NICE TO KNOW

Excel Data Table Utility

Lewis Coggins

recorded: February, 2011

In this recording, I'm going to demo an Excel tool called data tables. data tables have a lot of utility in conducting sensitivity analyses, asking what if questions: What if a parameter varies across some range? How does that impact a critical output parameter or an output of your model? The setup I'm going to use is basically a simple version of the model that I built in response to the week-three challenge about wabbits so this is the scenario where we have a deterministic model that is predicting the abundance of the wabbits, considering deaths from cats, as well as deaths from native reptiles. You all most likely built this model, and your model may look different than mine, but I'll just quickly walk you through my model.

Using Terry's coloring convention, I have colored this block of cells green because these are my input cells, and I have colored this block blue because the ending abundance is my key output cell. For my input variables, I have the initial abundance of rabbits, and I've named my cells. I have the Clutch Size, which is the number of young wabbits per each female wabbit. I have the number of cats that are present to inflict mortality on wabbits before they give birth and the rate at which cats kill wabbits; in this case, it's five, and the rate at which the native reptiles kill wabbits. And, in this case, that is the proportion of the total wabbits available to be killed by the reptiles. I also have this kind of intermediate calculation here, which is the deaths of wabbits from cats, before they give birth, and that is simply the number of cats times the cat death rate, and then I have the dynamics of my model represented here.

So I have the number of wabbits at the beginning of year one as the initial abundance. Then I have the deaths attributable to cats. And notice I have this “if” statement here, and I basically say, "Well, the number of wabbits that are killed by cats is equal to the cat-related deaths--this cell--if, and only if, the wabbit population exceeds that value. So if there are fewer than 200 wabbits, then the number of deaths from cats will just be the number of wabbits, otherwise it'll be 200, is all that if statement says. I make the assumption that the number of deaths from reptiles is dependent on this reptile death rate--this cell--the difference between the starting number of wabbits and then the number that were killed by cats. So basically I'm making the assumption that all the deaths from cats occurred, and then whatever is left a fraction of those, 2 percent, are killed by reptiles. So that's a simplifying assumption that I basically put in here. It's a sequence so first the cats kill wabbits. Then whatever is left the reptiles take a proportion of.

The number of females that give birth is the starting number of wabbits minus the deaths from cats, minus the deaths from reptiles, and then it's divided by 2 because these are wabbits of both sexes, and I just make an assumption that I have a sex ratio of 50-50 so I divide it by 2. The number of females giving birth is equal to that term only if deaths from cats and the deaths from reptiles don't exceed the total number of wabbits. If it does, then, of course, there won't be any females so I stick a 0 in here.

And then the other dynamics is that the wabbits in year two are equal to the wabbits in the previous year times the clutch size. Okay? So then I basically just copy those relationships down, and ultimately I have this Ending Abundance, which is the number of wabbits in the fifth year. And I've plotted that time series here, as you can see, again following up on advice from Tony. That's not necessarily the most informative way to produce results of this model, and, of course, this doesn't have any sensitivity analysis, so that's what I would like to do. And, I would like to use the data table to conduct my sensitivity analysis.
We would like to be able to have an idea of how ending abundance changes, for instance, as either our cat-control measures as we vary cats across 40 to smaller numbers or across our grickle-grass management, where we manage grickle grass and then impact the clutch size. What would an analysis like that look like? Well, like for instance, say we looked at various values of cats, and we wanted to know, well, what happens if we remove all the cats? What happens if we remove all but five? What happens if we don't remove any, if we leave it at 40? We would then like to be able to look at how the ending abundance changes. Right? That would be a sensitivity analysis to look at how this value changes as we put different values in here for the number of cats that impact the wabbit population. So you know we can do this by hand.

We can go in here, and we could put a 0. Okay. Now the ending abundance is 834. We can go in here and put a 5. Now the ending abundance is 744. We can put a 10. Now the ending abundance is 655. Or we can skip all the way to 40. The ending abundance is 119. That's fine, we can do that by hand. But what a data table will do is, it will repeatedly insert the values of this control parameter that you want to vary and then look at how the ending abundance changes. To make the data table function work, it requires a particular format. data tables can be used to evaluate changes along one parameter that's called a one-dimensional data table, or it can work across two parameters or two-dimensional data table. Let's look at the one-dimensional data table first. The way it works is, you define the values that you want to repeatedly insert along, for instance, a column. This can be done along a row, but I'm going to show it to you along a column. Then to the right, one column, and up one row from the very first value that you want to examine, you establish a link to the cell that you want in your spreadsheet that you want to keep track of. In this case, it'll be ending abundance. It'll be this cell right here. I'll just go ahead and type in the name. It's going to report ending abundance. Okay? I'm just going to delete these out of here for the moment.

So this is the format you need. You need a column that has your input values and then to the right, one column, and up one row from the first one you need to insert a link to the cell that you want to keep track of. Then what you do is, you invoke the data table function, but first you have to highlight this range of cells. Well, maybe one way to do it is start on the cell that you are keeping track of--in this case the ending abundance--and then highlight down to the last value that you want to insert. So highlight that range. Then you go to the Data Ribbon, and you go to the What If analysis, and you choose the function data table. And this dialogue box comes up. It's not really very informative, but I will tell you that it is asking you to specify either two other cells if it's a two-dimensional data table or one if it's a one-dimensional data table.

And what we want to specify is this column input cell because the values that we want to repeatedly substitute in occur in a column. These values occurring in this column, we want to be repeatedly substituted into this cell right here, the number of cats. So what it'll do is, it'll take zero cats inserted into that cell, and it'll report the ending abundance here with zero cats. And it'll take 5, put it into that cell, take this Ending Abundance report it in that cell and so forth. So once you've defined that, you say OK, and voila, it has done it for us. Let's just get rid of all these decimal places because we're not really interested in them. Sure enough. So you can see that, if there are lots of cats, 40, to kill wabbits, the ending abundance is going to be smaller than if there are no cats to kill wabbits.

So I'm going to plot this. If I plot out the results of this data table, this is what I see. I see that when we have zero cats, the Ending Abundance is large, and we can confirm that. Go here, type 0: 834. And when we have a lot of cats, the ending abundance is smaller, and we can confirm that: 119. Again, we just did a sensitivity analysis across that input value or actually this management policy. We chose different levels of management, and we were able to see what impact that had on the ending abundance.

Now we can do the same thing with a two-dimensional data table. Let's claim that we want to look at, again, the number of cats, and we also want to look at the clutch size. We want to evaluate different grickle-grass control policies. And we want to look at clutch sizes across a range from, let's say, 1.8 in increments of .05. So 1.85, and let's go ahead and make sure we can see, we can actually see what these values are. We want to go all the way up to 2.1 so we're going to look at these different clutch sizes. I'm going to go ahead and merge that and color this, and then I'm going to actually move cats down to here and merge this and center it and color this slightly different color. Yeah, how about that.

Okay. So we want to look at what happens when we vary our cat-control policy from 0 to 40 and our grickle-grass control policy such that we realize clutch sizes from 1.8 to 2.1. I'm starting to build the two-dimensional data table setup. It's pretty similar in that you set up the two different series across which you want to vary values, and then now, instead of the cell that you want to keep track of being here is where we put it before, it has to be up one from the column that contains your series of values and to the left one column from the range of values that's reported in the row. So here I'm going to say: = ending abundance, and I'm just going to get rid of those decimals and then to use the data table we start again on the cell that we are keeping track of in the data table, and we highlight across all the columns of our row variable and down all the rows of our column variable. So we highlight that range.

Then we go to Data, choose What If Analysis data table. Now we're going to fill out both of these. The values that occur in a row we want to be substituted in, in this case, for the clutch size so that would be cell B2. And the values that occur in a column we want to be substituted into the number of cats. So that's the basic setup. We will repeatedly put in .18 for clutch size, and the repeatedly put in all the values for cats, to fill out this column. Then we will put in 1.85 for clutch size and all the different values for cats. That'll fill out this column, and so forth. So we say OK, and there we have it.

This makes sense because the highest ending abundance is when we have the largest clutch size and the fewest number of cats. The lowest ending abundance is when we have the smallest clutch size and the largest number of cats. We would like to be able to make a plot of these results. There's a particular way to do this; it's not quite intuitive, but I want to show you how to make this plot in a way that'll save you a lot of time. First, I'm going to show you how to do it wrong. To do it wrong you would like to have a plot that has either a clutch size as different lines, like in this graph, and then cats as occurring as different X values. You kind of want to highlight both the X values and the different series values. So to do that you might think you highlight that Data Series and then go ahead and go Insert a line graph, like so. And here's what you get: you get this series actually, that is it's interpreting this column that you want to be different X values actually as a different Data Series. Of course, you can delete this, and you can kind of monkey around with this and change this to here. Now you got that far, but now what you have to do is, you have to go in here and name each one of these series so that would be the series of a clutch size of 2.1. That's a lot of work.

Here's the trick: Instead of doing that first, grab this Ending Abundance cell, and move it out of the way. Okay? Then highlight the same range like so. Go to Insert a line graph. Now you have what you wanted. You have the values of cats ranging across the different values on the X axis. You have the Ending Abundance on the Y axis, and each one of these lines is labeled appropriately as to what the clutch size is. And then what you can do is, you can put this back in here as you make changes to, for instance, some other value in here, like maybe starting abundance. Then your graph will get updated appropriately. So what I've done is, I've just put in the Y axes and the X axes for cats and Ending Abundance, and I've labeled the legend so that we know that each one of these different lines refers to a different clutch size.

And the last thing I did is, I went in here, and I typed what the assumed rate that wabbits were killed by cats so you recall that this data table is constructed under the assumption that each cat kills five wabbits per year. The reason I did that is because I want to use this setup I have to actually conduct a sensitivity analysis across an even wider range of parameters. So right now we're looking at one management policy as numbers of cats, another management policy as grickle-grass control, right? If we control grickle grass and make artificial rearing burrows we can get up to 2.1 wabbits per female. But that's under the assumption of how many wabbits are killed for each cat.

The problem told us that there could be as few as half a wabbit killed per each cat. So if I change this to 0.5, like so, that's how the ending abundance changes across these two different policies. In this case, cat control has very little impact on the performance or the numbers of wabbits in the end. We could save this. I'm going to copy this, and I'm going to paste it down here as a picture. That's ending abundance, if we assume that there's half a wabbit per cat killed. If we change this to 2.5, now we can see cat control has a bigger impact, and we can copy this and paste it here as a picture. So that's 2-1/2, and then now this one we can just move this back to 5 and change this to 5.

Okay. So now we've done a pretty good analysis of different control options for both cats and grickle-grass control across different assumptions of the numbers of wabbits killed per cat. If we want to be the most conservative, we might think about these results where, if cats are a problem, we should be controlling them. Alternatively, if in the future we get good evidence to suggest that, well, cats really they don't kill too many wabbits, then we might be in this situation. And then the policy of trying to catch cats and have them adopted actually doesn't impact the wabbit population much. There's two different types of data tables: a one-dimensional data table or a two-dimensional data table and how we can use them to do sensitivity analyses.

The last thing I wanted to point out is the Help function for the data table. So if I hit F1, it brings up Excel Help. I type *Data* (space) *Table*, calculate multiple results with a data table. This has a reasonably good Help file for both how to calculate what they call one-variable data tables, what I was calling a one-dimensional data table or two-variable data tables, and they kind of step by step walk you through how it's done. You can either have a data table that's column oriented, which is what I showed you, or you can have one that's row oriented, and sometimes for the architecture of your spreadsheet that might be helpful. And then it goes through how you do a two-variable data table, which we walked through also. You can have a look at that help if you want to look at some other examples of using a data table.

< 00:22:52 END >