Thesis Seminar
Sampling strategy for quantitative research

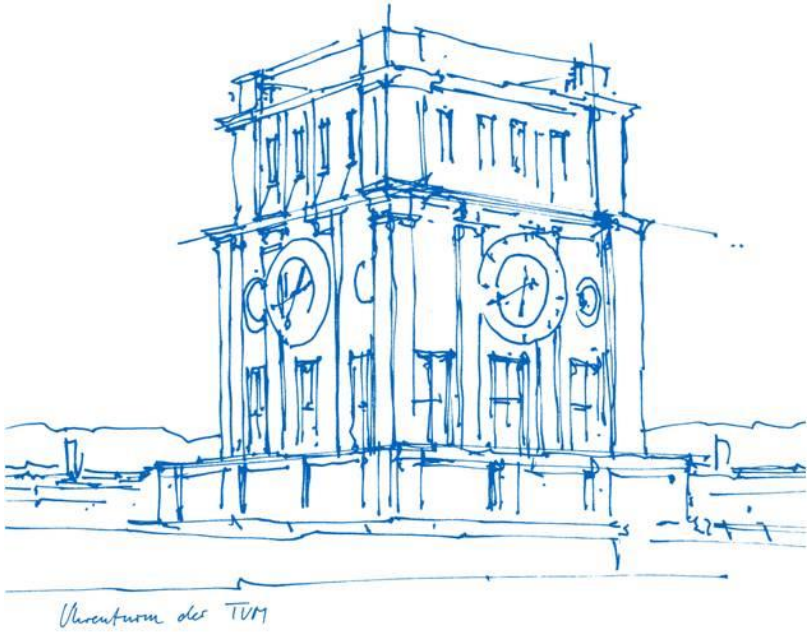
Dr. Theresa Treffers

Technical University of Munich

TUM School of Management

Chair for Strategy and Organization

Prof. Dr. Isabell M. Welpe



Data collection methods I

Primary data collection

- Data is collected to answer a (specific) research question
- Data collection through surveys, experiments, interviews, phone
- Population or sample

Secondary data collection

- Existing dataset is used to answer a research question
- Researcher has no influence on type, size, and quality of data

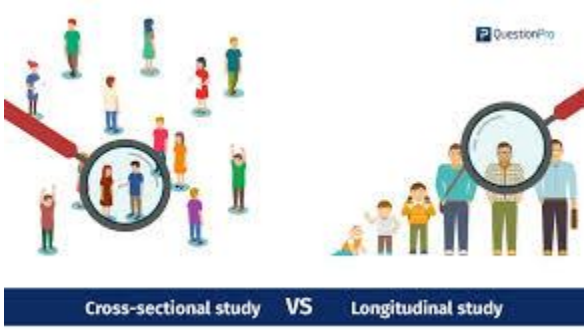
Data collection methods II

Cross-sectional data collection

- Status quo analysis
- Data collection at one point in time
- Cannot observe any change

Longitudinal data collection

- Data collection at several points in time
- Possible causal interpretations of relationships
- Observation of changes, developments, influences



Research question, population, and sample

What is your research question?

- How can this research question be translated into questions and answers?

Who can answer your research question?

- What‘s the population?
- Can you ask the population? Or do you have to draw a sample?
- If you can only ask a sample, how should you choose it?

Who is the population?

How can we collect data about unit and department leaders?

Hypothesis: There is a difference between unit leaders and department leaders in their job satisfaction.

How to test difference between means?
see Data Analysis

How to measure job satisfaction?

Consider the unit of analysis

- **Individuals:** e.g. How does personality facilitate innovative behavior?
- **Groups:** e.g. How does team composition facilitate innovative team behavior?
- **Organizations:** e.g. How does absorptive capacity facilitate innovative performance?
- **Interactions:** e.g. How does Leader-Member exchange influence task performance?
- **(Social) artifacts:** e.g. Does the tenor of announcements of product retraction affect media coverage about firms?

Theory informs the level of analysis (and vice versa)

Population vs sample

Population ≠ world population

Population = target population, i.e., all people that are interesting to answer your research question, e.g., innovation managers, all German-speaking students etc.

- It‘s very rare to collect data from the whole population.
- Usually, you draw a sample.
- How a sample is drawn is crucial for the generalization of the results.

Sampling strategies

Research question: How do Germans think about the elections in East Germany?

- **Random sampling**
  - Select subjects from all over Germany (registration office, phone book?)
- **Stratified random sampling**
  - Divide German man and women, then select subjects randomly from these two groups
- **Cluster sampling**
  - Select subjects from the five biggest cities in Germany.
- **Multi-stage sampling**
  - Divide population in clusters or layers. Clusters/layers are randomly drawn and subjects from these clusters/layers are randomly drawn.

Sample characteristics of representativeness

- Definition of population
- Sampling procedure
- Distribution of important attributes of population in the sample

What can be done to increase representativeness?

- Correct distortions caused by software programs
- Weight answers
- Draw conclusions with great care
- Non-response bias tests
- Find out why companies did not respond

Exercise: How to get a representative sample?

Please think about how to identify a representative sample for the following populations:

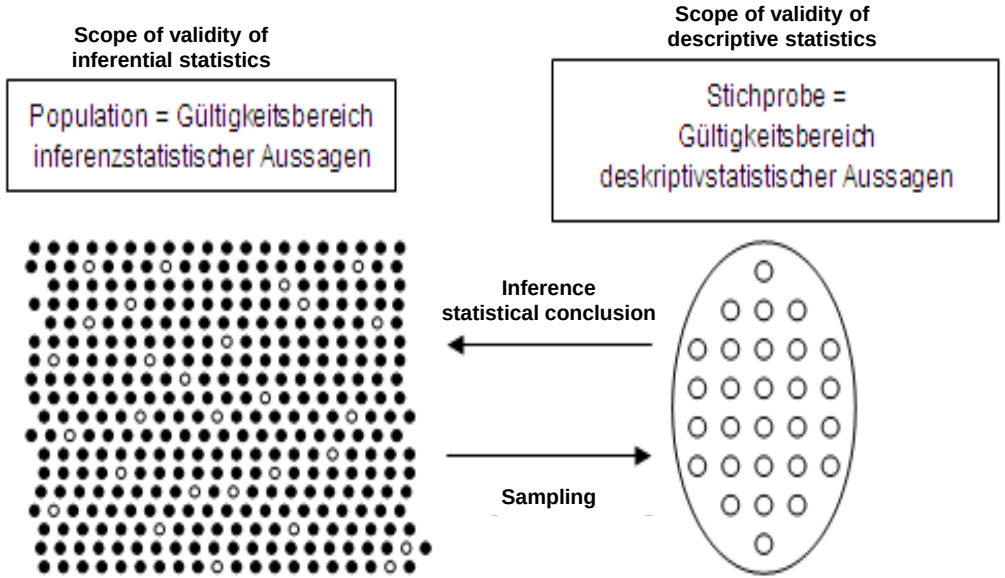
... venture capital funded companies

... Nanotechnology companies in Germany

... Executives

Question to you:

What do you respond to someone who says to you that your research is bad because you have used a so-called "convenience sample"



Sources of error in sampling

Random Error
(sampling error)

- probability calculation can handle this type of error
- Unavoidable error in sampling
- Will lower with larger sample size

Systematic Error
(non-sampling error)

- probability calculation can handle this non-error type
- Avoidable
- Will *not* lower with larger sample size

Beller (2004)

What is shown here?

Intelligence of men between 18 and 22 years.

- Unequal distribution across regions
- Be aware when sampling!