

# Grading criteria final project

EPITA 2023 - DUQUESNE

The project is due by 28/01/24 in jupyter notebook format.

The report should include the following sections:

## Introduction (5pts)

- What is the subject and context of the project?
- What is the source of the data?
- Who collected the data ?
- When was the data collected ?
- In what context/for what purpose was the data collected?
- What does the dataset contain? Describe the data, its type (continuous, categorical...)
- What is the size of the dataset?
- What is the dataset's license?

## Research question (2pts)

- Cite your research question
- How might the data you've chosen enable you to answer your question?
- Bonus: What are your hypotheses?

## For each of the 3 sub-questions (9 pts)

- Explain its relevance to your main question.
- What is your methodology for analysis?
- Produce at least one graphical representation (chart) per sub-question:
  - The code is clear and readable
  - The chosen graphical representation is consistent with the type of data and the analysis approach.
  - Graphic representation is complete
  - The graphical representation respects best practices
  - Graphical representation minimizes potential bias
- What analysis can you draw from the graphs? A detailed analysis is expected.
- What are your results?
- What are the limitations?

## **A conclusion (4pts)**

- Does your analysis enable you to answer your research question? Why or why not?
- What are the limitations and biases of your methodology?
- Do you have any sources that support or contradict your conclusions?
- What would be the next steps in your analysis if you had had more time?
- Can you think of another method than data analysis to answer your question? Expand.

Bonus: What did you learn? What would you have done differently if you had to do the project again?

For each part of the notebook I expect commented code and code that respects PEP8 as well as clear markdown text sections with proper english phrasing and minimal grammatical error.