Big Data Lab

Spark Installation

1. **IntelliJ Installation**

**1.1 From Source :-**

* Download IntelliJ **Community Version** from [here.](https://www.jetbrains.com/idea/download/other.html)

* Unzip the file
* Using the terminal go to the downloaded folder and cd into the bin folder.
* Run the following command to start the IDE

./idea.sh

**1.2. Installation using a shell script :-**

Running this shell script automates the complete installation process.

* Open the terminal and create a new file named install-intellij.sh
* Copy paste the below code into the new file and save it.

*#!/bin/sh*

cd ~/Downloads

wget "https://download.jetbrains.com/idea/ideaIC-2021.3.tar.gz?\_gl=1\*14ggh2a\*\_ga\*MTk0NjM1MjAzNS4xNjQwNTk3NTQz\*\_ga\_V0XZL7QHEB\*MTY0MDU5NzU0My4xLjEuMTY0MDU5NzgxOS4w&\_ga=2.226875417.589077562.1640597544-1946352035.1640597543"

mv "ideaIC-2021.3.tar.gz?\_gl=1\*14ggh2a\*\_ga\*MTk0NjM1MjAzNS4xNjQwNTk3NTQz\*\_ga\_V0XZL7QHEB\*MTY0MDU5NzU0My4xLjEuMTY0MDU5NzgxOS4w&\_ga=2.226875417.589077562.1640597544-1946352035.1640597543" "ideaIC-2021.3.tar.gz"

sudo tar -xvzf ideaIC-2021.3.tar.gz

cd ~/Downloads/

cd idea-IC-213.5744.223//bin

./idea.sh

After saving the file run the following command.

bash install-intellij.sh

1. **Setup IntelliJ for Spark**

* Navigate to Plugins on the taskbar and install Scala.

Graphical user interface, text, application, chat or text message

Description automatically generated

* Restart the IDE once the plugin is installed

1. **Creating Scala Project**

* Click on create a project, select scala and use sbt option(default) to create an sbt-based scala project.

Graphical user interface, application

Description automatically generated

* Project settings :-
* Select JDK : 1.8
* SBT version : 1.5.8, check download sources
* Scala version : 2.11.8, check download sources

Graphical user interface, text, application

Description automatically generated

**Note:** if the versions aren’t set properly the programs won’t run due to version mismatch errors.

* Open build.sbt file and add the following dependencies

val sparkVersion = "2.4.0"

libraryDependencies ++= Seq(

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

"org.apache.spark" %% "spark-sql" % sparkVersion,

)

* Graphical user interface, application, icon

  Description automatically generated When this icon appears in your project please click on it to sync your changes made in the build.sbt file.
* Once the build is successful, head over to src/main/scala. Click on *new Scala class*. Give a file name and create a Scala Object .Write the business logic here and run the file.

Graphical user interface, text

Description automatically generated

**Note:** If you are unable to create a new scala class then the IDE is downloading some important files from the internet, please look at the progress bar found near the bottom right corner of the screen. Once the progress bar finishes all processes and vanishes you will be able to create a new scala class/object.