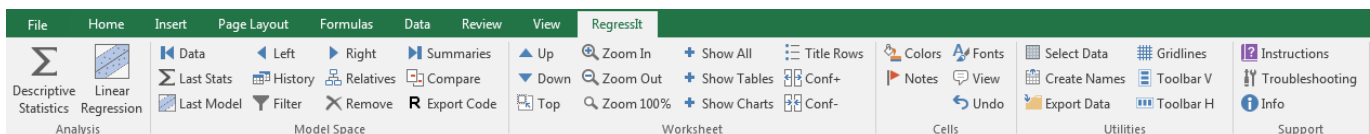


RegressIt

User Manual (Jan. 3, 2018)

RegressIt™ is a free Excel add-in for multivariate descriptive data analysis and regression analysis, first released to the public in 2014 at regressit.com. The 2018 version features a novel 40-button ribbon interface which controls every aspect of the generation and navigation and presentation of your data analysis, as well as providing access to internal documentation and teaching materials. It is not necessary to use the native Excel ribbon or scrolling and paging tools nor do any typing during a RegressIt session apart from optional customization of analysis names. The ribbon interface is designed so that the entire analysis can be driven efficiently not only with a mouse but also by tapping one's finger on a touchscreen, and the output files are easily viewable and navigable on smart phones. There are separate linear regression versions of the program for PC's and Macs, and they have identical features. A separate logistic regression version runs on PC's This document illustrates the use of the linear regression version.

RegressIt can be launched either by opening its file as needed or by installing it on the Excel add-ins menu. When it is running, it adds a new ribbon to Excel's top menu. The buttons on the ribbon are organized into 6 main groups, whose options are described below.



ANALYSIS

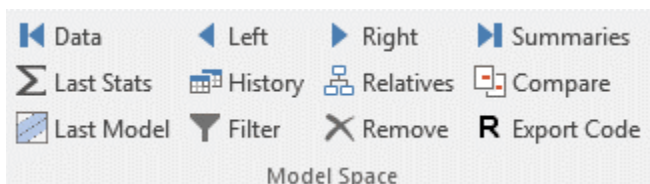


Descriptive Statistics: launches the descriptive stats procedure

Linear Regression: launches the regression procedure

From the dialog boxes of both procedures, you can launch the **variable transformation** procedure. *See the next section of this document, beginning on page 5, for details of how to use these procedures.*

MODEL SPACE NAVIGATION



VCR buttons: move one worksheet to the right or left, or jump to data sheet (leftmost sheet) or model summary sheet (rightmost sheet)

Last stats: toggle back and forth to the last-visited stats worksheet from wherever you are.

Last model: toggle back and forth to the last-visited model worksheet from wherever you are.

Note: these buttons only work if you use RegressIt's navigation tools to move around in the workbook, not if you jump to other worksheets by clicking on their tabs at the bottom of the screen,

History: activates a dialog box with a list of all worksheets in the file and identifying information about the analysis on each one. You can select one and jump directly to it from here. The list remains open while you do this so you can keep jumping around. This also provides a handy view of the audit trail: great for browsing through files produced by yourself or others to see what was done at what date and time on what computer.

Relatives: navigates through a sequence of regression models by parent-child relationships

See page 9 for examples of using the History and Relatives tools.

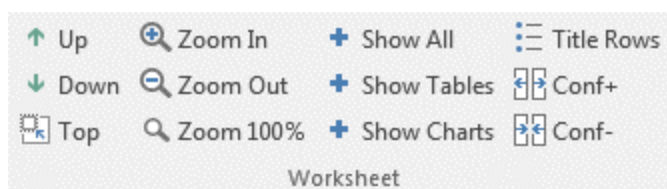
Compare: Hitting this button while positioned on the title row of any table or chart will reposition it at the top of the viewing area and do the same for all other model sheets, so that the same pieces of output will appear in the same relative positions on the screen when moving back and forth among model sheets to compare their results. Also, this tool together with the Top and Zoom tools can be used to re-set the entire workbook to a pristine viewing state, with cell A1 in the top left corner of the window, possibly with a different choice of zoom level to suit a new computer (e.g., for a presentation).

Filter: this executes the Excel filter tool. It can be used to filter and sort the coefficient table (say, on P-values) or to sort a table of squared correlations with the dependent variable (to pre-select independent variables) or to sort the residual table on error size or other stats there.

Remove: clicking this button while selecting some range of cells in the coefficient table for an existing regression model will cause those rows to be grayed-out, which de-selects them from the next regression model launched from that sheet. This is intended to be used in conjunction with the filter tool: sort the coefficient table on P-values and deselect a bunch of insignificant variables all at once. It can also be used on a descriptive statistics output sheet in which the “squared correlations with first variable” option was chosen. If you launch a regression model from that sheet, the variables listed there will be picked up by default unless some have been deselected with this tool. This is convenient for throwing out variables with insignificant correlations with the dependent variable when going from descriptive analysis to regression analysis, particularly when the set of candidate variables is very large.

Export R code: exports R code for the regression model on whose sheet you are positioned. There is a menu of options for table and chart output. If chart output is chosen, the chart titles in R include the model name from RegressIt. After hitting this button and clicking through the options, go to the command prompt in RStudio and hit <Paste><Enter> to run the code. You can also copy R code for regression models from R to RegressIt: just copy and paste the model specification (“...<lm(...)”) text string onto a cell in the Excel file and launch the regression procedure from that cell. *See pages 15-7 for an example.*

WORKSHEET NAVIGATION



Up/down: jump up and down by whole table or whole chart, so that the uppermost one is always correctly aligned with the top left corner of the screen. This is important for efficient viewing because it doesn’t require fine adjustment of the viewing area with the scroll bar. It also minimizes re-drawing speed, which is important for hi-res or editable charts. This feature is also very useful in live presentations.

Zoom in/out: zooms all worksheets at once in a series of predefined levels.

Top: jump to top left of current worksheet

Show all: toggle the display of all charts and tables at once by grouping or ungrouping of their rows. When they are toggled off while title rows remain displayed, you just see a list of all tables and charts on the sheet, and they can be selectively displayed from there.

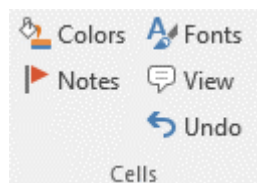
Show Tables: toggle the display of tables only

Show Charts: toggle the display of charts only

Title rows: toggle title rows on and off. Above each table or chart is a title row that includes the same model details (name, dep. variable, sample size) that are in chart titles. For audit trail purposes it is useful to include this extra line in the copied area when tables are individually copied to other documents.

Conf +/- : these buttons interactively increase or decrease the confidence level for confidence limits shown in tables and editable charts on the regression worksheet. (Fun for demonstrations.)

CELL DISPLAY



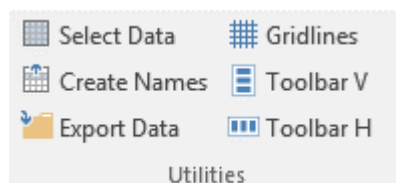
Colors: toggle fill colors on and off. This tool affects correlations and autocorrelations on the descriptive stats sheet, t-stats and standardized coefficients and correlations of coefficients on the model sheet, and coefficients in the model summary worksheet. There is a predefined scale of fill colors ranging from dark blue (very significantly positive) to dark red (very significantly negative).

Fonts: toggle font coding on and off. This tool affects the same cells as the coloring tool, but changes the color (shade of gray) and boldness of font rather than fill color as an indicator of significance (better for printing).

Notes: toggles red flags for cell comments on and off. When the flags are on, the comment will pop up next to the cell when you hover the mouse over it. There is always a cell comment behind the model name on the model sheet and behind the run time cell on the model summary sheet, containing more facts about the model (e.g., run time and name of user and file at run time, for audit trail purposes). The coefficient cells on the model summary worksheet also have comments with more details of each coefficient (t-stats, standardized values, etc.). On the model sheet, if the “teaching notes in cell comments” box was checked at run time, there are teaching notes behind the titles in the regression statistics table and coefficient table (R-squared and all that). *Customization of the teaching notes to fit an instructor’s own materials is something that could be done.*

View: Hitting this key while positioned in a cell that contains a comment (whether or not its flag is visible) will cause the comment to be displayed at the top center of the Excel window. This is the best way to view comments when the Excel window is not very wide, because otherwise the comments may extend off the right of the screen. Also, this is the only way to view comments on a touchscreen or tablet computer when not using a mouse, because hovering over a cell is not possible in that case.

UTILITIES



Select data: if you hit this button while positioned anywhere on the data worksheet, it will find and select the range containing the data.

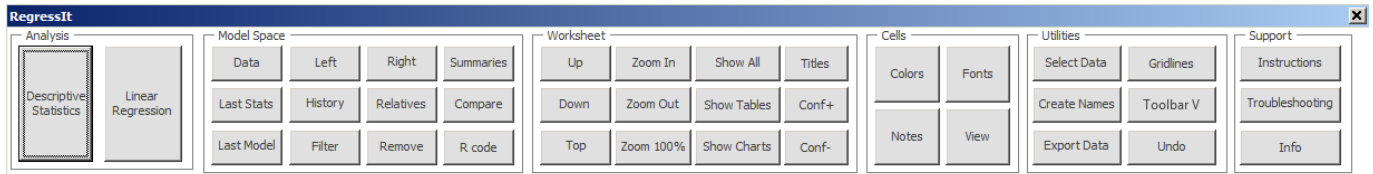
Create names: executes the create-name-from-selection command to assign variable names (= range names). Hit this immediately after hitting the “Select data” button when setting up a new file.

Export data: exports the contents of the data worksheet to a CSV file and writes R code to the clipboard for loading it into RStudio. There is an option to customize the file name and location when doing this. The default file name is the data worksheet name and default file location is the current file location. After hitting this

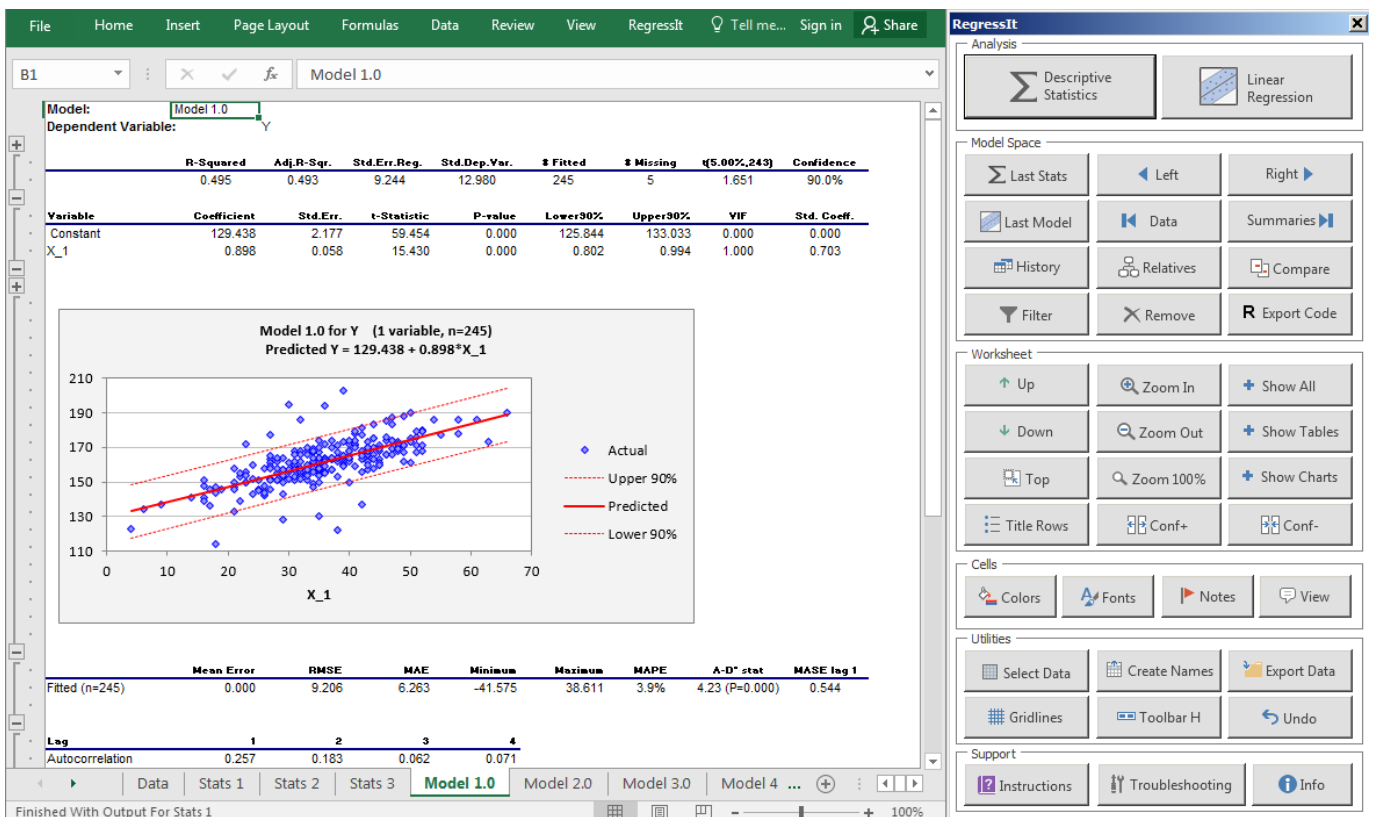
button, go to the command prompt in RStudio and hit <Paste><Enter> to run the code, which loads the data into a frame with the given name.

Gridlines: toggles gridlines and column headings on and off. This can be used for manually re-sizing or hiding rows and columns. (Normally this is not needed: the output is designed for easy reading by default.)

Toolbar H: launches the horizontal floating toolbar interface (below), which has the same button layout as the ribbon but can be positioned anywhere on the computer screen (even outside the Excel window), something else that can be useful for presentations. It is better than the ribbon for use on a touchscreen or tablet computer. You can put it at the bottom of the screen so that your hand won't block your view of the spreadsheet. The buttons are also a bit bigger than those on the ribbon, better for clumsy fingers.



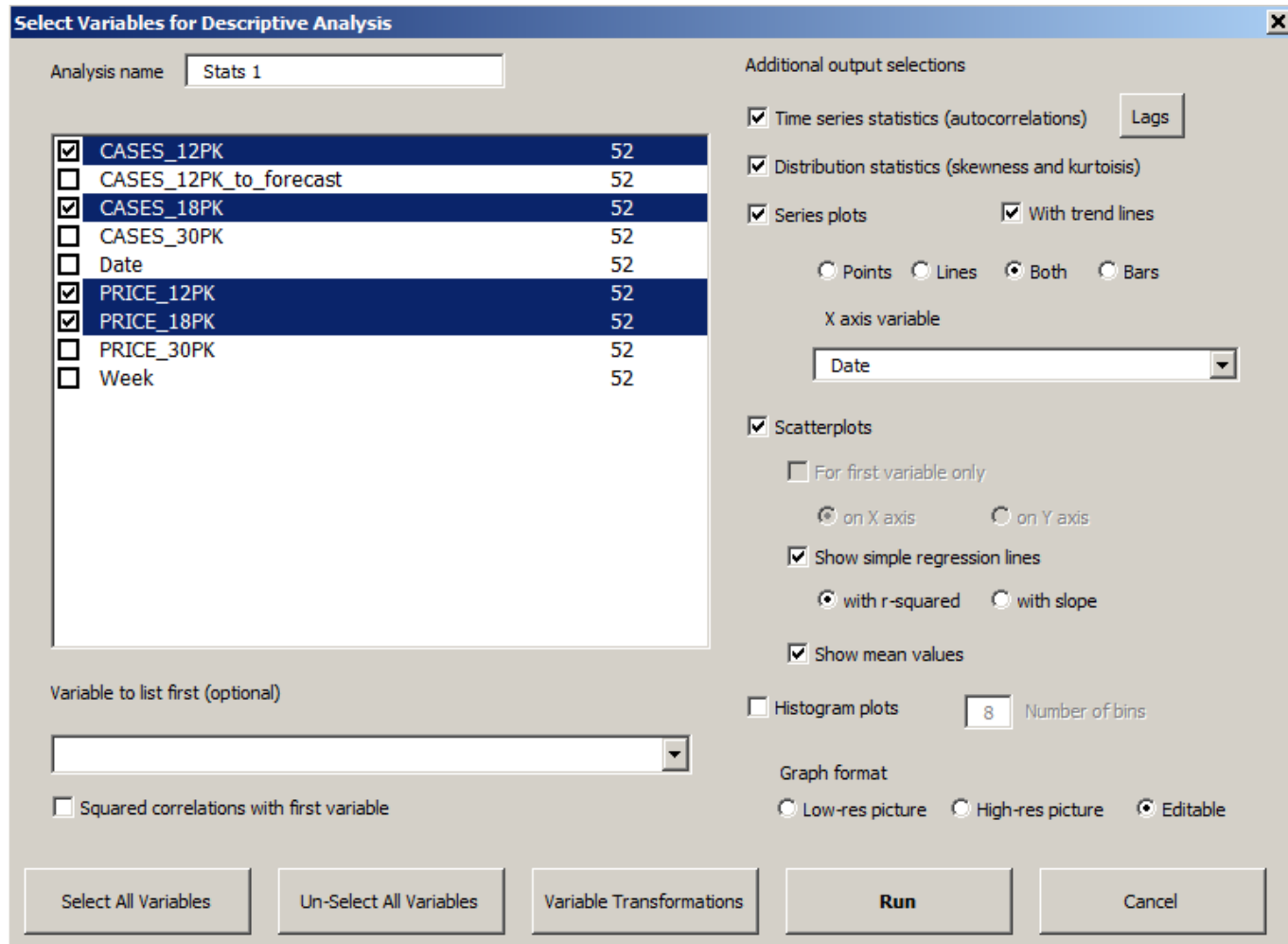
Toolbar V: launches the vertical floating toolbar interface, which can be positioned to the left or right of the Excel window as shown below. This is generally the best interface to use on a touchscreen or tablet computer because it leaves more vertical space for output while keeping your hand off to one side. It has even bigger buttons. This toolbar can also be activated by hitting Ctrl-G.



SUPPORT: These buttons at the right of the ribbon display message boxes with general comments on how to use the program, as well as program version information. You should check the website (regressit.com) periodically for updates, particularly if you have encountered technical issues.

PROCEDURE MENUS AND SAMPLE OUTPUT

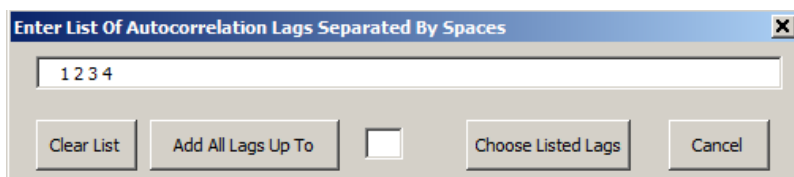
DESCRIPTIVE ANALYSIS PROCEDURE: This procedure always produces a descriptive stats table and correlation matrix. There are also options for autocorrelations, series plots, scatterplots, and histogram plots. *See pages 10 to 13 of this handout for screen shots of RegressIt output worksheets.* There is an option to choose the variable to be listed first, i.e. topmost in the stats table and correlation matrix and first in the ordering of charts that are produced. This is normally used when one of the variables is pre-identified as the dependent variable to be used in subsequent regression analysis. If the “squared correlations with first variable” option is chosen, the correlation matrix is replaced by a table of correlations and squared correlations versus the first variable, as a prelude to regression. Regression variables can be pre-selected from it by using the “Remove” tool on the ribbon to deselect those having only tiny correlations with the prospective dependent variable.



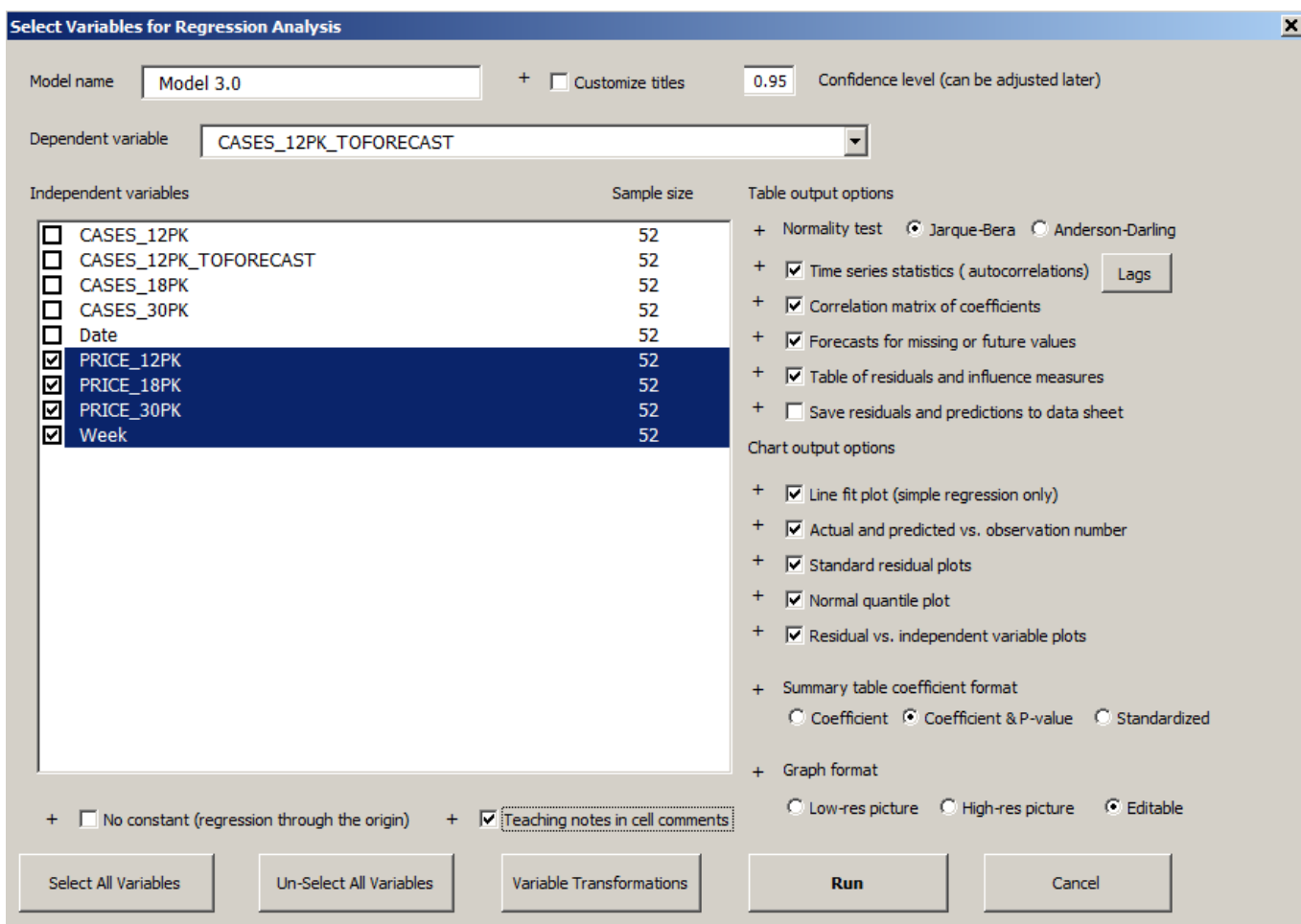
There are 3 graph format options. The choice of which one to use generally depends on sample size and presentation quality that is needed. It is possible to have a very large number of separate charts on the output worksheet (hundreds or even thousands—no kidding), although drawing speed and file size are limiting factors, particularly with large sample sizes. Editable graphs require much less storage space than pictures for small sample sizes, vice versa for big ones. However, they are continuously re-drawn when the worksheet is scrolled up and down, which can be problematic for large data sets, particularly on a Mac. The Up/Down tools on the RegressIt toolbar help to mitigate this issue, because they move up or down by whole chart, keeping the topmost one in alignment, without the need to fine-tune their vertical position using the scroll bar. Storage space for hi-res pictures also depends on sample size. They are fixed format, but they are not bitmaps: they contain the data, although you

can't touch it. Bottom line: for large sample sizes (many thousands) it is generally best to go with low-res picture format. It looks fine unless the graph is enlarged significantly (e.g. to fill a whole Powerpoint slide). Hi-res picture and editable charts remain razor-sharp under enlargement and are more efficient for smaller sample sizes. Another consideration in chart type selection is that on low-res and high-res charts, text and graphic elements scale up proportionally under enlargement, while on editable charts they don't.

LAGS FOR AUTOCORRELATIONS: This option applies to both descriptive statistics (autocorrelations of variables) and regression (autocorrelations of residuals). Autocorrelations can be consecutive or not. For example, for monthly time series data, you might choose only lag 1 and lag 12.



LINEAR REGRESSION PROCEDURE:



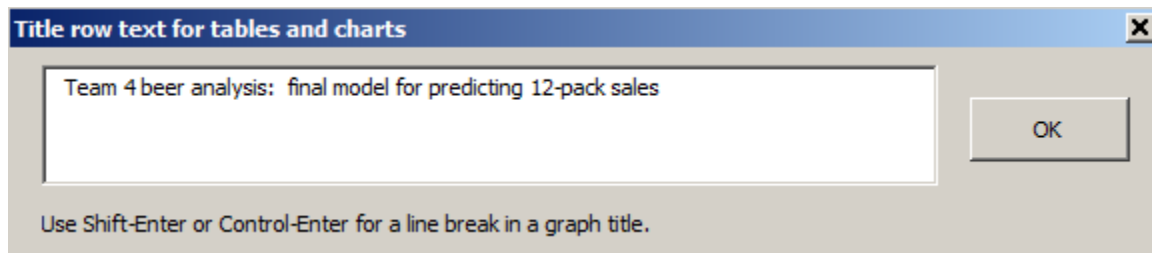
If you click on a plus sign, it will open a message box with a description of the model option to the right.

Standard text output (summary stats table, coefficient table, ANOVA table) is always produced, and there are many options for additional tables and charts. Forecasts can be generated for rows where the dependent variable is missing and all independent variables are present. Teaching notes that explain the key output statistics can be optionally embedded in the output worksheet as cell comments.

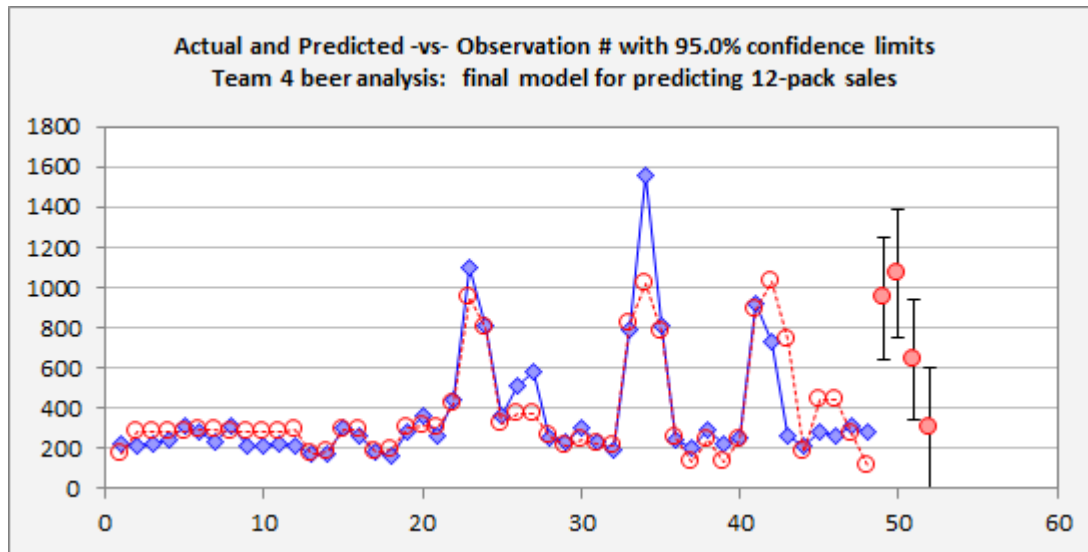
Model names for use in table and chart titles can be customized. If the name of the previous model was “12-pack model 3.2”, the default for the next one will be “12-pack model 3.3,” but it can be changed arbitrarily before running the model. The model name entered in this field will also become the output worksheet name.

The model information row in table and chart titles can also be fully customized if needed, independently of the worksheet name. The dialog box below appears next if the “Customize titles” box is checked. This feature is useful when preparing a final set of charts and tables for presentation or inclusion in a report. Students can be also asked to include identifying information about themselves or their team.

For example, when this custom title is entered...



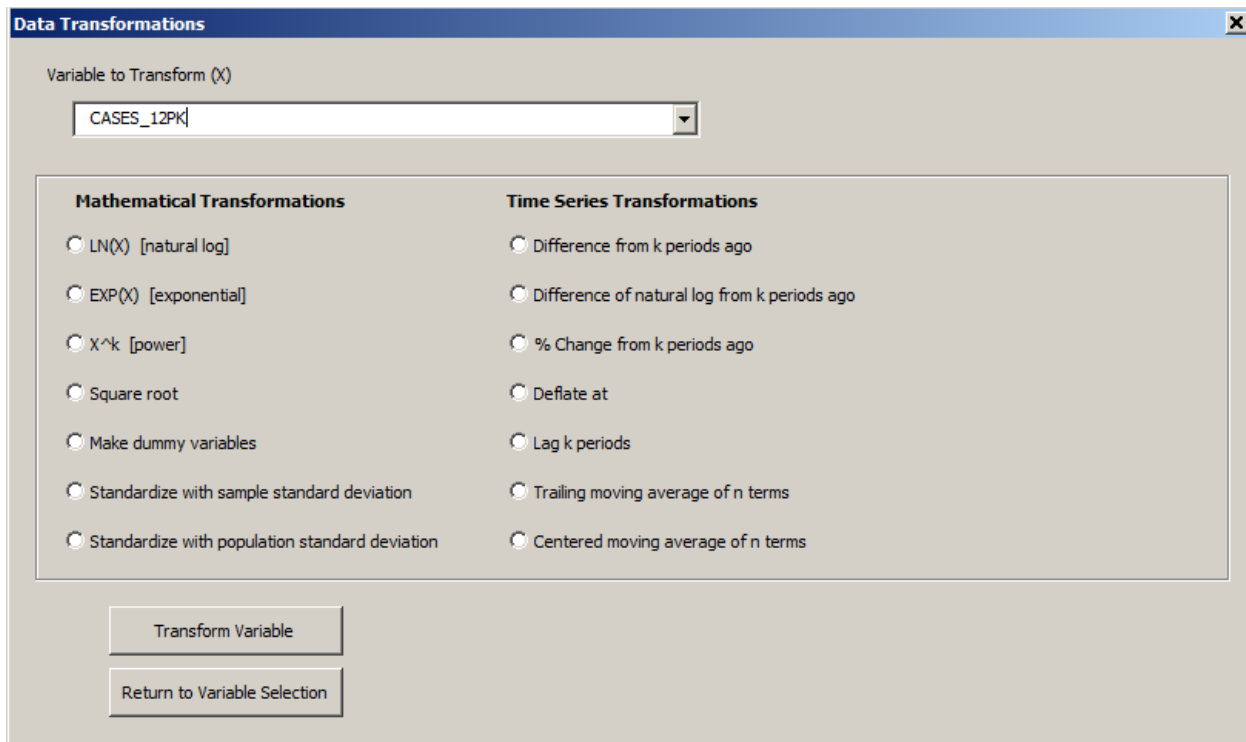
...the charts come out looking like this:



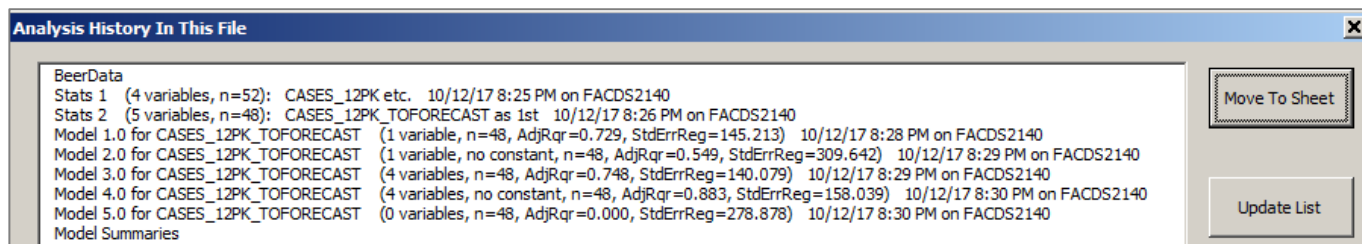
However, it is generally best to use the default chart titles for routine work, because they link back to the names of the worksheets in the file from which the charts were copied. It’s usually enough to just be thoughtful in assigning model names, and RegressIt’s default naming convention helps with this.

Point sizes on graphs are automatically scaled to suit the size of the data set.

VARIABLE TRANSFORMATION TOOL: This can be launched from either the descriptive stats or regression dialog box. It includes time transformations if the “Time Series Data” box is checked when it is launched. Transformed variables are automatically assigned descriptive names, e.g., the natural log of X is automatically named X.LN. X lagged by 1 period is automatically named X.LAG1, and so on. The transformation name is tacked on at the end rather than the beginning so that the original and transformed variables appear adjacent to each other in alphabetically sorted variable lists. Also, the new column of data for the transformed variable is inserted next to the column of the parent variable on the data worksheet.

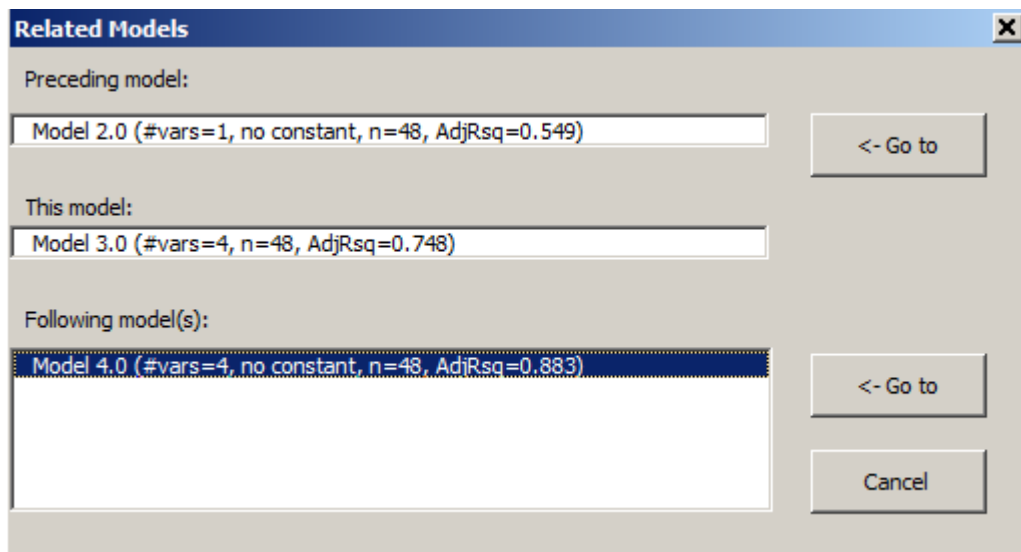


HISTORY TOOL: Provides a list of all worksheets in the file and allows you to jump directly to any of them. This box stays open on the screen until “Cancel” is hit, so you can visit a sequence of worksheets without having to re-launch it. On a PC it can be kept on top of the worksheet while you work. On a Mac you should position it outside the Excel window. The last sheet visited is remembered for use by the “Last Stats” and “Last Models” buttons, so that you can easily toggle back and forth between any two sheets.



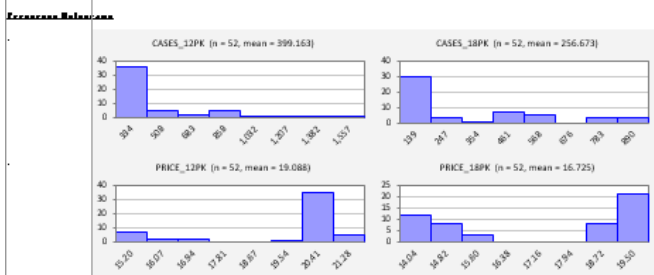
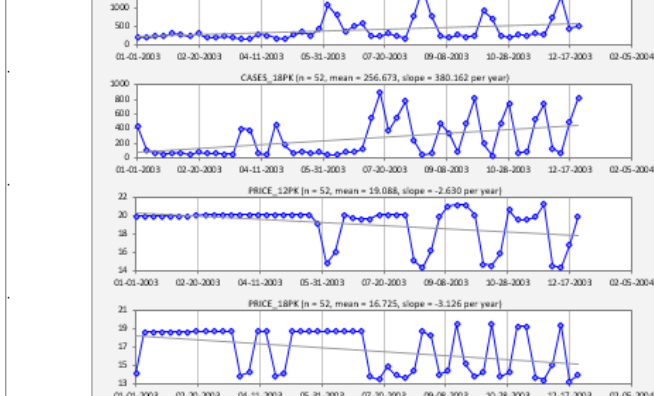
Audit trail information for regression models, such as run time and computer name, is stored here and in several other places in the workbook, both in text strings in cells and in cell comments. This is useful for managing your own individual work or teamwork by keeping track of the temporal sequence as well as which of your many computers you may have used to store related files. It also makes it hard for a student to hand in a file created by someone else for a homework assignment or a laboratory exercise. More detailed information about all regression models in the analysis sequence (error stats, coefficients, etc.) is stored on the Model Summaries worksheet in publication format, i.e. side-by-side columns that allow easy comparisons of model parameters. Information about all models previously created remains there even if their model worksheets are later deleted, whereas the history list is dynamically created to show only worksheets currently in the file.

RELATIVES TOOL: Allows you to jump to parents or children of the regression model whose sheet you are currently looking at, in order to easily retrace your steps or back up to an earlier model in order to go down another search path. (How does it work? Each regression model worksheet contains information about its parents and children, hidden in white font. Look in cells Q2, R2, etc.)

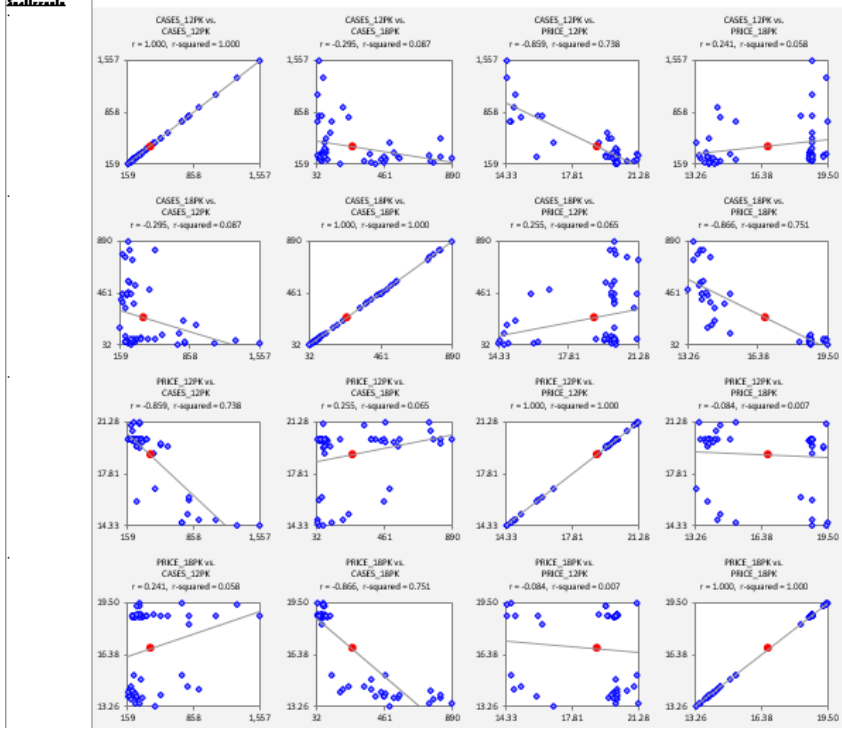


Varsta	EPK	Mean	Median	Std.Dev.	Max	Min	Skewness	Kurtosis
CASES_12PK	52	399.163	272.588	182.555	41.937	193.888	1.557	2.157
CASES_18PK	52	256.673	18.888	258.872	35.819	32.888	1.857	4.168
PRICE_12PK	52	19.088	19.388	2.888	8.238	14.338	-1.448	0.588
PRICE_18PK	52	16.725	16.858	2.411	8.334	13.218	-1.265	-1.381

Varsta	Lag 1	Lag 2	Lag 3	Lag 4
CASES_12PK	0.478	-0.335	-0.274	-0.387
CASES_18PK	0.348	-0.168	0.225	0.478
PRICE_12PK	-0.244	-0.487	-0.282	-0.254
PRICE_18PK	0.388	-0.141	0.175	0.438



Varsta	CASES_12PK	CASES_18PK	PRICE_12PK	PRICE_18PK
CASES_12PK	1.000			
CASES_18PK	-0.295	1.000		
PRICE_12PK	-0.859	0.255	1.000	
PRICE_18PK	0.241	-0.866	-0.084	1.000



Sample output worksheet from the **DESCRIPTIVE STATISTICS** procedure, showing descriptive stats, autocorrelations, series plots, histogram plots, correlations, and scatterplots.

Here, the series plots include optional trend lines and their slopes as well as an optional custom time axis scale.

Scatterplots include optional regression lines and center-of-mass points.

Color-coding has been turned on for the autocorrelations and correlations. This can be toggled on and off at any time if RegressIt is active.

The editable option was used, so that all 24 charts are native Excel charts that are subject to editing, although their descriptive titles and default formatting generally make this unnecessary.

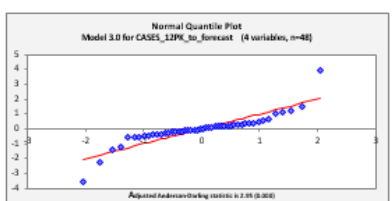
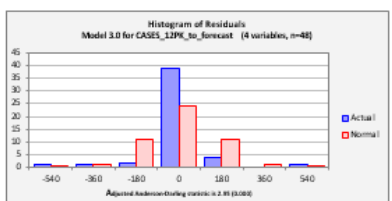
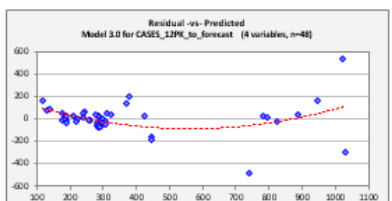
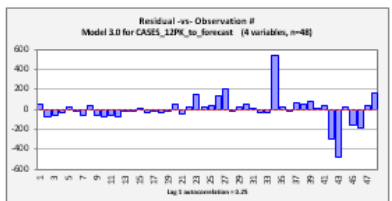
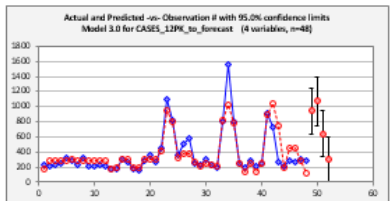
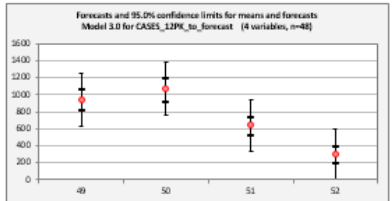
Model	R-Squared	Adj. R-Sq.	Err. Deg.	Std. Dev.	DF	Mean	2-Std. Err.	Conf. Int.
Model 3.0	0.703	0.748	140.873	278.878	48	4	2.847	35.8X

Variable	Coefficient	Std. Err.	t-Stat	P-Value	Lower 95%	Upper 95%	WIF	Std. Coeff
Constant	2,108	518.762	4.065	0.001	955.087	3,262	0.000	0.000
PRICE_12PK	-192.243	51.815	-3.710	0.000	-411.140	-73.346	1.231	0.330
PRICE_9PK	133.275	51.555	2.585	0.010	28.149	238.401	1.255	0.330
PRICE_3PK	3.242	27.323	0.118	0.900	-51.878	58.382	0.003	0.003
Week	2.297	1.558	1.474	0.146	-0.846	5.253	1.200	0.515

Mean Error	RMSE	MAE	MAPE	MAPE @ 0.5 Std.
-1.855E-02	192.181	77.619	476.787	552.658

Statistic	1	2	3	4
Autocorrelation	0.245	0.328	0.219	0.183
Std. Error of Forecast	1.582	0.187	1.431	0.725

Obs	Forecast	Std. Err.	Lower 95%	Upper 95%	Actual	Mean	Std. Dev.	Std. Err.	PRICE_1	PRICE_1	PRICE_1	Week
49	345.271	152.657	137.493	553.049	18.673	822.388	1.868	14.56	15.19	15.19	49	
50	1,878	155.355	755.288	1,000	18.154	315.458	1.288	14.55	15.49	15.15	50	
51	141.651	149.486	348.153	345.123	51.192	536.487	246.316	16.81	19.24	15.15	51	
52	382.191	149.884	1.473	662.788	51.823	193.224	485.848	19.86	19.32	15.15	52	



Sample output worksheet from the **LINEAR REGRESSION** procedure, including forecasts that were automatically generated for 4 missing values of the dependent variable at the end of the series.

Color coding has been toggled on for the t-stats, standardized coefficients, and residual autocorrelations. It would also be applied to the correlation matrix of coefficients if that had been included.

The residual table, if it had been included, would appear at the bottom and would include actual and predicted values, residuals, standardized residuals and their absolute values, leverage, and Cook's D. The filter tool on the ribbon can be used to easily sort the residual table on any of these statistics.

Model:	Model 3.0							
Dependent Variable:	CASES_12PK_to_forecast							
	R-Squared	Adj.R-Sqr.	Std.Err.Reg.	Std. Dev.	# Fitted	# Missing	(2.50%,43)	Conf. level
	0.769	0.748	140.079	278.878	48	4	2.017	95.0%

Variable	Coefficient	Std.Err.	t-stat
Constant	2,130	590.762	3.606
PRICE_12PK	-117.219	11.863	-9.881
PRICE_18PK	23.775	9.656	2.462
PRICE_30PK	3.242	27.923	0.116
Week	2.297	1.658	1.386

The t-statistic of an independent variable is its estimated coefficient divided by the coefficient's own standard error, i.e., its number of standard errors away from zero. The t-stat's value (which has the same sign as the coefficient) is an indicator of whether that variable has been found to have a measurably non-zero effect in explaining or predicting variations in the dependent variable, in the context of other variables included in the same model.

The t-statistic associated with any one variable is model-dependent. Its value may change, sometimes significantly, if other related variables are added or removed. Also, the t-stat of a variable whose true coefficient is non-zero tends to grow larger in magnitude as the sample size increases, because standard errors of coefficients grow smaller as the sample size increases.

A commonly used rule of thumb is that a variable's contribution to a model is not statistically significant if its t-stat is less than 2 in absolute value, i.e., if its estimated value is less than 2 standard errors away from zero. This is not a hard-and-fast rule, but as a practical matter the removal of a variable whose t-stat is less than 2 in magnitude will probably not increase the standard error of the regression by very much.

Whether a variable should be removed also depends on other considerations, such as the objectives of the analysis and whether there are other supporting arguments for its use in the presence of the other variables (if any). If the model specification is a priori unknown and the data have been collected in an ad hoc fashion, simpler is generally better.

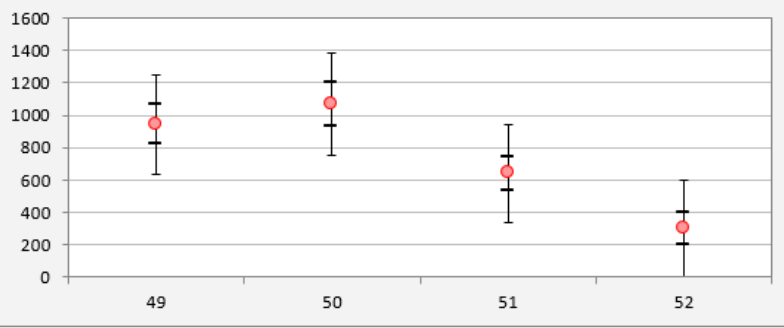
	Mean Error	RMSE	MAE
Fitted (n=48)	-1.066E-12	132.583	77.613

Lag	1	2	3
Autocorrelation	0.245	0.028	0.213
StdErrorsFromZero	1.682	0.187	1.431

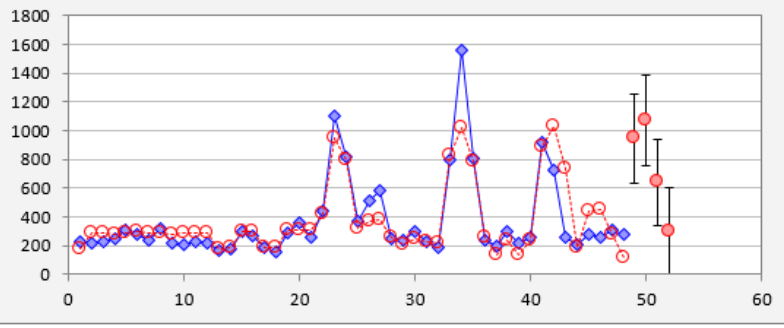
Obs#	Forecast	StErrFct	Lower95%F
49	945.271	152.657	637.409
50	1,070	155.959	755.208
51	641.661	149.486	340.193
52	302.131	149.084	1.473

PRICE_18PK	PRICE_30I	Week
15.13	15.19	49
19.43	15.19	50
13.26	15.19	51
13.92	15.19	52

Forecasts and 95.0% confidence limits for Model 3.0 for CASES_12PK_to_forecast



Actual and Predicted -vs- Observation # with 95.0% confidence limits Model 3.0 for CASES_12PK_to_forecast (4 variables, n=48)



TEACHING NOTES can optionally be included as cell comments behind the key statistics in the output.

They can be viewed by hovering the mouse over the cell (when its flag is turned on) or by clicking on the cell and then clicking the **View** button on the ribbon. The latter method will cause them to be displayed at the top center of the Excel window, rather than off to the side, better for small screens.

MODEL SUMMARIES WORKSHEET

Summary of Regression Model Results					
Dependent Var: CASES_12PK_TOFORECAST	Model 1.0	Model 2.0	Model 3.0	Model 3.0 (#vars=4, n=48, AdjRsq=0.748)	
Run Time	10/12/17 8:28 PM	10/12/17 8:29 PM	10/12/17 8:29 PM	Dependent variable = CASES_12PK_TOFORECAST	
# Fitted	48	48	48	Run time = 10/12/2017 8:29:43 PM	
Mean	369.688	369.688	369.688	File name = BeerDataToForecast12PK2a.xlsx	
Standard Deviation	278.878	461.326	278.878	Computer name = FACDS2140	
Number Of Variables	1	1	4	Program file name = RegressItPC_1012	
Standard Error of Regression	145.213	309.642	140.079	Version number = 2017.10.12	
R-squared	0.735	0.559	0.769	Execution time = 00h:00m:01s	
Adjusted R-squared	0.729	0.549	0.748	0.883	0.000
Mean Absolute Error	90.035	196.336	77.613	101.786	190.109
Mean Absolute Percentage Error	27.4%	50.8%	22.4%	31.0%	53.0%
Maximum VIF			1.290	0.000	
Normality Test	***	***	***	**	***
Mean Absolute Scaled Error	0.542 (lag 1)	1.183 (lag 1)	0.468 (lag 1)	0.613 (lag 1)	1.145 (lag 1)
Residual Autocorrelation	0.24 (lag 1)	0.53 (lag 1)	0.25 (lag 1)	0.32 (lag 1)	0.51 (lag 1)
Coefficients:	Model 1.0	Model 2.0	Model 3.0	Model 4.0	Model 5.0
Constant	2,785 (0.000)		2,130 (0.001)		369,688 (0.000)
PRICE_12PK	-125.087 (0.000)	17.774 (0.000)	-117.219 (0.000)	-88.147 (0.000)	
PRICE_18PK			23.775 (0.018)	42.924 (0.000)	
PRICE_30PK			3.242 (0.908)	88.093 (0.000)	
Week			2.297 (0.173)	3.517 (0.061)	

Example of a model summaries worksheet, which shows side-by-side model stats in journal-article format. Optional color coding of coefficients has been toggled on here too.

This sheet (which is located to the right of all analysis worksheets) is updated with every new regression model, and it provides an audit trail for all models that have been fitted in this file's lifetime, even those whose output worksheets may have been subsequently deleted. It is very easy to spin off large numbers of models with memory-intensive chart output, so the user will often want delete inferior ones in the course of a long analysis. However, it is good to maintain the entire audit trail, particularly to verify the efforts of students.

The cell comment behind the model name includes much more audit trail information, and cell comments behind the coefficient estimates provide more detail there.

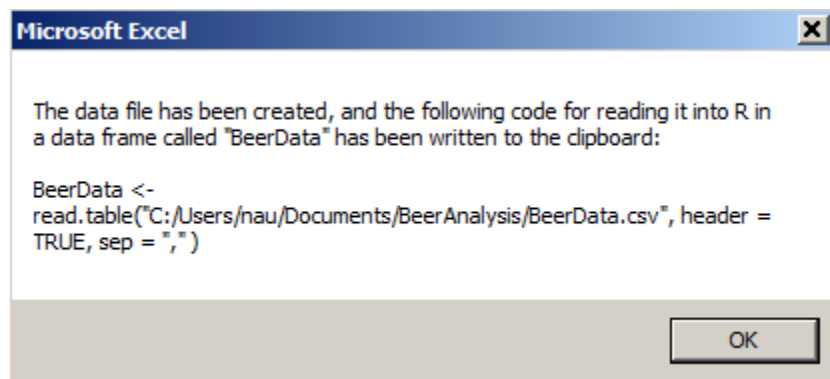
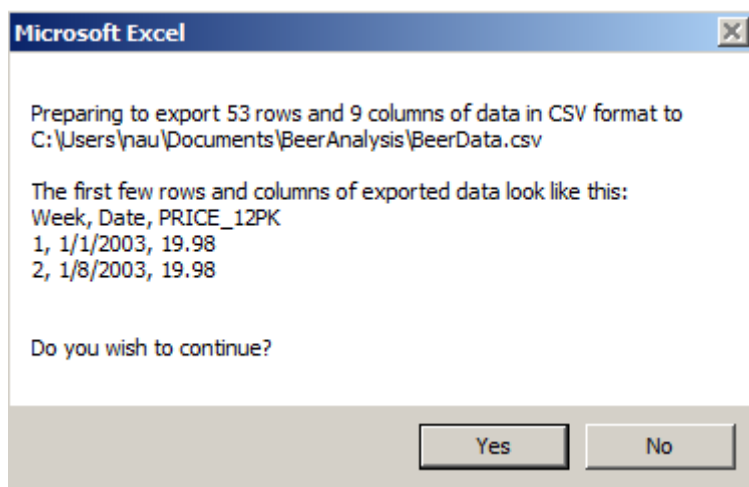
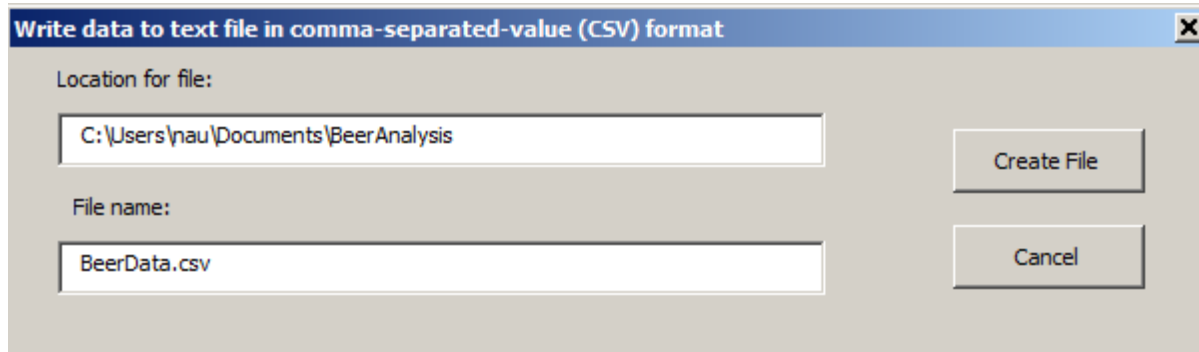
The user has the option at run time to store the coefficients in the table in coefficient plus P-value text format (which was done here), or else in numeric form, either standardized or unstandardized. Regardless of which option was chosen for display, the cell comments provide all the other information about each coefficient estimate.

Another feature is that any regression model whose summary output is stored here can be easily re-run. The model specifications are stored in the cell right above the model name and hidden in white font. If the regression button is hit while positioned on this cell, the model information will be read from it, and that model's worksheet will be moved to if it is present. If it is not present, the variable specifications will become the default selections for a new model, so the original one can be regenerated with a couple of clicks. The entire history of models in the workbook can be re-created from this single sheet in this way.

If models had also been fitted to some other dependent variable, their results would be stored in a different block of rows farther down. Thus, only models fitted to the same dependent variable are displayed side-by-side.

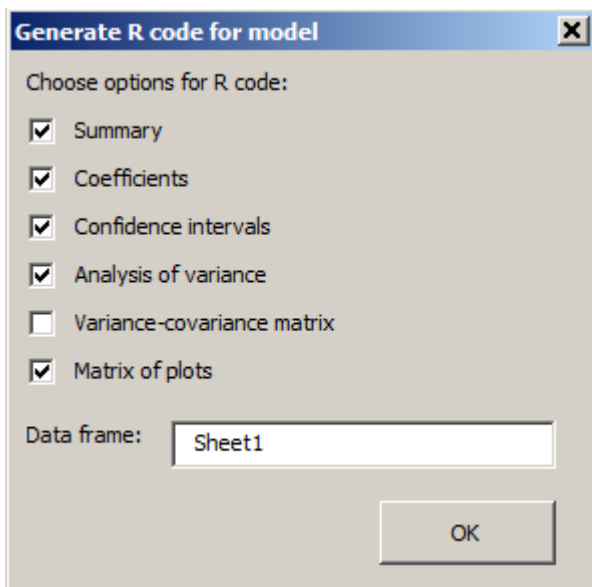
EXPORTING DATA AND MODELS TO R

The **Export Data tool** on the ribbon exports the entire contents of the data worksheet to a CSV file (a text file in comma-delimited format). The following screens appear next. A default file location (current folder) and file name (current data worksheet name) are offered, but they can be edited. The user is offered the opportunity to review what the data will look like in the file before it is created. At the same time, R code for reading the data from the file is copied to the clipboard. After hitting “Yes”, the user just needs to go to the command prompt in RStudio and hit <Paste><Enter> in order to run the code and load the data into a new frame with same name as the file name.

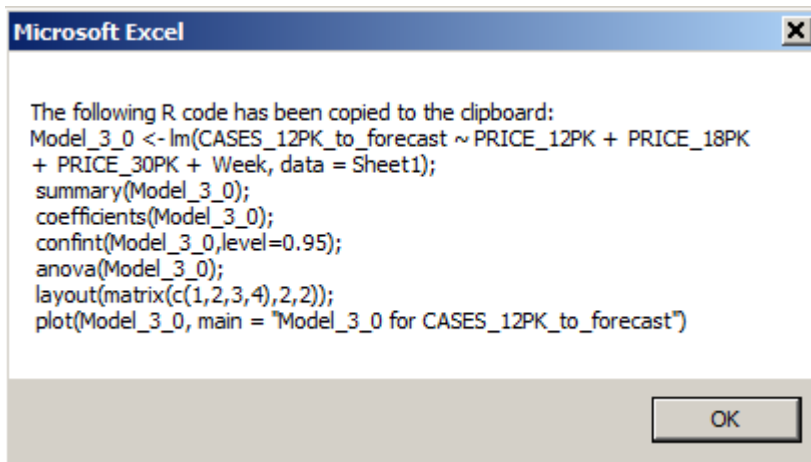


R CODE TOOL:

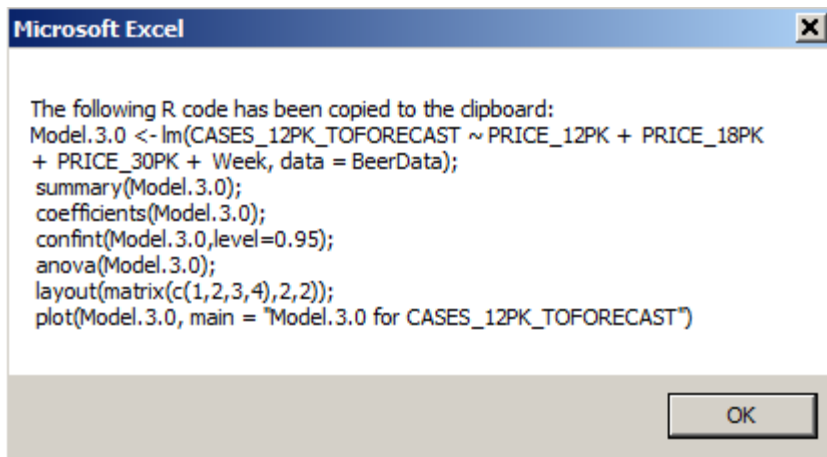
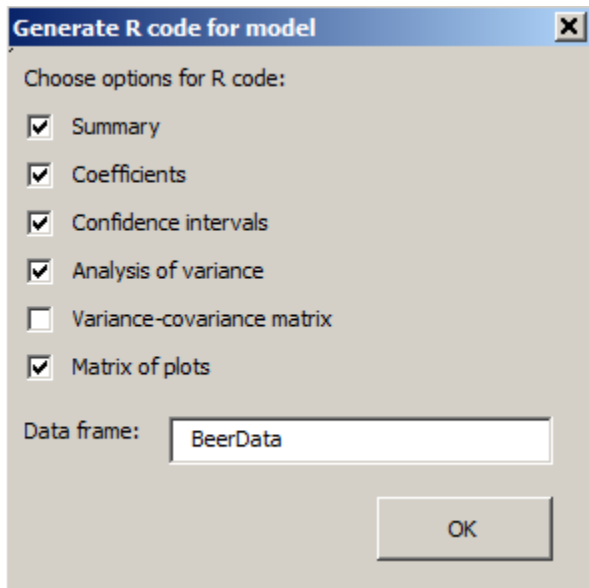
A number of output options can be chosen exporting R code for the model on whose worksheet you are currently positioned. These are the default options that are selected when you first hit the button:



The following screen appears next, allowing you to review the code that will be copied to the clipboard:



After hitting "Yes", you just need to go to the command prompt in RStudio and hit <Paste><Enter> to run the code. The default data frame name is the RegressIt data sheet name, although it can be changed. The RegressIt model name is included in chart titles produced in R--see the screen shots on the next page. Many more output options could easily be added.



It takes just a few mouse clicks and one hit of Enter to transfer the data from Excel to R, and another few mouse clicks and one hit of Enter to fit a model in Rstudio after fitting the model in RegressIt. Here's what you will see. No worries about typing or syntax errors!

```

Console ~/

R version 3.4.0 (2017-04-21) -- "You Stupid Darkness"
Copyright (C) 2017 The R Foundation for Statistical Computing
Platform: x86_64-w64-mingw32/x64 (64-bit)

R is free software and comes with ABSOLUTELY NO WARRANTY.
You are welcome to redistribute it under certain conditions.
Type 'license()' or 'licence()' for distribution details.

R is a collaborative project with many contributors.
Type 'contributors()' for more information and
'citation()' on how to cite R or R packages in publications.

Type 'demo()' for some demos, 'help()' for on-line help, or
'help.start()' for an HTML browser interface to help.
Type 'q()' to quit R.

> BeerData <- read.table("C:/Users/nau/Documents/BeerAnalysis/BeerData.csv", header =
TRUE, sep = ",")
> Model.3.0 <- lm(CASES_12PK_TOFORECAST ~ PRICE_12PK + PRICE_18PK + PRICE_30PK +
week, data = BeerData); summary(Model.3.0); coefficients(Model.3.0); confint(Model
.3.0,level=0.95); anova(Model.3.0); layout(matrix(c(1,2,3,4),2,2)); plot(Model.3.0,
main = "Model.3.0 for CASES_12PK_TOFORECAST")

```

Output produced in RStudio:

Notice that the charts have model-specific titles by default, matching those used in RegressIt.

