Install spark
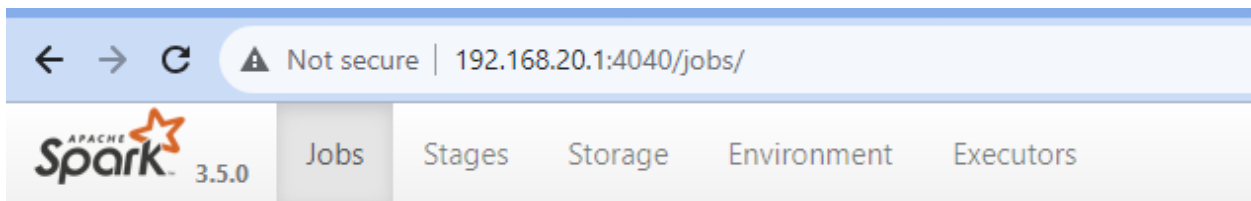https://www.oracle.com/java/technologies/downloads/#jdk17-windows

```
scala> val numbers = List(1,2,3,4,5)
numbers: List[Int] = List(1, 2, 3, 4, 5)
```

```
scala> val numbers = List(1,2,3,4,5)
numbers: List[Int] = List(1, 2, 3, 4, 5)

scala> val numbersDF = sc.parallelize(numbers).toDF("Id")
numbersDF: org.apache.spark.sql.DataFrame = [Id: int]

scala> numbersDF.show()
+---+
| Id|
+---+
|  1|
|  2|
|  3|
|  4|
|  5|
+---+
```

← → C   ⚠ Not secure | 192.168.20.1:4040/jobs/

Spark 3.5.0   Jobs   Stages   Storage   Environment   Executors

Spark Jobs (?)

User: thanhnga
Total Uptime: 12 s
Scheduling Mode: FIFO

▶ Event Timeline

Apache Spark 3 ✕ | Java Downloads ✕ | Untitled docume ✕ | Downloads ✕ | Spark: Trying to ✕ | Apache Spark 3 ✕ | Spark shell - Exe ✕ | +

← → C | ⚠ Not secure | 192.168.20.1:4040/executors/

Spark 3.5.0   Jobs   Stages   Storage   Environment   Executors   **Spark shell** application UI

Executors

▸ Show Additional Metrics

Summary

	RDD Blocks	Storage Memory	Disk Used	Cores	Active Tasks	Failed Tasks	Complete Tasks	Total Tasks	Task Time (GC Time)	Input	Shuffle Read	Shuffle Write	Excluded
Active(1)	0	0.0 B / 434.4 MiB	0.0 B	2	0	0	0	0	17 min (0.3 s)	0.0 B	0.0 B	0.0 B	0
Dead(0)	0	0.0 B / 0.0 B	0.0 B	0	0	0	0	0	0.0 ms (0.0 ms)	0.0 B	0.0 B	0.0 B	0
Total(1)	0	0.0 B / 434.4 MiB	0.0 B	2	0	0	0	0	17 min (0.3 s)	0.0 B	0.0 B	0.0 B	0

Executors

Show 20 entries   Search:

Executor ID	Address	Status	RDD Blocks	Storage Memory	Disk Used	Cores	Active Tasks	Failed Tasks	Complete Tasks	Total Tasks	Task Time (GC Time)	Input	Shuffle Read	Shuffle Write	Thread Dump	Heap Histogram
driver	192.168.20.1:58618	Active	0	0.0 B / 434.4 MiB	0.0 B	2	0	0	0	0	17 min (0.3 s)	0.0 B	0.0 B	0.0 B	Thread Dump	Heap Histogram