

## ASSIGNMENT – V

### TITLE :

Develop and Implement Applications Using Apache Spark.

NAME : Shinde Shubham Dnyandeve.

ROLL NO : 23107121.

CLASS : TY-B

### PROGRAM :

#### INPUT :

I am Student of VPKBIET. I am Studying in AI&DS Department.

#### CODE :

```
from pyspark import SparkContext

sc = SparkContext("spark://localhost:7077", "WordCount")

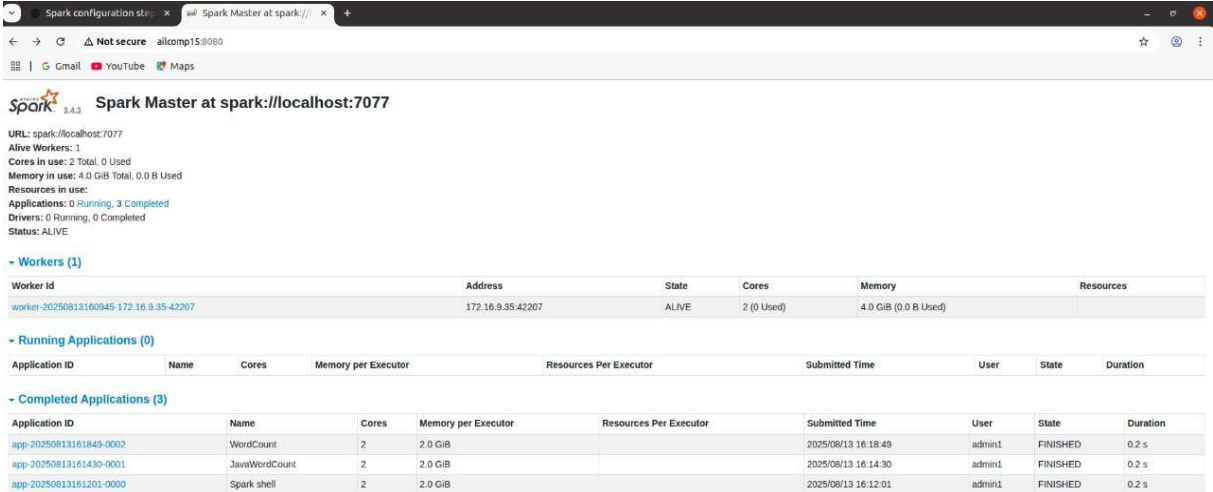
text_file = sc.textFile("/mnt/spark/README.md")

counts = text_file.flatMap(lambda line: line.split()) \
    .map(lambda word: (word, 1)) \
    .reduceByKey(lambda a, b: a + b)

# Save output
counts.saveAsTextFile("/home/admin1/wordcount_output")

sc.stop()
```

### OUTPUT :



The screenshot shows the Spark Master web interface at `spark://localhost:7077`. The interface displays the following information:

- URL:** `spark://localhost:7077`
- Alive Workers:** 1
- Cores in use:** 2 Total, 0 Used
- Memory in use:** 4.0 GiB Total, 0.0 B Used
- Resources in use:**
- Applications:** 0 Running, 3 Completed
- Drivers:** 0 Running, 0 Completed
- Status:** ALIVE

**Workers (1)**

Worker Id	Address	State	Cores	Memory	Resources
worker-20250813160945-172.16.9.35-42207	172.16.9.35:42207	ALIVE	2 (0 Used)	4.0 GiB (0.0 B Used)	

**Running Applications (0)**

Application ID	Name	Cores	Memory per Executor	Resources Per Executor	Submitted Time	User	State	Duration
----------------	------	-------	---------------------	------------------------	----------------	------	-------	----------

**Completed Applications (3)**

Application ID	Name	Cores	Memory per Executor	Resources Per Executor	Submitted Time	User	State	Duration
app-20250813161849-0002	WordCount	2	2.0 GiB		2025/08/13 16:18:49	admin1	FINISHED	0.2 s
app-20250813161430-0001	JavaWordCount	2	2.0 GiB		2025/08/13 16:14:30	admin1	FINISHED	0.2 s
app-20250813161201-0000	Spark shell	2	2.0 GiB		2025/08/13 16:12:01	admin1	FINISHED	0.2 s