

Reflection from working with traffic data - Traffic 1

Steps for responsible data science

- **Utilized Publicly Available Data:** Ensure transparency and reproducibility in our research.
- **Responsible Data Storage:** Sharing it only within our group to maintain confidentiality and comply with data protection guidelines.
- **Acknowledged Data Sources:** Properly cited all data sources to give credit and allow others to trace the origins of the data.
- **Data Cleaning and Documentation:** Recorded all processes used to edit or clean the data, reducing manipulation and enhancing transparency.

What needs more attention?

- **Enhancing Data Representativeness:** Ensure that our data accurately reflects all relevant categories, including various vehicle types, geographic regions, and time periods.
- **Mitigating Data Biases:** Address biases in our data, such as the underrepresentation of motorbikes. This requires actively seeking out additional data sources or adjusting our analysis methods to account for vehicle types that are not well-represented in our dataset.
- **Assessing Societal and Environmental Impact:** Evaluate the potential societal and environmental impacts of our research findings more thoroughly. Reflecting on how our work affects public health, policy decisions, and environmental sustainability is crucial.

Guidelines for responsible design and development

- **Avoid Premature Conclusions:** Do not hastily draw conclusions from the research. Be cautious when sharing findings with stakeholders, especially if the guidance could influence citizens' lives in the future.
- **Ensure Transparency and Reproducibility:** Ensure transparency in data handling by sharing your methods and making your results reproducible. This openness allows others to verify and build upon your work.
- **Obtain Consent Before Sharing Reports:** After the research, do not share the report without obtaining consent from all group members and ensuring compliance with data source agreements.

What would you do differently next time?

- **Expand the Scope of Analysis:** Due to time constraints, we only analysed certain highway parts and generalized our findings to the entire country. Next time, we would include a more comprehensive range of highway sections.
- **Consider Additional Influencing Factors:** Other factors that might influence our research were not thoroughly examined. In future research, we would incorporate these variables, recognizing that neglecting relevant factors can lead to incomplete or biased results.
- **Include Other Sectors Beyond Traffic:** Our study focused solely on the traffic sector. Next time, we would consider additional sectors that contribute to nitrogen pollution. A multidisciplinary approach aligns with responsible data science by offering a holistic view and avoiding oversights that could misinform stakeholders.