**INN Hotel Booking Cancellation Prediction Notebook**

This README provides a detailed overview of the INN Hotel Booking Cancellation Prediction notebook, presenting a comprehensive guide to the contents and significance of the notebook. The notebook comprises data analysis, visualization, data preprocessing, and predictive modeling steps aimed at predicting hotel booking cancellations effectively

**1. Project Overview**

The INN Hotel Booking Cancellation Prediction notebook is focused on predicting hotel booking cancellations to mitigate losses and optimize the booking process. It provides insights into historical booking data and leverages a predictive model to forecast potential cancellations.

**2. Data Analysis**

The notebook begins with an exploration of the provided data, including the past and new datasets. It conducts descriptive statistics analysis, outlier detection, and correlation analysis to gain a comprehensive understanding of the dataset.

**3. Data Visualization and Insights**

The notebook employs various visualizations and statistical insights to unravel patterns and relationships within the data. It explores the percentage of canceled bookings that were rebooked and identifies key insights about booking cancellations.

**4. Data Preprocessing**

A significant part of the notebook is dedicated to data preprocessing, encompassing data merging, missing value treatment, outlier handling, and encoding of categorical variables. Additionally, it covers train-test split and predictive modeling preparation.

**5. Predictive Modeling**

The notebook delves into predictive modeling using various algorithms such as Logistic Regression and presents a model validation process using classification reports and ROC AUC curves.

**6. Conclusion and Future Steps**

Finally, the notebook concludes by summarizing the key findings, model performance, and outlines future steps for enhancing the predictive capabilities and monitoring new data.

**7. Additional Information**

For additional details, code, and execution results, please refer to the notebook file "INN Hotel Booking Cancellation Prediction.ipynb".

**8. Code Overview**

The notebook first reads the provided data files including the past and new datasets to begin the analysis.

**Descriptive Statistics**

It provides an in-depth overview of the data statistics, presenting data distributions and insights into variables such as lead time, special requests, average price per room, and more.

**Data Visualization and Data Preprocessing**

The notebook includes various data visualizations such as pie charts and correlation plots to understand booking status, canceled bookings and the correlations between various features. Additionally, it covers data preprocessing steps such as dropping unnecessary columns, handling missing values, and treating outliers.

**Predictive Modeling and Evaluation**

The notebook proceeds to predictive modeling by employing algorithms like Logistic Regression and Decision Tree. It includes model validation, classification reports, and ROC AUC curve analysis for model performance assessment

The INN Hotel Booking Cancellation Prediction notebook encompasses a thorough analysis of historical booking data and predictive modeling to enhance the understanding and prediction of hotel booking cancellations, providing valuable insights for better decision-making.

If you have any queries or require further information, please refer to the notebook file "INN Hotel Booking Cancellation Prediction.ipynb" for a comprehensive analysis.