

Documentation for preprocessing_data.ipynb

This notebook performs **Data Pre-Processing** on the restaurant dataset and is labeled as **Lab 10 Task**.

Data Pre-Processing Steps

1. **Dealing with Null Values:**
 - o For numerical columns (Established_Year, Seating_Capacity, Average_Price_USD, Years_Operating), missing values (NaN) were filled with the **mean** of the respective column.
 - o For categorical columns, missing values (NaN) were filled with the **mode** (most frequent value) of the respective column.
2. **Dropping Unnecessary Columns:** The columns Chef, Signature_Dish, and Restaurant_Name were permanently dropped from the DataFrame.
3. **Changing Datatypes:** All columns in the DataFrame were explicitly converted to the **int64** data type.

Post-Processing Check

The final check using `df.info()` confirmed that the pre-processing was successful, with **no null values** and all 15 columns having the **int64** data type.