

AdTrans 2023-1-PL01-KA220-HED-000158917



# Methods of sociology and demography

## Research phase

Kristyna Rybova



Co-funded by  
the European Union



Co-funded by the European Union. Views and opinions expressed are however those of the author or authors only and do not necessarily reflect those of the European Union or the Foundation for the Development of the Education System. Neither the European Union nor the entity providing the grant can be held responsible for them.



# RESEARCH PHASE

Information preparation and organization

Research topic, research problem

Research objective

Research questions

Hypotheses

Variables

Research file

Methods

Data processing

Data interpretation

# RESEARCH PHASE

## Information preparation and organization

- Research has several stages that both follow each other and overlap in time - think, organize, plan!
- Research activities can last several weeks and months, and in larger research projects, even several years.
- It allows you to navigate the issue: you can find out what has already been researched on a given or similar issue, what was found, and how it was researched.
- The theoretical foundations necessary for specifying the research problem are obtained.
- Common mistakes
  - Underestimating the time required for research
  - Collecting data in the field that the researcher cannot process
  - Missing data to meet the goal

# STAGES OF QUANTITATIVE RESEARCH

Time interval	1	2	3	4	5	6	7	8
Determining the topic, research problem, objectives and questions, hypotheses	Dark Blue	Dark Blue	Dark Blue	Dark Blue	Light Blue	Light Blue	Light Blue	Light Blue
Information preparation	Dark Blue	Dark Blue	Dark Blue	Dark Blue	Light Blue	Light Blue	Light Blue	Light Blue
Preparation of research methods	Light Blue	Dark Blue	Dark Blue	Dark Blue	Dark Blue	Light Blue	Light Blue	Light Blue
Data collection and processing	Light Blue	Light Blue	Light Blue	Light Blue	Light Blue	Dark Blue	Dark Blue	Light Blue
Data interpretation	Light Blue	Light Blue	Light Blue	Light Blue	Light Blue	Light Blue	Dark Blue	Dark Blue
Publications (including research reports)	Light Blue	Light Blue	Light Blue	Light Blue	Dark Blue	Dark Blue	Dark Blue	Dark Blue

# RESEARCH PHASE

## Determining the research topic

- Research subject
- Errors in formulating the topic
  - Outside the area
  - Too wide
- Proposal of current topics (originality, usefulness, attractiveness, importance)
- Feasibility of research

# RESEARCH PHASE

## Determining the research topic/problem

- If I want to solve a problem, I need to know exactly what it is.
- A research problem as a set of questions with the following functions (Punch, 2008):
  - They organize the project and give it direction and coherence
  - They define the project and determine its boundaries
  - They keep the researcher focused on their own project goals
  - They point to the data that will be needed for the project
  - Manages data collection

# RESEARCH PHASE

## Research project

- WHAT? – goals, main tasks to achieve the goal, formulation of hypotheses
- WHERE? – to what population will we generalize the results – spatial definition
- WHEN? – research timeframe
- HOW? – collection methods and techniques
- WHO? – who will participate in the research
- WHY? – why is the research topic important?
- FOR HOW MUCH? – financial budget



# RESEARCH PHASE

## Research objective

- What do we want to achieve with the research investigation?

# RESEARCH PHASE

## Determining research questions

- Research questions specify the research topic
- Formulated in an interrogative sentence so that they can be answered at the end of the research investigation
- When formulating them:
  - Appropriate to our work (BP, DP, GAČR project, SGS, EU funds, etc.)
  - "what" and "how"
  - One or two main questions supplemented by a maximum of five sub-questions are enough.

# RESEARCH PHASE

## Hypothesis determination

- Hypothesis = scientific assumption (principle, basis); a statement whose validity is only assumed, but at the same time it is formulated in such a way that it can be confirmed or refuted.
- It is determined at the beginning of the research (it is not possible to start collecting data in the field and establish hypotheses during it)
  - Research problem → hypotheses
  - Theory → hypothesis → acceptance of the theory
  - Theory → hypothesis → refutation of the theory → new theory
- A hypothesis is by no means just any assumption, a guess, or a "blind shot"; it must be based on knowledge that is known about the phenomenon under study, or on the researcher's practical experience, which we confirm or refute.

# RESEARCH PHASE

## Hypothesis determination

- VP: What is the impact of an individual approach to clients on their satisfaction with the services provided?
- Hypotheses: H1: A doctor who has an individual approach to clients achieves higher satisfaction with the services provided by his/her clients than a doctor without an individual approach.

# RESEARCH PHASE

## Hypothesis determination – rules:

- 1. A hypothesis is a statement, expressed in a declarative sentence (on the contrary, the VP is in the interrogative form). At the end of the research, we must accept this statement (it is true) or refute it (it is not true).
- 2. A hypothesis expresses a relationship between two variables.
- 3. The hypothesis must be testable (empirically researchable) – its variables must be measurable or categorizable

# RESEARCH PHASE

Examples of hypotheses:

H1: Students of an elementary school teacher who has a non-directive teaching style have a more positive attitude towards their subject than those of a teacher who has a directive teaching style.

H2: We assume that adequate forms of cooperation between school and family are not always used.

# RESEARCH PHASE

Hypotheses most often express:

- differences – more, more often, stronger, differently

For example, second-grade students prefer school to fifth-grade students

- relationships – positive relationship, negative relationship, correlation

For example, there is a positive correlation between parental education and a child's intelligence.

- consequences – if – then, how – so, by what – thereby

For example, if a teacher increases the number of praises given to students, then students will learn more.

# RESEARCH PHASE

Variables

Character of variables

- distinguishable condition
- completeness condition
- condition of uniqueness

Variables

- Nominal
- Ordinal
- Cardinal
- different statistical procedures can be used for each type of variable



# RESEARCH PHASE

Data

- primary
- secondary

# NOT STATISTICS AS STATISTICS

## **Descriptive statistics**

Research questions, not hypotheses (or descriptive hypotheses)

Used to describe a research file

## **Inferential statistics**

Research hypotheses and their verification

It is used to generalize conclusions to the population

## **Primary analysis**

Own research, own data

## **Secondary analysis**

Using the results of previous research and their data

Open Access, Open data (Horizon 2020)

## **Meta-analysis**

# WHAT ABOUT THE DATA?

Coding

Cleaning

1st degree sorting

Character editing

2nd stage sorting

Statistical analyses (factor, cluster, correlation, etc.)

Generalization (hypothesis testing)

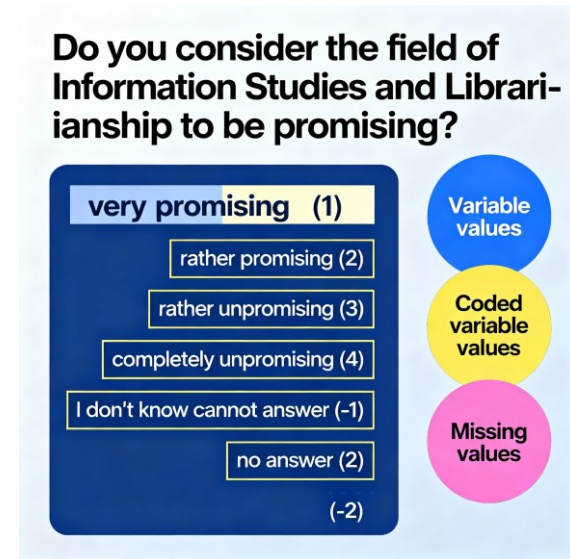
# CODING

## Data matrices

- Respondents = rows
- Variables = columns

## Working with the encryption key

- Directly in the field
- Online application – in the background
- Manual coding



Spokojenost s nabídkou kurzů						
	Velmi souhl. asím	Spíše souhl. asím	Ani souhl. asím ani nesouhl. asím	Spíše nesouhl. asím	Vůbec nesouhl. asím	Nevím / nemohu odpovědět
Povinné (A) kurzy mají logickou časovou posloupnost.	1	2	3	4	5	-1
Obsahy jednotlivých povinných (A) kurzů se nepřekrývají.	1	2	3	4	5	-1
Jsem spokojen/a s tematickou šíří nabídky povinně volitelných (B) kurzů.	1	2	3	4	5	-1
Jsem spokojen/a s počtem nabízených povinně volitelných (B) kurzů.	1	2	3	4	5	-1

# FIRST-LEVEL GRADING

For each character, the frequency of occurrence is identified

Output: **frequency table**

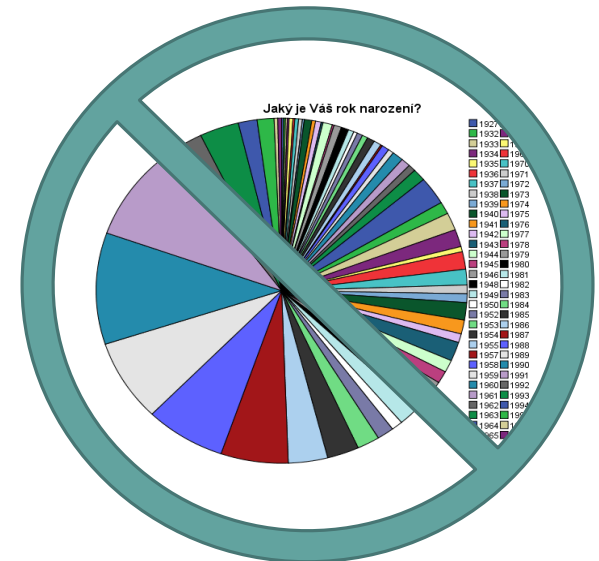
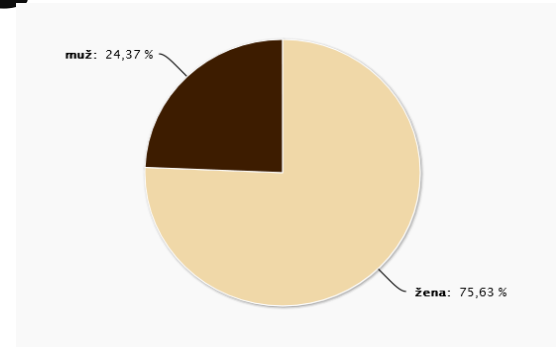
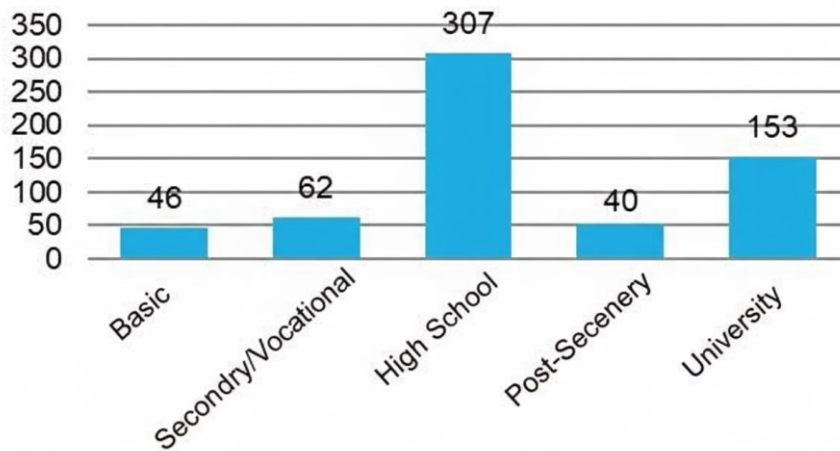
What is your education?				
		Response frequency	Relative frequency	Relative frequency
Valid values	Basic	46	7.5	7.6%
	Basic education/secondary school graduation	62	10.1%	10.2%
	High school with graduation certificate	307	50.1%	50.5%
	Post-secondary education VOŠ	40	6.5%	25.2%
Missing values (don't know, didn't answer)	<b>Total valid values</b>	608	99.0%	100.00
	Missing values	5	0.8%	
<b>Total</b>	613	5	100.00	

# FIRST-LEVEL GRADING

For each character, the frequency of occurrence is identified

Output: frequency plots

## What is your education?





# AVERAGE VALUES

Mode

Median

Arithmetic mean (+ standard deviation)



# CHARACTERS EDITIONS

Categorization

Coding open and semi-open questions

# SECOND-LEVEL GRADING

We are looking for connections between variables. Outputs: contingency tables

1. Relative frequencies are important to us.
2. We need to determine which variable we consider dependent and which independent.
3. If the independent variable is in columns, we compare column percentages. If the independent variable is in rows, we compare row percentages.

# DATA ANALYSIS TOOLS

## Online tools

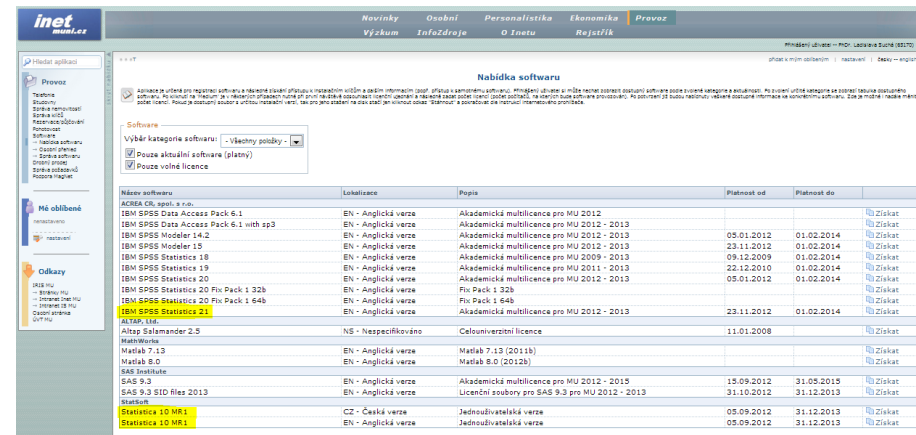
Survs, SurveyMonkey, Survio, ...

## Common spreadsheets

MS Excel, Open Office Calc or Google Spreadsheets

## Special desktop tools for statistical analysis

SPSS, Statistica, R, ...



The screenshot shows the 'Nabídka softwaru' (Software Offer) page on the inet website. It features a search bar, a filter for 'Všechny softwary' (All software), and a table of software products. The table includes columns for 'Název softwaru' (Software Name), 'Lokalizace' (Localization), 'Popis' (Description), 'Platnost od' (Valid from), 'Platnost do' (Valid until), and 'Získat' (Get).

Název softwaru	Lokalizace	Popis	Platnost od	Platnost do	Získat
ACRBA C&I, spol. s r.o.					
IBM SPSS Data Access Pack 6.1	EN - Anglická verze	Akademická multilicence pro MU 2012			Získat
IBM SPSS Data Access Pack 6.1 with sp3	EN - Anglická verze	Akademická multilicence pro MU 2012 - 2013			Získat
IBM SPSS Modeler 14.2	EN - Anglická verze	Akademická multilicence pro MU 2012 - 2013	05.01.2012	01.02.2014	Získat
IBM SPSS Modeler 15	EN - Anglická verze	Akademická multilicence pro MU 2012 - 2013	23.11.2012	01.02.2014	Získat
IBM SPSS Statistics 18	EN - Anglická verze	Akademická multilicence pro MU 2009 - 2013	09.12.2009	01.02.2014	Získat
IBM SPSS Statistics 19	EN - Anglická verze	Akademická multilicence pro MU 2011 - 2013	22.12.2010	01.02.2014	Získat
IBM SPSS Statistics 20	EN - Anglická verze	Akademická multilicence pro MU 2012 - 2013	05.01.2012	01.02.2014	Získat
IBM SPSS Statistics 20 Flo Pack 1 32b	EN - Anglická verze	Flo Pack 1 32b			Získat
IBM SPSS Statistics 20 Flo Pack 1 64b	EN - Anglická verze	Flo Pack 1 64b			Získat
IBM SPSS Statistics 21	EN - Anglická verze	Akademická multilicence pro MU 2012 - 2013	23.11.2012	01.02.2014	Získat
ALTAIP, s.r.o.					
Altap Šalamander 2.5	NS - Nespecifikováno	Celouniverzitní licence	11.01.2008		Získat
MathWorks					
Matlab 7.1.3	EN - Anglická verze	Matlab 7.1.3 (2011b)			Získat
Matlab 8.0	EN - Anglická verze	Matlab 8.0 (2012b)			Získat
SAS Institute					
SAS 9.3	EN - Anglická verze	Akademická multilicence pro MU 2012 - 2013	18.09.2012	31.05.2015	Získat
SAS 9.3 SID Files 2013	EN - Anglická verze	Licenční soubory pro SAS 9.3 pro MU 2012 - 2013	31.10.2012	31.12.2013	Získat
StatSoft					
Statistica 10 MRI	CZ - Česká verze	Jednoudavatelská verze	05.09.2012	31.12.2013	Získat
Statistica 10 MRI	EN - Anglická verze	Jednoudavatelská verze	05.09.2012	31.12.2013	Získat