THE VALUE OF ANALYTICS FOR BIG DATA IN BUSINESS INTELLIGENCE

FAROOK SHAIKH, DR. G.L. KHARADE

KLE SOCIETY'S JAGADGURU GANGHADHAR COLLEGE OF COMMERCE

Abstract

To predict trends, big data analytics uses real-time, raw data. The performance of a business and, eventually, its bottom line may be greatly impacted by the proper handling of this data. Big data analytics has been widely hailed as a revolutionary technological development in both academia and business. Using the right data or datasets has always been essential to a company's success. As the amount of data increases, so does the need for analytics. Successful businesses are getting a competitive edge through the analysis of big data. This study aims to demonstrate the value of big data analytics as a tool for business management.

Key Words: Difference between Traditional Data and Big Data, Organizations Performance, Using of Data Analytics by Business Management, Intelligence and analytics for Business, Framework for Analyzing Large Data Sets, etc.

Introduction

Big data and business analytics are essential for firms looking to gain a competitive advantage in today's market by leveraging various data sources like the Web, online transactions, social media, and more. Recent studies highlight the importance of big data analytics in improving organizational outcomes such as agility, innovation, and competitive performance. Experts in this field provide valuable insights and a competitive edge to organizations, with some considering big data analytics as the "fourth paradigm of science." It is also seen as a new frontier for innovation, competition, and productivity, with high-performing enterprises viewing it as a key differentiator and a path to growth. While information overload is a significant challenge in big data environments, companies that can effectively process and analyze large datasets can derive valuable insights from this wealth of information.

In the current era of rapid business information growth, the ability to access data quickly is crucial. Studies indicate that organizations must pinpoint areas where data-driven insights can be beneficial and undertake data analytics projects to leverage big data effectively. This requires developing the organizational capacity to plan, execute, and allocate resources strategically, as well as aggregating data to derive actionable insights. Executives skilled in knowledge collection, analysis, and utilization are now a key competitive advantage in business. The fundamental characteristics of big data are volume, velocity, and variety, with volume referring to large data quantities, velocity to the speed at which data is generated and shared on platforms like social media, and variety to data in diverse formats. The world produces 2.5 quintillion bytes of data daily, with 90% generated in the last two years alone. Data is expanding exponentially, surpassing the expertise of data analytics professionals. Big data presents three primary concerns for organizations: lack of organization, new opportunities, and cost-effective technologies.

This article highlights the significance of Big Data analytics as a valuable tool in business management. It discusses the concepts of Big Data and analytics, compares traditional data analysis with Big Data analysis methods, and explores the role of Big Data analytics in business management. Additionally, empirical research findings on the importance of applying Big Data in business are presented.

Big data refers to large, complex data sets that cannot be managed by traditional systems. Analyzed using modern tools, big data provides businesses with valuable insights for informed decision-making. It includes technologies for storing, analyzing, and managing data on a large scale to identify patterns and design intelligent solutions. Challenges related to big data include collection, storage, search, sharing, analysis, and visualization. Companies are using big data analytics to uncover new insights and improve business efficiency, gaining a competitive edge. This type of analytics involves processing large volumes of data from various sources to inform decision-making. Big data analytics has made progress in integrating design, behavior, and economic direction, with five key technical areas contributing significantly to analytics and business intelligence.

The text discusses various types of analytics, including analytics, mobile analytics, and network analytics, which are used to process large amounts of data quickly and understand patterns for better trading decisions. Big data analytics involves collecting, processing, cleansing, and analyzing large data sets to help businesses make use of big data. The articles focus on speculation, critical thinking, and the potential for rational opportunities through big data technology.

DIFFERENCE BETWEEN TRADITIONAL DATA AND BIG DATA

Big data analysis involves more than just data, including IT infrastructure and skilled staff, and aims to uncover hidden patterns and trends in large datasets. This process can reveal new relationships and insights that can improve business profitability and efficiency. While the costs associated with specialized software and experts may be offset by these benefits in the long run, there are distinct differences between traditional analytics and big data analysis. Figure 2 provides a summary of these variations.

Traditional Data	Big Data
Traditional data is generated in enterprise level.	Big data is generated outside the enterprise level.
Its volume ranges from Gigabytes to Terabytes.	Its volume ranges from Petabytes to Zettabytes or Exabyte's.
Traditional database system deals with structured data.	Big data system deals with structured, semi-structured, database, and unstructured data.
Traditional data is generated per hour or per day or more.	But big data is generated more frequently mainly per seconds.

Traditional Data	Big Data
Traditional data source is centralized and it is managed in centralized form.	Big data source is distributed and it is managed in distributed form.
Data integration is very easy.	Data integration is very difficult.
Normal system configuration is capable to process traditional data.	High system configuration is required to process big data.
The size of the data is very small.	The size is more than the traditional data size.
Traditional data base tools are required to perform any data base operation.	Special kind of data base tools are required to perform any database schema based operation.
Normal functions can manipulate data.	Special kind of functions can manipulate data.
Its data model is strict schema based and it is static.	Its data model is a flat schema based and it is dynamic.
Traditional data is stable and inter relationship.	Big data is not stable and unknown relationship.
Traditional data is in manageable volume.	Big data is in huge volume which becomes unmanageable.
It is easy to manage and manipulate the data.	It is difficult to manage and manipulate the data.
Its data sources includes ERP transaction data, CRM transaction data, financial data, organizational data, web transaction data etc.	Its data sources includes social media, device data, sensor data, video, images, audio etc.

When analyzing big data, the focus should be on identifying correlations and patterns that indicate something is happening, rather than explaining why it is happening. This means reversing the traditional method of using hypotheses and searching for arguments to verify them. Discovering unexpected correlations can stimulate assumptions. Big data analytics is based on processing real-time data from corporate environments, ensuring accurate and timely analysis results.

Organizations Performance

Organizational performance is crucial for an organization's survival in the market and meeting stakeholder expectations. It involves analyzing and measuring results against goals, comparing actual productivity to desired outcomes. Teeth highlighted that innovation, protecting intangible knowledge assets, and leveraging them are essential for higher

performance. Managers at different levels ensure resources are used effectively to achieve organizational goals through actions and activities.

Using of Data Analytics by Business Management

Businesses are facing challenges in understanding and implementing big data strategies, with only 12% having done so and 70% planning to start. Knowledge of customers, products, and rules is essential for success, and big data can help organizations compete and make informed decisions. Businesses globally are using big data for decision-making across various functions such as Supply Chain, Product Research and Development, Marketing Management, and Sales.

The application of big data analytics is crucial for creating business value, requiring expertise in processing and extracting insights from data. It is used to manage organizational assets, monitor processes, strengthen supply chains, improve automation, and drive business transformation. Organizational performance is measured against goals and stakeholder expectations, emphasizing innovation and leveraging knowledge assets. Big data analytics in business management helps organizations compete and make informed decisions, with only 12% currently implementing a strategy but 70% planning to do so. It can improve decision-making capabilities and organizational performance by providing clear and important information. The quality of analytical tools impacts decision-making and organizational performance, distinguishing high-performing organizations from low-performing ones. Big data analytics can benefit any organization by improving operational efficiency and competitive advantage.

Intelligence and analytics for Business

Business intelligence (BI) is crucial for companies to effectively utilize available data, including Competitive Intelligence, Customer Intelligence, Market Intelligence, and more. BI enhances business performance by identifying opportunities, highlighting threats, and improving decision-making. However, current BI solutions primarily focus on structured internal data, leaving valuable insights in unstructured and external data untapped. The continuous collection of vast amounts of data from various sources facilitated by computing and Internet technologies has created an imperfect reality that can lead to biased decisions.

Big data analytics is transforming business intelligence by allowing companies to utilize large amounts of data to enhance customer satisfaction, manage supply chain risks, gather competitive insights, make real-time business decisions, and optimize pricing strategies. Research indicates that retailers can increase market share and operating profits by 60% through effective use of big data. One key aspect of big data analytics is its role in creating knowledge, developing new management principles, and shaping the economy. The business intelligence process typically consists of five phases.

Data collection involves gathering information from different sources, both external (market data providers, industry analytics) and internal (Google Analytics, CRM, ERP). Data cleansing/standardization is the process of preparing collected data for analysis by verifying data quality and ensuring consistency. Data storage involves loading data into a data warehouse for future use. Data analysis is an automated process that applies quantitative and qualitative analytical techniques to convert raw data into valuable information. Reporting includes creating dashboards and visual representations of analytical results that users can interact with to derive actionable insights.

An Impact of Big Data in enhancing business value through business intelligence

the key benefits of big data analytics is that it can significantly increase a retailer's operating profit margin by 60% through gaining market share and leveraging detailed consumer data. In general, big data analytics offers five main advantages: increased visibility of relevant data, improved performance and adaptability through accurate data capture, better understanding of customer needs, complementing decision-making with valuable insights from automated algorithms, and creating new business opportunities, products, and services. Big data analytics plays a crucial role in generating knowledge, shaping new management principles, and driving economic growth. Specifically, it can significantly enhance supply chain management by improving efficiency, planning, inventory control, risk management, market intelligence, and personalized services in real-time. Additionally, big data analytics enables innovation in product and service ideas and facilitates collaboration between subsidiaries to optimize operational processes cost-effectively. Furthermore, it aids in the decision-making process by providing insights based on a better understanding of different decision contexts and the necessary information processing mechanisms. To successfully implement big data analytics for decision-making, companies must focus on reducing data ambiguity and diversity.

Framework for Analyzing Large Data Sets

The optimal big data architecture patterns for an organization will vary based on factors like the particular industry, the scale of the company, and the data needs. Nonetheless, there are some general principles that can be adhered to in order to guarantee that the big data reference architecture remains effective and valid. One recommended practice is to implement the Big Data Cloud architecture, which entails keeping all data in a centralized repository in its original, unrefined state. This approach provides enhanced flexibility and simplifies data access, as data can be processed and analyzed as required without undergoing lengthy and expensive cleaning and transformation processes. Additionally, employing a distributed file system, such as the HDFS framework in Big Data (Hadoop Distributed File System), is also considered a best practice for storing and processing data. Understanding an organization's data needs is crucial in designing an architecture that can effectively meet those requirements. For instance, processing a large amount of architectural and flow data models in real-time big data necessitates a Big Data Hive architecture with a streaming data platform like Apache Kafka. The challenges of volume, veracity, and variety make it difficult to analyze and draw conclusions from Big Data. To address this, companies require a new IT architecture optimized for efficient processing of big data, with cloud computing offering unlimited resources on demand. Apache Hadoop is a popular architecture for Big Data, allowing organizations to collect, process, and analyze data efficiently in the rapidly evolving field of Big Data Analytics.

Hadoop, along with tools like YARN, MapReduce, Spark, Hive, Pig, and NoSQL databases, is a distributed data infrastructure that distributes large data collections across multiple nodes in a core server cluster, eliminating the need for expensive custom hardware. This allows for more efficient processing and analysis of Big Data, making Hadoop the foundation for tasks like scientific analysis, sales and business planning, and handling large volumes of sensor data from IoT sensors. Major companies such as Yahoo, Facebook, Amazon, eBay, The New York Times, Chevron, and IBM utilize Apache Hadoop for their data processing needs.

Approach

Big data analytics in business management involves a systematic framework for deriving insights from large datasets to improve decision-making. It starts with defining business objectives, collecting and integrating data from various sources, and setting up a storage infrastructure. Data preprocessing ensures quality and consistency before exploratory analysis identifies patterns and correlations. Analytical techniques are then applied to develop predictive models for customer behavior, supply chain optimization, and risk assessment. Validated models lead to actionable insights and recommendations, which are integrated into business processes through collaboration. Continual monitoring and optimization ensure ongoing success. Data models are continuously updated to adapt to changes in the business environment, ensuring the effectiveness of data-driven strategies. This includes maintaining data security and compliance standards to protect sensitive information and maintain stakeholder trust.

Obstacles associated with this matter.

1. **Challenges of Big Data**

   Researchers have defined big data and highlighted its key features, focusing on the 4V data properties: volume, variety, velocity, and veracity. The challenges associated with these properties are significant. The volume challenge arises from the massive amount of data being generated, requiring innovative methods for storage and processing. The variety challenge stems from the diverse formats in which data is received, with a majority being unstructured and difficult to analyze. The velocity challenge relates to the speed at which data is produced, with real-time processing becoming increasingly important for applications like online shopping. The veracity challenge concerns the quality and accuracy of data, especially in scenarios where important decisions are being made. Ensuring data reliability is crucial, particularly in contexts like social networking sites where user opinions can impact outcomes.

   Key challenges in Big Data include:

   Storage: With vast amounts of data generated daily, the greatest challenge is storage (especially when the data is in different formats) within legacy systems. Unstructured data cannot be stored in traditional databases.

   Processing: Processing big data refers to the reading, transforming, extraction, and formatting of useful information from raw information. The input and output of information in unified formats continue to present difficulties.

   Security: Security is a big concern for organizations. Non-encrypted information is at risk of theft or damage by cyber-criminals. Therefore, data security professionals must balance access to data against maintaining strict security protocols.

Finding and Fixing Data Quality Issues: Many of you are probably dealing with challenges related to poor data quality, but solutions are available. The following are four approaches to fixing data problems:

- Correct information in the original database.

- Repairing the original data source is necessary to resolve any data inaccuracies.

- You must use highly accurate methods of determining who someone is.

Scaling Big Data Systems: Database sharing, memory caching, moving to the cloud and separating read-only and write-active databases are all effective scaling methods. While each one of those approaches is fantastic on its own, combining them will lead you to the next level.

Evaluating and Selecting Big Data Technologies

Companies are spending millions on new big data technologies, and the market for such tools is expanding rapidly. In recent years, however, the IT industry has caught on to big data and analytics potential. The trending technologies include the following:

- Hadoop Ecosystem
- Apache Spark
- NoSQL Databases
- R Software
- Predictive Analytics
- Prescriptive Analytics

Big Data Environments

In an extensive data set, data is constantly being ingested from various sources, making it more dynamic than a data warehouse. The people in charge of the big data environment will fast forget where and what each data collection came from.

Real-Time Insights

The term "real-time analytics" describes the practice of performing analyses on data as a system is collecting it. Decisions may be made more efficiently and with more accurate information thanks to real-time analytics tools, which use logic and mathematics to deliver insights on this data quickly.

Data Validation

Before using data in a business process, its integrity, accuracy, and structure must be validated. The output of a data validation procedure can be used for further analysis, BI, or even to train a machine learning model.

Healthcare Challenges

Electronic health records (EHRs), genomic sequencing, medical research, wearables, and medical imaging are just a few examples of the many sources of health-related big data.

Barriers to Effective Use of Big Data in Healthcare

- The price of implementation

- Compiling and polishing data

- Security

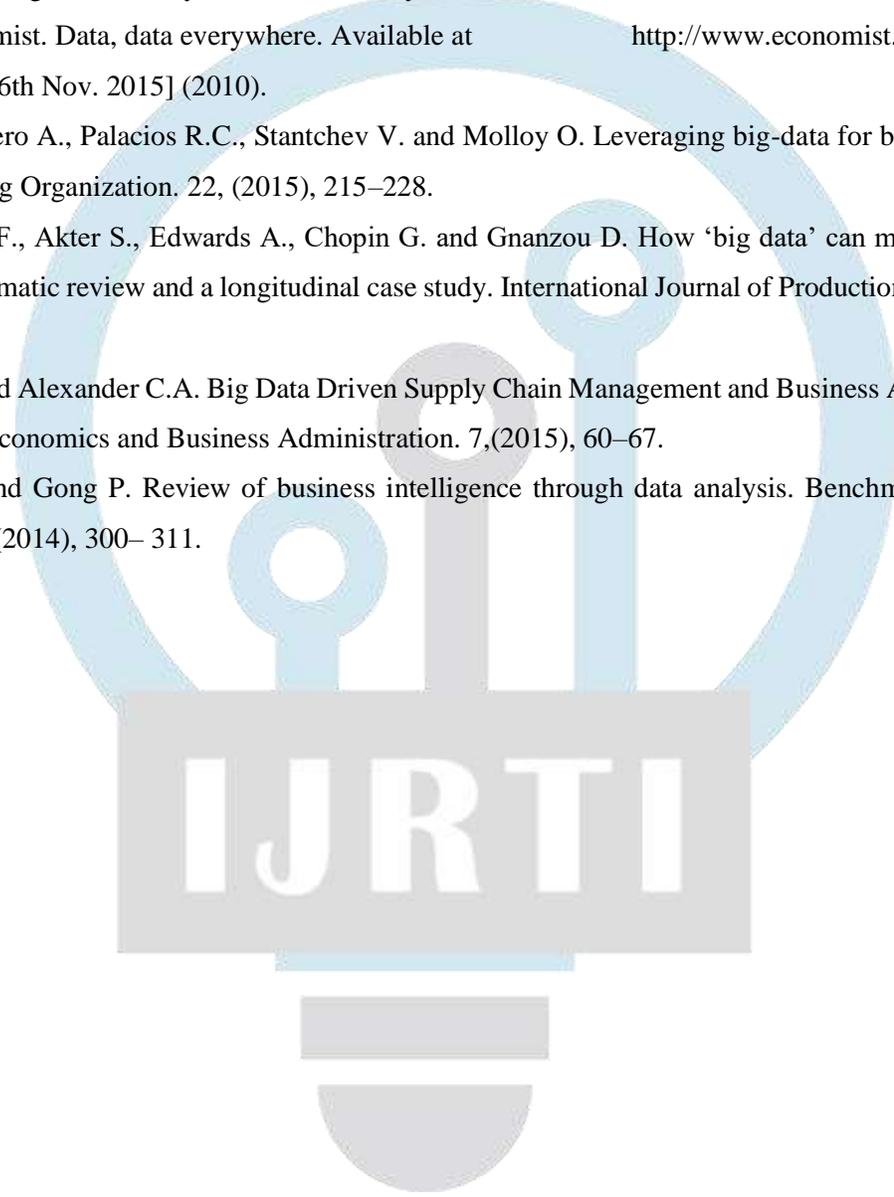- Disconnect in communication

Measuring Big Data for Success:

- **Data Quality KPIs:** Monitor metrics like accuracy, completeness, consistency, and timeliness to ensure data reliability.

- **Storage and Processing Costs:** Track the cost of storing and processing data to assess the efficiency of infrastructure and analytics pipelines.

- **Processing Time:** Measure the time it takes to analyze data to ensure real-time or near real-time insights.

- **Business Impact:** Evaluate the impact of big data analytics on key business outcomes, such as revenue growth, cost reduction, and improved customer satisfaction.

- **Return on Investment (ROI):** Assess the financial benefits of big data investments compared to the costs incurred.

- **Data Governance and Compliance:** Measure the effectiveness of data governance policies in ensuring data security, privacy, and compliance with regulations.

- **User Adoption and Engagement**: Track the adoption rate of big data analytics by different departments and users to ensure that insights are being used effectively.

REFERENCES

1. Ahmad W. and Quadri B.S.M.K. Big Data promises value: Is hardware technology taken onboard?. Industrial Management & Data Systems. 115(9), 2015.

2. Assunção M.D., Calheiros R.N., Bianchi S., Netto M.A.S. and Buyya R. Big Data computing and clouds: Trends and future directions. Journal of Parallel and Distributed Computing. 79–80 (2015), 3 – 15.

3. Avram (Olaru) M.G. Advantages and challenges of adopting cloud computing from an enterprise perspective. Procedia Technology. 12(2014), 529 – 534.

4. Davenport T.H. How strategists use 'big data' to support internal business decisions, discovery and production. Strategy & Leadership. 42(4), (2014), 45 – 50.

5. Erevelles S., Fukawa N. and Swayne L. Big Data consumer analytics and the transformation of marketing. Journal of Business Research. 69(2), (2016), 897-904.

6. García-Valls M., Biug T. and Lu C. Challenges in real-time virtualization and predictable cloud computing. Journal of Systems Architecture. 60(2014), 726–740.

7. Gandomi A. and Haider M. Beyond the hype: Big data concepts, methods, and analytics. International Journal of Information Management. 35(2015), 137–144.

8. Hashem I.A.T., Yaqoob I., Anuar N.B., Mokhtar S., Gani A. and Khan S.U. The rise of " big data" on cloud computing: Review and open research issues. Information Systems. 47(2015), 98 – 115.

9. Internetlivestats.com. Twitter Statistics. Retrieved November, 2015, from http://www.internetlivestats.com/twitter-statistics/ (2015).

10. Kambatla K., Kollias G., Kumar V. and Grama A. Trends in big data analytics. Journal of Parallel Distributed Computing. 74,(2014), 2561 – 2573.

11. Kimble, C. and Milolidakis, G. Big Data and Business Intelligence: Debunking the Myths. Global Business and Organizational

12. Kościelniaka H. and Puto A. BIG DATA in decision making processes of enterprises. Procedia Computer Science. 65,(2015), 1052– 1058.

13. Kowalczyk M. and Buxmann P. (2014). Big Data and Information Processing in Organizational Decision Processes: A Multiple Case Study. Business & Information Systems Engineering. 5,(2014), 267–278.

14. Leonard P. Customer data analytics: privacy settings for 'Big Data' business. International Data Privacy Law. 4, (2014), 53–68.

15. Marín-Ortega P.M., Dmitriyev V., Abilov M. and Gómez J.M. New Approach in Designing Business Intelligence Solutions in Era of Big Data. Procedia Technology.

16,(2014), 667–674. 16. Marshall A., Mueck S. and Shockley R. How leading organizations use big data and analytics to innovate. Strategy & Leadership. 43, (2015), 32–39.

17. Narayanan V. Using Big-Data Analytics to Manage Data Deluge and Unlock Real-Time Business Insights. Journal of Equipment Lease Financing. 32, (2014), 1–7.

18. Prescott M.E. Big data and competitive advantage at Nielsen. Management Decision. 52, (2014), 573–601.

19. Russom P. Big data analytics. (2011).

20. Saldaña J. The coding manual for qualitative researchers. 2nd ed. Los Angeles: Sage Publications. (2013)

21. Schomm F., Stahl F. and Vossen G. Marketplaces for data: an initial survey. SIGMOD Record. 42,(2013), 15–26.

22. Shahzad F. State-of-the-art Survey on Cloud Computing Security Challenges, Approaches and Solutions. Procedia Computer Science. 37,(2014), 357–362.

23. Sharma S., Tim U.S., Wong J., Gadia S. and Sharma S. A brief review on leading big data models. Data Science Journal. 13,(2014), 138–157.

24. Tan K.H, Zhan Y.Z.,Ji G., Ye F. and Chang C. Harvesting big data to enhance supply chain innovation capabilities: An analytic infrastructure based on deduction graph. International Journal of Production Economics. 165,(2015), 223–233.

25. Tankard C. Big data security. Network Security. (2012), 5–8.

26. The Economist. Data, data everywhere. Available at   http://www.economist.com/node/15557443 [Accessed 16th Nov. 2015] (2010).

27. Vera-Baquero A., Palacios R.C., Stantchev V. and Molloy O. Leveraging big-data for business process analytics. The Learning Organization. 22, (2015), 215–228.

28. Wamba S.F., Akter S., Edwards A., Chopin G. and Gnanzou D. How 'big data' can make big impact: Findings from a systematic review and a longitudinal case study. International Journal of Production Economics. 165,(2015), 234–246.

29. Wang L. and Alexander C.A. Big Data Driven Supply Chain Management and Business Administration. American Journal of Economics and Business Administration. 7,(2015), 60–67.

30. Xia B.S. and Gong P. Review of business intelligence through data analysis. Benchmarking: An International Journal. 21,(2014), 300– 311.