Applying Machine Learning to Analyze ATM Cash Withdrawal Patterns and Develop Optimization Model: Insights from Credit and Debit Card Transactions in India

Dr Riddhi Dave[1], Dr. Nili Shah[2], Rani Niral Shah[3], Dipti Gadhavi[4]

[1]Professor, Som-Lalit Institute of Management Studies, Ahmedabad
Mail ID: riddhidave@somlalit.org
ORCID iD: 0009-0002-6456-0288
[2]Assistant Professor, Som-Lalit Institute of Management Studies, Ahmedabad
Mail ID: nilishah@somlalit.org
[3]Assistant Professor, Som-Lalit Institute of Management Studies, Ahmedabad
Mail ID: ranishah@somlalit.org
[4]Assistant Professor, Som-Lalit Institute of Management Studies, Ahmedabad
Mail ID: dpgadhavi81@gmail.com

KEYWORDS

Machine learning, ATM, Credit card, Debit card, Clustering

ABSTRACT

Data analysis through Machine learning is becoming very pivotal and valuable for organizations. Banks are considered to be the backbone of our Economy. Banks in India are striving for resource optimization and cost cutting and use of ATMs has proved to be very cost effective strategy deployed by banks in India which benefits not only the banks but also customers. Indian scheduled commercial banks are reckoned globally for their best practices. The study has been undertaken with the objective of : Grouping the banks into clusters according to how their ATMs are used to withdraw cash using credit cards and debit cards; understanding the distribution of ATM cash withdrawals through credit cards and debit cards within each clusters and analyzing the relationship between the number of ATMs and ATM cash withdrawals made through credit and debit cards within each cluster, and to gain insights into how this relationship varies within and between clusters. In a nutshell, the research paper aims at analyzing the pattern of ATMs cash withdrawals of scheduled commercial banks in India done through credit card and debit cards through various machine learning techniques. The Viksit Bharat Mission envisions a developed India by fostering financial inclusivity, economic growth, and technological advancement. Optimizing the use of ATM machines in India can significantly contribute to these goals by ensuring efficient resource allocation and enhanced accessibility to banking services. The outcome of this research are expected to enable banks in framing strategies for optimizing their resources.

Introduction:

India's banks are using ATMs (Automated Teller Machines) more and more to streamline their business processes and improve client support.

Because ATMs are so widely used, banks are able to cut the operational expenses that come with having physical locations. Banks can cut expenses by minimizing the need for physical branches by encouraging clients to utilise ATMs for tasks like cash withdrawals, balance inquiries, and fund transfers. ATM usage enables banks to optimize their operations, reduce costs, enhance customer service, and drive strategic growth. By encouraging customers to use ATMs for their banking needs, banks can achieve operational efficiency and deliver a seamless banking experience to their customers.

When using an ATM, credit and debit cards are both frequently utilised to provide consumers with rapid access to cash and other financial services. When consumers use their credit cards rather than debit cards to make ATM withdrawals, banks stand to gain more. Banks receive payment from consumers using credit cards to make ATM withdrawals through transaction fees and interest charges. There is a fee associated with each credit card ATM transaction, which increases the bank's revenue. Interest is charged if payment is not made in full by the deadline, which raises the bank's profits even more.

However, since the money is taken out of the customer's account directly when they use their debit cards to make ATM withdrawals, banks are not charged interest in this scenario. Even while they can still receive a tiny fee from each ATM transaction, banks make less money on these transactions than they do on credit card purchases. So, be it thorough credit card or through debit card , banks tend to gain as , the more and more their ATMs are used.

By implementing advanced optimization techniques, such as linear programming and machine learning, banks can identify underutilized ATMs and reallocate resources to areas with higher demand. This not only improves financial infrastructure but also reduces operational costs, allowing banks to reinvest savings in digital infrastructure, rural outreach, and other developmental initiatives.

Review of Literature:

Ogbuji, C. N. et al. (2012) discuss that, one of the current alternatives to the labor-intensive, cascading transaction method carried out by what is often known as a paper-based payment instrument is the Automated Teller Machine (ATM). Customers of the bank can use an ATM to complete their financial transactions from nearly any ATM in the world. As a result, the ATM does various staffing duties as well as the conventional duties of bank cashiers. It functions electronically, and it responds to customer requests instantly. According to Kumbhar (2011), public and private sector banks have roughly the same views regarding the service quality dimension aside from cost effectiveness of automated teller machine services.

Pandlan et al. (2012) correctly conclude that survival in today's environment depends heavily on money. People can withdraw money from automated teller machines whenever they need to and in an emergency.

Michael (2001) discovered that bank switching is linked to the placement of ATMs, the quantity of ATMs, and the diversification of services offered. One kind of invention that can automatically take deposits, make withdrawals, move money between accounts, and collect bills is an automated teller machine (ATM). It emphasizes the quality of banking services and the relationships between the bank and its depositor (Komal & Singh 2009).

According to research by Joseph and Stone (2003), client satisfaction is significantly influenced by safe and convenient locations, an adequate quantity of automated teller machines, user-friendly systems, and ATM functioning.
On the other hand, Dilijonas, Krikscuiunen, Sakalauskas, and Simutis (2009) state that the main components of ATM service quality are a sufficient number of ATMs, safe and convenient locations, user-friendly systems, speedy, minimal error, highly untimed, cash backs, costing, and services covering. Shariq and London (2012) noted that consumers favor using ATMs over other online banking options.Tuli, Khatri, and Yadav (2012) found that the availability of machines and ease of use are the two main factors that encourage customers to use ATM services.
Tavsanli and Hamlaci (2021) used cluster analysis to map the Turkish banking industry and identify which banks are most likely to negatively impact the stability of the financial system. They divided Turkish banks into three clusters and discovered that the first cluster differed statistically significantly from the second and third clusters.
Gramatiy ,N, Volobueva, Y; Syurpita ,S (2021) found clustering as a very important technique and have created clustering of top five Ukrainian banks using selected financial parameters.

Objectives of the Study: The study has been undertaken with the following objectives.

1. Grouping the banks into clusters according to how their ATMs are used to withdraw cash using credit cards and debit cards.
2. To understand the distribution of ATM cash withdrawals through credit cards and debit cards within each cluster; visualizing the shape and spread of the data to identify differences or similarities between clusters.
3. To analyze the relationship between the number of ATMs and ATM cash withdrawals made through credit and debit cards within each cluster, and to gain insights into how this relationship varies within and between clusters.
4. To develop a model to optimize ATM allocation across clusters while maintaining current withdrawal efficiency levels.

The above objectives are believed to help banks in India to analyze how their ATMs are used for cash withdrawals through credit cards and debit cards as this analysis will enable banks to calibrate their strategies regarding optimizing ATM network management, boost consumer engagement, and improve service offerings.

Research methodology:

Reserve bank of India publishes every month the data titled ‘BANKWISE ATM/POS/CARD STATISTICS’. The data is available for all the types of commercial banks in India: Public Sector Banks, Private Sector Banks, Foreign Banks, Payment Banks and Small Finance Banks. However, through the review of literature it was found that there is no visible study done on how effectively different banks’ ATMs are used for cash withdrawals using credit cards and debit cards. For the study in this paper the above titled data by RBI for March 30th, 2023 has been used and analysis is made using this data. By applying clustering techniques to this data, banks that exhibit similar patterns in ATM cash withdrawals can be grouped together and banks can benchmark their performance against others in the same cluster.
To group the banks, K-means clustering technique is used as it is found that this technique works well for finding organic groupings in the data that are based on shared traits, such ATM usage trends. Furthermore as k-means clustering allows for the creation of mutually exclusive clusters, ensuring that each bank belongs to only one segment and hence ensuring the interpretability and

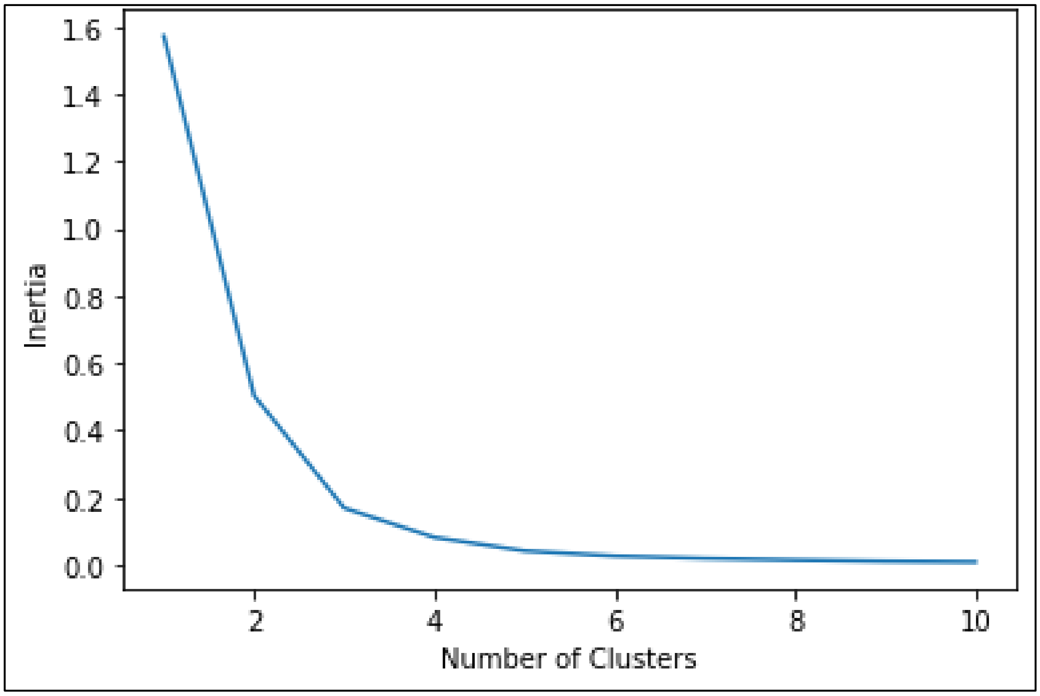
applicability of the clustering results for formulating cluster-specific strategies; the banks in the research here have been grouped using K-means clustering technique. Once the clustering is done machine learning tools available in python libraries –pandas, statsmodel, matplotlib and seaborn have been used for analysis.

Findings:

The researcher has employed the elbow method and silhouette score to determine the optimal number of clusters for K-means clustering.

The elbow method involves plotting the number of clusters against the sum of squared distances from each data point to its assigned cluster center. The "elbow point" on the plot represents the optimal number of clusters, where adding more clusters does not significantly reduce the sum of squared distances. Based on the data the following is the elbow curve created.

Chart 1: Elbow curve



(Source: Created by the Researchers)

The silhouette score, on the other hand, quantifies the quality of the clustering results. It measures how similar an object is to its own cluster compared to other clusters. A higher silhouette score indicates better-defined clusters. The following the silhouette score calculated by the research based on the analysis of the above elbow curve.

Table 1: Silhouette Score

No of clusters	Silhouette Score
3	0.821
4	0.806
6	0.682
8	0.675
10	0.652

(Source: Calculated by the Researchers)

Null Hypothesis (H0): There is no significant difference between the means of the clusters for the number of ATMs, ATM cash withdrawals through credit cards and ATM cash withdrawals through debit cards

Alternative Hypothesis (H1): There is significant difference between the means of the clusters for the number of ATMs, ATM cash withdrawals through credit cards and ATM cash withdrawals through debit cards.

Parameter	F statistic	P-value	Accept/Reject
Number of ATMs	0.597	0.554	Accept null Hypothesis
ATM cash withdrawals through credit cards	14.424	0.000	Reject null Hypothesis
ATM cash withdrawals through debit cards	892.859	0.000	Reject null Hypothesis

(Source: Calculated by the Researchers)

So, it can be concluded that there is no significant difference in the number of ATMs between the clusters but clusters show significant variation in cash withdrawals using credit cards and debit cards. This suggests that clustering is particularly effective in capturing cash withdrawal pattern using debit card and credit card indicative efficiency of clusters in using ATMs installed by them. As, can be seen from the above table the silhouette score is highest when the banks are grouped into three clusters. Post that K means clustering technique is applied and all the sixty scheduled commercial banks in India are grouped into the following three clusters based on the usage of their ATMs for cash withdrawals using credit cards and debit cards. The following table represents banks in each cluster from the clustering done above.

Table 2: Clustering of Banks based on K-means clustering technique

Clusters:	Banks:
Cluster 0	BANK OF INDIA
	BANK OF MAHARASHTRA
	CENTRAL BANK OF INDIA
	INDIAN OVERSEAS BANK
	PUNJAB AND SIND BANK
	UCO BANK
	BANDHAN BANK LTD
	CITY UNION BANK LTD.
	CSB BANK LTD.
	DCB BANK LTD
	DHANALAKSHMI BANK LTD
	FEDERAL BANK LTD
	IDBI BANK LTD

	IDFC FIRST BANK LTD
	INDUSIND BANK LTD
	JAMMU AND KASHMIR BANK LTD
	KARNATAKA BANK LTD
	KARUR VYSYA BANK LTD
	KOTAK MAHINDRA BANK LTD
	NAINITAL BANK LTD
	RBL BANK LTD
	SOUTH INDIAN BANK
	TAMILNAD MERCANTILE BANK LTD
	YES BANK LTD
	AMERICAN EXPRESS BANKING CORPORATION
	BANK OF AMERICA
	BARCLAYS BANK PLC
	CITI BANK
	DBS INDIA BANK LTD
	DEUTSCHE BANK LTD
	HSBC LTD
	SBM BANK INDIA LTD
	STANDARD CHARTERED BANK LTD
	WOORI BANK
	AIRTEL PAYMENTS BANK
	FINO PAYMENTS BANK
	INDIA POST PAYMENTS BANK
	JIO PAYMENTS BANK
	NSDL PAYMENTS BANK
	PAYTM PAYMENTS BANK
	CAPITAL SMALL FINANCE BANK LTD
	EQUITAS SMALL FINANCE BANK LTD
	ESAF SMALL FINANCE BANK LTD
	FINCARE SMALL FINANCE BANK LTD
	JANA SMALL FINANCE BANK LTD
	NORTH EAST SMALL FINANCE BANK LTD

	SHIVALIK SMALL FINANCE BANK LTD
	SURYODAY SMALL FINANCE BANK LTD
	UJJIVAN SMALL FINANCE BANK LTD
	UNITY SMALL FINANCE BANK LTD
	UTKARSH SMALL FINANCE BANK LTD
Cluster 1	STATE BANK OF INDIA
Cluster 2	**BANK OF BARODA**
	CANARA BANK
	INDIAN BANK
	PUNJAB NATIONAL BANK
	UNION BANK OF INDIA
	AXIS BANK LTD
	HDFC BANK LTD
	ICICI BANK LTD

(Source: Based on Calculation by the Researchers)

So, in a nutshell the following table represents no banks in each cluster as mentioned above
Table 3: Number of banks in each cluster based on K-means clustering technique

Cluster	No of Banks
0	51
1	1
2	8
Total	60

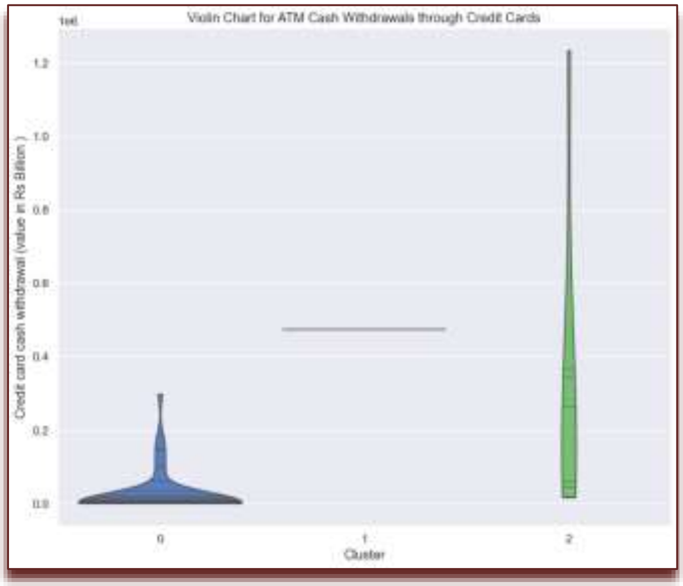
(Source: Based on Calculation by the Researchers)

The application of K-means clustering resulted in the formation of three distinct clusters based on the parameters: number of ATMs, cash withdrawals through ATMs using credit cards, and cash withdrawals through ATMs using debit cards**.** The clusters consist of 51 banks in Cluster 0, 1 bank in Cluster 1, and 8 banks in Cluster 2.

While the disparity in cluster sizes might seem unusual at first glance, it is a direct consequence of the underlying distribution of the data. K-means clustering aims to minimize within-cluster variance while maximizing the separation between clusters leading to the K-means clustering algorithm inherently grouping the banks with similar patterns together.

Once the clustering is done, for understanding the distribution of ATM cash withdrawals through credit cards and debit cards within each cluster; visualizing the shape and spread of the data to identify differences or similarities between clusters violin charts have been created each for ATM cash withdrawals through credit cards and debit cards within each cluster.

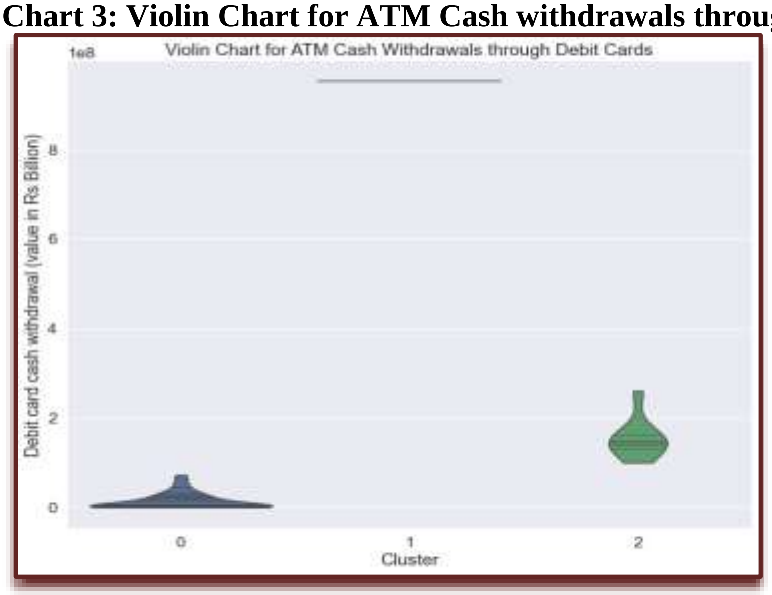
Chart 2: Violin Chart for ATM Cash withdrawals through credit cards.



(Source: Created by the Researchers)

As can be seen from above, the violin plot reveals distinct distributions of the ATM Cash withdrawal through credit cards across the three clusters .The violin for Cluster 0 indicates concentration of data points spread across specific values with low withdrawal amounts. The violin for cluster 1 appears as a straight line because of less variability of data points due to only one bank being part of the cluster 1. Also as can be seen in the chart above the amount of ATM cash withdrawal through credit cards in this cluster is more than cluster 0 and less than cluster 2. The violin for Cluster 2 shows a wider spread of data points. The wider body at lower values indicates a higher density of data points in that range, while the tail extending towards higher values suggests some transactions with significantly higher withdrawal amounts compared to Cluster 0 and 1.

Chart 3: Violin Chart for ATM Cash withdrawals through debit cards.

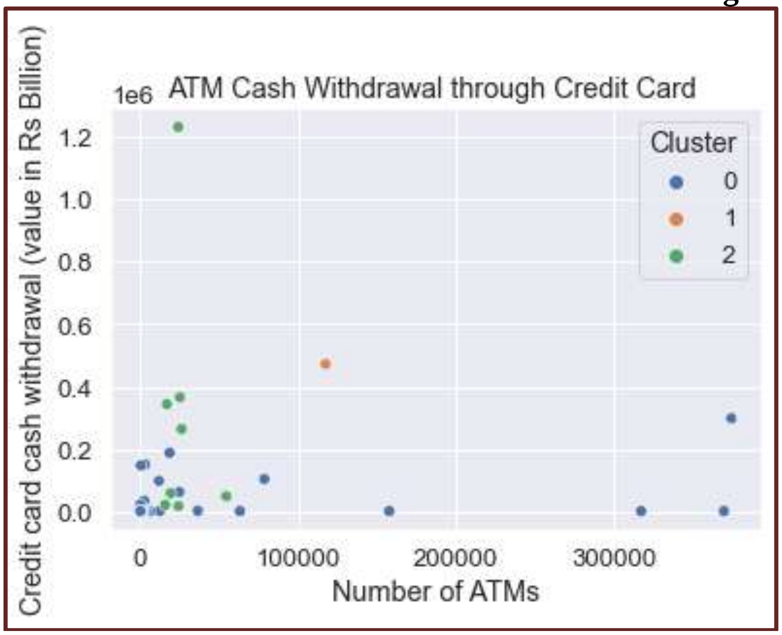


(Source: Created by the Researchers)

As can be seen from above, the violin plot reveals distinct distributions of the ATM Cash withdrawal through debit cards across the three clusters .The violin for Cluster 0 is narrow indicating a concentration of data points spread across specific values with low withdrawal amounts. The violin for cluster 1 appears as a straight line because of less variability of data points due to only one bank being part of the cluster 1. Also as can be seen in the chart above , the amount of ATM cash withdrawal through debits cards in this cluster is much more than both the clusters 0 and the cluster 2. The violin for Cluster 2 appears like a pot shape with narrow top and wide bottom showing more withdrawals of lower values and relatively less withdrawals of higher values.

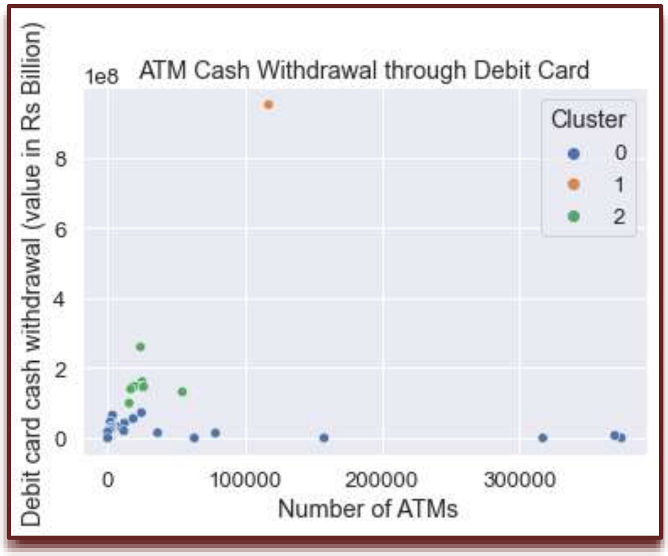
For analyzing the relationship between the number of ATMs and ATM cash withdrawals made through credit and debit cards within each cluster, and to gain insights into how this relationship varies within and between clusters scatter plots and heat maps are created.

Chart 4: Scatter Plot of ATM cash withdrawal through credit cards



(Source: Created by the Researchers)

Chart 5: Scatter Plot of ATM cash withdrawal through debit cards



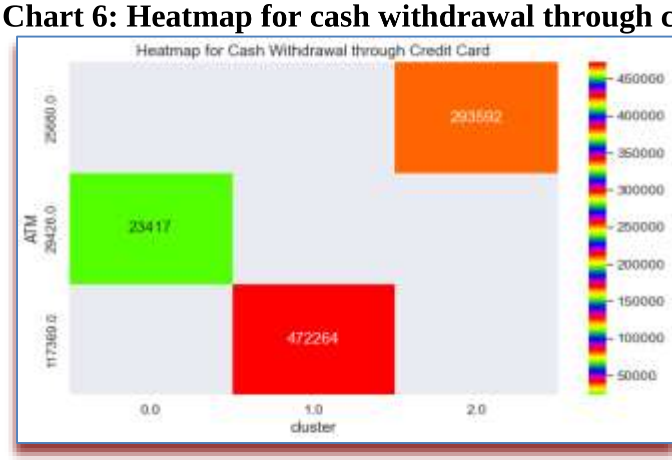
(Source: Created by the Researchers)

The above scatter plots shows cluster wise number of ATMs and cash withdrawals through credit cards and debit cards respectively. As can be seen from the chart, banks in cluster 0 have highest range as far as number of ATMs is concerned, however, the amount of withdrawals through credit cards scatter at relatively lower values showing lowest withdrawal per ATM**.** Whereas, banks in the cluster 2 can be seen generating more cash withdrawal through credit card per ATM. And for the cluster 1 it can be seen- per ATM cash withdrawal through credit card being more than cluster 0 and less than cluster 2. Hence, it can be concluded that banks in cluster 2 have the highest efficiency as far as credit card withdrawal is concerned as it generates the highest cash withdrawals per ATM compared to banks in other two clusters**.**

For, cash withdrawals through debit cards, as can been seen in the plot above, the cluster 1 can be seen generating more cash withdrawal through debit card per ATM. And for the cluster 2 it can be seen- per ATM cash withdrawal through debit card being more than cluster 0 and less than cluster 1. Hence, it can be concluded that cluster 1 has the highest efficiency as far as debit card withdrawal is concerned as it generates the highest cash withdrawals per ATM compared to banks in other two clusters**.**

To dwell deeper into the analysis. Heatmaps are created by the researcher as shown below

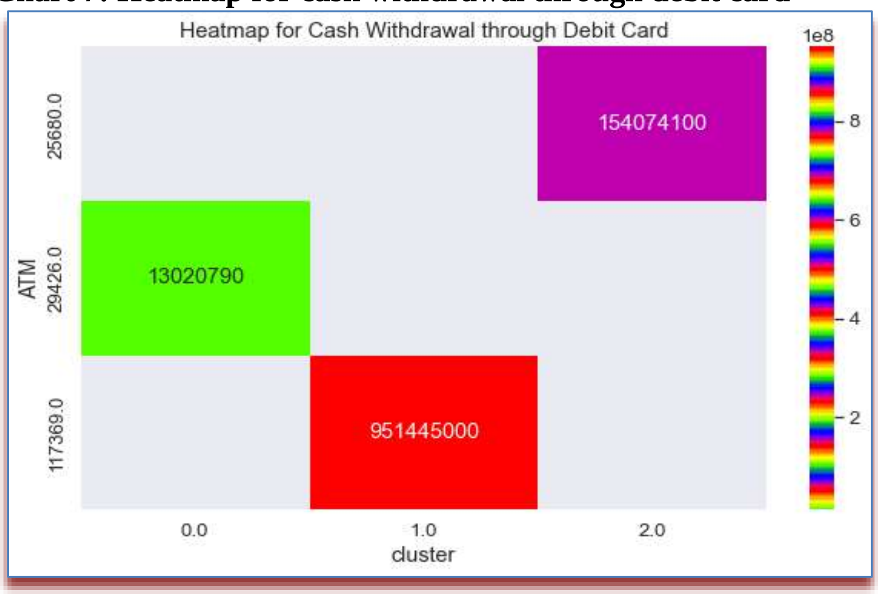
Chart 6: Heatmap for cash withdrawal through credit card



(Source: Created by the Researchers)

As can be seen in the chart above cluster 1 has the highest number of ATMs followed by cluster 0 and cluster 1 respectively. The value inside the boxes in the chart above shows the total value of transactions of cash withdrawal through credit cards for each cluster. As can be seen in the chart above the total value of withdrawals through credit cards is highest in cluster 1 and lowest in cluster 0. However, as indicated by the plotting of boxes within the Heatmap it is is interesting to note that though Cluster 2 has less overall value compared to cluster 1 for withdrawals through credit card, it has transactions with higher values, even though there are fewer of them compared to Cluster 1. Similarly, though cluster 0 has the lowest overall value, , it has transactions with higher values compared to cluster 1.

Chart 7: Heatmap for cash withdrawal through debit card



(Source: Created by the Researchers)
As can be seen in the chart above cluster 1 has the highest number of ATMs followed by cluster 0 and cluster 2 respectively. The value inside the boxes in the chart above shows the total value of transactions of cash withdrawal through debit cards for each cluster. As can be seen in the chart above the total value of withdrawals through debit cards is not only just highest in cluster 1 but also much more compared to cluster 0 and cluster 2. However, as indicated by the plotting of boxes within the Heatmap it is interesting to note that though Cluster 2 has less overall value compared to cluster 1 for withdrawals through debit card, it has transactions with higher values, even though there are fewer of them compared to Cluster 1. Similarly, though cluster 0 has the lowest overall value, , it has transactions with higher values compared to cluster 1.

Optimization Model: To achieve the goal of a developing model to optimize ATM allocation across clusters while maintaining current withdrawal efficiency levels a linear programming framework was used. The model's aim was to minimize the number of ATMs required, ensuring that the average withdrawal per ATM remained constant. The following are objective and constraints sets for the model

- **Objective**: To minimize the total number of ATMs required across all clusters.

Minimize $Z = \sum_{i=1}^{n} xi$
where xi represents the number of ATMs in cluster i

- **Constraints**:
    1. Total withdrawals in each cluster were required to remain constant or marginally reduced (±1%) to maintain transaction efficiency.

   $Wi \geq xi \times avg\_withdrawal\_per\_ATM$

    2. The number of ATMs in each cluster could not exceed the current allocation.

The optimization problem was solved using a numerical optimization technique that identifies the optimal solution while adhering to the constraints. This was implemented using Python's linear programming tools. The results provide the optimal allocation of ATMs for each cluster, ensuring operational efficiency without compromising withdrawal levels.
The following outcome shows the optimal number of ATMs per cluster based on the model developed above.

Table 4:	Optimal Number of ATMs per Cluster
Cluster 0	4175 ATMs
Cluster 1	117369 ATMs
Cluster 2	25680 ATMs
Total Reduced ATMs	25251

(Based on the optimization model developed by the Researcher)

Conclusion and Managerial Implication:
The analysis done in the study through clustering enables banks to introspect their performance within their own clusters and in comparison to other clusters which ultimately aiming for calibrating strategies for better profitability through better resource optimization. The clustering done in the study can help in optimizing resource allocation. For example, if several banks in a given cluster experience high withdrawal volumes at certain times, they can collaborate to ensure adequate cash availability without each bank needing to maintain excess reserves. Clustering done in the above study can also aid in risk management as banks in the same cluster may face similar risks related to cash management, security, or operational challenges so by sharing insights and best practices, they can collectively mitigate these risks more effectively.
It can be concluded that when more cash is withdrawn from these ATMs, especially through debit or credit cards, it indicates that more people are using the bank's services. This increased usage is a positive sign for the bank because it means their services are in demand.
From a resource optimization perspective, when more people use the ATMs to withdraw cash, the bank can better plan its cash management. It can ensure that the ATMs are stocked with enough cash to meet customer demands, reducing the chances of running out of cash and inconveniencing customers. Additionally, the bank can analyze withdrawal patterns to optimize the placement of ATMs and potentially reduce operational costs.
Optimized ATM networks enhance the efficiency of cash withdrawals, ensuring that both debit and credit card users have consistent and seamless access to financial services. Moreover, reducing redundant ATMs helps minimize carbon footprints, aligning with India’s sustainability goals.

Through such measures, the banking sector can play a pivotal role in fostering economic inclusion and empowerment, ensuring that the financial ecosystem becomes a robust pillar supporting the broader objectives of the Viksit Bharat Mission.

Bibliography:

- Gramatiy ,N, Volobueva, Y; Syurpita ,S (2021) The dynamics of the banking sector of the national economy. SEPS Journal ,2021
- Khan, M.A (2010). An Empirical Study of Automated Teller Machine Service Quality and Customer Satisfaction in Pakistani Banks. European J
- Kumbhar, V. M. (2011). Customers satisfaction in ATM services: An empirical evidences from public and private sector banks in India. Management Research and Practice, 3(2): 24-35.
- Kumbhar. (2011). Determinants of Customer's Satisfaction in ATM Services Setting: Empirical Evidences from India. Prajnan, XL (2). Joseph, M. and Stone, G. (2003). An empirical evaluation of US bank customer perceptions of the impact of technology on service delivery in the banking sector. International Journal of Retail & Distribution Management, 31 (4): 190-202.
- Mouthiho, L. (1992). Customer satisfaction measurements: prolonged satisfaction with ATMs. International Journal of Bank Marketing, 10 (7): 30-7
- Ogbuji, C. N., Onuoha, C. B., and Izogo, E. E. (2012).Analysis of the Negative Effects of the Automated Teller Machine (ATM) as a Channel for Delivering Banking Services in Nigeria. International Journal of Business and Management 7(7). Rao, P.S., Rajasekhar, D., and Ratnam,
- Malik Garima and Bansal Sanjay(2015) Impact of ATM service on customer perception and satisfaction of Indian Banks. The Asian Journal of Technology Management Vol 8.NO 2(2015)96-117.
- N.V. (2013). An empirical study of customers satisfaction in ATM services. International Journal of Management and Research & Business 4: 45-61. Shariq and Tondon. (2012). Technology for Banks in India Challenges. IBA Bulletin, XXIII (3): 23-45.
- Singh, S. (2011). The impact of ATM services on customer satisfaction in Indian banks. International Journal Financial Services Management, 5 (2).