

## In Previous Years Questions

Big data is characterized by four key dimensions known as the “4V’s”:

### 1. Volume

- This refers to the sheer amount of data being generated and stored.
- Big data deals with massive datasets that traditional data management tools cannot efficiently handle.
- Example: Social media platforms, sensor networks, and financial transactions generate enormous amounts of data, demonstrating the vastness of Big Data.

### 3. Variety

- This refers to the diverse range of data formats and structures present in big data.
- Unlike traditional data sets, big data can be structured, semi-structured, or unstructured, encompassing text, images, audio, video, and sensor data.
- This variety requires specialized tools and techniques for processing and analysis.
- Example: The variety of data in Big Data sets, such as social media posts, sensor measurements, and video recordings, highlight the need for specialized processing and analysis techniques.

### 2. Velocity

- This refers to the speed at which data is generated and processed.
- Big data often involves real-time or near real-time data streams that require immediate analysis and action.
- Example: Stock market data, traffic data, and streaming services illustrate the high

velocity nature of Big Data, where quick analysis and action are critical.

## 4. Veracity

- This refers to the accuracy, completeness, and trustworthiness of the data.
- Big data can be noisy, incomplete, and riddled with errors due to its vastness and diverse sources.
- Example: Data cleaning and pre-processing are crucial steps in Big Data analysis to ensure the veracity of the data and extract reliable insights.

### References:

- Practical big data analytics by Natraj Dasgupta
- "Data Science for Business" by Foster Provost and Tom Fawcett (2013)
- "Big Data: New Tricks for Econometrics" by James H. Stock and Mark W. Watson (2012)