**BÀI TẬP TƯƠNG TÁC VỚI PYSPARK**

**I. Set up Dataproc để tạo cluster:**

**1) Enable Dataproc Metastore: 🡪 Click enable the API**

**A screenshot of a computer

Description automatically generated**

**A computer screen shot of a computer

Description automatically generated**

**2) Create Dataproc:**

**A screenshot of a computer

Description automatically generated**

**A screenshot of a computer

Description automatically generated**

**A screenshot of a computer

Description automatically generated**

**A screenshot of a computer

Description automatically generated**

**A screenshot of a computer

Description automatically generated**

**II. Tương tác với PySpark:**

**1) Chọn vào cluster vừa tạo:**

**A screenshot of a computer

Description automatically generated**