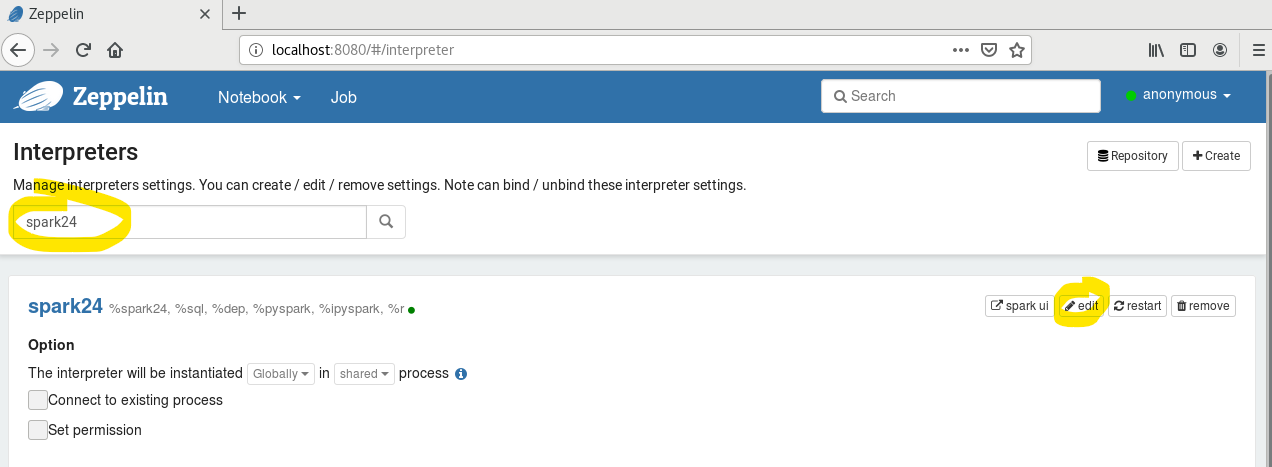
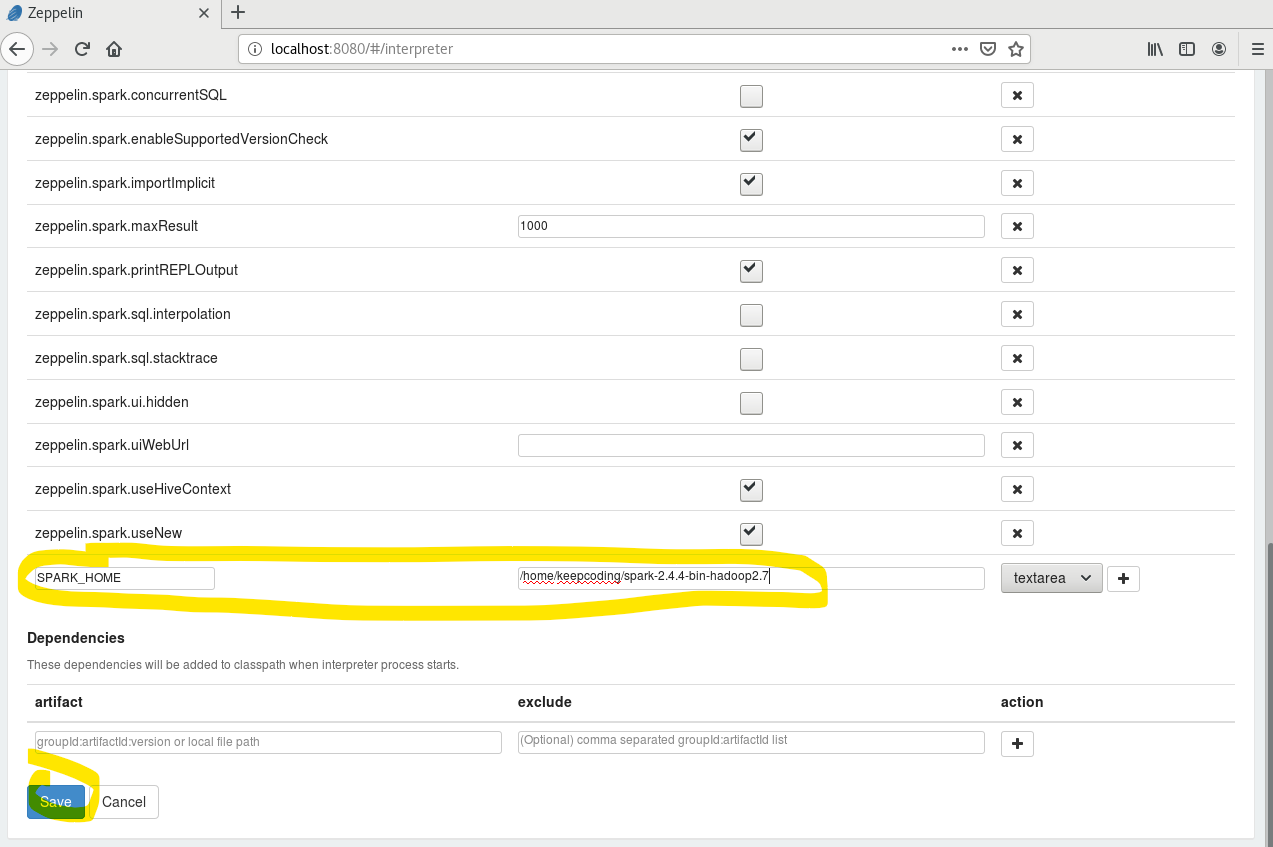


SAVE

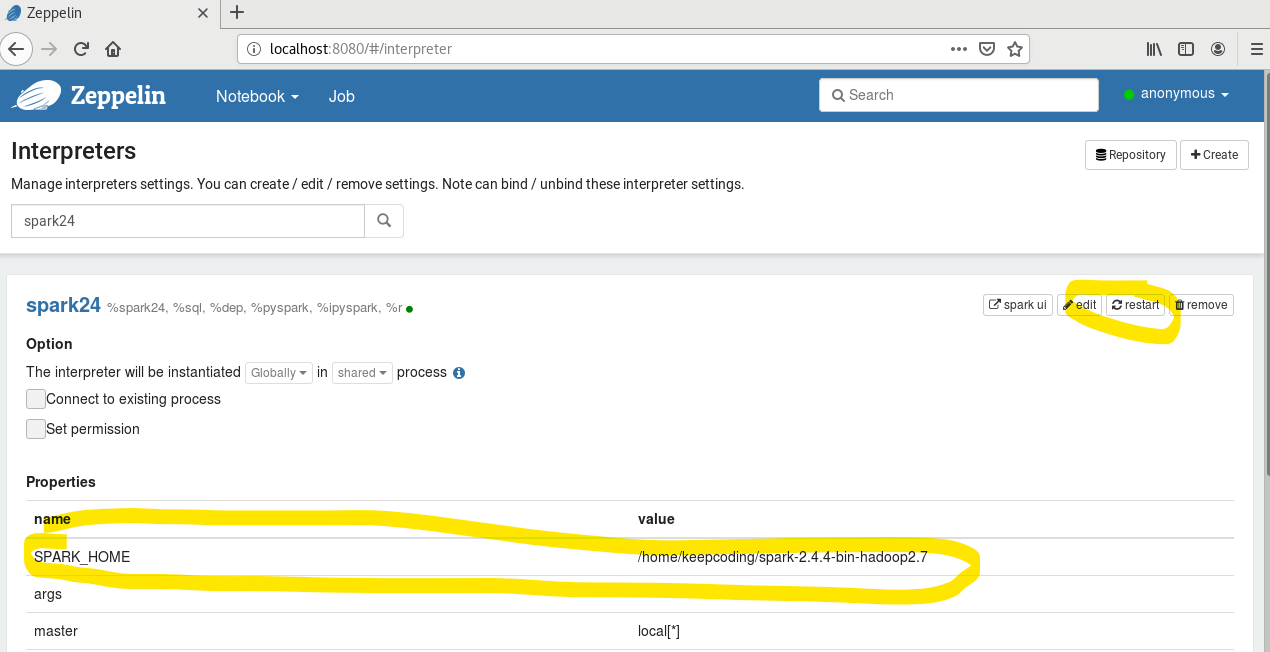
Ahora lo buscamos



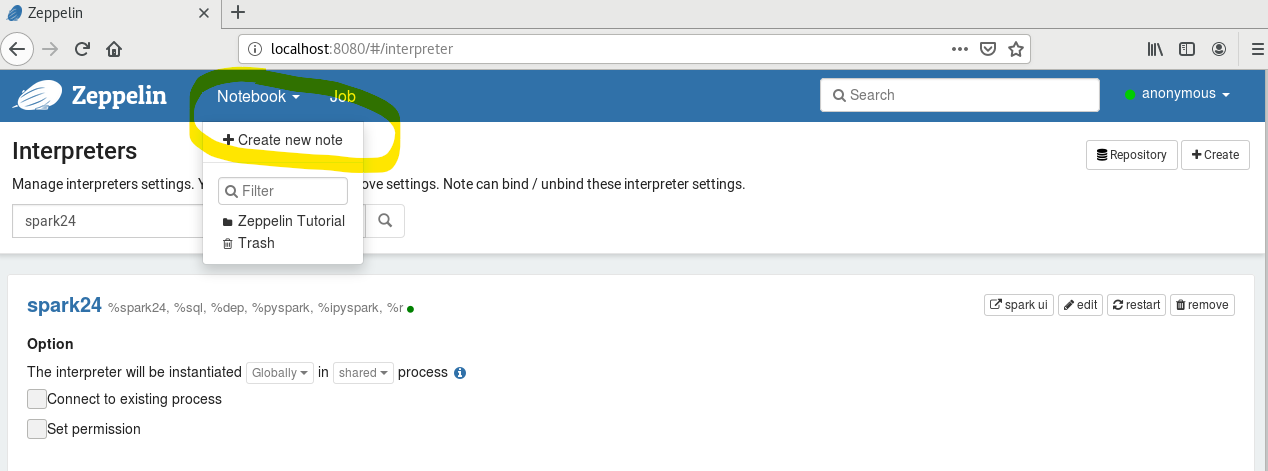
Dentro de properties añadimos la ruta donde tenemos nuestro spark

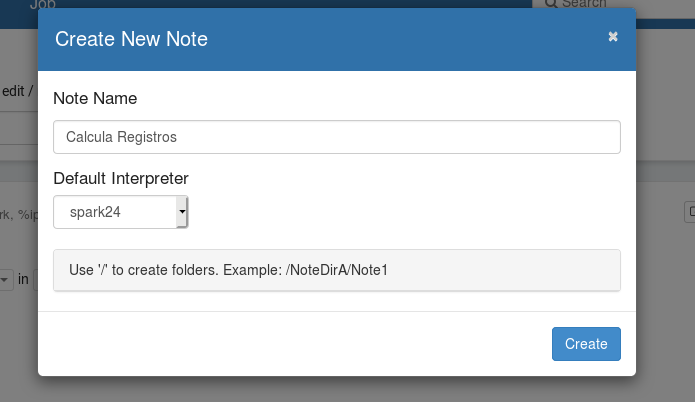


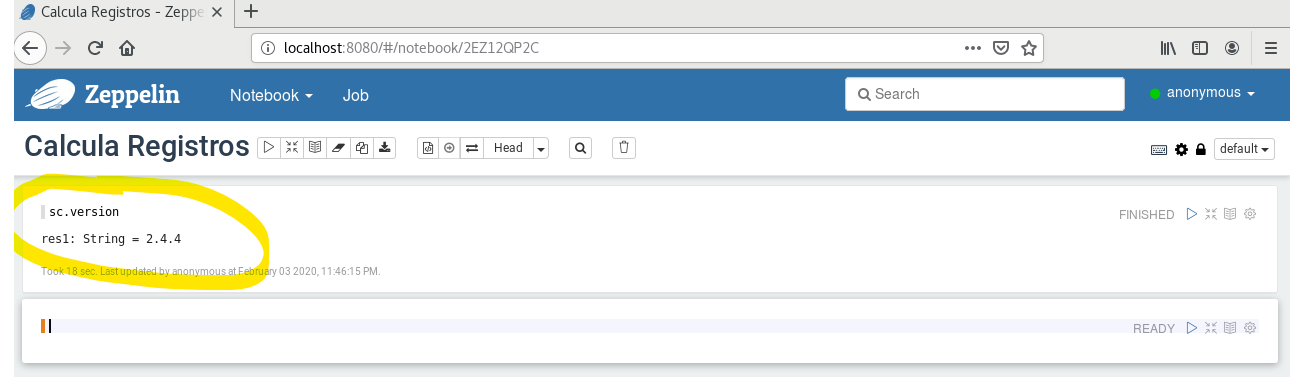
Vemos que ya está añadido en properties y restart para asegurarnos



Creamos new note en notebook para asegurarnos que esta todo bien







//importamos los paquetes necesarios

import org.apache.spark.sql.SparkSession

import org.apache.spark.sql.types.{IntegerType, StringType, StructType}

import org.apache.spark.sql.functions.\_

//definimos la SparkSession, punto de entrada a Spark SQL

val spark = SparkSession.builder()

.appName("PracticaContarRegistros")

.master("local")

.getOrCreate()

//creamos el schema del fichero csv con el que vamos a trabajar. Sabemos que son 4 campos:

// 1º identificador del usuario, 2º nombre del usuario, 3º edad del usuario, 4º nº de amigos del usuario

val schemaAmigos = new StructType().add("id", IntegerType, true)

.add("nombre", StringType, true)

.add("edad", IntegerType,true)

.add("amigos", IntegerType, true)

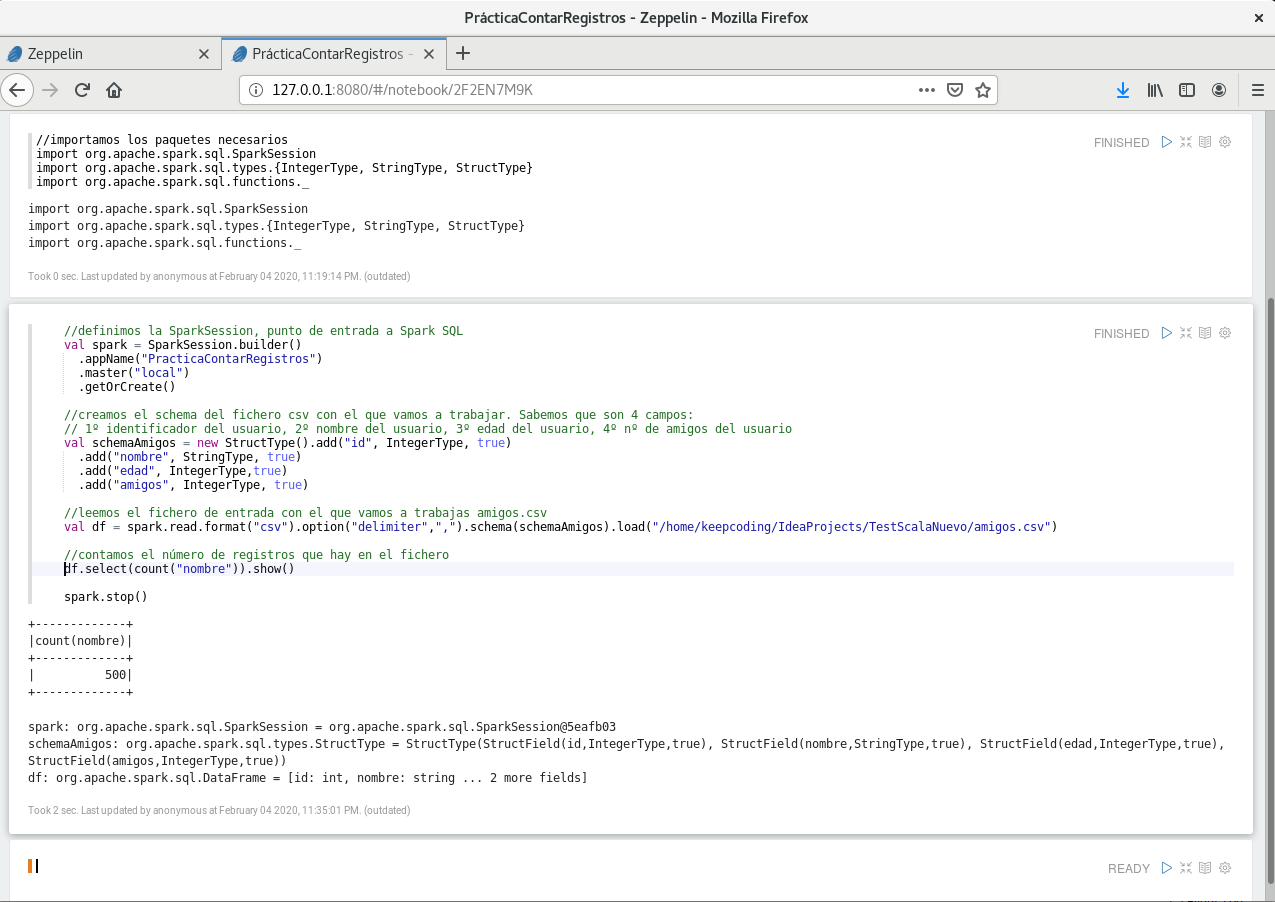
//leemos el fichero de entrada con el que vamos a trabajas amigos.csv

val df = spark.read.format("csv").option("delimiter",",").schema(schemaAmigos).load("/home/keepcoding/IdeaProjects/TestScalaNuevo/amigos.csv")

//contamos el número de registros que hay en el fichero

df.select(count("nombre")).show()

spark.stop()



Zookeeper

root@debian:/home/kafka/kafka\_2.11-2.4.0# bin/zookeeper-server-start.sh config/zookeeper.properties

kafka

root@debian:/home/kafka/kafka\_2.11-2.4.0# bin/kafka-server-start.sh config/server.properties

crear topic

root@debian:/home/kafka/kafka\_2.11-2.4.0# bin/kafka-topics.sh --create --zookeeper localhost:2181 --replication-factor 1 --partitions 1 --topic PracticaKafka

Created topic PracticaKafka.

productor

root@debian:/home/kafka/kafka\_2.11-2.4.0# cat bin/personal.json| bin/kafka-console-producer.sh --broker-list localhost:9092 --topic PracticaKafka

consumidor

root@debian:/home/kafka/kafka\_2.11-2.4.0# bin/kafka-console-consumer.sh --bootstrap-server localhost:9092 --topic PracticaKafka --from-beginning

