

## Let's use Technology Use Data from Cycle 14 of the General Social Survey with Fathom for a data analysis project

### Example: Who communicates by email most frequently?

#### Data Content:

This Technology Use dataset in Fathom contains 82 attributes for 1,000 randomly selected Canadians of 15 and over. The data were extracted from the General Social Survey cycle 14 on technology use.

#### Procedure:

#### To Access Technology Use data from Cycle 14 GSS

1. Start-up Fathom  
After you have installed Fathom from the CD, Fathom's icon should appear on your desktop. Double-click it and Fathom will start.
2. From within Fathom choose *Open* from the *File* menu.
3. Go to the *TechUse.ftm* document. It is within the Statistics Canada Data folder within the Sample Documents Folder on the Ontario Fathom Dynamic Statistics CD-ROM.

#### Inspecting Individual Cases

The box with the gold balls labelled *Respondents* is a collection. It contains information about 1,000 people or *cases*. These cases can be viewed one at a time or in a case table.

4. Click once on the collection (the box of gold balls) and move the mouse over the lower-right corner of the border. Click and drag the mouse down and to the right to "open" the collection.

As the collection opens you will see some gold balls. Each gold ball represents one person who took the survey. The number below each ball is the unique identification number for that person.

5. Double-click the first of the gold balls.
6. Click the comments tab at the top of the resulting inspector window. Read the text for a general description of the General Social survey and where to find more information about the survey. Question: How large was the total sample for this survey? \_\_\_\_\_

7. Click the cases tab near the top-left of the resulting inspector window.

The inspector for the collection displays the attributes and attribute values for the case. Note the arrows in the left-corner of the inspector. Next to the arrows it says, "1/1000," indicating that this is the first case out of 1000.

8. Click on the right-pointing arrow several times, to look at the values for other cases in the collection.

There are 82 attributes for each case. Each attribute corresponds to a question from the original survey.

9. You can drag the boundaries between the columns to make more room for attribute names and values.

### **Viewing Cases in a Case Table**

It can be helpful to view data in a table so that each row is one case and each column holds the value of one attribute.

10. Click once on the collection to select it.
11. Position the mouse over the case table icon in the tool bar. Click and drag a new case table into the document.
12. You can resize the case table by dragging its edges or corners.

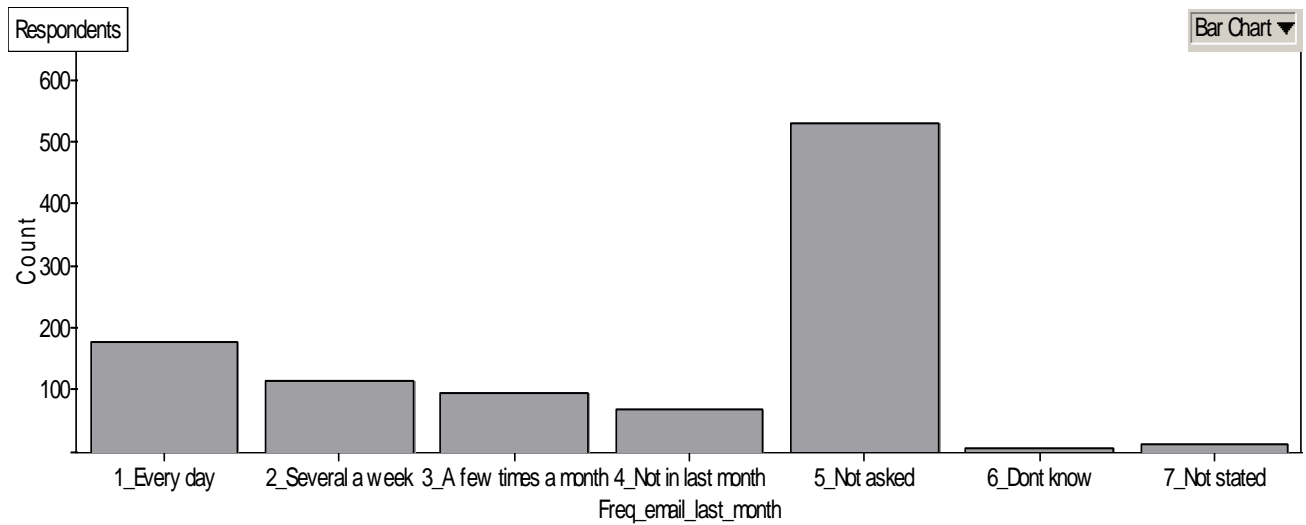
*Hint:* Keep your case tables small so that there is enough room for other objects such as graphs.

### **Making Graphs**

Graphs provide very powerful tools for exploring data. For the purpose of this analysis, we are interested in identifying which sex communicated by email most frequently in the past month?

13. Drag a graph from the tool bar into an empty place in the document.  
The graph comes up empty because Fathom doesn't know what attribute to plot.
14. Click and drag the attribute Freq\_email \_last\_month to the x-axis of the graph.

Since this attribute is categorical, Fathom creates a bar chart with the height of each bar proportional to the number of cases in each category.



count ( )

What does this graph tell you about frequency of communicating by email?

15. Make a second bar chart, this time with sex on the x-axis.

16. Click once on the bar labelled '05\_Not asked' in the first bar chart.

When you select cases in one view, Fathom shows them as selected in all views by shading the selected cases in red.

What does this graph tell you about frequency of communication by email in the last month by males and females?

Notice that because respondents who were “not asked” this question are still included in the graph, it is difficult to compare the proportions of communicating by email.

### Filter Cases

Sometimes we’re only interested in a subset of the available data. In this case, we want to examine the frequency of communication by email by sex of respondent, without being distracted by questions that were “not asked” of certain respondents.

With less activity in the graph, it may be easier to determine who uses the internet to communicate by email most frequently?

17. Click once on the chart of frequency of communication by email in the past month.

18. Choose *Add Filter* from the *Data* menu.

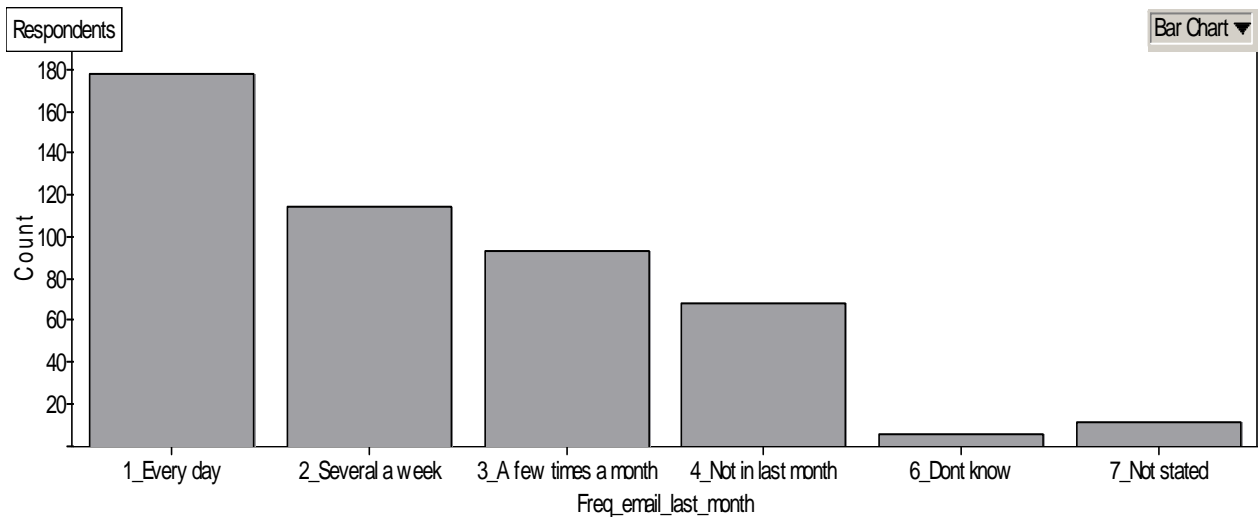
This brings up the formula editor. The expression you create here will filter out those cases for which the value of the expression is false.

19. Create the expression shown here to filter out the cases where the frequency of communication by email was not asked:

`¬includes (Freq_email_last_month, "asked")`

**Note:** the first symbol is what you get when you click the 'not' button on the pop-up keypad

- You can type the whole expression, or you can enter attribute names by double-clicking them in the list of attributes.
- When you type a left parenthesis, the right parenthesis appears automatically.
- When you type a double quote, its matching double quote appears automatically.
- Attribute names that are correctly entered appear in blue.
- There is no way to find the list of possible categories for an attribute.



`count ( )`  
`¬includes (Freq_email_last_month, "asked")`

## Ribbon Chart

Now that we have created our filter, let's make a ribbon chart since it is an excellent tool to make comparisons.

20. Hold down the mouse on the Bar Chart popup menu in the upper-right corner of the bar chart of frequency of communication by email and choose Ribbon Chart.

21. A ribbon chart allocates a strip along its length for each category. The width of the strip is proportional to the count. Since selection shows vertically in each strip, it is relatively easy to compare the proportion selected in different categories.

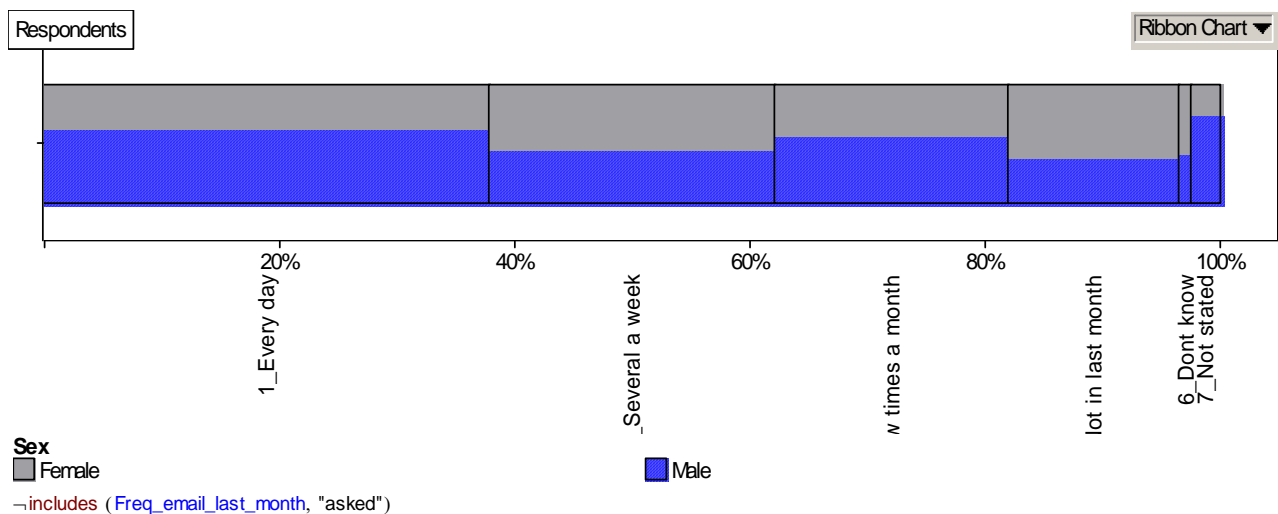
Can you note a difference in the frequency of communication by email by sex?

22. Get rid of the selection by clicking in white space in the graph.

23. Now, drag the sex attribute into the middle of the ribbon chart.

Dropping the sex attribute into the middle of a ribbon chart splits the each category and shows a legend.

Can you notice a difference in the frequency of communicating by email last month by sex of respondent?



## Displaying Summary Statistics

Sometimes it is important to have the actual numbers, not just a picture. Let's look at the information we generated using a ribbon chart, but this time in a summary table.

24. Drag a summary table from the tool bar and place it in an empty area of the document.

25. Drag the Freq\_email\_last\_month attribute to the empty column header of the summary table.

26. Resize the summary table so that you can see all of its contents.

Now let's break these results down according to sex.

27. Drag the Sex attribute onto the summary table's empty row header.

28. Resize the resulting table so that you can see all the data.

What can you conclude from this table?

This summary table is much like the bar chart we made in the previous section. It is fine for comparing counts, but not as good for comparing proportions. We can change that by changing the formula for the values computed in each cell.

29. Double-click on the formula displayed under the summary table.

This will open up Fathom's formula editor so that you can change what is computed.

34. Click on the plus mark to the left of the Special heading in the formula editor.

This should open up a list of "special" quantities that Fathom can compute.

30. Scroll-down the list you can see the *rowProportion* entry.

31. Double-click the *rowProportion* entry causing it to be entered into the formula pane.

32. Click the OK button at the bottom of the formula editor.

The summary table then reconfigures. Explain what the values in the cells represent.

### **Optional Enrichment Activities**

1. Make a graph showing the frequency of communication by email in other provinces by age of respondent and describe the pattern you see in this graph. Please provide a possible explanation for the shape in this distribution.
2. Make a graph showing the frequency of communication by email in another country by age of respondent and describe the pattern you see in this graph. Please provide a possible explanation for the shape in this distribution.

Originator: Sara Beattie, Statistics Canada  
File: Tech Use Lesson – Email use.doc  
Folder: MDM Treasure Trove/ Data / Tech Use  
Updated: May 3, 2004