Advances in Consumer Research
ISSN (Print): 0098-9258
ISSN(Online): 3079-1766

Vol. 2, Issue 4 (2025) https://acr-journal.com/

The 42 Vs of Big Data

Dr. P. Sobha Rani[1], Dr. T. Geeta Madhuri Naidu[2]

[1]Assistant Professor, Department of Information Systems, GITAM School of Business, GITAM Deemed to be University
[2]Assistant Professor, Department of Accounting and Finance , GITAM School of Business GITAM Deemed to be university

Cite this paper as: Dr. P. Sobha Rani, Dr. T. Geeta Madhuri Naidu, (2025) The 42 Vs of Big Data. *Advances in Consumer Research*, 2 (4), 2710-2719

KEYWORDS

Big Data, Data, 3Vs,4 Vs,1C,5 Vs,14 Vs, 17 Vs 42Vs

ABSTRACT

The phrase "big data" became popular in 2013 and was identified as a group of data sets that were too huge and complicated for traditional database management systems. The 3 Vs (volume, velocity, and variety) or 4 Vs (adding veracity and Complexity) define its challenges, surpassing organizations' data, storage, and computing capacities. In essence, big data encompasses the digitalization of various aspects of life, generating dynamic data on the Internet and the Web. This work explores the problems , techniques, and applications of big data, emphasizing future business intelligence (BI) architectures. Recognizing data as an "ocean of Universal Facts," The research explores the potential of big data for future technological advancements, serving as a crucial resource for students and professionals entering the field. Focusing on the 5 Vs attributes of data and domain-wise comparisons, it addresses applications in Wholesale Trade, Retail, Utilities, Education, Transportation, Banking, Communication, Manufacturing, Government, Healthcare, and more. Big data, with its challenges in storage, analysis, and visualization, is used with analytics to obtain knowledge and a competitive edge. The article gives a general overview of the types, content, architecture, technology, and attributes (Volume, Velocity, Variety, Value, and Veracity) of big data. The 7 Vs Ten Vs (Volume, Variety, Velocity, Veracity, Validity, Value, Variability, Venue, Vocabulary, and Vagueness) and (Volume, Velocity, Variety, Variability, Veracity, Value, and Visualization) of big data are also covered in this work. Big data is transforming sectors such as manufacturing, e-commerce, banking, and insurance, this paper identifies and defines fourteen characteristics, addressing the growing challenges and introducing three new characteristics for efficient handling of big data. along with exploring three new features , to enhance the efficiency of handling big data. Additionally, it presents an updated list of 42 Vs to provide a comprehensive framework for understanding big data and data science in the modern era.

1. INTRODUCTION

The "Data is an Ocean of Universal Facts" underscores the significance of data in making decisions, highlighting the necessity of accurate computation of inferential value through data analytics. This growing divide between data volume and derived information, now referred to as "BIG DATA", highlights the evolving landscape. Businesses strive for decisions driving growth and customer satisfaction, beginning with data collection, objective-setting, and informed decision-making through enhanced knowledge. Understanding data attributes is fundamental, forming the basis of knowledge and analysis. Big Data Analytics involves a pipeline from acquisition to interpretation. A clear understanding of data attributes is crucial, as it informs analysis and decision-making. This survey explores diverse attributes across business domains, comparing their significance and challenges.

Big data, a cluster of vast and diverse data sets, has been integral to information science and digital communication since the beginning of computing.Its growth is exponential, fueled by data creation from various sources such as social networking sites, web servers, mobile devices, and call centres. However, because of its enormous quantity, speed, and complexity, managing big data poses several difficulties. surpassing the capabilities of traditional database technologies. Many organizations grapple with massive data volumes produced by high-volume transactions, such as digital photos, online logs, sensors, and contact centres, crucial for enhancing operational efficiency and driving business success. The relentless expansion of big data, beyond the capacity of traditional software tools to record, organize, and handle within acceptable timeframes, underscores its dynamic nature. Data volumes are predicted to escalate exponentially, with projections indicating

a staggering 175-181 Zettabytes (10^{21} Bytes) growth in the next two years and even doubling in some estimations. This continuous evolution of big data underscores the pressing need for enhanced strategies in capturing, curating, handling, and analyzing this ever-expanding information landscape. The actual exponential expansion of big data volume over time emphasizes its status as a dynamic and evolving phenomenon.

2. EVOLUTION OF 17 VS OF BIG DATA

3 Vs of Big Data

Big data is a novel concept, and handling it presents some difficulties extend beyond just its sheer volume. According to the widely recognized 3 V's model, big data management confronts organizations with the expansion of three critical properties: volume, velocity, and variety. The 2001 study "3D Data Management: Controlling Data Volume, Velocity" by Meta Group and Variety," Doug Laney, a Gartner analyst, developed this concept, which still offers insightful information about the challenges of managing large amounts of data.

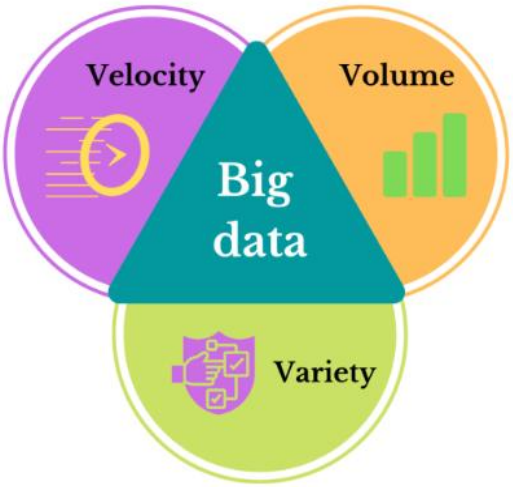

Fig 1: 3 Vs of Big Data

4 Vs of Big Data

SAS (Statistical Analysis System) has extended the traditional 3 V's model of big data by incorporating two more dimensions: complexity and variability. Additionally, The four Vs—volume, velocity, variety, and veracity—are how Oracle defines big data.This expansion reflects the growing recognition of the multifaceted nature of data management challenges.

Fig 2: 4 Vs of Big Data

5 Vs of Big Data

Oguntimilehin. A, proposed an expanded perspective on big data, introducing five key dimensions: Complexity, Value, Veracity, Volume, Velocity, and Variety.

Fig 3: 5 Vs of Big Data

7 Vs of Big Data

Volume, Velocity, Variety, Variability, Veracity, Value, and Visualization are the Seven Vs of Big Data Analytics is essential for crafting a robust strategy to handle extensive and intricate datasets. This comprehensive framework provides guidance for managing, analyzing, and deriving valuable business insights from large data sets. Mastering these dimensions is pivotal in developing an effective big data strategy that drives informed decision-making and maximizes organizational success.

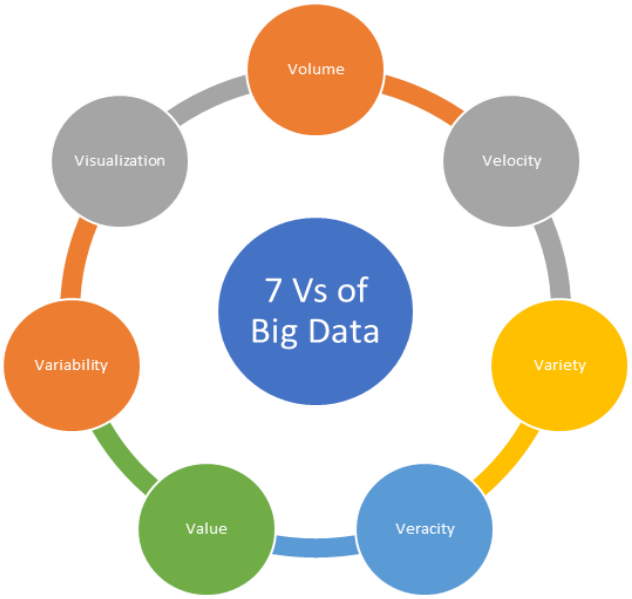

Fig 4: 7 Vs of Big Data

10 Vs of Big Data

In 2014, Kirk Born of Data Science Central elaborated on the idea of big data by defining it in terms of the 10 V's: volume, velocity, variety, veracity, value, validity, variability, venue, vocabulary, and vagueness.

Fig 5: 10 Vs of Big Data

14 Vs of Big Data

The comprehensive listing and definition of all **major characteristics of big data in the provided table offer a crucial research framework** for both scholars and practitioners, facilitating the effective management and utilization of big data resources. The exploration of these characteristics, encompassing the Fourteen V's and 1C as detailed in the table, is pivotal in shaping the trajectory of big data research. By delving into factors such as volume, velocity, value, and others, this study aims to streamline big data management, fostering its seamless integration into value-added applications and research endeavors.

S.No.	Characteristic	Elucidation	Description
1	**Volume**	Data size	Refers to the quantity of collected and stored data. The size is typically measured in terabytes (TB) or more.
2	**Velocity**	Speed of Data	Refers to the speed at which data is generated, transferred, and processed from the source to the destination.
3	**Value**	Importance of Data	Represents the potential business value or insights that can be derived from analyzing big data.
4	**Variety**	Type of Data	Involves different formats of data—structured, semi-structured, and unstructured like images, videos, and audio.
5	**Veracity**	Data Quality	Refers to the accuracy, trustworthiness, and reliability of the data being analyzed.
6	**Validity**	Data Authenticity	Ensures that the data used is correct, relevant, and appropriate for analysis and decision-making.
7	**Volatility**	Duration of Usefulness	Refers to how long the data remains valid or useful for analysis and decision-making.
8	**Visualization**	Data Representation	The process of representing data in visual formats like graphs, charts, and dashboards to make analysis more intuitive.
9	**Virality**	Spreading Speed	Defines the rate at which data is shared or spread across users and systems.

10	**Viscosity**	Lag of Event	Describes the delay between when an event occurs and when it is captured, processed, and described.
11	**Variability**	Data Differentiation	Refers to the inconsistency in data flow and how well the system differentiates between relevant and irrelevant data.
12	**Venue**	Different Platforms	Indicates that data originates from various platforms like personal systems, private and public clouds.
13	**Vocabulary**	Data Terminology	Refers to the standards, definitions, models, and structures used to describe data.
14	**Vagueness**	Indistinctness in Data	Points to unclear or ambiguous data that lacks precise meaning or interpretation.
15	**Complexity**	Correlation of Data	Highlights the challenge of managing data from multiple sources and identifying patterns, relationships, or changes over time.

Table 1: 14 Vs of Big Data Characteristics

17 Vs of Big Data

Many organizations today grapple with vast amounts of data, which inherently brings forth numerous challenges. Despite the plethora of proposed solutions, as discussed earlier, persistent issues remain. This underscores the need for a deeper comprehension of big data to effectively address its associated problems. Exploring existing literature on big data, this paper unveils three additional characteristics: verbosity, voluntarism, and versatility. These traits, alongside the established 14 V's and 1C, are envisaged to streamline big data management, enhancing its applicability in value-added contexts and research environments.

Fig 6: 14 Vs of Big Data

3. THE NEED FOR ADDITIONAL BIG DATA RESEARCH

The discourse surrounding big data has evolved significantly, reflecting the dynamic landscape of data management and analysis. S. Vikram Phaneendra and Madhusdhan Reddy noted the shifting paradigm from manageable data volumes handled by traditional RDBMS tools to the current era, where big data poses formidable challenges due to its distinct characteristics: Value, Variability, Velocity, Volume, and Variety. Dnvsl Indira and Kiran Kumar Reddy further elucidated the complexities of big data, highlighting its structured and unstructured nature, massive scale, and heterogeneity, emphasizing the need for advanced models for effective handling and transfer across networks. Wei Fan and Albert Bifet delved into extensive data mining, showcasing its transformative potential in extracting valuable insights from previously unattainable datasets. Experian disclosed that 97% of US businesses strive for comprehensive customer insights but need help with data management dilemmas, underscoring the critical need to tackle significant data challenges promptly.

Moreover, the escalating frequency of data breaches, as highlighted by the US government accountability office, necessitates robust data management strategies. In light of these discussions, it becomes evident that the conventional focus on the established characteristics (Fourteen Vs and a C) of big data needs to be revised. Recent developments have prompted the identification of new dimensions in big data characteristics. Most organizations today grapple with vast amounts of data, prompting the recognition of three new critical factors: verbosity, voluntarism, and versatility. These additional dimensions underscore the growing complexity of data management challenges and highlight the need for updated approaches to effectively navigate and leverage big data resources.

There's a pressing need to explore and define additional characteristics to enable more effective management and utilization. As we navigate the sophisticated analytics landscape in 2017, the introduction of The 42 Vs of Data Science and Big Data reflects this complex domain's evolving understanding and communication, providing a comprehensive framework to navigate its difficulties.

42 Vs of Big Data

In the expansive domain of Big Data, grasping its intricate layers is imperative to leverage its potential fully. The concept of the "42 Vs of Big Data" draws inspiration from Douglas Adams' renowned work, "The Hitchhiker's Guide to the Galaxy," where the enigmatic number "42" signifies the answer to life's ultimate question. Although no universally acknowledged roster of 42 Vs exists, this notion playfully underscores the vast array of considerations within the realm of Big Data.

Within this paper, I delve into a select subset of these Vs, concentrating on 17 pivotal dimensions that encapsulate crucial aspects of Big Data. By delving into these dimensions, I aim to offer profound insights into the remaining 25 Vs, namely Vivify, Vogue, Voice, Vane, Vanilla, Vantage, Valor, Varifocal, Varmint, Varnish, Vastness, Vaticination, Vault, Veer, Veil, Verdict, Versed, Version Control, Vet, Vexed, Viability, Vibrant, Victual, Virtuosity, Visibility, Vivify, Vol.and Vogue.These dimensions are intrinsic to the intricate process of managing and analyzing vast volumes of data. And These Characteristics are defined as:

Beyond the notion of the 42 Vs, this paper illuminates the foundational elements that underlie the Big Data landscape. Focusing on these 17 Vs provides a nuanced perspective, serving as a valuable resource for practitioners, researchers, and enthusiasts alike. And These 42 Vs Characteristics are defined as:

Big Data Characteristics	Description
Vane	Refers to the direction or course in which data trends are heading. Understanding the vane of data helps in predicting future patterns and making informed decisions.
Vanilla	Represents the raw, unprocessed form of data before any modifications or enhancements. It underscores the importance of accessing data in its purest state for accurate analysis.
Vantage	Denotes the advantageous position gained through insightful data analysis. Having a vantage point allows organizations to make strategic decisions and gain competitive advantages in their respective industries.
Valor	Signifies the courage and bravery required to tackle the challenges associated with Big Data. It emphasizes the need for resilience and determination in navigating complex data landscapes

Varifocal	Refers to the ability to adjust focus and perspective when analyzing data. Being varifocal enables individuals to zoom in on specific details or zoom out for a broader understanding, depending on the context.
Varmint	In the context of Big Data, a varmint represents any disruptive or problematic elements within datasets. Managing varmints involves identifying and mitigating factors that could negatively impact data quality or analysis.
Varnish	In Big Data refers to optimization and caching mechanisms used to enhance data access performance. Similar to varnish improving surfaces, data varnishing optimizes storage and retrieval processes, reducing latency by caching frequently accessed data. This boosts efficiency and scalability by lessening the load on backend servers, resulting in improved user experiences and operational efficiency within Big Data solutions.
Vastness	Highlights the sheer magnitude and scale of data available for analysis. Recognizing the vastness of data underscores the need for robust infrastructure and advanced analytics techniques to derive meaningful insights.
Vaticination	Refers to the act of forecasting or predicting future trends based on data analysis. Vaticination involves using historical data and advanced analytics to anticipate potential outcomes and make proactive decisions.
Vault	Represents the secure storage and protection of sensitive or confidential data. Maintaining a vault ensures data privacy and compliance with regulatory requirements, safeguarding against unauthorized access or breaches.
Veer	Signifies a sudden change or deviation in data trends or patterns. Understanding when data veers off course is crucial for detecting anomalies and adjusting analytical approaches accordingly.
Veil	Represents the concealment or obscuring of certain aspects of data for privacy or security reasons. Applying a veil ensures that sensitive information remains protected while still allowing for analysis and insights extraction.
Verdict	Refers to the conclusive decision or judgment derived from data analysis. A verdict indicates the outcome or findings resulting from thorough examination and interpretation of data.
Versed	Denotes being knowledgeable or experienced in handling and analyzing data. Being well-versed in data analytics methodologies and tools is essential for effectively extracting insights and driving informed decision-making.

Version Control	entails overseeing various data versions or iterations to guarantee accuracy and consistency. Throughout its existence, version control enables data integrity maintenance, modification tracking, and state reversion.
Vet	Signifies the process of thoroughly examining and verifying the accuracy and quality of data. Vetting involves conducting rigorous checks and validations to ensure that data is reliable and suitable for analysis
Vexed	Represents the challenges and complexities inherent in dealing with Big Data. Acknowledging the vexed nature of data encourages the development of innovative solutions and strategies to overcome obstacles
Viability	Refers to the feasibility and sustainability of Big Data initiatives. Assessing the viability of projects involves considering factors such as resource requirements, potential returns on investment, and alignment with organizational goals.
Vibrant	Describes data that is dynamic, active, and continuously evolving. Recognizing the vibrancy of data underscores the importance of real-time analytics and adaptive strategies to capture valuable insights.
Victual	Represents the nourishment or sustenance derived from data-driven insights. Just as victuals provide nourishment for the body, data-driven insights nourish decision-making processes, enabling organizations to thrive and succeed.
Virtuosity	Denotes the mastery or expertise in leveraging Big Data for strategic advantage. Achieving virtuosity in data analytics involves honing skills, leveraging advanced techniques, and continuously refining analytical approaches.
Visibility	Refers to the clarity and transparency of data and analytics processes. Improving visibility enables stakeholders to understand and trust the insights derived from data analysis, fostering informed decision-making and collaboration
Vivify	Signifies the act of bringing data to life or imbuing it with vitality and relevance. Vivifying data involves transforming raw information into actionable insights that drive meaningful outcomes and innovation
Vogue	indicates how well-liked or current a particular data analysis methodology or practice is. Understanding the popularity of data analytics enables businesses to stay up to date with new developments in technology and trends in order to stay competitive.

Voice	Denotes the expression or representation of data insights in a meaningful and compelling manner. Giving voice to data involves effectively communicating insights to stakeholders through visualization, storytelling, and other methods to drive understanding and action.

4. CONCLUSION

Big Data emerges as a vibrant and ever-expanding domain, driven by the exponential growth in data volume from diverse sources like mobile devices and call centers. While this paper has delved into fundamental attributes such as volume, velocity, and value, it also recognizes the encompassing breadth epitomized by the whimsical notion of the "42 Vs of Big Data." Inspired by the imaginative world of Douglas Adams, this concept playfully underscores the intricate layers of Big Data's complexity. Through this lens, we grasp the multifaceted nature of Big Data, prompting a deeper exploration of its nuances and possibilities.

REFERENCES

[1] Nagham Saeed*, Laden Husamaldin, “Big Data Characteristics (V’s) in Industry”, Iraqi Journal of Industrial Research (IJOIR), http://ijoir.gov.iq, DOI: 10.53523/ijoir Vol. 8, No. 1, June 14,2021.

[2] J. Johnson, S. Friend and H. Lee, "Big Data Facilitation, Utilization, and Monetization: Exploring the 3Vs in a New Product Development Process", Journal of Product Innovation Management, vol. 34, no. 5, pp.640-658, 2017. Available: 10.1111/jpim.12397 [Accessed 21 May 2021].

[3] J. Kepner, V. Gadepally, P. Michaleas, N. Schear, M. Varia, A. Yerukhimovich and R. K. Cunningham,"Computing on masked data: a high-performance method for improving big data veracity", 2014 IEEE HighIraqi Journal of Industrial Research, Vol. 8, No. 1 (2021)8.

[4] Performance Extreme Computing Conference (HPEC), 2014. Available: 10.1109/hpec.2014.7040946 [Accessed 21 May 2021].

[5] T. Nguyen, "A Framework for Five Big V’s of Big Data and Organizational Culture in Firms", 2018 IEEE International Conference on Big Data (Big Data), 2018. Available: 10.1109/bigdata.2018.8622377 [Accessed 21 May 2021].

[6] J. Debattista, C. Lange, S. Scerri and S. Auer, "Linked 'Big' Data: Towards a Manifold Increase in Big Data Value and Veracity", 2015 IEEE/ACM 2nd International Symposium on Big Data Computing (BDC), 2015. Available: 10.1109/bdc.2015.34 [Accessed 21 May 2021].

[7] M. Khan, M. Uddin and N. Gupta, "Seven V's of Big Data understanding Big Data to extract value", Proceedings of the 2014 Zone 1 Conference of the American Society for Engineering Education, 2014. Available: 10.1109/aseezone1.2014.6820689 [Accessed 21 May 2021].

[8] Vinaya Keskar, Jyoti Yadav, Ajay H. Kumar, “5V’s of Big Data Attributes and their Relevance and Importance across Domains”, International Journal of Innovative Technology and Exploring Engineering (IJITEE), ISSN: 2278-3075, Volume-9 Issue-11, September 2020.

[9] Muhammad Mashab Farooqi, Munam Ali Shah, Abdul Wahid,Adnan Akhunzada, Faheem Khan, Noor ul Amin, and Ihsan Ali. “Big Data in Healthcare: A Survey”, https://www.researchgate.net/publication/328859276, DOI: 10.1007/978-3-319-96139-2_14, January 2019.

[10] Reihaneh H. Hariri, Erik M. Fredericks and Kate M. Bowers, “Uncertainty in big data analytics: survey, opportunities, and challenges”, Journal of Big Data 6, 44 (2019). https://doi.org/10.1186/s40537-019-0206-3, 04 June 2019.

[11] J. Ranjan, "The 10 Vs of Big Data framework in the Context of 5 Industry Verticals", PRODUCTIVITY, vol.59, no. 4, pp. 324-342, 2019. Available: 10.32381/prod.2019.59.04.2.

[12] S.Vikram Phaneendra & E.Madhusudhan Reddy “Big Datasolutions for RDBMS problems- A survey,” In 12th IEEE/IFIP Network Operations & Management Symposium (NOMS 2010) (Osaka, Japan, Apr 19{23 2013).

[13] Jeffrey Ray, Olayinka Johnny, Marcello Trovati, Stelios Sotiriadis and Nik Bessis, “The Rise of Big Data Science: A Survey of Techniques, Methods and Approaches in the Field of Natural Language Processing and Network Theory”, Big Data Cogn. Comput.018, 2, 22; doi:10.3390/bdcc2030022, www.mdpi.com/journal/bdcc. Volume 2 Issue 3, 2 August 2018.

[14] Arockia Panimalar., Varnekha Shree., and Veneshia Kathrine. “The 17 V’s of Big Data”, International Research Journal of Engineering and Technology (IRJET), Volume: 04 Issue: 09 | Sep-2017, e-ISSN: 2395-0056, p-ISSN: 2395-0072.

[15] Oguntimilehin A., Ademola E.O. ‘‘A Review of Big Data Management, Benefits and Challenges,’’ Journal of Emerging Trends in Computing and Information Sciences, vol-5, pp-433-437, June 2014.

[16] Stephen Kaisler, Frank Armour, J.Alberto Espinosa and Wolliam Money “Big Data: Issues and Challenges Moving Forward,” Hawaii International Conference on System Sciences 46th, pp-995-1003, 2013.

fffff