

How To CODE SURVEY DATA

Support material to the
INSIGHT DISCOVERY REPORT[©]
PROFESSIONAL SURVEY DATA ANALYSIS
EVEN STUDENTS CAN AFFORD

Document: *How To Code Survey Data*
Product: Insight Discovery Report®
Date: September 30, 2017

Author: **MarketingStat GmbH**
+41 (0)61 401 60 55
info@MarketingStat.com

Disclaimer

MarketingStat GmbH abides by the ICC/ESOMAR International Code on Market and Social Research.
ESOMAR: European Society for Opinion and Market Research.
ICC: International Chamber of Commerce.

MarketingStat GmbH is Developer Member of the Association of Software Programmers, ASP

Organizing Survey Data

There are three steps to put the data collected with an online survey in the right shape to be analysed and produce the *Insight Discovery Report*:

1. Organize close-end and open-end questions
2. Create the Frame Code for each question (aka variable)
3. Convert respondent answers to codes

Tip: Store a copy of the original data before you begin putting the survey data in shape for the analysis.

Organizing close-end questions

From a list of possible answers close-end questions allow the respondent to choose one single answer. These answers must be organized in a single column with the question label in the first cell of the column.

The dataset you export from your online survey provider may look like the following image.

| Original Dataset | | | Q10 - Where do you find information about your work? | | | | | | | |
|------------------|--|--|--|-----|-----|-----|-----|-----|-----|-----|
| N. | Q2 - How many years have you been working in water supply? | Q13 - Are you a certified master of fountains? | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 |
| 1 | <1 year | Yes | | | | | | | | |
| 2 | 36 | No | | | 1 | 1 | | | 1 | 1 |
| 3 | 24 | Yes | | | 1 | 1 | | | | 1 |
| 4 | 4 | Yes | | | | | | | | |
| 5 | 12 | Yes | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 |
| 6 | 28 | No | 1 | | 1 | 1 | | | 1 | |
| 7 | 10 | No | | | | | | | | |
| 8 | 6 | Yes | | | | 1 | 1 | | | 1 |
| 9 | 5 | Yes | | | 1 | 1 | 1 | | | 1 |
| ... | ... | No | ... | ... | ... | ... | ... | ... | ... | ... |
| 382 | 11 | No answer | | | 1 | | | | | 1 |
| 383 | 3 | Yes | | | | | | | | |
| 384 | 4 | No answer | | | | | | | | |

Q2 and Q13 are close-end questions while Q10 is an open-end one. Both question types must be coded in a form the elaboration software can transform into meaningful contingency tables (crosstabs). To do so we create a *Frame Code*.

The Frame Code

The Frame Code instructs the software on how to handle each answer to each question. It looks like this:

| Code | Q2 - How many years have you been working in water supply? | Q13 - Are you a certified master of fountains? | Q10 - Where do you find information about your job? |
|------|--|--|---|
| 1 | Up to 10 years | Yes | Magazines |
| 2 | 11 to 20 years | No | Online |
| 3 | 21 to 30 years | No answer | Association |
| 4 | Over 30 years | | Fairs / Continuing Education |
| 5 | | | Flyers |
| 6 | | | Mailings |
| 7 | | | Supplier / Dealer |
| 8 | | | Colleagues |

It's all quite simple. Take Q2 for example. If the first interviewee's answer to Q2 is "Less than 1 year", according to the frame code this answer is substituted with a code "1" which covers all answers up to 10 years. The second interviewee's answer for Q2 is coded with a "4" and so on. The coded dataset looks like in the following image.

| N. | Q2 - How many years have you been working in water supply? | Q13 - Are you a certified master of fountains? | Q10 - Where do you find information about your job? | | | | | | | | |
|-----|--|--|---|-----|-----|-----|-----|-----|-----|-----|-----|
| | | | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | |
| 1 | 1 | 1 | | | | 1 | 1 | | | 1 | 1 |
| 2 | 4 | 2 | | | | 1 | 1 | | | | 1 |
| 3 | 3 | 1 | | | | | | | | | |
| 4 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 |
| 5 | 2 | 1 | 1 | 1 | 1 | | | | | | |
| 6 | 3 | 2 | 1 | 1 | 1 | | | | | | |
| 7 | 1 | 2 | | | | 1 | 1 | | | | 1 |
| 8 | 1 | 1 | | | | | | | | 1 | 1 |
| 9 | 1 | 1 | 1 | 1 | 1 | | | | | | |
| ... | ... | ... | ... | ... | ... | ... | ... | ... | ... | ... | ... |
| 382 | 2 | 3 | 1 | | | | | | | | |
| 383 | 1 | 1 | | | | | | | | | |
| 384 | 1 | 3 | | | | 1 | | | | | |

Organizing open-end questions

The treatment of open-end questions does not differ much from how we treated the close-end ones. The only difference is that an open-end question is coded over as many columns as codes. For instance, Q10 in the original dataset above is an open question with 8 codes. Each code takes a column. The first cell of each column above the respondent answers is occupied by the code number and above it there is the question label. Leave blank cells where there is no answer (make sure it is really blank. Pay attention, certain characters are invisible).

Coding answers

Consider using the Excel function VLOOKUP to turn close-end respondent answers into frame codes. This is possible in most cases but not always. Sorting data and using filters helps make the coding process less tedious.

**You don't have time to put your survey data in shape?
MarketingStat can do it for you.**