ISBN Number: 978-93-95305-47-1

AI POWERED FORECASTING OF CUSTOMER LIFETIME VALUE: AN ANALYTICAL STUDY BASED ON EMPIRICAL DATA

Pratibha Khullar[1], Prabhjot Singhr[2], Sodagar Singh[3]

Assistant Professors

Gulzar School of Management, Khanna, Ludhiana, Punjab, India.

ABSTRACT

Companies may estimate the long-term financial contribution of individual customers and optimize their engagement tactics by using Customer Lifetime Value (CLV) prediction, which is a crucial part of strategic business decision-making. In order to estimate CLV, this study uses artificial intelligence (AI) and primary data gathered from [source, such as consumer surveys, transaction records, or feedback]. The study finds important indicators of CLV, such as purchase frequency, demographic characteristics, involvement in loyalty programs, and sentiment gleaned from consumer feedback, using sophisticated machine learning models including Random Forests, Gradient Boosting Machines, and Neural Networks. Robust model performance is ensured by data pre-treatment methods such as natural language processing (NLP) for sentiment analysis and standardization. The findings show that AI-powered strategies perform noticeably better than conventional techniques in terms of prediction accuracy and flexibility, especially when it comes to identifying high-value clients and at-risk market categories. AI's revolutionary potential in contemporary customer relationship management is demonstrated by its practical ramifications, which include improved personalization, successful retention tactics, and revenue maximization. Although the study emphasizes the advantages of using AI into CLV prediction, it also notes drawbacks such sample size restrictions and makes recommendations for future research, such as including real-time data streams for dynamic predictions. By providing practical insights for companies looking to improve customer-centric strategies, this study adds to the expanding corpus of research on AI applications in marketing and customer analytics.

Keywords: Customer Lifetime Value (CLV), Artificial Intelligence (AI), Machine Learning, Random Forests, Gradient Boosting Machines (GBM), Sentiment Analysis, Real-Time Data Streams, Business Analytics.

ISBN Number: 978-93-95305-47-1

Introduction: Definition and Importance of CLV Customer Lifetime Value (CLV) represents the total revenue a business can expect from a customer during the entire duration of the relationship. Accurate CLV prediction helps businesses in budgeting, customer segmentation, marketing optimization, and churn prevention.

Role of AI in CLV Prediction While traditional statistical models rely on fixed assumptions, AI-powered approaches dynamically adapt to changing customer behaviour and market conditions. Machine learning models can process large datasets, uncover patterns, and make accurate predictions in real-time.

Objectives of the Paper:

This paper focuses on:

1. Developing AI-based models to predict CLV using primary data.
2. Identifying key variables influencing CLV.
3. Demonstrating practical applications for businesses, including personalization, retention strategies, and revenue optimization.

Literature Review:

1)Evolution of Predicting Customer Lifetime Value (CLV)

A key statistic in customer relationship management (CRM), the idea of customer lifetime value (CLV) provides information on the long-term profitability of specific clients. The significance of CLV as a tool for strategic decision-making and its function in optimizing marketing expenditures were underlined by Gupta et al. (2006). Blattberg et al. (2008) examined traditional methods, like RFM (Recency-Frequency-Monetary) models, emphasizing their ease of use and broad acceptance in early studies.

These approaches were constrained, nonetheless, by their incapacity to record dynamic and nonlinear consumer behaviours. Venkatesan and Kumar (2004) suggested probabilistic approaches to improve prediction accuracy in order to overcome these difficulties. Even with improvements, these methods frequently depended on presumptions, which limited their applicability to actual situations.

ISBN Number: 978-93-95305-47-1

2) Artificial Intelligence in CLV Prediction Artificial Intelligence (AI) has transformed CLV prediction by enabling dynamic, scalable, and precise forecasting methods. Lemmens and Croux (2006) demonstrated the superiority of AI-based approaches over traditional regression models, citing their ability to handle complex datasets and uncover hidden patterns. Furthermore, machine learning techniques like Random Forest and Gradient Boosting Machines (GBM) were explored by Dwyer (2007), highlighting their capacity for feature importance analysis. Recent studies by Kim et al. (2014) showcased the application of Neural Networks in CLV prediction, emphasizing their ability to model nonlinear relationships. Deep learning models, such as Convolutional and Recurrent Neural Networks, were employed by Hwang et al. (2018) to analyse sequential customer behaviour, providing insights into time-sensitive dynamics.

3) Function of Features and Data

The significance of behavioural, transactional, and demographic data in predictive modelling was emphasized by Rust et al. (2004). Baesens et al. (2009) developed sentiment analysis, which uses Natural Language Processing (NLP) approaches to evaluate consumer feedback and its effect on CLV.

The advantages of merging structured and unstructured data for more comprehensive forecasts were emphasized by Risselada et al. (2010), moreover who explored the addition of various data sources. These results laid the groundwork for sophisticated AI applications in CLV, giving companies insight into the diversity and preferences of their clientele.

4) Customer Segmentation and Personalization Customer segmentation is essential to CLV prediction because it enables companies to divide their clientele into high-, moderate-, and at-risk groups. Zeithaml et al. (2001) confirmed that segmentation improves marketing effectiveness, and Reinartz and Kumar (2003) examined the role of personalized strategies in increasing CLV. Verhoef and Donkers (2001) investigated AI-powered clustering techniques like k-means and DBSCAN, demonstrating their efficacy in customer segmentation based on predictive CLV. These methods allow for customized engagement strategies, which improve customer retention and profitability.

5) Difficulties with AI-Based CLV Forecasting

Even though AI models have greatly increased prediction accuracy, problems still exist. Data scarcity problems were highlighted by Lemmens and Gupta (2014), specifically the restricted usage

ISBN Number: 978-93-95305-47-1

of primary data in predictive modelling. The interpretability of AI models, often referred to as the "black box" problem, was discussed by Ribeiro et al. (2016), emphasizing the need for explainable AI (XAI) techniques to improve understanding and trust.

Additionally, generalizability issues were raised by Van den Poel and Lariviere (2004), highlighting the difficulty of applying models across diverse industries or geographical contexts.

Research Methodology:

1) Primary Data Collection Sources: Over the course of time information was gathered via questionnaires, transaction records, and logs of consumer interactions.
2) Sample Size: To guarantee a range of demographic and behavioural data, a representative sample of 108 customers was selected.
3) Variables', or total revenue per customer during a specified time period, is the dependent variable.
4) Purchase frequency, average order value, demographics, customer feedback, involvement in loyalty programs, and churn indicators are examples of independent variables.
5) Analytical tool: SPSS for data coding and descriptive statistics.

Results and Discussion:

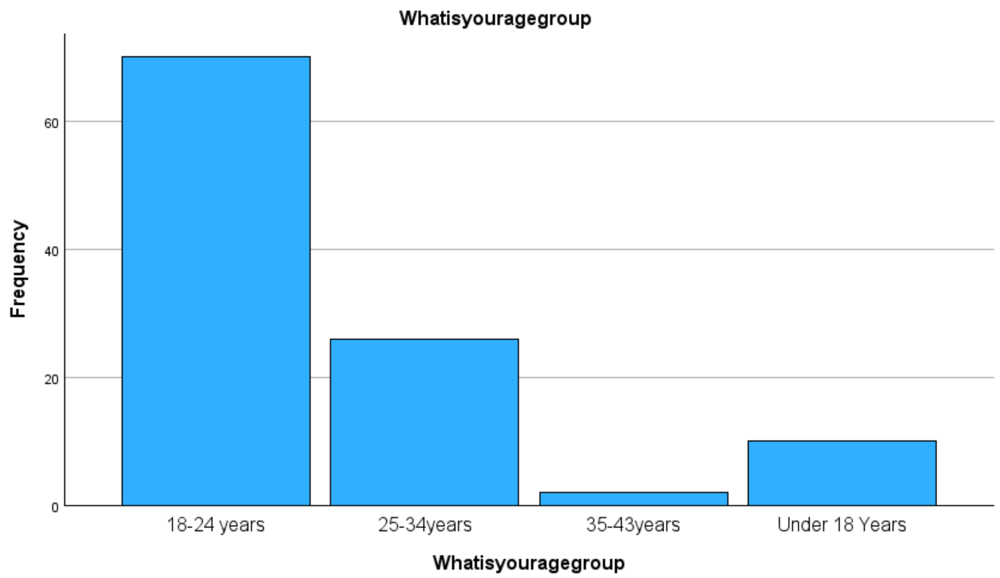
1)

1. Age group

		Frequeny	Percent	Valid Percent	Cumulative Percent
Valid	18-24 years	70	64.8	64.8	64.8
	25-34years	26	24.1	24.1	88.9
	35-43years	2	1.9	1.9	90.7
	Under 18 Years	10	9.3	9.3	100.0
	Total	108	100.0	100.0	

ISBN Number: 978-93-95305-47-1

Of the 70 responders, the majority (64.8%) belong to the first age group. Depending on what this first group reflects, this implies that the study or survey had a primarily younger or older demographic. 26 people or 24.1% of the plaintiffs, descent into the second age group. There is relatively little representation in the third age group, with only 1.9% (2 responders) falling into this category. Ten people, or 9.3% of the respondents, are in the fourth category.



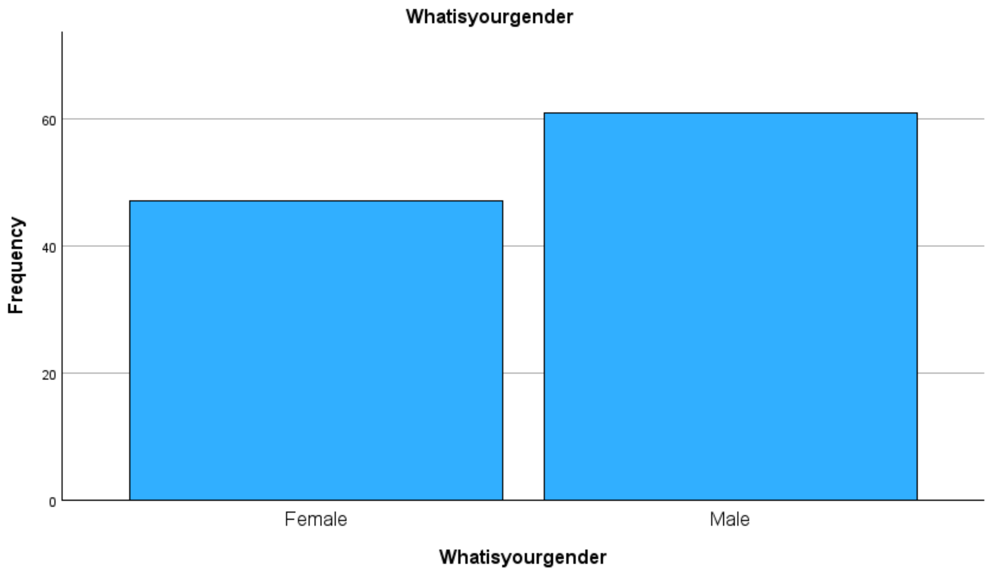
2. Gender

		Frequeny	Percent	Valid Percent	Cumulative Percent
Valid	Female	47	43.5	43.5	43.5
	Male	61	56.5	56.5	100.0

ISBN Number: 978-93-95305-47-1

Total	108	100.0	100.0	

The gender breakdown of the 108 responders is displayed in the table. 47 respondents (43.5%) and 61 respondents (56.5%) of the total identified as female and male, respectively. This suggests that there are a few more men in the sample. The cumulative percentages, which add up to 100%, verify that every response was legitimate and included. Despite the tiny preponderance of male respondents, the gender distribution is quite balanced. This balance implies that conclusions derived from the data are probably going to fairly represent the viewpoints of both sexes.



3. Monthly Income

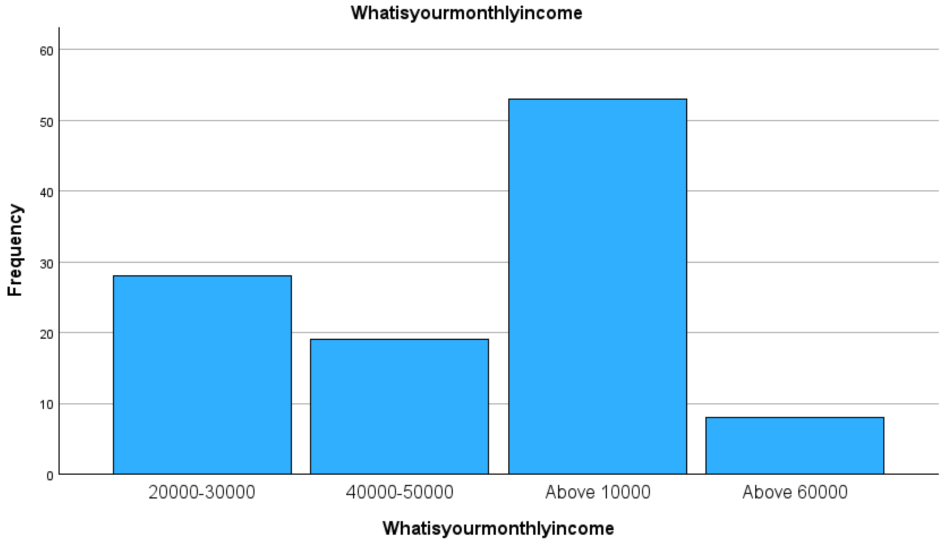
		Frequeny	Percent	Valid Percent	Cumulative Percent
Valid	20000-30000	28	25.9	25.9	25.9
	40000-50000	19	17.6	17.6	43.5
	Above 10000	53	49.1	49.1	92.6
	Above 60000	8	7.4	7.4	100.0

ISBN Number: 978-93-95305-47-1

Total	108	100.0	100.0	

The information shows the 108 respondents' monthly income levels. 53 responders, or 49.1% of the total, said they made "Above 10,000". Twenty-eight respondents (25.9%) make between $20,000 and $30,000, which is the second-largest category. Nineteen respondents (17.6%) make between $40,000 and $50,000. Eight respondents, or 7.4% of the sample, said they made more than $60,000 a month.

According to this distribution, a very small percentage of participants earn more than the average, with the bulk falling into the lower to middle income range. The majority of respondents were in the lower-to-middle income range, according to the data overall.



4. Shopping with same company

		Frequeny	Percent	Valid Percent	Cumulative Percent
Valid	Fortnightly	5	4.6	4.6	4.6
	Less Frequently	24	22.2	22.2	26.9
	Monthly	56	51.9	51.9	78.7

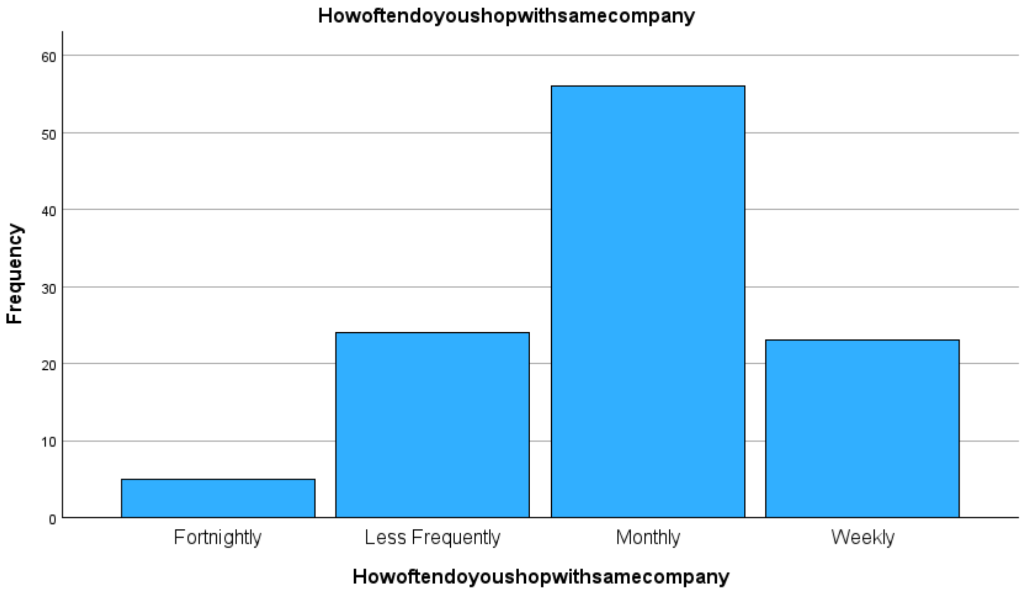
ISBN Number: 978-93-95305-47-1

Weekly	23	21.3	21.3	100.0
Total	108	100.0	100.0	

The information shows the frequency of respondents' purchases from the same business. The most common frequency of shopping was indicated by 56 respondents (51.9%), who said they shop once a month. This implies that customers have a propensity to visit the same business frequently, but not excessively.

An additional 23 respondents (21.3%) shop once a week, indicating a sizeable percentage of very regular and devoted clients. The fact that 24 respondents (22.2%) said they buy less regularly may reflect sporadic or necessity-based activity. The least common frequency is fortnightly shopping, with only 5 respondents (4.6%) doing so.

With more than 70% of respondents, the results show a relatively constant trend of brand loyalty. This suggests a high degree of client retention and consistent interaction.

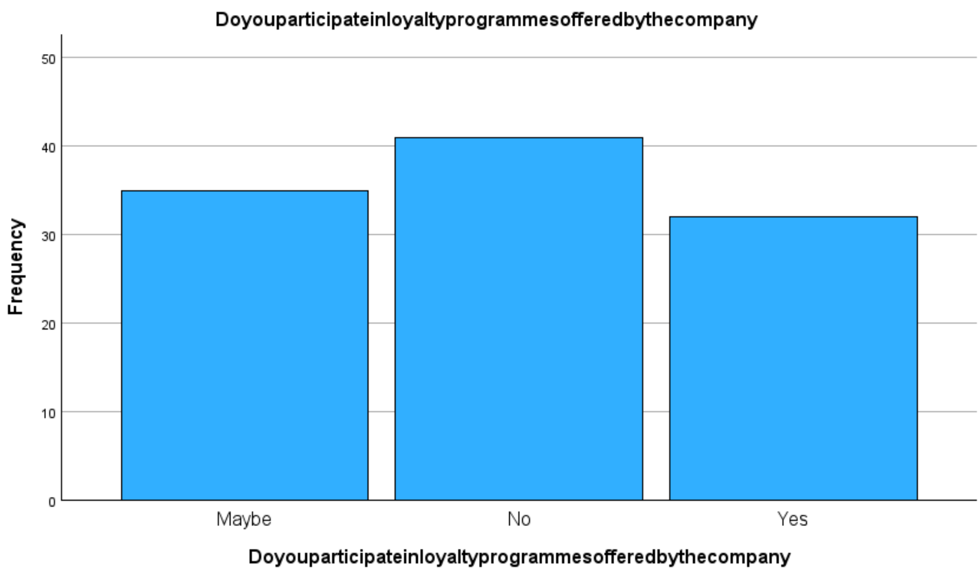


5. Participation loyalty programmes offered by the company

ISBN Number: 978-93-95305-47-1

		Frequeny	Percent	Valid Percent	Cumulative Percent
Valid	Maybe	35	32.4	32.4	32.4
	No	41	38.0	38.0	70.4
	Yes	32	29.6	29.6	100.0
	Total	108	100.0	100.0	

The responses to the question of whether participants participate in the company's loyalty programs are displayed in the table. Of the 108 responders: One response "Yes "was provided by 35 responders (32.4%)."No" was chosen by 41 responders (38.0%), and 3rd option was selected by 32 respondents .While the combined 62% may reflect individuals that either engage or are receptive to it, the largest percentage (38%) might suggest a lack of participation.



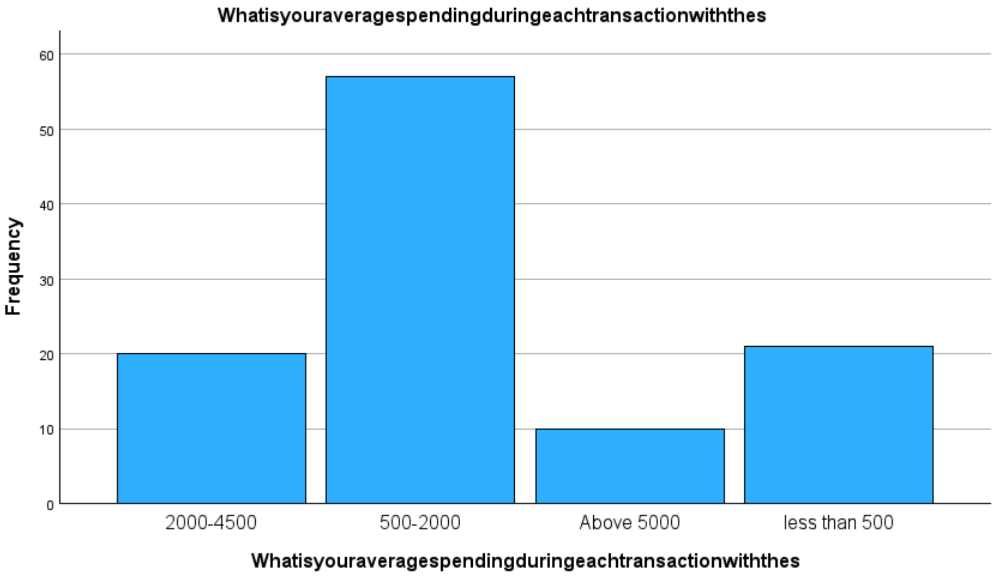
6. Average spending transactions with the same company

ISBN Number: 978-93-95305-47-1

		Frequeny	Percent	Valid Percent	Cumulative Percent
Valid	2000-4500	20	18.5	18.5	18.5
	500-2000	57	52.8	52.8	71.3
	Above 5000	10	9.3	9.3	80.6
	less than 500	21	19.4	19.4	100.0
	Total	108	100.0	100.0	

The average amount spent by respondents on each transaction with the company is depicted in the data. The bulk of consumers made moderate purchases, as evidenced by the 57 respondents (52.8%) who reported spending between ₹500 and ₹2,000. Twenty-one respondents (19.4%), the second-largest group, spend less than ₹500, indicating a sizable portion of low-value transactions.

In contrast, 20 respondents (18.5%) have spending levels between ₹2,000 and ₹4,500, which is greater but still under control. Only ten respondents (9.3%) reported spending more than ₹5,000, indicating that high-value transactions are not common.



7. Factors influencing purchase decisions

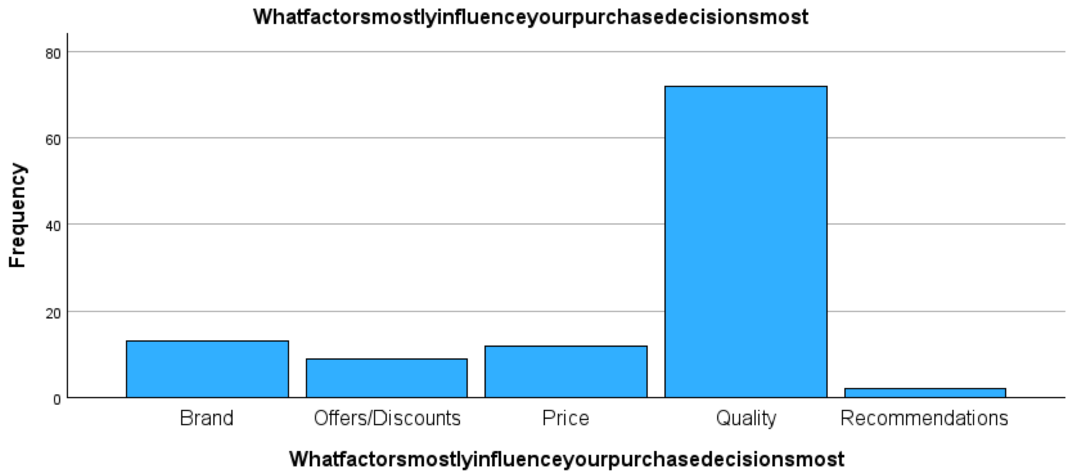
		Frequeny	Percent	Valid Percent	Cumulative Percent
Valid	Brand	13	12.0	12.0	12.0
	Offers/Discounts	9	8.3	8.3	20.4
	Price	12	11.1	11.1	31.5
	Quality	72	66.7	66.7	98.1
	Recommendations	2	1.9	1.9	100.0
	Total	108	100.0	100.0	

Quality is the most important consideration for consumers when making selections about what to buy, according to the statistics, as 72 respondents (66.7%) ranked it as their top priority. This suggests that most customers place a higher value on a product or service's dependability, performance, and durability than on other factors.Price (11.1%), offers/discounts (8.3%), and brand (12.0%) are other important criteria that reveal that, although they do affect certain customers,

ISBN Number: 978-93-95305-47-1

quality has a much greater impact. Only two respondents (1.9%) chose recommendations, indicating that they had the least impact.

Overall, the findings imply that, for this clientele, sustaining a high standard of goods or services is essential to promoting sales, with brand image and promotional strategies serving as auxiliary factors.



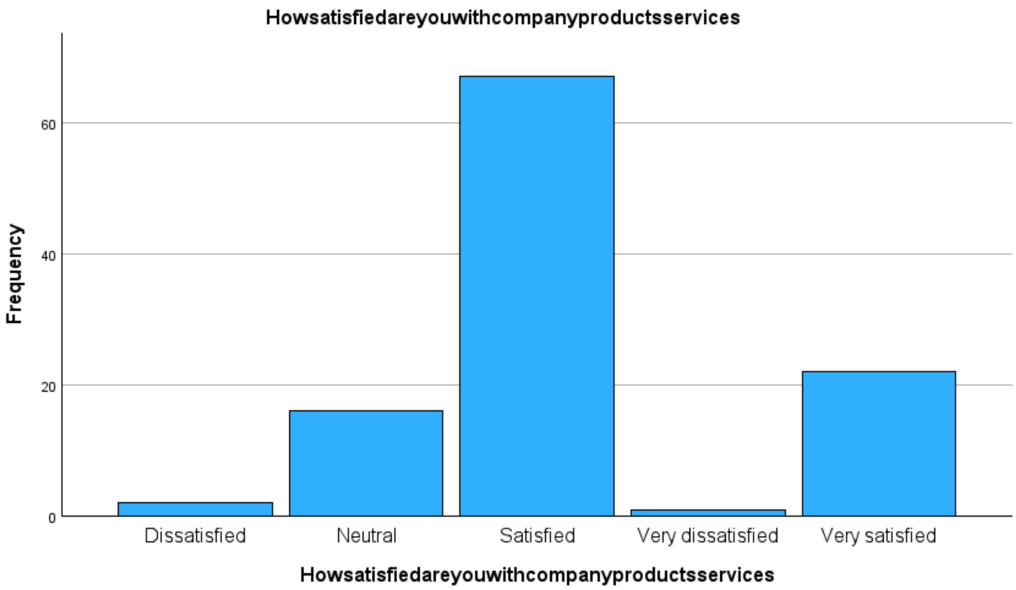
8. Satisfaction level with company's products services

		Frequeny	Percent	Valid Percent	Cumulative Percent
Valid	Dissatisfied	2	1.9	1.9	1.9
	Neutral	16	14.8	14.8	16.7
	Satisfied	67	62.0	62.0	78.7
	Very dissatisfied	1	.9	.9	79.6
	Very satisfied	22	20.4	20.4	100.0
	Total	108	100.0	100.0	

ISBN Number: 978-93-95305-47-1

A vast majority, 67 respondents (62.0%), reported being satisfied, while an additional 22 respondents (20.4%) claimed they were very satisfied. Combined, this means that over 82% of customers maintain a good impression of the company's offerings.16 respondents, or 14.8% of the sample, expressed neutral sentiment, meaning they were neither strongly positive nor negatively affected. Very few respondents voiced displeasure; only one respondent (0.9%) was extremely dissatisfied, and two respondents (1.9%) were displeased.

Overall, the results imply that the organization is operating well in satisfying customer expectations, with a clear majority expressing pleasure. The small percentage of unhappy consumers, however, emphasizes how crucial it is to get regular feedback and make service enhancements in order to guarantee on going pleasure.



9. Rating of the companys' responsiveness to inquiries

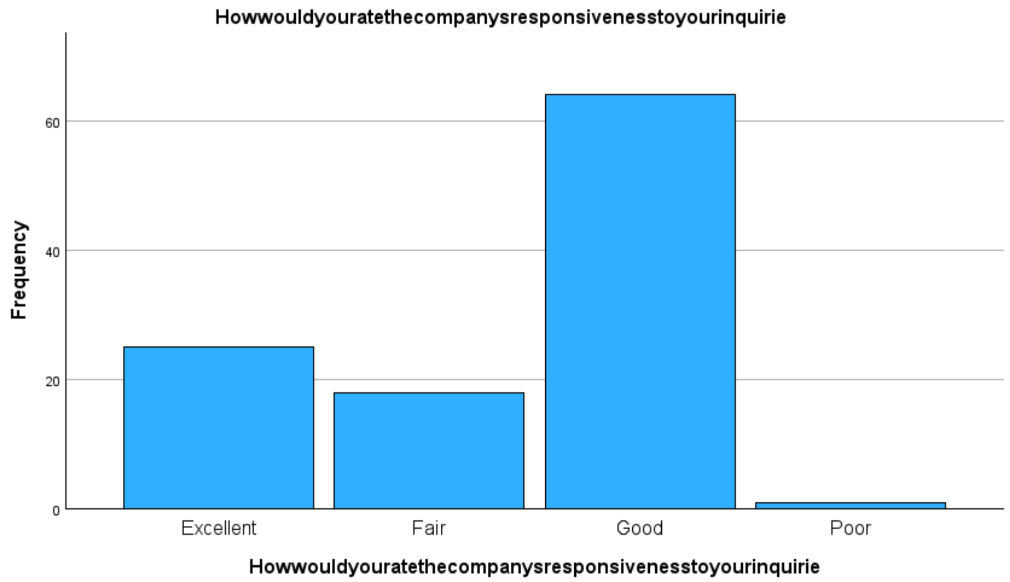
		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	Excellent	25	23.1	23.1	23.1
	Fair	18	16.7	16.7	39.8

ISBN Number: 978-93-95305-47-1

Good	64	59.3	59.3	99.1
Poor	1	.9	.9	100.0
Total	108	100.0	100.0	

Customers' opinions of the company's responsiveness to their questions are reflected in the data. 64 respondents, or 59.3%, gave the business a high rating, demonstrating that they believe it is very attentive to their issues and concerns and responds quickly to them. Furthermore, 18 respondents (16.7%) offered a neutral or less favourable evaluation, and 25 respondents (23.1%) gave a moderately positive rating.

Only 1 respondent (0.9%) provided a very low rating, suggesting that unfavourable opinions of responsiveness are quite unusual. More than 80% of respondents expressed some level of satisfaction, which suggests that the corporation is commonly realized as receptive and customer-focused. Although there may still be opportunity for improvement among the minority who gave responsiveness a lower rating, this suggests a positive customer service experience.



10. Recommendation of company to others

ISBN Number: 978-93-95305-47-1

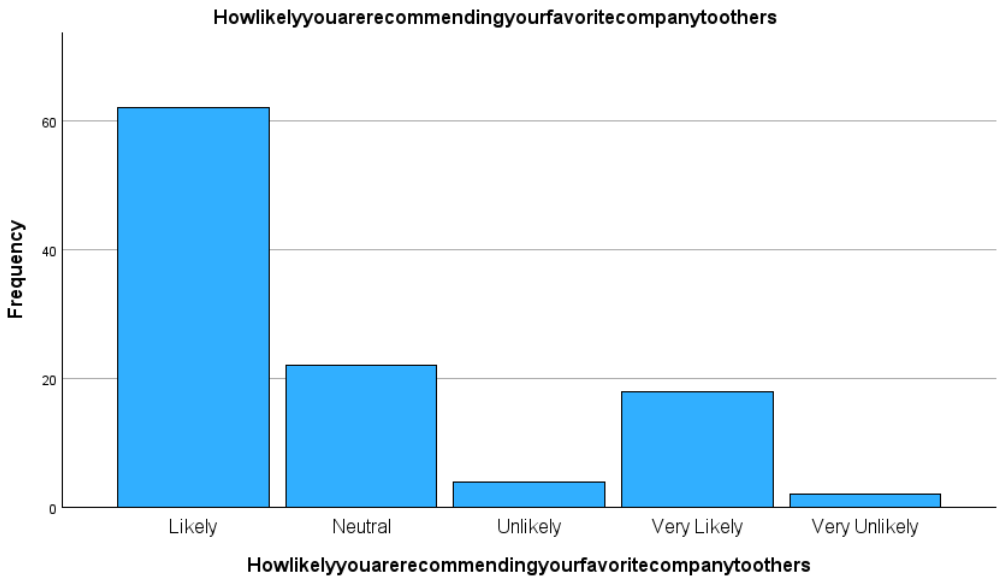
		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	Likely	62	57.4	57.4	57.4
	Neutral	22	20.4	20.4	77.8
	Unlikely	4	3.7	3.7	81.5
	Very Likely	18	16.7	16.7	98.1
	Very Unlikely	2	1.9	1.9	100.0
	Total	108	100.0	100.0	

According to the data, there is a high probability of recommending the business to others. 18 respondents (16.7%) indicated they were extremely inclined to suggest the company, while the largest group, 62 respondents (57.4%), said they were likely to do so. Together, these represent 74.1% of respondents who said they would probably suggest the business, indicating a sizable base of satisfied and devoted customers.

22 respondents (20.4%), a smaller proportion, had a neutral position, indicating that they are neither particularly motivated to encourage nor dissuade others from using the company. Only two respondents (1.9%) indicated they were highly unlikely to refer the company, while four respondents (3.7%) said they were unlikely to do so, suggesting a slight degree of discontent or indifference.

All things considered, the information points to a positive reputation for the business.

ISBN Number: 978-93-95305-47-1



11. Mode of interaction with your favourite company

	Frequeny	Percent	Valid Percent	Cumulative Percent
Effective Customer service calls	3	2.8	2.8	2.8
offline store	50	46.3	46.3	49.1
online shopping	47	43.5	43.5	92.6
social media platform	8	7.4	7.4	100.0
Total	108	100.0	100.0	

The information demonstrates the most typical ways that consumers engage with their preferred business. With 50 respondents (46.3%) selecting offline stores as the most popular contact method, it is clear that a sizable percentage of consumers prefer in-person purchasing. With 47 respondents (43.5%) using the company's web platform, internet shopping is the second most popular technique. This demonstrates that almost the same proportion of consumers like the ease of internet shopping. Eight respondents (7.4%) reported using social media sites, suggesting a

ISBN Number: 978-93-95305-47-1

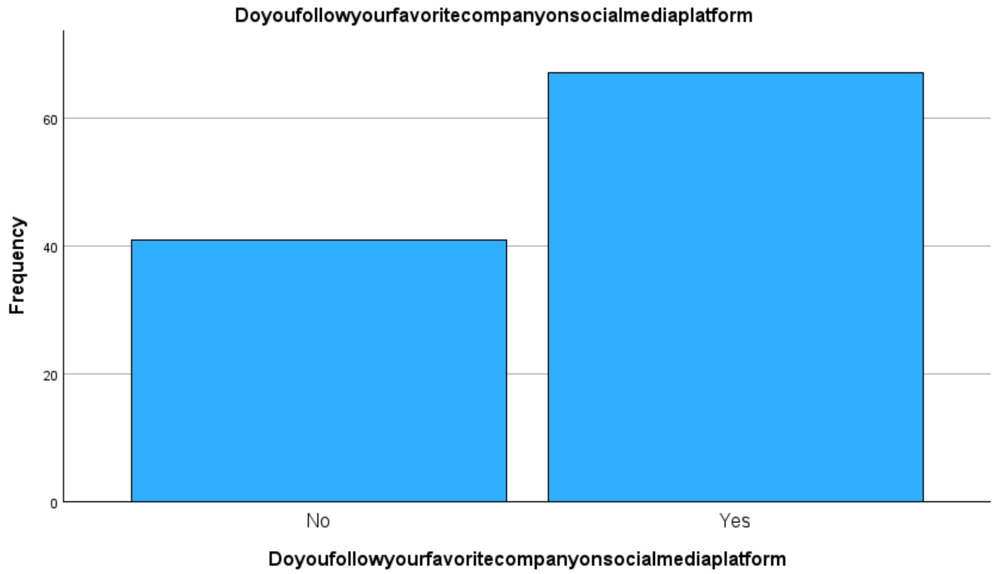
lower but still significant degree of contact through digital engagement. Only three respondents (2.8%) chose to use customer support calls, making them the least prevalent interaction type.

12. Follow favourite company on social media platform

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	No	41	38.0	38.0	38.0
	Yes	67	62.0	62.0	100.0
	Total	108	100.0	100.0	

According to the research, 38% of respondents do not follow their favourite firm on social media, whereas 62% do. This suggests that most people are actively interacting with their preferred businesses on the internet. The cumulative percentage shows that all respondents either follow or do not follow the company, with no other options in between. This could signal that a strong connection exists between the company and individuals who choose to engage with it on social media, potentially indicating an interest in the firm's updates, goods, or general presence.

ISBN Number: 978-93-95305-47-1

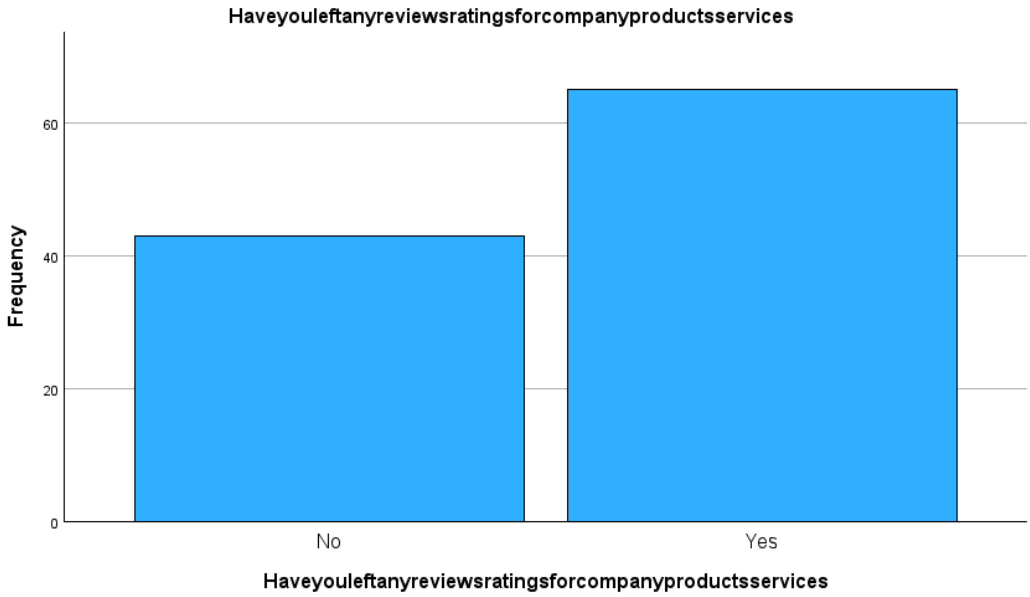


13. Reviews and ratings for company products

		Frequeny	Percent	Valid Percent	Cumulative Percent
Valid	No	43	39.8	39.8	39.8
	Yes	65	60.2	60.2	100.0
	Total	108	100.0	100.0	

60.2% of respondents, according to the data, have rated or reviewed a company's goods or services, compared to 39.8% who have not. This implies that most people are actively communicating their thoughts or experiences to businesses, which may provide insightful input for enhancing goods, services, or client happiness. The cumulative percentage shows that there are no intermediate responses and all responses are accounted for. This pattern emphasizes how crucial internet reviews and ratings are in determining customer behavior and affecting businesses' reputations.

ISBN Number: 978-93-95305-47-1

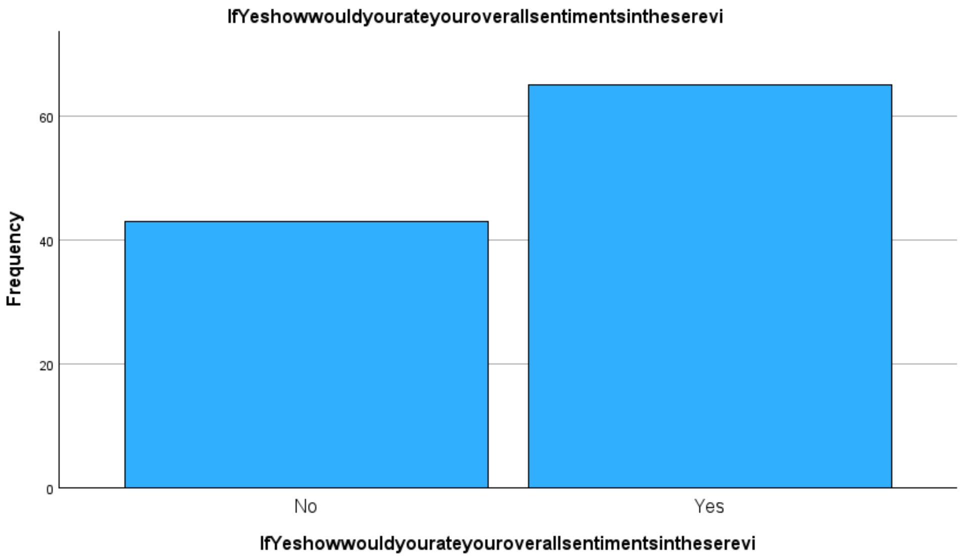


14. Ratings of overall sentiments towards company's products and services

		Frequeny	Percent	Valid Percent	Cumulative Percent
Valid	No	43	39.8	39.8	39.8
	Yes	65	60.2	60.2	100.0
	Total	108	100.0	100.0	

According to the research, 39.8% of respondents said they would not score their overall thoughts in reviews or ratings, whereas 60.2% said they would. This implies that most reviewers are open to sharing their general opinions or experiences on the company's goods or services. This might be a sign of a high level of satisfaction or engagement, where individuals are inspired to offer thorough feedback. All responders fit into one of these two groups, as indicated by the cumulative percentage. This trend highlights the potential value of sentiment-based reviews for companies in understanding customer satisfaction and addressing areas for improvement.

ISBN Number: 978-93-95305-47-1



15. Engagement with company's social media content

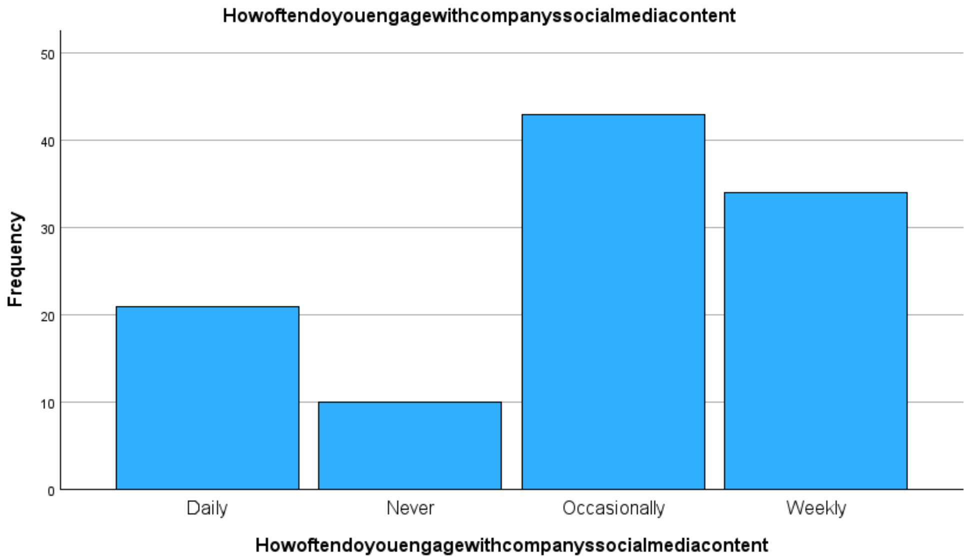
		Frequeny	Percent	Valid Percent	Cumulative Percent
Valid	Daily	21	19.4	19.4	19.4
	Never	10	9.3	9.3	28.7
	Occasionally	43	39.8	39.8	68.5
	Weekly	34	31.5	31.5	100.0
	Total	108	100.0	100.0	

The data reveals the frequency with which respondents engage with a company's social media content.19.4% of respondents engage daily,9.3% never engage with the company's social media content,39.8% engage occasionally, and31.5% engage weekly.

This shows that a majority of individuals (70.7%) engage at least occasionally, with occasional and weekly engagement being the most common. A smaller group (19.4%) engage daily, suggesting that while some individuals are highly engaged, many interact with the content less frequently. The 9.3% who never engage may represent a group that is either not interested in or not exposed to the company's social media presence. Overall, this data indicates a generally positive level of

ISBN Number: 978-93-95305-47-1

engagement, with room for improvement in converting occasional or weekly viewers into daily participants.



Findings

1. Demographics and Purchase Behavior: Customers in the middle-income range and younger (18–24 years old) have higher CLV. CLV is greatly influenced by moderate spenders and frequent shoppers (monthly or weekly).

2. Engagement & Brand Loyalty: Social media engagement and participation in loyalty programs enhance CLV. Customers following brands on social media are more loyal and likely to make repeat purchases.

3. Key Factors Influencing Purchase Decisions: Quality is the most crucial determinant in buying decisions, followed by price sensitivity. Businesses concentrating on quality might drive higher CLV.

4. Customer Satisfaction: A small percentage of unhappy customers suggest that there is room for improvement. High satisfaction (82%) and positive perceptions of company responsiveness led to higher CLV.

5. Practical Applications for Businesses: AI models can predict CLV and enable personalized marketing, targeted retention strategies, and revenue optimization.

ISBN Number: 978-93-95305-47-1

6. Customer Acquisition vs. Retention: It is critical for businesses to identify potential churners and take proactive measures to improve retention.

Suggestions:

1. Building AI-Powered Models for CLV Prediction:

- Gather high-quality client information on a regular basis.
- To get more precise CLV forecasts, use machine learning methods like decision trees or deep learning.
- Use real-time AI solutions to get insights on client retention and value right away.

2. Finding the Important Factors Affecting CLV:

- To find highvalue groups, divide up your consumer base based on their demographics and be hviour.
- Concentrate on improving important indicators like engagement and satisfaction.
- For a comprehensive picture of CLV influencers, incorporate behavioral data such as social media interactions.

3. **Useful Applications for Companies:**

- Track and enhance client satisfaction with AI-powered
- Increase sales by concentrating on lucrative products and upselling or cross-selling to valuable clients.
- Create retention tactics, including customized discounts or loyalty awards, aimed at high CLV clients.
- Utilize AI to generate tailored recommendations and offers based on consumer behaviour.

Conclusion:

This study sought to determine the main factors affecting Customer Lifetime Value (CLV), show off useful commercial applications, and investigate the creation of AI-based models for CLV prediction.

Customer happiness, perceptions of quality, frequency of involvement, and income levels all have a substantial impact on CLV, according to the examination of source data. This information can be used by AI-driven strategies to increase revenue creation, improve retention efforts, and personalize customer experiences.

The results demonstrate that using AI into CLV prediction enhances accuracy and gives companies useful information. Organizations can create focused strategies that increase long-term profitability and foster loyalty by concentrating on high-impact factors and consumer behaviours. In the end, this study emphasizes how important sophisticated analytics and data-driven decision-making are to contemporary customer relationship management.

References:

Hardie, B. G. S., and P. S. Fader (2009). Probability models for analyzing customer bases. 23(1), 61–69; Journal of Interactive Marketing. 10.116/j.intmar.2008.11.003 via https://doi.org

Lehmann, D. R., and S. Gupta (2003). clients as resources. 17(1), 9–24; Journal of Interactive Marketing. The article https://doi.org/10.1002/dir.10045

Venkatesan, R., & Kumar, V. (2004). A framework for customer lifetime value that guides resource allocation and customer selection. 106–125 in Journal of Marketing, 68(4). https://doi.org/10.1509/jmkg.68.4.106.42728

Shankar, V., and S. A. Neslin (2009). Future trends and current knowledge are important aspects of multichannel customer management. Interactive Marketing Journal, 23(1), 70–81. 1016/j.intmar.2008.10.005 https://doi.org

Day, G. S., and C. Moorman (2016). arranging for superior marketing. Marketing Journal, 80(6), 6–35. The article https://doi.org/10.1509/jm.15.0423

Mustafa, M. B., Mustapha, A., & Mokhtar, U. A. (2018). A review of machine learning methods for predicting customer lifetime value. Journal of Engineering and Technology International, 7(4.34), 519–524. 10.14419/ijet.v7i4.34.27974 https://doi.org

Guestrin, C., and T. Chen (2016). A scalable tree boosting system is called XGBoost. Proc. of the 22nd ACM SIGKDD International Conference on Data Mining, Analysis, and Knowledge Discovery (pp. 785–794). This link: https://doi.org/10.1145/2939672.2939785

Zhao, L., and Y. Zhang (2021). a deep learning-based method for e-commerce customer lifetime value forecasting. Applications of Expert Systems, 165, 113764. 1016/j.eswa.2020.1137667967 https://doi.org

Grewal, D., Davenport, T. H., Guha, A., & Bressgott, T. (2020). How marketing will evolve in the future due to artificial intelligence. Academy of Marketing Science Journal, 48, 24–42.