

ReadCSV.scala

```
import org.apache.spark.sql.SparkSession

object ReadCSV {

  def main(args: Array[String]): Unit = {

    val spark = SparkSession.builder()

      .appName("CSV Reader")

      .master("local[*]")

      .getOrCreate()

    val df = spark.read

      .option("header", "true")

      .option("inferSchema", "true")

      .csv("src/main/resources/data.csv")

    df.show()

    spark.stop()

  }

}
```

build.sbt

```
name := "ScalaCSVProject"

version := "0.1"

scalaVersion := "2.13.12"

libraryDependencies ++= Seq(

  "org.apache.spark" %% "spark-core" % "3.5.0",

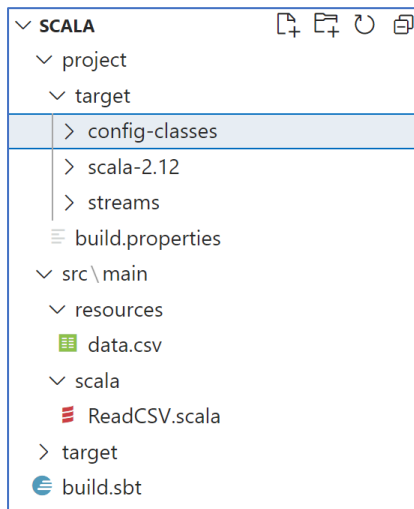
  "org.apache.spark" %% "spark-sql" % "3.5.0"

)
```

data.csv

```
id,name,age
1,John,25
2,Alice,30
3,Bob,28
```

OUTPUT:



```
C:\Windows\System32\cmd.e x + v
25/04/17 09:46:48 INFO DAGScheduler: Job 2 finished: show at ReadCSV.scala:15, took 0.168407 s
25/04/17 09:46:48 INFO CodeGenerator: Code generated in 25.9513 ms
+---+-----+
| id| name|age|
+---+-----+
|  1| John| 25|
|  2| Alice| 30|
|  3|  Bob| 28|
+---+-----+

25/04/17 09:46:48 INFO SparkContext: SparkContext is stopping with exitCode 0.
25/04/17 09:46:48 INFO SparkUI: Stopped Spark web UI at http://Vedpathak:4040
25/04/17 09:46:48 INFO MapOutputTrackerMasterEndpoint: MapOutputTrackerMasterEndpoint stopped!
25/04/17 09:46:48 INFO MemoryStore: MemoryStore cleared
25/04/17 09:46:48 INFO BlockManager: BlockManager stopped
25/04/17 09:46:48 INFO BlockManagerMaster: BlockManagerMaster stopped
25/04/17 09:46:48 INFO OutputCommitCoordinator$OutputCommitCoordinatorEndpoint: OutputCommitCoordinator s
topped!
25/04/17 09:46:48 INFO SparkContext: Successfully stopped SparkContext
[success] Total time: 52 s, completed 17 Apr, 2025 9:46:48 AM
25/04/17 09:46:49 INFO ShutdownHookManager: Shutdown hook called
25/04/17 09:46:49 INFO ShutdownHookManager: Deleting directory C:\Users\PC\AppData\Local\Temp\spark-d144c
be8-22d5-4b1d-9dc8-5d661f978198

C:\Users\PC\Desktop\scala>
```

```
C:\Windows\System32\cmd.e x + v
Microsoft Windows [Version 10.0.26100.3194]
(c) Microsoft Corporation. All rights reserved.

C:\Users\PC\Desktop\scala>sbt run
[info] welcome to sbt 1.10.11 (Oracle Corporation Java 1.8.0_411)
[info] loading project definition from C:\Users\PC\Desktop\scala\project
[info] loading settings for project scala from build.sbt...
[info] set current project to ScalaCSVProject (in build file:/C:/Users/PC/Desktop/scala/)
[info] compiling 1 Scala source to C:\Users\PC\Desktop\scala\target\scala-2.13\classes ...
[info] running ReadCSV
Using Spark's default log4j profile: org/apache/spark/log4j2-defaults.properties
25/04/17 09:46:29 INFO SparkContext: Running Spark version 3.5.0
25/04/17 09:46:29 INFO SparkContext: OS info Windows 11, 10.0, amd64
25/04/17 09:46:29 INFO SparkContext: Java version 1.8.0_411
25/04/17 09:46:30 WARN Shell: Did not find winutils.exe: java.io.FileNotFoundException: Hadoop bin direct
ory does not exist: C:\hadoop\bin\bin -see https://wiki.apache.org/hadoop/WindowsProblems
25/04/17 09:46:30 INFO ResourceUtils: =====
25/04/17 09:46:30 INFO ResourceUtils: No custom resources configured for spark.driver.
25/04/17 09:46:30 INFO ResourceUtils: =====
25/04/17 09:46:30 INFO SparkContext: Submitted application: CSV Reader
25/04/17 09:46:30 INFO ResourceProfile: Default ResourceProfile created, executor resources: Map(cores ->
name: cores, amount: 1, script: , vendor: , memory -> name: memory, amount: 1024, script: , vendor: , of
fHeap -> name: offHeap, amount: 0, script: , vendor: ), task resources: Map(cpus -> name: cpus, amount: 1
.0)
25/04/17 09:46:30 INFO ResourceProfile: Limiting resource is cpu
```