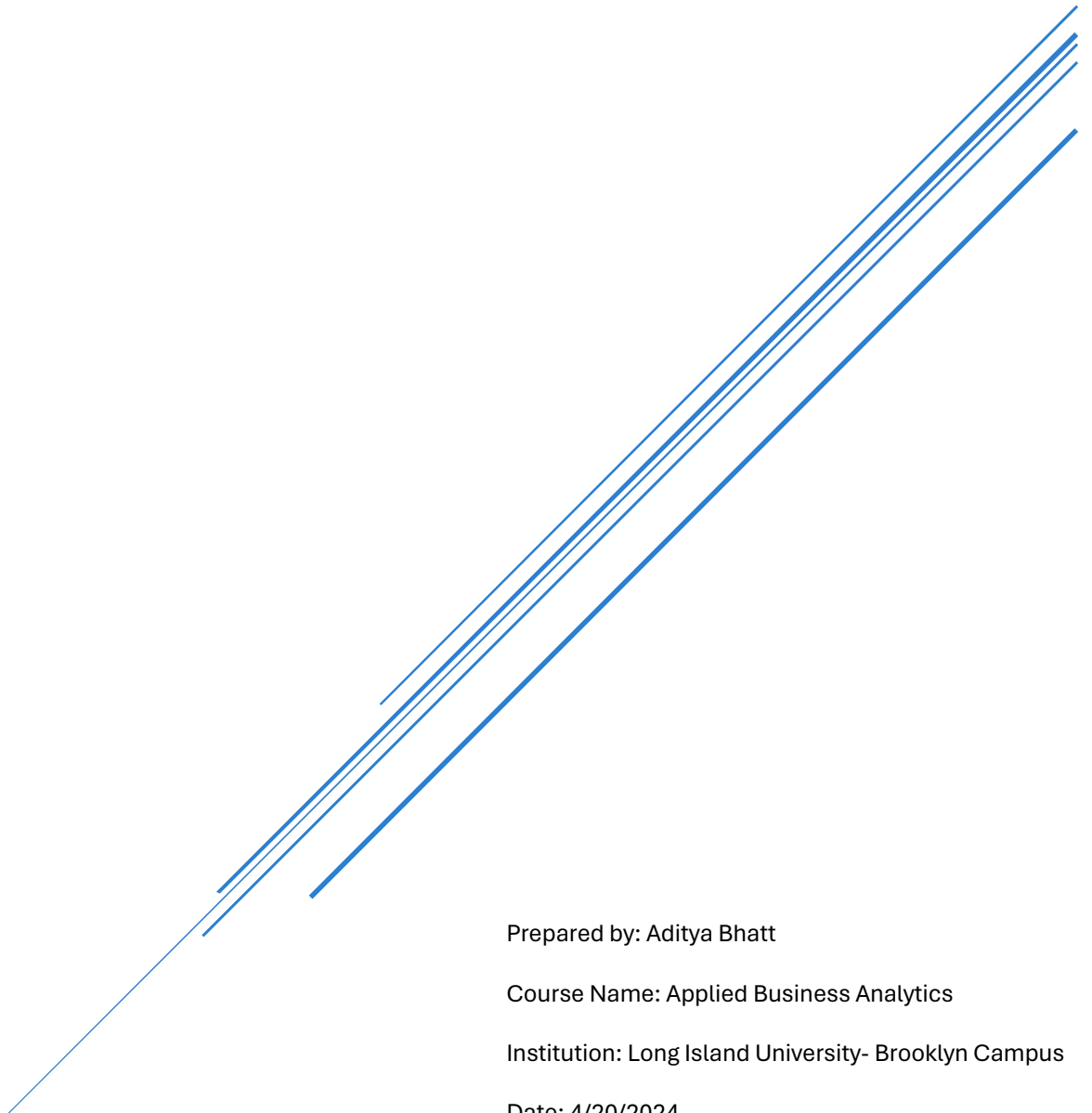Project Title: **Data-Driven Dining: Leveraging Bengaluru's Culinary Landscape for Restaurant Success**

Strategic Insights from Zomato Data Analysis

Prepared by: Aditya Bhatt

Course Name: Applied Business Analytics

Institution: Long Island University- Brooklyn Campus

Date: 4/20/2024

Appendix:

Important Aspect	Description
Abstract	Provides an overview of the project's objective to offer data-driven insights and strategies for aspiring restaurateurs in Bengaluru.
Dataset Description	Describes the Zomato dataset used, including details about the various features and their meanings.
Background	Provides context on Bengaluru's vibrant culinary landscape and the opportunities and challenges faced by aspiring restaurateurs.
Goals and Objectives	Outlines the primary goal of empowering restaurateurs with data-driven insights and the specific objectives, such as understanding the culinary landscape, identifying success factors, exploring demographic influences, and evaluating pricing strategies.
Data Selection Criteria	Explains the rationale behind selecting the Zomato dataset due to its comprehensive coverage of restaurants in Bengaluru and the availability of diverse information.
Data Cleaning	Mentions the importance of data cleaning by removing duplicates, NaN values, and performing necessary transformations.
Data Visualization	Includes various visualizations, such as heatmaps, column charts, and bar charts, to analyze and interpret the data, providing insights into factors like online ordering, table booking, location dynamics, and popular restaurant chains.
Data Prediction Models	Describes the use of linear regression, decision tree, random forest regression, and extra tree regression models to predict ratings and identify the most important features.
Recommendation Engine	Discusses the advantages and opportunities of implementing a recommendation engine for personalized customer experiences, increased engagement, and data-driven decision-making.
Sentiment Analysis	Highlights the value of sentiment analysis in gauging customer feedback and opinions, identifying areas for improvement, and tailoring offerings to meet preferences.

Important Aspect	Description
Conclusion	Summarizes the key considerations for new restaurant seekers, including service, ambiance, food quality, and value.

Abstract: This project offers a roadmap for aspiring restaurateurs seeking to navigate Bengaluru's dynamic food scene with confidence. Harnessing the power of data sourced from Zomato, this study unveils key insights into the factors influencing restaurant success, including location dynamics, cuisine preferences, pricing strategies, and thematic trends. By leveraging these data-driven insights, entrepreneurs can make informed decisions to establish and differentiate their restaurants effectively, catering to the diverse tastes and preferences of Bengaluru's vibrant population.

Dataset: https://www.kaggle.com/code/vahidehdashti/complete-zomato-dataset-eda

Data Definition:

1. url: The URL of the restaurant page on Zomato's website. (Text)

2. address: The physical address of the restaurant, including street name, locality, and city. (Text)

3. name: The name of the restaurant. (Text)

4. online_order: Indicates whether the restaurant accepts online orders (Yes/No). (Categorical)

5. book_table: Indicates whether the restaurant allows table booking (Yes/No). (Categorical)

6. rate: The average rating of the restaurant, typically on a scale of 1 to 5. (Numeric)

7. votes: The number of votes/ratings the restaurant has received. (Numeric)

8. phone: The contact phone number of the restaurant. (Text)

9. location: The locality or area where the restaurant is situated. (Text)

10. rest_type: The type of restaurant (e.g., Casual Dining, Fine Dining, Cafe, etc.). (Text)

11. dish_liked: Popular dishes liked by customers at the restaurant. (Text)

12. cuisines: The types of cuisines offered by the restaurant. (Text)

13. approx_cost(for two people): The approximate cost for two people to dine at the restaurant. (Numeric)

14. reviews_list: A list of reviews and ratings provided by customers. (Text)

15. menu_item: The menu items offered by the restaurant. (Text)

16. listed_in(type): The type of listing on Zomato (e.g., Buffet, Cafes, Delivery, etc.). (Text)

17. listed_in(city): The city in which the restaurant is listed. (Text)

Background: The culinary landscape of Bengaluru, often hailed as the gastronomic capital of India, is as diverse and vibrant as the city itself. With a population exceeding 12 million and a burgeoning dining scene boasting over 12,000 restaurants, Bengaluru offers a kaleidoscope of flavors, cuisines, and dining experiences to its residents and visitors alike.

In recent years, Bengaluru has witnessed a culinary renaissance fueled by a confluence of factors, including rapid urbanization, increasing disposable incomes, and a burgeoning food culture influenced by the city's cosmopolitan populace. This culinary renaissance has not only transformed Bengaluru into a melting pot of culinary diversity but has also presented aspiring restaurateurs with unparalleled opportunities and challenges.

For entrepreneurs venturing into the competitive world of restaurant ownership, understanding the intricacies of Bengaluru's dynamic food landscape is paramount to success. From choosing the right location to crafting a menu that resonates with discerning palates, every decision can make or break a restaurant's fortunes in this bustling metropolis.

To navigate this complex terrain effectively, aspiring restaurateurs require actionable insights and strategies grounded in data-driven analysis. This is where the Zomato dataset comes into play. By leveraging the rich repository of data provided by Zomato, including restaurant details, customer reviews, ratings, and menu offerings, entrepreneurs can gain invaluable insights into Bengaluru's culinary ecosystem.

Through a comprehensive analysis of the Zomato dataset, aspiring restaurateurs can unravel the mysteries of Bengaluru's culinary landscape, identifying key success factors, understanding demographic influences, evaluating pricing strategies, and uncovering emerging trends. Armed with these insights, entrepreneurs can make informed decisions at every stage of the restaurant journey, from conceptualization to execution, thereby maximizing their chances of success in Bengaluru's thriving restaurant industry.

In essence, the Zomato dataset serves as a treasure trove of information, illuminating the path to restaurant success in Bengaluru's dynamic and ever-evolving culinary scene. By harnessing the power of data-driven analysis, aspiring restaurateurs can turn their culinary dreams into reality, carving out a niche in the vibrant tapestry of Bengaluru's gastronomic landscape.

Goals and Objectives:

Primary Goal: Empower aspiring restaurateurs with data-driven insights and strategies to establish and differentiate their restaurants successfully in Bengaluru.

Objectives:

1. **Understand Bengaluru's Culinary Landscape:** Analyze data sourced from Zomato to gain insights into the diverse array of restaurants, cuisine preferences, thematic trends, reviews, and ratings prevalent in Bengaluru.

2. **Identify Key Success Factors:** Uncover the factors influencing restaurant success, including location dynamics, online ordering trends, booking preferences, customer reviews, and ratings.

3. **Explore Demographic Influences:** Investigate how demographic factors such as age, income, and cultural background impact dining preferences and restaurant choices in different localities of Bengaluru, reflected in reviews and ratings.

4. **Evaluate Pricing Strategies:** Analyze the relationship between pricing, menu offerings, customer reviews, and ratings to identify effective pricing strategies for restaurants in Bengaluru.

5. **Provide Actionable Recommendations:** Offer strategic recommendations based on data analysis, including insights from reviews and ratings, to assist aspiring restaurateurs in making informed decisions regarding location selection, menu development, pricing, and marketing strategies.

Data Selection Criteria:

The data for this project was sourced from Zomato, a leading restaurant discovery platform, due to its comprehensive coverage of restaurants in Bengaluru and the availability of diverse information such as restaurant details, customer reviews, ratings, and menu items. Zomato's dataset provides valuable insights into Bengaluru's culinary landscape, allowing for a comprehensive analysis of factors influencing restaurant success.

Data Cleaning: Before diving into coding, I ensure to clean the data by removing duplicates and NaN values, while also performing necessary transformations to optimize its usability and accuracy.

This preparatory step sets a solid foundation for efficient and effective data analysis or processing.

Data Visualization:

1. Heatmap:
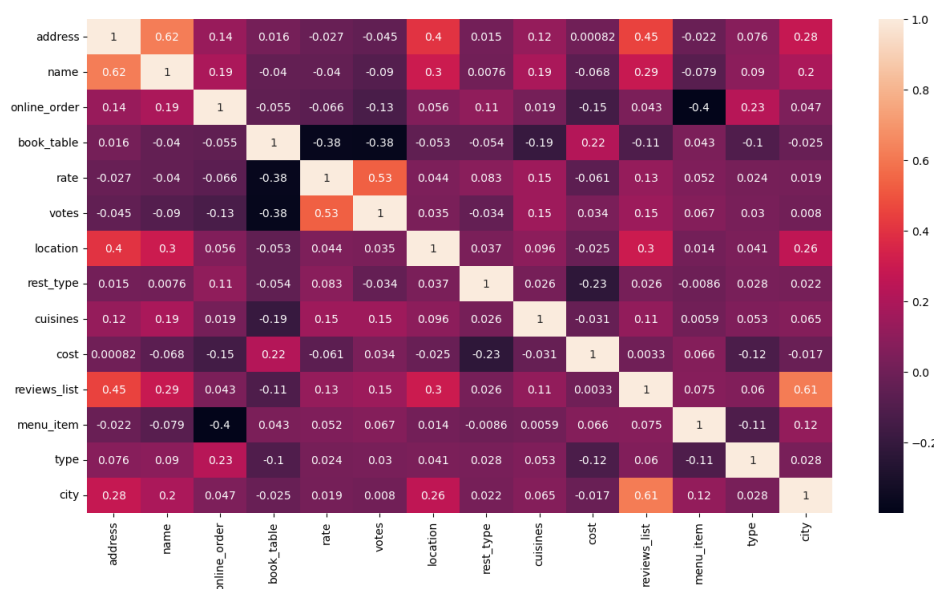
Based on the heatmap results Reviews and Rate are mostly positive correlation with all the features excepts online orders, orders, and booking tables.

2. Column Chart: Restaurants delivery online or Not:

Mostly Restaurants are delivery the foods online which gives a boost. The surge in online food delivery services has revolutionized the restaurant industry, offering unparalleled convenience to customers and a significant boost to businesses. With the click of a button, diners can now enjoy their favorite dishes from the comfort of their homes, driving remarkable growth in the digital food delivery market.

In response to the digital dining trend, many new restaurants are implementing innovative strategies to optimize their online ordering systems. From user-friendly mobile apps to seamless website interfaces, these establishments are capitalizing on the convenience of online food delivery to reach a wider audience and enhance customer satisfaction. By embracing technology, they're not just serving meals; they're delivering exceptional dining experiences right to customers' doorsteps.


Restaurants delivering online or Not

3. Column Chart: Restaurants allowing table booking or not:
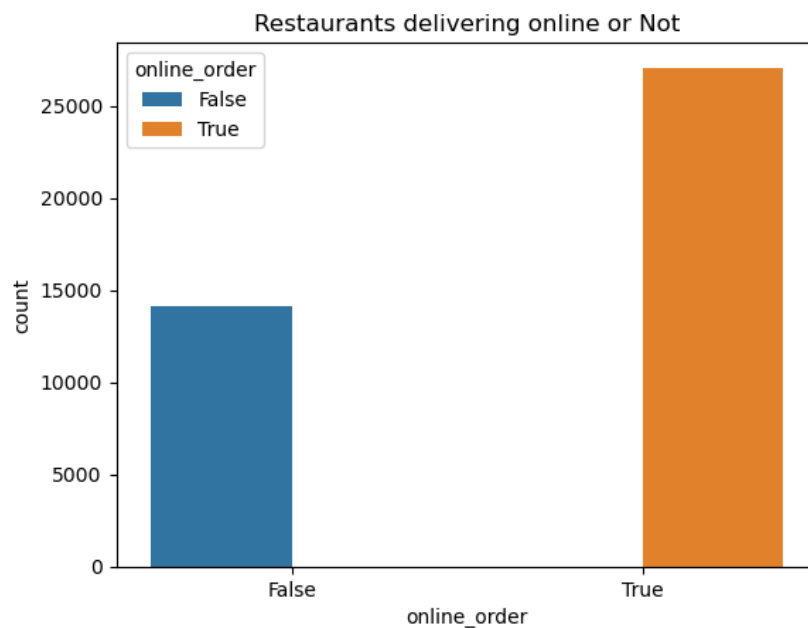
Restaurants allowing table booking or not
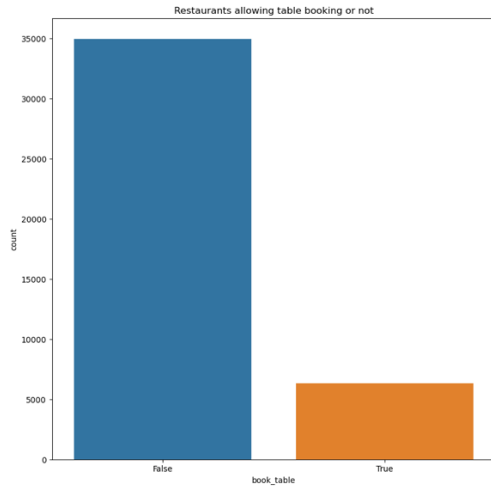
Based on the above chart mostly restaurants are not allowing table booking. New restaurants should carefully balance the benefits and drawbacks of table booking and first-come, first-served table service. Offering a mix of both can cater to different customer preferences while optimizing seating efficiency. Additionally, maintaining clear communication with customers about wait times and reservation policies is essential for managing expectations and ensuring a positive dining experience.

4. Column Chart: Table booking vs rate:
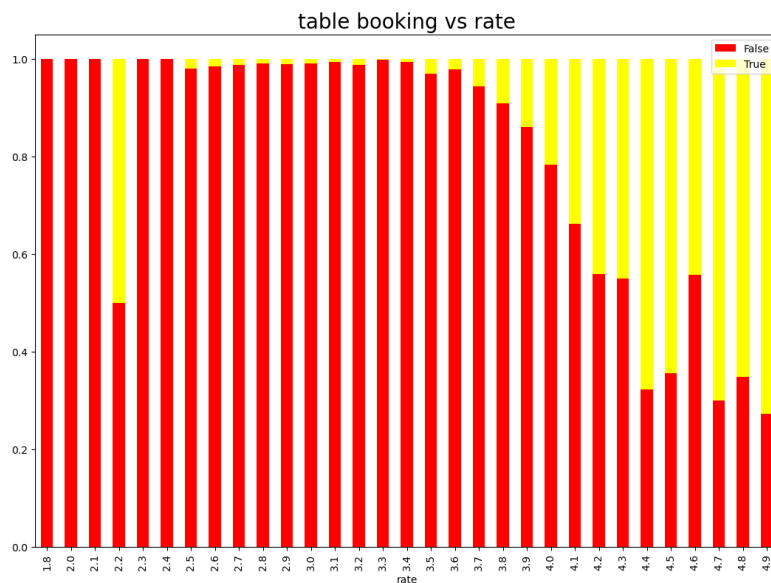
In the above chart we can observe that the restaurants who do 100 % table booking facilities that have low ratings. The restaurants who do online and first come first serve table options have good ratings.

New restaurants must keep in mind that both options they must work on as to satisfy the customer's need.

5. Column chart: City Vs Count:
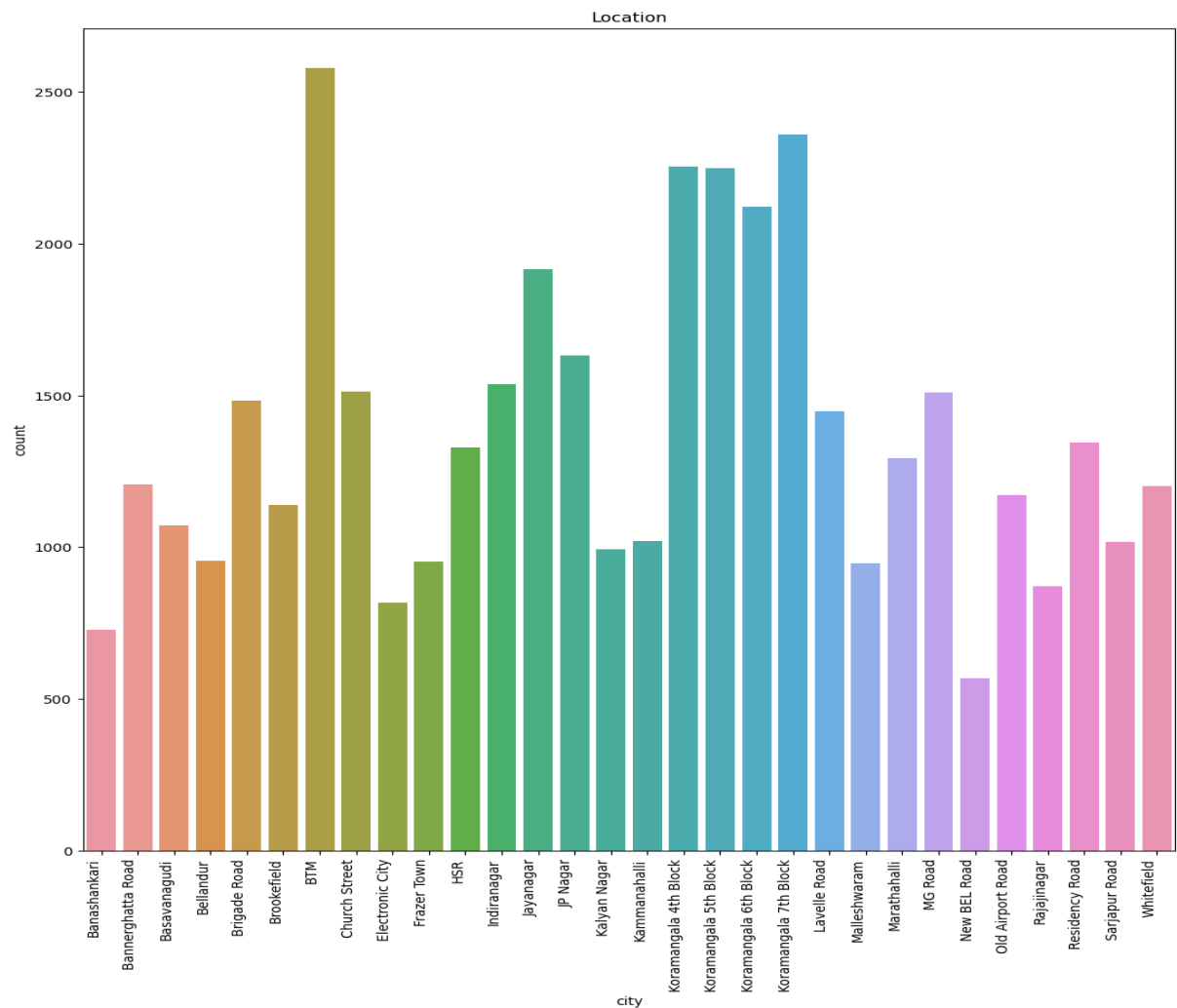
The data highlights BTM Koramangala as prime restaurant locations, suggesting high demand and potential for success. New restaurant seekers should prioritize thorough market research and strategic location selection to capitalize on bustling areas like BTM Koramangala, maximizing visibility and foot traffic for their establishments.

In addition to traditional data analysis, leveraging local insights and community feedback can offer a unique perspective on ideal restaurant locations. Understanding the demographic, cultural, and culinary preferences of the area can further guide new restaurant seekers in making informed decisions, ensuring their concept resonates with the surrounding community in a distinctive and impactful way.

6. Column Chart: Types of Service vs Count:

Type of Service
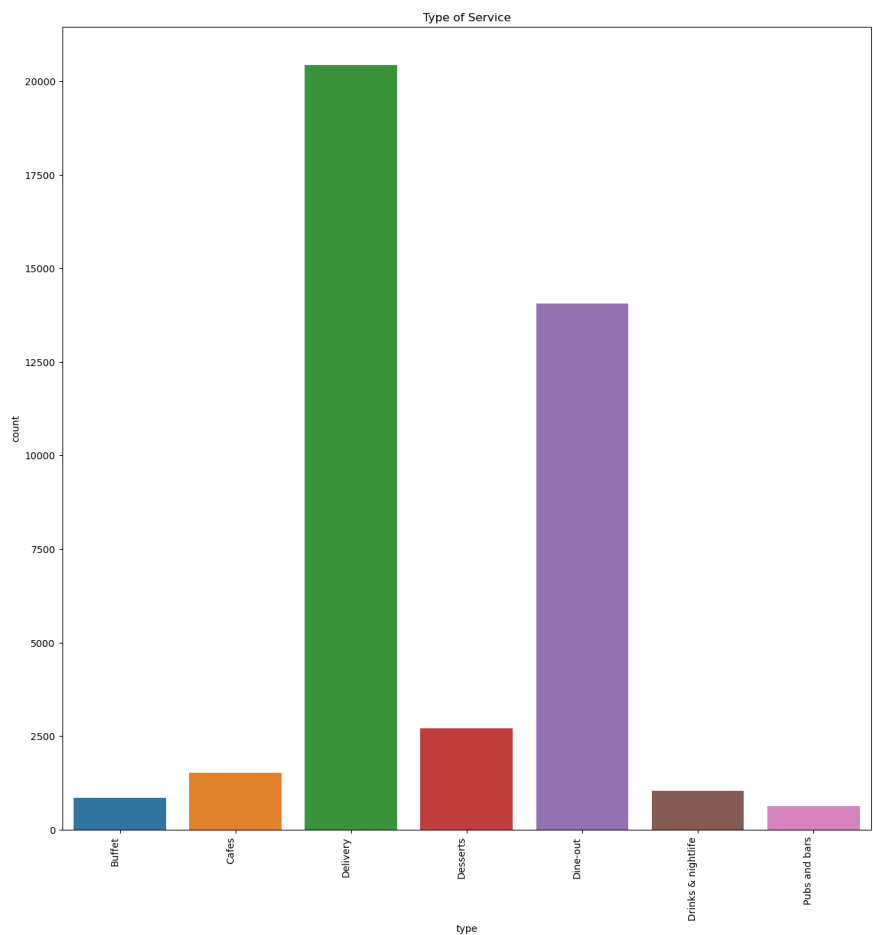
7. Column Chart: Type Vs Rate

Type - Rating
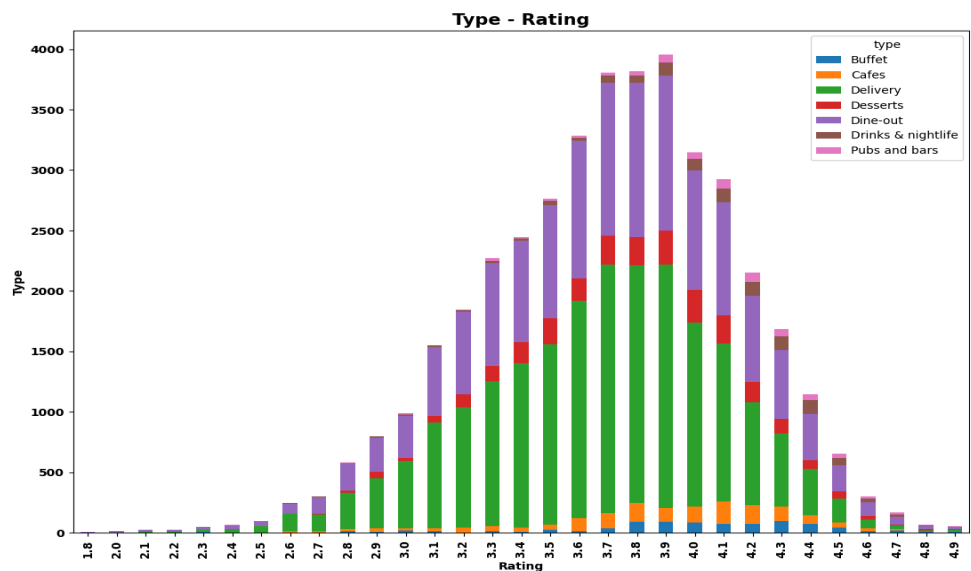
New restaurant hunters should remember to ensure that they offer all the services outlined in the aforementioned chart, with ratings falling between 3.3 to 4.2.

8. Bar Chart: Most famous restaurant chains in Bangalore

To improve our chances of success in the restaurant industry, it's advisable to study other restaurants to identify areas where they excel or where they may be lacking.

By observing successful establishments like CCD, as indicated in the chart below, new restaurant seekers can learn from their business strategies and adopt similar approaches to thrive in the market.

Most famous restaurant chains in Bangalore
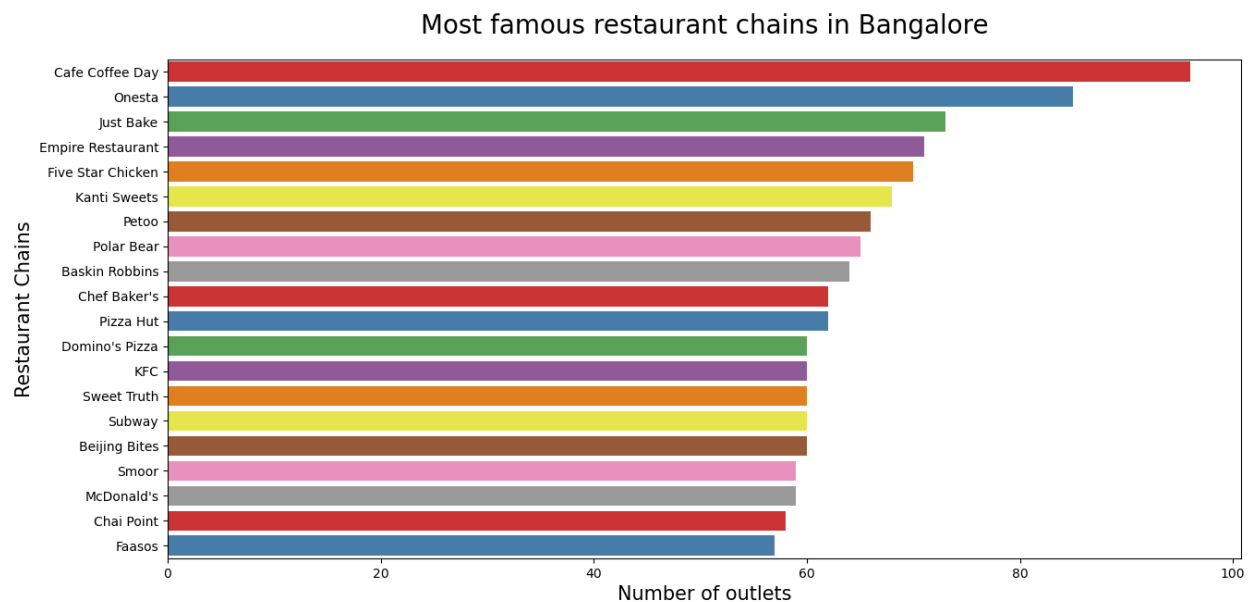
Data Predictions Models:

1. Lineear Regression: Testing R2: 27%

Most Important features:

rest_type - 0.002379

votes   - 0.000148

location -   0.000123

cuisines - 0.000054

menu_item -   0.000007

2. Decision Tree: Testing R2: 86%

Most Important features:

votes - 0.556168

cuisines - 0.139283

location - 0.112614

cost - 0.091277

rest_type - 0.065752

3. Random Forest Regression: Testing R2: 88%

Most Important features:

votes - 0.567011

cuisines - 0.134819

location - 0.103588

cost - 0.085976

rest_type - 0.068879

4. Extra Tree Regression: Testing R2: 94%

Most Important features:

votes - 0.346757

book_table - 0.172307

cuisines - 0.130355

location - 0.116156

cost - 0.115770

Recommendation Engine:

Focusing on a recommendation engine for a new restaurant can offer numerous advantages and opportunities for success:

1. **Personalized Customer Experience**: A recommendation engine can tailor suggestions based on individual preferences, enhancing the dining experience, and increasing customer satisfaction.

2. **Increased Engagement**: By providing personalized recommendations, restaurants can engage customers more effectively, leading to repeat visits and higher customer loyalty.

3. **Enhanced Marketing Efforts**: A recommendation engine enables targeted marketing campaigns, reaching customers with specific preferences and increasing the effectiveness of promotional efforts.

4. **Improved Decision-Making**: Restaurant owners can use recommendation engine data to make informed decisions about menu offerings, pricing strategies, and promotional activities.

5. **Competitive Edge**: Implementing a recommendation engine sets a restaurant apart from competitors, demonstrating a commitment to innovation and customer service.

6. **Data-Driven Insights**: By analyzing user interactions with the recommendation engine, restaurants gain valuable insights into customer preferences, helping them refine their offerings and improve business operations.

7. **Optimized Menu Design**: Understanding which dishes are most frequently recommended or ordered allows restaurants to optimize their menu design, highlighting popular items and removing less desirable ones.

8. **Efficient Resource Allocation**: Restaurants can use recommendation engine data to forecast demand, manage inventory, and allocate resources more efficiently, reducing waste and maximizing profitability.

9. **Cross-Selling and Upselling**: Recommendation engines can suggest complementary items or upgrades, increasing the average order value and generating additional revenue for the restaurant.

10. **Adaptability and Scalability**: Recommendation engines can evolve over time, incorporating feedback and adjusting recommendations based on changing customer preferences and market trends. This adaptability allows restaurants to remain relevant and competitive in a dynamic industry landscape.

I used below logic for my problem.

Term Frequency-Inverse Document Frequency

Term Frequency-Inverse Document Frequency (TF-IDF) vectors for each document. This will give you a matrix where each column represents a word in the overview vocabulary (all the words that appear in at least one document) and each column represents a restaurant, as before.

TF-IDF is the statistical method of evaluating the significance of a word in a given document.

TF — Term frequency(tf) refers to how many times a given term appears in a document.

IDF — Inverse document frequency(idf) measures the weight of the word in the document, i.e if the word is common or rare in the entire document. The TF-IDF intuition follows that the terms that appear frequently in a document are less important than terms that rarely appear. Fortunately, scikit-learn gives you a built-in TfIdfVectorizer class that produces the TF-IDF matrix quite easily.

From my dataset see the real example as per below.

->recommend('Pai Vihar')

TOP 6 RESTAURANTS LIKE Pai Vihar WITH SIMILAR REVIEWS:

	cuisines	Mean Rating	cost
Samosa Singh	North Indian, Mithai	3.60	250.0
Prasiddhi Food Corner	Fast Food, North Indian, South Indian	3.45	200.0
Shrusti Coffee	Cafe, South Indian	3.45	150.0
Shanthi Sagar	South Indian	3.44	250.0
Mayura Sagar	Chinese, North Indian, South Indian	3.32	250.0
Container Coffee	South Indian	3.11	200.0

Sentimental Analysis:

Sentiment analysis is valuable for new restaurant seekers as it allows them to gauge customer feedback and opinions effectively. By analyzing sentiment expressed in reviews and social media posts, restaurant owners can understand customer satisfaction levels, identify areas for improvement, and tailor their offerings to meet customer preferences. This insight helps new restaurants build a positive reputation, attract more customers, and ultimately succeed in a competitive market.
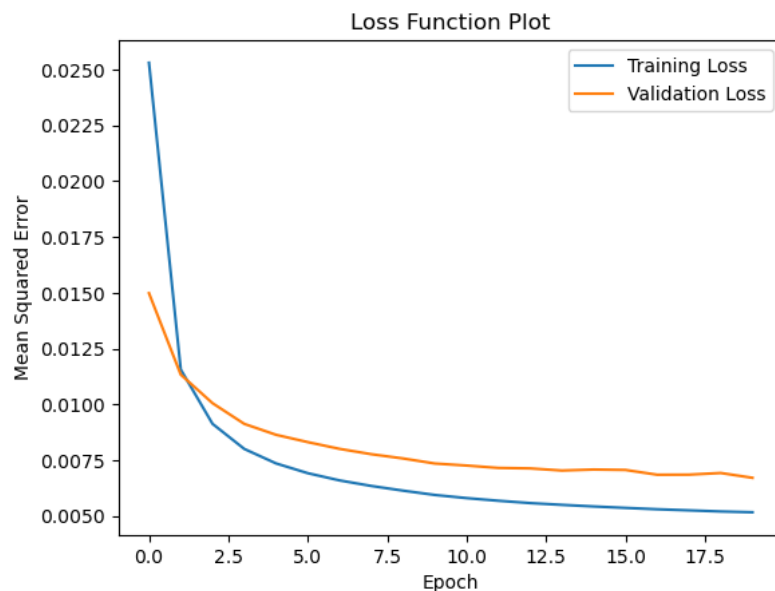
This graph appears to be a loss function plot typically used in machine learning to visualize the training and validation loss over the course of training a model. The x-axis represents the training epochs or iterations, while the y-axis shows the loss values.

There are two curves plotted:

1. The blue curve represents the training loss, which indicates the model's error or loss on the training data during the learning process.

2. The orange curve represents the validation loss, which shows the model's performance on a held-out validation dataset that was not used for training.

As the training progresses, we can see that both the training loss and validation loss decrease, indicating that the model is improving its performance on both the training data and the unseen validation data. However, after a certain number of epochs (around 10-12 based on the plot), the validation loss starts to increase slightly, while the training loss continues to decrease. This behavior is known as overfitting, where the model starts to memorize the training data too well, but its performance on unseen data (validation set) degrades.

Conclusion:

a new restaurant seeker should keep the following things in mind:

Service: Many restaurants seem to have high ratings for service, indicating that attentive and friendly staff can greatly enhance the dining experience. Good service appears to be a strong point for several establishments.

Ambiance: Ambiance, which likely includes factors like noise level, atmosphere, and overall vibe, is also highly rated across most restaurants. Creating an enjoyable ambiance is crucial for providing a pleasant dining experience.

Food: While ambiance and service are important, the quality of food should still be a top priority. The heatmap shows varying ratings for food across different restaurants, so researching and prioritizing establishments known for their culinary excellence is advisable.

Value: Some restaurants appear to offer better value or cost-effectiveness compared to others. For diners on a budget or seeking the best bang for their buck, considering the value aspect can be beneficial.

Thank you!

GitHub Link: https://github.com/adibhatt1997/BengaluruEatsInsights.git