the splanatory v Number of l	22.55165391 3.730811171 t Stat 3.609770008 0. 2.458598231 0. regression i	P-value 002003183 024307596 is significa	0.024307596 Lower 95% 2.212030636 0.227770645 ant (p - value) o predict the	8.372099798 2.903533703 e < 0.05).
Explanatory Variance And Summary Output Regression St. Multiple R R Square Adjusted R Square Standard Error Observations ANOVA Regression Residual Total Intercept Parking lot Significance Significance Summary Output Regression S Multiple R	distics 0.501392503 0.251394442 0.209805244 1.931530785 20 df 1 18 19 Coefficients Si 5.292065217 1.565652174 F values tells elves, each except ex	\$\$ 22.55165391 67.15460109 89.706255 andard Error 1.466039445 0.636806842 s us that the explanatory v	22.55165391 3.730811171 t Stat 3.609770008 0. 2.458598231 0. regression i	P-value 002003183 024307596 is significa	0.024307596 Lower 95% 2.212030636 0.227770645 ant (p - value) o predict the	8.372099798 2.903533703 e < 0.05).
Explanatory Variations Summary Output Regression St. Multiple R R Square Adjusted R Square Standard Error Observations ANOVA Regression Residual Total Intercept Parking lot Significance Significance Standard Error Observations ANOVA Regression Residual Total Intercept Parking lot Response Varial Explanatory Variations Summary Output Regression S Multiple R R Square	distics 0.501392503 0.251394442 0.209805244 1.931530785 20 df 1 18 19 Coefficients Si 5.292065217 1.565652174 F values tells elves, each except ex	\$\$ 22.55165391 67.15460109 89.706255 8andard Error 1.466039445 0.636806842 8 us that the explanatory v	22.55165391 3.730811171 t Stat 3.609770008 0. 2.458598231 0. regression i	P-value 002003183 024307596 is significa	0.024307596 Lower 95% 2.212030636 0.227770645 ant (p - value) o predict the	8.372099798 2.903533703 e < 0.05).
Explanatory Variations Summary Output Regression St. Multiple R R Square Adjusted R Square Standard Error Observations ANOVA Regression Residual Total Intercept Parking lot Significance Significance Standard Error Observations ANOVA Regression Residual Total Intercept Parking lot Response Varial Explanatory Variations Summary Output Regression S Multiple R R Square Adjusted R Square	distics 0.501392503 0.251394442 0.209805244 1.931530785 20 df 1 18 19 Coefficients Si 5.292065217 1.565652174 F values tells elves, each exception of the company of the	\$\$ 22.55165391 67.15460109 89.706255 8andard Error 1.466039445 0.636806842 8 us that the explanatory v	22.55165391 3.730811171 t Stat 3.609770008 0. 2.458598231 0. regression i	P-value 002003183 024307596 is significa	0.024307596 Lower 95% 2.212030636 0.227770645 ant (p - value) o predict the	8.372099798 2.903533703 e < 0.05).
Explanatory Van SUMMARY OUTPUT Regression St Multiple R R Square Adjusted R Square Standard Error Observations ANOVA Regression Residual Total Intercept Parking lot Significance Significance Standard Square SUMMARY OUTPUT Regression S Multiple R R Square Adjusted R Square Standard Error	distics 0.501392503 0.251394442 0.209805244 1.931530785 20 df 1 18 19 Coefficients Si 5.292065217 1.565652174 F values tells elves, each except ex	\$\$ 22.55165391 67.15460109 89.706255 8andard Error 1.466039445 0.636806842 8 us that the explanatory v	22.55165391 3.730811171 t Stat 3.609770008 0. 2.458598231 0. regression i	P-value 002003183 024307596 is significa	0.024307596 Lower 95% 2.212030636 0.227770645 ant (p - value) o predict the	8.372099798 2.903533703 e < 0.05).
Explanatory Van SUMMARY OUTPUT Regression St Multiple R R Square Adjusted R Square Standard Error Observations ANOVA Regression Residual Total Intercept Parking lot Significance Significance Standard Square SUMMARY OUTPUT Regression S Multiple R R Square Adjusted R Square Standard Error	distics 0.501392503 0.251394442 0.209805244 1.931530785 20 df 1 18 19 Coefficients Si 5.292065217 1.565652174 F values tells elves, each except ex	\$\$ 22.55165391 67.15460109 89.706255 8andard Error 1.466039445 0.636806842 8 us that the explanatory v	22.55165391 3.730811171 t Stat 3.609770008 0. 2.458598231 0. regression i	P-value 002003183 024307596 is significa	0.024307596 Lower 95% 2.212030636 0.227770645 ant (p - value) o predict the	8.372099798 2.903533703 e < 0.05).
Explanatory Var SUMMARY OUTPUT Regression St. Multiple R R Square Adjusted R Square Standard Error Observations ANOVA Regression Residual Total Intercept Parking lot Significance Significance Standard Explanatory Var SUMMARY OUTPUT Regression S Multiple R R Square Adjusted R Square Standard Error Observations ANOVA	distics 0.501392503 0.251394442 0.209805244 1.931530785 20 df 1 18 19 Coefficients Si 5.292065217 1.565652174 F values tells elves, each exchesion exchision exchesion exchasion exchesion exchasion exchasion exchesion exchasion excha	\$\$ 22.55165391 67.15460109 89.706255 20.636806842 8 us that the Explanatory v Number of l	22.55165391 3.730811171 t Stat 3.609770008 0. 2.458598231 0. regression is ariable does MLJ bedrooms, I	P-value 002003183 024307596 is significate s help us t R Parking lo	0.024307596 Lower 95% 2.212030636 0.227770645 ant (p - value 0 predict th	8.372099798 2.903533703 e < 0.05). he prices.
Explanatory Variable Response Varial Explanatory Variable R Response Varial Explanatory Variable R R Square Adjusted R Square Standard Error Observations ANOVA Regression Residual Total Intercept Parking lot • Significance A Square Standard Error Observations Multiple R R Square Adjusted R Square Adjusted R Square Standard Error Observations ANOVA Regression Regression	distics 0.501392503 0.251394442 0.209805244 1.931530785 20 df 1 18 19 Coefficients Si 5.292065217 1.565652174 F values tells elves, each ex colle: Price riable: Area, f Statistics 0.702259246 0.493168049 0.398137058 1.685711946 20 df 3	\$\$ 22.55165391 67.15460109 89.706255 20.636806842 30.636806842 30.636806842 30.636806842 30.636806842 30.636806842 30.636806842	22.55165391 3.730811171 t Stat 3.609770008 0. 2.458598231 0. regression i ariable does MLJ bedrooms, I	P-value 002003183 024307596 is significate s help us t R Parking lo	0.024307596 Lower 95% 2.212030636 0.227770645 ant (p - value 0 predict th	8.372099798 2.903533703 e < 0.05). he prices.
Explanatory Variable Explanato	distics 0.501392503 0.251394442 0.209805244 1.931530785 20 df 1 18 19 Coefficients Si 5.292065217 1.565652174 F values tells elves, each exchesion exchision exchesion exchasion exchesion exchasion exchasion exchesion exchasion excha	\$\$ 22.55165391 67.15460109 89.706255 20.636806842 30.6368	22.55165391 3.730811171 t Stat 3.609770008 0. 2.458598231 0. regression i ariable does MLJ bedrooms, I	P-value 002003183 024307596 is significate s help us t R Parking lo	0.024307596 Lower 95% 2.212030636 0.227770645 ant (p - value 0 predict th	8.372099798 2.903533703 e < 0.05). he prices.

Total

19

89.706255

	Coefficients	Standard Error	t Stat	P-value	Lower 95%	Upper 95%
Intercept	2.667893129	1.67836639	1.589577309	0.131492473	-0.890084674	6.225870932
Area (Sq ft)	0.059750485	0.154380203	0.387034628	0.70383001	-0.267520926	0.387021895
# of bedrooms	1.236821718	0.542442699	2.280096533	0.036650141	0.086894565	2.38674887
Parking lot	1.046580675	0.618622039	1.691793388	0.110065941	-0.264839463	2.358000813

- R^2 indicates that the fitted equation explains % of the variation in price.
- The overall $Significance\ F$ values of the MLR is significant ($p-value\ <\ 0.05$).
- But

• But,	
	p – value for Area
P-value Intercept 0.131492473 Area (Sq ft) 0.70383001 # of bedrooms 0.036650141 Parking lot 0.110065941	$H_0: eta_1=0$ $H_0: eta_1\neq 0$ $\therefore p-value=0.7>0.05$ null hypothesis cannot be rejected. Also, look at the confidence interval. b_1 can be anywhere between $[-0.26,\ 0.38]$. It includes 0, we cannot reject it.
	p-value for # of bedrooms
Lower 95% Upper 95%	$H_0: eta_2 = 0$ $H_0: eta_2 \neq 0$ $\therefore p-value = 0.03 < 0.05$ null hypothesis is rejected. $\therefore b_2 = 1.23 \text{ is a good estimate of } \beta_2.$ $p-value \text{ for Parking lot}$
	$H_0: eta_3 = 0$ $H_0: eta_3 \neq 0$ $\therefore p-value = 0.11 > 0.05$ null hypothesis cannot be rejected.
	Also, the confidence interval for b_3 is $[-0.26, 2.35]$.

- · Individually, area and parking lot were significant.
- It happened because there was a strong correlation among the explanatory variables.

	Area (Sq ft)	# of bedrooms	Parking lot	Price
Area (Sq ft)	1			
# of bedrooms	0.503295389	1		
Parking lot	0.434453051	0.274318858	1	
Price	0.467910037	0.605720798	0.501392503	1

• That's what's making partial slopes insignificant even when the marginal slopes were larger.

	Marginal Slope	Partial Slope
Area	0.327	0.059
# of bedroom	1.59	1.23
Parking lot	1.56	1.04

• Also the standard errors for the partial slopes are larger than the marginal slope.

	Marginal Slope	Partial Slope
Area	0.145	0.154
# of bedroom	0.12	0.54
Parking lot	0.63	0.628

VIF

- From our example, the explanatory variables in the MLR are turning to be insignificant.
- The explanatory variables aren't significant once we have taken account of the other explanatory variables.

 Partial slope conveys the unique variation explained by that particular explanatory variable. · However once you take account of the parking lot and the area, now the number of bedrooms do not offer anything unique that has already not been explained by these two variables. Similarly, once you take account of the area and the number of bedrooms, now the parking lot does not offer anything unique that has not already been explained by these two variables. • This is why the explanatory variables are turning out to be insignificant in MLR, but they are significant in the SLR. • This is the impact of explanatory variables being correlated. R-Square VIF VIF_SQRT Area 0.05975 0.348302 1.534452 1.23873 # of bedrooms 1.236822 0.257125 1.346122 1.160225 Parking lot 1.046581 0.192899 1.239002 1.113104 What signs do these Collinearity parameters show of collinearity: Signs of Collinearity 2. Marginal and partial slopes F-statistic R² increases less than we'd expect. 4. Standard errors for partial and marginal Slopes of correlated explanatory variables in the model change dramatically. slopes • The *F*-statistic is more impressive than individual *t*-statistics. 5. VIF • Standard errors for partial slopes are larger than those for marginal slopes. · Variance inflation factors increase. • We know that whenever we add an explanatory variable, R^2 is supposed to go up. • If there's a multi-collinearity, R^2 does not go up drastically, it goes up only fractionally. Value of Marginal slopes > Partial slopes Standard errors for Partial slopes > Marginal slopes What are some remedies for Collinearity collinearity? · Remedies for Collinearity Remove redundant explanatory variables. • Re-express explanatory variables (e.g., use the average of Market % Change and Dow % Change as an explanatory variable). • Do nothing if the explanatory variables are significant with sensible estimates. · Re-expressing explanatory variables mean we can combine the correlated explanatory variables to create another explanatory variable. Removing Explanatory Variables Issues After adding several explanatory variables to a model, some of those added and some of those originally present may not be statistically significant. • Remove those variables for which both statistics and substance indicate removal (e.g., remove Dow % Change rather than Market % Change).