



AI-First Data Mining Project Workflow

A step-by-step guide to using AI effectively throughout your project.



1. Project Ideation

Use AI to:

- Generate 5–10 project ideas
- Identify relevant public datasets
- Propose research questions & hypotheses
- Evaluate feasibility, scope, and required methods



2. Dataset Acquisition

Use AI to:

- Locate reliable dataset sources
- Summarize dataset structure (rows, columns, types)
- Anticipate potential issues (missingness, imbalance, noise)



3. Exploratory Data Analysis (EDA)

Use AI to:

- Write full EDA code (summary stats, plots, correlations)
- Identify missing data, outliers, anomalies
- Suggest feature engineering ideas



4. Preprocessing & Data Cleaning

Use AI to:

- Recommend appropriate cleaning steps
- Generate preprocessing pipelines (scaling, encoding, imputation)
- Validate data transformations & explain reasoning



5. Model Selection & Baselines

Use AI to:

- Recommend suitable algorithms for your task
- Generate baseline models (e.g., logistic regression, random forest, k-means variants)
- Explain algorithmic tradeoffs and methodological assumptions



6. Model Training & Evaluation

Use AI to:

- Produce training/evaluation code
- Suggest hyperparameter tuning strategies
- Interpret metrics and diagnose weaknesses



7. Visualization & Insight Extraction

Use AI to:

- Improve plot clarity and design
- Propose alternative visualizations that better communicate findings
- Translate statistical results into clear, accessible insights



8. Report Writing

Use AI to:

- Draft each required section:
 - Introduction
 - Methodology
 - Results & Discussion
 - Conclusion
- Refine clarity, conciseness, and narrative flow
- Check alignment with rubric expectations



9. Presentation Creation

Use AI to:

- Convert your report into a slide deck
- Generate speaker notes for each group member
- Simplify visuals for live explanation during your talk



10. AI Self-Audit Before Submission

Use AI to:

- Critique your methodology, assumptions, and conclusions
- Identify gaps, errors, or weak analysis
- Suggest improvements for clarity, correctness, and completeness



Deliverables

- **Project Pitch & 3 Peer Feedbacks**
- **7 Worklogs**

- **Draft 3-page Report**
- **Individual/Group Presentation**
- **Final 3-page Report**
- **Peer Evaluations**

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