

5

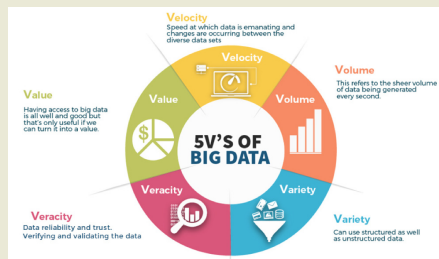
# Understanding The Vs of Big Data

– Manpreet Kaur

Research Scholar, SOMS, IGNOU, New Delhi

[ID https://orcid.org/0000-0002-3196-0376](https://orcid.org/0000-0002-3196-0376) [✉ mghai.89@gmail.com](mailto:mghai.89@gmail.com)

"Big Data" is a field of data science that investigates how huge amount of data can be separated and examined so as to gather information from them. There are five main characteristics of big data known as the "5 V's of Big Data" which help us to effectively understand the meaning of big data. These attributes of big data help the organisations to improve the value of its current information.



Source: <https://www.techentice.com/>

## Keywords

- Big Data
- Velocity
- Value
- Data Consistency
- Variety

## ARTICLE HISTORY

**Paper Nomenclature:** Scrutiny Tip (ST)

**Paper Code:** CYBNMV2N2FEB2020ST1

**Submission Online:** 06-Feb-2020

**Manuscript Acknowledged:** 07-Feb-2020

**Originality Check:** 10-Feb-2020

**Originality Test Ratio:** 0%

**Peer Reviewers Comment:** 13-Feb-2020

**Blind Reviewers Remarks:** 15-Feb-2020

**Author Revert:** 19-Feb-2020

**Camera-Ready-Copy:** 21-Feb-2020

**Editorial Board Citation:** 29-Feb-2020

**Published Online First:** 13-May-2020

## Introduction

The five 'Vs', viz. volume, variety, velocity, veracity, and value, are considered as the pillars of Big Data. Each of these have been briefly explained below:

- **Volume** refers to the *magnitude* of data that is gathered or produced by companies each day. The constant influx of huge amount of data can no longer be saved or analyzed using conventional data processing methods. The need for additional storage is one of the major issues business encounter when introducing big data practices. Thus, volume is the most important and distinctive feature of Big Data as it imposes specific requirements to all traditional technologies and tools

currently used. It is difficult to gather, manage, and analyze the huge amount of data with the existing IT infrastructure and tools; thus, these large quantity of data sets require new and innovative tools and approaches for collecting, storing, and analyzing data.

- **Variety** refers to the *diversity* of data. The data is generated from different sources and formats and contain multidimensional data fields. The data collected maybe structured, semi-structured or unstructured. Data also comprise images, videos and speech recordings in addition to conventional datasets. In today's world, the percentage of unstructured data is more than 80%. However, there are different ways to process these large amount of unstructured information, either through investing in big data analytical platforms or hiring data professionals.

- **Velocity** refers to the *speed* with which the data is generated, collected, analyzed and reprocessed. Since big data is gathered in such large quantities, it is not only the storage infrastructure which is required to process the information properly but also the speed at which it should be processed. Fast processing of information maximizes efficiency and yields better results for business than a huge amount of data that takes longer time to analyse. As the flow of data is continuous from various sources it becomes important to make available the right information at the right time.
- **Veracity** refers to the *authenticity* of data and its credibility. Big Data involves working with all degrees of quality since the volume, velocity and variety factors usually result in a shortage of quality but the amount of data

collected and the speed at which it generates mean nothing if that data isn't authentic and reliable. Thus, veracity emphasises the relevance of quality of data and the trustworthiness of source of data.

- **Value** refers to the *added value* that the collected data can provide to the companies. The core of value lies in obtaining meaningful insights from huge data sets. It emphasizes that it is not the quantity of data that speak by themselves and therefore it should be accompanied with technological tools and human minds to derive value from them. It emphasizes that the focus should not be on the aggregation of data and storage infrastructure but also on how worthy the collected data is for the company. Thus, it becomes important to do cost benefit analysis of collecting and analysing data to ensure that it adds value to the business.

### Reference

- <https://www.quora.com/Why-is-data-called-big-data>
- <https://softwareengineering.stackexchange.com/questions/378103/how-to-process-huge-amount-of-data-with-limited-processing-resources>
- [https://books.google.cz/books?id=nHc5DwAAQBAJ&pg=PA83&lpg=PA83&dq=%22huge+amount+of+data+with%22&source=bl&ots=GjoRHQB7WJ&sig=ACfU3U2MW29A\\_5Q-w9MsNmAb\\_uOt-9xkl0g&hl=cs&sa=X&ved=2ahUKEwj8stnEoLPnAhWPjQKQHVkSDOoQ6AEwBnoECAKQAQ](https://books.google.cz/books?id=nHc5DwAAQBAJ&pg=PA83&lpg=PA83&dq=%22huge+amount+of+data+with%22&source=bl&ots=GjoRHQB7WJ&sig=ACfU3U2MW29A_5Q-w9MsNmAb_uOt-9xkl0g&hl=cs&sa=X&ved=2ahUKEwj8stnEoLPnAhWPjQKQHVkSDOoQ6AEwBnoECAKQAQ)
- [https://simple.wikipedia.org/wiki/Data\\_mining](https://simple.wikipedia.org/wiki/Data_mining)
- <https://www.softwarereadinghelp.com/data-mining-examples/>
- <https://www.writing-skills.com/data-singular-or-plural>
- <https://www.quora.com/What-should-I-learn-to-be-able-to-work-with-a-huge-amount-of-data-with-Python-and-server-database-like-PostgreSQL>
- <https://ieeexplore.ieee.org/docuement/7528083?section=abstract>
  - <https://stackoverflow.com/questions/12196653/how-does-dictionary-search-a-particular-key-value-out-of-huge-amount-of-data?rq=1>
- <https://answers.microsoft.com/en-us/msoffice/forum/all/i-have-more-very->

large-ranges-of-data-the-data-is/d7e231d6-cb17-4fd3-94d3-a5e182f54-bcc

- [https://se.mathworks.com/matlabcentral/answers/366757-how-to-handle-a-huge-amount-of-data-with-surface-and-stair-plots-to-avoid-memory-spikes?s\\_tid=prof\\_contriblnk](https://se.mathworks.com/matlabcentral/answers/366757-how-to-handle-a-huge-amount-of-data-with-surface-and-stair-plots-to-avoid-memory-spikes?s_tid=prof_contriblnk)
- [https://books.google.com/books?id=VQ5FDwAAQBAJ&pg=PA87&lpg=PA87&dq=%22of+huge+amount+of+data%22&source=bl&ots=q-X1KecR59&sig=ACfU3U1t6cfp86QC4qks\\_IQwBl-d0iHwGA&hl=en&sa=X&ved=2ahUKEwjB\\_u\\_EoLPnAhXEs1kKHZ\\_RC1sQ6AEwCxoECAcQAQ](https://books.google.com/books?id=VQ5FDwAAQBAJ&pg=PA87&lpg=PA87&dq=%22of+huge+amount+of+data%22&source=bl&ots=q-X1KecR59&sig=ACfU3U1t6cfp86QC4qks_IQwBl-d0iHwGA&hl=en&sa=X&ved=2ahUKEwjB_u_EoLPnAhXEs1kKHZ_RC1sQ6AEwCxoECAcQAQ)
- [https://books.google.cz/books?id=FJEYi7FpFcC&pg=PA267&lpg=PA267&dq=%22of+data+the+data+is%22&source=bl&ots=XxPFoZbhO&sig=ACfU3U0z3-P42xNO2W6hYvpsqZRDfC\\_8Pw&hl=cs&sa=X&ved=2ahUKEwjR0tjEoLPnAhWQ-aQKHc0yBbgQ6AEwCxoECAgQAQ](https://books.google.cz/books?id=FJEYi7FpFcC&pg=PA267&lpg=PA267&dq=%22of+data+the+data+is%22&source=bl&ots=XxPFoZbhO&sig=ACfU3U0z3-P42xNO2W6hYvpsqZRDfC_8Pw&hl=cs&sa=X&ved=2ahUKEwjR0tjEoLPnAhWQ-aQKHc0yBbgQ6AEwCxoECAgQAQ)
- <https://books.google.cz/books?id=hxSjBQAAQBAJ&pg=PA614&lpg=PA614&dq=%22of+data+the+data+is%22&source=bl&ots=MnK3sKMWdl&sig=ACfU3U1TJGysjNwVsh32u2dDQPpDGiPR7Q&hl=cs&sa=X&ved=2ahUKEwjR0tjEoLPnAhWQ-aQKHc0yBbgQ6AEwCnoECAKQAQ>
- <https://www.geeksforgeeks.org/file-organization-in-dbms-set-1/>
- <https://stats.stackexchange.com/questions/399094/is-a-model-fitted-to-data-or-is-data-fitted-to-a-model>
- <https://stackoverflow.com/questions/50229517/how-can-i-create-an-array-from-a-huge-amount-of-data>
- <https://patents.google.com/patent/US20100179962A1/en>
- [https://books.google.com/books?id=zK2rDwAAQBAJ&pg=PA118&lpg=PA118&dq=%22of+huge+amount+of+data%22&source=bl&ots=vmoXosg6VW&sig=ACfU3U1kQWQXNFrYcMKhQMIQKNRfobBjJg&hl=en&sa=X&ved=2ahUKEwjB\\_u\\_EoLPnAhXEs1kKHZ\\_RC1sQ6AEwC3oECAsQAQ](https://books.google.com/books?id=zK2rDwAAQBAJ&pg=PA118&lpg=PA118&dq=%22of+huge+amount+of+data%22&source=bl&ots=vmoXosg6VW&sig=ACfU3U1kQWQXNFrYcMKhQMIQKNRfobBjJg&hl=en&sa=X&ved=2ahUKEwjB_u_EoLPnAhXEs1kKHZ_RC1sQ6AEwC3oECAsQAQ)
- <https://answers.sap.com/questions/441501/get-huge-amount-of-data-with-good-performance.html>
- <https://www.verdict.co.uk/public-health-big-data/>
- <https://stackoverflow.com/questions/11449325/approaches-to-fast-visualization-of-huge-amount-of-data>
- [https://books.google.cz/books?id=tlwDwAAQBAJ&pg=PT390&lpg=PT390&dq=%22huge+amount+of+data+with%22&source=bl&ots=OgOgC1ZCm&sig=ACfU3U2ix8QPq\\_KnUaoA5je8-7pYafLHOQ&hl=cs&sa=X&ved=2ahUKEwj8stnEoLPnAhWPjQKQHVkSDOoQ6AEwB3oECAgQAQ](https://books.google.cz/books?id=tlwDwAAQBAJ&pg=PT390&lpg=PT390&dq=%22huge+amount+of+data+with%22&source=bl&ots=OgOgC1ZCm&sig=ACfU3U2ix8QPq_KnUaoA5je8-7pYafLHOQ&hl=cs&sa=X&ved=2ahUKEwj8stnEoLPnAhWPjQKQHVkSDOoQ6AEwB3oECAgQAQ)
- <https://books.google.cz/books?id=qX5TDwAAQBAJ&pg=PA73&lpg=PA73&dq=%22huge+amount+of+data+with%22&source=bl&ots=YcDtSL4h0D&sig=ACfU3U3lpMTIzG8a2pnoNixtmuelC90RJ>

g&hl=cs&sa=X&ved=2ahUKEwj8stnEoLPnAhWPjQKQHVkSDOoQ6AEwBHoE-CACQAQ

- [https://books.google.cz/books?id=TxgXhfMmJ5sC&pg=PA106&lpg=PA106&dq=%22of+data+the+data+is%22&source=bl&ots=TBtRPIEwiB&sig=ACfU3U3PnJ\\_xkVthEN1xCnVjxiMDuWzFcg&hl=cs&sa=X&ved=2ahUKEwjR0tjEoLPnAhWQ-aQKHc0yBbgQ6AEwC3oECAsQAQ](https://books.google.cz/books?id=TxgXhfMmJ5sC&pg=PA106&lpg=PA106&dq=%22of+data+the+data+is%22&source=bl&ots=TBtRPIEwiB&sig=ACfU3U3PnJ_xkVthEN1xCnVjxiMDuWzFcg&hl=cs&sa=X&ved=2ahUKEwjR0tjEoLPnAhWQ-aQKHc0yBbgQ6AEwC3oECAsQAQ)
- <https://books.google.cz/books?id=hhtFDwAAQBAJ&pg=PA300&lpg=PA300&dq=%22huge+amount+of+data+with%22&source=bl&ots=tRwa8hertJ&sig=ACfU3U0LhmE4yBEnvZi8D8miMPmpScCA1w&hl=cs&sa=X&ved=2ahUKEwj8stnEoLPnAhWPjQKQHVkSDOoQ6AEwBxoE-CACQAQ>
- <https://www.quora.com/What-percent-age-of-data-is-unstructured>
- <https://toulouse2018.esof.eu/en/programme.html>
- <https://stackoverflow.com/questions/208399/ hashing-data-to-ensure-it-wasnt-corrupted-during-transfer>
- [https://books.google.com/books?id=eQR6DwAAQBAJ&pg=PA489&lpg=PA489&dq=%22of+huge+amount+of+data%22&source=bl&ots=OlB3MXz5k&sig=ACfU3U32kiABORxLiMXj0\\_sjaN3nhztf6A&hl=en&sa=X&ved=2ahUKEwjB\\_u\\_EoLPnAhXEs1kKHZ\\_RC1sQ6AEwDX-oECAgQAQ](https://books.google.com/books?id=eQR6DwAAQBAJ&pg=PA489&lpg=PA489&dq=%22of+huge+amount+of+data%22&source=bl&ots=OlB3MXz5k&sig=ACfU3U32kiABORxLiMXj0_sjaN3nhztf6A&hl=en&sa=X&ved=2ahUKEwjB_u_EoLPnAhXEs1kKHZ_RC1sQ6AEwDX-oECAgQAQ)
- [https://books.google.com/books?id=d\\_Y-DwAAQBAJ&pg=PA298&lpg=PA298&dq=%22of+huge+amount+of+data%22&source=bl&ots=Ownot5uBb0&sig=ACfU3U07r0Lj-VA4bW-GejyuV2HD3wyOOG&hl=en&sa=X&ved=2ahUKEwjB\\_u\\_EoLPnAhXEs1kKHZ\\_RC1sQ6AEwDHoECAYQAQ](https://books.google.com/books?id=d_Y-DwAAQBAJ&pg=PA298&lpg=PA298&dq=%22of+huge+amount+of+data%22&source=bl&ots=Ownot5uBb0&sig=ACfU3U07r0Lj-VA4bW-GejyuV2HD3wyOOG&hl=en&sa=X&ved=2ahUKEwjB_u_EoLPnAhXEs1kKHZ_RC1sQ6AEwDHoECAYQAQ)
- <https://www.enterrasolutions.com/blog/good-data-bad-data-big-data/>
- <https://www.linguee.jp/%E8%8B%B1%E8%AA%9E-%E6%97%A5%E6%9C%AC%E8%AA%9E/%E7%BF%BB%E8%A8%B3/a+amount+of+data.html>
- <https://www.linguee.ru/%D0%B0%D0%BD%D0%B3%D0%BB%D0%B8%D0%B9%D1%81%D0%BA%D0%B8%D0%B9-%D1%80%D1%83%D1%81%D1%81%D0%BA%D0%B8%D0%B9/%D0%BF%D0%B5%D1%80%D0%B5%D0%B2%D0%BE%D0%B4/amount+of+data.html>
- <https://www.linguee.jp/%E8%8B%B1%E8%AA%9E-%E6%97%A5%E6%9C%AC%E8%AA%9E/%E7%BF%BB%E8%A8%B3/to-manage+a+large+amount+of+data.html>
- <https://www.itee.uq.edu.au/developing-model-factors-affecting-effective-use-big-data>
- <https://www.coursehero.com/tutors-problems/Statistics-and-Probability/11603062-Analyze-the-data-to-ensure-that-it-meets-the-assumptions-of-the-test/>
- [https://www.stat.purdue.edu/~doerge/BIOLFORM/D/SPRING16/KatalWazidGoudar\\_2013.pdf](https://www.stat.purdue.edu/~doerge/BIOLFORM/D/SPRING16/KatalWazidGoudar_2013.pdf)



**Manpreet Kaur** is an Assistant Professor in Department of Commerce, Mata Sundri College for Women, University of Delhi. She is pursuing her PhD from School of Management Studies, IGNOU. She has more than six years of teaching experience. She has contributed several research papers and has participated in various National and International Conferences.

mghai.89@gmail.com

### Annexure I

Submission Date	Submission Id	Word Count	Character Count
10-Feb-2020	D63703068 (urkund)	542	8120

**Urkund Analysis Result**

**Analysed Document:** Internet of skills.docx (D63703068)  
**Submitted:** 2/10/2020 3:53:00 PM  
**Submitted By:** editorial.scholastic.seed@gmail.com  
**Significance:** 0 %

Sources included in the report:  
 Instances where selected sources appear: 0

*Note: The Cybernomics had used the urkund plagiarism [http://www.orkund.com] tool to check the originality.*



### Reviewers Comment

- Review 1:** In this paper the author has covered all the importance of the Big data in the terms of V’s volume, variety, velocity, veracity.
- Review 2:** In my opinion: The Big data is alternative one of person’s shorthand words, nonetheless this is one that Janice in Accounting.
- Review 3:** Every business, big or small, is managing a considerable amount of data generated through its various data points and business processes. At times, businesses are able to handle these data using excel sheets, access databases or other similar tools.



### Editorial Excerpt

This Article has 10% significance of plagiarism; the finding related to “**Understanding The Vs of Big Data**” in general words The Big data volume defines the ‘amount’ of data that is produced. The value of data is also reliant on on the scope of the data. Nowadays data is made from various sources in different formats organized and shapeless. In essence, when the media talk about Big Data, they’re not just speaking about vast sums of data that are potential gem troves of gen. After review comment it is obvious to take this object under “**Scrutiny Tip (ST)**” Category.

### Acknowledgement

Author is highly indebted to Scholastic Seed Inc & editorial team of Cybernomics, For making the write-up in the shape of an article.

### Disclaimer

All views expressed in this paper are my/our own. Some of the content is taken from open source websites & some are copyright free for the purpose of disseminating knowledge. Those some We/I had mentioned above in the references section and acknowledged/cited as when and where required. The author/s has cited their joint own work mostly, Tables/Data from other referenced sources in this particular paper with the narrative & endorsement has been presented within quotes and reference at the bottom of the article accordingly & appropriately. Finally some of the contents which are taken or overlapped from open source websites for the knowledge purpose. Those some of i/we had mentioned above in the references section.



### Citation

Manpreet Kaur  
 “Understanding The Vs of Big Data”  
 Volume-2, Issue-2, Feb 2020.  
 (www.cybernomics.in)

Frequency: Monthly, Published: 2020  
 Conflict of Interest: Author of a Paper  
 had no conflict neither financially  
 nor academically.

