**Exploring Zomato Restaurant Data: A Comprehensive Analysis**

**1. Introduction**

The restaurant industry has witnessed a significant transformation with the advent of online platforms like Zomato, which offer a plethora of dining options to consumers worldwide. Zomato, a popular online restaurant search and discovery service, provides users with information about restaurants, including their names, locations, cuisines, ratings, and more. This project aims to delve deep into the Zomato restaurant data to extract valuable insights and make predictions about two key aspects: the average cost for two people and the price range of restaurants.

**2. Problem Definition**

The primary objective of this analysis is to leverage machine learning techniques to predict the average cost for two people and the price range of restaurants based on various features such as location, cuisine, and ratings. By doing so, we aim to provide valuable insights for both consumers and restaurant owners. Consumers can benefit from knowing the expected cost of dining at a particular restaurant, while restaurant owners can gain insights into pricing strategies and customer preferences.

**3. Data Collection and Overview**

The datasets used for this analysis include Zomato.csv and country\_code.csv. The Zomato.csv file contains detailed information about restaurants, while the country\_code.csv file provides a mapping between country codes and country names. Upon loading the datasets, we conducted an initial overview to understand the structure and contents of the data.

**4. Data Analysis**

In this section, we performed exploratory data analysis (EDA) to gain insights into the Zomato restaurant data. We examined various aspects of the data, including the distribution of restaurants across different countries and cities, the diversity of cuisines offered, and the relationship between average cost, ratings, and other factors. Visualizations such as histograms, scatter plots, and heatmaps were used to illustrate key findings.

**5. EDA Concluding Remarks**

During exploratory data analysis, several interesting insights emerged:

* The dataset contains restaurants from various countries and cities, indicating its global coverage.
* Cuisine diversity is prevalent, with a wide range of cuisines offered by restaurants.
* The average cost for two people varies significantly across different locations and cuisines.
* There is a correlation between the average cost for two and the aggregate rating of restaurants, suggesting that higher-rated restaurants tend to be more expensive.

**6. Pre-processing Pipeline**

Before building machine learning models, the data underwent pre-processing to ensure its quality and suitability for modeling. The pre-processing pipeline included steps such as handling missing data, cleaning duplicates, and feature engineering. Techniques such as imputation, data cleaning, and feature extraction were applied to prepare the data for modeling.

**7. Building Machine Learning Models**

For predicting the average cost for two and price range, we experimented with various machine learning algorithms, including linear regression, decision trees, and random forests. The dataset was split into training and testing sets, and the models were trained on the training data. Performance metrics such as mean squared error and accuracy were used to evaluate the models' performance.

**8. Model Evaluation and Selection**

After evaluating the performance of different machine learning models, we selected the best-performing model for each prediction task. The selected models demonstrated reasonable accuracy in predicting the average cost for two and price range, enabling stakeholders to make informed decisions.

**9. Concluding Remarks**

In conclusion, the analysis of Zomato restaurant data provided valuable insights into the factors influencing the average cost for two and price range of restaurants. By leveraging machine learning models, we were able to make predictions with reasonable accuracy, empowering consumers and restaurant owners alike. Future research could explore additional features or advanced modeling techniques to further improve prediction accuracy and provide deeper insights into the restaurant industry.

**10. References**