Group 2: Feedback (14/09/25) Grade: 74/100

Assessment of Group 2's Airbnb Project Report

Substantive Question (5%) 4/5

Commendable: The business question is clearly framed around identifying the key drivers of Airbnb prices in New York City and developing a predictive model for optimal pricing. This is highly relevant to Airbnb's operations and provides a solid basis for both technical modelling and strategic recommendations.

Needs Further Work: The framing could have been enhanced by explicitly linking price optimisation to wider business challenges (e.g., regulation, competition from hotels, or consumer trust factors).

Rational/Arguments for the Business Question (5%) 3/5

Commendable: The rationale for applying a comparative modelling approach (Linear Regression, Random Forest, XGBoost, with both tuned and untuned versions) was logical and well-justified. The group demonstrated understanding that different models handle skewed, high-dimensional data differently, and tuning hyperparameters is essential to improve generalisation.

Needs Further Work: The rationale section would have benefited from more engagement with existing research on predictive modelling of real estate or tourism markets, to situate their work in a stronger academic context.

Expected Business Impact (5%) 4/5

Commendable: The proposal of a semi-autonomous pricing agent to provide dynamic, model-driven pricing recommendations to Airbnb hosts is innovative and grounded in their findings. The recommendations are clear and actionable, with a strong link to Airbnb's interest in optimising host revenues.

Needs Further Work: While the agent concept is sound, its practical limitations are underexplored. For example, the team could have considered risks such as seasonal events, outliers in demand, or the potential resistance of hosts to automated pricing.

Methodology (40%) 29/40

Knowledge and Understanding (10%) 8/10: The group demonstrated strong technical knowledge, explaining each step of the workflow, including cleaning, feature engineering, one-hot encoding, log transformation of prices, scaling, and cross-validation. They also correctly recognised the importance of addressing skewness and feature scaling.

Application of Knowledge and Understanding (10%) 8/10: The application of both Random Forest and XGBoost, with hyperparameter tuning, was appropriate and well executed. The choice of RMSLE and R^2 as metrics was well explained, and they demonstrated the value of moving beyond a baseline Linear Regression model.

Criticality (20%) 13/20: The group compared models rigorously, noting that the tuned XGBoost model outperformed others ($R^2 = 0.623$ vs. baseline 0.553). They also noted limitations in model reliability at price extremes. However, critical depth was missing

in discussing why specific models outperformed others and in reflecting on feature importance beyond headline numbers.

Needs Further Work: While the methodology is robust, the report did not interrogate the potential overfitting risk of tuned models in detail, nor did it reflect deeply on the interpretability trade-offs between ensemble models and simpler linear models.

Visualisations of Results (40%) 30/40

Commendable: The group used an extensive range of visualisations (price distributions, log transformations, scatter plots, correlation heatmaps, model performance comparisons, feature importance charts, borough-level pricing, prediction accuracy ranges, and executive dashboards). This was a strong element of the report.

Knowledge and Understanding (10%) 8/10: The visualisations show a clear understanding of how to present skewed distributions, feature importance, and comparative model accuracy.

Application of Knowledge and Understanding (10%) 8/10: The figures were well-integrated into the narrative and supported their conclusions effectively.

Structure and Presentation (20%)- 14/20: The presentation was polished, featuring executive-style dashboards and business-oriented charts that enhanced the report's readability for a non-technical audience.

Needs Further Work: Some visualisations (e.g., prediction scatter plots and pricing opportunity matrix) could have included more explicit interpretation in the text, rather than relying on the reader to infer insights. Linking these more clearly to the business recommendations would improve their impact.

Overall Presentation Style (5%) 4/5

Commendable: The report is professionally structured, with an effective executive summary, clear methodology, results, and recommendations. The writing style is accessible and appropriate for both technical and business audiences.

Needs Further Work: At times, the report is overly detailed in technical workflow (e.g., Python code, hyperparameter outputs), which risks losing an executive reader. Greater conciseness in the technical detail would strengthen accessibility.

General Feedback

This is an impressive report that demonstrates excellent technical execution with tuned ensemble models, alongside thoughtful business recommendations. The proposed semi-autonomous pricing agent is a strong example of applying technical findings to a business context. The group's use of visualisation was powerful. To reach higher distinction bands, the report would need deeper critical reflection on model interpretability, limitations, and external market factors.

Final Thoughts

The work is technically advanced, well-presented, and strategically relevant.