

1. En una máquina virtual realice la configuración de apache spark, puede guiarse en cualquier tutorial o el proporcionado por el docente.

url: <https://computingforgeeks.com/how-to-install-apache-spark-on-ubuntu-debian/>

```
alvaro@alvaro-VirtualBox: /opt/spark/sbin
alvaro@alvaro-VirtualBox:~$ cd /opt/spark/sbin
alvaro@alvaro-VirtualBox:/opt/spark/sbin$ ./start-master.sh
starting org.apache.spark.deploy.master.Master, logging to /opt/spark/logs/spark-alvaro-org.apache.spark.deploy.mast
er.Master-1-alvaro-VirtualBox.out
failed to launch: nice -n 0 /opt/spark/bin/spark-class org.apache.spark.deploy.master.Master --host alvaro-VirtualBo
x --port 7077 --webui-port 8080
full log in /opt/spark/logs/spark-alvaro-org.apache.spark.deploy.master.Master-1-alvaro-VirtualBox.out
```

192.168.100.37:8080
 Spark Master at spark://alvaro-VirtualBox:7077
 URL: spark://alvaro-VirtualBox:7077
 Alive Workers: 1
 Cores in use: 1 Total, 0 Used
 Memory in use: 1024.0 MiB Total, 0.0 B Used
 Resources in use:
 Applications: 0 Running, 0 Completed
 Drivers: 0 Running, 0 Completed
 Status: ALIVE

```
alvaro@alvaro-VirtualBox:/opt/spark/sbin$ ./start-slave.sh spark://alvaro-VirtualBox:7077
This script is deprecated, use start-worker.sh
starting org.apache.spark.deploy.worker.Worker, logging to /opt/spark/logs/spark-alvaro-org.apache.spark.deploy.work
er.Worker-1-alvaro-VirtualBox.out
alvaro@alvaro-VirtualBox:/opt/spark/sbin$
```

192.168.100.37:8080
 Spark Master at spark://alvaro-VirtualBox:7077
 URL: spark://alvaro-VirtualBox:7077
 Alive Workers: 1
 Cores in use: 1 Total, 0 Used
 Memory in use: 1024.0 MiB Total, 0.0 B Used
 Resources in use:
 Applications: 0 Running, 0 Completed
 Drivers: 0 Running, 0 Completed
 Status: ALIVE

Workers (1)

Worker Id	Address	State	Cores	Memory	Resources
worker-20221212005403-192.168.100.37-39785	192.168.100.37:39785	ALIVE	1 (0 Used)	1024.0 MiB (0.0 B Used)	

Con el shell podra ejecutar scala por defecto

```
alvaro@alvaro-VirtualBox: /opt/spark/bin
alvaro@alvaro-VirtualBox: /opt/spark/sbin
alvaro@alvaro-VirtualBox: /opt/spark/bin$ ./spark-shell
22/12/12 01:03:54 WARN Utils: Your hostname, alvaro-VirtualBox resolves to a loopback address: 127.0.1.1; using 192.168.100.37 instead (on interface enp0s3)
22/12/12 01:03:54 WARN Utils: Set SPARK_LOCAL_IP if you need to bind to another address
Setting default log level to "WARN".
To adjust logging level use sc.setLogLevel(newLevel). For SparkR, use setLogLevel(newLevel).
22/12/12 01:04:10 WARN NativeCodeLoader: Unable to load native-hadoop library for your platform... using builtin-jav
a classes where applicable
Spark context Web UI available at http://192.168.100.37:4040
Spark context available as 'sc' (master = local[*], app id = local-1670821454161).
Spark session available as 'spark'.
Welcome to

  ____      _
 / ___|  _ \| | | |
 \___ \| |_) | |_| |
  ___) | |_) | | | |
 |____|_|_|\___|_|_|_|

 version 3.3.1

Using Scala version 2.12.15 (OpenJDK 64-Bit Server VM, Java 11.0.17)
Type in expressions to have them evaluated.
Type :help for more information.

scala>
```

Instale Python para spark

```
alvaro@alvaro-VirtualBox: /opt/spark/bin
alvaro@alvaro-VirtualBox: /opt/spark/bin$ ./pyspark
Python 3.8.10 (default, Nov 14 2022, 12:59:47)
[GCC 9.4.0] on linux
Type "help", "copyright", "credits" or "license" for more information.
22/12/12 01:12:21 WARN Utils: Your hostname, alvaro-VirtualBox resolves to a loopback address: 127.0.1.1; using 192.168.100.37 instead (on interface enp0s3)
22/12/12 01:12:21 WARN Utils: Set SPARK_LOCAL_IP if you need to bind to another address
Setting default log level to "WARN".
To adjust logging level use sc.setLogLevel(newLevel). For SparkR, use setLogLevel(newLevel).
22/12/12 01:12:25 WARN NativeCodeLoader: Unable to load native-hadoop library for your platform... using builtin-jav
a classes where applicable
22/12/12 01:12:28 WARN Utils: Service 'SparkUI' could not bind on port 4040. Attempting port 4041.
Welcome to

  ____      _
 / ___|  _ \| | | |
 \___ \| |_) | |_| |
  ___) | |_) | | | |
 |____|_|_|\___|_|_|_|

 version 3.3.1

Using Python version 3.8.10 (default, Nov 14 2022 12:59:47)
Spark context Web UI available at http://192.168.100.37:4041
Spark context available as 'sc' (master = local[*], app id = local-1670821949347).
SparkSession available as 'spark'.
>>>
```