Step 1: Install Java (if not already)

Apache Spark needs Java. Install OpenJDK:

bash
CopyEdit
```bash
sudo apt update
sudo apt install openjdk-11-jdk
```

Verify it:

bash
CopyEdit
```bash
java -version
```

🔧 Step 2: Install Apache Spark

1. Download Spark (pre-built for Hadoop):

bash
CopyEdit
```bash
wget
https://dlcdn.apache.org/spark/spark-3.5.0/spark-3.5.0-bin-hadoop3.tgz
```

2. Extract and move to /opt:

bash
CopyEdit
```bash
tar -xvzf spark-3.5.0-bin-hadoop3.tgz
sudo mv spark-3.5.0-bin-hadoop3 /opt/spark
```

3. Set environment variables. Add to your ~/.bashrc:

bash
CopyEdit

```bash
export SPARK_HOME=/opt/spark
export PATH=$PATH:$SPARK_HOME/bin
export JAVA_HOME=/usr/lib/jvm/java-11-openjdk-amd64
```

Reload it:

bash
CopyEdit
```bash
source ~/.bashrc
```

✅ Step 3: Run `spark-shell`

bash
CopyEdit
```bash
spark-shell
```

🧠 Paste this logic inside `spark-shell`:

scala
CopyEdit
```scala
val numbers = List(-10, 0, 5, -3, 7)
val rdd = sc.parallelize(numbers)
val result = rdd.map { num =>
  if (num > 0) s"$num is Positive"
  else if (num < 0) s"$num is Negative"
  else s"$num is Zero"
}
result.collect().foreach(println)
```