Zomato Price Wizard: Forecasting Dining Costs and Price Ranges

**Introduction**

The Zomato Data Analysis project is a culinary exploration designed for avid food enthusiasts eager to Savor the best cuisines worldwide within their budget. It caters not only to the discerning taste buds but also aids in identifying value-for-money restaurants and understanding the diverse culinary landscapes across different countries.

**Data Storage**

The project harnesses the power of two datasets - Zomato.csv and country\_code.csv. The latter acts as a key enabler, providing crucial variables such as Country code and Country name. Meanwhile, Zomato.csv stands as a comprehensive repository, uniquely identifying each restaurant through its Restaurant Id and encompassing vital information such as Restaurant Name, City, Address, Locality, Cuisine offerings, Average Cost for two, and more.

**Problem Statement**

Our mission in this dataset is two-fold:

1. **Average Cost for Two:** The project seeks to unveil the average cost for two people, transcending currency differences to provide a universal understanding of dining affordability.
2. **Price Range:** Delving into the intricacies of pricing, the analysis aims to decipher the range of prices for the delectable offerings at various restaurants.

**Data Exploration and Analysis**

**1. Problem Definition**

The heart of the project lies in understanding and addressing the culinary aspirations and financial constraints of foodies worldwide. By defining the problem at its core, the analysis aims to strike a balance between providing delightful gastronomic experiences and respecting budget limitations.

**2. Data Analysis**

Our journey begins with a meticulous exploration of the datasets. The union of Zomato.csv and country\_code.csv sets the stage for an in-depth analysis. Key variables such as Cuisine offerings, Average Cost for two, and Price range act as focal points, guiding our exploration and shaping our understanding of the diverse culinary landscape.

# Code to load and merge datasets

import pandas as pd

import numpy as np

# Load Zomato dataset

zomato\_data = pd.read\_csv("https://github.com/dsrscientist/dataset4/raw/main/zomato.csv", encoding='latin1')

# Load country code dataset

country\_data = pd.read\_excel("https://github.com/dsrscientist/dataset4/raw/main/Country-Code.xlsx")

# Merge datasets

merged\_data = pd.merge(zomato\_data, country\_data, how='left', left\_on='Country Code', right\_on='Country Code')

**3. EDA Concluding Remarks**

The exploratory data analysis serves as the compass for our project, navigating through cuisine preferences, restaurant locations, and user ratings. The concluding remarks provide a holistic overview of the key findings, setting the foundation for subsequent predictions and actionable insights.

**Pre-processing Pipeline**

Ensuring the dataset's cleanliness and relevance is paramount. The pre-processing pipeline involves meticulous handling of missing values, standardizing currencies, and addressing categorical variables. By preparing the data meticulously, we ensure the robustness of subsequent predictions.

**Building Machine Learning Models**

**1. Predicting Average Cost for Two**

Harnessing the power of regression models, the analysis embarks on predicting the Average Cost for two. The pandas merge operation (pd.merge) proves instrumental in combining datasets, enhancing the predictive capabilities. Model evaluation metrics become our guiding light, offering insights into accuracy and effectiveness.

# Code for predicting Average Cost for Two

from sklearn.model\_selection import train\_test\_split

from sklearn.linear\_model import LinearRegression

from sklearn.metrics import mean\_squared\_error, r2\_score

**2. Predicting Price Range**

The spotlight then shifts to classification models, tasked with predicting the Price range. The intricate dance between features and predictions unfolds, providing a nuanced understanding of the factors influencing price categorization. Model accuracy and precision become crucial benchmarks.

**# Code for predicting Price Range**

from sklearn.ensemble import RandomForestClassifier

from sklearn.metrics import accuracy\_score, classification\_report

**Concluding Remarks**

In conclusion, the Zomato Restaurant Data Analysis project is not just an exploration of cuisines and their costs; it's a journey that unravels the intricate tapestry of global dining. The predictions and insights garnered pave the way for informed decisions by both food enthusiasts and industry stakeholders, shaping the future of dining experiences.

**Submission**

The culmination of this project is housed in the dedicated GitHub repository, where the MS-Word Document File encapsulates the entire journey. From the initial problem definition to the intricacies of predictive modelling, the submission offers a comprehensive narrative of the Zomato Data Analysis endeavour.