DIT590 Research Methods & Technical Writing

Lecture 3: Survey

Richard Berntsson Svensson
http://www.rbsv.eu/courses/rmtw

Re-cap

- Purpose statement
  - Purpose, intent objective
  - Describe, understand, The relationship between comaprison of
- Research Questions
  - State RQ and/or hypotheses
  - What, how
  - Null hypotheses

Survey method

Surveys

- A common goal of survey research is to collect data representative of a population
- Information gathered from the survey is used to generalize findings from a drawn sample back to the population
  - within the limits of random error

Errors in Surveys

- **Sampling error and biases** included by the sample design
  - The difference between a sample a the population form which it is selected
  - Unlikely to end up with a truly representative sample
- **Non-sampling errors** caused by problems in data collection and processing
  - *Overcoverage*: Inclusion of data from outside the population
  - *Undercoverage*: Sampling frame does not include elements in the population
  - *Measurement error*: e.g. when respondents misunderstand a question, or find it difficult to answer
  - *Processing error*: Mistakes in data coding
  - *Non-response*: Failure to obtain complete data from all selected individuals

Representative Sampling

- Sample = *representative* when it is an accurate proportional representation of the population under study
- A truly representative sample requires every person in a population to have had an equal chance of being chosen to participate in the survey = *randomization*
  - Interview every 100th person who show up during the time you were collecting data
  - Walk up to people outside the building – only include those that happen to pass this place during this time
  - Unintentionally be biased who you question – unconsciously choose not to question people who look preoccupied, busy, or unfriendly people
- Computer-list of the population – randomly generate a list of 2 percent
  - Still might not be representative
  - Correct proportion of included groups
- *Stratification* – truly representative sample of the population
- Stratify a population – decide what sub-categories of the population might be statistically significant – randomly generate certain percentages of sub categories

Survey sample size

- Procedures used in how sample size is calculated should always be reported
- Allowing the reader to make his/her own judgments whether they accept the researcher's assumptions and procedures
- Census – sample
  - Many studies based on entire population **census data** achieve low response rates
  - Using adequate sample along with high quality data collection efforts produce more reliable, valid, and generalizable results

Surveys

- Surveys work best with standardized question where it is possible to be confident that the questions mean the same thing to different respondents
- The requirement is that you know what kind of information you want
- Often carried out for descriptive purposes
  - How are voters' views related to age, gender, income, region of the country?

The instrument: Questionnaire

- A set of common questions laid out in a standard and logical form to record individual respondent's attitudes and behavior.

- **Questions:** framing of the precise questions that are asked. Does the question elicit a useful and unbiased response?

Type of questions

- Structured or fixed response, *closed ended* questions
- Partially-structured questions
- Non-structured or *open ended* questions
- It is important to understand when and how to use these questions when designing your survey

Open-ended questions

- Invite free ranging responses
- Suited for small and qualitative surveys
  - When exploring new ideas and no pre-conceived ideas about what to expect from the respondents

+ Useful for obtaining a deep understanding of the respondents' views and behaviors

- Views and opinions may be difficult to capture precisely

Closed questions

- Offers the respondent a closed set of responses from which to choose.
- Norm in quantitative studies
  - When you have a thorough understanding of the responses so that you can appropriately develop the answer choices
  - When you are not trying to capture new ideas or thoughts from the respondents.

- Important to ensure correct response codes otherwise respondents' will use the dustbin category of 'others'.

Partially structured questions

- A partial list of answer choices, but may still have some doubt or uncertainty about possible responses

Example

- Why did you join the SW Engineering Program? (Please select all that apply)
  - () My parents asked me to join
  - () I wanted to get into a university program with good relationships with the industry
  - () I thought it would help me to improve my employability
  - () I knew other students who were studying SW Engineering
  - () Other, please specify_____

Example Question

What brand of computer do you own?

A. IBM
B. Apple

Problems are…

- What if the respondents does not own a computer?
- What if he/she owns a different brand of computer?
- What if he/she owns both a PC and an Apple?

- **How do you correct this problem?**
  1. Do you own an IBM PC (circle: Yes or No)
  2. Do you own an Apple computer (circle: Yes or No)
  3. What brand of computer do you own? (check all that apply)
     - Do not own a computer____
     - IBM PC____
     - Apple____
     - Other____

Questionnaire design: Biased and neutral questions

Bad Question: Leading

- Do you think that the new cafeteria lunch menu offers a better variety of healthful foods than the old one?
  - () Yes
  - () No

Good Question: Neutral

- How do you feel about the new cafeteria lunch menu compared to the old one?
  - () The new menu offers a better variety of healthful foods
  - () The old menu offers a better variety of healthful foods
  - () The selections are similar
  - () No opinion

Questionnaire design

- Flow of questions is critical to good interview
  - Easy to answer questions should be put early
  - Interesting questions first
  - Non-threatening (can put people off)
  - Ensure anonymity
  - Respondents should be thanked for their time and effort

Pre-testing!

- The only good question is a pre-tested one!
- Does the respondents understand the questions in the same manner?
- Try it out on a few people
- One way is to discuss the questionnaire with your peers, colleagues (reviewing)
- Conduct a pilot study (recommended)

Discuss!

Think of a time when you have been asked to fill in survey, either online or on paper.

What factors influenced your responses?

Surveys

- Data are usually collected through
  - Questionnaires
  - Interviews
- Qualitative (open-ended questions)
- Quantitative (closed questions)
- Two basic types of surveys
  - Cross-sectional
  - Longitudinal

Basic steps in Survey Research

Organization - determine who is to be sampled and what is to be learned about the sample

Questionnaire Design - based on the goal of the survey, prepare questions and arrange in a logical order to create the survey questionnaire

Sampling - develop a repeatable plan to randomly choose a sample capable of meeting the survey's goals - select a sample

Data Collection - develop a plan for contacting the sample and collecting data from participants - carry out

Data Processing - enter collected data into computer and check for accuracy

Analysis - compile and disseminate the results

Components of a Survey Method Plan

- Identify the purpose of survey research
- Indicate why survey is preferred data collection procedure
- Indicate cross-sectional/longitudinal survey
- Specify form of data collection (questionnaires/interviews)
- Identify the population in study, size, means of identifying individuals in population
- Identify sampling design
- Identify selection process for individuals (random/convenience…)
- Identify whether stratification of population
- Discuss procedure for selecting sample
- Indicate number of people in study and procedures used to compute this number (i.e. large enough sample from the population)
- Instrumentation – efforts to establish validity (content, predictive, construct)
- Include sample items from the instrument (or entire) for reader
- Indicate major content sections, e.g. cover letter, items, and scales used to measure items on the instrument
- Indicate pretest of instrument, pilot testing, focus groups
- Indicate steps for administering and follow up to ensure high respons rate
- Relate variables, research questions and items on instrument
- Present steps involved in data analysis

Carrying out a survey - steps

- 1. Initial design and planning
  - Purpose? RQs? What is feasible?
  - How many? How much? Who? Where? When?
  - Population?
  - Sampling frame
- 2. Designing the questionnaire
  - Interview-based or self-completion?
  - Survey questions designed to fit goals of the research
  - Do the respondents understand the questions? Link RQ to SQ:s
- 3. Pre-testing it
  - Run a pilot
  - Revise the survey and procedure according to feedback
- 4. Final design and planning
  - Editing. Ensure there are no spelling mistakes
- 5. Data collection
  - Codes to identify responses, e.g. ‘1’ for female and ‘2’ for male
- 6. Analysis and reporting
  - What does the analysis mean? What is your interpretation? Validity issues?

Question Design

A poorly designed questionnaire renders results meaningless

- Make sure the question asks only one clear thing
- Ask questions the respondent can accurately answer
- Questions should be relevant
- Use short items (easily read, understood, and answered quickly)
- Avoid negative items
- Avoid biased items and terms
- Avoid slang, jargon, and technical terms (unless warranted)
- Whenever possible, develop consistent response methods
- Make questions as impersonal as possible
- Sequence questions from the general to the specific
- Closed questions - try to develop exhaustive and mutually exclusive response alternatives
- Place questions with similar content together in the survey instrument.
- Make the questions as easy to answer as possible
- Use very clear definitions for unique and unusual terms
- Use an attractive format that conveys a professional image
- Make sure that questions measure what they are supposed to – pilot test, focus

Advantages

- Useful in describing the characteristics of a large population
- Can be administered from remote using mail, email or telephone
- Very large samples are feasible, making the results statistically significant even when analyzing multiple variables
- Many questions can be asked about a given topic giving considerable flexibility to the analysis
- Similar data can be collected from groups then interpreted comparatively
- Economy in data collection due to the focus provided by standardized questions - only questions of interest to the researcher are asked, recorded, codified, and analyzed

Disadvantages

- Depend on subjects' motivation, honesty, memory, and ability to respond
- Although randomly sampled participants errors due to non-response may exist
- People who choose to respond on the survey may be different from those who do not respond, thus biasing the estimates
- Answer-choices could lead to vague data sets, e.g. the choice "moderately agree" may mean different things to different subjects, and to anyone interpreting the data for correlation

Validity Threats

- Whether one can draw meaningful and useful inferences from scores on the instrument
  - Content
  - Predictive/concurrent
  - Construct

[Creswell]

- Background of the sample is a central influence
- Hypothesis guessing
- The instrumentation were misunderstood by the subjects
- Willingness to participate
- Generalize the findings

To do …

- Find someone to work with
  - Notify course responsible
- Work on Assignment 1
- Theory: Read CRES: 8, [Expe], [SiSE], [ReSE]
- Attend Wednesday's lecture about Software Experiment