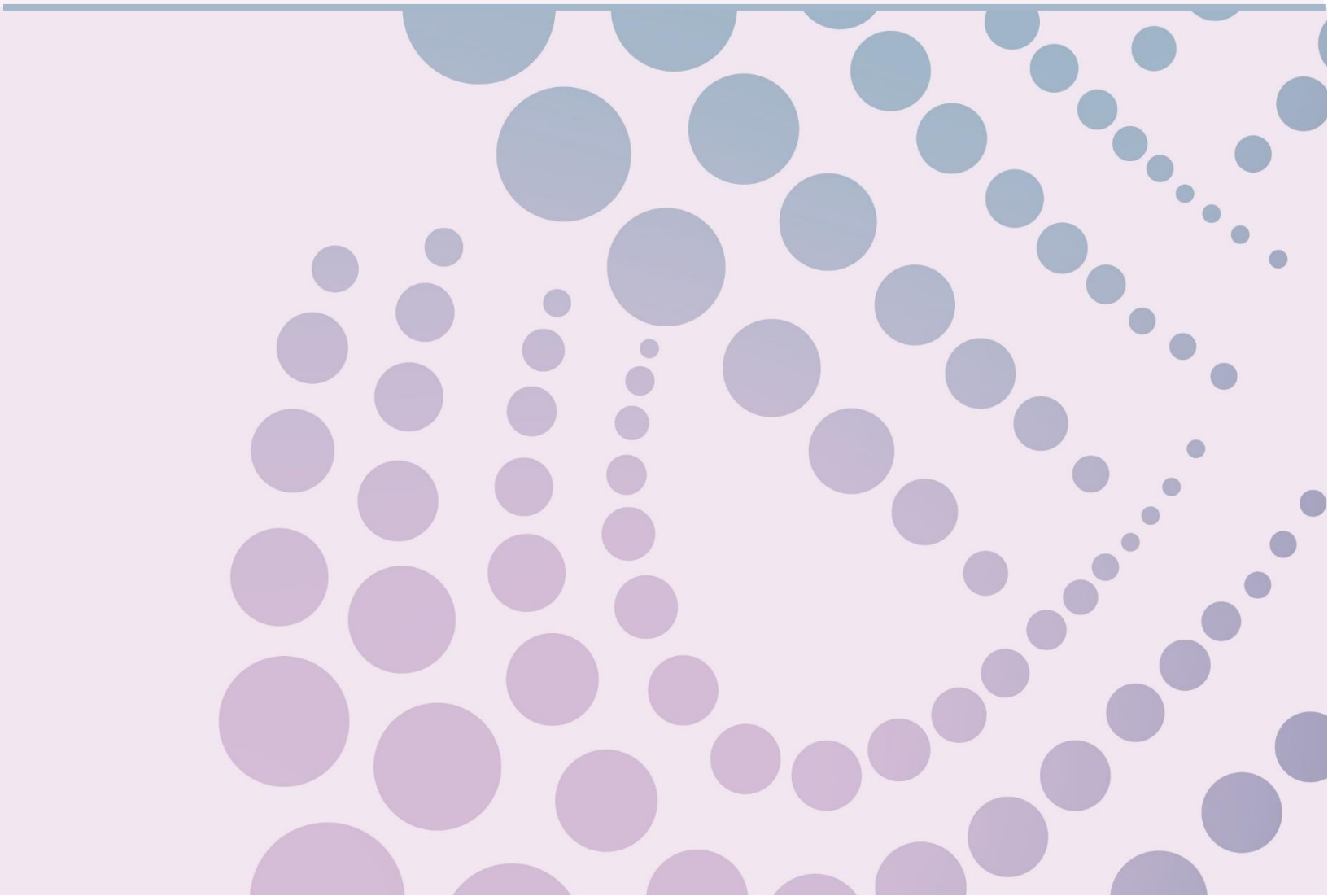
Using Surveys to Evaluate Impact

REFERENCE MANUAL

LAURIE REEDMAN



Overview

This reference document was created to familiarise the reader with the steps in doing an online survey to evaluate the impact of their work while placing minimum response burden on respondents, as well as how to analyse and interpret the survey responses. It is not a complete guide to survey methodology or data analysis. Included here are tips and pointers, and hyperlinks to reference websites. Examples are based on real data from an undisclosed post-event feedback survey. The author hopes you find this reference manual helpful.

About the author

Laurie Reedman earned an honours Bachelor of Mathematics, majoring in statistics at the University of Waterloo in Canada. During her 30-year career at Statistics Canada as a Methodologist, Laurie served as the Chief of the Quality Secretariat for over a decade. Under Laurie's leadership the Quality Secretariat undertook numerous initiatives aimed at improving the quality of statistical products and the processes which produce them. Laurie was enthusiastically involved internationally, regularly presenting at conferences and leading outreach and capacity building activities at the national statistics offices of several Caribbean and African countries. As Assistant Director of the Methodology Innovation Centre, Laurie led an interdisciplinary team to create a Framework and Guidelines for responsible machine learning processes at Statistics Canada, and she served as the Methodology Branch lead on the Data Literacy Initiative, whose objective was to produce self-serve training material on data literacy topics.

Queen Mary University of London

TABLE OF CONTENTS

OBJECTIVES AND DELIVERABLES

Being very clear about why you are doing the survey, what you are trying to learn from respondents, and what the final results should look like will be instrumental in designing the survey questions. The following list includes what you should be thinking about before beginning to develop the survey questions.

- Make clear objectives about what you want to measure
  - What story do you want to tell?
  - What have you been mandated to measure?
  - Is it a one-time survey or will it be repeated?
  - How do you want the results segmented?
    - Given your target population of respondents, what are the domains of interest and do you reasonably expect enough responses in each one to facilitate doing comparisons; be realistic
- Draft the summary tables and charts that will appear in your deliverables
  - At this point there will be no data but making titles and headings for summary tables will help you clarify your objectives
- Set metrics
  - In your analysis – what levels are ok, what levels mean you need to react somehow, do something
- Set benchmarks
  - If doing repeated surveys, which metrics do you want to track through time
  - How do you expect metrics to behave over time, how will you know when things are ok versus when to react
- Determine the format for deliverables, for example:
  - Report
  - Data visualization, infographic, static or interactive
  - Presentation
  - Blog
- Draft an outline or skeleton of all deliverables, indicating where your findings and analytical results will fit in

Reference Websites:
All of these websites offer advice on how to set objectives. The first two suppose private sector business objectives but the theory is applicable to public sector.
Survey Monkey
Customer Thermometer
Council of Europe, Venice Commission

REPEATED SURVEYS

One objective of a post-service delivery survey is often to find out what your participants gained or learned, or what impact or effect the service or program had on their subsequent activities. This is a comparison between what they knew or were doing before and after the training or program implementation. To do this, you would survey your participants with the same set of questions before and after the intervention. This doesn't have to be all the questions in the questionnaire. At the beginning you are only looking for baseline information about the topics where you are looking for an increase in knowledge or for activities that were undertaken as a result of the intervention.

One way to do this is to ask on a scale of 1 to 5, with 1 being low and 5 being high, how knowledgeable participants think they are about each topic you are interested in, or how confident they feel about their ability to do each activity you are interested in. Administer these questions before and after the training. Then do a one tailed t-test (see the section T-test for difference of means in this reference manual) to see if there is a statistically significant increase in self-reported knowledge/confidence level after the training. One can only say that a difference between two values is statistically significant if a hypothesis test has been performed, such as a Chi-squared test (see the section Chi-squared hypothesis test in this reference manual). Two values could be different, but the difference might not be statistically significant but rather the result of random fluctuations.

If measuring increase in knowledge or abilities is one of your objectives, you need to survey at multiple points in time with the same questions.
What if you intend for your participants to gain knowledge or abilities, but your analysis reveals that this didn't happen, or not to the extent you intended? At what point should action be taken? Think ahead to what signals you need to see for post-service-delivery decision making, and then ensure that your survey questions will provide the data needed to fuel those signals.

CRAFTING THE QUESTIONS

These are “rule of thumb” recommendations for any type of survey, whether the goal is to seek feedback, level of satisfaction, or impact.

- Start with an overall satisfaction question
- Followed by more detailed questions getting deeper into what you are interested in learning from respondents
- Ask demographic and sensitive questions towards the end, to limit the risk of people not answering the whole questionnaire if they are not comfortable answering these questions. While you might be tempted to put these questions at the start of the survey or to make them ‘mandatory’ questions that mean the respondent cannot submit the survey until the question is answered, focus group testing at Statistics Canada found that if you make sensitive questions mandatory you risk people refusing to do the whole survey. But if they’ve gone along answering everything else then they are more likely to continue asking whatever you ask them
    - Offer a “prefer not to say” response for sensitive questions
- Use wording that has already been tested or used previously that will be well understood
- Spell out acronyms and define technical terms, or simply don’t use them
- Avoid double-barreled questions (asking two questions in one)
- Keep it short – 5-10 minutes max for respondents to complete
    - Can’t give you a magic number of questions here – it depends on how long it takes to answer your questions.
    - Test your questions on your colleagues and ideally people who represent your target participants to make sure the questions are well understood and answerable
- Don’t ask anything that is not part of your objectives
    - Don’t waste people’s time on questions you do not intend to use
    - Think about what answers you want, and then turn that into questions
- Offer a “none” or “N/A” option where appropriate
- Remember to say “thank-you” at the end of the survey

Jisc and Microsoft Forms have built-in question formats, making it easy to build a survey and facilitating basic analysis and visualization. The types of questions are:

Choice questions

- Respondent chooses one or several responses from a list
    - Use key words, and as much as possible, put key words at the front of each category or item in the list

    - You can allow only one response (yes/no) or multiple (tick all that apply)
    - Option to add other/please specify where respondent can type free text

Scale, rank or rating questions

- Use familiar scales, for example Likert 1-5
    - Be consistent with the direction of the scale (for example on the Likert scale, 1 is a strong negative and 5 is a strong positive)
- Jisc calls this a “grid” question

Date and time questions

- Respondent types digits or chooses a date from a calendar
    - Format is imposed

Free text questions

- Respondent types whatever they want
    - Spontaneous or freely expressed responses can give much more valuable data than pre-defined responses. However free text is more difficult to summarize and analyze. Use free text questions judiciously.

Logic in survey questions refers to allowing different sequencing of questions depending on the answers. For example, if someone did not participate in a particular activity then they would skip any questions related to that activity. When you build surveys in both Jisc and Forms you have the ability to build in logic of skip patters that automatically move respondents through the correct sequence of questions.

Additional things Jisc can do:

- Show progress bar
- Pre-populate information you already have about respondents (for say a repeated survey)
- Screen people out (trespassers, tourists, legitimately out of scope respondents) depending on their responses to certain questions
- Download as CSV (Excel)
    - Uncoded – you get the actual text in Excel cells, comma delimited where multiple responses are allowed; needs parsing and coding to do analysis
    - Coded – responses are turned into indicator variables where multiple responses are allowed; ordinal where only one response is allowed; ready for numerical analysis

Things that are challenging in Jisc

- Date and time – tricky to get it to set the time, you can’t just type it in
- How to add “other, please specify” is not obvious

    - I think you have to give the response option “other” and then use logic, if they click that one, then provide a free text question for them to write in their answer
- Two pages of the help are not the same, some has more info (I think there was a new version recently, things look a little different in YouTube videos)

Recall that if you want to measure knowledge gained or impact of training or a service, you will need to ask the same questions at the start and at the end of the training or service. Not the entire questionnaire, just questions measuring knowledge or impact.

Reference Websites:
Pew Research, Writing Survey Questions
Zapier, Writing Effective Surveys

DATA COLLECTION

Things to plan and execute:

- How to contact your intended respondents – do you have email addresses
- Dates and duration of collection period, how long will you leave the questionnaire “active”
- Draft the invitation to do the survey
    - Why their response matters to you
    - What’s in it for them, incentive
    - Reassurance about security, anonymity (if applicable)
    - Timeframe
- Decide on follow-up measures
    - Will you send reminders, at what time intervals
    - What response rate is acceptable, overall or by domains of interest
- Before the actual collection period, test that the platform and questionnaire function as expected
- Consider doing the survey in real time, while participants are still engaged

DATA ANALYTICS

Once the data is collected, there are four major steps to making your findings understandable, consumable, and reliable. The steps are:

1. Data cleaning
2. Data analysis
3. Data visualization
4. Data interpretation

The activities in each step will be described below, along with recommended software products.

DATA CLEANING

Assume that all data starts out “dirty”; in other words, having errors or gaps. Despite your best efforts at crafting the questions and designing the survey, something may still need cleaning. If you go straight into analysis without first looking at the data, you risk arriving at false conclusions. Data cleaning is making the data say what the survey respondents meant it to say. It is not intended to alter or hide the truth; rather the opposite, it is intended to reveal the truth that might have been obscured by typos, misunderstandings, or a flaw in the questionnaire logic.

That said, online survey tools such as Jisc and Microsoft Forms do a great job of providing templates for your survey that go a long way towards ensuring that the data is clean. For example, built-in logic or skip patterns ensure that respondents are presented with the correct sequence of questions depending on their answers, and you can avoid having missing values by specifying which questions require answers before the respondent can “submit”.

A good practice for the data cleaning step when using Jisc and Forms would be to look at the summary graphs and charts of the responses and see if they make sense.

Another purpose for data cleaning is to organize or format the data so that you can do data analysis on it. For example, creating indicator or dummy variables, or parsing out key words from strings of text.
Data cleaning can be done using a programming language, a spreadsheet, or with a software product designed specifically to analyze survey data. Here’s what I recommend:

- If you have software such as SPSS, SAS, SPlus, R, Python, Power BI or Tableau, and you like using it, go ahead and use that. Save your programs so you can use them again or at least use parts of them again.
- Excel can handle everything you need to do for data cleaning, and if you save sequences of actions in macros, you can use them again.

The examples in this reference manual use Excel.

WHAT TO DO WITH FREE TEXT RESPONSES

If you have write-in responses, what analysis will you want to do with these? Are there few enough that you can simply read them and report them as anecdotal information? Or will you want to summarize them, by hand or with automation?

Both Jisc and Forms provide a list of the free text responses in their automatically generated analysis. You can scan through the responses and pick out key words. To better understand what people are trying to tell you, export to Excel or copy/paste the free text responses to a word processing software and work there. Use the search function to find instances of key words or phrases, and count how often they occurred. If you have access to more sophisticated tools such as machine learning algorithms, you can use sentiment analysis to interpret and summarize the free text. If you have qualitative research skills, performing a thematic analysis of the textual responses would also be appropriate.

Important: by including free text questions, you're asking respondents to slow down, think about your question, and type a response. In return, it's your responsibility to gather, interpret and reflect their thoughts in the story you tell from this data. Be prepared to give free text data some extra attention.

PREPARING MULTIPLE RESPONSE DATA FOR ANALYSIS

Everyone loves "tick all that apply" questions. But dealing with the answers is not trivial. Most survey software exports multiple responses into a single cell with a comma between each response (comma separated values, or CSV format). Jisc goes one step further and has the option to export survey data in a coded CSV format, meaning the responses are coded to numerical values, and answers to multiple response questions are separated out into unique cells. Here let's look at an example for preparing comma separated values for doing analysis.

Example 1: Excel SEARCH with ISNUMBER and SUMPRODUCT functions

Q13. Which elements of the training did you find most useful? Tick all that apply. Response categories include Lectures, Case studies, Slides from sessions, Q&A, and a few others. The survey data was exported from Jisc into Excel as CSV but not coded, and the first few responses from Q13 looked like this:

Q13
Lectures
Lectures, Case studies
Case studies, Slides from sessions, Q&A
Lectures, Q&A

Probably you asked this question because you might change or drop the training elements that participants found *least* useful. So you need to count all the responses to know which ones were most and least useful. But how do you do that with the data like this? You could do it by hand, but that will be very cumbersome if you have more than a handful of respondents. What you really need here is indicator variables, sometimes called dummy variables, one for each possible response, with the value 1 if the person ticked that response, and the value 0 if they didn't. You want your data to look like this:

Lectures	Case studies	Slides from sessions	Q&A
1	0	0	0
1	1	0	0
0	1	1	1
1	0	0	1

(Aside: Jisc data exported as coded CSV looks like this.)

Now a simple tabulation reveals that lectures were cited 3 times, Slides from sessions was cited only once, and Case studies and Q&A were tied with 2 citations each. You could use a histogram (frequency graph) to visualize this data. You could do a Chi squared test to see if the difference in popularity between Lectures and Slides from sessions is statistically significant. The data cleaning step is transforming the raw data from multiple response questions into indicator variables that can be fed into data analysis or turned into data visualizations.

In Excel there's an elegant way to go right from having multiple responses all in a single cell, to having indicator variables for each possible response.

The Excel function to do this uses the SEARCH to search for text in a phrase, together with the ISNUMBER function that turns the response from SEARCH (a number if the text is found, an error if it is not) into the value TRUE if the text is found, and the value FALSE if it is not. It looks like is:

=ISNUMBER(SEARCH("text you are looking for",cell with multiple responses))

Watch this video to see how to use this function. Then watch this video, which tells you how to use the SUMPRODUCT function with a double negative to turn the TRUE values into 1s and the FALSE values into 0s.

Then you put all three functions together into one statement:
=SUMPRODUCT(-- (ISNUMBER(SEARCH("text you are looking for",cell with multiple responses))),cell with the value 1)

So all you have to do is:

1. To the right of your column of multiple responses (let's say they're in column A), make a column of all 1s. So column B, all the way down, put a 1 for every row where you have data in column A.

2. Now in the next few columns to the right, label one column per possible response with the actual text of the response, or a shortened version of it. These are your indicator variables. In this example your indicator variable column labels are "Lectures", "Case studies", "Slides from sessions", and "Q&A".

3. In the first cell of the first indicator variable (cell C2 in this example) type the formula =SUMPRODUCT(-- (ISNUMBER(SEARCH("Lectures",$A2))),$B2). If you typed it correctly it should put a 1 in cell C2 if Lectures were listed in cell A2, and a 0 otherwise.

4. Once you get the formula correct in cell C2, copy/paste the same formula all the way down column C. Check that the 1s and 0s are in the right rows.

5. Copy the formula from cell C2 to cell D2. Now you are in the indicator variable for Case studies. Edit the formula to search for "Case studies" instead of "Lectures". Check that the other parameters are still $A2 and $B2. We use the dollar sign so they won't change. Check that the formula is working in cell D2, then copy it all the way down column D.

6. Repeat until you have all your indicator variables. Now you're finished cleaning the multiple response data. It's ready for analysis.

DATA ANALYSIS

Typical summary statistics

- Mean, median, mode
- Frequency distribution

- Cross tabulation
- Test of hypotheses (Chi squared test)
- Test of significant difference in the mean of two groups (t-test)

Data visualisations

- Table
- Histogram
- Bar chart
- Line chart
- Pie chart

Jisc and Forms can do some basic analysis as soon as you start getting responses. You can see frequency displayed in bar charts, pie charts and tables. You can scroll through free-text responses. That might be fine but probably you'll want to do more analysis. You can export from Jisc into PDF or into a format that can be easily used in Excel or in a statistical analysis software such as SPSS. Forms has a button taking you directly to the data in Excel.

Data analysis software and tools

Any statistical software will do all the summary statistics you could possibly want, and a whole lot more. Statistical software is not typically the best choice for data visualizations. Examples are SAS, SPSS, SPlus, R. If you have one, use it; if you don't, no worries.

There are software packages marketed specifically as survey data analysis tools. I haven't tried any but you could check them out. A couple I found are MAXQDA which is not free but is specifically for analyzing questionnaire data, particularly qualitative and mixed mode data, and Qualtrics which is designed for survey data analysis. If you don't have budget or resources to learn a new software, no worries.

There is nothing you can't do in Excel. You can do all of the data cleaning, summary statistics and even the visualizations in Excel, although sometimes data visualization in Excel is cumbersome.

If you want something less cumbersome than Excel for data visualizations and even some of the data manipulation, then I recommend investigating a business intelligence software such as Power BI or Tableau. These two are industry leaders. They are also a great choice for program monitoring and evaluation. Power BI desktop can easily import data from Excel and is great for data visualizations and dashboards.

Possibly the best option for statistical analysis is Minitab. It looks a lot like Excel but is more intuitive and easier for doing analysis and visualizations. You can find Minitab tutorials on Youtube.

All the following examples in this section have been performed in Excel.

MEAN, MEDIAN AND MODE

Mean is the average – the total of all the values, divided by the number of values.

Median is the middle value – put all values in order smallest to largest, and the median is the one in the middle. If you have outliers or extreme values in your data, then median is more representative than the mean because it will stay in the middle, unlike the mean which will get pulled or biased towards extreme values.
Mode is the value that occurs most frequently.

The first three and a half minutes of this video give a good refresher.

Excel has functions for each of them. The one for mean is called AVERAGE; the one for median is called MEDIAN, and for mode, you can specify that your data is single mode (MODE.SNGL) or multi modal (MODE.MULT).

FREQUENCY DISTRIBUTIONS AND HISTOGRAMS

Numerical data is data that is numbers, such as counts of things, temperature, measurements. Categorical data is names of things, such as which workshops you've attended in the past, employment status, educational background. When you have categorical data, a nice way to summarize it is to count the instances of the different categories, which is called a frequency distribution. If you graph it, it's called a histogram. Power BI and Minitab can make histograms quite easily. Excel can make histograms, but it takes two steps: first you make a table with the count (frequency) of each category, then you make a bar chart from the counts.

Example 2: Histogram in Excel from indicator variables

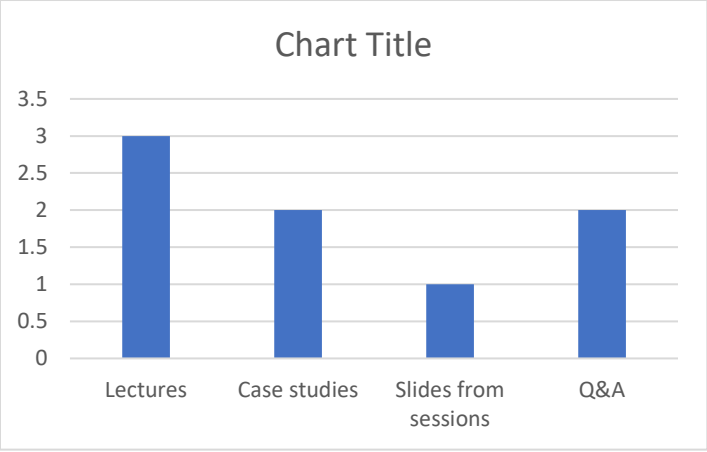
Let's use Q13 again, which elements of the training the participants found most useful, the "tick all that apply" data that we turned into indicator variables in example 1.

Lectures	Case studies	Slides from sessions	Q&A
1	0	0	0
1	1	0	0
0	1	1	1
1	0	0	1

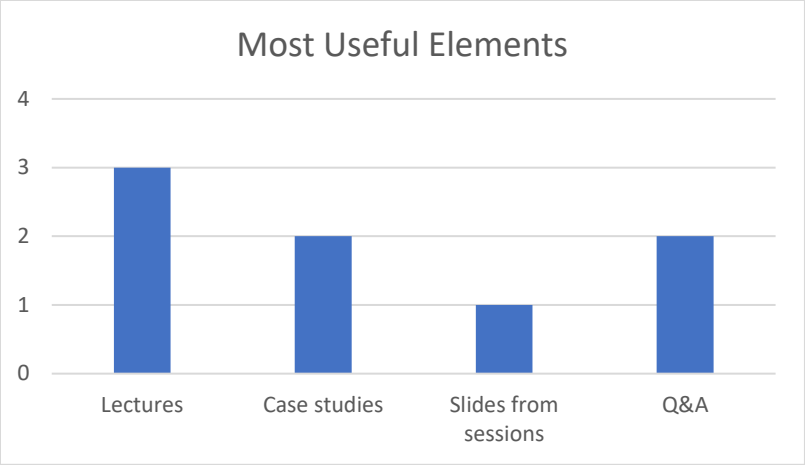
In Excel, insert a blank row between the row with the column labels and the first row of 1s and 0s. In this new row, calculate the total of the 1s and 0s in the column with a function such as =SUM(A3:A6)
The result will look like this:

Lectures	Case studies	Slides from sessions	Q&A
3	**2**	**1**	**2**
1	0	0	0
1	1	0	0
0	1	1	1
1	0	0	1

Now select just the two top rows, the column labels and the sums, and select insert and a bar chart, to get this:



Edit the title to something appropriate. Right click the vertical axis then click Format axis and change Major units to 1 (or some other whole number that makes sense given the highest count in your data) and change the Minor units to blank. The result in this example will look like this:



Example 3: Histogram in Excel from categorical data

Q5. Which of the following best describes your core activities at the time you attended the Masterclass training?

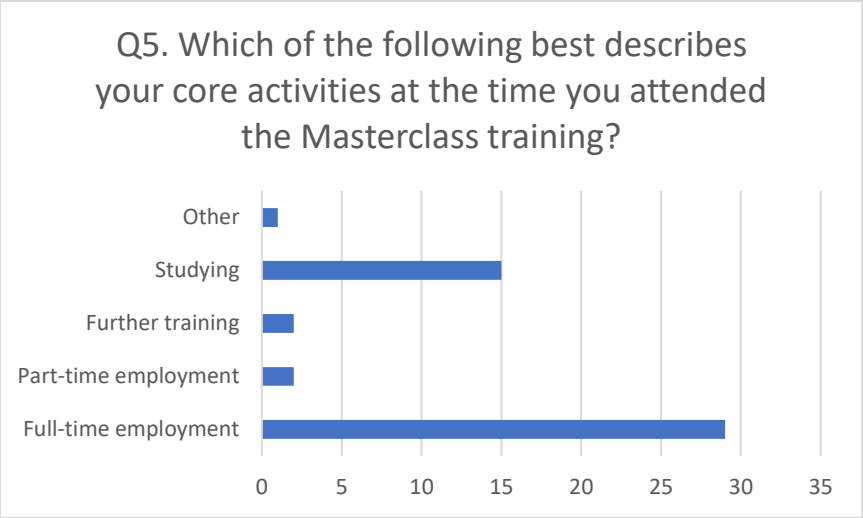
Respondents had to choose one of the following activities:

- Studying
- Part-time employment
- Full-time employment
- Further training
- Other

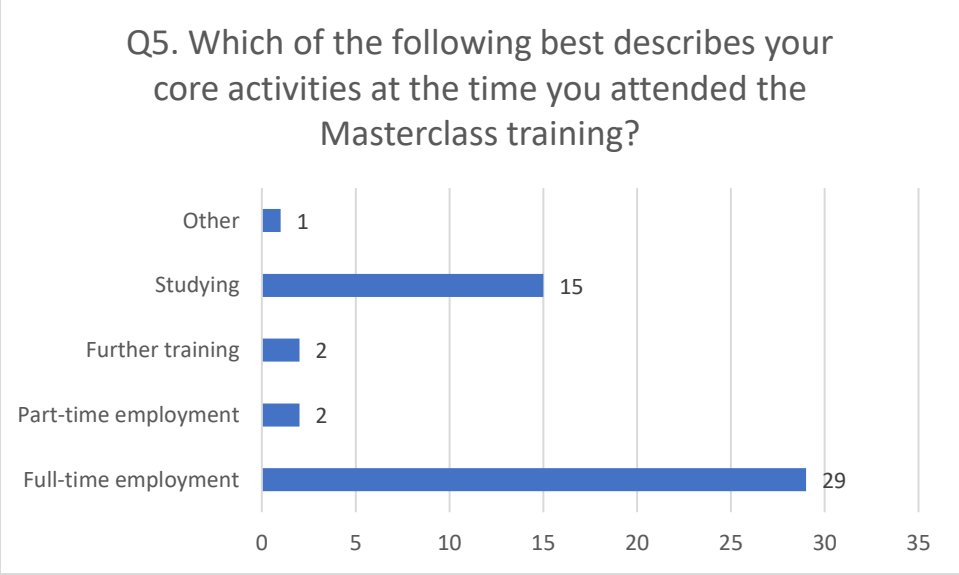
Because this is a “choose only one” type of question, not tick all that apply, we don’t need to make indicator variables, we simply need to count the number of occurrences of each activity. We can do this in Excel with a pivot table.

Highlight the column of responses and on the Insert tab, click Pivot table. Choose to place the pivot table in a new worksheet. In the pivot table fields dialogue box, click the column label, which is the text of Question 5. You will see “Count of 5. Which of the …” in the Sum Values box (bottom right). That’s good, you want it there, and you want it to say Count of, not Sum of. You ALSO need it in the Columns box, so grab it from the field name box and drag it to the Columns box. Now the pivot table should have all the activities as column labels, and all the counts. This is exactly the same type of table that you created from your indicator variables, by adding the sums of the columns directly underneath the column labels. Close the pivot table fields dialogue box. It would be great if you could make your bar chart directly from the pivot table, but unfortunately you can’t. Excel knows that there are calculations in behind the pivot table and it does not want to graph that. So highlight the pivot table and copy it, then click somewhere else

on the sheet and click paste special, and choose “values”. So you just made an identical table but with just the values, not the pivot table structure. Now you have an opportunity to edit the column headings if you like (remove “count of” for example) and put the columns in whatever order you want. Now highlight your nice table and on the insert tab choose bar chart.



Some of the labels are long so I chose a horizontal bar chart this time. You can click on any one of the bars, then right click, and choose Add data labels, to get the exact count on each bar.



CROSS TABULATIONS

A cross tabulation is just a 2-way table.

Example 4: Cross tab or 2-way table in Excel

For this example let's use Q4. Are you from an under-represented group, and Q18. How inclusive and diverse did you find the training. An Excel pivot table is a great way to get a 2-way table of these variables.

	4. Are you from an under-represented group?			
18. How inclusive was the training	No	Yes	Prefer not to say	Grand Total
Not at all inclusive	0	2		2
Somewhat inclusive	8	12	1	21
Very inclusive	11	11		22
Grand Total	19	25	1	45

Here is an awesome tutorial on using Excel pivot tables.

CHI SQUARED HYPOTHESIS TEST

The Chi-square test of independence determines whether there is a statistically significant relationship between categorical variables. It is a hypothesis test that answers the question – do values of one categorical variable depend on the value of other categorical variables?

Example 5: Chi squared test in Excel

Let's use the cross tabulation of questions 4 and 8 from example 4.

In a hypothesis test, the null hypothesis is that there is no relationship between two variables, so knowing the values of one variable does not help you predict the values of the other variable. We calculate the test statistic at a certain significance level, alpha, usually 0.05. If the p-value of our test statistic is less than alpha then we reject the null hypothesis and conclude that there is indeed a relationship. Conversely, if the p-value of our test statistic is greater than alpha then we do not reject the null hypothesis and therefore conclude that there is no relationship.

In example 4 we had 19 people who were from an under-represented group, 25 who were not, and 1 person who preferred not to say. For simplicity, let's remove that one person from the table and do the Chi squared test to see if there is a difference between participants from an under-represented group and those not from an under-represented group, in how inclusive they found the training to be.

We start with the same cross tab but this time, without the person who preferred not to say. This is the observed distribution of responses.

	4. Are you from an under-represented group?		
18. How inclusive was the training	No	Yes	Grand Total
Not at all inclusive	0	2	2
Somewhat inclusive	8	12	20
Very inclusive	11	11	22
Grand Total	19	25	44

Our assumption, or the null hypothesis, is whether or not participants were from an under-represented group had no bearing on how inclusive they found the training to be. What we want to test is if the distribution of responses to the question "How inclusive did you find the training?" that we observed is the same as the distribution we would expect if it made no difference whether they were from an under-represented group or not (or in other words if the null hypothesis were true).

To get started we need the percentages of total number of people who gave the 3 different responses to question 18. 2 out of 44 or 4.5% of respondents said "Not at all inclusive", 20 out of 44 or 45% of respondents said "Somewhat inclusive", and 22 out of 44 or 50% of respondents said "Very inclusive". These percentages are shown in the leftmost column in the table below. Next, we need to predict how many people would have given each of the 3 possible responses to question 18, if it didn't matter whether they were from an under-represented group or not. To do this, we multiply the percentages we just calculated by the total number of respondents who reported not being from an under-represented group (19 people) and from an under-reported group (25 people).

The predicted counts are shown below:

Q18 responses	Distribution of Q18 responses, all respondents	Predicted count of No's	Predicted count of Yes's
Not at all inclusive	0.045454545	0.86363636	1.13636364
Somewhat inclusive	0.454545455	8.63636364	11.36363636
Very inclusive	0.5	9.50000000	12.50000000

Even though we're talking about counts of people, we keep all the decimal places in this calculation. What this table tells us is that if it didn't matter whether people were from an under-represented group or not, 0.86 of those who were not from an under-represented group would have said that the training was not at all inclusive, and 1.14 of those who were from an under-represented group would have said that the training was not at all inclusive, and so on.

Next we use a built-in function in Excel called CHITEST to calculate the test statistic we need to determine if the responses we observed are statistically the same as or different from what we predicted under the null hypothesis. The test statistic in this case came out at 0.36. We compare this to our significance level, or alpha, which we chose to be 0.05. Our test statistic is greater than our significance level. This means that our predicted counts are not significantly different from our observed counts, statistically speaking. We cannot reject the null hypothesis in this case, or in other words, we conclude that how inclusive participants found the training to be was not impacted by whether or not they belong to an under-represented group.

If we had simply looked at the counts in the first table we might have been tempted to conclude that participants from an under-represented group found the training to be less inclusive because there were more "not at all inclusive" and "somewhat inclusive" responses than among those not from an under-represented group. This example demonstrates the importance of using statistically sound test statistics rather than just the eye-ball method for drawing evidence based conclusions.

Here is a video on how to do a Chi squared test in Excel.

T-TEST FOR DIFFERENCE IN MEANS

A t-test is used to determine if there is a statistically significant difference in the means of two groups. This is very handy when your data has measurements rather than counts. You can watch this YouTube video to see how to do a t-test.

DATA VISUALIZATIONS

Common types of data visualizations are:

- Table
- Histogram
- Bar chart
- Line chart
- Pie chart

Here's a really interesting video exploring how we interpret a pie chart.

You can put several data visualizations together with short text to tell a story. Infographics can be static, like a poster, or can be interactive, with links to the data so that the user can change parameters and see a change in the visualization. Here are some examples on the Statistics Canada website. The population clock is a fun one.

Here's a technical guide to creating compelling data visualizations.

Here's a basic introduction to data visualizations.

DATA INTERPRETATION / STORYTELLING

If you're clear about your objectives for doing the survey, the story will tell itself once you have cleaned the data and then start doing your analysis and making data visualizations.

This video talks about aligning your data story with what your audience needs to know. This video shows how to use data to tell a story, using Excel and PowerPoint.

HOW TO APPLY THIS TO PROGRAM EVALUATION DATA

Program evaluation is an assessment of the results of a program to date, its current state and future risks. It's standard practice to evaluate ongoing programs on a regular basis to determine if funding will continue or if changes to program approach or management structure are required. The following are common elements of program evaluation:

- Stakeholders – how do they perceive the program and its impacts.
- Benefits – what are the outcomes, is the program doing what it was intended to do.

- Financials – financial management and control; this may focus on variance from planned budget and cost/benefit analysis.
- Risk – the effectiveness of risk management and review of issues and unintended consequences. A review of current risks may be relevant to deciding the future of the program.
- Quality – the quality of deliverables and services including tangible things such as reliability and intangible factors such as usability.
- Schedule – delivery compared to committed dates.
- Recommendations – analysis of gaps, root causes and an action plan to correct problems.

Doing a feedback survey at regular intervals following service delivery could feed into program evaluation. Design what questions the program evaluation should answer, then work backwards from these objectives to see what data you need to collect. Decide what format stakeholders and decision makers will want to see the program evaluation, and at what time intervals. For example, will they want a dashboard that is updated monthly, or daily, or a PowerPoint presentation quarterly, …. Automate as much as possible to minimize time spent fiddling with the data and visuals.
Example of a program evaluation dashboard: