NAAN MUDHALVAN – DATA ANALYTICS WITH TABLEAU

PROJECT TITLE: INDIAN FOOD EDA

TEAM MEMBERS

1. DHIVAKARAN K
2. RAVIRAJAN S R
3. VIGNESH S
4. BARATHKUMAR B

INTRODUCTION

The aim of this project is to perform Exploratory Data Analysis (EDA) on a dataset related to Indian food. Exploratory Data Analysis is a crucial step in data analysis that helps in understanding the dataset, discovering patterns, and extracting insights. In this project, we will explore various aspects of Indian cuisine, such as ingredients, recipes, regional variations, and popularity.

PROJECT DESCRIPTION

The project aims to analyse different Indian dishes and infer about the time taken for preparing each dish with help of Exploratory Data Analysis (EDA) on a dataset created based on Indian food varieties, thus pictorially representing the preparation process as slow, medium, and fast.

PROJECT OVERVIEW

Exploratory Data Analysis (EDA) on Indian Food

DATA COLLECTION

The first step is to collect a comprehensive dataset on Indian food. This dataset can be obtained from various sources such as recipe websites, food blogs, culinary books, or public repositories. The dataset should ideally include information such as the name of the dish, ingredients, cooking methods, region, popularity, and any other relevant attributes.

DATA CLEANING AND PREPROCESSING

Once the dataset is obtained, it needs to be cleaned and pre-processed before analysis. This involves handling missing values, removing duplicates, standardizing formats, and transforming the data into a suitable structure for analysis. For example, ingredient names and measurements might need to be standardized, and textual data may require text cleaning techniques like removing stop words, punctuation, and converting to lowercase.

EXPLORATORY DATA ANALYSIS (EDA)

After cleaning the dataset, the EDA process begins. The analysis can involve various steps and techniques, including:

1. Descriptive Statistics: Calculating basic statistics such as mean, median, mode, range, and standard deviation of different attributes. This provides an initial overview of the dataset.

2. Data Visualization: Creating visual representations of the data to identify patterns, trends, and relationships. This can be done using charts, graphs, heatmaps, scatter plots, and geographical maps. For example, visualizing the distribution of recipes across different regions or plotting the popularity of dishes over time.

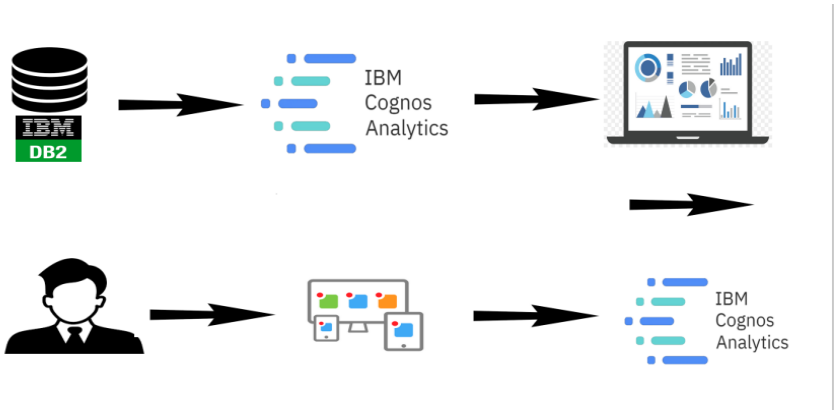
3. Ingredient Analysis: Analyzing the most used ingredients, their frequency, and variations across different regions. This can help identify key ingredients that define Indian cuisine and understand regional preferences.

4. Regional Variations: Investigating the diversity of Indian cuisine across different regions and states. This can include analyzing the prevalence of specific dishes, spices, cooking styles, and regional specialties.

5. Popularity and Ratings: Analyzing the popularity and ratings of Indian dishes or cuisines. This can involve examining user reviews, ratings, social media mentions, or any available popularity metrics to identify popular dishes or trends.

6. Nutritional Analysis: Exploring the nutritional content of Indian dishes, such as calories, macronutrients, and common ingredients contributing to specific nutritional values. This can help understand the health aspects of Indian cuisine.

TECHNICAL ARCHITECTURE



PROJECT FLOW

To accomplish this, we have to complete all the activities listed below, Define Problem / Problem Understanding

- Specify the business problem or Business requirements
- Literature Survey
- Social or Business Impact.

- **Data Collection & Extraction from Database**
  - Collect the dataset
  - Connect IBM DB2 with IBM Cognos

- **Data Preparation**
  - Prepare the Data for Visualization

- **Data Visualizations**
  - No of Unique Visualizations

- **Dashboard**
  - Responsive and Design of Dashboard

- **Story**
  - No of Scenes of Story

- **Report**
  - Creating a report

- **Performance Testing**
  - Amount of Data Rendered to DB ‘
  - Utilization of Data Filters
  - No of Calculation Fields
  - No of Visualizations/ Graphs

- **Web Integration**
  - Dashboard and Story embed with UI With Flask

- **Project Demonstration & Documentation**
  - Record explanation video for project end to end solution

- Project Documentation - Step by step project development procedure

Milestone 1: Define Problem / Problem Understanding

Activity 1:

Specify the business problem - Refer Project Description

Activity 2:

Business requirements:

The business requirements for this project would likely include:

- Data collection:

  The first requirement is to collect relevant dataset from Kaggle based on the region, ingredients, cuisines, courses, and diet in order to compare food varieties from different regions of India and the World.

- Data cleaning and preparation:

  The collected data must be cleaned and processed to ensure it is suitable for analysis. This may involve removing irrelevant information, correcting inconsistency, and missing values, and transforming the data into a format that is compatible with the analysis tools.

- Data analysis:

  The data must be analyzed to uncover meaningful insights. This could involve using techniques such as descriptive statistics, regression analysis, and data visualization to gain a deeper understanding of the data.

- Report creation:

  The insights and findings from the data analysis must be presented in a comprehensive report that includes visualizations and data tables. The

report must be well organized and easy to understand, with clear and concise explanations of the results.

Activity 3: Literature Survey

LITERATURE REVIEW: EXPLORATORY DATA ANALYSIS (EDA) ON INDIAN FOOD

INTRODUCTION

Exploratory Data Analysis (EDA) has become an essential tool in understanding and extracting insights from large datasets. In the context of Indian food, EDA plays a crucial role in unraveling the complexities of this diverse and rich culinary heritage. This literature review explores existing research and studies that have employed EDA techniques to analyze Indian food datasets, shedding light on the ingredients, regional variations, popularity, and nutritional aspects of Indian cuisine.

DIETARY ASSESSMENT AND INDIAN CUISINE ANALYSIS USING KNN AND EDA by Pearl Ahuja, Diksha Solani, Amita Goel, Nidhi Sengar, Vasudha Bahl (2021)

Description: In this paper, an exertion has been made to analyse the various food items for assessment of most favoured cuisine in every region of India along with recommendation of various dishes preferable to that region using KNN and EDA.

FOOD NUTRITIONAL ANALYSIS AND EDA by Atishya Mahesh Jain, KrishnaPriya B, Seno Sunil, Helen Grace Jikku, Sheethal Kj (2022)

Description: This project deals with the usage of Exploratory Data Analysis (EDA) in order to carry out nutritional analysis, thus finding the nutrient composition, quality, and contamination of a food variety.

ZOMATO DATA WITH EDA, GEOSPATIAL AND SENTIMENT ANALYSIS by Priyadharshini R (2021)

Description: The main objective of this project is to find insights and get some idea about the restaurants in Bangalore using different techniques like EDA, Geospatial and Sentiment Analysis.

EDA AND VISUALIZATION OF INDIAN STREET FOOD: A DATA-DRIVEN PERSPECTIVE by Das and Mohanty (2022)

Description: This study employs EDA and visualization techniques to analyze data on Indian street food. The authors explore the regional variations, popularity, and ingredients of street food across different Indian cities. Through visualizations, they highlight the street food culture and its unique culinary offerings, uncovering hidden gems and trends.

EDA OF POPULAR INDIAN DISHES ON SOCIAL MEDIA by Patel and Shah (2020)

Description: This study utilizes EDA to analyze the popularity of Indian dishes on social media platforms. The authors collect data from various social media sources and explore metrics such as user ratings, reviews, and mentions. They identify popular dishes, trends, and emerging flavors, providing insights into the changing preferences of consumers.

CONCLUSION

The literature review demonstrates the growing interest in applying EDA techniques to explore Indian food datasets. These studies have provided valuable insights into the ingredients, regional variations, popularity, and nutritional aspects of Indian cuisine. EDA has proven instrumental in unravelling the intricacies of Indian food, shedding light on the diversity, flavours, and cultural significance associated with this culinary heritage. Future research can further leverage EDA to explore new dimensions of Indian food, such as the impact of globalization, dietary preferences, and the fusion of traditional and modern culinary practices.

Activity 4: Social or Business Impact

1. Social Impact:

   Health and Nutrition Awareness: EDA on Indian food datasets facilitates the exploration of nutritional aspects, enabling individuals to make informed dietary choices. By analysing calorie content, macronutrients, and ingredient compositions, EDA promotes health and nutrition awareness. This knowledge empowers individuals to adopt healthier eating habits and make informed decisions about their diet.

2. Business Model/Impact:

   Restaurant Menu Planning: EDA on Indian food datasets helps restaurants and food establishments understand customer preferences and adapt their menus accordingly. By analysing popular dishes, regional variations, and emerging trends, businesses can tailor their offerings to meet consumer demands. This improves customer satisfaction and drives business growth.

CONCLUSION

EDA on Indian food datasets has a profound social and business impact. It promotes the preservation of culinary heritage, encourages cultural exchange, and raises awareness of health and nutrition. In the business realm, EDA drives menu planning, product development, marketing strategies, and supply chain optimization.

Milestone 2: Data Collection & Extraction from Database

Activity 1: Collect the dataset

Please use the link to download the dataset:-

https://www.kaggle.com/datasets/kanishk307/6000-indian-food-recipes-dataset

Activity 1.1: Understand the data

Check the below link out to understand the dataset in detail:-

https://www.kaggle.com/datasets/kanishk307/6000-indian-food-recipes-dataset

Activity 2: Connect IBM DB2 with IBM Cognos Explanation video link:

https://drive.google.com/file/d/1VzWCqOKKUrJrYpPc6TUwYTaF19o6ni5r/view?usp=sharing

Milestone 3: Data Preparation

Activity 1: Prepare the Data for Visualization

Preparing the data for visualization involves cleaning the data to remove irrelevant or missing data, transforming the data into a format that can be easily visualized, exploring the data to identify patterns and trends, filtering the data to focus on specific subsets of data, preparing the data for visualization software, and ensuring the data is accurate and complete. This process helps to make the data easily understandable and ready for creating visualizations to gain insights into the performance and efficiency.

Activity 1.1 : Preparing a Data Module

https://www.kaggle.com/datasets/kanishk307/6000-indian-food-recipes-dataset

Milestone 4: Data Visualization

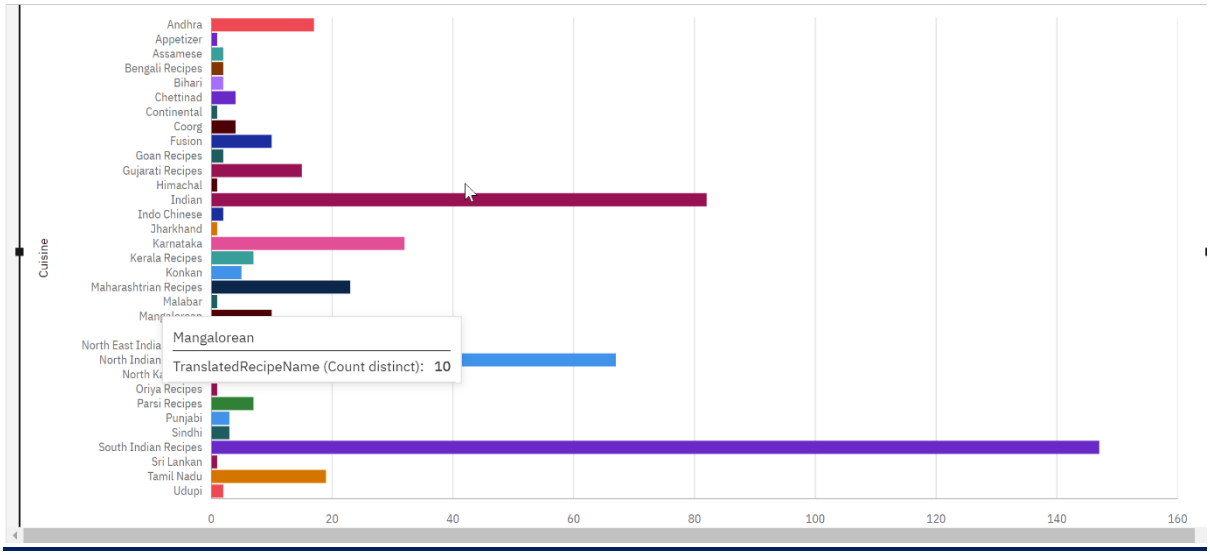
Data visualization is the process of creating graphical representations of data in order to help people understand and explore the information.
The goal of data visualization is to make complex data sets more accessible, intuitive, and easier to interpret. By using visual elements such as charts, graphs, and maps, data visualizations can help people quickly identify patterns, trends, and outliers in the data.

Activity 1: No of Unique Visualizations

The number of unique visualizations that can be created with a given dataset. Some common types of visualizations that can be used to analyze the Kaggle data include bar charts, line charts, heat maps, scatter plots, pie charts, maps etc. These visualizations can be used to compare performance, track changes over time, show distribution, and relationships between variables.

Activity 1.1: Value by Area?

Explanation video link:
https://drive.google.com/file/d/1hEPDA0TUO4XOVob5JhB5hbrSbAS92Nx/view?usp=sharing



Milestone 5: Dashboard

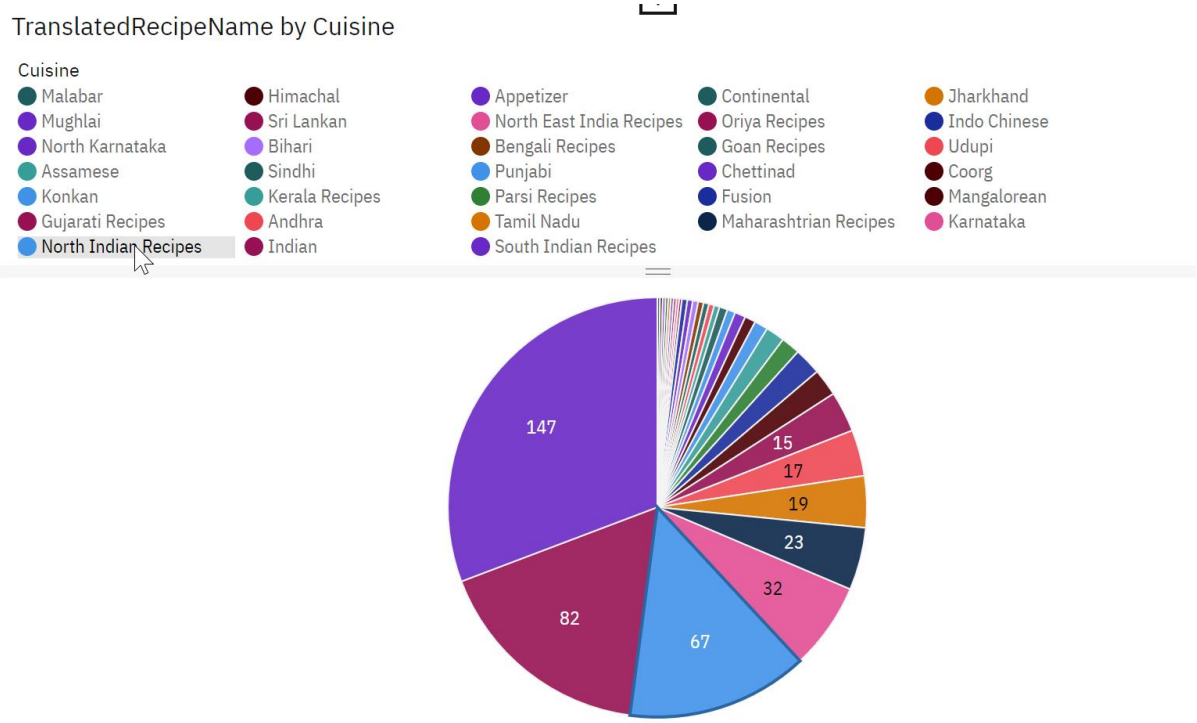
A dashboard is a graphical user interface (GUI) that displays information and data in an organized, easy-to-read format. Dashboards are often used to provide real-time monitoring and analysis of data and are typically designed for a specific purpose or use case. Dashboards can be used in a variety of settings, such as business, finance, manufacturing, healthcare, and many other industries. They can be used to track key performance indicators (KPIs), monitor performance metrics, and display data in the form of charts, graphs, and tables.

Activity 1: Responsive and Design of Dashboard

The Indian Food EDA Dashboard offers a comprehensive and interactive visual representation of the insights derived from exploring a rich dataset on Indian cuisine. This dashboard provides a user-friendly interface that allows users to delve into the diverse flavours, regional variations, popular dishes, and nutritional aspects of Indian food. With a range of interactive visualizations and data exploration tools, the dashboard facilitates a deeper understanding and appreciation of Indian culinary traditions.

Explanation video link:

https://drive.google.com/file/d/1l9dlN-BsRgnS3weGFS-fknZ5nr2Qbb1E/view?usp=drive_link



Milestone 6: Report

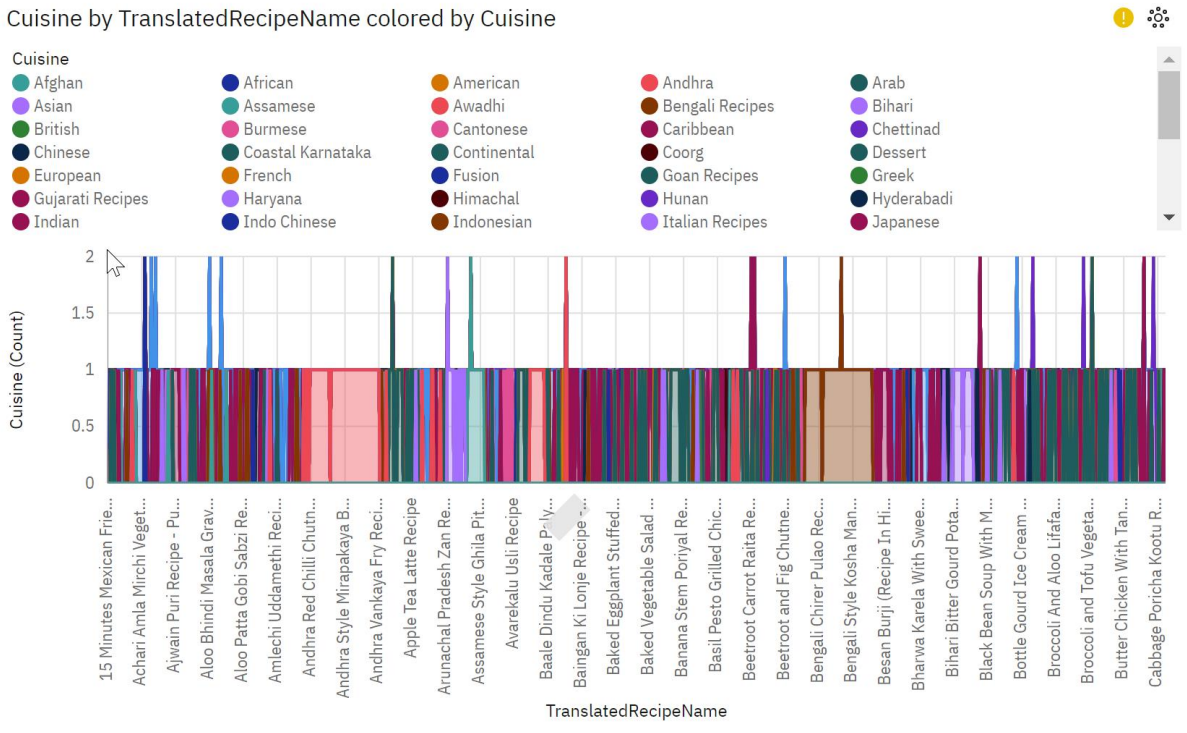
A report is a document that presents information in a specific format and layout, usually based on data from a database or other data source. A report in IBM Cognos can contain various elements, such as tables, charts, graphs, and images, as well as text and data elements, and it is designed to be used by business users to help them better understand their data and make informed decisions. There are several different types of reports available in IBM Cognos, including list reports, crosstab reports, chart reports, and report studio reports, among others. The type of report that you choose will depend on the specific needs and requirements of your organization, as well as the data that you need to present.

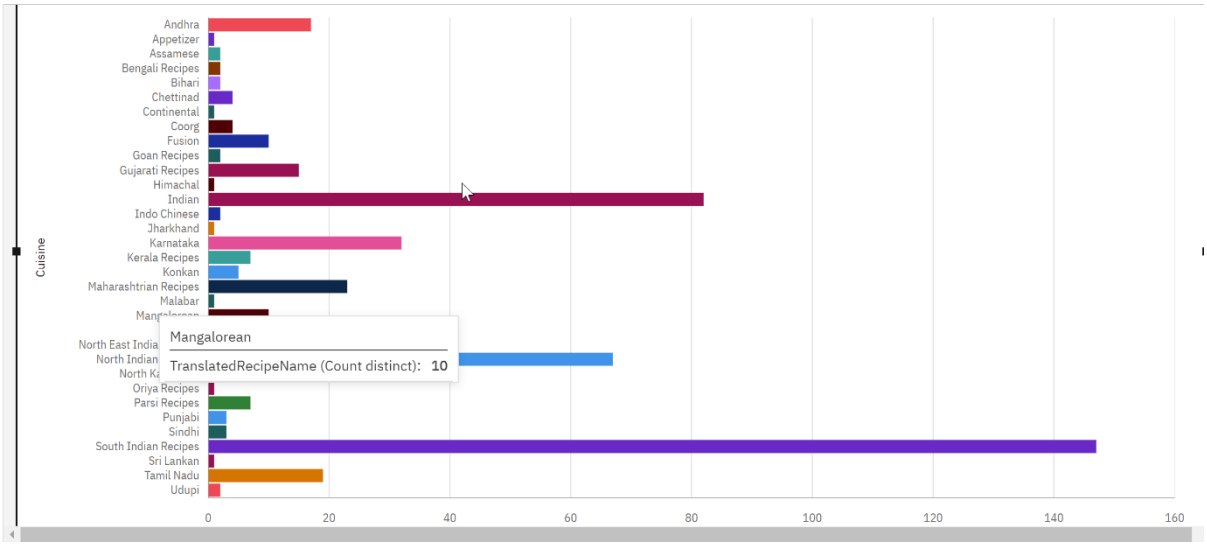
Activity 1: No of Scenes of Report

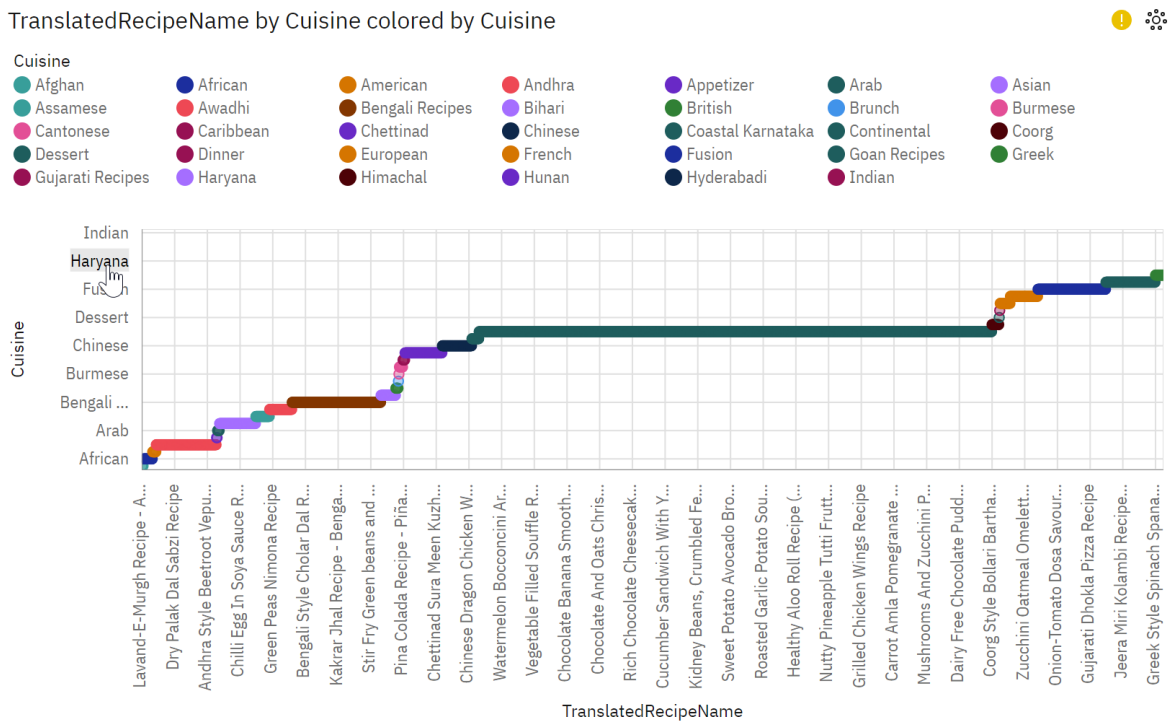
This report presents the findings and insights derived from an in-depth Exploratory Data Analysis (EDA) on a comprehensive dataset of Indian food. The aim of this analysis is to explore the diverse flavours, regional variations, popular dishes, and nutritional aspects of Indian cuisine. By employing EDA techniques, we unravel the complexities of Indian food, shedding light on its cultural significance, culinary traditions, and the impact of regional diversity.

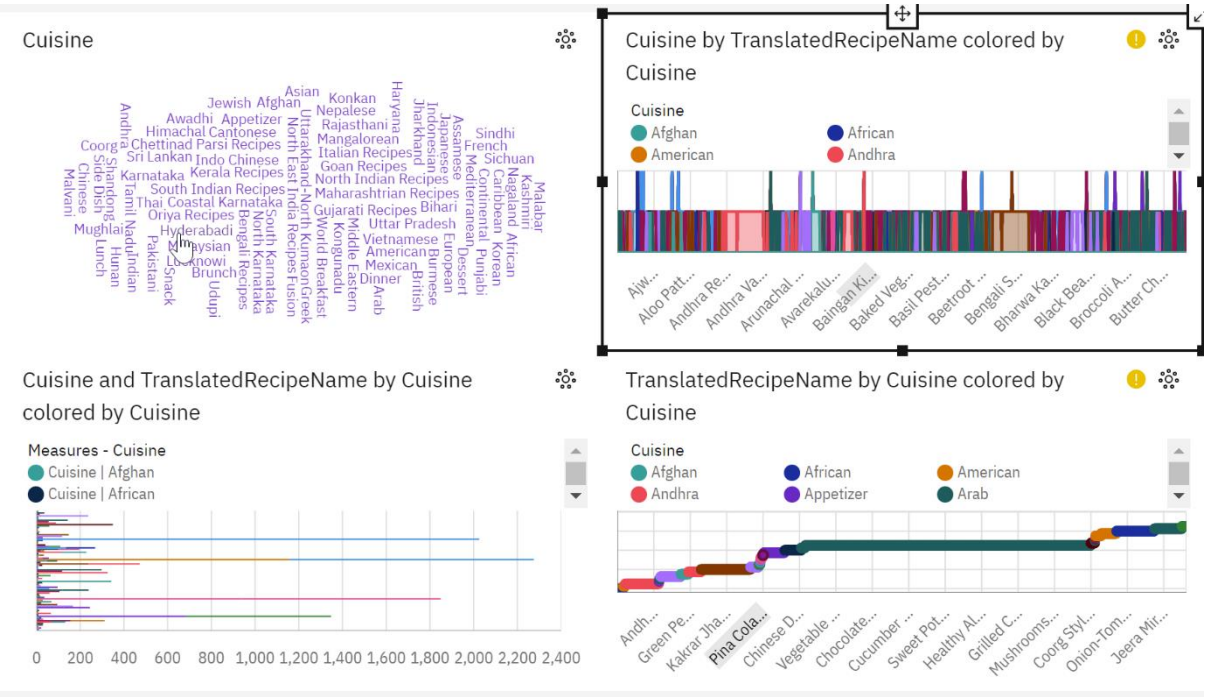
Explanation video link:

https://drive.google.com/file/d/1l9dlN-BsRgnS3weGFS-fknZ5nr2Qbb1E/view?usp=drive_link




Cuisine
Andhra
Appetizer
Assamese
Bengali Recipes
Bihari
Chettinad
Continental
Coorg
Fusion
Goan Recipes
Gujarati Recipes
Himachal
Indian
Indo Chinese
Jharkhand
Karnataka
Kerala Recipes
Konkan
Maharashtrian Recipes
Malabar
North East India
North Indian
Oriya Recipes
Parsi Recipes
Punjabi
Sindhi
South Indian Recipes
Sri Lankan
Tamil Nadu
Udupi
Mangalorean
TranslatedRecipeName (Count distinct): 10
0
20
40
60
80
100
120
140
160


TranslatedRecipeName by Cuisine colored by Cuisine
Cuisine
Afghan
African
American
Andhra
Appetizer
Arab
Asian
Assamese
Awadhi
Bengali Recipes
Bihari
British
Brunch
Burmese
Cantonese
Caribbean
Chettinad
Chinese
Coastal Karnataka
Continental
Coorg
Dessert
Dinner
European
French
Fusion
Goan Recipes
Greek
Gujarati Recipes
Haryana
Himachal
Hunan
Hyderabadi
Indian
Cuisine
Indian
Haryana
Fusion
Dessert
Chinese
Burmese
Bengali ...
Arab
African
Lavand-E-Murgh Recipe - A...
Dry Palak Dal Sabzi Recipe
Andhra Style Beetroot Vepu...
Chilli Egg In Soya Sauce R...
Green Peas Nimona Recipe
Bengali Style Cholar Dal R...
Kakrar Jhal Recipe - Benga...
Stir Fry Green beans and ...
Pina Colada Recipe - Piña...
Chettinad Sura Meen Kuzh...
Chinese Dragon Chicken W...
Watermelon Bocconcini Ar...
Vegetable Filled Souffle R...
Chocolate Banana Smooth...
Chocolate And Oats Chris...
Rich Chocolate Cheesecak...
Cucumber Sandwich With Y...
Kidney Beans, Crumbled Fe...
Sweet Potato Avocado Bro...
Roasted Garlic Potato Sou...
Healthy Aloo Roll Recipe (...
Nutty Pineapple Tutti Frutt...
Grilled Chicken Wings Recipe
Carrot Amla Pomegranate ...
Mushrooms And Zucchini P...
Dairy Free Chocolate Pudd...
Coorg Style Bollari Bartha...
Zucchini Oatmeal Omelett...
Onion-Tomato Dosa Savour...
Gujarati Dhokla Pizza Recipe
Jeera Miri Kolambi Recipe...
Greek Style Spinach Spana...
TranslatedRecipeName



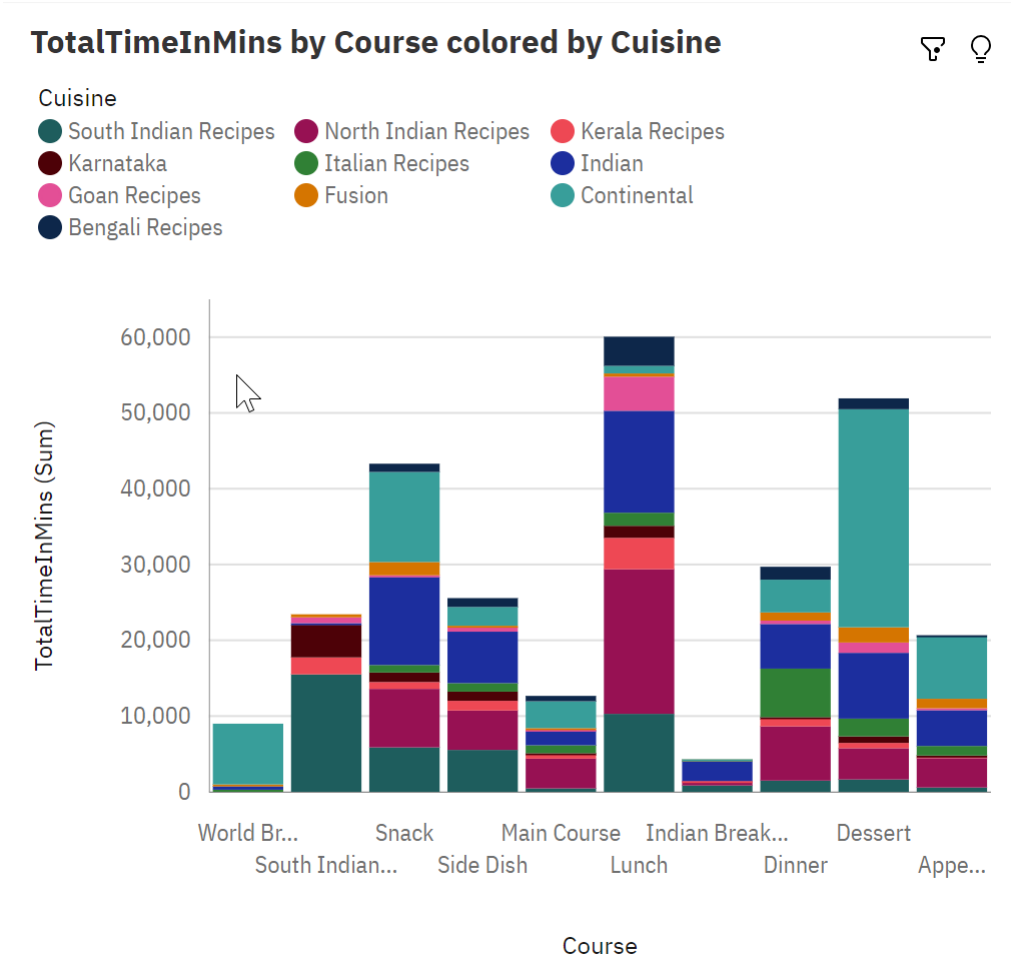
Milestone 7: Story

A data story is a way of presenting data and analysis in a narrative format, with the goal of making the information more engaging and easier to understand. A data story typically includes a clear introduction that sets the stage and explains the context for the data, a body that presents the data and analysis in a logical and systematic way, and a conclusion that summarizes the key findings and highlights their implications. Data stories can be told using a variety of mediums, such as reports, presentations, interactive visualizations, and videos.

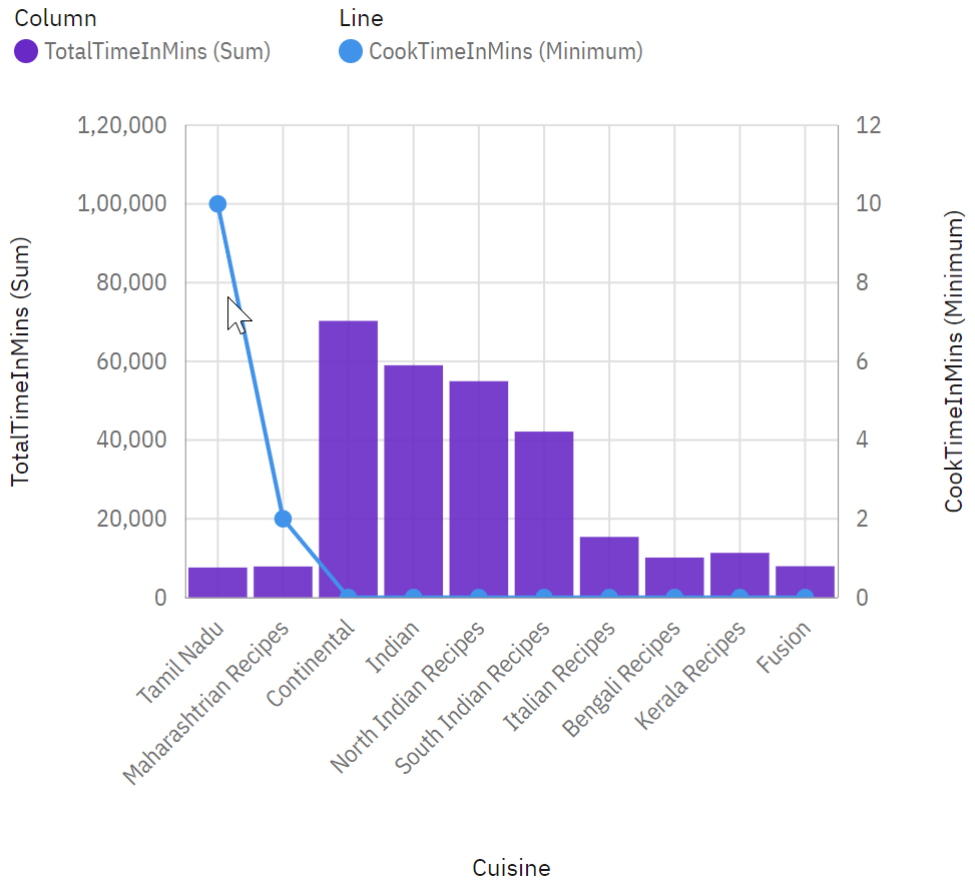
Activity 1: No of Scenes of Story

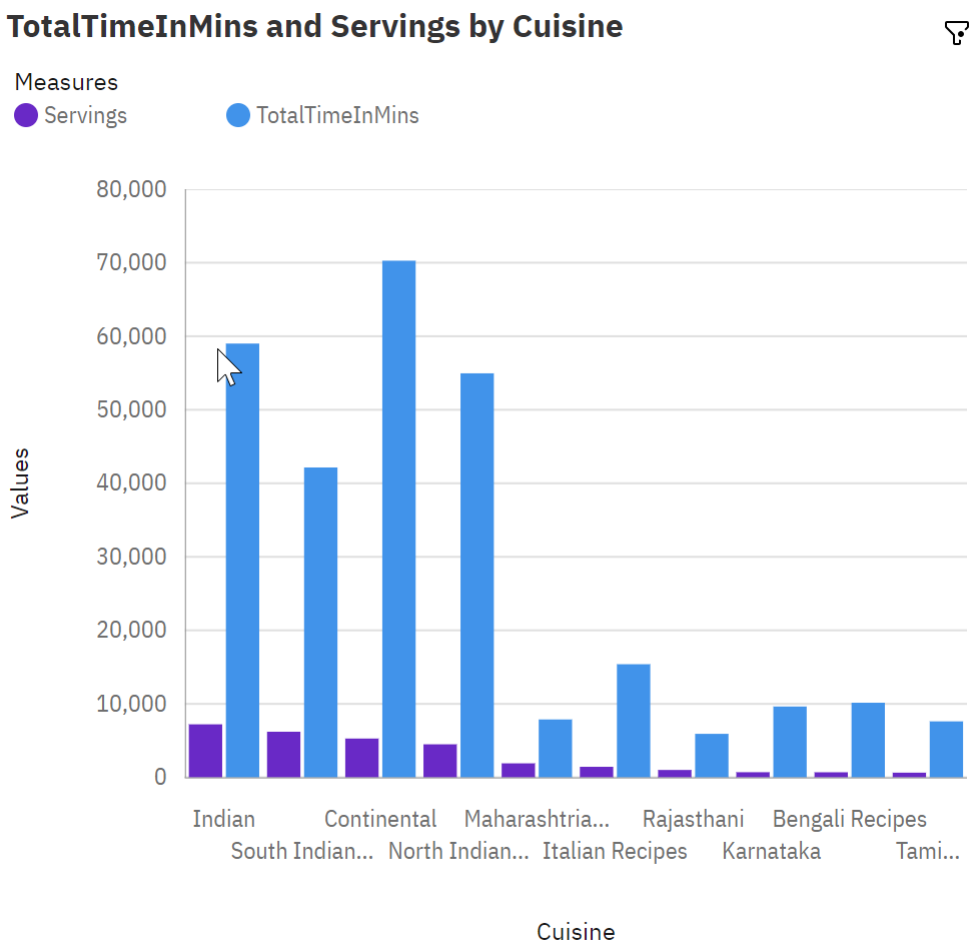
Embark on a captivating culinary journey through the rich and diverse flavors of India with our interactive story based on Exploratory Data Analysis (EDA) of Indian food. Immerse yourself in the vibrant world of Indian cuisine as we uncover the secrets hidden within a comprehensive dataset. From savoring regional delights to exploring popular dishes and understanding the nutritional aspects, this story takes you on an unforgettable gastronomic adventure.

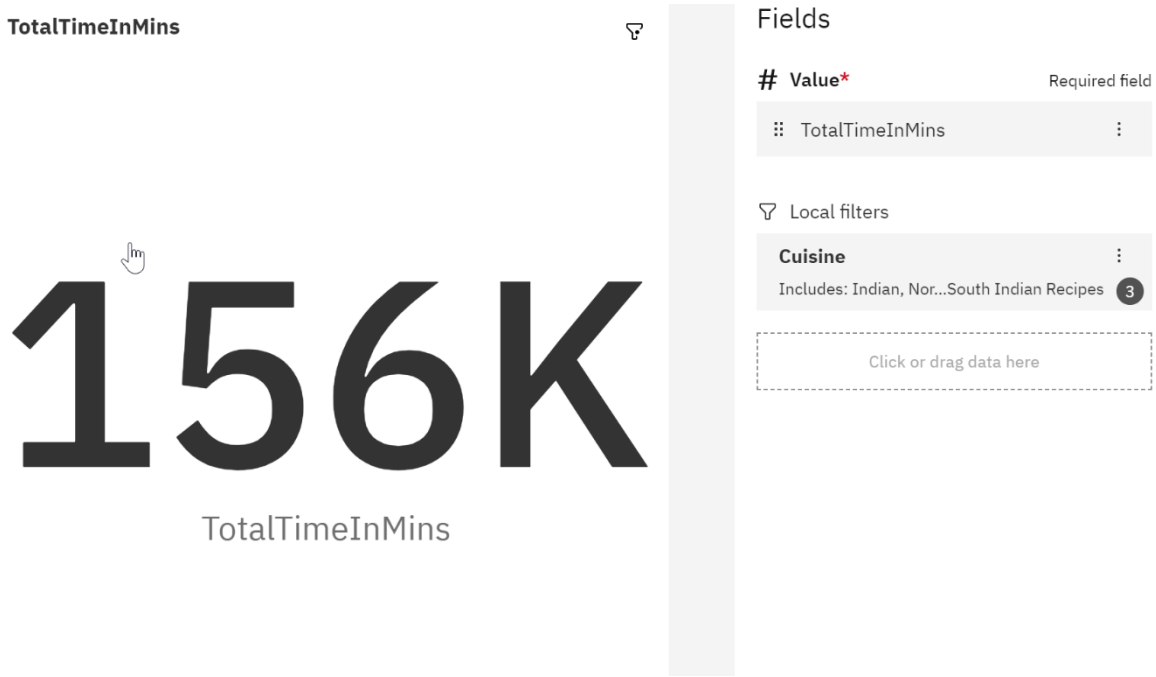
Explanation video link:
https://drive.google.com/file/d/1l9dlN-BsRgnS3weGFS-fknZ5nr2Qbb1E/view?usp=drive_link

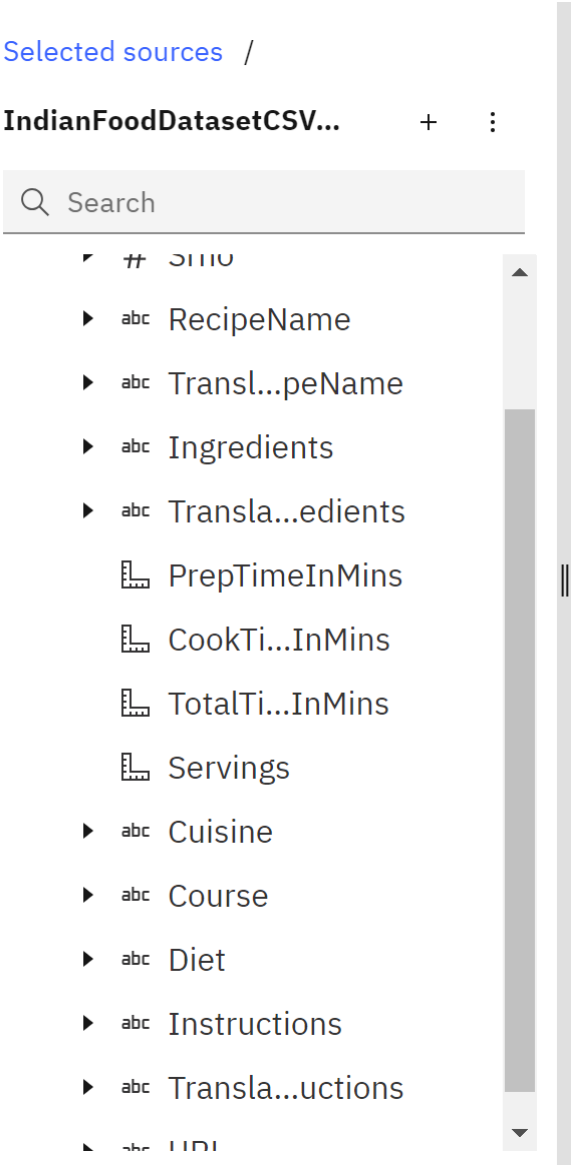

TotalTimeInMins by Course colored by Cuisine
Cuisine
South Indian Recipes
North Indian Recipes
Kerala Recipes
Karnataka
Italian Recipes
Indian
Goan Recipes
Fusion
Continental
Bengali Recipes
TotalTimeInMins (Sum)
60,000
50,000
40,000
30,000
20,000
10,000
0
World Br...
South Indian...
Snack
Side Dish
Main Course
Lunch
Indian Break...
Dinner
Dessert
Appe...
Course

CookTimeInMins and TotalTimeInMins by Cuisine




TotalTimeInMins and Servings by Cuisine
Measures
Servings
TotalTimeInMins
80,000
70,000
60,000
50,000
40,000
30,000
20,000
10,000
0
Values
Indian
South Indian...
Continental
North Indian...
Maharashtria...
Italian Recipes
Rajasthani
Karnataka
Bengali Recipes
Tami...
Cuisine

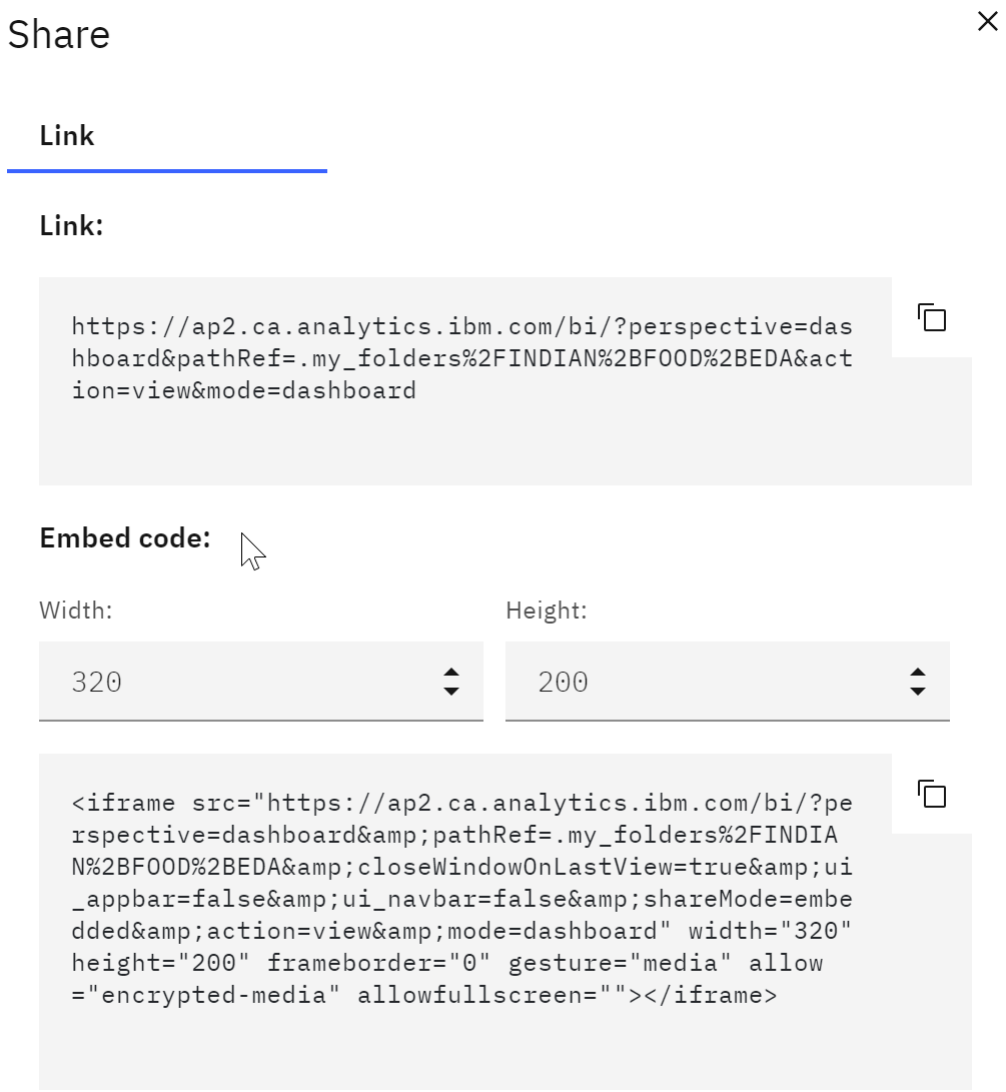

TotalTimeInMins
156K
TotalTimeInMins
Fields
Value*
Required field
TotalTimeInMins
Local filters
Cuisine
Includes: Indian, Nor...South Indian Recipes
3
Click or drag data here



Milestone 9: Web integration

Publishing helps us to track and monitor key performance metrics, to communicate results and progress. help a publisher stay informed, make better decisions, and communicate their performance to others. Integrating dashboard/reports/stories to web

Step 1: Go to Dashboard/story/report, click on share button on the top ribbon



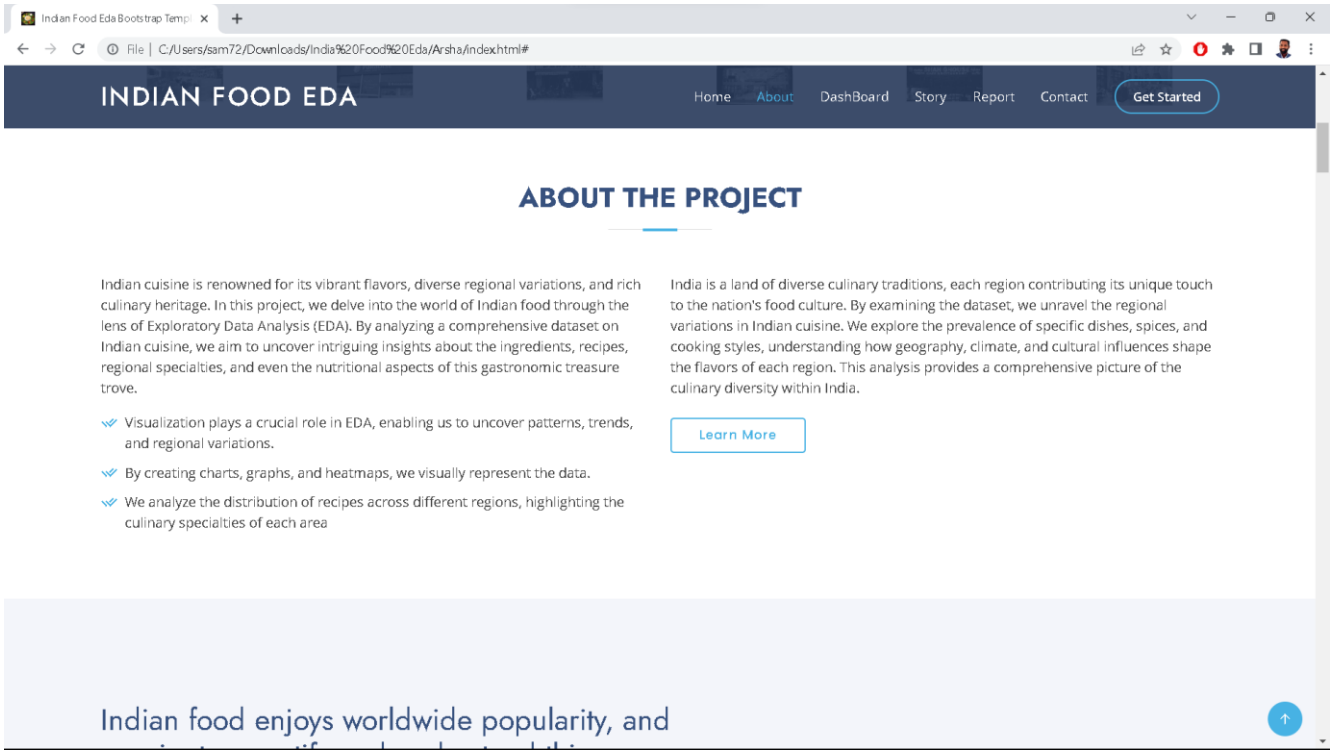
Note: You can also change the width and height of the dashboard/story/report as you like.

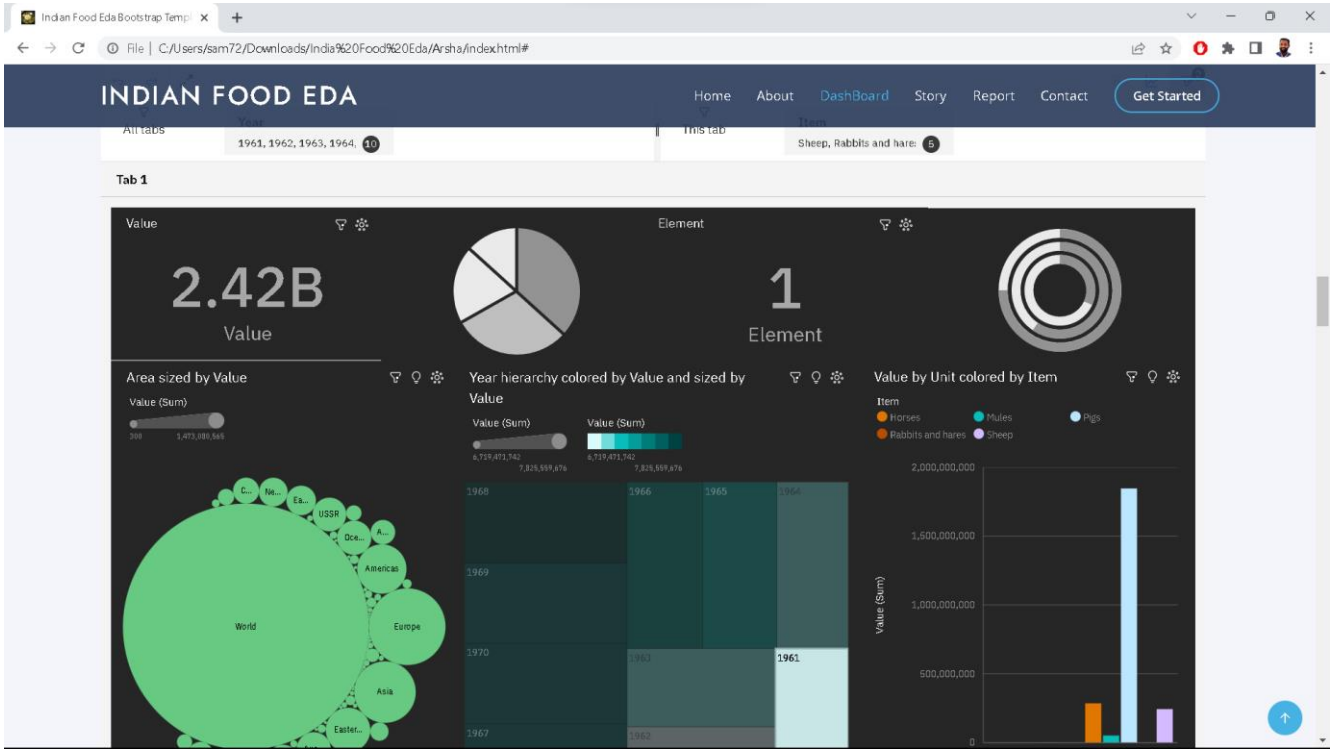
Activity 1: Dashboard and Story embed with UI With Flask

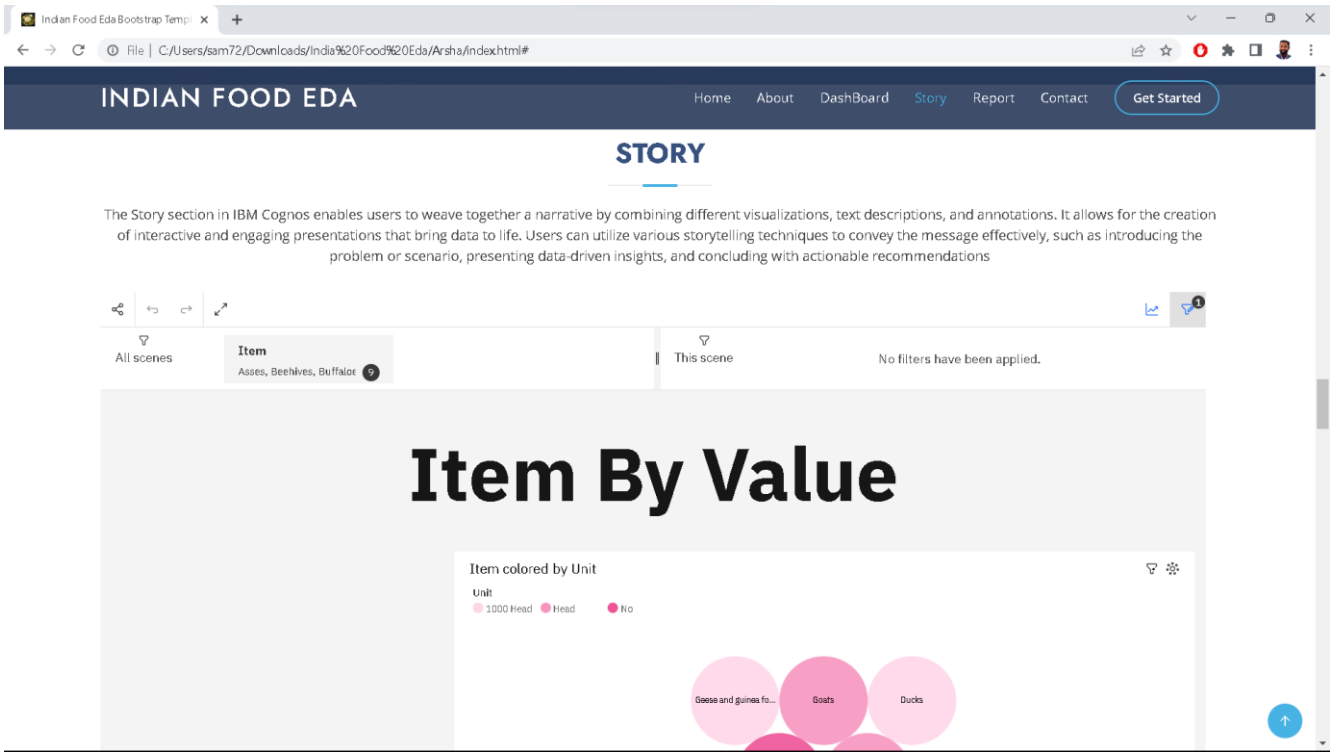
Explanation video link:
https://drive.google.com/file/d/1xnH1CK-8fbdRWwH6U1v8S1_cVV2n5w9e/view?usp=drive_link

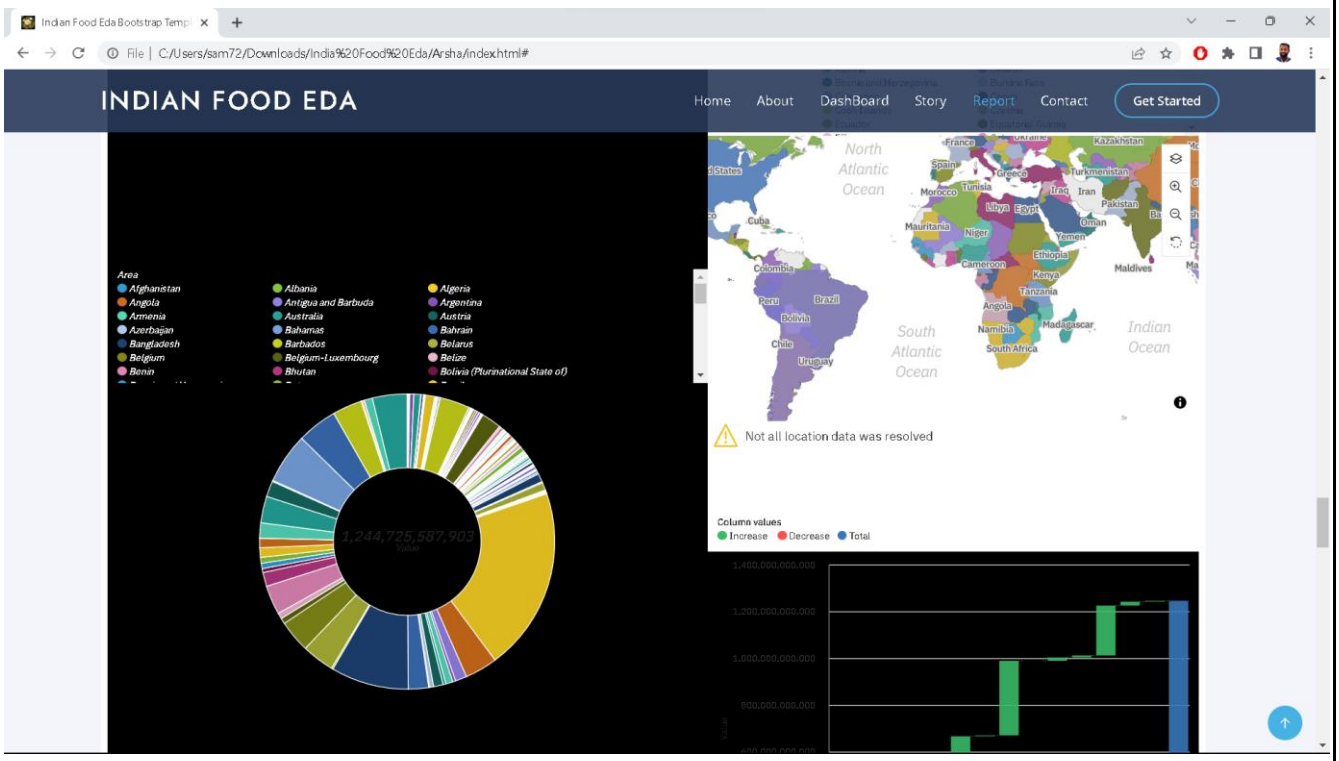
Flask Deployment :- * Running on http://127.0.0.1:1200/ (Press CTRL+C to quit)

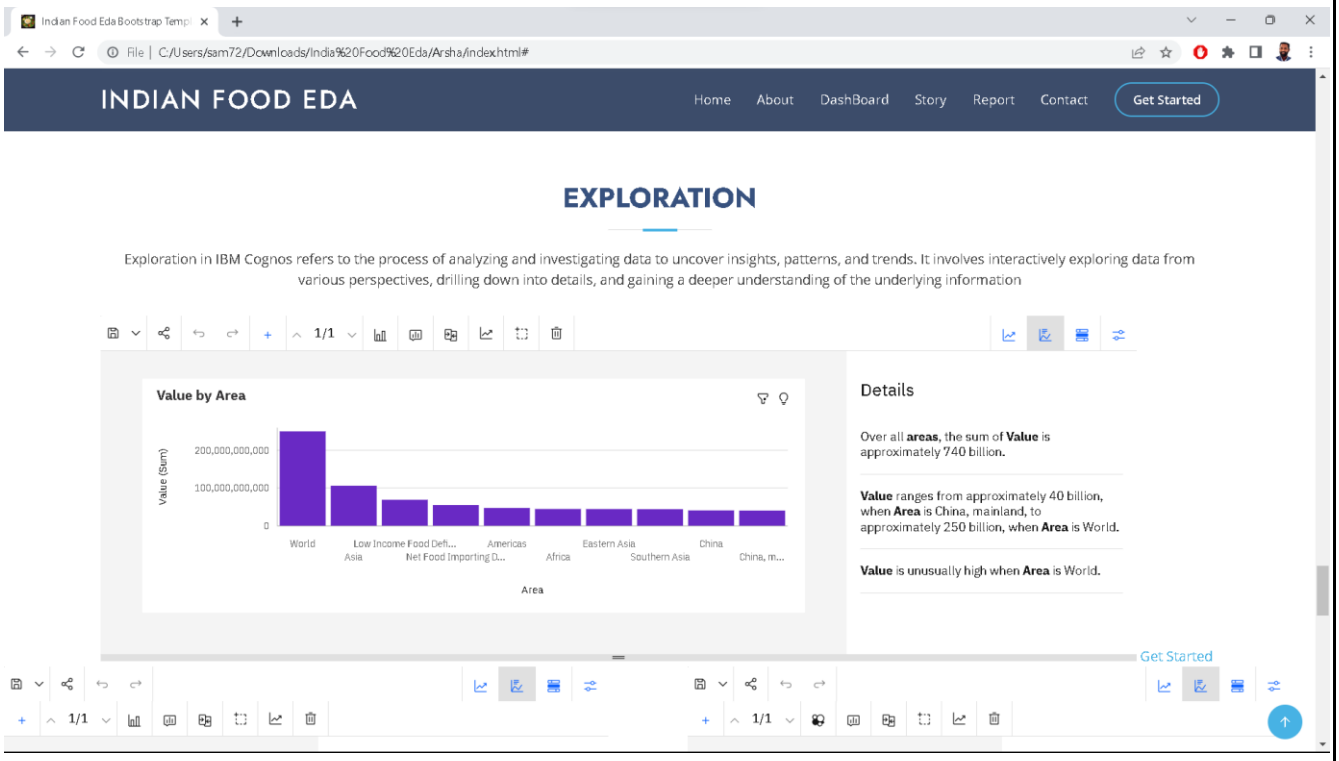

Indian Food Eda Bootstrap Temp
File | C:/Users/sam72/Downloads/India%20Food%20Eda/Arsha/index.html#
INDIAN FOOD EDA
Home
About
DashBoard
Story
Report
Contact
Get Started
Unveiling the Flavors, Regional Delights, and Popularity of Indian Cuisine
Nutritional Aspects of Indian Cuisine through Exploratory Data Analysis
Get Started
Watch Video

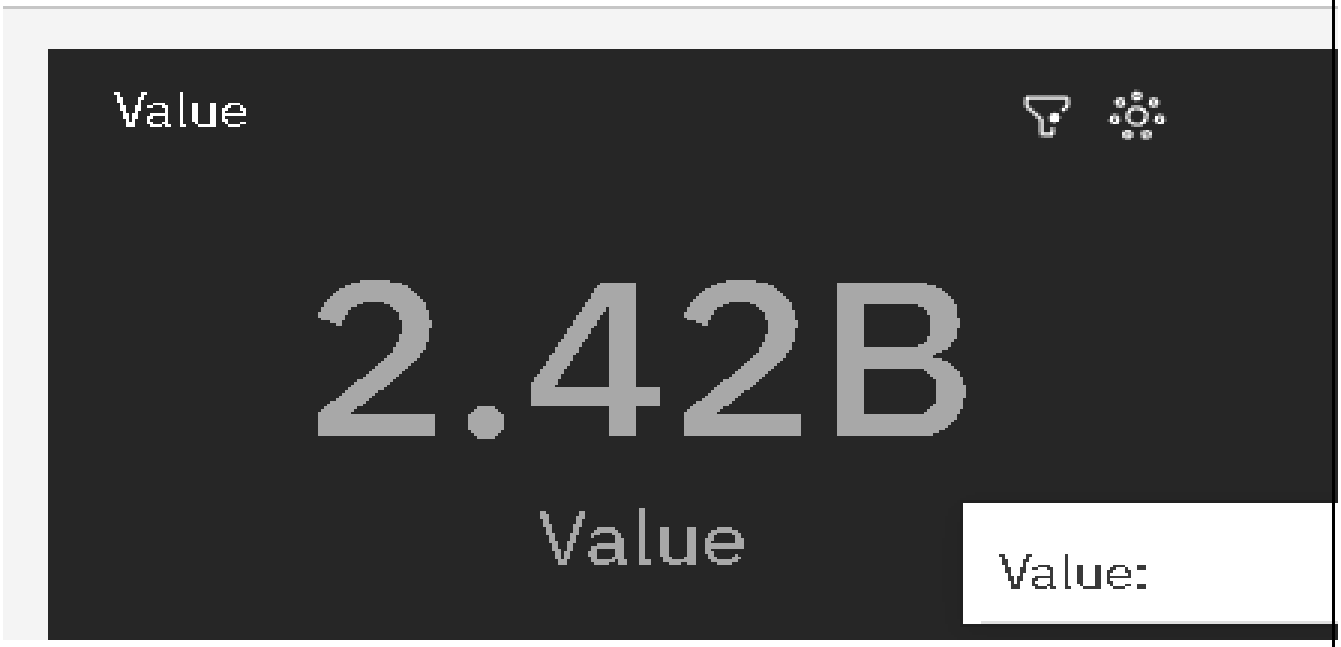

Indian Food Eda Bootstrap Temp
File | C:/Users/sam72/Downloads/India%20Food%20Eda/Arsha/index.html#
INDIAN FOOD EDA
Home
About
DashBoard
Story
Report
Contact
Get Started
ABOUT THE PROJECT
Indian cuisine is renowned for its vibrant flavors, diverse regional variations, and rich culinary heritage. In this project, we delve into the world of Indian food through the lens of Exploratory Data Analysis (EDA). By analyzing a comprehensive dataset on Indian cuisine, we aim to uncover intriguing insights about the ingredients, recipes, regional specialties, and even the nutritional aspects of this gastronomic treasure trove.
Visualization plays a crucial role in EDA, enabling us to uncover patterns, trends, and regional variations.
By creating charts, graphs, and heatmaps, we visually represent the data.
We analyze the distribution of recipes across different regions, highlighting the culinary specialties of each area
India is a land of diverse culinary traditions, each region contributing its unique touch to the nation's food culture. By examining the dataset, we unravel the regional variations in Indian cuisine. We explore the prevalence of specific dishes, spices, and cooking styles, understanding how geography, climate, and cultural influences shape the flavors of each region. This analysis provides a comprehensive picture of the culinary diversity within India.
Learn More
Indian food enjoys worldwide popularity, and


Indian Food Eda Bootstrap Temp
File | C:/Users/sam72/Downloads/India%20Food%20Eda/Arsha/index.html#
INDIAN FOOD EDA
Home
About
DashBoard
Story
Report
Contact
Get Started
All tabs
1961, 1962, 1963, 1964,
This tab
Sheep, Rabbits and hares
Tab 1
Value
2.42B
Value
Element
1
Element
Area sized by Value
Value (Sum)
Year hierarchy colored by Value and sized by Value
Value (Sum)
Value (Sum)
Value by Unit colored by Item
Item
Horses
Mules
Pigs
Rabbits and hares
Sheep
2,000,000,000
1,500,000,000
1,000,000,000
500,000,000
Value (Sum)
World
Europe
Asia
Americas
USSR
1968
1966
1965
1969
1970
1961
1967


Indian Food Eda Bootstrap Temp
File | C:/Users/sam72/Downloads/India%20Food%20Eda/Arsha/index.html#
INDIAN FOOD EDA
Home
About
DashBoard
Story
Report
Contact
Get Started
STORY
The Story section in IBM Cognos enables users to weave together a narrative by combining different visualizations, text descriptions, and annotations. It allows for the creation of interactive and engaging presentations that bring data to life. Users can utilize various storytelling techniques to convey the message effectively, such as introducing the problem or scenario, presenting data-driven insights, and concluding with actionable recommendations
All scenes
Item
Asses, Beehives, Buffalos
This scene
No filters have been applied.
Item By Value
Item colored by Unit
Unit
1000 Head
Head
No
Geese and guinea fo...
Goats
Ducks


Indian Food Eda Bootstrap Temp
File | C:/Users/sam72/Downloads/India%20Food%20Eda/Arsha/index.html#
INDIAN FOOD EDA
Home
About
DashBoard
Story
Report
Contact
Get Started
Area
Afghanistan
Albania
Algeria
Angola
Antigua and Barbuda
Argentina
Armenia
Australia
Austria
Azerbaijan
Bahamas
Bahrain
Bangladesh
Barbados
Belarus
Belgium
Belgium-Luxembourg
Belize
Benin
Bhutan
Bolivia (Plurinational State of)
1,244,725,587,903
Value
North Atlantic Ocean
South Atlantic Ocean
Indian Ocean
Not all location data was resolved
Column values
Increase
Decrease
Total


Indian Food Eda Bootstrap Temp
File | C:/Users/sam72/Downloads/India%20Food%20Eda/Arsha/index.html#
INDIAN FOOD EDA
Home
About
DashBoard
Story
Report
Contact
Get Started
EXPLORATION
Exploration in IBM Cognos refers to the process of analyzing and investigating data to uncover insights, patterns, and trends. It involves interactively exploring data from various perspectives, drilling down into details, and gaining a deeper understanding of the underlying information
Value by Area
Value (Sum)
200,000,000,000
100,000,000,000
0
World
Asia
Low Income Food Defi...
Net Food Importing D...
Americas
Africa
Eastern Asia
Southern Asia
China
China, m...
Area
Details
Over all areas, the sum of Value is approximately 740 billion.
Value ranges from approximately 40 billion, when Area is China, mainland, to approximately 250 billion, when Area is World.
Value is unusually high when Area is World.
Get Started



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.