



**“ROLE OF BIG DATA ANALYTICS TO GAIN PRODUCT AND
MARKETING EDGE”**

Ramalakshmi.V¹.

Associate Professor

Krupanidhi Degree College

#12/1, Chikkabellandur, Carmelaram Post,
Varthur Hobli, Bangalore – 560035, India.

Madhur Pahwa²,

CEO,

Simple Lecture,
Bangalore, India.

ABSTRACT

This paper presents the usage of big data to gain product and marketing edge with special reference to retail sector. Big data is not been used everywhere in every sector but it plays a vital role now a days in many business segment. In retail sector big data can play an important role.

METHODOLOGY – *In this paper quantitative research approach is used. The quantitative information is provided by survey, which can be analyzed statistically. Every respondent with similar background are asked the same series of questions are the main phenomena of quantitative research. The large numbers of questionnaire involves with this approach and it is structured. The researcher investigates for verbal or written response of his/her questions or statements in this study. For assembling data about individual preference, prediction, past events, and private behaviors, survey is the most efficient method. The utmost strength of a survey method is inventiveness. It is a convenient approach to gather information and the most cost-effective way in a lot of other situations.*

***FINDINGS** – Big Data analytic user get better insight in business as predictive analytics provides those opportunities to identify their potential customers, better pricing and segmentation can be done, consumer loyalty is increased. One of the major issues regarding analytics is retailers feel there are retailers may feel as if there are too many analytical options to explore from so the retailers need to selectively target investments in analytics based on their strategies and industry positions. User's interest in Hadoop Distributed file system is high although they are rarely used by organizations. OLAP is the tool used by most of the organizations today and the trend would continue as OLAP makes multidimensional data exploration fast and intuitive. Users are not sure of the outcome of the big data analytics in their organization. It is so mainly because most vendors have barely begun to prove their software's effectiveness, so it's still early days for this market and even users are gradually getting to know the benefits big data analytics can provide to their business.*

KEY WORDS:

Big Data, Business intelligence, Cloud computing, Online Analytic Process, Retail sector

INTRODUCTION:

Big data is a buzzword, or catch-phrase, used to describe a massive volume of both structured and unstructured data that is so large that it's difficult to process using traditional database and software techniques. In most enterprise scenarios the data is too big or it moves too fast or it exceeds current processing capacity.

While the term may seem to reference the volume of data, that isn't always the case. The term big data, especially when used by vendors, may refer to the technology (Which includes tools and processes) that an organization requires handling the large amounts of data and storage facilities. The term big data is believed to have originated with Web search companies who had to query very large distributed aggregations of loosely-structured data.

Big data is a popular term used to describe the exponential growth and availability of data, both structured and unstructured. And big data may be as important to business and society – as the Internet has become. Why? More data may lead to more accurate analyses. More accurate analyses may lead to more confident decision making. And better decisions can mean greater operational efficiencies, cost reductions and reduced risk.

STATEMENT OF THE PROBLEM

The size of the data is growing day by day with the exponential growth of the enterprises; today enterprises are exploring data to get a better insight. For decision making in an organizations, there is a continuous need for processing and analysing large datasets. Being in a digital era lot of digital data like shifting thoughts and opinions shared on Twitter, social relationships, purchasing habits, photos, videos etc. Data is generated from the many sources in the form of structured as well as unstructured form. The data thus exceeds the capacity and capabilities of conventional storage, reporting and analytics systems, it demands new problem- solving approaches. While the organizations are collecting more data than ever the data tends to reside in silos that makes it difficult to integrate. Another aspect of big data is the cost to both store and process these massive amounts of data. With the advent of powerful computing, advanced database technologies, wireless data, mobility and social networking, it is now possible to bring together and process big data in many beneficial ways. The retail in industry in specific is showing an increasing use of data from various sources about customers, their shopping behavior, current trends and conversations on social media. The study is carried out in retail industry because of the huge data being generated and used by the firms for analysis. Firms have now access to the various information tools, but the path between the technologies and the way companies can gain competitive advantage is not explored much.

OBJECTIVES OF THE STUDY

Primary objective

The main objective of the study is to identify the prospects of Big Data Analytics in retail Sector.

Secondary objectives: The secondary objectives of the study are to:

- To understand the initial applications, kinds of data, and approaches that enterprises are employing for their Big Data initiatives
- To identify the drivers for Big Data Analytics in retail sector
- To identify the technologies used for handling Big Data in retail sector
- To analyze the benefits of Big Data analytics in retail sector
- To identify the challenges in increasing the adoption of Big Data Analytics in retail sector

Scope of the Study

Subject scope: Big Data has turned out to be a predictable part of the retail industry. A lot of problems may occur while using big data in developing countries like India.

The scope of this study is to spot the problem and prospects of big data in the retail industry. To identify the type of data being used, problems faced while implementing and using the big data in retail in a developing country like India.

Geographic Scope: For this study the data was collected from various retail organizations in India.

Important Terms/ Operational Definitions

Big Data

Big data is a popular term used to describe the exponential growth and availability of data, both structured and unstructured. And big data may be as important to business – and society – as the Internet has become.

Big Data Analytics

Big data analytics refers to the process of collecting, organizing and analyzing large sets of data to discover patterns and other useful information. Not only will big data analytics help you to understand the information contained within the data, but it will also help identify the data that is most important to the business and future business decisions. Big data analysts basically want the knowledge that comes from analyzing the data.

In computing, online analytical processing, or OLAP is an approach to answering multi-dimensional analytical (MDA) queries swiftly. OLAP is part of the broader category of business intelligence, which also encompasses relational database, report writing and data mining. Typical applications of OLAP include business reporting for sales, marketing, and management reporting, business process management (BPM), budgeting and forecasting, financial reporting and similar areas, with new applications coming up, such as agriculture. The term OLAP was created as a slight modification of the traditional database term Online Transaction Processing ("OLTP").

MPP

The speed, scale and rapid growth of today's information exceed what even the largest of supercomputers could ever handle alone, which is why Massively Parallel Processing (MPP) systems are becoming the norm. Vertica delivers a simple, yet highly robust and scalable MPP

solution for the masses with linear scaling and native high availability on industry standard hardware.

E-tailers

The sale of goods and services through the Internet, Electronic retailing, or e-tailing, can include business-to-business and business-to-consumer sales. E-tailing revenue can come from the sale of products and services, through subscriptions to website content, or through advertising. It is a play on the words retail and e-commerce.

Methodology

Empirical research is the way of gaining knowledge by means of direct and indirect observation or experience. Evidence was collected and analyzed quantitatively and qualitatively.

Sampling; Simple random sampling method is used for the study.

Population

Type of industry selected for the survey was IT and Retail sectors.

Sample Size

The sample for the study would be limited to 90 respondents.

Sampling Unit: For this study the respondents were from various retail organizations in India to understand problems and prospects in retail sector.

Type of Sampling

It is of vital importance that the sample selection should be representative of the population. Thus the results obtained from the sample can be generalized to the larger population. Simple random sampling method was used for the study.

Research Instruments Used

The research instrument used for the study was questionnaire. Two point scales, five point scale and multiple option questions were used for collecting the data to identify the problems faced and prospects of the users using big data analytics.

Sources of Secondary Data

The secondary data for this was obtained from company magazines and brochures, website, newspapers, internet, text books, reports and other promotional materials. In addition, for gathering primary data from the relevant system users, a survey was conducted through structured questionnaires.

The data for this study is collected from statements about privacy policy, acceptable use policy, terms of use and service level agreements available from the websites providing information on Big Data Analytics.

Plan of Analysis

The data collected from various sources will be tabulated and represented using the percentage and ranking method. Graphical tools like bar diagrams, pie charts etc., will be used to illustrate the tabulated data pictorially. Inferences and interpretation of the data will be done on the basis of tabulated data. The research process involved the following steps.

- A literature review was undertaken
- Questionnaire was constructed and piloted.
- The population and sampling procedure were established and methods of data collection and analysis were determined.

Introduction to Big Data

Big data is a buzzword, or catch-phrase, used to describe a massive volume of both structured and unstructured data that is so large that it's difficult to process using traditional database and software techniques. In most enterprise scenarios the data is too big or it moves too fast or it exceeds current processing capacity.

While the term may seem to reference the volume of data, that isn't always the case. The term big data, especially when used by vendors, may refer to the technology (which includes tools and processes) that an organization requires handling the large amounts of data and storage facilities. The term big data is believed to have originated with Web search companies who had to query very large distributed aggregations of loosely-structured data.

Big data is a popular term used to describe the exponential growth and availability of data, both structured and unstructured. And big data may be as important to business – and society – as the Internet has become. Why? More data may lead to more accurate analyses.

More accurate analyses may lead to more confident decision making. And better decisions can mean greater operational efficiencies, cost reductions and reduced risk.

Big Data Defined

As far back as 2001, industry analyst Doug Laney (currently with Gartner) articulated the now mainstream definition of big data as the three Vs: volume, velocity and variety **Volume:** Many factors contribute to the increase in data volume. Transaction- based data stored through the years. In the past, excessive data volume was a storage issue. But with decreasing storage costs,

other issues emerge, including how to determine relevance within large data volumes and how to use analytics to create value from relevant data.

Velocity: Data is streaming in at unprecedented speed and must be dealt with in a timely manner. RFID tags, sensors and smart metering are driving the need to deal with torrents of data in near-real time.

Reacting quickly enough to deal with data velocity is a challenge for most organizations.

Variety: Data today comes in all types of formats. There are structured, numeric data in traditional databases. Information created from line-of-business applications. Email, video, audio, stock ticker data and financial transactions are unstructured text documents

Managing, merging and governing different varieties of data is something many organizations still grapple with. At SAS, they consider two additional dimensions when thinking about big data:

Variability: In addition to the increasing velocities and varieties of data, data flows can be highly inconsistent with periodic peaks. Is something trending in social media? Daily, seasonal and event-triggered peak data loads can be challenging to manage. Even more so with unstructured data involved.

Complexity: Today's data comes from multiple sources. And it is still an undertaking to link, match, cleanse and transform data across systems.

However, it is necessary to connect and correlate relationships, hierarchies and multiple data linkages or your data can quickly spiral out of control.

ANALYSIS

The primary data was collected from sample of 90 respondents from retail industry across India, chosen by simple random sampling method. In order to collect the necessary data, online questionnaire were circulated to identify the prospects of Big Data Analytics in Retail industry.

Table 1 showing details about the use of Big Data Analytics

Access to relevant, precise and timely data			
No	Response	No. of Respondents	Percentage
1	Minimal	0	0
2	Less than adequate	45	50
3	Adequate	27	30

4	More than Adequate	9	10
5	World Class	9	10
	Total	90	100

Source: Field survey

Interpretation: From the above table it can be analyzed that majority of the respondents (50%) felt that big data provides less than adequate sources to access relevant, precise and timely data where as 30% says it is an adequate source, 10% believe that it is a more than adequate source and 10% says it's a world class source.

This could be because there are many vendors as well as many types of analytical options such as OLAP, predictive analytics, Business Intelligence, decision analytics etc and choosing the right data and analytical approach is quite essential to get the relevant, precise and timely data for any organization using big data analytics to gain competitive advantage.

Table 2 showing data sources providing valuable insight to organization

S.no.	Response	No of respondents	Percentage
1	Document, Images, Video , E-mail	18	20
2	Click stream	27	30
3	Social Media data	18	20
4	Data from sensor devices	18	20
5	Phone conversation	9	10
6	Weblogs	0	0
	Total	90	100

Source: Field survey

Interpretation: Majority of the respondents feel Click-stream data provides valuable insight to the organization as most of the retail stores have online store as well, as a result a lot of click stream data is generated. It is used determine the effectiveness of the site as a channel-to-market by quantifying the user's behavior while on the Web site. It is used to keep track of what pages the user lingers on, what the user puts in or takes out of their shopping cart, and what items the user purchases.

Table 3 showing details about biggest adoption and implementation barrier to BDA

SL.NO	BARRIERS	NO. OF RESPONDENTS	PERCENTAGE
1	Identifying what data to collect	0	0
2	Implementing the right solution to accurately analyze	36	40
3	Capacity to capture and analyze data from all functional units	18	20
4	Lack of trained staff	18	20
5	Lack of funding	0	0
6	Current database lacks in- db analytic	9	10
7	Others	9	10
	Total	90	100

Interpretation: One of the major issues regarding analytics as felt by retailers is that are too many analytical options as well as vendors to explore from so implementing the right solution is the biggest challenge. Retail industry using BDA need to selectively target investments in analytics based on their strategies and industry positions to gain the kind of advantage they are looking for in their business

Table 4 Showing details about usage of Technology by the organization to handle Big Data

SL.NO.	TECHNOLOGY	NO OF RESPONDENTS	PERCENTAGE
1	Hadoop	0	0
2	In- Memory computing	18	20
3	Text Mining	0	0
4	Predictive analytics	18	20
5	No SQL	0	0
6	OLAP tools	36	40
7	Map Reduce	0	0
8	Others	18	20
	Total	90	100

Source: Field survey

Interpretation: From the analysis it is evident that user's interest in Hadoop Distributed file system is high although they are rarely used by organizations. OLAP is the tool used by most of

the organizations today and the trend would continue as OLAP makes multidimensional data exploration fast and intuitive.

Table 5 Showing Competitive advantage of Big Data Analytics in retail industry

S.No	Particulars	No. of Respondents
1	Predicts sales trends	83
2	Deliver assortments based on customer insight	86
3	Matching inventory and customer requirements	76
4	Create and execute strategies to address potential outcomes	73
5	By determining the optimal response times to encourage a sale	87
6	By developing precise, targeted marketing campaigns	68
7	Providing highly personalized shopping experiences	67
8	Anticipate customer purchasing behaviour	87
	Total No. of Respondents	90

Source: Field survey

Interpretation: Big data analytics has its competitive advantage in retail sector in many ways. But from the study it has been analyzed it is highly effective to Anticipate customer purchasing behaviour and to determining the optimal response times to encourage a sale which can be easily done with the help of click stream analytics.

Table 6 Showing Key drivers of analytics in retail sector

S.No	Particular	No. of Respondents	Percentage
1	Growing volumes and complexity of customer data	36	40
2	Need for faster and more relevant information	36	40
3	Need to maximize the value of customer data	0	0
4	Increasing competition	0	0

5	To gain deeper insights on customer behavior	18	20
6	Move towards creating personalized shopping experience	0	0
Total		90	100

Source: Field survey

Interpretation: From the study it is seen that Key drivers of Big Data analytics in retail sector are Growing volumes and complexity of customer data and Need for faster and more relevant information because faster decisions help the company gain major revenue benefits.

FINDINGS

- Most of the people are aware of big data analytics as there is a vast change in technology most of them are aware of what benefits their organization can get using this latest technology and to get better and efficient results in retail sector and this can be seen from the survey. The level of penetration of usage of Big Data is small, deeper it goes in the organization more. BDA is used to attain better and efficient result
- Big data can be changed to world class or more adequate data by the use of more advances and efficient technology
- Click-stream act as the most valuable insight to the organization, other source can also act as a valuable insight if used in a better way.
- Most of the online retailers have suffered from a scalability crisis where big data was more of a curse than a blessing. There's no doubt that big data presents technical challenges due to its volume, variety, and velocity. Data volume alone is a big concern for some organizations. Though there are some problems associated with the detailed analyses of big data, online retailers can discover new facts about their customers, markets, partners, costs, and operations then use that information for business advantage. Big Data analytic user get better insight in business as predictive analytics provides them opportunities to identify their potential customers, better pricing and segmentation can be done, consumer loyalty is increased
- One of the major issues regarding analytics is retailers feel there are retailers may feel as if there are too many analytical options to explore from so the retailers need to selectively target investments in analytics based on their strategies and industry positions. User's interest in Hadoop Distributed file system is high although they are rarely used by organizations. OLAP is

the tool used by most of the organizations today and the trend would continue as OLAP makes multidimensional data exploration fast and intuitive.

- Users are not sure of the outcome of the big data analytics in their organization. It is so mainly because most vendors have barely begun to prove their software's effectiveness, so it's still early days for this market and even users are gradually getting to know the benefits big data analytics can provide to their business.
- Organizations use big data to target the audience the most but it can also be used in a better way to understand customer demand and to change the pricing strategy
- Big data has a big use in the different sector, in retail it is used to track revenue trends and monitor customer behavior. It can also be used in much better way to maintain customer loyalty
- Most of the organizations don't have any proven business case but big data can have many business cases with high return on investment if used in an efficient and productive way
- Most of the organizations use stand alone big data application whereas some use it with ERP or Business Processes, Big data application itself can play a vital role if used in a better way

SUGGESTIONS

- Advanced analytics and big data together can be used together to provide maximum benefits to the retail sector as the big data is a special enterprise asset that merits leverage and advance analytics provides that leverage to the enterprise.
- The barriers of big data analytics must be unknown and understood well in advance such as inadequate staffing or skills, lack of business support etc so that the enterprise can be prepared to take counter steps to deal with these barriers.
- The newest analytics platforms like Hadoop, Map Reduce, no-SQL, Mongo DB, Public and private clouds, SAAS, complex event processing etc should be considered as it would satisfy new requirements for highly diverse data types, querying, outsourcing of analytics and real time analytics.

Need to give priority to in-memory computing as it would increase the query response speed, data visualization for faster decision and it's always better to rely on SQL as it is a natural fit to analytics. Organizations should periodically reevaluate its current portfolio of analytic databases and tools for various reasons such as to handle diverse big data, to satisfy modern requirements and as the technology keeps becoming obsolete and are replaced by newer better ones.

- Retail Organizations should focus on customer centric outcomes as that would enable them to better understand the customer needs, their buying behavior, provides better insights to generate enhanced products, improve brand performance and drive customer loyalty and satisfaction.
- Retailers should focus on gaining the specific skills needed within the organization, especially those that will increase the organizations ability to analyze unstructured data and visualize its analysis to make it more usable to business executives.
- Retail leaders should invest in centralized databases and focusing on data hygiene and analytics which will help them gain better insight into their customers that wasn't possible even a few years ago. Online retailers can use predictive analytics to predict the revenue from certain products in the next quarter. With such information available a merchant can better manage its inventory costs and avoid key out-of-stock products.
- Although there obvious benefits of Big Data Analytics retailers can only get the those benefits if they gain executive support, educate people in the organization and integrate the right technology. Once internal resources are up to date including human knowledge and technology assets –Big data analytics possibilities are vast. Retail enterprises should identify a small business unit to conduct a pilot run so that the enterprise can give a clear value proposition.

CONCLUSION

The main aim of the study has been to identify the prospects of Big Data Analytics in retail industry. In these times of economic uncertainty and decreasing margins, retailers must improve their approach to driving traffic and sales in to their business. With increased mobility shoppers today can easily research competitive offerings. And with increased social media adoption they are also likely to report their experience, both positive and negative. Big Data Analytics can be right solution to tackle the issues mentioned and to increase the productivity of the business.

This study was conducted to understand the adoption rates, implementation and technology barriers for adoption, to understand the initial applications, kinds of data, and approaches that enterprises are employing for their Big Data initiatives, Identify the drivers for Big Data Analytics in retail sector, Identify the technologies used for handling Big Data in retail sector and Analyze the benefits of Big Data analytics in retail sector. From the study it can be concluded that retail sector need to embrace the power of big data analytics and implement solutions using a complete strategy. In the rush to adopt big data analytics many organizations are using big data analytics for the first time and hence seems to be confused with its offerings. By enabling

effective customer segmentation, location based pricing, predictive analytics, bid data solutions provide tremendous opportunities to help retailers gain the competitive advantage desired by them.

REFERENCES

1. Adomavicius, G., and Tuzhilin, A. 2005. "Toward the Next Generation of Recommender Systems: A Survey of the State-of- the-Art and Possible Extensions," IEEE Transactions on Knowledge and Data Engineering
2. Anderson, C. 2004. "The Long Tail," WIRED Magazine, Associated Press. 2012. "Columbia U Plans New Institute for Data Sciences," July
3. Barabási, A. 2003. *Linked: How Everything Is Connected to Everything Else and What it Means for Business, Science, and Everyday Life* , New York: Plume. Connections (21),
4. Bettencourt, L. M.A., Cintrón-Arias, A., Kaiser, D. I., and Castillo-Chávez, C. 2006. "The Power of a Good Idea: Quantitative Modeling of the Spread of Ideas from Epidemiological Models," *Physica A* (364),
5. Bitterer, A. 2011. "Hype Cycle for Business Intelligence," Gartner, Inc., Stamford, CT
6. Blei, D. M. 2012. "Probabilistic Topic Models," *Communications of the ACM* (55:4),
6. Bloomberg Businessweek. 2011. "The Current State of Business Intelligence," V. Batagelj, V., and Mrvar, A. 1998.
7. Pajek: A Program for Large Network Analysis," *Connections* (21), D. Allemang and J. Hendler. *Semantic web for the working ontologist: effective modeling in RDFS and OWL*. Morgan Kaufmann, 2008.
8. *International Journal on Semantic Web and Information Systems*, 5(3):1–22, 2009.
9. J. Hendler. *Broad data: Exploring the emerging web of data*. *Big Data* 1(1), 2013.
10. A. J. Hey, S. Tansley, K. M. Tolle, et al. *The fourth paradigm: data-intensive scientific discovery* Microsoft Research Redmond, WA, 2009.