The Problem of Multicollinearity:

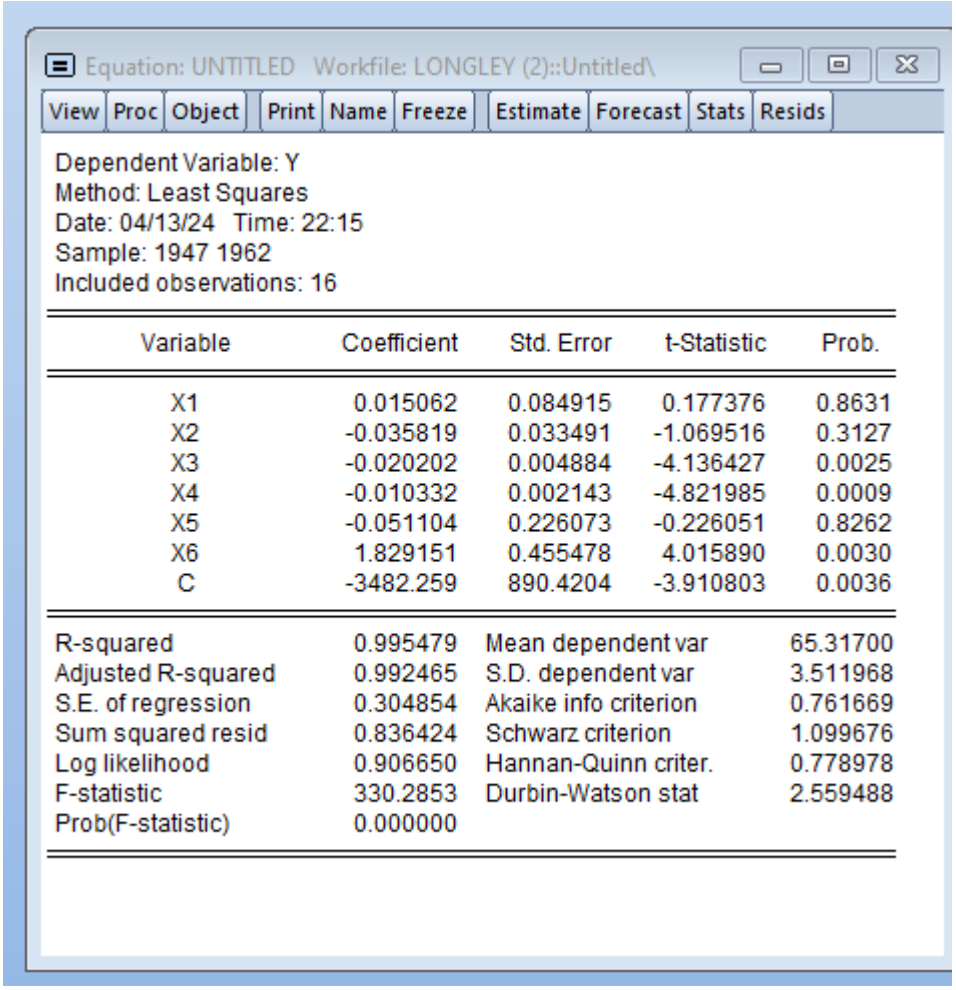
Definition:

Multicollinearity refers to a statistical phenomenon in which two or more independent (predictor) variables in a multiple regression model are highly linearly correlated. This high correlation makes it difficult to determine the individual effect of each variable on the dependent variable because their effects are not distinct.

Detection Techniques:

Utilizing Variance Inflation Factor (VIF) and correlation matrices to identify multicollinearity.

Using Correlation Matrices


Equation: UNTITLED Workfile: LONGLEY (2)::Untitled\
View Proc Object Print Name Freeze Estimate Forecast Stats Resids
Dependent Variable: Y
Method: Least Squares
Date: 04/13/24 Time: 22:15
Sample: 1947 1962
Included observations: 16
Variable Coefficient Std. Error t-Statistic Prob.
X1 0.015062 0.084915 0.177376 0.8631
X2 -0.035819 0.033491 -1.069516 0.3127
X3 -0.020202 0.004884 -4.136427 0.0025
X4 -0.010332 0.002143 -4.821985 0.0009
X5 -0.051104 0.226073 -0.226051 0.8262
X6 1.829151 0.455478 4.015890 0.0030
C -3482.259 890.4204 -3.910803 0.0036
R-squared 0.995479 Mean dependent var 65.31700
Adjusted R-squared 0.992465 S.D. dependent var 3.511968
S.E. of regression 0.304854 Akaike info criterion 0.761669
Sum squared resid 0.836424 Schwarz criterion 1.099676
Log likelihood 0.906650 Hannan-Quinn criter. 0.778978
F-statistic 330.2853 Durbin-Watson stat 2.559488
Prob(F-statistic) 0.000000

Equation: UNTITLED Workfile: LONGLEY01::Untitled\

View | Proc | Object | Print | Name | Freeze | Estimate | Forecast | Stats | Resids

Dependent Variable: Y
Method: Least Squares
Date: 04/17/20 Time: 15:45
Sample: 1947 1962
Included observations: 16

Variable	Coefficient	Std. Error	t-Statistic	Prob.
X1	15.06187	84.91493	0.177376	0.8631
X2	-35.81918	33.49101	-1.069516	0.3127
X3	-20.20230	4.883997	-4.136427	0.0025
X4	-10.33227	2.142742	-4.821985	0.0009
X5	-51.10411	226.0732	-0.226051	0.8262
X6	1829.151	455.4785	4.015890	0.0030
C	-3482259.	890420.4	-3.910803	0.0036

R-squared	0.995479	Mean dependent var	65317.00
Adjusted R-squared	0.992465	S.D. dependent var	3511.968
S.E. of regression	304.8541	Akaike info criterion	14.57718
Sum squared resid	836424.1	Schwarz criterion	14.91519
Log likelihood	-109.6174	Hannan-Quinn criter.	14.59449
F-statistic	330.2853	Durbin-Watson stat	2.559488
Prob(F-statistic)	0.000000		

Equation: UNTITLED Workfile: LONGLEY (2)::Untitled\

View | Proc | Object | Print | Name | Freeze | Estimate | Forecast | Stats | Resids

Dependent Variable: Y
Method: Least Squares
Date: 04/13/24 Time: 22:33
Sample: 1947 1962
Included observations: 16

Variable	Coefficient	Std. Error	t-Statistic	Prob.
X1	0.015062	0.084915	0.177376	0.8631
X2	-0.035819	0.033491	-1.069516	0.3127
X3	-0.020202	0.004884	-4.136427	0.0025
X4	-0.010332	0.002143	-4.821985	0.0009
X5	-0.051104	0.226073	-0.226051	0.8262
X6	1.829151	0.455478	4.015890	0.0030
C	-3482.259	890.4204	-3.910803	0.0036

R-squared	0.995479	Mean dependent var	65.31700
Adjusted R-squared	0.992465	S.D. dependent var	3.511968
S.E. of regression	0.304854	Akaike info criterion	0.761669
Sum squared resid	0.836424	Schwarz criterion	1.099676
Log likelihood	0.906650	Hannan-Quinn criter.	0.778978
F-statistic	330.2853	Durbin-Watson stat	2.559488
Prob(F-statistic)	0.000000		

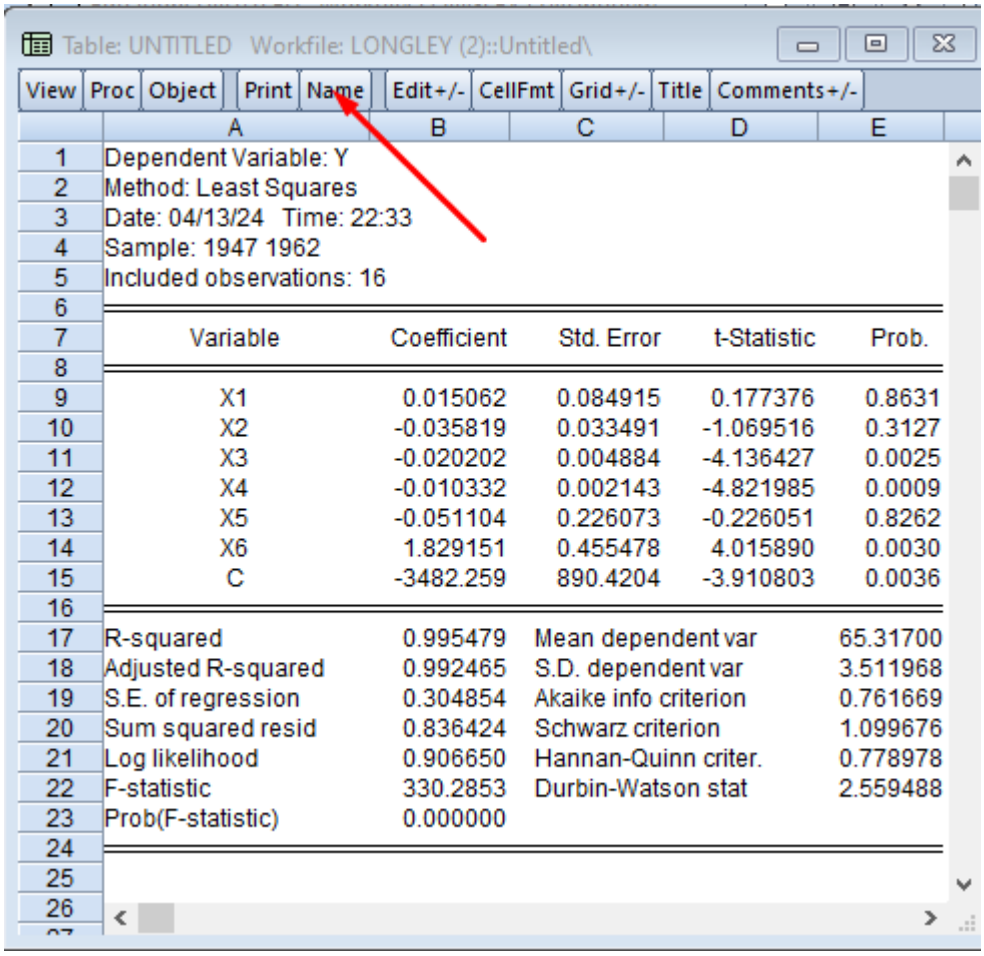
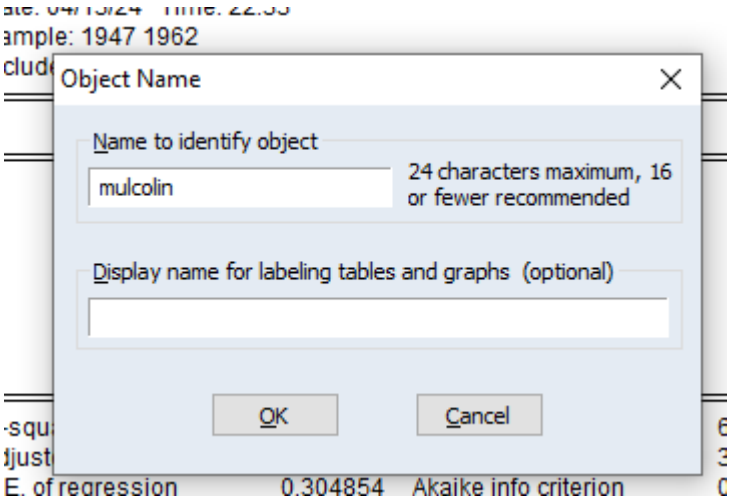
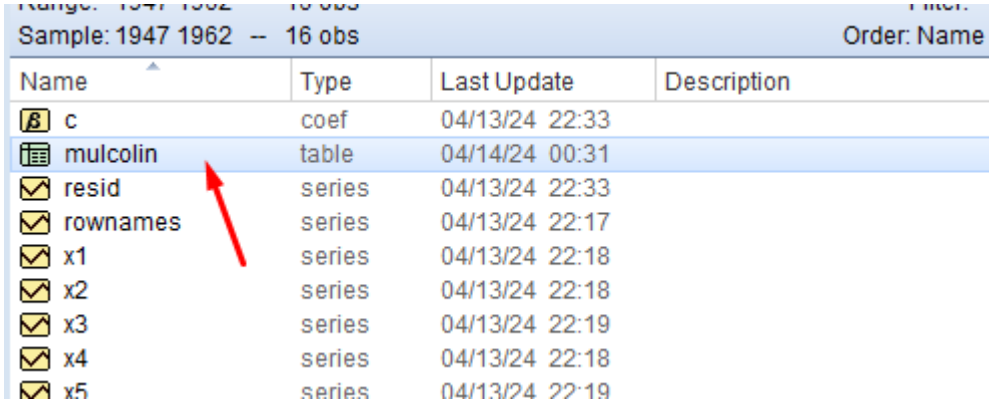
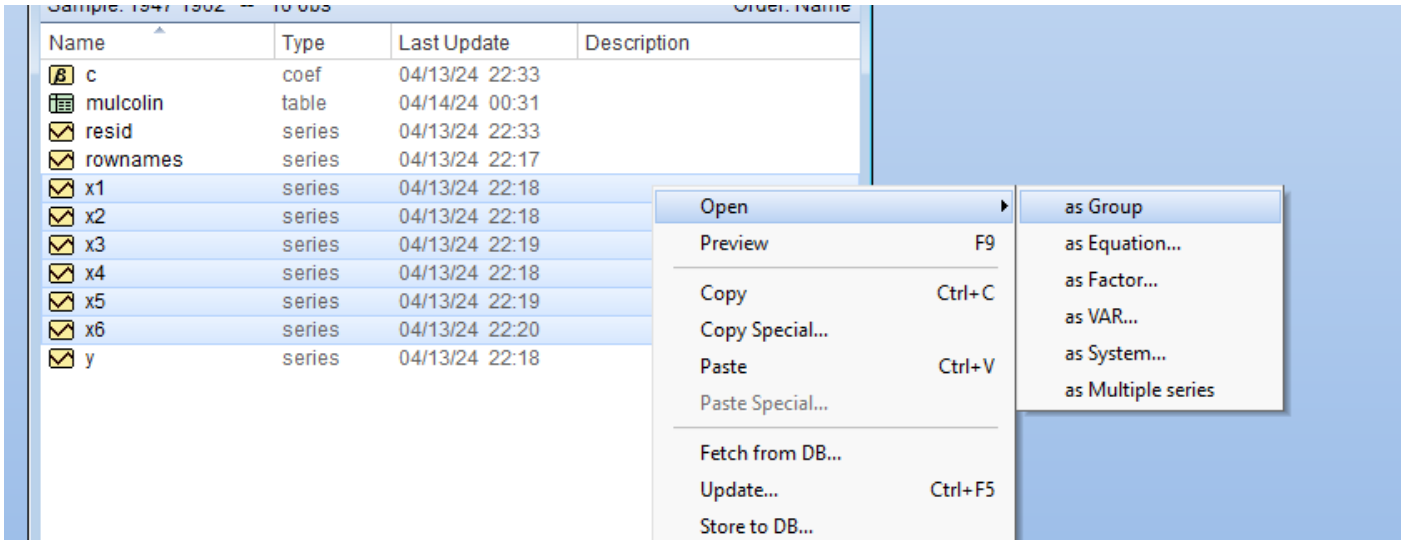
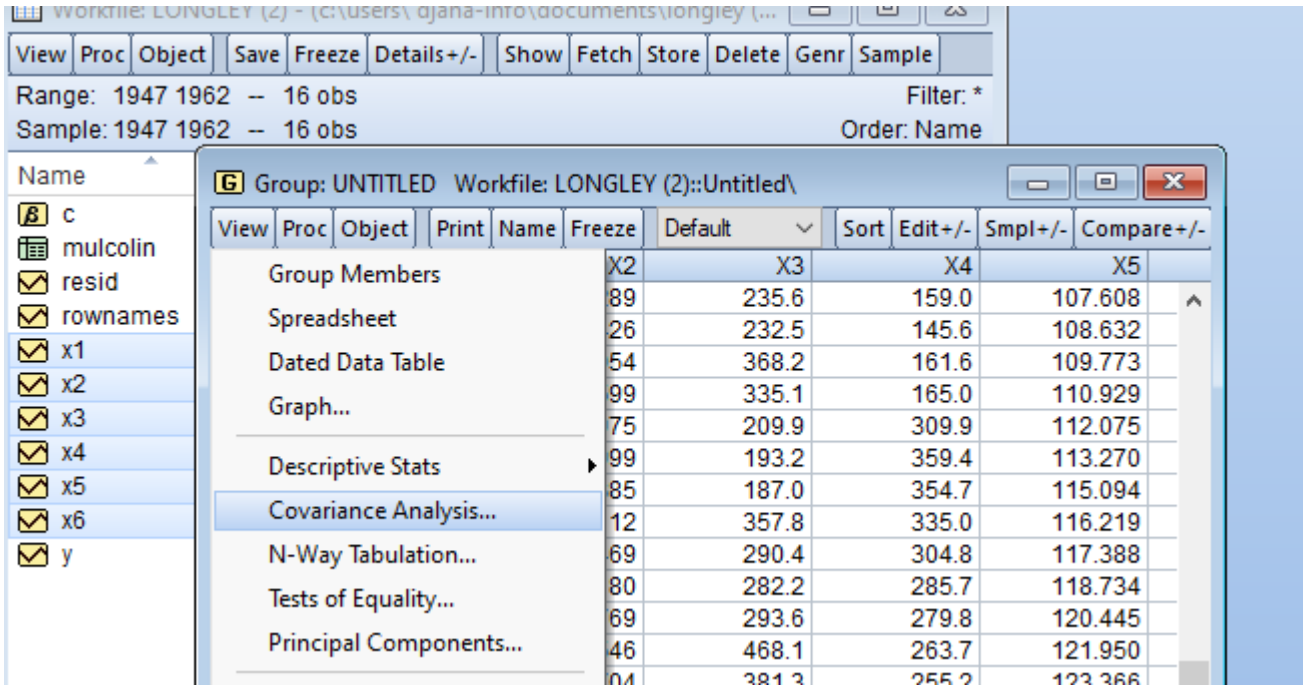


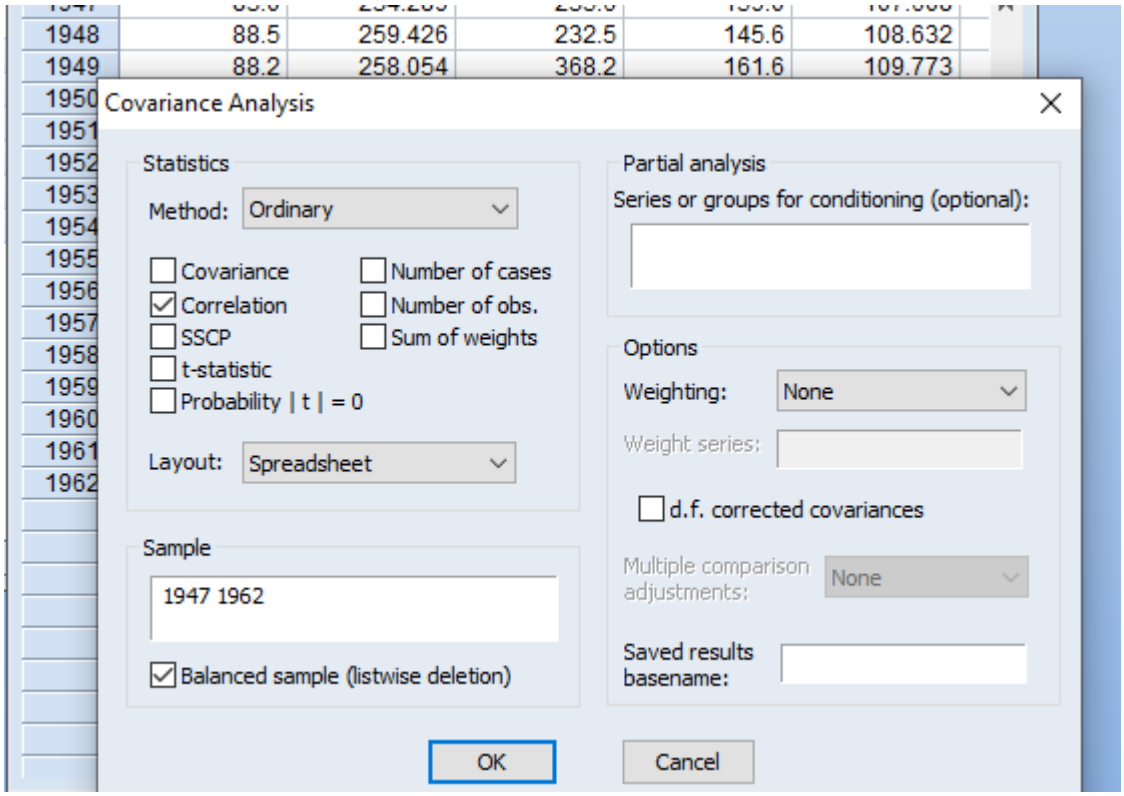
Table: UNTITLED Workfile: LONGLEY (2)::Untitled\
View Proc Object Print Name Edit+/- CellFmt Grid+/- Title Comments+/-
Dependent Variable: Y
Method: Least Squares
Date: 04/13/24 Time: 22:33
Sample: 1947 1962
Included observations: 16
Variable Coefficient Std. Error t-Statistic Prob.
X1 0.015062 0.084915 0.177376 0.8631
X2 -0.035819 0.033491 -1.069516 0.3127
X3 -0.020202 0.004884 -4.136427 0.0025
X4 -0.010332 0.002143 -4.821985 0.0009
X5 -0.051104 0.226073 -0.226051 0.8262
X6 1.829151 0.455478 4.015890 0.0030
C -3482.259 890.4204 -3.910803 0.0036
R-squared 0.995479 Mean dependent var 65.31700
Adjusted R-squared 0.992465 S.D. dependent var 3.511968
S.E. of regression 0.304854 Akaike info criterion 0.761669
Sum squared resid 0.836424 Schwarz criterion 1.099676
Log likelihood 0.906650 Hannan-Quinn criter. 0.778978
F-statistic 330.2853 Durbin-Watson stat 2.559488
Prob(F-statistic) 0.000000

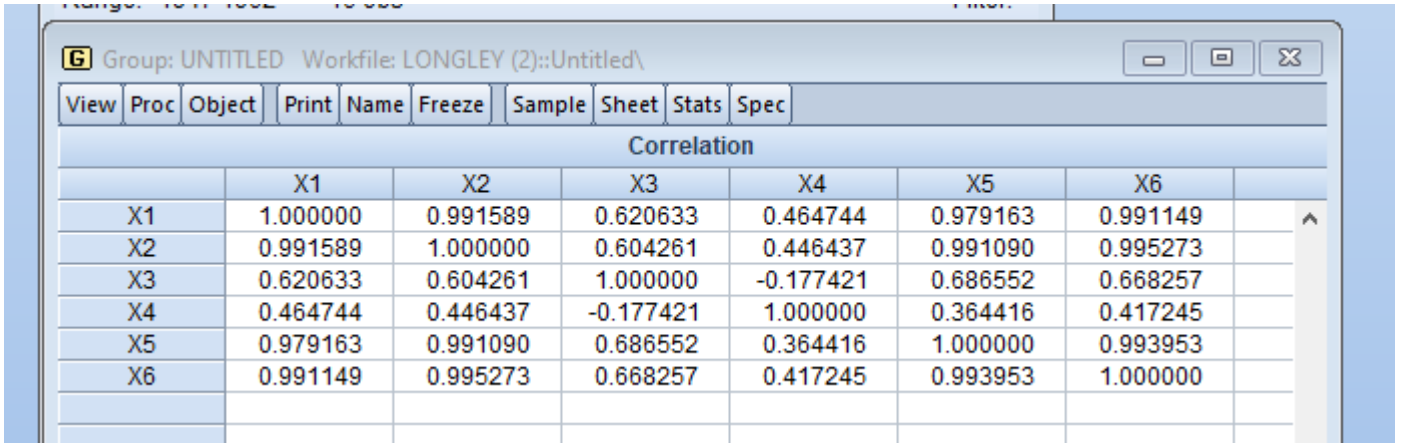

ample: 1947 1962
Object Name
Name to identify object
mulcolin
24 characters maximum, 16 or fewer recommended
Display name for labeling tables and graphs (optional)
OK
Cancel
E. of regression 0.304854 Akaike info criterion


Sample: 1947 1962 -- 16 obs
Order: Name
Name
Type
Last Update
Description
c coef 04/13/24 22:33
mulcolin table 04/14/24 00:31
resid series 04/13/24 22:33
rownames series 04/13/24 22:17
x1 series 04/13/24 22:18
x2 series 04/13/24 22:18
x3 series 04/13/24 22:19
x4 series 04/13/24 22:18
x5 series 04/13/24 22:19

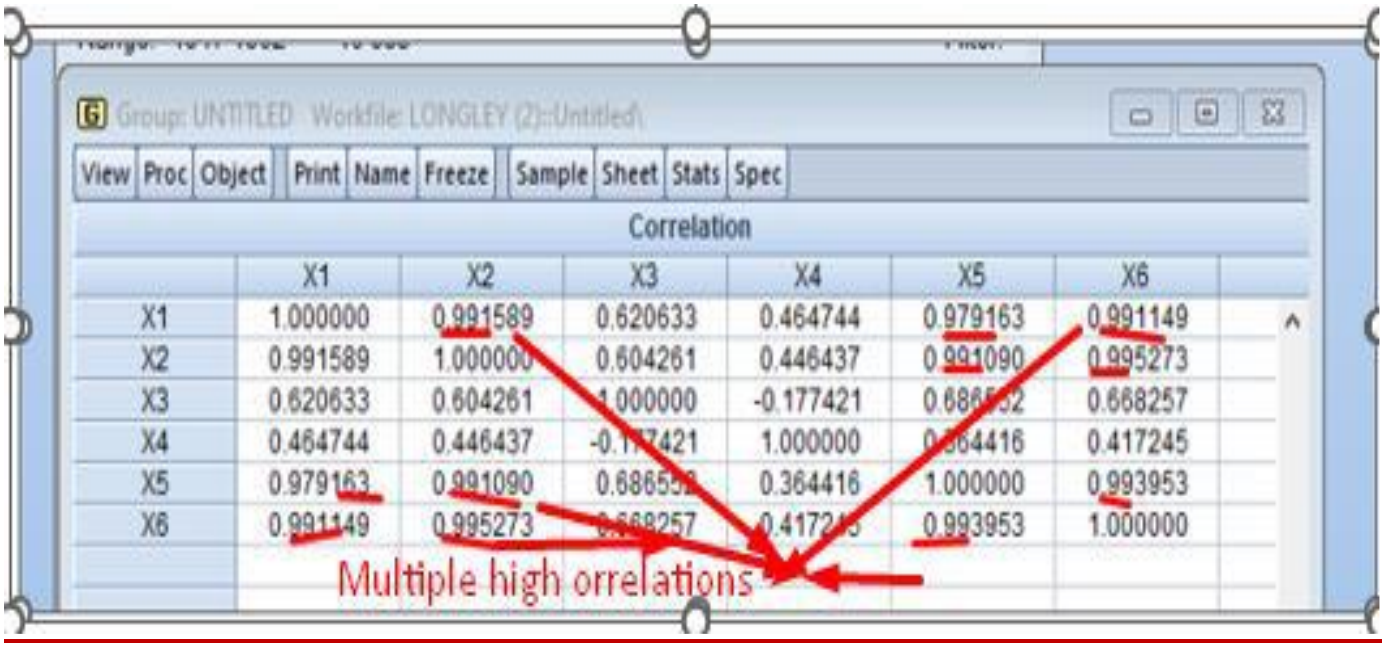

Name
Type
Last Update
Description
c coef 04/13/24 22:33
mulcolin table 04/14/24 00:31
resid series 04/13/24 22:33
rownames series 04/13/24 22:17
x1 series 04/13/24 22:18
x2 series 04/13/24 22:18
x3 series 04/13/24 22:19
x4 series 04/13/24 22:18
x5 series 04/13/24 22:19
x6 series 04/13/24 22:20
y series 04/13/24 22:18
Open
Preview F9
Copy Ctrl+C
Copy Special...
Paste Ctrl+V
Paste Special...
Fetch from DB...
Update... Ctrl+F5
Store to DB...
as Group
as Equation...
as Factor...
as VAR...
as System...
as Multiple series


View Proc Object Save Freeze Details+/- Show Fetch Store Delete Genr Sample
Range: 1947 1962 -- 16 obs
Filter: *
Sample: 1947 1962 -- 16 obs
Order: Name
Name
c
mulcolin
resid
rownames
x1
x2
x3
x4
x5
x6
y
Group: UNTITLED Workfile: LONGLEY (2)::Untitled\
View Proc Object Print Name Freeze Default Sort Edit+/- Smpl+/- Compare+/-
Group Members
Spreadsheet
Dated Data Table
Graph...
Descriptive Stats
Covariance Analysis...
N-Way Tabulation...
Tests of Equality...
Principal Components...
X2 X3 X4 X5
89 235.6 159.0 107.608
26 232.5 145.6 108.632
54 368.2 161.6 109.773
99 335.1 165.0 110.929
75 209.9 309.9 112.075
99 193.2 359.4 113.270
85 187.0 354.7 115.094
12 357.8 335.0 116.219
69 290.4 304.8 117.388
80 282.2 285.7 118.734
69 293.6 279.8 120.445
46 468.1 263.7 121.950





	X1	X2	X3	X4	X5	X6
X1	1.000000	0.991589	0.620633	0.464744	0.979163	0.991149
X2	0.991589	1.000000	0.604261	0.446437	0.991090	0.995273
X3	0.620633	0.604261	1.000000	-0.177421	0.686552	0.668257
X4	0.464744	0.446437	-0.177421	1.000000	0.364416	0.417245
X5	0.979163	0.991090	0.686552	0.364416	1.000000	0.993953
X6	0.991149	0.995273	0.668257	0.417245	0.993953	1.000000


Group: UNTITLED Workfile: LONGLEY (2)::Untitled\
View Proc Object Print Name Freeze Sample Sheet Stats Spec
Correlation
X1 X2 X3 X4 X5 X6
X1 1.000000 0.991589 0.620633 0.464744 0.979163 0.991149
X2 0.991589 1.000000 0.604261 0.446437 0.991090 0.995273
X3 0.620633 0.604261 1.000000 -0.177421 0.686552 0.668257
X4 0.464744 0.446437 -0.177421 1.000000 0.364416 0.417245
X5 0.979163 0.991090 0.686552 0.364416 1.000000 0.993953
X6 0.991149 0.995273 0.668257 0.417245 0.993953 1.000000
Multiple high orrelations