**BUSAN 302**

**Group Project Proposal**

**Analysis of European AirBnB’s using   
Machine Learning**

**Fri 8-10 Group B**

**Dataset chosen: AirBnB**

| Name | UPI |
| --- | --- |
| Lance Nillo | lnil547 |
| Danny Weng | dwen293 |
| Nancy Wu | nwu928 |
| Daniel Chua | dchu804 |
| Winson Liang | wlia752 |

# Problem

How can we effectively assess the appropriate rental rates for AirBnB listings in European cities while considering the significance of various factors influencing pricing?

Our main problem revolves around using the given dataset extracted from AirBnB listings in European cities to develop a comprehensive predictive pricing strategy/model that will enhance both profitability and guest satisfaction. In order to create an effective machine learning algorithm to perform a prediction analysis, we must fully leverage the range of different factors such as, room-type, geographical location/proximity, reputation, facilities and indices, which influence pricing (Zhang et al., 2017). Other problems we can consider:

* What are the significant factors that influence a guest’s satisfaction rating?
* Does location and city-central/metro proximity have an underlying effect on the listing prices of an AirBnB?
* What is the incremental cost associated with increasing the number of bedrooms by 1?
* In what way do the different indices factor into the AirBnB pricing?

**Type of Machine Learning algorithm/model:** Supervised

With the AirBnB dataset, we will utilise a supervised learning approach. The dataset is well structured, with respective labeled feature variables, giving us information on a range of characteristics regarding the rental property. Furthermore, there is an appropriate target feature, being the rental price. This is appropriate for supervised learning algorithms as they are used to learn/train from labelled input and output data, allowing predictive or classification decisions based on this knowledge.

**Name of Machine Learning algorithm/model:** Neural Network

The Machine Learning algorithm we will use is a neural network. A neural network is a supervised algorithm, which maps input features to an output. The mapping process used by neural networks involves multiple connected layers in the network containing artificial neurons. Each of these neurons is a mathematical process unit (Georgevici & Terblanche, 2019). When used in unison with all other neurons, the relationship between the input features and output can be learned, and therefore enables the neural network to capture the complex non-linear relationships between the features and rental prices (Georgevici & Terblanche, 2019). By training and capturing these complex relationships between features and rental prices, neural networks offer a robust and accurate tool for predicting AirBnB rental prices.

# Dataset Summary

AirBnB is a vacation rental company based on the principle of peer-to-peer (P2P) lending or home sharing — connecting people who want to rent out their homes to travellers for short-term stays.

The dataset is an extraction of the AirBnBs in a few cities in Europe. Its attributes can be roughly grouped into 6 categories:

* **Price:** Comprises one floating point attribute, Price, that denotes the price of the AirBnB.
* **Features:** Variables that describe the type of AirBnB. This comprises five attributes including Person Capacity, Multiple Rooms, Business, Cleanliness Rating, and Bedrooms
* **Proximity:** Distance from a certain location. This comprises two attributes including City Center (km) and Metro Distance (km)
* **Indices:** Measures of value of the AirBnB. Comprises 4 attributes, including Attraction Index, Normalised Attraction Index, Restaurant Index, and Normalised Restaurant Index
* **Room type**: Describes what type of room it is, using boolean values. There are three attributes including Entire home/apt, Private room, and Shared room
* **Guest** **Rating**: Provides details about the guest stay experience. This includes one attribute: Guest Satisfaction.

# Columns of Interest

Studies conducted on AirBnB listing prices, hotel accommodation, and rentals have shown that reputation, ratings, facilities, and distance are some of the most important price determinants (Zhang et al., 2017). Based on these studies, our project will be focusing on the following columns of interest:

Our response variable will be **Price**. This will represent the core of our prediction efforts and we will rely of these feature variables that capture key aspects of AirBnB listings:

* **Bedrooms**: The number of bedrooms in a property can give us insight into its size and capacity, which can influence its pricing.
* **City Center (km)/ Metro Distance (km)**: The proximity of a listing to a City Centre or Metro station can affect attractiveness and accessibility for guests, thus impacting the rental pricing.
* **Restaurant Index**: Determines the availability of dining restaurants in the vicinity, which can impact demand and therefore the rental price.
* **Person Capacity**: How many people it can comfortably accommodate, another key factor which influences the rental pricing, and could affect its maintenance.
* **Cleanliness Rating**: The cleanliness and maintenance level of a property can greatly impact guest satisfaction, and satisfied guests are more likely to lead to higher prices. Higher Prices will generally mean cleaner properties as well.
* **Guest Satisfaction**: This is a vital aspect of AirBnB Rentals as properties with "Excellent" ratings can potentially boost demand. Guests may be more forgiving with high rental pricing if the property is rated at an Excellent rating, as consumers tend to believe that high price is an indicator of better quality (Verma & Gupta, 2004).

# References

Georgevici, A., & Terblanche, M. (2019). Neural networks and deep learning: a brief introduction. *Intensive Care Medicine*, *45*(5), 712–714. https://doi.org/10.1007/s00134-019-05537-w

Verma, D. P. S., & Gupta, S. S. (2004). Does higher price signal better quality? *Vikalpa*, *29*(2), 67–78. https://doi.org/10.1177/0256090920040206

Zhang, Z., Chen, R. J. C., Han, L. D., & Yang, L. (2017). Key Factors Affecting the Price of Airbnb Listings: A Geographically Weighted Approach. *Sustainability, 9*(9), 1635. https://doi.org/10.3390/su9091635