

# How Many Vs are there in Big Data?

Chuck Cartledge

Thursday 18<sup>th</sup> February, 2016 at 16:23

## Contents

<b>1 Introduction</b>	<b>1</b>
<b>2 Big Data Vs</b>	<b>1</b>
<b>3 Conclusion</b>	<b>3</b>

## List of Tables

1 Big Data Vs. . . . .	2
------------------------	---

## 1 Introduction

Doug Laney is widely credited with originating the first big Vs of Big Data: volume, variety, and velocity. Since then the number of Vs has grown and grown. I will list all the Vs that I can find in chronological order until I get tired, to see if there is a pattern in their growth.

## 2 Big Data Vs

Table 1 chronicles some of the Big Data Vs.

Table 1: Big Data Vs.

Num.	Year	V	Definition	Source
1	2001	Variety	... no greater barrier to effective data management will exist than the variety of incompatible data formats, non-aligned data structures, and inconsistent data semantics.	[3, 5]
2	2001	Velocity	E-commerce has also increased point-of-interaction (POI) speed and, consequently, the pace data used to support interactions and generated by interactions.	[3]
3	2001	Volume	E-commerce channels increase the depth/breadth of data available about a transaction (or any point of interaction).	[3]
4	2013	Validity	... is the data correct and accurate for the intended use.	[1, 4, 5, 6, 8]
5	2013	Value	How to determine the prescriptive value of data?	[1, 2, 4, 7, 8, 9]
6	2013	Variability	Many options or variable interpretations can confuse interpretation.	[1, 2, 5, 7, 9]
7	2013	Veracity	...to the biases, noise and abnormality in data.	[1, 2, 4, 6, 8, 9]
8	2013	Viability	...can the data be analyzed in a way that makes it decision-relevant?	[2, 5]
9	2013	Virility	... Defined by some users as the rate at which the data spreads; how often it is picked up and repeated by other users or events.	[9]
10	2013	Viscosity	... used to describe the latency or lag time in the data relative to the event being described.	[9]
11	2013	Visibility	... the state of being able to see or be seen - is implied. [4, 8, 5]	
12	2013	Visualization	Making all that vast amount of data comprehensible in a manner that is easy to understand and read. With the right analyses and visualizations, raw data can be put to use otherwise raw data remains essentially useless.	[7]

(Continued on the next page.)

Table 1. (Continued from the previous page.)

Num.	Year	V	Definition	Source
13	2013	Volatility	... how long is data valid and how long should it be stored.	[5, 6]
14	2014	Vagueness	... confusion over the meaning of big data (Is it Hadoop? Is it something that weve always had? Whats new about it? What are the tools? Which tools should I use? etc.)	[1]
15	2014	Venue	... distributed, heterogeneous data from multiple platforms, from different owners systems, with different access and formatting requirements, private vs. public cloud.	[1]
16	2014	Vocabulary	... schema, data models, semantics, ontologies, taxonomies, and other content- and context-based metadata that describe the datas structure, syntax, content, and provenance.	[1]
17	2015	Vincularity	... it implies connectivity or linkage.	[5]
18	2015	Visible	We live in an increasingly visual world and the statistics of increase in the number of images and videos shared on the Internet is staggering.	[5]
19	2015	Vitality	... criticality of the data is another concept that is crucial and is embedded in the concept of Value.	[5]

### 3 Conclusion

Doug Laney’s initial 3Vs were based on his experiences in a business mergers and acquisitions environment. The terms and ideas were easy to understand and to relate to. Many, many people have hopped on the “V” band wagon, and will probably continue to for the foreseeable future. From a practical stand point; if your data won’t fit in one machine, or the machine takes too long to return an answer, you have a Big Data problem, regardless of many Vs apply.

## References

- [1] Kirk Borne, *Top 10 big data challenges - a serious look at 10 big data vs*, <https://www.mapr.com/blog/top-10-big-data-challenges-%E2%80%93-serious-look-10-big-data-v%E2%80%99s>, 2014.
- [2] Seth Grimes, *Big data: Avoid 'wanna v' confusion*, <http://www.informationweek.com/big-data/big-data-analytics/big-data-avoid-wanna-v-confusion/d/d-id/1111077?>, 2013.
- [3] Doug Laney, *3d data management: Controlling data volume, velocity and variety*, META Group Research Note **6** (2001).
- [4] Rob Livingstone, *The 7 vs of big data*, <http://rob-livingstone.com/2013/06/big-data-or-black-hole/>, 2013.
- [5] Rajiv Maheshwari, *3 vs or 7 vs - whats the value of big data?*, <https://www.linkedin.com/pulse/3-vs-7-whats-value-big-data-rajiv-maheshwari>, 2105.
- [6] Kevin Normandeau, *Beyond volume, variety and velocity is the issue of big data veracity*, <http://insidebigdata.com/2013/09/12/beyond-volume-variety-velocity-issue-big-data-veracity/>, 2013.
- [7] BI Staff, *Why the 3vs are not sufficient to describe big data*, <https://datafloq.com/read/3vs-sufficient-describe-big-data/166>, 2013.
- [8] University of Technology Staff, *The 7 vs of big data*, <http://mbitm.uts.edu.au/feed/7-vs-big-data>, 2013.
- [9] Bill Vorhies, *How many vs in big data the characteristics that define big data*, <http://data-magnum.com/how-many-vs-in-big-data-the-characteristics-that-define-big-data/>, 2013.