A

PROJECT REPORT

On

Zomato (Bangalore Restaurant Explorer)

Submitted in partial fulfillment of completion of the course

ADVANCE DIPLOMA IN IT NETWORKING & CLOUD COMPUTING

NSTI (W) NOIDA

2023-2024

Submitted By:

Krati Shukla

Kamal

ACKNOWLEDGEMENT

In humble gratitude, we acknowledge the divine blessings that have illuminated our path and guided us towards the successful culmination of our endeavours. We extend our heartfelt appreciation to the individuals who have been instrumental in our journey. First and foremost, we express our deepest gratitude to our esteemed guide and mentor, Mrs. Deepika Singh, Mrs. Mala Mishra, Miss Ankita Shukla, and Mr. D.A. Guruvulu. Their unwavering support and invaluable guidance have served as our compass during critical moments. We also extend our sincere thanks to our dedicated Edunet Mentor, whose cordial support and insightful suggestions have enriched our learning experience.

We acknowledge the generosity of our respected Head of the division, Mrs. Shashi Mathur (GD NSTI(W) NOIDA), for granting us access to the facilities that facilitated our work. We also extend our appreciation to the other esteemed faculty members who have contributed to our knowledge and growth. We express our gratitude to all who have played a pivotal role in our journey, and we are indebted for the opportunities and support we have received.

ADIT TRADE IN NSTI (W) NOIDA

The Advanced Diploma in IT and Cloud Computing program, a collaborative effort between NSTI (W) Noida and Edunet Foundation, is a comprehensive course designed to provide students with advanced skills in information technology and cloud computing. It covers diverse topics, including computer networking, database management, virtualization, cloud technologies, and cybersecurity. Practical labs, workshops, and real-world projects offer hands-on experience, preparing students for success in the dynamic IT industry. Graduates will possess a solid foundation in IT fundamentals and cloud computing, positioning them as highly sought-after professionals in the field.

Project Requirements

Project Name	Zomato (Bangalore Restaurant Explorer)
Language Used	Python
Editor	Jupyter Notebook, Google Collab
Web Browser	Google Chrome, Microsoft Edge

Team Composition and Workload Division

Krati Shukla	Data Analysis, Synopsis
Kamal	Data Analysis, Synopsis

Tables of Content

INTRODUCTION

Zomato (Bangalore Restaurant Explorer) The project aims to analyse and visualize restaurant data from Zomato in Bangalore, providing valuable insights into the city's culinary landscape. By leveraging the Zomato API or dataset, we will explore various facets of the restaurant scene, including cuisine types, average ratings, pricing, and popular locations.

This project is all about checking out and understanding the restaurants in Bangalore using information from Zomato. We want to find out cool things about the food scene in the city. We'll use Zomato's data to look at different aspects like what kinds of food are popular, how well restaurants are rated, how much they cost, and where the hotspots are.

PROBLEM STATEMENT

- **Data Retrieval :**
- **Problem:** Obtaining accurate and up-to-date restaurant details from Zomato's API or dataset may pose challenges, including data completeness and consistency.

- **Data Exploration :**
- **Problem:** Understanding the structure and content of the restaurant data could be challenging. Identifying key features and patterns might require effective exploratory data analysis techniques.

- **Rating and Price Analysis:**
- **Problem:** Determining whether there is a correlation between restaurant ratings and prices, and understanding the factors influencing customer ratings.

- **Cuisine Trends:**
- **Problem:** Analysing and interpreting the popularity of different cuisines in Bangalore, uncovering trends, and understanding customer preferences.

- **Location-based Insights:**
- **Problem:** Identifying areas with a high concentration of restaurants and understanding the factors contributing to certain areas being restaurant hotspots.

- **Interactive Visualization:**
- **Problem:** Creating visually appealing and informative charts and graphs to effectively communicate the findings from the data analysis

- **Predictive Modelling :**
- **Problem:** If opting for predictive modelling, addressing challenges related to selecting relevant features, model accuracy, and ensuring the model's practical utility.

- **Information Dissemination:**
- **Problem:** Summarizing and communicating the key findings in an engaging and informative way for a diverse audience, including those interested in exploring Bangalore's food scene.

GOALS

- **Data Retrieval:**
- **Goal:** Retrieve comprehensive and up-to-date restaurant details from Zomato's API or dataset, ensuring data accuracy, completeness, and consistency.

- **Data Exploration:**
- **Goal:** Conduct exploratory data analysis to understand the structure and content of the restaurant data. Identify key features, patterns, and potential insights that will guide further analysis.

- **Rating and Price Analysis:**
- **Goal:** Determine the correlation between restaurant ratings and prices. Identify and analyse factors influencing customer ratings, helping to understand the relationship between cost and perceived quality.

- **Cuisine Trends:**
- **Goal:** Analyse and interpret the popularity of different cuisines in Bangalore. Uncover trends in cuisine preferences and customer choices, providing insights into the diverse culinary landscape.

- **Location-based Insights:**
- **Goal:** Identify areas in Bangalore with a high concentration of restaurants. Explore factors contributing to specific locations being restaurant hotspots, aiding in understanding the city's dining dynamics.

- **Interactive Visualization:**
- **Goal:** Create visually engaging and informative charts and graphs to effectively communicate the project's findings. Develop interactive visualizations that enhance user understanding of the restaurant data and insights.

- **Predictive Modelling :**
- **Goal:** If opting for predictive modelling, develop a model that accurately predicts restaurant-related outcomes. Select relevant features and assess the model's practical utility in providing valuable insights into restaurant performance.

- **Information Dissemination:**
- **Goal:** Summarize and communicate key findings in an engaging and informative manner suitable for a diverse audience. Create a compelling narrative that showcases interesting aspects of Bangalore's food scene, potentially recommending noteworthy restaurants.

SOFTWARE SPECIFICATION

- **Data Retrieval:**
- **Software Specification:**
- Zomato API or Zomato dataset for obtaining restaurant details.
- Python programming language for data retrieval and manipulation.
- Requests library for making API calls if using the Zomato API.

- **Data Exploration:**
- **Software Specification:**
- Jupyter Notebooks for interactive and exploratory data analysis.
- Python libraries such as Pandas, NumPy, and Matplotlib for data manipulation and visualization.

- **Rating and Price Analysis:**
- **Software Specification:**
- Python for data analysis using Pandas and NumPy.
- Matplotlib and Seaborn for visualizing rating and price distributions.
- Statistical analysis tools if needed .

- **Cuisine Trends:**
- **Software Specification:**
- Python for data analysis using Pandas.
- Matplotlib, Seaborn, or Plotly for visualizing cuisine trends.

- **Location-based Insights:**
- Software Specification:
- Geospatial libraries like Folium for mapping restaurant locations.
- Matplotlib or Plotly for visualizing location-based insights.

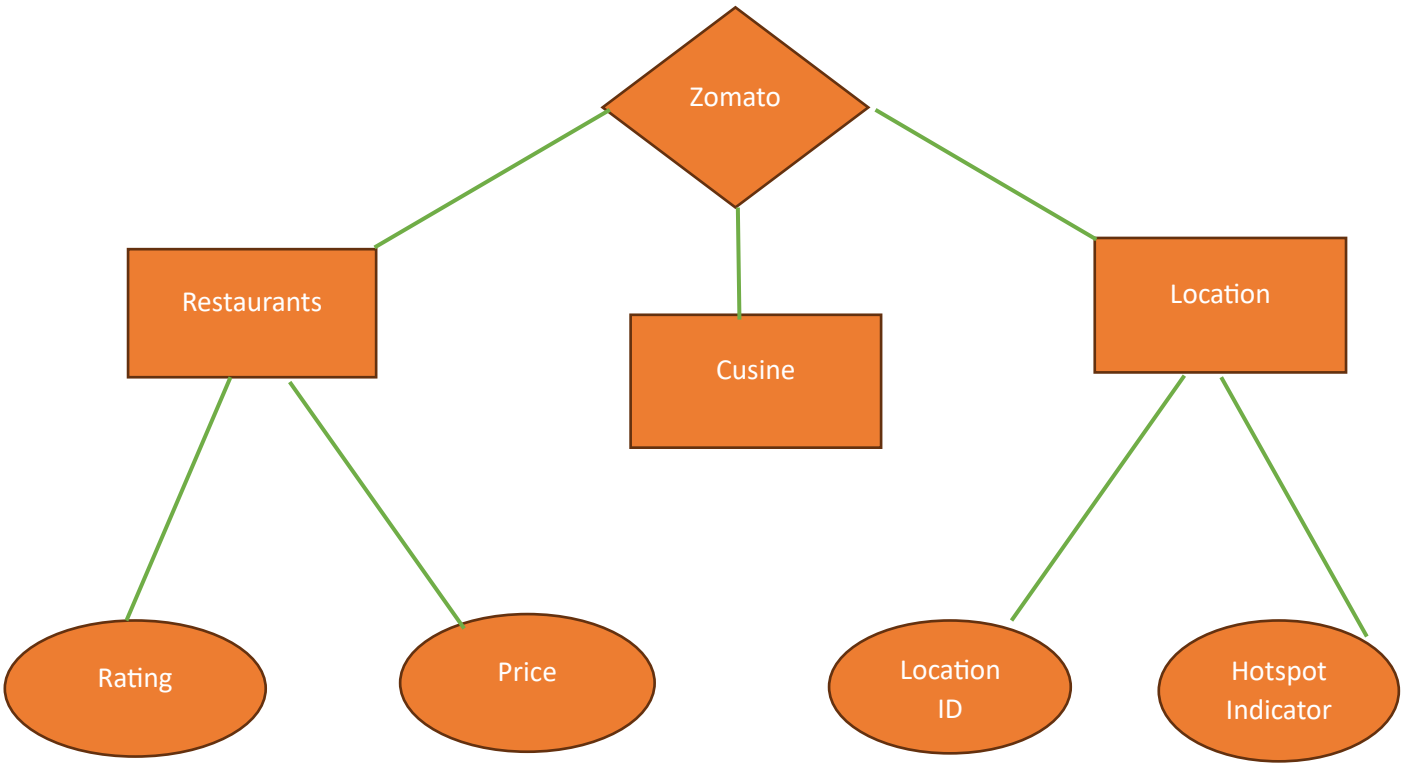
- **Interactive Visualization:**
- **Software Specification:**

- Plotly or Bokeh for creating interactive visualizations.
- Jupyter Dash or Streamlit for building interactive dashboards.

- **Predictive Modeling :**
- **Software Specification:**
- Scikit-learn or TensorFlow for developing predictive models.
- Jupyter Notebooks for model development and evaluation.

- **Information Dissemination:**
- **Software Specification:**
- Jupyter Notebooks for creating a cohesive narrative.
- Markdown for creating documentation.

E-R DIAGRAM



.

OVERVIEW

The **" Zomato Bangalore Restaurant Explorer"** project is designed to uncover the intricacies of Bangalore's culinary landscape by delving into Zomato's extensive restaurant data. The primary objective is to extract comprehensive and up-to-date details from Zomato's API or dataset, ensuring accuracy and completeness. Through a meticulous data exploration phase using Jupyter Notebooks and Python libraries such as Pandas, NumPy, and Matplotlib, the project aims to unravel key features, patterns, and potential insights embedded within the restaurant data.

Subsequent analyses will focus on determining correlations between restaurant ratings and prices, identifying influential factors on customer ratings, and discerning the popularity of different cuisines in Bangalore.

Location-based insights will uncover areas with high concentrations of restaurants, shedding light on the dynamics of Bangalore's dining scene. Interactive visualizations, created with tools like Plotly, Bokeh, and Jupyter Dash or Streamlit, will be employed to make the project's findings visually appealing and accessible. Optionally, predictive modeling using Scikit-learn or TensorFlow may be implemented to forecast restaurant-related outcomes.

The project will culminate in a cohesive narrative presented in Jupyter Notebooks, employing Markdown for documentation and possibly utilizing Voila or NBConvert for enhanced interactivity, summarizing the key findings and recommendations for an engaging and diverse audience.

SOURCE CODE

Importing libraries

In [1]:
```
# import required libraries
import pandas as pd
import numpy as np
import matplotlib.pyplot as plt
import seaborn as sns
plt.style.use('dark_background')
```

In [2]:
```
df=pd.read_csv('zomato.csv') #reading csv file , loading a dataset.
```

In [34]:
```
# print Dataset
df
```

Out[34]:

	url	address	name	online_order	book_table	rate	votes	phone	location	rest_type
0	https://www.zomato.com/bangalore/jalsa-banasha...	942, 21st Main Road, 2nd Stage, Banashankari, ...	Jalsa	Yes	Yes	4.1/5	775	080 42297555\r\n+91 9743772233	Banashankari	Casual Dining
1	https://www.zomato.com/bangalore/spice-elephan...	2nd Floor, 80 Feet Road, Near Big Bazaar, 6th ...	Spice Elephant	Yes	No	4.1/5	787	080 41714161	Banashankari	Casual Dining

In [10]:
```
df.describe()
```

Out[10]:

	votes
count	51717.000000
mean	283.697527
std	803.838853
min	0.000000
25%	7.000000
50%	41.000000
75%	198.000000
max	16832.000000

In [11]:
```
df.info()
```

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 51717 entries, 0 to 51716
Data columns (total 17 columns):
 #   Column                       Non-Null Count  Dtype
---  ------                       --------------  -----
 0   url                          51717 non-null  object
 1   address                      51717 non-null  object
 2   name                         51717 non-null  object
 3   online_order                 51717 non-null  object
 4   book_table                   51717 non-null  object
 5   rate                         43942 non-null  object
 6   votes                        51717 non-null  int64
```

Listed in (city) and location, both are there , lets keep only one.

In [42]:
```
df=df.drop(['listed_in(city)'], axis =1)
df.head()
```

Out[42]:

	name	online_order	book_table	rate	votes	location	rest_type	cuisines	Cost2plates	Type
0	Jalsa	Yes	Yes	4.1	775	Banashankari	Casual Dining	North Indian, Mughlai, Chinese	800	Buffet
1	Spice Elephant	Yes	No	4.1	787	Banashankari	Casual Dining	Chinese, North Indian, Thai	800	Buffet
2	San Churro Cafe	Yes	No	3.8	918	Banashankari	Cafe, Casual Dining	Cafe, Mexican, Italian	800	Buffet
3	Addhuri Udupi Bhojana	No	No	3.7	88	Banashankari	Quick Bites	South Indian, North Indian	300	Buffet
4	Grand Village	No	No	3.8	166	Basavanagudi	Casual Dining	North Indian, Rajasthani	600	Buffet

In [43]:
```
df['Cost2plates'].unique()
```

Out[43]:
```
array(['800', '300', '600', '700', '550', '500', '450', '650', '400',
       '900', '200', '750', '150', '850', '100', '1,200', '350', '250',
       '950', '1,000', '1,500', '1,300', '199', '80', '1,100', '160',
       '1,600', '230', '130', '50', '190', '1,700', '1,400', '180',
       '1,350', '2,200', '2,000', '1,800', '1,900', '330', '2,500',
       '2,100', '3,000', '2,800', '3,400', '40', '1,250', '3,500',
       '4,000', '2,400', '2,600', '120', '1,450', '469', '70', '3,200',
       '60', '560', '240', '360', '6,000', '1,050', '2,300', '4,100',
```

Cleaning Cuisines Column

In [5]:
```
cuisines = df['cuisines'].value_counts(ascending = False)
cuisines_lessthan100 = cuisines[cuisines<100]

def handle_cuisines(value):
    if(value in cuisines_lessthan100):
        return 'others'
    else:
        return value

df['cuisines'] = df['cuisines'].apply(handle_cuisines)
df['cuisines'].value_counts()
```

Out[5]:
```
others                                  26460
North Indian                             2913
North Indian, Chinese                    2385
South Indian                             1828
Biryani                                   918
                                        ...
South Indian, Chinese, North Indian       105
Italian, Pizza                            105
North Indian, Mughlai, Chinese            104
South Indian, Fast Food                   104
North Indian, Chinese, Seafood            102
Name: cuisines, Length: 70, dtype: int64
```

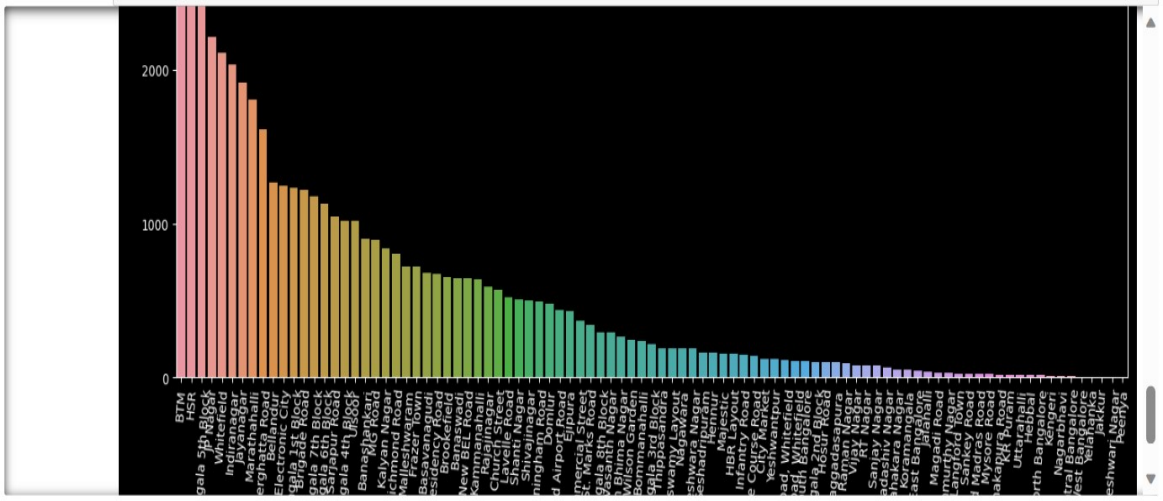
count plot of various location

In [23]:
```
plt.figure(figsize=(16, 10))
order = df['location'].value_counts().index  # Specify the order based on value counts
x = sns.countplot(data=df, x='location', order=order)
plt.xticks(rotation=90)
```

Out[23]:
```
(array([ 0,  1,  2,  3,  4,  5,  6,  7,  8,  9, 10, 11, 12, 13, 14, 15, 16,
        17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 31, 32, 33,
        34, 35, 36, 37, 38, 39, 40, 41, 42, 43, 44, 45, 46, 47, 48, 49, 50,
        51, 52, 53, 54, 55, 56, 57, 58, 59, 60, 61, 62, 63, 64, 65, 66, 67,
        68, 69, 70, 71, 72, 73, 74, 75, 76, 77, 78, 79, 80, 81, 82, 83, 84,
        85, 86, 87, 88, 89, 90, 91, 92]),
 [Text(0, 0, 'BTM'),
  Text(1, 0, 'HSR'),
  Text(2, 0, 'Koramangala 5th Block'),
  Text(3, 0, 'JP Nagar'),
  Text(4, 0, 'Whitefield'),
  Text(5, 0, 'Indiranagar'),
  Text(6, 0, 'Jayanagar'),
  Text(7, 0, 'Marathahalli'),
  Text(8, 0, 'Bannerghatta Road'),
  Text(9, 0, 'Bellandur'),
  Text(10, 0, 'Electronic City'),
  Text(11, 0, 'Koramangala 1st Block'),
  Text(12, 0, 'Brigade Road'),
```

SAMPLE SCREENSHOTS

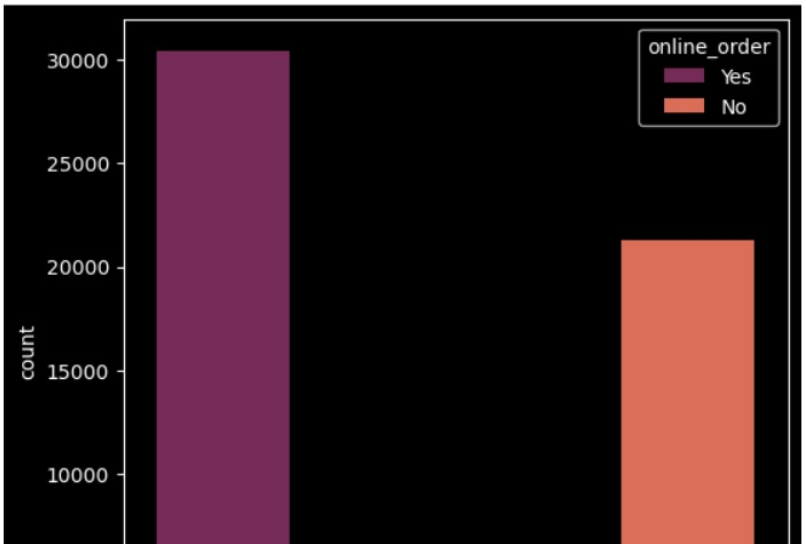
In [48]:
```
plt.figure(figsize=(16, 10))
order = df['location'].value_counts().index  # Specify the order based on value counts
x = sns.countplot(data=df, x='location', order=order)
plt.xticks(rotation=85)
```



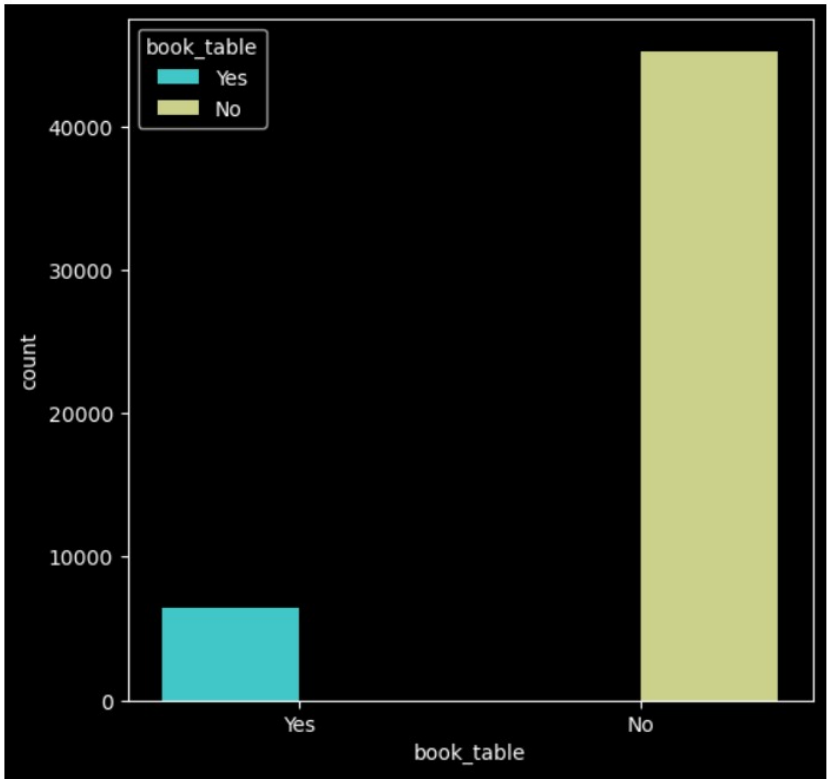
Visualizing Online Order

In [27]:
```
# Visualizing Online Order
plt.figure(figsize=(6, 6))
sns.countplot(data=df, x='online_order', palette='rocket', hue='online_order')
```

Out[27]: <Axes: xlabel='online_order', ylabel='count'>

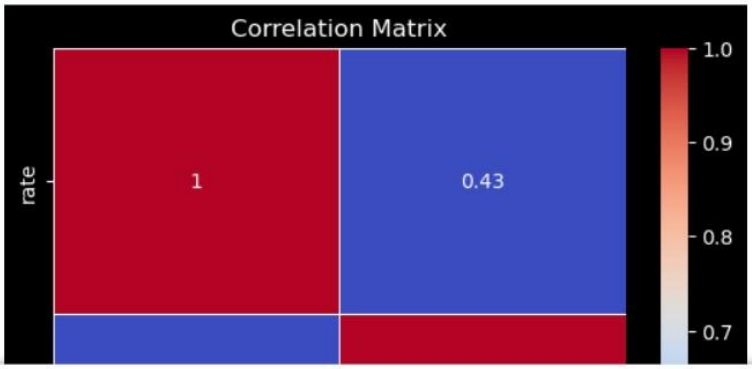


Out[28]: <Axes: xlabel='book_table', ylabel='count'>

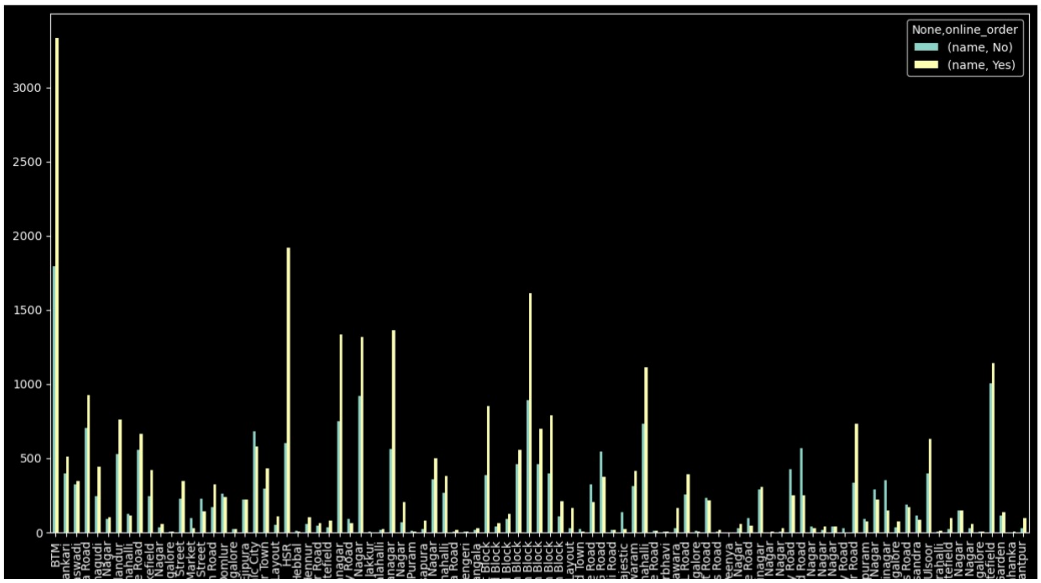


In [42]:
```
correlation_matrix = df[['book_table', 'rate', 'votes']].corr()
sns.heatmap(correlation_matrix, annot=True, cmap='coolwarm', linewidths=.5)
plt.title('Correlation Matrix')
plt.show()
```

```
C:\Users\kriti\AppData\Local\Temp\ipykernel_12772\706494751.py:1: FutureWarning: The default
me.corr is deprecated. In a future version, it will default to False. Select only valid colu
ic_only to silence this warning.
  correlation_matrix = df[['book_table', 'rate', 'votes']].corr()
```



Out[22]: <Axes: xlabel='location'>



```
In [51]: #ploting the heatmap for correlation
         plt.figure(figsize=(18,10))
         Corr_hmeq=df.corr()
         correlation_heatMap = sns.heatmap(Corr_hmeq, annot=True)
```

```
C:\Users\kriti\AppData\Local\Temp\ipykernel_20072\3223015564.py:3: FutureWarning: The default value of numeric_only in DataFram
e.corr is deprecated. In a future version, it will default to False. Select only valid columns or specify the value of numeric_
only to silence this warning.
  Corr_hmeq=df.corr()
```

CONCLUSION

The " **Zomato Bangalore Restaurant Explorer**" project has successfully unveiled a rich tapestry of insights into the city's vibrant culinary scene through a meticulous analysis of Zomato's restaurant data. The journey began with the comprehensive retrieval of accurate and up-to-date information, followed by a deep dive into the data's structure and content using exploratory data analysis techniques. Key findings emerged as the project progressed, revealing intriguing correlations between restaurant ratings and prices, shedding light on influential factors that shape customer perceptions. The exploration of cuisine trends provided a nuanced understanding of the diverse culinary landscape in Bangalore.

Location-based insights illuminated restaurant hotspots, offering valuable knowledge about the geographical dynamics of the city's dining establishments. The incorporation of interactive visualizations enhanced the project's communicative power, providing an engaging and visually appealing representation of the discovered patterns and trends. Optionally, predictive modeling offered the potential to anticipate restaurant-related outcomes, adding a layer of forward-looking analysis to the project.In the final stages, the project culminated in a compelling narrative, presented in Jupyter Notebooks with Markdown documentation. This narrative not only summarized the key findings but also provided a bridge between complex data analyses and a diverse audience, making the insights accessible and interesting.

The project's success lies in its ability to transform raw data into a story that paints a vivid picture of Bangalore's food culture, offering valuable recommendations and insights for both local enthusiasts and visitors exploring the city's culinary offerings.

REFRENCES

- https://www.kaggle.com/datasets

Thank You