Team Project Evaluation Form

Purpose: This confidential form helps ensure each member's efforts are assessed fairly. Please complete all sections candidly and individually; your feedback informs both individual grades and project learnings.

---

1. Identifying Information

- **Your Name:** Yue
- **Student ID:** 22014164
- **Group Name/Project Title:** NZ House Price Prediction System

---

2. Self-Assessment of Contribution

2.1 Activity Breakdown (must total 100%)

Category	Your % Contribution	Brief Evidence (one sentence)
Project Planning & Coordination	15%	Coordinated infrastructure decisions and deployment timeline
Research & Literature Review	5%	Researched cloud platforms and feature engineering best practices
Data Collection & Analysis	20%	Built ETL pipeline integrating 4 data sources (house, CPI, GDP, OCR)
Experimentation	10%	Implemented unified training framework for 3 model types
Design & Development	35%	Developed complete data pipeline, feature engineering, and Streamlit app
Writing & Documentation	10%	Created README, technical documentation, and team contribution docs
Presentation & Communication	0%	N/A - to be completed
Team Support & Collaboration	5%	Set up shared Supabase environment enabling team collaboration

Total	**100%**

2.2 Overall Contribution (1=low, 5=high): ☑ 5

2.3 Your Primary Strengths (up to three)

1. **Infrastructure & System Design**: Built production-grade data pipeline and deployment architecture that enabled efficient team collaboration

2. **Full-Stack Development**: Delivered end-to-end solution from ETL to interactive web application with integrated ML models

3. **Problem-Solving & Tool Selection**: Identified Supabase as optimal AWS alternative, overcoming collaboration barriers while maintaining cloud-native architecture

---

3. Peer Assessment

Teammate 1

Name & ID: Xuanhui - 18044666

% Contribution by Category:

- Project Planning & Coordination: 30%
- Research & Literature Review: 25%
- Data Collection & Analysis: 20%
- Experimentation: 25%
- Design & Development: 0%
- Writing & Documentation: 0%
- Presentation & Communication: 0%
- Team Support & Collaboration: 0%

Overall Rating (1–5): ☑ 5

Brief Evidence: Led project initiation, selected topic, identified and integrated external economic indicators (CPI, GDP, OCR), conducted comprehensive EDA justifying feature inclusion, and performed extensive hyperparameter optimization for tree-based models.

Teammate 2

Name & ID: Ming - 22014164

% Contribution by Category:

- Project Planning & Coordination: 0%
- Research & Literature Review: 15%
- Data Collection & Analysis: 0%

- Experimentation: 40%
- Design & Development: 35%
- Writing & Documentation: 10%
- Presentation & Communication: 0%
- Team Support & Collaboration: 0%

Overall Rating (1–5): ☑ 5

Brief Evidence: Specialized in tree-based models, implemented Random Forest and expanded CatBoost analysis, conducted comparative model evaluation, and performed feature importance analysis specific to ensemble methods.

Teammate 3

Name & ID: Zhengyang - _____

% Contribution by Category:

- Project Planning & Coordination: 0%
- Research & Literature Review: 15%
- Data Collection & Analysis: 0%
- Experimentation: 35%
- Design & Development: 35%
- Writing & Documentation: 15%
- Presentation & Communication: 0%
- Team Support & Collaboration: 0%

Overall Rating (1–5): ☑ 5

Brief Evidence: Specialized in linear models, implemented Ridge Regression with cross-validation, explored L1/L2 regularization techniques, conducted feature selection for linear assumptions, and documented model trade-offs.

---

4. Discrepancy Check

Please compare your self-assessment (Section 2) to your peers' evaluations (Section 3). If any teammate's ratings or percentages differ by >20% or >2 points, explain possible reasons:

No significant discrepancies expected. All team members contributed substantially within their specialized domains. The distribution reflects our deliberate division of labor:

- Xuanhui: Project leadership, external data integration, and initial modeling
- Ming & Zhengyang: Specialized model development (tree-based vs. linear)
- Yue: Infrastructure, data engineering, and production deployment

Each role was critical to project success and represented approximately equal overall effort despite different activity distributions.

---

5. Team Dynamics & Reflection

5.1 Collaboration Quality (choose one)

☑ **a.** Tasks were shared equitably; I would gladly work with this team again.

5.2 Key Successes (3 elements)

1. **Clear Division of Labor**: Early agreement on tree-based vs. linear model specialization prevented duplication and ensured comprehensive model coverage

2. **Centralized Feature Store**: The `feature_house` table design eliminated feature inconsistencies, prevented data leakage, and enabled fair model comparisons across all team members

3. **Production Deployment**: Successfully deployed a live application that integrates all team contributions into a cohesive, user-friendly system accessible to non-technical users

5.3 Main Challenges & Mitigations

Challenge: Initial AWS account limitations and complexity threatened to block cloud deployment and team collaboration

How addressed: Pivoted to Supabase as a lightweight AWS-backed alternative, providing equivalent S3 and RDS functionality with simplified team access, enabling rapid development without infrastructure overhead

5.4 Suggestions for Future Projects

- Establish shared infrastructure and data standards early (feature store, naming conventions) before individual model development begins

- Implement CI/CD pipeline with automated deployment to streamline updates and ensure code quality before production

---

6. Final Comments (Optional)

Use this space to highlight any additional observations about contribution, teamwork, or outcomes.

Final comments:

This project demonstrated the value of treating ML system development as a software engineering challenge, not just modeling exercises. By investing in infrastructure (ETL, feature store, unified training framework, deployment), we transformed what could have been scattered notebooks into a production-grade system that:

1. Enables continuous data updates and model retraining
2. Provides interactive exploration for stakeholders
3. Maintains feature consistency across all models
4. Supports easy addition of new models or features

The team worked exceptionally well, with each member taking ownership of their domain while remaining flexible to support others. Our decision to use Supabase over raw AWS services exemplifies pragmatic engineering—choosing tools that maximize team productivity rather than blindly following enterprise patterns.

The live deployment at https://app-test-qxq5b9dukmh7yw6xuyufxc.streamlit.app/ stands as evidence of our collective success in delivering not just models, but a complete, usable system.