

Ex. No. 1	Building a Basic Spark Application
Youtube Link	https://youtu.be/rotevpYMdTI
Date of Exercise	6.10.25

AIM

To build a basic Spark application that creates a DataFrame, applies transformations such as filtering and column selection, and displays the results.

Procedure:

- 1. Install and configure Apache Spark.**
- 2. Launch Spark (PySpark shell / Jupyter / VS Code).**
- 3. Create a SparkSession.**
- 4. Create a DataFrame from a list of records.**
- 5. Apply DataFrame operations:**
- 6. filter() to select specific rows**
- 7. select() to choose specific columns**

Display the result using show().

Program:

```
from pyspark.sql import SparkSession

# Step 1: Create Spark session

spark = SparkSession.builder.appName("BasicSparkApp").getOrCreate()

# Step 2: Create DataFrame

data = [
    ("Alice", 22, "CSE"),
    ("Bob", 25, "ECE"),
    ("Charlie", 23, "CSE"),
    ("David", 21, "MECH")
]

columns = ["Name", "Age", "Department"]

df = spark.createDataFrame(data, columns)

# Step 3: Filter and select

result = df.filter(df.Department == "CSE").select("Name", "Age")

# Step 4: Display output

result.show()

# Stop Spark session

spark.stop()
```

Output:

```
+-----+---+
| Name | Age |
+-----+---+
| Alice| 22|
| Charlie| 23|
+-----+---+
```

Result :

A basic Spark application was successfully created