DAY 1 (09/01/26)– Platform Setup & First Steps

🗓️ 09/01/2026

What I learned

- Why Databricks is preferred over Pandas and Hadoop for large-scale data
- Basics of Lakehouse architecture
- Databricks workspace components (Workspace, Compute, Data Explorer)
- How Databricks is used by companies for large-scale analytics
- Ran my first PySpark commands

What I worked on

- Created a Databricks Community Edition account
- Explored workspace navigation and compute setup
- Created my first Databricks notebook
- Executed basic PySpark DataFrame operations

📓 **Notebooks**

- notebooks/day1_databricks_basics.ipynb

Practice

```
data = [("iPhone", 999), ("Samsung", 799), ("MacBook", 1299)]

df = spark.createDataFrame(data, ["product", "price"])

df.show()

df.filter(df.price > 1000).show()
```