# MACHINE LEARNING IN BUSINESS

**MIS710 – Assignment 2 – Part B: Business Report**

# Enhancing Restaurant Performance on FoodieBay: Data-Driven Insights and Predictive Models


**Class Group:** Wednesday (15:00 – 16:30)

**Tutor:** XXX

**Students:** Cam Ha Nguyen

DEAKIN
BUSINESS
SCHOOL

## Table of Contents

# Executive Summary

**Problem:** FoodieBay needed insights into factors influencing restaurant ratings to improve user satisfaction and platform credibility.

**Approach:** FoodsAnalytics employed data analysis and machine learning techniques to predict restaurant ratings, providing actionable recommendations for improvements.

**Findings:**

1. Offering table booking improves ratings.
2. Online ordering availability has minimal impact.
3. Review rankings matter more than average cost.
4. Different restaurant types have varying ratings and costs.
5. City-wise, ratings remain relatively consistent.

**Recommendations:** FoodieBay should employ the Random Forest Regressor model for rating predictions and use data-driven insights to propose improvements to partner restaurants. This approach can boost user satisfaction, strengthen partnerships, and potentially increase engagement and revenue.

## 1. Introduction

In the world of data-driven decisions, the FoodsAnalytics and FoodieBay partnership signifies transformative potential. Guiding our collaboration is the Business Analysis Core Concept Model (BACCM), a framework essential to grasp its fundamental elements.

**Need:** FoodieBay's Business Development team seeks insights into the factors influencing restaurant ratings to elevate the platform's restaurant performance.

**Value:** Enhanced understanding of ratings directly impacts user satisfaction and platform credibility, aiding decision-making for both FoodieBay and FoodsAnalytics.

**Stakeholders:**

- Primary:
    - FoodieBay: aiming to improve partner restaurant performance and user experience.
    - FoodsAnalytics: seeking data-driven insights to offer valuable consulting services.
    - Users: expecting reliable restaurant ratings and experiences.
    - Partner Restaurants: interested in factors affecting their ratings.
- Secondary:
    - Regulatory Authorities: monitoring data privacy and compliance.
    - Competitors: may be affected by innovations and improvements made by FoodieBay.

**Solution:** FoodsAnalytics employs data analysis and machine learning to derive strategic recommendations for improving restaurant ratings. These recommendations benefit FoodieBay and its partner restaurants, enhancing user satisfaction.

**Change:** This initiative transforms data analysis and decision-making practices, potentially influencing user behavior, such as increased trust in ratings and greater platform engagement.

**Context:** Within the dynamic restaurant aggregator industry, this analysis leverages a dataset from FoodieBay, focusing on Indian partner restaurants, and capitalizes on the collaborative expertise of FoodsAnalytics and FoodieBay.

## 2. Insights from Exploratory Data Analysis (EDA)

### 2.1. Data Quality

The FoodieBay dataset initially contained 40,130 rows and 17 columns, covering various aspects of restaurant-related information. Key data cleansing steps were taken to prepare the dataset for analysis. Columns deemed irrelevant were removed, missing data was addressed, and no duplicate entries were found. Outliers were retained as they were relevant to user-generated metrics. Feature engineering introduced two columns to enhance data insights. The final dataset for analysis and model development consisted of 30,699 rows and 14 columns.

## 2.2. Investigation into Essential Questions

### 2.2.1. Question 1: Number of restaurants offering table booking and its impact on ratings

**Distribution of Table Booking Availability:**

Figure 1 reveals that a substantial majority of the restaurants in the dataset (87.12%) do not offer table booking services, while the remaining 12.88% do provide this feature.

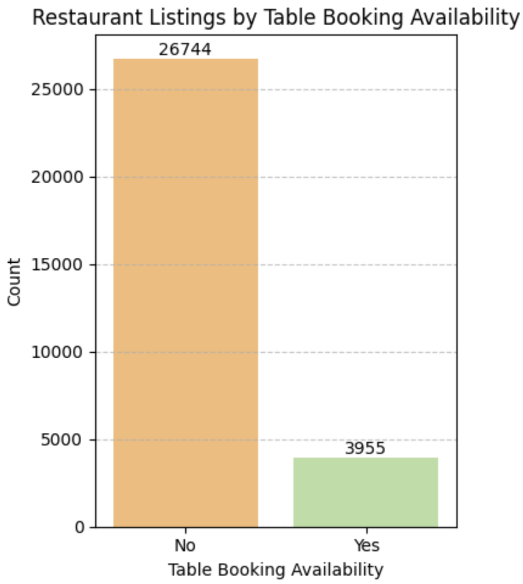


*Figure 1 - Restaurant listings by table booking availability*

**Impact on Restaurant Ratings:**

Figure 2 demonstrates that restaurants offering table booking tend to exhibit:

- Reduced rating variability, with a narrower rating range (approximately 2.2 to 4.9) compared to those without table booking (1.8 to 4.9).
- Higher mean and median ratings, indicating better average ratings.
- A more consistent rating range, as the middle 50% of restaurant ratings falls between 4.0 and 4.3, while for restaurants without table booking, the range is wider at approximately 3.3 to 3.9.

These findings suggest that table booking availability may significantly enhance user satisfaction and trust in restaurant ratings on the FoodieBay platform.

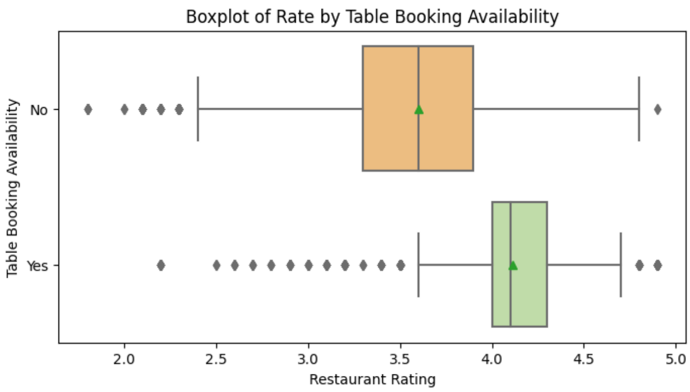


*Figure 2 - Restaurant rating distribution by table booking availability*

### 2.2.2. Question 2: Number of restaurants offering online ordering and its impact on ratings

**Distribution of Table Booking Availability:**

Approximately 70.22% of restaurants offer online ordering services, while 29.78% do not (Figure 3).

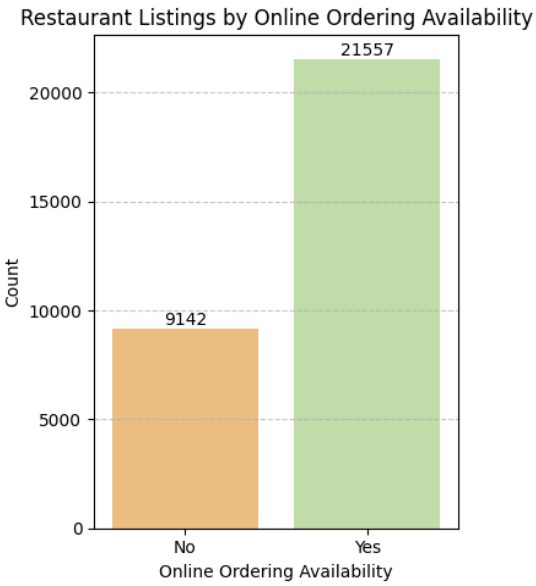


*Figure 3 - Restaurant listings by online ordering availability*

**Impact on Restaurant Ratings:**

Figure 4 indicates that the availability of online ordering services does not result in substantial differences in restaurant ratings on FoodieBay. Both categories - restaurants without online ordering services and those offering online ordering - exhibit:

- Similar mean and median ratings, with values centered around 3.6 to 3.7.
- Comparable rating ranges, as restaurants without online ordering have a range of about 3.3 to 3.9, while those with online ordering span approximately 3.4 to 4.0.

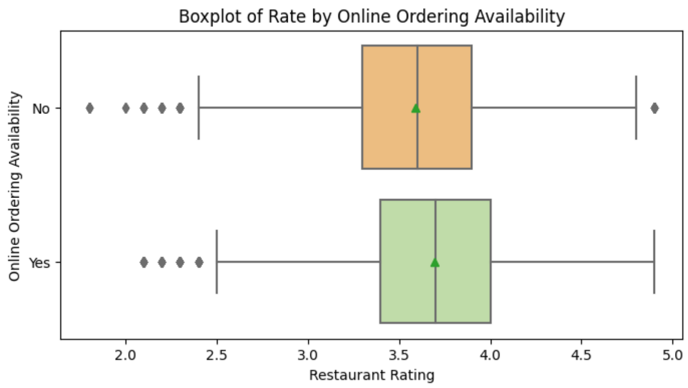


*Figure 4 - Restaurant rating distribution by online ordering availability*

### 2.2.3. Question 3: Effect of average cost for two and average customer review ranking on ratings

**Effect of Average Cost for Two on Restaurant Rating**

In Figure 5, the scatter plot reveals an R-squared value of 0.1479, indicating that about 14.79% of restaurant rating variance is linked to changes in average cost for two. With a slope coefficient of 0.0005, a unit increase in average cost for two (in INR) corresponds to a 0.0005 unit increase in restaurant rating. These findings suggest a statistically significant yet modest relationship, where small cost increases yield slight rating improvements, lacking substantial impact.

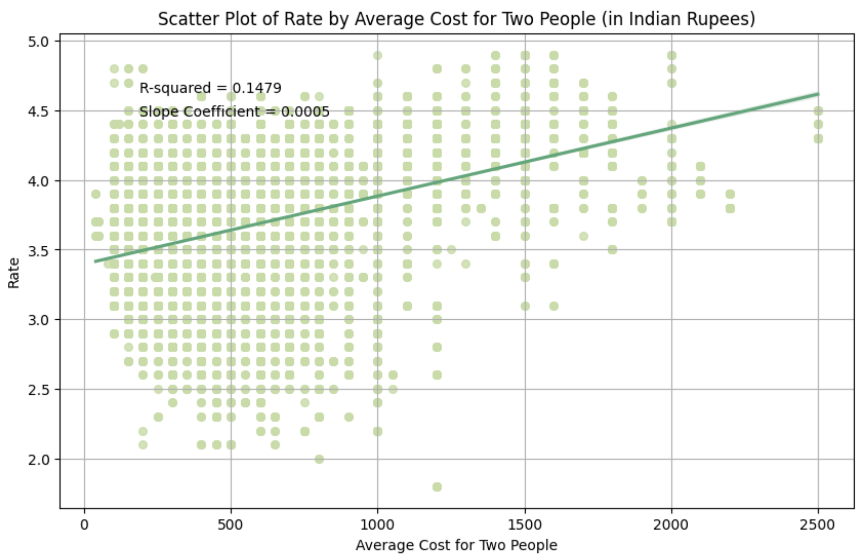


*Figure 5 - Scatter plot of restaurant rating by average cost for two people*

**Effect of Average Review Ranking on Restaurant Rating**

In Figure 6, the scatter plot shows an R-squared value of 0.2258, indicating that around 22.58% of restaurant rating variance relates to changes in average review ranking. With a slope coefficient of 0.2399, a unit increase in average review ranking yields a 0.2399 unit increase in restaurant rating. This underscores the greater influence of average review ranking on ratings compared to the average cost for two, leading to more significant rating enhancements.

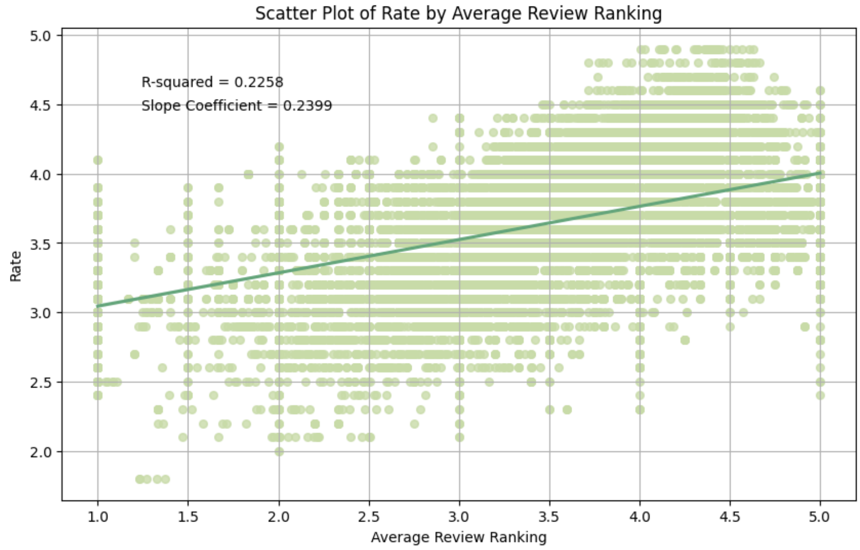


*Figure 6 - Scatter plot of restaurant rating by average review ranking*

### 2.2.4. Question 4: Variation in ratings and average cost for two among restaurant types

An overwhelming majority, comprising 90.28% of restaurants on FoodieBay, fall into the 'Dine-out' and 'Delivery' categories, with a smaller percentage of 6.35% representing cafes and dessert establishments. The remaining 3.37% consists of 'Buffet,' 'Pubs and Bars,' and 'Drinks & Nightlife' services (Figure 7)

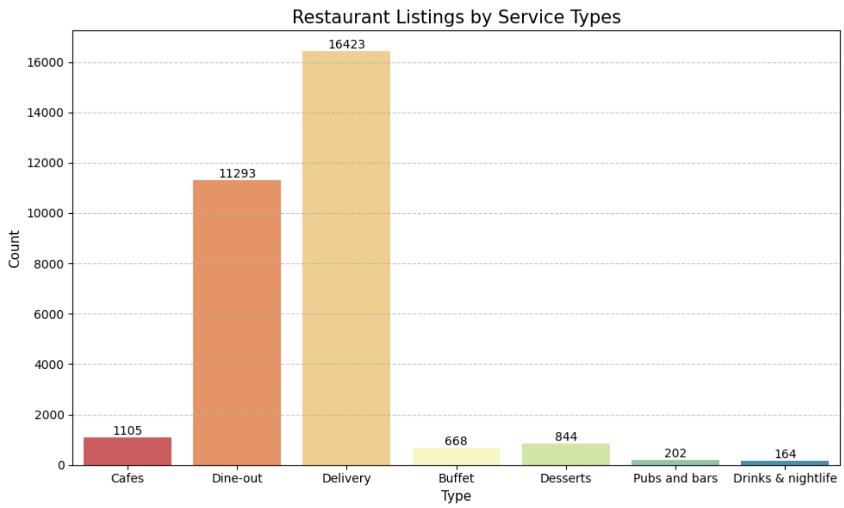


*Figure 7 - Restaurant listings by service types*

**Analysis of Variation in Ratings by Restaurant Type:**

Ratings vary by restaurant type (Figure 8):

- 'Buffet', 'Drinks & Nightlife', and 'Pubs and Bars' lead with the highest mean and median ratings, signifying consistent exceptional dining experiences due to their variety and ambiance focus.

- 'Cafes' and 'Desserts' maintain strong average ratings of 3.74 and 3.83, respectively, exceeding the all-restaurant average of 3.66.
- 'Dine-out' and 'Delivery' exhibit lower averages and the widest rating ranges, from 1.80 to 4.90 for 'Dine-out' and 2.00 to 4.90 for 'Delivery,' reflecting diverse customer experiences in these categories

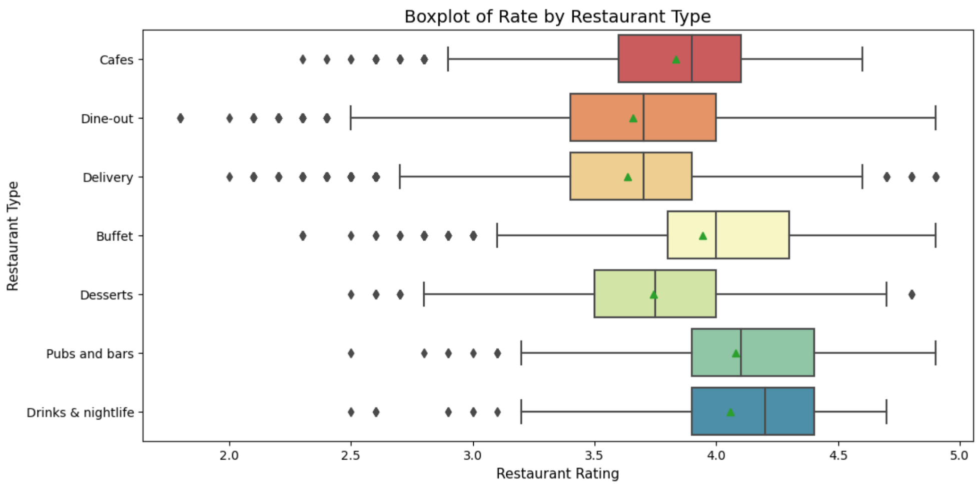


*Figure 8 - Distribution of restaurant rating by restaurant type*

**Analysis of Variation in Average Cost for Two by Restaurant Type**

From Figure 9, we can observe that:

- 'Drinks & Nightlife' (INR 1294.82) and 'Pubs and Bars' (INR 1277.48) feature the highest mean costs, highlighting premium ambiance and pricing.
- Buffet restaurants mainly range from INR 800 to INR 1500, as indicated by the interquartile range.
- 'Dine-out' and 'Delivery' restaurants fall in the mid-range, with average costs of INR 559.13 and INR 503.07 for two, respectively.
- Cafes (INR 617.74) and Desserts (INR 488.15) offer more affordable dining options, with an emphasis on cost-effectiveness and enjoyable coffee and desserts.

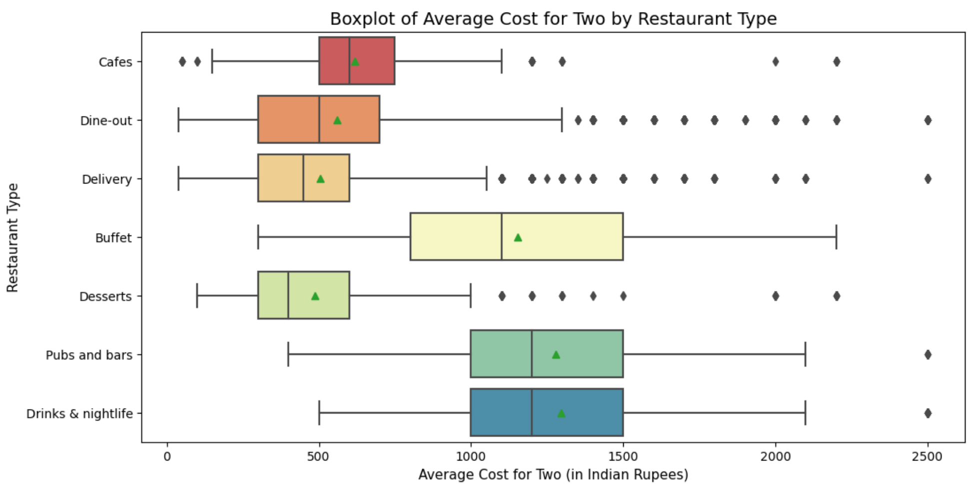

Figure 9 - Distribution of average cost for two by restaurant type

### 2.2.5. Question 5: Additional insights about restaurants, cities, and cuisines

**Restaurant Listings and Ratings Across Indian Cities**

Figure 10 highlights the leading locations in restaurant listings on FoodieBay, with BTM, Koramangala 7th Block, Koramangala 4th Block, Koramangala 5th Block, and Koramangala 6th Block emerging as the top five. Remarkably, BTM boasts over four times the number of restaurants compared to New BEL Road, which has the lowest restaurant listing count.

However, these cities exhibit relatively similar restaurant ratings, with median ratings ranging from 3.5 to 3.8 (Figure 11).

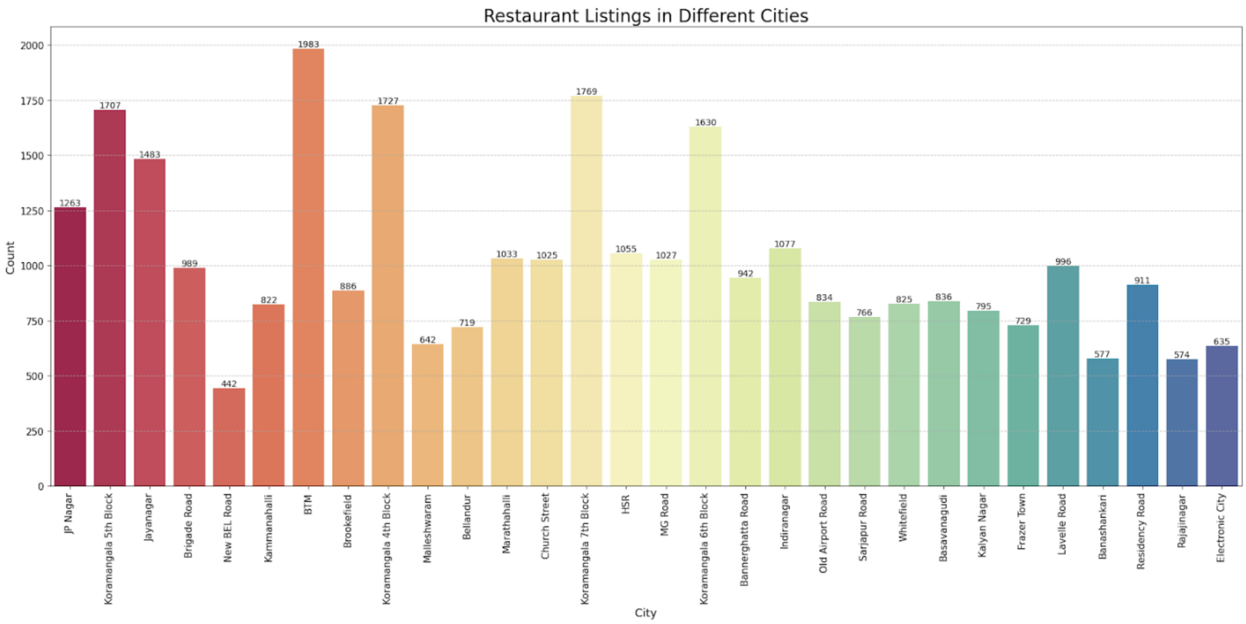


Figure 10 - Restaurant listings in different cities

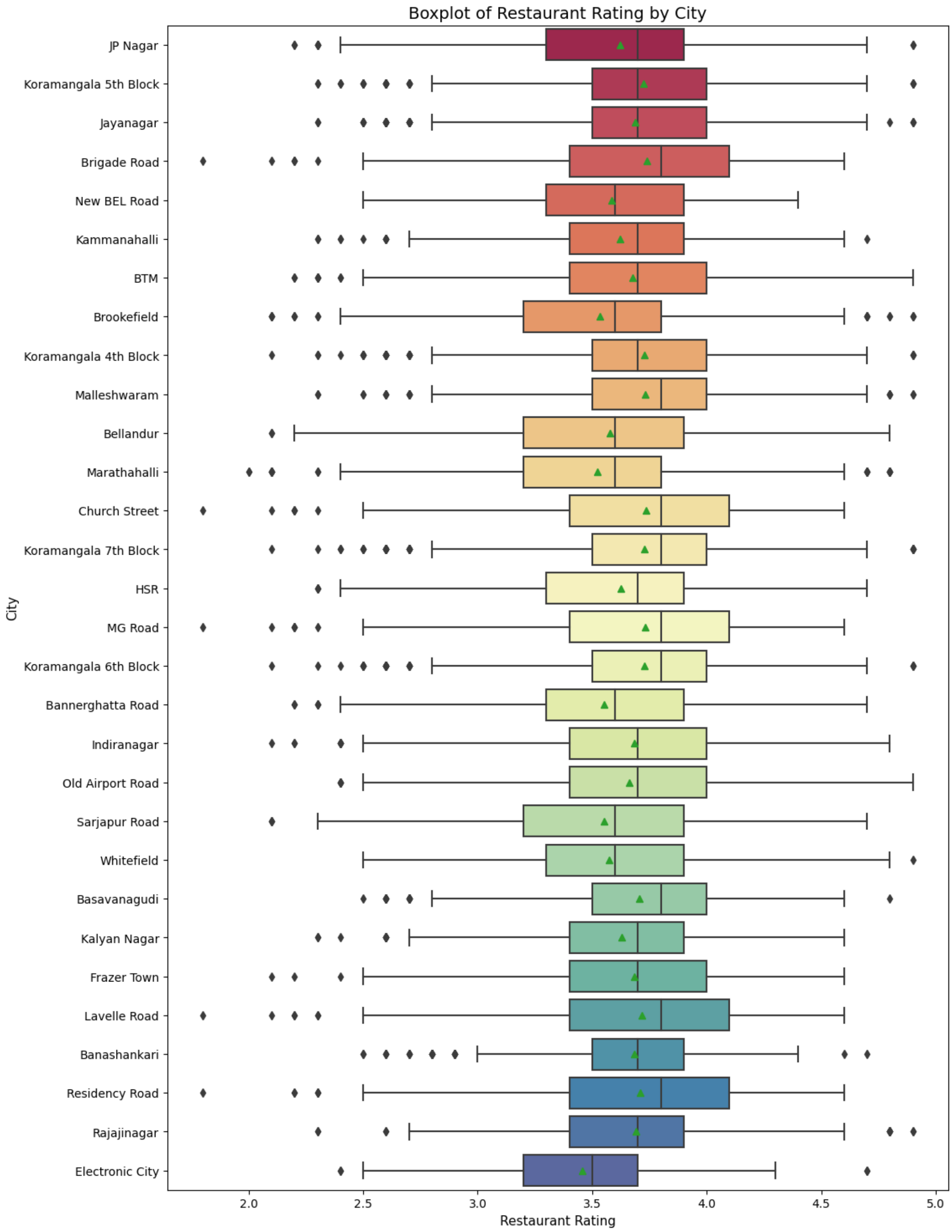


*Figure 11 - Distribution of restaurant rating by Indian cities*

**Cuisine Preferences and Top Dishes on FoodieBay**

Within our dataset of 30,699 restaurants, North Indian cuisine leads with 14,502 establishments, followed by Chinese, South Indian, Fast Food, and Biryani cuisines (Figure 12). In terms of preferred dishes, Pasta, Burgers, and Biryani stand out as top choices among FoodieBay users (Figure 13). These insights can inform menu offerings and marketing strategies for restaurants on the platform.

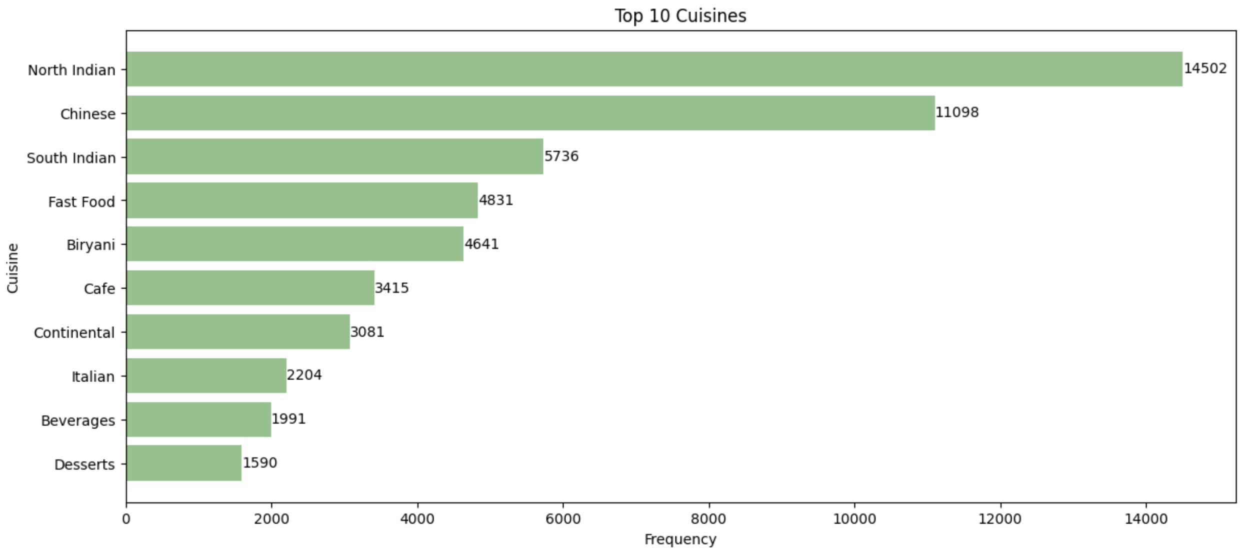


*Figure 12 - Top 10 cuisines on FoodieBay*

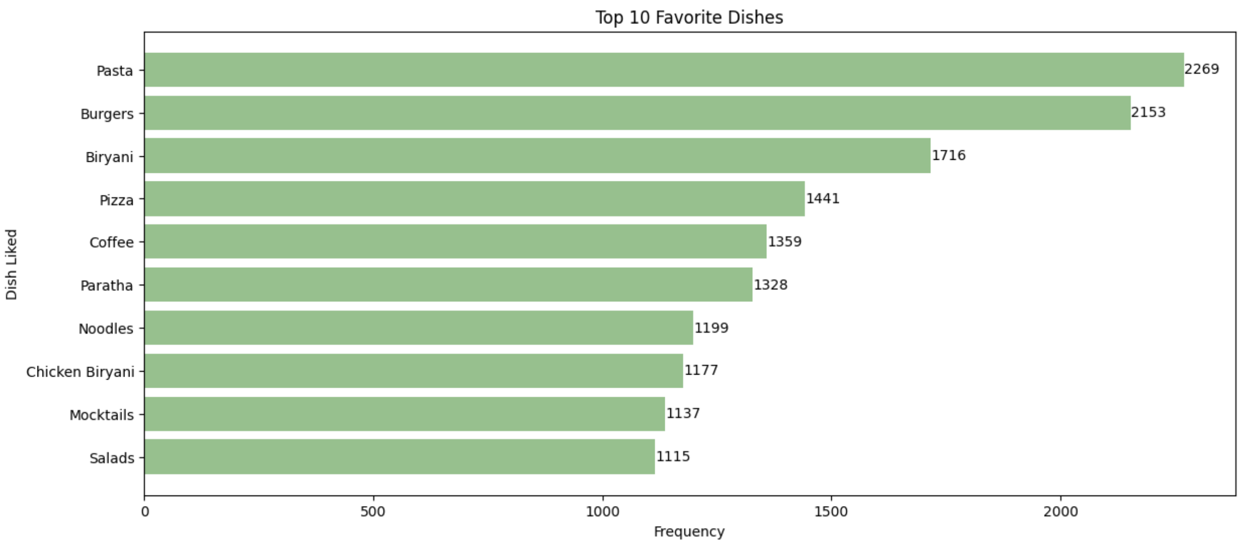


*Figure 13 - Top 10 favorite dishes in FoodieBay restaurants*

## 3. Proposed Machine Learning Solution

**Selected Machine Learning Model**

We recommend the Random Forest Regressor as it aligns perfectly with the business case and objectives outlined in the BACCM model. It provides a robust solution for predicting continuous numerical value like restaurant ratings, essential for enhancing user satisfaction and platform credibility on FoodieBay.

The model uses several predictors, including the number of dishes liked, average review ranking, votes, table booking availability, and average cost for two, to make these predictions. These factors are essential components that contribute to the model's ability to estimate restaurant ratings accurately.

**Model Performance and Discussion**

The Random Forest Regressor exhibits strong predictive performance with a low Root Mean Square Error (RMSE) of 0.149 (±0.003), indicating close alignment between predicted and actual ratings. The R-squared ($R^2$) value of 0.869 (±0.005) implies that approximately 86.9% of the variability in restaurant ratings can be effectively explained by the selected predictors in the model.

Besides strong predictive capability, the Random Forest Regressor is also less prone to overfitting than the tested Decision Tree. However, its complexity may pose challenges in explaining it to stakeholders. Additionally, for larger datasets in the future, training and utilizing Random Forest models may require significant computational resources.

## 6. Recommendations and Conclusions

**Recommendations for Business Applications**

- Enhanced User Experience: FoodieBay should leverage the predictive power of the Random Forest Regressor model to offer personalized restaurant recommendations, improving user satisfaction.
- Support for Partner Restaurants: Offer data-driven insights (e.g., highlighting the benefits of table booking for improved ratings), aiding them in enhancing and sustaining high ratings.
- Quality Control: Use the model to proactively identify and address issues in restaurants, ensuring consistent service quality.

**Benefits to Stakeholders and Value Proposition**

- FoodieBay: FoodieBay stands to improve user satisfaction, boost platform credibility, and strengthen relationships with partner restaurants. This translates to increased user engagement and, potentially, higher revenues.
- Users: FoodieBay users benefit from more tailored restaurant recommendations and a more reliable rating system, enhancing their dining experiences.
- Partner Restaurants: Partner restaurants can leverage insights to enhance their offerings and improve their ratings, potentially attracting more customers.

**Recommendations for Further Improvements**

- Data Enrichment: Continuously update data for accuracy and relevance.
- Feedback Loops: Establish mechanisms for dynamic data and collaboration with partner restaurants.
- Advanced Models: Explore advanced machine learning techniques for improved accuracy and insights.

In conclusion, data analysis and machine learning can elevate FoodieBay's user experiences, empower partner restaurants, and strengthen its position in the industry. These data-driven insights can lead to more informed strategic decisions, potentially improving user satisfaction and platform credibility.