# Azure Quick Steps — Key Vault, Storage, Data Factory, Databricks, Synapse

Azure Big Data Workflow

## 1) Login to Azure

* Open **https://portal.azure.com**.
* Sign in using your Azure account credentials.

## 2) Azure Key Vault

* In the search bar, type **Key vaults**.
* Click **+ Create** and fill in the Resource group, Name, and Region.
* Once created, open the vault and go to **Secrets**.
* Add a new secret by clicking **+ Generate/Import**.

*Purpose:* Key Vault is used to securely store passwords, keys, and sensitive information.

## 3) Storage Account & Container

* Search for **Storage accounts** and click **+ Create**.
* Provide the required details and create the account.
* After creation, open it and select **Containers**.
* Create a new container, give it a name, set access level, and save.

*Purpose:* Storage accounts hold files such as CSVs or other datasets.

## 4) Azure Data Factory

* Search **Data factories** → **+ Create** → open the Data Factory → **Launch studio**.
* Go to **Manage** → **Integration runtimes** → create a **Self-hosted** runtime.
* Link it to your Storage account.
* In **Author**, create a pipeline to copy or move data between sources.

*Purpose:* Data Factory is used to move and transform data.

## 5) Azure Databricks

* Search **Azure Databricks** → open workspace → **Launch Workspace**.
* In **Compute**, create a cluster.
* Create a notebook to work with data.
* Demonstrate: mount the CSV file, remove unwanted columns, delete unnecessary rows, display 10 rows, and show data for **12-Jul-2019**.

*Purpose:* Databricks is for data analysis and processing.

## 6) Azure Synapse Analytics

* Search **Synapse workspaces** → open → **Open Synapse Studio**.
* Create a SQL script using **Serverless SQL**.
* Query the CSV from storage and filter for **12-Jul-2019**.

*Purpose:* Synapse is for running queries and analytics over data.