

EatWell Data Analysis Report

A detailed analysis of restaurants in Bangalore to understand their culinary need in order to establish a new food chain in the city.

Link to the Dataset:

https://drive.google.com/file/d/1ASAtajtCIbmFvmGsdDghN-fryZ9nUND-/view?usp=drive_link

Content:

- * Problem Statement
- * Objective
- * Exploratory Data Analysis
- * Data Visualization
- * Interpretation/ Insights from the data
- * Conclusion

Problem Statement:

Embark on a culinary journey with 'Eat Well', a visionary chain restaurant startup in Bangalore. Our mission is to revolutionize the dining experience by offering delectable dishes tailored to the city's diverse palate. Through this project, we aim to develop captivating and interactive dashboards using Power BI, showcasing key insights into Bangalore's food preferences. These dashboards will serve as a beacon, guiding our data-driven decision-making process, and highlighting critical business metrics with a visually appealing flair.

Objective:

Eat Well, a startup chain restaurant, aims to enhance its operations and customer satisfaction through data-driven decision-making. The company seeks to develop multiple charts using data from various sources, including restaurant URLs, addresses, names, online ordering availability, table booking options, ratings, votes, contact information, locations, restaurant types, popular dishes, cuisines, average cost for two people, reviews, menu items, and categories and cities where restaurants are listed. The primary goal is to gain insights into customer preferences, popular cuisines, and restaurant trends in Bangalore, enabling the company to make informed decisions to improve customer satisfaction and optimize its offerings.

Exploratory Data Analysis:

Importing Required Libraries

```
import numpy as np
import pandas as pd
import matplotlib.pyplot as plt
import seaborn as sns
import warnings
warnings.filterwarnings('ignore')
```

Loading the Dataset

```
df = pd.read_csv(r"C:\Users\vamsh\Downloads\Data Analysis Projects\Data Analysis with Python\archive (1).zip", encoding='latin-1')
df
```

Cross...											
3	https://www.zomato.com/bangalore/addhuri-udupi...	1st Floor, Annakuteera, 3rd Stage, Banashankar...	Addhuri Udupi Bhojana	No	No	3.7/5	88	+91 9620009302	Banashankari	Qui	Bit
4	https://www.zomato.com/bangalore/grand-village...	10, 3rd Floor, Lakshmi Associates, Gandhi Baza...	Grand Village	No	No	3.8/5	166	+91 8026612447/r/n+91 9901210005	Basavanagudi	Cast	Dini
...
51712	https://www.zomato.com/bangalore/best-brews-fo...	Four Points by Sheraton Bengaluru, 43/3, White...	Best Brews - Four Points by Sheraton Bengaluru...	No	No	3.6 /5	27	080 40301477	Whitefield	B	
51713	https://www.zomato.com/bangalore/vinod-bar-and...	Number 10, Garudachar Palya, Mahadevanur...	Vinod Bar And Restaurant	No	No	NaN	0	+91 8197675843	Whitefield	B	

- **Statistical Analysis of Data:**

```
df.describe()
```

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

```
df.isnull().sum()
```

url	0
address	0
name	0
online_order	0
book_table	0
rate	7775
votes	0
phone	1208
location	21
rest_type	227
dish_liked	28078
cuisines	45
approx_cost(for two people)	346
reviews_list	0
menu_item	0
listed_in(type)	0
listed_in(city)	0
dtype:	int64

```
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
7   phone                                    50509 non-null  object
8   location                                 51696 non-null  object
9   rest_type                               51490 non-null  object
10  dish_liked                              23639 non-null  object
11  cuisines                                51672 non-null  object
12  approx_cost(for two people)              51371 non-null  object
13  reviews_list                            51717 non-null  object
14  menu_item                               51717 non-null  object
15  listed_in(type)                          51717 non-null  object
16  listed_in(city)                          51717 non-null  object
dtypes: int64(1), object(16)
memory usage: 6.7+ MB
```

- There are null values in multiple columns of the data, these null values has to be handled before performing the analysis as they deviate the results and cause mis interpretations.

- **Handling the Missing Values:**

```
# Handling Null Values
df = df.fillna(method='ffill')
df.head()
```

```
df.isna().sum()
```

```
url          0
address      0
name         0
online_order 0
book_table   0
rate         0
votes        0
phone        0
location     0
rest_type    0
dish_liked   0
cuisines     0
approx_cost(for two people) 0
reviews_list 0
menu_item    0
listed_in(type) 0
listed_in(city) 0
dtype: int64
```

- The null values are replaced by forward values using method= 'ffill'.
- Similarly the features without importance have been dropped and the cleaned data is used for analysis.

```
df.columns = df.columns.str.strip()
df.columns

Index(['url', 'address', 'name', 'online_order', 'book_table', 'rate', 'votes',
      'phone', 'location', 'rest_type', 'dish_liked', 'cuisines',
      'approx_cost(for two people)', 'reviews_list', 'menu_item',
      'listed_in(type)', 'listed_in(city)'],
      dtype='object')
```

```
df = df.rename(columns={'approx_cost(for two people)': 'app_cost_for_two'})
```

```
df = df.drop(['url', 'address', 'phone'], axis =1)
```

```
df.head()
```

	name	online_order	book_table	rate	votes	location	rest_type	dish_liked	cuisines	app_cost_for_two	reviews_list	menu_item	listed_in(type)
0	Jalsa	Yes	Yes	4.1	775	Banashankari	Casual Dining	Pasta, Lunch Buffet, Masala Papad, Paneer Laja...	North Indian, Mughlai, Chinese	800	['(Rated 4.0', 'RATED in A beautiful place to ...	[]	Buffet
1	Spice Elephant	Yes	No	4.1	787	Banashankari	Casual Dining	Momos, Lunch Buffet, Chocolate Nirvana, Thai G...	Chinese, North Indian, Thai	800	['(Rated 4.0', 'RATED in Had been here for din...	[]	Buffet

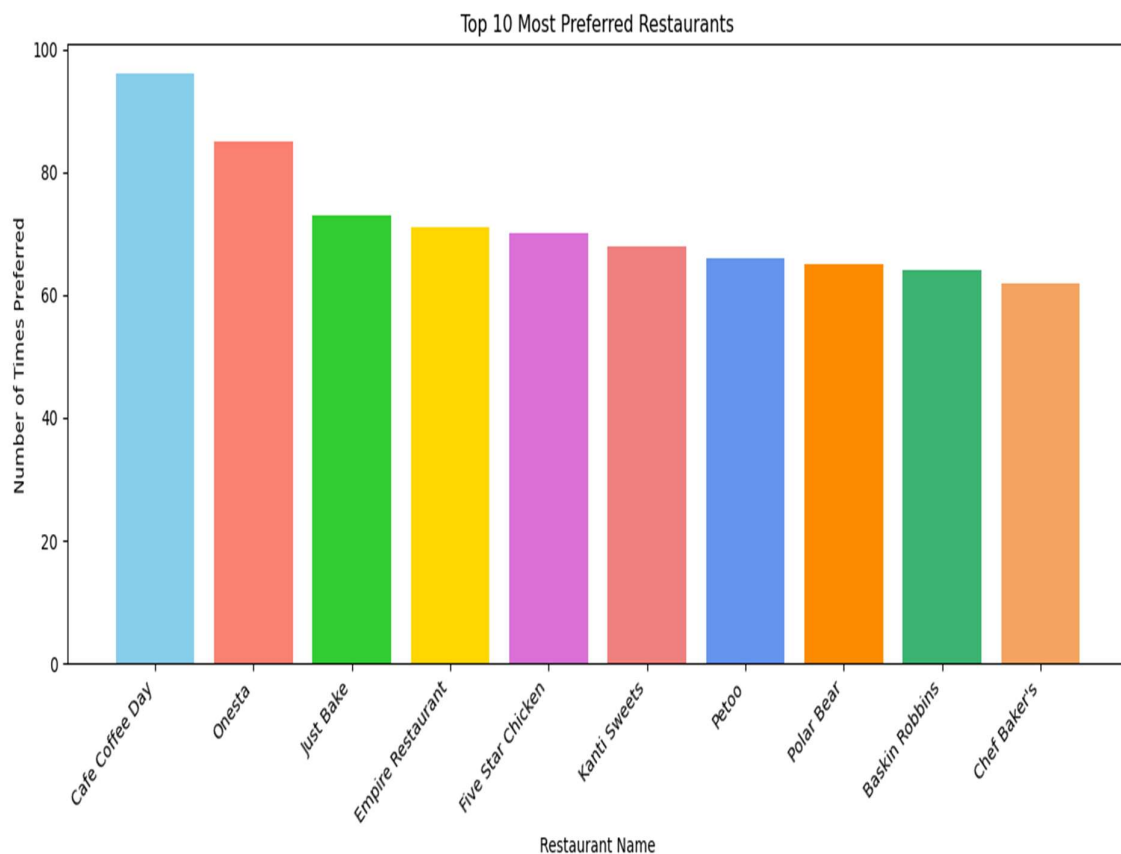
Data Visualization:

- **Top 10 Most Preferred Restaurants**

```
top_restaurants = df['name'].value_counts().head(10)

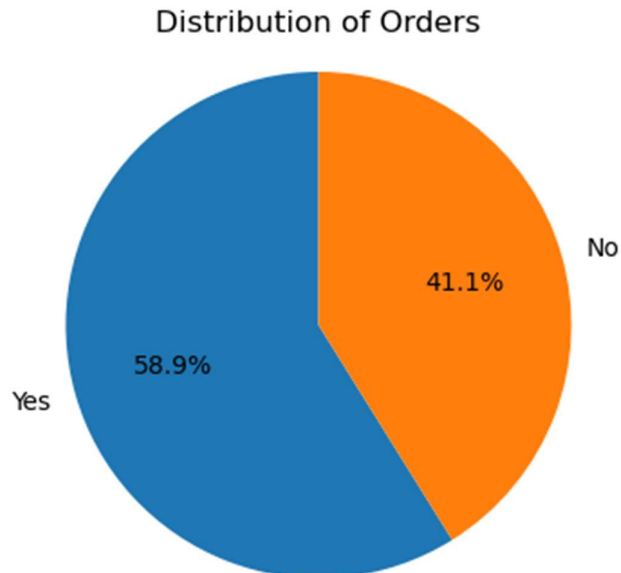
colors = ['skyblue', 'salmon', 'limegreen', 'gold', 'orchid', 'lightcoral', 'cornflowerblue', 'darkorange', 'mediumseagreen', 'sandybrown']

plt.figure(figsize=(12, 6))
bars = plt.bar(top_restaurants.index, top_restaurants.values, color=colors)
plt.xlabel('Restaurant Name')
plt.ylabel('Number of Times Preferred')
plt.title('Top 10 Most Preferred Restaurants')
plt.xticks(rotation=45, ha='right')
plt.tight_layout()
plt.show()
```



- **Ordering Dynamics: Exploring the Most Preferred Order Mode**

```
online_orders = df['online_order'].value_counts()
plt.figure(figsize = (6,4))
plt.pie(online_orders, labels = online_orders.index, autopct = '%1.1f%%', st
plt.title('Distribution of Orders')
plt.axis('equal')
plt.show()
```

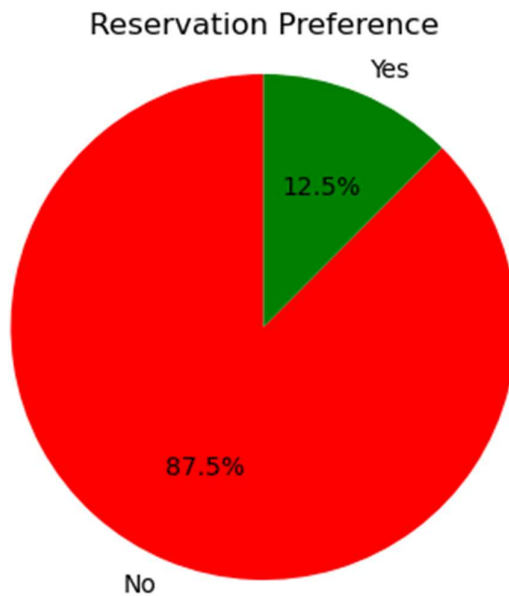


➤ **Interpretation:**

- 58.9% of the customers order food online, the reasons may vary from eating at comfort zone, escaping the Bangalore traffic, coupons and huge discounts on orders.
- 41.1% of the customers dine at the restaurants as they might be social gatherings, prefer it the old way of enjoying the meal with family, live in the neighbourhood, etc.

- **Reservation Preference: Book Table:**

```
book_table = df['book_table'].value_counts()
colors = ['red', 'green']
plt.figure(figsize = (6,4))
plt.pie(book_table, labels = book_table.index, autopct = '%1.1f%%', startangle=90)
plt.title('Reservation Preference')
plt.axis('equal')
plt.show()
```

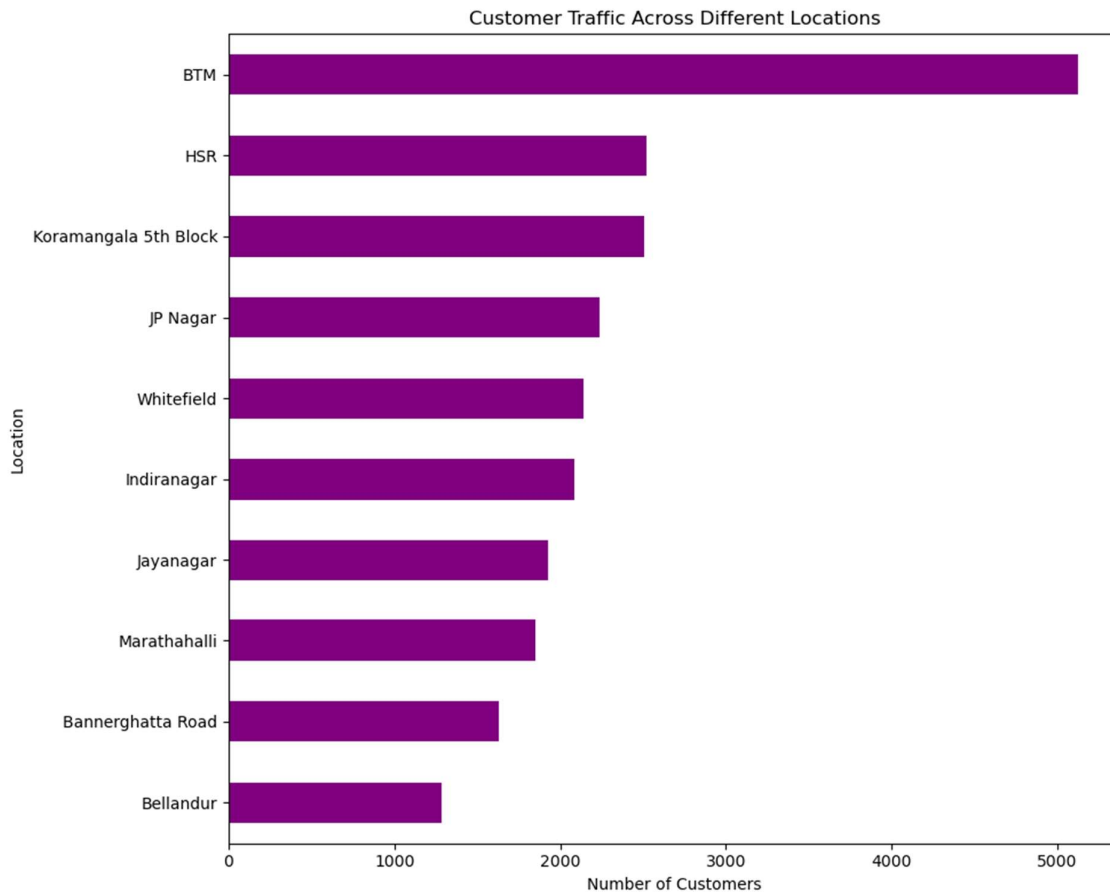


Interpretation:

- 87.5% of the customers are unwilling to book the table, the reasons may be unavailability of time, improper services, lack of space, prefer to eat at home, expensive etc.
- 12.5% customers are booking the table indicate social gathering like meeting friends, party, celebrations, dining with executives or family, company meetings, etc.
- This gives us a view that customers prefer to book tables only on special occasions so investment in furniture, decorations and other exterior may not be much useful. This investment can be diverted to hiring good chef, including more items in menu, providing quicker and affordable home services, improving quality of food.

- **Analysing Customer Traffic Across Different Location:**

```
location_traffic = df['location'].value_counts().head(10)
plt.figure(figsize=(10, 8))
location_traffic.sort_values().plot(kind='barh', color='purple')
plt.xlabel('Number of Customers')
plt.ylabel('Location')
plt.title('Customer Traffic Across Different Locations')
plt.tight_layout()
plt.show()
```

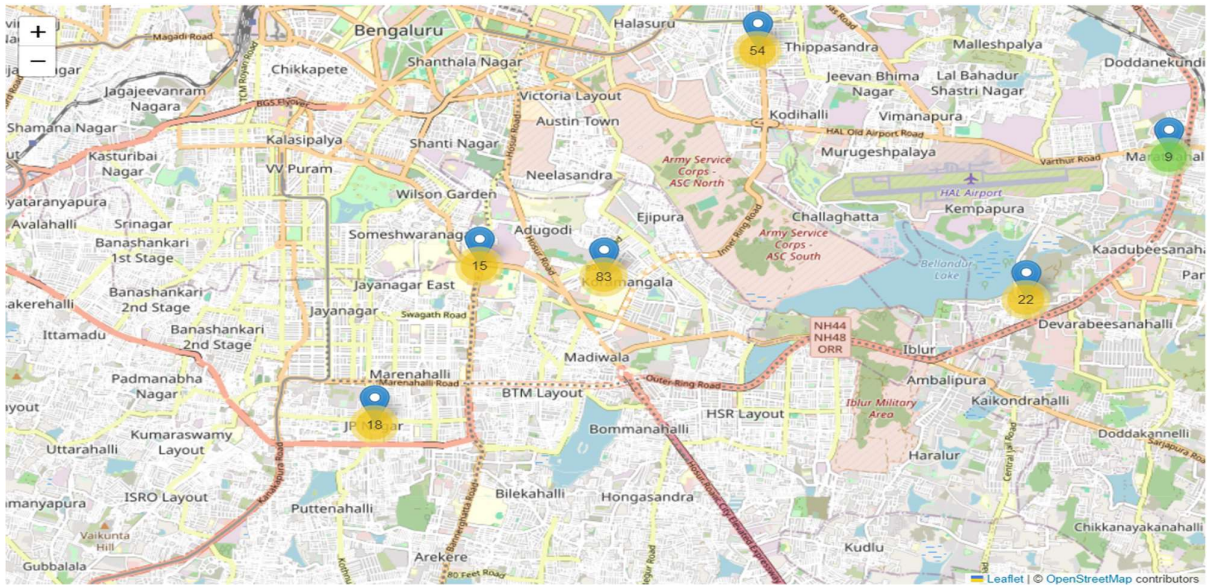


Interpretation

BTM, HSR, Koramangala 5th Block, JP Nagar, Whitefield are the the TOP-5 location with the highest customer traffic. The reasons for this are:-

- The locations contains public attractions like Malls, Theatres, Parks, Government Buildings, etc.
- They are easily accessible as they are present in the heart of the city -- easy connection.
- Presence of IT Parks and housing colonies play a major role in the huge customer density.
- Word of Mouth for the good services and quality of food attracts even more customers in the region.

- Geographical Location of Various Restaurant in Bangalore City



- Top Rated Restaurants:

```
import re

def clean_name(name):
    cleaned_name = re.sub(r'^a-zA-Z0-9\s', '', name)
    return cleaned_name.strip()

df['Name of the Restaurant'] = df['name'].apply(clean_name)

df['rate_numeric'] = df['rate'].str.extract('(\d+\.\d*)').astype(float)

avg_ratings = df.groupby('Name of the Restaurant')['rate_numeric'].mean().sort_values(ascending=False).round(2)
topRatedRestaurants = avg_ratings.head(10)
topRatedRestaurants = pd.DataFrame(topRatedRestaurants)
topRatedRestaurants = topRatedRestaurants.rename(columns={'rate_numeric': 'Average Ratings'})
topRatedRestaurants
```

Average Ratings	
Name of the Restaurant	
Sant Spa Cuisine	4.90
Byg Brewski Brewing Company	4.90
Asia Kitchen By Mainland China	4.90
Punjab Grill	4.87
Belgian Waffle Factory	4.84
Flechazo	4.80
The Pizza Bakery	4.80
OG Variar Sons	4.80
Biergarten	4.77
Barbecue by Punjab Grill	4.75

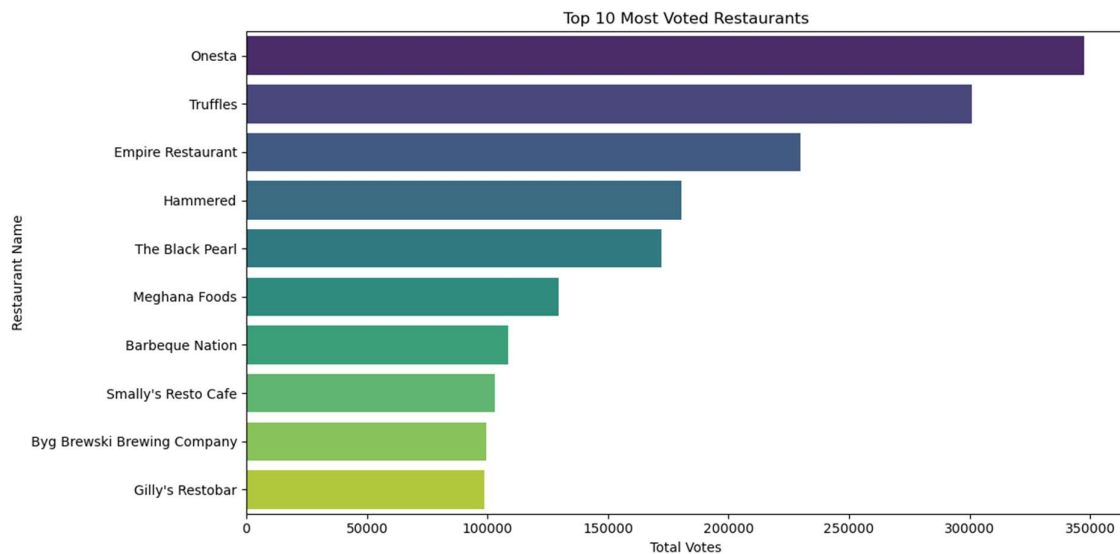
- **Customer Recommended Restaurant:**

```
total_votes = df.groupby('name')['votes'].sum().sort_values(ascending=False)

most_voted_restaurants = total_votes.head(10)
most_voted_restaurants
```

```
name
Onesta                347520
Truffles              301059
Empire Restaurant     229808
Hammered              180602
The Black Pearl       172122
Meghana Foods         129557
Barbeque Nation       108425
Smally's Resto Cafe   102877
Byg Brewski Brewing Company 99531
Gilly's Restobar      98808
Name: votes, dtype: int64
```

```
plt.figure(figsize=(12, 6))
sns.barplot(x=most_voted_restaurants.values, y=most_voted_restaurants.index, palette='viridis')
plt.xlabel('Total Votes')
plt.ylabel('Restaurant Name')
plt.title('Top 10 Most Voted Restaurants')
plt.tight_layout()
plt.show()
```



- **Highly Preferred Locations:**

```
import re

def extract_numeric_rate(rate):
    try:
        return float(re.search(r'\d+\.\d+', rate).group())
    except AttributeError:
        return None

df['numeric_rate'] = df['rate'].apply(extract_numeric_rate)
avg_rating_by_location = df.groupby('location')['numeric_rate'].mean().sort_values(ascending=False)
top_avg_rating_by_location = avg_rating_by_location.head(10)
top_avg_rating_by_location
```

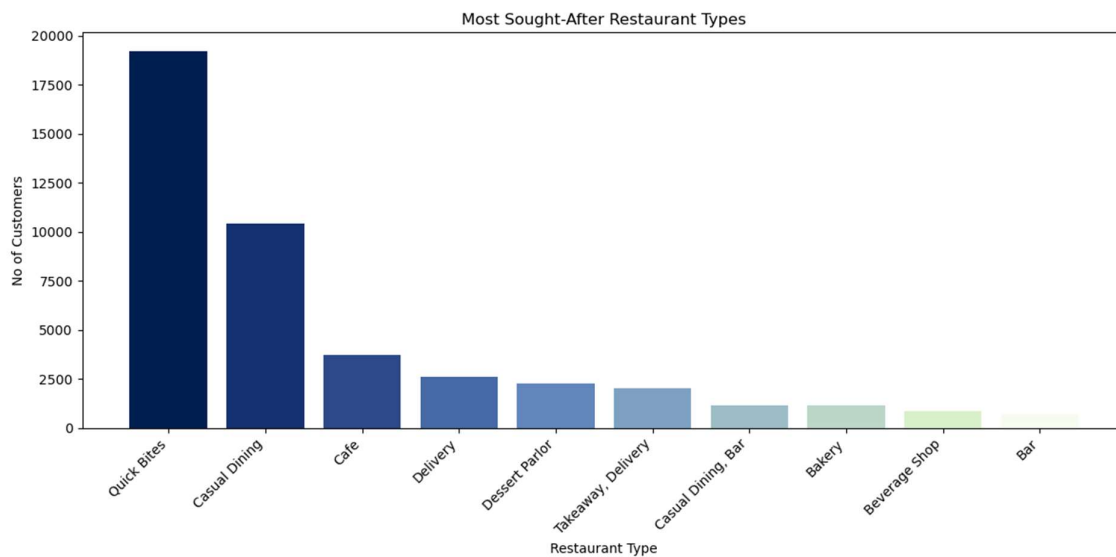
```
location
Lavelle Road      4.108527
St. Marks Road    4.001136
Koramangala 5th Block  3.979992
Church Street      3.970796
Koramangala 3rd Block  3.960377
Sankey Road        3.948148
Cunningham Road     3.893469
Koramangala 4th Block  3.858087
Rajarajeshwari Nagar  3.850000
Sadashiv Nagar      3.840323
Name: numeric_rate, dtype: float64
```

These are the highly rated/ preferred locations for food. Establishing a Food outlet in these locations boosts the reputation of the product and leads to high sales.

- **Popular Restaurant Categories:**

```
rest_type = df.rest_type.value_counts().head(10)
colors = ['#002051', '#163172', '#2d498a', '#466aa5', '#6285bb',
          '#7fa1c1', '#9dbcc5', '#bcd6c8', '#d9f1ca', '#f7fcf0']

plt.figure(figsize=(12, 6))
plt.bar(rest_type.index, rest_type.values, color=colors)
plt.xlabel('Restaurant Type')
plt.ylabel('No of Customers')
plt.title('Most Sought-After Restaurant Types')
plt.xticks(rotation=45, ha='right')
plt.tight_layout()
plt.show()
```

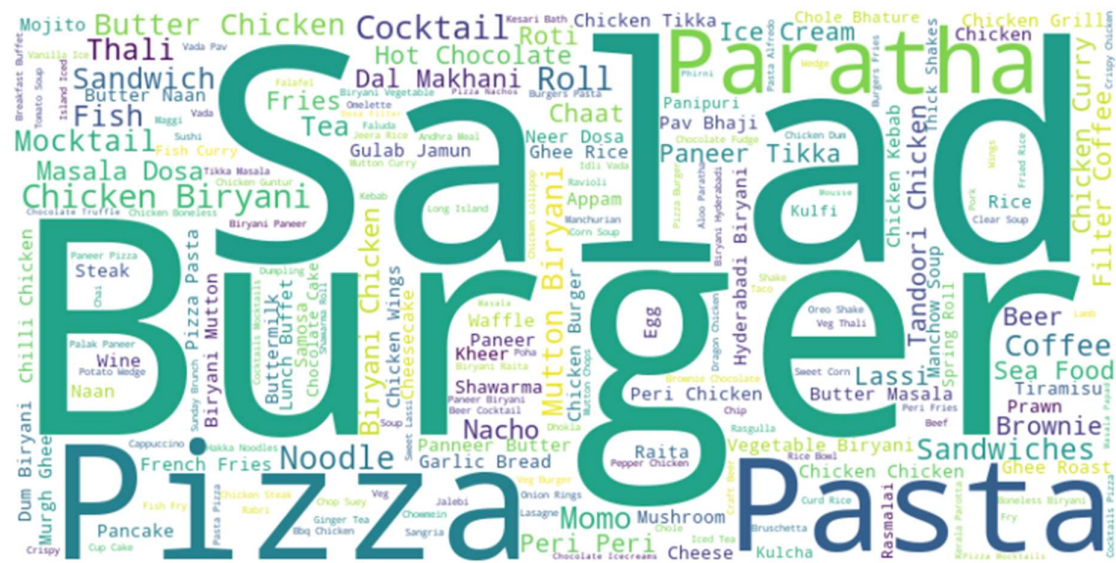


- **Most Ordered Dishes:**

```
from wordcloud import WordCloud

text = ' '.join(dishes['dish_liked'])
wordcloud = WordCloud(width=800, height=400, background_color='white').generate(text)

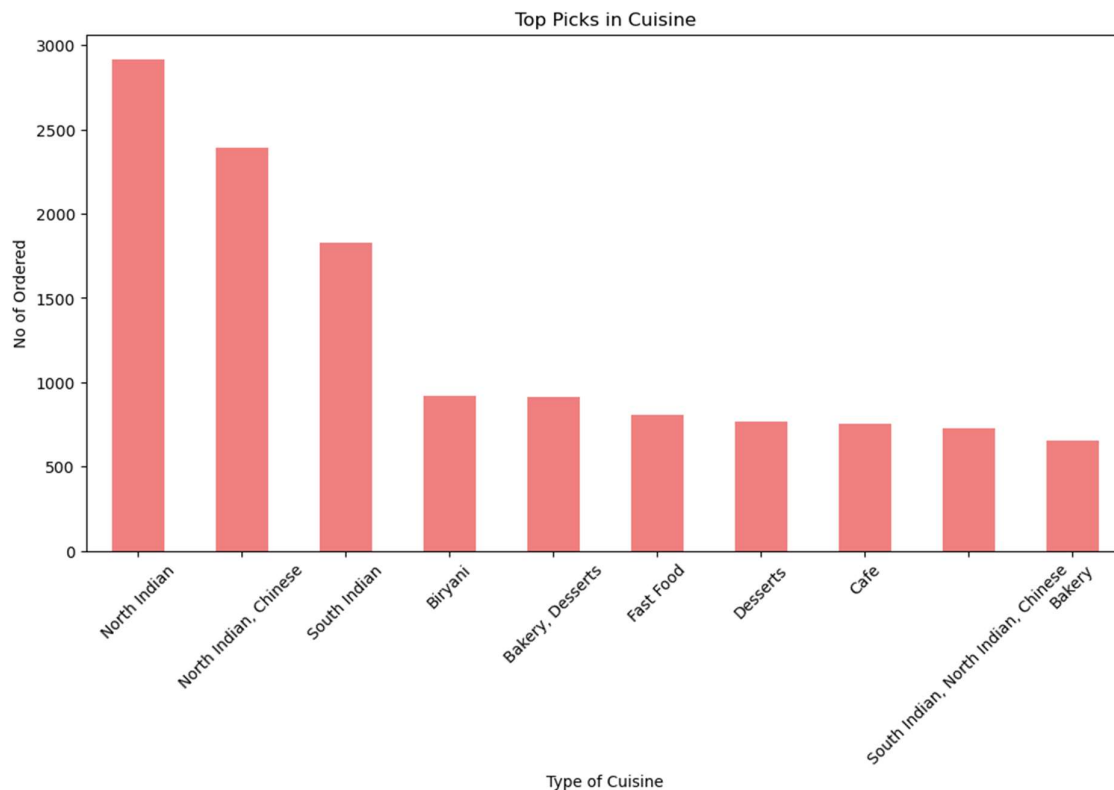
plt.figure(figsize=(10, 6))
plt.imshow(wordcloud, interpolation='bilinear')
plt.axis('off')
plt.show()
```



- Burger, Salad, Pizza, Pasta, Paratha, Chicken Biryani are the top dishes ordered by the customers.
- Fast food are in great demand than regular meals.

- **Popular Cuisines Among Customers:**

```
cuisines = df.cuisines.value_counts().head(10)
plt.figure(figsize=(12, 6))
cuisines.plot(kind='bar', color='lightcoral')
plt.title('Top Picks in Cuisine')
plt.xlabel('Type of Cuisine')
plt.ylabel('No of Ordered')
plt.xticks(rotation=45)
plt.show()
```



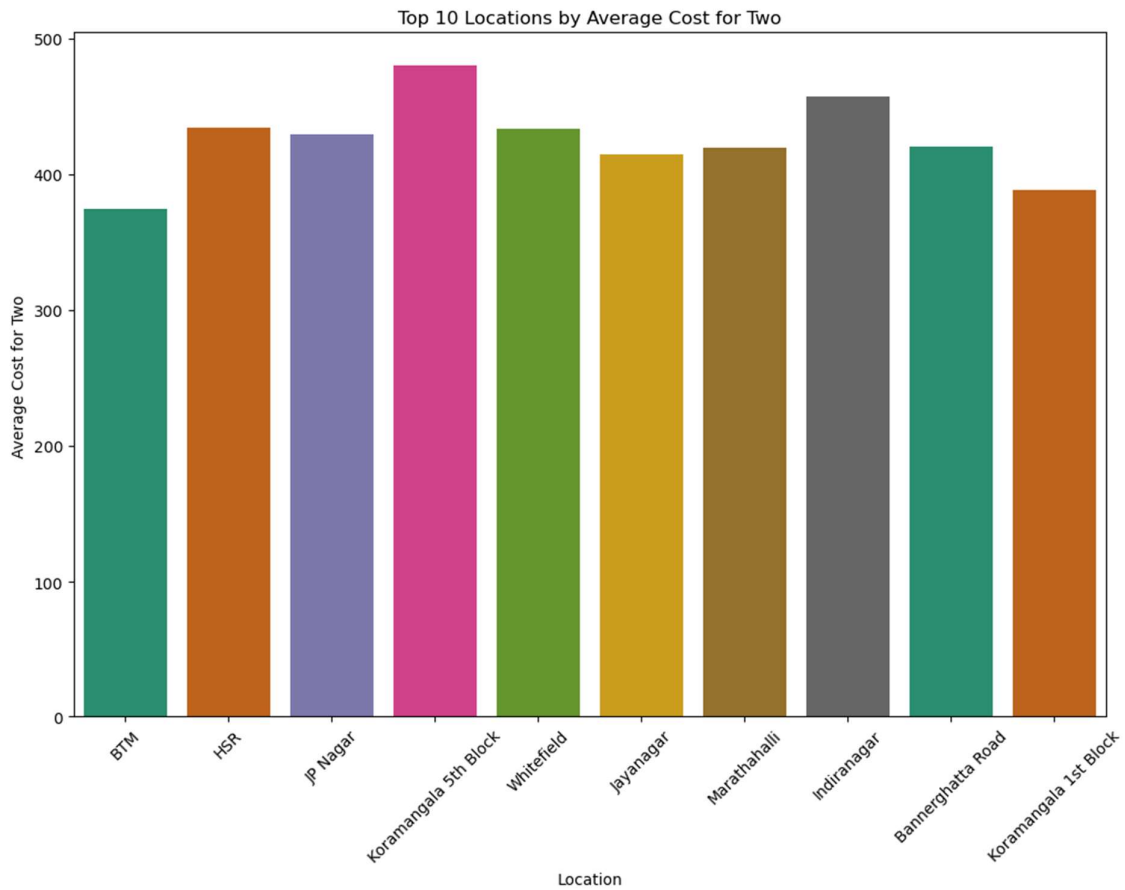
Insights:

- North Indian, Chinese, South Indian Cuisines are in holding the high foot in the lives of the people from Bangalore.
- This is followed by Biryani, Bakery, Desserts, Fast foods, etc.
- Overall, the data suggests a diverse taste in cuisine types, with a strong preference for North Indian, Chinese, and South Indian cuisines, as well as bakery and dessert items.

- **Average Cost for Two in Top 10 Locations:**

```
df['app_cost_for_two'] = pd.to_numeric(df['app_cost_for_two'], errors='coerce')
df = df.dropna(subset=['app_cost_for_two'])
top_10_locations = df['location'].value_counts().head(10)
df_top_10 = df[df['location'].isin(top_10_locations.index)]
average_cost = df_top_10.groupby('location')['app_cost_for_two'].mean().loc[top_10_locations.index]

plt.figure(figsize=(12, 8))
sns.barplot(x=average_cost.index, y=average_cost, palette='Dark2')
plt.xlabel('Location')
plt.ylabel('Average Cost for Two')
plt.xticks(rotation=45)
plt.title('Top 10 Locations by Average Cost for Two')
plt.show()
```



➤ The average cost for two to order food in top 10 locations in Bangalore is more than RS. 400.

- **Common Menu:**

```
common_menu = df['menu_item'].value_counts()
common_menu = pd.DataFrame(common_menu)
common_menu.drop(common_menu.index[0], inplace = True)
common_menu.head(10)
```

	count
menu_item	
[Butter Chicken Pizza', 'Bombay Veggie Burger', 'Tawa Paneer Burger', 'Oh So Cheesy Burger', 'Barbecue Chicken Burger', 'Peri Peri Chicken Burger', 'BBQ Lamb', 'Cheesy Chicken', 'Crunchy Ferrero', 'Lots of Nuts', 'Kadhai Paneer Pizza', 'Chilli Paneer and Mushroom Pizza', 'Chipotle Veggie Pizza', 'Creamy Cheese Pizza', 'Southern Veggie Korma Pizza', 'Popeye, The Corny Man Pizza', 'Paneer Tikka Pizza', 'Desi Margherita Pizza', 'Roasted Paneer with Mustard Pizza', 'Bangalore Express Pizza', 'Italian Chaska Pizza', 'Mayo and Cheese Pizza', 'Classic Margherita', 'Cajun Hawaiian Pizza', 'Butter Chicken Pizza', 'Lasooni Bhuna Murg Pizza', 'Chilly Chicken Pizza', 'Hot Garlic Chicken Pizza', 'Murg Barbecue Pizza', 'Mediterranean Mutton Keema Pizza', 'Super Green Burger', 'Quinoa Black Bean Burger', 'Bombay Veggie Burger', 'Tawa Paneer Burger', 'Oh So Cheesy Burger', 'Barbecue Chicken Burger', 'Chipotle Lamb Burger', 'Crumb Fried Fish Burger', 'Peri Peri Chicken Burger', 'BBQ Lamb', 'Cheesy Chicken', 'Picnic Chicken Burger', 'Korean Grilled Chicken with Kimchi', 'Kiddy Kat Strawberry', 'Kiddy Kat Chocolate', 'Bubblegum', 'Oreo Shake', 'Caramel Shake', 'Chocolate Shake', 'Strawberry Shake', 'Vanilla Shake', 'Cold Coffee', 'Virgin Mojito', 'Iced Tea', 'Hot Chocolate', 'Cappuccino', 'Latte', 'Espresso Shot', 'Brownie Shake', 'Red Velvet Cheesecake Shake', 'Oreo Cheesecake Shake', 'Salted Caramel Popcorn', 'Coconut Crumble', 'Coffee Crunch', 'Bira', 'Thanda Paan', 'Sticky Toffee Pudding', 'Blueberry Cheesecake', 'Nutella Cheesecake', 'Caramel Custard with Strawberry or Figs', 'Red Velvet Cheesecake', 'Tiramisu', 'Zesty Vanilla and Topsy Brownie', 'Oreo Explosion', 'Chocolate Almond Crumble', 'Crunchy Ferrero', 'Stoner's Chocolate Decadence', 'Brownie Bash', 'Chocolate Lava', 'Dark Chocolate Caramel', 'Chocolate Overload', 'Lots of Nuts', 'Peanut Butter Crunch', 'Nutty Fudgy', 'Lemon Pistachio Biscotti', 'Fruilicious Ice Cream', 'Strawberry Tease Ice Cream', 'Mango Tango Ice Cream', 'Lychee Lovers Ice Cream', 'Monkey Business Ice Cream', 'Apple Crumble Ice Cream', 'Blueberry Bliss Ice Cream']	
[Avil Milk', 'Oreo Shake', 'Chocolate Shake', 'Royal Falooda', 'Fruits Salad with Ice Cream', 'Fruits Salad with Ice Cream', 'Blackcurrant Ice Cream', 'Spanish Delight Ice Cream', 'Chillout Special Ice Cream', 'Chocolate Ice Cream', 'Butterscotch Ice Cream', 'Pista Ice Cream', 'Strawberry Ice Cream', 'Vanilla Ice Cream', 'Royal Falooda', 'Strawberry Falooda', 'Chillout Special Falooda', 'Lemon Juice', 'Mint Lemon Juice', 'Ginger Lemon Juice', 'Lemon Soda', 'Mint Lemon Soda', 'CAP Juice', 'Kiwi Juice', 'Carrot Juice', 'Pomegranate Juice', 'Watermelon Juice', 'Muskmelon Juice', 'Anjeer Juice', 'Pure Strawberry Juice', 'Papaya Juice',	

- Most followed Common Menu across various restaurants in Bangalore.

Conclusion:

From the analysis performed the following points can be shared:

- The people of Bangalore prefer to order food online, so the suggestion goes to EatWell is to build it in a way that it can serve more online customers.
- The reservation of tables is found quiet less this tells to concentrate more on quality of food than the exteriors.
- The customer Traffic in BTM is quite high so this can be the best place to set up the service, as busy locations attract more and regular customers and decrease the loss of food leftovers.
- The average cost of food in premium restaurants is above RS.400 so make sure that food is within the affordable price of customers.
- Bangalore is the IT hub of India, the city consists of people all over India and they prefer to have North Indian cuisines, Chinese Cuisines and South Indian cuisines. They prefer to have quick food than lengthy meals so a food chain that serves in Comfortable environment in the highly rated food location with the menus mentioned above has a high rate of success.