```
!pip install pyspark
```

```
Requirement already satisfied: pyspark in /usr/local/lib/python3.11/dist-packages (3.5.4)
Requirement already satisfied: py4j==0.10.9.7 in /usr/local/lib/python3.11/dist-packages (from pyspark) (0.10.9.7)
```

```
from pyspark.sql import SparkSession
```

```
spark = SparkSession.builder.appName("Colab_Spark").getOrCreate()
```

```
data = [(1, "Ayush", 25), (2, "Aryan", 30), (3, "Rahul", 35)]
columns = ["id", "name", "age"]

df = spark.createDataFrame(data, columns)
df.show()
```

```
+---+-----+---+
| id| name|age|
+---+-----+---+
|  1|Ayush| 25|
|  2|Aryan| 30|
|  3|Rahul| 35|
+---+-----+---+
```

```
filtered_df = df.filter(df.age > 25)
filtered_df.show()
```

```
+---+-----+---+
| id| name|age|
+---+-----+---+
|  2|Aryan| 30|
|  3|Rahul| 35|
+---+-----+---+
```

```
spark.stop()
```