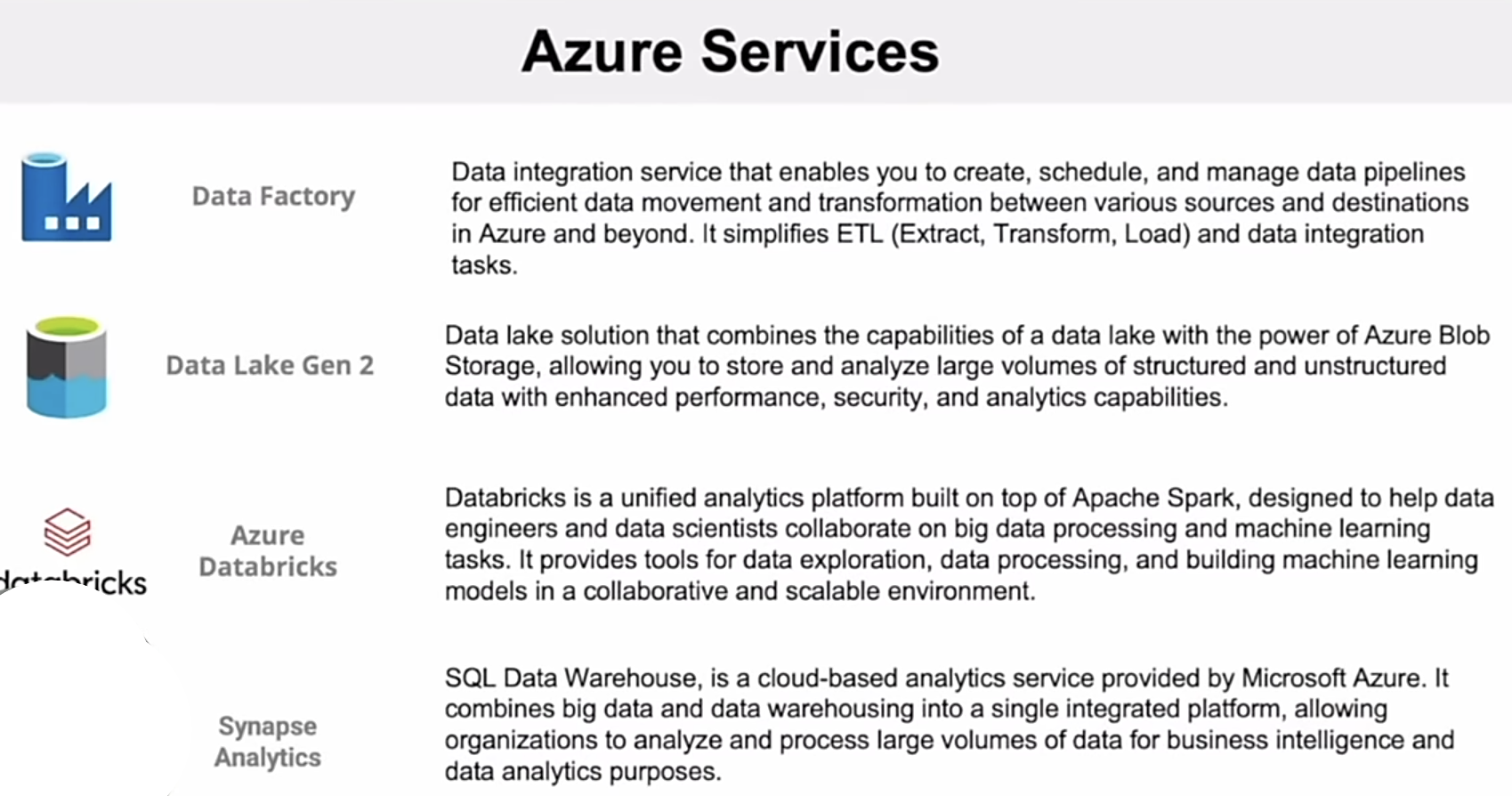
Une image contenant texte, diagramme, carte, capture d’écran

Description générée automatiquement

* Create an account with email (First Level)
* With that account, possibility to create multiple subscriptions (Free trial, eng. subscription, …). So based on the budget, you can allocate different subscription plans to different departments.
* Inside a subscription, you have a resource group (logical grouping of different resources such as data factory, data lakes, apps analytics, Azure database, …)



* Azure blob: the object storage. When you store a file on Azure blob, it is considered as an entire object (so inside a bucket, each file is considered as an object)

**Create an account**

* Azure cloud > First link

<https://azure.microsoft.com/fr-fr/>

* Click on Free Account / Start with Azure > Try Azure freely
* Create a Microsoft Account (if you don’t have yet)
* Create your Azure profile (Make sure you have a credit card)
* Access to the Azure portail

<https://portal.azure.com/>

Une image contenant texte, conception, capture d’écran

Description générée automatiquement

**Set up our data lake (e.g. folder in our container ???)**

* Create a storage account: will allows you to store object data
* Create a new resource group: tokio-olympic
* Give a name to the storage account: should be unique in all Azure worldwide (e.g.: tokyoolympicdatavad)
* Select the nearest region for example
* Choose the performance
* Specify redundance: if you want for e.g. to replicate the data across different data centers or region
* Click on next
* Activate the hierarchical Namespace: all data stored will be available in a hierarchical format (as you could have seen it on your computer and not as an object)
* Click on next
* Click on next
* Click on next
* After the validation, click on “Create”
* Click on “Go to resources”
* At the left window, we are interested in “data storage” > “containers”: store the data as an object
* Click on “containers” > “create container” > Give a name (e.g.: tokyo-olympic-data) > “create”
* Click on our container “tokyo-olympic-data” > “add directory” > Give a name (e.g.: raw-data) > “save”
* Create another directory > Give a name (e.g.: transformed-data) > “save”

**Use Data Factory for data ingestion (source to raw-data)**

* In the search bar, search for “Data factories” > Open it in a new tab > “Create data factory”
* Select our resource group (tokio-olympic) > Give a name (e.g. tokyo-data-dt)
* Select region and version > Next until the end > Create > “Go to resources” > “Launch Studio”

Left panel:

* Home: where all data factors will be showed

Une image contenant texte, Police, Bleu électrique, nombre

Description générée automatiquement

* Author: where we create the pipeline

Une image contenant texte, Police, nombre, ligne

Description générée automatiquement

* Monitor: monitor each part of the pipeline (error, …)

Une image contenant texte, capture d’écran, Police, nombre

Description générée automatiquement

* Manage: configure things

Une image contenant texte, capture d’écran, Police, nombre

Description générée automatiquement

**Start building the pipeline**

* Click on Author > “+” button > Pipeline > Pipeline > Right side, Give a name: (e.g.: data-ingestion)

Left side: we have the available activities (these are operations we want to perform): many activities available that we can integrate in our data factory

* Click on “Move and transform” > Drag and drop “Copy Data” (to copy data from the source to our data storage) > Below “Source”