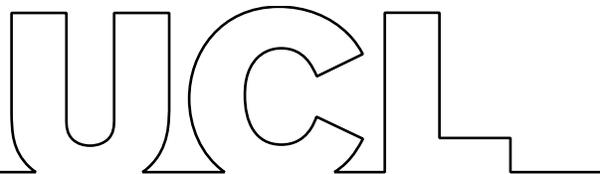


Evaluation and Feedback Survey Methods

REFERENCE MANUAL

LAURIE REEDMAN



Overview

This reference document was created to familiarise the reader with the steps in doing an online survey to seek feedback, satisfaction or impacts from participants of a service you have delivered, such as training or workshops. It is not a complete guide to survey methodology or data analysis. Included here are tips and pointers, hyperlinks to reference websites, and examples drawn from UCL Public Policy’s Masterclass Series Training Impact Survey. The author hopes you find it helpful.

CONTENTS

OBJECTIVES AND DELIVERABLES2

CRAFTING THE QUESTIONS5

ONLINE SURVEY PRODUCTS6

DATA COLLECTION.....7

DATA ANALYTICS8

DATA CLEANING.....8

WHAT TO DO BEFORE YOU EVEN START DATA CLEANING..... 8

MISSING VALUES, OR VALUES THAT DEFY LOGIC 9

FORMATTING ISSUES..... 10

WHAT TO DO WITH WRITE-IN RESPONSES 10

PREPARING MULTIPLE RESPONSE DATA FOR ANALYSIS..... 12

DATA ANALYSIS 14

MEAN, MEDIAN AND MODE 15

FREQUENCY DISTRIBUTIONS AND HISTOGRAMS 15

CROSS TABULATIONS..... 18

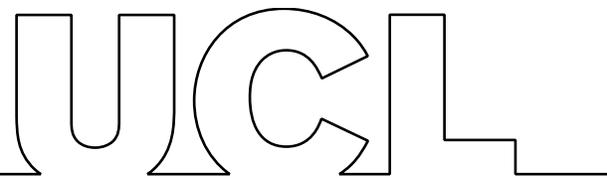
CHI SQUARED HYPOTHESIS TEST 18

T-TEST FOR DIFFERENCE IN MEANS..... 20

DATA VISUALISATIONS 20

DATA INTERPRETATION20

HOW TO APPLY THIS TO PROGRAM EVALUATION DATA21



OBJECTIVES AND DELIVERABLES

- Make clear objectives about what you want to measure
 - 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, are there thresholds that will trigger you to take an action?
- Set benchmarks
 - If you are conducting repeat 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 visualisation, 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](#)
- [European Public Administration Network](#)

For example, the purpose of the Public Policy Masterclass Series Training Impact Survey is stated as:

To explore why people went down particular routes and whether there are patterns in engagement.

To really hone the objective of the Training Impact Survey, turn these phrases into specific things that can be measured:

© 2022 University College London. All rights reserved. This resource has been added to UCL Public Policy's resources to support your studies. You may print and/or download a single copy for your personal, educational use. Further redistribution of teaching materials, including making copies available on the internet, is **not** permitted.



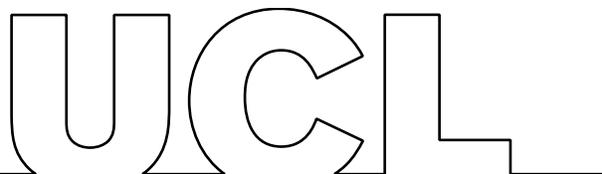
- **Why** people went down particular routes ...
 - If looking for causality, that's not easy; if looking for differences, for example people with these characteristics did A while people with these other characteristics did B, that can be done with a cross-tab (two-way table) and a Chi squared test for significance
- **Went down particular routes**
 - By routes do you mean doing certain things post-training? So again it could be a cross-tab of types of people versus different outcomes
- **Whether there are patterns in engagement**
 - By patterns do you mean that one group of people report doing certain engagement activities while another group of people report doing different engagement activities? Could do this with another cross-tab, or comparing histograms of engagement activities for different groups.
 - The questionnaire has a lot of categorical questions (where the respondents ticked one or more responses). A good way to display these is in a frequency chart, also known as a histogram. If there are fewer than 6 categories, a pie chart is a good option. For example for questions asking are you an ECR, are you from an under-represented group and how inclusive did you find the content, you could use a pie chart.

Let's say you want to know which elements of the training your participants found most useful, and you want to know if the responses are different for people who self-identify as ECR versus those who do not. How would you do this?

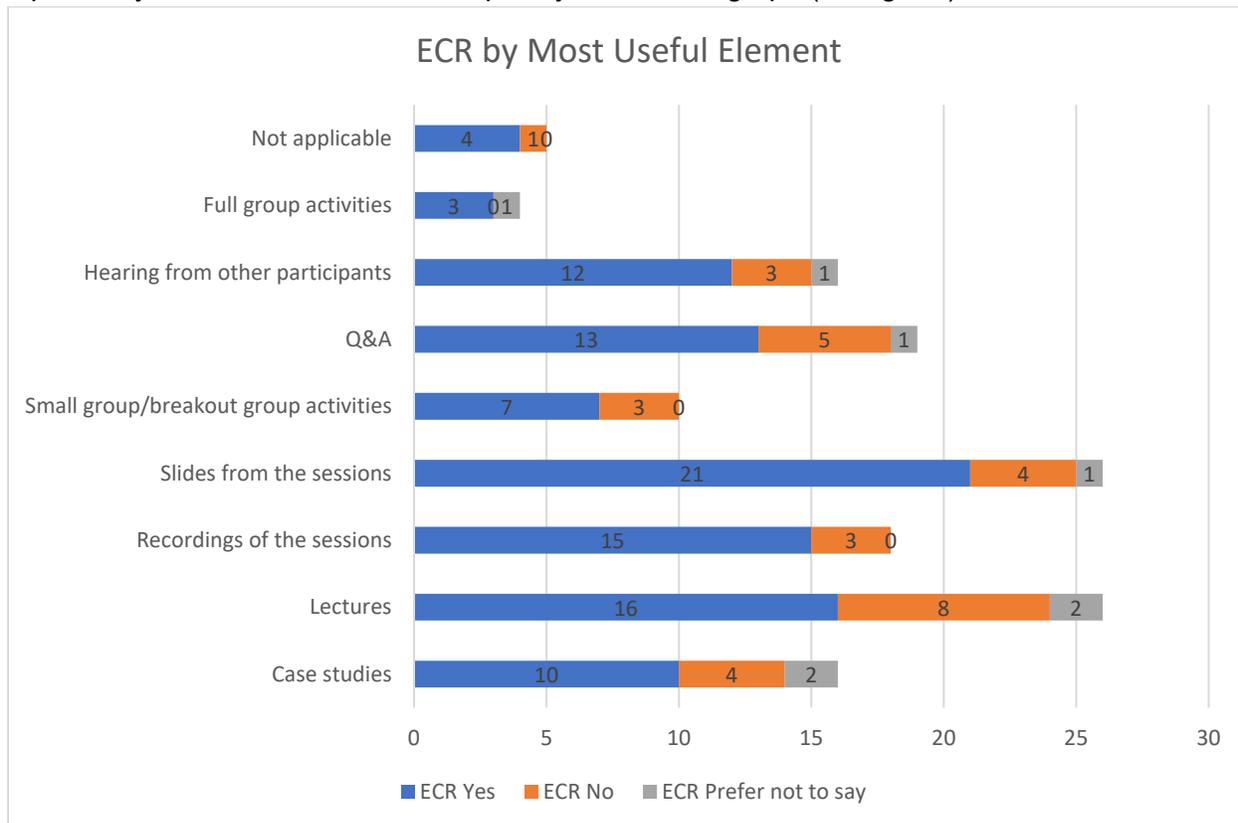
Option 1, you could do a 2-way table, or cross-tab:

Elements of training	Count of ECR who found this element useful	Count of non-ECR who found this element useful	Total respondents who found this element useful
Lectures	16	8	26
Case studies	10	4	16
Slides from sessions	21	4	26
Q&A	15	5	19
Etc.			

This might be difficult to interpret. It has the facts, but it isn't very intuitive.

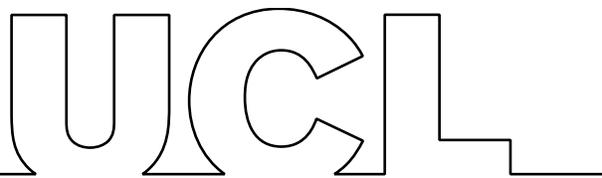


Option 2, you could do a stacked frequency distribution graph (histogram):



This tells you that slides from sessions and lectures were the most useful overall, but for some reason non-ECR participants did not find the slides particularly useful. Is this worth exploring? Now remember, here you are seeing actual data. At the time of setting objectives, you won't have the data yet. You have to think in advance, what will be important for you and your audience to see? Are you prepared to change the training elements that are least appealing, or only appeal to one segment of the participants but not to another? We set objectives so that we don't waste participants' time answering questions that we are not interested in analysing, and we don't waste our audience's attention looking at analysis that is not going to lead to action. This particular analysis is informative to the design of future training. Is that one of your objectives?

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 training had on their subsequent activities. This is a comparison between what they knew or were doing before and after having participated in your training. To do this, you would survey your participants with the same set of questions both at the beginning and at the end of the training. 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 because of skills and knowledge gained through the training.

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 to see if there is a statistically significant increase in self-reported knowledge/confidence level after the training. See the Data Analysis section of this reference manual for an explanation on how to conduct a one tailed t-test.

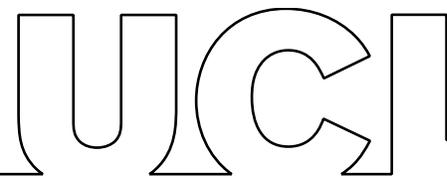
If measuring increase in knowledge or abilities is one of your objectives, you need to survey at multiple points in time with the same questions.

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
- 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
 - 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
- Avoid double-barreled questions (asking two questions in one)
- Keep it short – 5-10 minutes max for respondents to complete
 - There's no magic number of recommended questions – 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
- If asking about satisfaction on a scale, 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)
- Remember to say “thank-you” at the end of the survey
- Consider analysis when designing questions
 - Limit free text responses, they will need to be parsed and assigned to categories
 - Use key words and keep them to the front end of phrases
 - This will facilitate summarising frequency of responses
 - For “tick all that apply” questions, don’t use the same punctuation within a response as you will use to delimit responses in the exported data (this will be fully explained in the Example for Data Analysis)

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](#)

ONLINE SURVEY PRODUCTS

Here we assume you are doing an online survey, of people who have participated in a service you delivered, such as training or workshops. You do not want anyone outside of your participant population to be able to complete the survey, and you want your participants to answer only once. There are many online survey products available, some for free. [Jisc](#) and [Survey Monkey](#) are good candidates. Below are some things to consider if you are choosing an online survey product, and also things you should discuss with your online survey provider.

- Data security
 - Look for security of the instrument itself (nobody can tamper with the questions) and security of the collected data (nobody can hack into the data)
 - Look for secure access, so you control who is able to respond, and so that they can only respond once
 - Consider if anonymous response is needed, and if so, is it possible
 - What are the provider’s protocols on retention and destruction of your materials (collection instrument and data) in the host environment
 - Is there secure transfer between the host environment and your environment
 - Check with the legal or communications department of your organisation, they may recommend pre-approved providers and/or tools
- Cost and licensing
- Accessibility
 - Do you require multilingual versions of the questions
 - What options are available for vision impaired audiences
 - What other accessibility features might be needed



- Facilitating development
 - What features are offered in terms of testing, modifying, updating, versioning, archiving the questions
 - What support is offered in terms of design of questions, expertise in collecting sensitive data
- Real-time support
 - What support does the online survey provider offer during development and during the collection period
 - What is your risk mitigation plan
- Facilitating data analysis
 - How and in what format is the data exported and transferred to your environment
 - Do you have options in the formatting of datasets and data variables
 - Take into account the types of questions (for example multiple response, write-in); how can data formatting facilitate the analysis
 - If you don't understand why this is important then involve your data analyst in this part

Reference Websites:

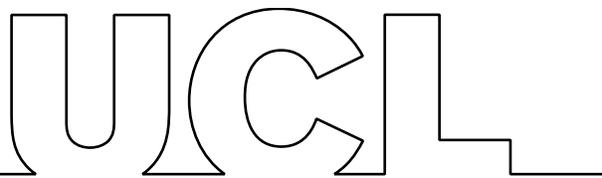
Both of these websites offer advice on how to choose an online survey tool

- <https://thenextscoop.com/choose-online-survey-tool-for-business/>
- <https://survicate.com/website-survey/tool/>

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
- Is the host or survey service provider (for example, in UCL's case, Jisc) available if there are problems, or if you have questions
 - How would you contact them if you needed to, what is their service standard on response time
- Before the actual collection period, test that the platform and questionnaire function as expected



- How will data be exported and transferred to your work environment

Ensure that all your data collection requirements can be met by the online survey product you intend to use. Discuss details with the online survey provider.

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 visualisation
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”. Despite your best efforts at crafting the questions and designing the survey, something is going to 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.

Another purpose for data cleaning is to organise 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.

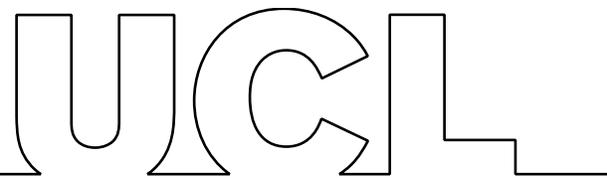
Data cleaning can be done using a programming language, a spreadsheet, or with a software product designed specifically to analyze survey data. Here’s what I recommend:

- If you have software such as SPSS, SAS, SPlus, R, Python, 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 BEFORE YOU EVEN START DATA CLEANING

- Import the data from the data collection product or website into your working environment



- Use filenames, time stamps and file sizes to verify that you got all your data, without duplicates or missing files
- Check that you have all the variables you are expecting; typically, you should have one variable per question on your survey, maybe one variable per response if there are questions with multiple responses, as well as variables holding identification information (if you asked participants for their email address for example).
- Your employer may have a security protocol on the safe storage of personal identification data; if so, be sure to follow that
- Once you are satisfied that you have all the data together in one place, before you do anything else, make a copy and keep it separate from where you are working. This way if the data gets corrupted you have a clean backup and can start over.

MISSING VALUES, OR VALUES THAT DEFY LOGIC

For each variable, check for:

- Missing values, if there are, is it ok, if not ok, can you impute
- Range of values – if there are values outside the expected range, are they ok, if not ok, can you impute

If you have questions that are logically linked, do the responses make sense

Examples from UCL Public Policy's Masterclass Series Training Impact Survey

- The survey questionnaire did not allow respondents to leave a question blank, so missing values were not an issue
- Most questions had categories to choose from, so values out of range were also not a problem; when respondents were able to write in a response, it was a phrase, so again there was no expected range
- There were questions that follow a logical flow, and there were some errors detected.

EXAMPLE 1: ERROR IN RESPONDENT'S LOGIC

Q1. Which UCL Faculty do you belong to? Select N/A if you are not at UCL.

Q2. Which organisation are you based at? Select N/A if you are at UCL.

So everybody should have either something in Q1 and N/A in Q2, or N/A in Q1 and something in Q2. But we got things like this:

Q1. Medical Sciences; Q2. UCLH.

UCLH is part of UCL, therefore we replaced UCLH in Q2 with N/A.

EXAMPLE 2: FLAW IN SURVEY QUESTIONS

Q7. Alongside your main professional activity, are you engaged in other activities related to public policy? (Check all that apply).

© 2022 University College London. All rights reserved. This resource has been added to UCL Public Policy's resources to support your studies. You may print and/or download a single copy for your personal, educational use. Further redistribution of teaching materials, including making copies available on the internet, is **not** permitted.



Q7a. If you chose “Other” in Q7, please specify:

Q7 asks “are you”, to which the answer should be “yes” or “no”, however the only answers respondents could choose from were a list of possible activities, with “Other” as the last category. There was no category for “no” or “none of the above”. So everyone who was not engaged in other activities chose “Other” as their response to Q7, and then typed “none” or “no” or “N/A” in Q7a. This was corrected by changing the response from “Other” to “None” in Q7 for everyone who had “none” or “no” or “N/A” in Q7a, and then changing the Q7a responses for those same respondents to blanks.

FORMATTING ISSUES

If you have numeric variables, ensure they are formatted appropriately for whatever mathematical operations you might want to do, for example mean, median, mode calculation, standard deviation, statistical analysis. Keep decimal places if appropriate.

If you have categorical variables, ensure they are all formatted the same way. Be particularly careful with spelling, capitalisation, spaces, punctuation. This is most important if people type in their own responses. Using drop down menus minimises the risk of typos and inconsistent formatting.

WHAT TO DO WITH WRITE-IN 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 summarise them, by hand or with automation?

A super simple way to find key words is on the Home tab, the Editing group, click Find and select, then click Find. In the dialogue box type the text you are looking for. This is a really quick and easy way to get an idea which key words were mentioned frequently. I recommend doing this first, and it might be all you need.

If you have many write-in responses, then you will probably want to use some sort of automation to find where key words are mentioned, or to find frequently used words or phrases. Machine learning algorithms can do this, but if you don’t have machine learning resources, you can do a simple version in Excel.

The Excel SEARCH function will tell you where in a cell certain text starts, or it will return an error if the text is not in the cell. The FIND function will do exactly the same thing, except that it is case sensitive. SEARCH is not case sensitive, so it is better to use SEARCH.

EXAMPLE 3: EXCEL SEARCH FUNCTION TO FIND KEY WORDS IN TEXT

© 2022 University College London. All rights reserved. This resource has been added to UCL Public Policy’s resources to support your studies. You may print and/or download a single copy for your personal, educational use. Further redistribution of teaching materials, including making copies available on the internet, is **not** permitted.



Q17. Was there anything from the content delivered during the training that was unclear, challenging, or could be improved?

Let's say you want to know if anyone mentioned anything about writing, or engagement.

Column A has all the write-in responses. Let's put an indicator for the text "writ" so it will pick up both "write" and "writing" in column B, and an indicator for the text "engage" in column C.

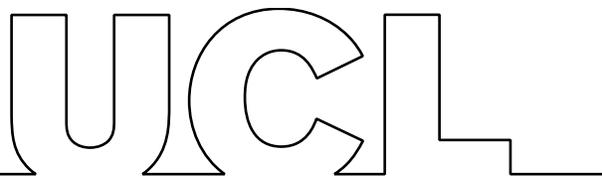
The table below is a few made-up responses to Q17 in column A, the formula to detect the text "writ" in column B, and the formula to detect the text "engage" in column C.

	Column A	Column B	Column C
Row 1	Q17 write-in responses	WRITING	ENGAGEMENT
Row 2	I would have liked more detail on writing for policy audiences	=search("writ",A2)	=search("engage",A2)
Row 3	Material on stakeholder engagement was too vague	=search("writ",A3)	=search("engage",A3)
Row 4	Presenter was totally engaged with participants, it was great!	=search("writ",A4)	=search("engage",A4)
Row 5	The written material was poorly organised	=search("writ",A5)	=search("engage",A5)

The table below shows the results of the formulas in the above table.

	Column A	Column B	Column C
Row 1	Q17 write-in responses	WRITING	ENGAGEMENT
Row 2	I would have liked more detail on writing for policy audiences	35	#VALUE!
Row 3	Material on stakeholder engagement was too vague	#VALUE!	25
Row 4	Presenter was totally engaged with participants, it was great!	#VALUE!	23
Row 5	The written material was poorly organised	5	#VALUE!

The text "writ" was found in row 2 starting at the 35th character and in row 5 at the 5th character, and the text "engage" was found in row 3 starting at the 25th character and in row 4 starting at the 23rd character.



You could use conditional formatting (Home tab, conditional formatting pane, highlight cells rules option) to highlight cells whose value is greater than zero, which are the rows where the text you searched for was found.

Here's a [video](#) you might find helpful.

PREPARING MULTIPLE RESPONSE DATA FOR ANALYSIS

Everyone loves “tick all that apply” questions. But dealing with the answers is not trivial. Let's look at an example.

EXAMPLE 4: 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, and the responses from Q13 looked like this:

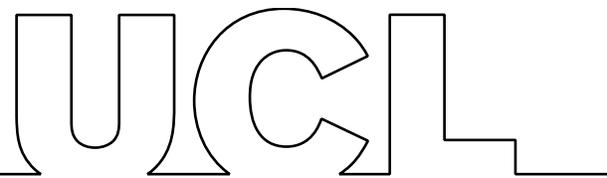
Q13
Lectures
Lectures, Case studies
Case studies, Slides from sessions, Q&A
Lectures, Q&A

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 can't. 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

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 visualise 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 visualisations.



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 that we saw in the last example 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.

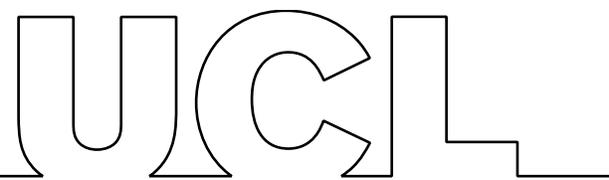
Then check this [page](#) out, which tells you how to use the SUMPRODUCT function with a double negative to turn the TRUE values into 1s and the FALSE values into 0s.

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

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 visualisations. 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 analysing 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 visualisations in Excel, although sometimes data visualisation in Excel is cumbersome.

If you want something less cumbersome than Excel for data visualisations 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 visualisations 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 visualisations. Here are some [Minitab tutorials](#).

All the examples here are in Excel.



MEAN, MEDIAN AND MODE

You learned these calculations in grade school. 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

When you have categorical data, a nice way to summarise 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 5: 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 4.

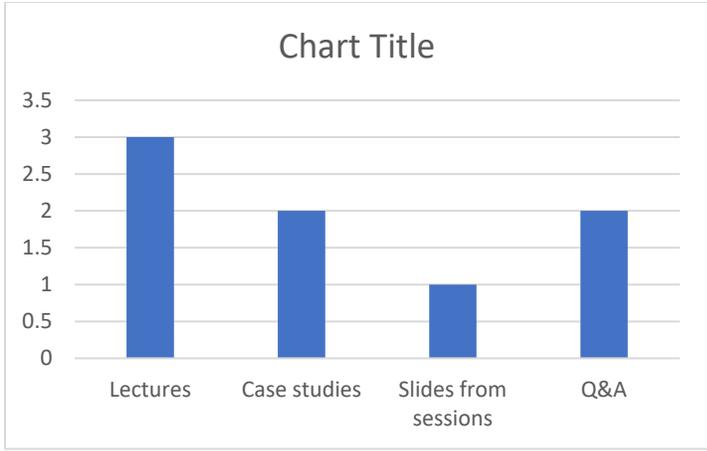
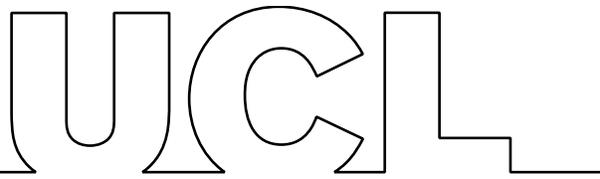
Lectures	Case studies	Slides from sessions	Q&A
1	0	0	0
1	1	0	0
0	1	1	1
1	0	0	1

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)

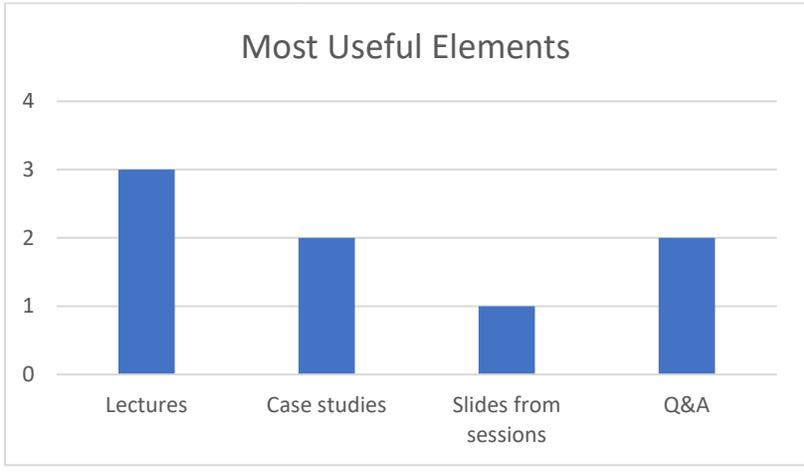
The result will look like this:

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 6: HISTOGRAM IN EXCEL FROM CATEGORICAL DATA

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

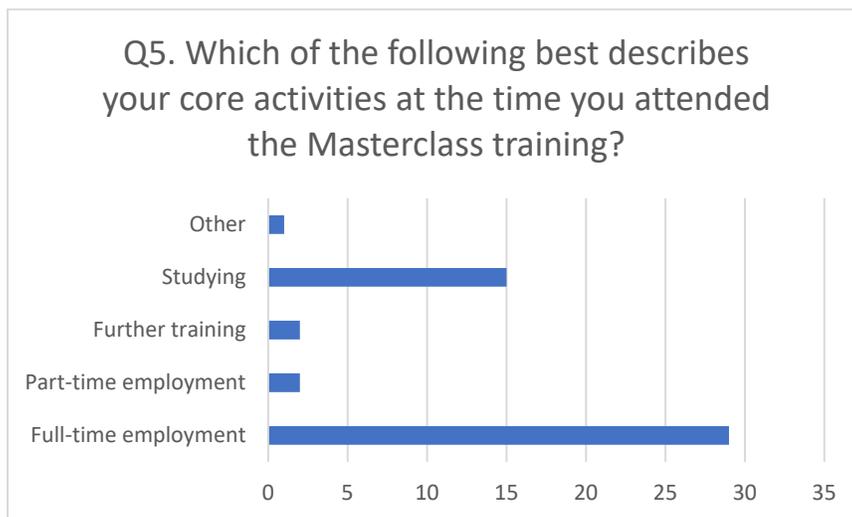
Respondents had to choose one of the following activities:

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

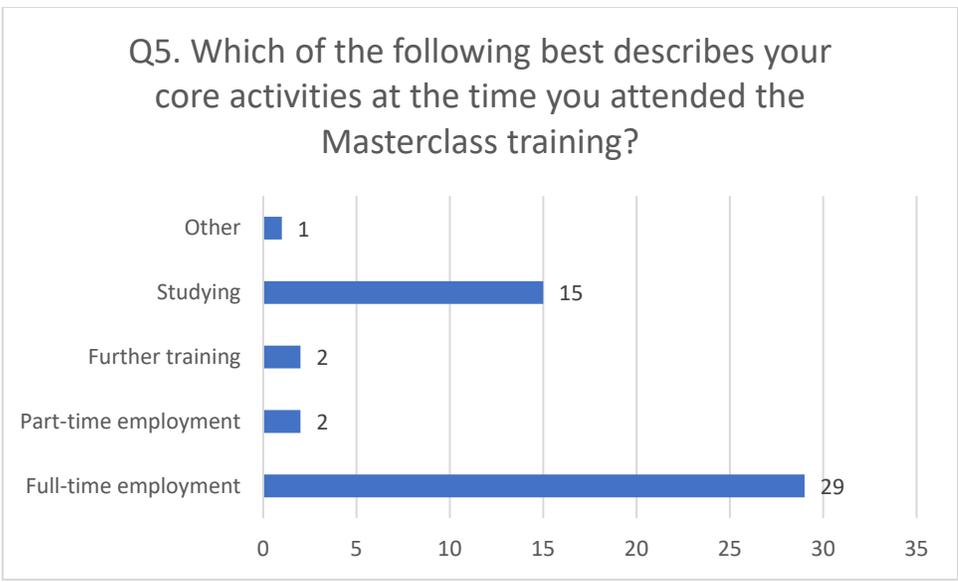
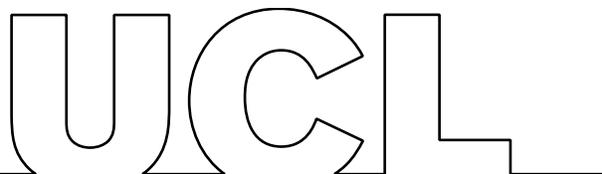


Because this is a choose one type of question, not tick all that apply, we don't need to make indicator variables, we simply need to count the number of occurrences of each activity. We can do this in Excel with a pivot table.

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 7: 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.

18. How inclusive was the training	4. Are you from an under-represented group?			
	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

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 8: CHI SQUARED TEST IN EXCEL

Let's use the cross tabulation of questions 4 and 8 from example 7.

In a hypothesis test, the null hypothesis is that there is no relationship, so knowing the values of one variables 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 conclude that there is no relationship.

In example 7 we had 19 people who were from an under-represented group, 25 who were not, and 1 person who preferred not to say. Let's remove that one person from the table and do the Chi squared test. Then we can see if the result changes if we assume that person actually was, and actually was not from an under-represented group.

We start with the same cross tab without the person who preferred not to say. This is the observed distribution.

18. How inclusive was the training	4. Are you from an under-represented group?		
	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

What we want to test is if the distribution that we observed is the same as the distribution we would expect if there was no relationship (or in other words if the null hypothesis were true).

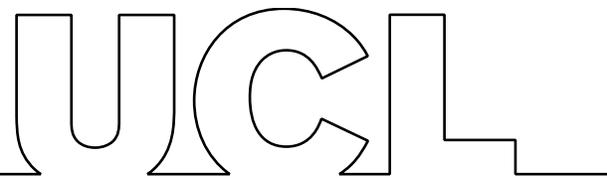
To get started we need the percentages of total people who gave the 3 different responses to question 18. 2 out of 44 or 4.5% of respondents said "Not at all inclusive", 20 out of 44 or 45% of respondents said "Somewhat inclusive", and 22 out of 44 or 50% of respondents said "Very inclusive". We multiply these percentages by the total number of respondents who reported being from an under-represented group and not from an under-reported group to get the distribution of responses to question 18 that we would expect if there was no relationship between the two questions. The predicted counts are shown below:

0.863636364	0.090909091
8.636363636	11.36363636
9.5	12.5

The function to perform the Chi squared test in Excel is:

=CHITEST(observed values, expected values), and it returns a p-value.

© 2022 University College London. All rights reserved. This resource has been added to UCL Public Policy's resources to support your studies. You may print and/or download a single copy for your personal, educational use. Further redistribution of teaching materials, including making copies available on the internet, is **not** permitted.



In our case the p-value is much smaller than 0.05, leading us to reject the null hypothesis and conclude that there is a relationship. Respondents from an under-represented group found the training to be less inclusive than respondents not from an under-represented group.

Even if we put the one person who preferred not to say if they were from an under-represented group or not in the “yes” or “no” column, the p-value remains well below 0.05, so our conclusion remains the same.

Video [tutorial](#) on doing 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. UCL’s Public Policy Masterclass Series Training Impact Survey did not have any measurement questions, so you can watch these two YouTube videos to see how to do a t-test.

Video [tutorial](#) on hypothesis testing step-by-step.

Video [tutorial](#) on two means t-test in Excel.

DATA VISUALISATIONS

Common types of data visualisations 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 visualisations 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 visualisation. 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 visualisations.

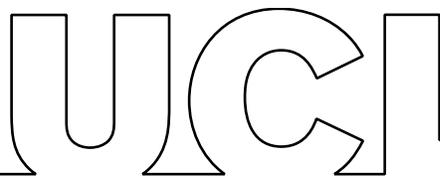
Here’s a [basic introduction](#) to data visualisations.

DATA INTERPRETATION

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 visualisations.

This [video](#) talks about aligning your data story with what your audience needs to know.

© 2022 University College London. All rights reserved. This resource has been added to UCL Public Policy’s resources to support your studies. You may print and/or download a single copy for your personal, educational use. Further redistribution of teaching materials, including making copies available on the internet, is not permitted.



This [video](#) shows how to use data to tell a story, using Excel and PowerPoint.

HOW TO APPLY THIS TO PROGRAM EVALUATION DATA

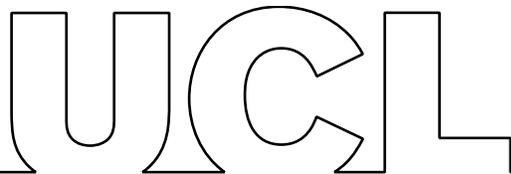
Program evaluation is an assessment of the results of a program to date, its current state and future risks. It's standard practice to evaluate ongoing programs on a regular basis to determine if funding will continue or if changes to program approach or management structure are required. The following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

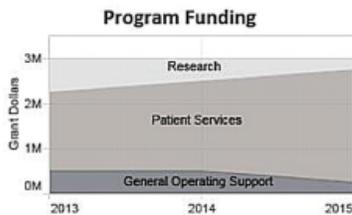
Here is a link for some [further reading](#).

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 minimise time spend fiddling with the data and visuals.

Example of a program evaluation dashboard:



SAMPLE EVALUATION DASHBOARD: RURAL HEALTH CARE ACCESS PROGRAM



Patients Served

Population	2013	2014	2015
Adults	9,081	10,039	10,257
Children	4,239	4,274	4,419

2016 YTD: 1,081 adults and 417 children served

