■ Exploratory Data Analysis (EDA) – Zomato Dataset

■ Objective

Perform Exploratory Data Analysis (EDA) on the Zomato dataset using Pandas, Matplotlib, and Seaborn to extract insights, identify patterns, and visualize relationships in the data.

■ Dataset

File:**zomato.csv**

Contains details about restaurants such as:
- Restaurant ID, Name, Location (Country, City, Locality)
- Cuisines offered
- Price range and Average cost for two
- Online delivery availability
- Ratings, Votes, and Reviews

■■ Steps Performed

1. Data Loading
2. Data Summary
3. Univariate Analysis
4. Bivariate Analysis
5. Correlation Analysis
6. Categorical Insights

■ Key Insights

- Most restaurants fall in the 3.0 – 4.0 rating range.
- Votes are moderately positively correlated with ratings.
- Indian cuisine is among the most popular.
- Online delivery shows differences in ratings compared to restaurants without delivery.
- Price range has a weaker effect on ratings compared to other factors.

■■ Tools & Libraries

Python 3, Pandas, Matplotlib, Seaborn

■ Code Explanation

1. Import Libraries & Load Dataset

- Imported Pandas for datamanipulation.
- Imported Matplotlib and Seaborn for visualizations.
- Loaded dataset zomato.csv with encoding 'latin-1'.

2. Basic Data Exploration

- Displayed columnnames,dataset info, first 5 rows, and shape.
- Generated summary statistics with .describe().
- Checked for missing values and duplicate rows.

3. Univariate Analysis

- Histogram of aggregate ratings.
- Countplot of restaurants by country.
- Boxplot for price range distribution.

4. Bivariate Analysis

- Boxplot:Pricerange vsAggregate rating.
- Scatterplot: Votes vs Aggregate rating.

5. Correlation Analysis

- Correlation heatmap ofnumericalcolumns.
- Pairplot of Aggregate rating, Votes, and Price range.

6. Categorical Insights

- Top10 cuisines barchart.
- Boxplot: Online delivery vs Ratings.

7. Observations

- Ratings mostly between3.0 – 4.0.

- More votes generally mean higher ratings.
- Indian cuisine is very common.
- Online delivery influences ratings.
- Price range effect is weaker.

■ Deliverables

- Jupyter Notebook: **task-5.ipynb**
- Summary Report: **task-5-summary.pdf**
- Dataset: **zomato.csv**