Modern Data WareHouse

on-premises data
(Oracle, SQL, SAP,
Teradata, file shares)

cloud data

Saas data
(Salesforce, Dynamics)

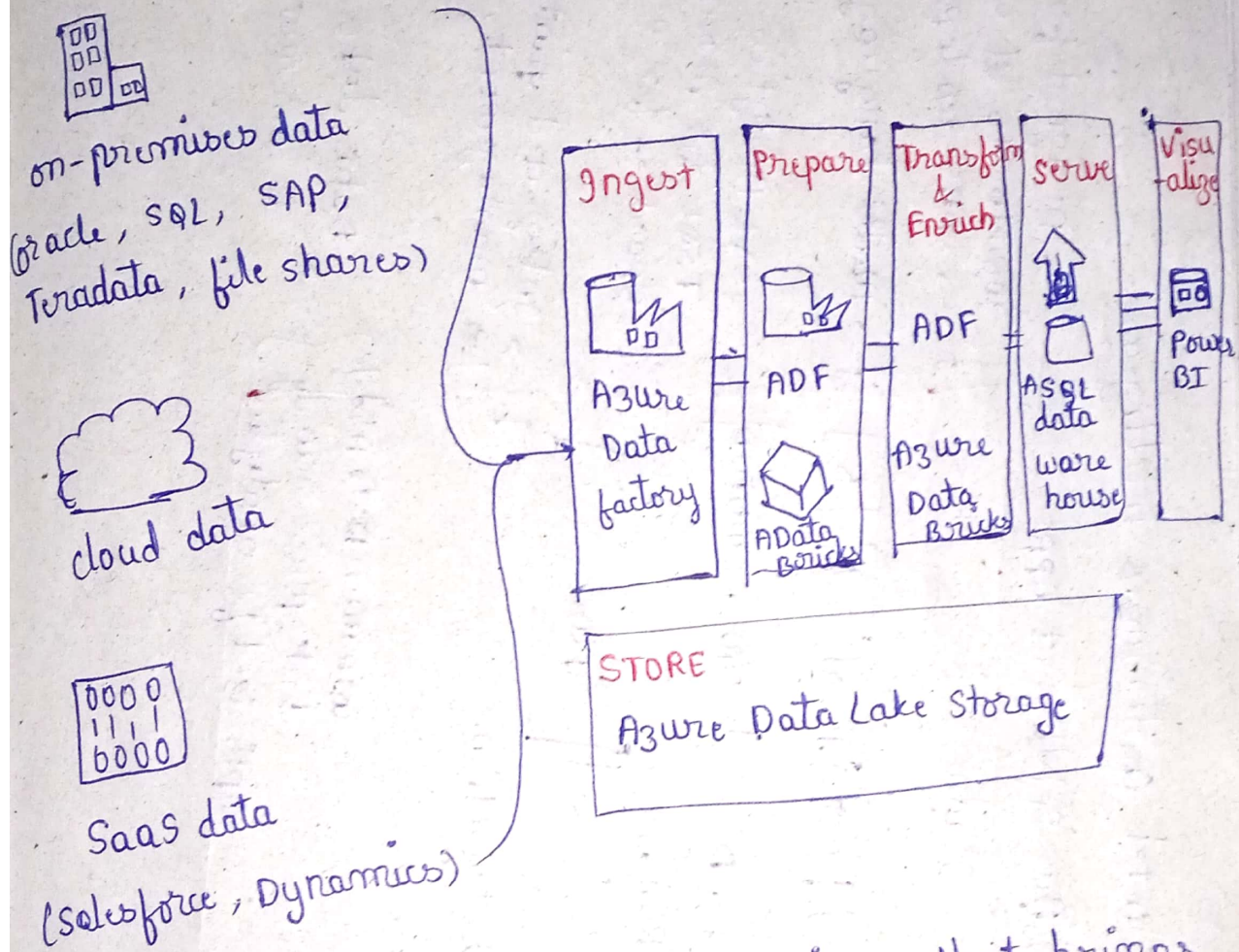| Ingest	Prepare	Transform & Enrich	Serve	Visu -alize
Azure Data factory	ADF AData Bricks	ADF Azure Data Bricks	ASQL data ware house	Power BI

STORE
Azure Data Lake Storage

* It is a limitless analytics service, that brings together enterprise data warehousing & Big data analytics.

* It allows to query data using (i) Serverless on-demand
(ii) provisioned resources.

* It brings these two worlds together with a unified experience to [ingest prepare manage serve] data for immediate BI and ML needs.

Azure Synapse Analytics (Architecture)

Integrated data platform for BI, AI, and continuous intelligence

→ Designed for analytics workloads at any scale

AI/ML/IOT, BI/Bus Intelligent APPS }

→ SAAS developer experiences

Synapse Analytics

experience → Synapse Analytics Studio

Platform

Language → multi languages are used

| Management | SQL | Python | .NET | Java | Scala | R |

Form factors

Security → PROVISIONED | ON-DEMAND

SQL Analytics :- offers
{ TSQL for Processing
Spark :- for " with R, Python, Scala, .NET }

Analytics Runtimes

| Monitoring | SQL | | Spark |

Metastore | Data Integration

→ Integrated platform services for management, security etc.

Azure Data Lake Storage

{ Common Data Model? → Data lake integrated & common data
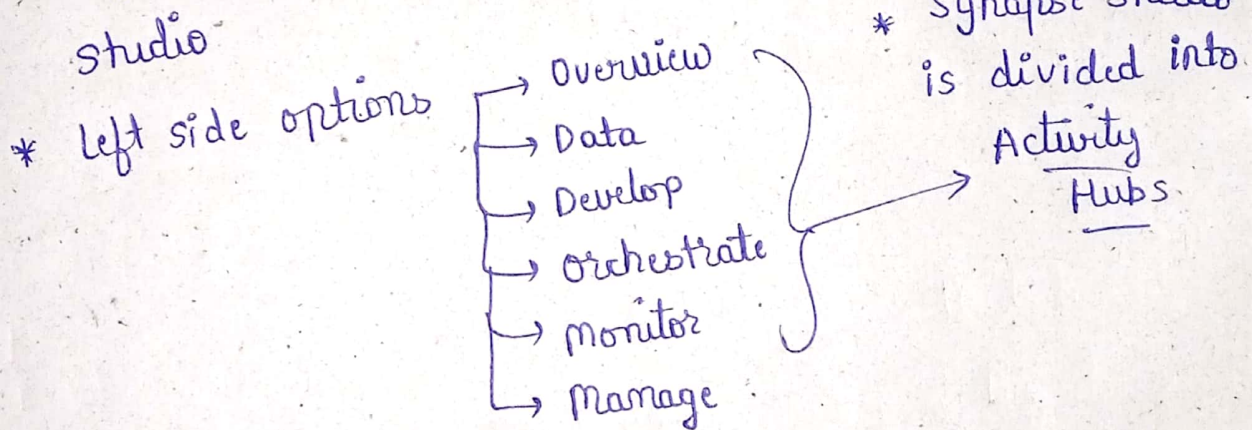Enterprise Security } → Model

Optimized for analytics

Connected Services (Third party Integration)

1) Azure Data Catalog
2) ADLS
3) Azure Data share
4) Azure Databricks
5) " HD Insight
6) Azure ML
7) Power BI

Provisioning Synapse Workspace

* Provide necessary details & click on Create
* After deployment, go to resource and Launch studio

* Left side options
  - Overview
  - Data
  - Develop
  - Orchestrate
  - Monitor
  - Manage

* Synapse studio is divided into
  Activity
  Hubs

Azure Synapse Analytics - SQL Analytics

Focus areas
1) Best in class price per performance
2) Industry - leading industry
3) Workload aware query execution
4) Data flexibility
5) Developer productivity

What workloads are NOT suitable?

1) Operational workloads (OLTP)

* High frequency reads and writes
* Large number of singleton Selects
* High volumes of single row inserts.

2) Data preparations

* Row by row processing needs
* Incompatible formats (xml)