Creating Tables

Tutorial

Time: 20 minutes
Skill level: Low
Editions: Reader, Basic, Professional

Learning objective(s)

⇒ Creating tables using questions

Activities

1. Click in the **File** menu and select **Open Existing Project …**
2. Select `Tutorial 1.Q` which will be in `c:\Program Files\Q\Examples` (unless Q has been installed in a different location on your computer). Your screen will now look like Figure 1.

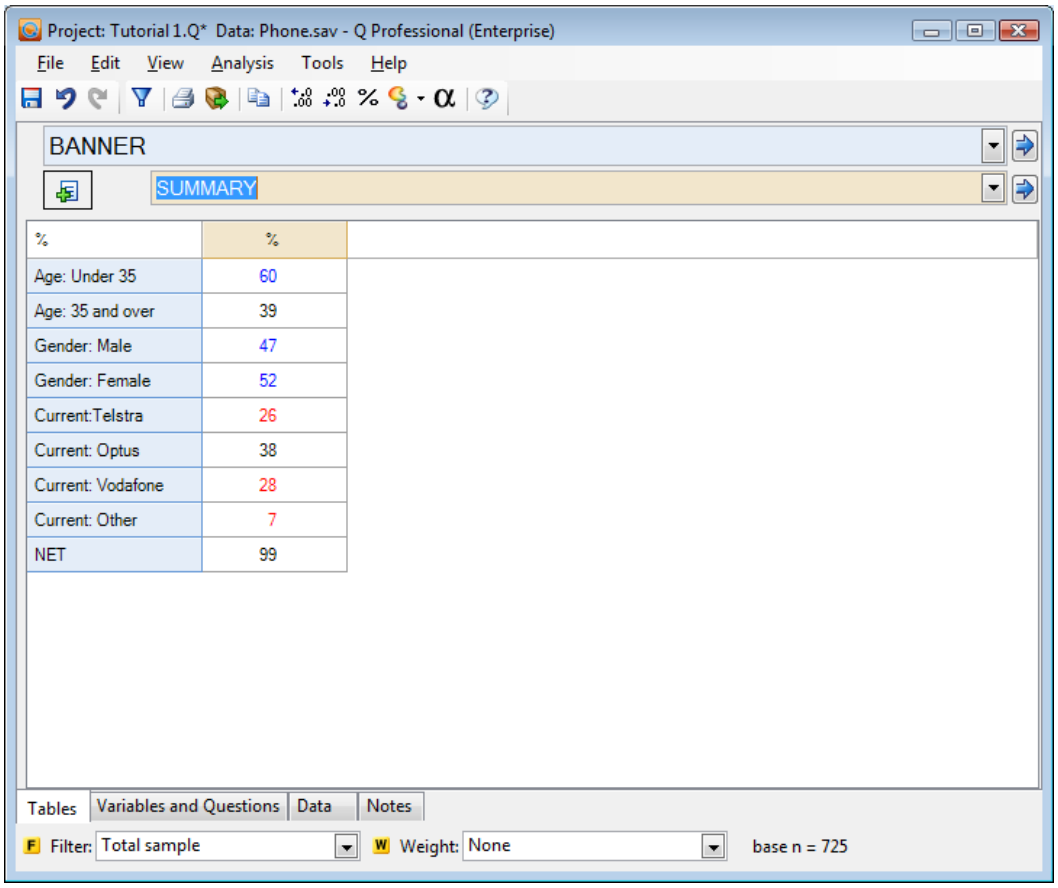

Figure 1. **SUMMARY** of BANNER

3. This project contains data on mobile phone ownership and usage. This table shows the percentage of the sample that fall into each of the categories. For example, 60% of the sample is aged less than 35 years and 26% are currently signed up to Telstra for their mobile phone usage.

4. If you add up the numbers in the table, they add up to more than 100%. This is because the categories are not mutually exclusive (e.g., somebody can be both male and under 35).
5. If you look at the NET at the bottom of the table, you will see that it is 99% (rather than the 100% one would more commonly see when a total is shown). This is the net of the categories in the questions. It does not add up to 100% because there are some respondents in the database who are not included in any of the categories in this question (due to having missing data on the age and gender questions).
6. Click on the blue drop-down and select **Q6. Aided Awareness**. The table shown is for aided awareness of mobile phone companies. This table shows the percentage of people who are aware of each of the mobile phone companies (e.g., 90% of people have heard of Optus).
7. Click on the brown drop-down and select **BANNER.** You have now created what is commonly known as a *crosstab* (see Figure 2). By default, this table is presenting column percentages (this is indicated on the top left of the table). That is, the table shows the percentage of people in each demographic group (columns) who are aware of the mobile phone companies (rows).

Q6. Aided Awareness

BANNER

Column %	Age		Gender		Current Telco				
	Under 35	35 and over	Male	Female	Telstra	Optus	Vodafone	Other	NET
AAPT/Cellular One	24	19	24	21	24	22	20	28	22
New Tel	11	6	8	10	6	11	10	14	9
One-tel	40	37	37	40	43	42	33	34	39
Optus	92	89	91	91	89	98	87	80	91
Orange (Hutchison)	55	51	55	52	51	58	45	70	53
Telstra (Mobile Net)	87	90	86	90	97	89	79	86	88
Virgin Mobile	46	22	40	33	31	40	35	46	36
Vodafone	91	69	86	80	70	82	99	72	83
Other 1	6	6	7	6	5	7	3	22	6
Other 2	0	1	1	1	1	0	0	2	1
Don't know	1	0	1	0	1	0	1	0	0
NET	100	99	100	100	100	100	100	100	100

Figure 2. Aided awareness by BANNER

8. The blue drop-down is used to select the question that will appear in the rows. The brown drop-down is used to determine how you wish to breakdown the results of that question. Through the brown drop-down, you can choose **RAW DATA** (i.e., the underlying data used to construct tables), a **SUMMARY** of the information (e.g., the percentages in each category) or select another question to see the interrelationship between two questions. If you have a two dimensional question (defined later), you will only get a choice between **RAW DATA** and **SUMMARY**. If your data is text (e.g., verbatim responses from an open-ended question), you will only have the option of viewing it as **RAW DATA** – see the tools for managing text variables described in the *Q Reference Manual*.

9. The red and blue coloring of the numbers can be used to quickly identify interesting results. For example, 46% of people aged less than 35 years are aware of Virgin Mobile, while only 22% of those aged 35 and older are aware of this brand. A similar age skew exists for Vodafone, where 91% of people aged less than 35 are aware of the brand versus 69% of those 35 and older.
10. Q automatically conducts tests of statistical significance on all the tables and highlights significant results. See the *Quick Start Manual* for more information.
11. From the blue drop-down select **Brand Health Grid**. Note that the brown drop-down automatically reverts back to **SUMMARY**. (This is because the brand health grid is a two dimensional question and cannot be crossed with another question.)
12. Click on the blue drop-down, type `Q2` and select **Q20. Image**. You will see a table with brand names in the rows and attributes in the columns. Using the red and blue font coloring as a guide, you can see that Optus and Telstra perform well on many dimensions including reliable, good coverage, conveniently located stores and meeting communication needs. Optus performs particularly well on being friendly, having the best phones and I like them, whereas Telstra is performing poorly on friendly, low prices and fashionable, suggesting an image overhaul might be beneficial in some areas.
13. Sometimes this table might take a while to appear because Q has performed a large number of computations. This table contains the tabulation of 180 different variables.
14. Select **Q24. Past-time hours spent** in the blue drop-down. This table shows the average number of hours people spent on various activities on weekdays and weekends.
15. Press the  button once to increase the number of decimal places shown. Your table should look like Figure 3.
16. Q automatically shows averages when a numeric question is selected in either, but not both of, the blue or brown drop-downs. Correlations are shown when numeric questions appear in both menus.

Average	Weekday hours spent	Weekend hours spent	SUM
making personal calls	.26	.51	.77
on home phone	.56	.80	1.36
watching TV	2.07	3.02	5.08
on PC internet access	1.64	1.60	3.24
accessing Internet from Mobile	.02	.01	.03
smoking	.19	.40	.59
reading newspaper	.36	.72	1.08
reading magazine	.25	.28	.53
listening to radio	1.66	1.55	3.20
SUM	7.00	8.89	15.88

Figure 3. Table: Numeric data

17. The first column shows the average number of hours spent undertaking each pastime on weekdays (i.e., 0.26 means 26% of an hour, or 15 minutes and 36 seconds). The second column shows the average hours spent on each pastime on the weekends. The last column shows the sum of the first two or, in other words, the total number of hours spent on each pastime.
18. From the blue drop-down select **Q7. Company currently with** and remove the decimal places using the  button. Note that there are lots of categories with very small percentages. Also note that the base n or sample size of 717 is shown at the bottom of the screen near the right (i.e., "base n = 717; total n = 725; 8 missing"). The "8 missing" indicates that amongst all the observations there are 8 with missing data for this table.
19. Now select **Q7. Company currently with (collapsed)**. This is a different version of the same question, but 3 people have been automatically filtered (people who said "Don't know" or who were not on contracts) and the smaller brands have been merged.
20. In the brown drop-down select **BANNER**.
21. As discussed in step 7, this table shows column percentages by default. However, this can be modified. Click your right mouse button anywhere on the table and click on **Statistics - Cells**. You should now see a range of statistics to select from (see Figure 4).

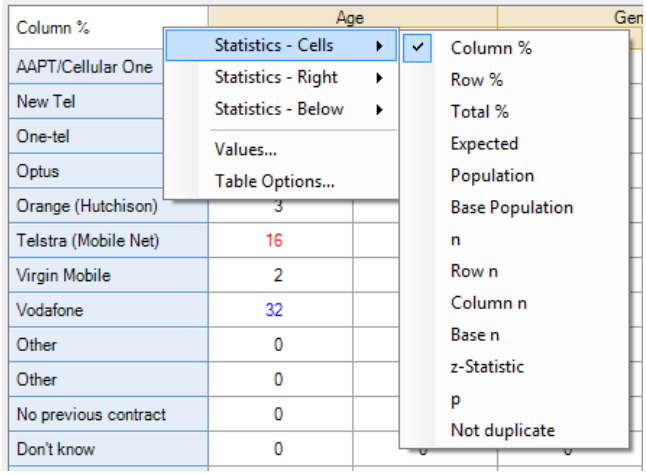

Figure 4. *Selecting Statistics*

22. Click on **n**. An additional row of data will be created which shows you the number of people in each cell (e.g., there are a total of 432 people in the sample aged under 35 and of these, 195 are Optus customers, which equates to 45% of all people aged under 35 in the sample).
23. Right clicking on the table allows you to select different statistics. The precise series of data available to be viewed depends upon the type of questions being viewed. When you are looking at more complicated tables, additional statistics will be available.
24. Press the % button at the top of the screen. Any numbers in the tables that are percentages will now be shown with a percentage sign (%) next to them.
25. By default, Q does not show the percentage sign next to percentages. This is because academics have worked out that the more text to appear on a screen, the less likely researchers are to identify interesting results.
26. Right click on the tables and deselect **n**. Right click again and select **Row %** from the **Statistics – Cells** options. While you learnt earlier that 45% of people aged under 35 were Optus customers, you can now see that 71% of the Optus customers in the sample are aged under 35 (i.e., 195/276).
27. In the blue drop-down select **Time (categorical)**. You will now see a table showing a variety of time-related measures.
28. Now change the question selected in the blue drop-down to **Time (in weeks)**. This shows the same data, but it has been turned into a numeric question to facilitate analysis. It shows that, for example, the average member of the sample had been with their current phone carrier for 111 weeks.
29. It is often useful to convert questions with ordered categories, such as **Time (categorical)**, into numeric questions, such as **Time (in weeks)**. It usually increases statistical power (i.e., makes you more likely to find significant differences) and permits the easy assessment of relationships with other questions

in the survey, as is illustrated in the next step.

30. In the brown drop-down select **BANNER**. As one of the questions is numeric and one categorical, Q presents the information in terms of an average, permitting comparison of averages by the groupings of the categorical question. You can see that, for example, people aged under 35 have been with their carrier for an average of 94 weeks, while people aged 35 and over have been with their carriers for 139 weeks on average.
31. When one variable is numeric, the color-coded significance tests highlight differences between the categories of the categorical question. That is, the color-coding tells us that older people have significantly longer relationships with their current supplier than younger people, and Telstra customers have longer relationships with Telstra than other customers have with their respective phone companies.
32. If you look at the bottom of the screen you will note that it shows that the **base n** varies from 371 to 716. Right click on the table and select **Statistics – Right** and then select **Base n**. Note that the sample sizes differ – this is because some questions, such as time until contract expires, were only asked for a subset of the sample.
33. Go to **Statistics – Cells** and select **Column n**. This shows us the sample sizes actually used in the computation of the averages. That is, the data file contains data on "With current company" for 714 people all up, but 434 of these people are in the Under 35 age category and have been used to compute the average for people aged under 35 (278 people are aged 35 and over and there are 2 people for whom we have data on "With current company" but did not provide their age).
34. Now select **Time (in weeks)** in the brown drop-down (i.e., it should be selected in both) and add a decimal place. If it is not already showing, go to **Statistics – Cells** to select **Correlation**. You have just created a *correlation* matrix*,* which is perhaps the most difficult to interpret of all of Q's outputs (To facilitate interpretation, deselect **Base n**).
35. Each number in the table is a *correlation*, which is an index in the range of -1 to 1. A value of 0 indicates that there is no relationship between two variables. A value of less than 0 – indicates a *negative correlation* (i.e., on average, higher values of one variable equate to lower values of the other and vice versa). A value of more than 0 indicates a *positive correlation* – higher values of one variable on average equate to higher values of another. The further the correlation from 0, the stronger the relationship between the variables.
36. Looking at this table reveals that there is a weak negative correlation between time with current company and time until current contract expires (i.e., not surprisingly, on average, the longer one has been with one's current carrier, the sooner the contract will expire), while the longer people have had a phone the longer they

have been with their current carrier (i.e., there is a moderate positive correlation between "With current company" and "Since first phone").