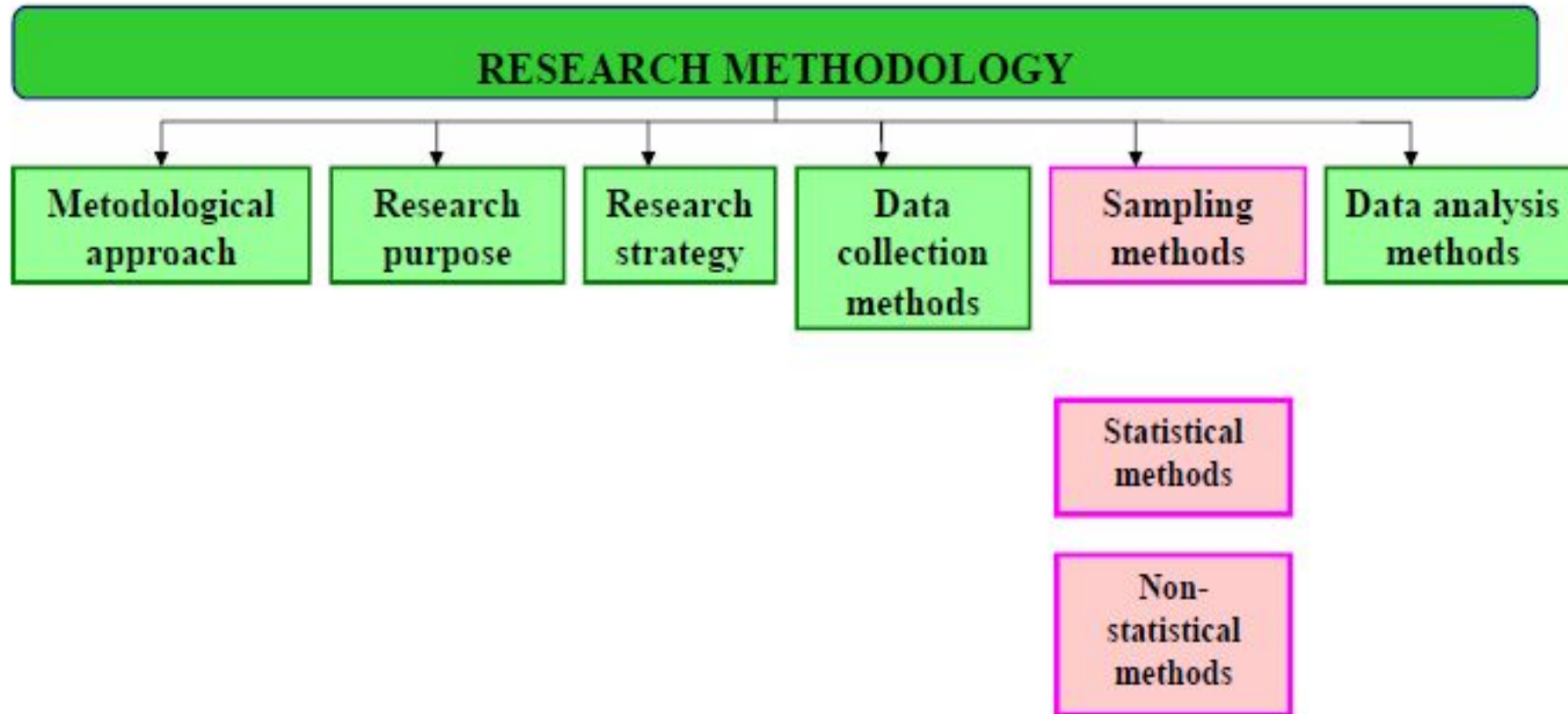# Sampling

Sampling

Table of content

- population, sample, representative sample
- sampling methods:
  - statistical (probability sampling)
  - non-statistical (non-probability sampling)
- sample size
- sampling plan

Population

- group of people (e.g. customers, consumers, business partners) or objects (companies, shops, institutions) that have something in common
- total group from which information is needed

Example

- All Facebook users over 18 years who live in the Czech Republic and are Star Wars fans.
- All Czech people over 15 years old who have travelled to Croatia for a holiday during the last 2 years.
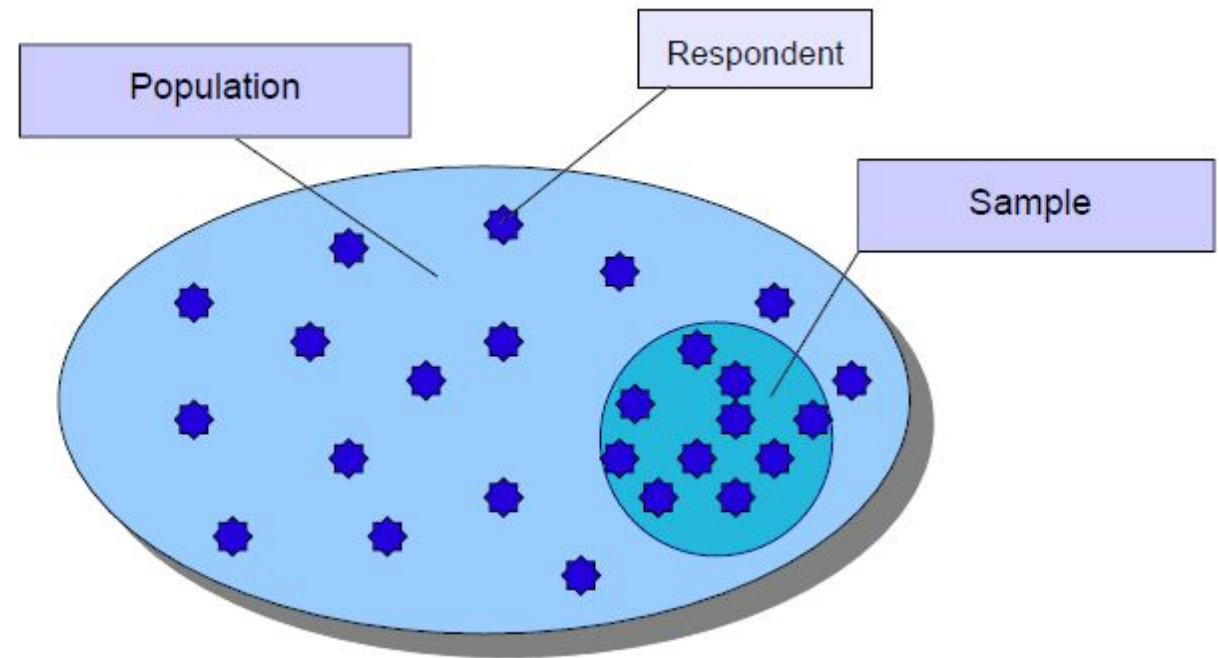- All students of gastronomy high schools in the Liberec Region.

Census

- studies involve the population overall
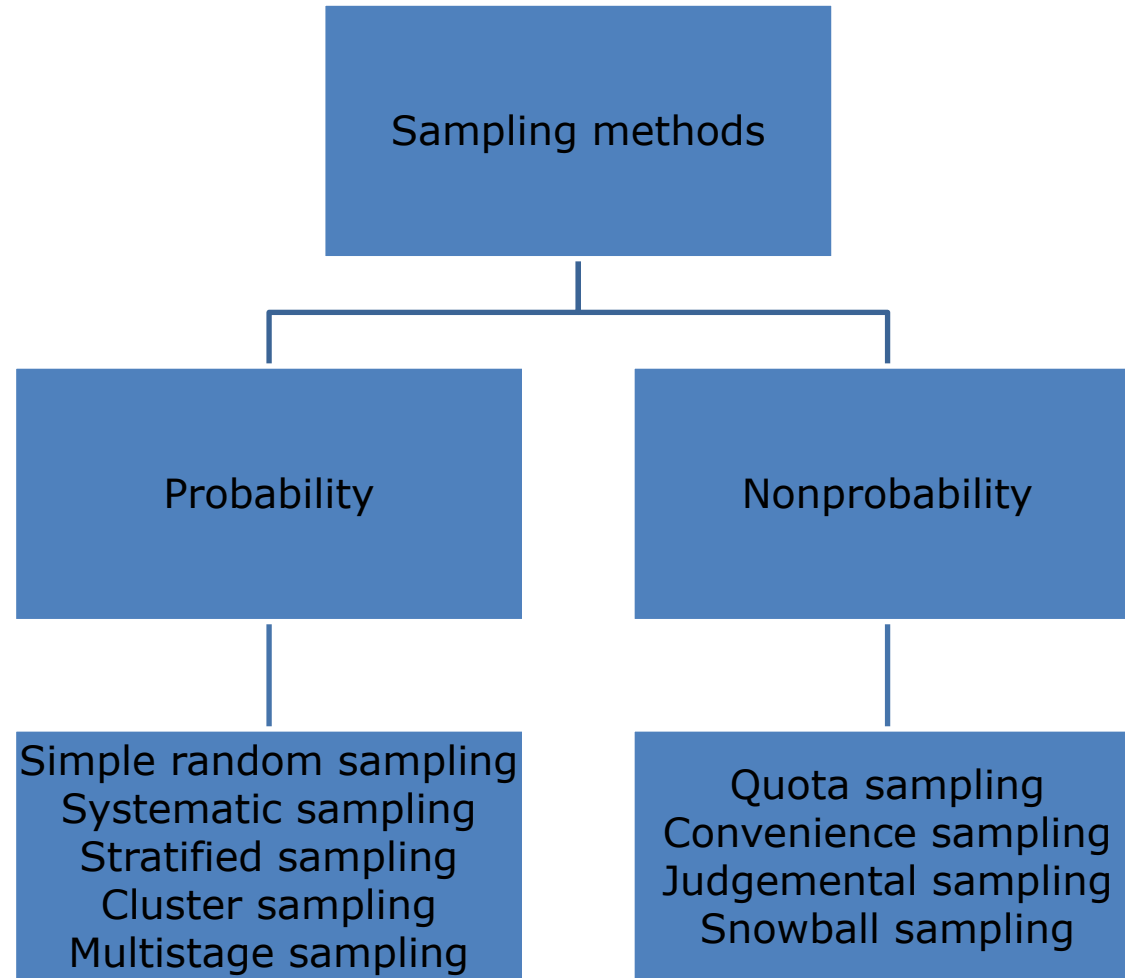
- expensive, rather slow

Sample

- subset of a larger population of interest

- **respondent** – source of required information

- **sampling frame**

- sample must be representative!

Process of sampling

1. Population definition
2. Sampling frame
3. Sampling methods
4. Sample size
5. Sampling plan
6. Choosing a sample (contacting and interviewing respondents)

Sampling methods

```
                    ┌─────────────────────┐
                    │  Sampling methods   │
                    └─────────────────────┘
                              │
                ┌─────────────┴─────────────┐
        ┌───────────────┐           ┌───────────────┐
        │  Probability  │           │ Nonprobability│
        └───────────────┘           └───────────────┘
                │                           │
    ┌───────────────────────┐   ┌───────────────────────┐
    │ Simple random sampling│   │    Quota sampling     │
    │  Systematic sampling  │   │ Convenience sampling  │
    │  Stratified sampling  │   │ Judgemental sampling  │
    │   Cluster sampling    │   │  Snowball sampling    │
    │  Multistage sampling  │   └───────────────────────┘
    └───────────────────────┘
```

Probability sampling

- based on a chance

- the most accurate

- the sample is randomly drawn from the list of the population

- all respondents have the same chance of being selected in the sample

- the results are statistically sounder

- problem: very few lists of respondents exist in marketing

Probability sampling

Advantages

- the most accurate
- the less biased
- every element has the same chance to be drawn
- ideal for statistical processing

Disadvantages

- sampling frame needed
- difficult for data collection
- What if selected elements don't want to participate?

Probability sampling

- Simple random sampling
- Systematic random sampling
- Stratified sampling
- Cluster sampling
- Multistage sampling

Simple random sampling

- table of random numbers, lottery drawing
- suitable for a homogenous population

Examples: secondary schools, companies in the automotive industry, customers from a database
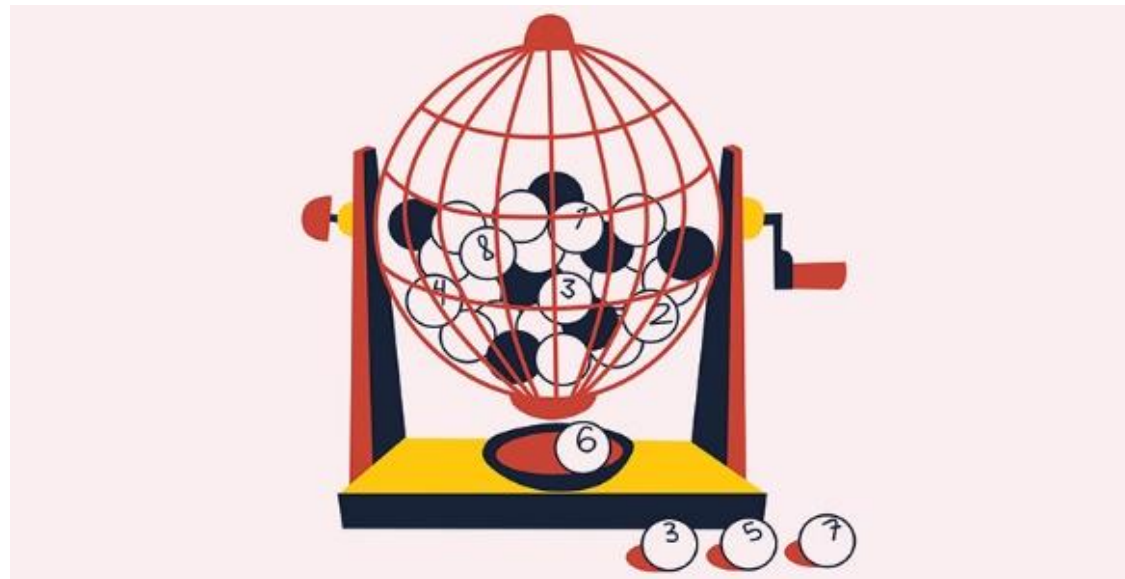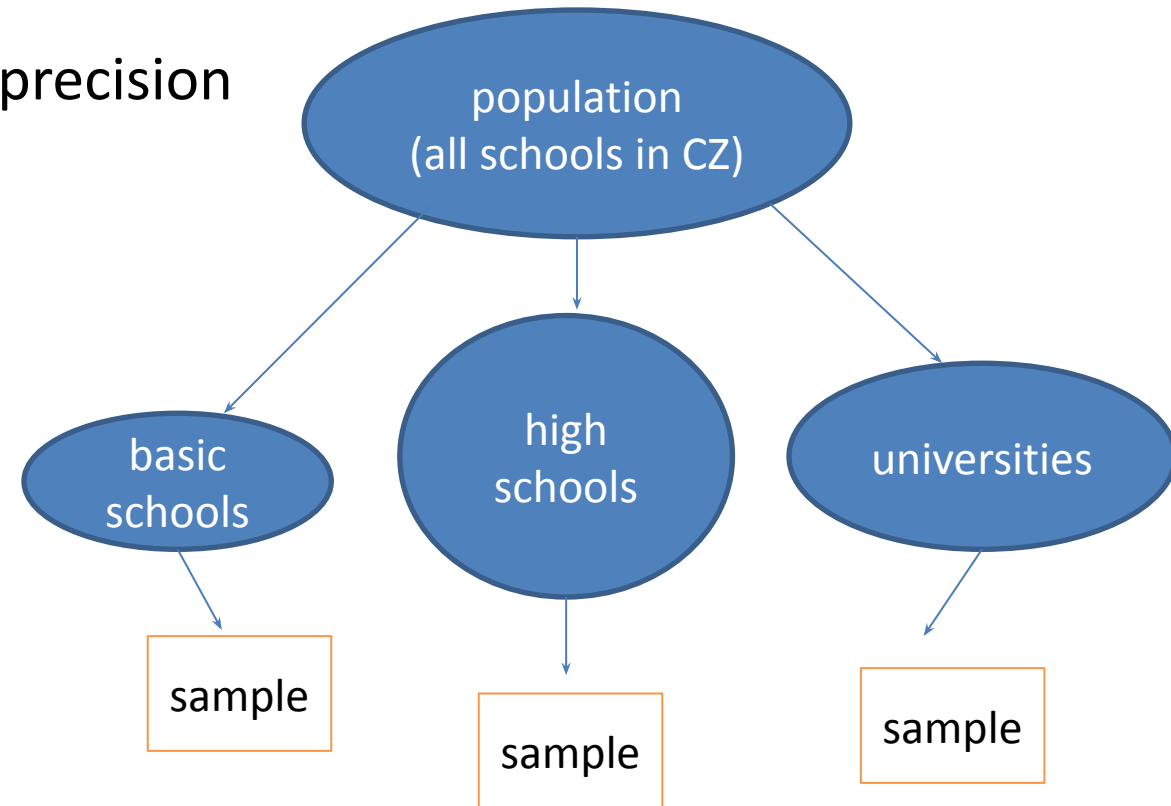
Systematic sampling

- the first respondent is randomly selected, further within the same intervals (every n^{th} member of the population)

- sampling without the biased choice

- formula for the interval = size of the population / size of the sample

Population

Sample (every 3rd)
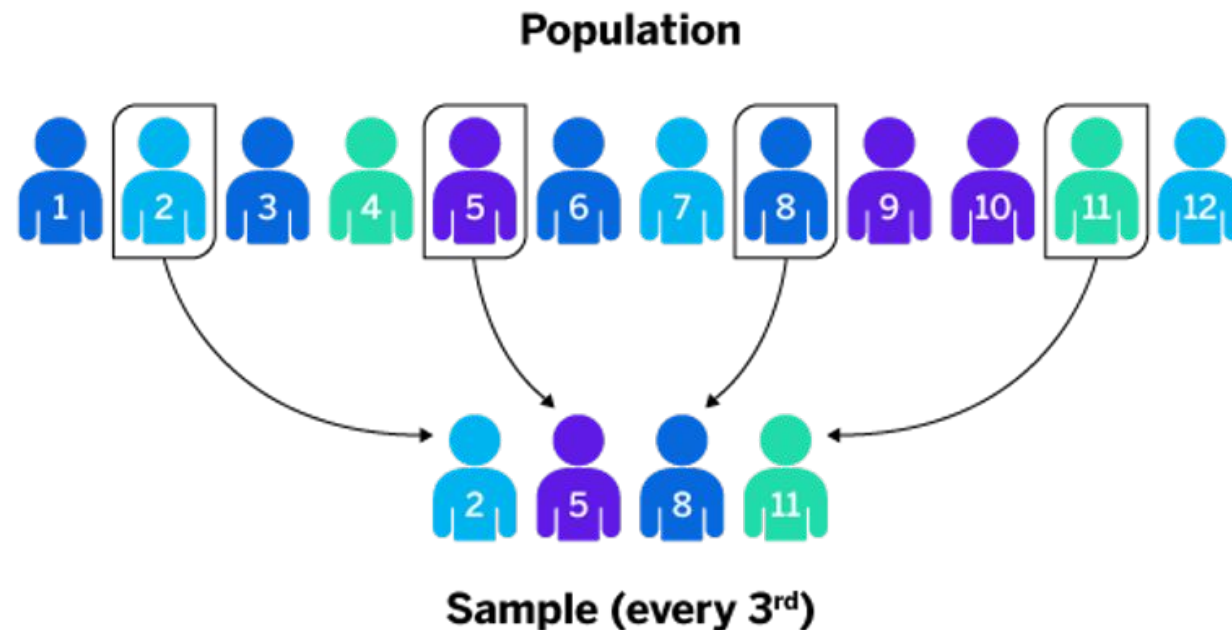
Stratified sampling

- two steps:
  1. population divided into homogenous sub-populations (= strata) – each stratum more or less equal on some characteristic
  2. simple random elements are chosen from each group - proportionate or disproportionate
- objective: to reduce costs without loss of precision

Example: customers in towns and villages

```
          population
        (all schools in CZ)
         /      |        \
  basic      high      universities
  schools    schools
    |          |            |
  sample     sample       sample
```

Cluster sampling

- two steps:
  1. population divided into sub-groups (= clusters - formed as heterogeneous)
  2. random sample of cluster(s) is selected

- objective: to increase sampling efficiency by decreasing costs

- example: village, school class, town, interest group
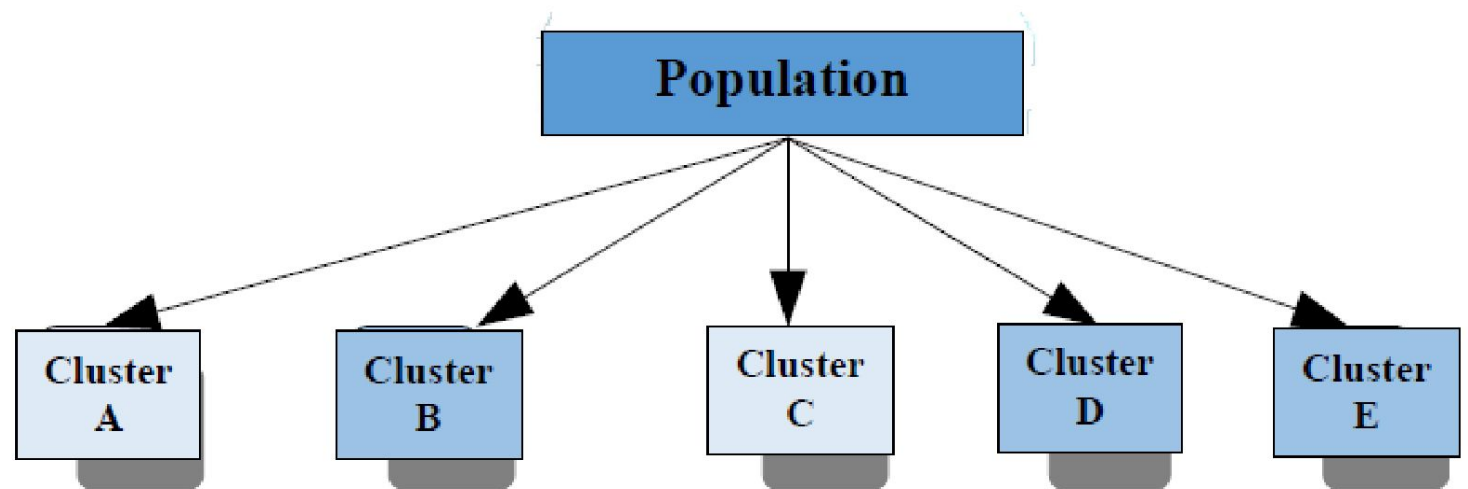
Multistage sampling

- similar to cluster sampling in more steps:

  1. subsampling within the clusters

  2. random sampling within the clusters

- objective: to increase sampling efficiency by decreasing costs
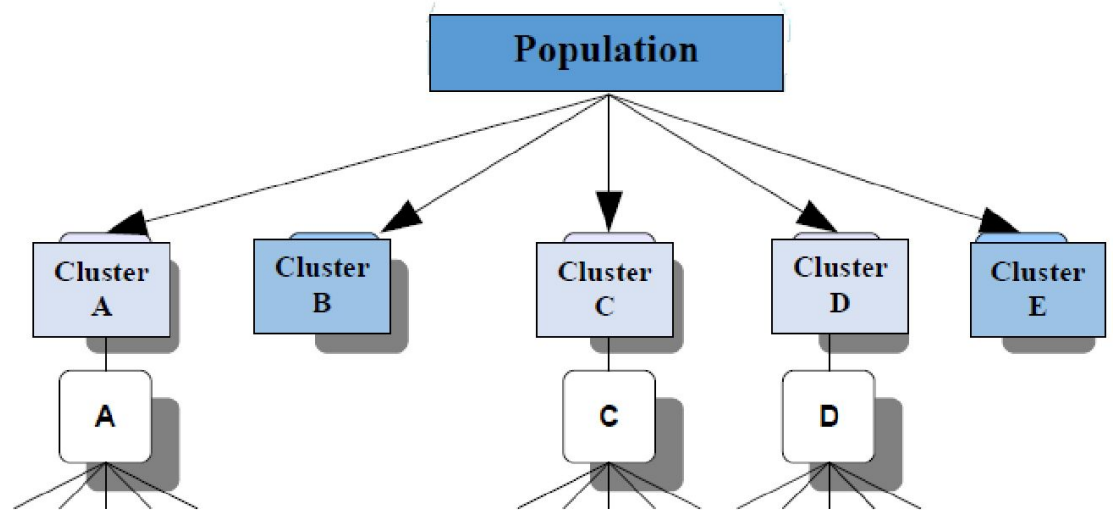
Non-probability sampling

- not based on a chance

- the sample is dependent on human judgement (by needs of the research, experience of experts)

- each unit of the population has a different chance of being selected

- representativeness cannot be guaranteed

- sometimes complicated to apply statistical procedures

Non-probability sampling

Advantages

- the sampling frame is not necessary
- faster
- cheaper
- higher response rate (sometimes)

Disadvantages

- representativeness is not guaranteed (it can be guaranteed on condition of good knowledge of the population)
- high sampling error – the larger the sample, the lower the sampling error
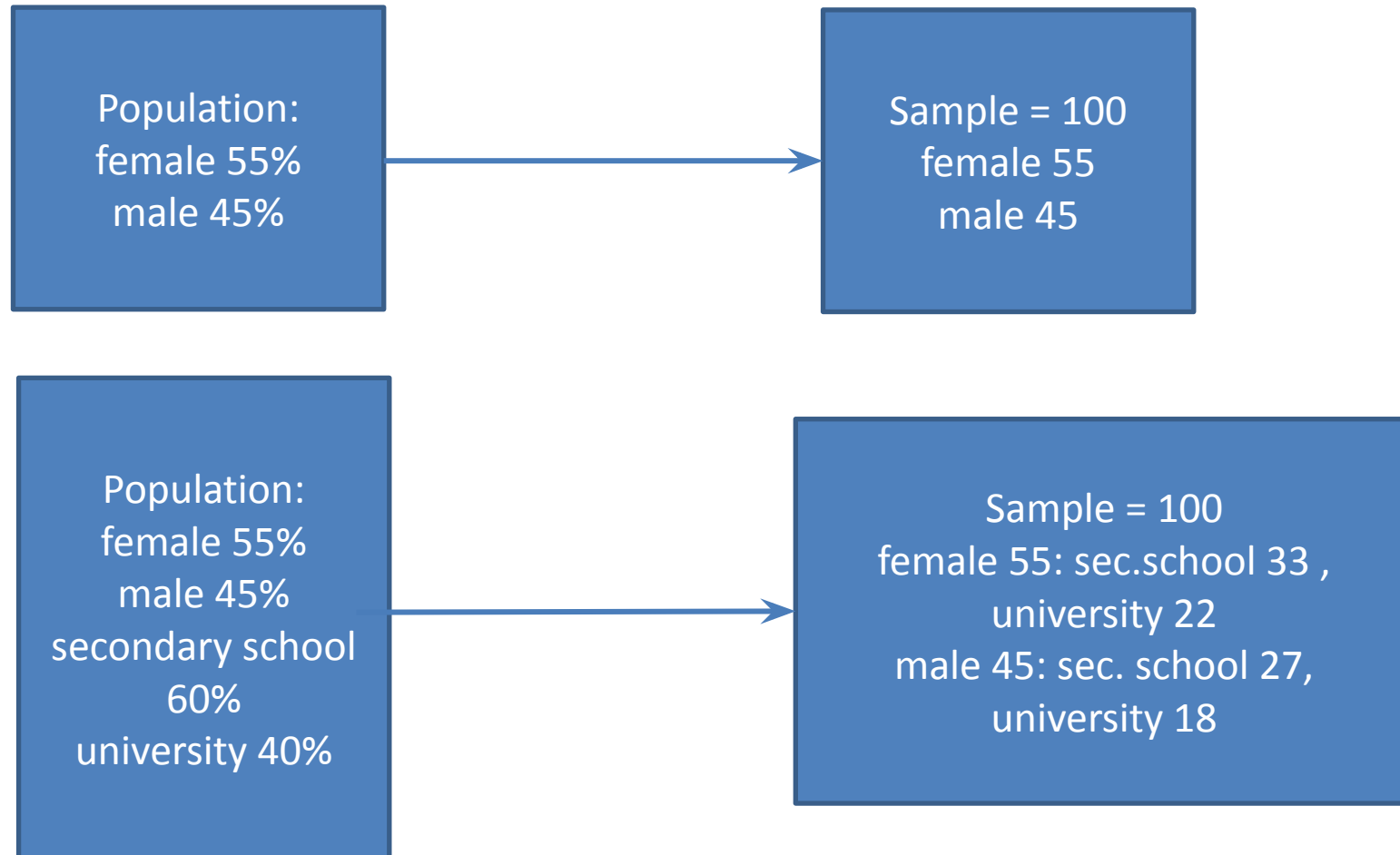
Non-probability sampling

- Quota sampling
- Convenience sampling
- Judgmental sampling (purposive)
- Snowball sampling

Quota sampling

- controlling the number of respondents by one or more **criteria**, **quotas** regarding demographics (e.g. age, income), specific attitudes (e.g. satisfaction level), specific behaviour (e.g. frequency of purchase)
- the similarity between quotas and population
- frequently used
- **knowledge of the population is necessary**
- similar to stratified sampling
- Example: opinion of the Czech population, customers of a brand (without a database)

Quota sampling

Population:
female 55%
male 45%

Sample = 100
female 55
male 45

Population:
female 55%
male 45%
secondary school 60%
university 40%

Sample = 100
female 55: sec.school 33 ,
university 22
male 45: sec. school 27,
university 18

Convenience sampling

- sample is based on ease of access
- respondents are convenient for the researcher to reach (availability at a given time, geographical proximity, willingness to participate, friends…)
- mall intercepts
- used for pilot testing

Judgemental sampling

- purposive sampling
- selection according to an experienced researcher's belief
- sampling error, low representativeness
- cheaper, less time-consuming
- example of application: pedestrians in the streets, tourists at tourist sights, friends and coworkers

Snowball sampling

- referral sampling
- selection of initial group => interview => nomination of others in target population
- used: if it is rare and unique to define the target population
- examples of application: university students, scientists, car fans, fishermen

Snowball Sample
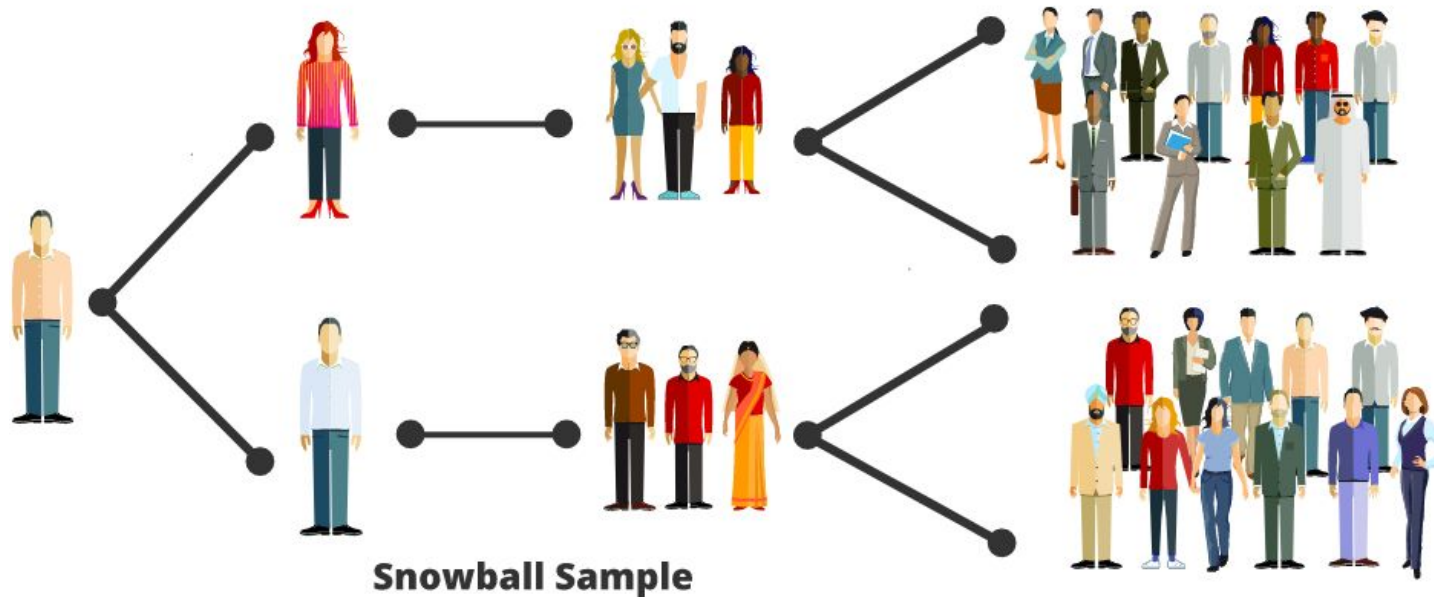
Sample size

Statistical methods

- formulas
- information about the population is necessary

Non-statistical methods

- by judgement or experience of the researcher
- by cost or budget, we can afford or want to spend on the research
- by the size of the sample of the similar research conducted in the past
- by access to respondents

Sampling plan

- **technical and personnel requirements**

- the most appropriate location for conducting a survey (households, companies, sites for questioners)

- schedule for the data collection (days and daytime, time span)