MEMBERS & ROLES

Team Member	Role
Sai Akshith Bandari	Analytical Lead
Leslie Buckholtz	PM & Conceptual Design Lead
Cibi Siddarth	Technical Lead

Milestone #3 Summary

In our previous assignments, we presented an enhanced analysis framework for evaluating Airbnb investment opportunities in the San Diego market. This framework incorporated qualitative and simulated data features to improve decision-making accuracy.

Review of Data Presented

1. **Sentiment and Safety Features**: We integrated sentiment polarity scores derived from listing descriptions and a simulated crime score to identify safe zones, both positively influencing the investment scoring.

2. **Revised Scoring Framework**: A new investment score out of 90 points was introduced, including criteria such as price, availability, guest rating, proximity to downtown, sentiment polarity, and safety. Listings scoring 60 or above were considered highly investible.

3. **Visual Insights and Future Enhancements**: Key pricing drivers and high investment zones near downtown and coastal areas were identified using XGBoost and heatmaps. We plan to incorporate real-world crime data, additional features, advanced models, and enhanced visualizations for deeper analysis.

4. **Data Preparation and Features**: The dataset was filtered for San Diego listings and included features like bedroom count, guest reviews, and engineered variables such as price per accommodate and distance to downtown. Missing review scores were imputed, and listings with null prices or locations were excluded.

5. **Model Training and Results**: Four regression models were tested—Linear Regression, Random Forest, XGBoost, and Neural Network—with XGBoost performing best, achieving an R^2 of 0.81 and the lowest RMSE and MAE, making it the preferred model for price prediction.

6. **Investment Scoring System**: A composite score up to 100 was created based on predicted price, availability, guest rating, and proximity to downtown. Listings scoring 80 or above were highlighted as prime investment candidates.

7. **Visualization and Next Steps**: High-scoring listings were mapped interactively, showing concentration near downtown and coastal areas. The dataset was exported

for Power BI dashboard integration, with plans for expanding the model to other cities and enhancing features and visualizations.

Detailed Plans for Power BI Representation

We are now focused on building the Power BI output to create end-user visualizations to support gaining deeper insights and understanding the Airbnb investment using three major report layers:

1. **ROI Section**: This section focuses on price distribution and revenue trends. We provide detailed visualizations of price distributions across different neighborhoods, highlighting areas with the highest potential returns. Additionally, we analyze revenue trends over time to identify seasonal patterns and other factors influencing revenue.

   - **Average Daily Rate (ADR)**: Measures the average income earned per occupied room per day.

   - **Revenue per Available Room (RevPAR)**: Calculates the revenue generated per available room, regardless of whether it is occupied.

   - **Occupancy Rate**: Indicates the percentage of available rooms that are occupied over a specific period.

   - **Gross Operating Profit (GOP)**: Represents the total revenue minus the operating expenses.

   - **Return on Investment (ROI)**: Measures the profitability of the investment relative to its cost.

2. **Location Section**: This section covers availability, geographic trends, and neighborhood clusters. We create interactive maps to visualize the availability of listings in different areas, along with geographic trends that may impact investment decisions. Neighborhood clusters highlight areas with similar characteristics and investment potential.

   - **Geographic Distribution of Listings**: Visualizes the concentration of listings in various geographic areas.

   - **Neighborhood Popularity**: Indicates the most popular neighborhoods based on rental performance and availability.

3. **Listing and Host Review Section**: This section provides review score breakdowns and analyzes host and amenities. We present detailed breakdowns of review scores for each listing, along with insights into host performance and the amenities offered. This analysis will help identify top-performing hosts and listings with the best amenities.

   - **Average Review Score**: Tracks the average rating given by guests for each listing.

   - **Review Volume**: Measures the total number of reviews received by each listing.

   - **Host Response Rate**: Indicates the percentage of guest inquiries that are responded to by hosts.

   - **Host Rating**: Measures the average rating given to hosts by guests.

   - **Amenity Score**: Tracks the availability and quality of amenities offered by each listing.