**SPARK UI**

A screenshot of a computer

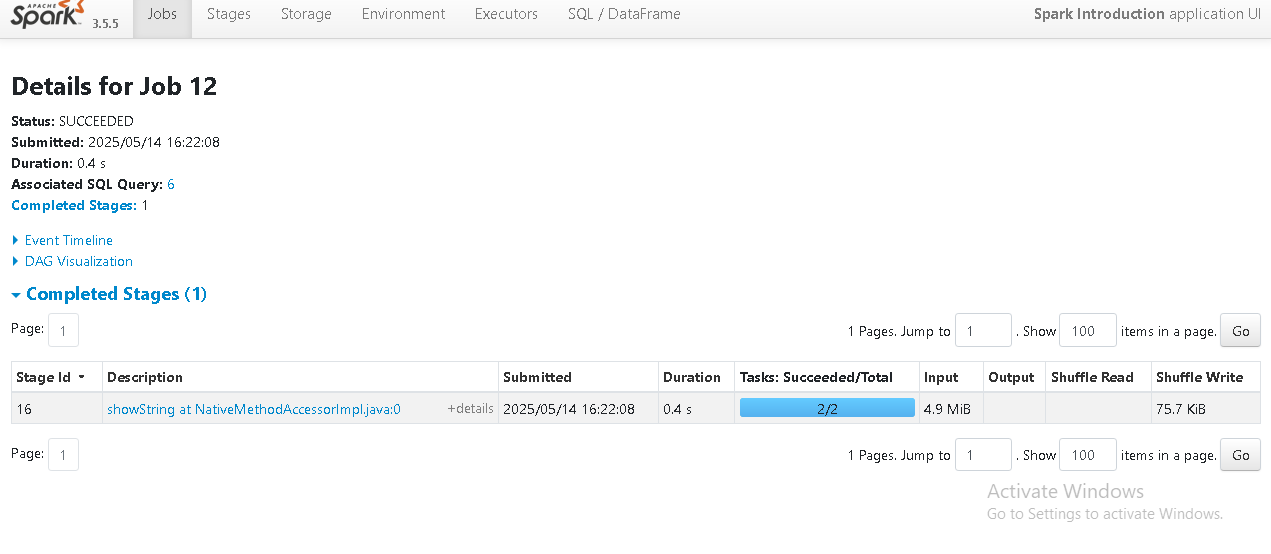
AI-generated content may be incorrect.

**This is the Apache Spark Web UI (version 3.5.5) showing the "Jobs" tab under the "Spark Introduction" application.**

**It displays completed Spark jobs, showing Job IDs 10 to 13, all involving the method showString in Java.**

**Each job includes details like submission time, duration, number of succeeded stages, and task success rates.**

**All jobs completed successfully, with some tasks/stages being skipped, indicating Spark optimization or cached results**



This is the detailed view of Job 12 from the Apache Spark UI (version 3.5.5), showing the job **succeeded** with a duration of **0.4 seconds**.

It was submitted on **2025/05/14 at 16:22:08** and is associated with **SQL Query: 6**

Only one stage (Stage ID: 16) was involved, executing showString at a native method with **2/2 tasks succeeded**.

The job processed **4.9 MiB of input data**, with **75.7 KiB shuffle write**, indicating minimal data shuffling.

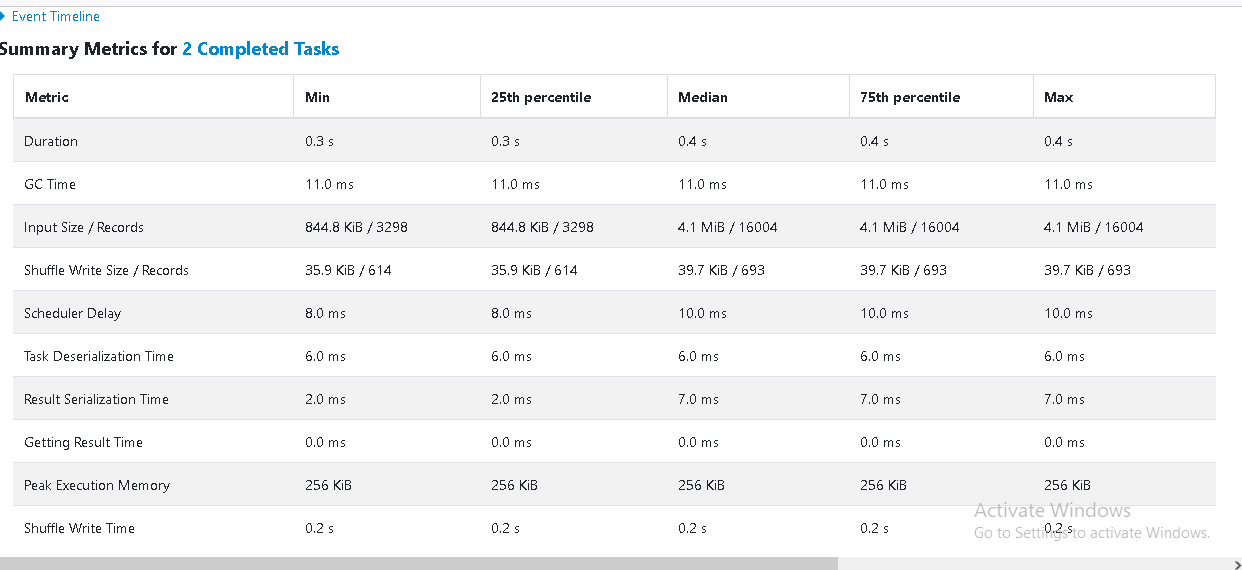


The Directed Acyclic Graph (DAG) shows the execution plan of Stage 16 in a Spark job.

process begins with **Scan csv**, indicating data is read from a CSV file.

Next, **WholeStageCodegen (1)** is used, meaning Spark optimized the query execution using whole-stage code generation.

Finally, an **Exchange** node appears, which typically represents a **shuffle** or repartitioning step in Spark.



The task duration ranged from **0.3 to 0.4 seconds**, with consistent **GC time of 11 ms** across tasks.

**Input size** varied significantly, from **844.8 KiB / 3,298 records** to **4.1 MiB / 16,004 records**.

**Shuffle write size** was between **35.9 KiB and 39.7 KiB**, with a consistent record count of **614–693**.

Memory usage was stable at **256 KiB**, and **task (de)serialization and scheduler delays** were minimal, indicating efficient execution.