BASIC METHODS OF SCIENTIFIC RESARCH

DATA PROCESSING, ANALYSIS, EVALUATION



DATA ANALYSIS

- Collecting data \rightarrow large number of data
- <u>Analysis:</u>
 - Logical analysis \rightarrow drawing conclusions
 - Statistical analysis
 - Purification of data
 - Arranging data into datasets, tables
 - Qualitative data numerical value (0, 1); establishing categories
 - Quantitative data
 - graphs
 - Analysis, statistical evaluation



ARRANGING DATA, PREPARATION FOR DATA ANALYSIS

• Checking the procedure and methods of data collection

- Collecting data questionnaire, lab notebook etc.
- Data arrangement:
 - Was everything correctly recorded?
 - Are there any obviously false data?
 - Is all necessary information included?
 - Excluding wrong answers, false data

- Purification of data
 - Are obtained data within the acceptable range?
 - Data that are not correct, should be corrected or left out:
 - Something is missing
 - False according to common sense (duplication, impossible data)
 - Mistake in recording data



ARRANGING DATA, PREPARATION FOR DATA ANALYSIS

Arranging data into rows

- Descriptive datasets (data arranged next to each other)
- Comparative / grouping datasets
 - Data arrangement:
 - qualitative,
 - quantitative,
 - according to time

Depicting data on various scales

• Nominal variable:

There is no difference between data regarding their size or order (e.g. listing places of birth)

Ordinal variable:

Data are arranged in some kind of order

• Interval variable:

Data differ in their values



ARRANGING DATA, PREPARATION FOR DATA ANALYSIS

Preparing datasets and tables

I. Identifying variables and related characters

- Topic is divided into measurable categories
- Characters (<u>attributes</u>, concepts that characterize the variable) belong to each variable
- e.g. Attendance of a class:
- <u>Variable</u>: "frequency of absence"
- <u>attributes:</u> "never"; "once weekly"; "once monthly"; "multiple times in a month"; "does not attend class at all"

- 2. Which scale can be used to depict the dataset derived from a variable
 - order
 - interval
 - proportion etc.
- 3. Coding data points
 - Each piece of information receives a numerical value
 - questionnaire:
 - Questions with closed end:
 - e.g. Answer "A": 01; answer "B": 02
 - Questions with open end: answers of many kind, classify into categories, categories receive a code



• Data analysis: study of data arranged, prepared; evaluation of data

• Qualitative analysis:

- Analysis of characters (not numerical data)
- E.g. observations, interviews "non-numerical evaluation"
- Data are not transformed to numerical values; we describe a phenomenon, feature, compare it with a theory, search for relationships

Quantitative analysis:

• Analysis with numerical data – usually with a statistical test



Descriptive statistics:

- order
- frequency distribution
- variance
- mean

Mathematical statistics:

- Searching for correlations
- Analysis of differences
- Researcher himself
- Help of a statistician



Descriptive statistics:

Arranging data into order, groups:

e.g. Size order, alphabetical order, according to gender

- Frequency distribution:
 - How the elements of the sample are distributed between various data groups?
 - <u>Relative frequency</u>: what is the ratio of a possible result?

- variance
- Mean:
 - In a group: answers given by a person – where are they located in relation to the mean value of the other persons
 - Middle value, extreme values
 - median
 - <u>modus</u>: the most frequent data among the data of the sample; or: the middle value of the group with the highest frequency



Mathematical statistics:

- Study of relationships:
 - Independent variable dependent variable
 - Do variables have an effect on each other? Are they related to each other?

- Chi-square test:
 - Test if the value of a variable depends on the value of the other variable
 - Test for the presence and strength of relationship between 2 nominal or ordinal variable
 - Relationship between 2 variables: correlation coefficient
 - a value between -1 and 1



PRESENTATION OF DATA

- Diagram: two-dimensional presentation of data
 - based on coordinate system
 - not based on coordinate system
 - line diagram, bar diagram,, histogram, pie diagram etc.



PRESENTATION OF DATA

Classification of poisonous plants according to toxic substances



Cyanogenic glycosides
Other glycosides
Phenolics
Plant acids
Polyacetilenes
Alkaloids
Proteins, lectins
Terpenes

Survey in nurseries and kindergartens, Pécs, Hungary