

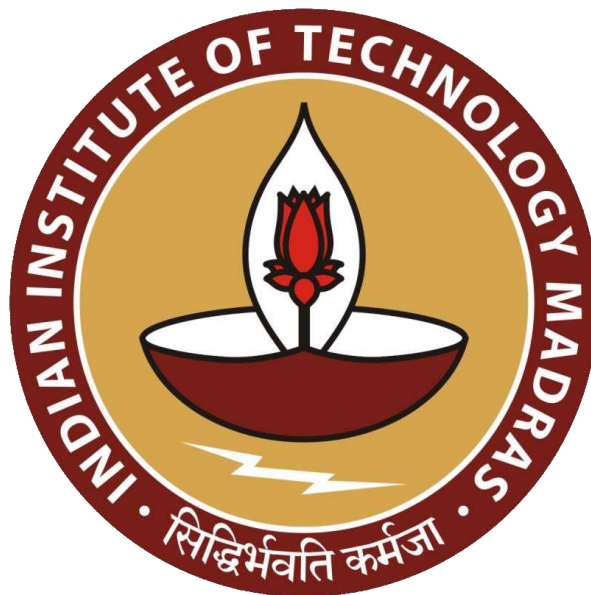
Enhancing Gym Business Sustainability: Overcoming Challenges and Boosting Customer Engagement

Final Report for the BDM Capstone Project

Submitted by

Name: Harishankar Nagar

Roll Number: 21f2000681



IITM Online BS Degree Program,

Indian Institute of Technology, Madras, Chennai

Tamil Nadu, India, 600036

Contents

1	Executive Summary	3
2	Detailed Explanation of Analysis Process/Method	4
2.1	Segmenting the Data	4
2.2	Removing Missing Values	5
2.3	Eliminating Redundant Data	5
2.4	Validating the Data.....	6
2.5	Creating Necessary Subsets for Analysis.....	6
2.6	Data Visualisation	6
2.7	Hypothesis Testing	7
3	Results and Findings	7
3.1	Findings From Customer Data.....	7
3.2	Findings From Feedback Data	13
3.3	Insight From Expenditure Data.....	15
4	Interpretation of Results and Recommendation	16

1 Executive Summary

In this capstone project, I have engaged with Mr. Amit Chaudhary, a multi-business owner, who is facing challenges in his gym business “Muscle Factory Gym”. Amit Chaudhary’s business in the college setting is associated with general shops and bakery/canteen. He started a process of diversification of the business in another direction by opening a gym. The current issues for the gym are mainly due to the establishment of a new and much better gym inside the college compound. The intense competition has reduced customer base and revenues which are made worse by lack of experience in the gym sector and high dependency on college students.

To tackle these concerns, a data-driven approach was adopted. This approach entailed gathering and carefully refining three crucial sets of data:

- **Customer Data:** It includes the information on customers. Package type, package duration, amount paid, date of start, some demographics variables like gender, and age with membership id and mobile number.
- **Customer Feedback Data:** Identifying the pain points is important for understanding customer satisfaction levels as well as improvements required in services facilities and general customer experience.
- **Renewal Data:** The dataset comprises of data from the customers who have decided to renew their subscriptions. In order to extract the revenue Information, this will be employed.
- **Expenditure Data:** Expenditure data provides a panoramic view that supports effective expense control, appropriate budgeting, and efficient utilization of resources.

Through a thorough examination of this data, I was able to uncover valuable observations and important measurements. These included the number of new members, the monthly revenue generated, the overall revenue, and the characteristics of our customers. It is worth highlighting that the analysis revealed a significant drop in customer retention during the later months.

After conducting an in-depth analysis of customer data, revenue trends, and feedback, several key recommendations have been identified to address the challenges faced by the business.

2 Detailed Explanation of Analysis Process/Method

The collected datasets encompass valuable information; however, they are not devoid of flaws. Challenges arise from problems like absent values, repetitive information, and inaccuracies in data format. To obtain significant insights, it is crucial to commence a data transformation procedure to tackle these shortcomings and improve the overall dataset quality. The analysis process involved several key steps to transform raw data into meaningful insights. Here's a detailed breakdown of each step:

- Segmenting the Data
- Removing Missing Values
- Eliminating Redundant Data
- Validating the Data
- Creating Necessary Subsets for Analysis
- Data Visualization
- Hypothesis Testing

2.1 Segmenting the Data

The dataset called "Package_renewal" presented an exceptional challenge. It brought together details regarding subscription renewals across several months, leading to data entries that only had information for the specific month in which a customer renewed their subscription. In all other months, the dataset contained empty or blank values in different columns. Furthermore, the dataset included data about the payment amount made by customers during registration. This extra information posed a difficulty when analysing solely the data related to renewals. However, by segregating the data into separate monthly files, I have made the dataset more manageable and appropriate for analysis.

The feedback survey incorporates textual data, including opinions and suggestions, which lacks a defined format. To address this, I have distinguished this unstructured data from the well-organized information. This segmentation yields two distinct types of data: one suitable for visualization and another specifically geared towards understanding customer opinions. This strategic separation enhances the manageability of the feedback data, transforming it from a potentially messy dataset into a more engaging and insightful resource.

2.2 Removing Missing Values

After dividing the data into monthly files, I noticed the existence of absent values, which had been passed down from the original "Package_renewal" dataset. To guarantee the integrity and accuracy of the data, I decided to eliminate the missing values from the renewal files, as it was essential to possess thorough and dependable data to analyse subscription renewals.

About the "Customer Information" dataset, I identified the absence of values in crucial columns like "Age", "Address", "Amount", and "Package_type." Nevertheless, if I were to eliminate rows with missing values in this dataset, it would result in the loss of valuable information about the amount paid during registration and the specifics regarding the subscription start date. To confront this obstacle, I resolved to substitute the missing values in these columns with the term "unknown." This enabled me to retain the records and associated information, even in situations where specific details were unavailable.

After dividing the feedback data into segments, I observed that the text section had some empty spots. Nevertheless, since my main objective does not revolve around using these sections for visualizing or analysing statistics, the absence of information in those sections is not a worry. The purpose of these sections is simply to gain a better understanding of customer opinions.

2.3 Eliminating Redundant Data

After effectively dividing the information and addressing the missing values, the next obstacle was to eliminate repetitive data from the monthly files. These recurring entries contained details about the amount paid during both registration and renewal. Additionally, there was no obvious way to differentiate which rows corresponded to renewal transactions and which ones were associated with registration. To overcome this problem, I utilized the "Customer_Id" from the "Customer Information" dataset. This unique identifier enabled me to connect and cross-check the payments with specific customers. By doing so, I could confidently remove the payments made during registration, ensuring that only the relevant renewal transactions were kept for analysis. This approach not only streamlined the dataset but also improved its accuracy and suitability for analysing payment trends related to renewal, while still preserving the necessary context provided by the "Customer Information" dataset.

2.4 Validating the Data

After eliminating unnecessary information and using `Customer_Id` to distinguish, the subsequent vital stage was to guarantee the accuracy of the dataset by conducting extensive data validation. Throughout this procedure, I diligently examined the data and detected small inaccuracies, such as discrepancies in the date column. A particular problem that arose was the improper sequence of dates. For example, I noticed that "Dec-23" came before "January-23" in the data, which was incorrect. To rectify such errors related to dates, I implemented the essential corrective actions to establish a coherent and sequential order in the date column.

2.5 Creating Necessary Subsets for Analysis

After successfully verifying the information and resolving any issues related to accuracy, the subsequent phase in the analysis process entailed generating various valuable divisions from the dataset. These divisions played a crucial role in acquiring more profound understandings and comprehending diverse facets of the information.

- **Revenue Dataset:** Gathering data on income from fresh registrations, renewals, and overall revenue.
- **Monthly Present Contingency Table:** Assessing categorical variables such as Gender, Address, and Package_type to identify recurring patterns.
- **Attrition Rate:** Computing the rate at which customers disengage each month.
- **New Customer Count for Each Month:** Monitoring the number of new customers acquired to analyse trends.
- **Active Flag:** To monitor which customer is still active. It will be helpful in analysing the association of different variables to active status.

These divisions greatly facilitated comprehensive analysis and informed decision-making, taking into consideration revenue, customer demographics, attrition, and patterns of customer acquisition.

2.6 Data Visualisation

In the following phase, I created visual representations illustrating income information obtained from fresh sign-ups and renewals, as well as the distribution of active clientele every month. These representations were categorized by sex, package kind, and location to detect noticeable trends and patterns. The purpose of plotting income information based on various

demographic factors like sex, package type, and location is to reveal any noticeable patterns or trends that could offer valuable insights. This approach enables a detailed comprehension of how these factors could influence income generation and customer activation.

Moreover, contingency tables were established to investigate the connections between categorical factors and the status of customer activation. This statistical method assists in identifying potential dependencies or independence between factors, shedding light on any significant relationships that may exist.

2.7 Hypothesis Testing

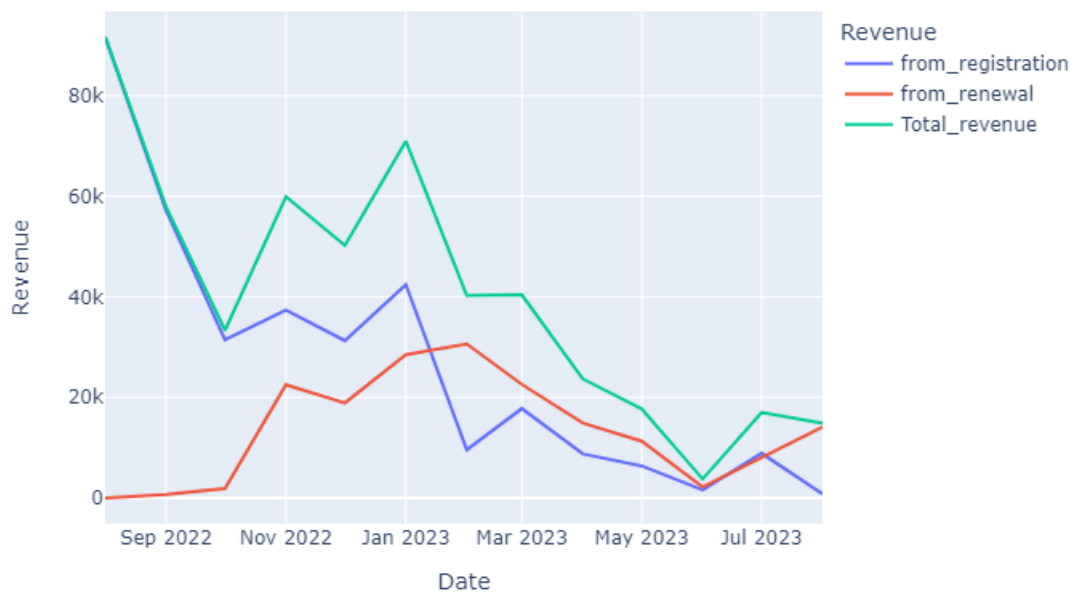
To evaluate the connection between being active and demographic factors like gender, location, and package type, we initially opted for the chi-square test. However, the data did not meet the requirement of the chi-square test, which mandates that each cell count must be 5 or higher. Consequently, we adopted an alternative statistical approach and utilized Fisher's exact test. Fisher's exact test is particularly suitable for cases where the dataset is small and cell counts might be below the threshold of 5. This test offers a robust analysis, enabling reliable evaluations of associations in scenarios where the chi-square test may not be applicable. By acknowledging and addressing the specific characteristics of the dataset, the use of Fisher's exact test guarantees the validity of the statistical analysis, providing accurate insights into the relationships between being active and the demographic factors under consideration.

3 Results and Findings

3.1 Findings From Customer Data

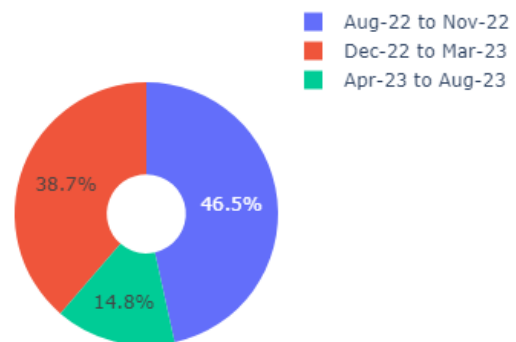
The line chart (Figure 1) depicting the gym's revenue trends reveals a notable decline post-January, with a pronounced dip in March. Peaks in revenue coincide with significant events: the gym's opening generated heightened interest due to a lack of nearby fitness facilities, while a January spike aligns with the surge in New Year's resolutions. This analysis underscores the impact of external factors on revenue fluctuations and offers valuable insights for strategic decision-making. Recognizing the contextual drivers behind these patterns enables the formulation of targeted interventions during downturns and the optimization of opportunities presented during peak periods.

Figure 1 Revenue Plot



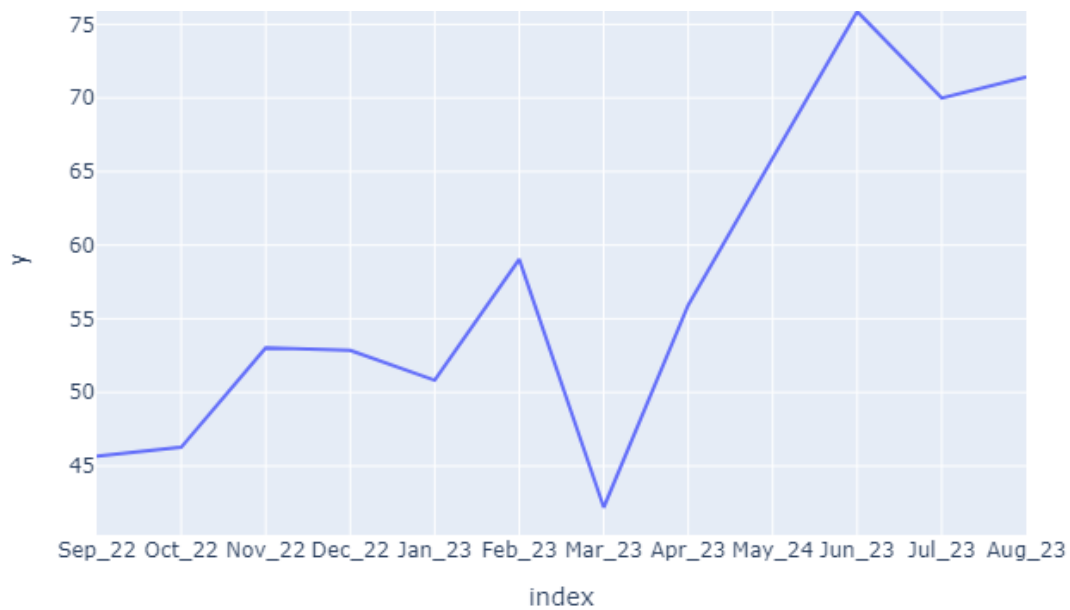
The gym's revenue collection pattern across distinct time periods (Figure 2) sheds light on its financial dynamics. The initial four months from Aug-22 to Nov-22 marked a robust start, accounting for 46.5% of the total revenue at ₹2,42,880.0. Subsequently, from Dec-22 to Mar-23, revenue remained substantial, contributing 38.7% at ₹2,01,800.0. However, a noticeable decline was observed in the last period, Apr-23 to Aug-23, where revenue dropped to ₹77,100.0, comprising 14.8% of the total.

Figure 2 Revenue of Different Time Periods



These attrition rates (Figure 3) represent the percentage of customers who discontinued their memberships or became inactive during each respective month. Analysing these rates over time can provide insights into trends. The noticeable uptick in attrition rates post-March 2023, with a significant spike exceeding 75% in June, is a critical observation signalling a substantial increase in customer churn during this period. Such a sharp rise in attrition rates suggests a pivotal turning point that demands a closer examination of underlying factors contributing to customer disengagement.

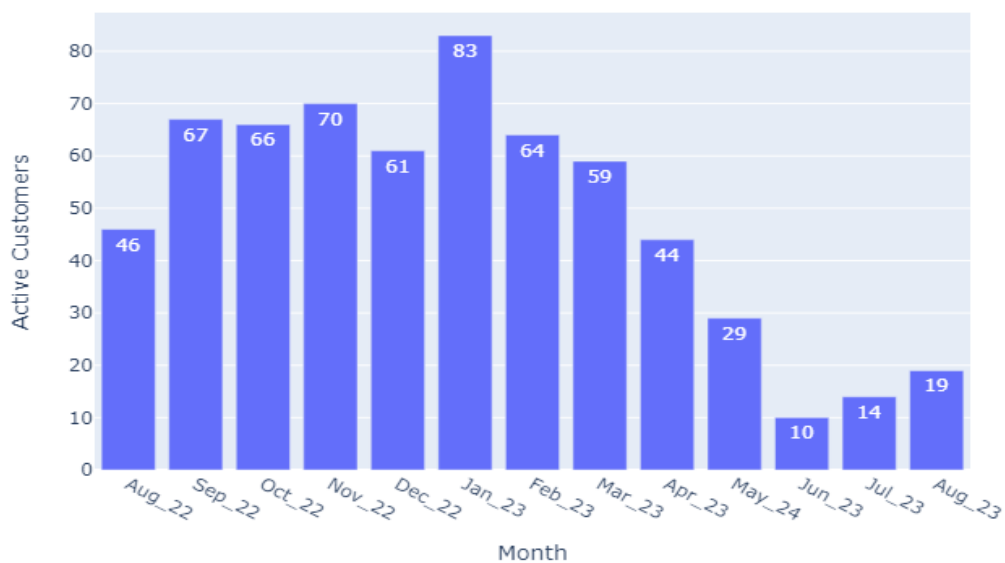
Figure 3 Attrition Rate for Each Month



Analysing the active number of customers based on various conditions, such as gender, location, and gym subscription package type, provides a deeper understanding of the factors influencing customer engagement. By plotting these variables, I gain insights into the gym's demographic dynamics and member preferences.

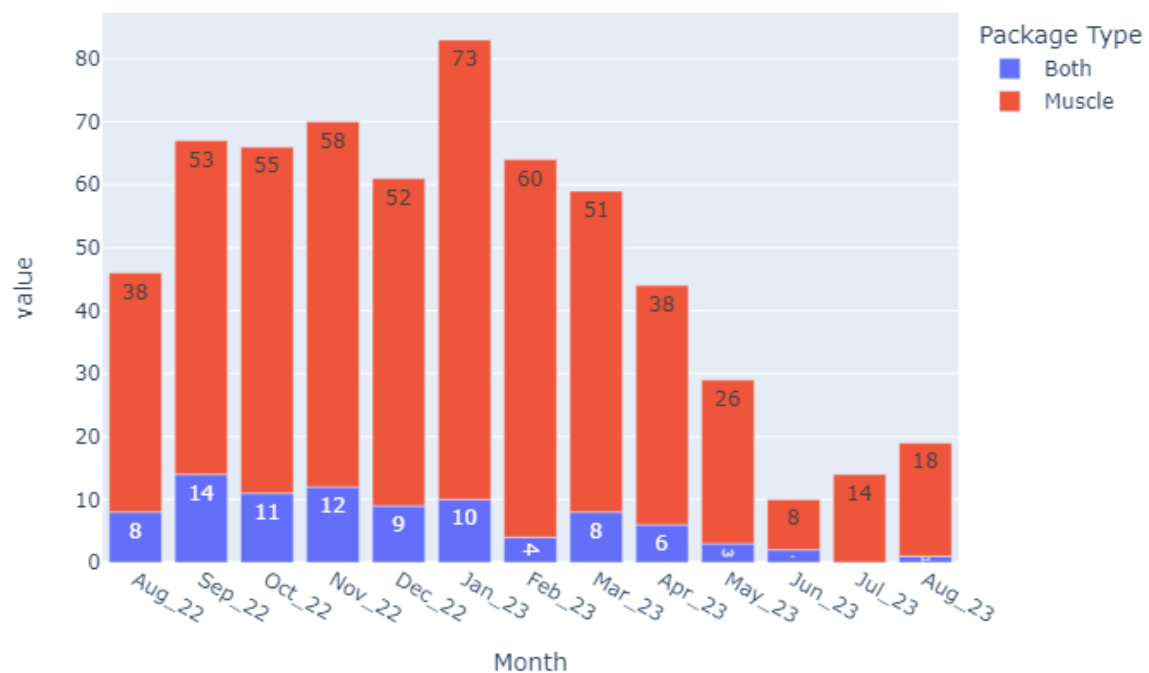
The bar plot of monthly active customers (Figure 4) gives the information of active customers in each month. providing a clear and easily interpretable snapshot of the gym's monthly performance.

Figure 4 Bar Plot of Monthly Active Customers



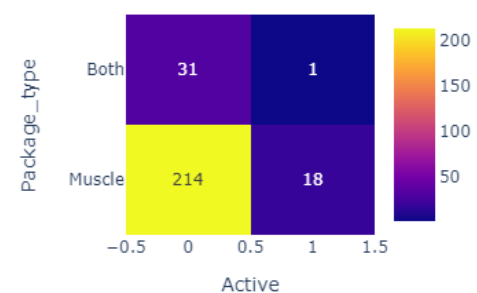
The stacked bar chart (Figure 5) that displays the number of customers who are actively engaged, categorized by the type of package they have chosen (Muscle and Both), offers a compelling visual insight into the preferences of customers. The clear contrast in the heights of the stacked bars vividly emphasizes that the Muscle package is significantly more popular than the Both package. The visual representation of customer distribution in these two categories unmistakably underlines the strong preference for the Muscle package among our active customers. This discovery not only sheds light on the current composition of our customer base but also presents an opportunity for targeted marketing or strategic adjustments to increase the appeal and adoption of the Both package, ultimately leading to a more balanced and diverse customer portfolio.

Figure 5 Stacked Bar Chart: Active Customers by Package Type



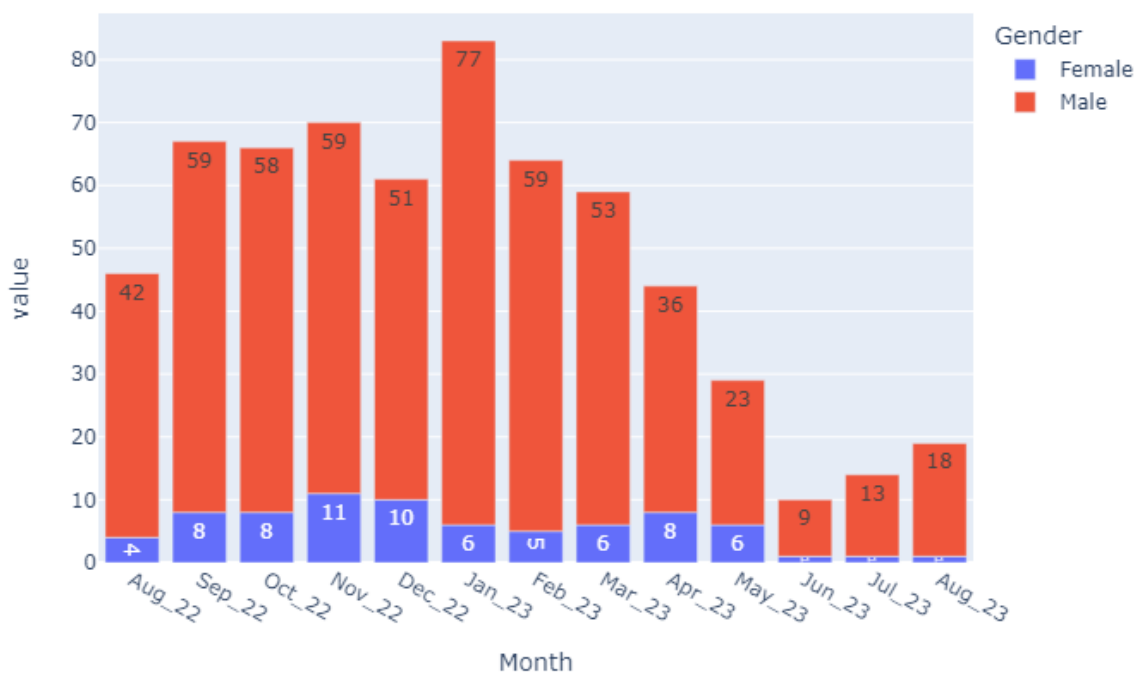
The Fisher's exact test was employed to examine the association between package type (Muscle and Both) and the active status of customers (1 for active, 0 for inactive). The obtained p-value of 0.485 suggests that this observed association is not statistically significant at a conventional significance level (e.g., 0.05). Therefore, based on the Fisher's exact test, there is insufficient evidence to conclude a significant association between package type and customer activity status.

Figure 6 Package Type and Customer Activity Contingency



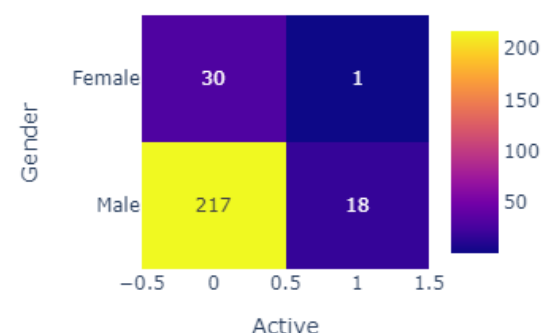
The stacked bar chart (Figure 7) displaying the breakdown of active customers by gender reveals a striking discrepancy between the number of active males and females. The graphical depiction emphasizes that the male population constitutes a substantially larger proportion of the engaged customer base, while the contribution from females seems relatively restricted. This observation prompts further investigation into the factors influencing gender-specific involvement and underscores the potential for focused efforts to bolster female participation.

Figure 7 Stacked Bar Chart: Active Customers by Gender



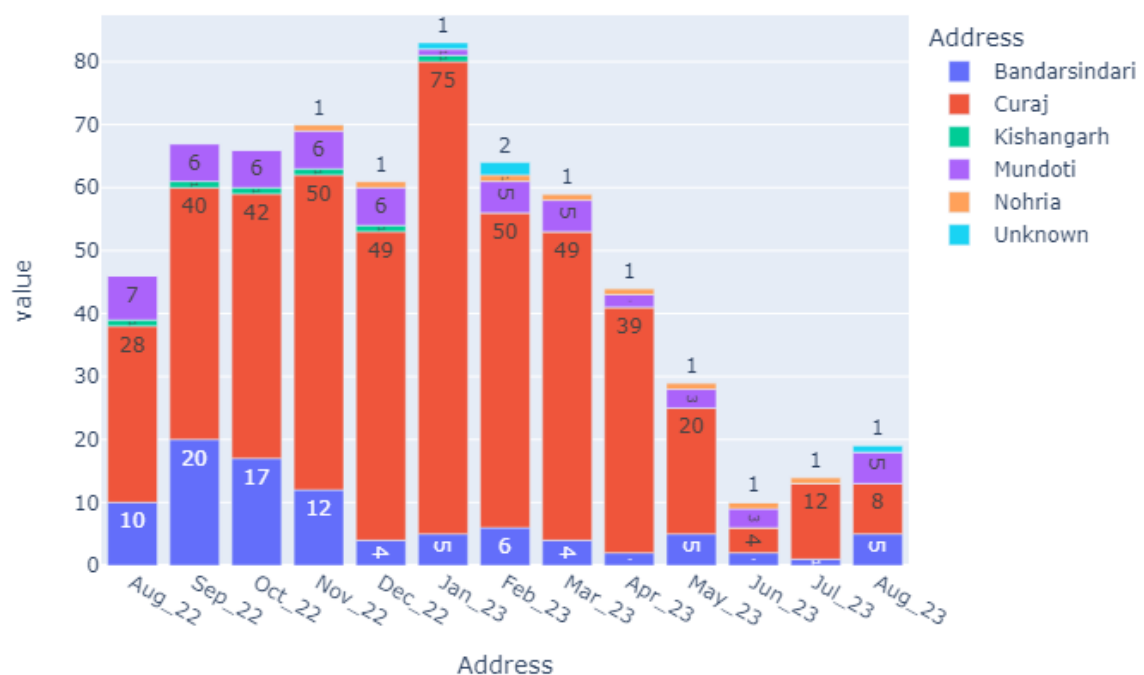
The Fisher's exact test was applied to assess the association between gender and the active status of customers. The obtained p-value of 0.708 suggests that there is no statistically significant association between gender and customer activity status at a conventional significance level (e.g., 0.05). Therefore, based on the Fisher's exact test, there is no compelling evidence to conclude that the likelihood of a customer being active is significantly different between genders. So, gender does not appear to be a significant factor influencing customer activity status.

Figure 8 Gender and Customer Activity Contingency



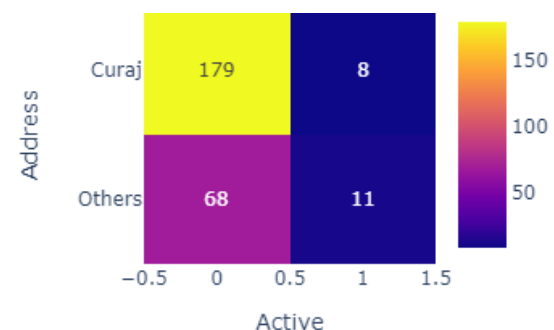
The customer activity chart for the location (address) (Figure 9) has revealed an intriguing revelation: a notable cluster of active customers originating from Curaj (Central University of Rajasthan). Astonishingly, the information indicates that the customer population from Curaj greatly exceeds the total count of customers from other areas. This finding emphasizes Curaj as the dominant customer base for the gym, emphasizing its crucial role in shaping the gym's overall monthly activity patterns.

Figure 9 Stacked Bar Chart: Active Customers by Address



In the next step, I wanted to assess whether the location of the customer is associated with their active status or not. However, the contingency table did not meet the requirements for conducting the chi-square and Fisher's exact tests. As a result, I decided to simplify the location column by categorizing it into two groups: "Curaj" and "Others" (representing all other places). The chi-square test was then applied to evaluate the association between customer activity status and location. This statistical test examined whether the observed distribution of active and inactive customers across "Curaj" and "Others" was significantly different from what would be expected by

Figure 10 Address and Customer Activity Contingency



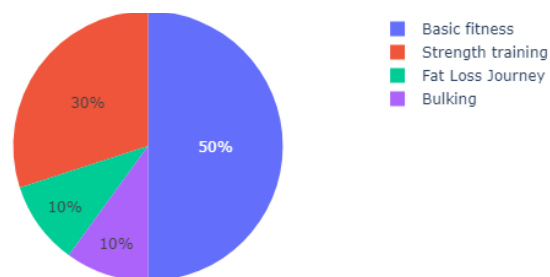
chance. The calculated chi-square statistic of 6.4045, with 1 degree of freedom, yielded a p-value of 0.0114. The p-value falling below the standard significance level (e.g., 0.05) indicates a statistically significant association between customer activity status and location. The findings imply that the likelihood of a customer being active is not uniform across locations. Specifically, the distinction between "Curaj" and "Others" is noteworthy, suggesting that location influences customer activity status.

3.2 Findings From Feedback Data

The feedback survey received a minimal response, as only 20 individuals out of around 250 gym members participated. This presents a challenge when attempting to draw comprehensive conclusions. The small number of respondents may not accurately represent the diverse range of perspectives and experiences among all the members. However, the insights gathered from these 20 participants can still provide valuable qualitative feedback.

The investigation regarding the reasons why customers decide to join the gym provided enlightening outcomes, exposing a wide

Figure 11 Pie Chart of Joining Reason

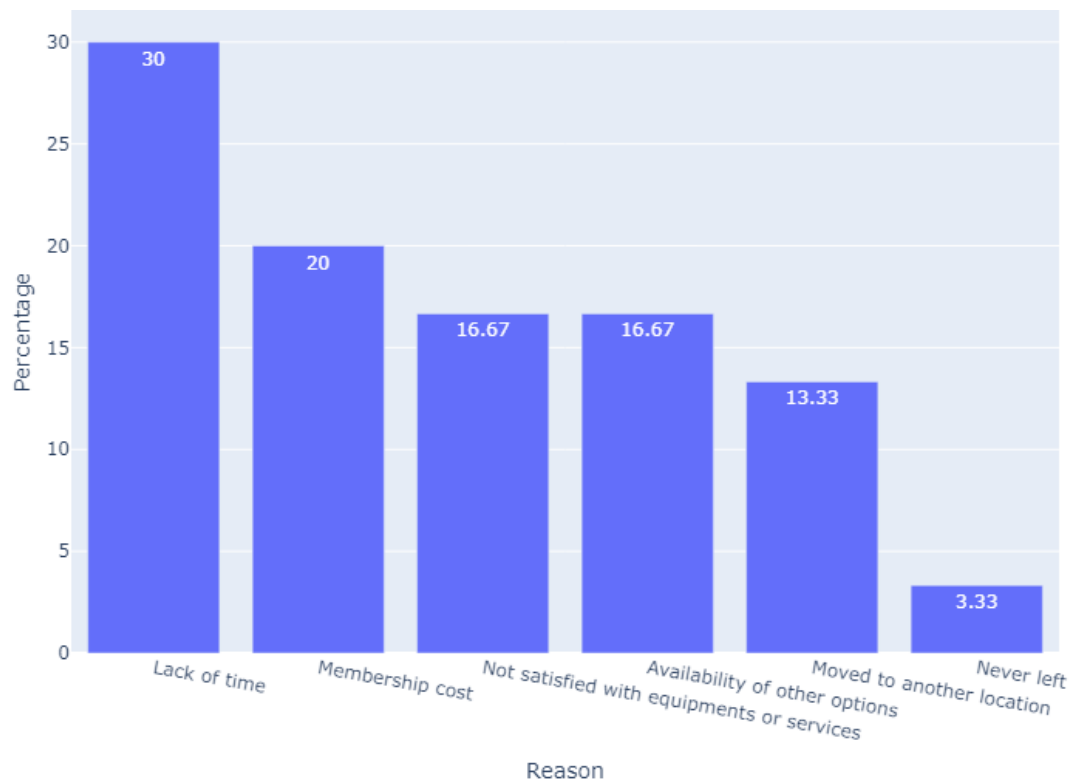


array of motivations. A noteworthy half of the participants mentioned fundamental fitness as their main aspiration, shedding light on a general inclination towards holistic well-being. Furthermore, 30% conveyed a particular emphasis on

bolstering their physical power, accentuating a longing to develop muscle mass and enhance overall strength. Each of the remaining 10% of individuals indicated undertaking a weight loss journey or pursuing a goal of augmenting their physique. These discoveries highlight the significance of customizing gym services and programs to accommodate diverse objectives, offering valuable insights for effectively engaging and satisfying customers.

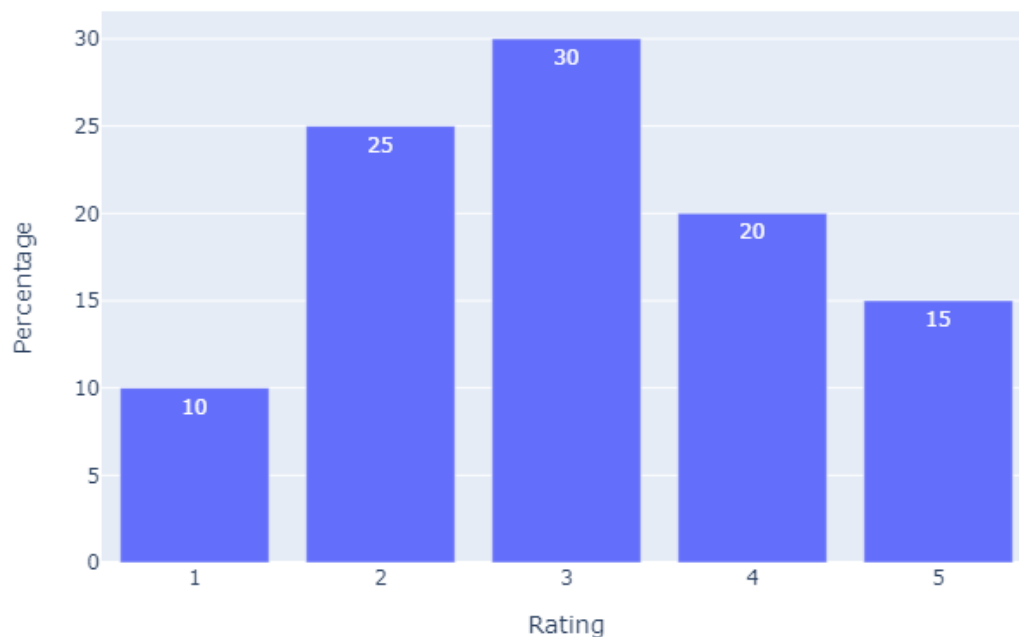
The data on customer reasons for leaving services (Figure 12) reveals key insights into the factors influencing customer attrition. The most prevalent reasons include "Lack of time" (30%) and "Membership cost" (20%), emphasizing the need to consider time constraints and pricing strategies. Issues related to service satisfaction and competition are also significant contributors, with "Not satisfied with equipment or services" and "Availability of other options" both at 16.67%.

Figure 12 Percent Bar Chart of Leaving Reasons

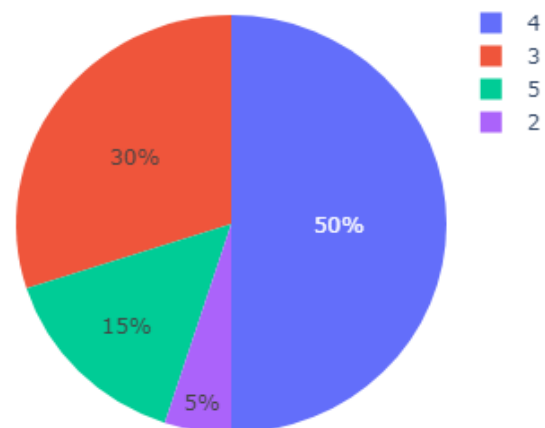


The distribution of ratings regarding pricing (Figure 13) offers a distinct representation of the degree to which customers are content. Remarkably, 65% of customers provided a rating of 3 or below, indicating their dissatisfaction with the pricing framework. Conversely, a mere 35% of clients assigned a rating of 4 or 5, implying that a smaller section found the pricing to be agreeable or pleasing. This disparity in ratings highlights a significant proportion of customers expressing their discontentment with the pricing, which suggests the need for improvement or modification in the pricing strategy to harmonize more effectively with customer expectations and augment overall contentment.

Figure 13 Bar Chart of Price Rating (in Percentages)



The overall experience ratings from *Figure 14 Overall Experience* customers reflect a positive sentiment, with 65% providing ratings of 4 or 5 out of 5. This majority indicates a high level of satisfaction and contentment with the gym services and facilities. Additionally, 30% of customers rated their experience as a 3 out of 5, signifying a more neutral sentiment, suggesting that their experience was satisfactory but possibly with room for improvement. Only a minimal 5% of customers rated their experience as 2 or below, indicating a relatively small portion of individuals expressing dissatisfaction.



The feedback on disliked features and lacking amenities in the gym provides valuable insights into areas that may require attention and improvement. Some recurring themes include dissatisfaction with the loud and uninspiring music, concerns about trainers and associated costs, and issues related to the availability of heavy dumbbells, free squat racks, and incline benches. Notably, the absence of certain equipment like treadmills, calf raise seated machines, and normal squat machines is mentioned, suggesting a desire for a more comprehensive range of workout options. Additionally, customers express a need for better facilities such as changing rooms, washrooms, and even amenities like hot and cold-water coolers. The call for additional services includes requests for diet charts, trainer support, and designated areas for warm-up, cool down exercises, stretching, and proper guidance. Addressing these concerns could contribute to enhancing the overall gym experience and satisfaction of the clientele.

3.3 Insight From Expenditure Data

The expenditure details of the gym reveal a comprehensive breakdown of costs, with ₹15,00,000.00 allocated to equipment and accessories, ₹5,00,000.00 to infrastructure, ₹2,34,900.00 for expenses between August and April, and ₹24,400.00 for costs incurred from May to August. The total expenditure amounts to ₹22,59,300.00. From August 2022 to April 2023, the gym's average expenditure amounted to ₹26100. However, the running cost decreased significantly in the months between May 2023 and August 2023 to ₹6100. This decline in average expenses can be attributed to the departure of the gym's trainer. The absence of the trainer may have contributed to customer attrition. Additionally, as of the date when I

gathered this information, the gym's total revenue collection reached ₹521780.0, indicating a substantial **deficit of ₹1737520**.

Purpose	Amount
Equipment and Accessory	₹ 15,00,000.00
Infrastructure	₹ 5,00,000.00
August-April	₹ 2,34,900.00
May-August	₹ 24,400.00
Total	₹ 22,59,300.00

4 Interpretation of Results and Recommendation

After conducting a thorough analysis, it becomes evident that there was a significant decrease in both revenue and the number of active customers following March 23. Further examination revealed that while gender and package type were not directly linked to customer attrition, there was a notable discrepancy between the number of male and female customers, as well as between those with the "Muscle" package and the "Both" package. Additionally, it was observed that customer location played a role in gym attrition. A bar plot depicting the count of active customers by location revealed that Curaj had the largest customer base but experienced a substantial drop in active customer count after March 23.

Through personal conversations with customers and feedback forms, the main reasons for customer attrition were identified. The most frequently mentioned reason was a lack of time, followed by concerns about the cost of gym membership. Customers also expressed dissatisfaction with the gym's equipment, specifically noting the unavailability of certain items. Upon further investigation into the cost factor, it was discovered that the Muscle plan had a monthly membership fee of 1000rs, while the "Both" package (which included both cardio and muscle training) cost 1700rs per month. These prices were significantly higher compared to other gyms in the city offering similar facilities. Some customers pointed out that despite paying 1000 rupees for the muscle package, they were unable to use the treadmill without paying an additional 700rs. Others mentioned the lack of basic amenities such as changing rooms, washrooms, and even simple conveniences like hot and cold-water coolers.

Additionally, many customers from Curaj (Central University of Rajasthan) stated that the university had recently renovated its own gym, which offered free facilities such as treadmills and an air-conditioned environment to its students.

During the feedback survey regarding the university gym, students from Curaj reported various issues, including problems with timing, equipment quality, and relatively limited time slots for usage (each student was allocated a 45-minute time slot).

Based on the data and feedback from customers, I have some recommendations for the business. Here are a few suggestions:

- **Revise Pricing Strategy:** It may be beneficial to reconsider the gym's pricing strategy in order to be more competitive in the local market. Evaluate the pricing of different packages, particularly the "Muscle" and "Both" packages, to ensure they align with customer expectations and local market standards. Offering transparent and inclusive pricing can improve the perception of value and attract a wider customer base.
- **Provide One-Week Free Personal Training to New Members:** Attract new customers by offering a complimentary one-week personal training session to those who join. This not only incentivizes potential customers, but also allows them to experience the benefits of personalized guidance. This trial period can help demonstrate the expertise of the trainer, address any initial concerns, and potentially convert trial users into long-term members.
- **Address Trainer Availability Issue:** It is of utmost importance to prioritize the prompt filling of vacant trainer position. A key aspect in delivering consistent and dependable customer support lies in having a competent and accessible training staff. It is crucial to regularly evaluate the staffing requirements and guarantee the presence of a well-prepared team to aid members.
- **Improve Equipment Availability:** Respond to customer concerns regarding the availability of specific equipment by ensuring a well-maintained and diverse range of gym equipment. This can enhance the gym's appeal, attracting customers with specific fitness preferences and needs.
- **Capitalize on Curaj Gym's Weaknesses:** Address the identified issues with the university gym by positioning “Muscle Factory Gym” as an appealing alternative. Provide flexible time slots to accommodate different schedules, ensure top-notch equipment, and implement effective crowd control measures. Highlight these advantages in the marketing strategy to attract university students and staff who may be dissatisfied with the limitations of the current university gym.
- **Improve Amenities and Facilities:** Make investments in essential amenities like changing rooms, restrooms, and water coolers to enhance the overall customer

experience. Attending to these fundamental needs can greatly contribute to customer satisfaction and retention. Furthermore, explore the possibility of offering complimentary services, such as access to treadmills, to make memberships more attractive.

- **Organize Gym Competitions:** Foster a sense of community and motivation by arranging fitness competitions within the gym. Events like weightlifting competitions, push-up challenges, or even fitness challenges can create a positive and engaging atmosphere. Not only do these events provide a platform for members to showcase their accomplishments, but they also promote camaraderie among the members. Consider providing incentives or recognition for winners to encourage participation.
- **Expand Marketing Efforts to Nearby Villages:** To diversify the customer base, broaden marketing efforts beyond the immediate vicinity. Target neighbouring villages through advertising campaigns that emphasize the gym's facilities, competitive pricing, and any special promotions. Offering promotional packages to residents in nearby villages could attract new members and contribute to a more diverse customer demographic.
- **Enhance Communication and Conduct Customer Surveys:** Strengthen communication channels with both current and potential customers. Regularly update customers about new facilities, promotions, and improvements. Utilize social media and other platforms to raise awareness and encourage engagement. Conduct frequent surveys to gather ongoing feedback from customers. Utilize the insights gained to continuously improve services, address emerging concerns, and ensure that the gym's offerings align with evolving customer preferences.

By incorporating these recommendations, the gym can strive to create a more inclusive, appealing, and competitive fitness environment, thus reducing attrition and fostering sustained growth.

Thank You