

The Three Vs of Data: Volume

Volume refers to the amount of data being generated. You should think in terms of gigabytes, terabytes, and petabytes. Many systems and applications are just not able to store, let alone ingest or process that much data.

Many factors contribute to the increase in data volume. This includes transaction-based data stored for years, unstructured data streaming in from social media, and the ever-increasing amounts of sensor and machine data being produced and collected.

There are problems related to the volume of data. The storage cost is an obvious issue. Another problem is filtering and finding relevant and valuable information in large quantities of data that often contains non-valuable (or even corrupted!) information.

You also need a solution to analyze data fast enough, in order to maximize the business value today, and not just next quarter, or next year.

