## **Big data and analytics**

Data comes in all types of forms and formats. When we talk about big data, we're referring to large volumes of data. Data from weather systems, communications systems, imaging platforms, and many other scenarios generate large amounts of data. This amount of data becomes increasingly hard to make sense of and make decisions around. The volumes are so large that traditional forms of processing and analysis are no longer appropriate.

Open source cluster technologies have been developed, over time, to try to deal with these large data sets. Microsoft Azure supports a broad range of technologies and services to provide big data and analytic solutions. Some of the most common big data and analytic service types in Azure are Azure SQL Data Warehouse, HDInsight, and Data Lake Analytics.

## **Azure Synapse Analytics**

[Azure Synapse Analytics](https://docs.microsoft.com/en-us/azure/sql-data-warehouse" \t "https://docs.microsoft.com/en-us/learn/modules/identify-azure-solutions/az-portal) (formerly Azure SQL Data Warehouse) is a limitless analytics service that brings together enterprise data warehousing and big data analytics.

## **Azure HDInsight**

[Azure HDInsight](https://azure.microsoft.com/services/hdinsight" \t "https://docs.microsoft.com/en-us/learn/modules/identify-azure-solutions/az-portal) is a fully managed, open-source analytics service for enterprises. It is a cloud service that makes it easier, faster, and more cost-effective to process massive amounts of data. HDInsight allows you run popular open-source frameworks and create cluster types such as [Apache Spark](https://docs.microsoft.com/en-us/azure/hdinsight/spark/apache-spark-overview" \t "https://docs.microsoft.com/en-us/learn/modules/identify-azure-solutions/az-portal),[Apache Hadoop](https://docs.microsoft.com/en-us/azure/hdinsight/hadoop/apache-hadoop-introduction" \t "https://docs.microsoft.com/en-us/learn/modules/identify-azure-solutions/az-portal),[Apache Kafka](https://docs.microsoft.com/en-us/azure/hdinsight/kafka/apache-kafka-introduction" \t "https://docs.microsoft.com/en-us/learn/modules/identify-azure-solutions/az-portal), [Apache HBase](https://docs.microsoft.com/en-us/azure/hdinsight/hbase/apache-hbase-overview" \t "https://docs.microsoft.com/en-us/learn/modules/identify-azure-solutions/az-portal),[Apache Storm](https://docs.microsoft.com/en-us/azure/hdinsight/storm/apache-storm-overview" \t "https://docs.microsoft.com/en-us/learn/modules/identify-azure-solutions/az-portal),[Machine Learning Services](https://docs.microsoft.com/en-us/azure/hdinsight/r-server/r-server-overview" \t "https://docs.microsoft.com/en-us/learn/modules/identify-azure-solutions/az-portal). HDInsight also supports a broad range of scenarios such as extraction, transformation, and loading (ETL); data warehousing; machine learning; and IoT.

## **Azure Data Lake Analytics**

[Azure Data Lake Analytics](https://azure.microsoft.com/services/data-lake-analytics" \t "https://docs.microsoft.com/en-us/learn/modules/identify-azure-solutions/az-portal) is an on-demand analytics job service that simplifies big data. Instead of deploying, configuring, and tuning hardware, you write queries to transform your data and extract valuable insights. The analytics service can handle jobs of any scale instantly by setting the dial for how much power you need. You only pay for your job when it is running, making it more cost-effective.