**PYSPARK setup in Intellij**

**Perquisites for setup:**

1. Intellij IDE
2. Python plugin in Intellij
3. Java
4. Apache spark module

**Apache Spark module:**

Apache Spark version : 2.4.6

<https://spark.apache.org/downloads.html>

<https://www.apache.org/dyn/closer.lua/spark/spark-3.0.0/spark-3.0.0-bin-hadoop2.7.tgz>

**Steps to pyspark in Intellij:**

1. Check python plugin is installed in Intellij IDE in below path

Configure -> Plugins

A screenshot of a cell phone

Description automatically generated

A screenshot of a social media post

Description automatically generated

1. Create new project in Intellij

A screenshot of a cell phone

Description automatically generated

1. Select python in list and Project SDK should be python. Then click next.

A screenshot of a cell phone

Description automatically generated

Give the project name and project location. Then click Finish.

A screenshot of a social media post

Description automatically generated

1. Go to File -> Project Structure - > Modules

A screenshot of a social media post

Description automatically generated

1. Click Add Content Root in Module page to add Pyspark modules and Py4j.

A screenshot of a social media post

Description automatically generated

1. Add python folder from spark module(which we have downloaded)

(D:\Users\BoddulaR\Downloads\spark-2.4.6-bin-hadoop2.7\python)

A screenshot of a cell phone

Description automatically generated

1. Add Py4j zip file from spark module(D:\Users\BoddulaR\Downloads\spark-2.4.6-bin-hadoop2.7\python\lib\py4j-0.10.7-src.zip)

A screenshot of a social media post

Description automatically generated

1. After adding python folder and Py4j folder click “Apply” and “OK”
2. PySpark completed. Create new python file to write code.
3. Add winutils.exe file in “D:\Users\BoddulaR\Downloads\spark-2.4.6-bin-hadoop2.7\bin folder”

(Download from <https://github.com/steveloughran/winutils/blob/master/hadoop-2.7.1/bin/winutils.exe> )

1. Add HADOOP\_HOME = D:\Users\BoddulaR\Downloads\spark-2.4.6-bin-hadoop2.7 in environment variables.
2. Restart your Intellij IDE