DATAWISE
Intelligence. Insights. Infinity.

Optimization of ATM Network Using Geo-Spatial Data Analysis Techniques

+91 40 40204837

info@mydatawise.com

Datawise
Hyderabad, India

Banking sector in India is witnessing an era of reforms with moves such as demonetization which have pushed digitalization in banking sector. With consistent thrust on making the economy cashless, digital modes of banking such as ATMs, Internet Banking, and Unified Payment Interfaces are gaining significance rapidly.

With the increasing card usage in all sections of society, banks have to spend significant amount of resources on setting up ATM networks across geographies. It is very crucial to understand the existing network of ATMs as well as the expected ATM utilization and impact on network efficiency before adding new ATMs to the existing network.

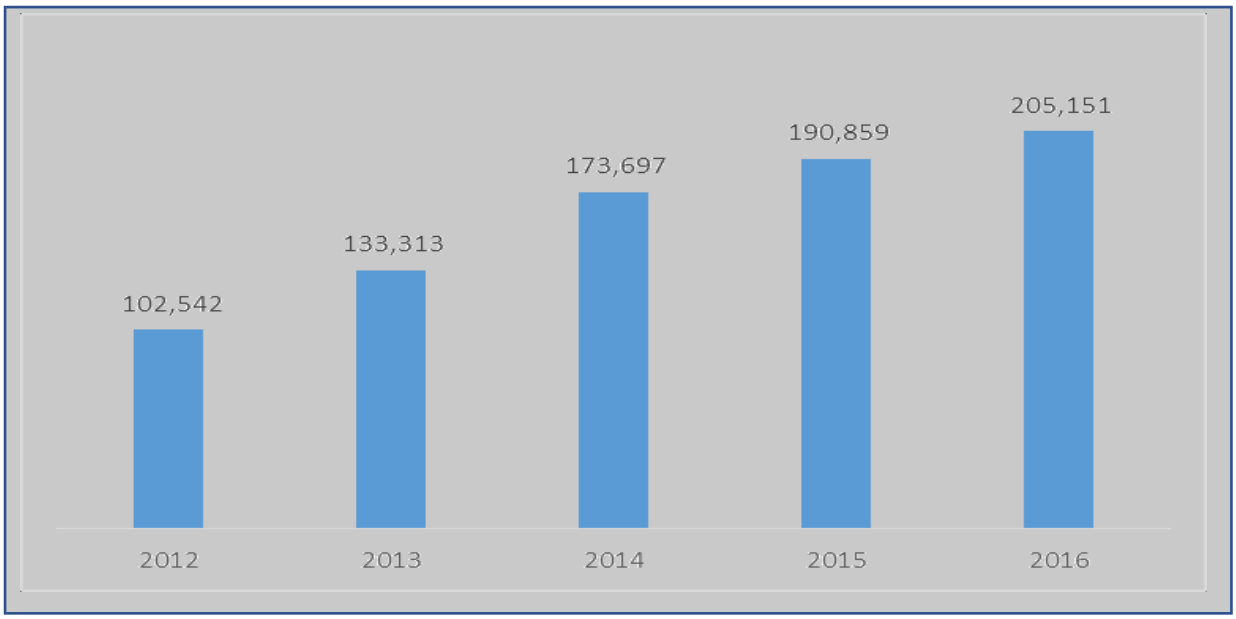
This paper aims at devising a methodology to establish an optimum ATM network using geo- spatial analysis, assisting in establishing new ATMs based on the blind spots in current system as well as predicting the expected usability of proposed ATMs. The model can be further refined to identify the expected usage patterns, customer footfall, and propensity to purchase or usage. Geo spatial layer on top of exiting network of ATMs mapped to the population levels increases the usability and effectiveness of the proposed

BACKGROUND

With increasing focus on Digital India, the banking sector in India is moving towards branchless banking. Direct impact of this step is the decrease in the number of branch visits of customers and a simultaneous increase in the usage of facilities such as Internet Banking, Mobile banking, ATMs, etc.

According to RBI data, the total number of ATMs in India has increased by 121% from 2012 to 2016, whereas the number of bank branches has only increased by 41%. Many private and public banks in India have started strategizing on reducing their branches to reduce their servicing costs to their customers. For example, In May 2016 HSBC took a strategic decision to shut half of its retail branches in India to improve the overall efficiency of the bank.

Figure 1: Total Number of ATMs in India



Despite this increase in the number of ATMs, the per capita availability of ATMs in India is only 18 for every 100,000 of population compared to 44 for every 100,000 of population globally.

As of March-2016, average number of transactions per ATM in India is 3,495 per month. The following table shows the banks with highest and lowest number of transactions per ATM per month:

Table 1 Top five and bottom five banks with respect to the number of transactions per month at ATMs

Bank name	Number of transactions per month
Top 5 banks	
BANDHAN BANK	15,643
STATE BANK OF TRAVANCORE	8,179
STATE BANK OF HYDERABAD	7,607
INDIAN BANK	6,446
STANDARD CHARTERED BANK LTD	6,114
Bottom 5 banks	
PUNJAB AND SIND BANK	687
RATNAKAR BANK LIMITED	724
THE LAXMI VILAS BANK LTD	760
DHANALAKSHMI BANK LTD	1,090
INDUSIND BANK LTD	1,163

Although the number of ATMs is increasing rapidly, it should be noted that there is significant difference between the utilization of ATMs of various banks. While the average amount spent per ATM in 2016 is INR 11.14 Million, the lowest and highest utilizations vary between INR 2.35 Million per ATM to INR 34.33 Million per ATM. Due to this transformation of customer behavior from traditional physical banking to digital, optimum utilization of existing ATMs and proper positioning of new ATMs is a key strategic priority for all the banks in India.

There are revolutionary steps being taken in the Telecom industry by major players such as Airtel, to identify the most suitable location for setting up a mobile tower so as to optimize the utilization of the tower. Similarly, the positioning of ATMs for banks plays a significant role in underutilization or overutilization of an ATM. But no widely accepted methodology exists in banking industry to set up ATMs or to identify the blind spots ideal for setting up new ATMs.

There are a number of banks which have very high number of ATMs but relatively lesser number of transactions per ATM and vice versa. The major reason behind this mismatch between the number of transactions and the number of ATMs is the inefficient distribution of ATM network across key locations which leads to either underutilization or overutilization.

This paper aims at developing a method to analyze the impact of demographic factors and geographic factors on the ATM network.

ATM NETWORK EFFICIENCY AND GEO-SPATIAL DATA ANALYSIS

Geo Spatial data is the information about any object specified in numerical values based on its geographical location or coordinates. The geographical coordinates used for research purpose are latitude and longitude which are one of the most widely accepted criteria for determining geographical location of an object or a place.

Geo spatial analysis is the collection, visualization, and analysis of historical data to describe the impact of change in location of an object based on geographical coordinates. It is a systematic approach to apply statistical or other analytical techniques to any data which has geographical information or a spatial dimension attached to it. The typical steps involved in analyzing geo-spatial data with latitude and longitude coordinates include formation of groups among data, identification of center points for reference, calculation of distance from reference points, relative mapping of distances, overlay of relevant parameters such as utilization, efficiency, demographic data, etc.

An effective ATM network is one which shall cater to the maximum number of account holders of the bank and should be able to service newer geographies having higher user concentration. Hence calculating ATM efficiency is not as direct as any other usage pattern or efficiency measures can be. A few important parameters to be considered can be the concentration of banking population in an area, concentration of ATMs in an area, number of clusters of ATMs in a city or locality, distance between clusters of distribution of the existing network, etc. Geo-spatial analysis plays a key role for a problem of such nature which requires mapping of locations to distribution and usage.

DATA FOR ANALYSIS

The major data points used for geo-spatial analysis are as follows and have been explained in detail in subsequent paragraphs:

a) Latitude of ATMs

b) Longitude of ATMs

c) Geographical ward wise distribution of Hyderabad based on Census 2011 data

d) Number of Households in each ward

e) Number of households availing banking service

f) Area of each ward

Reserve Bank of India issues the ATM and POS statistics at a monthly frequency which includes the number of ATMs, number of transactions, and number of cards of banks at an aggregated level. This data for August 2016 was used to find out the banks with highest number of ATMs in India.

From the list of banks, the top 2 banks with largest ATM networks, ICICI bank and HDFC bank were selected for the analysis and comparison. The addresses and geographical locations of each ATM of these banks based on Geo-codes were extracted from their respective websites. For extraction of ATMs, each branch was taken as a reference and all ATMs in a radius of 10 km was extracted along with address and geo code. This resulted in a total of approximately 300,000 rows because of overlaps in areas and duplicates. The data was further cleaned to remove duplicates and missing data to arrive at a unique list of ATMs. This list was further cleaned to arrive at a list of ATMs in Hyderabad and Ranga Reddy districts. The ATMs outside Hyderabad city limits were eliminated from this list to arrive at the concise list of ATMs in Hyderabad for both ICICI and HDFC bank. ICICI bank has a total of 620 ATMs and HDFC bank has 321 ATMs in Hyderabad which is a statistically significant sample for conducting analysis and can later be extrapolated to use for any geography.

Clustering as a technique helps in dividing the entire data points of a dataset into smaller number of partitions known as clusters. K-means clustering is a method of clustering which helps to form a number of clusters so as to minimize the distance of each data point of the cluster from the center of cluster. K-means clustering was used for dividing entire list of ATMs into clusters.

The total number of clusters were tested using k-means and an elbow curve, which helped to determine the ideal number of clusters as 3 for the sample under consideration. However, for the sake of political mapping of the city, the number of clusters were standardized to 8. The same set of center points was fed as an input foe k-means clustering of both banks to ensure uniformity in geographical distribution.

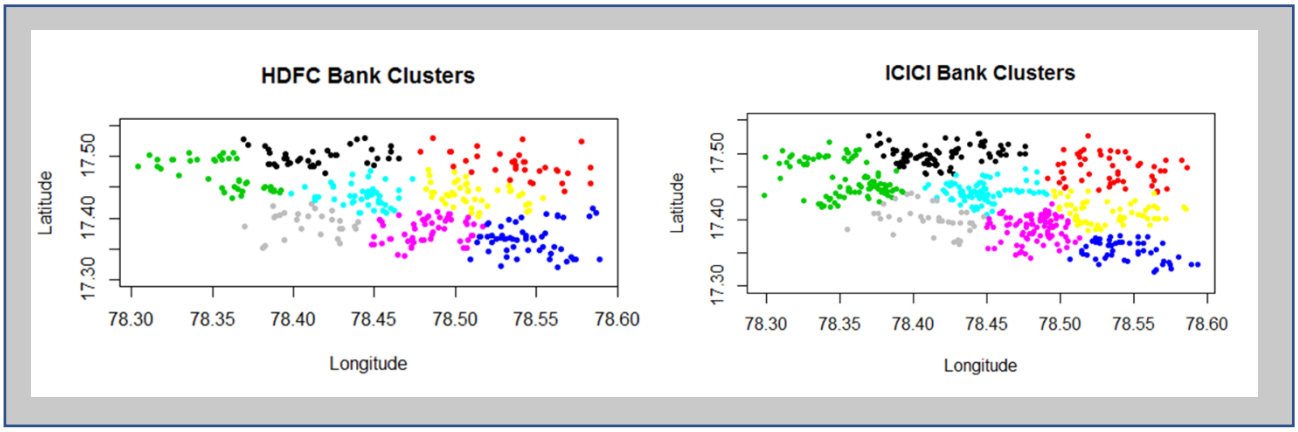
The entire Hyderabad population and area was divided into 22 areas and each area was mapped to 8 clusters. The HH12 data from Census 2012, was used to find out the total number of households as well as the number of households availing banking services in each of the areas.

DETERMINING THE ATM NETWORK EFFICIENCY

The concept of efficiency of ATM networks is based on the assumption that the usability of an ATM network is largely dependent on the expected coverage of area by ATM as well as the population in a particular area. With this in mind, the proposed efficiency calculation approach takes into consideration the area of a cluster, the total number of households in the cluster, total number of households availing banking services in the cluster and the number of ATMs in a cluster.

In the first iteration, the total number of ATMs were distributed into 10 clusters. After the first iteration, the outliers were eliminated so as to eliminate errors in clustering arising out of average radius of each cluster. The number of clusters were also reduced to 8 based on the distribution of ATMs and their average radius. Number of ATMs in each cluster and the area of each cluster was used to find out the ATM concentration in each cluster.

Figure 2 Clusters of HDFC and ICICI Bank ATMs for Hyderabad



The entire area of Hyderabad city was covered with respect to the Tehsils in Hyderabad city limits. All 24 Tehsils were regrouped to 22 and mapped to corresponding clusters based on the nearest cluster to each Tehsil. HH 12 census dataset provides the total number of households and number of households availing banking services in each Tehsil also known as sub-district. This was used to calculate the number of ATMs per household. To make the calculation more insightful, the number of ATMs per household availing banking services was calculated. This is precisely useful since the usability of an ATM largely depends on the banking population and not on the overall population.

The number of banking households per 100 square meters was calculated since a square kilometer area would have reduced the accuracy of prediction. In the final stage, the number of ATMs per banking household per 100 sq. km. was calculated which can be considered as the efficiency of the ATM network in the selected locality or cluster. Comparing the efficiency of ATM network in any cluster with the overall efficiency of ATM network would help to identify the areas of high importance which will add more value to the efficiency of ATM network distribution and also areas having lower efficiencies in terms of the network distribution. The heat map generated combined with these efficiencies can also be used to identify the blind spots in the existing ATM network.

ANALYSIS

A total of 610 ATMs of ICICI Bank and 298 ATMs of HDFC bank in Hyderabad city limits were used for the research purpose. It can be observed that though there is a significant difference in the number of ATMs of both the banks, the distribution of ATMs geography wise is almost similar.

The total geography of Hyderabad city has been divided into 8 clusters for granular level analysis. 22 sub-districts of Hyderabad have been further mapped to the 8 clusters. Following table helps in identifying the clusters, sub-districts associated with each cluster, the total number of household at level and the number of households availing banking services at each level.

Table 2 List of all clusters and associated districts with demographical data for Hyderabad

Sr. No.	Sub-district	Cluster number	Total number of households in sub-district	Households availing banking services in sub-district	Total number of households in cluster	Households availing banking services in cluster
1	Balanagar	1	141,412	96,633	141,412	96,633
2	Golconda	2	46,360	25,909	181,058	101,356
3	Nampally	2	37,789	22,223		
4	Asifnagar	2	96,909	53,224		
5	Malkajgiri	3	103,209	77,193	196,929	138,930
6	Maredpalle	3	45,622	31,471		
7	Tirumalagiri	3	48,098	30,266		
8	Hayathnagar	4	55,782	34,584	121,365	62,262
9	Bandlaguda	4	65,583	27,678		
10	Ameerpet	5	16,370	12,947	48,955	34,562
11	Secunderabad	5	32,585	21,615		
12	Saroornagar	6	134,926	89,135	341,760	193,295
13	Saidabad	6	74,080	49,649		
14	Bahadurpura	6	87,453	35,117		
15	Charminar	6	45,301	19,3		
16	Yelal	7	9,967	6,415		

Sr. No.	Sub-district	Cluster number	Total number of households in sub-district	Households availing banking services in sub-district	Total number of households in cluster	Households availing banking services in cluster
17	Himayathnagar	7	25,222	18,738	257,520	170,342
18	Shaikpet	7	63,878	37,854		
19	Khairatabad	7	70,896	48,223		
20	Musheerabad	7	87,557	59,112		
21	Amberpet	8	37,809	26,177	138,220	94,862
22	Uppal	8	100,401	68,685		
Total			1,427,209	892,242	1,427,209	892,242

ICICI bank has 35% more ATMs as compared to HDFC bank in Hyderabad. ATM density can be calculated as the number of ATMs per square km in a particular geography. It represents how dense the network of ATMs is and also gives an idea about concentration of ATMs.

The following tables represent the distribution of ATM networks of both banks cluster-wise and the ATM density in each cluster.

Table 3 Distribution of ATM network for HDFC bank

Bank Name	Cluster No.	Sub-districts	Average distance from centre in Km	Count of ATMs	Area of cluster in Sq. Km	ATM Density
HDFC	1	Balanagar	2.88	46	26.01	1.77
HDFC Bank	2	Golconda, Nampally, Asifnagar	3.36	39	34.62	1.13
HDFC Bank	3	Malkajgiri, Marredpally, Tirumalagiri	3.6	48	40.47	1.19
HDFC Bank	4	Hayathnagar, Bandlaguda	3.15	63	33.8	1.86
HDFC Bank	5	Ameerpet, Secunderabad	2.25	50	14.63	3.42
HDFC Bank	6	Saroornagar, Saidabad, Bahadurpura, Charminar	2.73	64	21.21	3.02

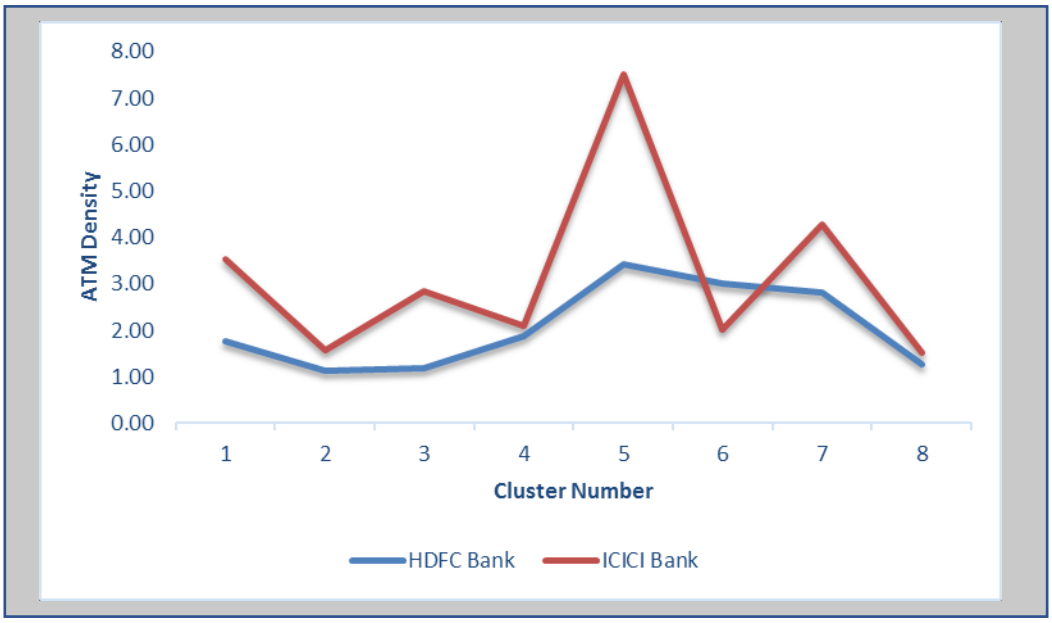
Bank Name	Cluster No.	Sub-districts	Average distance from centre in Km	Count of ATMs	Area of cluster in Sq. Km	ATM Density
HDFC Bank	7	Yelal, Himayathnagar, Shaikpet, Khairatabad, Musheerabad	2.7	58	20.51	2.83
HDFC	8	Amberpet, Uppal	2.75	31	24.33	1.27
Total/ Average			**2.93**	**399**	**26.95**	**2.06**

Table 4 Distribution of ATM network for ICICI bank

Bank Name	Cluster No.	Sub-districts	Average distance from centre in Km	Count of ATMs	Area of cluster in Sq. Km	ATM Density
ICICI Bank	1	Balanagar	2.88	92	26.01	3.54
ICICI Bank	2	Golconda, Nampally, Asignagar	3.28	54	34.62	1.56
ICICI Bank	3	Malkajgiri, Marredpally, Tirumalagiri	3.58	115	40.47	2.84
ICICI Bank	4	Hayathnagar, Bandlaguda	3.4	71	33.8	2.1
ICICI Bank	5	Ameerpet, Secunderabad	2.06	110	14.63	7.52
ICICI Bank	6	Saroornagar, Saidabad, Bahadurpura, Charminar	2.46	49	21.21	2.03
ICICI Bank	7	Shaikpet, Khairatabad, Musheerabad	2.4	88	20.51	4.29
ICICI Bank	8	Amberpet, Uppal	2.82	37	24.33	1.52
Total/ Average			**2.86**	**610**	**26.95**	**3.17**

The average ATM density for ICICI bank is higher by 35% which is exactly the same percentage difference between number of ATMs of ICICI and HDFC bank. Cluster number 5, which includes sub-districts Ameerpet and Secunderabad have the highest ATM density for both the banks. This can be attributed to the fact that these areas are one of the business and education hubs of Hyderabad, so the resident population of this cluster is lower as compared to the actual number people in the cluster.

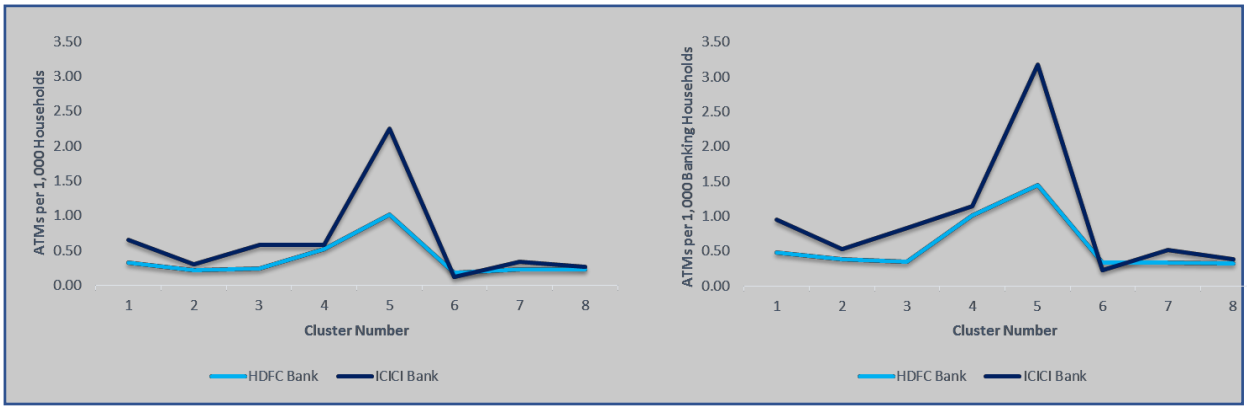
Figure 3 Cluster wise ATM Density of HDFC and ICICI Bank



It can be observed that the average radius of a cluster and the ATM density have inverse correlation except a few exceptions such as cluster number 3 for ICICI bank which has higher radius as well higher ATM density.

The above analysis given an idea about the distribution of ATM networks of both the banks with only one aspect which is the geographical aspect. It is important to analyze the impact of overlaying demographical data over the geographical data. We will use the total number of households as well as the number of households availing banking services for the further calculations. It is very important to consider the number of banking households in an area as the total population in an area can be higher, for example, in slum areas, but the number of people actually utilizing the banking services or ATM services can be very low.

Figure 4 Comparison of number of ATMs per 1,000 Households and per 1,000 banking households



It is evident from the graphs that the number of ATMs per 1,000 banking households is high for both the banks in all the clusters as compared to the number of ATMs per 1,000 households. The highest increase in number of ATMs per 1,000 households to 1,000 banking households for ICICI is 0.94 for cluster 5, whereas for HDFC, it is 0.43 for cluster number 4. Cluster number 4 has highest increase because of the fact that it is the business and educational hub. The average increase in number of ATMs per 1,000 households to per 1,000 banking households is 57.5% for HDFC bank and 52.2% for ICICI bank. This re-emphasizes the assumption of ATM network distribution that it should be dependent upon the banking population and not the general population.

IMPLICATIONS

There has been some work which has been done on network optimization that has considered methods such as time series, and regression. However, the use of geo-spatial analytics for network optimization has not been explored because of lack of data as well as methods to be able to do so.

This alternate model using geo-spatial analysis for determining ATM network efficiency will be useful not only for banks under study, but can also be used across banking industry as a whole, to improve reach. The recent move of demonetization undertaken by the Indian government is expected to give a significant boost to the usage of debit or credit cards. With increasing number of cards, requirement of number of ATMs is also expected to increase steeply. With expected steep demand for ATMs, models for optimization of ATM network distribution will play a pivotal role for every bank in short term as well as long term.

Easily interpretable visualizations of existing network along with identification of blind spots will help the bank in taking strategic decisions more effectively and in much lesser time. This can be beneficial to increase the reach of existing network, effective positioning of new ATMs and will ultimately lead to improved operational and financial performance of the bank through optimum utilization of ATM network.


The accuracy of the presented methodology is based on publicly available data points. This can further be made more accurate and precise by using exact pin pointed location of every ATM. Data received directly from banks will be more accurate and updated and can magnify the effectiveness of implications by many folds.
Also, by replacing the population data with the actual customer base of a bank across various geographies, the ATM distribution can be designed for any bank and can yield more

About DATAWISE®

DATAWISE® offers a suite of products and solutions suited to the needs of various situations and industries. Solutions provided for one customer are not necessarily suitable for others, and readers are advised to use their own judgment regarding the suitability of these solutions to their business needs.

DATAWISE®'s business analysis services support the full spectrum of clients' needs with services directed mainly at helping companies discover opportunities for improvement through use of analytical capabilities. We offer analytical services in the following areas:

Strategic Analytics: Alignment of strategic intent with actual work, requiring strategic analytics to answer key decision support questions such as whether to enter into a new segment of business or not, whether to reach new customers or not, and other go, no-go decisions.

Behavioral Analytics: Assistance in determining the 'why' and 'how' of a customer behavior (rather than the 'what') in order to ensure that marketing plans yield the desired results through capturing customer events and actions over time and using these stored interactions to determine typical behavior and deviations from that behavior.

Tactical Analytics: Tactical analytics models that we deploy are typically short-term in nature, and are focused on answering immediate questions rather than aligning to a longer-term goal.

Predictive Analytics: We created complex multi-dimensional models that collate data generated from several interaction points to create models that enable the prediction of future events to help identify of both risks and opportunities.

DATAWISE® has also developed proprietary analytics models OPTLIOX™, CREST™, Infinity™ and DATTAB™, catering to specific customer needs.

• Hyderabad • Delhi • Mumbai • Bangalore • Jaipur • USA

www.mydatawise.com mail at info@mydatawise.com