Amazing International Airlines Inc.

Data Mining Project Guidelines

Fall Semester
2025-09-08

Project Overview

Project at a Glance

- **Client**: Amazing International Airlines Inc. (AIAI)
- **Objective**: Customer segmentation for personalized marketing
- **Duration**: Full semester project with 3 deliverables
- **Team Size**: 3-4 students per group
- **Total Points**: 100 points (30 + 60 + 10)
- **Methodology**: CRISP-DM framework

Timeline & Deliverables

Deliverable	Components	Points	Focus
EDA	Notebook + Report + Poster	30	Data Understanding
Clustering	Notebook + Report + Video	60	Modeling & Strategy
Discussion	Group discussion	10	Individual assessment

Project Introduction

Amazing International Airlines Inc. (AIAI)

- **Challenge**: Design personalized services and marketing strategies
- **Context**: Highly competitive airline industry
- **Opportunity**: Leverage data-driven customer segmentation
- **Dataset**: 3-year loyalty program and flight activity data

Customer Segmentation Perspectives

- **Value-based segmentation**: Economic contribution grouping
- **Behavioral segmentation**: Purchasing habits and travel behaviors
- **Demographic segmentation**: Age, occupation, and attribute patterns
- **Goal**: Integrate perspectives into comprehensive marketing strategy

Project Structure & Methodology

CRISP-DM Methodology

- **Business Understanding**: Industry challenges and segmentation needs
- **Data Understanding**: Dataset exploration and quality assessment
- **Data Preparation**: Feature engineering and preprocessing
- **Modeling**: Multi-perspective clustering analysis
- **Evaluation**: Validation and business interpretation
- **Deployment**: Strategic recommendations and implementation

Learning Objectives

Students will gain proficiency in:

- Applying unsupervised learning techniques for customer segmentation
- Conducting multi-perspective segmentation analysis
- Translating technical insights into actionable business strategies

Deliverable 1: Exploratory Data Analysis (30 points)

Key Tasks

- Conduct descriptive statistics and visualizations
- Assess data quality issues and clustering reliability
- Identify preliminary behavioral signals
- Develop and justify engineered features

Components Overview

- **Component 1**: Jupyter Notebook (5 points)
  - File: `GroupXX_EDA_Code.ipynb`
  - Clean, documented code with systematic workflow
- **Component 2**: Report (15 points)
  - File: `GroupXX_EDA_Report.pdf`
  - Maximum 5 pages using CRISP-DM framework
- **Component 3**: Infographic Poster (10 points)
  - File: `GroupXX_EDA_Poster.pdf`
  - A3 size, executive-focused communication

Deliverable 2: Clustering Analysis (60 points)

Key Tasks

- Prepare dataset for modeling
- Apply multiple clustering perspectives
- Compare at least two clustering approaches per perspective
- Propose final merged segmentation solution
- Integrate insights into coherent framework

Components Overview

- **Component 1**: Jupyter Notebook (5 points)
  - File: `GroupXX_Clustering_Code.ipynb`
  - Multiple algorithms implemented and compared
- **Component 2**: Report (35 points)
  - File: `GroupXX_Clustering_Report.pdf`
  - Maximum 10 pages with strategic recommendations
- **Component 3**: Video Presentation (20 points)
  - 2-4 minute executive-level presentation
  - Clear segment profiles and actionable strategies

General Policies & Guidelines

Group Composition

- **Maximum**: 4 students per group (recommended: 3)
- **Enrollment**: All students must register on Moodle
- **Deadline**: Group formation before first delivery
- **Changes**: Not recommended after first delivery

AI Tools Usage

- **Permitted**: With full disclosure required
- **Documentation**: AI Usage Statement in annexes
- **Responsibility**: Students accountable for all content
- **Requirement**: Original analysis must exceed AI contribution
- **Explanation**: Students may be asked to explain AI-generated content

Mandatory Annexes

All reports must include:

- **AI Usage Statement**: Specific tools and usage documentation
- **Contribution Statement**: Individual student contributions
- **Responsibility Statement**: Group accountability certification

Assessment & Penalties

Grading Distribution

- **Deliverable 1**: 30 points total
  - Jupyter Notebook: 5 points
  - Report: 15 points
  - Infographic Poster: 10 points
- **Deliverable 2**: 60 points total
  - Jupyter Notebook: 5 points
  - Report: 35 points
  - Video Presentation: 20 points
- **Discussion Component**: 10 points

Penalty System

- **Late Delivery**: 10% penalty per day
- **Guideline Deviations**: Additional deductions may apply
- **Quality Standards**: All components must meet passing standards
- **Plagiarism Check**: All reports undergo plagiarism verification

Bonus Components (Optional)

Bonus Components (Optional)

Deliverable 1 Bonus Options

Up to 20% additional credit

Deliverable 2 Bonus Options

Up to 20% additional credit

Success Strategies

Report

- **Focus**: Insights and interpretations, not just descriptions
- **Structure**: Follow provided frameworks strictly
- **Business Context**: Connect technical findings to strategic value
- **Evidence-Based**: Support all claims with data analysis
- **Executive Communication**: Accessible to non-technical stakeholders

Technical Implementation

- **Code Quality**: Clean, documented, reproducible workflows
- **Multiple Methods**: Compare different clustering approaches
- **Validation**: Use appropriate metrics and statistical testing
- **Integration**: Merge perspectives into coherent solution
- **Business Translation**: Connect technical results to actionable strategies

Questions & Support

Getting Help

- Course materials and documentation
- Office hours and TA support
- Moodle discussion forums
- Group collaboration and peer learning

Final Reminders

- **Originality**: Ensure all work represents group analysis
- **Documentation**: Maintain clear records of all processes
- **Timeline**: Plan deliverables well ahead of deadlines
- **Quality**: Prioritize depth over breadth in analysis
- **Business Value**: Always connect findings to AIAI strategic needs