Universidad Técnica Particular de Loja

Nombre: Nahomi Astrid Cabrera Picoita

Fecha: 8 de junio de 2024

# Índice

[Índice 1](#_Toc169173585)

[DESCARGAR WORKBENCH 1](#_Toc169173586)

[DESCARGAR NET BEANS 2](#_Toc169173587)

[Conexión a NetBeans usando JPA 5](#_Toc169173588)

[INSTALACIÓN DE ZEPPELIN 6](#_Toc169173589)

[Instalar los modulos LBS 7](#_Toc169173590)

[Instalar paquetes esenciales y de utilidad: 8](#_Toc169173591)

[Instalar Apache Spark: 9](#_Toc169173592)

[Instalar Git y SBT: 11](#_Toc169173593)

[Instalar Apache Zeppelin: 14](#_Toc169173594)

[Configurar variables de entorno: 16](#_Toc169173595)

[Aplicamos cambios: 17](#_Toc169173596)

[Ingresamos a: 18](#_Toc169173597)

[Dudas: *¿Qué es Apache Spark?* 18](#_Toc169173598)

[¿Qué es Apache Zeppelin? 19](#_Toc169173599)

[Comandos utilizados en clases: 19](#_Toc169173600)

## DESCARGAR WORKBENCH

sudo apt update

sudo apt install mysql-server

sudo systemctl status mysql

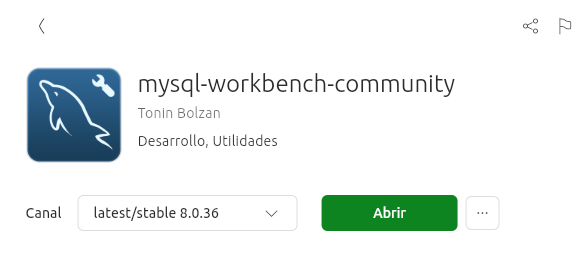
sudo mysql --version

sudo mysql

ALTER USER 'root'@'localhost' IDENTIFIED WITH mysql\_native\_password BY 'tuclave';

exit;

(Descarga desde la tienda de aplicaciones Workbench)



## DESCARGAR NET BEANS

sudo apt update

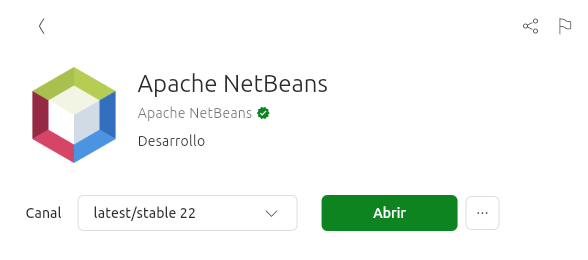
sudo apt install -y openjdk-17-jdk

java -version (Verifica la version instalada)

sudo add-apt-repository "deb http://archive.ubuntu.com/ubuntu focal universe"

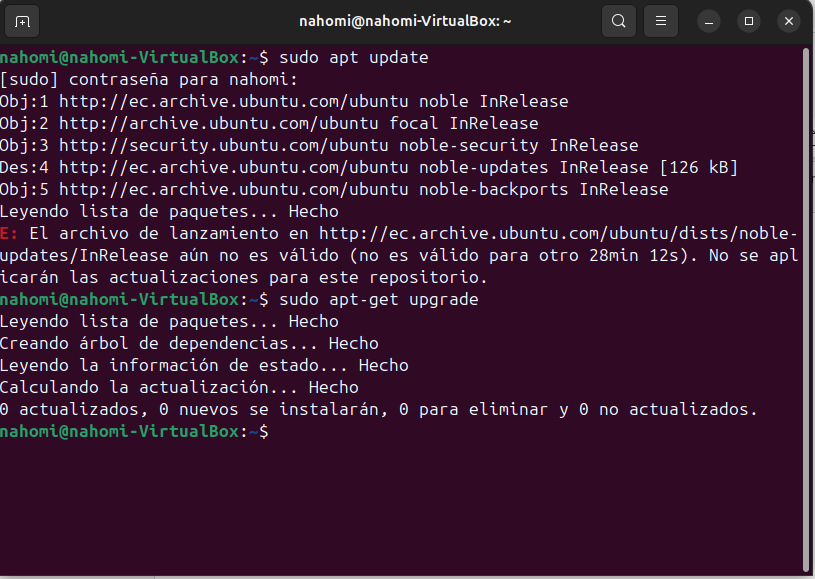
sudo apt update

(Descarga desde la tienda de aplicaciones Apache NetBeans)



No muestro capturas de esta instalación porque no tomé capturas en ese momento.

Pero algo así se vería:



Versiones instaladas de MySQL y OpenJDK

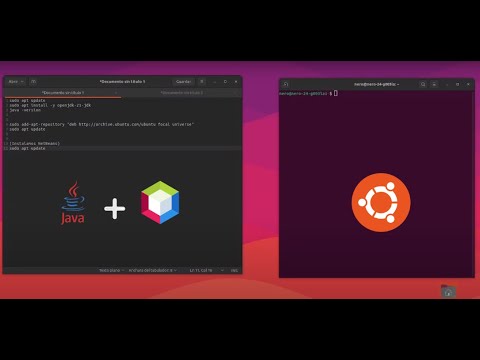


Envío links de los videos que seguí:

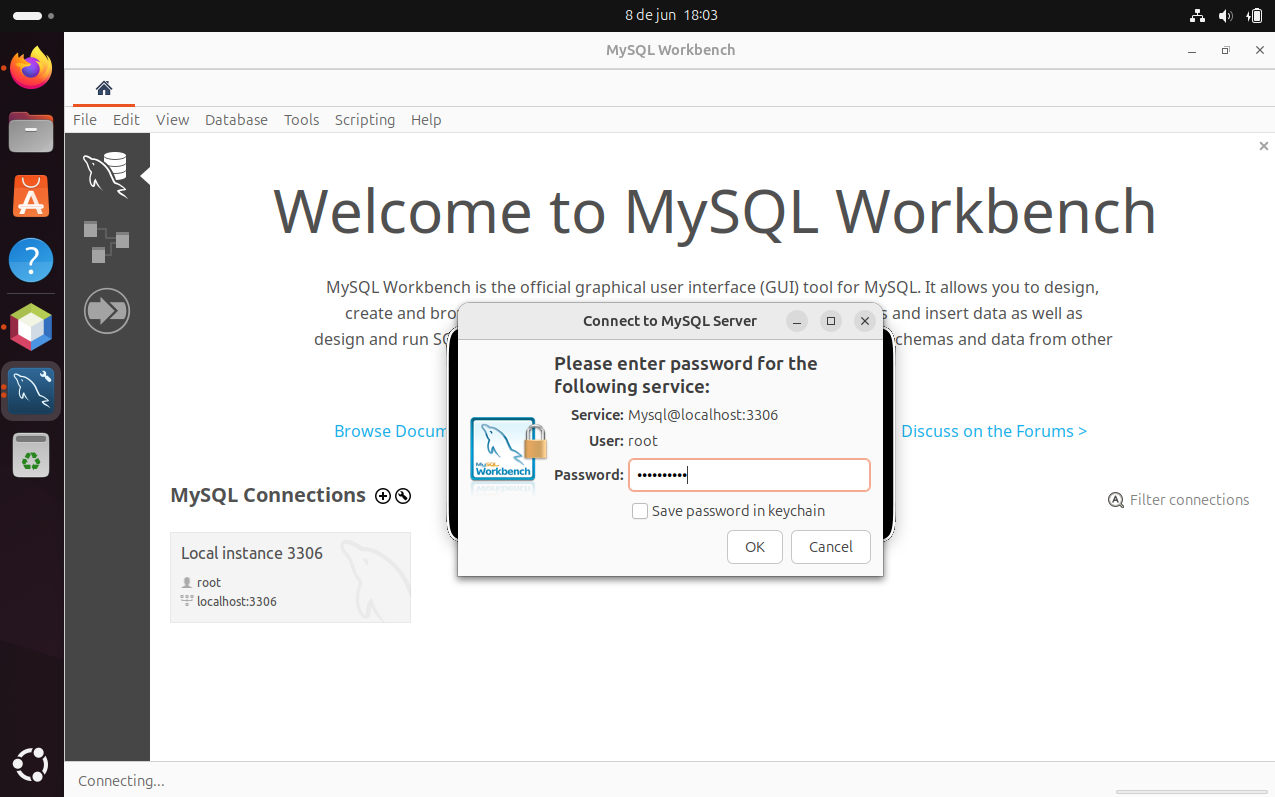
<https://youtu.be/3PKm03KXsHw?si=OzXqbg0eqIDa_WSp>

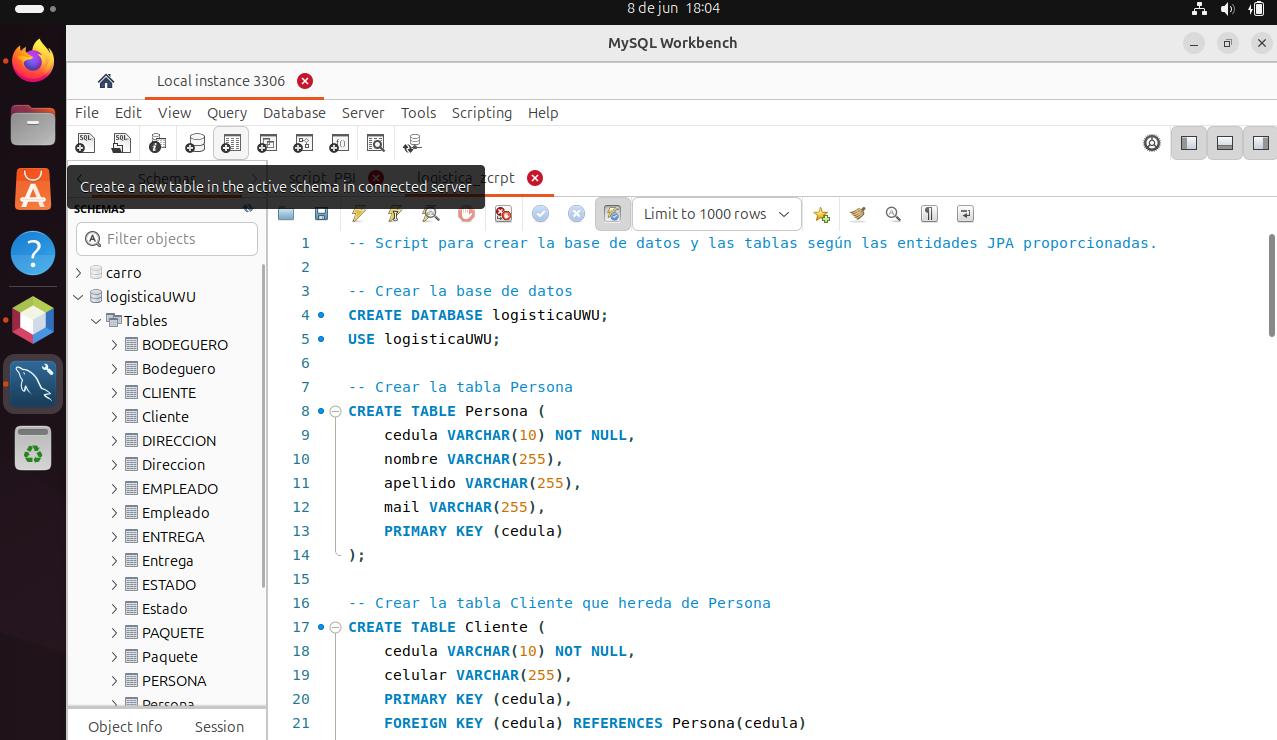
[](https://youtu.be/3PKm03KXsHw?si=OzXqbg0eqIDa_WSp)

<https://youtu.be/M-PSm1LYGhE?si=fUbh9tguyuuKZ3Ml>

[](https://youtu.be/M-PSm1LYGhE?si=fUbh9tguyuuKZ3Ml)

Capturas:





[Regresar a índice](#_Índice)

## Conexión a NetBeans usando JPA

Las clases JPA Controllers son las mismas, y las entidades también. Solo cambia que la librería utilizada dejó de ser jakarta y se usa javax.

El main es bastante sencillo, solo creo una persona y la busco, y ya.

***Main***

public class main {

public static void main(String[] args) {

EntityManagerFactory emf = Persistence.createEntityManagerFactory("configurar\_Herramientas\_PU");

EntityManager em = emf.createEntityManager();

Persona p = new Persona("1150057501", "Nahomi", "Cabrera", "nacabrera3@utpl.edu.ec");

PersonaJpaController pc = new PersonaJpaController();

em.getTransaction().begin();

em.persist(p);

em.getTransaction().commit();

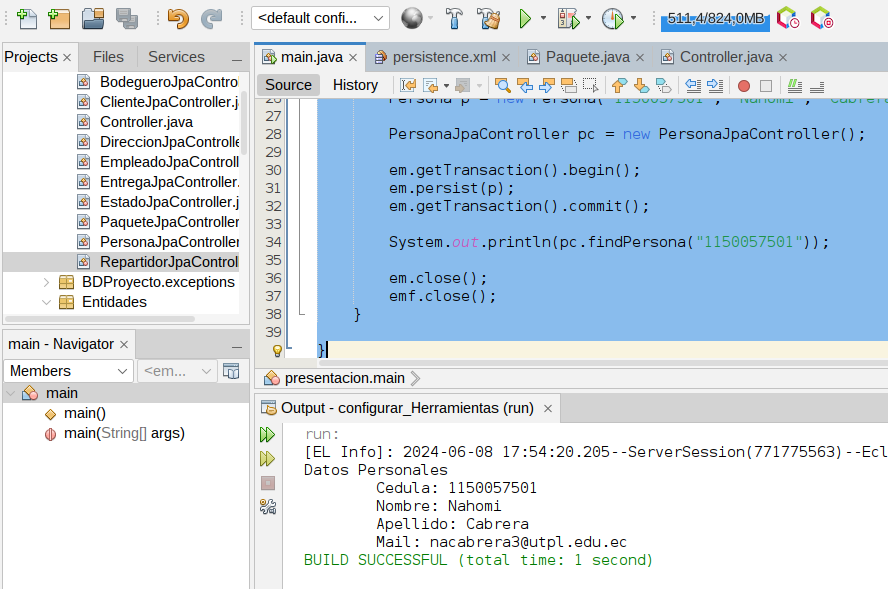
System.out.println(pc.findPersona("1150057501"));

em.close();

emf.close();

}

}

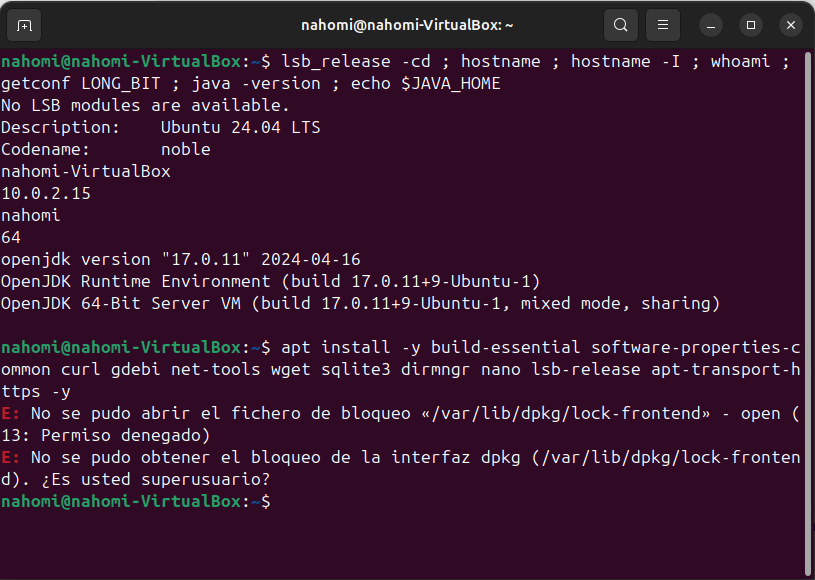


[Regresar a índice](#_Índice)

## INSTALACIÓN DE ZEPPELIN

lsb\_release -cd ; hostname ; hostname I ; whoami ; getconf LONG\_BIT ; java -version ; echo $JAVA\_HOME

apt install -y build-essential software-properties-common curl gdebi net-tools wget sqlite3 dirmngr nano lsb-release apt-transport-https -y

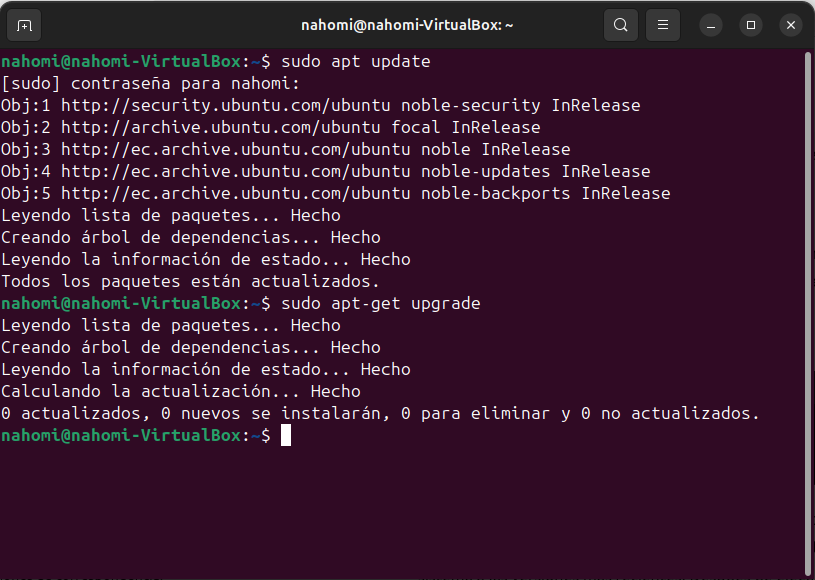


Instalo los módulos de LSB con los siguientes comandos:

Actualizo los paquetes disponibles e instalados:

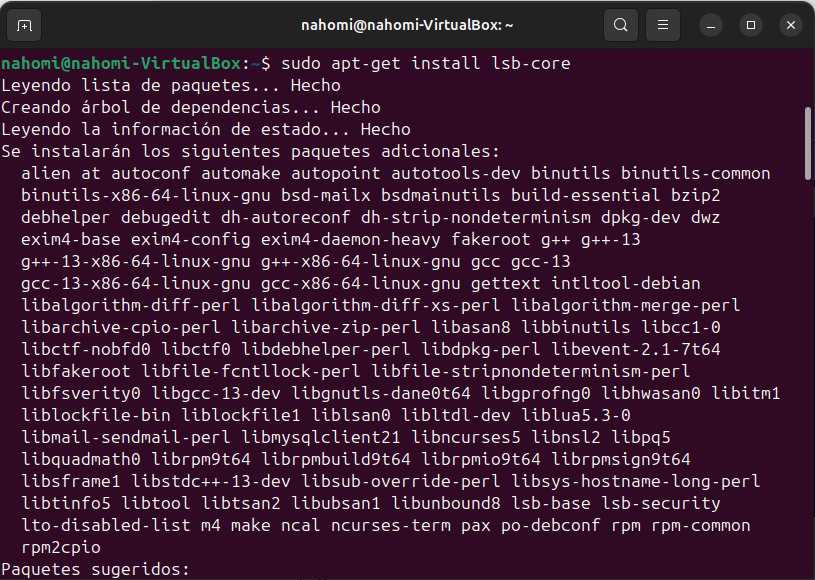
sudo apt update

sudo apt upgrade



Instalar los modulos LBS

sudo apt-get install lsb-core

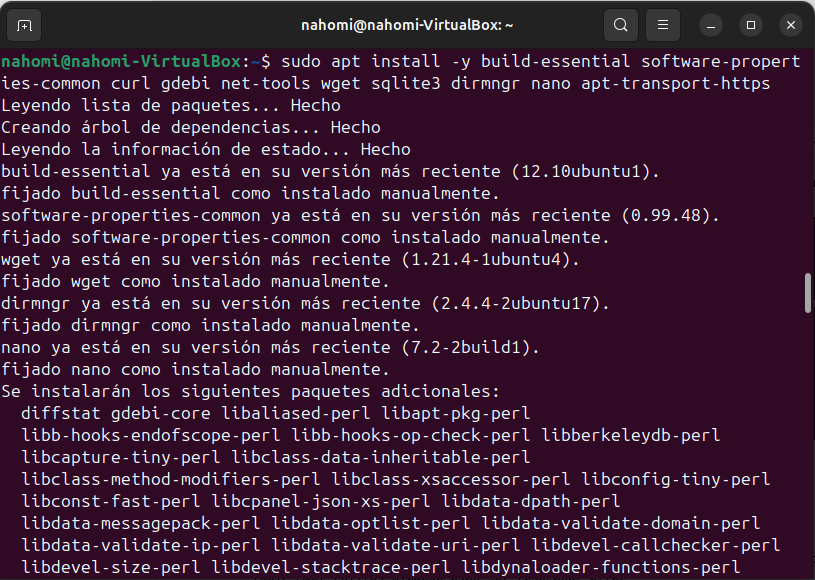


Le sale un mensaje que les advierte que usará determinado espacio y que si está seguro de continuar. Escribe S o Y. Y espera a que se instale todo.

[Regresar a índice](#_Índice)

### Instalar paquetes esenciales y de utilidad:

sudo apt install -y build-essential software-properties-common curl gdebi net-tools wget sqlite3 dirmngr nano lsb-release apt-transport-https

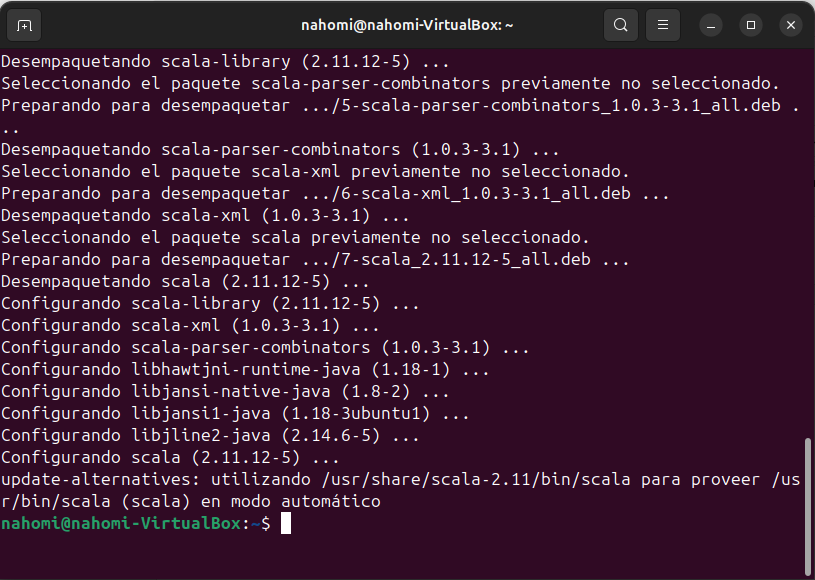
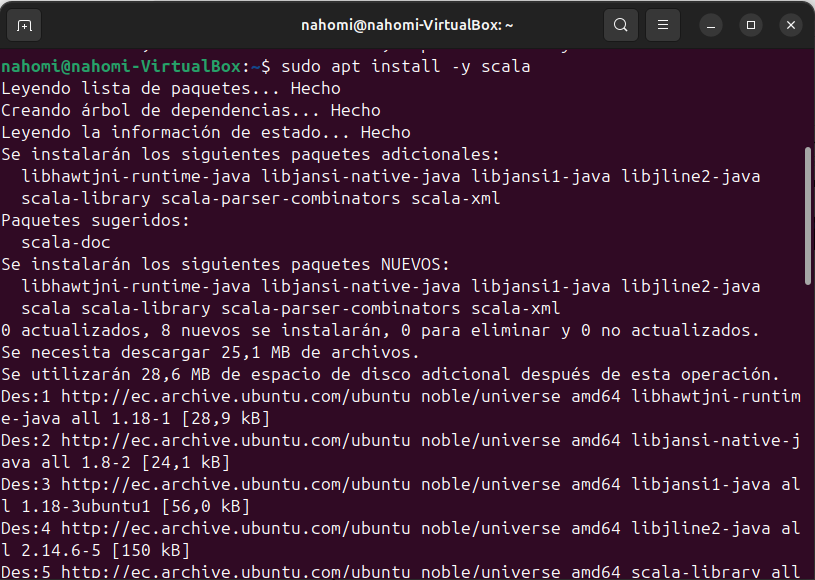


Espera a que se instale y tiene fe en el computador.

Listo, seguimos.

Ahora, instalamos Scala:

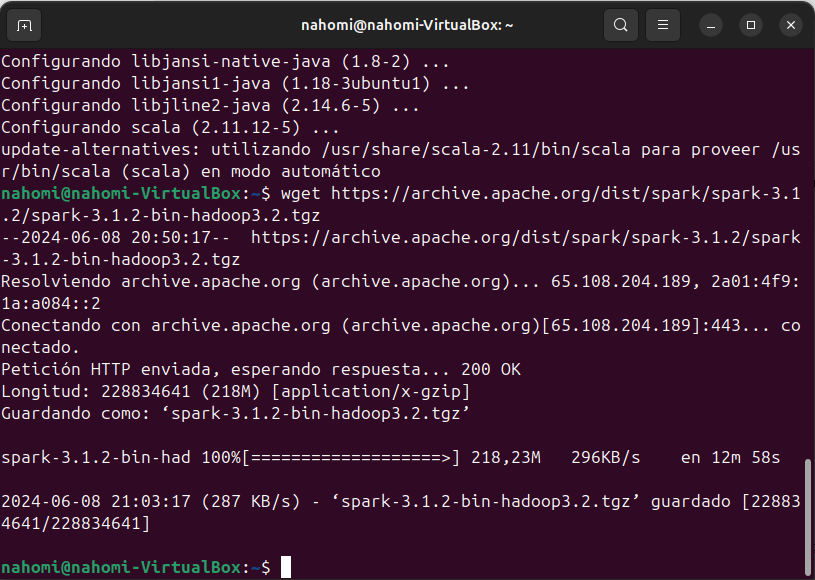
sudo apt install -y scala



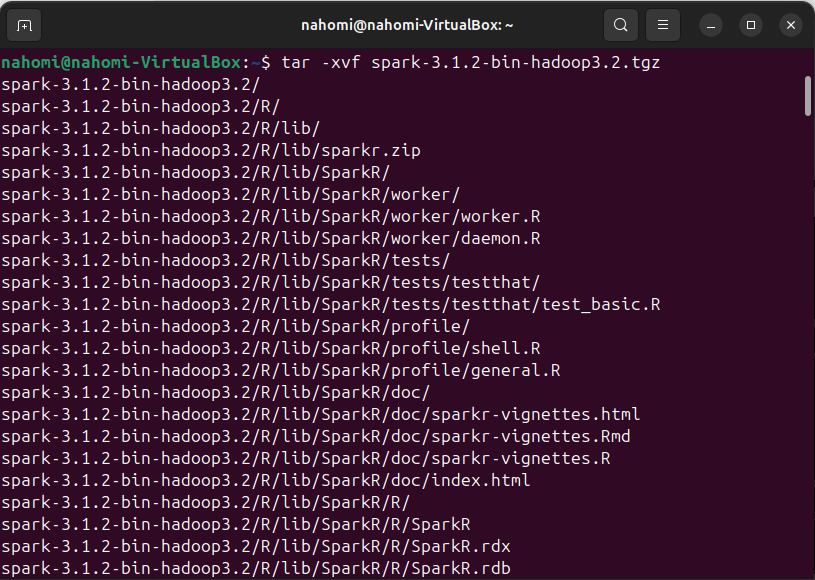
[Regresar a índice](#_Índice)

### Instalar Apache Spark:

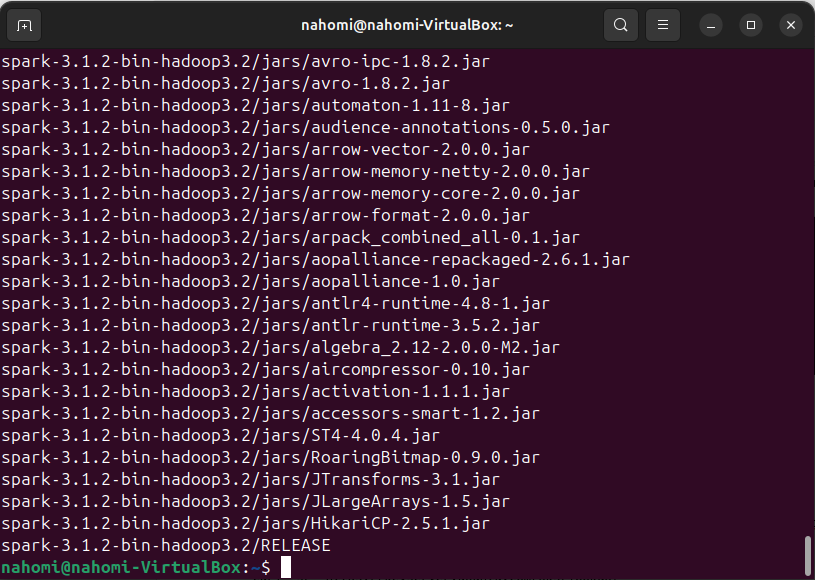
wget https://archive.apache.org/dist/spark/spark-3.1.2/spark-3.1.2-bin-hadoop3.2.tgz



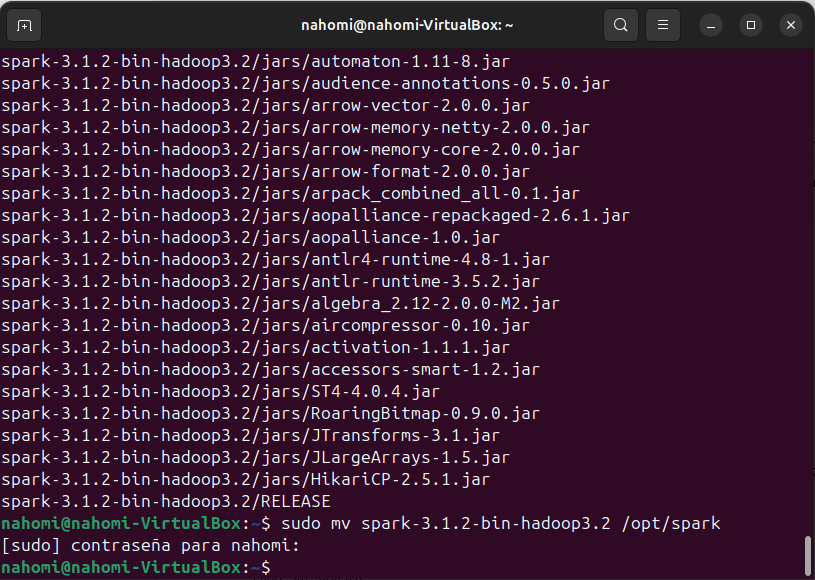
tar -xvf spark-3.1.2-bin-hadoop3.2.tgz



Te salen muchas líneas



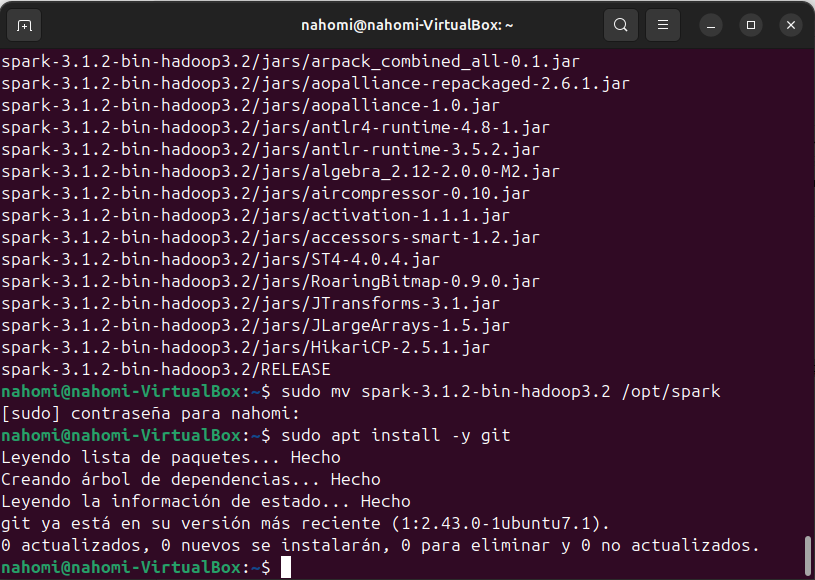
sudo mv spark-3.1.2-bin-hadoop3.2 /opt/spark



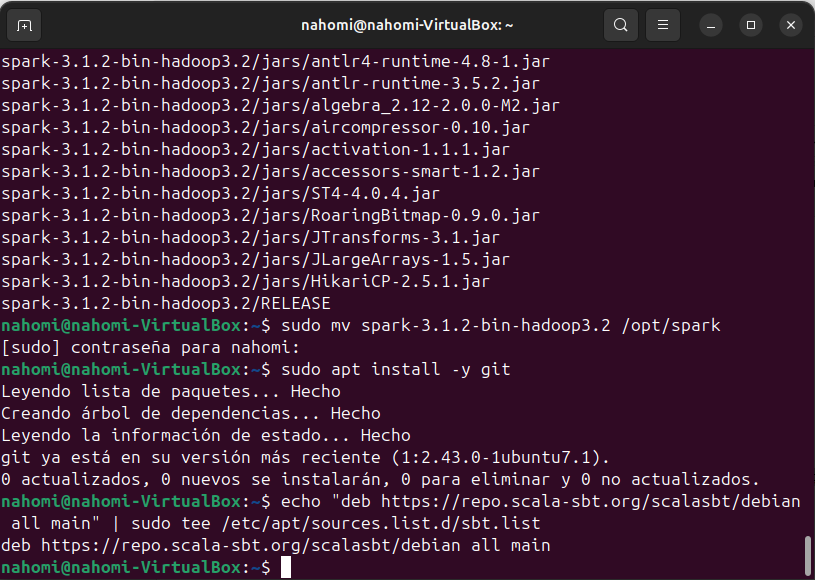
[Regresar a índice](#_Índice)

### Instalar Git y SBT:

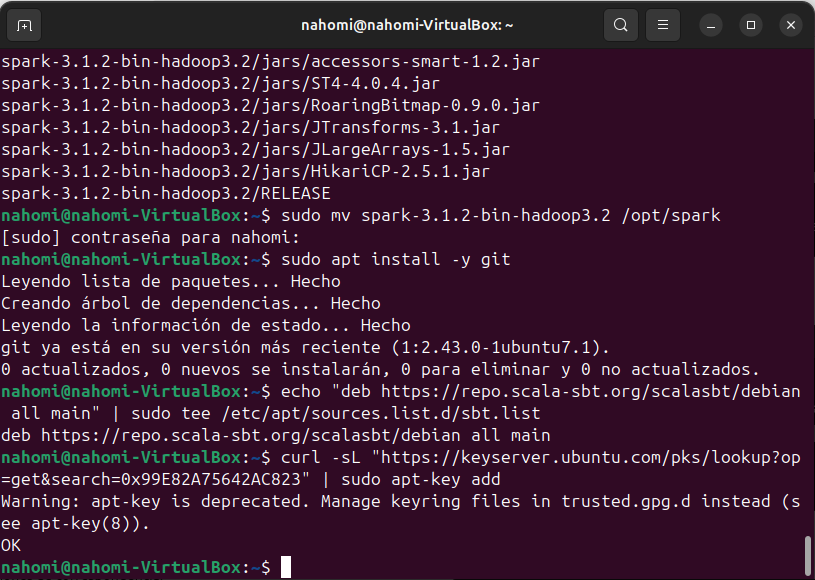
sudo apt install -y git



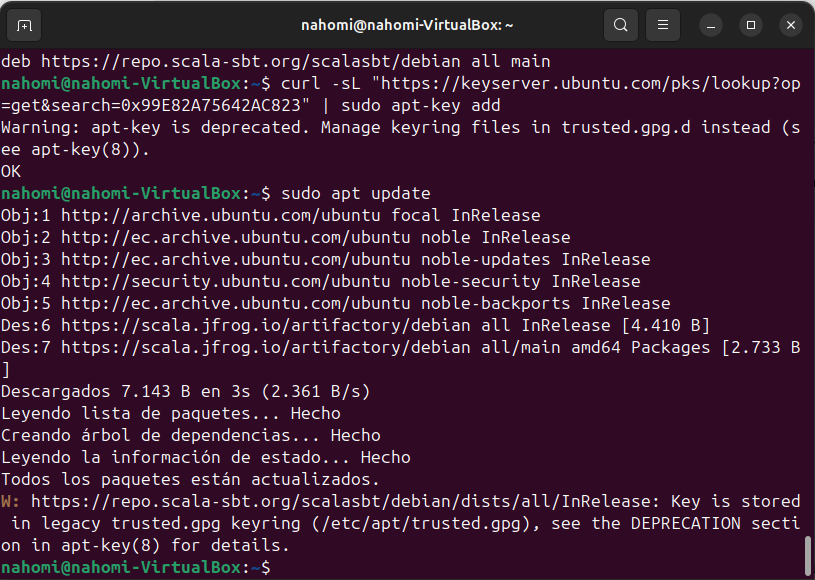
echo "deb https://repo.scala-sbt.org/scalasbt/debian all main" | sudo tee /etc/apt/sources.list.d/sbt.list



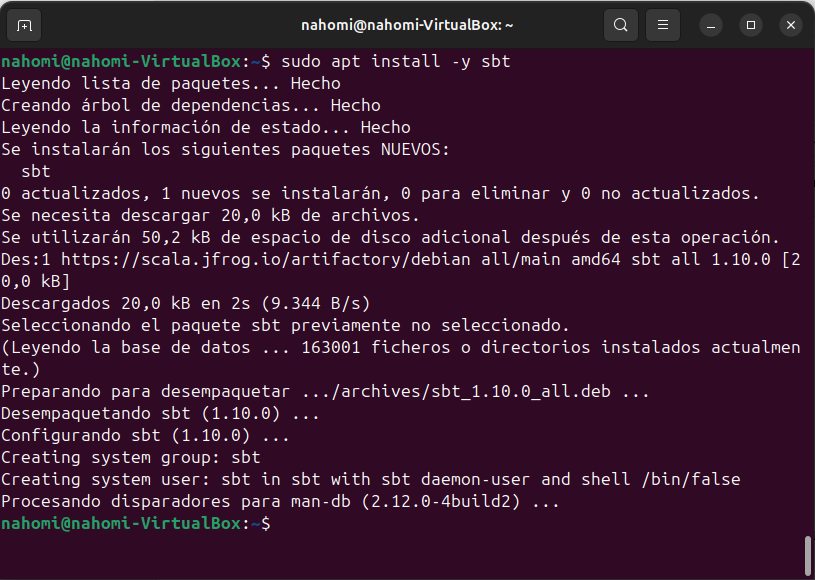
curl -sL "https://keyserver.ubuntu.com/pks/lookup?op=get&search=0x99E82A75642AC823" | sudo apt-key add



sudo apt update



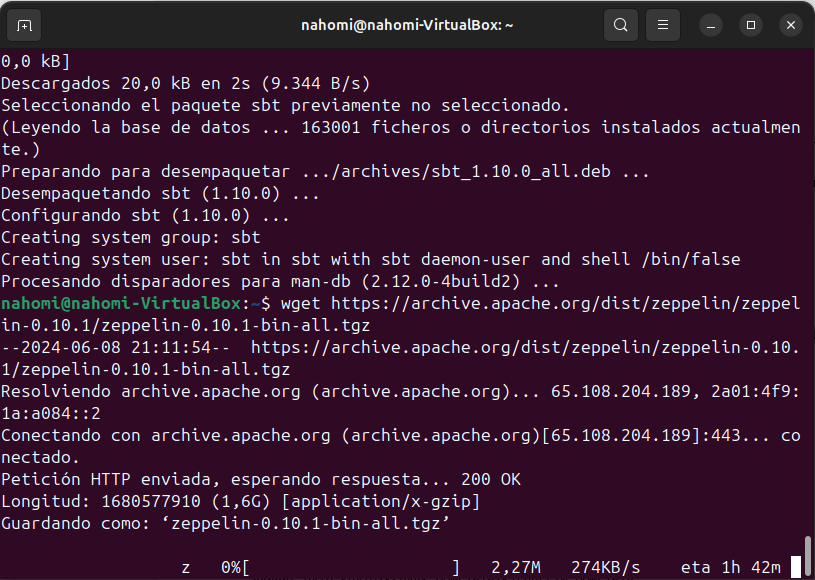
sudo apt install -y sbt

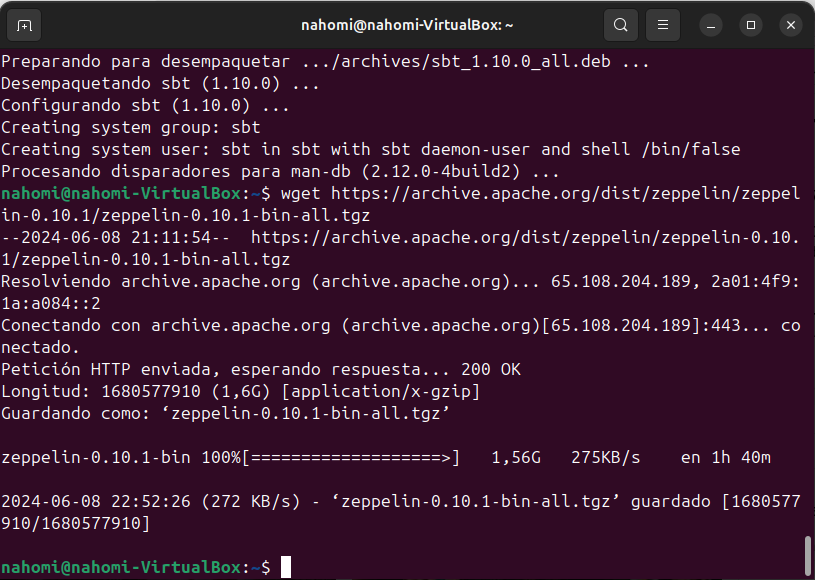


[Regresar a índice](#_Índice)

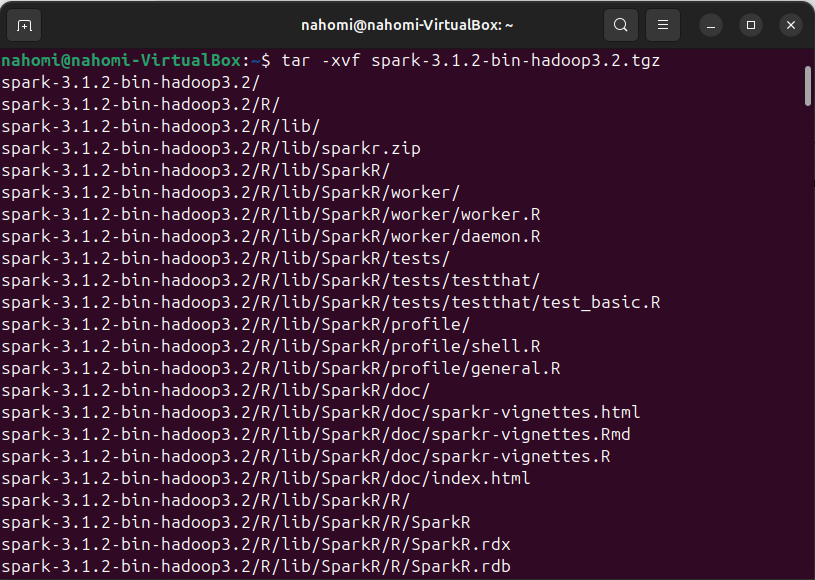
### Instalar Apache Zeppelin:

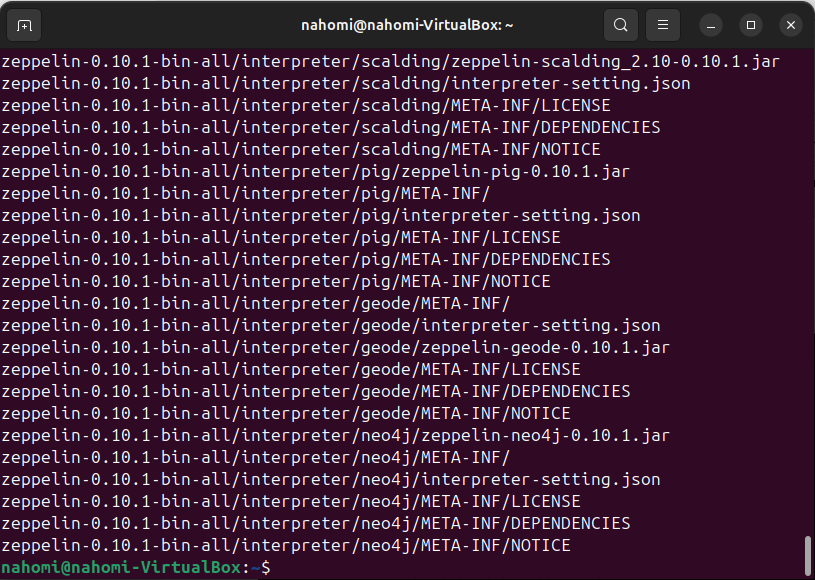
wget <https://archive.apache.org/dist/zeppelin/zeppelin-0.10.1/zeppelin-0.10.1-bin-all.tgz>



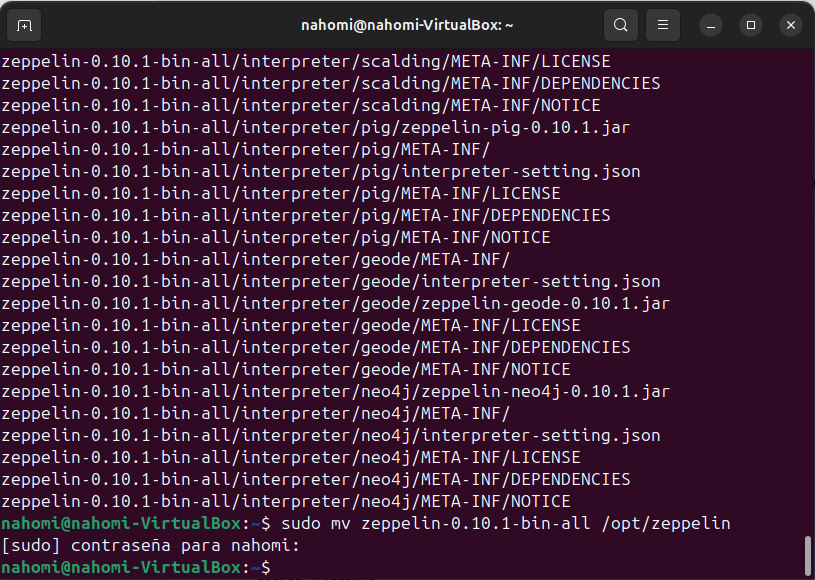


tar -xvf zeppelin-0.10.1-bin-all.tgz





sudo mv zeppelin-0.10.1-bin-all /opt/zeppelin



[Regresar a índice](#_Índice)

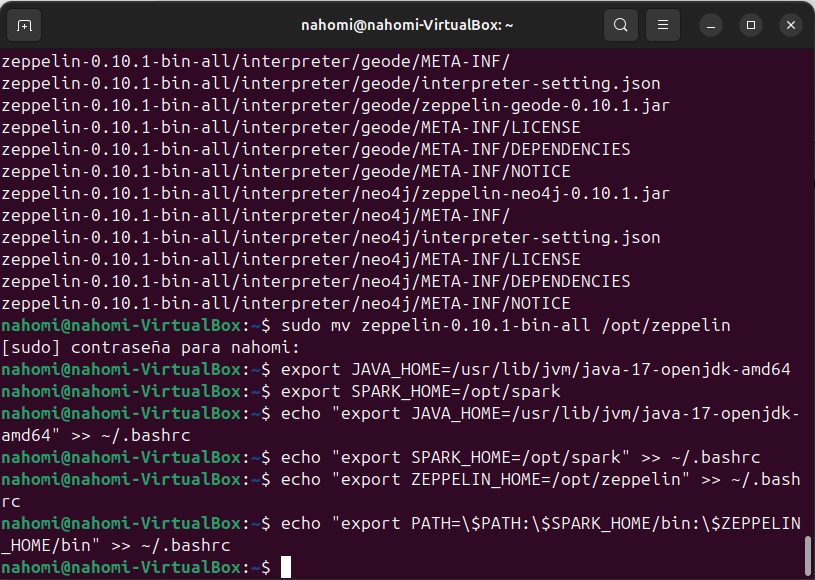
### Configurar variables de entorno:

echo "export JAVA\_HOME=/usr/lib/jvm/java-17-openjdk-amd64" >> ~/.bashrc

echo "export SPARK\_HOME=/opt/spark" >> ~/.bashrc

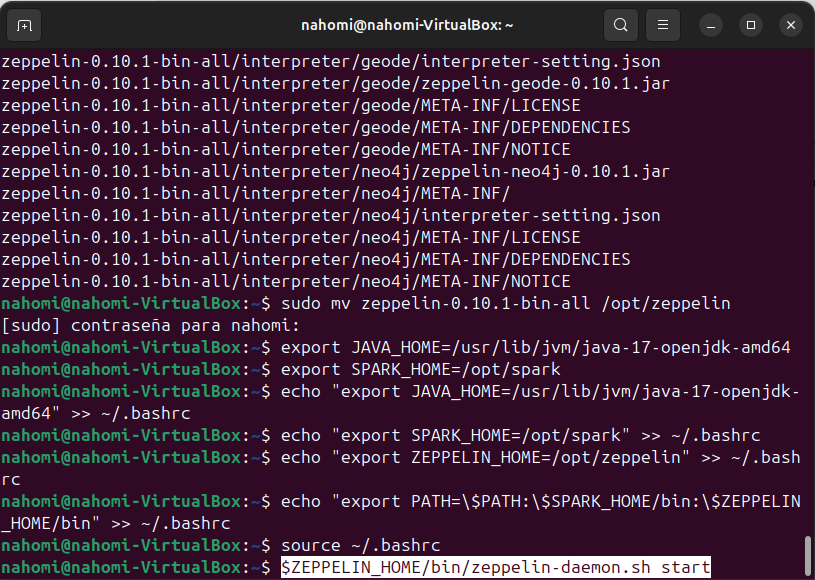
echo "export ZEPPELIN\_HOME=/opt/zeppelin" >> ~/.bashrc

echo "export PATH=\$PATH:\$SPARK\_HOME/bin:\$ZEPPELIN\_HOME/bin" >> ~/.bashrc



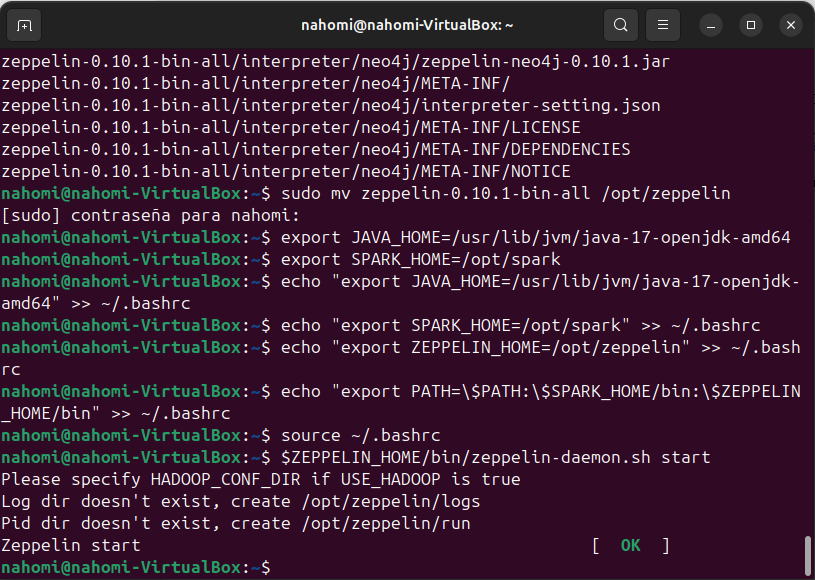
### Aplicamos cambios:

source ~/.bashrc



Iniciamos Apache Zeppelin:

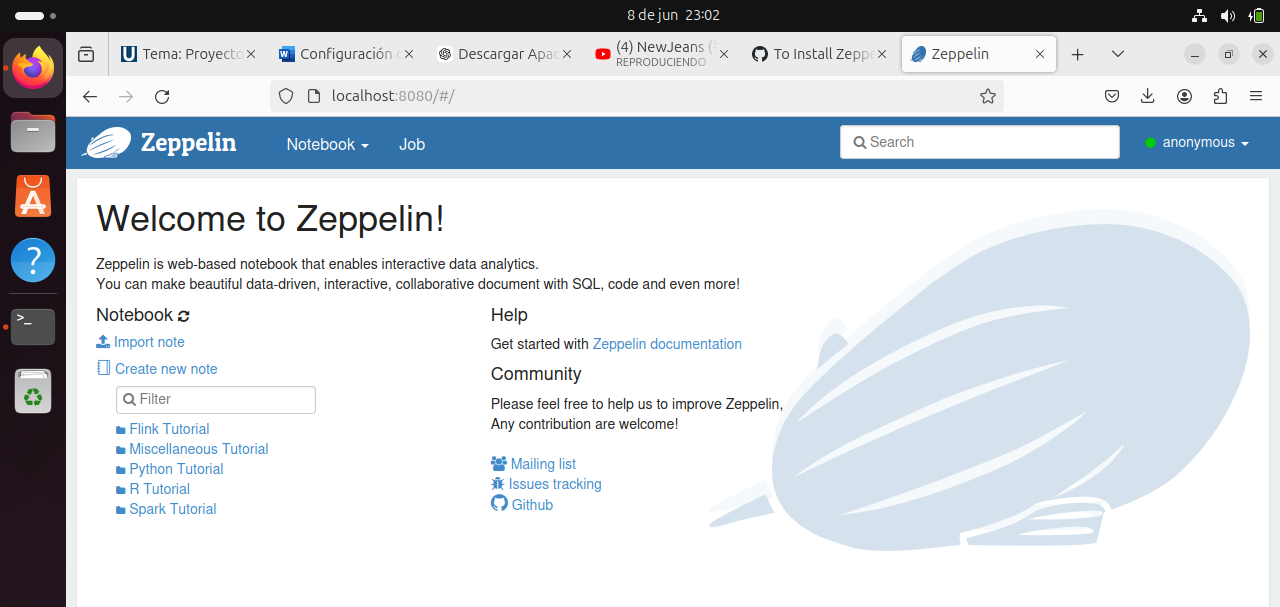
$ZEPPELIN\_HOME/bin/zeppelin-daemon.sh start



[Regresar a índice](#_Índice)

### Ingresamos a:

<http://localhost:8080>

¡Y LISTO!

[Regresar a índice](#_Índice)

### Dudas: ¿Qué es Apache Spark?

Apache Spark es un framework de procesamiento de datos en clúster de código abierto diseñado para el procesamiento rápido de grandes volúmenes de datos.

¿Qué es Apache Zeppelin?

Apache Zeppelin es una herramienta web-based notebook que puede ser utilizada como un entorno de desarrollo interactivo (IDE) para varios lenguajes de programación y frameworks de análisis de datos.

Apache Zeppelin es una herramienta poderosa para el análisis de datos interactivo, la visualización y el desarrollo de código en múltiples lenguajes. Es especialmente útil en entornos de Big Data debido a su integración con Apache Spark y otros frameworks de procesamiento de datos.

[Regresar a índice](#_Índice)

# Comandos utilizados en clases:

“pwd”, este comando sirve para ver en dónde estás ubicado en ese momento.

“ls -al”, este comando lista los archivos y sus respectivos permisos.

“cd ‘nombreCarpeta’/” te ubica en la carpeta con el nombre buscado.

“sdk list java”, hay que comprobar si no da error este comando, sino hay qye instalar sdk man.

“curl -s “https://get.sdkman.io”| bash”, este comando sirve para instalar sdkman.

“sdk install scala 2.12.1”, esta es la versión que debemos tener scala.

“wget https://archive.apache.org/dist/spark/spark-3.2.1/spark-3.2.1-bin-hadoop3.2.tgz”, instalar spark en su versión 3.2.1. (actualizar).

“$ ./zeppelin-daemo.sh start” iniciar el zeppelin