

Azure Synapse Vs. Azure Data Factory Vs. Databricks

Azure Synapse :

Azure Synapse is unified analytics service that includes capabilities for data integration, data warehousing and analytics. For integration it supports over 95 connectors that allow to gather data from multiple sources. After collection the data can be stored in data warehouse.

Synapse's numerous features to manage end-to-end big data workflow have quickly made it popular among analytics professionals. Its ability to handle unstructured data stored in a data lake makes it even more powerful. Synapse also supports analyst-friendly languages like T-SQL, Python, Scala, Spark SQL, and .Net.

Azure Data Factory :

Azure Data factory is fully managed, serverless data integration service. It has more than 90 connectors to collect data from different sources. We can transform data without writing single code of line in Azure Data Factory. Since it is a no-code platform, it enables non-technical professionals to gather and transform data efficiently.

Azure Data Factory also supports Git and CI/CD to incrementally build ETL/ELT pipelines. With Azure Data Factory, you can even monitor your pipelines without its no-code capabilities.

Databricks :

Azure Databricks is a unified, open analytics platform for building, deploying, sharing, and maintaining enterprise-grade data, analytics, and AI solutions at scale. It allows to process structured, semi-structured and non-structured data. It provides code compatibility with Python, R, Scala and Java making it flexible to use.

Unlike Synapse (which is mainly for BI/analytics) and ADF (mainly for ETL orchestration), Databricks focuses on scalable data processing and building intelligent models. It also integrates well with Azure Data Lake, Event Hubs, Synapse, and MLflow for end-to-end workflows.

Differences between Azure Synapse, Azure Data Factory and Azure Databricks

Feature / Service	Azure Synapse	Azure Data Factory	Azure Databricks
Purpose	Data warehouse & analytics	Data integration & orchestration	Big data processing & ML
Data Handling	Structured/semi-structured	Structured & unstructured	Any type (structured, unstructured, streaming)
Language	SQL (T-SQL)	GUI + limited SQL	Python, R, Scala, SQL
Use Case	BI, dashboards, ad-hoc queries	ETL pipelines, data movement	AI/ML, advanced analytics, big data
Typical Users	Data analysts, BI developers	Data engineers	Data scientists, ML engineers