INDIAN FOOD EXPLORATION DATA ANALYSIS

Presented by

KEERTHIVASAN N	(611220104070)
MEIROOPASRI R	(611220104118)
POORNASRI S	(611220104119)
REVATHI A	(611220104122)

UNDER THE GUIDANCE OF

Mr. J. MURUGESAN (IT)

ASSISTANT PROFESSOR,

DEPARTMENT OF COMPUTER SCIENCE AND ENGINEERING

CSE

CSE

INTRODUCTION

- Data analytics is a method of applying quantitative and qualitative techniques to analyze data, aiming for valuable insights. With the help of data analytics, we can explore data and we can even draw conclusions about our data.

- 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.

ABSTRACT

- 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 unraveling the intricacies of Indian food, shedding light on the diversity, flavors, 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.

LITERATURE SURVEY

PAPER-I

TITLE: Exploratory Data Analysis of Indian Cuisine Ingredients
Exploratory Data Analysis of Indian Cuisine Ingredients

AUTHOR: Gupta et al, 2018

This study focuses on analyzing a dataset of Indian cuisine ingredients using EDA techniques. The authors explore the frequency and distribution of ingredients, identifying the most commonly used spices, herbs, and vegetables. They also investigate the regional variations in ingredient preferences, highlighting the distinct flavor profiles across different Indian states.

LITERATURE SURVEY

PAPER-II

TITLE: Exploring Regional Variations in Indian Food: An EDA Approach

AUTHOR: Sharma and Jain ,2019

This research delves into regional variations in Indian food by employing EDA techniques on a dataset comprising recipes from different states. The authors examine the prevalence of specific dishes, ingredients, and cooking methods within each region. They uncover regional specialties, uncovering the unique culinary traditions that define Indian cuisine.

LITERATURE SURVEY

PAPER-III

TITLE: EDA of Popular Indian Dishes on Social Media"

AUTHOR: Patel and Shah ,2020

CSE

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.

LITERATURE SURVEY

PAPER-IV

TITLE : Nutritional Analysis of Indian Food: An EDA Perspective

AUTHOR: Chatterjee and Saha ,2021

CSE

Focusing on the nutritional aspects of Indian cuisine, this research conducts EDA on a dataset comprising nutritional information of Indian dishes. The authors analyze the calorie content, macronutrient distribution, and common ingredients contributing to specific nutritional values. They highlight the balance and diversity of Indian food, providing valuable insights for health-conscious individuals.

DRAWBACKS IN EXISTING SYSTEM

- The user interface may not be the most intuitive, we feel that there is room for improvement.
- Its data visualization features require configuration to integrate with third-party tools.
- Errors and fake analysis of literacy rate may be occur.
- Lack of access to computers and the internet or who limited digital literacy skills may face disadvantages in education and the job market.

PROPOSED SOLUTION

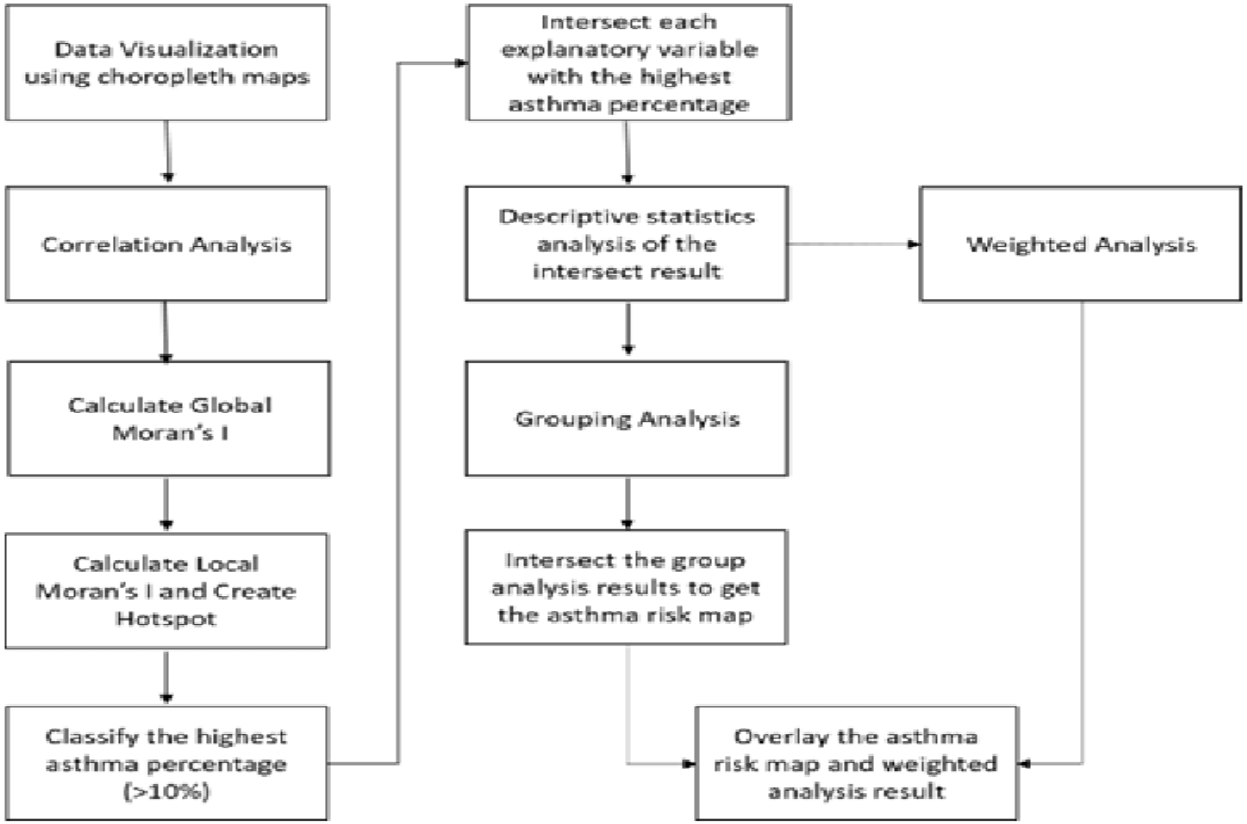
- Indian food EDA can be a valuable resource for both enthusiasts of Indian cuisine and professionals in the food industry..
- It can help users discover new recipes, restaurants, and food trends while offering insights into the rich culinary traditions of India.
- In our effort, we have tried to predict the Literacy Rates of each state using a reduced set of features.
- The scalability of our project is very huge, especially covering the people on the different types of food and Ingriediants

ADVANTAGES

- Lower costs—reduces maintenance due to complete report coverage and a zero-footprint environment.
- Faster results—shorten reporting time due to seamless integration and adaptive authoring.
- Improved decision-making—reports and dashboards present data in easily-understood formats.
- Ability to use a variety of charts—crosstabs, bar or 3D bar, pie or doughnut, line, gauge, funnel, scatter, dot density, waterfall, and so forth.

CSE

DATAFLOW DIAGRAM



SYSTEM SPECIFICATION

HARDWARE USED:

Processor -AMD/INTEL

RAM -4GB

Hard Disk -256 GB

SOFTWARE USED:

Language - HTML, CSS and Javascript

Framework - Flask

Package Manager & Build Tool - PIP

Database - IBM_DB2

MODULES

1. Home

- Home page of the Project.

2. About

- About the Project Web Page.

3. Analysis Page

- Dash Board Page
- Report Page
- Stories Page

MODULE DESCRIPTION

1. Home:

User will know about the site using the home page. It contains other navigation page like services page, team page, about page and analysis page. It act like starting page of the website.

2. About :

In this, it contains about the services provided my the web site. It also contains what food rate analysis have done, past analysis rate surveys , food service details etc...

MODULE DESCRIPTION

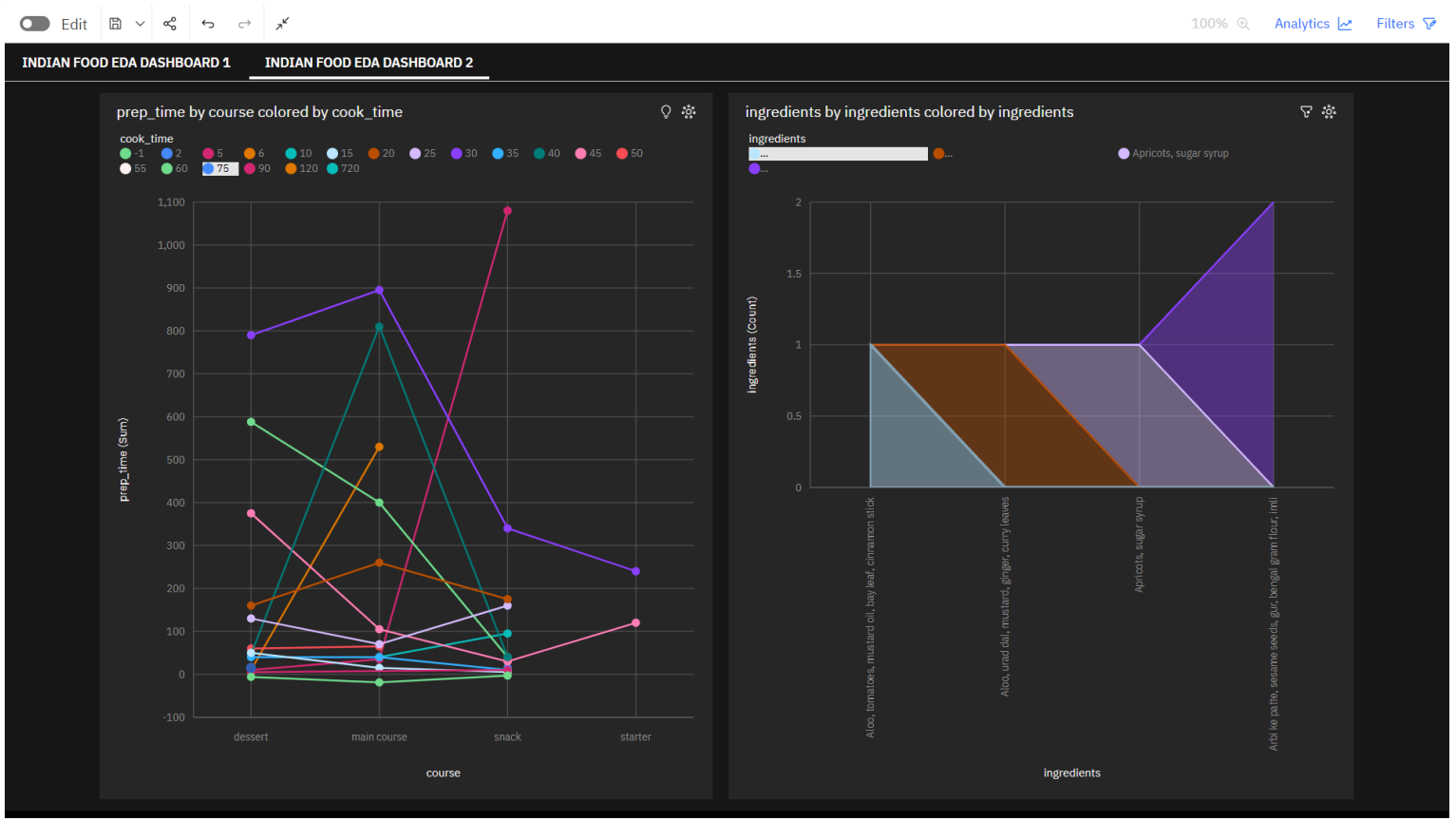
3. Analysis Page : It contains about dash boards ,report and stories of the literacy rate analysis, their rate percentage , food based analysis rate , state and area based food analysis.

Dashboards : Here we can see the visualization of the survey which is more understandable in this project.

Reports : Here we will see about the progress made by the project and working of the process

RESULT AND DISCUSSION

DASHBOARDS

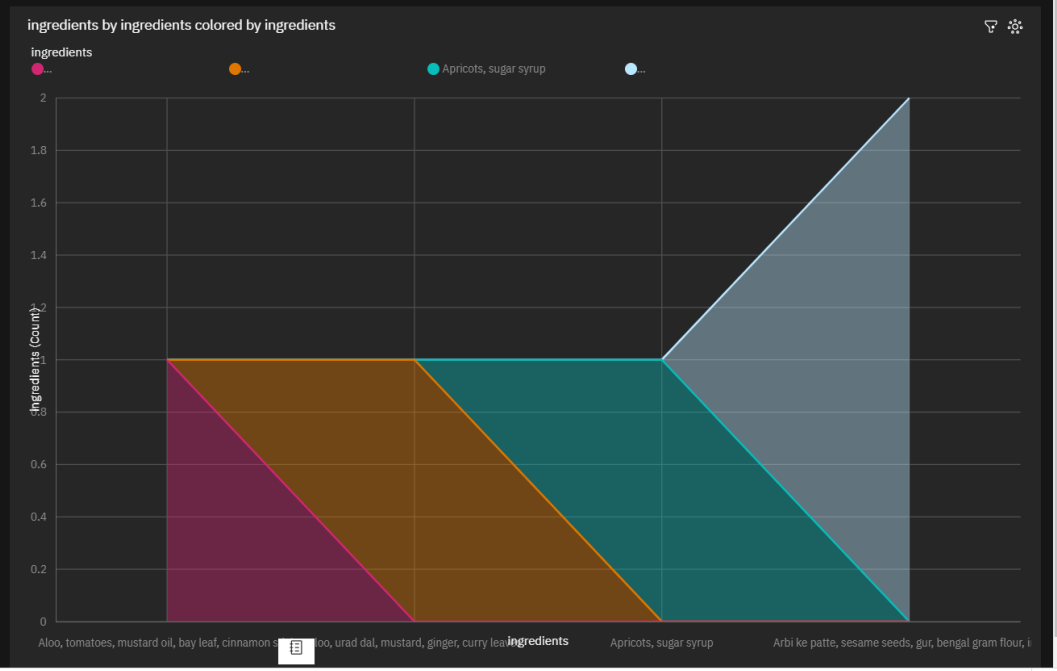

Edit
100%
Analytics
Filters
INDIAN FOOD EDA DASHBOARD 1
INDIAN FOOD EDA DASHBOARD 2
prep_time by course colored by cook_time
cook_time
-1 2 5 6 10 15 20 25 30 35 40 45 50
55 60 75 90 120 720
prep_time (Sum)
1,100
1,000
900
800
700
600
500
400
300
200
100
0
-100
dessert
main course
snack
starter
course
ingredients by ingredients colored by ingredients
ingredients
Apricots, sugar syrup
ingredients (count)
2
1.5
1
0.5
0
Aloo, tomatoes, mustard oil, bay leaf, cinnamon stick
Aloo, urad dal, mustard, ginger, curry leaves
Apricots, sugar syrup
Arbi ke patte, sesame seeds, gur, bengal gram flour, imli
ingredients

CSE

Edit

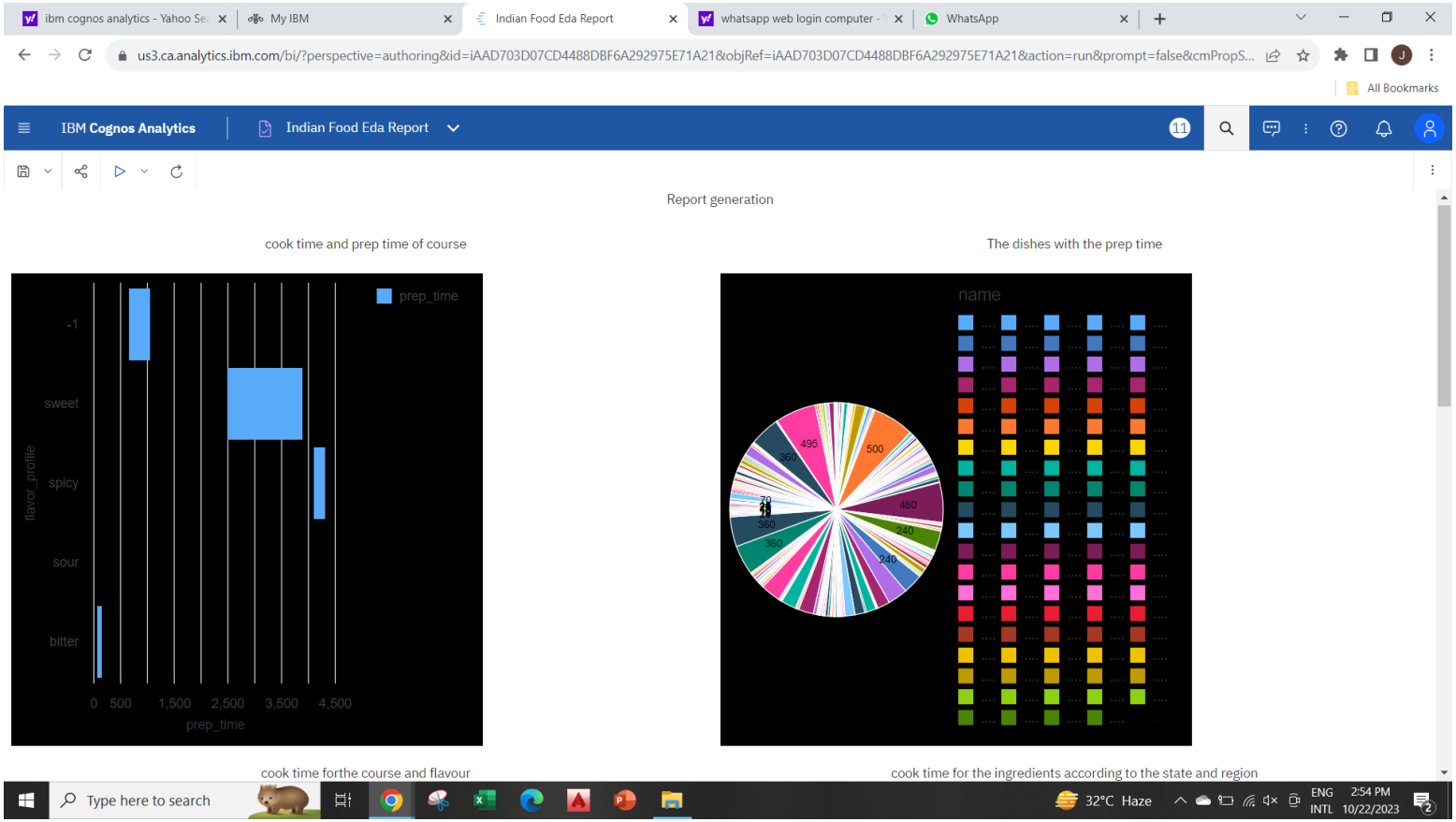
Analytics Filters

- Arbi ke patte, sesame seeds, gur, bengal gram flour, imli is the most frequently occurring category of ingredients with a count of 2 items with ingredients values (40 % of the total).
- The total number of results for ingredients, across all ingredients, is 5.



Prev scene Next scene Scene 5 of 9 0:01.4 0:05.0

RESULTS & REPORTS

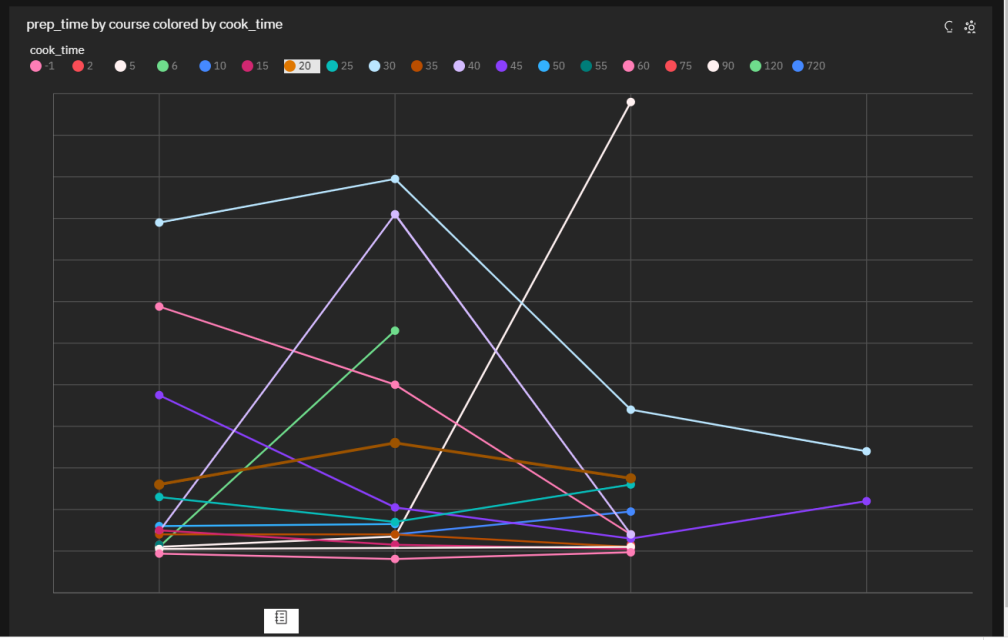


CSE

Edit

Analytics Filters

- Cook_time 30 has the highest total prep_time due to state Punjab.
- prep_time is unusually high when the combinations of course and cook_time are snack and 90, main course and 30, main course and 40 and dessert and 30.
- prep_time is unusually high when cook_time is 30.
- prep_time is unusually low when course is starter.
- snack has a prep_time of over a thousand for cook_time 90.
- 90 has a prep_time of over a thousand for state -1.
- Punjab state accounted for 36% of main course prep_time compared to 1% for dessert.
- cook_time 30 has the highest prep_time at 2265, out of which course main course contributed the most at 895.
- state Punjab has the highest prep_time at 1560, out of which cook_time 30 contributed the most at 725.
- state Punjab has the highest prep_time at over 1500, out of which course main course contributed the most at over a thousand.
- course main course prep_time from cook_time 30 is 895, whereas dessert is only 790.
- course main course has the highest total prep_time due to cook_time 30.
- course main course has the highest total prep_time due to state Punjab.
- Over all values of course and cook_time, the sum of prep_time is nearly eight thousand.
- The summed values of prep_time range from -19 to over a thousand.
- For prep_time, the most significant value of cook_time is 30, whose respective prep_time values add up to almost 2500, or 28.5 % of the total.
- For prep_time, the most significant values of course are main course, dessert, and snack, whose respective prep_time values add up to over 7500, or 95.5 % of the total.



Prev scene Next scene Scene 7 of 9 0:02.5 0:05.0

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.

- The dataset is uploaded to the database using the IBM cloud and the database is connected to the IBM Cognos Analytics tool. In the IBM Cognos Analytics tool, the data module option is selected and the dataset is pre-processed

CSE

THANK YOU