Hotel Booking Analysis and Prediction

Objective :- The goal of this project is to analyze hotel booking data to gain insights into guest behaviors and trends, and to build predictive models to improve hotel management strategies, such as predicting cancellations.

Key Tasks :-

Data Import and Preprocessing:

1. Import and read the hotel booking dataset.
2. Perform data cleaning by handling missing values, removing irrelevant columns (e.g., agent and company), and filtering out anomalous entries (e.g., bookings with zero adults, children, and babies).

Exploratory Data Analysis (EDA):

1. Identify the distribution of guests' countries of origin.
2. Analyze room pricing patterns to determine how much guests pay per night.
3. Identify the busiest months for hotel bookings.
4. Determine which months have the highest average daily rates.
5. Analyze booking patterns concerning weekdays and weekends.

Feature Engineering:

1. Create useful features for machine learning models, such as encoded categorical data and handled outliers.
2. Select important features using correlation analysis and univariate analysis.

Machine Learning Model Building:

1. Build and evaluate various machine learning models to predict hotel booking cancellations.
2. Compare model performance using metrics like accuracy and confusion matrix.
3. Perform cross-validation to ensure the robustness of the models.

Visualization:

1. Use visualization tools (e.g., Plotly) to create interactive maps and charts for better insight into the data.

Tools and Technologies :-

1. Python Libraries: Pandas, NumPy, Matplotlib, Seaborn, Plotly, Scikit-learn
2. Models: Logistic Regression, Naive Bayes, K-Nearest Neighbors, Random Forest, Decision Tree

Expected Outcomes :-

1. Comprehensive insights into guest demographics and booking behaviors.
2. Accurate predictive models to forecast booking cancellations, aiding in better resource management and strategic planning for hotels.