

AI-First Data Mining Project: Rubrics and Schedule (Master Handout)

This is the **master handout** for the AI-First team project. It specifies (1) project requirements, (2) grading rubrics, and (3) the submission schedule.

Related AI-First handouts (read these)

These files are posted alongside this handout on Moodle:

- **AI-First Workflow Guidelines** (policy + “non-negotiables”): `AI_First_Guidelines.md`
- **AI-First Project Workflow** (step-by-step workflow): `AI_First_Workflow_For_Project.md`
- **Student Prompt Library** (prompts for each stage): `AI_First_Prompt_Library_for_Projects.md`
- **Project Idea Generation Guide** (how to generate and pitch ideas):
`AI_First_Generating_Project_Ideas.md`
- **Project Idea Thread Example** (example pitch + peer feedback):
`AI_First_Generating_Project_Ideas_Example.md`

Non-negotiable requirements (AI-First)

1. **Baseline model required.** You must include a baseline and compare your main model(s) against it.
2. **Oral defense required.** Be prepared to explain your work and answer questions during the final exam period.
3. **AI Use Appendix required (final report).** Your final report must include an **AI Use Appendix** that discloses how AI was used and what you validated. Use the template in **AI-First Workflow Guidelines**: `AI_First_Guidelines.md` .

Overview: what each group will do

Group projects allow students to apply data mining techniques to a real-world dataset. Each group will:

1. Identify a dataset relevant to a problem of interest.

2. Formulate a clear research question or hypothesis.
3. Apply appropriate data mining methods to analyze the dataset.
4. Present findings and insights in a **10-minute** presentation.
5. Submit a concise **3-page** report summarizing the project.

Guidelines for the project

1) Dataset selection

- Choose a publicly available dataset or create one (with instructor approval).
- Ensure the dataset contains enough data to justify analysis. You need a sufficient number of features to discover interesting facts.
- Provide a brief description of the dataset in the report and presentation.

2) Problem definition

- Clearly define the research question or hypothesis your project addresses.
- Justify why the problem is important and relevant.

3) Methodology

- Explain the data preprocessing steps (e.g., cleaning, normalization, feature selection).
- Justify the choice of data mining techniques applied (e.g., clustering or classification methods).
- Describe the evaluation metrics used to assess the methods.
- Include a **baseline model** and explain why it is a reasonable baseline.

4) Analysis and results

- Present key findings from your analysis.
- Use appropriate visualizations to support your results (e.g., graphs, tables, charts). Emphasize informative rather than “cool” visualizations.
- Discuss any limitations or challenges faced.

5) Presentation

- Create a **10-minute** presentation to summarize your project.
- Use clear and engaging visuals.

- Allocate time for all group members to contribute to the presentation.

6) Report

- The report should be **3 pages** (excluding references and appendices).
- Structure the report with the following sections:
 - **Introduction:** Problem definition and dataset description.
 - **Methodology:** Data preprocessing and data mining techniques used (including your baseline).
 - **Results and Discussion:** Key findings and their implications.
 - **Conclusion:** Summary and future work.
- Use proper formatting (12-point font, 1-inch margins, single spacing).
- Include an **AI Use Appendix** (required). See: `AI_First_Guidelines.md`

Grading rubric for group (21 points)

1) Dataset and problem definition (3 pts)

- **3:** Dataset is well-chosen, and the problem is clearly and effectively defined.
- **2:** Dataset is relevant but not well-justified; the problem is vague.
- **0–1:** Dataset is inappropriate or missing; the problem is unclear.

2) Methodology (5 pts)

- **5:** Appropriate data mining techniques applied and well-justified; preprocessing is thorough; includes a clear baseline.
- **3–4:** Somewhat appropriate techniques with partial justification; preprocessing is minimal; baseline exists but is weak or under-explained.
- **0–2:** Techniques are inappropriate or unjustified; preprocessing is absent; baseline is missing or not meaningful.

3) Analysis and results (5 pts)

- **5:** Results are insightful, well-presented, and supported by clear visualizations.
- **3–4:** Results are somewhat clear but lack depth or sufficient visualizations.
- **0–2:** Results are unclear or unsupported by analysis.

4) Presentation (4 pts)

- **4:** Presentation is clear, engaging, and well-structured; all group members contribute.
- **2–3:** Presentation is somewhat clear but lacks structure or balanced group participation.
- **0–1:** Presentation is unclear, poorly organized, or lacks group participation.

5) Report (3 pts)

- **3:** Report is concise, well-organized, and follows formatting guidelines (**including required AI Use Appendix**).
- **2:** Report is somewhat organized but exceeds or falls short of page requirements, or misses required elements.
- **0–1:** Report is disorganized or fails to meet guidelines.

6) Draft report (1 pt)

- **1:** A draft report was submitted on time.
- **0:** Draft report missing or late.

Additional notes

- **Group size:** 5 students.
- **AI is encouraged**, but you are responsible for the accuracy of all work submitted.
- **Project deadline:**
 - Presentations: **Final Exam Period**
 - Report submission: **Final Exam Day (by midnight)**
 - Oral defense occurs during the final exam period.

Individual project idea generation guidelines

- Create a descriptive and concise project title. Avoid vague or overly broad titles. Examples:
 - **Good:** “Corn Crop Yields Across the World”
 - **Bad:** “Data Mining Madness”
- Write a concise project pitch (**150–200 words**) using the template in **AI First Generating Project Ideas** (Step-2: Prepare Your Short Pitch).

See also the **AI First Generating Project Ideas Example** document.

- Provide feedback for **3** other project ideas. The template in **AI First Generating Project Ideas** (Step-4: Peer Feedback) is recommended.

See also the **AI First Generating Project Ideas Example** document.

See: `AI_First_Generating_Project_Ideas.md` and
`AI_First_Generating_Project_Ideas_Example.md`

Individual project idea generation grading rubric (3 pts)

- **1:** Project description is comprehensive with some specific details, and the title is informative.
- **1:** Data set is described including how it will be obtained (with links provided) or how it will be generated.
- **1:** Provide substantive feedback on **3** other students' ideas (questions and/or constructive comments).

Guidelines for individual worklogs

- **Purpose:** Each student is required to submit regular, individual worklogs to document their contributions to the project. The worklog ensures accountability and helps track progress.
- **Frequency:** Update the worklog weekly with detailed entries of tasks completed, challenges faced, and plans for the next week.
- **Content requirements:** Each entry should include:
 - **Tasks completed:** List specific contributions for the week.
 - **Challenges:** Note any difficulties encountered and how they were addressed.
 - **Next steps:** Outline planned activities for the next week.
 - **Time spent:** Record approximate hours spent on the project each week. (Expect to work **3 hours per week**.)
- **Format:** Worklogs are entered in Moodle weekly and can be read by all group members.

Grading rubric for individual worklogs (2 pts)

- **2:** Individual tasks completed were listed. Difficulties encountered were described. The next week's planned activities were outlined. Worked at least **3 hours**.
- **1:** Worked less than **3 hours**.

Grading rubric for individual presentation (1 pt)

- **1:** Presentation is clear, engaging, and well-structured.

Peer evaluation rubric for group members (3 points)

Each student will evaluate their group members on the following categories.

1) Group member contributions (1 point)

- **1:** Actively contributes to the project by completing assigned tasks and supporting team efforts.
- **0:** Contribution is minimal or tasks are frequently incomplete.

2) Group member collaborations (1 point)

- **1:** Works well with the group, communicates effectively, and resolves conflicts constructively.
- **0:** Does not collaborate effectively or creates conflict within the group.

3) Group member reliability (1 point)

- **1:** Consistently meets deadlines and follows through on commitments.
- **0:** Frequently misses deadlines or fails to complete assigned tasks.

Schedule

Week	Due date	Max score	Item	Grade type
04	Mon, Jan 05	3 pts	P0: Project Idea Generation	Individual
05	Fri, Jan 16	2 pts	P1: Work Log 1	Individual
06	Fri, Jan 23	2 pts	P2: Work Log 2	Individual
07	Fri, Jan 30	2 pts	P3: Work Log 3	Individual
08	Fri, Feb 06	2 pts	P4: Work Log 4	Individual
09	Fri, Feb 13	2 pts	P5: Work Log 5	Individual

Week	Due date	Max score	Item	Grade type
10	Fri, Feb 20	2 pts	P6: Work Log 6	Individual
10	Fri, Feb 20	1 pt	P9: Draft Report	Group
Final Exam	To be determined	2 pts	P7: Work Log 7	Individual
Final Exam		1 pt	P8: Individual Presentation	Individual
Final Exam		4 pts	P10: Group Presentation	Group
Final Exam		3 pts	P11: Report	Group
Final Exam		13 pts	P12: Project	Group
Final Exam		3 pts	P13: Peer Evaluation	Individual
	Total	42 pts		

AI Use Disclosure: OpenAI's ChatGPT 5.2 was extensively used to brainstorm ideas, create drafts and check for inconsistencies. (Yosi Shibberu)