



## Final Project

**Objective:** The final project is an opportunity for you to apply the data science skills you have learned in this course to a real-world dataset of your choice. You will analyze the data, derive insights, and present your findings in a comprehensive report and presentation.

### Project Proposal (Assignment 3)

Submit a brief proposal outlining:

- The dataset you plan to use.
- The questions you aim to answer.
- The methods and tools you anticipate using.

*You can use the document - INFO212 - Final Project Proposal example as reference.*

It is mandatory that you validate and get the Professors approval of your Project Proposal until the due date (Proposal Due).

*See your Professors as a potential customer that you want to sell your project.*

In the Project Proposal component, you will need to:

- Submit the Project Proposal document (via moodle), with clear project objectives, datasets to use, research questions, Methodology, Tools and deliverables. (70 of 100 marks)
- Project Presentation and validation. (30 of 100 marks)

### Project Requirements

- **Dataset:** Select a dataset that is complex enough to demonstrate your data analysis skills.
- **Tools:** Use Python, or any other data science tools we have covered in the course.
- **Analysis:** Perform exploratory data analysis, statistical analysis, and predictive modeling.
- **Visualization:** Create meaningful visualizations to support your findings.
- **Report:** Write a detailed report explaining your methodology, analysis, and conclusions.
- **Presentation:** Prepare a 5-minute presentation to showcase your project to the class.



## Project Submission

- Report: a PDF document of your report.
- Code: Include your Jupyter Notebook with comments and datasets or links for the dataset that you used.
- Presentation: Submit your presentation slides in PDF or PowerPoint format.

## Evaluation Criteria

- Data Understanding: Clarity in the selection and understanding of the dataset. (10 marks)
- Methodology: Appropriateness of the methods used for analysis. (10 marks)
- Analysis: Depth and accuracy of the analysis performed. (10 marks)
- Visualization: Quality and clarity of the visualizations created. (20 marks)
- Report: Organization, writing quality, and thoroughness of the report. (20 marks)
- Presentation: Effectiveness of the presentation. (30 marks)

## Important Dates

- Proposal Due: 5-Jun-2024 (Assignment 3)
- Final Date Due: 18-Jun-2024
- Presentation Dates: 19-Jun-2024

## Rules

- The total marks are 100.
- Work in group of 4 (max). University Academic Honesty Rules and Procedures will be adhered to strictly.
- Zero tolerance will be given to academic fraud of any sort, including copying content from the internet and/or other assignments.
- Please name your assignment file as INFO212-finalproject-group-number.zip, with the required deliverables.
- Submit your assignment file via moodle until the due date.
  - Assignments will not be accepted beyond that point. Missing work will earn a zero grade.
- Images must be clear and legible. Assignments will be judged on the basis of code, visual appearance, grammatical correctness, and quality of writing, as well as their contents. Please make sure that the text of your assignments is well-structured, using paragraphs, full sentences, and other features of well-written presentation. Text font size should be at least 11 points.

# INFO 212: Data Science Programming I



- Presentation will follow the group numbers order.

## Tips for Success

- Start early and manage your time effectively.
- Keep your code organized and well-commented.
- Practice your presentation multiple times.
- Be prepared to answer questions about your work.

## Questions?

If you have any questions or need guidance, please feel free to reach out your professors during office hours or via email. However, questions must be sent up to 72 hours of the deadline set for the submission of the assignment. Questions sent after the 72 hours deadline, will not be answered.