**Assignment: Airbnb Rental Price Prediction**

**Objective**

You are given the **Airbnb NYC dataset**. Your task is to **predict the rental price (price column)** of an Airbnb listing based on various features such as neighborhood, room type, number of reviews, and availability.

**Assignment Tasks**

**Part A: Data Understanding & Cleaning**

1. Load the dataset in **Pandas** and explore it (.head(), .info(), .describe()).
2. Identify and handle **missing values**.
3. Remove or treat **outliers** in the price column (hint: listings with price = 0 or extremely high values).
4. Encode categorical variables like neighbourhood\_group and room\_type.

**Part B: Exploratory Data Analysis (EDA)**

1. Create a **histogram of prices** – does the distribution look normal or skewed?
2. Show the **top 10 neighbourhoods** with the highest average prices.
3. Use a **heatmap (correlation matrix)** to check relationships between numerical features.
4. Plot **average price vs room\_type** using a bar chart.

**Part C: Feature Engineering & Scaling**

1. Select useful features (location, room type, reviews, availability, etc.) for prediction.
2. Apply **different scaling methods** (StandardScaler, MinMaxScaler) and compare their effect.

**Part D: Model Training & Evaluation**

1. Split the dataset into **train/test sets** (e.g., 80/20).
2. Train and evaluate the following regression models:

* Linear Regression
* Ridge & Lasso Regression
* Decision Tree Regressor
* Random Forest Regressor
* Support Vector Regressor (SVR)
* Gradient Boosting / XGBoost

1. Evaluate models using:

* **MAE (Mean Absolute Error)**
* **MSE (Mean Squared Error)**
* **RMSE (Root Mean Squared Error)**
* **R² Score**

**Part E: Comparison & Conclusion**

1. Create a **comparison table** of model performance (MAE, RMSE, R²).
2. Write a short **conclusion**:

* Which model performed best?
* Why might that model be more suitable for this dataset?
* How could you improve predictions further (feature engineering, outlier handling, advanced models)?