Re-Educating Digitisation
Visitor Research

RED
Visitor Research I

- I. Introduction -

Author:
Prof. Dr. Tibor Kliment

RED
Re-Educating
Digitisation

Rheinische
Hochschule
Köln

Tibor Kliment

Since 2001 Professor for
Empirical Media and
Culture Marketing at
Univ. of Applied Sciences Cologne (RFH)

- Before audience research for public broadcasters for (ARD, RTL), head of department at EMNID Inst./Media Department, Managing Director at a Consulting Institute IFEP

- Focussing on marketing research, audience development, media management, communication planning

- More than 30 years of experience in applied marketing-/media-/audience research

RED
Re-Educating
Digitisation

Rheinische
Hochschule
Köln

Topics of the workshop I

Section 1

I. Introduction
- Participant introduction
- Why we do visitor research

II. The Research process
- From the questions to the data:
Preparing, planning and conducting research

III. Designing questions and questionaires I
- **Types and functions of questions, scaling, do´s and don´ts**

Section 2

IV. Designing questions and questionaires II
- Development of questionaires, dramaturgy

V. Theme packages in visitor studies
- Demographics, visitor/visit information, ticketing/distribution, communication

VI. Reflecting on participant´s research and projects

Section 3

VII. Overview about research methods: Quantitative and qualitative methods

VIII. Preparing the next workshop
- Organizing the visitor research project in Turin
- Developing the questionaire for Turin

A short introduction:

1. Please introduce yourself shortly: Name, company business, size of staff

2. Have you done visitor research in your institution already?
   - Since when?
   - How often?
   - What were main topics?

3. What did you do with the results

4. What were the benefits/deficits for your festival from these studies?

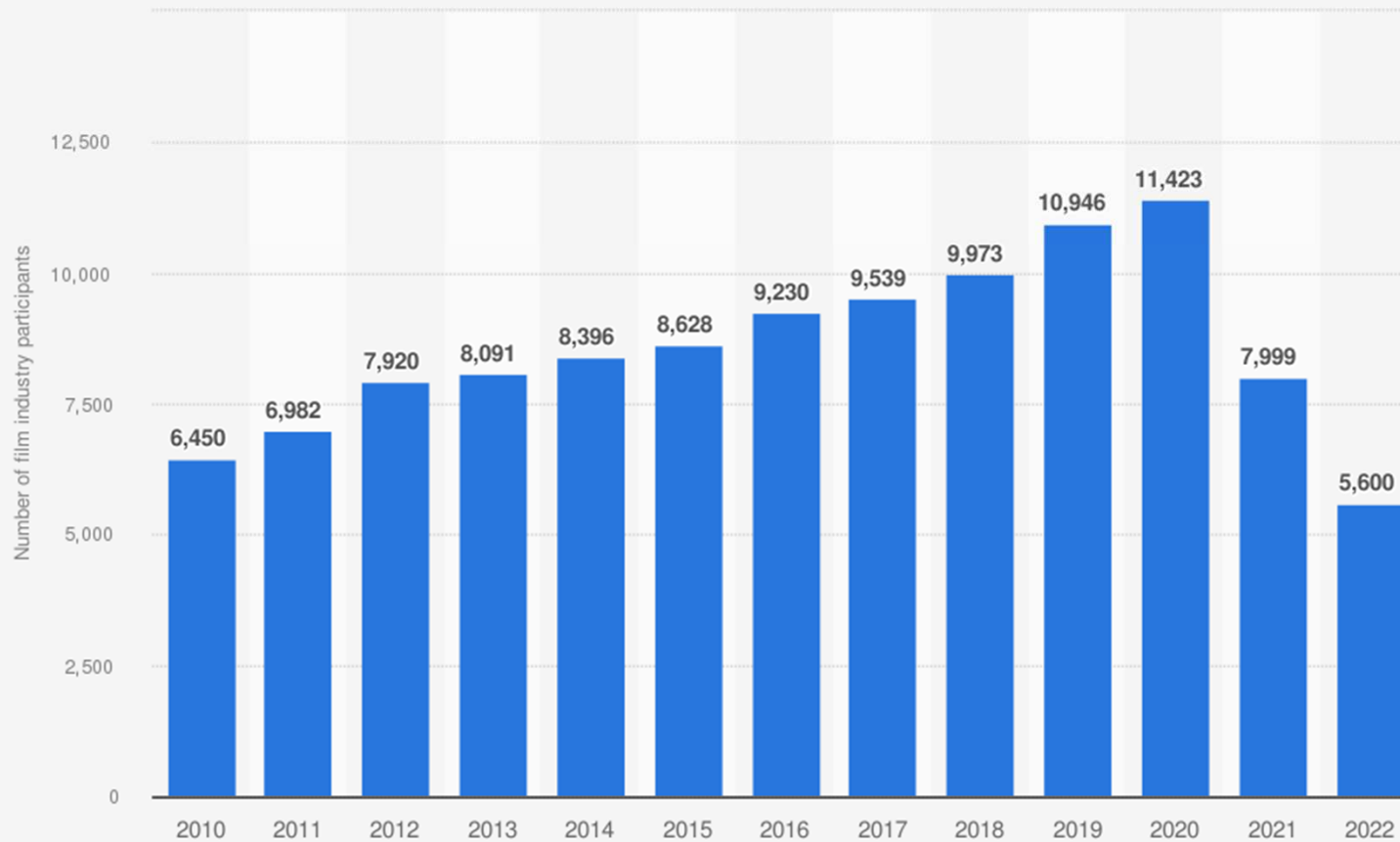5. Are there special subjects you would like to discuss in this workshop?

General trend: Culture is under pressure

- The importance of visitor research has increased due to the growing challenges that face cultural institutions:
    - They today compete for their visitors' financial and time budgets with a strongly growing variety of leisure activities; streaming services enhance competition dramatically
    - Growing legitimacy pressure and decreasing public funding
    - Corona has caused in some countries a lasting decrease in audience numbers in all areas (in others not)
    - Economic decline and inflation reduces financial budgets on the side of visitors, sponsors and media companies
- The social differentiation in cultural participation has increased: age, gender, life-styles, and people's formal education shape attendance. Target groups become smaller are much more difficult to grasp than in the 1990´s.
- Adressing the younger audience is the most urgent challenge: Today younger people are less likely to turn to high culture offerings than formerly.
- It is not self-evident that the younger generations will turn to cultural offers at some point as their life cycle progresses.
- The culture sector predominantly lives by the generation 50+

Visitor numbers after COVID (I)

Number of film industry participants at the Berlin International Film Festival in Germany from 2010 to 2022

Number of film industry participants

Year	Participants
2010	6,450
2011	6,982
2012	7,920
2013	8,091
2014	8,396
2015	8,628
2016	9,230
2017	9,539
2018	9,973
2019	10,946
2020	11,423
2021	7,999
2022	5,600

Source
Berlinale
© Statista 2022

Additional Information:
Germany; 2010 to 2022

Co-funded by
the European Union

RED
Re-Educating
Digitisation
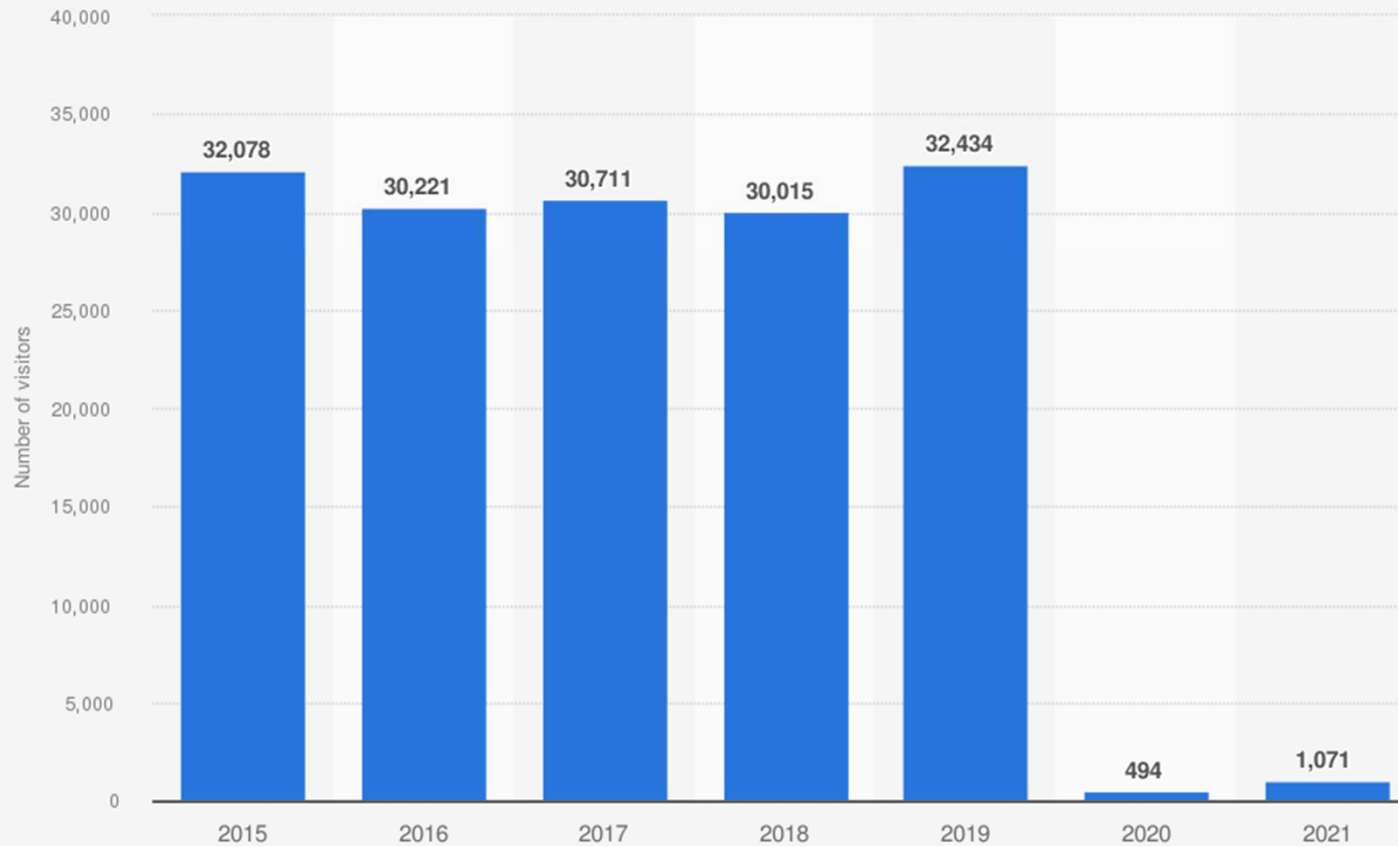
Rheinische
Hochschule
Köln

Visitor numbers after COVID (II)

Number of visitors to the Sodankylä Film Festival in Finland from 2015 to 2021

Visitor numbers after COVID (III)

Cinema visiting frequency compared to situation prior to Covid pandemic in France in 2022

- More often 9%
- Less often 39%
- No change 52%
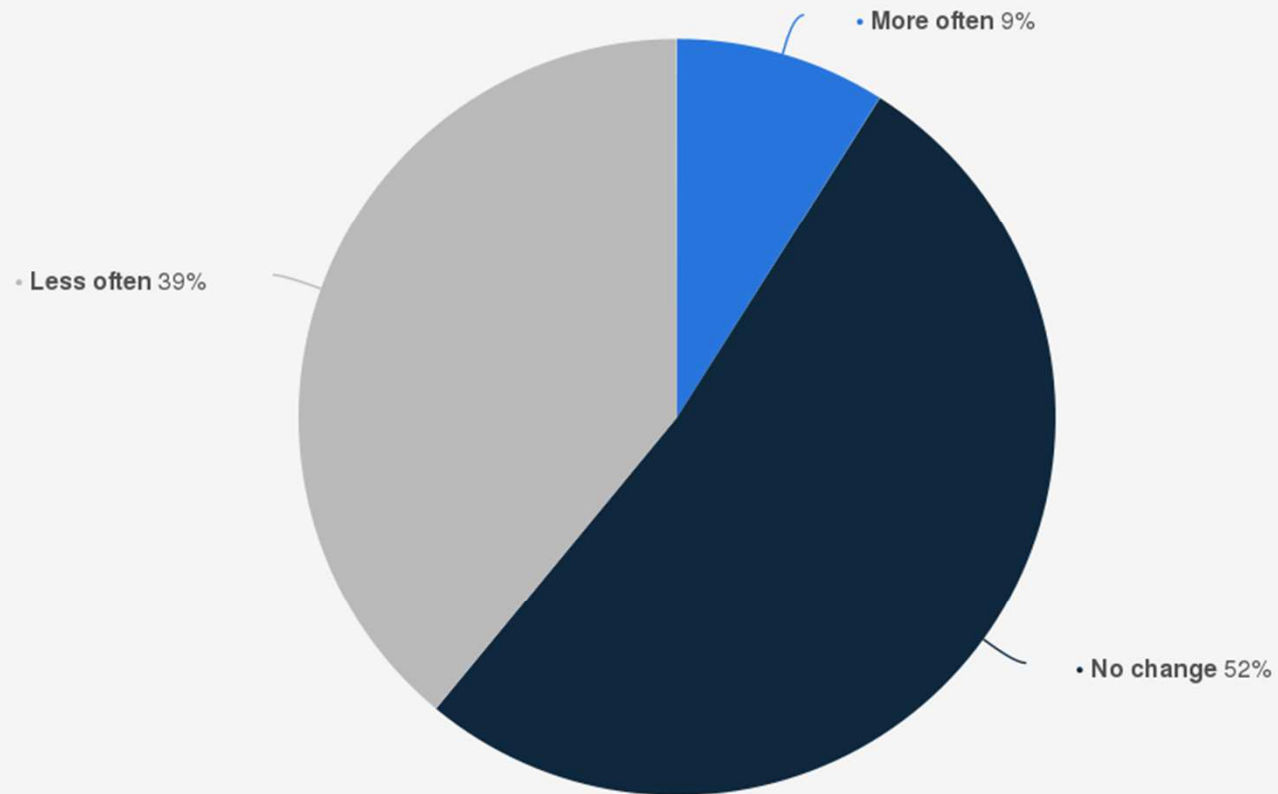
Source
Odoxa
© Statista 2022

Additional Information:
Belgium; France; Germany; Italy; Spain; United Kingdom; April 27 to 28, 2022**; 1,005*; 18 years and older

Visitor numbers after COVID (IV)

Anzahl der Besuche des Zurich Film Festivals von 2010 bis 2022

Anzahl der Besuche

Jahr	Besuche
2010	39.500
2011	51.000
2012	58.000
2013	71.000
2014	79.000
2015	85.000
2016	90.500
2017	98.300
2018	104.000
2019	117.000
2020	68.000
2021	102.000
2022	137.000

Quelle
Zurich Film Festival
© Statista 2022

Weitere Informationen:
Schweiz
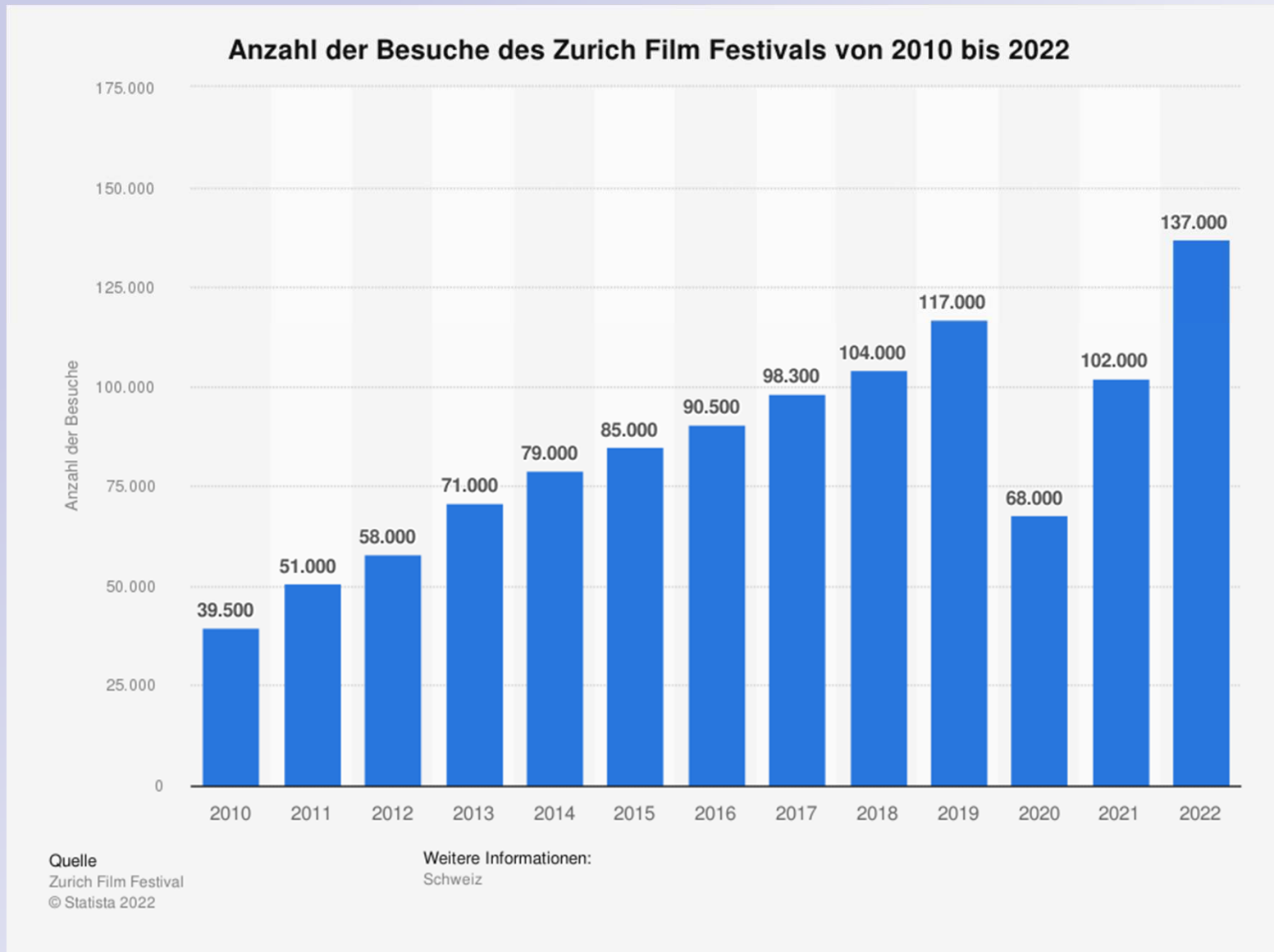
- *„Business saying: What you don´t know, can kill you!"*

Why we do visitor research?

We want to increase ticket turnover and the financial share that is contributed by the visitors. How can we do that?

We want transparency about our visitors and their perspectives on the festival. We want to know who we -don't - reach.

We can show our sponsors the impact of their sponsoring activities, enabling them to give more money, and using their sponsorships in a targeted manner.

Visitors research is helpful für all marketing decision, such as pricing, bundling offers, optimization of service, selection of communication channels.

Visitor research is fun for us, it surprises or confirms, arouses our curiosity. It makes (mis)successes tangible for us.

We simply want professional as well as private visitors to feel comfortable with us. What is the service needed? Do we have to rethink our program offer?

We want to prove the acceptance of our work and move away from sole visitor statistics or pure key figure systems.

Political sponsors need to legitimize the tax money they spent on us. We help them by showing the acceptance of the festival among our audience and in the city, besides visitor statistics.

We to know where our strengths and weaknesses lie from the perspective of our professional visitors.

Private sponsors want to know, who is coming to us. We give them figures about our audience and show them interesting target groups.
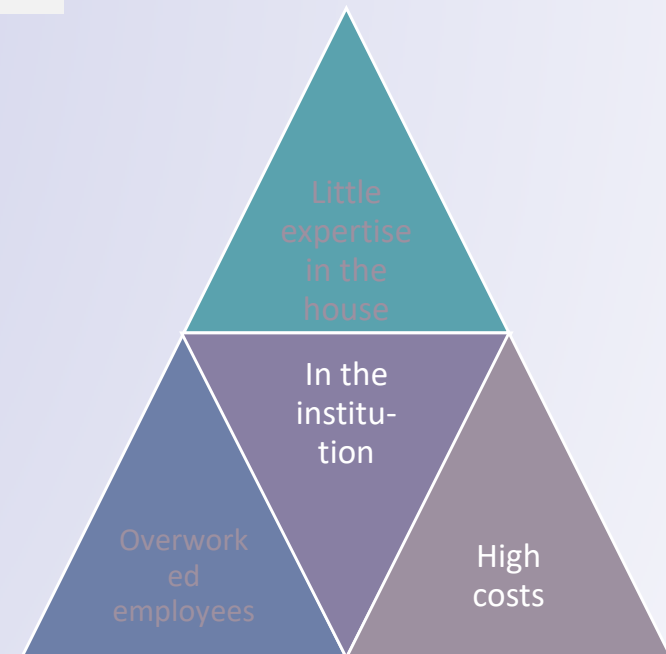
RED
Re-Educating Digitisation

Rheinische Hochschule Köln

But visitor research is not a foregone conclusion…

In a german study on experiences of cultural institutions with their visitor research: 40% show very positive experiences, 53% promised more, 7% were disappointed (Paatsch 2015).

Common problems of visitor research:

- Lack of clarity or disagreement about what exactly should be studied : Audience, promotion, ticketing, service, programme issues? Everything?

- Overloaded questionnaires, imprecisely elaborated questions, lack of scientific standards

- Respondent numbers are too small and/or not representative in the results.

- Results are little meaningful, patchy and not implemented, partially implemented or implemented too late.

- Lack of marketing knowledge as an neccessary backgroung für usefull research

- In the institution no one feels really responsible for implementing the results

Little expertise in the house

In the institu-tion

Overwork ed employees

High costs

RED
Re-Educating
Digitisation

Rheinische
Hochschule
Köln

What is the outcome for you:

After experiencing this workshop:

- You know about the entire research process in doing visitor research

- You know about the most important methods of data collection, especially interviewing techniques

- You know how to develop valid questionaires

- You know what questions to pose for what purpose

After experiencing all workshops:

- You are able to plan and manage the entire research process and you know how to choose the right research design for what purpose

- You can do data analyses including - to some extend - statistical analyses

- You know how to interpßret and present the results

- You know how to combine marketing concepts (in the area of pricing, distribution, communication, product development etc.) with visitor research

➢ **For that a rather continuous participation in the workshop is strongly recommended**

Benefits: What is the outcome for your *festival*?

What you may gain in the end:

- **Precise knowledge about your visitor structure and target groups**
- **Precise knowledge about visitors` motifs, needs, why they come and why not**
- **Can reduce barriers to the visit, gain more visitors, make them pay more**
- **Make better and more efficient campaigning**
- **Make more turnover from ticketing and funding**
- **Save money because of more loyal visitors and less expensive campaigning**

➢ **For that a continuos research is strongly necessary!**

End of this section!

RED
Visitor Research I

- II. The Research Process -

Author:
Prof. Dr. Tibor Kliment

Prerequisites for successful visitor research and evaluation

1. Openness and willingness
- First and foremost, there must be a willingness to learn and work with the results (also willingness to change things…)
- Courage to critically reflect on one's own work and possibly admit mistakes
- Analyses as an opportunity to learn and further develop services and programs
- Developing a visitor-oriented institution…

2. Inclusion of all stakeholders
- The entire team of the cultural institution should be involved in the studies, including top level and those those with constant contact with visitors (box office, supervisors etc.)
- All management levels should fully support the studies and create acceptance

3. Sufficient time and precise planning
- Sufficient time is needed for successful studies: to prepare, conduct, evaluate and to implement their results
- Even for small-scale studies, at least two months should be allowed in any case

4. Regular use
- Carry out surveys not only selectively, but on a regular basis
- Is the only way to gain a comprehensive and meaningful picture of the users of the various offers or to check the improvements

5. Sufficient resources
- An investigation takes time and - always - costs money
- The costs vary and depend on the *objectives and conditions* of the study, according to the *time and effort* required for the research, the chosen *survey method* and the *personnel* required for the different phases of the study

6. Competent staff
- If the investigation is carried out internally by your own staff, appropriately qualified employees are needed for the various steps of the investigation
- Qualifications basically include *method skills*, *experience* in this field of research, a little knowledge of *statistics*, knowledge of *project management*
- It is not advisable to do a complete study in addition to the usual workload staff already has

7. Serious engagement with results
- When dealing with the results, it is important that they are actually used by the institution
- Results should be taken seriously, those responsible should be prepared to take the results into account in their work
- Studies that are only carried out because they are imposed from above or outside of the festival are under pressure to legitimise and of little use

Before starting the research process: Make brief preliminary information on the study (description of services) that includes:

- Objectives of the study
- Issues of interest (if necessary prioritised)
- Addressees of the study
- Possible timeframe
- Available budget

Also further information may be needed about:

- Mission statement (if available), annual reports
- Event programme, performance announcements
- Planning status, (research may take place before, during or after the festival)
- Visitor statistics, ticket statistics
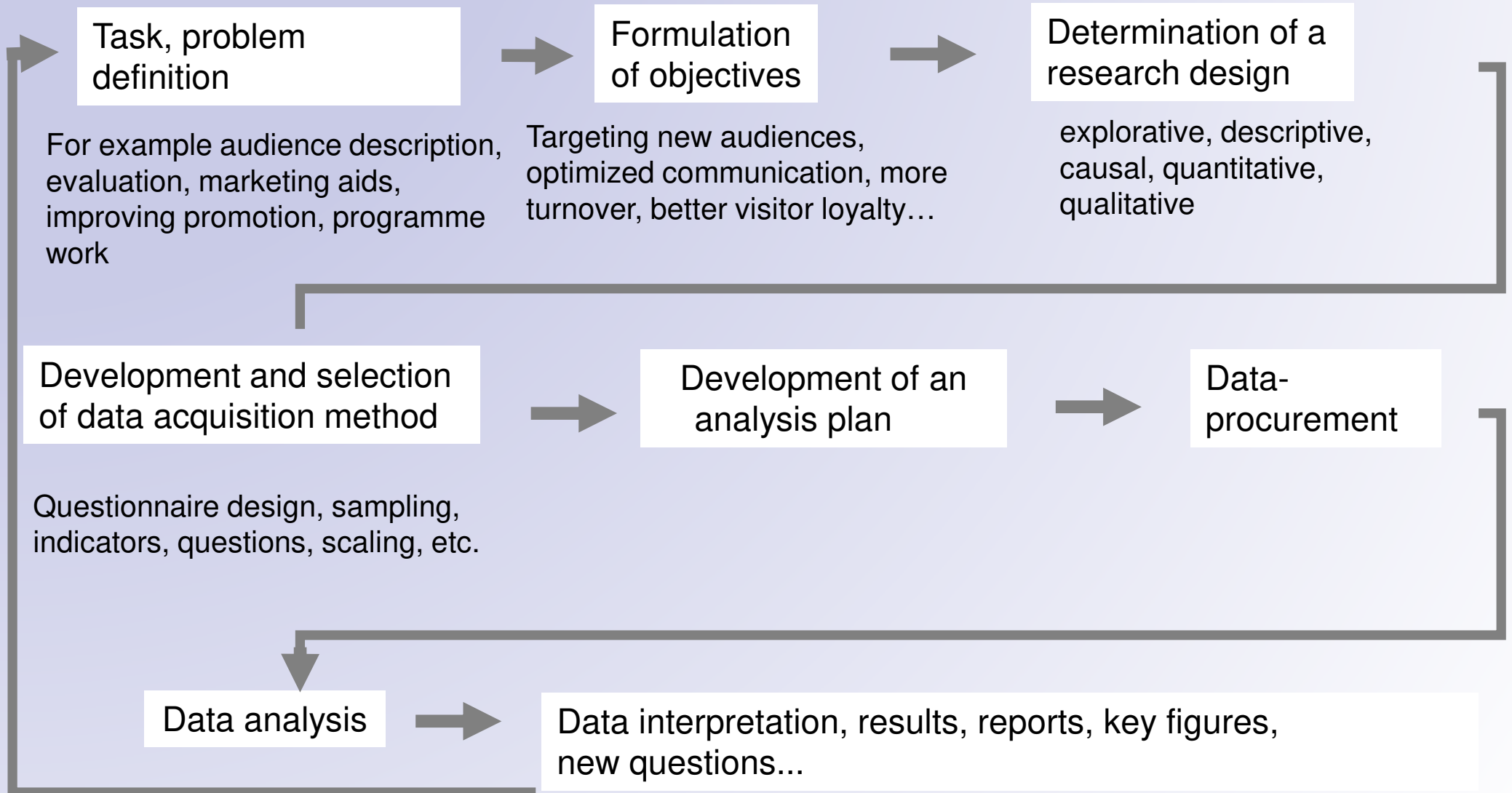- Existing studies from other sources

Recommendation for a transparent investigation process

Joint kick-off event about what to be investigated with all staff and stakeholders to present the project and eliminate possible ambiguities and fears
■ Continuous information of all stakeholders about the progress of the study (e.g. via internet, intranet, newsletter)
■ Definition of verifiable milestones
■ Discussions to consider different perspectives of stakeholders
■ Joint final presentation and discussion of the results in the investigated institution
■ Joint discussion on derived implementations, such as offers and strategies

Overview of the research concept - from the research question to the data

Task, problem definition

For example audience description, evaluation, marketing aids, improving promotion, programme work

Formulation of objectives

Targeting new audiences, optimized communication, more turnover, better visitor loyalty…

Determination of a research design

explorative, descriptive, causal, quantitative, qualitative

Development and selection of data acquisition method

Questionnaire design, sampling, indicators, questions, scaling, etc.

Development of an analysis plan

Data-procurement

Data analysis

Data interpretation, results, reports, key figures, new questions...

STEPS of the complete research process in more detail

STEP 1: Detecting problems/formulation of objectives/solutions for the problem

STEP 2: Determination of the form of the visitor research

STEP 3: Clarification of resources

STEP 4: Secondary research in the institution

STEP 5: Decision on quantitative/qualitative approach

STEP 6: Choice of survey method/construction of the survey instrument

STEP 7: Determination of the form of examination

STEP 8: Pretest and revision

STEP 9: Conducting the survey/fieldwork

STEP 10: Recording the results

STEP 11: Evaluation and analysis

STEP 12: Dealing with the results

RED
Re-Educating
Digitisation

Rheinische
Hochschule
Köln

Step 1: Formulation and limitation of the objectives

Basis for any investigation is to locate problems in the institution, that can possibly be solved by audience research.
Formulate goals precisely and limit them for processing. Therefore, sufficient time should be given to this very crucial step.
The objectives should be defined jointly by those responsible for the festival.

Therefore, it should be clarified in this step:
- What is the problem to solved? What should we do/know to solve that problem?
- Derived from that, what are the objectives of the study?
- How should the object of investigation be narrowed?
- What is the focus of the investigation: Visitor structures, visitor behaviour, effects of programs, visitor motivation, satisfaction with the services offered, assessments of the service, analysis of the competition, planning communication etc.?

What is the extend of the examination?
- Is a one-time examination sufficient or should changes and effects also be recorded over time?
- When defining the objectives, it is important to concentrate on the most important ones and not to address all questions at once in one study.

How do I generally arrive at my questions or research expectations?

For example:

- The institution has a certain self-image as mission, vision, mission, brand, etc. Does the market reflect this?

- External target agreements with politics, sponsors or internal agreements as guidelines

- Service expansions or improvement measures were undertaken in certain areas and are to be examined in terms of effects

- Examine effects of service reductions or savings

- Evaluate regular innovations (special offers, new programme, etc.)

- Assess external shocks with massive effects in all areas ("corona")

- Visitor research as a constant "voice of the audience", commentator, analyst, forecaster or early warning!

How do I arrive at my research questions? For example:

- We currently have a great selection of movies. Have we had more first-time visitors as a result? What about regular visitors?

- Our visitors become constantly older? What is the reason for that?

- The new season has begun: Why is visitor satisfaction with the current programme worse than last year?

- We have strengthened the service staff, improved orientation in the festival and upgraded the gastronomy. Is this reflected in visitor satisfaction?

- We spend a lot of money on advertising and PR without knowing exactly whether it is going into the right channels. Should we better go to other media? And which ones?

- How is the newly relaunched website perceived by users?

- We want to justify the need for more financial resources to political committees and stakeholders. How is this possible?

- Unfortunately, the ticket prices had to be increased recently. Is a different audience coming now?

- We want to give proof for our private sponsors, that their activities have an impact? How can we do that?

The last step, the implementation of the results, should already be considered when formulating and limiting the objectives

Goals must be formulated "SMART
(cf. Klein 2011: Kultur-Marketing, Munich):

✓ Specific: Is the target content clearly defined?
✓ Measurable: How can you tell if the goal has really been achieved?
✓ Achievable: Is the achievement of the goal actually feasible? Is it within the sphere of influence of the persons responsible? Is the goal challenging but not overwhelming?
✓ Realistic: Is it feasible?
✓ Time-scaled/terminated

Example:
• **Objective:** Target group younger people between 18 and 29 years
• **Time reference:** One calendar year
• **Example target:** The range of offers and services for younger people increases. The share of younger people within the next calendar year is therefore to increase by 25%
• **Area:** Young people from city of the festival

STEP 2: Determining the form of visitor research

Visitor can be defined as studies that focus on visitors/users/audiences of cultural institutions.

Visitor/non-visitor research
- Visitor research focuses on visitors/users/audiences of the institution
- The field of non-visitor research also offers potential: Non-visitor studies are surveys outside the institution, focuses other topics than visitor research
- In addition to target groups that have not yet been reached, "non-visitors" can also be analysed within visitor research (for example when subscriptions are cancelled)

Example objectives of non-visitor studies:
- Information on catchment areas, level of awareness and image of the cultural institution; associations and expectations of potential target groups; reduction of barriers to visits

Methods:
- Oral, postal, telephone or on-line surveys

- Research is primarily used **retrospectively**
- But: Research can be used at different points in time throughout the work process:
- Provides information during the **conception, design, implementation and revision** of a festival and are a basis for decision-making.

	Terms	Goal	Content of research	Content of planning
in the run-up/ planning process	Ex-ante Research	Up-front Research	Improve planning; support project goal setting; feasibility studies/ needs assessment.	Prior knowledge, interest, expectations, attitudes of potential target groups	Planning of programme, marketing concept, complete set-up
during/ design process	On-going Research	Formative Research	Improvement of implementation processes; accompanying research; early control and correction possibility	Reactions of test persons to designs/ prototypes; user-friendliness of design alternatives	Checking texts, guidance systems, interactive elements, titling, arrangement of objects
closing	Ex-post Research	Summative Research	Final success/effectiveness control; review of target achievement; experience for future projects	Judgements on the overall offer and specific aspects, user structures	Overall assessment of an event, exhibition, guided tour after the visit

RED Re-Educating Digitisation

Rheinische Hochschule Köln

STEP 3: Clarification of resources

- In general, the use of resources that it is worthwhile in the long term to invest in well-founded and targeted studies.
- It is not only the institutions themselves that are called upon to do this, but it should also be the task of the funding agencies
- It is not only about resources for conducting the studies, but also for implementing the results

Who is conducting the study?
Sufficient resources and competent staff were described as prerequisites for successful studies

Internal or external implementation?
- Is the study conducted internally, i.e. with the institution's own staff?
- Or is an external researcher brought in?
- Is a combination of internal and external implementation possible?

	Internal evaluation	External evaluation
Advantages	■ rapid implementation ■ lower costs ■ Expertise ■ direct transposition	■ Independence/objectivity ■ Methodological competence ■ Comparisons possible ■ Influence and credibility of the evaluation higher
Disadvantages	■ lack of methodological competence ■ "Operational blindness ■ Lack of time	■ higher costs ■ Fear/defensive reactions of those affected ■ Lack of expertise

How much work is estimated? The following sample estimate can serve as a rough guide for calculating the workload for studies

Typical for less extensive study in terms of workload:
- Goal: E.g. to gain insights into visitor structures and evaluations at events
- Conception: 2-3 weeks
- Preparation/execution of fieldwork: 2 weeks
- Analysis: 3-4 weeks

TIP – ENLARGE YOUR RESOURCES:

- In order to be able to conduct solid studies, a fixed part of the budget should be planned in advance

- For the realisation of studies, you can also check whether you can link up with other studies (surveys by city/tourism marketing), in which the festival can place some specific questions

- Co-operation with other cultural institutions in the city can also be useful to enable studies (e.g. museums in a city, theatres in a region)

- Cooperation with universities

- Raise third-party funds for studies - film/cinema associations can promote studies; sponsoring is also conceivable

RED
Re-Educating
Digitisation

Rheinische
Hochschule
Köln

STEP 4: What knowledge is already available?

Is there pre-existing data in the institution?

- Before collecting your own data, it makes sense to research what existing knowledge the institution can draw on.

- This can be information from the institution itself, from other institutions or from general studies.

Possible sources:

- Previous studies carried out in the institution (to compare the results)

- Visitor and ticket statistics for the analysis of visitor numbers, visitor structures, ticket types, usage times, etc.

- Analyses of website / app usage can be analyzed by a variety of tools

- Other sources like guest books, interviews with front-end staff etc.

- Research by external institutions at state or EU-level, film assoc., cinemas etc.

STEP 5: Decision on quantitative or qualitative approach

Depends on the interest in knowledge: If a larger number of data is to be collected, statistical results and direct comparisons are intended, a quantitative approach makes sense.

On the other hand, a qualitative approach serves to gain deeper insights into a rather unknown field of investigation.

Quantitative studies
- Examine a broad data base
- Record frequencies and numerical parameters directly measurable
- Based on standardized methods and statistical evaluations
- The goals are to investigate into well defined questions and to make generaliseable and comparable statements

Qualitative studies
- Qualitative approaches aim at only few cases
- Characteristics are openness, process character and flexibility of the research.
- Assumptions are to be gained first, not verified. The aim of the research is to gain deeper insights into complex interrelationships, decision-making processes and interactions

Combination of quantitative and qualitative approaches
- Combination of quantitative and qualitative approaches is often recommended, as not all aspects to be investigated can be recorded numerically (e.g. satisfaction or effects).
- For example: Qualitative studies are carried out at the beginning of research processes in order to gain knowledge about a rather unknown topic.
- A quantitative study can build on this basis. Or a qualitative study can be carried out after quantitative study to gain insights in more detail.

STEP 5: Decision on quantitative or qualitative approach II

Examples of quantitative parameters:

Attendance and usage statistics, number of performances visited, admission income, number of media reports, amount of third-party funds raised, utilisation of educational programmes, number of cooperation partners, proportion of first-time and repeat visitors, structure of the audience…

Examples of qualitative parameters:

Needs, ratings and opinions of visitors, attitudes of stakeholders, barriers of visit, likes and dislikes about the festival, patterns of perceptions onto the festival, ...

STEP 6: Choice of survey method and construction of the survey instrument

The choice of survey method is a fundamental decision in the research process.

There is no "best" method, but it depends on the questions to be investigated and your ressources

Criteria for a proper choice are:

- Which survey method is suitable for my research objectives?
- Which survey method is suitable for the interviewing conditions given?
- Which survey method is suitable within the scope of my ressources?
- Which survey method is suitable for my research knowledge?

Types of survey methods

Written surveys:
- Less time-consuming than oral interviews and very cost-effective
- No influence of an interviewer, but there is little control of the filling situation and no follow-up with the respondents possible

Types:
- Questionaires distributed/displayed (paper or digital version)
- posted
- put online/sent by mail

Oral surveys/interviews:
- More time-consuming, require trained interviewers
- Makes it possible to gather more in-depth opinions and give help to the respondent
- Interviews can be much longer. motivation to participate is higher
- Expensive and time consuming
- However, the influence of the interviewer must be taken into account and controlled if possible..

Types:
- face-to-face
- telephone
- group discussion

Very often used: Personally distributend, but self-completion questionaires

STEP 7: Determination of the form of examination

When is the examination?
Based on the objectives of the study, it must be decided whether a single data collection is sufficient (cross-sectional survey) or whether multiple surveys are necessary

Multiple surveys can analyse differences at different points in time (longitudinal survey).

Cross-sectional survey
Is the standard form of survey: Asking different visitors about the same content (e.g. at the end of a movie performance)

Longitudinal survey
Trend Study: Asking different visitors about the *same content* at the beginning of the festival and at the end to record differences in perception.

Example: Survey the same people before visiting the festival, when leaving and two weeks afterwards to collect short and long term effects of the visit. (trend survey)
Monitoring of success of an advertising campaign at different times

Panel study: Information is collected from the same persons at different times

Example: Survey the same people when leaving a performance and two weeks later to collect short and long term effects of the visit

STEP 8: Pretest and revision

- A pretest is a test of the survey instrument/the survey procedure before the actual fieldwork to ensure comprehensibility and functionality
- Gives important hints and possibilities for improvement
- Especially useful with new research questions, unknown target groups or new methods

To be clarified at the pretest:

- Required duration of the survey? Is the survey instrument too long?
- Do test subjects drop out of the study?
- Are the questionaires protocolls complete?

- Was the instrument used by the test persons as planned?
- Is the survey instrument understandable? Do filters work properly?
- Are question formulations and answer options comprehensible?
- Are the answer options complete?

- What is the appropriate organizational procedure?
- Is the location suitable?
- How are the respondents addressed? Are there refusals?
- What materials are needed?

It makes sense to carry out the pretest first with 5-15 surveys or 1-2 qualitative interviews. Ask the test persons/respondents for feedback

STEP 9: Conducting the survey/fieldwork

The appropriate and targeted implementation of the survey (fieldwork) depends on the chosen approach and methods (qualitative vs. quantitative)

Basic tips for fieldwork:

- Participation in a survey should be made as easy and pleasant as possible for the respondents

  – If you do a after-visit survey make sure people have time for a survey (e.g. after the performance but before going to the cloakroom)

  – Objectives and duration of the survey should be explained to the respondents, and the importance of their participation made clear

- Make sure a quiet place is chosen for the interview and that sufficient seating and writing facilities are available. To motivate respondents, small gifts are useful.

- In order to provide a representative sample of respondents distribute the interviews over <u>all days of the festival and over different times per day</u>. Make a precise timetable and continuously review it to ensure that the sample is conducted in a targeted manner.

If extra (external) staff is deployed for the field work, they should be trained before and consulted regularly on the progress and possible ambiguities

- The response rate of surveys (i.e. the number of completed questionnaires) cannot always be calculated exactly in advance.
It depends strongly on the number of visits and the motivation of the respondents to participate.
Buffers should always be planned in order to be able to extend the survey period if needed.

STEP 10: Recording the results/data entry

Should be carried out carefully in order to avoid errors in the analyses.
A data check is necessary, as well as sufficient knowledge of the recording software. It may be helpful to have a manual that explains all the important steps in the collection process.

Collection of quantitative data from a standardized questionaire:
- Completeness: E.g. it can make sense to sort out questionnaires from about 50% of questions that have not been answered.
- Plausibility: E.g. information from different questions may be checked for consistency, such as age and employment status
- Validity: E.g. age limits of respondents, questionnaire completed by two persons, incorrect answers, answer patterns (like those later discussed when talking about errors in formulating questions)

The following points should be regarded when doing the data entry:
- It is useful to number all the questionaires continuously and put the numbers in the data set in order to be able to allocate them later
- Create a code plan (see example)
- If necessary, record additional information: e.g. date, time, week of the day, language of the respondents etc.
- Clarify: Type of recording of open answers, other information, missing information

STEP 10: Recording the results

Collection of *qualitative* data:

Transcription of the data (interviews) always necessary, for which binding transcription rules are defined, such as:

- First line: Interview number, date, place, interview duration
- How is the respondent?
- Number, if there are several respondents in a group discussion
- Pauses of at least 5 seconds = "{...}".
- Unintelligible passages = "[...]"
- Omit repetitions of words
- Emphasis of terms = underlining of the term
- Vocalisations in brackets, e.g. "(laughs)" or "(sighs)".

Technical devices may facilitate the transcription process (use of mobiles, specialized dictation programmes, usage of MS Office and Windows 11 etc.)

RED
Re-Educating
Digitisation

Rheinische
Hochschule
Köln

STEP 10: Recording the results

Programmes for the collection and analyses of quantitative data:

MS Excel

- Very common, usually existing programme from Microsoft Office
- Possibility of recording the data in tabular form
- To be taken into account: Data collection can become confusing with larger amounts of data,
- Excel does not offer an automatic control function
- Evaluation options: Counting frequencies, forming totals, filtering according to specific data, creating diagrams, etc.

SPSS V27.0

- Reference at https://www.ibm.com/software/de/ analytics/spss/, cost-intensive licences
- Comprehensive statistics programme, which is necessary for extensive evaluations
- Requires a little training effort
- Possibility of data collection in tabular form
- Possibilities for comprehensive evaluation and statistical procedures, in particular also for multi-variate evaluations (in which different questions are related to each other)
- Possibilities to create diagrams (limited design possibilities)

STEP 10: Recording the results

Programmes for the collection and analyses of quantitative data:

PSPP

- PSPP is a free alternative to the statistical software SPSS and offers a similar range of functions for performing statistical analyses (supports descriptive statistics, T-tests, variance analysis, linear and logistic regression, correlation, cluster analysis, reliability and factor analysis and parameter-free statistics)
- However, not all SPSS functions are included in PSPP
- PSPP is open source: Avoids the restrictions of the SPSS licence, but pays attention to compatibility
- Freeware is suitable for institutions, who do not receive a costly SPSS licence as an alternative for scientific analyses.
- System requirements: PSPP is available for Windows, Mac OS X and Linux. The freeware is included in the packages of distributions such as Debian and Ubuntu.

Programmes for online surveys

- Very large market of providers for online surveys
- Include restricted possibilities for evaluation and presentation data
- Differences in handling, costs, inclusion of advertising
- Providers, among others www.soscisurvey.de (free of charge for public education, without advertising), www.umfrageonline.com (free trial period), www.surveymonkey.de (free of charge for small surveys) etc.

STEP 11: Evaluation and analysis (I)

Evaluation
The evaluation is the important step to be able to process the data sets according to the formulated objective. A basic distinction here is between quantitative and qualitative data.

Evaluation of quantitative data:
- Counting in percent (never absolute numbers), tables
- Calculation of arithmetic mean/average, comparison of mean values
- Further statistical methods: Correlation analyses using cross-tabulations as the most important multivariate form of evaluation (= evaluation of several questions)

Presentation in tables and diagrams:
- Most common chart types: Column, bar, pie, line chart; do not use too many different types.
- Necessary information: meaningful superscript, size of the population/missing information, units of measurement (in %, absolute numbers), reference to multiple answers

Evaluation of qualitative data:
- Evaluation programmes e.g. Atlas.ti, MAXQDA
- Qualitative content analysis = transcribed data are summarised and structured; categories are developed from the material.

STEP 11: Evaluation and analysis (II)

Data analysis

- The analysis of the results are very individual, depending on the objectives of the study

- In each case, it must be analysed specifically which research questions were to be answered and to what extent this could not be done.

- The methodological procedure, the actual research conditions with functioning and possibly also non-functioning aspects should be reflected.

- It is important to deal transparently with possibly distorting influences (e.g. renovation measures during the survey, change of interviewers), deviations from the planned research procedure (e.g. change of the survey period) and to point these out in the evaluation.

- The interpretation of the results can become more meaningful, especially through comparison with other results.

STEP 11: Evaluation and analysis (III)

How do I evaluate my results later?

It is advantageous to formulate <u>concrete </u>questions or hypotheses for the values <u>before looking at the data for the first time</u>!

– **Set precise expectations for the data, if possible also for the subgroups:**

  - *Why is the value the way it is?*
  - *Why is it not higher or lower?*
  - *Why is it unchanged, even though we...*
  - *Why has the value (not) changed over time?*
  - *Why is the value different for a comparable institution?*
  - *What does that mean for us? Etc. ...*

– **Think about what would be "good", "acceptable" or "unsatisfactory" for you/the institution.**

– **Question the values. Try to find explanations, especially if data turn out differently than expected.**

– **Note contextual conditions that may have affected the results during the interview phase**

Quote:
"For a ship that does not know which port it wants to go to, any wind is the wrong one!"

RED
Re-Educating
Digitisation

Rheinische
Hochschule
Köln

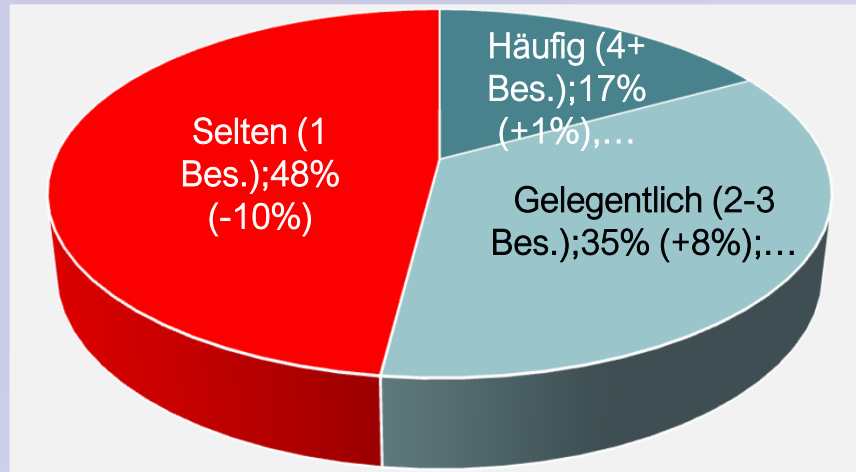### Some simple tips for creating presentations

- Presentations force to **focus on content** (topic, area, specific questions)
- Graphical representations can **illustrate** complex issues in a **simple way**
- The visualisation of facts or contexts causes them to be remembered **better and longer**
- An argument supported by graphs or charts is more suggestive than tables or texts, and often difficult to refute in argumentation
- They offer **possibilities for manipulation** (choice of axis divisions, reference base, pictograms, etc.)
- Observe the same presentation principles as above for the tables
- Ensure that the design is **graphically "attractive"** (preferably not from the dashboard but via PPT or similar).
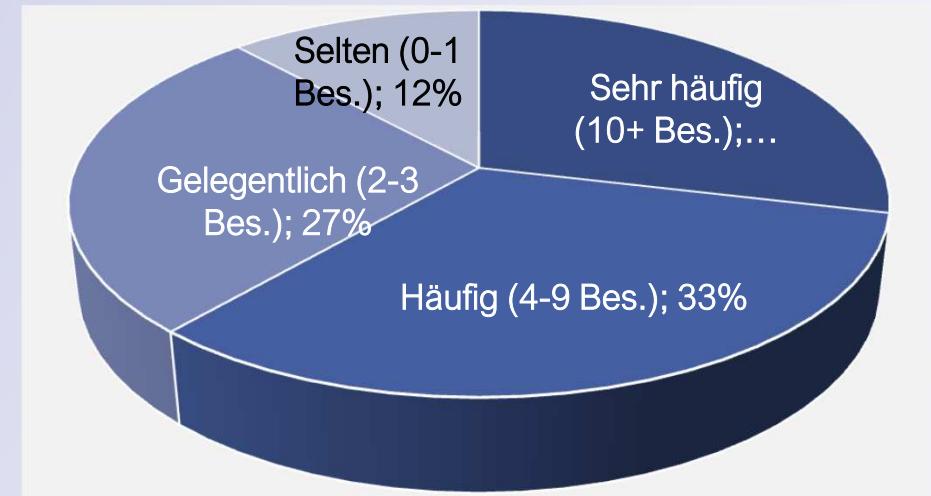
Examples Presentation ->

Fig. X: Frequency of visits to FoXXX and in general
(in percent of respondents)
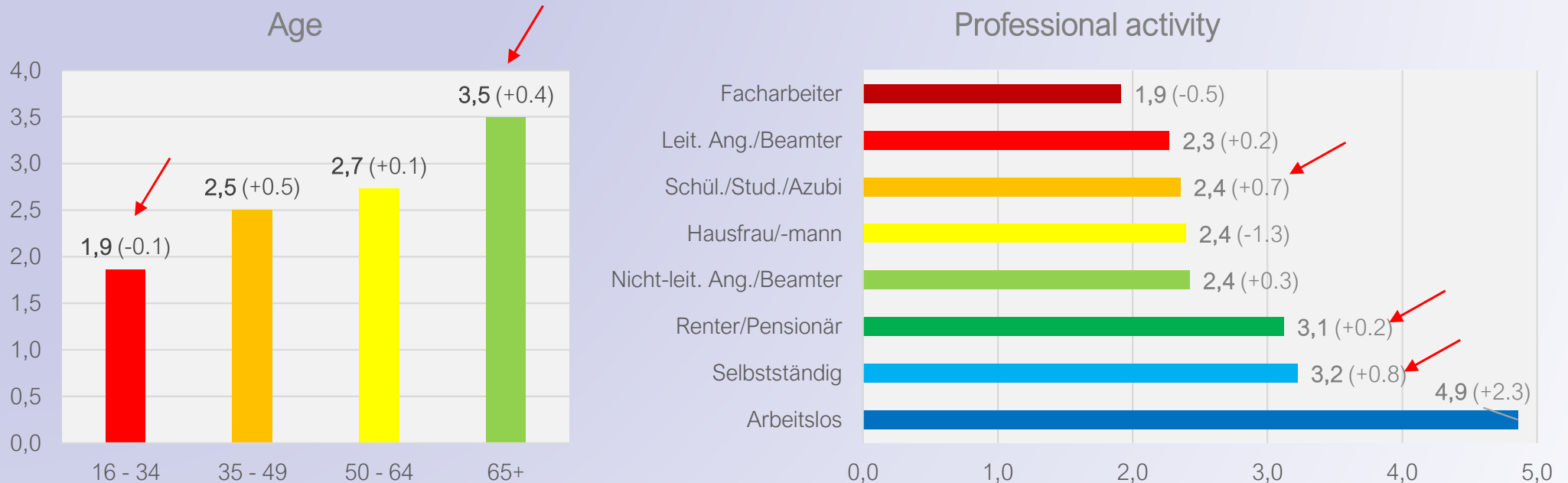
Frequency of visits
(Number of visits per year)

Häufig (4+ Bes.);17% (+1%);…

Selten (1 Bes.);48% (-10%)

Gelegentlich (2-3 Bes.);35% (+8%);…

Frequency of visits in general
(Number of visits per year)

Selten (0-1 Bes.); 12%

Sehr häufig (10+ Bes.);…

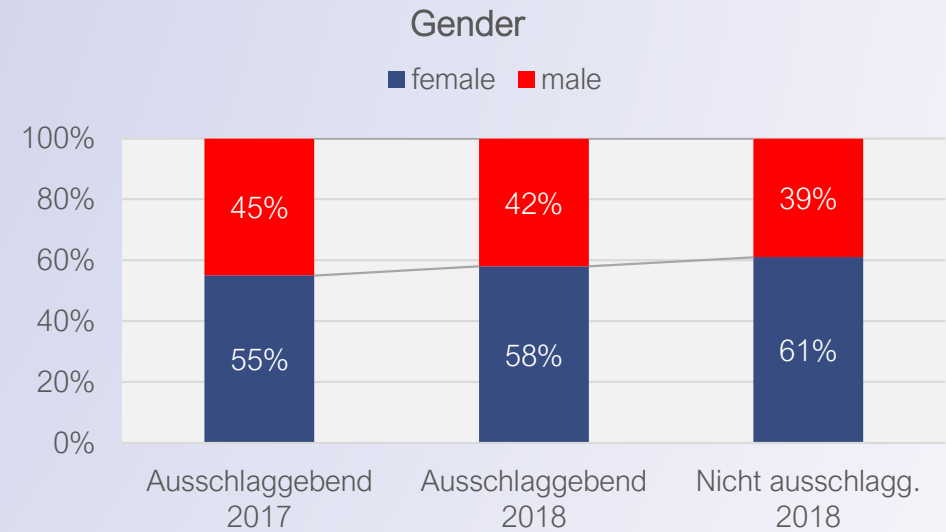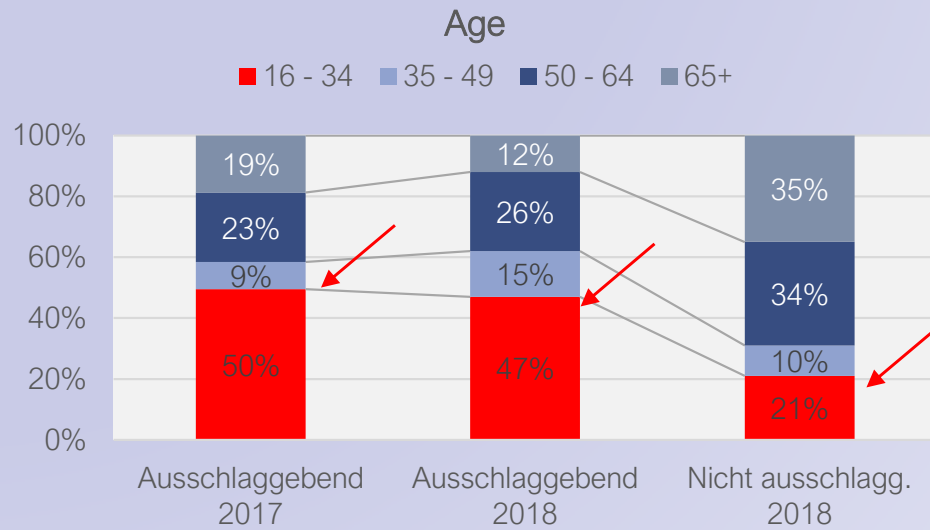Gelegentlich (2-3 Bes.); 27%

Häufig (4-9 Bes.); 33%

- Half of the visitors (48%) come about once a year. A good third visits the museum occasionally, 17% of visitors come frequently. On average, visitors come to the museum about 2.7 times/year, which is not only comparatively high, but also an increase compared to the previous survey (2.4 visits). Visits have thus increased by 13%. This is still due to the museum's "fans", who come to the Folkwang 8.5 times per year.

- Other museums are visited very frequently by 29% and frequently by another 33%. On average, visitors go to other museums about 7.6 times/year, whereby this value is also pulled up here by some, in part extremely frequent visitors (in some cases well over 100 visits per year).
- The visit here and the visit to other museums are stat. connection. The following applies: frequent visitors to the museum do not always also frequently go to other museums and vice versa. In other words, the museum also attracts visitors who would otherwise hardly go to the museum. Free admission is probably the reason for this.

Fig. X: Average frequency of visits, by demographic group (mean number of visits in the last 12 months)

Age

- 16 - 34: 1,9 (-0.1)
- 35 - 49: 2,5 (+0.5)
- 50 - 64: 2,7 (+0.1)
- 65+: 3,5 (+0.4)

Professional activity

- Facharbeiter: 1,9 (-0.5)
- Leit. Ang./Beamter: 2,3 (+0.2)
- Schül./Stud./Azubi: 2,4 (+0.7)
- Hausfrau/-mann: 2,4 (-1.3)
- Nicht-leit. Ang./Beamter: 2,4 (+0.3)
- Renter/Pensionär: 3,1 (+0.2)
- Selbstständig: 3,2 (+0.8)
- Arbeitslos: 4,9 (+2.3)

- There are clear differences between the various age groups: The younger visitors up to the age of 34 are the least frequent visitors, with just under 2 visits per year, whereas the older visitors are very loyal customers, with 3.5 visits per year. The above-average visit frequency is mainly due to them, whereby they have even increased their visit frequency compared to the previous survey (3.1 visits).
- With regard to occupational status, pensioners and the self-employed come to the house particularly often, and executives the least often.
- Significant increases were achieved in the middle age groups as well as pupils/students/trainees and the self-employed.
- The different visit frequencies for age and occupation show that the visit barriers are located both before the museum visit and - albeit weaker - when using the exhibition offer in the museum.

Fig. X: Comparison of visitors for whom free admission was the decisive factor, by demographics
(Basis: By subgroups, in percent of respondents)

Age

■ 16 - 34 ■ 35 - 49 ■ 50 - 64 ■ 65+

	Ausschlaggebend 2017	Ausschlaggebend 2018	Nicht ausschlagg. 2018
65+	19%	12%	35%
50 - 64	23%	26%	34%
35 - 49	9%	15%	10%
16 - 34	50%	47%	21%

Gender

■ female ■ male

	Ausschlaggebend 2017	Ausschlaggebend 2018	Nicht ausschlagg. 2018
male	45%	42%	39%
female	55%	58%	61%

Education

■ Further schooling without A-levels
□ A-Levels/Fachabitur
■ Study

	Ausschlaggebend 2017	Ausschlaggebend 2018	Nicht ausschlagg. 2018
Study	59%	64%	65%
A-Levels/Fachabitur	29%	24%	20%
Further schooling without A-levels	6%	3%	9%
	6%	6%	6%

- In 2017, as in 2018, around 50% of the group for whom free admission was a decisive factor consisted of young visitors up to the age of 34. In the comparison group, for whom free admission was insignificant in 2018, the same proportion is only 21%.
- In terms of education and gender groups, the price-sensitive group clearly includes more high school graduates and more male visitors. However, the differences to the non-price-sensitive visitors decrease over time.
- However, a strong *socio-structural* change in the audience due to free admission still exists, or has become somewhat less so.

Fig. Y: Structure of new visitors, in combined representation by place of residence and age (all blocks add up to 100%, in percent of respondents)
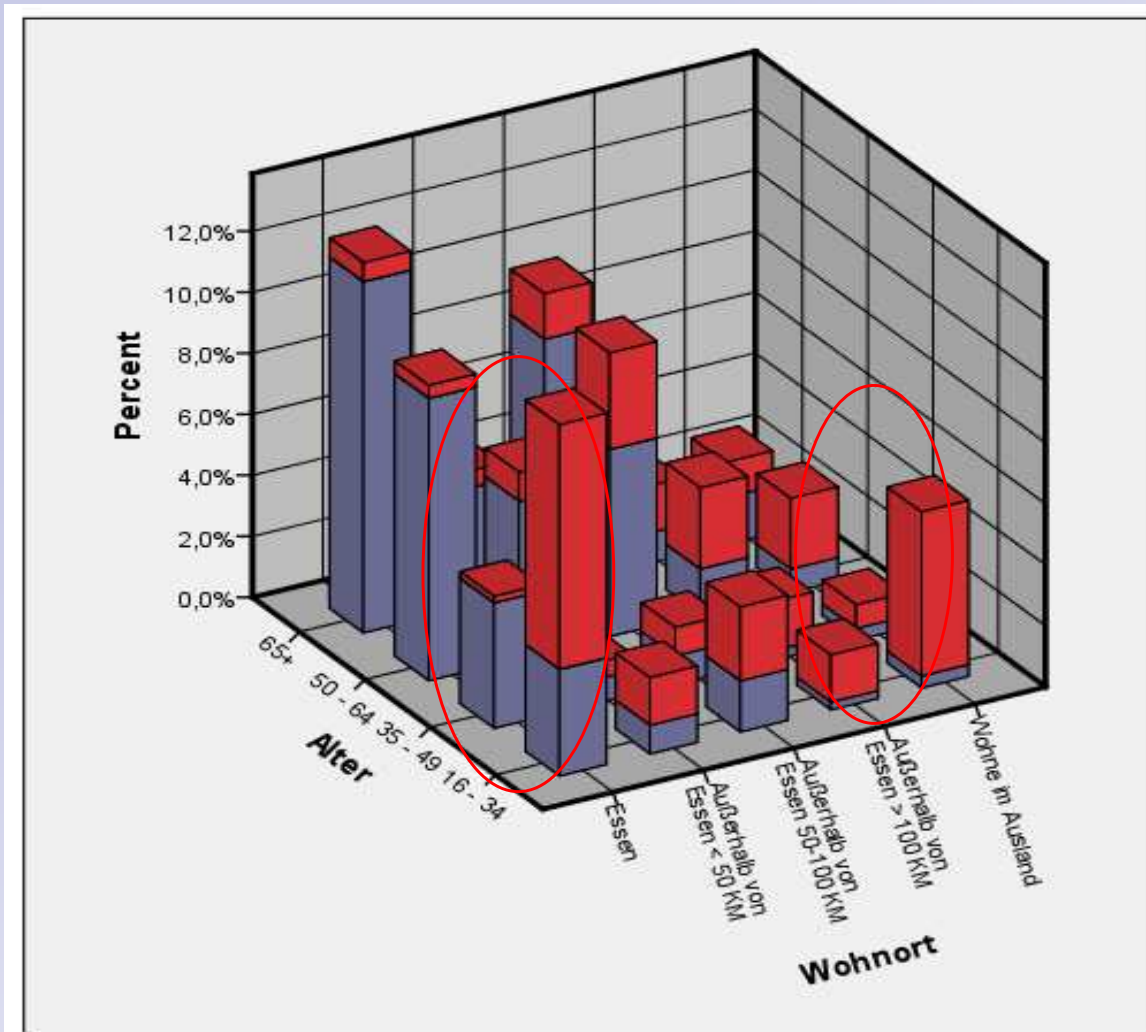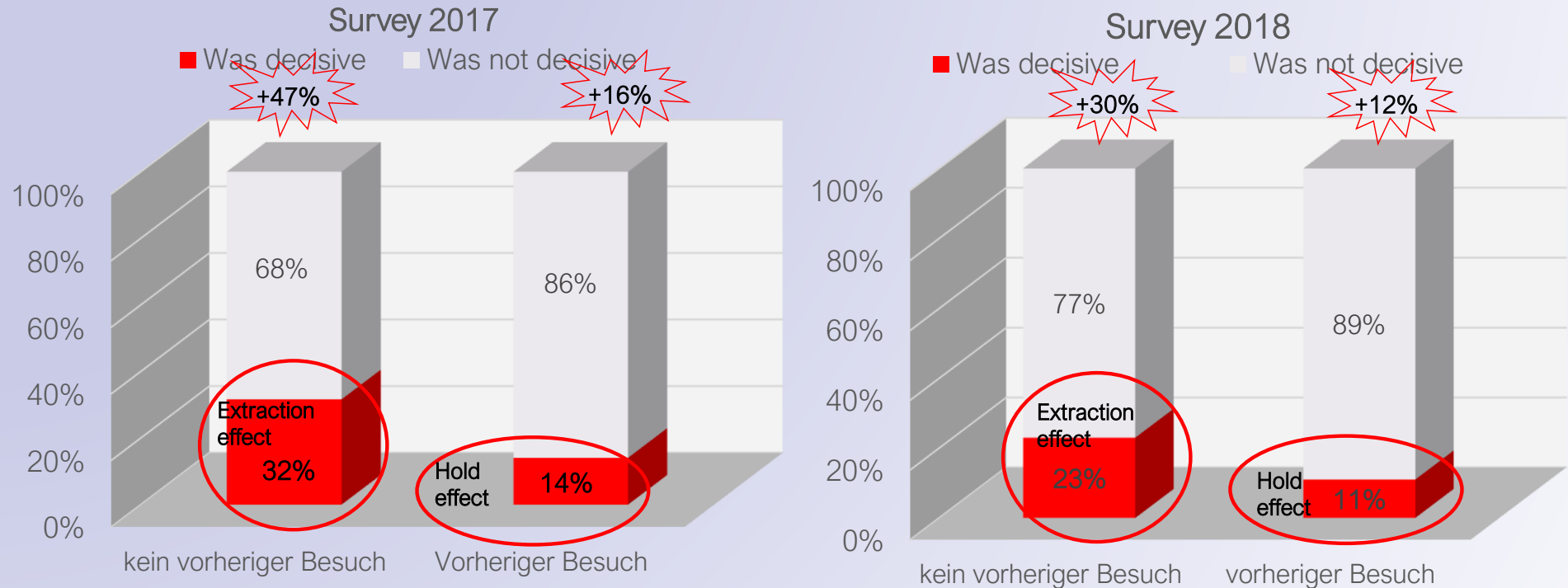

- In the combined consideration of the three characteristics, it becomes clear that the largest proportion of new visitors consists of the 16-34 age group from Essen.

- A second, also larger group of new visitors is recruited from the youngest age group who have travelled from abroad.

- The other new visitors are distributed relatively evenly across all age groups from the immediate and wider vicinity of Essen and the rest of Germany.

Fig. X: The importance of free admission for visitor development
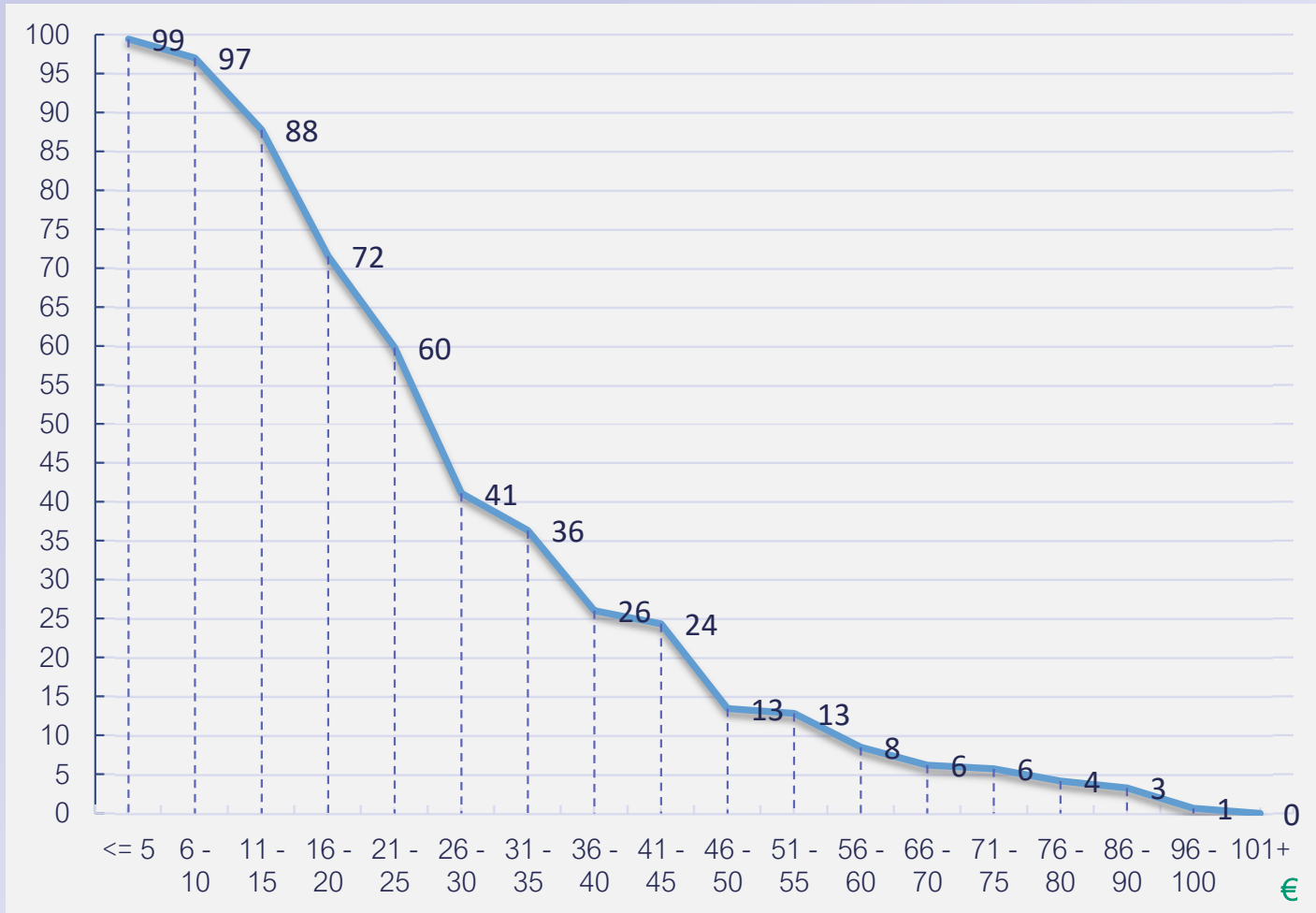(Basis: All respondents, in percent of subgroups)

Survey 2017

■ Was decisive   ☐ Was not decisive

+47%   +16%

- 100%
- 80%   68%   86%
- 60%
- 40%
- Extraction effect
- 20%   32%   Hold effect   14%
- 0%

kein vorheriger Besuch   Vorheriger Besuch

Survey 2018

■ Was decisive   ☐ Was not decisive

+30%   +12%

- 100%
- 80%   77%   89%
- 60%
- 40%   Extraction effect
- 20%   23%   Hold effect   11%
- 0%

kein vorheriger Besuch   vorheriger Besuch

- Not all new visitors come because of the free admission. In the current survey, 38% of visitors were coming to the Folkwang for the first time since the introduction of free admission ("new visitors"); in the 2017 survey, this figure was 27%, which corresponds to an increase of 40% in "new visitors" compared to the previous survey.
- Of these new visitors, free admission was the decisive factor for 23% of visitors to the Museum Folkwang in 2018. The "winning effect" of free admission is still 30% in this group. Among the old visitors, free admission was the decisive factor for 11% to continue visiting the Folkwang. The "retention effect" of free admission is 12% in this group.
- In the 2017 survey, an additional 8% of visitors came to the museum due to the free admission. In the 2018 survey, the share was 9%. Free admission thus continues to bring new visitors to the museum.
- Overall, almost 90% of the higher visitor numbers due to free admission are due to the increased visits by "old visitors", and about 10% are due to the new visitors who have come since the introduction.

The role of the ticket price: willingness to pay
(Basis: respondents interested in visiting a French theatre in the next 12 months (70%)). 12 months (70%))

% of respondents

99 97 88 72 60 41 36 26 24 13 13 8 6 6 4 3 1 0

<= 5 | 6 - 10 | 11 - 15 | 16 - 20 | 21 - 25 | 26 - 30 | 31 - 35 | 36 - 40 | 41 - 45 | 46 - 50 | 51 - 55 | 56 - 60 | 66 - 70 | 71 - 75 | 76 - 80 | 86 - 90 | 96 - 100 | 101+

€

STEP 12: Dealing with the results

- A decisive step is the implementation of the results of the study

- The institution should have formulated and prioritise the need for action to be derived before the study; an external evaluator can make recommendations

- It is important to go back to the goals defined at the beginning and the interest in the findings: What should be investigated, what was the problem to be solved?

- The basis for implementing the results is the acceptance of the results and the willingness to change

- Sufficient time is crucial - this applies not only to time for study design

- All stakeholders should be involved, the research process should be made transparent to them

- A **recipient-oriented report of the results** as well as a presentation and discussion with all staff members are important: Results should be prepared in different depths and adapted to different departments and interests.

- A summary is always useful

- Comparisons with other studies at one's own institution or in comparable institutions can be helpful

- Festivals should communicate more with each other about their studies and publish results.

STEP 12: Dealing with the results

Operate - depending on your resources - a reporting system appropriate to the addressees:

- ✓ Data documentation in the form of table volumes that can be called up for reference by anyone from anywhere
- ✓ Topic-specific individual reports, overarching annual reports (incl. summaries)
- ✓ Regular presentations with selected results for specific occasions, topics areas in the institution or topics, (temporal density depending on sample sizes)

- ✓ **Try new ways of doing things:**
  -> E.g. no data diggers in reports but visualise research results in videos

STEP 12: Dealing with the results

TIP:
- It is recommended to appoint someone in the institution which is mainly responsible for the monitoring and implementation of the implementation process
- She is responsible for locating the results in the institution and for introducing them to the various departments, e.g. during meetings and decisions.

How are results communicated and published?

- There is in most cases a critical attitude towards publications
- Advisable to control publications through one's own channels, to communicate the context of the studies with the objectives and framework conditions.
- Often, further questions may arise from examination results
- If research is carried out regularly changes that have been initiated can also be reviewed. It makes sense to refer to this further examination ($\rightarrow$ Prerequisites for regular application)
- It is useful to communicate results to all relevant stakeholders, in particular to **sponsors, other funding partners, local politics**, **tourism agencies** in the city etc.

Time schedule for a visitor research project
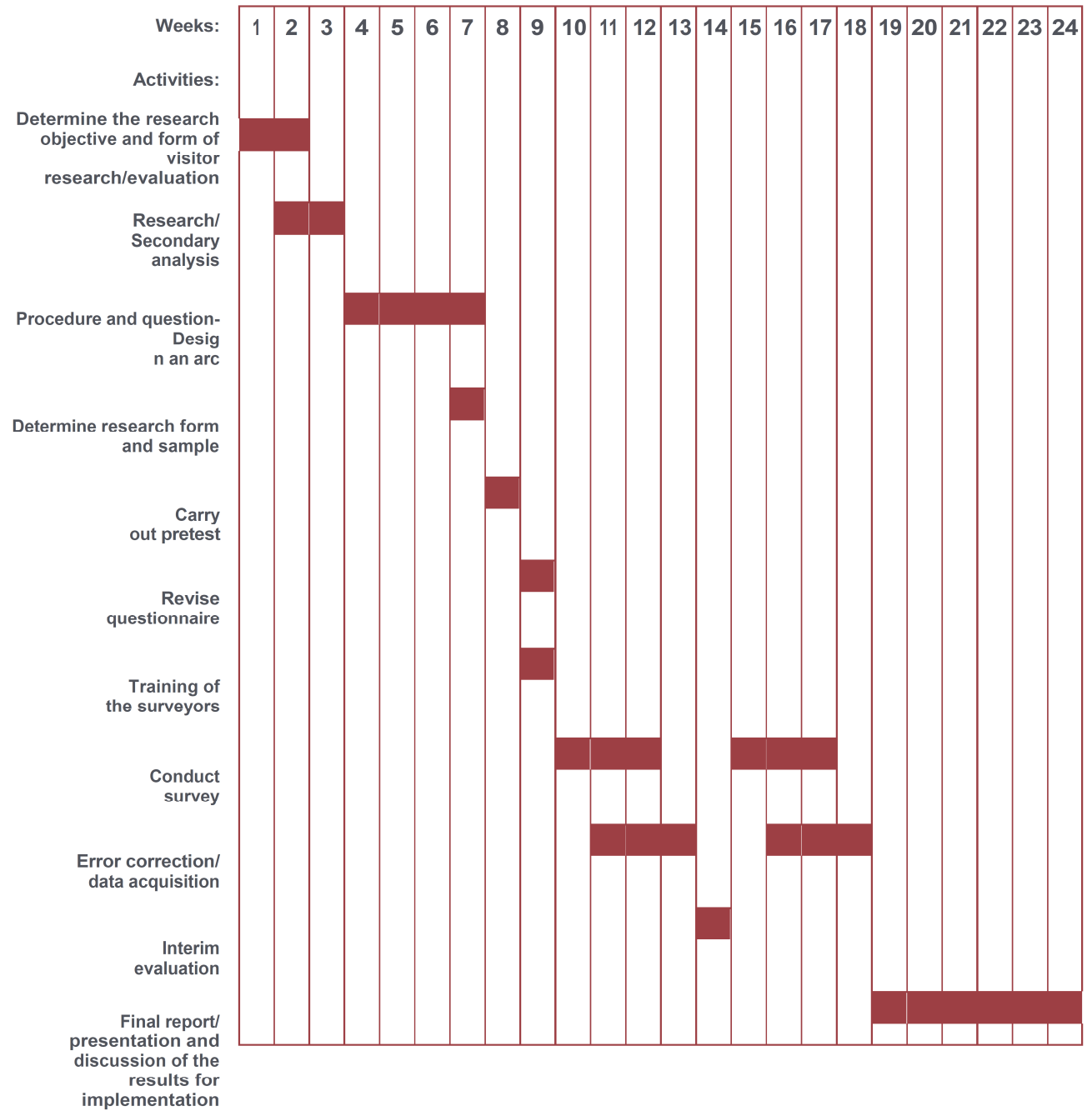
Weeks:	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24
Activities:																								
Determine the research objective and form of visitor research/evaluation	■																							
Research/Secondary analysis		■																						
Procedure and question-Design an arc				■																				
Determine research form and sample							■																	
Carry out pretest								■																
Revise questionnaire									■															
Training of the surveyors									■															
Conduct survey										■	■				■	■								
Error correction/data acquisition											■	■				■	■							
Interim evaluation														■										
Final report/presentation and discussion of the results for implementation																			■	■	■	■	■	■

End of this section!

RED
Visitor Research I

- III. Designing Questions
and Questionaires -

Author:
Prof. Dr. Tibor Kliment

And now:

Scientific questioning:

„The art of asking questions"

RED
Re-Educating
Digitisation

Rheinische
Hochschule
Köln

Respondent

is the object of study itself

**=
direct approach**

but:
- Often not possible
- Unfavourable, higher costs

represents object of study

**=
indirect approach**

But: It may not be possible to draw conclusions about the object(s) of investigation, as the respondent does not represent them.

Example: Acceptance of a new movie

Interview visitors   **Ask Expert, Film Distributors, Cinema Owners**

What happens in an interview?

Respondent must ...

- Understand the question semantically (what exactly is it about?) and pragmatically (what exactly does the interview want to know?)

- Find or generate relevant information in memory

- Format the answer and put it into the desired answer scheme

- Respondent must not but does: Edit answer (i.e. "adjust" to present oneself "correctly" to the interview/…er) -> undesirable (measurement error)

⇒ Questioning can change the respondent by influencing his mind, attitudes, induce learning effects… -> Interview can thus create its own artificial reality (falsification of results...) -> threat to representativeness

⇒ Examples: Product interest, usage interest, product judgement …

Therefore target group orientation is extremely important

Most important technique to increase the ability to provide information and at the same time major problem in the construction of questionnaires

Adjust to the target group of the persons regarding:

- Information level on the topic (laymen, amateur vs. professional)
- Expressiveness (Education)
- Time disposition of the respondent (respondent has to sacrifice time and usually does not receive anything in return)
- Environment of the survey
- Interest in the survey
- Kind of survey (written, oral, digital, paper etc.)

Types of questions

- There is a variety of question types and presentation options.
  Which one you choose depends largely on what statements and conclusions you want to make based on the answer.

- Before deciding on a particular format, one should ask oneself what the answers will be, what kind of data one will receive and whether one can draw the necessary conclusions from this data.

- Basically, a distinction is made between open and closed questions:

Level of standardisation

Standardisation refers to:

- Determining the wording of the question

- Order of the questions

- Preset answer options

Prerequisite for high standardisation:

-> Test questionnaire for comprehensibility, response behaviour etc.

Aim of a high level of standardisation: Extensive comparability of the survey results, independence from the individual interviewer

Questionnaire design: Open and closed questions

Open questions:
- No fixed answer categories
- Interviewer or respondent writes down answer to question verbatim

- Assignment to categories only takes place during evaluation

Closed questions:

- Pre-determined response categories

- Shapes:
    - Answer specifications "yes/no
    - Questions with alternative answers
    - Ticking for scale questions
    - Dialogue questions: Agree with statements of people in comic form with speech bubble

Mixed forms:
- Within a closed question "Other and namely ..." as a response category

- Ask open question - interviewer assigns answer to given categories (interviewer qualification!)

Open questions

- This type of question is particularly suitable if you want to have a large number of different assessments and for explorative questions, i.e. if you do not know which answer is possible:

Do you have a suggestion for something that we could do better?

_____

_____

- Open questions are particularly useful when opinions, attitudes, suggestions, etc. are of interest, where the likely answers are difficult to assess beforehand.

- Open questions are not suitable for surveys with large numbers of participants. The amount of data obtained is too extensive.

- This results in difficulties in the evaluation, since the answers to open questions are hardly comparable and a suitable processing would take a lot of time.

- Open questions must be assigned numerical codes, since analysis programs mostly only can process numbers

Advantages and disadvantages of open questions

Advantage of open questions:

- All facets can be captured

- Has a positive effect on the respondent's motivation

Disadvantage of open questions:

- Considerable analysis effort (therefore keep to a minimum for larger samples): Form categories, code answers and assign them to all questionnaires.

- It is often useful to use open questions in pre-surveys to find out what answers there are in the first place. From this, answer categories for closed questions can be generated, which can be used in the main survey.

Closed questions

- Closed questions are questions where the answer alternatives are given. They can be answered quickly and easily by the participant.

- This type of question has the advantage of making evaluation very easy. The participants can be compared or grouped.

To be noted for closed questions (frequent source of error):

- All possible answers must be included in given categories

- Non-overlapping answers (disjunctive classes)

- Unambiguousness of classes

- Exclude or explicitly allow multiple answers

The simplest form of a closed question is a dichotomous question in which there are exactly two answer options (usually yes / no).

Example:

Would you buy product XY again?
O *Yes*
O *No*

In a closed question, a participant can also hoose the answer that suits him or her from several possible answers.

Example:

I am satisfied with the service:
O *agree strongly*
O *agree*
O *neither agree nor disagree*
O *do not agree*
O *strongly do not agree*

What´s better: *Questions* or *statements* in questionaire?

Example

- *"Do you consider yourself a sociable person?„*
  yes / partly / no

vs.

Does this apply to you?

- "I consider myself a sociable person"

Agree completely / agree mostly / rather don´t agree / Disagree completely

Statements have the advantage that more differentiated alternatives can be given!

All interesting/possible answer alternatives should be listed and, there should be no need to obtain additional answers.
If you are not sure, you have the possibility to choose an semi-open answer option. Introduce "other" or "don't know"

I am satisfied with the product:
o *agree strongly*
O *agree*
O *neither agree nor disagree*
O *do not agree*
O *strongly do not agree*
O *don´t know*
O o*ther*

Number of given answer alternatives depends on how detailed you want to evaluate. It can often be useful to specify several alternatives and later group together e.g. "strongly agree" and "agree" both to be taken as agreement.

Likert rating scale

Rating procedure
A rating scale is an allocation scale in which the respondent's
person a certain characteristic expression of a stimulus with
must assign a scale value, which is selected from a given response
scale.

Example of a rating scale:
How do you like the design of brand X?
very good good medium poor very poor

1	2	3	4	5

Advantages of the rating procedures:

- easy constructability
- simple manageability
- easy comprehensibility

Co-funded by
the European Union

RED
Re-Educating
Digitisation

Rheinische
Hochschule
Köln

Determination of the response dimension - options

- Frequency (rarely - frequently; 1x a week - 7x a week)

- Intensity (not at all - very much)

- Probability (definitely not – definitely)

- Agreement (not at all true - completely true)

Construct a Likert scale, to capture attitudes toward the following:

television advertising
fast food
environmental protection
slimming diets
influencers
current policy of the german government

Rating scale construction

One-pole or two-pole ratings?
- *single-pole rating scales:* Intensity query (low-high)
- *bipolar rating scales:* diametrical pair of opposites (beautiful-ugly)
- *Practice:* often single-pole rating scales

Open or closed ratings?
- *Open ratings:* without answer specifications
- *Closed ratings:* with response specifications

Degree of gradation?
The more graded a rating scale is, the more difficult it is for the respondent to clearly assign a feature to a characteristic.
the more difficult it is for the respondent to clearly allocate a characteristic
to a particular scale value
- *Problems:* Discriminatory capacity of the respondent decreases;
- Reliability of the data values obtained decreases
- *Practice:* four- to seven-level scales

Other scales as a special form of questions

I like
very much

I like
not at all

pure continuum

vote
fully

vote
not

1   2   3   4

unipolar scales with visible coding and verbal extreme point description

Applies

Does not apply

1   2   3   4   5
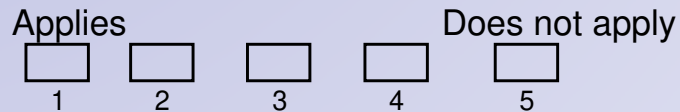
Unipolar scale with verbal description of all expressions

very
good

good

unde
cided

not
good

Not good
at all
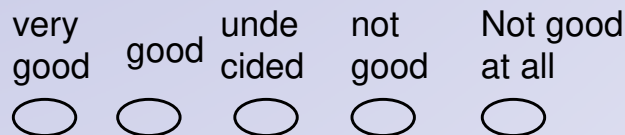
Bipolar scale with visible coding and verbal extreme point description

unsympathetic

sympathetic
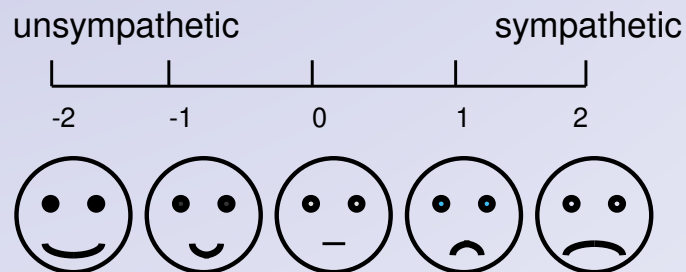
-2   -1   0   1   2

Kunin's face scale

Source: based on Berekoven, Eckert, Ellenrieder (1999), p. 75

Response specifications:
Type of scaling (e.g. degree of agreement)

- Scale of „1 / 2 / 3 / 4 / 7 / 5" versus -2 / -1 / 0 / 1 / 2 makes a difference
- Similar: *"Agree ... . Do not agree"* vs. *"Agree - Disagree"*
- Number of answer categories (generally between four and seven, in fact most often five)
- Rating (very important . . . not at all important) or ranking (most important . . . least important)
- Opinions: "Don't know" default?
- Different wording length of the alternatives
- Use of judgmental words or justification in the answer
- Inequality of positive and negative response alternatives
- Certain emphasis of words by the interviewer in an oral interview
- Suggestive questions: „Don't you agree, that .....?"

Number of answer specifications / answer categories : Even number

When determining the number and type of given answer categories, one should proceed very deliberately. A distinction is made between two basic approaches : Even number of categories

6. What of the following does the Cologne Cathedral mean to you? The Cathedral is…

Please one tick per row!

	strongly agree	mostly agree	disagree	strongly disagree
6.1 an important historic point of interest	O_1	O_2	O_3	O_4
6.2 a place of religious retreat	O_1	O_2	O_3	O_4
6.3 a place of art	O_1	O_2	O_3	O_4
6.4 an ambassador of Christian Faith	O_1	O_2	O_3	O_4
6.5 a place of veneration of relics of the Holy Three Kings	O_1	O_2	O_3	O_4
6.6 the most beautiful place in Cologne	O_1	O_2	O_3	O_4

There is no middle position here that can be used as a neutral fallback. The respondents are forced to at least give a tendency in their judgement.

Advantage: the answers can be dichotomised later, i.e. one can make a reduction in the information of the data (e.g. agreement - disagreement).

Disadvantage: A person who really wants to answer the question neutrally is forced to lean towards one side here. The respondent might feel constrained in his/her answering behaviour.

Solution: You may add a neutral "don't know" option

Number of answer specifications / answer categories : Odd number

An odd number of levels implies the existence of a "middle category".

5. Which of the following do you in general associate with the Cologne Cathedral?
Please one tick per row!

	strongly agree	mostly	neither nor	mostly	strongly agree	
5.1 inspiring	O_2	O_1	O_0	O_{-1}	O_{-2}	boring
5.2 repellent	O_2	O_1	O_0	O_{-1}	O_{-2}	hospitable
5.3 emotionally close	O_2	O_1	O_0	O_{-1}	O_{-2}	emotionally distant
5.4 warmhearted	O_2	O_1	O_0	O_{-1}	O_{-2}	cold
5.5 dark	O_2	O_1	O_0	O_{-1}	O_{-2}	bright

- Advantage: the respondent can express an undecided opinion

- Disadvantage: Odd scales bear the risk that the middle value is ticked more often than average, i.e. no positive or negative tendency is recognisable.

- In addition, an uneven number of answer categories is associated with the disadvantage that the middle category is chosen even if the respondent does not want to or cannot answer the question at all.

Solution: This disadvantage can be avoided by offering an abstention option ("don't know" category). Care should be taken to ensure that the number of categories to the right and left of the middle is the same to avoid bias.

Default response classes I

- Many subjects are then more willing to give an answer because they do not have to reveal themselves in detail

- Example: Question about income (e.g. in Germany), also weight or age (e.g. for women in the USA).

- Caution: Do not choose classes that are too large, otherwise too much information will be given away.

- Only the class midpoints are used in the calculation (problem of open final classes)

Default response classes II

- A famous example of this is the question about the number of hours a person streams videos per day:

- If one gives as choices "0.5 hours", "0.5 - 1.5 hours" and "more than 1.5 hours", the participants will conclude that the middle category corresponds to the average and the majority will place themselves there

- If one were to add more categories with higher numbers, the average would shift and the participants would watch TV "longer"

- If you have no information about the distribution of answers, you should ask an open question and builds classes later in the course of the analysis

- Check carefully before wether the interviewing-software you are working with allows for this

Multiple Answers

Are one or more answers possible? If the latter, specify by a
note in the interview: „Multiple answers possible"!

Examples:

- *What do you expect from a well-designed website?*

- *How often do you use the following media?*

- *What features do you like about this festival?*

4. Below a few reasons for visiting the Cathedral are listed. Why did you come to the Cologne Cathedral today?
Multiple responses possible!

O_1 To church service/devotions

O_2 For confession

O_3 To see the architecture

O_4 To see the artwork/the relics in the Cathedral

O_5 To be close to God/for praying

O_6 As part of a pilgrimage

O_7 Was part of an organized program

O_8 To rest/find peace

O_9 Wanted to show the Cathedral to others

O_{10} Other reasons

Item - Battery

You can arrange questions in tables or matrices. This is particularly useful if several aspects play a role in a subject area or if answer categories are often repeated.

In the following are a few reasons you might go to a festival. I go ...				
	Strongly agree	Mostly agree	Disagree	Strongly disagree
... to see an outstanding movie	O_1	O_2	O_3	O_4
... because someone convinced me	O_1	O_2	O_3	O_4
... I wanted to see a certain star	O_1	O_2	O_3	O_4
... to do something with others	O_1	O_2	O_3	O_4

- This question type is used when several statements on one topic are to be evaluated.be evaluated. This makes the presentation clearer and easier to handle.
- However, the tables should not become too long (contain too many statements), as this can be tiring when filling them out.

Items and response dimension must fit each other, i.e. not:

- *"Do you have difficulty climbing stairs?"*
  O rarely   O sometimes   O always

- No mixing of response dimensions, so don't:
  O not at all   O somewhat   O sometimes   O often   O always

- All items of a item-battery must have the same answer alternatives, i.e. not:

  "Do you often feel dizzy? " O yes O no
  "Do you suffer from weather sensitivity?" O never O sometimes O always

How to develop an Item – Battery?

- Items are derived from existing theories or scientific literature.
  Example: Read about theories about film selection or film consumption

- Use existing questionnaires from other surveys (adjective lists, rating scales, interview questionnaires, etc.);
  they can serve as a source of ideas for the formulation of questionnaire statements

- Own experiences, everyday observations can be used to generate questionnaire items

- Preliminary qualitative, open-ended research with affected persons can be conducted

- Interviews with experts can be used to derive questionnaire items

Semantic differentials / polarity profiles

- Semantic differentials or polarity profiles are questions where the participant has to decide which of two opposing statements he (rather) agrees with.
- They are particularly suitable for interrogating settings. When designing them, it is important to ensure that the two poles are really opposite.

5. Which of the following do you in general associate with the Cologne Cathedral?
Please one tick per row!

	strongly agree	mostly	neither nor	mostly	strongly agree	
5.1 inspiring	O_2	O_1	O_0	O_{-1}	O_{-2}	boring
5.2 repellent	O_2	O_1	O_0	O_{-1}	O_{-2}	hospitable
5.3 emotionally close	O_2	O_1	O_0	O_{-1}	O_{-2}	emotionally distant
5.4 warmhearted	O_2	O_1	O_0	O_{-1}	O_{-2}	cold
5.5 dark	O_2	O_1	O_0	O_{-1}	O_{-2}	bright

Fractionation of answers

During fractionation, the respondent is presented with two stimuli presented at the same time. She is then asked to provide a numerical To give an estimate of the ratio between the two stimuli, the is present in a particular property.

Example:

Compare the artistic quality of different film festivals with the festival in Cannes as a benchmark: For Cannes, assume a value of 1.0."

Ratio of sharpness compared to A:

Festival	Answer
Cannes	1.00
Venice	0.5
Locarno	2.0
...	0.75

Constant summation scale

Constant sum methods provide the respondent with a certain number of points to be distributed among the stimuli in such a way as to reflect their relative importance in relation to the respective trait characteristic.

Example :
"Distribute 100 points among the following five festivals in such a way that it results in the extent to which you rate the individual festivals.

Festival	Answer
Venice	20
Cannes	25
Locarno	10
Zurich	5
...	40
	100

Rating scale

- Rating procedure:
  Respondents should make an assessment of all alternatives.

Which is your favorite type of movies? Tell me wether you like it very much, somewhat, somewhat less, not at all.

- action
- adventure
- crime
- thriller
- drama
- family
- documentary
- comedy
- musical/dance
- sci-fi
- horror
- western
- war
- epics
- other

Ranking scale

Ranking procedure: Questions where a ranking or hierarchy is formed are suitable for comparisons between different answer options.

Respondents are asked to rank

Which is your favorite type of movies?

- action
- adventure
- crime
- thriller
- drama
- family
- documentary
- comedy
- musical/dance
- sci-fi
- horror
- western
- war
- epics
- other

Questionnaire design: Functional questions

☐ Often less important in terms of content, but indispensable for the process and success of the survey:

- *Introductory questions* (icebreakers) should reduce inhibitions to answer, arouse interest. They should be easy to answer, but not banal or completely irrelevant

- *Buffer questions* (transition questions) are used when context or spillover effects are to be avoided or at least reduced

- *Concluding questions* end the content part of the survey

- *Controll questions* check the validity of answers

- *Filtering* and *forking issues ->*

Function of filter questions

- Goal: Exclude certain question complexes for some of the test persons -> save time, omit unimportant items, keep motivation high

- But: Filters must be easy to recognise:

    - for face-to-face interviews, interview training necessary

    - clear layout for written questionnaires in PAPI

    - for online (or telephone) interviews predefined by software (optimal)

Questionnaire design: Functional question filter

Question 1: Do you smoke?

    yes   no

Question 2:
What do you smoke?
☐ electr. cigarette
☐ Cigarettes
☐ Cigars / Cigarillos
☐ Whistle
☐ Other, <u>namely</u>

Question 3: How many per day?

Question 4: Do you drink alcohol?

Questionnaire design: Functional question bifurcation

Question 1: Do you smoke?

☐ yes

☐ no

Question 2:
Did you smoke in the past?

Question 3:
What do you smoke/smoked?
☐ electr. cigarette
☐ Cigarettes
☐ Cigars / Cigarillos
☐ Whistle
☐ Other, <u>namely</u>

☐ yes no

Question 4: How many per day?

Question 5: Do you drink alcohol?
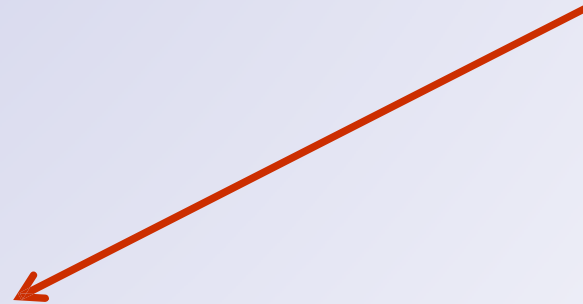
Exercise task (1)

A cinema house would like to find out the customers' attitudes towards movies. The owner has drafted a questionnaire for a written survey and asks for expert opinion.

Questions:

1*. It never happens to me that I am not informed about the latest film trends:*
        yes / partly / no

2. *I am already looking forward enthusiastically to the coming movie season:*
        yes / no

3. *There is no better business than film business: yes / no*

4 *I consider dealing with film issues is not only unimportant, but ultimately also a waste of time and therefore harmfull:*
                        Applies / partly applies / does not apply

Exercise (2)

5. Going to the cinema is an attribute of dynamic people: yes / no

6. I like reading movie magazines because I like to go to the movies:
strongly agree / agree / disagree

7. How do you judge this cinema and the movie you have seen? Please express your
opinion in school grades from (1) Very Good to (6) Unsatisfactory.

8. When I buy ticket, I like to pay with:

> *- with EC card*
> *- with VISA Gold Card*
> *- cash*
> *- with American Express*
> *- with PIN*

Typical errors in the wording of the question

Error…

On the questions	For the answers	In the formulation
• Too general • Several thoughts per question • Reference to current events • No explanations given/ examples for unfamiliar issues • Questions without reference to use; exception: Icebreaker or functional questions • Putting the analysis into the question itself	• Unclear answer reference: place, time and context unclear for respondent • Inconsistent, incomplete or overlapping answer specifications • Unbalanced answer specifications • Forgetting "does not apply" or "do not know" when needed	• Ambiguous/unclear • Suggestive/ stereotypical • Too long/complex • Too extreme • Negations • Foreign words, dialect expressions, words with double meanings (homographs) • Unbalanced formulation

Rules/tips for formulating questions I

- As a rule: Simple formulations adapted to the *language* and *knowledge level* of the respondents
- Questions should be short (especially for telephone or oral interviews)
- No foreign words
- No double negatives or no negation at all in the question
  Ex. From a survey: Not using streaming services for watching movies means that you are not up to date as a consumer. Do you not agree? yes / partly/ no
- Neutral wording, balanced including all alternatives given in the answers
- Questions not too complicated, straightforward and simple
  Ex. from a survey: "How much money do you spend per capita in your household on media usage including all subscriptions, ticket purchases, special item purchases (e.g. books, magazines, CD) and purchases for technology devices (tv, cellular, tablett etc.)?"

Rules/tips for formulating questions II

- Avoid statements that can be interpreted in more than one way
  (*"How do you like the german chancellor Olaf Scholz?"*)
- Avoid statements that are irrelevant to the subject being researched
- Avoid statements that can be affirmed either by almost everyone or by almost no one (means question should not be too easy or too heavy)
- Select statements that utilise the full range of the scale of judgement of interest
- Choose simple, clear, direct language
- Statements should be short and rarely contain more than twenty words
- Each statement should contain only one complete thought
- Items must be applicable to all persons who may enter the sample
- Every possible answer (even a negative one) must be clearly interpretable

Rules/tips for formulating questions III

- No multidimensionality in individual questions:

– *Smokers should pay higher taxes because they are partly to blame for high medical costs,* *versus*

– *Smokers should be/are...?*

- Specify the reference frame (e.g. time or other) as precisely as possible:

– *How much do you spend on media per month / paital etc.often did you go to the movies in the last 12 months?*

General: Provide all answer options exhaustively ("Other" only as a stopgap).

Answer options unidimensional, must not overlap → Negative example next page

From: "Fit from 50 through proper nutrition" (part of the BMFMG's "fit in old age" campaign)

How many meals do you eat a day?

- I eat irregularly, often only 1 or 2 times a day
- When I don't know what to do, my way quickly leads into the kitchen
- I eat 4 to 5 times a day, evenly throughout the day

How much fluid do you consume per day?

- The most I can manage is 1 litre a day
- I hardly ever drink because I'm never thirsty
- I must drink 1.5 litres a day

Rules/tips for formulating questions IV

Answer specifications as explicit as possible (especially for questions about behaviour):

- *Rather not: How often do you go to a movie?*

  Regularly - now and then - rarely - never

- *But rather: How often have you gone to a movie in the past 12 months?*

  Serveral times a week / ...a month / ...a year / once a year / more rare / not at all

- *But: There is possibly influence on the answer specifications! Therefore: How often have you gone to the film festival in the past 7 days?*

  _____times

Rules/tips for formulating questions V

Delicate or sensitive questions, e.g. about socially undesirable behaviour or behaviour that is generally not talked about with strangers (e.g. extreme political opinions, money, attitude towards corona measures or such opinions that are strongly against the public debate).

How to cope with:

- *Reference to wide distribution: "Many people nowadays leave their mobile phones on in the cinema. How about you: Have you ever ... "*

- *Downplaying: "Leaving your mobile phone on at the cinema usually doesn't bother anyone. Have you ever ... "*

- *Catch you off guard: "When was the last time you left your mobile phone on in the cinema?"*

Better option in personal interview: Use self-completers (= answers not directly to interviewer, no interviewer effect!)

Rules/tips for formulating questions VI

➢ Avoid monotonous response behaviour with item batteries:

The following are a few reasons you might go to a movie. I go the movie...				
Please mark each line once!	Strongly agree	Mostly agree	Disagree	Strongly disagree
… to learn about films / educate myself	O_1	O_2	O_3	O_4
… for aesthetic enjoyment	O_1	O_2	O_3	O_4
… to meet other people interested in movies	O_1	O_2	O_3	O_4
… for entertainment/leisure	O_1	O_2	O_3	O_4
… to do something with others	O_1	O_2	O_3	O_4

Rules/tips for formulating questions VII

Item

"I have no interest in movies "

Critique

The item itself does not contain a double negative.
However, if it is negated, there is one overall:
Someone with interest in movies must answer that they have "not no" interest in movies; this is too complicated and leads to errors in the answer.

Rules/tips for formulating questions VII b

Item

Critique

"Given the quality of this festival, is this comparable to others you know?"

Comparable it is in any case, the question should aim at the result of the comparison.

Even if respondents understand the question as asking whether the quality is "as good" as that of others, a "no" would not be interpretable, it could be mean "better" or "worse".

"Have you considered to come more often to this festival than in recent times?"
Yes / No

This item is ambiguous. Both an affirmative and a negative answer would not be interpretable.

Rules/tips for formulating questions VII c

Item

"How often have you been to the movies in the last year?"

Critique

The time specification is problematic with this item. If this question is asked in November 2022, the respondent may understand the last year as either the calendar year 2021, the year 2022 or the period Nov. 2021 to Nov. 2022 (better "in the past 12 months")

RED
Re-Educating
Digitisation

Rheinische
Hochschule
Köln

Rules/tips for formulating questions VII d

Item	Critique
Have you taken use of admission discounts to the festival?	This item assumes the right to use of an admission discount. A "no" can stand either for not being entitled to do this or the lack of knowledge about or lack of interest. Here, it must first be clarified with a filter question whether the respondent is entitled to use a discount at all in order to then ask about usage and knowledge of discount.
Have you any problems when selecting the movies you like to see by the program website?	This question does not apply to all potential respondents, as many people do not use the program site, or select systematically (e.g. come with others that do the selection, choose movie spontaneously or at random etc.)

Rules/tips for formulating questions VII e

Item

"How do you get to the festival regularly?"
- by car
- on foot
- by bike
- by public transport

Critique

This item is in every respect unsuitable for measuring popularity of public transport.
Apart from the fact that it presupposes that the respondent comes regularly, owns a car, a bike and that public transport exists in the place of residence, the choice of means of transport depends also on the distance between the place of residence and the festival.
Someone who lives closely to the festival does not use public transport.

Source of error: Question wording

The question wording can cause large differences in the results:
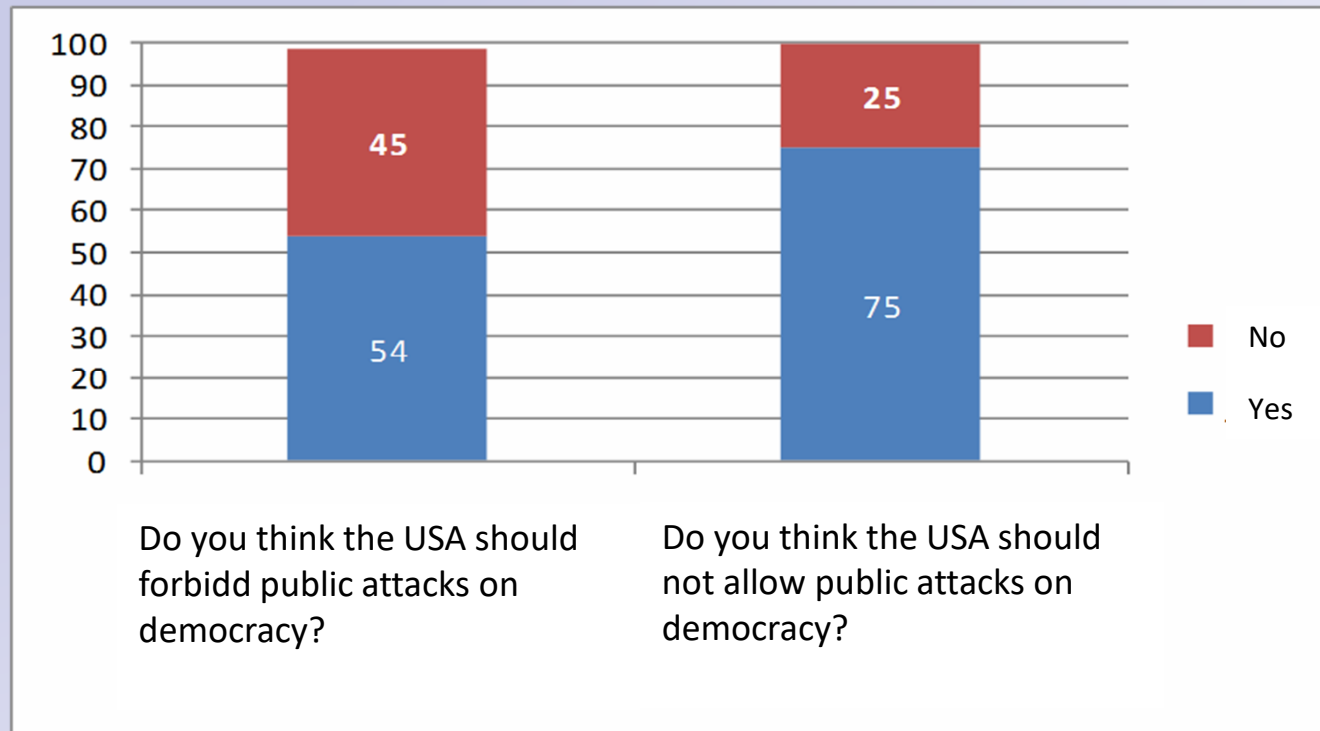
Do not formulate questions that the respondent tends to give a particular answer because of the wording or the way it is presented – event if you were in favour of a certain result

Instrument effects I: Errors in question wording

- Obviously suggestive questions or only single terms

- Choice of single words: "Do you think ....should ban?" vs. " ....should not allow"

- No specification of alternatives (unbalanced formulation):
  "Are you for public subsidies for art institutions?" vs. *"Are you for public subsidies for art institutions or are you against?"*

- Building in justifications for the answer in the question itself: „
  „Are you for public subsidies for art institutions, because art is urgently needed just today?"

- Too complex question wording: Taken from a survey among students (by the „Student Service" of the university of Cologne):

  "Please tell me how much money you have available in total per month, i.e. support from parents incl. rent subsidies, BaFöG, scholarships, student loans, governmental child benefit, regular support from other people (e.g. grandparents or other relatives), your own work and other sources?"

Sources of error:
Response set and lack of opinion

Def. „Response Set" = Content-independent response tendency

Respondents always tick the middle (or penultimate, etc.) category in a questionnaire with several measurement items

- Participants tend to take a middle position ("partly/partly", "neither...nor") regardless of the question, "maybe") to be ticked

- To counteract this tendency, it is important to formulate the items very precisely and accurately.

- In addition, one could consider using a scale with more levels (e.g. a seven- instead of a five-level scale), as this would allow for more precise differentiation.

- Or use an even number of categories with no point in the middle

8. On a scale of 10 (highly recommend) to 1 (wouldn't recommend at all), how likely are you to recommend visiting the Cologne Cathedral? You can choose any number in between.
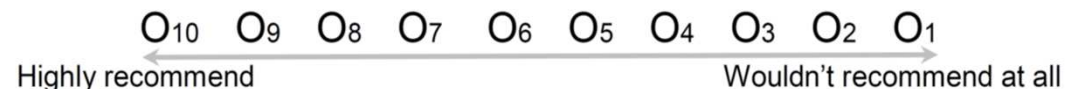
O_{10} O_9 O_8 O_7 O_6 O_5 O_4 O_3 O_2 O_1

Highly recommend   Wouldn't recommend at all

Sources of error:
Response set and lack of opinion

- Def. „Response Set" = Content-independent response tendency

  - Response tendency or acquiescence - example:

- There are people who tend to agree in principle with a statement regardless of its content ("yes-say" - tendency, (content-independent approval tendency).

- In order to exclude these distortions of the questionnaire, one should include some controll questions. These control questions are semantically rotated, i.e. formulated negatively.

- One would phrase some of the items negatively and have these new statements rated.
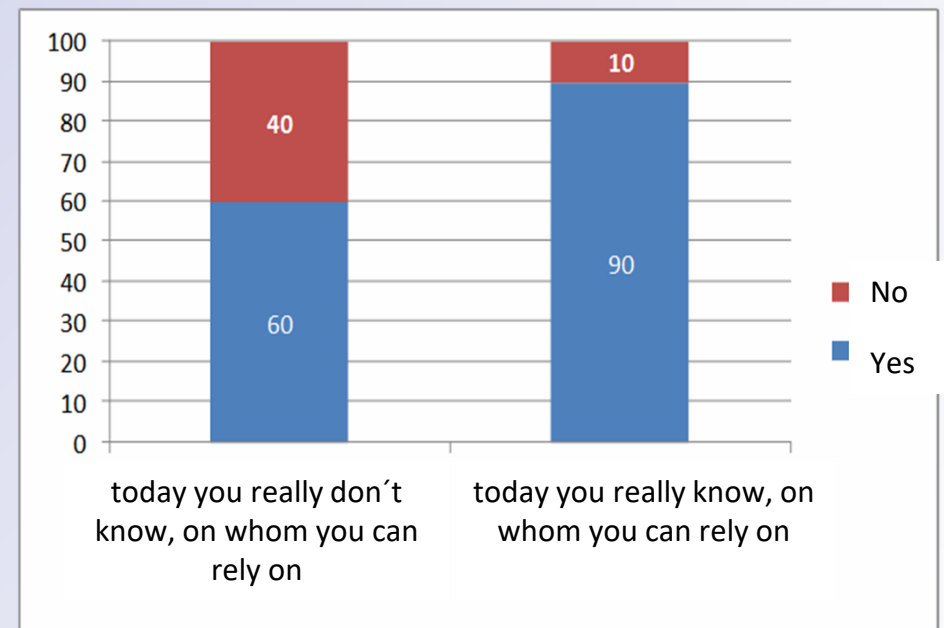
Countermeasures for response set and lack of opinion

- Response set
  - ☐ Item batteries
  - ☐ Vary response categories
  - ☐ Avoid middle category
- Response trend
  - ☐ Item battery should contain both negative and positive phrases
  - ☐ Formulate alternative answers

- Lack of opinion

  - ☐ Filter: "Do you know the problem", "Do you have an opinion on ...?"
  - ☐ Categories "do not apply" "do not care" or "do not know

Exercise: Go through the questionaire carefully and find out the mistakes!…

Questionnaire design: Aim of a good design

- A well-constructed questionnaire should serve to avoid errors and bias in response behaviour

- It should support the respondents (and possibly the interviewers)

- Three aspects are of particular importance:

  - The visual design (layout)

  - The content structure

Questionnaire design: Layout

- Orientation towards the target group (age, readability etc.)

- The respondents with a self-completion questionaire usually have no help from interviewers, so the questionnaire must be *self-explanatory*

- The layout should radiate seriousness, importance and ease of use

- Indications of the type of answer marking, assistance

- Filter guidance must be made clear

- Question together with suggested answers on same page

- At the end: Express thanks once again

Questionnaire design: Dramaturgy / structure

- ☐ Order of questions in thematic blocks (modules)

  - Start with easy, general questions, that motivate the respondent; collect „Yes"-answers („Icebreaker"); crucial for the respondent´s motivation

  - For very different topics: *Transitions*

  - Within the modules from easy to difficult questions

  - Demographics always at the end

- ☐ Module sequence (tension curve)

  - Beginning easy, middle, than hard questions, end socio-demographic

  - Do *funneling within the blocks* (from general to specific oder vice versa)

- ☐ For "*sensitive* issues"

- Tricky questions always at the end: When respondent drops out, a large part of the FB is already filled in

- *Funnel questions*: You start a topic with the more harmless questions and ask the trickier questions afterwards: slow habituation.

Questionnaire design: Socio-demographic information

- Such as gender, age, school-leaving qualification, field of study, personal income, household income, household size, marital status, children, information on the family of origin or migration background.

- Life-style

- Exclude objective information that could in principle also be obtained in other ways.

- With demographics there are often standard instruments available (questions), but they have to be adapted to the specific conditions respectively to the target group and the research interest.
Example: Measuring income, viewing habits

Error source question sequence and context

Halo effect

- Questions can radiate to other questions, one also speaks of sequence effect or context effect.

- These can play a role, for example, in questions about general life satisfaction when positive (or negative) memories are activated.

- Radiation effects can also be intentionally employed

Countermeasures

- Testing in the pretest (e.g. with questionnaire splits)
  = Two groups are presented with different versions of the questionnaires in different order.

- Widely divide issues affected

Example: Context effect in attractiveness measurement

The positioning of a question - relative to neighbouring questions and in the questionnaire as a whole - influences the response behaviour

How attractive or unattractive do you rate this person?

unattractive   attractive

Example:
Context effect in attractiveness measurement

2,55   2,40   1,88

Instructions, directions and guidance for participants

Besides the questions, the instructions for the participants play an important role in the construction of the questionnaire.

Each questionnaire should contain information in which the following information can be found:

- General purpose of the survey
- Responsible organiser of the survey
- Time required for processing
- When online or mail survey: Closing date
- Contact person with telephone number and email address for queries
- Information on how anonymity or confidentiality will be ensured
- What to do with the completed questionaire
- A thank you!

One should never assume that an answer format is self-explanatory, otherwise one runs the risk of getting incorrectly completed questionnaires or that participants do not complete the questionnaire at all.

The easiest way to clarify this question is a pretest. If ambiguities arise here, an answer type must be explained.

Comments and suggestions to the questionaire

At the end of the questionnaire there may be space for comments and suggestions from the participants.

This is:

- Polite,

- helpful in revealing possible misunderstandings, and

- it also gives a kind of relief to the very involved respondents.

For example: If a participant expresses that they find the questions incomprehensible, caution should be exercised when scoring that participant's answers and the individual should be omitted from the overall scores.

End of this section !

RED
Visitor Research I

- IV. „Theme Packages II" -

Author:
Prof. Dr. Tibor Kliment

RED
Re-Educating
Digitisation

Rheinische
Hochschule
Köln

Research questions come first

Participant quote: "...I'm actually concerned with questions that I don't have, insights that I cannot draw and data I am suspecting."

RED Re-Educating Digitisation

Rheinische Hochschule Köln

Overview of the research concept - from the research question to the data

Problem definition → **Formulation of objectives** → **Determination of a research design**

For example audience description, evaluation of measures, improving promotion, better programme work

Targeting new audiences, optimized communication, more turnover, better visitor loyalty…

explorative, descriptive, quantitative, qualitative

Development and selection of data acquisition method → **Development of an analysis plan** → **Data-procurement**

Questionnaire design, sampling, indicators, questions, scaling, etc.

Data analysis → Data interpretation, results, reports, key figures, new questions...

Theme Packages

Theme packages and their meanings:

1. **Audience structure / context of the visit**

2. **Visiting behaviour / interests / decision making**

3. **Visitor expectations, motives and satisfaction**

4. **Distribution / Ticketing / Pricing**

5. **Communication**

6. **Effects on Tourism**

Selection of certain packages in a survey is strongly needed!

1. Audience structure / context of the visit

Always the fundamental basis for the description of the audience are the socio-demographic data:

- **This includes visitor characteristics such as age, gender, income, professional status, household size.**

- **Further standard information concerns personal geographical data such as the place of residence or the nationality of the visitors, evt. migration.**

Socio-demographic audience data can be supplemented by data on the audience's wider social environment and context of visit:

- **The composition of the group with which the visitor comes**
- **Who decides for the visit**

1. Who visits: Audience structure / context of the visit (I)

1. What is your gender?

- Female
- Male
- Diverse

2. What year were you born? *Please write full year*: _____
   -> Alternatively: How old are you? *Please write*: _____

3. Where do you live?
(1) In XY
(2) Within 50 km
(3) …
(4) Outside the country (please specify): _____

-> Alternatively: Specify post code

4. What level of general education have you completed'
(1) Up to 10th year
(2) Secondary school without A-levels/high school diploma
(3) A-levels/high school diploma
(4) University Degree (Technical College /University)
(5) Other
-> Alternatively: Years of school/university visit

- **Demographic questions are a general and indispensable basis of any visitor analysis (despite other concepts such as lifestyles, soc. milieus etc.)**

- **Age, gender, occupation and formal education, among others, are important indicators for measuring diversity in the audience.**

- **Age in particular is still one of the most important criteria for describing visitors:**

  Classic proxy variable, shaped by age cycle, cohort and period effects. Also by spec. values and needs, resource endowment etc.

- **Note: The age of the visitors is slightly overestimated in the year of birth-question.**

RED Re-Educating Digitisation

Rheinische Hochschule Köln

1. Who visits: Audience structure / context of the visit (II)

5. What is your personal monthly income?

(1) 0 - 500 €
(2) 501 - 1.000 €
(3) 1.001 - 1.500€
(4) …
Categories should have same size!
Make them not too big and not too small!
Income is always a delicate question!
Other wordings possible (e.g. money available etc.)

6. What is your professional status?
(1) Managerial level employee/Senior Civil Servant
(2) Salaried employee/Civil servant
(3) Self-employed
(4) Currently unemployed
(5) Schoolchild/Student/Trainee
(6) Pensioner
(7) Housewife/Househusband
(8) Other

- **Important für pricing policy of the festival**
- **Analysis of time budgets**
- **Analysis of disadvanteged groups**
- **Reach of target groups**
- **Assessing needs, wants and motives of the audience**
- **Analysis of the composition of the audience**
- **Most important for typing the visitors for further analysis (tabels)**

Locate the demograhics always at he end!

Migration status

8. What is your nationality?

7. Could you please tell us if one (or more) of the following points is true for you or your parents?
(1) You yourself or at least one of your parents was born in a country other than Germany
(2) You yourself or at least one of your parents is a naturalised german citizen

9. Who are you here with today?
More than one answer possible!

(1) I am here alone
(2) with my partner
(3) with colleagues
(4) with relatives, friends, acquaintances
(5) in an organised group
(6) other

10. And who has decided the visit?
(1) Me
(2) Others (partner, colleagues etc.)
(3) We decided together
(4) Can´t specify

Most demographic characteristics develop their analysis potential in connection with other visitor characteristics!

- **Important für pricing policy of the festival**
- **Analysis of target groups**
- **Assessing needs, wants and motives of the audience**
- **Analysis of the composition of the audience**
- **Most important for typing the visitors for further analysis (tabels)**

Gathering information on

- **the reasons for or interests in the visit**
- **size and nature of the social environment of the visitors**
- **the decision-making processes for the visit (women often play a decisive role)**
- **design of group offers with regard to content/topics, group sizes, composition and pricing**

Visiting behaviour / interests / decision making

Data on audience behaviour should always be collected. This includes questions on the frequency of visits, what was seen, first/repeated visit etc.

Visiting behaviour / interests / decision making

Are you visiting our festival for the first time today?

(1) yes, first time
(2) no, was already here
(3) do not know

How often did you visit this festival in the past three years, including this visit? (Only one visit per festival!)
Please write_____

How often do you visit/will you visit in this festival week?

(2) Once
(3) 2-3 times
(4) 4-5 times
(5) 6+ times
(6) do not know
(7) not specified

- **Important for typing the visitors for further analysis**
- **Clarifies the extent to which there is a balance between first-time and regular visitors Purpose: High proportion of repeat visitors signals bonding effect and good visit experience; share of first-time visitors signals hurdles with the facility**
- **Analysis of the communication channels and success by visitor types**

A possible target is a ratio of 1/3 first-time and 2/3 repeat visitors, but this depends on many aspects (the context, the program etc.)

- **Informs about binding effect of the festival**
- **Enables typing and, if necessary, prioritisation of visitors**
- **Also enables the analysis of "non-visitors" to some extend**
- **Important for pricing and combined offers**
- **But: Problem of overreporting in the data**

Visiting behaviour / interests / decision making

Which of our movies/events/offers was the reason for your visit here today?

Include permanent and special events of the festival

…

…

Which of our movies/events/offers did you see/use during your visit today?
…
…

Question for

- **the use of offers,**
- **topic interests and non-interests,**
- **specificity of needs,**
- **visitor retention**

Particularly interesting in combination with:

- **demographics,**
- **visit frequencies,**
- **residence,**
- **communication media etc.**

Analyses as above, in addition:

- **The comparison between both questions is interesting:
  What was the occasion and what was actually visited?**

- **Gives clues for composite offers/ticketing**

Visiting behaviour / interests / decision making

Which is your favorite type of movies?
Please mark each row („yes"/"no").

- action
- adventure
- crime
- thriller
- drama
- family
- documentary
- comedy
- musical/dance
- sci-fi
- horror
- western
- war
- epics
- Other

Which is your favorite type of music?
Please mark each row („yes"/"no").

…
…

Question for

- the use of offers,
- topic interests and non-interests,
- specificity of needs,
- visitor retention

Particularly interesting in combination with:

- demographic data,
- visit frequencies,
- residence,
- communication media etc.

In addition:

- Gives clues for program work, and
- the media campaign,
- interesting in combination with other variables (demographics etc.)

RED Re-Educating Digitisation

Rheinische Hochschule Köln

Visiting behaviour / interests / decision making

What specifically prompted you to visit us today? How important were the following points for you?
In the following, I will name various aspects and you please tell me on a scale of 1 to 5 how important they are for you!
(1) "very important",
(2) "rather important",
(3) "partly/ partially",
(4) "rather unimportant" and
(5) "very unimportant".
(6) "I cannot judge"

- Genre of the movie
- Topic/content
- Major artist / performer
- Author
- Composer
- Music
- Reputation of the event/festival/production
- Location of the venue

Relevant question on the issues:

- **Motivation of the visitors**

- **Interesting with table breakdown: Which visitors have which information needs (e.g. first-time visitors vs. regular visitors, age-groups)?**

- **Important for comparing the occasion fo the visit with what the visitor has actually seen/visited**

- **Communication (what to highlight in the campaign, for whom?)**

Visiting behaviour / interests / decision making

After seeing the movie, what will you you talk about to friends regarding the movie in a positive light?

(1) Lead actor 1
(2) Lead actor 2
(3) The sound
(4) The story
(5) Subplots
(6) Dialogs
(7) The Music
(8) The comfort in the theatre hall

Relevant question on the issues:

- **Impact of word of mouth**

- **Interesting with table breakdown: Which visitors have which information needs (e.g. first-time visitors vs. regular visitors, age-groups)?**

- **Communication (what am I highlighting in the media, for whom?)**

Visiting behaviour / interests / decision making

How long did your visit in XY (or in the building and outside of XY) last?

Only the currently visited venue is meant here, NOT several venues of a festival spread across the city!

In the building:
(1) 0- 30 min
(2) > 30 min-1 hr.
(3) > 1 hr.-1.5 hr.
(4) > 1.5 hrs.-2 hrs.
(5) > 2 hrs.-2.5 hrs.
(6) > 2.5 hrs -3 hrs
(7) > 3 hrs.-3.5 hrs.
(8) > 3.5 hrs.-4 hrs.
(9) > 4 hrs.-4.5 hrs.
(10) > 4.5 hrs.-5 hrs.
(11) > 5 hrs -5.5 hrs
(12) > 5.5 hrs.-6 hrs.
(13) > 6 hrs
(14) not yet completed
(15) is not visited
(16) do not know
(17) not specified

- **Indicator of the extent/type of interest in the content offer/mode of attendance**
- **Indicator for the quality of stay in the facility**
- **Notes on the importance of services in the facility**
- **Meaningful in combination with demographic data**

Visiting behaviour / interests / decision making

Did you plan your visit today in advance or did you spontaneously decide to go to the festival today?

(1) planned - just wanted to go here
(2) planned - visit combined with other activities combined
(3) spontaneous
(4) do not know
(5) others decided
(6) not specified

Similar interpretation as in previous slide for question:

- **Binding force/attractiveness**
- **Typification/prioritisation of visitors**

Indicator e.g. for

- **the kind of interest in the visit,**
- **the nature of decision-making, and**
- **interest in the facility/offer in general**

Visiting behaviour / interests / decision making

Have you visited the following cultural and leisure activities in the last 12 months (not including today's visit)?

(1) Film screenings/ cinema
(2) Libraries
(3) Adult education centre
(4) Literary events/ readings
(5) -
(6) -
(7) -

... in the area of stages/music offers:
(8) Speech theatre/acting performances
(9) Musical, revue, variety and show performances
(10) (Word) Cabaret/ Comedy Events
(11) Ballet, dance theatre performances
(12) Opera performances
(13) Classical music concerts
(14) Jazz/new music/sound art concerts, experimental music concerts
(15) Other concerts (rock, pop, hip hop, metall, electro, folk etc.)
(16) Clubs

... and in the area of museums/exhibitions:
(17) Natural history, science and technology exhibitions
(18) Cultural, historical and archaeological exhibitions
(19) Art exhibitions/ galleries
(20) Memorials/ Places of Remembrance

(21) do not know
(22) not specified

Identification of:

- **Highlight cultural leisure interests and lifestyle of the respondents**

- **Identify potentials for joint offers / possible cooperation partners for the institution**

- **Clues for cooperation with suitable institutions (program, communication etc.)**

- **Findings about the context of the respondents' place of residence**

RED
Re-Educating
Digitisation

Rheinische
Hochschule
Köln

Visiting behaviour / interests / decision making

Which means of transport do you usually use to get to
the festival? Put appropriate filtering questions before!
(1) Car
(2) Taxi
(3) Coach
(4) Motorbike
(5) Bus/train
(6) on food
(7) Bicycle
(8) Neither, but_____

How would you rate the overall accessibility of
the festival?
(1) Very good
(2) Good
(3) Neither good nor bad
(4) Bad
(5) Very bad
(6) Does not concern me
(7) Cannot answer

If you have marked "bad" or "very bad": Would
you please clarify your opinion?_____

Identification of:

- **Use of means of transport**

- **Satisfaction with public transport and other means to come to the festival**

- **Basis for talks with lokal administration**

- **Interesting for segmentation for place of living, age etc.**

Visitor motives, expectations and satisfaction

3. Visiting behaviour / motives / decision making

Some general remarks on the meaning of audience satisfaction (I)

- **Visitor satisfaction** is important because the **impact** of a festival visit, the **propensity for further visits**, the **dissemination of the experience among peers** etc. depend largely on it.

- Visitor satisfaction is more about the **short-term effects** than the long-term effects on attitudes towards the festival. For example, "satisfaction" can also arise when a festival-goer feels stimulated or excited by surprises, new designs, provocation and unusual representations.

- Satisfaction with a performance is also not only the result of the actual core performance of the festival - the "content´s quality" - but also of a **multitude of additional services.**

- These include, among other things, the staff, the atmoshere, the seating, sound, the food and beverage services, the toilettes, cloakroom etc.

- For this reason, it is important to find out about satisfaction with the core performance of the festival and also with additional services from the visitors' point of view.

- Finally: If it is found often, that those in charge of the cultural offer and its audience have very different expectations and also that there is a high level of dissatisfaction among the visitors.

Some general remarks on the meaning of audience satisfaction (II)

- However, **"satisfaction" or "dissatisfaction"** are emotional states - they can change relatively quickly (or can be changed relatively quickly by the festival)

- Satisfaction results from a **comparison** of the **expectations a visitor has** of a visit to the festival with the **experience** he or she has of a specific visit:

- **Experience** ⟷ **Expectations**

⬇

Satisfaction

-> If his expectations are not met, the visitor is dissatisfied. The absolute level of expectations is less relevant than achiving a positive or zero difference

The expectations of a festival´s visit is shaped by
- one's own experiences in the past,
- general needs/desires,
- information received by exchange of information with other visitors or through the theatre's information policy,
- expectations that relate to the performance of a festival and "additional services" such as the friendliness of the staff, the information policy, the intermission buffet.

RED Re-Educating Digitisation

Rheinische Hochschule Köln

3. Visiting behaviour / motives / decision making

Some general remarks on the meaning of audience satisfaction (III)

- **Respondent is often unaware of his/her expectations and how they influence their evaluation**
- **Or that he can only express vaguely what he feels.**
- **He also is often not aware of his or her satisfaction.**
- **It therefore is important to ask very specifically about the expectations and especially the evaluations of the individual performance elements of a festival in both areas, the service and the program offer.**

3. Visiting behaviour / motives / decision making

How do the performances meet your expectations? Please mark with a cross!

(Please one cross per line!)	meets...			
	fully	predomi nantly to	less to	not at all
Varied game plan	1	O	O	O
Sophisticated productions	O	O	O	O
Excellent music performances	O	O	O	O
Impressive movies	O	O	O	O
Gripping performances	O	O	O	O
Great ambience	O	O	O	O

Relevant question on the issues:

- **Motivation of the visitors**

- **Interesting with table breakdown: Which visitors have which evaluations (e.g. first-time visitors vs. regular visitors, age-groups)?**

- **What could be improved?**

- **Communication (what am I highlighting in the media, for whom?)**

Visiting behaviour / interests / decision making

After seeing the movie, would you rate the film?

(1) Excellent
(2) Very good
(3) Good
(4) Fair
(5) Poor

After seeing the movie, would you recommend others to see it at a theatre?

(1) Definitely
(2) Probably
(3) Might/might not
(4) Probably not
(5) Definitely not

After seeing the movie, did it meet your expectations?

(1) Better than expected
(2) About what I expected
(3) Not as good as I expected,

…

Question for

- **the use of offers**
- **topic interests and non-interests**
- **specificity of needs**
- **visitor retention**
- **impact of word of mouth**

Particularly interesting in combination with:

- **demographic data**
- **visit frequencies**
- **residence**
- **communication media etc.**

Analyses as above possible, in addition:

- **The comparison between both questions is interesting: What was the occasion and what was actually used?**

- **Possible notes for composite offers/ticketing**

In the following there will be mentioned several aspects of our **service.**
Please tell me on a scale of 1 to 5 how satisfied or dissatisfied you are in this regard.

(Please one cross per line!)	very happy	quite	less	not at all satisfied	I cannot say
Subscription system	O_1	O_2	O_3	O_4	O_5
Ticket sales at the box office (evening/theatre box office)	O_1	O_2	O_3	O_4	O_5
Ticket sales via the Internet	O_1	O_2	O_3	O_4	O_5
Cash desk opening hours	O_1	O_2	O_3	O_4	O_5
Evening staff (cloakroom, sales, ticket contr.)	O_1	O_2	O_3	O_4	O_5
Catering/restaurant	O_1	O_2	O_3	O_4	O_5

- **Differentiated evaluation of the service as well as the content offer. Service determines up to 50% of the visitor overall experience.**

- **Every visitor has different requirements: Particularly meaningful in breakdown tables with age groups, regular/occasional/first-time visitors, etc.**

- **However, only the satisfaction with, not the importance of the factors is surveyed! This can be checked by other questions. Or by statistical analysis or by checking open answers regarding possibilities for improvement.**

- **Moreover, satisfaction judgements are strongly influenced by situational factors and moods.**

-> Therefore, time series comparisons are particularly useful here!
-> These satisfaction questions are recommended for use as standard questions

3. Visiting behaviour / motives / decision making

What evaluations do you <u>generally </u>associate with the festivals´ program?
(Please one cross per line!)

	strongly agree		neither nor		strongly agree	
15.1 alive	2	1	0	-1	-2	**monotonous**
15.2 innovativ	2	1	0	-1	-2	**traditional**
15.3 important	2	1	0	-1	-2	**unimportant**
15.4 adapted	2	1	0	-1	-2	**provocative**
15.5 service-oriented	2	1	0	-1	-2	**not serviceori.**
15.6 expensive	2	1	1	-1	-2	**cheap**

Relevant question on the issues:

- **Important for assessing program and service**

- **Description of the visitors**

- **Interesting with table breakdown: Which visitors have which evaluations (e.g. first-time visitors vs. regular visitors, age-groups)?**

- **How could image be improved?**

- **Communication (what am I highlighting in the media, for whom?)**

RED Re-Educating Digitisation

Rheinische Hochschule Köln

And how satisfied were you overall in terms of your visit today?

(1) very satisfied
(2) rather satisfied
(3) partly/partly
(4) rather dissatisfied
(5) very dissatisfied

This kind of question generates in most cases very positive results: 1 + 2 together rate up 80-90%; due to social desirability it overestimates visitor satisfaction

How likely are you to recommend a visit to our festival to a friend, family member or colleague?

(1) very likely
(2) probably
(3) neither likely nor unlikely
(4) unlikely
(5) very unlikely
(6) do not know
(7) not specified

Possible to type visitors and analyse them on different characteristics in the table breakdown:

- **Age groups, men/women or first-time/repeat visitors, industrial vs. private visitors etc.**

- **They are also suitable as a dependent characteristic to determine the influence of individual service offers or other the program (e.g. via stat. regression analyses).**

- **Open-ended satisfaction questions are psychologically important for respondents**

- **Willingness to recommend is mostly more reliable and valid than satisfaction questions: Fluctuate less and are less situationally influenced.**

The NET PERFORMANCE INDICATOR

On a scale of 10 (highly recommend) to 1 (wouldn't recommend at all), how likely are you to recommend visiting the festival? You can choose any number in between.

Explanation: The "promoters" are the active advertisers for the festival (scale points 1-2). "Passives" are neither harmful nor useful (scale points 3-4). The "critics" (also detractors) have a negative attitude and harm.

Why the "Net-Performance Indicator"?

The willingness to recommend is closely linked to visitor satisfaction, loyalty and return visits. It is also stronger than "like statements", less influenced by the situation and less susceptible to fluctuation. And it requires a stronger "public" commitment.

The NPI is an established performance measure in marketing (benchmarks possible).Calculation: NPI = Proportion of Promoters - Proportion of Critics

O_{10} O_9 O_8 O_7 O_6 O_5 O_4 O_3 O_2 O_1

Highly recommend   Wouldn't recommend at all
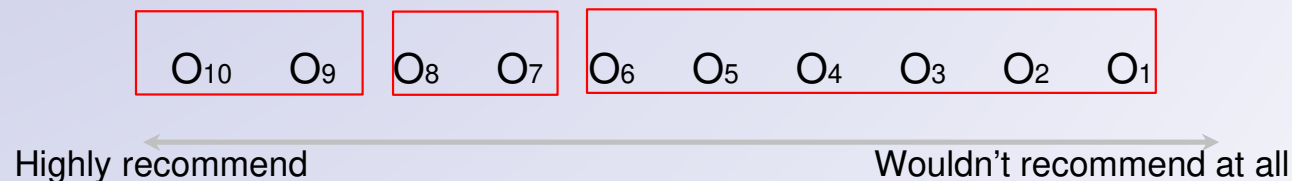
What is the main reason for this judgement? Please note!

_____

_____

Example: Evaluation of the festival by the „*Net-Performance Indikator*" (NPI)
(Shares in % by respondents)

Passive
11%

Critics
7%
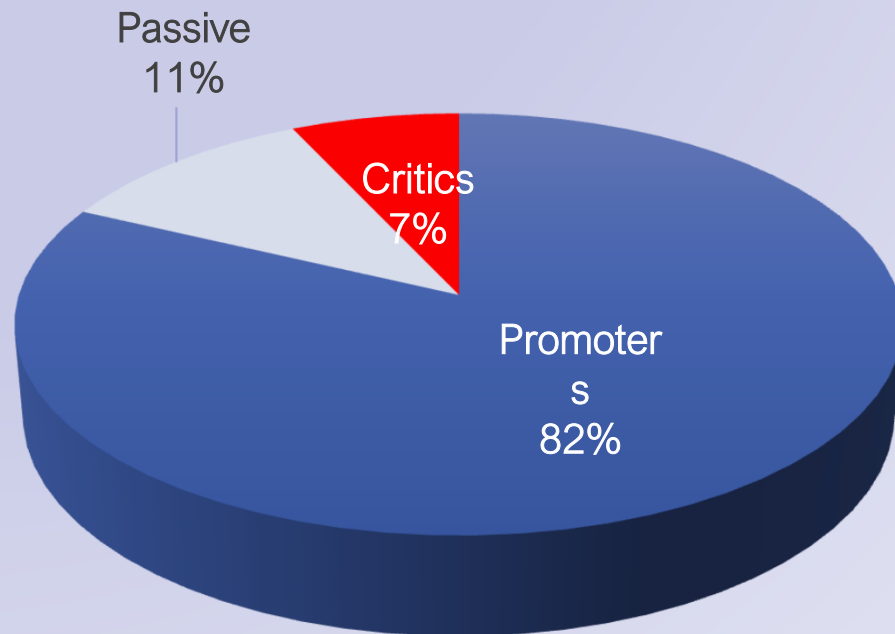
Promoter
s
82%

Here 82% of visitors are promoters, 11% are passives and only 7% are critics.

The NPI is therefore 75

NPI of the theatre by gender, age and place of residence

(NPI by target groups
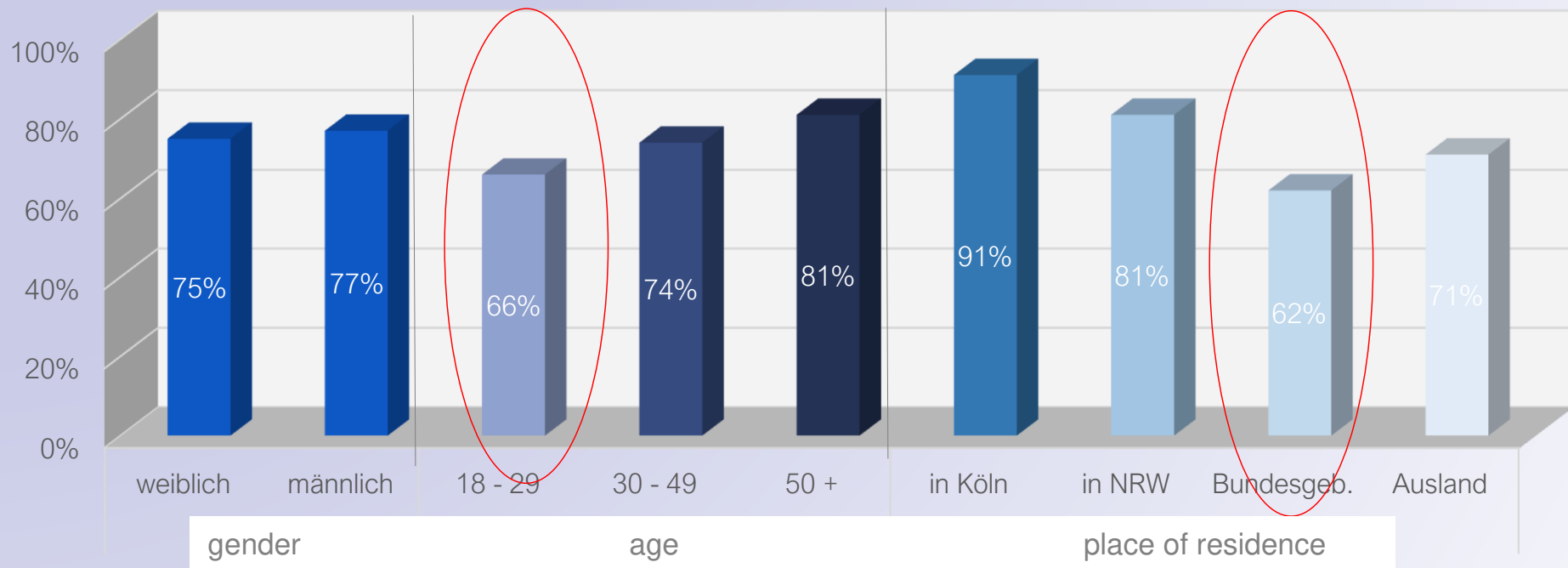
- **The willingness to recommend a visit significantly with age. The youngest age group is comparatively critical with an NPI of 66, whereas the older visitors are very satisfied with an NPI of 81. The differences between men and women are negligible.**

- **The NPI among visitors to Cologne is extremely high at 91. With increasing distance, the willingness to recommend decreases more strongly (62 in the federal territory), which is a rather poor value. In contrast, foreign visitors are again more satisfied (NPI=71). Values in the range of 70 and below normally signal the need for action!**

What of the following does apply to you? (Please one cross per line!)	meets...			
I go to the theatre to...	fully	predomin antly	less	not at all
...relax	O	O	O	O
…educate myself	O	O	O	O
...have a say about art	O	O	O	O
...get food for thought	O	O	O	O
...relax myself	O	O	O	O
... experience aesthetic pleasure	O	O	O	O
... entertain me/have fun	O	O	O	O
...do something together with others	O	O	O	O

Behind this are the following motivational concepts:

- Escapism
- Learning/education
- entertainment
- Cultural enjoyment
- Social activities
- "Status work"/being able to have a say

These concepts represent fundamental, stable and lasting motivs that are independent from certain experiences

Why are you here today? Please state the main reason:

_____

After that you can give other important reasons

(B) I was curious about what to expect here
(C) I want to learn something new
(D) The visit is useful for my hobbies or work
(F) I would like to keep my professional knowledge up to date
(G) Here I can find distance from the hectic everyday life
(I) I enjoy the environment or atmosphere here
(J) I enjoy a good time here with others
(M) One should have been here
(N) I want to experience something or have fun

Scale: fully – predominantly – less - not at all

Behind this are the following motivational concepts as mentionend above:

- Escapism
- Learning/education
- entertainment
- Cultural enjoyment
- Social activities
- "Status work"/being able to have a say

Motivation represents longlasting attidues and are the basis for the evaluation of the festival

4. Thematic area Distribution, Ticketing, Pricing

4. Thematic area Distribution, Ticketing, Pricing

- These include **ticketing,** the **long-term commitment to a** subscription and the **pricing policy of** a festival.

- **Ticketing** refers to all issues related to ticket sales. This includes the use of, preference for and satisfaction with the **theatre's own sales outlets** and **sales agents** (travel agencies, ticket agencies).

- Such information allows for an overview of the **sales network** with regard to the wishes of the visitor and helps to identify shortcomings in the area of theatre ticket sales.

- Furthermore, it is about obtaining information about the buying behaviour of the visitors, e.g. when the tickets are bought or with which means of payment the tickets are paid.

- This information helps to use available resources (e.g. staff) in a more targeted way, for example by adapting box office hours to visitor demand or by creating or modifying ticket reservation systems.

- **The pricing policy** includes questions about the current or new design of the pricing structure, about the possibilities for pricing policy, for example, "pain thresholds" of the audience in the case of (necessary) price increases. This can result in a price system differentiated according to visitor groups.

Where did you buy your ticket?
Question refers only to interviewee's ticket, not to potential escorts!

(1) I have an annual pass / have subscription (or equivalent offer of the institution)
(2) at the box office
(3) ordered/acquired by telephone
(4) ordered/acquired online
(5) ordered/acquired in writing
(6) (not ordered but) purchased at advance booking office
(7) from a visitor organisation
(8) via tour operator, was in the travel package
(9) received as a gift / won
(10) Tourist association
(11) Hotel/accommodation: e.g. Voucher Hotel
(12) Information stands of the institution
(13) have free entry/free ticket
(14) other: _____
(15) someone else bought the ticket for me
(16) do not know
(17) not specified

- **Central question on the evaluation and management of distribution channels, distribution partners or distribution locations**

- **Particularly meaningful in combination with demographic data, place of residence or the characteristic of when the booking was made**

When did you buy your ticket for today's movie/event/for the entire festival?

(1) Today
(2) Yesterday
(3) 2-3 days ago
(4) 4-7 days ago
(5) 8-14 days ago
(6) More than 14 days ago, but within the last month
(7) Longer than one month ago
(8) I cannot say
(9) I did not have to buy a ticket e.g. (professional pass).
(10) I cannot say

Which means of transport did you use to get to our cinema/festival today?

- **Analysis of the decision-making and planning behaviour of visitors**

- **Particularly meaningful in combination with demographic data and decision-making behaviour for visiting**

- **Helpful for price differentiations**

- **Helpful for the timing of communication measures**

Transport choice:

- **Relevant for the planning of services or composite offers**

- **Meaningful in combination with demographic data (age), place of residence, etc.**

RED Re-Educating Digitisation

Rheinische Hochschule Köln

Willingness to pay

Question refers only to interviewee's ticket, not to potential escorts!

How would you rate the ticket prices of the festival?
1. Appropriate
2. Somewhat high
3. Much too high
4. Too low
5. Not sure
6. Don´t know the price

If you could choose, what would be the maximum amount that you would be willing to pay for admission?

€ _____

- **Central questions for price calculation**
- **Helpfull for price differentiation for certain target groups**
- **Helpful for setting a dynamic pricing**
- **Can help to increase significantly turnover**
- **Can help to gain successfully target groups**

Some remarks on „price elasticity"

What Is Price Elasticity of Demand?

- Price elasticity of demand is a measurement of the change in the consumption of a product in relation to a change in its price.

- The prices of some goods are very <u>inelastic</u>. That is, a reduction in price does not increase demand much, and an increase in price does not hurt demand, either. For example, gasoline has little price elasticity of demand.

- If the <u>quantity demanded</u> of a product changes greatly in response to changes in its price, it is elastic. That is, the demand point for the product is stretched far from its prior point. This is often for cultural offers the case.

- As a rule of thumb, if the quantity of a product demanded or purchased changes more than the price changes, then the product is considered to be elastic (for example, the price goes up by 5%, but the demand falls by 10%).

- If the change in quantity purchased is the same as the price change (say, 10% ÷ 10% = 1), then the product is said to have unit (or unitary) price elasticity.

- Finally, if the quantity purchased changes less than the price (-5% demanded for a +10% change in price), then the product is deemed inelastic.

- Many factors determine the demand elasticity for a product, including price levels, the type of product or service, income levels, and the availability of any potential substitutes.

Example: Museum in Cologne: Price elasticity of ticket demand (price-sales function) (in %)

Visitors in ten thousand

Extremely price-sensitive area of demand.

A price increase of €1 results in an almost 50% reduction in demand from potential visitors.

$y = 0,2135x^2 - 10,591x + 123,22$

$R^2 = 0,9689$
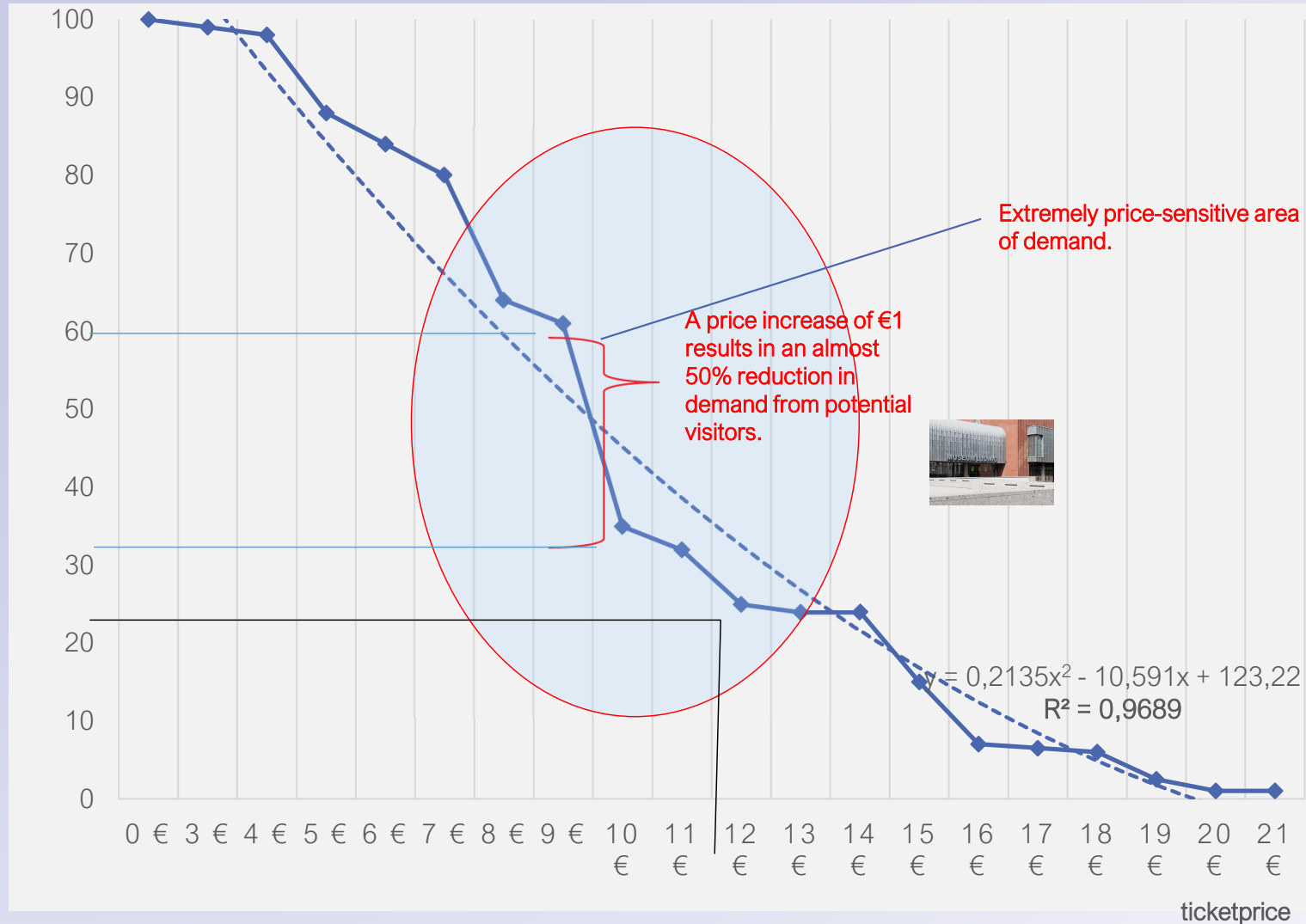
ticketprice

- With a ticket price of 12 € on average, about 250 TSD visitors were willing to pay this price.

- At the same time, it can be seen that the price sensitivity of visitors is considerable in the range of 7-12 €. Example: A price increase from 9 € to 10 € reduces the demand from approx. 620 thousand to over 320 thousand potential visitors. In other words, a price increase of 11% reduces the demand of potential visitors by almost 50%.

- The price sensitivity of visitors is very high.

Example: A Museum in Cologne: Extrapolated sales development for alternative prices (in percent of respondents)

Visitors in ten thousand

Example:
Price reduction from 12 €
auf 9 €

At a price of 9 €, up to 600 visitors could come. This results in a turnover of 5.4 million €. This makes a potential increase in turnover of 60%

250 TSD assumed vistors pay a price of 12 on average

$y = 0{,}2135x^2 - 10{,}591x + 123{,}22$
$R^2 = 0{,}9689$

With a current ticket price of € 12 on average and 250,000 visitors, this results in a revenue of around € 3 million.

ticketprice

- How does the turnover develop when the ticket price is reduced? A price reduction from the current average of €12 paid by visitors to, for example, €9, may result in an increase in turnover from approx. €3 million to approx. €5.4 million.

- The price reduction thus results in a 60% increase in turnover!

- Considerations in the direction of a price differentiation downwards seem worthwhile! In any case, a price reduction means significant visitor and turnover gains!

Example for price differentiation: Free theaters in Cologne and willingness to pay
(Basis: respondents interested in visiting a free theatre in the next 12 months

% of respondents

Additional turnover at prices of 15 €

Turnover at average price with 25 €

Price reductions attract additional visitors and increase turnover

Additional turnover at prices of approx. 40 €

Price changes remain revenue neutral

Additional turnover at prices of 70 €

Price increase also increase turnover

99 97 88 72 60 41 36 26 24 13 13 8 6 6 4 3 1 0

<= 5 | 6 - 10 | 11 - 15 | 16 - 20 | 21 - 25 | 26 - 30 | 31 - 35 | 36 - 40 | 41 - 45 | 46 - 50 | 51 - 55 | 56 - 60 | 66 - 70 | 71 - 75 | 76 - 80 | 86 - 90 | 96 - 100 | 101+

€
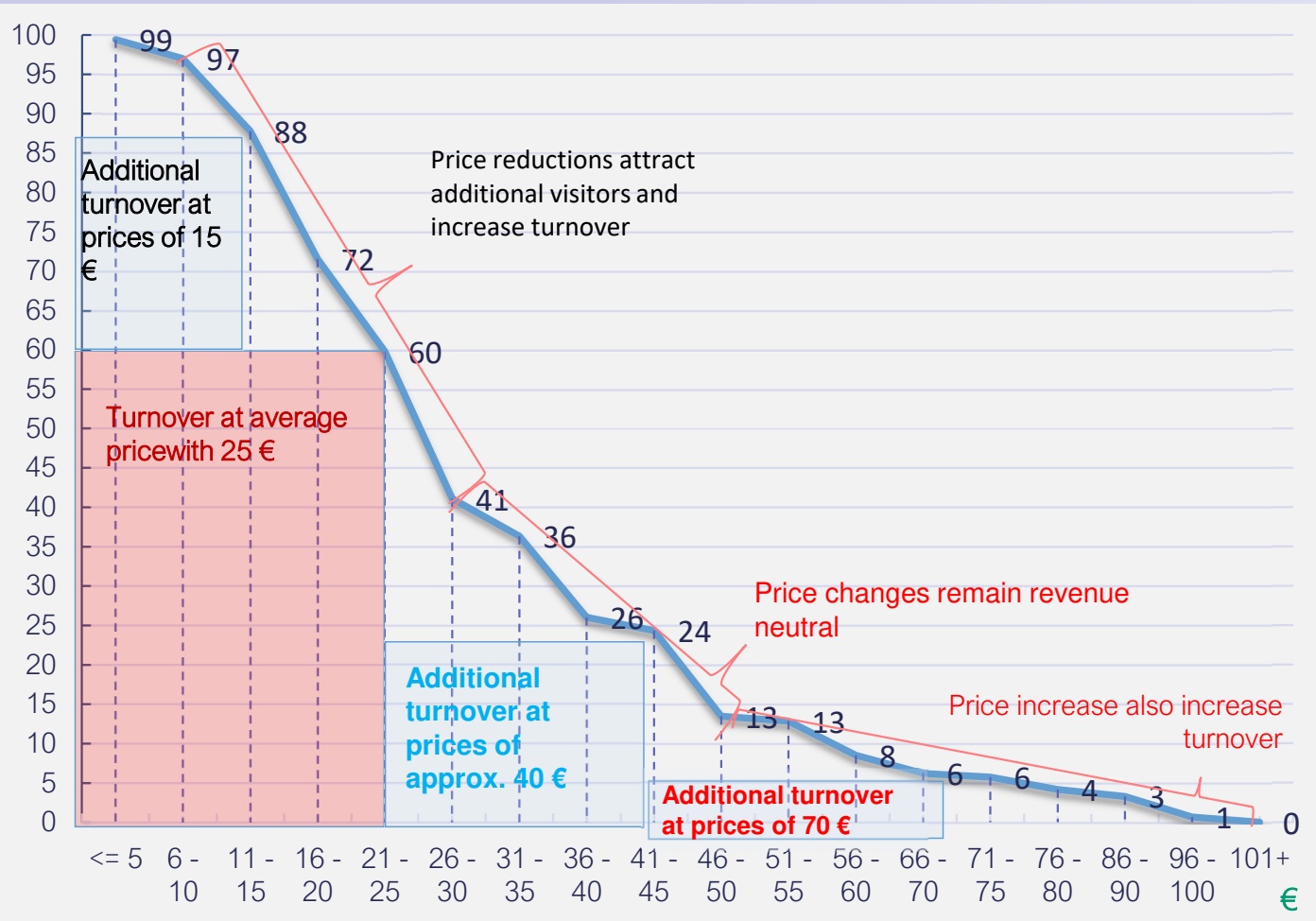
5. Thematic area „Communication"

Analysis of Communication Channels and „Sources of Interest"

Is a central question for the evaluation of the festival's communication activities: Fist clarify your communication aims!
Then:

1. Which channels are most effective and cost-efficient?

2. Which target groups use which channels?

3. By using statistical analysis you can find out: Which channels have the biggest impact on campaign awareness, sympathy, visiting the festival

4. Which channels in future should be sorted out or strengthened?

How did you find out about the movie/event you visited today?

Communication channels of the festival
(1) Flyers, information brochures, event and festival calendars, postcards, etc. (print) of the institution/festival.
(2) Posters, billboards, outdoor advertising of the institution/festival
(3) Website, blog of the institution/festival
(4) Social media channel (e.g. Facebook) of the institution/festival
(5) Messenger message (e.g. WhatsApp) of the institution/festival
(6) App of the institution/festival
(7) Newsletter of the institution/festival
(8) Mailing of the institution/festival
(9) Other communication channels of the institution/festival: _____

Tourism communication channels
(10) Local tourist information (also tourism centre, spa administration, tourist association)
(11) Tourism information online (city portal, federal state portal, German National Tourist Board)
(XX) tourist booking portals (e.g. AirbnB, Booking.com, Expedia, Trivago, GetYourGuide)
(12) Travel guide/travel magazine, book on place of interview/region
(13) Tour operator/travel agency
(14) Hotel, accommodation
(15) Special combined tourist offers (Berlin-Stars Sparbuch, Berlin Welcome Card, RAV: Echt Bodensee Card , etc.)
(16) Other tourism communication channels: _____

Further communication channels
(17) Newspaper/journal/magazine (print, online)
(18) Radio/TV
(19) other website/blog (not of the institution/festival)
(20) Social media channel (not of the institution/festival)
(21) App (not of the institution/festival, not social media)
(22) Information from another cultural institution (e.g. newsletter/e-mail)
(xx) Notes from city administration (e.g. cultural office)
(23) Other additional communication channels: _____

Other
(24) personal recommendation (by family, friends, acquaintances, colleagues, school, studies)
(25) I already know/was here before
(26) I did not see or hear any information before my visit
(27) I can no longer say
(28) Other: _____
 (29) do not know
(30) not specified15

But be aware of the following problems:

1. **The answers result from a combination of the respondent's media use and the intensity of the festival's communication efforts:**
    - **Is a perceived communication activity due to the intensive media use of the respondent, or**
    - **to the activity of the institution? Or both?**

    Conversely, is a non-perceived advertisement

    – **due to the lack of respondent´s use of the channel referred, or**
    – **is it due to too little activity by the festival? Or both?**

    -> These questions can be tested by questioning the general media usage

2. **Perception of the advertising messages also depend strongly on the interest of the respondent**

3. **Positive responses also possible due to confusion of sender/source (concerns especially small institutions where this happens frequently)**

On the problem of assessing the success of communication measures:

	Communication activity of the institution on the channel given	No communication activity of the facility on the channel
Use of the channel by the respondent	• OK • But: Problem of the attribution of the success of the measure Note underreporting due to unaided query form	• If recognised by the respondent confusion or memory distortion possible
No use of the channel by the respondent	• No effect due to wrong choice of medium or too little activity of the institution	Irrelevant

But be aware of the following questions (II):

2. **How do you find out contact points:**

   Passive recall: **Giving a list where the respondent can find the channels he uses**

   Active recall: **<u>Unaidedly</u> collects the respondent's remembered contact points with the institution:**
         -> Underestimation of the perceived communication
          communication measures by the respondent! Especially with push media!

3. **Perception of the advertising messages also depend strongly on the interest of the respondent**

4. **Positive responses also possible due to confusion of sender/source (concerns especially small institutions where this happens frequently)**

Which of the following do you usually use to inform yourself / inform yourself about cultural offers/events?

Communication channels of the festival
(1) Flyers, information brochures, event and festival calendars, postcards, etc. (print) of the institution/festival.
(2) Posters, billboards, outdoor advertising of the institution/festival
(3) Website, blog of the institution/festival
(4) Social media channel (e.g. Facebook) of the institution/festival
(5) Messenger message (e.g. WhatsApp) of the institution/festival
(6) App of the institution/festival
(7) Newsletter of the institution/festival
(8) Mailing of the institution/festival
(9) Other communication channels of the institution/festival: _____

Tourism communication channels
(10) Local tourist information (also tourism centre, spa administration, tourist association)
(11) Tourism information online (city portal, federal state portal, German National Tourist Board)
(XX) tourist booking portals (e.g. AirbnB, Booking.com, Expedia, Trivago, GetYourGuide)
(12) Travel guide/travel magazine, book on place of interview/region
(13) Tour operator/travel agency
(14) Hotel, accommodation
(15) Special combined tourist offers (Berlin-Stars Sparbuch, Berlin Welcome Card, RAV: Echt Bodensee Card , etc.)
(16) Other tourism communication channels: _____

Further communication channels
(17) Newspaper/journal/magazine (print, online)
(18) Radio/TV
(19) other website/blog (not of the institution/festival)
(20) Social media channel (not of the institution/festival)
(21) App (not of the institution/festival, not social media)
(22) Information from another cultural institution (e.g. newsletter/e-mail)
(xx) Notes from city administration (e.g. cultural office)
(23) Other additional communication channels: _____

Other
(24) personal recommendation (by family, friends, acquaintances, colleagues, school, studies)
(25) I already know/was here before
(26) I did not see or hear any information before my visit
(27) I can no longer say
(28) Other: _____
 (29) do not know
(30) not specified15

What exactly were you looking for on our website ?
Multiple answers possible.

Group 1 (helpful):
(1) Find out what is currently on offer in general
(2) Arrival information
(3) Opening hours
(4) Prices
(5) Find out more about a movie/performance
(6) Find out more about workshops
(7) Find out more about a (special) event (e.g. conferences, party)
(8) Find out more about the festival in general
(9) Search for information on memberships/friends groups/support associations
(10) Find out more about the gastronomic offer

Group 2 (works):
(11) Buy tickets
(12) Shop in the online shop
(13) Listen to/watch programme previews (e.g. embedded YouTube clips, trailers)
(14) Listen to/watch/download movies in media library (view/listen to entire works)
(15) Use (other) digital resources (e.g. collections, exhibits, archives, library photos).
(16) Contact the festival/artists
(17) Communicate my opinion/ideas to the festival

(18) Other: _____
(19) do not know
(20) not specified

Helpful for identifying:

- **Central: Evaluation of the User-XPerience**

- **Relevant content on the page**

- **Typical search behaviour of visitors and corresponding navigation requirements**

- **Compare ratings in the breakdown of different target groups**

- **Evaluation of the page possible in combination with follow-up questions on general media use (which social media, messenger, websites etc.)**

Frequency of use of information sources for cultural events in Cologne, by age (Basis: All respondents, subgroups in %)

30 - 59 years

Source	frequent	now and then
Gedr. Infos d. Veran.	18%	33%
Soziale Netzwerke	19%	13%
Homep. der Veranst.	31%	29%
Plakate/Aussenw.	27%	41%
Homep. Stadt Köln	23%	26%
Radio und TV	21%	31%
Tipps v. Freund./Bek.	46%	35%
Stadtmagazine	15%	22%
Örtl. Tagespr.	38%	20%

0% 10% 20% 30% 40% 50% 60% 70% 80% 90%

60 + years

Source	frequent	now and then
Gedr. Infos d. Veran.	30%	30%
Soziale Netzwerke	10%	7%
Homep. der Veranst.	23%	23%
Plakate/Aussenw.	26%	34%
Homep. Stadt Köln	20%	15%
Radio und TV	41%	26%
Tipps v. Freund./Bek.	42%	36%
Stadtmagazine	14%	16%
Örtl. Tagespr.	67%	15%

0% 10% 20% 30% 40% 50% 60% 70% 80% 90%

18 - 29 years

Source	frequent	now and then
Gedr. Infos d. Veran.	14%	33%
Soziale Netzwerke	43%	7%
Homep. der Veranst.	31%	25%
Plakate/Aussenw.	27%	40%
Homep. Stadt Köln	15%	19%
Radio und TV	20%	25%
Tipps v. Freund./Bek.	42%	34%
Stadtmagazine	11%	18%
Örtl. Tagespr.	27%	17%

0% 10% 20% 30% 40% 50% 60% 70% 80% 90% 100%

The comparison of the age groups reveals greater differences in the use of the media:

- **While classic media such as TV, radio, etc. play a major role among the older respondents, with the local daily press standing out, the organisers' homepages and social networks are in the foreground among the younger respondents.**
- **The urban. Internet services are used more by older people, which can be attributed to the generally greater interest in culture.**
- **Outdoor advertising, jerked information such as flyers and the like are used in all groups.**
- **Personal tips from friends and acquaintances, which are an important source of information for 80%, are consistently at the top of the list.**

6. Effects on Tourism

Only to tourists place of residence:

Are you staying overnight in… (interview location) today?

(1) Yes
(2) No, I am currently staying in (please enter location)
(3) No, I am in here without an overnight stay (staying at home).
(4) do not know
(5) not specified

How many nights in total will you spend there?

Nights_____

Is today's visit to our festival the reason for the visit?

(1) main reason for the visit here… (place of interview)
(2) also reason for the visit, but mainly other places in the city
(3) no reason for the visit
(4) do not know
(5) not specified

- **Tourism issues are an argumentation aid for festivals to use the data to economically legitimise their work**

- **May be useful to legitimise additional financial help**

- **Image related questions are also very useful for the city, tourism agencies or funding companies (may give donors, bringadditional visitors etc.**

Please mark a circle in each line!	Applies fully	Predomi- nantly	Less	Not at all	Don't know
Exciting	O	O	O	O	O
Flagship for the city of Turino	O	O	O	O	O
Essential for me	O	O	O	O	O
Important for the cultural life here	O	O	O	O	O
Innovative	O	O	O	O	O_0

End of this section!

Types of indicators

- *Definitional indicators:*
  By definition, set by the researcher

- *Correlative + inferential indicators:*
  Constitutive of the concept due to probability statements (e.g. „working motivation")

Level of experience and knowledge / observability

low
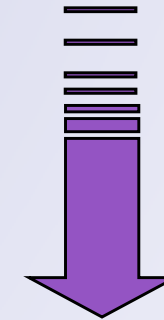
high

Overview of the research concept - from the research question to the results

Task, problem definition

For example audience description, evaluation, marketing aids, improving promotion, programme work

Formulation of objectives

Targeting new audiences, optimized communication, more turnover, better visitor loyalty…

Determination of a research design

explorative, descriptive, causal, quantitative, qualitative

Development and selection of data acquisition method

Questionnaire design, sampling, Indicator formation, questions, scaling, etc.

Development of an analysis plan

Data-procurement

Data analysis

Data interpretation, results, reports, key figures, new questions...

Now a little exercise

Make a...

dimensional analysis for the following concepts and develop indicators that may express the results in figures:

1. *Quality of Google's search results*

2. *„Customer satisfaction"*

Building Indicators

- *Indicators:* Observable measures that support a question ("Make it measurable„)

- In order to make the concepts to be analysed measurable, they must be translated into suitable parameters (indicators) (= operationalisation)

- *Operationalisation:* Instructions with the help of which certain *indicators* are assigned to a question -> develop an appropriate definition (nominal, realdef.)

- It is important to ensure that the indicators are simple and clear, plausible and based on what is necessary to measure

- You need theories, assumptions or good arguments to establish the connection between the question and the indicators.

- *Concept of multiple indicators:* Each construct should be operationalised by several indicators!

- Example: Analysis of „visitor motives", "advertising impact", „success of a library"

Are you employed?

(1) Employed (= gainfully employed, including self-employed)
(2) Apprenticeship
(3) In retraining
(4) Student/schoolchild
(7) Retired, pensioner, in early retirement
(8) Currently unemployed/unemployed
(9) Housewife/ husband
(10) Other
(11) don't know
(12) Not specified

- **Analysis of over-/underrepresented groups in your audience**
- **Reach of target groups**
- **Important for selecting communication channels and services**
- **Assessing needs, wants and motives of the audience**
- **Analysis of the communication success**
- **Most important for typing the visitors for further analysis (tabels)**

Example: The use of multiple indicators

Concept
"Advertising Impact"

Indicator

"Actual Visit"

Ticket sales

Indicator

"Campaign Awareness"

Feature: Aided recall

Indicator

"Campaign Sympathy"

Characteristic: Question with Likert scale

Indicator

"Intention to visit"

Characteristic: Formulated intention to visit in the questionnaire
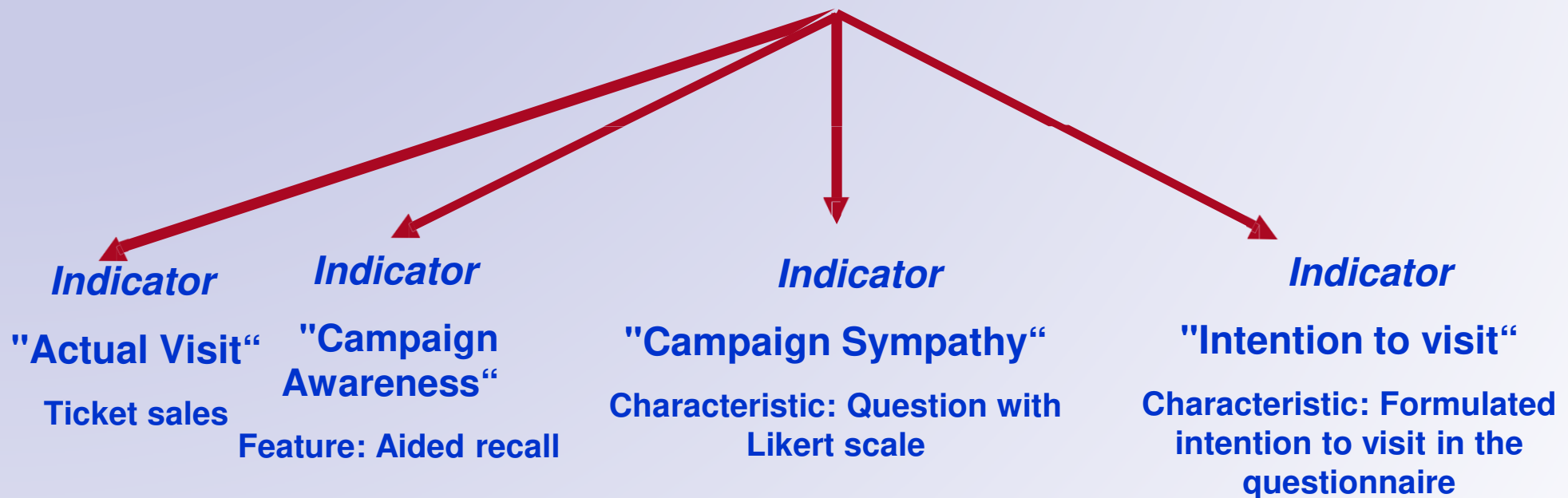
Different dimensions of the same concept

Exercise II

Your festival employed influencers on different social media to adress a younger target group.

In a research you want to answer the question, wether this had an impact resp. was a success.

Please develop indicators and a research design. Be sure not to mix up the effects from other media you might have also used!

Exercise II

1. What does 'soc. media', "influencers", advertising impact' mean? (Definition of terms ('concept specification', indicators, operationalisation)
   - Number of followers
   - Mainly professional
   - Regularly prepared content for followers, regular contacts, expertise, company cooperation ...

2. What exactly could the impact look like? (hypothesis formation)
   - Reason for buying the product, is "cool", identification figure, agreement in requirements, trust, use

3. How can we ensure we capture the influence (accurately, and only) of influencers? (Research design)
   Where first contact, customer journey, ...

4. How do collect the data? E.g. by questionnaire [in person or by phone] or even online survey? (Survey method)

5. From which persons (units of investigation) should data be collected (sample)?

6. How is the data evaluated?

Example: Target dimensions and quantitative indicators for measuring *"success"* of public libraries
(cf. Library Index BIX of the German Library Association)

Target dimensions:

- **Offers: Indicators:** Number of media per inhabitant, audience area, staff, computer workstations, internet services, events

- **Use: Indicators:** Library visits per inhabitant, virtual visits, borrowing, lending/medium, opening hours

- **Efficiency: Indicators:** Media budget per loan, staff hours and visits per opening hour, expenditure per visit

- **Development:** Indicators: renewal rate, training rate, investments per inhabitant

RED
Visitor Research II

- Introduction -

Author:
Prof. Dr. Tibor Kliment

Topics of the visitor research workshop II

Section 1

I. **Introduction**
   - Topics, welcome of new participants

II. **Theme packages in visitor studies II**
   - Demographics, visitor/visit information, ticketing/distribution, Communication

III. **Introduction into data analysis with SPSS and statistical basics**
   - User surface of SPSS, data management
   - Introduction into basic statistical analyses
   - Building tables and interpreting results

Section 2

III. **Introduction data analyses with SPSS and statistical basics**

IV. **Data analyses of the See You Sound Festival visitor study**
   - Preparing the data mask, data entry of questionaires
   - Data checking, cleaning and modification
   - Doing first data analyses

Section 3

V. **Continuing Data analyses of the See You Sound Festival visitor study**
   - Hands on data analyses and data presentation: Frequencies, Means, dispersion, crosstabs

VI. **Introduction into other types of interviewing**

What is the outcome for you:

After experiencing this workshop:

- **You know how to develop professional questionaires**

- **You know about specialized data analyses programs**

- **You can do basic statistical analyses**

- **You know how to interpret and present the results**

- **You know how to implement the results into your festival business…**

After experiencing all workshops:

- **You are able to plan and manage the entire research process and you know how to choose the right research design**

- **You know how to combine marketing concepts (in the area of pricing, distribution, communication, product development etc.) with visitor research**

- **Al together: Your audience development strategy is lifted onto another level, you have more success with your festival and your life will be happier**

RED
Visitor Research II

- Introduction into Statistics and SPSS -

Author:
Prof. Dr. Tibor Kliment

RED
Re-Educating
Digitisation

Rheinische
Hochschule
Köln

Overview: What is in this set of slides?

➢ What is SPSS? General introduction about what you can do with SPSS and why SPSS is important

➢ Introduction into the user surface of SPSS

➢ How to get it in: Introduction into data entry module

➢ How to get it right: Introduction into data checking and transforming

➢ What is behind? Introduction into some basic statistical terms

➢ When it gets exciting: Doing the first steps in data analyses

What is SPSS?

- SPSS stand for "Statistical Product and Service Solutions".

- It is a proprietary softwarte for data analysis and statistical modelling

- It was developed in 1968 and has been operated by IBM since 2009

- (One of) The most popular statistical package(s), which can perform highly complex data, data transformation and analysis with simple instructions

- Current version is V.29. Our version used here is V.16. However it is more than sufficient for all our purposes and beyond

Why is SPSS important?

- SPSS enables data to be collected, organised and interpreted

- It offers a simple and intuitive user interface that allows users to perform complex statistical analyses without having in-depth statistical knowledge

- SPSS is an industry standard in many fields, including social sciences, health sciences, business and especially in market research

What can you do with SPSS?

- Enter, organise, verify and clean data

- Describe many statistical figures (mean, median, deviation etc.)

- Do crosstabs that are ready for presentations

- Perform statistical tests, including t-test, ANOVA

- Perform causal analyses, like regression analyses, correlation and more

- Create numerous graphical representations, including bar charts, line graphs and histograms, 3-dim diagramms

- Interpret and report results

Overview of the SPSS interface

- **Data Editor:** Here data can be entered and organised

- **Ouput Viewer:** Here the results of statistical analyses and graphic are displayed

- **Syntax Editor:** Here you can write and execute SPSS code

- **Menu bar:** Here you can access the various functions of SPSS

- **Toolbar:** Here you can quickly call frequently used functions

Basic Structure of data entry

- The default window has the data editor

- There are two different windows:

  - Data View

  - Variable View
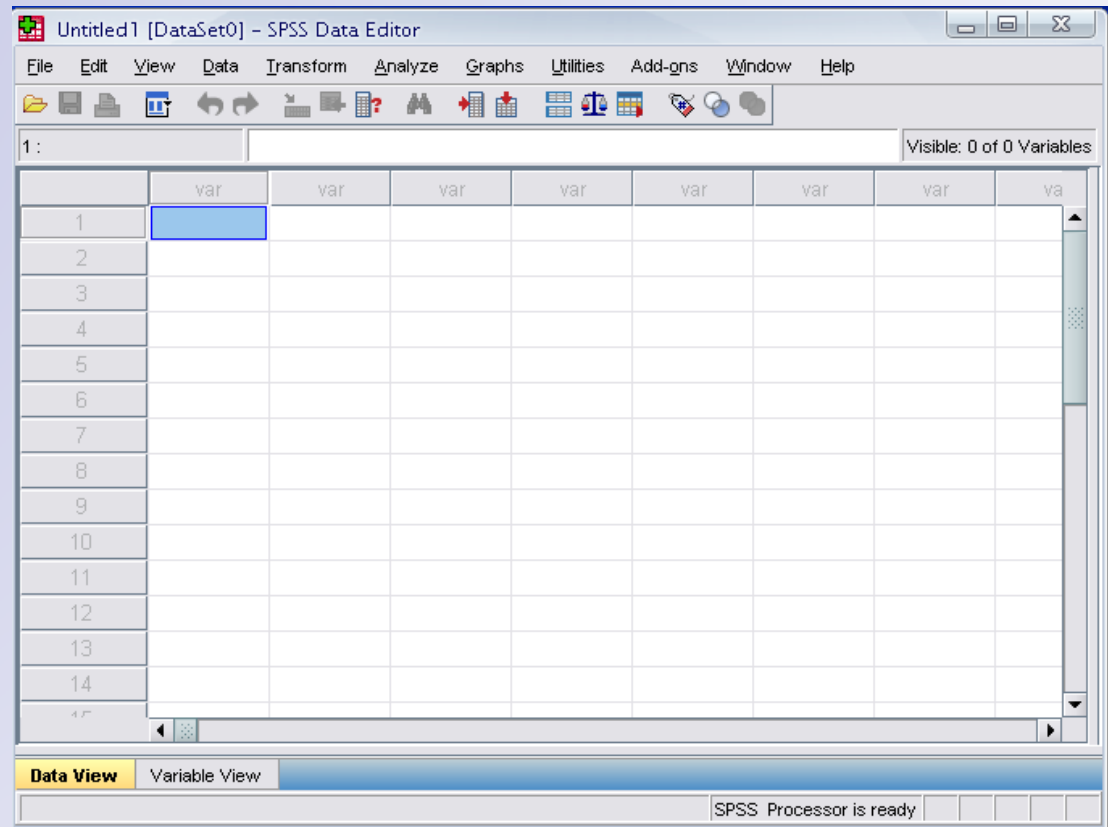
1. Data view window
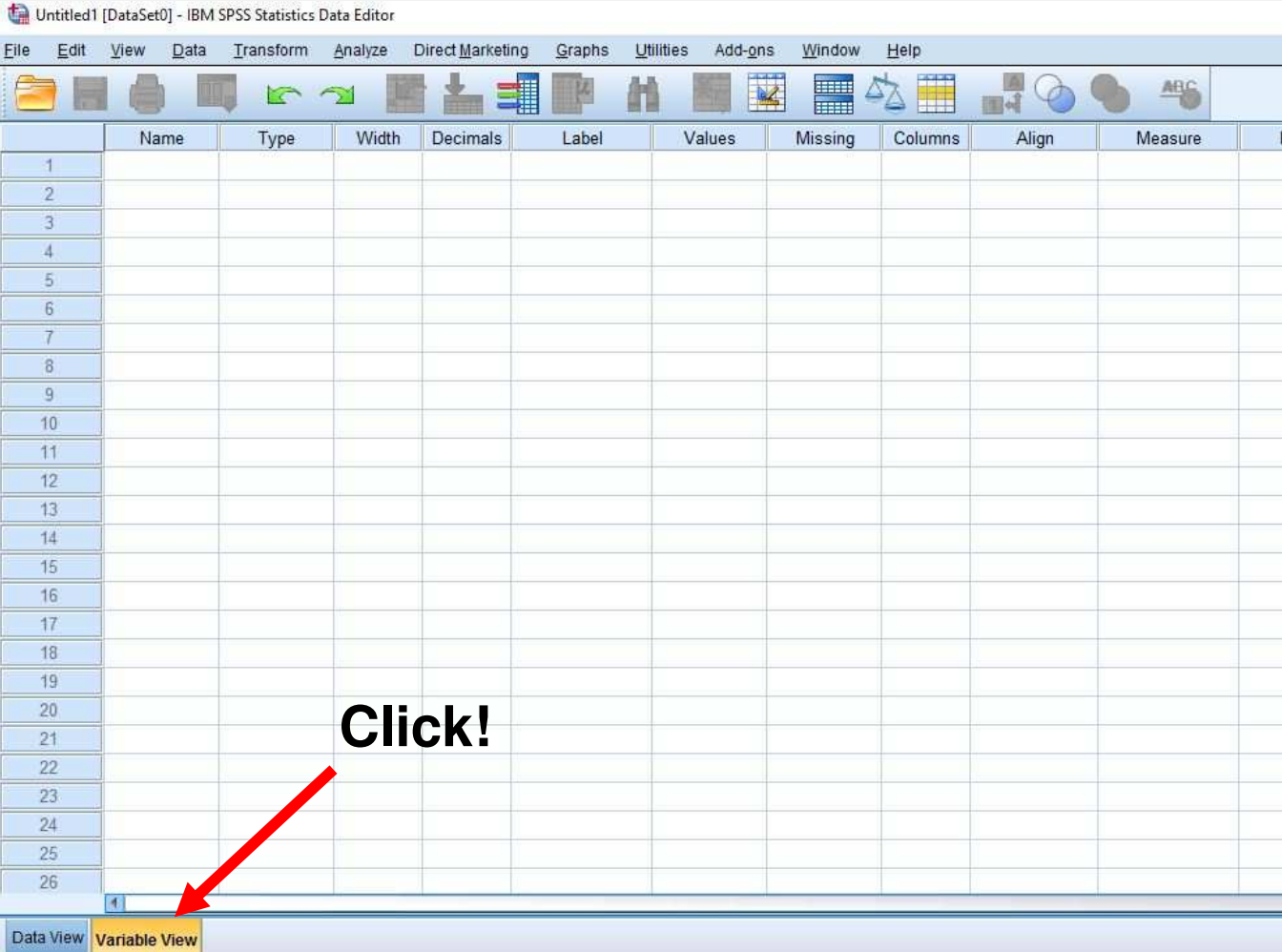
- First sheet when you open the data editor
- This page contains the data table, structured into cases and variables
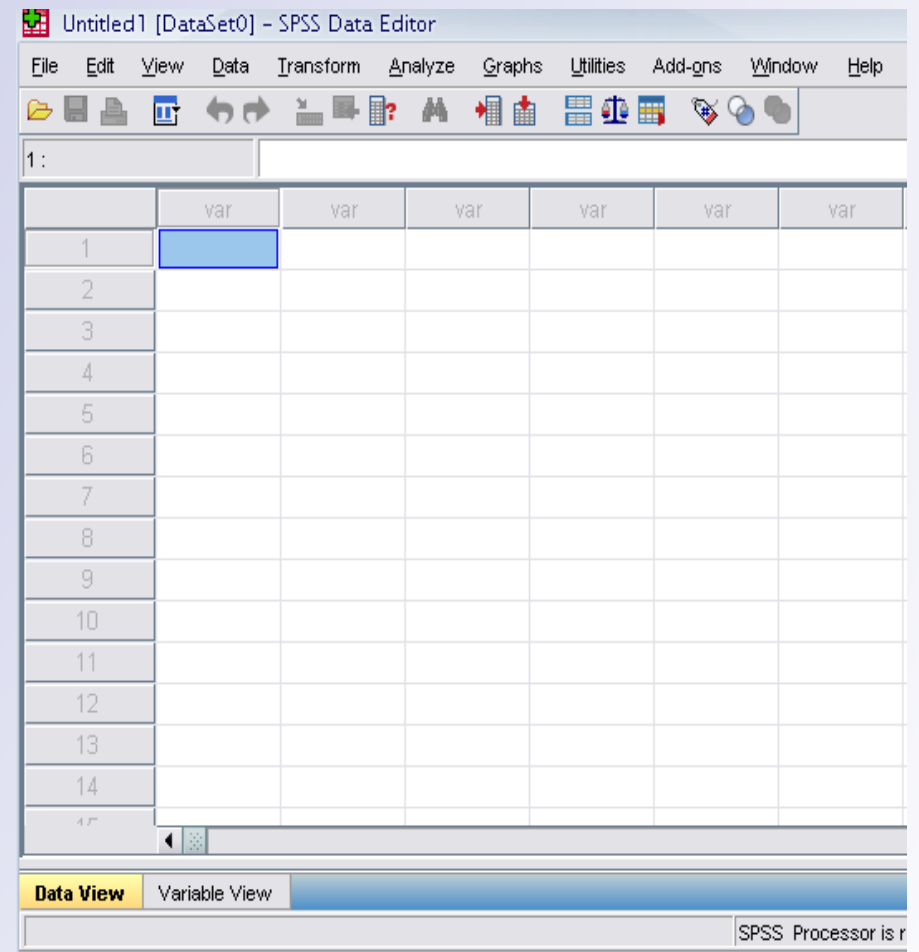
2. Variable view

- Contains information about the data, that is stored with the data set

  - Name

  - Type

  - Width

  - Decimals

  - Label

  - Values

  - Missing

Click!

2. Variable view

Variables: Something that contains different expression/values

- **Values are the properties of the statistical units that are of interest**
- **As a rule, variable is a characteristic of a unit, not the unit itself.**
- **We also call the statistical units a characteristic carrier.**

Haarfarbe: Blond
Augen: Blau
Größe: 187 cm
Gewicht: 87 kg
Bearer of traits
Schuhgröße: 43
Geschlecht: männlich
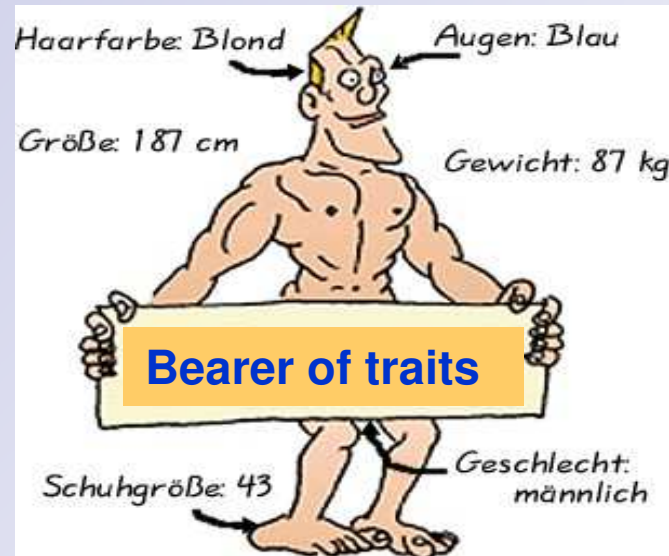
- **Variables are assigned to each statistical unit, they have different values**

2. Variable view

Variables and value expressions / values

Characteristic values are all possible values that a variable can assume.

Characteristic values X $x_1, x_2, \ldots, x_k$

Ex: Religios affiliation may be Catholic, Protestant, ...

When a varible has only one value, it is no variable any longer for SPSS, and it should be removed from the data file

In SPSS there are variables, variables names, values and value labels

RED
Re-Educating
Digitisation

Rheinische
Hochschule
Köln

Variable view – variable name

- Each variable must have a unique name of not more than 8 characters

- It has to start with a letter, usually V…

- Try to give shart meaningful variable names; usually adopt the specifications from questionaire

- Describing the characteristics, for example – Age

Variable view - Type

Click on the 'Type' box

Internal formats:

- Numeric

- String (alphanumeric)

- Date (for time series)

Output formats:

- Comma

- Dot

- Scientific notation

- Dollar

- Custom currency

- Date

Variable view - Width

- Allows you to determine the numer of characters SPSS will allow to be entered for the variable
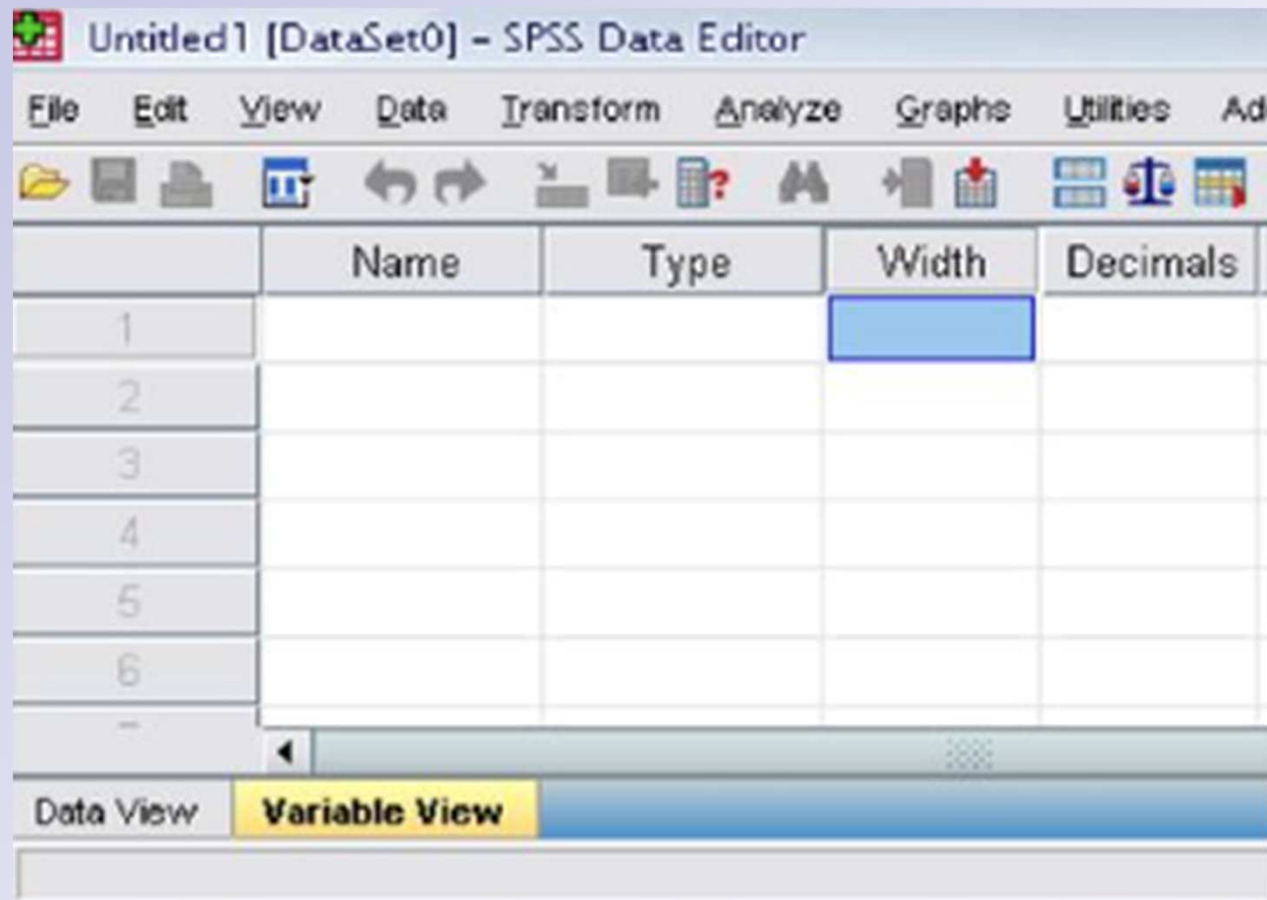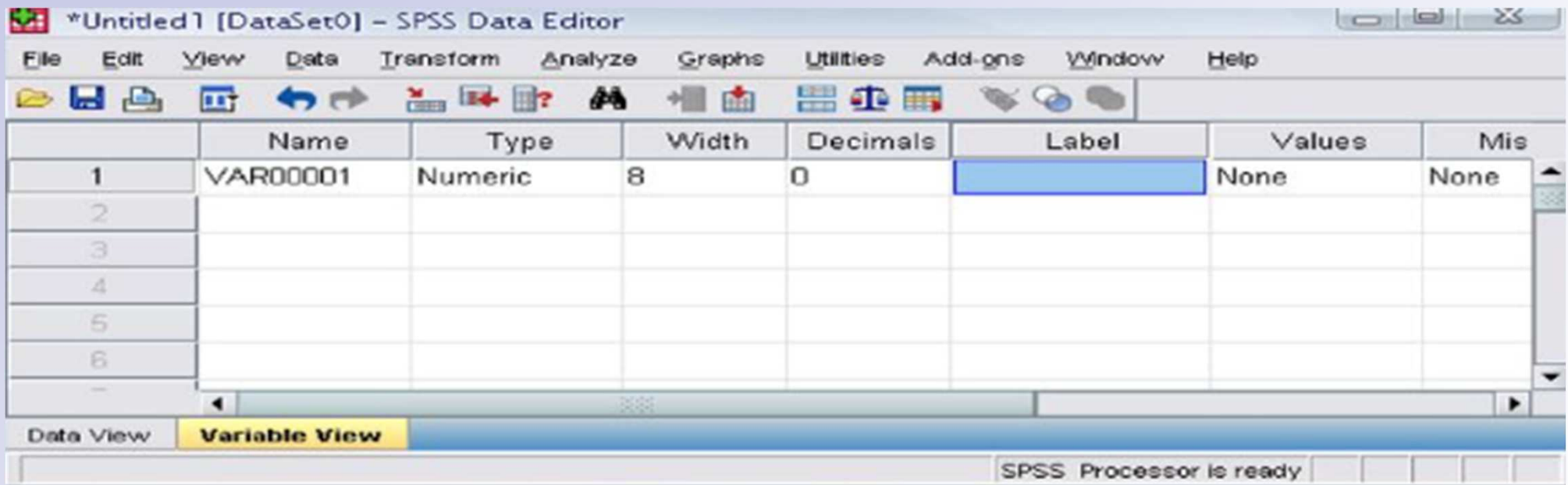
Variable view - Decimals

- Numbers of decimals

- Has to be less than 16 or equal to 16

- Usually try to avoid decimals
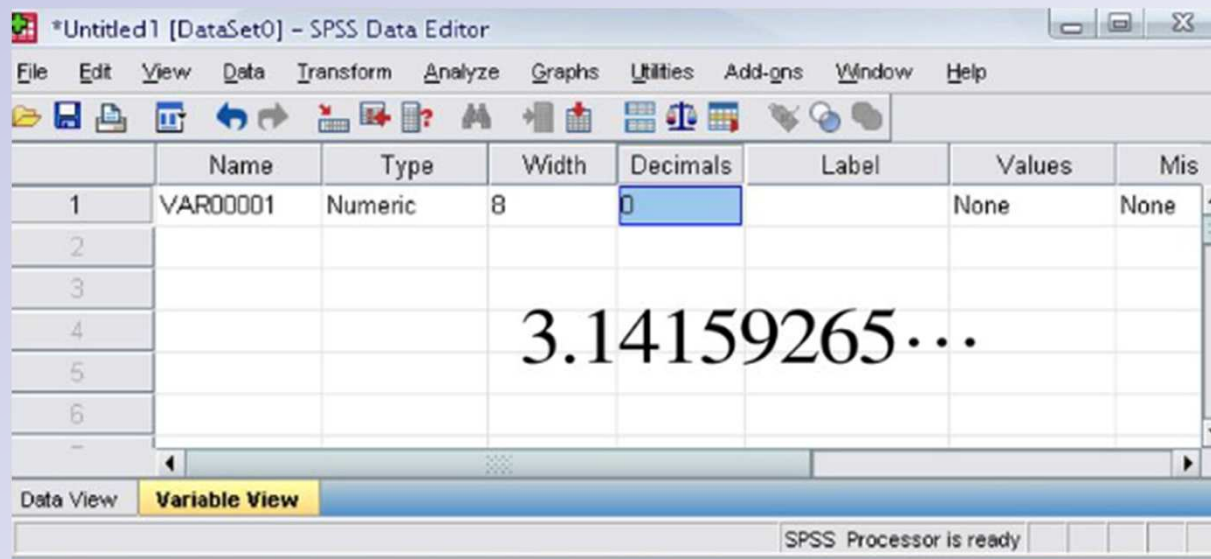
Variable view - Label

- Descriptors for the variables, Maximum 255 characters
- Variable name describes the variable in more detail; it can be long as you like and contain any signs, but try to keep it as short as possible for practical reasons
- In the output analyses only the variable label appears

Variable view - Values

- Is used and to suggest which numbers represent which categories when the variable represents a category
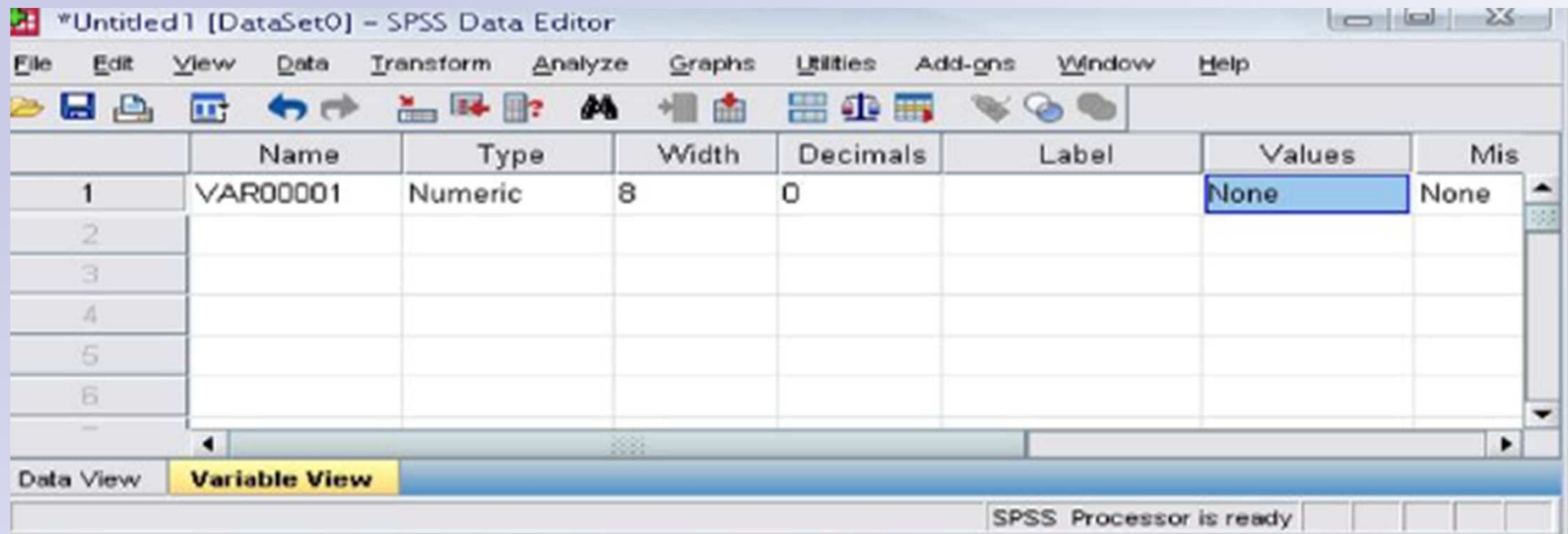
Variable view – defining the Value Labels

- Click the cell in the value column as shown below

- You can put up to 60 characters for the value and the label

- After defining the values click 'Add' and then 'Ok'
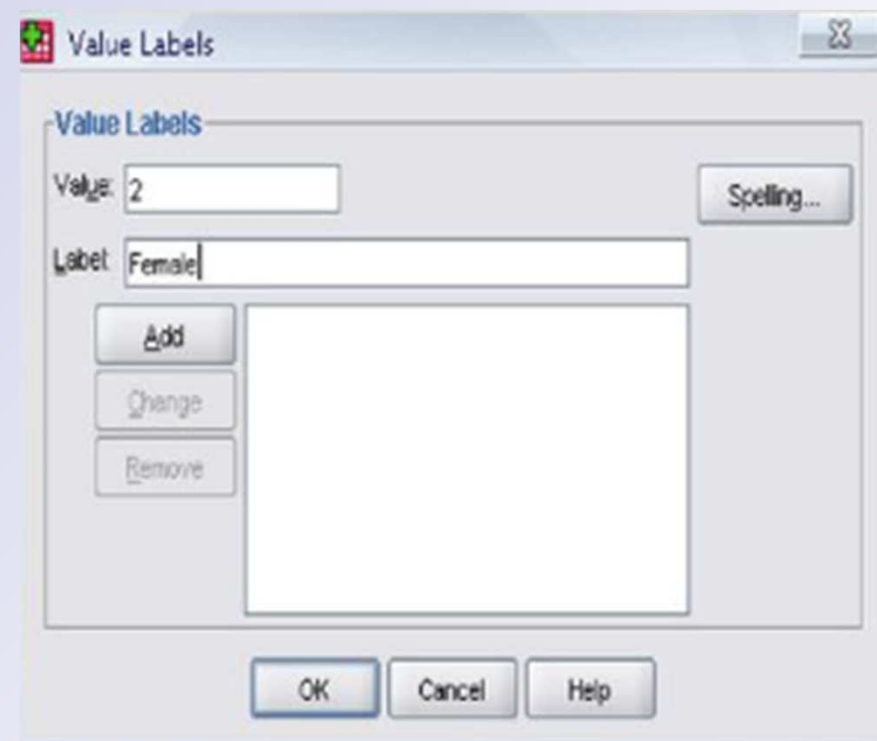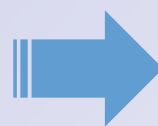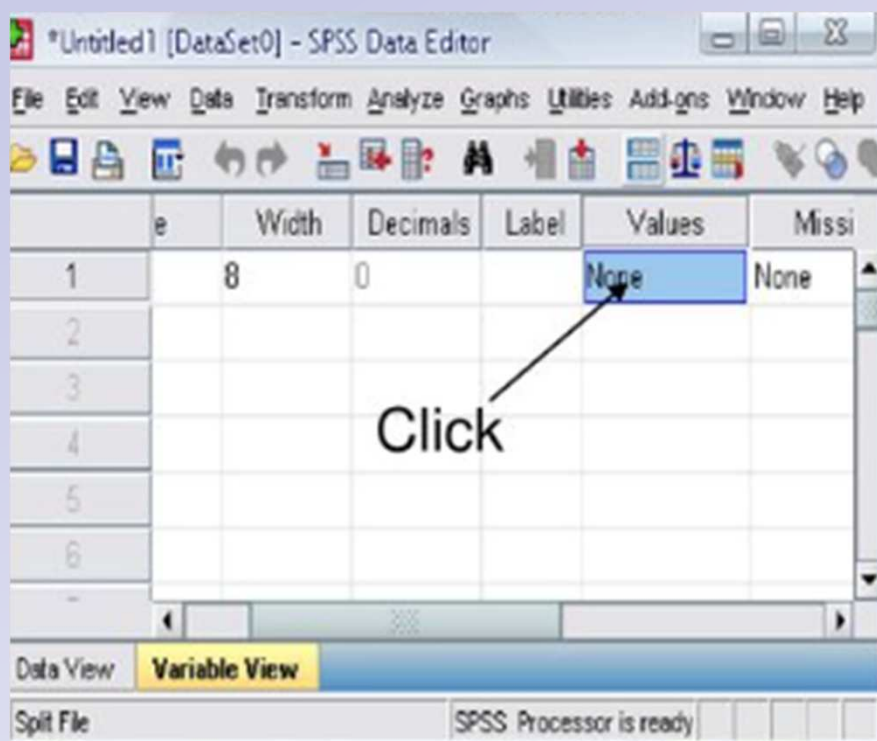
Variable view – Missing

- Defines missing values, the values are excluded from some analysis

- Different types of missings:

  - does not apply,

  - filter questions,

  - no response,

  - can´t specify, do not know

- Options: Up to three discrete missing values possible, SPSS automatically sets system missings (sysmis) with a dot

- A range of missing values plus one discrete missing value
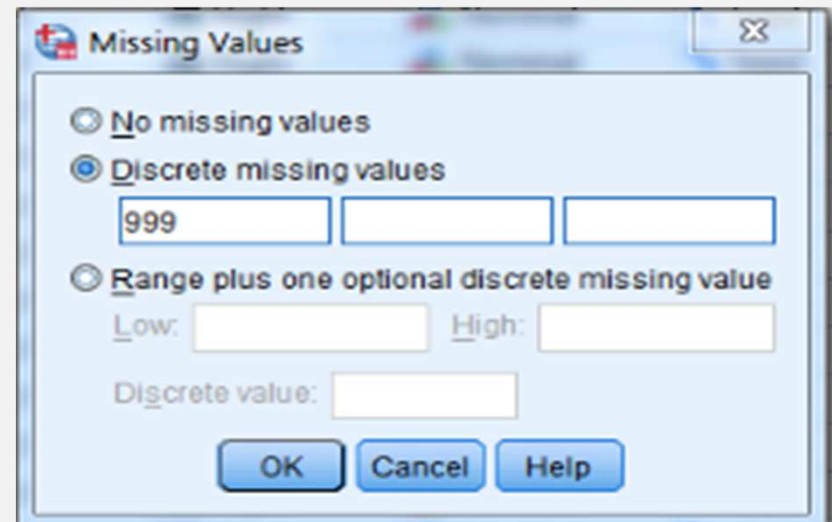
Variable view – Measure

- Defines the type of variable

- Very important (!), because certain procedures and types of analyses are dependent on what ist specified here

- SPSS automatically assigns a certain type of variable, but it is sometimes mistaken

- You always have to check for the right variable assignment

- There are three different types of variables possible for assignment: nominal, ordinal and scale

Measure: Nominal scaled variables

- A variable can be treated as **nominal** when ist values represent categories with no intrinsic ranking.

- The classification of units of enquiry is done with regard to their *possession* or *non-possession of* a certain characteristic expression.

- The individual characteristic expressions cannot be differentiated *according to rank* - i.e. they cannot be placed in *an order*, they are merely *designations of categories*.

- The nominal scale represents the *lowest* level of measurement.

- Example of nominal variables include for instance region, zip code or gender of individual or religous affilation.

- The nominal scale should always be coded by the researcher in order to ease the analysis process, for example; 1 = Male, 2 = Female

Measure: Ordinal scaled variables

- A variable can be treated as **ordinal** when its values represent categories with some intrinsic ranking;*"greater/smaller"* statements can be made between the characteristic values

- I.e. the individual characteristic values can be ranked *in* order.

- Examples of ordinal variables include a degree of satisfaction among the consumers, preference from very high to very low, and degree concern towards a certain issue.

- Assign numeric codes to represent the degree of something among respondents. For example 1=Highly satisfied, 2=satisfied, 3=neutral, 4=dissatisfied, 5=highly dissatisfied

Measure: Metric or intervall scale

- A variable can be treated as scale when its **values represent ordered categories with a meaningful metric**, so **that distance comparisons between values are appropriate**

- Examples of scale variables include age in years, income in thousands of dollars, money spent on a visit or number of visitors per festival

  - **Interval scaled variables:** Here, the expressions can not only be ranked, but the *exact distances* between the expressions can be specified. In addition, the distances are *always the same.*

  - Ex.: *time calculation, temperature in Celsius or in Fahrenheit, intelligence measurement*

  - **Ratioscaled variables:** Represents the highest level of measurement! Here there is an *absolute* (natural) zero point in the range of values, therefore: Statements about relations are permissible!

  - E.g.: *temperature in Kelvin, number of visits, age, income, money spent on tickets*

- Both measurement levels are on **metric measurement** level, since metric measurement level allows arithmetic arithmetic operations to be performed.

- In addition: **Variables at ordinal scale level are often treated as (quasi-)metric in *practice* so that arithmetic operations, such as averaging, can be carried out**. However, this is not permissible in principle.

Measure: Metric or intervall scale

Scale level	Equality/ Diversity;	meaningful relations and operations		
		bigger/smaller; Order, rank	Differences/ Distances	Ratios (Ratio-form numbers)
nominal	yes	no	no	no
ordinal	yes	yes	no	no
interval	yes	yes	yes	no
ratio	yes	yes	yes	yes

Measures: Examples

Little group work: Assign the right scale level!

1. Gender

2. Place of living

3. Income

4. Number of visits

6. Product price

7. No. of a questionnaire

8. Proportion of women in a festival

9. Daily mean temperature

10. BA predicate

11. Formal education

12. Visitor satisfaction

13. Professional status

14. Age / age classified

15. Net performance indicator

16. Postcode

Exercise: Creating a data file and assigning the proper data descriptions from the SYS - Questionaire

Output viewer

- When you perform any command in SPSS, the Output Viewer window opens automatically and displays a log of the actions taken and the associated output

- Is where the results of statistical analysis are shown,

- Also any command executed through the drop-down menus or syntax will be printed to the Output Viewer (includes opening, closing, or saving a data file)

- If an Output Viewer window is not open when a command is run, a new Ouput Viewer window will automatically be created

2. Output Viewer

- The Ouput Viewer window is divided into two sections/frames
- **Left frame** contains an outline of the content in the Output Viewer
- This Outline is useful when you have run many SPSS commands and need to locate a particular section of output easily
- The **right frame** contains the actual output
- Clicking on an item in the left frame you will jump to that content in the right frame
- Items that have been selected in the right frame are indicated by a red arrow and a box drawn around the content.
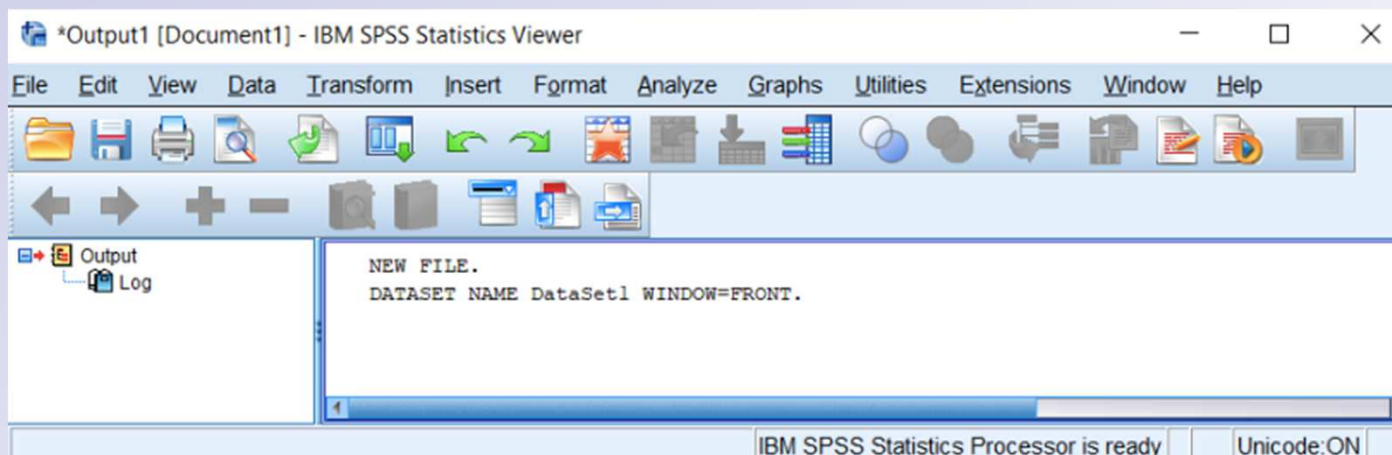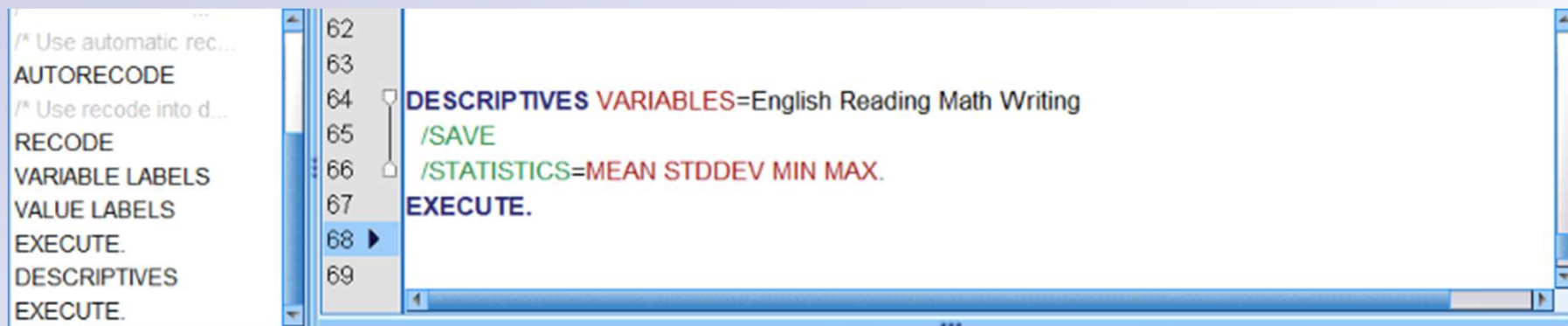
3. Syntax Editor

- SPSS syntax is a programming language unique to SPSS
- It can be used as an alternative to the drop-down menus for data manipulation and statistical analyses
- Users can write, debug and execute SPSS syntax
- To open a new Syntax Editor window, **click "File → New → Syntax"**
- The syntax editor offers much more options than the menues do
- Very usefull for executing repeated comand or jobs
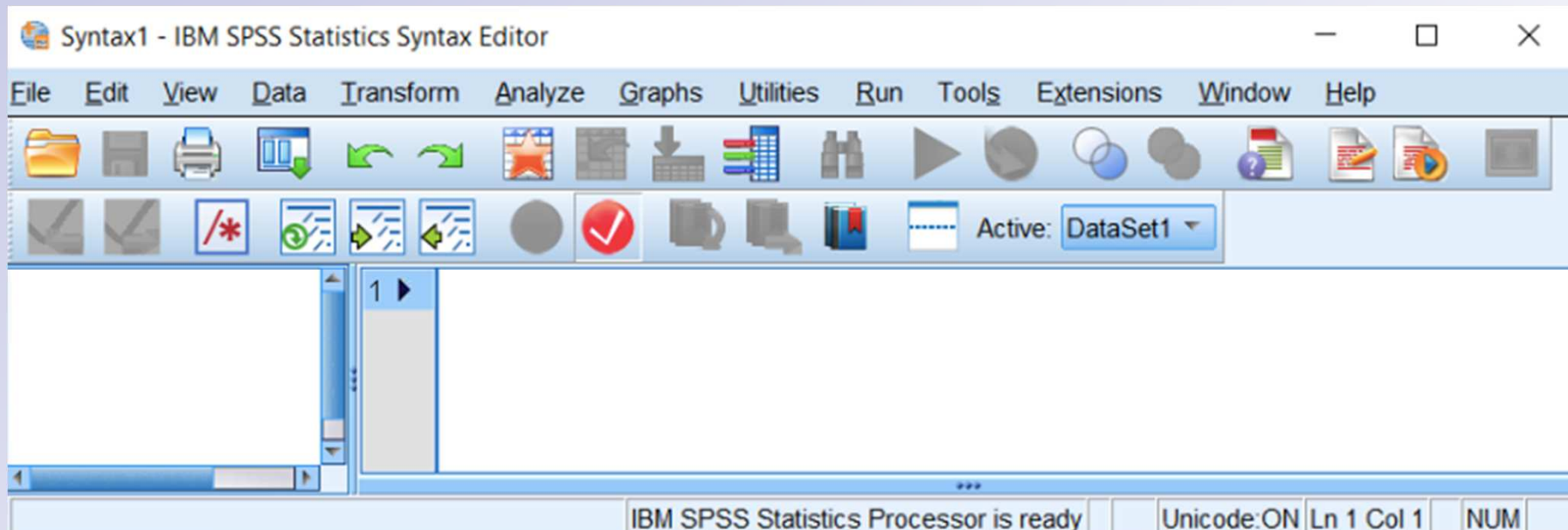
Syntax Editor

- The right panel of the Syntax Editor window is where your syntax is entered.
- The left panel of the Syntax Editor window shows an outline of the commands in your Syntax, and can be used to navigate within your code.
- You can jump to a specifix part of your code by clicking on the command in the left panel
- This feature is useful for showing the start and end points of a command, escpecially if the command is longer than one line
- Can be saved as an *.sps file by clicking "File → Save" or "File → Save as" within the Syntax Editor window

Save the data

- All data in the different windows can be saved in specific files with specific endings; for this you have to be in the referred window

- Most important is the **data file**, that always ends with a *.sys

- Never forget to save, otherwise you loose the work done; SPSS asks always for saving the files when closing the programm

- Having made modifications in the data, syntax or output, it is very useful to save the data under different file names (an easy way is to use the date of the day)

Summary

- SPSS is a proprietary software for data analysis and statistical modelling
- It is important because it provides a simple and intuitive user interface that allows users to perform complex statistical analyses
- SPSS lets you enter, organise, analyse and interpret data, including statistical tests, graphing and more
- The SPSS interface consists of a Data Editor andVariable View: Displays variable definition information, including defined variables and value labels, data types etc.
- Important is the output viewer, that presensts results of analyses and referred commands
- SPSS is costly, but there are „used" license codes legally available in the internet
- ➢ We will later switch to a free open source programm (PSPP), which is however not as powerful as SPSS but provides a similar surface

Co-funded by
the European Union

RED
Re-Educating
Digitisation

Rheinische
Hochschule
Köln

Data checking and modifying

- Before we can work with the data, they have to be checked before

- Data failure is due to different reasons

- There are several procedures available in SPSS for data checks:

  - Validation

  - Identify duplicate cases/values

  - Identify unusual cases/values

  - Frequency command (!): Make a frequency command of all variables and look for unsusual, missing or impossible values

  - Remove those cases/values from the data set

Exercise: Find duplicate cases in Cologne data file

Recoding variables

Recoding can be used in different situations:

- To categorize a variable (for example, you may want to use people's age to create different age groups, to merge categories, to change string into numerical data)

- Want to change the order of values when displaying data or doing analyses with two or more variables

In SPSS, there are three basic options for recoding variables:

- Recode into **Different Variables**

- Recode into **Same Variables**

- **DO IF Syntax**

Each of these options allows you to re-categorize an existing variable. Recode into different Variables and DO IF Syntac create a new variable without modifiying the original variable, while Recode into same Variable will permantenly overwrite the original variable.

In general, it is best to recode a variable always into a different variable so that you never alter the original data and can easily access the original data if you need to make different changes later on.

Recoding variables – Recode into *Different Variable*

1. From the menu, choose '**Transform**' → 'R**ecode into Different Variables**'. The 'Recode into different variables' dialog box will appear.

2. Select the variable you want to recode

3. In the '**Output Variable**' area, enter the name for the new variable and click '**Change**'

4. Click '**Old and New Value**' to specify how to recode values.

5. Specifiy an old value and a new value. Click '**Add**' to place the specification into the **Old → New** list. In this example, the age variable is recoded into four age groups (below 20,21 to 40, 41 to 60,61 and older.)

6. Click '**Continue**' and return to the previous dialog box.

7. Click '**OK**'.

Recording Variables
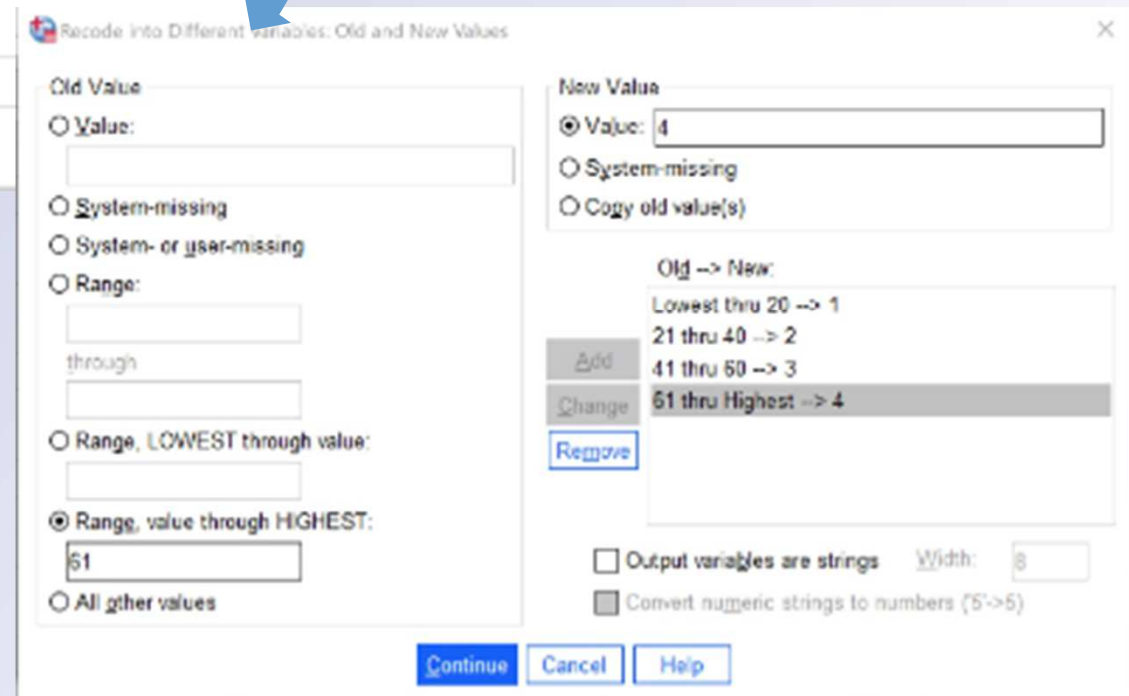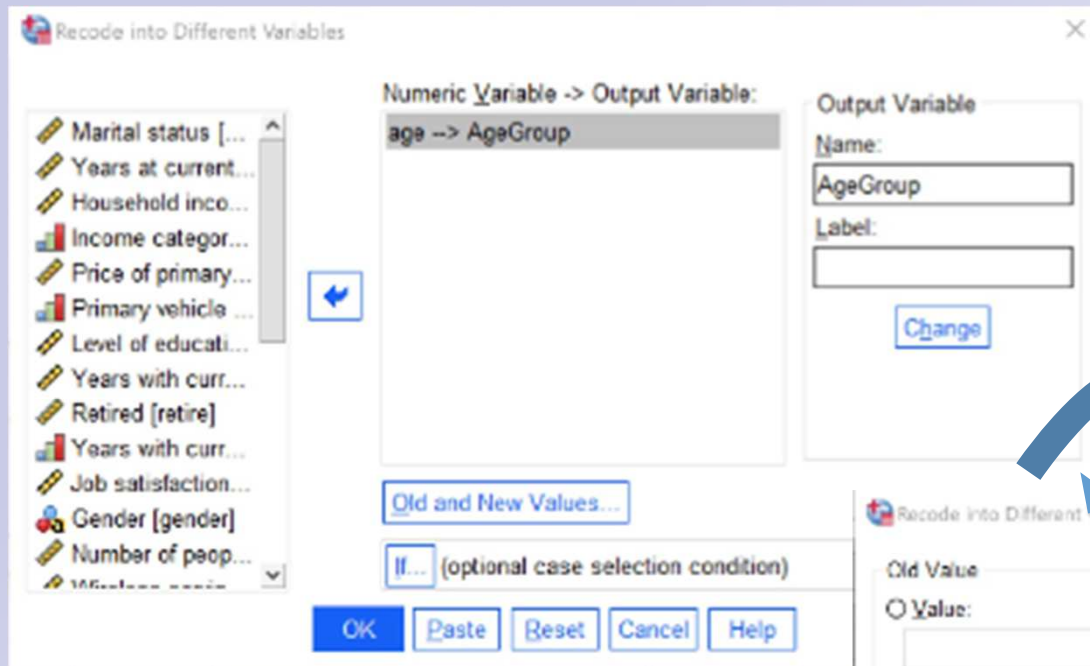
Recode Variables – Recode into *Same Variable*

- Works the same way as 'Recode into different variables' (transform → Recode intro Same Variables), except for that any changes made will permanently alter the original variable.

- The original values will be replaced by the recoded values

- In general, it is good practice not to recode into the same variable because it overwrite the original variable.

- If you ever need to use the variable in its original form (or wanted to check your steps), that information would be lost

Recode Variables – DO IF Syntax

- Performs similarly to the recode procedures, but allows for more control over specifying numeric ranges.

- If you want to perform an recode under certain conditions, or if you want to perform a recoding based on more than one variable, you'll need to use 'DO IF-ELSE IF Syntax'.

- This is given, if you want to recode a variable under certain conditions

Recode Variables to build categories – Some general rules

How many categories are formed or how category boundaries are defined? There are many suggestions in the literature that are of formal or qualitative in nature.

- If you have a lot of data, you can create more categories than if you have a small data set. Rules of thumb for the number of categories to be formed:

- **k = $\sqrt{n}$ for n ≤ 100**
- A more precise formula is: **k = 1 + 3.32 x log n** (Sturges formula) where log is the logarithm to the base 10.
- General: k = 10 for n ≈ 100, k = 13 for n ≈ 1,000 and k = 16 for n ≈ 10,000.

Our example: n = 130. Then k = 1 + 3.32 - log 130 = 1 + 3.32 x 2.11 = 8.018 ≈ 8.

$a = b^x \Leftrightarrow x = \log_b$ Example: 3 is the logarithm of 8 to the base 2; $\log_2 8 = 3$

The logarithm indicates the exponent by which one must exponentiate the base b to obtain the numerus a.

Recode Variables – Some general rules

- Often the number of categories is selected in order to have catchy and **meaningful category limits.**

- In an example with net income, the category limits follow each other in steps of 500; we then have 9 categories instead of 8.

- One should **avoid cutting up areas** in which characteristic values appear in clusters by a class boundary or in which the majority of cases is in

- Do not let them appear in clusters at the **edge of a class**.

- The categories must be clearly defined and **mutually exclusive** (disjunctive), i.e. each measured value can only be put into one category.
  No duplicate values for limits (e.g. 0 - under 500 €, 500 to under 1,000 €, 1,000 to under 2,000 etc.)

Recode Variables – Some general rules

- The categories must be **exhaustive**, i.e. each measured value must be assignable to a single category.

- One should form **intervals of equal width**, with round numbers as interval centres or with round numbers as interval limits (e.g. income 1,500 to 2,000 €).

- **Avoid classes that are open to one side**. If, however, the minimum and maximum are unknown or the cell populations are too small, the lowest class must remain open at the bottom and the highest class must remain open at the top (e.g. top category of the variable net income: "4.000 and more").

- Here, however, the question is how the class centers are to be determined for further calculations.
  Here you use the **midpoint of the class specified**, for open classes use the boundary given

- **Build classes so that in each class there are enough cases for building crosstabs (min 100 cases)**

Generate a new variable by compute

- Generating variables in SPSS is simple, especially if you want to generate a new variable from an already existing variable.

- Compute is generally used, when you want to combine two or more variables

- Reseachers often generate new variables that are copies of current ones if they want to change or recode the date, while also keeping the original data so it is not lost.

- There is variety of possibilities of creating new variables .

Count Values – First Step

- Use '**Transform**' and click on '**Count Values within Cases**' to call up the count dialog.

- Indicate a target variable that will receive the counts and specify a label for it.

- Consider which values are to be counted (you can count more than one value)
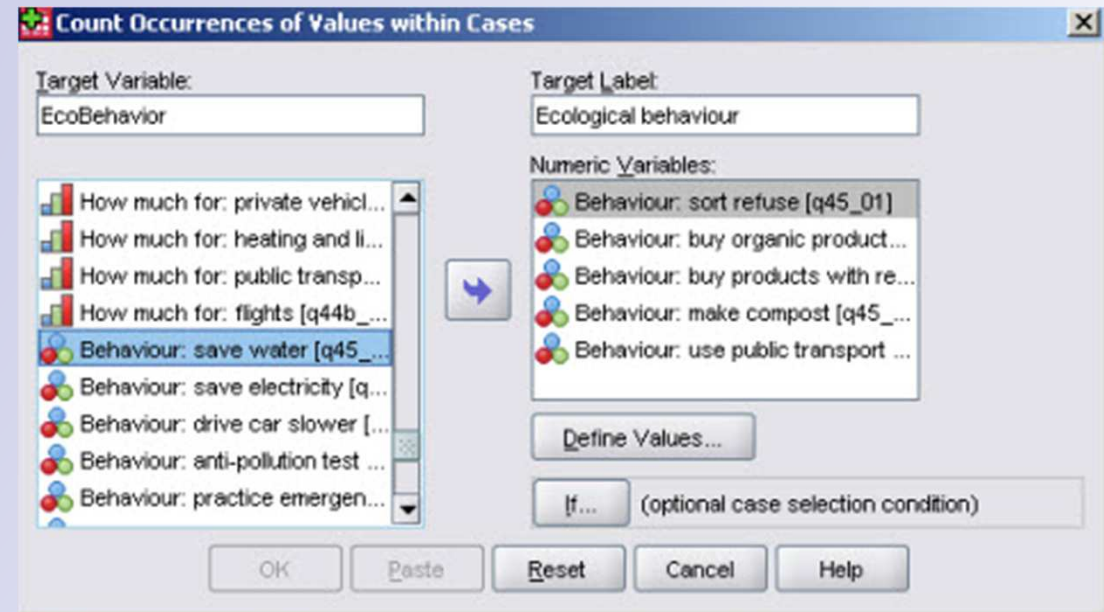
Count Values – Second Step

- Select the value(s) you want to count. You can count specific values, ranges of values, but also Missing values and SYSMIS. Add all values you want to count. (In our example we are counting 1s in all the variables)

- Note: counting never produces missing values

Co-funded by
the European Union

RED
Re-Educating
Digitisation

Rheinische
Hochschule
Köln
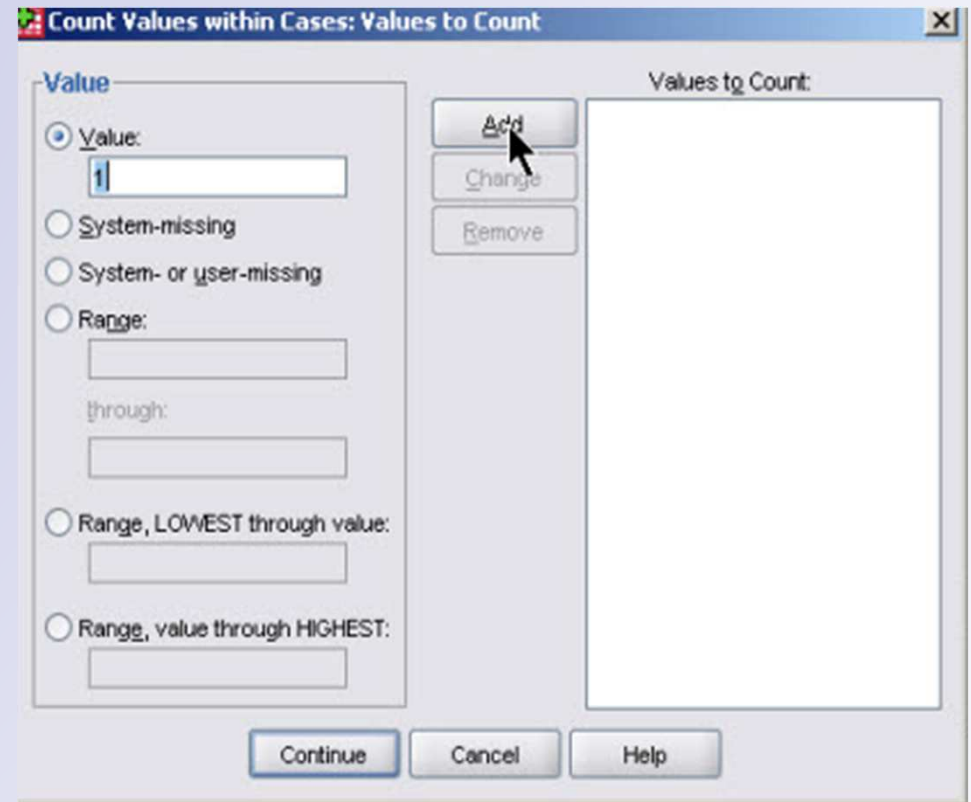
Saving the file

- The file should always be saved in order to save the work that has been done to date

- In order to save the data you have to be in the data view window

- Click 'File' and after that's 'Save as'

- You can save the file in different forms by clicking 'save as type'

- The saved data should always end with 'sav' in order to open it in SPSS again

- To finish saving the file click 'save'

Exercise:

- **Build classes** for Education and Age in the data set

- **Compute** an new variable **"LeisActs"** that **sums up** the degree of leisure activities v2.1 – v2.8. Recode the new variable into another variable **"LeisAct_3" into** 3 categories for high, mid and low activity and assign appropriate labels. Do a frequency command and interpret the result. Take care of missing values!

- Find out whether there are people who do not do a single activity!

- Compute a new variable from a combination of gender and migration background by using an IF – Command.
  Do a frequency command and interpret the result.

Introduction into some basic statistical data analyses with SPSS

Univariate analyses: Doing Frequencies (distribution table)

- Gives you an idea about the distribution of one variable in a short glace
- Most **common and important mode of data analyses** in SPSS
- Shows you absolute figures, percentages, valid percentages (excludes missings) and cumulative percentages

The data analysis – Frequencys: important core variables

Frequencies command can be used to make

- **countings**

- **and also to determine quartiles, percentiles,**

- **measures of central tendency (mean, median, and mode),**

- **measures of dispersion (range, standard deviation, variance, minimum and maximum),**

- **measures of kurtosis and skewness, and create histograms.**

Example: Evaluation of ticket prices and willingness to pay at a museum in Cologne
(shares in percent of respondents, mean values per subgroup)

Evaluation of the ticket prices

- a bit to high 33%
- too low 1%
- Much to high 4%
- weiß nicht 8%
- appropriate 54%
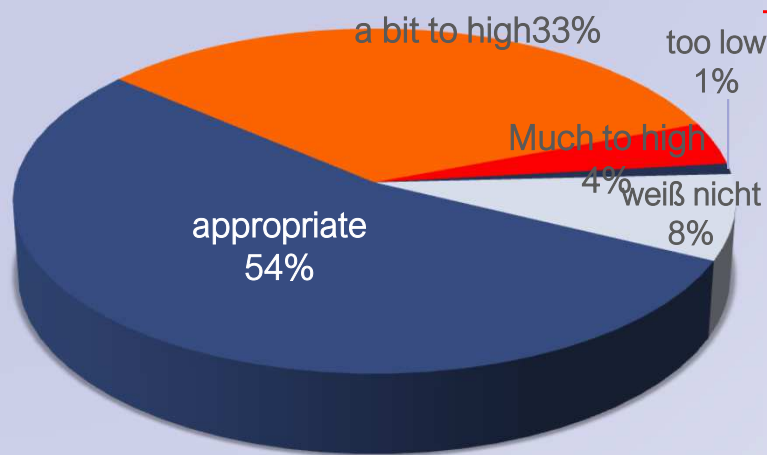
Average willingness to pay

€

12,8	8,2	6,6	7,9	10,5
reasonable	somewhat too high	much too high	don't know	average
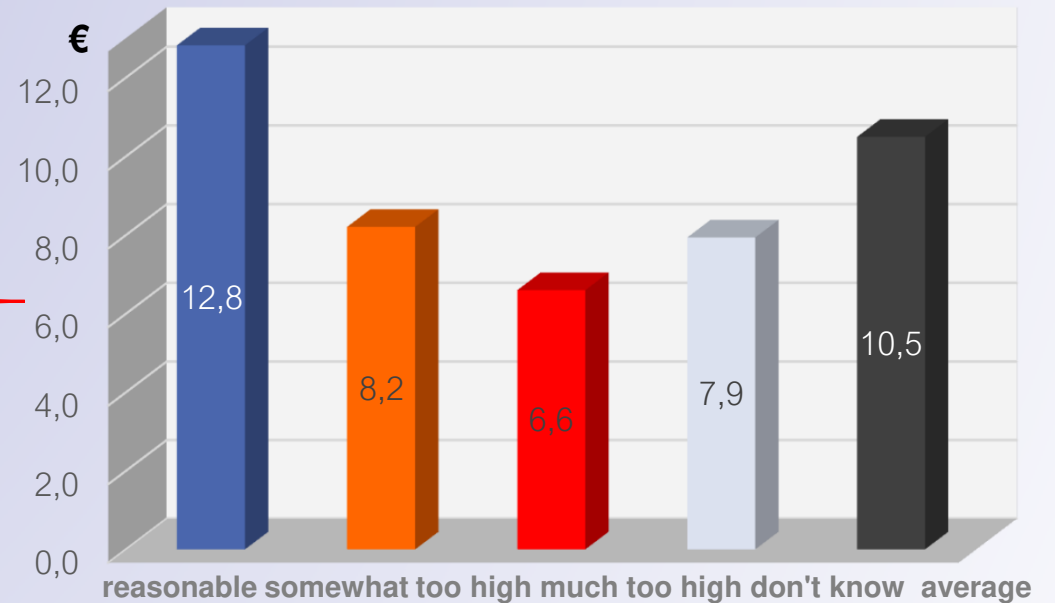
- • A good half of the visitors consider the ticket prices to be reasonable (54%), while just under 40% of those surveyed thought they were somewhat or significantly too high. A share of just under 10% found the prices too low or did not know them. The median willingness to pay is €10.50, which is 22% lower than the ticket price for full-payers. This shows a relatively high level of dissatisfaction with the pricing.

- • To assess what "reasonable" or "too high" actually mean, the breakdown by willingness to pay helps: while "reasonable" at €12.8 is very close to the actual ticket price of full-payers, "a little too high" already implies a significantly lower willingness to pay (€8.2 on average), while "much too high" only implies the willingness to pay about 50% of the ticket price.

A short introduction into statistical measures

- While frequency counts provide information about the entire distribution of a variable, statistical measures provide information about specific properties of a distribution.

- Measures serve to **condense information**, because they provide compressed information about the characteristic properties of a distribution with only one **single number**.

- Statistical measures are to describe a univariate distribution.

They are divided into two groups:

- **Location measures** (also: measures of central tendency, centrality values)

- **Measures of dispersion** (also: dispersion measures)

What are central statistical measures?

Measures of centrality	Measures of dispersion
Mode ($'X'_M$) (read X across M)	Range (R)
Median ($'X'_Z$) (read X across Z)	(mean) quartile distance (QA/QMA), abs. Deviation
Arithmetic mean ($'X'$) (read X across)	Variance (s^2), Standard Deviation (s)
Geometric mean ($'X'_G$) (read X across G)	
Harmonic mean ($'X'_H$) (read X across H)	

RED Re-Educating Digitisation

Rheinische Hochschule Köln

A brief introduction into some statistical measures

- While **location measures** provide information about the **centrality of a distribution**, i.e. reflect the typical (mode), the central (median) or the average (arithmetic mean) value of a distribution,

- The **dispersion measures** indicate the extent to which the data of a distribution deviate from these "typical" values.
  In this way, they measure the degree of heterogeneity of a distribution and show how well or how poorly the location measures represent a distribution:

  – If we have a relatively **homogeneous distribution**, i.e. only very few measured values deviate from the centrality values (= low dispersion), then the **centrality values are very good representatives** of the distribution.

  – If instead we have a very **heterogeneous distribution**, i.e. the measured values deviate quite strongly from the centrality values (= high dispersion), then the **centrality values do not represent the distribution** very well.

What is the Mode ($´X´_M$)? (Read X across m)**?**

The 'MO' mode $´X´_M$

- The simplest positional measure is the mode or modal value
- To determine the modal value, one asks the simple question *"Which value occurs most frequently?"* It is applicable from nominal measurement level!

- The mode is the value that occurs most frequently in a data set
- There can be multiple modes in a record or there can be no mode at all
- If serval values share the greatest frequency of occurrence, each of them is a mode
- The frequencies producure reports only the smallest of such multiple modes

RED Re-Educating Digitisation

Rheinische Hochschule Köln

Examples of the Mode

X_i	h_i	
1	1	
2	**12**	
3	10	
4	5	
5	2	
7	1	
Total	n = 31	

Here h = 12 with x_i = 2; $`X´_M$=2
Since we only have one modal value here, it is a unimodal distribution.
There may not be a clear modal distribution!

Caution: Do not confuse with the h_i value; for categorical (Nominal, ordinal) characteristics, the mode is the most frequent measured value x_i . For cardinal characteristics, it is the measured value itself.

The concept of mode is consistently applicable to nominal, ordinal as well as cardinal/scale features. For classified data, one speaks of modal class instead of mode.

Robustness: The value of the mode is not influenced by outliers or extreme values for cardinal characteristics

What is the „Mean" Value (´X´)? (Read X across)

- Mean is the sum of all data divided by the number of data

- There are different types of means, including **the Arithmetic Mean (AM)**, the **Geometric Mean** and the **Harmonic Mean**

- **-> For our purposes in most cases the AM is important**

- Formula for the **Arithmetic Mean**: **´X´ = (x_1 + x_2 + x_3 + ... + x_n) / n**

Examples of the Arithmetic Mean

- Data set: 1,2,3,4,5

- Number of data (n): 5

- Arithmetic mean (AM): (1+2+3+4+5) / 5 = 3

x_i	h_i	f_i
1	1	0,03
2	12	0,39
3	10	0,32
4	5	0,16
5	2	0,06
7	1	0,03
'x`=22/7=3,66	n = 31	

In case of a frequency distribution the following formula is used:

$$\grave{A}M = \frac{(1 \times 1) + (2 \times 12) + (3 \times 10) + (4 \times 5) + (5 \times 2) + 7 \times 1)}{31} = 2,97$$

$$\grave{A}M = 1 \times 0,03 + 2 \times 0,38 + 3 \times 0,32 + 4 \times 0,16 + 5 \times 0,06 + 7 \times 0,03 = 2,97$$

The arithmetic mean is generally calculated for multiple frequencies as follows (the h_i or f_i are the weights):

$$\grave{A}M = \frac{h_1 . x_1 + h_2 . x_2 + h_3 . x_3 + ... + h_n . x_n}{n} = 1/n \sum_{i=1} h_i . x_i = \sum_{i=1} f_i . x_i$$

Frequency of visits to independent theatres, by visitors
(Basis: respondents who have been to independent theatres in the last 12 months, mean value of visits per subgroup)

Age

1,80 — 1,62
1,32
1,56
1,68

1,30

,80

,30

-,20

18 - 29   30 - 44   45 - 59   60 +

Formal education*

1,80
1,60
1,40   1,28
1,20
1,00
,80
,60
,40
,20
,00

1,45

1,68

bis weiterf. Schule ohne Abitur   Abitur/Fachabitur   Studium

Mean differences statistically significant!

Gender

1,50   1,41
1,65

1,00

,50

,00

männlich   weiblich
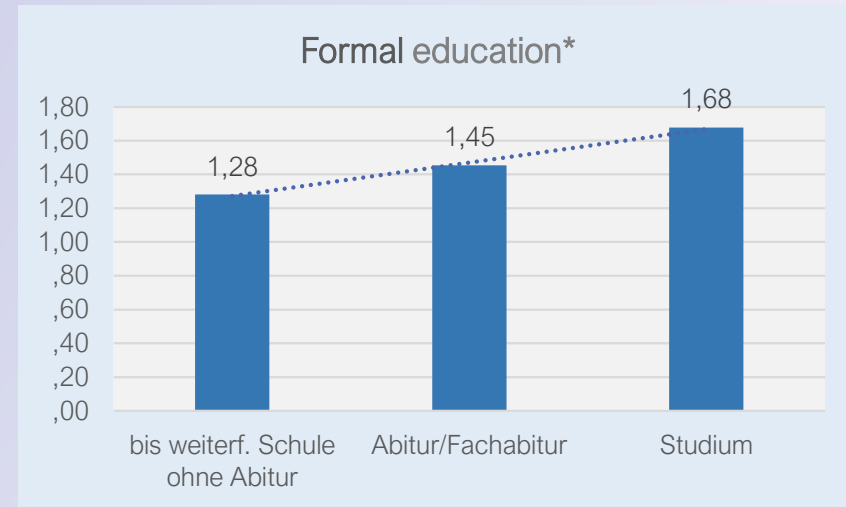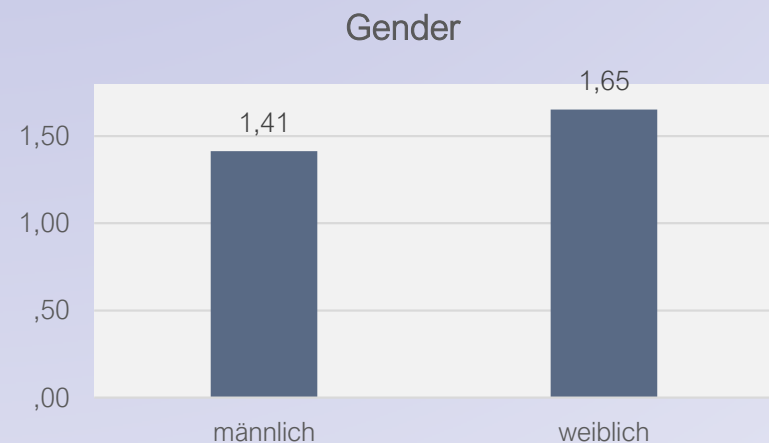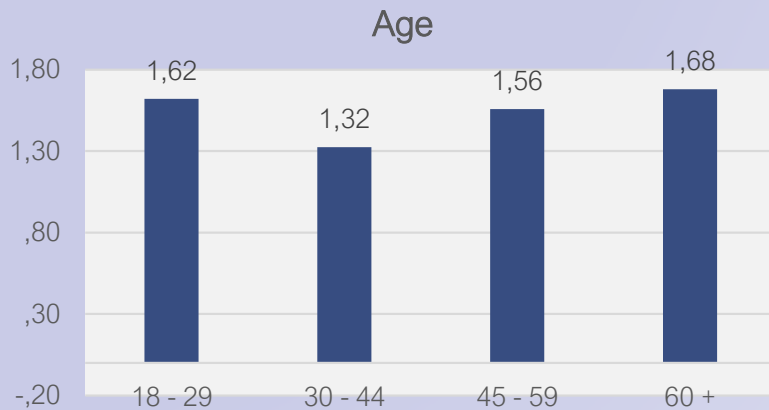
The analysis based on the existing visitors reveals to what extent the people who have found their way to the independent theatres find an offer that suits them. It therefore does not show the selection before the visit, but the selection of the audience on the basis of the visit experiences and offers within the theatre:

- For age and gender, the differences in visit frequency are not statistically significant, even though the older and female visitors show a slightly higher visit frequency.
- Only in the case of education does it remain the case that academics are more attracted to what is on offer. Not only do they go to an independent theatre in greater numbers, they also come back more often. Visitor loyalty is greatest here.
- Conclusion: The hurdles before visiting independent theatres are more selective than the programme itself!

RED Re-Educating Digitisation

Rheinische Hochschule Köln

Example: Mean willingness to pay, by age, place of residence and frequency of visits
(mean value in € per subgroup)
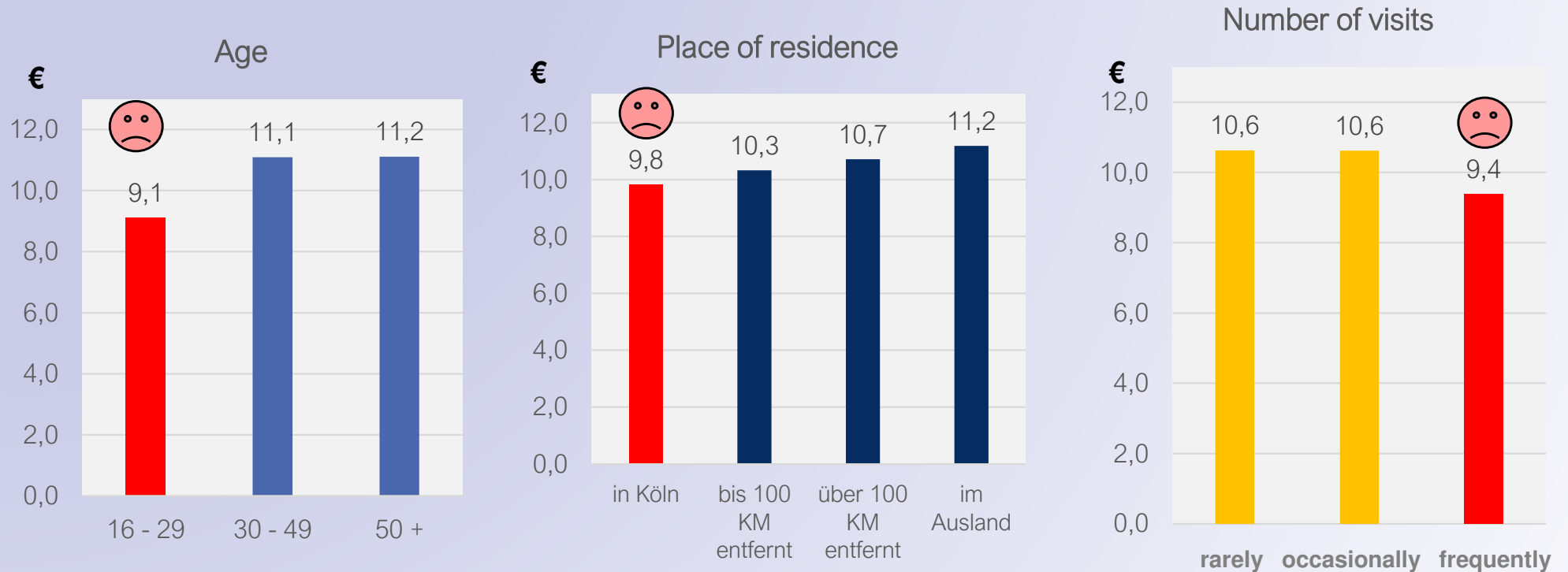
- Willingness to pay varies significantly with age, place of residence and frequency of visits. At € 9.1, the willingness to pay of 16-29 year-olds is 30% below the actual price. On average, the visitors from Cologne also consider an amount of less than 10 € to be appropriate. The particularly frequent visitors would obviously also like to see a reduction in the ticket price. For the long-distance visitors, on the other hand, the ticket price is less relevant, the willingness to pay is greater (but also only relatively little).
- Remarkably, visitors to the different exhibitions (permanent exhibition only, special exhibition only and both) do not differ in their willingness to pay (not shown).

Co-funded by
the European Union

RED
Re-Educating
Digitisation

Rheinische
Hochschule
Köln

What is the „Median" ($\acute{X}\acute{}_Z$)?

- The median is the value that lies in the middle of a sorted data set

- The median '$X\acute{}_Z$ is the value that halves a series of measurements ordered by size. It is applicable from ordinal measurement level to scale.

- The median '$X\acute{}_Z$ is that characteristic value of a characteristic X which 50 % of all characteristic values fall below or at most reach and which 50 % of all characteristic values exceed or at least reach.

- If the data set has an **odd number of values**, the median is simply the value in the middle

- If the data set has an **even number of values**, the median is the artithmetic mean of the two values in the middle

- The median is a measure of central tendency not sensitive to outlying values (unlike the mean, which can be affected by a few extremly high or low values).

- Is well suited for skewed data, i.e. little susceptible to outliers and extreme values

- This requires an ordinal or metric (cardinal) scale for determination and interpretation

Example of the Median (odd number of values)

a) Median calculation for <u>odd number of</u> cases (`M'$_{ZO}$)

The median here is *the measured value of the middle case of an ordered series of measured values.* Therefore, first order the data according to size, as in this ordered master list:

1,2,2,2,3,3,3,4,4,4,4,5,5,7,7 ; n = 15 (= odd number)

The median - the middle case - is calculated with the following formula:

$$(n + 1) / 2$$

$$\text{Here: } (15 + 1) / 2 = 8$$

Attention: Not 8, but the *measured value of the 8th. case* is the median;

Therefore M$_{ZO}$ = 4. As can also be seen from the cumulated absolute frequencies, the 16th case belongs to measured value 3.

Examples of the Median (even number of values)

a) Median calculation for _even number of cases_ (`M'$_{ZE}$)

For an even number of cases, the median is the halved value of the middle two cases. First, the middle two cases are calculated with:
$X_{n/2}$ = 1st. middle case and $X_{n/2+1}$ = 2nd. middle case

_Ex.:_ 1,2,2,2,2,3,3,3,4,4,4,4,5,5,7,7 ; n = 16 (= even number)

16 / 2 = 8th. case (= reading 3); (16 / 2) +1 = 9th. case (= reading 4)
'M'$_e$ = ½ (3 + 4) = 3.5

The median is now nothing other than the _Arithmetic Mean_ of the middle two measured values:
In general, the formula looks like this: 'M'$_{ZE}$ = 1/2 ($X_{(n/2)}$ + X)$_{(n/2)+1}$

Note again that the two middle cases [n/2 or (n/2)+1] are determined, which must then be replaced by their corresponding measured values in the case of non-cardinal data. For cardinal data, it is the value itselve.

How to calculate the Median of a set of data in SPSS

1. Click 'Analyses' → 'Descriptives Statistics' → 'Frequencies'
2. Move the variable for which you wish to calculate the median into the right-hand column
3. Click the 'Statistics button', select 'Median' under 'Central Tendency' and then press 'Continue'
4. Click 'OK' to perform the calculation
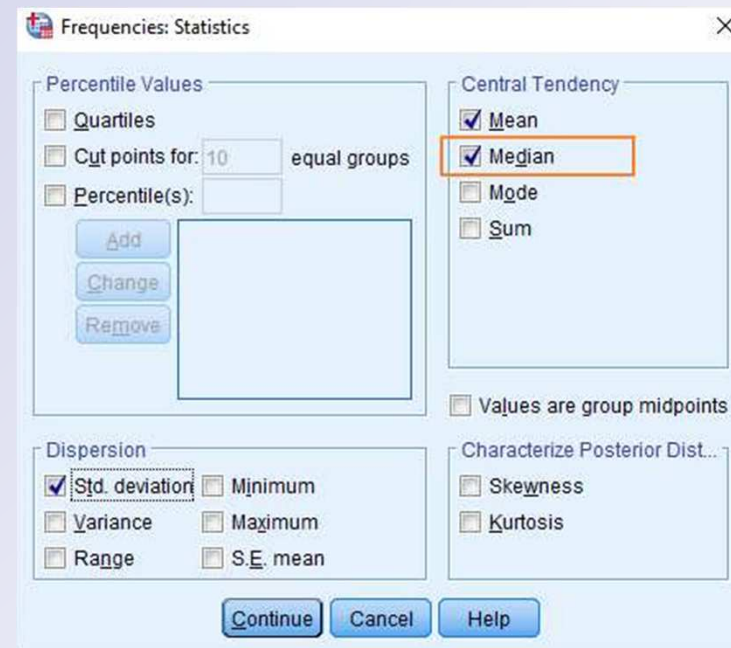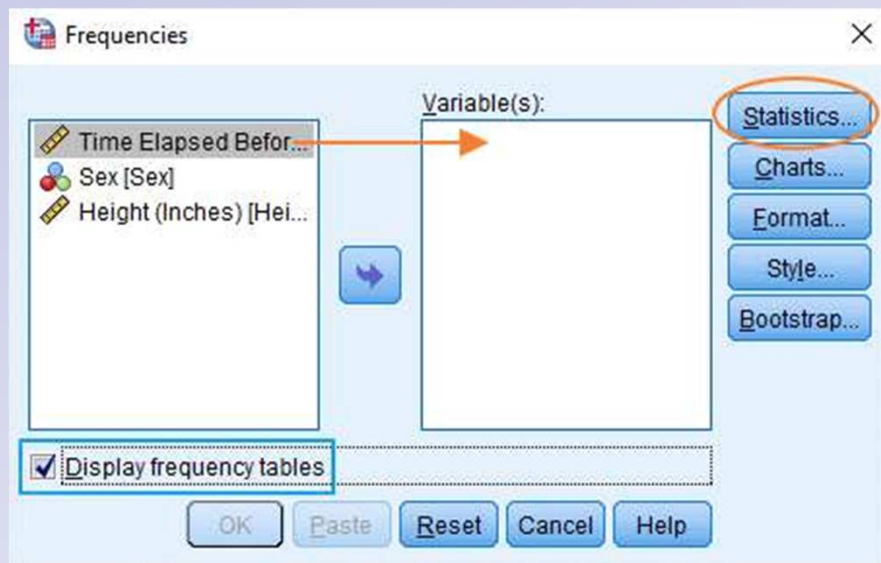
Measures of centrality in comparison

Positive Skew: Mode, Median, Mean. Symmetrical Distribution: Mean, Median, Mode. Negative Skew: Mean, Median, Mode.
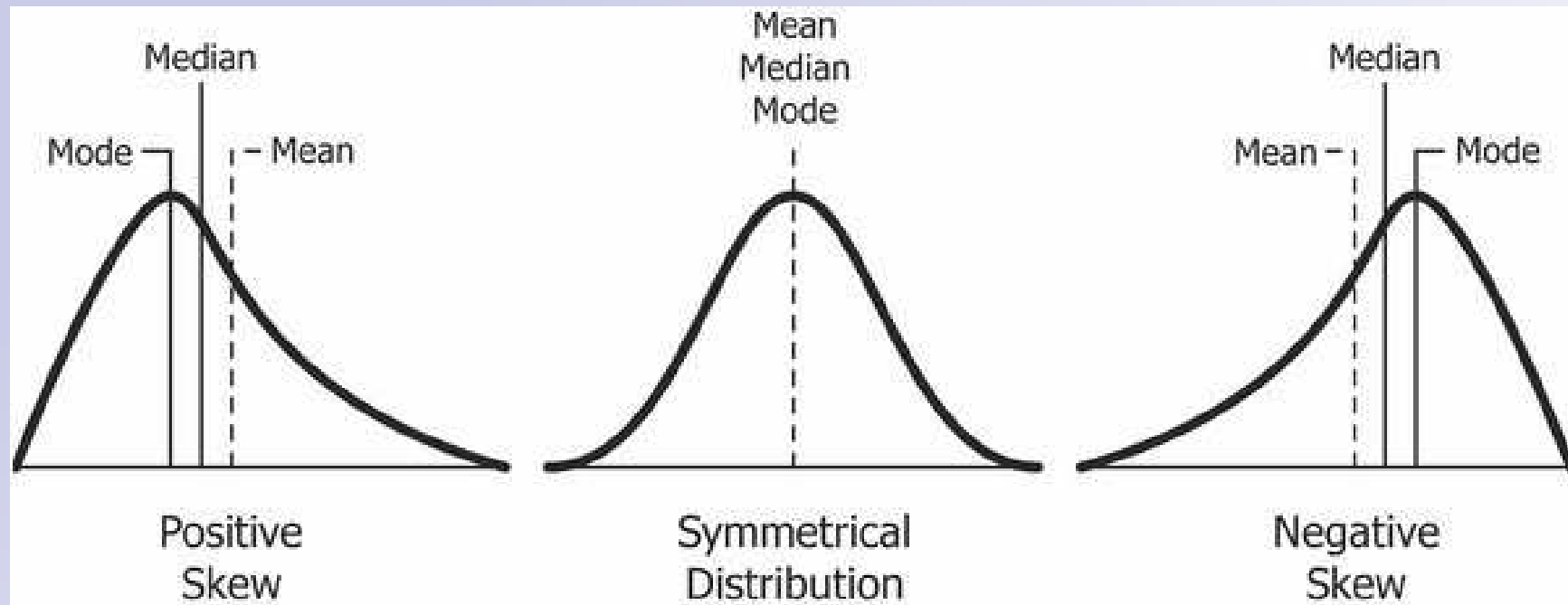
Exercise:

- **Calculate all measures of centrality for age, v2_1 and the newly computed variable about the degree of leisure activity ("LeisAct")**

- **Evaluate the skewness and symmetry of the three variables**

What are statistical measures and measures of dispersion?

SPSS has two primary options for calculating statistics:

- **Descriptives:** For basic statistics – mean, median, range and standard deviation
- **Frequencies:** Additional options – quartiles, percentiles and more dispersion statistics

Statistical measures → **Descriptives**
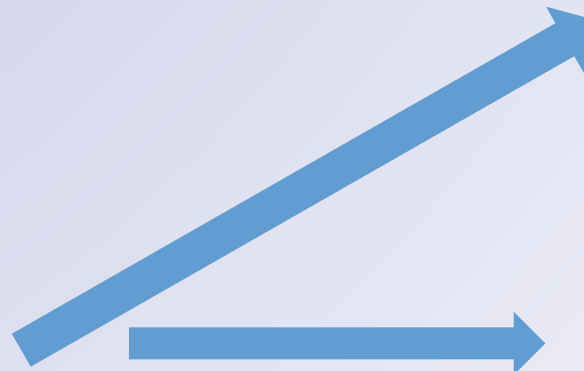
Measures of dispersion → **Frequencies**

Descriptives

- The Descriptives procedure displays univariate summary statistics for several variables in a single table and calculates standardized values (z scores).

- Variables can be ordered by the size of their means (in ascending or descending order), alphabetically, or by the order in which you select the variables (the default)

Types of Measures of Dispersion

- While mean values represent typical values of a sample, measures of dispersion are supposed to indicate whether the characteristic values are close to the mean value or at a greater or lesser distance from it.

- They "qualify" the quality of the position measure

- However, distances can only be measured for cardinal characteristics. Therefore, measures of dispersion are useful to define especially for cardinal characteristics - for ordinal characteristics also the quartile distance.

There are several types of measures of dispersion, including:

- Variance (s^2)

- Standard deviation (s)

- Interquartile range (IQR)

What is Variance?

- The variance is a measure of how far each value in the data set deviates from the mean value
- It shows us how "broadly" or "narrowly" a data set is distributed
- Formula for the variance: $s^2 = \sum(x_i - ´X´)^2 / n$
- Represents the basis for further calculations, e.g. in regression or analysis of variance
- A high variance shows us, that the mean value does not represent the data very well
- It is only applicable for scale data, not for ordinal or nominal

Example:

- Data set: 1,2,3,4,5
- Mean value (´X`) = 3
- Variance s^2: $[(1 - 3)^2 + (2 - 3)^2 + (3 - 3)^2 + (4 - 3)^2 + (5 - 3)^2] / 5 = 2$

What is Standard Deviation?

- The standard deviation is the square root of the variance
- It gives the unit of the measurement that corresponds to the data referred (e.g. centimeter, kilometer, euro, whereas variance has no dimension)
- Formula for the standard deviation: $s = \sqrt{s^2}$
- For normally distributed values, approx. 67% of all measured values lie in the intervall described by the mean +- 1 of the standard deviation
- Standard deviation describes the average deviation from the mean
- A high s shows, that the mean value does not represent the data very well
- It is only applicable for scale data, not for ordinal or nominal data

- Data set: 1,2,3,4,5
- Mean value (´X´): 3
- Variance (s^2): 2 (as described above)
- Standard deviation (s): $\sqrt{2} = 1.41$

What is the "Interquartile Range" (IQR)?

- The interquartile range is the difference between the 75th and 25th percentile of a data set; it encompasses **the middle 50% of the sample**
- An interquartile range is a measure of where 50% of the respondents/values lie
- The interquartile range formula is the first quartile subtracted from the third quartile:

$$IQR = Q_3 - Q_1$$

- The IQA is also robust to outliers and strong deviations

- Data set: 1,2,**5**,6,7,**9**,**12**,**15**,**18**,19,27, n: 11
- Q_2: Median: 9
- Q_1: 5 (median in the lower half of the data)
- Q_3: 18 (median of the upper half of the data)
- Interquartile Range: 18 – 5 = 13

Exercise:

- **Calculate all measures of dispersion for age, v2_1 and the newly computed variable about the degree of leisure activity ("LeisAct")**

- **Evaluate the degree of dispersion**

The data analysis – Box-Plot

- The boxplot (also box diagram or box graph) combines different scatter measures and position measures and displays them graphically.

- A boxplot shows all values of the five-point summary (minimum, first quartile, median, third quartile and maximum)

- These metrics are very robust to outliers and deviations from the normal distribution

- Boxplot therefore provide quick insights into the distribution of your data regardless of how it is distributed

- It is used to display the distribution of an ordinal or metric variable

- Boxplot also mark possible outliers

The data analysis – Box-Plot

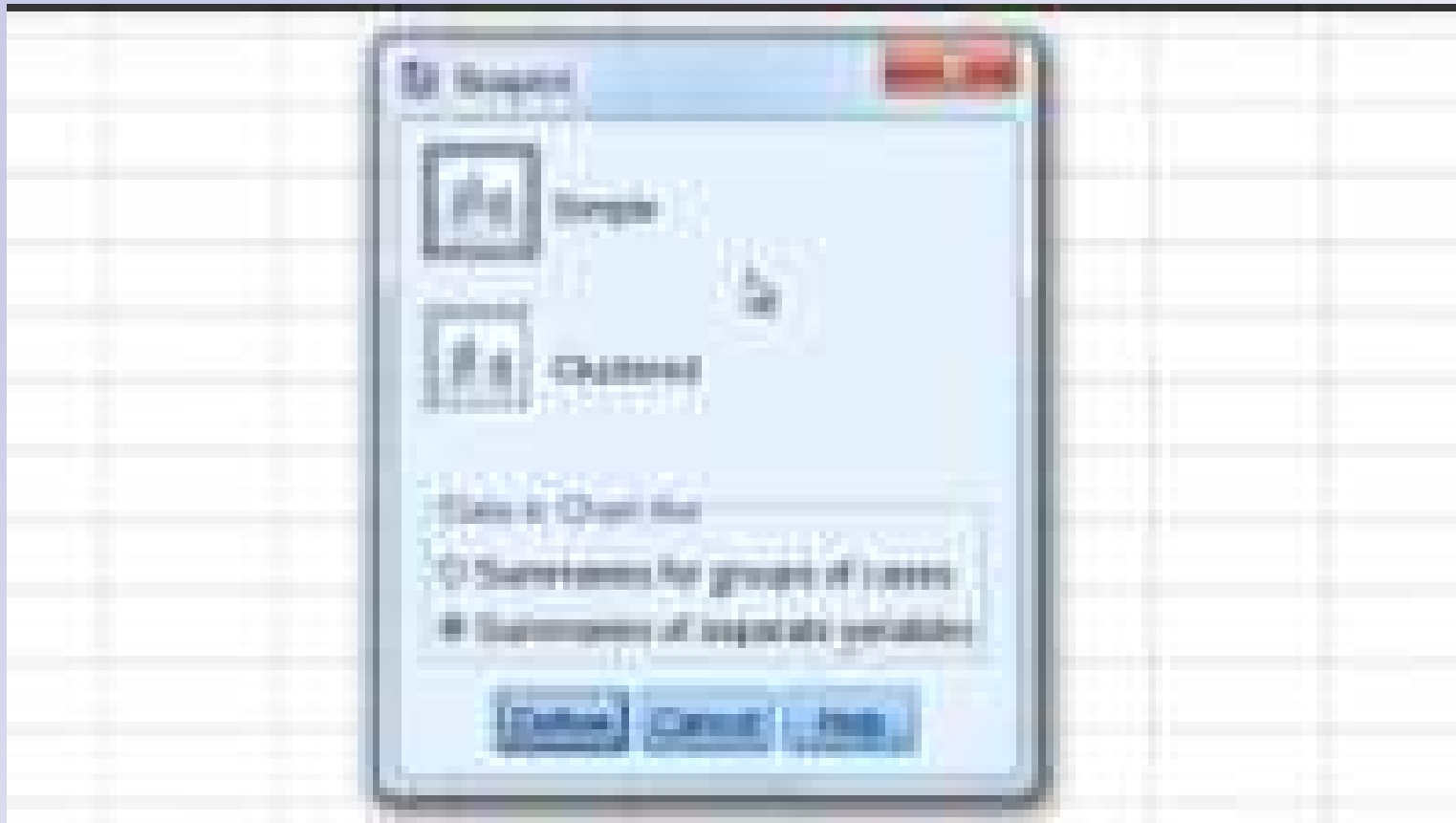## How to create and interpret a boxplot (1/2)

The data analysis – Box-Plot

How to create and interpret a boxplot

The Data Analysis – Box-Plot

How to create and interpret a boxplot (2/2)

https://www.youtube.com/watch?v=X4hSSu5oTj4&ab_channel=QuantitativeSpecialists

Co-funded by
the European Union

RED
Re-Educating
Digitisation

Rheinische
Hochschule
Köln

Exercise:

- **Calculate a boxplot for the newly computed variable about the degree of leisure activity ("LeisAct")**

- **Caculate a boxplot for LeisAct within age groups**

Introduction into statistical Bivariat Analyses: Crosstabs

Definition:

- A crosstab is a special type of table that relates one variable to another variables or more variables. The crosstab shows us how often certain combinations of variables occur.

- It is only applicable to categorical variables.

Why are crosstabs important for data analysis?

- Can help you visualize the **differences between subgroups**

- Shows you **possible relationships between two (or more) variables**

- Makes it easier to interpret data, which is especially benefical for researchers who have limited knowledge of statistical analysis

Crosstabs – rows, columns and table dimension

- Crosstabs are a simple and very efficient tool do realize relationships between variables

- Statistical knowledge is no prerequisite for doing analyses

- The dimensions of the crosstab refer to the **number of rows** and **columns in the table** (the "total" rows/column are not included.)

- The table dimensions are reported as R x C, where R is the number of categories for the row variable and C is the number of categories for the column variable.

- Additionally, a "square" crosstab is one in which the row and column variables have the same number of categories.
Tables of dimensions 2x2, 3x3, 4x4 etc. are all square crosstabs.

Crosstabs – Understanding Row, Column and Total Percent

A typical 2x2 crosstab has the following construction:

- The letters a,b,c and d represent what are called 'cell counts'.
- a = number of observations corresponding to row 1 AND column 1.
- b = number of observations corresponding to row 1 AND column 2.
- c = number of observations corresponding to row 2 AND column 1.
- d = number of observations corresponding to row 2 AND column 2.

	Column 1	Column 2	Row totals
Row 1	a	b	$a+b$
Row 2	c	d	$c+d$
Column totals	$a+c$	$b+d$	$a+b+c+d$

Crosstabs – Understanding Row, Column and Total Percent

By adding a, b, c and d, we can determine the total number of observations in each category and in the table overall.

- **Row sum** of row (i.e., total number of observations in row 1): a + b

- Row sum of row 2 (i.e., total number of observations in row 2): c + d

- **Column sum** of column 1 (i.e., total number of observations in column 1): a + c

- Column sum of column 2 (i.e., total number of observations in column 2): b + d

- **Total sum** (i.e., total number of observations in the table): n = a + b + c + d

The row sums and column sums are referred to as 'marginal frequencies'.

When you are describing the composition of your sample, it is often useful to refer to the proportion of the row or column that fell within a particular catefory. This can be achieved by computing the 'row percentages' or 'column percentage'

Crosstabs – Understanding Row, Column and Total Percent

Notice that when computing row percentage, the denominators for cells a,b,c, d are determined by the row sums (here, a + b and a + d). The percentage in the 'row totals' column here must equal 100%.

	Column 1	Column 2	Row totals
Row 1	a	b	$a + b$
Row 1 %	$a / (a + b)$	$b / (a + b)$	$(a + b) / (a+b) = 100\%$
Row 2	c	d	$c + d$
Row 2 %	$c / (c + d)$	$d / (c + d)$	$(c + d) / (c + d) = 100\%$
Column totals	$a + c$	$b + d$	$a + b + c + d$
% of total	$(a + c) / (a + b + c + d)$	$(b + d) / (a + b + c + d)$	$(a + b + c + d) / (a + b + c + d) = 100\%$

Crosstabs – Understanding Row, Column and Total Percent

Notice that when **total percentages** are computed, the denominators for all of the computations are equal to the **number of oberservations** in the table, i.e. a + b + c + d.

	Column 1	Column 2	Row totals
Row 1	a	b	$a + b$
% of total	$a/(a+b+c+d)$	$b/(a+b+c+d)$	$(a+b)/(a+b+c+d)$
Row 2	c	d	$c + d$
% of total	$c/(a+b+c+d)$	$d/(a+b+c+d)$	$(c+d)/(a+b+c+d)$
Column totals	$a + c$	$b + d$	$a + b + c + d$
% of total	$(a+c)/(a+b+c+d)$	$(b+d)/(a+b+c+d)$	$(a+b+c+d)/(a+b+c+d) = 100\%$

Components of a crosstab

<u>Example: Summarizing the relationships between Variables:</u>

- Let's use different aspects of the Crosstabs procedure to investigate the relationship between class rank and living on campus

- There are several variables relating to this question:

  - Rank: Class rank – Freshmen, Sophomore, Junior, Senior

  - LiveOnCampus: Do you live on campus? – Yes/No

  - State: Are you an in-state or out-of-state student? – In State, Out of state

Crosstabs – Understanding Row, Column and Total Percent

Simple Crosstabs: Output

- The second table Class Rank: Do you live on campus? (Crosstabulation) contains the crosstab

- We can quickly observe information about the interaction of these two variables:

- Many more freshman (1st. year student) lived on-campus (100) than off-campus (37)

- About an equal number of sophomore (students in 2nd. year) lived off-campus (42) versus on-campus (48)

- Far more juniors (3rd. year students) lives off-campus (90) than on-campus (8)

- Only one (1) senior (4th. year students) lives on-campus, the rest lived off-campus (62)

Class rank * Do you live on campus? Crosstabulation

Count

		Do you live on campus?		
		Off-campus	On-campus	Total
Class rank	Freshman	37	100	137
	Sophomore	42	48	90
	Junior	90	8	98
	Senior	62	1	63
Total		231	157	388

Crosstabs – Row, Column and Total percentage

Output – Row percents:

- If the row variable is '*RankUpperUnder*' and the column variable is '*LiveOnCampus*', then the row percentages will tell us what percentage of the upperclassmen or what percentage of the underclassmen live on campus

- Variable RankUpperUnder will determine the denominator of the percentage computation

Class Rank * Do you live on campus? Crosstabulation

			Do you live on campus?		Total
			Off-campus	On-campus	
Class Rank	Underclassman	Count	79	148	227
		% within Class Rank	34.8%	65.2%	100.0%
	Upperclassman	Count	152	9	161
		% within Class Rank	94.4%	5.6%	100.0%
Total		Count	231	157	388
		% within Class Rank	59.5%	40.5%	100.0%

Crosstabs – Row, Column and Total percentage

Output – Column percents:

- If the row variable is *'RankUpperUnder'* and the column variable is '*LiveOnCampus*',

  then the column percentage will tell us what percentage of the individuals who live on

  campus are upper (Junior/Senior) or underclassmen (Fresh/Sopho)

- Variable '*LiveOnCampus*' will determine the denominator of the percentage camputation

Crosstabs – Row, Column and Total percentage

Output – Total percentage:

- If the row variable is '*Rank Upper / Under*' and the column variable is 'LiveOnCampus', then the total percentage tells us what proportion of the total is within each combination of '*RankUpper / Under*' and '*LiveOnCampus*'

- The overall table size determines the denominator of the percentage computations

Class Rank * Do you live on campus? Crosstabulation

			Do you live on campus?		Total
			Off-campus	On-campus	
Class Rank	Underclassman	Count	79	148	227
		% of Total	20.4%	38.1%	58.5%
	Upperclassman	Count	152	9	161
		% of Total	39.2%	2.3%	41.5%
Total		Count	231	157	388
		% of Total	59.5%	40.5%	100.0%

Displaying data in crosstabs

Graphical representation (e.g. bar chart)

How often do you do the following things?

	Going to the cinema	Going to a art house cinema	Visiting film festivals	Going to a concert	Going to the theatre
never	23%	48%	68%	29%	49%
rarely	34%	32%	19%	37%	30%
occasionally	34%	17%	10%	26%	18%
frequently	9%	3%	3%	8%	3%

■ frequently ■ occasionally ■ rarely ■ never
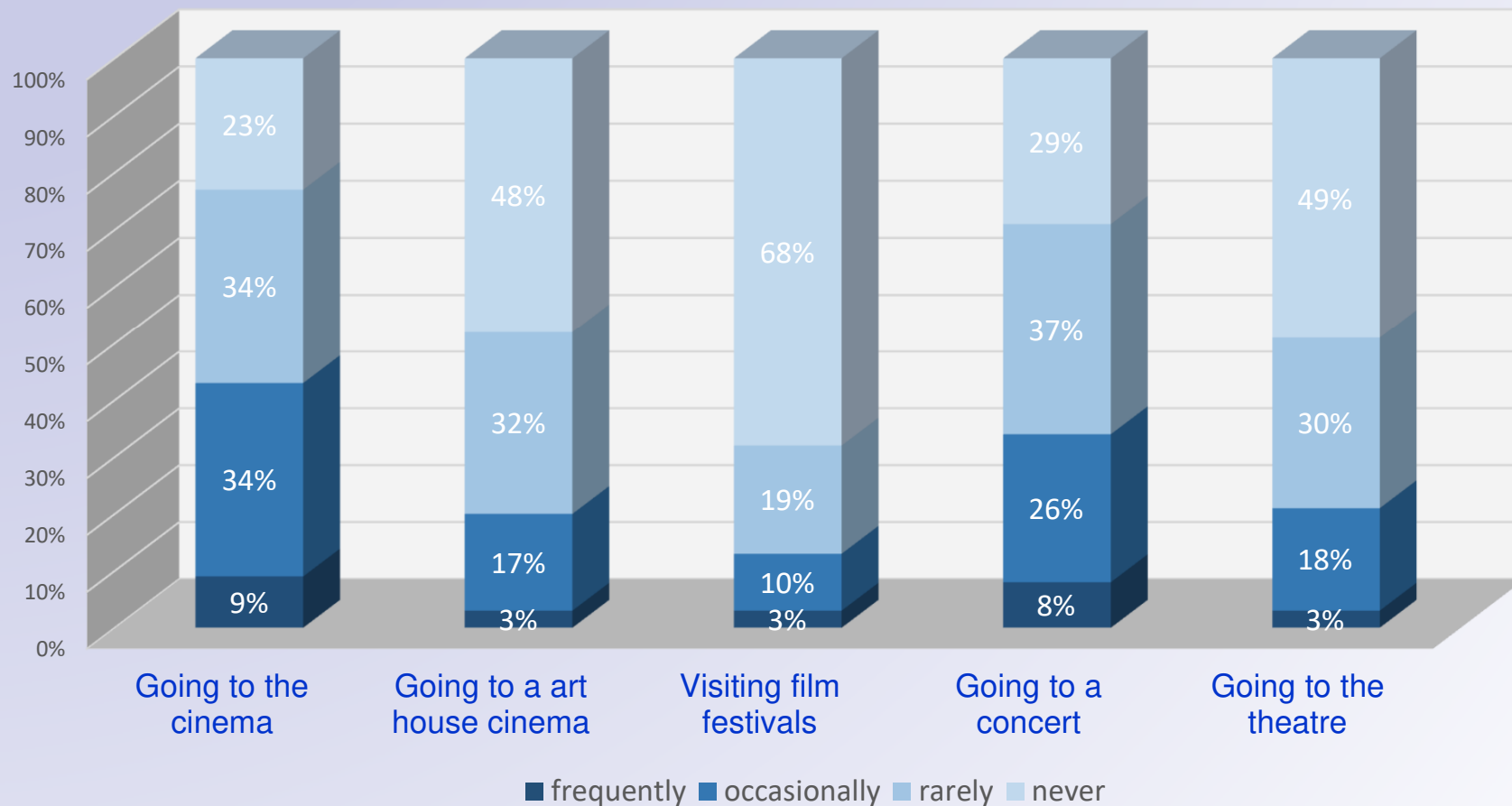
Some simple rules when doing crosstabs

Example Film_Survey Cologne: Attending Film Festival X Age

Have the following in mind:

- **Follow the KISS – Rule (keep it simple and stupid)**

- **Verify the table first by taking a look at the absolut numbers (total of each row or column should be 80 minimum)**

- **Think before doing tables, what you want to analyse / to see from the table**

- **Show only one kind of result in one table: Row percent, column percent, total percent**

- **Show the percents preferably always in column percent**

- **If you think, there may be an independent and a dependent variable, put the independent variable always in the *head* of the table**

Bad example: Avoid this!

Attending film festivals * Age Crosstabulation

			Age 18 - 34	Age 35 - 49	Age 50 - 64	Age 65 +	Total
Attending film festivals	Frequent	Count	9	4	3	1	17
		% within Attending film festivals	52,9%	23,5%	17,6%	5,9%	100,0%
		% within Age	5,1%	3,2%	2,3%	,8%	3,1%
		% of Total	1,6%	,7%	,5%	,2%	3,1%
	Occasionally	Count	25	18	5	5	53
		% within Attending film festivals	47,2%	34,0%	9,4%	9,4%	100,0%
		% within Age	14,2%	14,5%	3,8%	4,1%	9,6%
		% of Total	4,5%	3,3%	,9%	,9%	9,6%
	Rare	Count	38	28	26	12	104
		% within Attending film festivals	36,5%	26,9%	25,0%	11,5%	100,0%
		% within Age	21,6%	22,6%	19,7%	9,9%	18,8%
		% of Total	6,9%	5,1%	4,7%	2,2%	18,8%
	Never	Count	104	74	98	103	379
		% within Attending film festivals	27,4%	19,5%	25,9%	27,2%	100,0%
		% within Age	59,1%	59,7%	74,2%	85,1%	68,5%
		% of Total	18,8%	13,4%	17,7%	18,6%	68,5%
Total		Count	176	124	132	121	553
		% within Attending film festivals	31,8%	22,4%	23,9%	21,9%	100,0%
		% within Age	100,0%	100,0%	100,0%	100,0%	100,0%
		% of Total	31,8%	22,4%	23,9%	21,9%	100,0%

Doing Crosstabs with other data analyses software

Creating crosstabs using excel

Step 1: Enter the Dataset into Excel

	A	B	C	D	E
1	**Team**	**Position**	**Points**		
2	A	Guard	12		
3	A	Guard	19		
4	A	Forward	22		
5	A	Forward	24		
6	A	Forward	17		
7	A	Center	29		
8	B	Guard	32		
9	B	Guard	33		
10	B	Guard	19		
11	B	Forward	9		
12	B	Center	8		
13	B	Center	14		
14					
15					

Step 2: Create the Crosstab

Click the **Insert** tab along the top ribbon and then click the **PivotTable** button

File   Home   Insert   Draw   Page Layout   Formulas   Data   Review

PivotTable   Recommended PivotTables   Table   Pictures   Shapes   Icons   3D Models   SmartArt   Screenshot   Ge My

Tables   Illustrations

L30   fx

	A	B	C	D	E	F	G
1	**Team**	**Position**	**Points**				
2	A	Guard	12				
3	A	Guard	19				
4	A	Forward	22				
5	A	Forward	24				
6	A	Forward	17				
7	A	Center	29				
8	B	Guard	32				
9	B	Guard	33				
10	B	Guard	19				
11	B	Forward	9				
12	B	Center	8				
13	B	Center	14				

Creating crosstabs using excel

In the new window that appears, select the range that contains the data as the **Table**/**Range** and choose any cell you'd like in the **Existing Worksheet** to place the crosstab. We'll choose cell **E2**

	A	B	C
1	Team	Position	Points
2	A	Guard	12
3	A	Guard	19
4	A	Forward	22
5	A	Forward	24
6	A	Forward	17
7	A	Center	29
8	B	Guard	32
9	B	Guard	33
10	B	Guard	19
11	B	Forward	9
12	B	Center	8
13	B	Center	14

Create PivotTable ? ✕

Choose the data that you want to analyze
- ◉ Select a table or range
  - Table/Range: Sheet1!$A$1:$C$13
- ○ Use an external data source
  - Choose Connection...
  - Connection name:
  - ○ Use this workbook's Data Model

Choose where you want the PivotTable report to be placed
- ○ New Worksheet
- ◉ Existing Worksheet
  - Location: Sheet1!$E$2

Choose whether you want to analyze multiple tables
- ☐ Add this data to the Data Model

OK   Cancel

Creating crosstabs using excel

Step 3: Populate the Crosstab with Values

Once you click OK, a new window on the right side of the screen will appear.

Drag the **Team** variablel to the **Rows** area, the **Position** variable to the **Columns** area, then the **Position** variable again to the **Values** ares as follows:
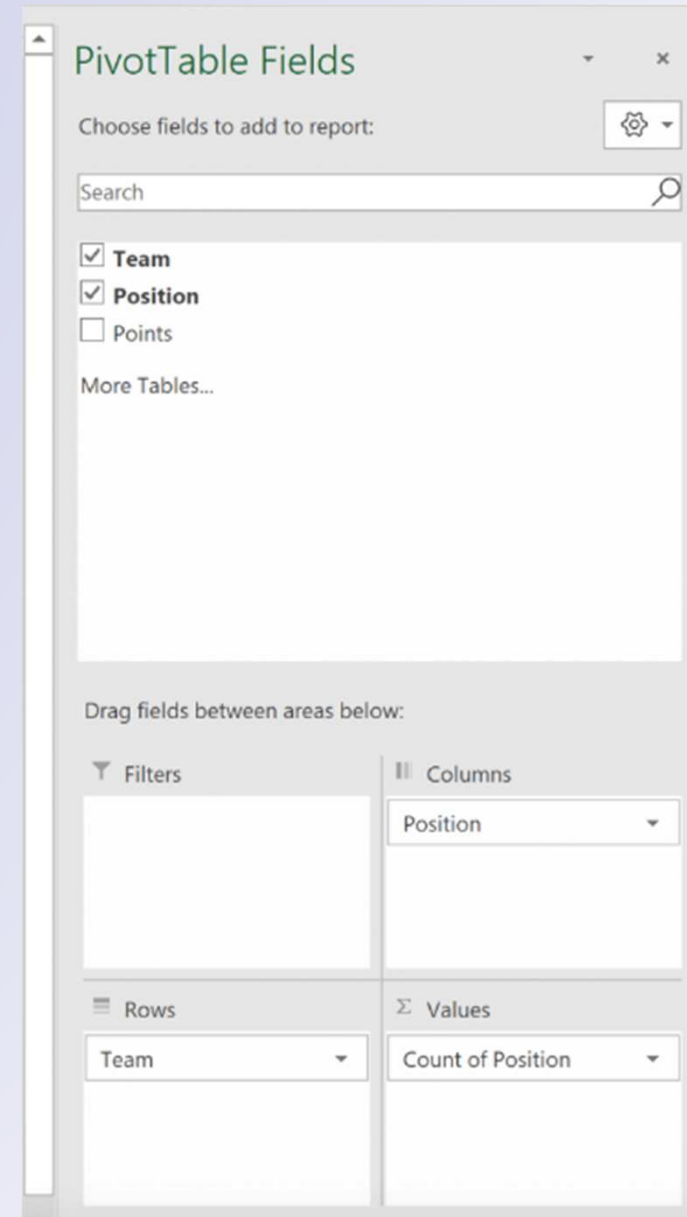
Creating crosstabs using excel

Once you do so, the following crosstab will appear in the cell that you specified:

	A	B	C	D	E	F	G	H	I
1	Team	Position	Points						
2	A	Guard	12		Count of Position	Column Labels ▾			
3	A	Guard	19		Row Labels ▾	Center	Forward	Guard	Grand Total
4	A	Forward	22		A	1	3	2	6
5	A	Forward	24		B	2	1	3	6
6	A	Forward	17		Grand Total	3	4	5	12
7	A	Center	29						
8	B	Guard	32						
9	B	Guard	33						
10	B	Guard	19						
11	B	Forward	9						
12	B	Center	8						
13	B	Center	14						
14									

Creating crosstabs using excel

Step 4: Interpret the Crosstab

Row totals:

A total of 6 players are on team A and a total of 6 players are on team B

Column Totals:

- A total of 3 players have a position of Center
- A total of 4 players have a position of Foward
- A total of 5 players have a position of Guard

Individual Cells:

- 1 player has a position of Center on team A
- 3 players have a position of Foward on team A
- 2 players have a position of Guard on team A
- 2 players have a position of Center on team B
- 1 player has a position of Forward on team B
- 3 players have a position of Guard on team B

Creating crosstabs using Google Sheets

Step 1: Enter the following dataset into Google Sheets that shows information for variour basketball players

	A	B	C	D	E
1	**Team**	**Position**	**Points**		
2	A	Guard	12		
3	A	Guard	19		
4	A	Forward	22		
5	A	Forward	24		
6	A	Forward	17		
7	A	Center	29		
8	B	Guard	32		
9	B	Guard	33		
10	B	Guard	19		
11	B	Forward	9		
12	B	Center	8		
13	B	Center	14		
14					

Creating crosstabs Google Sheets

Step 2: Click the "insert" tab along the top ribbon and then click "Pivot table" from the dropdown menu
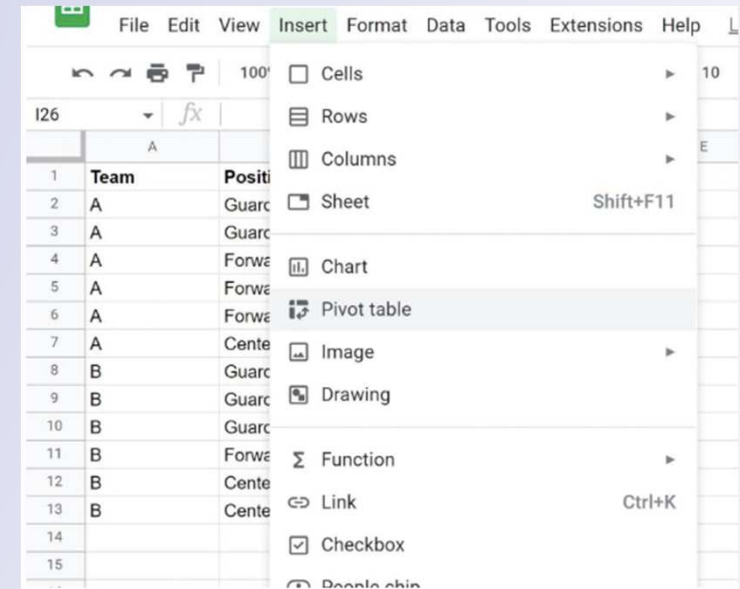
In the new window that appears, enter **Sheet1!A1:C13** as the Data range and **Sheet1!E1** as the Insert to location, then click **Create**:
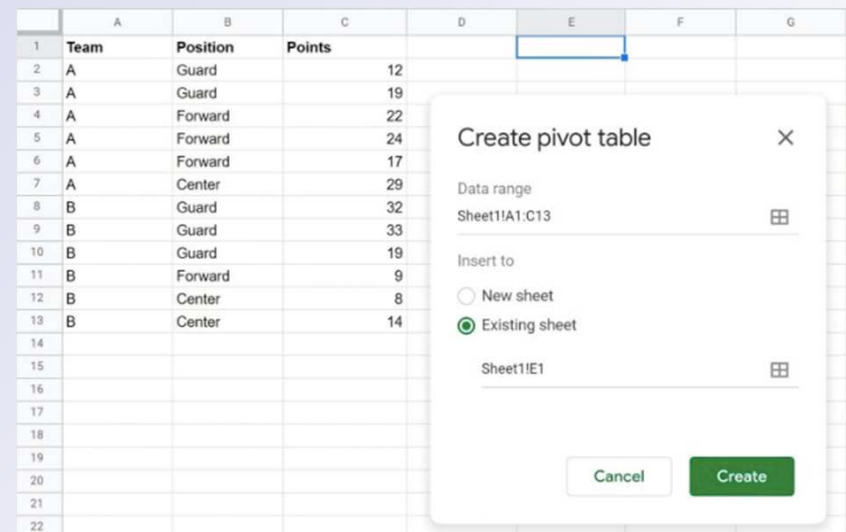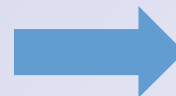
Creating crosstabs Google Sheets

Step 3: Once you click **Create**, a new Pivot table editor panel will appear on the right side of the screen.

Choose Team for the **Rows**, Position for the Columns and Points for the **Values**:

Once you do so, the following crosstab will appear in the cell that you specified:

RED
Re-Educating
Digitisation

Rheinische
Hochschule
Köln
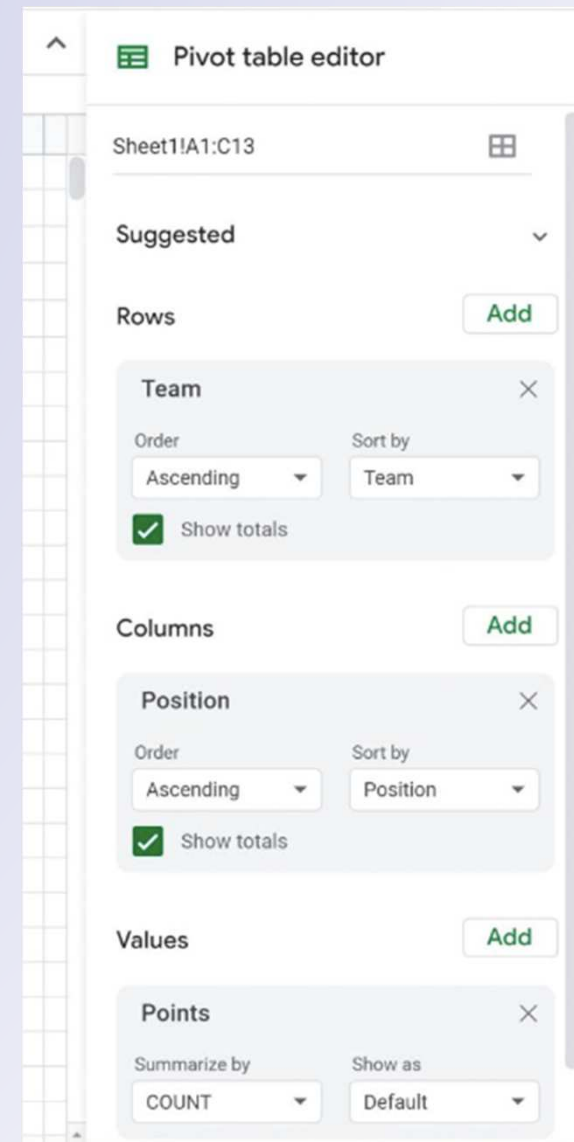
Creating crosstabs Google Sheets

	E	F	G	H	I
COUNT of Poin	Position				
Team	Center	Forward	Guard	Grand Total	
A	1	3	2	6	
B	2	1	3	6	
Grand Total	3	4	5	12	

Step 4: Interpret the Crosstab

Row Grand Total:
•   A total of 6 players are on team A and a total of 6 players are on team B
Column Grand Totals:
•   A total of 3 players have a position of Center; A total of 4 players have a position of Forward and a total of 5 players a position of Guard
Individual Cells:
•   1 player has a position of Center on team A, 3 player have a position of Forward on team A and two players have a position of Guard on team A.
•   2 players have a position of Center on team B, 1 player has a position of Forward on team B and 3 player have a position of Guard on team B.

End of this section!

RED – Workshop
Visitor Research II

- Types of Interviewing -

Prof. Dr. Tibor Kliment
Bukarest, November 2022

RED
Re-Educating
Digitisation

Rheinische
Hochschule
Köln

Common forms of questioning

low | | | | | | **Level of standardisation** → high

low

EDP

high

Exploratory interviews	Group discussion	Personal interviews	Telephone interviews	Written survey
without Content Analysis	**with pers. Moderator**	**PAPI** Paper and Pencil Interview	**TI** Telephone Interview	**Online Survey**
with Content Analysis	**Chat forum**	**CAPI** Computer Assisted Personal Interview	**CATI** Computer Assisted Telephone Interview	**CSAQ** Computerized Self Administered Questionnaire

Surveys with varying degrees of standardisation

- Narrative interviews

  - Theme is set globally
  - Questions are not fixed, but are developed developed during the survey
  - Order of questions is open

- Semi-standardised surveys

  - Interviewer guide available
  - Order of questions and wording not determined
  - Use of many open questions

- Standardised surveys

  - Formulation is fixed
  - Sequence is fixed
  - Response options are fixed

"qualitative"

"quantitative"

Number of participants in the survey

Individual survey

- Each person is interviewed individually
- Influence of sources of interference on the measurement results: Telephone, family members, co-workers, etc.

Oral group interviews/group discussion

- Several people are interviewed at the same time
- Influencing measurement results through peer pressure or opinion leaders:

Written group surveys

- Several people receive a questionnaire and fill it out (usually under supervision) on site (e.g. school class surveys, survey in a cinema).
- Advantages: situation can be controlled (with supervision), response rate usually high, queries are possible; disadvantage: individual disruptions can affect the entire survey process

Development of survey types
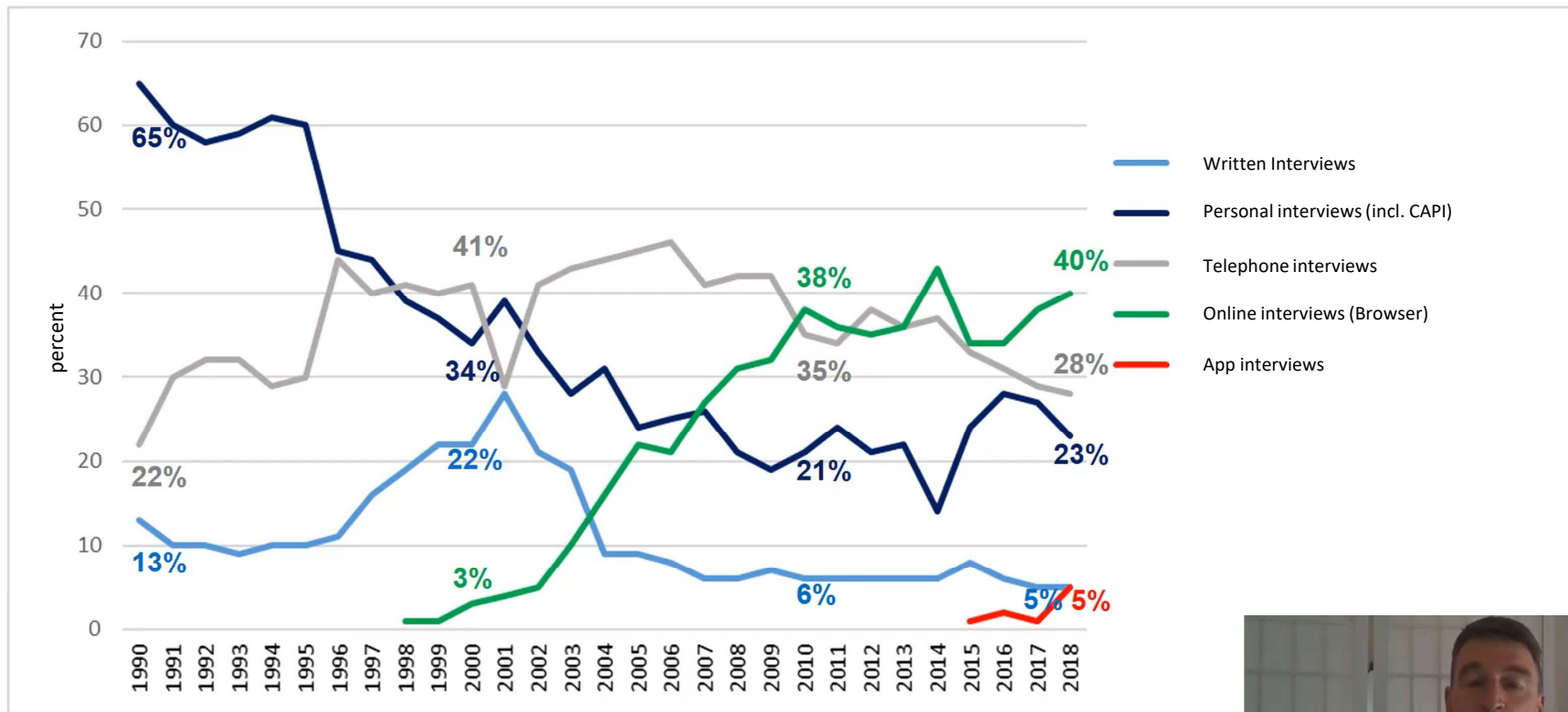
Quelle: https://www.adm-ev.de/die-branche/mafo-zahlen/

Standardised personal interview (1)

- Standardisation ensures extensive comparability of responses

- Interviewer records responses in the presence of the respondent

- In addition to questions, instructions to the interviewer also possible

- Templates are possible

- Implementation via

    - PAPI

    - Trays

    - Mobile phones

    - Notebooks

    - Mobile stations

Standardised personal interview (2)

- Much more time-consuming and expensive than written surveys

- Possible to clarify ambiguities, ask follow-up questions and obtain more in-depth results (e.g. on interests, motives, expectations, associations)

- Motivation to participate is higher than in written surveys

- Influence of the interviewer/situation on the interviewee and vice versa (Influences include language, appearance, age, gender and clothing of the participants, response shifts according to the gender of the interviewer and interviewee have been proven)

- Prejudices on the part of the interviewer towards the interviewees are also possible, which influence questions and writings down

Standardised personal interview (3)

Important for oral interviews:

■ Use of trained interviewers (required competences are e.g. friendliness and politeness, empathy, serious appearance, reliability, flexibility, quick comprehension, clear language, exact knowledge of the topic, questionnaire and interview rules, etc.).

■ Clear rules for different situations (e.g. addressing the interviewee, providing information, conducting the interview etc.).

■ Interviewer may appear to be an employee of the institution under study, as this may influence the answers (e.g. respondents may not be as honest about their satisfaction with the institution).

■ Use of several interviewers is to reduce interviewer effects

■ If possible, no third parties should be present during interviews (e.g. accompanying persons can be provided with programme booklets, catalogues, a free drink in the café etc.)

Standardised personal interview (2)

Advantages:

- Use even with complex questions that the interviewee could not answer at first or without the help of the interviewer
- Respondent can be claimed for longer than in a telephone interview
- Experience shows that the limit is 45 to 60 minutes (in-house interviews).
- Visit of the interviewer is usually announced (=> less stress)
- Personal interview more varied and less strenuous
- Visualisation techniques (e.g. product samples), lists with answer specifications possible

Disadvantages

- Accessibility of the test persons partly very difficult
- Costs usually significantly higher than with other methods
- Interviewer effects

Phone interviews (28% market share)

Interviewers conduct interviews from telephone studios with the support of on-screen interview masks.

Quelle: Koch/Gebhardt/Riedmüller 2016: Marktforschung, S. 53

Telephone survey (1)

CATI = Computer Assisted Telephone Interviewing

- Computer establishes connection to the test person (number generation, dialling, repeating the call etc.)

- Questions of the questionnaire are read from the screen and answers are directly entered into the computer

- Interim evaluation is possible at any time

- Automatic number generation possible (fixed network and mobile phone)

- Share of fixed lines still high (over 80% of HH in 2018 in Germany)

- But: Investment in hardware and software required

Telephone survey (2)

Advantages

- Contact with the respondent only via the voice of the interviewer (anonymity)
- Queries possible in case of comprehension problems
- Comparably low costs
- Interviewer can help and motivate

Disadvantages:

- Limited length = approx. 20 minutes max.
- No visualisation aids possible
- Difficulty in reaching some groups of people (answering machines, high mobility, etc.)
- Large number of persons without an entry in the telephone book (if applicable, generation of telephone numbers or use of mobile numbers)

Oral interviews by phone

■ Take into account respondents' reservations, low participation (especially if the client of the study is not known).

■ Difficult to reach, no complete sample (mainly because of mobile phones, missing entries in telephone directories).

■ Recommended: Use with existing telephone numbers or computer generated numbers; if telephone interview has been announced in advance and a suitable time has been agreed.

Examples of telephone interviews:

• Telephone interviews with subscribers,

• Workshop participants,

• Members of the Circle of Friends etc.

Standardised written survey (1)

(as online survey, but also f-t-f or by post)

Written surveys

Self-completion written questionnaires are the most frequently used survey instruments
Comparatively inexpensive and time-saving.
Influence of an interviewer on the respondent is eliminated

To be noted for written surveys:
- Questionaire must be self-explanatory (e.g. clear terms, references to filter questions and multiple answers, clear instructions on how to fill in the form etc.)
- Compact, clear and appealing design of the questionnaire
- Sufficient font size and legibility
- Reference to guarantee of anonymity
- Depending on the target groups of the survey, several variants, e.g. simpler questionnaires for younger respondents.
- If required, in several languages (note: correct translation e.g. of educational qualifications, school grades).

Standardised written survey (2)
(as online survey, but also f-t-f or by post)

Communication is limited to the questionnaire in written form (digital or haptic)

Careful questionnaire design urgently needed:

- Clear filter guides
- Clear questions and answer specifications
- Good readability (font size!)
- Comprehensibility for all subjects
- Layout adapted to target group

Standardised written survey (3)

Careful wording of the cover letter/introductory text:

- Why was this particular recipient selected (all in target group, random selection, special competence, ...)?

- Where does the address come from (customer file, all companies in XY sector, random selection from telephone book, ...)?

- What is the aim of the research? E.g. scientific interest as in Thesis, improvement of the product offer for customers, analysis of the market potential for a new product, ...

- How much time does it take to complete the questionnaire?

- Is there anything in return for completing it (gift, results, ...)?

Standardised written survey (4)

- Who commissioned the study (university, company, market research institute, ...)?

- What happens to the data (anonymous use according to data protection rules (DSGVO), storage, forwarding, ...)

- To whom and how (e-mail, letter, if necessary in a return envelope, ...) should the questionnaire be returned?

- Who can the recipient contact if there are still questions (contact person with telephone number)?

Advantages and disadvantages of written surveys (5)

Advantages:

- Interviewer effect not applicable

- Good for use in highly standardised surveys with a large number of respondents (requires experience and/or pretests)

- Significantly lower than the costs (only printing costs, postage costs, possibly return postage) for personal and telephone interviews, especially for online surveys.

- Quick implementation (online)

Disadvantages:

- No comprehension questions possible

- Survey situation (except for supervised group surveys) not controllable

- Response rates usually low: Between 1% and 50% depending on the topic, target group and activation technique

- Risk of occurrence of a systematic error due to biased response (if response no longer represents a random sample of questionnaire recipients) => representativeness endangered!

Written questionnaires that are distributed/displayed (6)

- Important rule: Never (!) just lay out questionnaires, as no representative picture of the visitor structure is possible (i.e. distortion by self-selection of respondents)

- Rather actively approach visitors and distribute questionnaires; important to reach a significant number of respondents.

- Take into account the time of completion: questions on judgements are influenced by the previously seen performance (*"Halo- resp. "Priming-Effect"*)

- Important: Supply writing and seating facilities, return boxes for anonymous drop-off.

- Small gifts for the respondents significantly increase the motivation to participate

- If digital questionnaires, e.g. as tablet stations or distribution via QR-Codes:
  The suitability for the target group must be considered; drop out-rate may be higher

- If interviews conducted by a station, it should be supervised by selecting and addressing the respondents and providing them with assistance).

Written questionnaires sent by post (7)

- Addresses must be available
- Possibility to reach population/non-visitors also outside the facility
- Larger scope of the questionnaire possible
- Personal letter with information about the survey and participation requirements, if necessary also a reminder letter at a later date
- Free returns important
- Recommended: Incentives through gifts or (record contact details separately)
- Insert in bulletin to all households (if no representative sample is to be addressed)

Some examples:

- A survey of a city's offerings,
- A survey of a festival´s subscribers,
- A survey of ticket purchasers

Written questionnaires placed online (8)

- Must be less elaborate and short
- Check the target group of the study, especially consider differences in age; also consider the accessibility of the target group.
- Important: Clarity of the questionnaire, progress indicator
- No overloading with multimedia elements, few technical requirements
- Check functionality with different internet browsers and end devices (mobile view)
- Control options if required: Passwords, personalised one-time links

Examples of written questionnaires posted online:
- Survey via pop-up questionnaire among users of a homepage,
- survey of newsletter subscribers,
- survey of facebook / Instagram followers of the festival´s account,
- survey of members of the Circle of Friends via e-mail...

Be aware of severe problems with representativity!

Forms CSQ (Computerised Self Administered Questionaires)

Pool of notebooks, tablets, mobile phones or stations

- Briefing of the test persons by supervisors

- Subjects enter answers themselves

Web survey

- E-mail with attachment

- E-mail with link to website/host
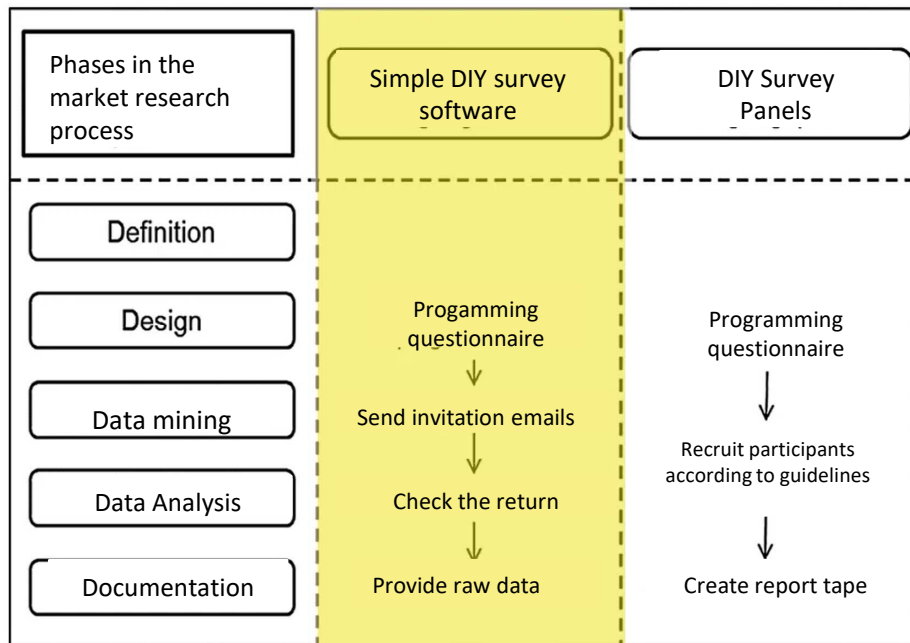
- On the website itself

- In the social network

Prerequisites:

- Access to the computer (still somewhat different due to age)

- Ensuring representativeness

Online-interviews (40% market share)

16% of respondents for online surveys are recruited from customer lists. Anonymous online surveys ("riversampling") are the exception (6%). Online access pools, from which samples are drawn and surveys initiated, are used for the most part (76%).

Phases in the market research process	Simple DIY survey software	DIY Survey Panels
Definition		
Design	Progamming questionnaire ↓	Programming questionnaire ↓
Data mining	Send invitation emails ↓	Recruit participants according to guidelines ↓
Data Analysis	Check the return ↓	
Documentation	Provide raw data	Create report tape

Quelle: Koch/Gebhardt/Riedmüller 2016: Marktforschung, S. 143

Adventages and disadvantages of online-Interviews by Access-Pools)

	Advantages	Disadvantages
Cover	➢ Younger (female) subjects easily accessible	- Older and peak target groups hardly reachable
Process	➢ fast execution ➢ Easy integration of image and video material	- Concerns about data security (sensitive issues)
Response	➢ Answers available at very short notice ➢ Dynamic plausibility checks	- Incentive systems lead to scale bias (social desirability) - Risk of chatbot manipulation
costs	➢ very favorable implementation costs	

Quelle: Koch/Gebhardt/Riedmüller 2016: Marktforschung, S. 53

Survey apps (5% market share)

Participants in a mobile panel download a special survey app onto their smartphone or tablet and can thus take part in surveys designed specifically for mobile devices.
In addition to panel set-up and panel maintenance, panel entertainment also plays a major role in order not to become dependent on extrinsic motivation
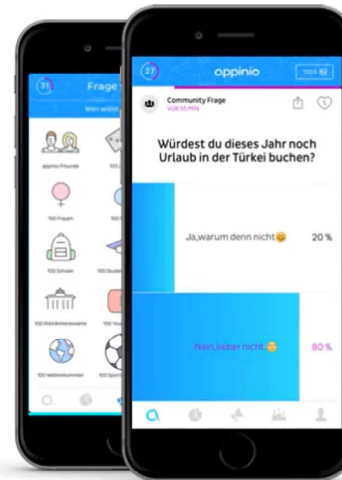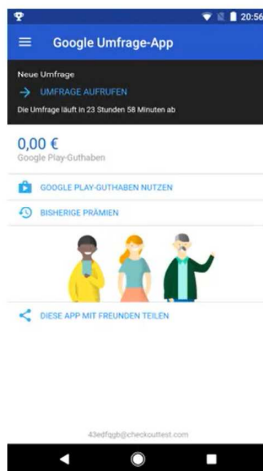
Quelle: Altobelli 2017: Marktforschung, S. 65

Other methods that can be implemented quickly (I)

Information on the audience can also be collected with less complex methods. These can be particularly useful for directly ascertaining the visitors' opinions.

- Useful to get direct feedback from visitors or to determine atmospheric images
- But usually less reliable results

Sources:

- Analyses of data from the ticket system, booking statistics or subscriber data
- Evaluations of existing data from former studies of the institution
- Staff that has visitor contact (for brief interview/observation)
- Visitors' books
- Noteboard wall/blackboard/stickers
- Mystery test/test visitors
- Communication on social media channels (e.g. by programmable API interfaces)

Other Methods that can be implemented quickly (II)

The staff who is in direct contact with the visitors can be included systematically:

- Regular query of brief information at the cash desk *("How did you find out about this event?", "Where do you live?" etc.)*

- Staff at the box office, cloakroom or supervisors can systematically observe visitors with regard to certain aspects using a short observation sheet *("Do the visitors find their way? Are the toilets found immediately? etc.)"*

- Visitor books: Entries should be systematically checked and, if necessary, reactions of the institutions can be made visible there

Simple ways of asking visitors about their satisfaction:

- Note boards at the end of a visit to an event; opinions can be expressed which are also seen by other visitors

- Even faster judgements can be recorded by the audience affixing small stickers on a grading scale on the wall. Here, too, a regular check of the answers is recommended.

RED
Re-Educating
Digitisation

Rheinische
Hochschule
Köln

	Written individual survey	Individual oral survey
Des-cription	• The visitors fill out a questionnaire before, during or after their visit in paper form, on a tablet or at a terminal in the museum. • The questions are usually closed, but there may be a few open questions.	• Visitors are interviewed by an interviewer before, during or after their visit by means of a questionnaire. • Closed questions, but also a small number of open questions are possible.
Fits best when…	• ... you want to find out something about your visitor structure/behavior/attitudes with little effort and at low cost • …you are looking for answers to specific questions	• ... you want to find out about your respondent structure or, if necessary, you are looking for more differentiated answers to specific questions. • …you want to have as much control as possible over the survey situation.
Stren-ghts	• Large number of participants possible • When the questionnaires are handed out: Selection of the participants can be controlled • Generates reliable data • Low costs for interviewer • No influence of the interviewer	• Large number of participants possible, but small samples can also be helpful • Selection of participants can be controlled • Generates very reliable data • Enquiries from both sides possible • Good control of the survey situation
Weak-nesses	• Self-selection of the participants may occur • No possibility for demands • Questionaires must be rather short • No control of the survey situation (e.g. incomplete questionnaires, use of filters)	• Predetermined answers bring only limited information • Depending on the number of participants & duration of the study, many interviewers necessary • Interference of the participants by the interviewer

Types of research designs

- Do not confuse research designs with the data collection techniques previously presented!

- Within the research designs, all data collection techniques can be used, e.g. observational experiment or survey panel, or different methods can be combined with each other

- Research designs (discussed here):
    - Case study
    - Cross-sectional or broad-based study
    - Panel/Trend
    - Qualitative research

Cross-sectional study

- Several objects of investigation (persons, companies, etc.) are examined at one point in time

- Most frequently represented in market research

Results of cross-sectional studies:

- Descriptive statements

- Trend statements (the more-the more statements)

- Causal statements (as with experiments) only possible with the use of assumptions

- Information content is usually less deep than in case studies (depending on the method used)

- Representation deadline permissible if applicable

Trend studies

Trend studies:
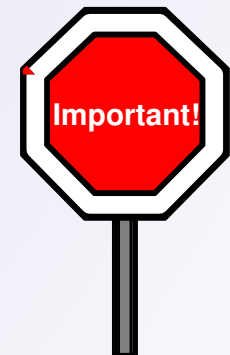- Different respondents
- On the same topics, same questions
- Investigated at different time points

Data collection techniques:
- Telephone, written or personal interview
- Observation (observations at the point of sale)

Danger of confusion:
- Surveys that are carried out in several waves but with the *same* people each time -> Panel
- So-called „Omnibus studies", which are carried out regularly but with *changing* persons and topics

Important!
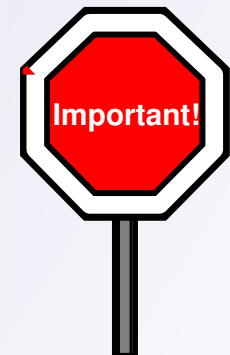
Panel studies

At the panel:

- The same respondents
- On the same topics, same questions
- Investigated at different time points

Data collection techniques at the panel:

- Telephone, written or personal interview
- Observation (GFK television research)
- Problem of keeping panels stable for a longer period of time ("panel maintenance", e.g. GFK television research panel).
- Methodological distortions due to the panel survey
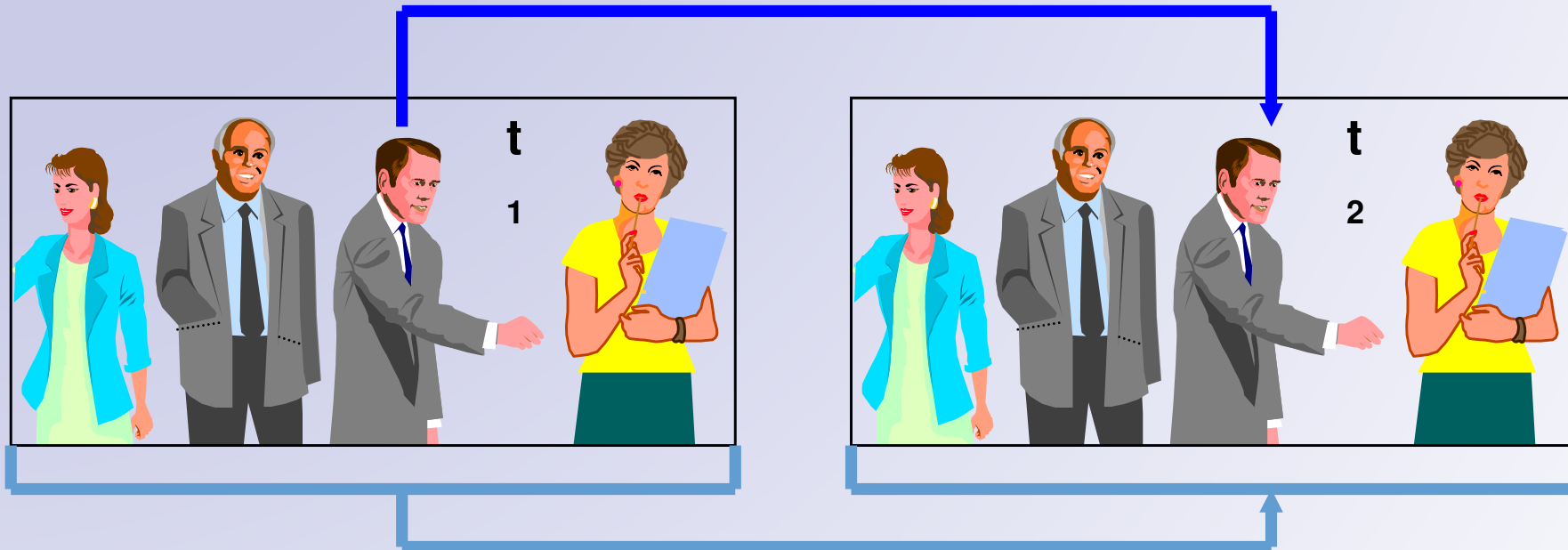
Attention - Danger of confusion:

- Surveys conducted with the same samples but changing topics (bus – studies or online access panels)

Important!

Example „Going to the film festival in October and December".

	Survey in October	Survey in December	December		
Will go	870	900	850	Will go	
			20	will not go	
Will not go	130	100	50	Will go	
			80	will not go	
Total	1000	1000	900	Will go	
			100	will not go	

Interindividual changes

Intraindividual changes

Interpretations

- Interpretation 1:

  All visitors in december also had an intention to go in october, and in addition 30 non-visitors from october could be "convinced". The 100 non-visitors in december are exactly the same 100 people who did not want to go in October.

- Interpretation 2:

  50 people have changed their mind: 20 willing visitors from october do not go after all and 50 unwilling visitors from october do go after all.

Observations (I)

Certain information is not sufficiently captured by surveys. This concerns, among other things, the actual behaviour of visitors at an event or in an exhibition.

Example targets:
- Visitor behaviour in bigger venues: time spent in front of reception, signaling, walking routes through the building etc.

Forms of observation: Overt or covert observation?
open= observed persons are informed about the study in advance
covert= observed persons are not informed about the study and are uninfluenced; it is recommended to inform them afterwards (e.g. via notice board, subsequent interview)

Stationary= observation at a "station", where several people are observed; e.g. in a cinema;
Course= Observation of the course of a visit/ tour etc.
Observations are more likely to be complex and demanding if comparable results are to be obtained.
Method can yield results that surveys cannot, such as the exact time and behaviour spent in certain areas

Observations (II)

■ Use of trained observers (required competences are e.g. systematic and precise approach, flexibility, quick comprehension, required exact knowledge of the observation protocol and the conditions on site etc.).

■ Systematic selection of the observed visitors according to random principle or quota specifications

■ Systematic recording of the aspects to be observed in protocols: previously defined, uniform, comparable

■ After each observation, the protocol should be gone through again in detail and supplemented before the next observation takes place.

■ Use of digital possibilities, such as GPS devices, navigation devices, Eye Tracking, (example: eMotion research project with tracking technology, measurement of heart rate and skin conductance, see www.mapping-museum-experience.com)

■ Observers should not influence visitor behaviour

■ In the case of covert observation, it should be determined how to respond to queries from the observed persons, if necessary.

■ Danger of misinterpretation (e.g. staying in front of a text for a long time does not necessarily mean that it was read completely out of interest).

■ Combination with surveys recommended (→ see example observation sheet)

RED
Re-Educating
Digitisation

Rheinische
Hochschule
Köln

Example of a standardized observation protocoll

Observation sheet no: _____   **Name Observer:** _____

Date: _____   **Time:** _____

Visitor volume: ☐ little   ☐ normal   ☐ mu ch

Data on the observed person:

☐ male   ☐ female

☐ alone   ☐ accompanied b y _____ People

Age (estimated): _____

☐ Audio guide   ☐ No audio guide

Entry time in room: _____

Clock Reading time Text 1: _____   ☐ Not read

Reading time text 2: _____   ☐ Not read

Dealing with interactive element:

☐ looks at element   ☐ readsinstructions

☐ tries unsystematically   ☐ triessystematically

☐ other _____

Notes: _____

_____

_____

Exit time from room: _____
Clock

	Loud thinking	Information from stationary observation
Des-cription	• Visitors are asked to verbalise their experience during their tour in the festival building, or on a website or other visual material	• Observe the behaviour of visitors at a particular point, object, media station, text etc. • You can also measure the length of stay • Technical aids can be used for this purpose
Fits best when…	• … you can follow the thought processes live during your visit • …you want to receive feedback on the visit experience as directly as possible	• … you would like to learn more about the attraction and holding power of certain areas • Pursue specific questions about use (how do people use the display, ticket machine etc.?)
Stren-ghts	• Visitor's thoughts can be recorded largely unfiltered • Selection of the participants can be controlled, control of the survey situation • Can be combined with technical devices (e.g. eye-tracking systems)	• Records actual behaviour of the participants • Large-scale studies possible, but small samples can also be helpful • Selection of the respondents can be controlled
Weak-nesses	• Requires willingness & self-confidence of the participants to express their impressions directly • Can be costly (with many respndents and technical procedures)	• Can disturb visitors • Does not capture the thoughts and evaluations of the participants • Can be costly (with many respondents and technical procedures)

Other methods that can be implemented quickly (I)

Information on the audience can also be collected with less complex methods. These can be particularly useful for directly ascertaining the visitors' opinions.

- Useful to get direct feedback from visitors or to determine atmospheric images
- But usually less reliable results

Sources:

- Analyses of data from the ticket system, booking statistics or subscriber data
- Evaluations of existing data from former studies of the institution
- Staff that has visitor contact (for brief interview/observation)
- Visitors' books
- Noteboard wall/blackboard/stickers
- Mystery test/test visitors
- Communication on social media channels (e.g. by programmable API interfaces)

Other Methods that can be implemented quickly (II)

The staff who is in direct contact with the visitors can be included systematically:

- Regular query of brief information at the cash desk *("How did you find out about this event?", "Where do you live?" etc.)*
- Staff at the box office, cloakroom or supervisors can systematically observe visitors with regard to certain aspects using a short observation sheet *("Do the visitors find their way? Are the toilets found immediately? etc.)"*
- Visitor books: Entries should be systematically checked and, if necessary, reactions of the institutions can be made visible there

Simple ways of asking visitors about their satisfaction:

- Note boards at the end of a visit to an event; opinions can be expressed which are also seen by other visitors
- Even faster judgements can be recorded by the audience affixing small stickers on a grading scale on the wall. Here, too, a regular check of the answers is recommended.

	Sticky note wall survey	Content analysis of guest books
Des-cription	• Visitors are asked to leave stickers or cards on a wall with comments, remarks, hints on a particular question.	• Entries in guest books, where visitors can spontaneously & freely leave their thoughts following a visit • Can be systematically evaluated
Fits best when…	• … you want to get a general feel for how your offers are affecting visitors, or if you want to identify specific problems.	• … you like to get an impression of how your offers are perceived by visitors • Identify specific problems without collecting data specifically for this purpose
Stren-ghts	• Simple & cheap • Low inhibition threshold because comments from other visitors are visible • Can bring out new, unintended aspects • Interest in visitor opinions visible in the festival rooms	• No effort for the survey • Easy & cheap to evaluate • Can bring out new aspects, as visitors can express themselves freely
Weak-nesses	• Strong self-selection of the participants • If necessary, especially positive or especially negative feedbacks • Provides limited information, as usually only one input question is asked • No possibility of follow-up questions for both sides • Link to further information on the person not possible	• Very strong self-selection of participants • Many general entries • Positive feedback is more frequent than average • No possibility of questions • Data cannot be linked to further information about the visitors

	Content analysis from ratings on the internet	**Evaluation of information from cash & booking systems**
Des-cription	• Entries by visitors on online rating portals and in social media are systematically evaluated	• Data collected automatically during ticket sales & bookings are systematically analysed (e.g. number of discounted tickets, tickets for public, industrials, journalists, tickets for certain offers, tickets sold at certain times/days
Fits best when...	• ... ... you would like to get an impression of how your offers are perceived by visitors or identify specific problems without collecting data specifically for this purpose.	• ...you would like to gain an initial overview of your visitors and their usage behaviour without having to collect data for this purpose.
Stren-ghts	• No effort for the survey • Simple & cheap • Brings light concrete problems that can be quickly solved • It is possible to react to evaluations or to ask for further information	• No effort for the survey • Easy to analyse u • Can be recorded over longer periods of time • Developments can be tracked u almost all visitors are recorded • No self-selection
Weak-nesses	• Strong self-selection of the participants • Many general entries • Especially positive or negative feedbacks • Data cannot be linked to further information about the visitors	• Only provide statistical key data • Data cannot be linked to further information about the visitors

Qualitative Research Methods

Research is generally based on *qualitative* or *quantitative* methods

➤ **Quantitative', better:** research with standardised procedures: Procedure planned from the beginning, survey instruments including response specifications precisely defined.

➤ **Qualitative', better:** interpretative-reconstructive, explorative, less standardised research: procedure in the research process flexible; above all: data collection either with open question and answer formats or in the form of (as far as possible) direct protocols.

RED Re-Educating Digitisation

Rheinische Hochschule Köln

Examples of different research designs:

"Interpretative research methods":

- This method is based on the fundamental idea that the meaning of actions, persons, situations and relationships is produced in the symbolically mediated process of communication.

- Perceptions and interpretations arise and change on the basis of individual knowledge, values, emotions, experiences etc. Insofar as these are individually different, the perception of the individual also differs.

- For research, this means that understanding is only possible from the perspective of the actors. This requires procedures that are as open as possible.

➢ Conclusion: Research must capture the individual active perceptions and constructions of reality in the consciousness of the actors. This requires open, qualitative methods.

Definition:

"Qualitative research is the attempt to reflect social contexts and views of these contexts in terms of concepts, behaviour, perceptions and narratives of the people who live in them." (Mayering)

Central aspects of <u>qualitative</u> research

- The perspective
    - is that of the person interviewed/observed
    - is dynamic instead of static
    - is multi-layered; there is not only one "truth".
    - is in context; the individual is socially embedded
- Research designs are
    - Mostly natural, i.e. in the normal environment
    - Flexible and reactive to the situation
- Research questions
    - The "why?" and "how?" are examined, not the "how many?" or "how strong?"

Central aspects of qualitative research

- Data and data collection:

  The data are rich in in depth information, but concern only a few units of study

- Intentionally selective sampling

- Reactive and interactive; flexible methods that respond to the context and the individual.

- Analysis and interpretation:

  · Deep understanding of the interrelationships

  · Respect for the uniqueness of cases

  · The aim is to show the diversity and range of the interrelationships

  · Theory and hypothesis building

Comparison of the methods

Qualitative methods

- **Openness,"Sensitizing concepts"**
- **Building „types", obtained from individual cases,**
- **Small samples according to theoretical sampling**
- **Exploration of new or complex phenomena**
- **Capturing deep interrelationships**

Quantitative methods

- **Very clear questions and hypotheses**
- **Terms clearly defined**
- **Inference from sample to population → large random samples**
- **Measurement of the same characteristics across all units**

Comparison of the methods

Qualitative methods

- Participant observation
- In-depth interviews
- Focus groups
- Qualitative content analysis

Quantitative methods

- Survey research and similar data collection
- Quantitative experiments
- Quantitative content analysis

Which methods do we use when?

Choice of methods

Before you choose a qualitative or quantitative method, you need to be aware of the nature of your research interest

What do I want to know?

Does the research interest have a precise question, hypotheses, well defined and achievable population?

Are the results intended to make a statement about this population? Are questions such as "How many?" or "To what extent?" in the foreground?

Then *quantitative methods are* the right choice.

Is hypothesis building the goal of the research? Is little known about the visitors and their social context?

Should reasons, motifs and attitudes for behaviour be explored in more depth? Are questions like "Why?" and "How?" in the foreground?

Then *qualitative methods are* the right choice.

Application of qualitative research

- Vaguely defined or "new" research fields

- Little prior knowledge in established research fields (experience of 3-d movies, use and experience of augmentend reality)

- Complex processes: Behaviour, motivations, decisions, motifs

- Sensitive topics

- Difficult study units (e.g. children, youths)

- Difficult to reach population, few cases (experts, industrials. executives)

- Explorative research in cooperation with quantitative research
  e.g. on the definition of research questions, on question comprehension in surveys, etc.)

Examples of the differences between qualitative and quantitative research ... Examples

a. TV consumption: Differences between watching linear TV and Netflix, reasons for choice

b. Hurdles to going to the movies, museum, theatre

c. Reaseons for watching a certain series

Question:

➤ What would be reasons for choosing a quantitative or qualitative method?

➤ What different kind of insights do you expect?

Qualitative oral interviews

Face-to-face oral interviews are also possible in a qualitative approach

Qualitative interviews:

■ Open and flexible approach: Use of a little/non-standardised guideline, formulation and order of questions not finally determined, open questions.

■ Interviews are recorded and transcribed

■ Case studies, rather small number of respondents

■ Common form: Problem-centred interview, expert interview

Examples of qualitative interviews: Problem-centred interviews with selected respondents on the role of cultural visits in the life course; expert interviews of industrials, of cultural actors or of cultural journalists.

Group discussion as a special form of oral interview

In group discussions, experts or representatives of different target groups exchange views on an object of investigation.
Critical discourse with different points of view and the interaction of the participants.
The aim is to reveal opinion-forming processes and attitudes.
Are particularly useful in less developed subject areas.
They are also increasingly used in cultural development planning as a form of citizen participation.

Note for group discussions:
- Group size usually 5-12 people
- Trained facilitator / moderator (required competences are e.g. knowledge of the topic under investigation, objectivity, flexibility; ability to initiate, maintain, steer the discussion).
- Use of leading questions and stimuli
- Recording and transcription
- Numerous different formats, especially for larger groups (e.g.

"World Café" with changing small groups at several stations to answer specific questions or "Fishbowl" with a discussion in an inner circle and other participants sitting around the circle.)

Examples of group discussions: Group discussion with film professionals / industrials, journalists and cultural politicians on the planning of a new festival; group discussion with citizens on attitudes towards cultural offers.

	In-depth interview	Focus group discussion
Des-cription	• Selected visitors are interviewed individually in detail. • The questions are always open-ended.	• Selected visitors are interviewed orally in groups • There is usually a flexible guide for the interview • The questions are essentially open
Fits best when…	• … you want a deeper picture, e.g. the personal attitudes, evaluations, backgrounds, wishes, etc. of your visitors. • …you may want to find out more about your visitors or seek answers to specific questions.	• … you want to gain more in-depth information about the perspectives of a specific target group on specific issues • …you want to learn about the pros and cons
Stren-ghts	• Flexible conversation possible • can bring differentiated, diverse & unintended information • Selection of participants can be controlled Questions from both sides possible • Control of the survey situation	• Flexible conversation possible • Can bring diverse & unintended information • Group dynamics does stimulate the discussion Selection of participants can be controlled
Weak-nesses	• May be difficult to motivate visitors for a longer interview. • Can be time consuming and demanding. • If necessary, interference of the participants by the interviewer	• Can be time-consuming and demanding • If necessary, mutual influence of the participants • Participants may not dare not to speak in groups • Group may exert a pressure of opinion on certain members

Thank you for the attention