

Installation

Mr DIATTARA Ibrahima

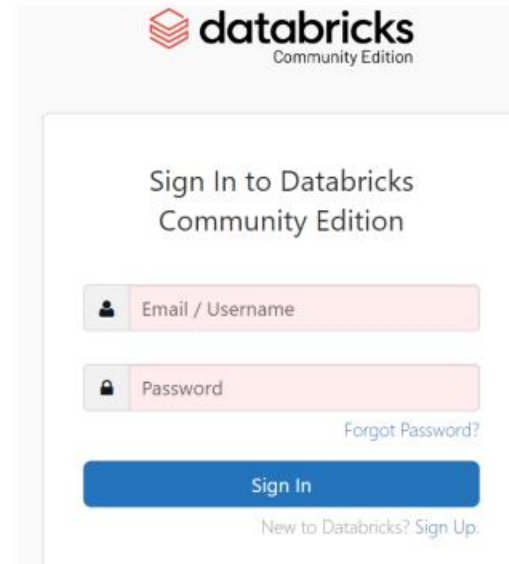
Ouverture Compte Databricks

Databricks vous offre la possibilité de créer:

- ❖ Machine
- ❖ Notebook (Spark Scala , pySpark, R, ..)
- ❖ Un système de fichier DBFS
- ❖

Lien pour ouvrir un compte : <https://community.cloud.databricks.com/login.html>

<https://community.cloud.databricks.com/login.html>



The screenshot shows the login interface for Databricks Community Edition. At the top, the Databricks logo and 'Community Edition' text are displayed. Below this, the heading 'Sign In to Databricks Community Edition' is centered. There are two input fields: 'Email / Username' with a person icon and 'Password' with a lock icon. A 'Forgot Password?' link is positioned to the right of the password field. A blue 'Sign In' button is located below the input fields. At the bottom, a link for 'New to Databricks? Sign Up.' is visible.

Create Compute Databricks

Compute

All-purpose compute

Job compute

Q Filter compute you have acce...

Created by ▾

Create compute

State ⌵

Name

Runtime

Active me...

Active cor...

Active DB...

Source

Creator

Notebooks



No compute

Create compute to run workloads from your notebooks and jobs. [Learn more about best practices for compute configuration](#)

Create compute

Comment installer IntelliJ ?

IntelliJ IDEA est un environnement de développement intégré(Integrated Development Environment - IDE)

Il y a deux versions :

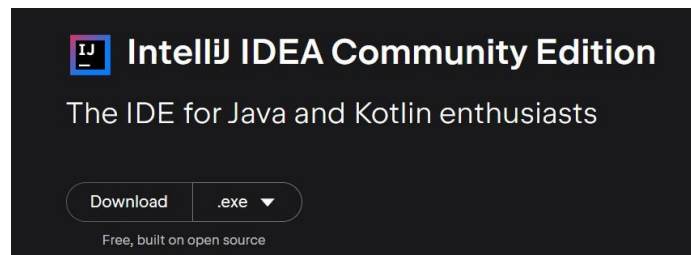
- ❑ IntelliJ IDEA Ultimate : Free 30-day trial
- ❑ IntelliJ IDEA Community Edition

Pour obtenir une version gratuite, vous devez choisir celle qui est Community que pour vous trouverez sur

Pour windows <https://www.jetbrains.com/idea/download/?section=windows>

Pour Mac <https://www.jetbrains.com/idea/download/?section=mac>

Pour Linux <https://www.jetbrains.com/idea/download/?section=linux>



Scala | Java

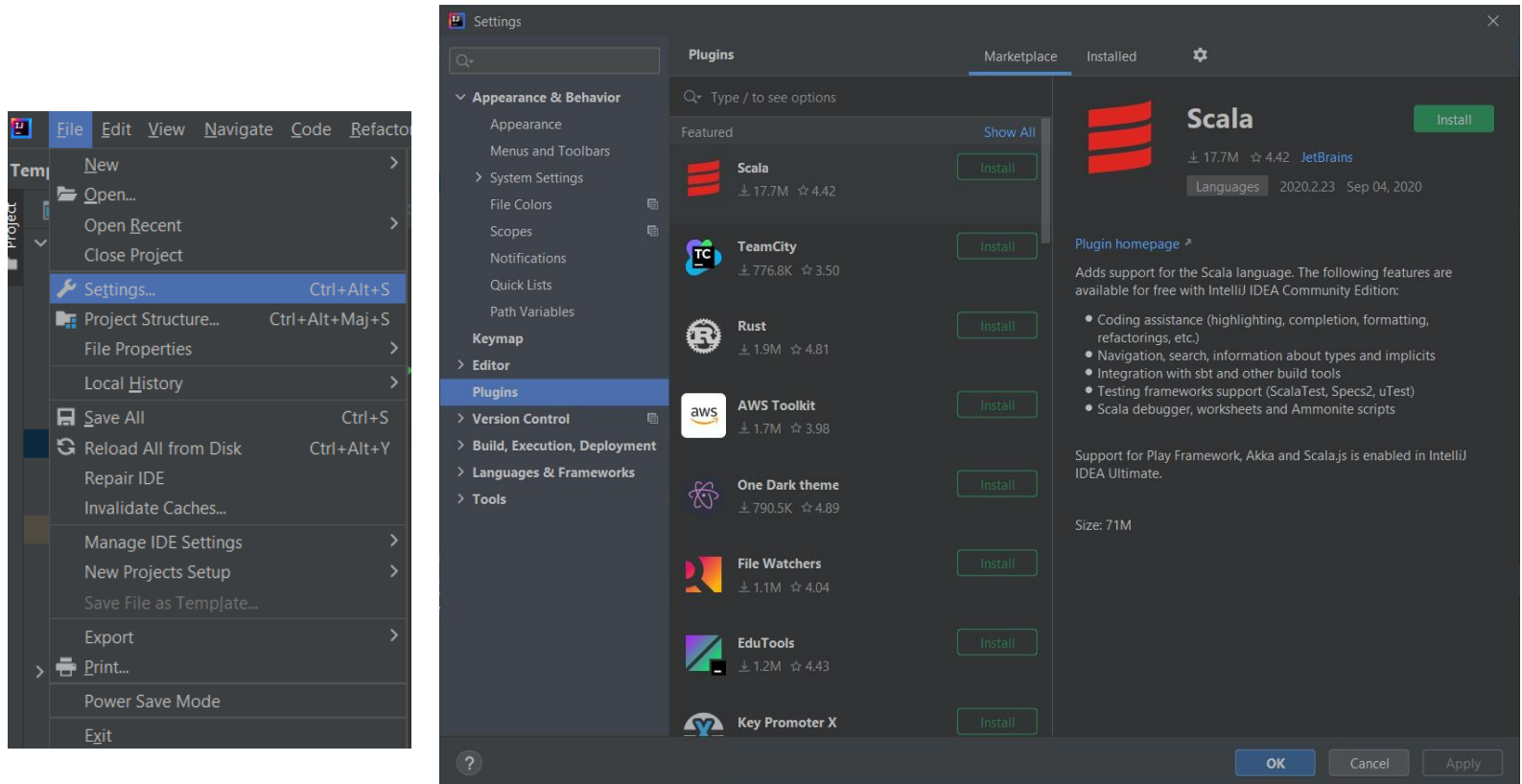
Dans ce cours on va choisir

- ❖ Scala 2.12.15

- ❖ Java 1.8 (version corretto-1.8.0 disponible dans notre IDE)

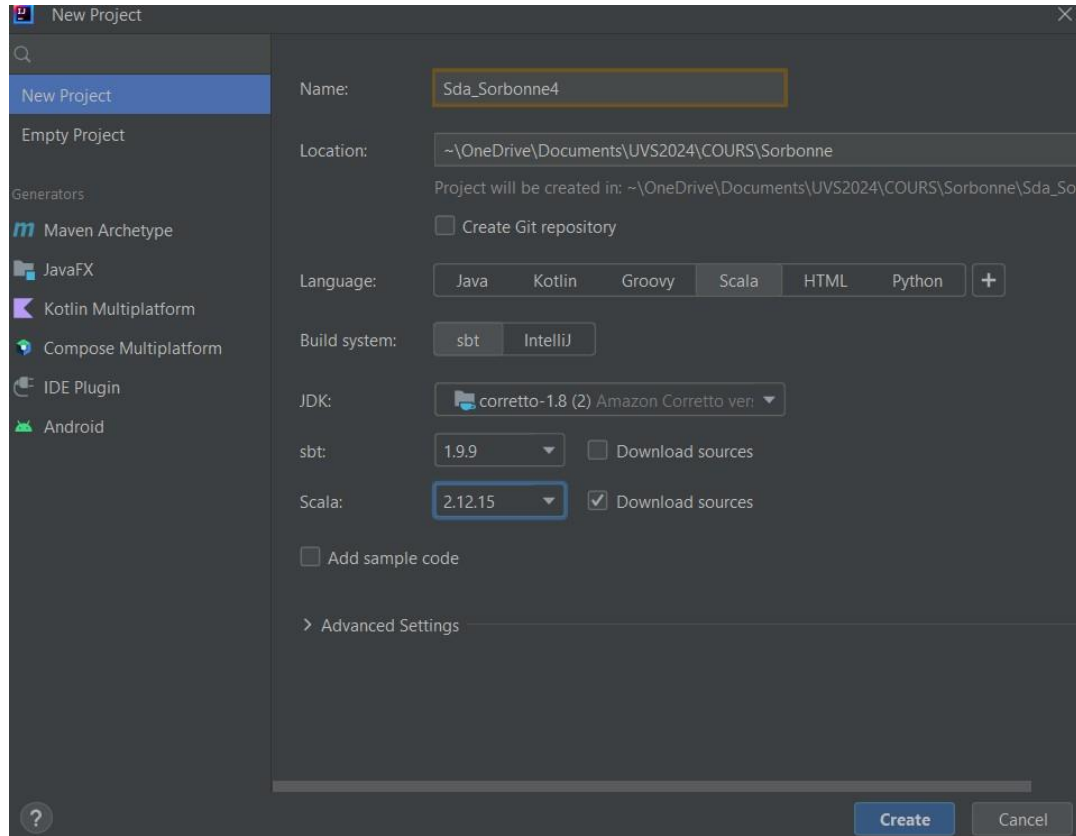
Scala Plugins

Pour pouvoir utiliser le langage Scala , il faut rajouter le plugin dans l'IDE



Créer un projet Sbt Scala

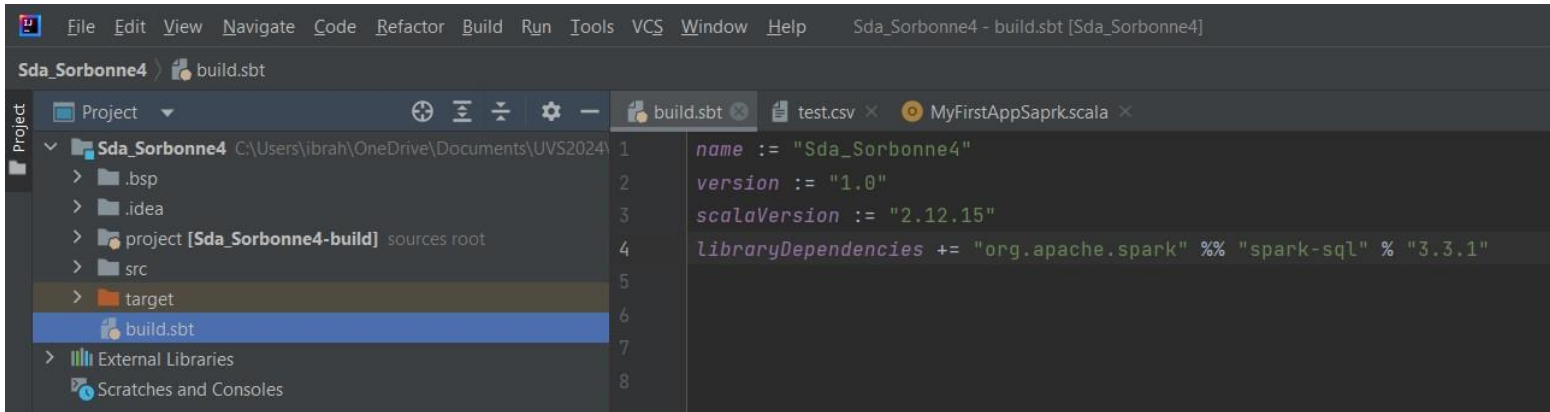
Aller dans IntelliJ -> Fille -> New project



Patiencez le temps que l'IDE résolve les dépendances

Editer le fichier **build.sbt**

SBT est un outil de construction open source pour les projets Scalae, similaire à Maven



```
1  name := "Sda_Sorbonne4"
2  version := "1.0"
3  scalaVersion := "2.12.15"
4  libraryDependencies += "org.apache.spark" %% "spark-sql" % "3.3.1"
5
6
7
8
```

Warning: Faire un refresh de votre sbt afin de ramener la dépendance Spark

Lien du code

https://github.com/idiattara/sda_2024/blob/main/sbt.build

Créer Votre premier Appli Spark

Lien du code

https://github.com/idiattara/sda_2024/blob/main/MyFirstApp.scala

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

object MyFirstApp{
  def main(args: Array[String]) {
    val csvFile="src/main/resources/datain/test.csv"

    val spark = SparkSession.builder.appName( name = "SimpleApplication")
      .config("spark.master", "local")
      .getOrCreate()

    val df = spark.read
      .format( source = "csv")
      .option("header", "true") // Si le fichier CSV contient un en-tête
      .load(csvFile)

    df.show()
  }
}
```