



CATTLE. ANTIMICROBIAL. LAND. FINANCIAL.

Microsoft Excel for the Veterinary Practitioner: Pivot tables and the =VLOOKUP function

W. Isaac Jumper, DVM



MISSISSIPPI STATE UNIVERSITY™
COLLEGE OF VETERINARY MEDICINE

Microsoft Excel for the Veterinary Practitioner: Pivot tables and the =VLOOKUP function

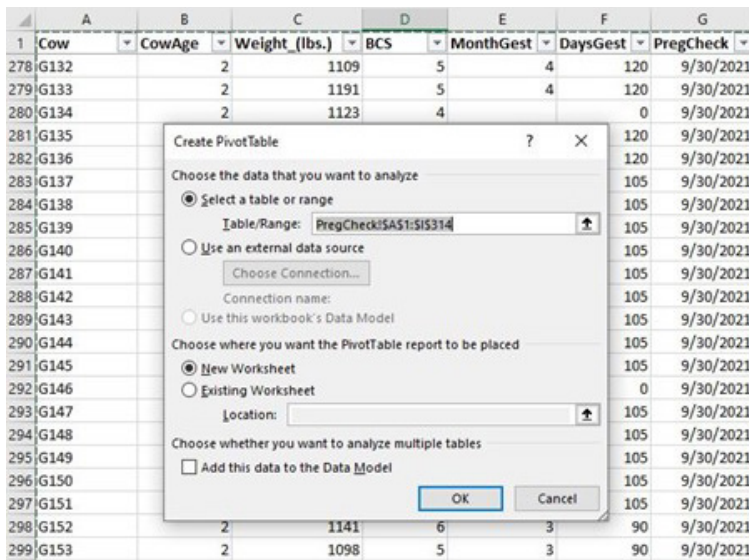
W. Isaac Jumper, DVM

Introduction:

Pivot tables are powerful tools used to create comparisons, summaries, and stratifications of data. Pivot tables offer the user a multitude of sorting and filtering capabilities that would be cumbersome or confusing by other means. The =VLOOKUP function allows users to find data in different locations (e.g. other columns with the same spreadsheet or columns in separate spreadsheets) and merge it with existing data. The objective of this article is to describe the use of a pivot table and the =VLOOKUP formula when analyzing animal health data in Microsoft Excel.

Adding a pivot table

Pivot tables allow the user to sort large amounts of data by categorical variables of the user's choosing. For example, the user might want to know how many cows in a cow-calf data set are open, 5 years of age or older, and in a body condition score of 4 or less. This type of question can be easily answered using a pivot table. Similarly, contingency tables may be created to facilitate risk factor analysis, such as calculating relative risk of neonatal calf scours based on age of dam.



A pivot table is based on data within one spreadsheet. Prior to creating a pivot table, it's important to ensure there are no blank rows or columns within your data. These blank rows or columns can prevent all the data within the spreadsheet from being incorporated into the pivot table. To create a pivot table, select any cell within the data (i.e. ensure a blank cell outside the data is not selected). With any cell containing data selected, go to the "Insert" tab on the ribbon at the top of the screen, and click on the "PivotTable" button. After clicking the "PivotTable" button, the menu shown in Figure 1 will appear, and the entire data set that will be incorporated into the pivot table will be selected with sliding dashed green lines. Typically, there are no changes that need to be made to this menu. The menu shown in Figure 1 indicates that the data contained in the range of cells from A1 to I314 within the "PregCheck" spreadsheet will be used to build the pivot table,

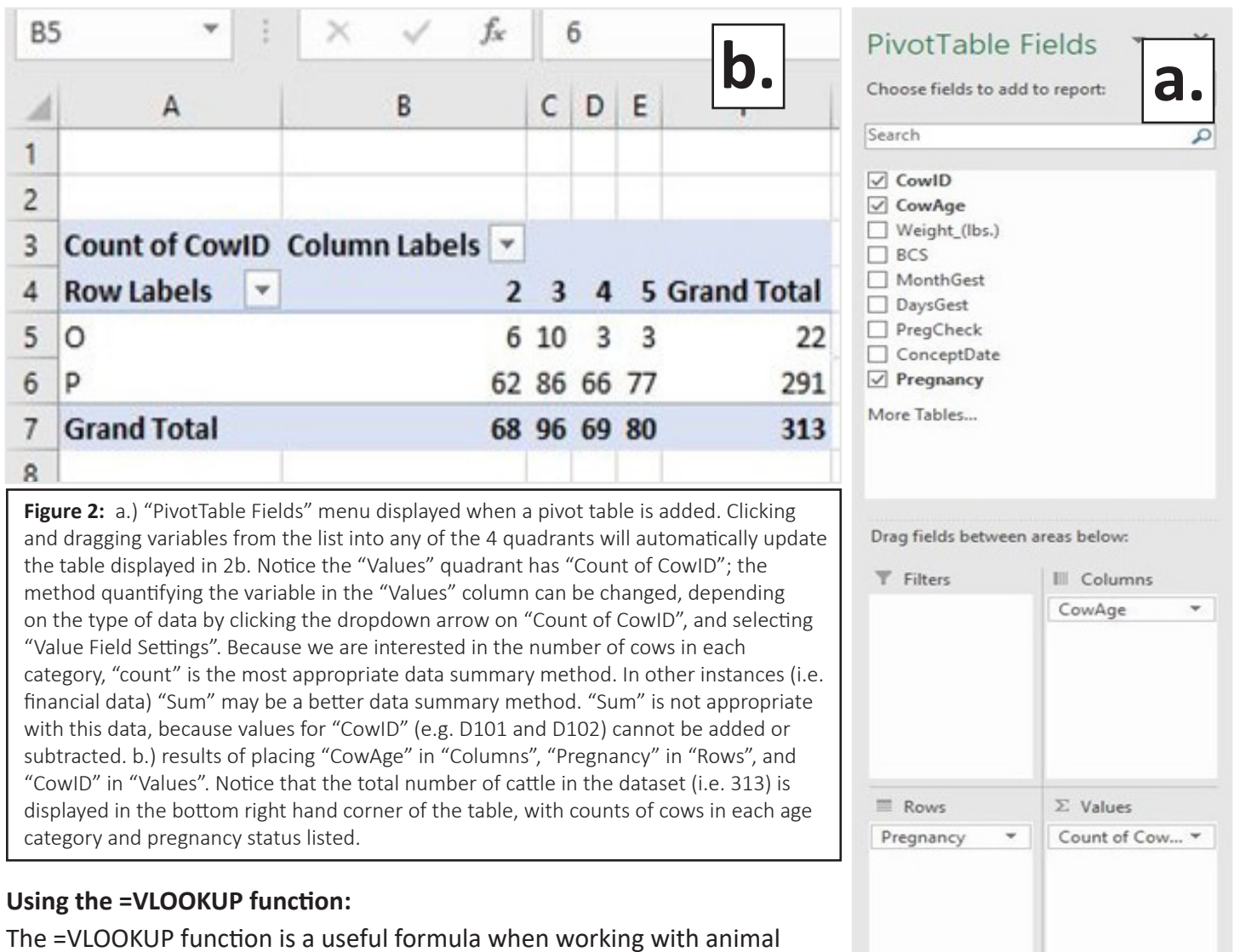
Figure 1: After clicking on "PivotTable" under the "Insert" tab, the menu above will be displayed. The user should ensure the correct data is selected, and that the pivot table will be placed in the desired location once build, before clicking on "OK".

and once built, the pivot table will be placed in a new worksheet. At this point, the user can click "OK", and a new spreadsheet will be opened in the workbook containing the pivot table.

When a pivot table is inserted into a workbook, a new spreadsheet will be inserted into the workbook. This spreadsheet is initially blank, but a "PivotTable Fields" menu is open by default on the right side of screen. The first box under the "PivotTable Fields" will contain a list of variables made up of all the column headings in the dataset used to generate the pivot table. Below this list is a 2-by-2 square table with the following quadrants: Filters, Columns, Rows, and Values. To begin populating the pivot table, click and drag any of the

desired variables from the list of column headings into one of the four quadrants listed. For example, Figure 2a shows the components of a contingency table created to display the number of open or pregnant cows within each age group. As the user places variables into any of the four quadrants under “PivotTable Fields”, the spreadsheet will automatically update to produce results. Figure 2b shows the results of the variables placed in each quadrant in Figure 2a.

Filters may be applied to pivot tables by placing the desired filter variable in the “Filter” quadrant of the “PivotTable Fields”. For example, the user might be interested in filtering the data shown in Figure 2b by Body Condition Score (BCS) to determine the distribution of body condition scores among pregnant and open cows of each age category. Pivot tables are a great way to sort and filter data to answer specific questions such as “How many 2 and 3-year-old cows in a body condition score of 4 were open?”



Using the =VLOOKUP function:

The =VLOOKUP function is a useful formula when working with animal health data. There are often situations where the user needs to find a value in a range of data, but manually searching for that value is not plausible. Large datasets may prevent the user from quickly and easily finding a value by manually scrolling through the data. For example, Figure 3a displays data that may be collected during calving season. As calves are born, they are recorded and numbered numerically by order of birth. In Figure 3, each calf has the ID of their dam in the adjacent column. The =VLOOKUP function allows you to search a range of cells, find a specific value, and return a related value

in an adjacent column. For example, data collected at weaning may appear as shown in Figure 3b. If the user desired to know the ID of the dam of each calf in the weaning sheet without manually search through the calving sheet, the =VLOOKUP formula can be used to search the calving sheet for the ID of the dam of each calf in the weaning sheet. The function requires four pieces of information as shown in Figure 3b. The first portion of the formula is the lookup value. In this example, the value we want to search for is the Calf ID, J290. The table array is the second piece of information needed by the formula. This is the range of cells that the will search when looking for the lookup value, J290. The

	A	B	C
1	CalfID	DamID	CalfBirthDate
2	J001	G103	2/5/2021
3	J002	G101	2/6/2021
4	J003	G105	2/11/2021
5	J004	G120	2/16/2021
6	J005	G113	2/19/2021
7	J006	G116	2/19/2021
8	J007	G122	2/19/2021
9	J008	G114	2/19/2021
10	J009	G131	2/19/2021

array for this example is the calving spreadsheet shown in Figure 3a. Thus far, the formula will search for J290 in the sheet. The third piece of information column index number. This is the number column containing the information that will be returned by the formula. In this case, damID is the desired value to be returned by the formula, and damID is in the second column (when moving from left to right) in the table array. Lastly, the function asks if the user wants an exact match or an approximate match. In this case, we are interested in an exact match. So, the =VLOOKUP function will search for the value J290 in the calving sheet within the range A1:C314, and it will return the adjacent value in the second column range A1:C314 (i.e. the ID of the dam). The same process could be used to return the birth date of calf shown in Figure 3a if the user

	A	B	C	D	E	F
1	CalfID	WW	DamID			
2	J071	471	F128			
3	J267	481	D169			
4	J081	526	E114			
5	J222	584	G106			
6	J264	495	G154			
7	J303	423	D112			
8	J044	604	F153			
9	J290	448	=VLOOKUP(A9,Calving!\$A\$1:\$C\$314,2,FALSE)			
10	J077	653				

Figure 3: a.) Commonly collected calving season data. Note the calf ID's are arranged in numerical order by date of birth. b.) Weaning weight data located in a spreadsheet separate from calving data. The =VLOOKUP formula allows the user to search data in the calving spreadsheet to fill in dam ID in the weaning spreadsheet.

needed days of age at weaning to calculate values such as average daily gain.

Using the =VLOOKUP function can be useful when data from one spreadsheet needs to be combined with data from another spreadsheet. This formula makes searching for values easier and more efficient than manually searching the spreadsheet.

Conclusion

Pivot tables and the =VLOOKUP function are useful tools to the veterinary practitioner working with animal health and production data. Pivot tables allow the user to create contingency tables to sort the data and explore relationships. The =VLOOKUP function allows the user to search and retrieve data in a quick and efficient manner, rather than manually searching large datasets. Collectively, these two aspects of Microsoft Excel offer the practitioner useful tools when working with animal health and production data.



This work is supported by the Agriculture and Food Research Initiative Competitive Grants Program grant no. 2018-69003-28706 from the USDA National Institute of Food and Agriculture. Any opinions, findings, conclusions, or recommendations expressed in this publication are those of the author(s) and do not necessarily reflect the view of the U.S Department of Agriculture.