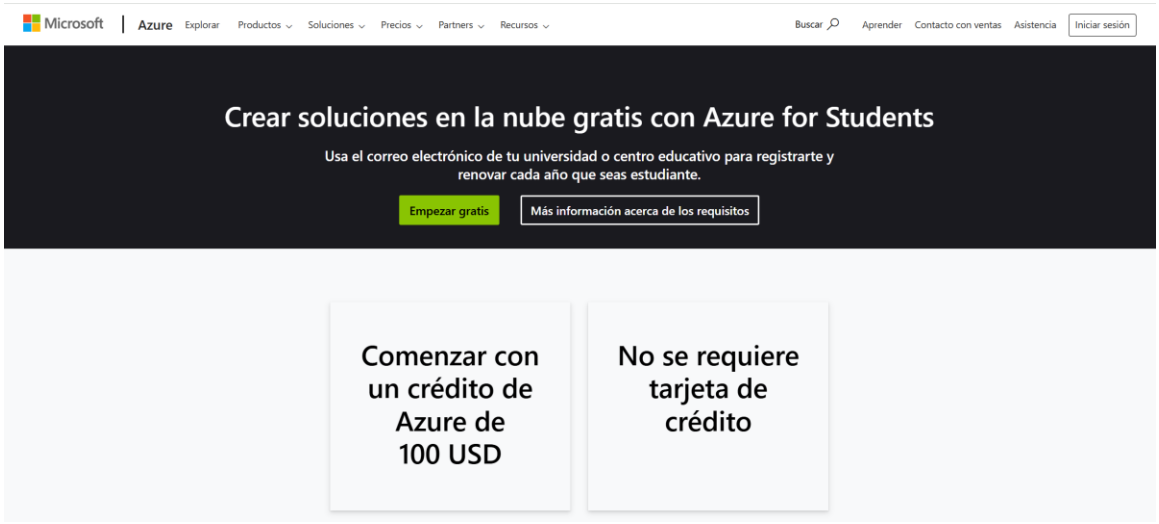
Universidad Católica Boliviana "San Pablo"
Unidad Académica Cochabamba
Departamento de Posgrado
Ing. Mauricio Alejandro Quezada Bustillo

DML-004 - Guía de Laboratorio 1

Esta guía paso a paso te ayudará a crear tu cuenta en Azure como estudiante de la Universidad Católica Boliviana, configurar tu entorno de trabajo en Azure Machine Learning Studio, crear un compute instance, y cargar un dataset desde Kaggle para trabajar con el SDK v2.

PARTE 1 – CREAR CUENTA DE ESTUDIANTE EN AZURE FOR STUDENTS

1. Ir a https://azure.microsoft.com/en-us/free/students/

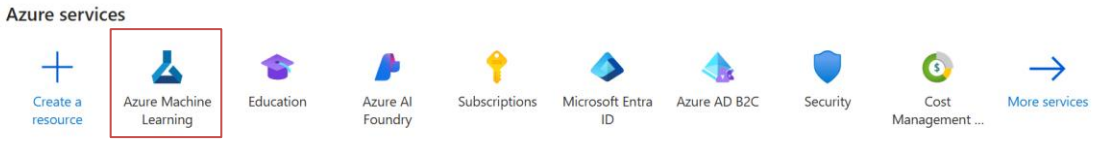


2. Presionar 'Start free' o 'Empieza gratis'.
3. Iniciar sesión con tu correo institucional (terminado en @ucb.edu.bo).
4. Completar el formulario con tus datos reales.
5. Verificar tu cuenta mediante el código enviado a tu correo.
6. Listo: ahora tienes $100 en créditos para usar en Azure por 1 año.

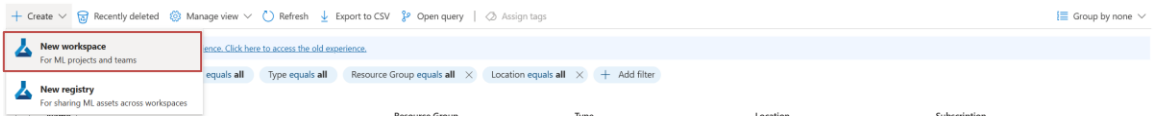
Universidad Católica Boliviana "San Pablo"
Unidad Académica Cochabamba
Departamento de Posgrado
Ing. Mauricio Alejandro Quezada Bustillo

PARTE 2 – CREAR UN WORKSPACE EN AZURE MACHINE LEARNING

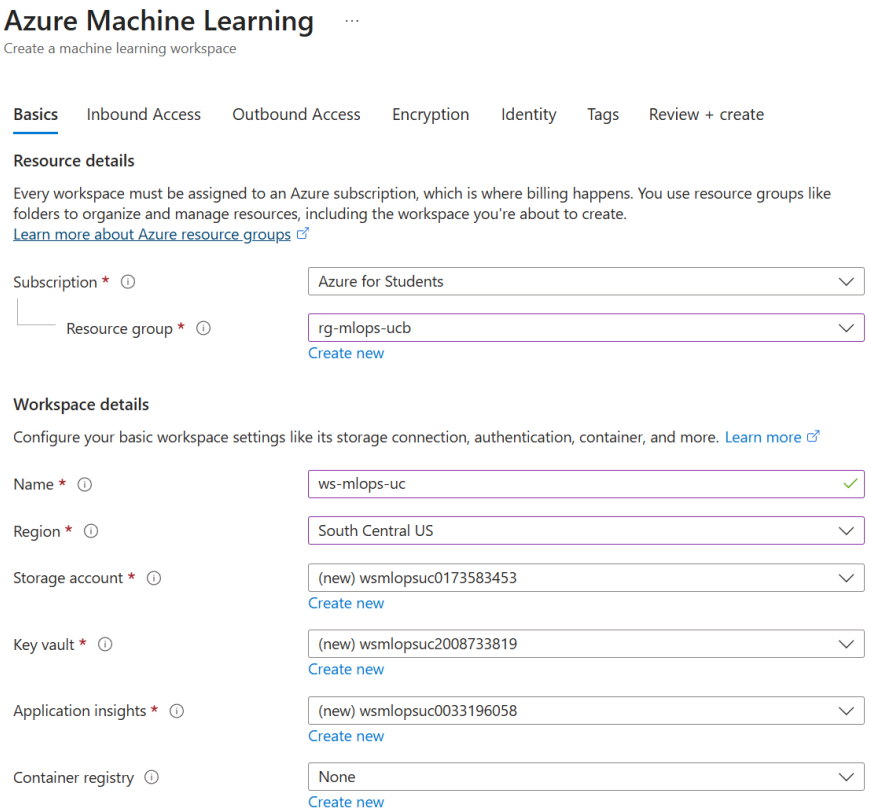
1. En el portal de Azure (https://portal.azure.com) buscar 'Machine Learning'.



2. Seleccionar 'Create'.



3. Elegir tu suscripción de Azure for Students.
4. En Resource Group, crear uno nuevo llamado rg-mlops-ucb.
5. Workspace Name: ws-mlops-ucb.
6. Region: seleccionar una región disponible (ej. South Central US).
7. Presionar 'Review + Create' y luego 'Create'.

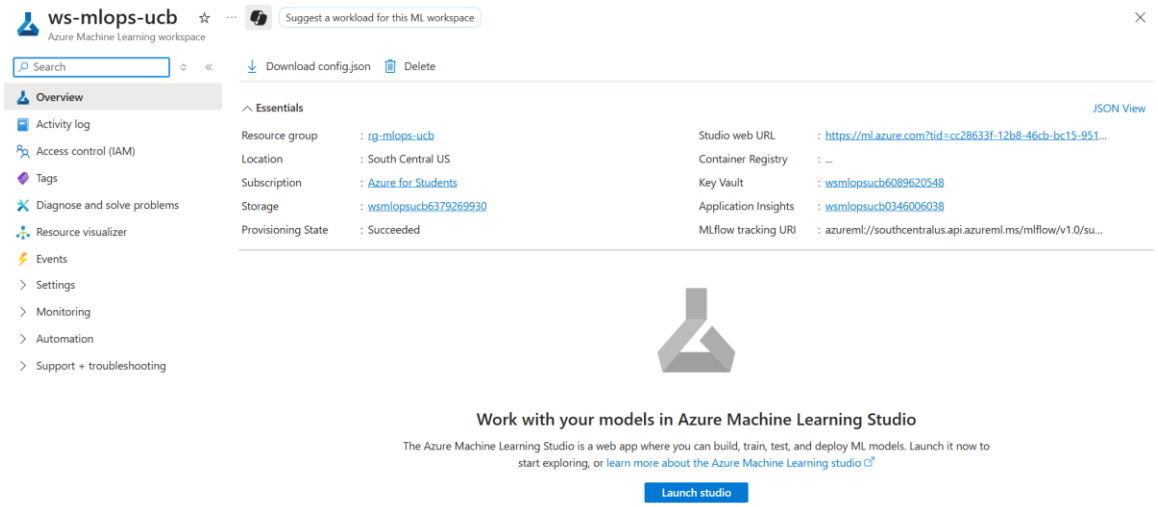


8. Esperar a que finalice el despliegue.

Universidad Católica Boliviana "San Pablo"
Unidad Académica Cochabamba
Departamento de Posgrado
Ing. Mauricio Alejandro Quezada Bustillo

PARTE 3 – ENTRAR A AZURE MACHINE LEARNING STUDIO

1. Ir a https://ml.azure.com.
2. Seleccionar el workspace creado (ws-mlops-ucb).
3. Verificar que puedas acceder al panel principal.

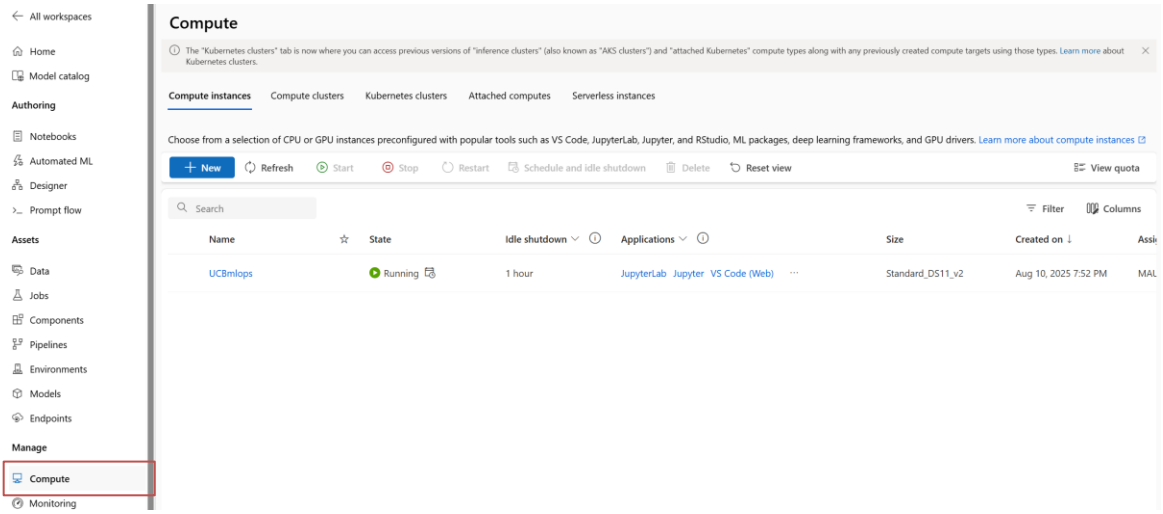


PARTE 4 – DESCARGAR DATASET DESDE KAGGLE

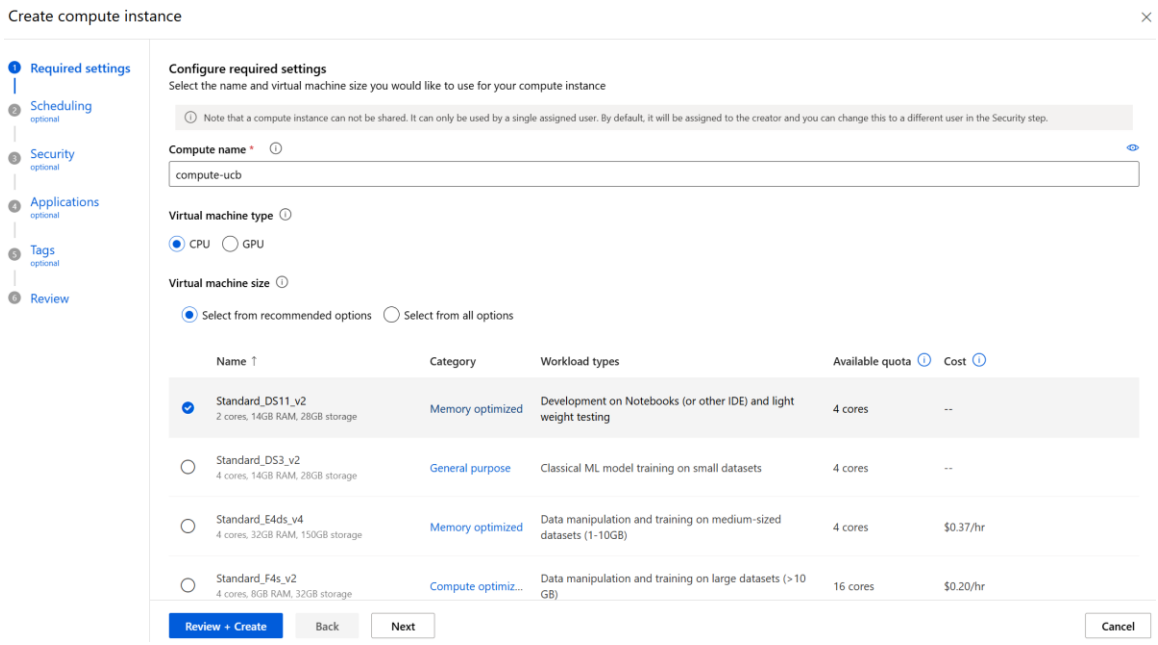
1. Ir a https://www.kaggle.com y registrarte (puede ser con Google o correo UCB).
2. Buscar el dataset 'Heart Disease' https://www.kaggle.com/datasets/johnsmith88/heart-disease-dataset?resource=download.
3. Descargar el archivo ZIP.
4. Extraer el archivo CSV en tu computadora.

PARTE 5 BIS – CREAR UN COMPUTE INSTANCE

1. En Azure Machine Learning Studio, ir a la sección 'Compute'.



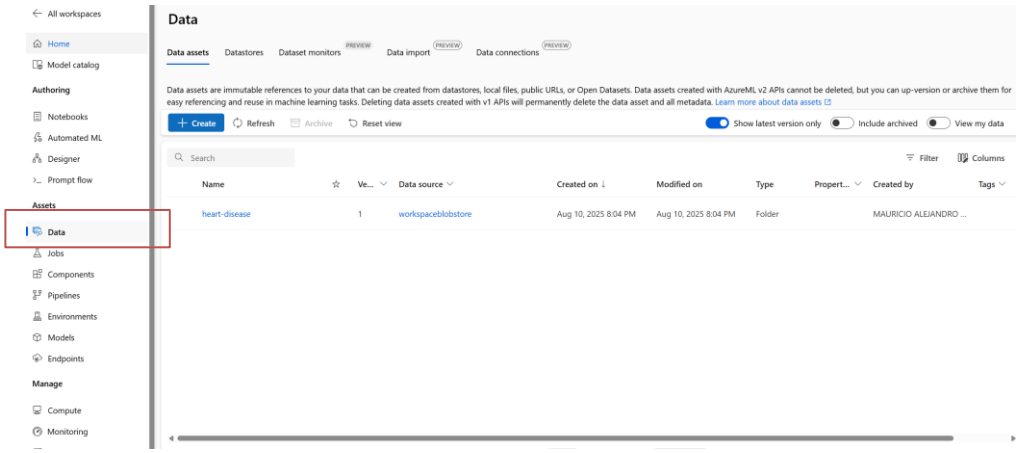
2. Seleccionar la pestaña 'Compute instances'.
3. Presionar '+ New'.
4. Configurar:
   - Nombre: compute-ucb
   - Tamaño recomendado: STANDARD_DS11_V2 (o más pequeño para ahorrar créditos).
   - Región: la misma que el workspace.
5. Presionar 'Create' y esperar que esté en estado 'Running'.



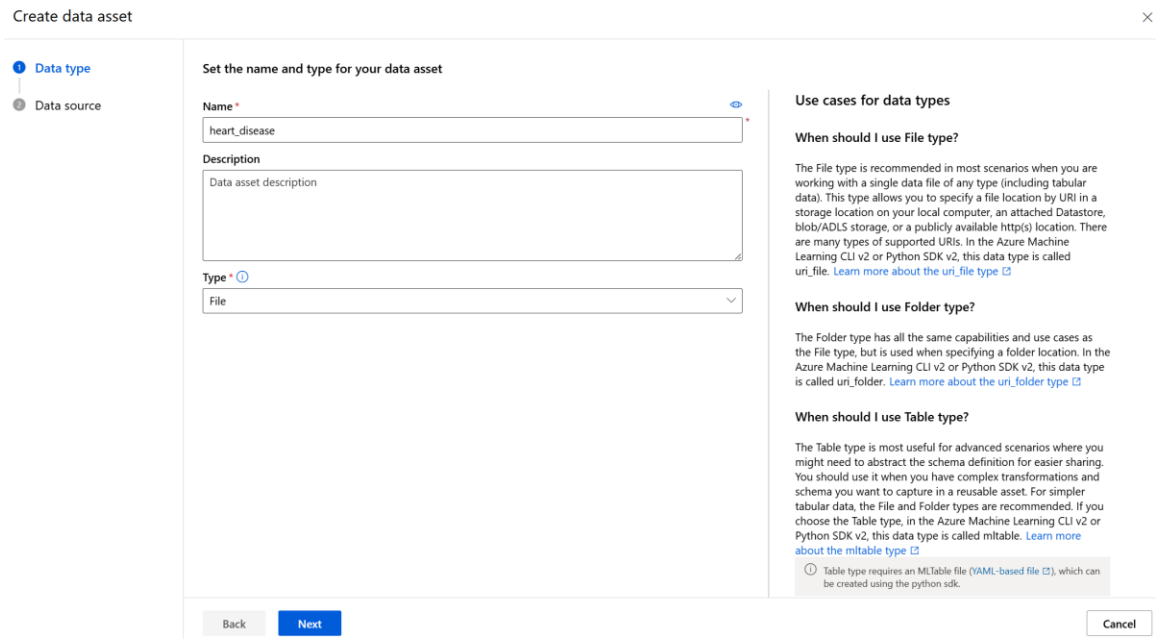
6. IMPORTANTE: al terminar, detenerlo para no gastar créditos.

PARTE 6 – SUBIR DATASET COMO DATA ASSET EN AZURE ML

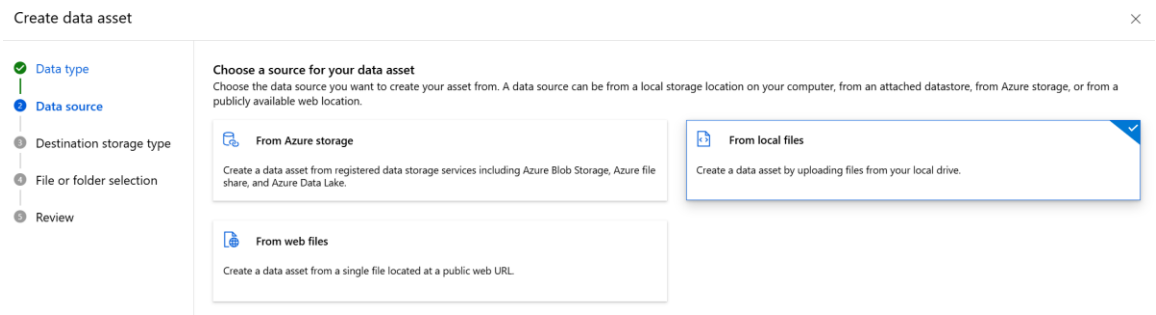
1. En Azure ML Studio, ir a la sección 'Data'.
2. Clic en '+ Create'.



3. Nombre: heart-disease
4. Tipo de archivo: File



5. Seleccionar 'From local files'.

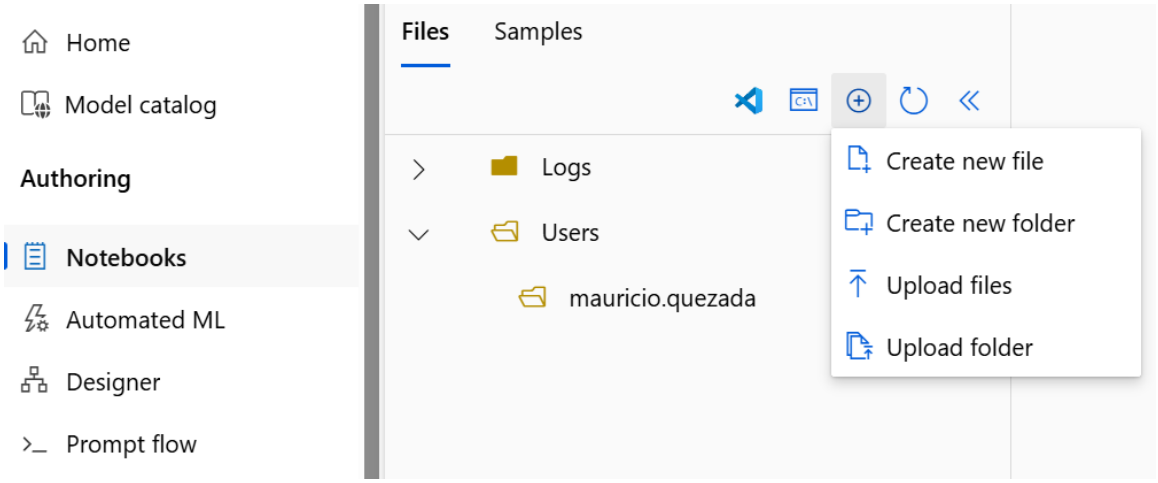


6. Cargar el CSV descargado de Kaggle.
7. Guardar.

Universidad Católica Boliviana "San Pablo"
Unidad Académica Cochabamba
Departamento de Posgrado
Ing. Mauricio Alejandro Quezada Bustillo

PARTE 7 – ABRIR NOTEBOOK Y CARGAR DATASET CON SDK V2

1. Crear un nuevo Notebook



2. Iniciar código con pandas

Código de ejemplo:

```
import pandas as pd
from azure.ai.ml import MLClient
from azure.identity import DefaultAzureCredential

ml_client = MLClient.from_config(credential=DefaultAzureCredential())
data_asset = ml_client.data.get('heart-disease', version='1')
df = pd.read_csv(data_asset.path)
df.head()
```

Este código conecta tu notebook con Azure ML, obtiene el dataset 'heart-disease' y lo carga en un DataFrame de Pandas.