**Big Data Programming**

**CSEE5590-0004/490-0004**

**Lab 3 Report**

Team : 6

Chandrasekhar Pentakota (class ID : 9)

Mustafa Bilal (class ID : 15)

**Links :**

<https://github.com/chanduhub/BigDataProgramming/wiki>

<https://youtu.be/8tPUolm89YY>

https://github.com/bmian93/BigDataProgramming/tree/master/BigDataProgramming/Lab3

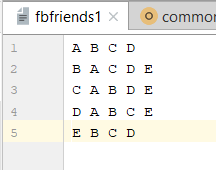
**Task 1**

**Objective :** To find Facebook common friends using Apache Spark framework and store the results to a file for easy lookup.

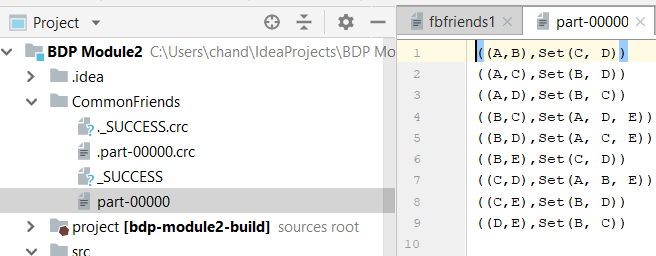
**Feautures :** This model can work on huge amounts of data of facebook profiles and find the mutual friends between the users.

**Steps :**

**Input –**

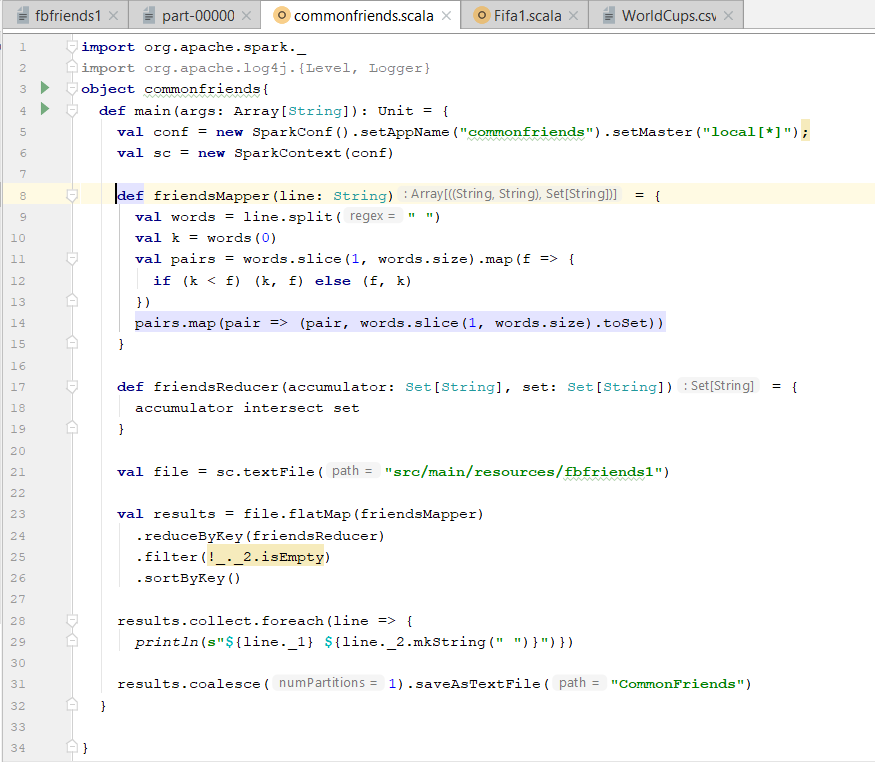
****

**Output –**

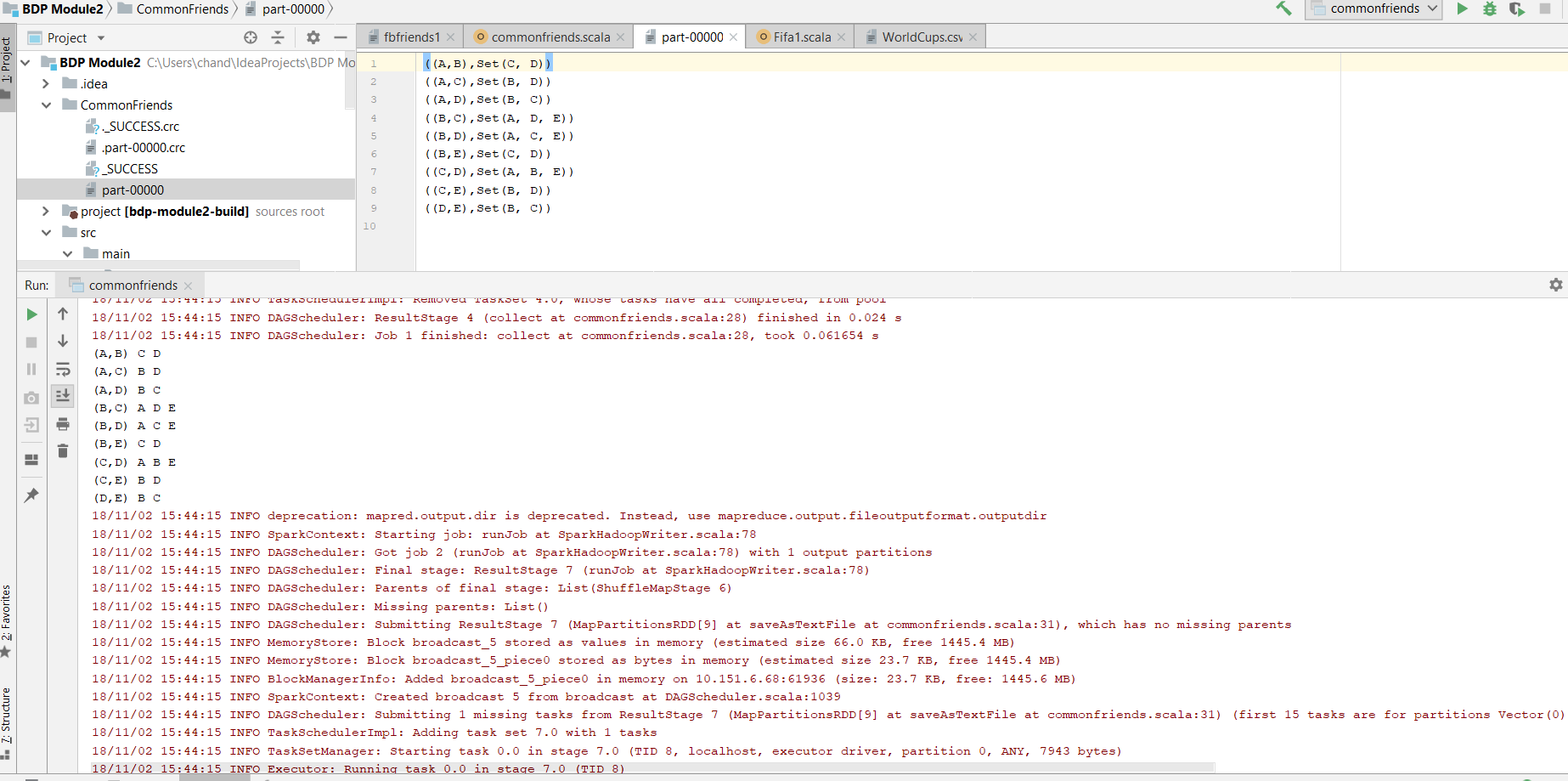
****

Here in input left side to arrow are the users and right side ones are his friends and are bidirectional that means if A is a friend of B then B is a friend of A. And in output file, we can see the mutual friends of any two users.

**Code :**

****

**Output :**

****

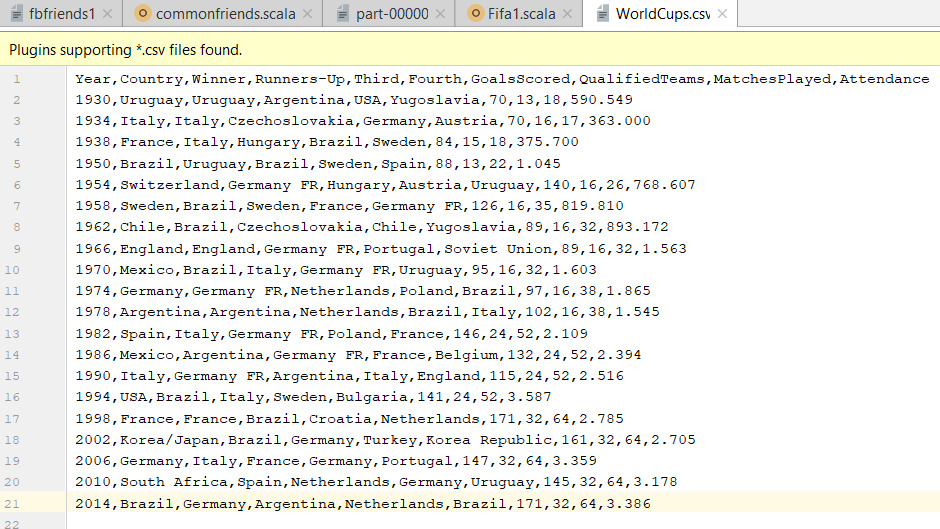
**Task 2**

**Objective :** To analyse FIFA World Cup data using spark DataFrame and spark RDDs.

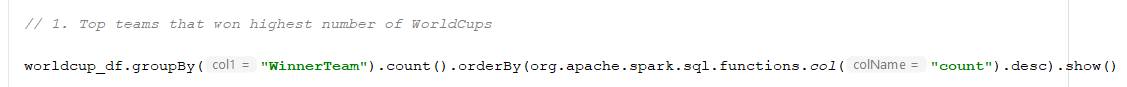
**Feautures :** This program can work on huge amounts of log data or any files and bring insights from them.

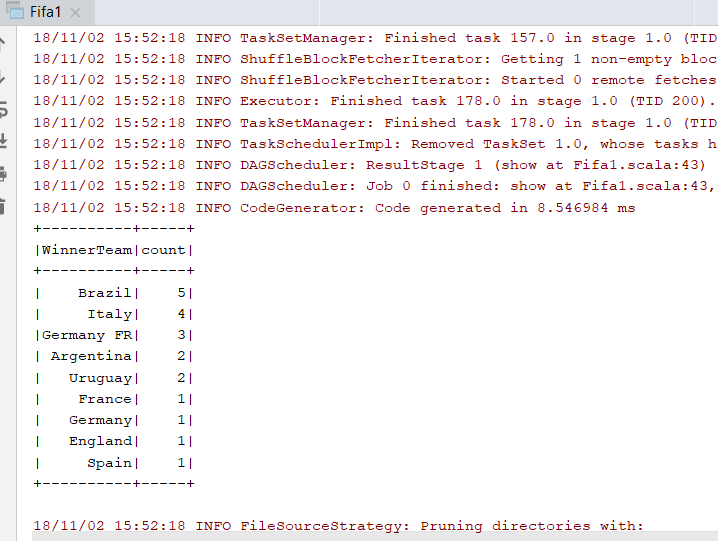
**Steps :**

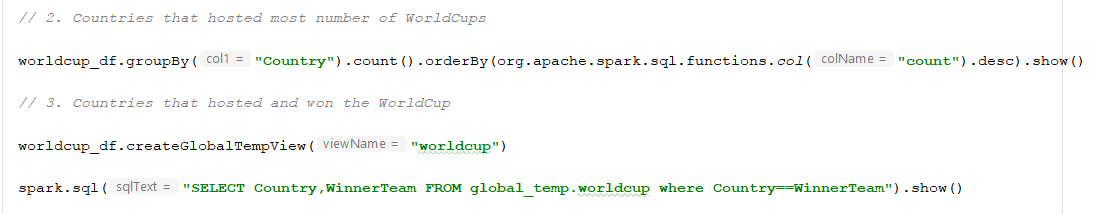
**Input –**

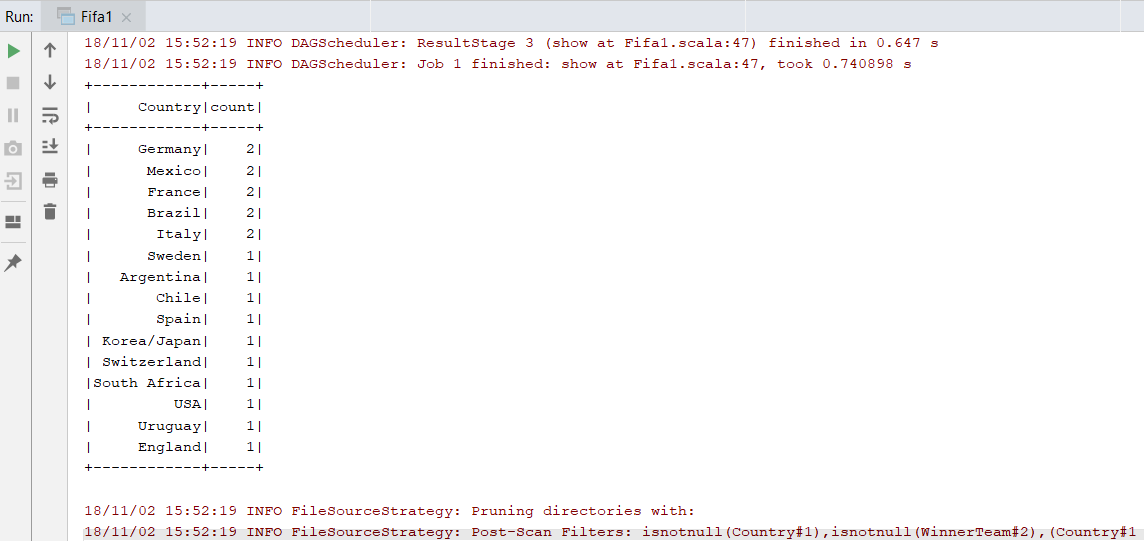
****

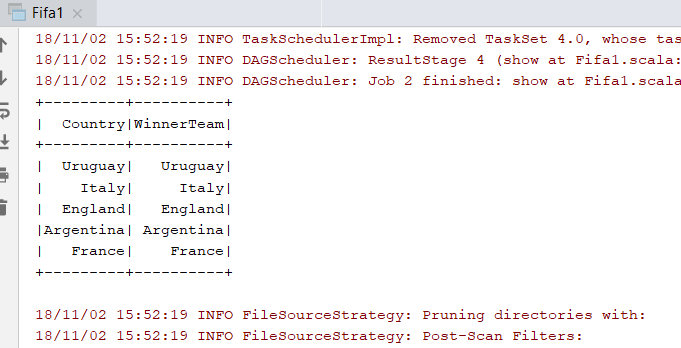
**Output –**

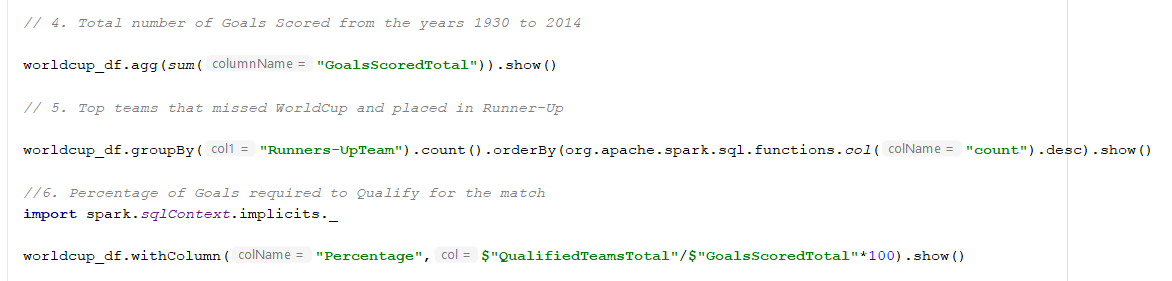
****

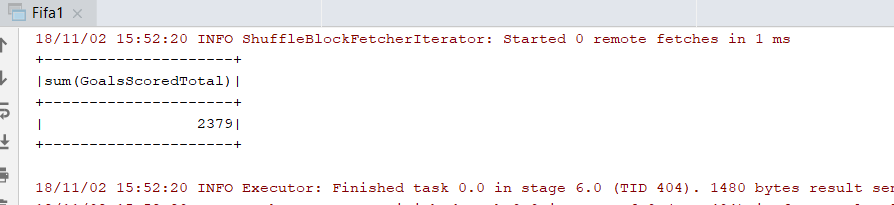
****

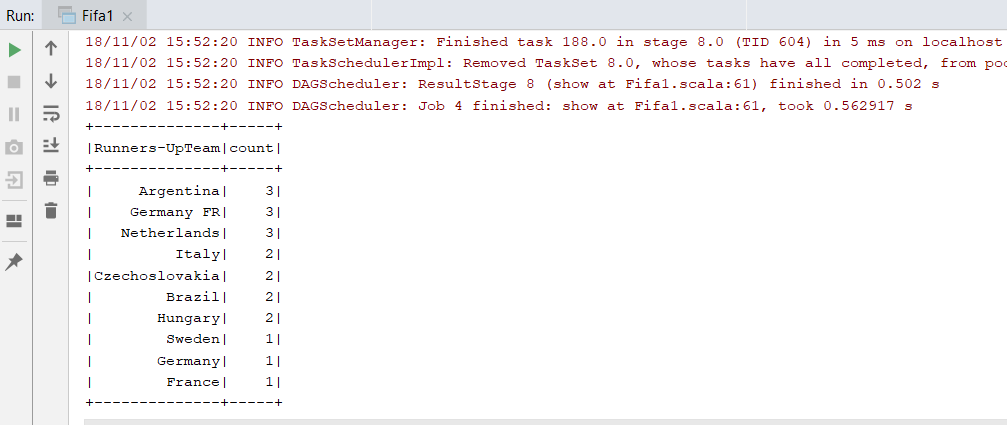


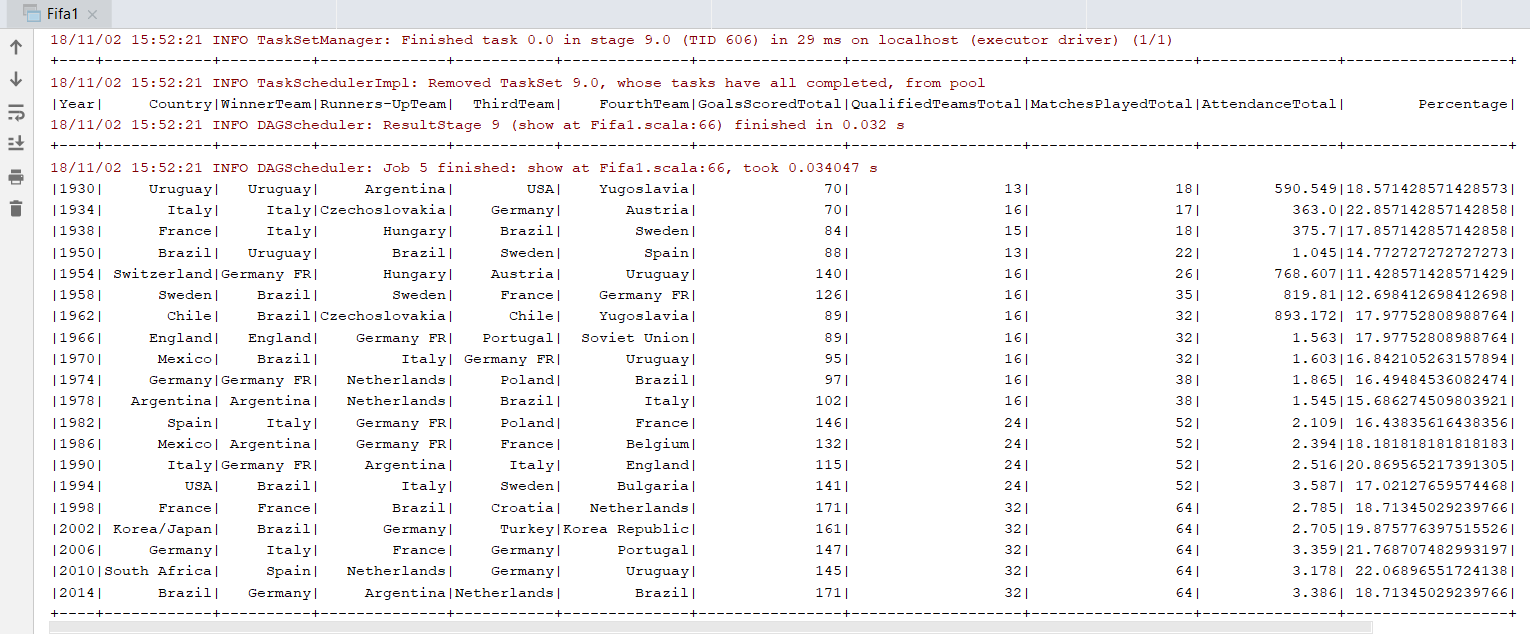


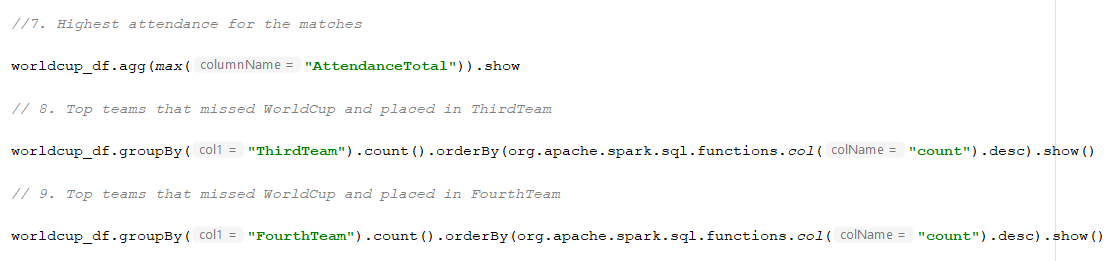


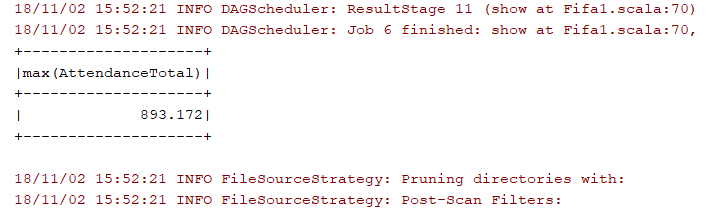


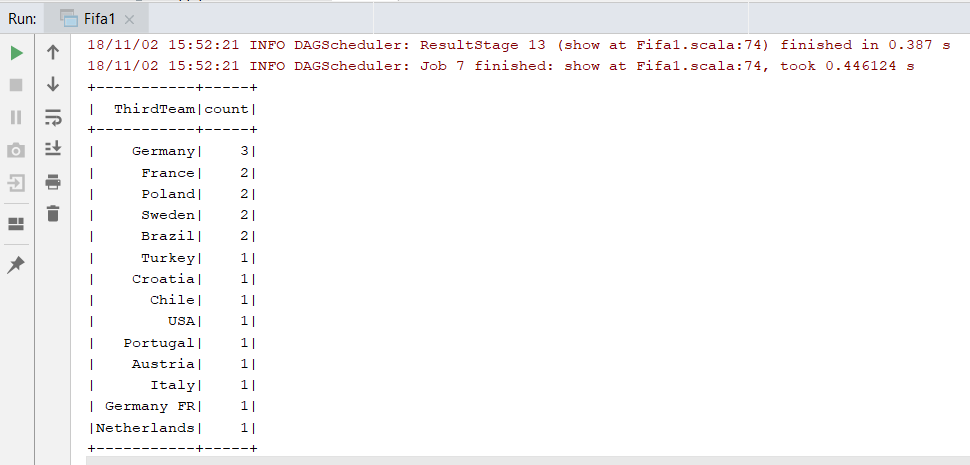


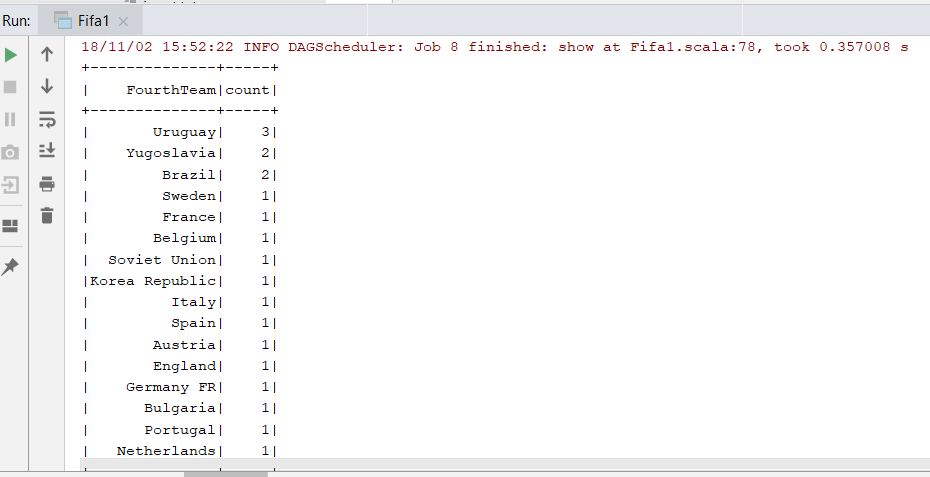


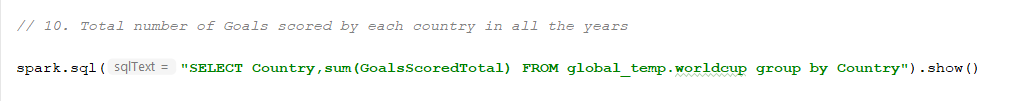


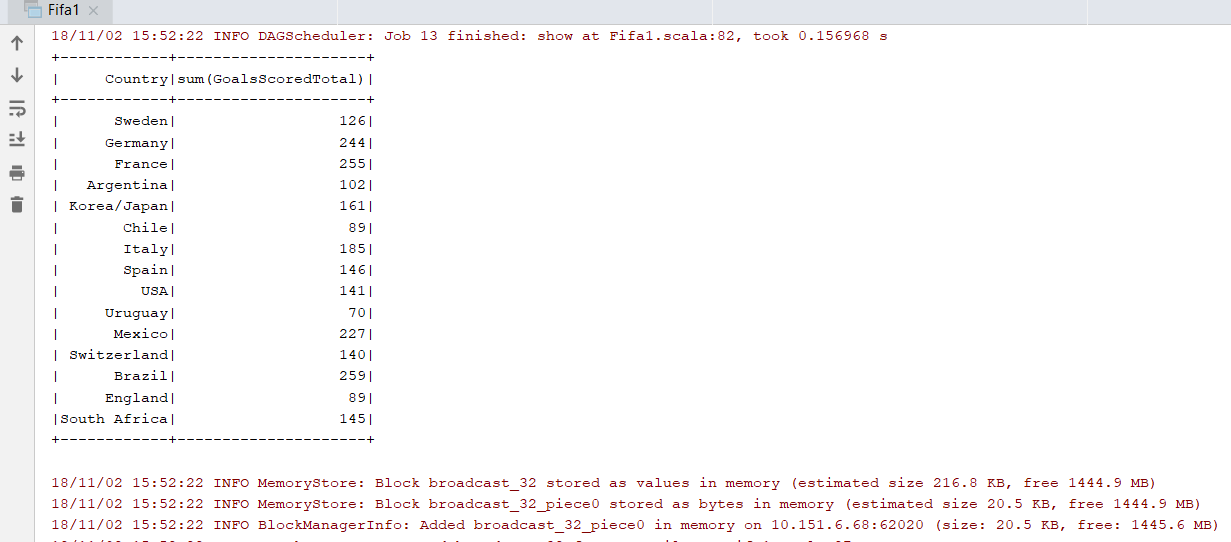


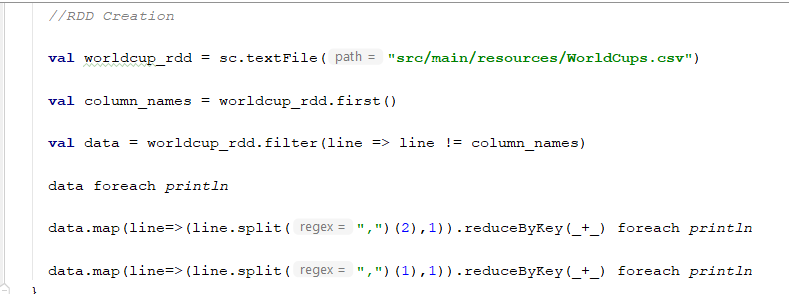




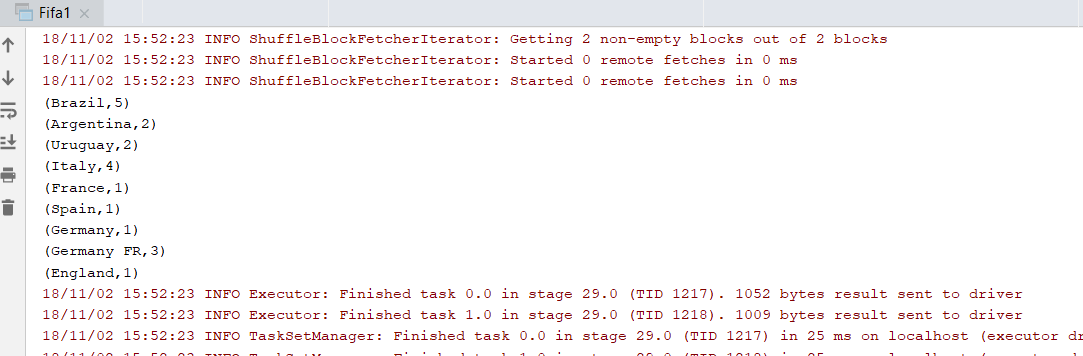


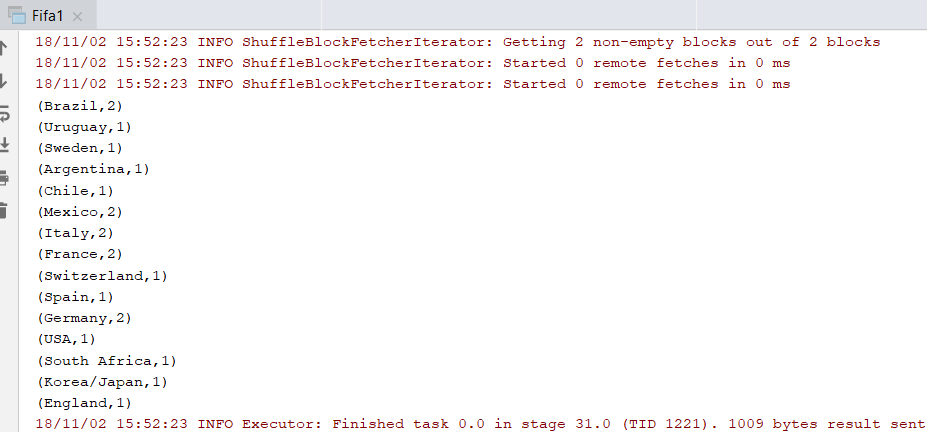




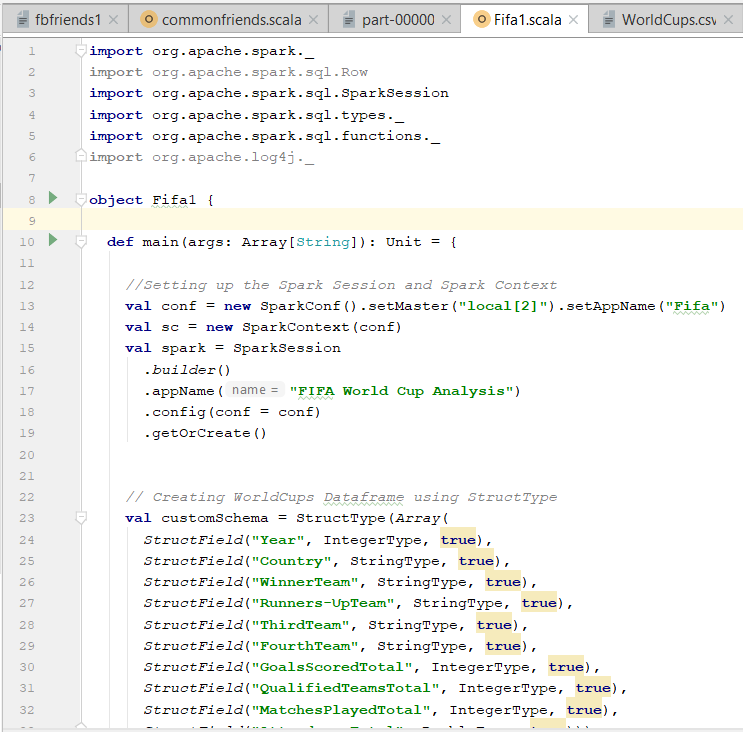


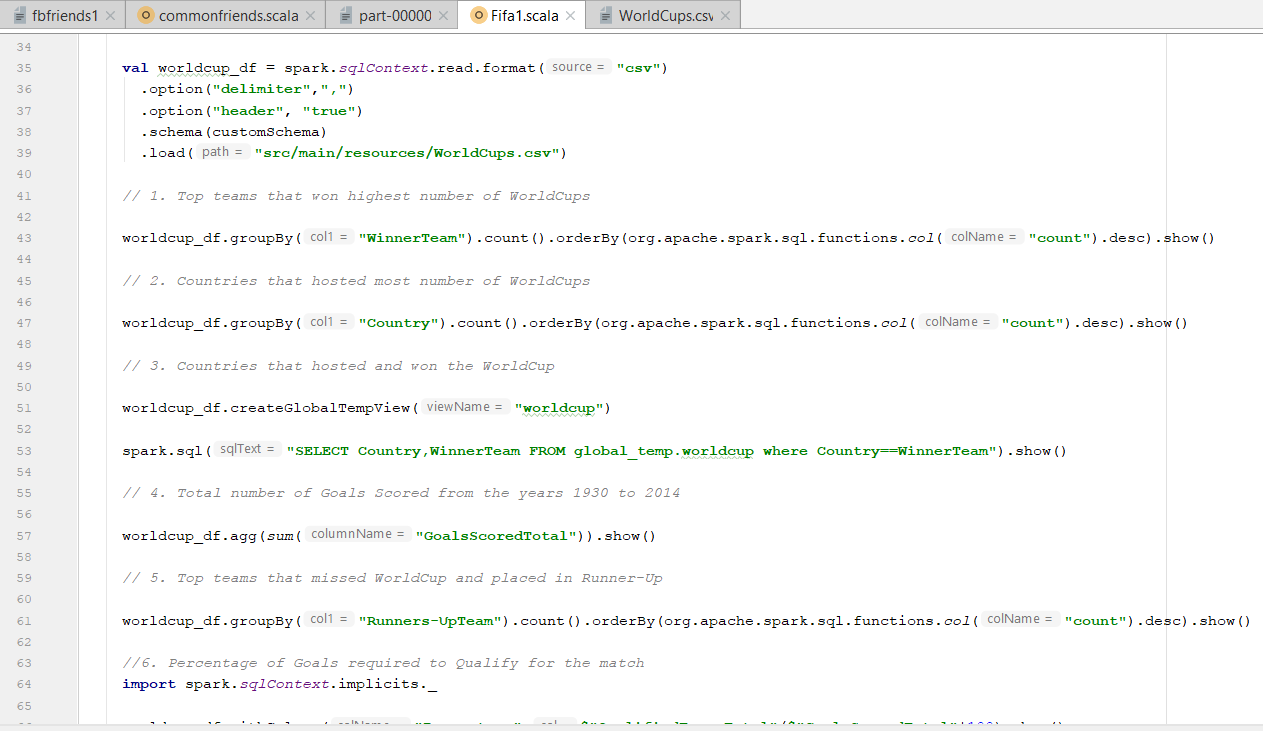


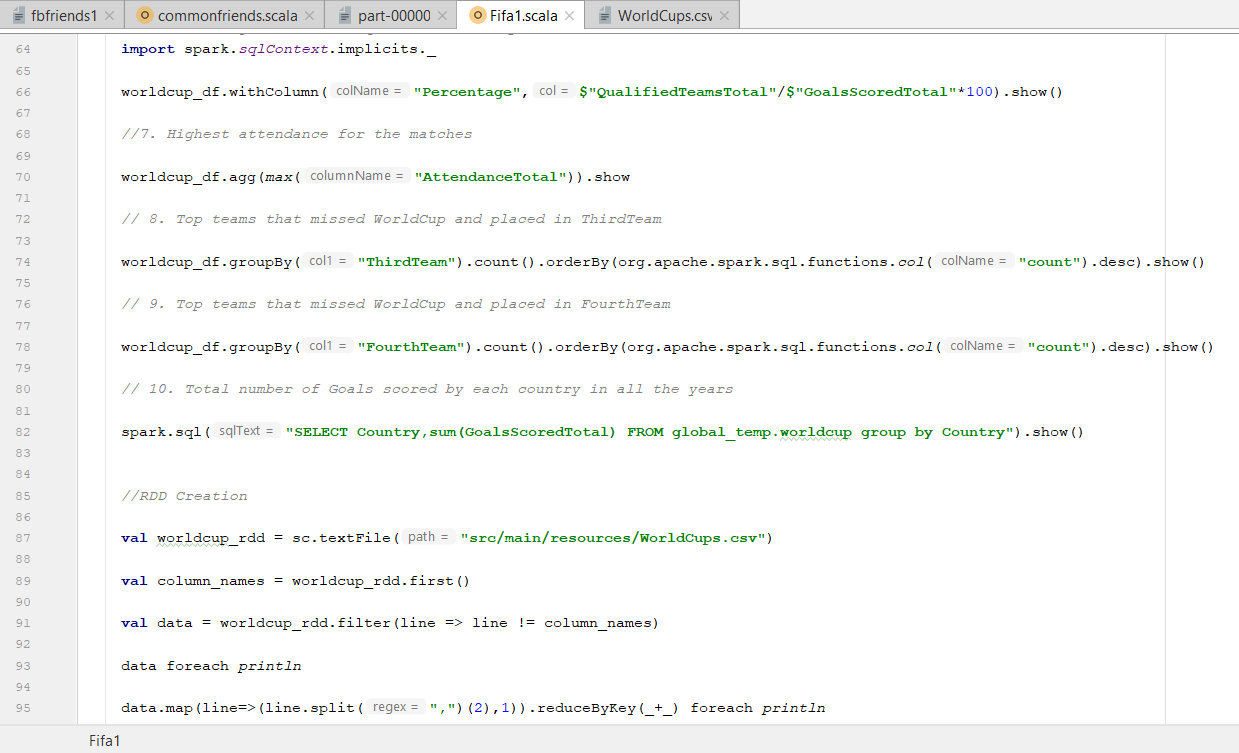




**Code :**

****

****

****