

4. General recommendation

Short presentation of literature review and notes on data collection

MUNI
SCI

4. Task 2

Prepare the first version of the literature review (LR)

- Read carefully notes from this presentation
- Look through recommended sources (internet, library, etc.)
- Design the type and structure of your literature review
- Write at least one paragraph for each section

Prepare short presentation (10 minutes)

Deadline: April 10,

Possible content of the presentation:

- Remind the title of Your thesis and the main aim
- What LR structure did you choose?
- What are the main sources of information?
- Approximately how many articles have you gone through?
- Present one paragraph as an example

MUNI
SCI

4. Data collection for Your diploma thesis

- Geography is a **synthetic science**, in contrast to e.g. Chemistry (=experimental science)
- Primary and secondary data sources
 - Primary – Your fieldwork, measurements etc.
 - Secondary – already existing sources, databases
- Qualitative and quantitative data (Quantitative and qualitative statistics)
- Statistical methods are not the only methodology, however:
 - It is the way how to present Your data (results) in objective way
 - The results of your data processing can be verified
 - The professional public will understand Your data analysis results
- A layman may not understand Your data analysis or may even misinterpret them



MUNI
SCI

4. Data collection for Your diploma thesis

- Primary qualitative and quantitative data collection approaches have their **own rules and recommended practices**

Qualitative vs. Quantitative data.

| Data type | What is it? | Methodology |
|--------------|---|--|
| Quantitative | Information that can be measured and written with numbers. This type of data claims to be credible, scientific and exact. | Surveys, tests, existing databases |
| Qualitative | Information that cannot be measured. It may involve multimedia material or non-textual data. This type of data claims to be detailed, nuanced and contextual. | Observations, interviews, focus groups |

<https://paperpile.com/g/data-for-your-thesis/>

- Garbage in, garbage out



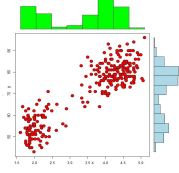
MUNI
SCI

4. Data collection for Your diploma thesis

Exploratory Data Analysis (EDA)

- A summary of descriptive statistics methods that precede the actual statistical processing.
- Its goal is to verify some properties of the input data set, which are necessary prerequisites for the own statistical methods of processing.
- Exploratory analysis focuses on graphical and tabular representation of data

Example: Our task is to find out if there is a relationship between two variables



EDA result: Be careful, your data contains two distinct subsets
(Your data are not normally distributed)

MUNI
SCI

4. Data collection for Your diploma thesis

Exploratory Data Analysis (EDA) will tell you whether:

- Are there any peculiarities in your data?
- Does your data have missing or outliers?
- Does your data meet the prerequisites for using the intended processing methods?
- Does your input data need to be transformed somehow?

MUNI
SCI

4. Notes about data collection

- Be in the permanent contact with Your supervisor
- Be clear about what data you will need
- Check data accessibility (w.r.t. quality, resolution, volume (big data), etc.)
- Check data readability (numerous formats)
- The necessary data processing requires suitable tools
- In particular, spatial data analysis tools (e.g. doing maps) are characterized by their specific data format requirements
- You will use data not only for analysis, but You will also need to present results of analysis (graphs, maps, ...)
- The composition of graphic outputs, such as the name, scale, legend, etc., are recommended to be processed according to knowledge from cartographic lectures and exercises.

MUNI
SCI

4. Task 3

Prepare a short overview about available data sources and about Your data collection.

Prepare short presentation (10 minutes)
Deadline: April 24,

Possible content of the presentation:

- What kind of data will you need?
- Will I have to create my own primary data or can I use existing sources?
- What is the data structure and format?
- Do I know the tools with which I will be able to analyze the data?
- Show us an example of your data

MUNI
SCI