

# Java-Success.com

Prepare to fast-track, choose & go places with 800+ Java & Big Data Q&As with lots of code & diagrams.

[Home](#) [Why? ▾](#) [300+ Java FAQs ▾](#) [300+ Big Data FAQs ▾](#) [Courses ▾](#)[👤 Membership ▾](#) [Your Career ▾](#)[Home](#) > [bigdata-success.com](#) > [Tutorials - Big Data](#) > [TUT - Spark Scala on Zeppelin](#) >

01A: Spark on Zeppelin – Docker pull from Docker hub

## 01A: Spark on Zeppelin – Docker pull from Docker hub

 Posted on [July 22, 2018](#)

**Pre-requisite:** Docker is installed on your machine for Mac OS X (E.g. \$ brew cask install docker) or Windows 10. [Docker interview Q&As](#).

### What is Apache Zeppelin?

Zeppelin is a web based notebook to execute arbitrary code in Scala, SQL, Spark, etc. You can mix languages. Apache Zeppelin helps data analysts, data scientist, and business users to get better understanding of data. As described below you can quickly explore data, create visualizations and share

### 300+ Java Interview FAQs

300+ Java FAQs



16+ Java Key Areas Q&amp;As



150+ Java Architect FAQs



80+ Java Code Quality Q&amp;As



150+ Java Coding Q&amp;As



### 300+ Big Data Interview FAQs

300+ Big Data FAQs



Tutorials - Big Data

TUT -  Starting Big Data

TUT - Starting Spark &amp; Scala

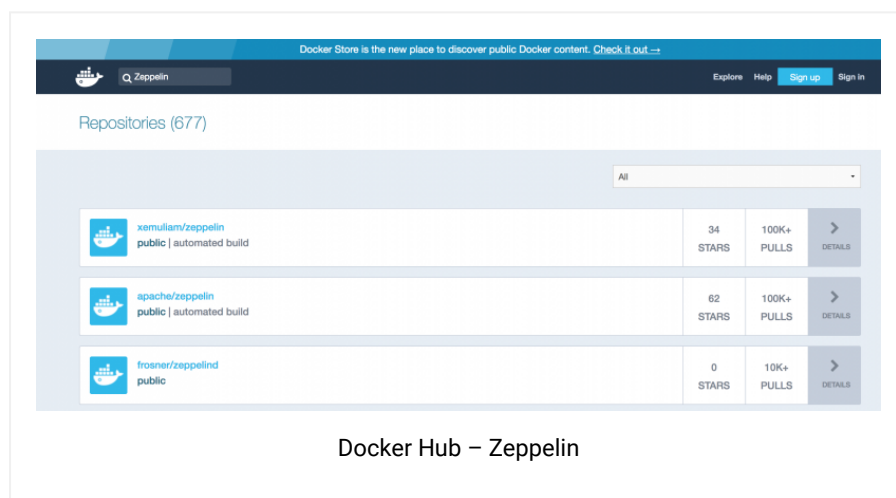
their insights, as web pages, with various stakeholders. For example

- 1) Prepare data using Shell by say downloading files with curl/wget, and then inject to HDFS.
- 2) Perform data analytics with Spark (i.e Scala) or pyspark (i.e. Python).
- 3) Perform simple visualizations in SQL.
- 4) Export the results with Shell, and publish to create graphs.

## How to install Apache Zeppelin on Docker

**Step 1:** Go to the **Docker Hub** <https://hub.docker.com/>, which is the repository for the images that you can pull create isolated containers.

**Step 2:** Search for “Zeppelin”.



**Step 3:** Select “**apache/zeppelin**”. Click on “**Dockerfile**” and inspect what is getting installed FYI.

TUT - Starting with Python

TUT - Kafka

TUT - Pig

TUT - Apache Storm

TUT - Spark Scala on Zeppelin

TUT - Cloudera

TUT - Cloudera on Docker

TUT - File Formats

TUT - Spark on Docker

TUT - Flume

TUT - Hadoop (HDFS)

TUT - HBase (NoSQL)

TUT - Hive (SQL)

TUT - Hadoop & Spark

TUT - MapReduce

TUT - Spark and Scala

TUT - Spark & Java

TUT - PySpark on Databricks

TUT - Zookeeper

## 800+ Java Interview Q&As

300+ Core Java Q&As



300+ Enterprise Java Q&As



150+ Java Frameworks Q&As



120+ Companion Tech Q&As



Tutorials - Enterprise Java



Click on “Build details” to get the version or tag. For example “0.8.0” or 0.7.3.

**Step 4:** Pull this from the docker hub, and build the image with the following command.

```
1 $ docker pull apache/zeppelin:0.8.0
2
```

This may take several minutes to download and create an image. Once done check the image with

```
1 $ docker images
2
1 REPOSITORY          TAG
2 apache/zeppelin     0.8.0
3
```

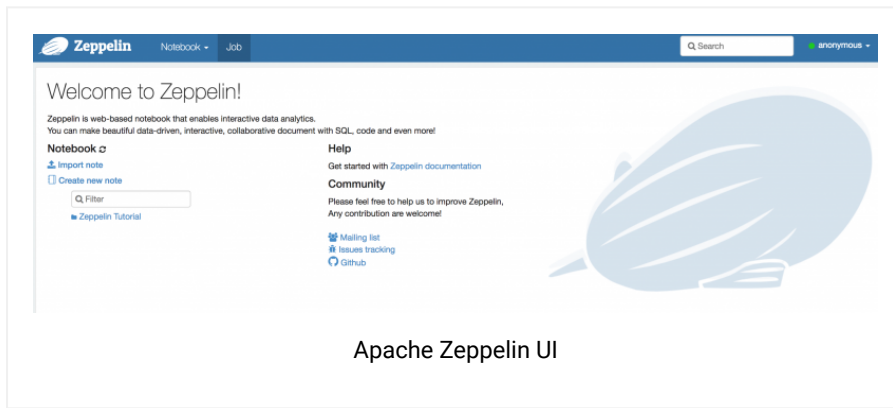
**Step 5:** Run the above image to create a container with the following command.

```
1 $ docker run -it -p 8080:8080 apache/zeppelin:0.8.0
2
```

You can open another terminal, and check if the container is up and running with:

```
1 $ docker ps
2
1 CONTAINER ID        IMAGE               COMMAND
2 b0461edaa8f0       apache/zeppelin:0.8.0  "/usr/l
3
```

**Step 6:** Go to a browser and type: “http://localhost:8080”.



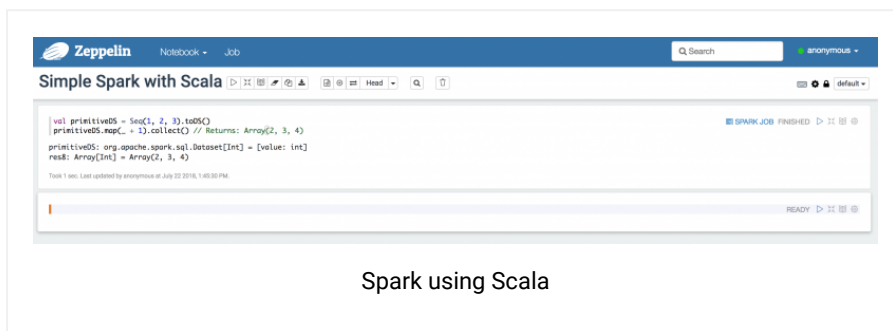
**Step 7:** Select the link “Create new note”, and name it “Simple Spark with Scala” and select the **interpreter** as “spark”.

Type the following simple Spark code to add 1 to the given set of numbers.

```
1 val primitiveDS = Seq(1, 2, 3).toDS()
2 primitiveDS.map(_ + 1).collect() // Returns: Array(2, 3, 4)
3
```

Press the play button, and the output will be:

```
1 Array(2, 3, 4)
2
```



◀ 01B: Spark on Zeppelin – custom Dockerfile

02: Spark on Zeppelin – read a file from local file system ▶

## Disclaimer

The contents in this Java-Success are copyrighted and from EmpoweringTech pty ltd. The EmpoweringTech pty ltd has the right to correct or enhance the current content without any prior notice. These are general advice only, and one needs to take his/her own circumstances into consideration. The EmpoweringTech pty ltd will not be held liable for any damages caused or alleged to be caused either directly or indirectly by these materials and resources. Any trademarked names or labels used in this blog remain the property of their respective trademark owners. Links to external sites do not imply endorsement of the linked-to sites. [Privacy Policy](#).