NAAN MUDHALVAN PROJECT

ANNA UNIVERSITY COLLEGE OF ENGINEERING PATTUKOTTAI

DEPARTMENT OF COMPUTER SCIENCE ENGINEERING

TOPIC: INDIAN FOOD EDA

TEAM ID : 3CFDCF419BD3A95BBA14C9531550BCBC

Project submitted by,

TEAM LEADER:

Arunkumar.K – 822120104008

Team members

Muthukumar .K– 822120104027

Balaji.-822120104009

Abishek.K. -822120104002

ABSTRACT:

Consumers are showing an increasing interest in obtaining additional benefits from food. Biofunctional foods have emerged on the market as a possible solution for such demands. Their composition and biological activity are such that – beyond their inherent nutritional value – promote beneficial effects on one or more target functions/organs in the body. Some of the biologically active compounds which are naturally found or added in functional foods (biofunctional constituents) are already associated with approved – by the European Food Safety Authority, EFSA – health claims. Scientific evidence supports chemical, toxicological and technological characterization of bioactive compounds as well as detailed description of their biological mechanisms. Biofunctional food is a modern approach towards a healthier diet which needs strong scientific support. Food scientists are keys elements for the development of this new serious of products and need to be constantly updated on the latest research findings. The e- Food science is an Erasmus + project which aims to use the potential of Europe's human and social capital originating from higher education institutes and the food industry, in order to develop innovative training material based on selected research findings originated from the participating institutions and also to exchange and transfer knowledge and know-how in food science and technology education in Europe. The objectives will be achieved by designing, developing and pilot testing freely accessible online educational material, for a common group of modules intended for

current and potential food. The present work covers, apart of some theoretical approaches, specific analytical methods and test experiments for the incorporation in the biofunctional quivers of the food scientist and consumer of ingredients, originated from Hippophaes rhamnoides added in wine-based products.

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 OVERVIEW:

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.

BRAINSTORM & IDEA PRIORITIZATION

Our country India is well known for its diversity in culture as well as for its large number of varieties in dishes.

This project is all about analysis of Indian-Food-Dishes data:

- Regional analysis of Indian dishes.
- Statewise analysis of Indian dishes.
- Flavor i.e sweet, sour, spicy of different Indian dishes.
- Veg/Non-veg analysis of Indian dishes.

Tools & techniques used in this project:

- A good dataset having all the relevant informations regarding Indian dishes.
- Python pandas library to convert the raw data into workable data frames.
- Python matplotlib & seaborn library for data visualization

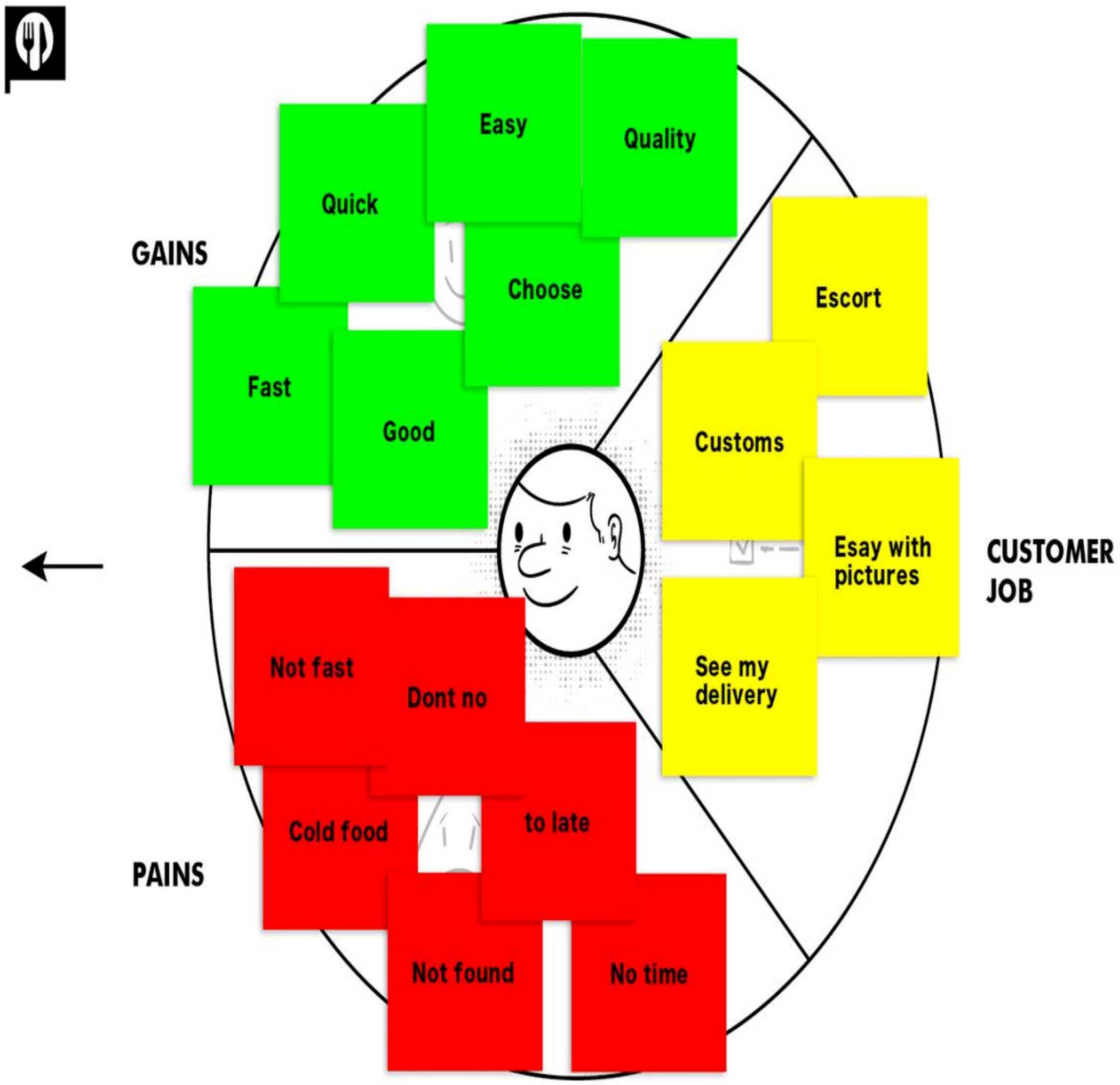
Downloading the Dataset:

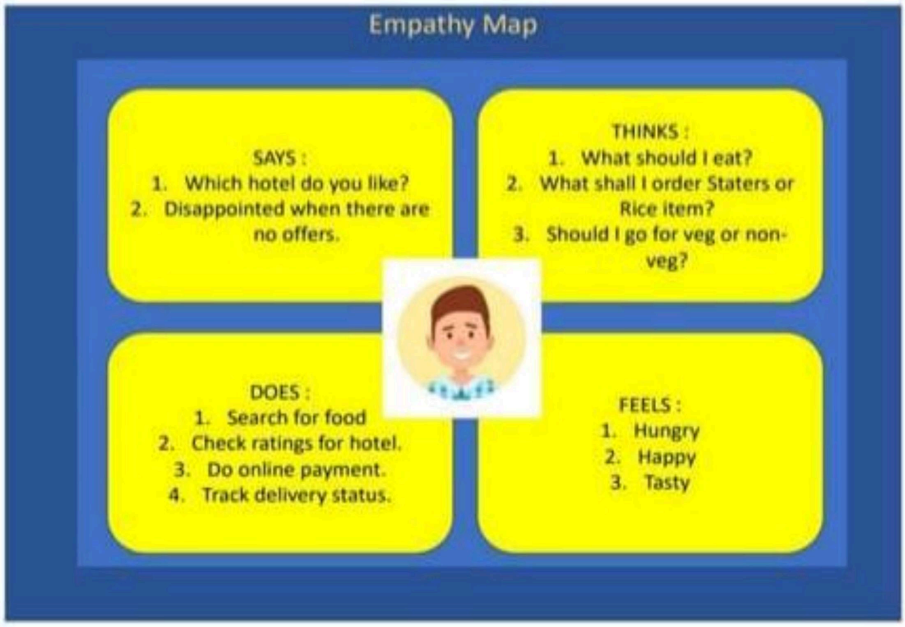
- Using opendatasets library
- Providing the URL to download the dataset
- To download the dataset .download method is used

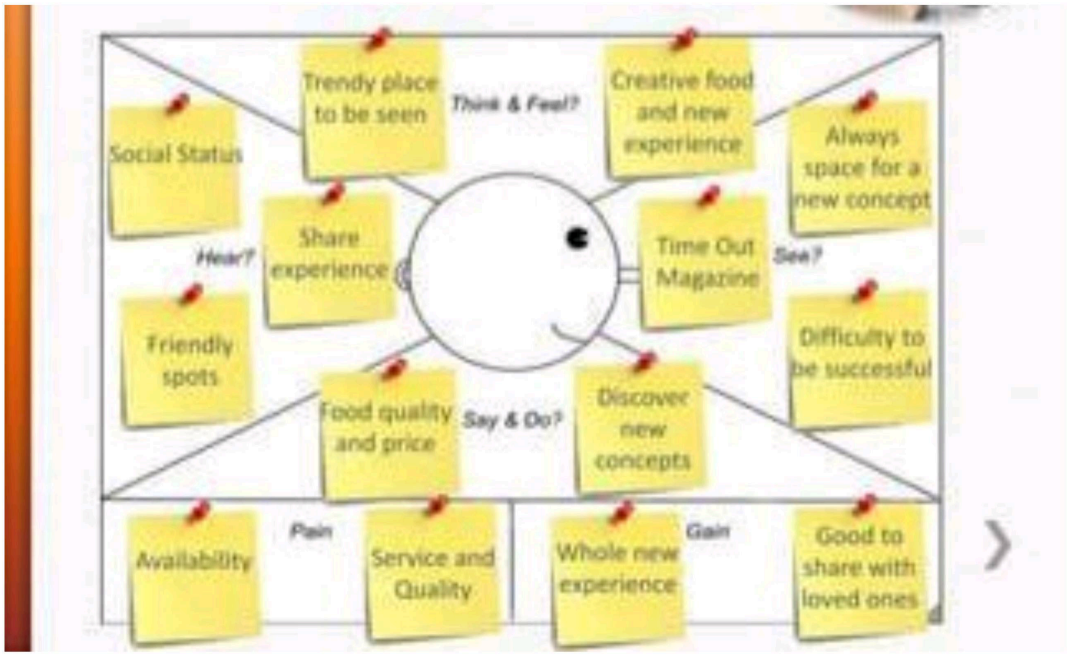
Literature review

The Popularity in classification of Indian Food is gaining slowly due to the awareness of food and health among people.As indicated by the World Health Organization (WHO) [5,14], more than 1.9 billion adults (18 years above) were overweight.It is terribly stunning to understand that 13% of the total populace includes both women and men (15% women and 11% men) are overweight.In reality, some of individualsover the globe are suffered from overweight, which has doubled since 1980. As a result, it shows that foodhas played animportant role in fitness of an individual.

EMPATHY MAPS

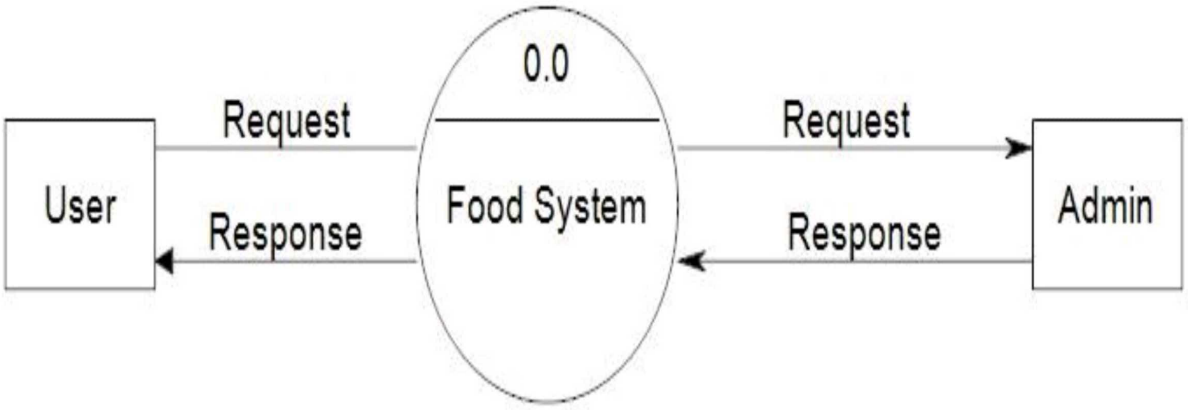



Empathy Map
SAYS :
1. Which hotel do you like?
2. Disappointed when there are no offers.
THINKS :
1. What should I eat?
2. What shall I order Staters or Rice item?
3. Should I go for veg or non-veg?
DOES :
1. Search for food
2. Check ratings for hotel.
3. Do online payment.
4. Track delivery status.
FEELS :
1. Hungry
2. Happy
3. Tasty

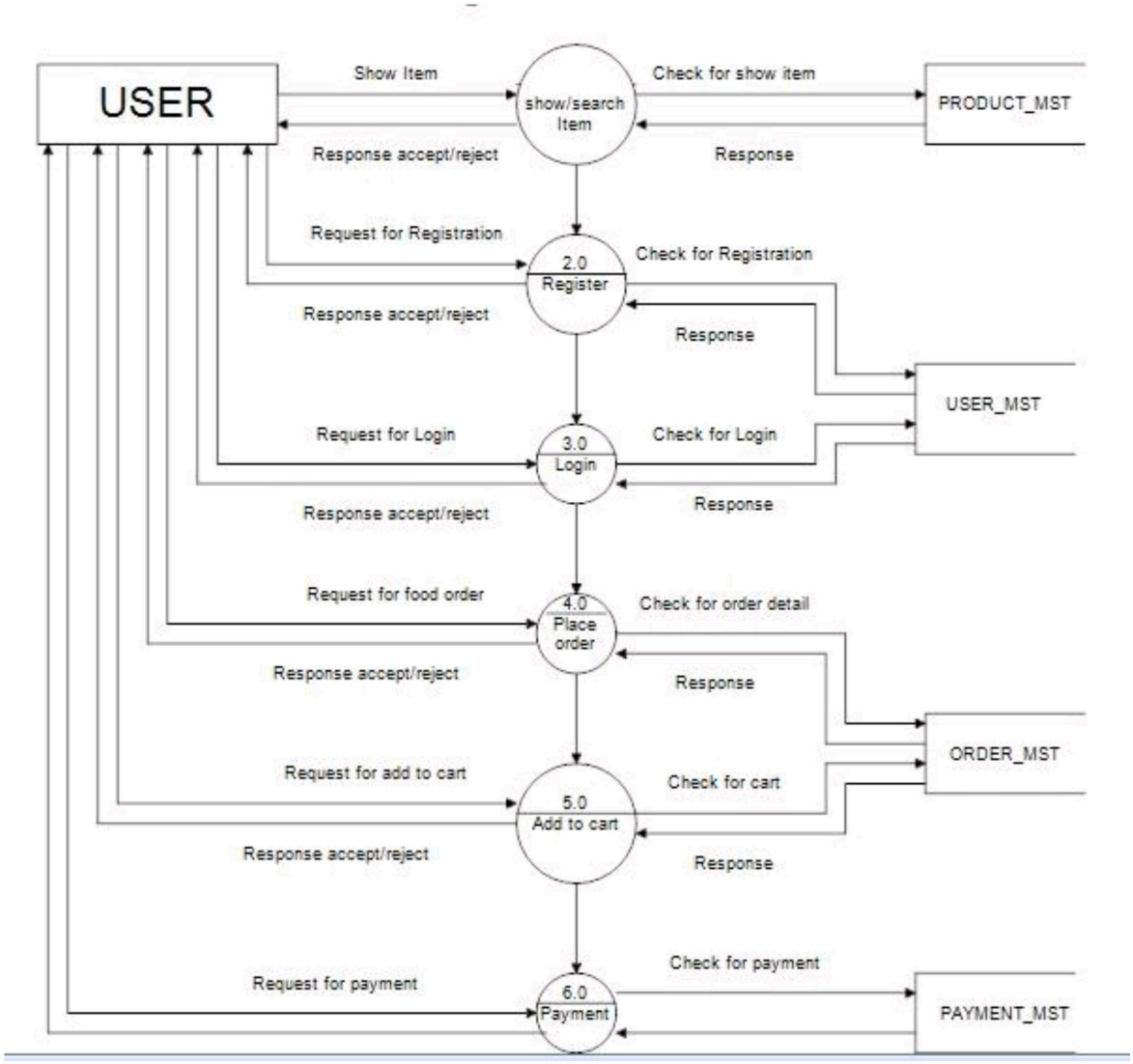

Trendy place to be seen
Think & Feel?
Creative food and new experience
Social Status
Always space for a new concept
Share experience
Time Out Magazine
Hear?
See?
Friendly spots
Difficulty to be successful
Food quality and price
Say & Do?
Discover new concepts
Pain
Gain
Availability
Service and Quality
Whole new experience
Good to share with loved ones

DATA FLOW DIAGRAM & USER STORIES

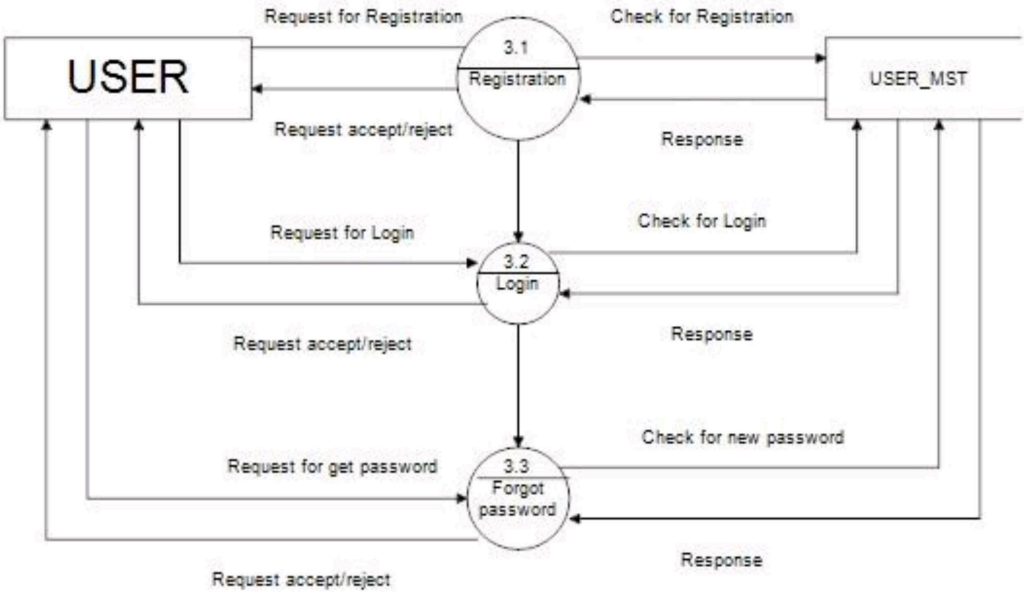
DFD level1:



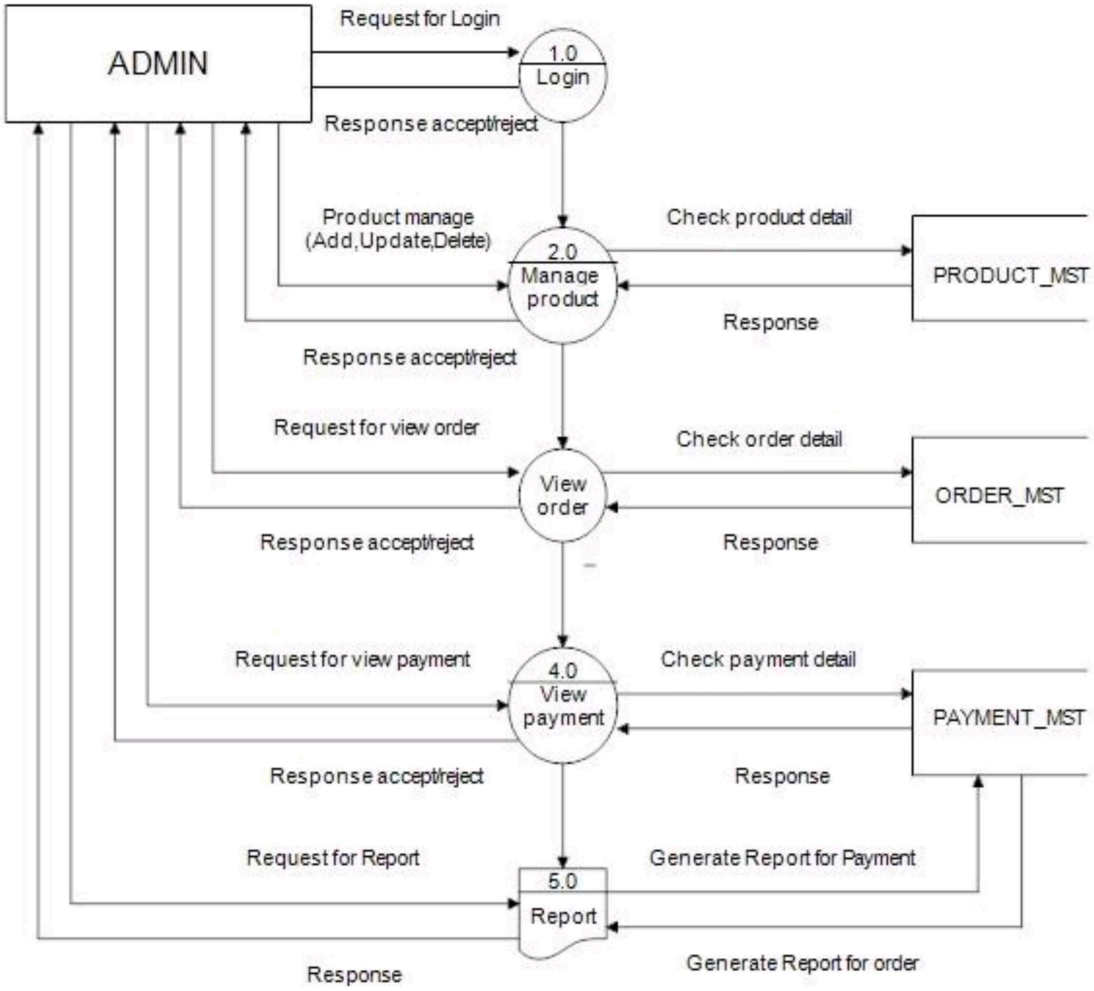
DFD level2:


USER
Show Item
show/search Item
Check for show item
PRODUCT_MST
Response accept/reject
Response
Request for Registration
2.0
Register
Check for Registration
Response accept/reject
Response
USER_MST
Request for Login
3.0
Login
Check for Login
Response accept/reject
Response
Request for food order
4.0
Place order
Check for order detail
Response accept/reject
Response
ORDER_MST
Request for add to cart
5.0
Add to cart
Check for cart
Response accept/reject
Response
Check for payment
Request for payment
6.0
Payment
PAYMENT_MST

DFD level3:


USER
3.1
Registration
USER_MST
Request for Registration
Check for Registration
Request accept/reject
Response
Request for Login
Check for Login
3.2
Login
Request accept/reject
Response
Check for new password
Request for get password
3.3
Forgot password
Response
Request accept/reject

DFD level4:


ADMIN
Request for Login
1.0
Login
Response accept/reject
Product manage
(Add,Update,Delete)
2.0
Manage
product
Check product detail
PRODUCT_MST
Response
Response accept/reject
Request for view order
View
order
Check order detail
ORDER_MST
Response accept/reject
Response
Request for view payment
4.0
View
payment
Check payment detail
PAYMENT_MST
Response accept/reject
Response
Request for Report
5.0
Report
Generate Report for Payment
Response
Generate Report for order

SOLUTION ARCHITECTURE

Solution Architecture:

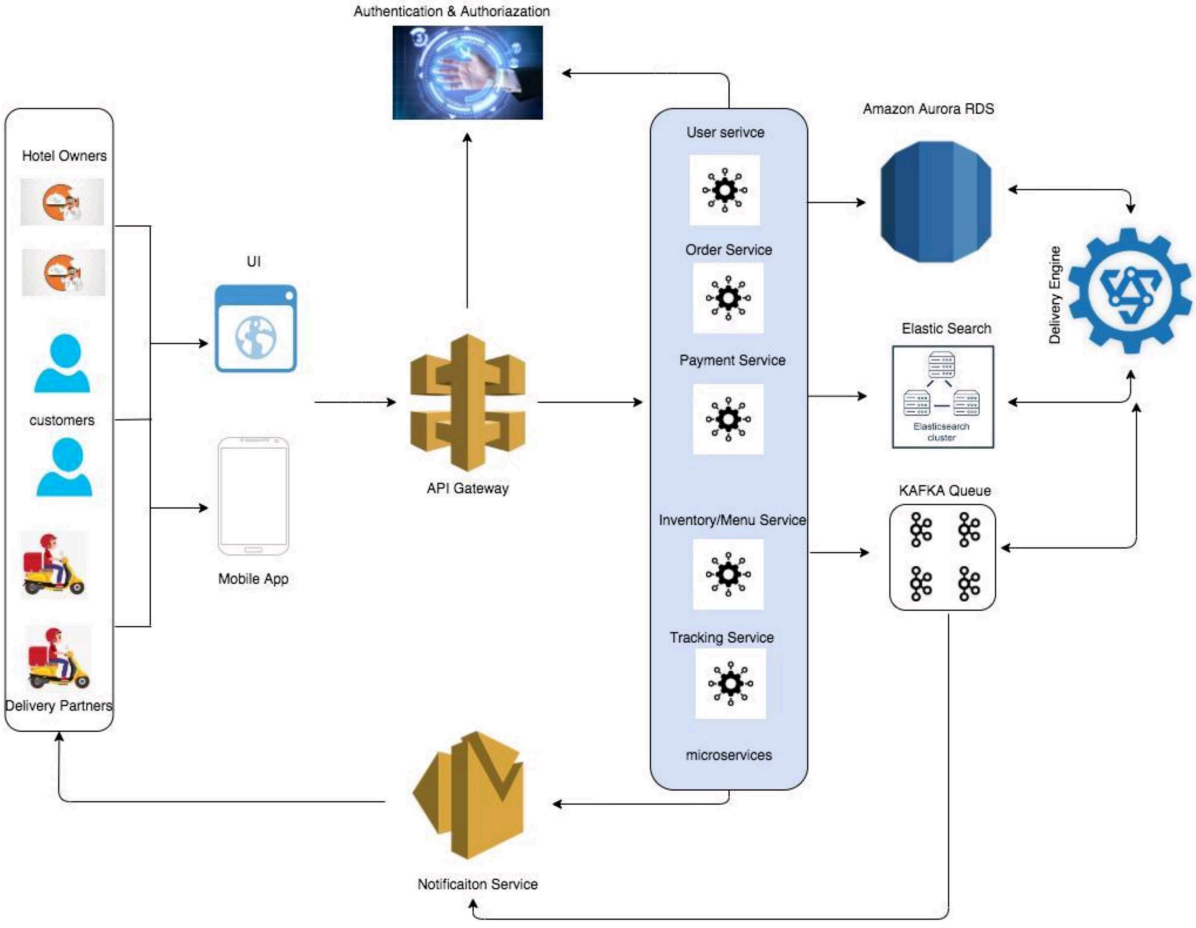
Solution architecture is a complex process – with many subprocesses – that bridges the gap between business problems and technology solutions. Its goals are to:

- Find the best tech solution to solve existing business problems.
- Describe the structure, characteristics, behavior, and other aspects of the

Software to project stakeholders.

- Define features, development phases, and solution requirements.
- Provide specifications according to which the solution is defined, managed,

And delivered.


Authentication & Authoriazation
Hotel Owners
customers
Delivery Partners
UI
Mobile App
API Gateway
User serivce
Order Service
Payment Service
Inventory/Menu Service
Tracking Service
microservices
Amazon Aurora RDS
Elastic Search
Elasticsearch cluster
KAFKA Queue
Delivery Engine
Notificaiton Service

Project Planning Template

Product Backlog, Sprint Schedule, and Estimation (4 Marks)

Use the below template to create product backlog and sprint schedule

Sprint	Functional Requirement (Epic)	User Story Number	User Story / Task	Story Points	Priority	Team Members
Sprint-1	Registration	USN-1	As a user, I can register for the application by entering my email, password, and confirming my password.	2	High	
Sprint-1		USN-2	As a user, I will receive confirmation email once I have registered for the application	1	High	
Sprint-2		USN-3	As a user, I can register for the application through Facebook	2	Low	
Sprint-1		USN-4	As a user, I can register for the application through Gmail	2	Medium	
Sprint-1	Login	USN-5	As a user, I can log into the application by entering email & password	1	High	
	Dashboard					

Project Tracker, Velocity & Burndown Chart: (4 Marks)

Sprint	Total Story Points	Duration	Sprint Start Date	Sprint End Date (Planned)	Story Points Completed (as on Planned End Date)	Sprint Release Date (Actual)
Sprint-1	20	6 Days	24 Oct 2022	29 Oct 2022	20	29 Oct 2022
Sprint-2	20	6 Days	31 Oct 2022	05 Nov 2022		
Sprint-3	20	6 Days	07 Nov 2022	12 Nov 2022		
Sprint-4	20	6 Days	14 Nov 2022	19 Nov 2022		

Velocity:
Imagine we have a 10-day sprint duration, and the velocity of the team is 20 (points per sprint). Let's calculate the team's average velocity (AV) per iteration unit (story points per day)

$$AV = \frac{sprint\ duration}{velocity} = \frac{20}{10} = 2$$

Project Design Phase-II
Technology Stack (Architecture & Stack)

Technical Architecture:

The Deliverable shall include the architectural diagram as below and the information as per the table1 & table 2

Example: Order processing during pandemics for offline mode

Guidelines:

1. Include all the processes (As an application logic / Technology Block)
2. Provide infrastructural demarcation (Local / Cloud)
3. Indicate external interfaces (third party API's etc.)
4. Indicate Data Storage components / services
5. Indicate interface to machine learning models (if applicable)

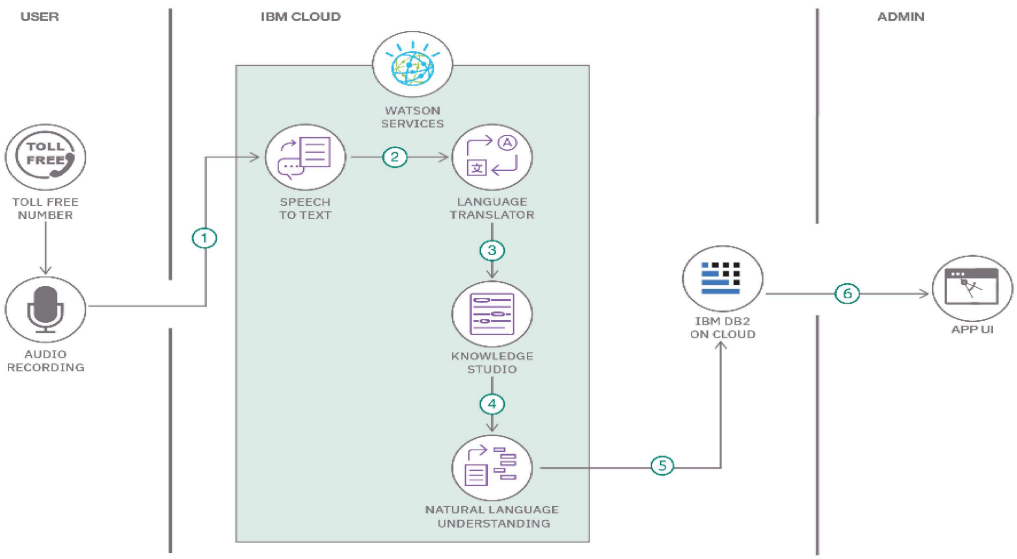

USER
IBM CLOUD
ADMIN
WATSON SERVICES
TOLL FREE NUMBER
SPEECH TO TEXT
LANGUAGE TRANSLATOR
AUDIO RECORDING
KNOWLEDGE STUDIO
NATURAL LANGUAGE UNDERSTANDING
IBM DB2 ON CLOUD
APP UI
1
2
3
4
5
6

Table-1 : Components & Technologies:

S.No	Component	Description	Technology
1.	User Interface	How user interacts with application e.g. Web UI, Mobile App, Chatbot etc.	HTML, CSS, JavaScript / Angular Js / React Js etc.
2.	Application Logic-1	Logic for a process in the application	Java / Python
3.	Application Logic-2	Logic for a process in the application	IBM Watson STT service
4.	Application Logic-3	Logic for a process in the application	IBM Watson Assistant
5.	Database	Data Type, Configurations etc.	MySQL, NoSQL, etc.
6.	Cloud Database	Database Service on Cloud	IBM DB2, IBM Cloudant etc.
7.	File Storage	File storage requirements	IBM Block Storage or Other Storage Service or Local Filesystem
8.	External API-1	Purpose of External API used in the application	IBM Weather API, etc.
9.	External API-2	Purpose of External API used in the application	Aadhar API, etc.
10.	Machine Learning Model	Purpose of Machine Learning Model	Object Recognition Model, etc.
11.	Infrastructure (Server / Cloud)	Application Deployment on Local System / Cloud Local Server Configuration: Cloud Server Configuration :	Local, Cloud Foundry, Kubernetes, etc.

Table-2: Application Characteristics:

S.No	Characteristics	Description	Technology
1.	Open-Source Frameworks	List the open-source frameworks used	Technology of Opensource framework
2.	Security Implementations	List all the security / access controls implemented, use of firewalls etc.	e.g. SHA-256, Encryptions, IAM Controls, OWASP etc.
3.	Scalable Architecture	Justify the scalability of architecture (3 – tier, Micro-services)	Technology used

S.No	Characteristics	Description	Technology
4.	Availability	Justify the availability of application (e.g. use of load balancers, distributed servers etc.)	Technology used
5.	Performance	Design consideration for the performance of the application (number of requests per sec, use of Cache, use of CDN's) etc.	Technology used

project development phase

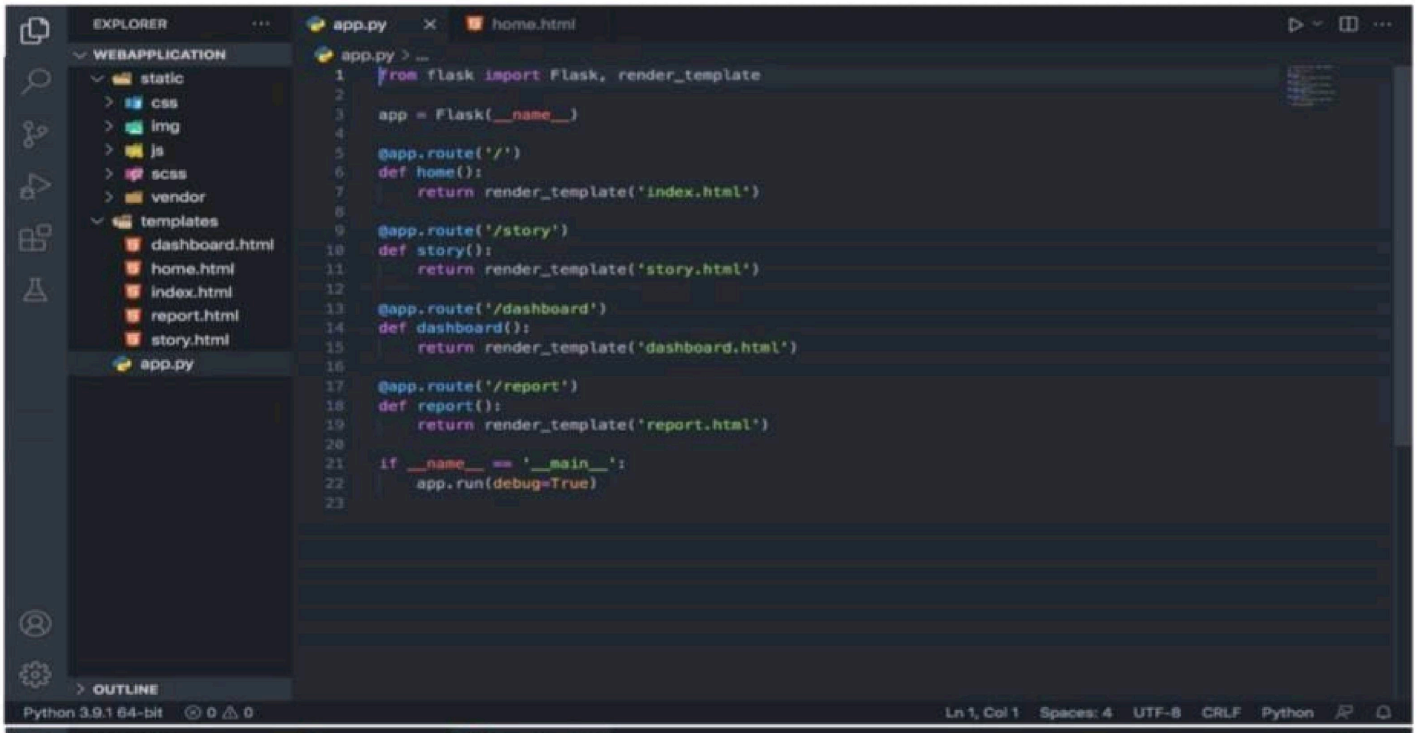
```
import numpy as np
import matplotlib.pyplot as plt
import seaborn as sns
from wordcloud import WordCloud, STOPWORDS

sns.set(rc= {'figure.figsize': (12,8)})
plt.style.use('ggplot')

from plotly.offline import init_notebook_mode, i
plot
import cufflinks
cufflinks.go_offline()
cufflinks.set_config_file(world_readable=True, t
heme='pearl')
import plotly.graph_objs as go
import plotly
import plotly.express as px
import plotly.figure_factory as ff
```

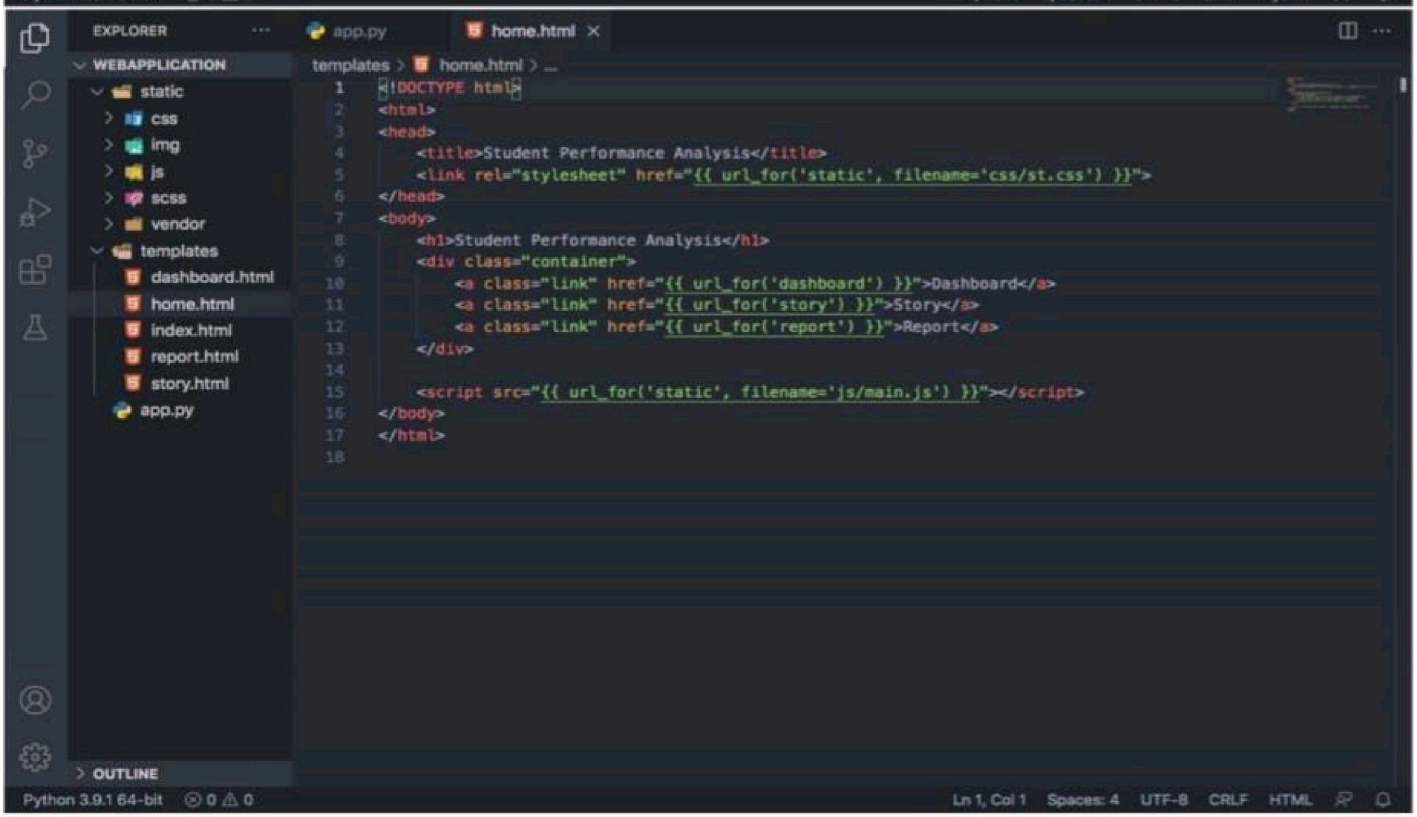
Wordcloud of Ingredients for Non Vegetarian Meals

In [35]:

```
nonveg = df.loc[df.diet == 'non vegetarian', 'in
gredients']

text2 = get_text(nonveg)

stopwords = set(STOPWORDS)
wc = WordCloud(background_color= 'black', stopwo
rds= stopwords,
               width=1600, height=800)

wc.generate(text2)
plt.figure(figsize=(20,10), facecolor='k')
plt.axis('off')
plt.tight_layout(pad=0)
plt.imshow(wc)
plt.show()
```

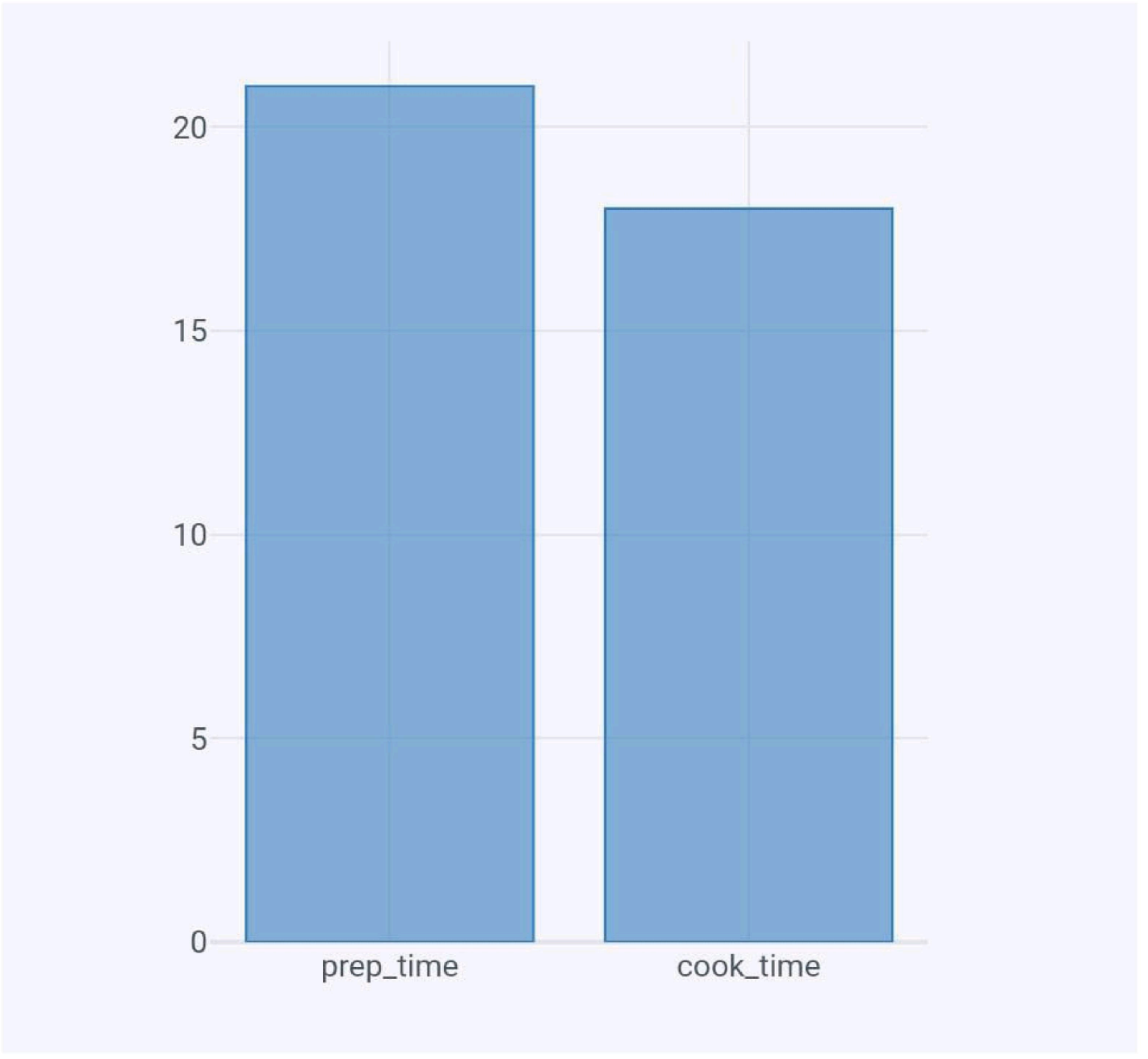
```
pie = df.region.value_counts()
pie_df = pd.DataFrame({'index':pie.index, 'value
s': pie.values})
pie_df.iplot(kind='pie', labels= 'index', values
= 'values', hole= .5, title="Value counts: regio
n")

pie2 = df.course.value_counts()
pie_df2 = pd.DataFrame({'index':pie2.index, 'val
ues': pie2.values})
pie_df2.iplot(kind='pie', labels= 'index', value
s= 'values', hole= .5, title="Value counts: cour
se")

pie3 = df.flavor_profile.value_counts()
pie_df3 = pd.DataFrame({'index':pie3.index, 'val
ues': pie3.values})
pie_df3.iplot(kind='pie', labels= 'index', value
s= 'values', hole= .5, title="Value counts: flav
or profile")

pie4 = df.diet.value_counts()
pie_df4 = pd.DataFrame({'index':pie4.index, 'val
ues': pie4.values})
pie_df4.iplot(kind='pie', labels= 'index', value
s= 'values', hole= .5, title="Value counts: die
t")
```


EXPLORER
app.py
home.html
WEBAPPLICATION
static
css
img
js
scss
vendor
templates
dashboard.html
home.html
index.html
report.html
story.html
app.py
OUTLINE
Python 3.9.1 64-bit
Ln 1, Col 1
Spaces: 4
UTF-8
CRLF
Python
from flask import Flask, render_template
app = Flask(__name__)
@app.route('/')
def home():
return render_template('index.html')
@app.route('/story')
def story():
return render_template('story.html')
@app.route('/dashboard')
def dashboard():
return render_template('dashboard.html')
@app.route('/report')
def report():
return render_template('report.html')
if __name__ == '__main__':
app.run(debug=True)

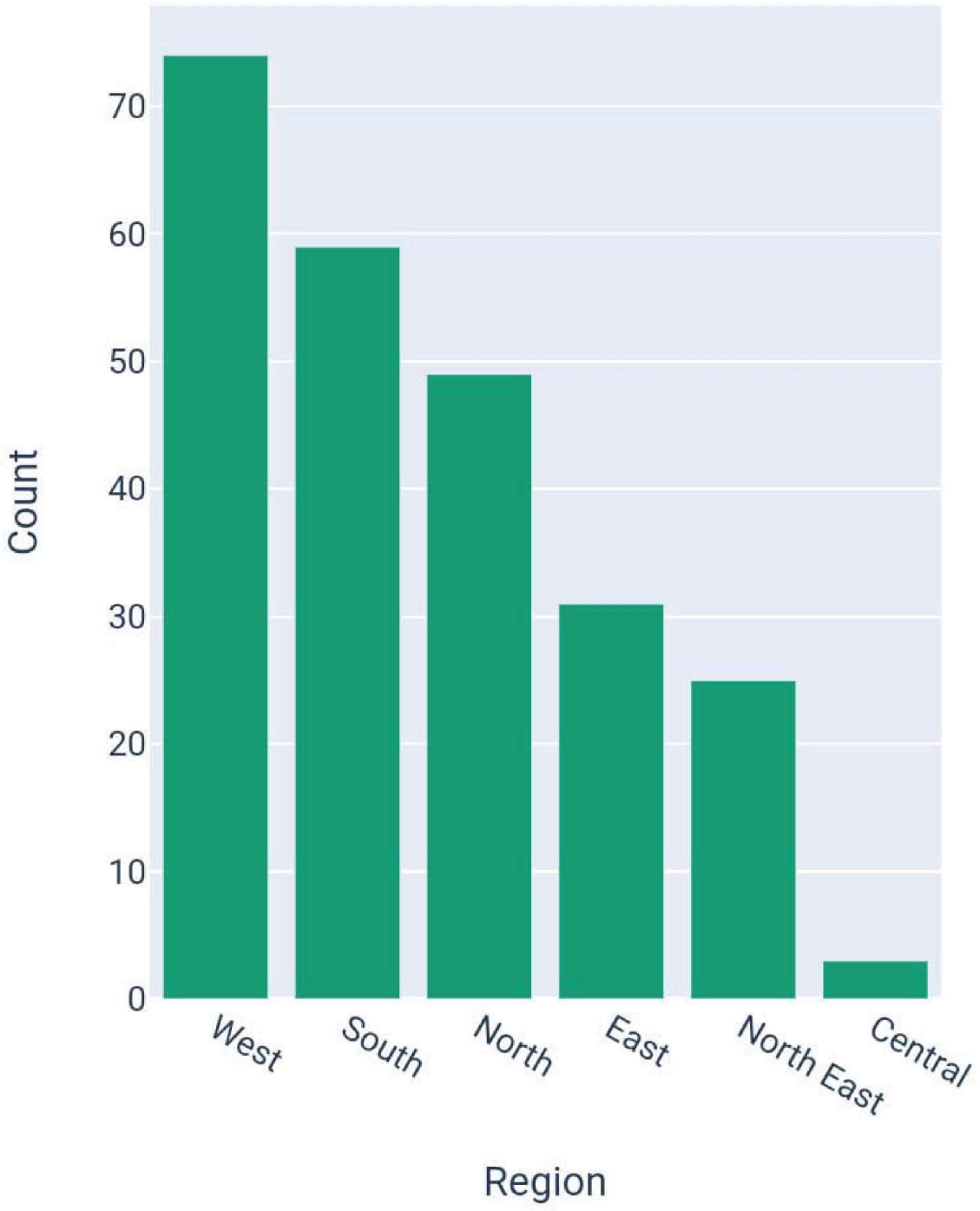

EXPLORER
app.py
home.html
WEBAPPLICATION
static
css
img
js
scss
vendor
templates
dashboard.html
home.html
index.html
report.html
story.html
app.py
OUTLINE
Python 3.9.1 64-bit
Ln 1, Col 1
Spaces: 4
UTF-8
CRLF
HTML
templates > home.html > ...
<!DOCTYPE html>
<html>
<head>
<title>Student Performance Analysis</title>
<link rel="stylesheet" href="{{ url_for('static', filename='css/st.css') }}">
</head>
<body>
<h1>Student Performance Analysis</h1>

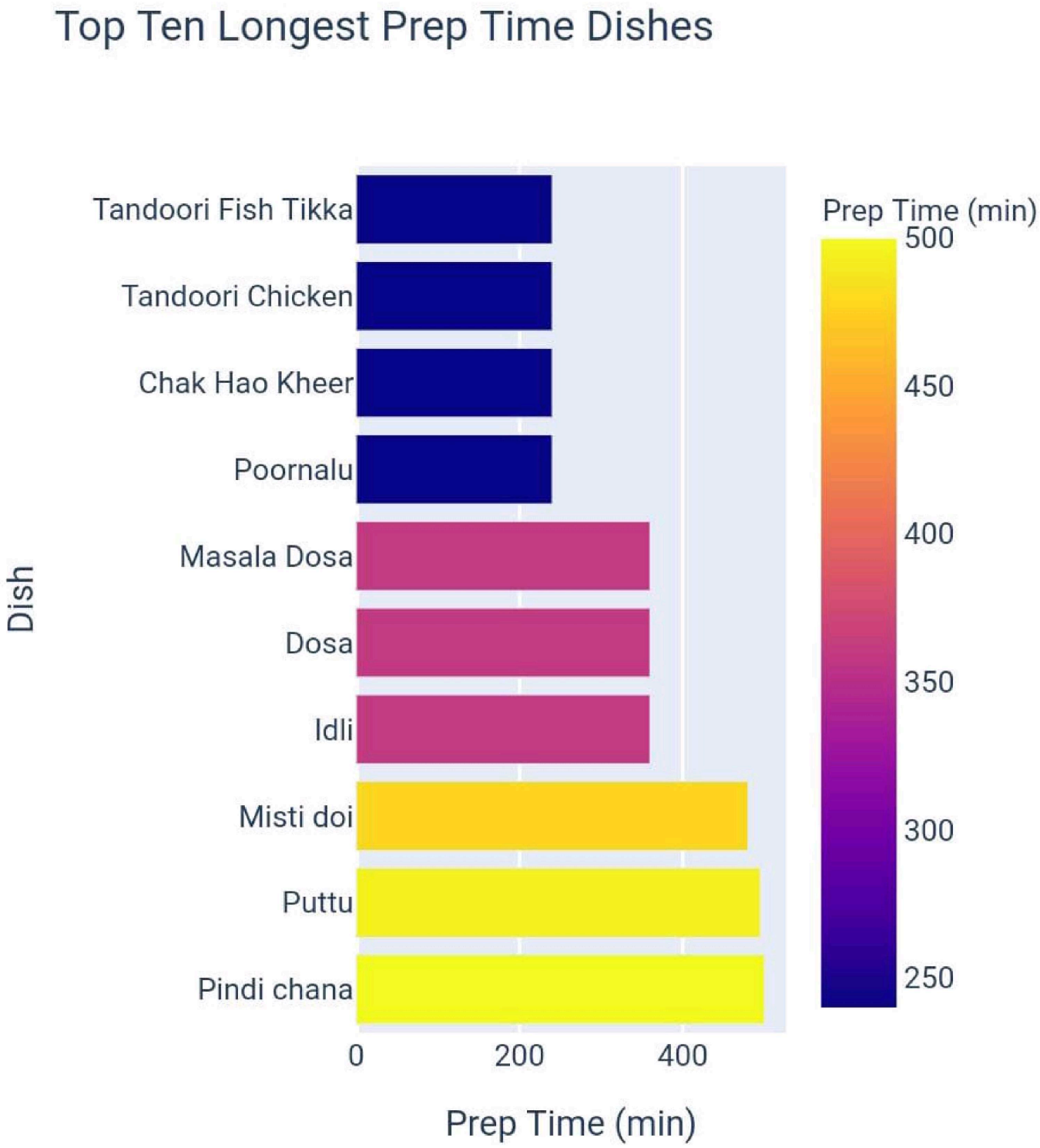
Dashboard
Story
Report

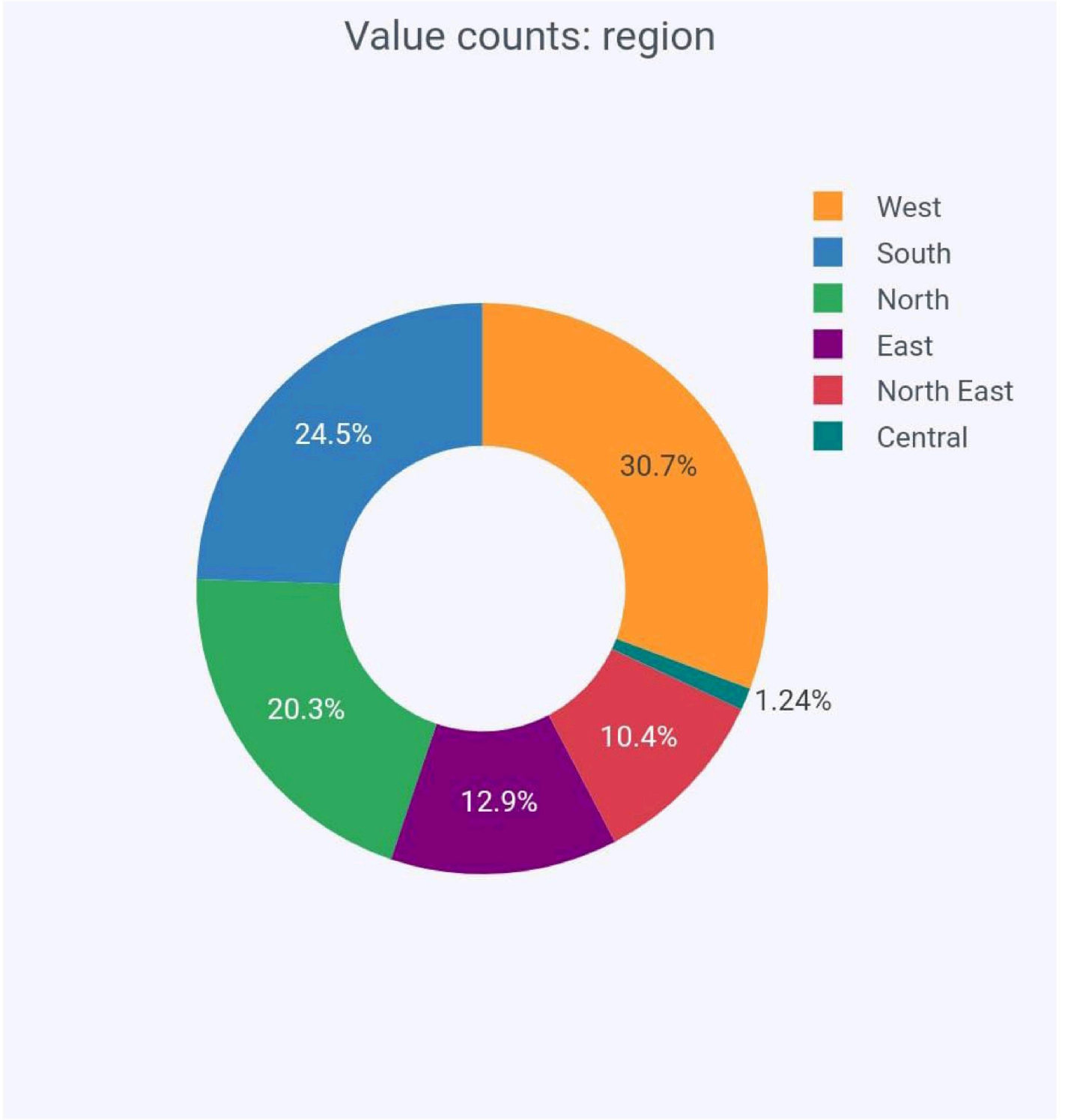
<script src="{{ url_for('static', filename='js/main.js') }}"></script>
</body>
</html>

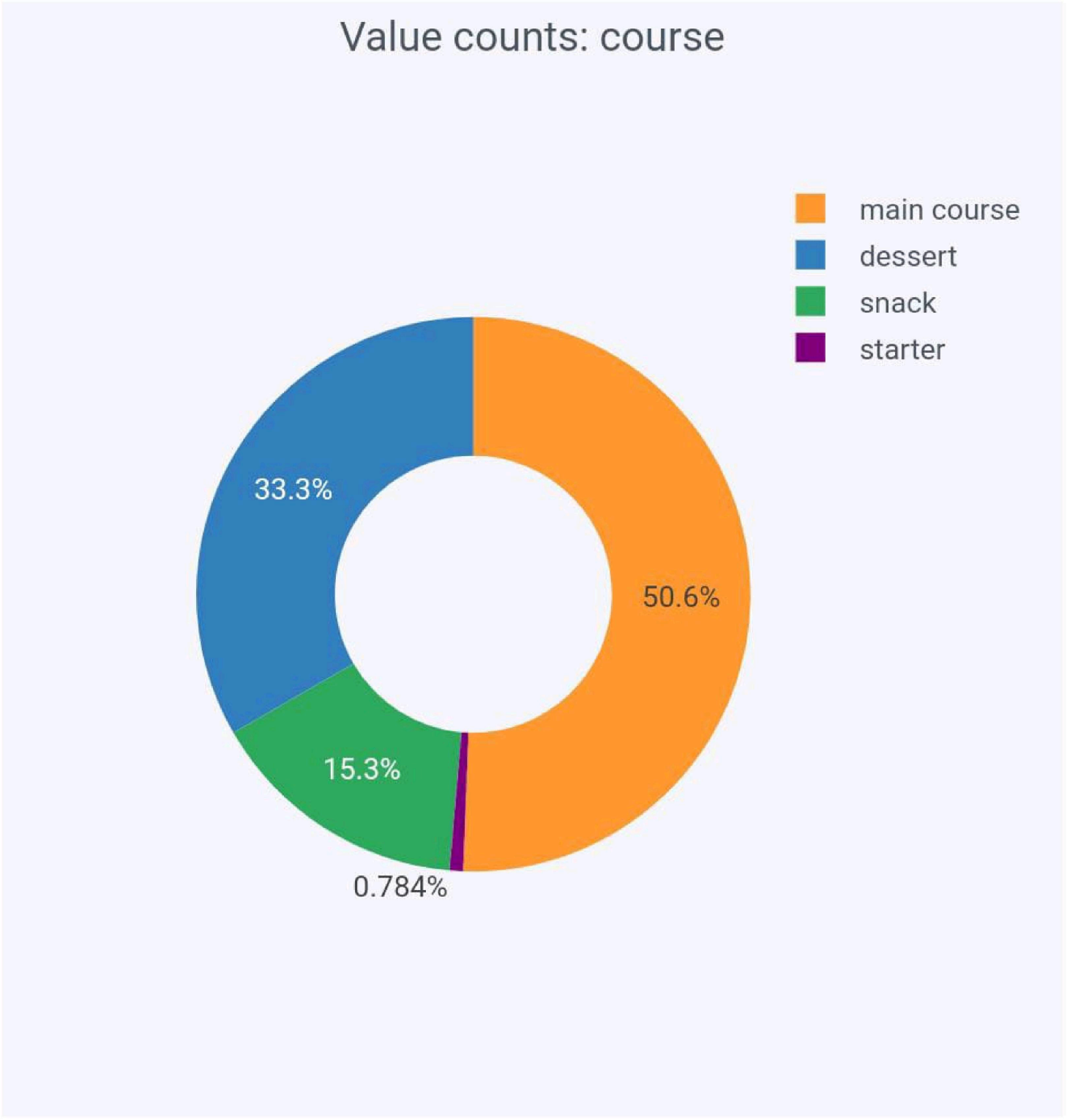

20
15
10
5
0
prep_time
cook_time

Region Count

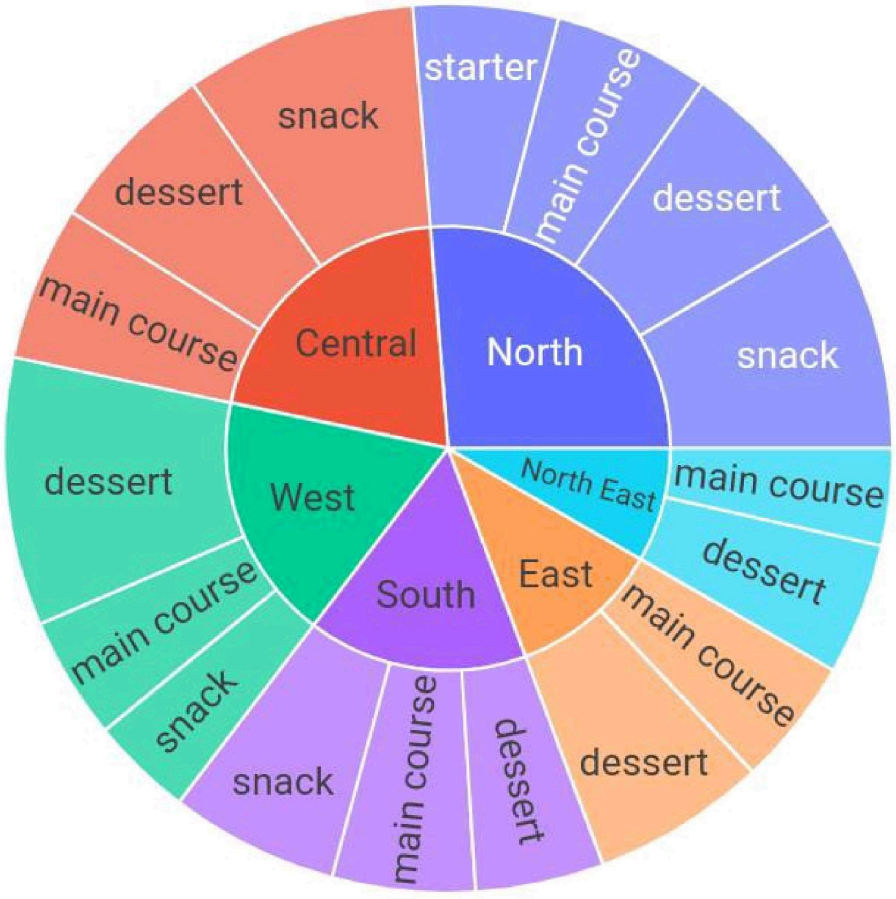



Top Ten Longest Prep Time Dishes
Tandoori Fish Tikka
Tandoori Chicken
Chak Hao Kheer
Poornalu
Masala Dosa
Dosa
Idli
Misti doi
Puttu
Pindi chana
Dish
0
200
400
Prep Time (min)
Prep Time (min)
500
450
400
350
300
250

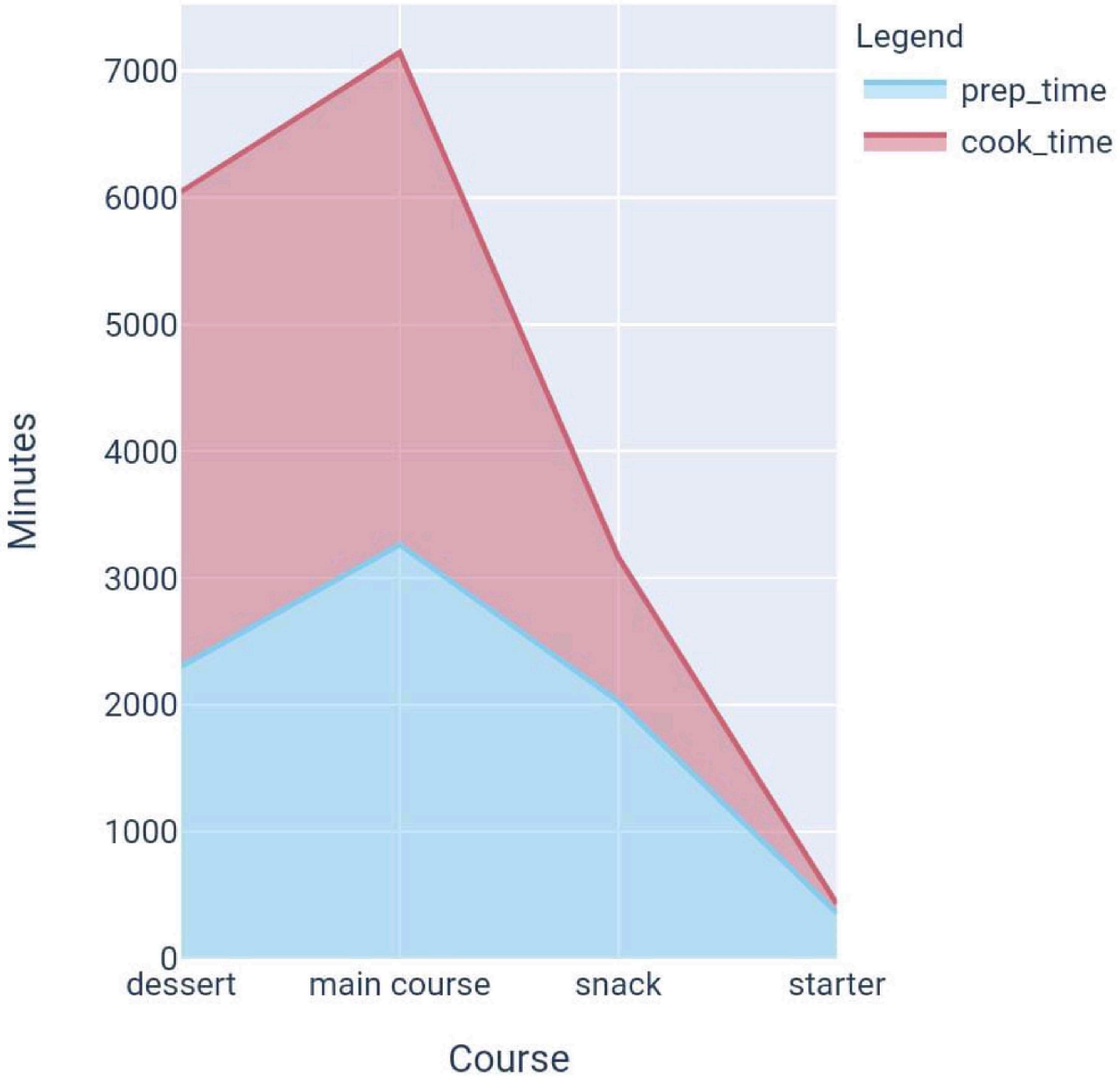

Value counts: region
West
South
North
East
North East
Central
30.7%
24.5%
20.3%
12.9%
10.4%
1.24%

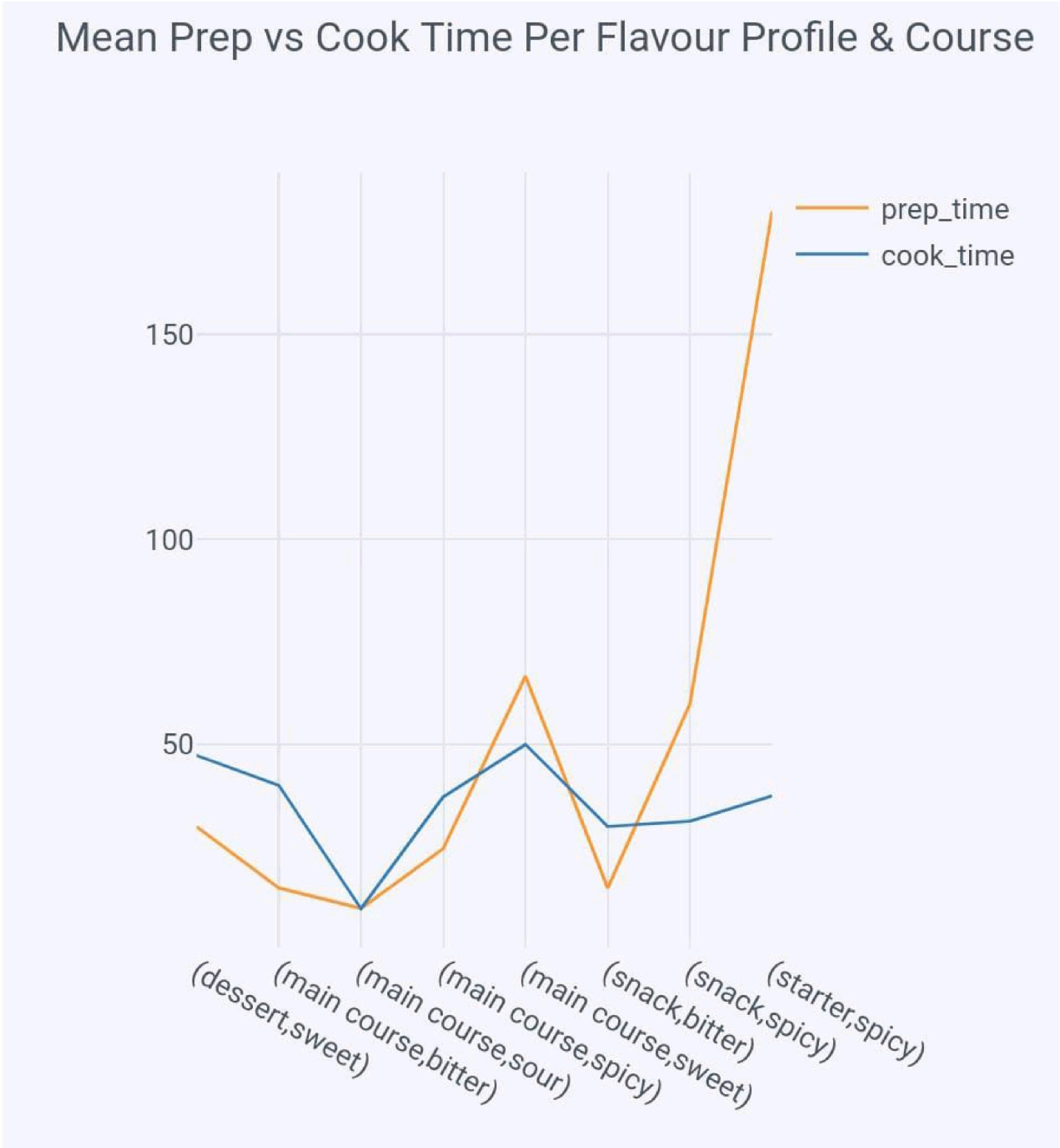

Value counts: course
main course
dessert
snack
starter
33.3%
50.6%
15.3%
0.784%

Cook Time


starter
main course
dessert
snack
snack
dessert
main course
Central
North
North East
main course
dessert
East
main course
dessert
South
dessert
main course
snack
snack
main course
dessert
West

Sum of Cook vs Prep Time per Course




Mean Prep vs Cook Time Per Flavour Profile & Course
prep_time
cook_time
150
100
50
(dessert,sweet)
(main course,bitter)
(main course,sour)
(main course,spicy)
(main course,sweet)
(snack,bitter)
(snack,spicy)
(starter,spicy)

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.

SOURCE CODE:

Github link: MuthuKumarkumar / indian-food-eda-