

Formal Requirements: BI-Project

You are required to submit a portfolio of no more than 12 pages that documents your BI project. The structure of your documentation must follow the CRISP-DM structure.

- The 12-page limit applies only to the main text.
- You are encouraged to include supporting plots and tables (which do not count toward the page limit).
- Do not include code or raw code output in the main body of the portfolio. Instead, describe your methods and results.
- You may include your code as a supplemental appendix (e.g. a pdf export of your notebook).
- The final submission must be a single PDF file, uploaded via Blackboard by August 15th, 2025.
- Use standard academic formatting: 11–12 pt font, 1.5 line spacing, and a clear, readable layout.

Further, you will present your project in a (max.) 15 min. presentation.

- Focus on your decisions and learnings!
- The presentations must be submitted on July 15th and will be presented on the 16th and 17th. Attendance on these dates is mandatory.

For your project work:

- You must use Python for your analysis.
- You may use any Python library (e.g., pandas, matplotlib, scikit-learn, seaborn, etc.).
- If you use external libraries that are not part of the standard Python distribution, cite them explicitly in your portfolio.

Your portfolio should follow the CRISP-DM phases, covering the following:

1. **Business Understanding** (~2 pages)
 - Provide context and relevant domain insights
 - Clearly state the goal of your BI project
2. **Data Understanding** (~4 pages)
 - Describe the dataset used
 - Explore and describe features
 - Discuss data quality (e.g., missing values)
 - Highlight patterns or trends
 - Include visualizations to support your data understanding
3. **Data Preparation** (~2 pages)
 - Detail how you selected and filtered your data
 - Explain any transformations or feature engineering steps (including missing data handling)
4. **Modeling** (~2 pages)
 - Describe two different models selected for analysis
 - Describe how each model was implemented
5. **Evaluation** (~1 page)
 - Present evaluation metrics for both models
 - Compare model performance and explain results
6. **Deployment Options & Future Outlook** (~0.5 pages)
 - Discuss how your model could be deployed in a real-world context
 - Suggest possible extensions or improvements
7. **Conclusion** (~0.5 pages)