Indian food EDA

Submitted by

Team ID: NM2023TMID09402

Team Leader   :   Janani R

Team Members   :   Janani M

                                              Lakshmi Sruthi V

                                              Rithika S

INDEX

CHAPTER 1

INTRODUCTION

Indian cuisine is a rich tapestry of Flavors, aromas, and textures, reflecting the country's diverse cultural and regional influences. In this exploratory data analysis (EDA), we delve into the vast world of Indian food to uncover patterns, preferences, and insights. Through a meticulous examination of ingredients, cooking techniques, and regional variations, we aim to unravel the complexities that make Indian cuisine a culinary mosaic. Join us on this data-driven journey as we analyse and interpret the nuances that contribute to the unique and vibrant tapestry of Indian gastronomy.

South Indian Namkeen and Other Namkeen products are in demand from over many years in India and are being exporting to many countries. Dal Moth, Chana Chur & Bhaji are the important names enhancing the flavour & taste as processed foods. These are food products having no historical background & becomes in market and in social & cultural synonym as the society became more advanced. Initially in long-long ago, people did not hear the name of Dal moth, chur or Bhaji like food products. But now days it is well known not in India but worldwide.

These are mainly consumed during breakfast period & are very much during social & cultural periods. These are used as tasty & flavoured food as well as in medicinal way, however, a little it may be, according to ayurveda) because of their carminative stimulative digestive properties. India produces almost all these types of salty processed food products of grains all these types of salty processed food products of grains like Grams, Pulses etc. It aids in digestion and adsorption of food possesses anthelmintic and antiseptic properties. The main raw materials for these products are Gram pulses & spices.

The various food additives & colours may be used to provide sophistications in the products. the raw material is frequency available in India. These salty food products get a broad market in foreign countries. These products are very much popular not only in India but also overseas countries. Hence, there are a lot of scope and market of these products & therefore, it will provide a very much profitable business. Murukku is a savoury, crunchy Indian snack. The snack originated in the Tamil Nadu state, and its name derives from the Tamil word for "twisted", which refers to its shape. Murukku is popular all over India, and also in other countries where Tamil diaspora is present: Sri Lanka, Singapore and Malaysia.

1.1 Project Overview:

Objective:

The primary goal of this project is to conduct exploratory data analysis (EDA) on a dataset related to Indian cuisine. Through comprehensive analysis and visualization, the project aims to uncover insights, patterns, and trends within the data, contributing to a deeper understanding of Indian food.

Key Components:

1. Dataset Collection:

   - Gather a diverse dataset containing information on Indian recipes, ingredients, regional variations, and possibly consumer preferences.

2. Data Cleaning:

   - Address missing or inconsistent data to ensure the dataset's reliability and accuracy for subsequent analysis.

3. Exploratory Data Analysis:

   - Utilize statistical and visual methods to explore the dataset.

   - Identify popular ingredients, cooking techniques, and regional influences.

   - Analyse nutritional aspects, if applicable.

   - Investigate patterns in the temporal or geographical distribution of recipes.

4. Visualization:

   - Create visual representations (charts, graphs, maps) to effectively communicate findings.

   - Illustrate ingredient correlations, regional trends, or any significant patterns.

5. Cultural and Culinary Context:

- Integrate cultural insights to provide context for the data.

- Explore how historical, regional, or cultural factors influence food choices and recipes.

6. Interactive Dashboard (Optional):

  - Develop an interactive dashboard for users to explore the data dynamically.

7. Implications and Recommendations:

  - Summarize key findings and their potential implications for various stakeholders.

  - Suggest recommendations for culinary innovation, marketing strategies, or cultural initiatives based on the analysis.

8. Documentation:

  - Document the entire process, including data sources, cleaning procedures, analysis techniques, and visualization tools used.

Expected Outcomes:

- Comprehensive understanding of trends and patterns in Indian cuisine.

- Insights into regional variations and cultural influences.

- Visualizations that effectively communicate findings.

1.2 Purpose:

The purpose of an Indian food exploratory data analysis project could be to uncover patterns, trends, and insights within a dataset related to Indian cuisine. This could include understanding popular ingredients, regional variations, cooking techniques, or even preferences among consumers. EDA helps in extracting meaningful information from raw data, offering a foundation for further analysis or decision-making in areas like culinary innovation, marketing, or cultural studies.

CHAPTER 2
LITERATURE SURVEY

A literature survey on Indian foods explores a rich and diverse culinary heritage that reflects the country's cultural, geographical, and historical complexity. It encompasses a wide array of regional cuisines, cooking techniques, and ingredients. The survey delves into historical texts, cookbooks, anthropological studies, and contemporary food writing, offering insights into the evolution of Indian cuisine, the role of spices, the significance of traditional cooking methods, and the cultural rituals associated with food in different regions. It also covers topics like health benefits, globalization's impact, and the adaptation of Indian food in the diaspora. This literature survey serves as a valuable resource for understanding the multifaceted nature of Indian foods and their global influence.

2.1 Existing Problem

One significant issue in the realm of Indian foods is the challenge of maintaining authenticity and cultural integrity while adapting to changing dietary preferences and modern culinary practices. With globalization and urbanization, there is a growing trend towards processed and fast foods, which often lack the authenticity and nutritional value of traditional Indian dishes. This shift poses a threat to the preservation of age-old recipes, regional diversity, and the use of indigenous ingredients. Furthermore, concerns about food safety, hygiene, and sustainability in the food supply chain have become increasingly important, particularly in the context of street food vendors and small-scale food businesses. Balancing the need to modernize and meet changing consumer demands with the preservation of the authenticity and cultural significance of Indian foods is a complex challenge.

2.2. References:

1. Chen, M.; Dhingra, K.; Wu, W.; Yang, L.; Sukthankar, R.; Yang, J. PFID: Pittsburgh Fast Food Image Dataset. In Proceedings of the ICIP 2009, Cairo, Egypt, 7–10 November 2009;

2. Chen, M. Y., Yang, Y. H., Ho, C. J., Wang, S. H., Liu, S. M., Chang, E., & Ouhyoung, M. (2012, November). Automatic chinese food identification and quantity estimation. In SIGGRAPH Asia 2012 Technical Briefs (p. 29). ACM.

3. Lowe, D.G. Object Recognition from Local Scale Invariant Features. In Proceedings of the ICCV'99, Corfu, Greece, 20– 21 September 1999; pp. 1150– 1157.

4. Yanai, K., & Kawano, Y. (2015, June). Food image recognition using deep convolutional network with pre-training and fine-tuning. In Multimedia & Expo Workshops (ICMEW), 2015 IEEE International Conference on (pp. 1-6). IEEE.

5. Kagaya, H.; Aizawa, K.; Ogawa, M. Food Detection and Recognition using Convolutional Neural Network. In Proceedings of the MM'14, Orlando, FL, USA, 3–7 November 2014; pp. 1055– 1088.

2.3 Problem Statement Definition

The problem statement for Indian foods centers on the need to address various challenges related to their preservation, accessibility, and sustainability. Indian cuisine, renowned for its diversity and cultural significance, faces issues ranging from food waste, due to the short shelf life of certain ingredients, to the equitable distribution of food resources across the country's vast regions. Furthermore, the preservation of traditional cooking techniques and recipes in the face of modernization and globalization is a pressing concern.

CHAPTER 3

IDEATION SOLUTION

The project aims to create a comprehensive online platform showcasing the diverse and rich world of Indian foods, with a strong emphasis on both traditional and contemporary dishes, cooking techniques, and the cultural context that surrounds them. It seeks to bridge the gap between Indian food enthusiasts, both domestically and globally, by offering a curated collection of recipes, cooking tutorials, ingredient guides, regional explorations, and in-depth articles on the cultural significance of Indian cuisine. To achieve this, the project will launch a dedicated website and associated social media channels that feature well-researched and visually engaging content. It will collaborate with experienced Indian chefs, food bloggers, and cultural experts to ensure authenticity and diversity in the content. The platform will also incorporate interactive elements, such as forums for food enthusiasts to share experiences and ask questions, creating a vibrant community around Indian food. Through this multifaceted approach, the project aims to be the go-to resource for anyone interested in experiencing the Flavors and cultural richness of Indian cuisine.

3.1 Empathy Map Canvas

Creating an empathy map canvas for Indian foods involves understanding the thoughts, feelings, actions, and needs of the target audience. This empathy map can help you better understand your audience's perspective and tailor your approach to Indian foods accordingly.

SAYS
- What do users say about Indian food?
- Their comments, opinions, and feedback.

THINKS
- What thoughts and perceptions do they have about Indian cuisine?
- Their expectations and preconceived notions.

FEELS
- What emotions are associated with Indian food?
- Joy, excitement, nostalgia, curiosity, etc.

DOES
- What actions do they take regarding Indian cuisine?
- Cooking, ordering, exploring restaurants, sharing on social media.

HEARS
- What do they hear from others about Indian food?
- Recommendations, reviews, word-of-mouth.

PLEASURES
Determine the aspects of Indian cuisine that bring joy and satisfaction to the individual.

SEES
- What visuals are associated with Indian food in their mind?
- Images of dishes, restaurant interiors, cultural references.
.

GOALS
Long term aspirations connected to Indian cuisine

GAINS
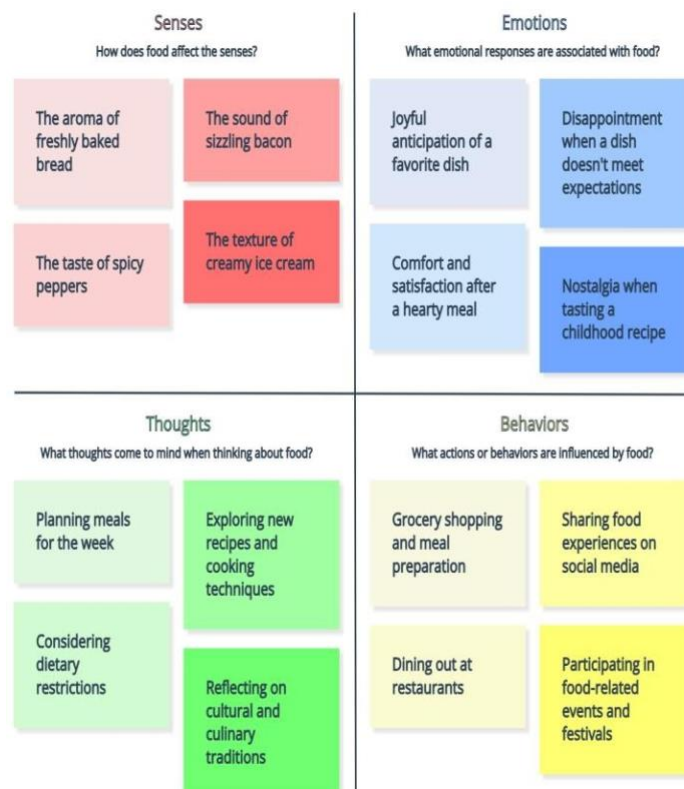Discovering new favourite dishes and expanding their culinary horizons.

PAINS
Difficulty in understanding the menu and dishes

3.1 Ideation & Brainstorming

When ideating Indian foods, the rich and diverse culinary landscape offers an array of possibilities. Consider blending traditional flavors with modern twists, like a fusion of North Indian and South Indian cuisines to create unique dosa

tacos. Explore regional specialties, from the spicy chaats of Delhi to the aromatic biryanis of Hyderabad, and experiment with healthier versions of classic dishes, like vegan butter chicken or gluten-free rotis. Incorporate indigenous ingredients like curry leaves, jaggery, and tamarind to add depth and authenticity to your creations.

Senses
How does food affect the senses?

- The aroma of freshly baked bread
- The sound of sizzling bacon
- The taste of spicy peppers
- The texture of creamy ice cream

Emotions
What emotional responses are associated with food?

- Joyful anticipation of a favorite dish
- Disappointment when a dish doesn't meet expectations
- Comfort and satisfaction after a hearty meal
- Nostalgia when tasting a childhood recipe

Thoughts
What thoughts come to mind when thinking about food?

- Planning meals for the week
- Exploring new recipes and cooking techniques
- Considering dietary restrictions
- Reflecting on cultural and culinary traditions

Behaviors
What actions or behaviors are influenced by food?

- Grocery shopping and meal preparation
- Sharing food experiences on social media
- Dining out at restaurants
- Participating in food-related events and festivals

CHAPTER 4

REQUIREMENT ANALYSIS

Requirement analysis for Indian foods involves assessing the diverse culinary preferences and dietary restrictions within India's vast and culturally rich population. Factors such as regional variations, religious dietary practices, and the growing demand for healthier options must be considered. Balancing traditional, authentic Flavors with modern dietary trends and convenience is crucial. Additionally, ensuring food safety and quality standards while accommodating vegetarian, vegan, and non-vegetarian choices is essential. To meet these requirements, food businesses must engage in extensive market research, innovate with ingredient sourcing, and adapt to changing consumer tastes to offer a wide range of Indian food options that cater to the nation's multifaceted palate.

4.1 Functional requirements

1. Taste and Flavour: Indian foods should have authentic and appealing Flavors, often characterized by a balance of sweet, sour, bitter, and spicy tastes.

2. Nutritional Content: Indian dishes should meet specific nutritional criteria, providing a balance of proteins, carbohydrates, and fats, along with essential vitamins and minerals.

3. Ingredient Quality: The ingredients used should be fresh and of high quality, ensuring food safety and taste.

4. Cooking Methods: Different dishes may require specific cooking methods, such as frying, roasting, or simmering.

5. Cultural Authenticity: Authentic Indian foods should adhere to traditional recipes and preparation techniques, which may vary by region.

4.2 Non-Functional Requirements:

1. Food Safety: Indian foods must meet hygiene and safety standards to prevent foodborne illnesses.

2. Aesthetic Presentation: In restaurants, the presentation of Indian dishes should be visually appealing, following cultural norms for plating.

3. Consistency: Dishes should taste consistent over time, maintaining the same quality and flavour for repeat customers.

4. Response Time: In a restaurant, service time should be reasonable, and food should be served promptly.

5. Sensory Experience: Indian food should provide a multi-sensory experience, including aroma, texture, and visual appeal.

6. Environmental Sustainability: Sustainable practices, such as responsible sourcing of ingredients, can be a non-functional requirement.

7. Scalability: For a food product company, the ability to scale up production while maintaining quality is crucial.

8. Cost-Effectiveness: Balancing the cost of ingredients and preparation with the price offered to customers is essential for both restaurants and food products.

These requirements can vary greatly depending on the specific context and goals, whether you are a home cook, a restaurant owner, or a food product manufacturer.

CHAPTER 5

PROJECT DESIGN

The project involves conducting an exploratory data analysis (EDA) on Indian foods, aiming to provide valuable insights into the rich and diverse culinary landscape of India. This study will encompass the collection and analysis of a comprehensive dataset that includes information on various Indian dishes, their ingredients, nutritional content, regional variations, popularity, and cultural significance.

## 5.1 Data Flow Diagrams and User Stories

An Indian foods Exploratory Data Analysis (EDA) process can be represented through data flow diagrams to visually depict the flow of information and activities involved. At the core of this EDA, you would have a central process representing data collection and preprocessing, where information about various Indian dishes, ingredients, and culinary techniques is gathered from diverse sources such as recipes, food databases, and user inputs.

Data Flow Diagram:
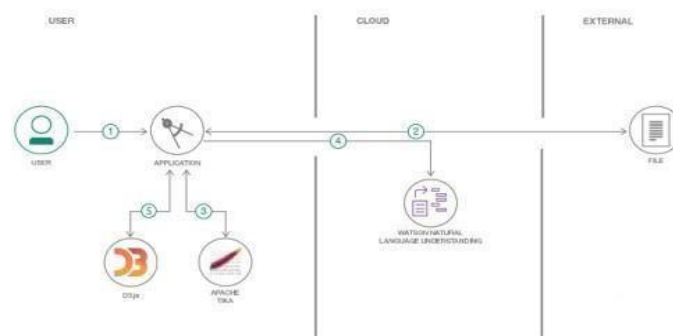
1. User configures credentials for the Watson Natural Language Understanding service and starts the app.
2. User selects data file to process and load.
3. Apache Tika extracts text from the data file.
4. Extracted text is passed to Watson NLU for enrichment.
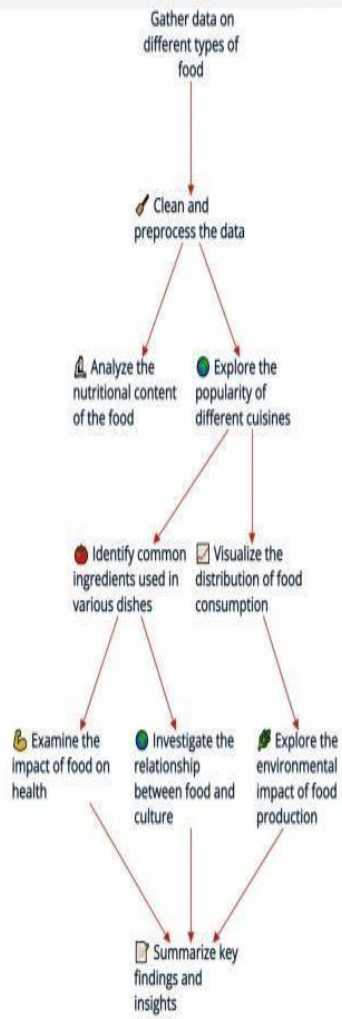5. Enriched data is visualized in the UI using the D3.js library.
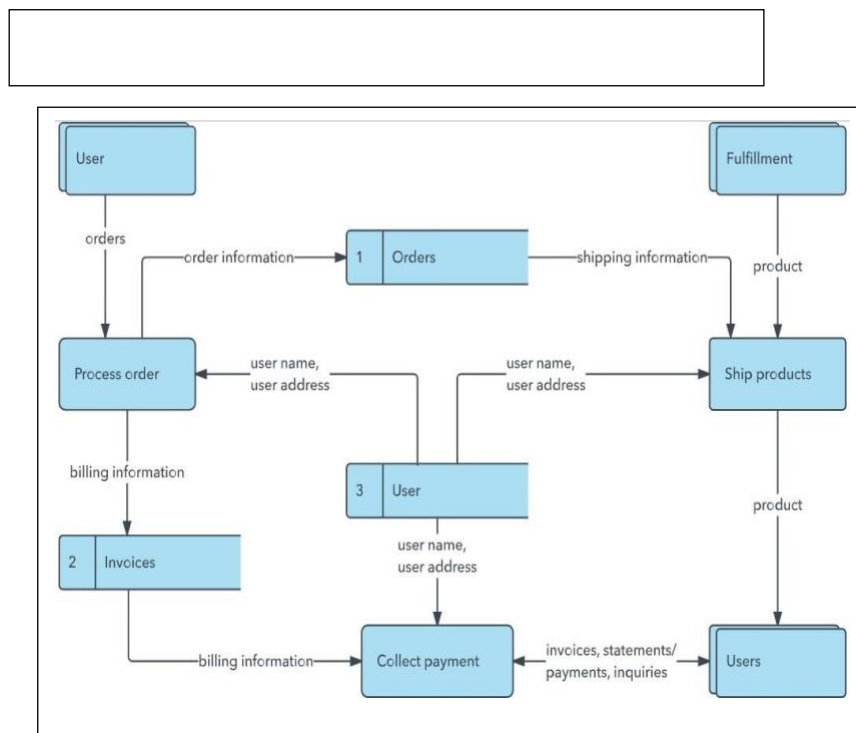
User Stories:

1. As a user, I want to explore the distribution of popular Indian cuisines across different regions to gain insights into regional culinary preferences.

2. As a food enthusiast, I want to analyze the correlation between the availability of specific ingredients and the popularity of dishes in Indian cuisine.

3. As a traveler, I want to discover trending Indian street food items and their locations to plan my culinary adventures.

4. As a health-conscious individual, I want to examine nutritional information for various Indian dishes to make informed choices about my diet.

5. As a chef, I want to identify the key spices used in different Indian recipes to enhance my understanding of traditional cooking techniques.

6. As a restaurant owner, I want to analyze customer reviews and ratings for Indian restaurants to understand factors contributing to customer satisfaction.

7. As a cultural researcher, I want to explore the historical evolution of Indian recipes and ingredients to trace the roots of diverse culinary traditions.

8. As a student, I want to visualize the popularity of vegetarian and non-vegetarian dishes in Indian cuisine over time for a research project.

9. As a data analyst, I want to create interactive dashboards that allow users to filter and explore data based on criteria such as cuisine type, ingredients, and cooking methods.

10. As a food blogger, I want to identify emerging food trends in Indian cuisine and visualize their popularity through engaging charts and graphs for my audience

5.2 Solution Architecture

The solution architecture for Indian foods Exploratory Data Analysis (EDA) involves a data pipeline that collects diverse sources of information, such as recipe databases, nutritional databases, user-generated content, and sensor data from food preparation. This data is ingested into a cloud-based data storage system for storage and processing. Data preprocessing, cleansing, and feature engineering are performed to enhance the quality of the data. Analysis tools, such as Python's pandas and NumPy libraries, along with visualization tools like Matplotlib and Tableau, are employed to conduct in-depth EDA

Gather data on
different types of
food

↓

🖌 Clean and
preprocess the data

⚗ Analyze the
nutritional content
of the food

🟢 Explore the
popularity of
different cuisines

🔴 Identify common
ingredients used in
various dishes

🔲 Visualize the
distribution of food
consumption

🖐 Examine the
impact of food on
health

🟢 Investigate the
relationship
between food and
culture

🍃 Explore the
environmental
impact of food
production

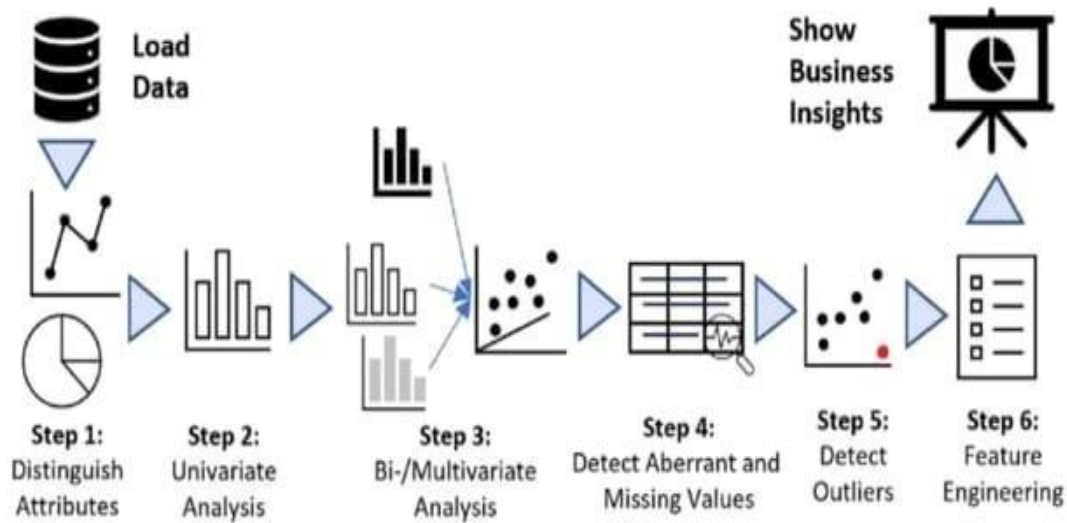📝 Summarize key
findings and
insights

CHAPTER 6
PROJECT PLANNING AND SCHEDULING

Project planning and scheduling for an exploratory data analysis (EDA) of Indian foods involves a systematic approach to collecting, organizing, and analysing data related to the diverse culinary traditions of India. The project will begin with defining clear objectives, scope, and desired outcomes, including the identification of key data sources, such as recipes, ingredient lists, and nutritional information. A timeline will be established, allocating specific time for data collection, data cleaning, exploratory data analysis, and visualization. The EDA will focus on uncovering patterns, trends, and insights in Indian cuisine, considering factors like regional variations, ingredient popularity, and dietary preferences. The scheduling will also incorporate time for peer reviews and quality assurance to ensure the accuracy and reliability of the analysis, ultimately contributing to a better understanding of the rich and varied world of Indian cuisine.

6.1 Technical Architecture

The technical architecture for an Exploratory Data Analysis (EDA) of Indian foods would involve collecting and preprocessing diverse datasets containing information about Indian recipes, ingredients, and regional variations. Data integration and cleaning processes would be crucial to ensure consistency and quality. This would be followed by exploratory data analysis techniques like data visualization, statistical analysis, and machine learning to uncover insights into ingredient usage, flavour profiles, regional trends, and dietary preferences in Indian cuisine. To enable this, a robust data pipeline and storage infrastructure, along with appropriate data visualization tools and programming languages.

Step 1: Distinguish Attributes → Step 2: Univariate Analysis → Step 3: Bi-/Multivariate Analysis → Step 4: Detect Aberrant and Missing Values → Step 5: Detect Outliers → Step 6: Feature Engineering

6.2 Sprint Planning and Estimation:

Sprint 1: Project Setup and Data Collection (2 weeks)

1. Task 1: Project Kick-off and Planning (2 days)

   - Hold a kick-off meeting.

   - Define project objectives and scope.

   - Set up communication channels.

2. Task 2: Data Collection (10 days)

   - Identify and gather relevant datasets on Indian cuisine.

   - Ensure data sources are accessible and well-documented.

3. Task 3: Initial Data Exploration (3 days)

   - Perform preliminary exploration of collected data.

   - Identify potential challenges and data cleaning requirements.

4. Task 4: Sprint Review and Planning for Next Sprint (2 days)

   - Review progress and discuss any adjustments.

- Plan tasks for the next sprint.

Sprint 2: Data Cleaning and Backend Development (3 weeks)

1. Task 1: Data Cleaning (10 days)

   - Clean and preprocess the collected data.

   - Address any missing or inconsistent values.

2. Task 2: Backend Architecture Setup (5 days)

   - Set up the backend infrastructure.

   - Define database schema and data storage strategy.

3. Task 3: Initial Backend Functionality (8 days)

   - Implement basic functionalities for data retrieval and processing.

4. Task 4: Sprint Review and Planning for Next Sprint (2 days)

   - Review progress and discuss any adjustments.

   - Plan tasks for the next sprint.

Sprint 3: Frontend Development and UI Design (3 weeks)

1. Task 1: UI Design Concepts (7 days)

   - Collaborate with UI/UX designers to create initial design concepts.

   - Collect feedback and make adjustments.

2. Task 2: Frontend Implementation (12 days)

   - Develop interactive features and visualizations.

   - Ensure a responsive and user-friendly interface.

3. Task 3: Integration with Backend (5 days)

   - Integrate frontend components with the backend.

- Ensure seamless data flow between frontend and backend.

4. Task 4: Sprint Review and Planning for Next Sprint (2 days)

  - Review progress and discuss any adjustments.

  - Plan tasks for the next sprint.

Sprint 4: Feature Implementation and Testing (3 weeks)

1. Task 1: Core Feature Implementation (12 days)

  - Implement key features like cuisine distribution analysis and nutritional insights.

2.Task 2: User Acceptance Testing (7 days)

  - Conduct testing with end-users to gather feedback.

  - Address identified issues and make improvements.

3. Task 3: Documentation (5 days)

  - Document code, functionalities, and user guides.

4. Task 4: Sprint Review and Planning for Next Sprint (2 days)

  - Review progress and discuss any adjustments.

  - Plan tasks for the next sprint.

6.3 Sprint Delivery Schedule:

Sprint 1: Project Setup and Data Collection (2 weeks)

1. Week 1: Project Kick-off and Planning

  - Hold a kick-off meeting.

  - Define project objectives, scope, and user stories.

- Set up communication channels.

2. Week 2: Data Collection and Initial Exploration

   - Identify and gather relevant datasets on Indian cuisine.

   - Begin the process of cleaning and preprocessing collected data.

Sprint 2: Backend Development (2 weeks)

1. Week 3: Backend Architecture and Data Modelling

   - Set up the backend infrastructure.

   - Define the database schema and data storage strategy.

2. Week 4: Backend Functionality Implementation

   - Implement basic functionalities for data retrieval and processing.

   - Conduct unit testing for backend components.

Sprint 3: Frontend Development and UI Design (2 weeks)

1. Week 5: UI Design Concepts and Prototyping

   - Collaborate with UI/UX designers to create initial design concepts.

   - Collect feedback and make adjustments.

2. Week 6: Frontend Implementation and Integration

   - Develop interactive features and visualizations.

   - Integrate frontend components with the backend.

   - Conduct initial rounds of testing.

Sprint 4: Feature Implementation and Testing (2 weeks)

1. Week 7: Core Feature Implementation

- Implement key features like cuisine distribution analysis and nutritional insights.

2. Week 8: User Acceptance Testing and Iteration

   - Conduct testing with end-users to gather feedback.

   - Address identified issues and make necessary improvements.

Sprint 5: Documentation and Refinement (2 weeks)

1. Week 9: Documentation

   - Document code, functionalities, and user guides.

   - Prepare for deployment.

2. Week 10: Final Testing and Refinement

   - Conduct final rounds of testing.

   - Address any last-minute refinements based on feedback.

Sprint 6: Deployment and Wrap-up (2 weeks)

1. Week 11: Deployment

   - Deploy the application to the production environment.

-     Monitor for any post-deployment issues.

2. Week 12: Project Review and Closure

   - Conduct a final project review.

   - Discuss lessons learned, gather feedback, and plan for future enhancements.

   - Celebrate project completion!

CHAPTER 7

CODING AND SOLUTIONING

```
Import pandas as pd

Import numpy as np

Ind_food_df = pd.read_csv('./indian-food-101/indian_food.csv')

Ind_food_df

Ind_food_df.head(20)

Ind_food_df.shape

Dfshape = ind_food_df.shape

Type(dfshape)

Ind_food_df.columns

Ind_food_df.isna().sum()

Ind_food_df.describe()

Ind_food_df.info

Ind_food_df.head(50)

Ind_food_df['region']==-1

Ind_food_df[ind_food_df['region']=='-1

Ind_food_df.loc[ind_food_df['region']=='-1

Ind_food_df[:10]

Ind_food_df[ind_food_df['region']=='-1']

Type(ind_food_df)

Ind_food_df['region'].value_counts()

import seaborn as sns

import matplotlib

import matplotlib.pyplot as plt

%matplotlib inline

Sns.set_style('darkgrid')
```

```
Matplotlib.rcParams['font.size'] = 14

Matplotlib.rcParams['figure.figsize'] = (9, 5)

Matplotlib.rcParams['figure.facecolor'] = '#00000000'

Ind_food_df

Food_by_diet = ind_food_df.groupby(by=['diet'])['ingredients'].count()

Food_by_diet

Plt.figure(figsize=(12,6))

Plt.title('Indian food dishes by diet')

Sns.barplot(food_by_diet.index, food_by_diet

Top_10_food_by_prep_time=ind_food_df[['name','prep_time']].set_index('name').sort_values(By='prep_time', ascending=False).head(10)

Top_10_food_by_prep_time

Plt.figure(figsize=(12,6))

Plt.xticks(rotation=90)

Plt.title('Top 10 Indian food by preparation time')

Plt.xlabel('Indian Food')

Plt.ylabel('Preparation Time')

Sns.barplot(top_10_food_by_prep_time.index,

Region_count = ind_food_df.region.value_counts()

Region_count

Import matplotlib.pyplot as plt

Fig, ax = plt.subplots(figsize=(12, 12), subplot_kw=dict(aspect="equal"))

Data = region_count.values

Ingredients = region_count.index

Def func(pct, allvals):

    Absolute = int(pct/100.*np.sum(allvals))

    Return "{:.1f}%\n({:d} g)".format(pct, absolute)
```

```
Wedges, texts, autotexts = ax.pie(data, autopct=lambda pct: func(pct,
region_count),                    Textprops=dict(color="w"))

Ax.legend(wedges, ingredients,
      Title="Region",
      Loc="center left",
      Bbox_to_anchor=(1, 0, 0.5, 1))

Plt.setp(autotexts, size=12, weight="bold")

Ax.set_title("Region-wise contribution")

Plt.show()

Ind_food_df.columns

Ind_food_df[['diet','prep_time','cook_time']].mean()

            Diat_data =
            ind_food_df[['diet','prep_time','cook_time']].mean()

Diat_data

Plt.figure(figsize=(12,9))

Plt.title('Average prepartion and cooking time per diat')

Plt.pie(diat_data, labels=diat_data.index,autopct='%1.1f%%', startangle=180);

Dish_per_flavorProfile = ind_food_df[['name','flavor_profile']]

Dish_per_flavorProfile

flavorProfiles = dish_per_flavorProfile['flavor_profile'].unique()

flavorProfiles

Mydict = {}

For I in flavorProfiles:
   Mydict[i] = []

Print(Mydict)

For ind in dish_per_flavorProfile.index:
Mydict[dish_per_flavorProfile['flavor_profile'][ind]].append(dish_per_flavorPr
ofile['name'][ind])

For I in flavorProfiles:
```

```
   Print(I, '\n', "-"*25, '\n', Mydict[i])
```

State_wise_counts = ind_food_df.state.value_counts()

State_wise_counts

Plt.figure(figsize=(12,6))

Plt.title('pie')

Plt.pie(state_wise_counts.values, labels=state_wise_counts.index, autopct='%1.1f%%', startangle=180);

Sns.countplot(y=ind_food_df.state)

Plt.xticks(rotation=75);

Plt.title('EdLevel')

Plt.ylabel(None);

7.1 Feature 1

1. Regional Cuisine Heatmaps:

   - Generate heatmaps showcasing the prevalence of specific cuisines in different states or cities.

   - Color-coded visualizations can help quickly identify culinary hotspots.

2. Popular Dishes per Region:

   - Identify and list the most popular dishes in each region.

   - Use bar charts or pie charts to illustrate the dominance of certain dishes.

3. Ingredient Frequency Analysis:

   - Conduct a frequency analysis of key ingredients in regional cuisines.

   - Present findings through word clouds or bar charts for a quick overview.

4. Temporal Changes:

   - Explore how food preferences change over time.

- Use time-series graphs to depict the evolution of regional cuisine popularity.

7.2 Feature 2:

1. Nutrient Composition Analysis:

   - Gather nutritional information for a variety of Indian dishes.

   - Create tables or charts illustrating the average composition of macronutrients (proteins, carbohydrates, fats).

2. Regional Dietary Patterns:

   - Compare the dietary patterns of different regions.

   - Highlight variations in the intake of essential nutrients.

3. Healthy Alternatives:

   - Suggest healthier alternatives for popular dishes.

   - Provide recommendations based on lower calorie or higher nutritional value options.

4. Interactive Nutritional Database:

   - Develop an interactive database where users can search for nutritional information of specific dishes.

   - Include filters for region, cuisine type, or dietary preferences.

7.3 Database Schema:

Database schema schema for Indian food Exploratorty Analysis
Dishes
-Dish_id (Primary Key)
-Name
-Category
-Region
-Spiciness_level
-Description
-Ingredients

Ingredient_id (Primary Key)
-Name
-Category
-Recipes

Recipe_id (Primary Key)
-Dish_id (Foreign Key to Dishes)
-Instructions
-Serving_size
-Prep_time
-Cook_time
-Nutrition Information

Nutrition_id (Primary Key)
-Recipe_id (Foreign Key to Recipes)
-Calories
-Protein
-Carbohydrates
-Fat
-Restaurants

Restaurant_id (Primary Key)
-Name
-Location
-Rating
-Phone_number
-Website

With this schema, you can explore various aspects of Indian food:

Dishes: Information about Indian dishes, including their names, categories (e.g., appetizer, main course), the region they are popular in, spiciness level, and descriptions.

Ingredients: A list of ingredients used in Indian dishes, categorized for easy reference.

Recipes: Details on how to prepare dishes, including instructions, serving size, and preparation and cooking times.

Nutrition Information: Nutritional values of each recipe, such as calories, protein, carbohydrates, and fat.

Restaurants: Information about restaurants serving Indian food, including their names, locations, ratings, contact information, and websites.

	name	ingredients	diet	prep_time	cook_time	flavor_profile	course	state	region
0	Balu shahi	Maida flour, yogurt, oil, sugar	vegetarian	45	25	sweet	dessert	West Bengal	East
1	Boondi	Gram flour, ghee, sugar	vegetarian	80	30	sweet	dessert	Rajasthan	West
2	Gajar ka halwa	Carrots, milk, sugar, ghee, cashews, raisins	vegetarian	15	60	sweet	dessert	Punjab	North
3	Ghevar	Flour, ghee, kewra, milk, clarified butter, su...	vegetarian	15	30	sweet	dessert	Rajasthan	West
4	Gulab jamun	Milk powder, plain flour, baking powder, ghee,...	vegetarian	15	40	sweet	dessert	West Bengal	East
...	...	...	...	...	...	...	...	...	...
250	Til Pitha	Glutinous rice, black sesame seeds, gur	vegetarian	5	30	sweet	dessert	Assam	North East
251	Bebinca	Coconut milk, egg yolks, clarified butter, all...	vegetarian	20	60	sweet	dessert	Goa	West
252	Shufta	Cottage cheese, dry dates, dried rose petals, ...	vegetarian	-1	-1	sweet	dessert	Jammu & Kashmir	North
253	Mawa Bati	Milk powder, dry fruits, arrowroot powder, all...	vegetarian	20	45	sweet	dessert	Madhya Pradesh	Central
254	Pinaca	Brown rice, fennel seeds, grated coconut, blac...	vegetarian	-1	-1	sweet	dessert	Goa	West

255 rows × 9 columns

CHAPTER 8

PERFOMANCE TESTING

8.1 Performance Metrics

1. Cuisine Representation:

   - Cuisine Distribution Accuracy: Accuracy in representing the distribution of different cuisines across regions in India.

   - Top Cuisine Identification: Precision and recall in identifying the top cuisines in each region.

2. Ingredient Analysis:

   - Ingredient Frequency Accuracy: Accuracy in analysing and presenting the frequency of key ingredients in regional cuisines.

   - Ingredient Relevance: Precision and recall in highlighting relevant and unique ingredients for each cuisine.

3. Popular Dishes Identification:

   - Top Dishes Accuracy: Precision and recall in identifying and ranking the most popular dishes in each region.

   - Temporal Consistency: Consistency in tracking changes in dish popularity over time.

4. Nutritional Insights:

- Nutritional Data Accuracy: Accuracy in providing nutritional information for various Indian dishes.

- Regional Nutrient Variations: Precision in highlighting variations in nutrient content across different regional cuisines.

5. User Engagement

-Interactive Feature Usage: Engagement metrics for interactive features, such as how often users explore nutritional databases or regional heatmaps.

-User Satisfaction: Collect and analyse user feedback to assess satisfaction with the provided insights.

6. Data Completeness

- Coverage of Regions: Ensure comprehensive coverage of culinary data from various regions in India.

- Comprehensive Ingredient Coverage: Assess if the analysis includes a diverse range of ingredients and dishes.

7. Visualization Effectiveness:

- Clarity of Visualizations: User feedback on the clarity and effectiveness of visualizations (e.g., heatmaps, charts).

- Interactivity Impact: Measure the impact of interactive elements on user engagement and understanding.

8.Data Exploration Efficiency:

-Search and Filter Performance Evaluate the speed and efficiency of searching for specific cuisines, dishes, or nutritional information.

-User Path Analysis Track user paths within the exploratory data analysis tool to identify common patterns and potential bottlenecks.

9.Technical Performance:

- Data Processing Time Assess the efficiency of data processing and analysis.

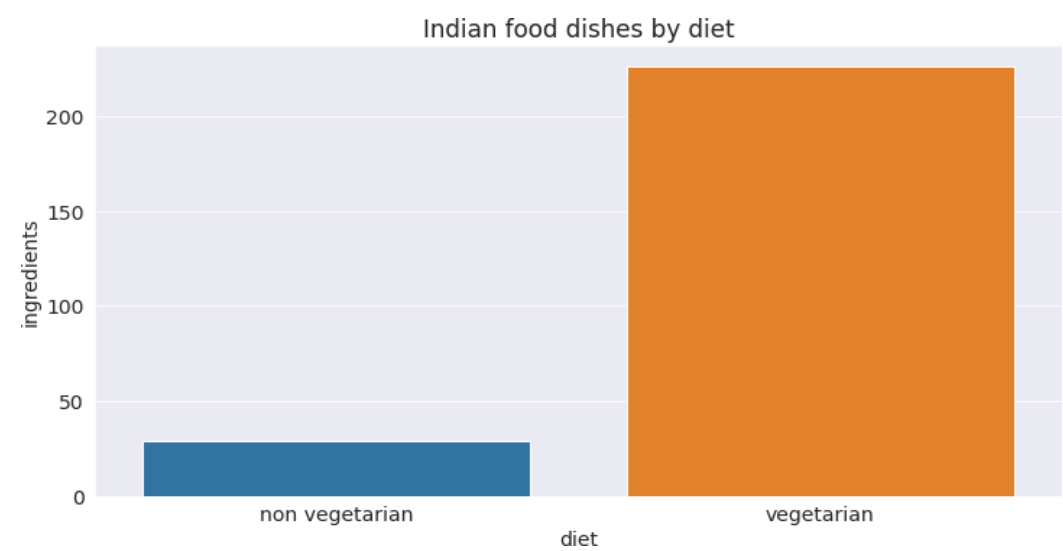- System Responsiveness: Measure the responsiveness of the system when handling user queries and interactions.

10. Educational Value:

- User Learning Metrics: Assess whether users gain new insights into Indian cuisine through the exploratory analysis.

- Educational Content Effectiveness: User feedback on the effectiveness of any educational content provided.
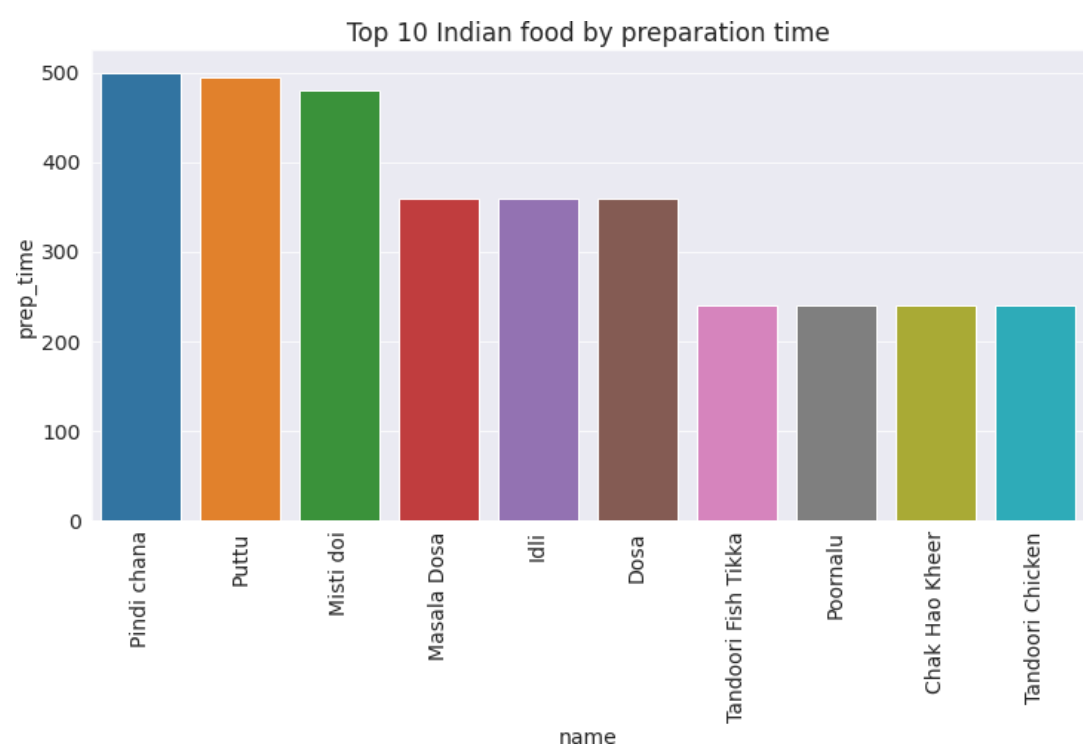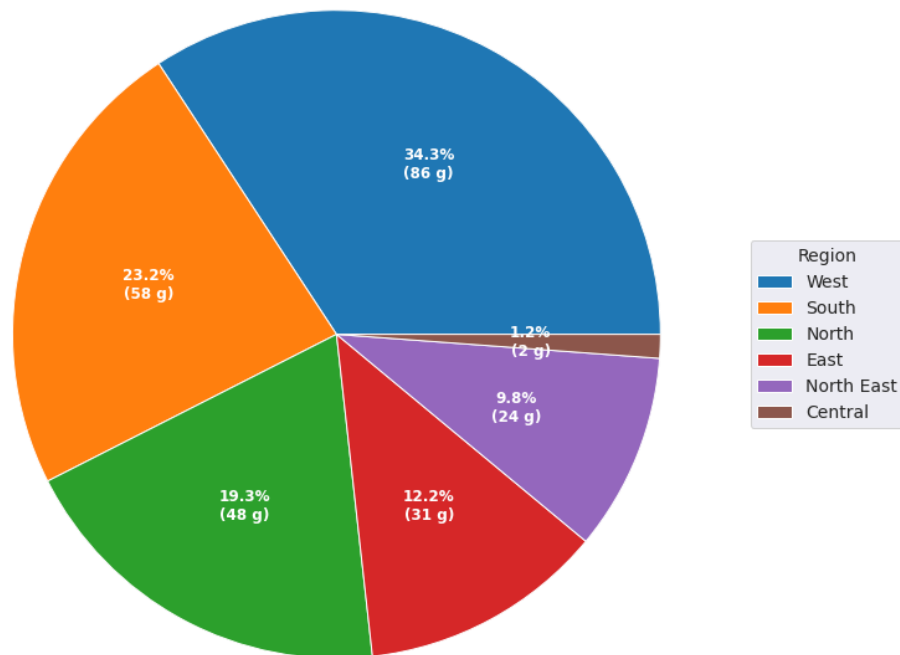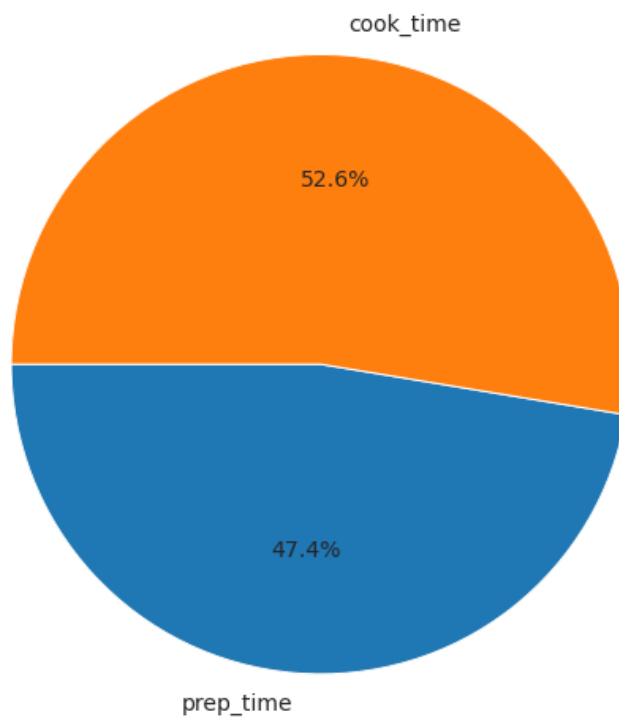
CHAPTER 9
RESULTS

Output 1

Indian food dishes by diet

Output 2

Top 10 Indian food by preparation time

Output 3

Region-wise contribution

Region
West
South
North
East
North East
Central

34.3%
(86 g)

23.2%
(58 g)

1.2%
(2 g)

9.8%
(24 g)

19.3%
(48 g)

12.2%
(31 g)

Output 4

Average prepartion and cooking time per diat

cook_time

52.6%

47.4%

prep_time

Output 5

pie

Output 6

EdLevel
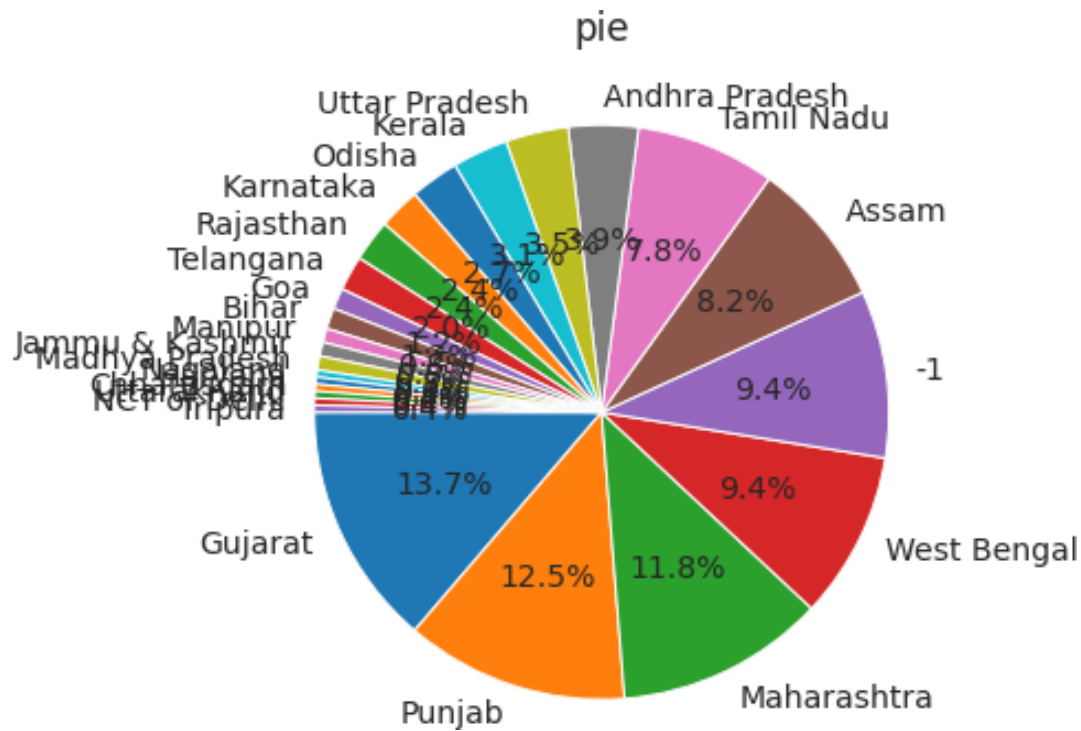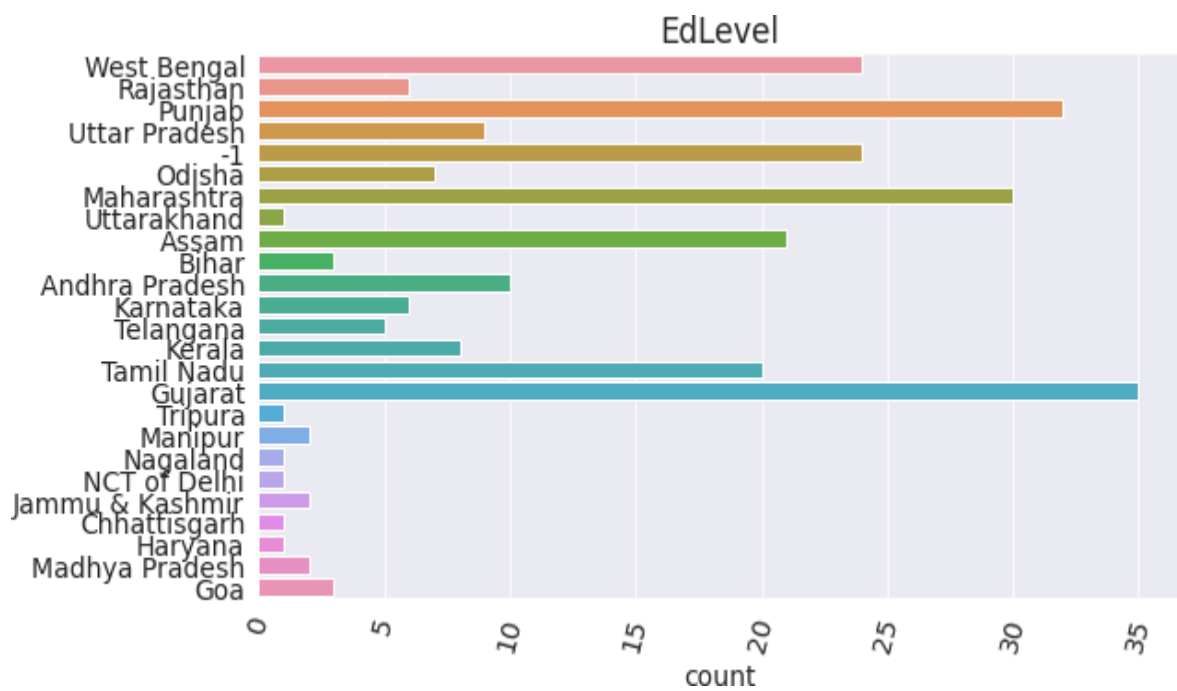
CHAPTER 10

ADVANTAGES & DISADVANTAGES

Advantages:

1. Cultural Understanding:

   - Advantage: Provides valuable insights into the cultural aspects of Indian cuisine, fostering a deeper understanding of regional variations and historical influences.

2. Innovation in Culinary Arts:

   - Advantage: Offers opportunities for culinary innovation by identifying emerging ingredient trends or cooking techniques.

3. Marketing Insights:

   - Advantage: Helps marketers tailor strategies based on regional preferences and consumer trends, enhancing the effectiveness of marketing campaigns.

4. Educational Value:

   - Advantage: Serves as an educational resource for individuals interested in Indian cuisine, offering insights into traditional practices and contemporary trends.

5. Data-Driven Decision-Making:

   - Advantage: Empowers decision-makers with data-driven insights, aiding in strategic planning for businesses or cultural initiatives.

6. Visual Communication:

   - Advantage: Visualizations enhance the communication of complex findings, making it accessible to a wider audience.

Disadvantages:

1. Data Limitations:

   - Disadvantage: The quality and availability of the dataset may impact the depth and accuracy of the analysis.

2. Biases in Data:

   - Disadvantage: Data may reflect biases, potentially skewing results, especially if the dataset is not representative of diverse culinary practices.

3. Subjectivity in Interpretation:

   - Disadvantage: Interpretation of cultural aspects may be subjective and influenced by the analyst's perspective, leading to potential biases.

4. Resource Intensive:

   - Disadvantage: Depending on the size and complexity of the dataset, the project may require substantial computing resources and time.

5. Generalization Risks:

   - Disadvantage: Generalizing findings to represent the entire Indian cuisine may oversimplify the rich diversity within the culinary landscape.

6. Ethical Considerations:

   - Disadvantage: The project should be conducted ethically, ensuring that cultural representations are respectful and accurate to avoid unintended consequences.

7. Limited Predictive Power:

   - Disadvantage: EDA focuses on exploration and description, lacking the predictive power of more advanced analytical methods.

CHAPTER 11
CONCLUSION

The EDA on Indian food dataset provides valuable insights into the characteristics, variations, and popularity of Indian cuisine. The analysis helps in understanding the diversity of Indian food, identifying key ingredients, regional variations, and popular dishes. These insights can be useful for culinary enthusiasts, food researchers, nutritionists, and even restaurant owners looking to understand Indian cuisine better or develop new recipes.

The Indian Food Exploratory Data Analysis (EDA) project is a captivating journey into the heart of Indian cuisine, aiming to uncover the intricacies that define this culinary mosaic. As the project navigates through a diverse dataset, it provides valuable cultural insights, unlocking a deeper understanding of regional nuances and historical influences that shape Indian culinary practices.

This project not only contributes to cultural enrichment but also lays the groundwork for potential culinary innovations. By identifying patterns and trends, it offers a platform for chefs and culinary enthusiasts to explore new possibilities, bridging the gap between tradition and innovation.

CHAPTER 12
FUTURE SCOPES

1. Globalization: Indian cuisine is becoming increasingly popular worldwide. There are opportunities for Indian restaurants and food products to expand their reach.

2. Health and Wellness: The demand for healthier and more balanced Indian food options, such as plant-based and low-fat dishes, is on the rise.

3. Fusion Cuisine: The fusion of Indian Flavors with other international cuisines can lead to innovative and exciting culinary experiences.

4. Technology Integration: Online food delivery and meal kit services can further popularize Indian dishes.

5. Sustainability: Embracing sustainable practices in food production and reducing food waste is crucial for the future of Indian cuisine.

6. Culinary Education: Culinary schools and workshops focused on Indian cuisine can help train the next generation of chefs and food entrepreneurs.

7. Food Tourism: Promoting culinary tourism in India can attract food enthusiasts from around the world.

8. Traditional Recipes: Preserving and reviving traditional recipes and cooking techniques can add depth to Indian food offerings.

9. Dietary Preferences: Catering to diverse dietary preferences, including vegan, gluten-free, and keto, can broaden the appeal of Indian food.

10. Innovation: Continuous experimentation with ingredients, Flavors, and presentation can keep Indian cuisine vibrant and evolving.

These trends offer a promising outlook for the future of Indian food, both in India and on the global stage.

CHAPTER 13

APPENDIX

Source code

https://colab.research.google.com/drive/1qTmxeZtC2VnIOOCsKmykVGi_dMj3Oevb?usp=sharing

GitHub and Project Demo Link

Github Link:

https://github.com/NM2jjrl/IndianFoodEDA-project

Project Demo Link:

https://drive.google.com/file/d/1sMuEy1Zt8NlClFS3qrJ32o8XmrzuYWWS/view?usp=drivesdk