# LEARNING OBJECTIVES

In Exercise 1 we looked at variables one at a time. In order to explore relationships between two nominal variables, we can use contingency tables (called Pivot Tables in Excel) to summarize the relationship between these variables.

* Learn techniques for analyzing nominal data.
* Create contingency tables to explore potential relationship between variables.
* Rephrase findings using natural language.

# DIRECTIONS

The next set of tasks involves the creation and use of Pivot Tables to help us summarize and analyze the relationship between variables in our data set.

We will start with a simple cross-tabulation table to report the relationship between the variables. In this example, we will examine three variables from the PEW Study pertaining to relationships, society, and gender (your variable names may differ).

# Preparing your Data

First, we will prepare our data for the analysis. Open your dataset workbook from Exercise 1.

* Make a copy of the **Working Sheet**. Rename it as **Society**.
* Make sure your worksheet only has one row of labels for each column.
* Delete all but five variables: **respid, RELATIONSHIPS, GOODFORSOCIETY** and **SEX**.
* You’re going to delete all the values for “Don’t Know” or “Refused to Answer” (in this dataset, this would usually be the values 8, 9, To clarify, you are not going to delete the entire case – just the indicator.

Some of the indicators remaining in the dataset that were volunteered by respondents are difficult to interpret. For example, in the question about relationships, the response “both equally” is perfectly valid, but may be difficult to interpret when comparing to responses such as “strengthens” or “weakens.” Sometimes it is advisable to look at a simplified subset of the evidence, and then add back more complexity when conducting your analysis. For the purposes of this exercise, we’re going to eliminate some responses.

* Delete indicators 3, 4, 5 for the **RELATIONSHIPS** variable, and delete indicator 3 for the **GOODFORSOCIETY** variable.
* Delete any cases that now have a blank for **GOODFORSOCIETY** (delete the entire row). This is going to be our outcome variable, so any case that is missing this value is irrelevant to our current analysis.
* Likewise, delete any cases that now have a blank for **RELATIONSHIPS**.
* This will reduce the dataset to 379 cases.

The table below references what variables you should have left.

|  |  |
| --- | --- |
| **RELATIONSHIPS**  (pial9) | Thinking about your relationships in general... OVERALL, would you say that communicating online with friends and family.   1. Generally STRENGTHENS those relationships, OR 2. Generally WEAKENS those relationships? |
| **GOODFORSOCIETY**  (pial11) | Overall, when you add up all the advantages and disadvantages of the internet, would you say the internet has mostly been a GOOD thing or a BAD thing for society?   1. Good thing 2. Bad thing |
| **SEX**  (sex) | RECORD RESPONDENT SEX   1. Male 2. Female |

# Create the Pivot Table

After you have finished all the above steps for preparing the data, you can create the pivot table.

* Place your cursor at cell A1 of the **Society** sheet.
* Go to the Insert tab and click on the PivotTable icon.
* The Create PivotTable Window appears, and the Table/Range box should default to your complete data set.
* Make sure the **New Worksheet** option is selected and click **OK**.
* A blank Pivot Table appears on the left of the sheet, with a Field List of items on the right that you can choose for your cross-tabulation. Drag the **SEX** button to the **Column Label** area and the **GOODFORSOCIETY** button to the **Row Label** area below the Field List.
* Drag the **Response Number** button to the **Values** area below the Field List (bottom right). Notice that the Response Number button automatically changes to Sum of Response Number. You need a Count of the Response Numbers to get the right numbers for each cell in the data section of the table, so click on the Sum of Response Number button and select Value Field settings, in order to see the options for summarizing this field.
* Your Pivot Table should now be populated on a new sheet.
* Rename this sheet with the name **Society Pivot**.

# Change Labels in the pivot

Take a look at the pivot table you have created and think about how it could be improved. Once again, the row and column labels are not very informative. This time we will use a different approach to fix this problem – we will create new text fields in the **Society** sheet and then use these fields in the pivot table layout.

* Change to your **Society** sheet and insert a column next to the SEX column. Call it **SEX Description** (or SEC Desc for short). Use the VLOOKUP function to fill this column with the correct text values for the codes in the **SEX** column. [HINT: Remember, make sure that the cell addresses for the Table array in the Lookup formula are **absolute** addresses.]
* Do the same for the **GOODFORSOCIETY** column (abbreviate the column label so you don’t have an excessively long name – for example, you could call it **SOC Desc**).
* Now go back to your **Society Pivot** and refresh the data by clicking on the Refresh icon on the Pivot Table Tools options tab. You should see your new field names appearing in the Field list. You may have to use the Change Data Source icon to expand the range of columns for your data.
* Drag the SEX button off the Labels area and drag the **SEX Desc** button onto the Labels area. Do the same for the **GOODFORSOCIETY** button.
* It would also be more useful if we could see percentages of the figures in the contingency table. Select the **Value Field** settings again and choose the **Show Values As** tab. Select the ‘**% of column’** option. Click on the Number Format button, choose Percentage format and set the decimal places to zero. Click on OK twice and you should see percentage figures in the pivot table.

# Creating Charts

It is often preferable to display data graphically, as this can be easier for the reader to assimilate. It is easy to prepare a chart from a pivot table.

* Click on your pivot table, and then click on the PivotChart button on the PivotTable Tools options tab.
* Experiment with options for formatting the chart to display the data as usefully as possible.

# Questions

1. In terms of “eye-balling” the pivot table, what does the evidence suggest about the potential relationships gender and whether or not someone thinks the Internet is mostly a good or bad for society?
2. Read the percentages across the rows. How would you characterize the proportion of women who characterize the Internet as bad for society? 1 in 3? 1 in 10? How does this compare to the proportion of men who characterize the Internet as bad for society?
3. Without doing the calculation, what would happen to the column percentages if we added back all the cases who responded “Some of both” (recorded with a 3) to the question about being good for society? This would an additional row in the table. Would this make the distribution between men and women appear more similar or more different? (If you get stumped, go ahead and create the table to check and see).

When you have finished examining these two variables, set up a similar analysis with the variable pertaining to whether a subject believes that online communication strengthens or weakens a relationship. Create your column of value labels, and then set up your pivot table. Remember to examine the data both in actual numbers and as percentages.

# Some More Questions

1. Which variable is a stronger predictor of whether someone considers the Internet to be good or bad for society? Gender or the perception of how online communication affects relationships? How would you compare the distribution of these latter two variables to the first pair you looked at?
2. Without doing the calculation, what would happen to the column percentages if we added back all the cases who responded “Both equally” or “Neither” (recorded with a 3 or 4) to the question about relationships?

# Submitting your work

To complete the exercise, submit your workbook with sheets for Valid Cases, Society, and Society Pivot that shows your pivot table.

Include an **answer sheet** where you record your answers to questions 1-5.