# Projects - Lab 1

Angelica Lo Duca

# 1. Start from a Clear Question

A project starts from a concrete question whose answer we search in data.

EXAMPLE
Do lower speed limits reduce deaths per accident in Europe (2010–2024)?

TASK
Write your own question.

TIPS
- One sentence only.
- Include time frame.
- Include geographic unit.
- Underline measurable elements.

# 2. Translate Question into Variables

Every question implies measurable variables.

**EXAMPLE**
- Independent → Speed limit (km/h)
- Dependent 1 → Number of accidents
- Dependent 2 → Number of deaths
- Control → Population

**TASK**
Identify variables in your question.

**TIP**
- List all variables.
- Classify them (independent / dependent / control).

# 3. Search Multiple Data Sources

One dataset is never enough. Different datasets capture different dimensions.

EXAMPLE
- ISTAT Accidents Victims
- European cities with speed limitis
- Social cost of an accident
- ...

TASK
Find at least 3 datasets.

TIPS
- Use targeted Google queries.
- Use open data portals.
- Save links in a shared document.

# 4. Primary vs Secondary Sources

- Primary = original administrative/raw data.
- Secondary = interpreted or summarized data.

EXAMPLE
- ISTAT → Primary
- EU annual safety report → Secondary

TASK
Classify your datasets.

TIPS
Create a table: Dataset | Primary/Secondary | Why | Limits.

# 4. License

Data are not automatically free to use. Every dataset has a license that defines:

- Who can use it
- Attribution requirements
- Redistribution limits
- Commercial use restrictions

## EXAMPLE
- ISTAT → usually CC-BY (attribution required)
- Government reports → public but may have limits

## TASK

For each dataset:

1. Find the license / terms of use
2. Identify restrictions
3. Add a column in your table: *Dataset | License | Restrictions*

# 6. Raw vs Aggregated Data

- Raw data allow flexible analysis.
- Aggregated data limit granularity.

EXAMPLE
- Accidents → raw data.
- EU fatalities → country-year total (aggregated).

TASK
Evaluate aggregation level.

TIP
- Specify unit of analysis.
- Describe what information is lost due to aggregation.

# 7. Build a Derived Variables

New indicators emerge from combining datasets.

EXAMPLE

Deaths per accident = Deaths / Accidents

Deaths per 100k inhabitants = Deaths / Population

TASK

Identify at least one derived metric.

# 8. Summary

Summary table with all your datasets and what they measure.