# Hive and PySpark Example

## Step 1: Create the Hive Table and Insert Data

Run these Hive commands in your Hive CLI:  
  
CREATE TABLE customers (  
 id INT,  
 name STRING,  
 age INT  
);  
  
INSERT INTO customers (id, name, age) VALUES   
 (1, 'John Doe', 30),  
 (2, 'Jane Smith', 25);

## Step 2: Query the Hive Table with PySpark

Now, use this PySpark code to connect to the Hive table and query it:  
  
from pyspark.sql import SparkSession  
  
# Step 2.1: Create Spark session with Hive support  
spark = SparkSession.builder .appName("PySpark Hive Example") .enableHiveSupport() .getOrCreate()  
  
# Step 2.2: Query the Hive table  
df = spark.sql("SELECT \* FROM customers")  
  
# Step 2.3: Show the result  
df.show()  
  
# Step 2.4: Stop the Spark session (optional)  
spark.stop()