

DECISION ANALYSIS

Species of Int Model, Part 1

Terri Donovan

recorded: January, 2010

Today we going to see how you would handle the species of Int problem on a spreadsheet. What we're going to do is to look at the Word document that you received from class. And we're going to just paste in the values and the data, into the spreadsheet directly from the Word document.

The information we need is in Table 1 and Table 2. In Table 1 we have twelve different sites and we have fourteen different species. And the data themselves are 0's and 1's, and they are in this big grid. Where a 1 means the species was present, and a 0 means the species was absent from a site. So Site #1 contains species 2, 3, 8, 11 and so on.

Let's go ahead and copy those values and put them in our spreadsheet. Just copy and open up a new spreadsheet. We're going to paste those values into cell B6. So what you can see here are the sites. Go down column B and the data are all in column C. The first thing we need to do is to split the data up into their columns. And so we'll do that by selecting the data. And then we'll go up to the Data tab and then press the Text to Columns button and then use wizard to split the data out. And when we are finished we can resize all of our row heights, so that our spreadsheet is nice and clean.

So let's go ahead and label cell B5, 'Sites'. And then our species are, 1, 2, all the way out to 14. And we'll label those, 'Species'. And those are a heading, so let's go ahead and select cell C4 through P4. And merge and center those and give them some color to indicate it's a heading. Our Sites are given 1 through 12. Our Species are 1 through 14. And our data are contained in this grid here.

Now each of these Sites also has a cost, so let's head back to our document. And we'll pick up the price, and I don't know what the units are, but you can call them whatever you would like. And again we'll copy the price values. And head back to our spreadsheet and just paste those in. And this time we're going to choose Paste / Paste. And again, [select Data] use the Text to Columns Feature to split the data out. And now these are the prices for each of these Sites. So Site #1 costs 10 units, Site #2 costs 6 units and so on.

What we want to do is to get those down in column R, so that we can use that information more succinctly. So let's go ahead and copy those. Choose cell R6. And this time choose Paste / Transpose. And that is the cost. So let's label the cell R5, 'Cost'. Give that a heading. We no longer need these data so we'll go ahead and delete them.

Now one thing we wanted to do was just to identify how many species each site has. So we'll start by entering the words, 'Total Species' in cell Q5. And we'll just enter a simple equation that will count how many species Site #1, Site #2, Site #3, Site #4 have and so on.

PAUSE the video: What equation would you use?

And can just simply sum the values in the data grid to get this answer. Double-click to fill those numbers down and we see that Site # 1 has 4 species, and Site #2 has 4 species, Site #3 has 8 species and so on. And to just give us a little bit more room, let's select cells and columns C through P. And we'll shrink those down a bit to give us a little more space to work in.

One thing that Tony suggested that we do is to calculate the number of species per unit cost. We can do that in column S. And our species per cost is just a simple formula that is going to tell us which site is the best bang for their buck in terms of species conservation - just presence and absence of species.

PAUSE the video: What formula would you use here?

It's just going to be the total species divided by the cost of that site. And we can drag that number down. Standardize the decimal points to two decimal points. If you can only choose one site, which site would you choose to conserve? You choose the site that has the most number of species per unit cost. And that would lead you to selection of Site #4.

But our problem was a little bit different than that. We asked if you could choose any of these sites, how many would you need to choose in order to protect all of the species. If we just selected Site #4, we would conserve 6 species, but we would be missing the protection of the other species. So we need to think about this problem a bit differently now.

One way we can do this is to set up what's called the 'Choice Matrix'. So let's enter the word, 'Choose?', with a question mark in cell A5. And what we're going to do is use cells A6 through A17 as model inputs. So we'll color those green. We're going to let you the modeler enter the

number 1, if you are going to choose that site and the number 0 if you are not going to choose that site. So these can only be binary values. For an example, let's go ahead and fill this in. This would be the choice of Sites number 1, 4 and 5. If we were to actually purchase Sites 1, 4, 5, we need to then summarize what our conservation result is actually going to be - as a result of this choice.

What we are going to do is enter some equations down in this section of the spreadsheet that will compute how many species are actually conserved given this choice matrix. Let's start off by using this area as a heading. Let's label these cells, 'Number of Sites Conserved Given the Choice Matrix'. In this cell, C19, enter an equation that will tell us how many Sites with species number 1 conserved at, given this choice matrix over here.

Let's just walk through this. We chose Site #1. Well, a species number 1 did not occur there. We chose Site #4, species 1 did not occur there. We chose Site #5, species 1 did occur there, and so a number 1 would be representative of this answer. We don't want to do this piece by piece, instead we want to write an equation that will automatically calculate this, so that when we change this choice matrix these numbers will be automatically updated.

PAUSE the video: What equation would you use?

One way to do it is to write a long equation that would multiply the choice matrix. So Site #1 times the presence or absence of species 1 in Site #1, plus, our second site #2 times the presence or absence of species 1 in the second site #2, plus, and so on. I'm going to go ahead and enter this full equation out, and you can see the pattern. And what's happening here is that if the site is chosen. Well two things have to happen. First the site must be chosen and the species must be present in order for this term, any of these terms, to result in the number 1. And since we're adding those piece by piece we should get an answer that is reflective of stepping through things as we had done by hand. I'm going to stop right there just for the sake of illustration. So it's a nice way to see the pattern by clicking on the formula bar itself. And if we had carried this through we would have found that we would have still ended up with the number 1 because we never selected any of these other sites.

Now there is another function in Excel that is very handy and does this in one equation. And that is called, the 'Sum Product' function. There it is SUMPRODUCT. It returns the sum of the products of corresponding ranges or arrays. And we have two arrays. Our first array is the choice array. And our second array is the presence and absence of species 1 across the twelve different sites. We need to anchor the choice array because we're going to copy this formula

through for the other species as well. And when we press OK, it gives us our answer. Let's go ahead and center that.

Now we can drag this equation over and what this is telling us, is given these choices, we would end up conserving species 1 at one site, species 2 at two sites, species 3 at two sites, species 5 at no sites. These numbers are critical because our challenge now is to come up with a portfolio here, a suite of sites that we choose to buy, that will conserve all fourteen species. But under the constraint that we're going to minimize cost. If we had an unlimited budget, of course we could just choose all of these sites for conservation. And we would conserve every single species at all the sites that they exist in.

So how are you going to deal with this portfolio problem here? You can enter the number 1 or 0 in any of these cells. And there are a lot of different ways for us to evaluate. How many ways? Well, let's calculate that. Since there are two options, 1 or 0. At each of the twelve sites, we take two and raise it to the 12th. And that tells us there are 4096 options of flipping switches on and off among these twelve different sites. That's a lot to play around with. That includes the option where everything is selected, and it includes the option where none of the sites are purchased.

We have a set of options that we can make. Given our options we want to do two things. We want to make sure that all fourteen of these species are conserved. And we want to also do this for the lowest cost possible.

So let's add two outputs now to our spreadsheet. Let's calculate the 'Total Species Conserved', given this portfolio, and we'll enter that in cell D1. This will be a model output. And for this we just need to count how many species occur in at least one site.

PAUSE the video: What formula would you use here?

There are a few ways to go here, but one of the easier formulas is the COUNTIF formula. The COUNTIF formula has a range and a criteria. We're starting on the range. We're going to select the values in cells C19 through P19. Enter a comma. And the criteria is that we want to find values that are greater than or equal to 1. That needs to be in a quotation mark because we have this greater than symbol and that turns it into a text field. When you are finished press Enter and your values are computed.

Given a choice matrix, we can know how many species we actually can conserve given those

choices. Then we can play around with our values and see how those numbers actually change. There is an example of a portfolio that conserves all 14 species.

But we have another problem, too. And that was we wanted to do this with a minimal cost. And so that means we need to calculate the cost of this portfolio. And we'll do that in cell D2. Again, this will be a model output. We need to enter an equation in cell D2, that would compute the total cost given this portfolio. Can you think of a formula that would do that quickly?

PAUSE the video: What formula would you use?

Right, we're back to the Sum Product function. We'll go up to SUMPRODUCT equation. And in this case, the first array is our choice matrix. And the second array is the cost. These are two different arrays. What's going to happen in this, is that if this site is selected we add the cost for that site. So here we would be adding 10 plus 6. You would not be adding 15, because we did not choose site 3. We'd be adding 5. We'd be adding 17. And we would be adding 11 and 8. And that result is provided by this equation.

I should have done this a little bit earlier, but while I'm thinking about it. I'm going to just simply add a little bit of formatting, under the Conditional Formatting. We're going to highlight cells that are equal to 1. And give them a light red fill with dark red text, just to help us identify which sites have which species more easily.

Here is one solution, all 14 species are saved at a total cost of 57 units. Can you do better than that?

Well, there are 4096 possibilities here. And you certainly don't want to do this by hand. One way that you can do that is to use a tool in Excel, called Solver. You want to make sure Solver is installed on your spreadsheet. Go up to the Office button, and then drop down to Excel Options. Then choose the Add ins link. And then say Manage, Excel Add ins, and press Go. And make sure your Solver [Add in] option is clicked on. And then just press OK. And you are ready to go.

You can find Solver under the Data tab. And it's located - on my sheet anyway - to the far right. When we press Solver a dialog box appears. And we're going to be filling out this dialog box. In this case, let's just think about what we want to do. We want to be able to change the Choice matrix. So those are going to be the cells we change.

We want to set a target cell which is the Total Cost. So we're going to choose the total cost, and we're going to minimize that. So we want to make some choices that will provide us with a minimum cost. And so the changing cells values are in cells A6 through A17. But we have some constraints right, we can't just say, you know, choose everything or choose nothing.

We first need to make sure that all 14 species are conserved. And second we need to tell Solver that the numbers 1 and 0 are the only acceptable values in these green boxes. Let's do that first.

Let's add a new constraint. And we'll choose our choice matrix, and then we'll choose our drop-down arrow. And then choose the option called B I N or binary. And that tells the spreadsheet that these values can be either 0 or 1. Let's add our second constraint. And that was that all the species must occur on at least one site. So these numbers need to be greater than or equal to 1. And that is all the constraints that we need. So we'll double check our dialog box. And when we are ready we'll go ahead and press Solve. In this case Solver found a solution. And we're going to keep those solutions. And we'll press OK.

You can see a couple of things here. The first is that it did find a solution. It chose sites 2, 4, 6 and 7. The total cost was 23. That is quite a bit less than our first option which also worked. So this is the optimal solution.

You can play around with all 4096 of these combinations. But this particular combination will satisfy the constraints that you'll save all 14 species for the lowest possible price.

The other thing you will notice is that Solver took a few liberties in making these not really binary. And that has to do with some of the settings in Solver, which we'll go through in a separate video. For our case this is fine. It's not going to matter too much, but if you want to you can change the Data, so that you display 0 decimal points. And we should also go ahead and make that a number. I wonder if Solver would do better if it is formatted as a number. That is something you can try.

At this point you found a solution. You've used the spreadsheet to evaluate 4096 different portfolio options for site selection. And you used Solver to find that portfolio that minimizes the total cost, but meets your goal of conserving all 14 species on at least one site.

< 00:22:11 END >