Project Design Phase-I
Solution Architecture

Date	30 October 2023
Team ID	NM2023TMID03207
Project Name	Indian Food EDA
Maximum Marks	4 Marks

Solution Architecture:
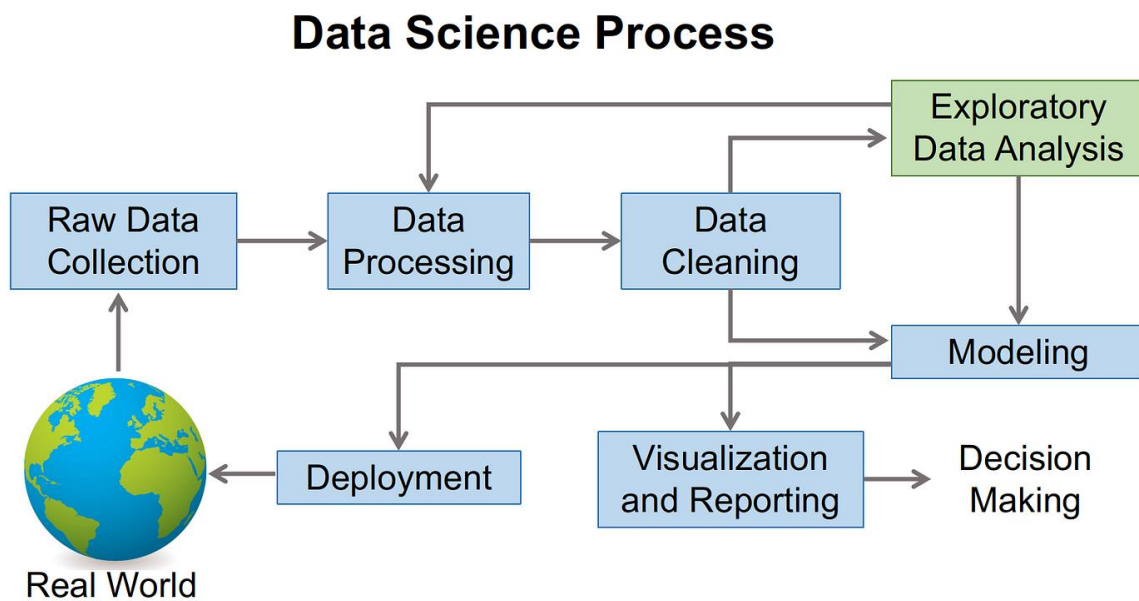
Solution architecture for Indian food EDA (Exploratory Data Analysis) involves defining the technical infrastructure, tools, and processes to analyze and gain insights from data related to Indian cuisine. EDA is a critical step in understanding and optimizing the Indian food domain, from restaurant management to recipe creation. Here's a high-level description of the architecture:

1. **Data Collection:**
   - Data sources: Collect data from various sources, including restaurant databases, food delivery platforms, recipe websites, and user-generated content.
   - Data types: Gather structured and unstructured data, such as menu items, ingredients, customer reviews, ratings, images, and more.

2. **Data Ingestion:**
   - Data Extraction: Extract data from diverse sources using ETL (Extract, Transform, Load) processes.
   - Data Storage: Store the data in a scalable and reliable database system, such as a relational database or a NoSQL database.

3. Data Preprocessing:
   - Data Cleaning: Clean and preprocess the data to handle missing values, outliers, and inconsistencies.
   - Data Transformation: Transform data into a suitable format for analysis, including encoding categorical variables, normalizing numerical data, and feature engineering.

4. **EDA Tools and Technologies:**
   - Programming Language: Utilize languages like Python and R for data analysis.
   - Data Analysis Libraries: Leverage data analysis libraries like Pandas, NumPy, Matplotlib, Seaborn, and Plotly.
   - Visualization Tools: Create interactive visualizations using tools like Tableau or Power BI.
   - Statistical Analysis: Apply statistical tests and techniques to uncover insights.

5. Exploratory Data Analysis:
   - Descriptive Statistics: Compute summary statistics, such as mean, median, mode, and standard deviation, for key food attributes.
   - Data Visualization: Create plots, charts, and graphs to visualize trends, patterns, and correlations within the data.
   - Hypothesis Testing: Test hypotheses related to Indian food preferences, ingredients, and regional variations.
   - Customer Segmentation: Segment customers based on their preferences and behavior.
   - Sentiment Analysis: Analyze customer reviews to understand sentiment and identify areas for improvement in the food industry.

6. Reporting and Insights:
   - Dashboard Creation: Develop interactive dashboards to present EDA findings in a user-friendly manner.
   - Key Metrics: Identify and track key performance indicators (KPIs) for Indian food-related businesses, such as customer satisfaction, popular dishes, and geographical trends.
   - Recommendations: Generate actionable recommendations for restaurants, chefs, and food delivery platforms based on EDA results.

7. Continuous Monitoring:
   - Set up monitoring mechanisms to continuously collect and analyze data to keep up with changing food trends and customer preferences.
   - Regularly update the EDA process as new data becomes available.

8. Scalability and Security:
   - Ensure that the architecture can scale to handle large volumes of data.
   - Implement security measures to protect sensitive data and comply with data privacy regulations.

9. Collaboration and Documentation:
   - Collaborate with domain experts, data scientists, and stakeholders to validate findings and refine analysis.
   - Document the entire EDA process, including data sources, preprocessing steps, and analysis techniques.