Welcome!

Module 6

– Data collection and analysis

What is evaluation?
Clarifying the concepts in the evaluation neighbourhood.
Main types of evaluation.
Evaluation process.
Evaluation stakeholders.

Based on D. Svoboda and EPDET

Group work!

Task 6-A – Your data

Brainstorming

What can be used as data in an evaluation?

Data

Data sources

- Existing information
  - Demographics, census and other databases, public records, media, registries, research, other evaluations, big data
- People
  - Stakeholders (beneficiaries; target groups, donors, program staff, managers, administrators; parliamentarians and policy makers; collaborators, etc.)
- Observations
  - Observed behavior, practices, activities (community meetings, etc.), events

Primary – Secondary data

- Primary Data:
  - data not collected previously
  - data collected for the purpose of identifying or answering specified evaluation questions
  - questionnaires, interview and focus group protocols, etc., designed specifically for the evaluation
- Secondary Data:
  - previously collected data
  - data collected not specifically to address the evaluation questions
  - databases, media reports, documents, registries, specific survey items collected for other purposes, but useful to the evaluation

Two types of data (and mixed)

- **Quantitative**
  - For example, survey, observation, social media, document analysis, etc.
- **Qualitative**
  - For example, interview, focus group, observation, journaling / Vlogs / photolanguage, storyboards, talking circles – brainstorming, social media, document analysis, etc.

Things to think about when collecting data

- **Breadth or depth**
- **Structured or unstructured**
- **Obtrusive or non-obtrusive**
- **One-time or recurring**

Surveys – Questionnaires

- Terminology is often used interchangeably
  - **Questionnaire:** series of specific questions and/or statements accompanied by a set of responses from which to choose
  - **Survey:** method or process for gathering information; may include questionnaires (hard copy and Internet), interviews in which the questionnaire questions are asked to individuals

Survey

Exercise 6-1 – Prepare a questionaire

Tips in developing survey

- Review other surveys – what worked?
- Provide explanation of the purpose of the survey
- Items – questions
  - Pilot test new item and response categories
  - Vocabulary has to be simple
  - Questions should be understood in the same way by all respondents
    - Example: Do you consider lack of adherence to be a big problem?

Tips in developing survey

- Avoid mixing positive and negative terms
- Use the same format
  - Statements versus questions (all items worded similarly; statements are often perceived as more neutral)
  - Grammar – same structure and tense (present, past, future)
  - Length – long items are likely to be seen as different
- Use a readable font

Tips in developing survey

- Provide sufficient „white space" to
  - differentiate item sets,
  - change in response categories
- Ensure that the respondent is able to answer the item
  - Provide a screening question if needed
  - Example: *Are you familiar with xxx…?* Before going on to ask about opinions about *xxx*
- Do not try and hide sensitive items, but don't put them at the beginning either

Writing good questions

- A good question is understood easily and correctly by all respondents
- A good question elicits the kind of information the researcher is looking for:
  - **Poor quality Q**: When did you move to Ankara, Turkey?
    - Answer: In 2005
    - Answer: When I was 23 years old
    - Answer: After I finished high school
  - **High-quality Q**: In what YEAR did you move to Ankara, Turkey?

Writing good questions

- Ask about first hand experience
  - **Poor quality Q**: Do you think that there is child labour in your village?
  - **High quality Q**: Have you seen child labour being used in your village?
- Ask one question at a time
  - **Poor quality Q**: Last week, did you attend school and do your assignment?
  - **High quality Q**: Last week, did you attend school?

Developing response categories

	Strongly Disagree	Disagree	Neutral	Agree	Strongly Agree	N/A
Likert Forced Choice	O	O		O	O	
Likert Neutral Choice	O	O	O	O	O	
Likert Forced Choice (N/A)	O	O		O	O	O
Likert Neutral Choice (N/A)	O	O	O	O	O	O

Response measurement scale

- Fixed-response formats
  - Dichotomous
    - Yes – No / True – False
  - Likert scaling – rating scale
    - Strength of agreement
    - Frequency of self-reported behavior
  - Multiple choice
    - Check all that apply
    - One correct response among a set of „foils"/„distractors"
      - Plausibility of foils/distractors must not reveal the correct response
  - Rank order
    - Use a reasonable number of objects to rank

Developing response categories

- Response categories
  - Ensure that the response categories allow respondents to answer
  - Clearly label response categories
    - Avoid ambiguous labels
      - most of the time versus a lot of the time
    - Define what is meant
  - Use a reasonable number of response options
  - Use an appropriate and non-influential range of response categories

Example: influence of response options

Low Frequency Choices		High Frequency Choices	
Up to ½ hour	7.4%	Up to 2½ hours	62.5%
½ to 1 hour	17.7%	2½ to 3 hours	23.4%
1 to 1½ hours	26.5%	3 to 3½ hours	7.8%
1½ to 2 hours	14.7%	3½ to 4 hours	4.7%
2 to 2½ hours	17.7%	4 to 4½ hours	1.6%
More than 2½ hours	16.2%	More than 4½ hours	0.0%

83.8%

37.5%

Schwartz et al. (1985). Response scales: Effects of category range on reported behavior and comparative judgments. Public Opinion Quarterly, 49, 388-395

Other methods - quantitive

- Observation (e.g., behavioral counts – duration, frequency, workflow, etc.)
- Transaction logs (e.g., social media and website activity, etc.)
- Highly structured interviews
- Simulation and experiments
- Documents
- Secondary data

When to use qualitative data

- Available theory may be
  - Inaccurate
  - Insufficiently developed
  - Inappropriate or biased
- Novel phenomena
  - Need for description and exploration
- Need for more comprehensive understanding
  - Meaning and tacit knowledge
  - Processes
- Development of causal explanation

Qualitative data collection tools

- Interview
- Focus group
- Observation
- Journaling / Vlogs / photolanguage
- Storyboards – story-telling
- Talking circles – brainstorming
- Social media
- Artistic expression
- Simulation and role-play
- Document analysis

Interviews & Focus Groups

- **Interviews**
  - Two interviewers, one respondent
    - One interviewer serves as note taker
    - Resource intensive, but often required
- **Group Interviews** (2 to 3 individuals)
  - Ask everyone the same questions
- **Focus Groups** (typically 6 to 10 individuals)
  - Not group interviews
  - Group interactions are key
  - Many formats (usually highly structured)
  - Examining common themes
  - Not easy to do well

Observation

- **Full participation** – immersion into the setting, program, culture – least common
- **Participant observer** – individuals has some role in the phenomena/process he/she observes, but also a clear role as an evaluator – somewhat common
- **Non-participant** – objective observer – most common
  - In person (obtrusive, non-obtrusive)
  - Video
  - Audio

Triangulation

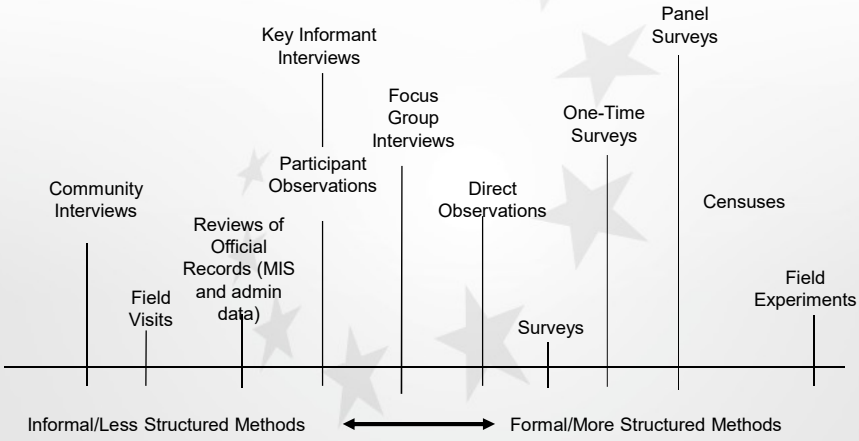
Use more **sources of information**, analyze findings for consistency

E.g., program staff, government officials, beneficiaries

Use more **data collection instruments**, analyze findings for consistency

E.g., interviews, focus groups, questionnaires, existing data, expert panels

Data collection methods



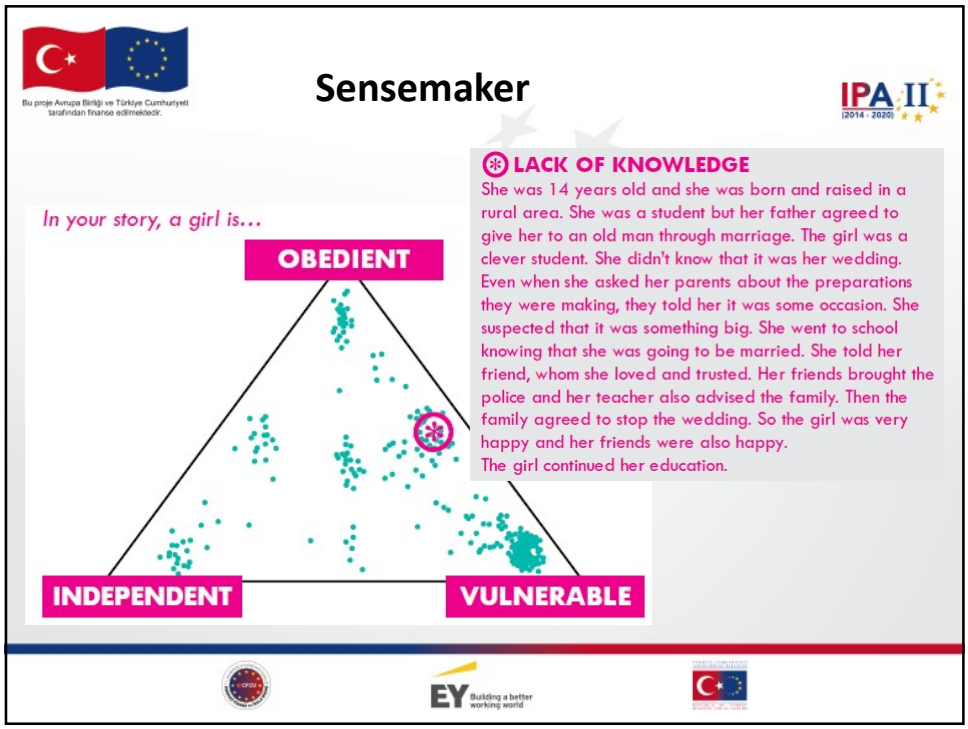
Narrative

Exercise 6-2 – Write your story

WHAT IS SENSEMAKER®?

SenseMaker® is a narrative-based research methodology that enables the capture and analysis of a large quantity of stories in order to understand complex change. It is a form of meta analysis of qualitative data that bridges a gap between case studies and large-sample survey data. The approach offers a methodological breakthrough for recognising patterns and trends in perceptions, behaviours and relationships.

SenseMaker® draws upon anthropology, complexity theory and neuroscience. The methodology involves collecting large numbers of short stories that, together, create a nuanced picture of a given topic, in the same way that many pixels come together to produce a clear image. SenseMaker® can also assist with early recognition of weak signals of changing social dynamics, and with the identification of emerging or outlier factors that deviate from normal trends and patterns.



Brainstorming

What is it „big data“?

Big Data

How can these be used in evaluation?

Sampling

- Is it possible to collect data from the entire population? (census)
  - If so, we can talk about what is true for the entire population
  - Often we cannot (time/cost)
  - If not, we can use a smaller subset: a SAMPLE

Sampling

https://www.youtube.com/watch?v=pTuj57uXWlk

Sampling

https://www.youtube.com/watch?v=EZrP_av3cmA

Sampling concepts

- Population (the total set of units)
- Census (collection of data from an entire population)
- Sample (a subset of the population)
- Sampling Frame (list from which to select your sample)
- Sample Design (methods of sampling: probability or non-probability)

Random sampling

- Lottery, each unit has a equal chance of being selected
- Can make estimates about the larger population based on the subset
- Advantages:
  - eliminates selection bias
  - able to generalize to the population
  - often cost-effective

Random sampling

- Simple random sample
- Random-start and fixed-interval sample
- Stratified random sample
- Random cluster sample
- Multistage random sample

Non-random sampling

- Can be more focused
- Can help make sure a small sample is representative
- Cannot make inferences to a larger population (you cannot generalize if you do not randomize…)

Types of non-random samples

purposeful (judgment)	set criteria to achieve a specific mix of participants (also quota sampling)
snowball (referral chain)	ask people who else you should interview
convenience	whoever is easiest to contact or whatever is easiest to observe
voluntary sampling	researcher allows cases/respondents to get involved voluntarily

Determining the sample size

- Statistics are used to estimate the probability that the sample results are representative of the population as a whole
- Evaluators must choose how confident they need to be
- Generally use 95% confidence level
  - 95 times out of 100, sample results will accurately reflect the population as a whole
- The higher confidence level, the larger sample needed

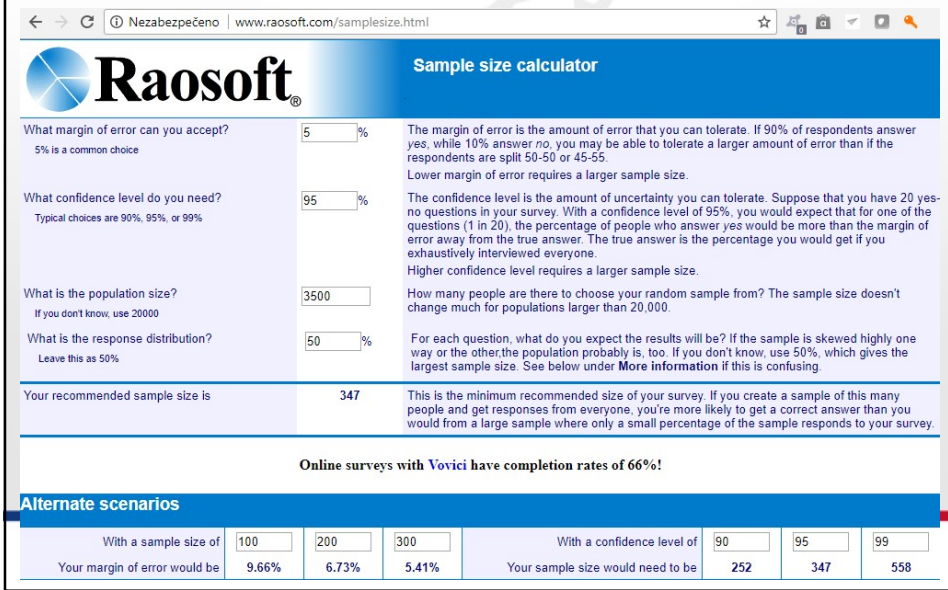
Sample size and population

- By increasing sample size, you increase accuracy and decrease margin of error
- The smaller the population, the larger the needed ratio of the sample size to the population size
- Aim for is a 95% confidence level and a ±5% margin of error (sampling error)

Sample sizes for large populations

Margin of Error	Confidence Level		
	99%	95%	90%
± 1%	16,576	9,604	6,765
± 2%	4,144	2,401	1,691
± 3%	1,848	1,067	752
± 5%	666	384	271

Sample size calculator



Nezabezpečeno | www.raosoft.com/samplesize.html

Raosoft®

Sample size calculator

What margin of error can you accept? 5% is a common choice	5 %	The margin of error is the amount of error that you can tolerate. If 90% of respondents answer *yes*, while 10% answer *no*, you may be able to tolerate a larger amount of error than if the respondents are split 50-50 or 45-55. Lower margin of error requires a larger sample size.
What confidence level do you need? Typical choices are 90%, 95%, or 99%	95 %	The confidence level is the amount of uncertainty you can tolerate. Suppose that you have 20 yes-no questions in your survey. With a confidence level of 95%, you would expect that for one of the questions (1 in 20), the percentage of people who answer *yes* would be more than the margin of error away from the true answer. The true answer is the percentage you would get if you exhaustively interviewed everyone. Higher confidence level requires a larger sample size.
What is the population size? If you don't know, use 20000	3500	How many people are there to choose your random sample from? The sample size doesn't change much for populations larger than 20,000.
What is the response distribution? Leave this as 50%	50 %	For each question, what do you expect the results will be? If the sample is skewed highly one way or the other,the population probably is, too. If you don't know, use 50%, which gives the largest sample size. See below under **More information** if this is confusing.
Your recommended sample size is	347	This is the minimum recommended size of your survey. If you create a sample of this many people and get responses from everyone, you're more likely to get a correct answer than you would from a large sample where only a small percentage of the sample responds to your survey.

Online surveys with Vovici have completion rates of 66%!

Alternate scenarios

With a sample size of	100	200	300	With a confidence level of	90	95	99
Your margin of error would be	9.66%	6.73%	5.41%	Your sample size would need to be	252	347	558

Group work!

Task 6-B – Analyse your data

Qualitative analysis

- Best used for in-depth understanding of the intervention
- Answers questions like:
  - What are some of the difficulties faced by staff?
  - Why do participants say they dropped out early?
  - What is the experience like for participants?

Contrast with quantitative analysis

- Used to answer questions like:
  - What are the mean scores for the different groups of participants?
  - How do participants rate the relevance of the intervention on a scale of one to five?
  - How much variability is there in the responses to the item?
  - Are the differences between the two groups statistically significant?

Analyzing qualitative data

- Non-numerical data collected as part of the evaluation:
  - E.g. open-ended interviews, written documents, focus groups transcripts
- May use content analysis to identify common words, phrases, themes and patterns
- Also progressive focusing - ongoing analysis in which themes emerge

Quantitative data: using statistics

- Quantitative data are numerical and analyzed with statistics
  - **descriptive statistics**: used to describe and analyze data collected about a quantitative variable
  - **inferential statistics**: used with random sample data by predicting a range of population values for a quantitative variable

Exercise for reflection

Exercise 6-3 – Reflection of Module 6

Module 6 – Take aways

- Data are key inputs for evaluation. Without availability of reliable data, coclusions may not be valid.
- Don‘t forget to triangulate your data sources and collection methods.