Internship Project Summary

Title: Restaurant Rating Analysis using Python  
Organization: Cognifyz Technologies  
Internship Level: Level 2 and Level 3  
Tools Used: Python, Pandas, Seaborn, Matplotlib, Scikit-learn, Jupyter Notebook

Objective

To analyze restaurant data and build a predictive model to estimate restaurant ratings based on various features such as cuisine, price range, delivery options, and more.

### Tasks Completed

**1. Data Cleaning and Preprocessing**

* Removed missing values
* Encoded categorical variables such as "Has Online delivery" and "Rating text"
* Dropped irrelevant columns like Restaurant ID, Address, and Locality

**2. Exploratory Data Analysis (EDA)**

* Plotted histograms, barplots, and boxplots to understand rating distribution
* Compared average ratings across cities and cuisines
* Analyzed rating trends by price range, online delivery availability, and table booking options

**3. Feature Engineering**

* Created new binary features for online delivery and table booking
* Mapped and encoded categorical variables for modeling

4. Predictive Modeling

* Built a Random Forest Regressor to predict aggregate ratings
* Split dataset into training and testing sets (80/20)
* Evaluated performance using Mean Absolute Error (MAE), Root Mean Squared Error (RMSE), and R² score

Key Insights

* Restaurants offering table booking and online delivery tend to have slightly better ratings
* Certain cuisines like North Indian and Italian are rated higher on average
* Higher price range restaurants often receive better ratings, though not always consistently
* There is a positive correlation between the number of votes and the rating-

File Descriptions:

1. Intern\_Level1.ipynb

Content:Exploratory Data Analysis (EDA) and basic data cleaning on a restaurant-related dataset.

Tasks Covered:

- Understanding dataset structure

- Handling missing values

- Encoding categorical variables

- Generating visualizations (bar plots, value counts)

- Deriving insights from features like rating, online delivery, and price range

2. Intern\_level2 and level3.ipynb

-Advanced data analysis and modeling tasks, including comparisons and prediction.

- Tasks Covered:

- Feature engineering (e.g., creating binary columns)

- Group-based analysis (e.g., ratings by cuisine or price range)

- Model building using machine learning algorithms like Random Forest

- Evaluation using metrics like R² and MAE

- Visualization of results

Instructions to Run:

1. Open the notebooks in Jupyter Notebook , JupyterLab , or VS Code with Jupyter extension

2. Ensure the dataset files (if used) are in the same directory.

3. Install necessary Python packages before running