



# What is Big Data?


Create an AWS Account




**Explore Free Analytics Offers**  
View free offers for Analytics services in the cloud



**Check out Analytics Services**  
Innovate faster with the most comprehensive set of Analytics services



**Browse Analytics Trainings**  
Get started on Analytics training with content built by AWS experts



**Read Analytics Blogs**  
Read about the latest AWS Analytics product news and best practices

- What is big data?
- Why you may need big data?
- How does big data work?
- How can AWS support your big data requirements?

## What is big data?

Big data can be described in terms of data management challenges that – due to increasing volume, velocity and variety of data – cannot be solved with traditional databases. While there are plenty of definitions for big data, most of them include the concept of what's commonly known as “three V’s” of big data:

- Volume:** Ranges from terabytes to petabytes of data
- Variety:** Includes data from a wide range of sources and formats (e.g. web logs, social media interactions, ecommerce and online transactions, financial transactions, etc)
- Velocity:** Increasingly, businesses have stringent requirements from the time data is generated, to the time actionable insights are delivered to the users. Therefore, data needs to be collected, stored, processed, and analyzed within relatively short windows – ranging from daily to real-time

## Why you may need big data?

Despite the hype, many organizations don't realize they have a big data problem or they simply don't think of it in terms of big data. In general, an organization is likely to benefit from big data technologies when existing databases and applications can no longer scale to support sudden increases in volume, variety, and velocity of data.

Failure to correctly address big data challenges can result in escalating costs, as well as reduced productivity and competitiveness. On the other hand, a sound big data strategy can help organizations reduce costs and gain operational efficiencies by migrating heavy existing workloads to big data technologies; as well as deploying new applications to capitalize on new opportunities.

## How does big data work?

With new tools that address the entire data management cycle, big data technologies make it technically and economically feasible, not only to collect and store larger datasets, but also to analyze them in order to uncover new and valuable insights. In most cases, big data processing involves a common data flow – from collection of raw data to consumption of actionable information.

**Collect.** Collecting the raw data – transactions, logs, mobile devices and more – is the first challenge many organizations face when dealing with big data. A good big data platform makes this step easier, allowing developers to ingest a wide variety of data – from structured to unstructured – at any speed – from real-time to batch.

**Store.** Any big data platform needs a secure, scalable, and durable repository to store data prior or even after processing tasks. Depending on your specific requirements, you may also need temporary stores for data in-transit.

**Process & Analyze.** This is the step where data is transformed from its raw state into a consumable format – usually by means of sorting, aggregating, joining and even performing more advanced functions and algorithms. The resulting data sets are then stored for further processing or made available for consumption via business intelligence and data visualization tools.

**Consume & Visualize.** Big data is all about getting high value, actionable insights from your data assets. Ideally, data is made available to stakeholders through self-service business intelligence and agile data visualization tools that allow for fast and easy exploration of datasets. Depending on the type of analytics, end-users may also consume the resulting data in the form of statistical “predictions” – in the case of predictive analytics – or recommended actions – in the case of prescriptive analytics.

### The Evolution of Big Data Processions

The big data ecosystem continues to evolve at an impressive pace. Today, a diverse set of analytic styles support multiple functions within the organization.

**Descriptive analytics** help users answer the question: “*What happened and why?*” Examples include traditional query and reporting environments with scorecards and dashboards.

**Predictive analytics** help users estimate the probability of a given event in the future. Examples include early alert systems, fraud detection, preventive maintenance applications, and forecasting.

**Prescriptive analytics** provide specific (prescriptive) recommendations to the user. They address the question – What should I do if “x” happens?

Originally, big data frameworks such as Hadoop, supported only batch workloads, where large datasets were processed in bulk during a specified time window typically measured in hours if not days. However, as time-to-insight became more important, the “velocity” of big data has fueled the evolution of new frameworks such as [Apache Spark](#), [Apache Kafka](#), [Amazon Kinesis](#) and others, to support [real-time and streaming](#) data processing.

## How can AWS support your big data requirements?

Amazon Web Services provides a [broad and fully integrated portfolio](#) of cloud computing services to help you build, secure, and deploy your big data applications. With AWS, there's no hardware to procure, and no infrastructure to maintain and scale, so you can focus your resources on uncovering new insights. With new capabilities and features added constantly, you'll always be able to leverage the latest technologies without making long-term investment commitments.

[Learn more about AWS big data platform and tools »](#)

### Immediate Availability

Most big data technologies require large clusters of servers resulting in long provisioning and setup cycles. With AWS you can deploy the infrastructure you need almost instantly. This means your teams can be more productive, it's easier to try new things, and projects can roll out sooner.

### Broad & Deep Capabilities

Big data workloads are as varied as the data assets they intend to analyze. A broad and deep platform means you can build virtually any big data application and support any workload regardless of volume, velocity, and variety of data. With 50+ services and hundreds of features added every year, AWS provides everything you need to collect, store, process, analyze, and visualize big data on the cloud. [Learn more about the AWS big data platform.](#)

### Trusted & Secure

Big data is sensitive data. Therefore, securing your data assets and protecting your infrastructure without losing agility is critical. AWS provides capabilities across facilities, network, software, and business processes to meet the strictest requirements. Environments are continuously audited for certifications such as ISO 27001, FedRAMP, [DoD SRG](#), and PCI DSS. Assurance programs help you prove compliance with 20+ standards, including HIPAA, NCSC, and more. Visit the [Cloud Security Center](#) to learn more.

### Hundreds of Partners & Solutions

A large partner ecosystem can help you bridge the skills gap and get started with big data even faster. Visit the [AWS Partner Network](#) to get help from a consulting partner or choose from many tools and applications across the entire data management stack.


### Big Data Solutions at AWS

Let us help you solve your big data challenges. Leave the heavy lifting to us, so you can focus more time and resources on the goals of your business or organization.


Learn more about [AWS big data solutions »](#)

Get started with big data analytics on AWS by [creating an account](#) today.


## Next Steps on AWS



**Check out additional product-related resources**  
[View free offers for Analytics services in the cloud »](#)



**Sign up for a free account**  
Instant get access to the AWS Free Tier.  
[Sign up »](#)



**Start building in the console**  
Get started building in the AWS management console.  
[Sign in »](#)

### Learn About AWS

- What is AWS?
- What is Cloud Computing?
- AWS Accessibility
- AWS Inclusion, Diversity & Equity
- What is DevOps?
- What is a Container?
- What is a Data Lake?
- What is Artificial Intelligence (AI)?
- What is Generative AI?
- What is Machine Learning (ML)?
- AWS Cloud Security
- What's New
- Blogs
- Press Releases

### Resources for AWS


- Getting Started
- Training and Certification
- AWS Solutions Library
- Architecture Center
- Product and Technical FAQs
- Analyst Reports
- AWS Partners

### Developers on AWS

- Developer Center
- [SDKs & Tools](#)
- .NET on AWS
- Python on AWS
- Java on AWS
- PHP on AWS
- JavaScript on AWS

### Help

- Contact Us
- Get Expert Help
- File a Support Ticket
- AWS re:Post
- Knowledge Center
- AWS Support Overview
- Legal
- AWS Careers



Amazon is an Equal Opportunity Employer: [Minority](#) / [Women](#) / [Disability](#) / [Veteran](#) / [Gender Identity](#) / [Sexual Orientation](#) / [Age](#).

Hi, I can connect you with an AWS representative or answer questions you have on AWS.

Need more info? Highlight any text to get an explanation generated with AWS generative AI.