

A. Overview

This tutorial provides instructions for analyzing data downloaded from the Gorongosa WildCam Lab for the *Measuring Biodiversity In Gorongosa* activity. In this tutorial you will investigate the richness, evenness, and Shannon Diversity Index of the different vegetation types in Gorongosa National Park.

B. Make a Pivot Table

For research questions in which you need to add up the number of animals in a particular category (e.g., number of animals in each season), creating a **pivot table** is a quick and easy way to analyze your data.

1. Open the spreadsheet you downloaded from WildCam Lab and immediately save it as an Excel file (.xls or .xlsx).
2. Click the **Insert** tab and select **Pivot Table** (Figure 1).

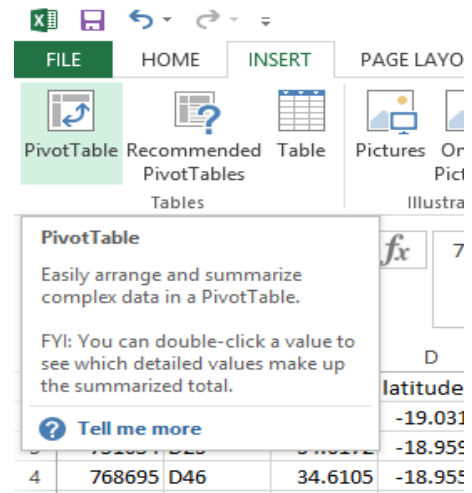


Figure 1: Inserting a pivot table.

3. Click on the button next to **Table/Range**. Navigate to your data and click on the letter “A” to highlight the first column. With the left mouse button held down, drag the cursor to the last column to highlight the entire data set. Under “Choose where you want the PivotTable report to be placed,” select **New Worksheet**. Click **OK** (Figure 2).

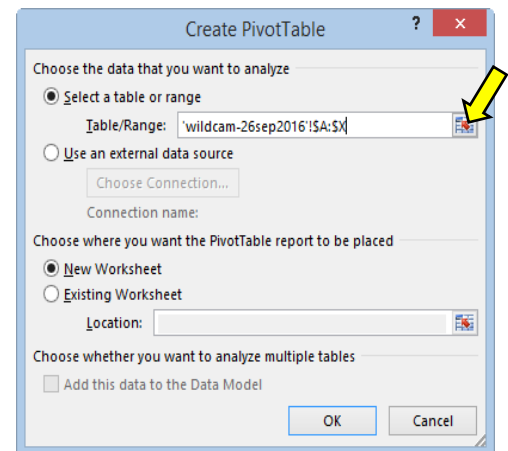


Figure 2: Selecting the range of data.

C. Build the Correct Pivot Table

A pivot table can be used to group data by certain variables. In this example, the numerical variable, “species_count,” will be grouped by two categorical variables, “veg_type” and “species.” To build the pivot table, you will follow the steps explained below to place the grouping variables (veg_type and species) in the ROWS field and the variable being grouped (species_count) in the VALUES field.

- In order to create a table that organizes the number of animals in each species for each vegetation type, you will need to select “veg_type”, “species” and “species_count” fields in the list of **PivotTable Fields**. (Figure 3).

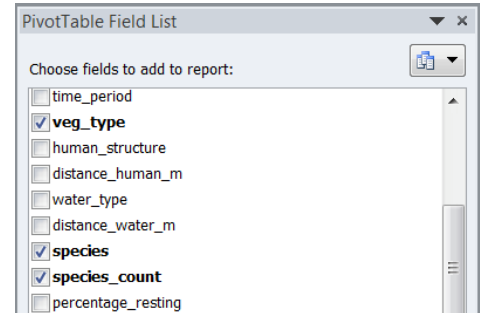


Figure 3: Selecting the fields to use in building the pivot table.

- By default, the fields you select will appear in the ROWS field. Drag the “species_count” variable to the VALUES field (Figure 4). This will produce a table with the veg types and species in one column and the species_count in the other column.

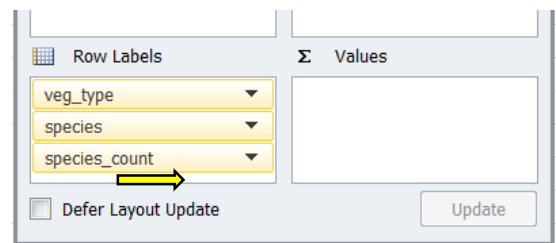


Figure 4: Moving the species_count variable to the VALUES field.

- In the VALUES column, make sure that the setting for your variable (species_count) is appropriate for the dependent variable you have chosen. In this case, we want to sum the variable “species_count.” The default is to count the number of observations (photos), but we want to total the number of animals observed in each photo. Click on **Count of species_count** in the VALUES column and click on **Values Field Settings**. Select “Summarize Values By” **Sum** (Figure 5).

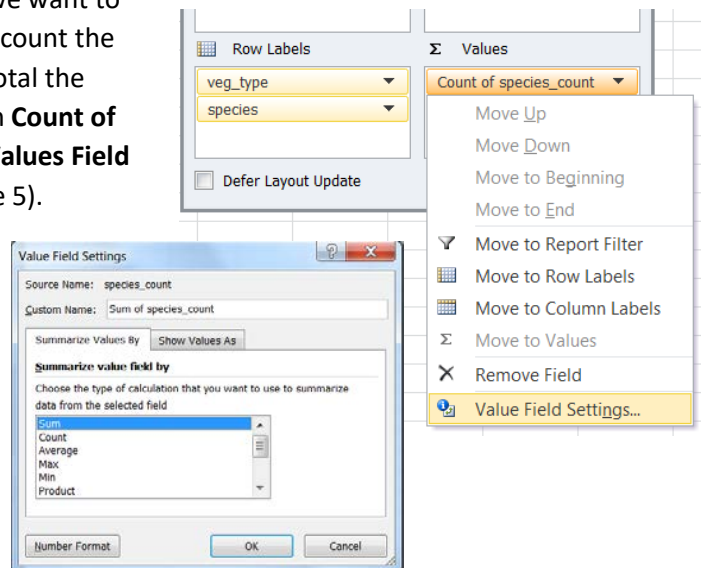


Figure 5: Choosing how to summarize the value field.

D. Calculate Richness

You now have a table with each species listed under the three vegetation types in the left column and the total number of animals in the column to the right. To calculate species richness, you simply have to count the number of species within each vegetation type and record it in the table in question 6 of your worksheet. There is an Excel formula that will make this easy for you.

- In any blank cell to the right of the pivot table, use the formula =COUNTIF(range, criteria) to count how many species are in the Grassland vegetation type. This formula will count how many cells that you select contain any text.

When you type the formula, in the "range" section, highlight the cells that contain all the species under the Grassland category. In the "criteria" section, type "*", which means count the cell if it contains any text. Write the value in the table in question 6 of your worksheet.

| Row Labels | Count of species_count | | |
|----------------------|------------------------|--|--|
| Floodplain Grassland | 13731 | | |
| Aardvark | 9 | | |
| Baboon | 3342 | | |
| Bird (other) | 17 | | |
| Buffalo | 83 | | |
| Bushbuck | 2739 | | |
| Bushpig | 52 | | |
| Caracal | 2 | | |
| Civet | 41 | | |
| Duiker | 297 | | |
| Elephant | 1157 | | |
| Genet | 7 | | |
| Ground Hornbill | 1 | | |
| Hippopotamus | 85 | | |
| Honey Badger | 21 | | |
| Human | 501 | | |
| Impala | 241 | | |
| Jackal | 1 | | |
| Kudu | 5 | | |
| Lion (female) | 3 | | |
| Lion (male) | 2 | | |
| Mongoose | 12 | | |

| Species Richness Table | |
|----------------------------|--------------------------|
| Vegetation Type | Richness |
| Floodplain Grassland | =COUNTIF(A4:A36,"*") |
| Limestone Gorge | COUNTIF(range, criteria) |
| Miombo Woodland | |
| Mixed Savanna and Woodland | |

Figure 6: Calculating richness for the Grassland vegetation type.

- Repeat the steps above for other three vegetation types and record your results in the table in question 6 of your worksheet. When you are done, you can clear the values from the cells you used to do the calculation.

E. Calculate the Shannon Diversity Index

You will use the pivot table to calculate the Shannon Diversity Index for each vegetation type. Remember that the Shannon Diversity Index is represented by the letter "H" and the formula is: $H = -\sum (p_i * \ln(p_i))$

Where:

- p_i = abundance of a species / total abundance
- \ln = natural log

- Create the following column headers next to the pivot table in columns C through F: **p_i , $\ln(p_i)$, $p_i \times \ln(p_i)$, H.**

| Row Labels | Sum of species_count | | | | |
|----------------------|----------------------|-------|------------|-----------------------|---|
| Floodplain Grassland | 18639 | p_i | $\ln(p_i)$ | $p_i \times \ln(p_i)$ | H |
| Aardvark | 9 | | | | |
| Baboon | 5695 | | | | |
| Bird (other) | 29 | | | | |
| Buffalo | 345 | | | | |
| Bushbuck | 2841 | | | | |
| Bushpig | 69 | | | | |

Figure 7: Adding column labels to the spreadsheet

10. Select the cell in the "pi" column next to the first species in the table. You will use the formula =species abundance/total abundance. For the species abundance, select the cell for the "Sum of Species count" of the first species in the grassland vegetation type. For the total abundance, select "Sum of Species count" of all animals in the grassland vegetation type. Repeat this for all species in all of the vegetation types.

| | A | B | C | D |
|---|----------------------|----------------------|--------|--------|
| 1 | | | | |
| 2 | | | | |
| 3 | Row Labels | Sum of species_count | | |
| 4 | Floodplain Grassland | 18639 | pi | ln(pi) |
| 5 | Aardvark | 9 | =B5/B4 | |
| 6 | Baboon | 5695 | | |
| 7 | Bird (other) | 29 | | |
| 8 | Buffalo | 345 | | |

Figure 8: Calculating pi.

11. Select the cell in the "ln(pi)" column next to the first species in the table. You will use the formula =LN(pi). Type =LN() and, with your cursor between the parentheses, select the cell that corresponds to the pi value for that species (the cell directly to the left). Copy this cell and paste it into the remaining cells within that vegetation type. This will repeat the formula using the cell to the left of it for the pi value. Repeat for each of the vegetation types.

| Row Labels | Sum of species_count | | |
|----------------------|----------------------|----------|---------|
| Floodplain Grassland | 18639 | pi | ln(pi) |
| Aardvark | 9 | 0.000483 | =LN(C5) |

Figure 9: Calculating ln(pi).

12. Select the cell in the "pi x ln(pi)" column next to the first species in the table. You will use the formula =pi*ln(pi). Type the "=" sign, then select the cell that corresponds with pi for that species (two cells to the left), type the "*" symbol (which is the multiplication symbol) and then select the cell corresponding to ln(pi) for that species (the cell directly to the left). Copy this cell and paste it into the remaining cells within that vegetation type. This will repeat the formula using the cell to the left of it for the pi and ln(pi) values. Repeat for each of the vegetation types.

| pi | ln(pi) | pi x ln(pi) |
|----------|----------|-------------|
| 0.000483 | -7.63579 | =C3*D3 |

Figure 10: Calculating pi x ln(pi).

13. Select any cell in the "H" column that is within the first vegetation type. Use the formula =-SUM(pi*LN(pi)) to calculate H for the Grassland vegetation type. To do this, type =-SUM() into the cell. With your cursor inside the parentheses, select all the cells in the "pi x ln(pi)" column for that vegetation type. Use the same formula for the other vegetation types. These values are the Shannon Diversity Indices for each vegetation type. Record your results in the table in question 6 of your worksheet.

| | Floodplain | 18639 | pi | ln(pi) | pi x ln(pi) | H |
|--------------|------------|----------|----------|----------|--------------|---|
| Aardvark | 9 | 0.000483 | -7.63579 | -0.00369 | =-SUM(E3:E7) | |
| Baboon | 5695 | 0.305542 | -1.18567 | -0.36227 | | |
| Bird (other) | 29 | 0.001556 | -6.46572 | -0.01006 | | |
| Buffalo | 345 | 0.01851 | -3.98947 | -0.07384 | | |
| Bushbuck | 2841 | 0.152422 | -1.8811 | -0.28672 | | |

Figure 11: Calculating H.

F. Calculate Evenness

You will use the Shannon Diversity Index and richness values to calculate evenness for each vegetation type.

Remember that Evenness is represented by the letter "E" and the formula is: $E = H / H_{MAX}$

Where:

H = Shannon-Weiner Diversity Index

H_{MAX} = the highest possible diversity for that sample (calculated by $\ln(\text{richness})$)

14. Create a table in a new worksheet with four rows for the four vegetation types and four columns for: "H", "Richness", " H_{MAX} ", and "Evenness". Fill in the H and Richness values you calculated in parts D and E.

| Vegetation Type | H | Richness | H_{MAX} | Evenness |
|----------------------------|---|----------|-----------|----------|
| Floodplain Grassland | | | | |
| Limestone Gorge | | | | |
| Miombo Woodland | | | | |
| Mixed Savanna and Woodland | | | | |

Figure 12: Table to calculate evenness.

15. In the first cell in the H_{MAX} column of the table, use the formula $H_{MAX} = \ln(\text{Richness})$. Type $=\ln()$ and, with your cursor between the parentheses, select the cell that corresponds to the richness value for the Grassland. Repeat this for all vegetation types.

| Vegetation Type | H | Richness | H_{MAX} |
|----------------------------|-------------|----------|------------|
| Floodplain Grassland | 2.075889588 | 33 | $=\ln(C3)$ |
| Limestone Gorge | 2.026096475 | 34 | |
| Miombo Woodland | 2.409250273 | 25 | |
| Mixed Savanna and Woodland | 2.364909414 | 43 | |

Figure 13: Calculating H_{MAX} .

16. In the first cell in the Evenness column of the table, use the formula $\text{Evenness} = H / H_{MAX}$. Type "=" and select the cell that corresponds to the H value for the Grassland, then type "/" and select the cell that corresponds to the H_{MAX} value for the Grassland. Repeat this for all vegetation types. Record your results in the table in question 6 of your worksheet.

| Vegetation Type | H | Richness | H_{MAX} | Evenness |
|----------------------------|-------------|----------|-------------|----------|
| Floodplain Grassland | 2.075889588 | 33 | 3.496507561 | $=B3/D3$ |
| Limestone Gorge | 2.026096475 | 34 | | |
| Miombo Woodland | 2.409250273 | 25 | | |
| Mixed Savanna and Woodland | 2.364909414 | 43 | | |

Figure 14: Calculating Evenness.

Note: The version of Excel used in this tutorial is Excel 2013 for PC. Not all features and functions may work the same way in other versions of Excel.