docker

노트북:	DE		
만든 날짜:	2018-07-19 오후 5:43	**수정한 날짜:**	2018-07-21 오후 3:26
작성자:	yyyura@naver.com		

docker

```
ubuntu@ip-172-31-26-63:~$ docker run hello-world
Unable to find image 'hello-world:latest' locally
latest: Pulling from library/hello-world
9db2ca6ccae0: Pull complete
Digest: sha256:4b8ff392a12ed9ea17784bd3c9a8b1fa3299cac44aca35a85c90c5e3c7afacdc
Status: Downloaded newer image for hello-world:latest

Hello from Docker!
This message shows that your installation appears to be working correctly.

To generate this message, Docker took the following steps:
 1. The Docker client contacted the Docker daemon.
 2. The Docker daemon pulled the "hello-world" image from the Docker Hub.
    (amd64)
 3. The Docker daemon created a new container from that image which runs the
    executable that produces the output you are currently reading.
 4. The Docker daemon streamed that output to the Docker client, which sent it
    to your terminal.

To try something more ambitious, you can run an Ubuntu container with:
 $ docker run -it ubuntu bash

Share images, automate workflows, and more with a free Docker ID:
 https://hub.docker.com/

For more examples and ideas, visit:
 https://docs.docker.com/engine/userguide/
```

이미지 설치

- docker pull ubuntu:16.04 # ubuntu:16.04 이미지 설치

이미지 목록 출력

- docker images

```
ubuntu@ip-172-31-26-63:~$ docker pull ubuntu:16.04
16.04: Pulling from library/ubuntu
3620e2d282dc: Pull complete
ef22f5e4b3b2: Pull complete
99f229f854da: Pull complete
4fe433abe16a: Pull complete
c9b72a16d85e: Pull complete
Digest: sha256:14066a391d902c386d6164d44ade3460ba044abcdf8df88b0ff79a6f635be8d3
Status: Downloaded newer image for ubuntu:16.04
ubuntu@ip-172-31-26-63:~$ docker images
REPOSITORY          TAG                 IMAGE ID            CREATED             SIZE
ubuntu              16.04               e13f3d529b1a        2 days ago          115MB
hello-world         latest              2cb0d9787c4d        8 days ago          1.85kB
```

컨테이너 목록 출력

- docker ps -a

```
ubuntu@ip-172-31-26-63:~$ docker ps -a
CONTAINER ID        IMAGE               COMMAND             CREATED             STATUS                          PORTS               NAMES
3979e4276f9e        hello-world         "/hello"            About an hour ago   Exited (0) About an hour ago                        serene_dubinsky
```

이미지로부터 -> 컨테이너 생성하기

- $ docker run [OPTIONS] IMAGE [COMMAND]
- ubuntu:16.04 이미지를 이용하여 yura_container란 이름의 container를 만들어
- bash 창을 실행하고 bash shell에 접속한다
- 이때 host의 10001 포트와 container의 10001번 포트가 연결된다

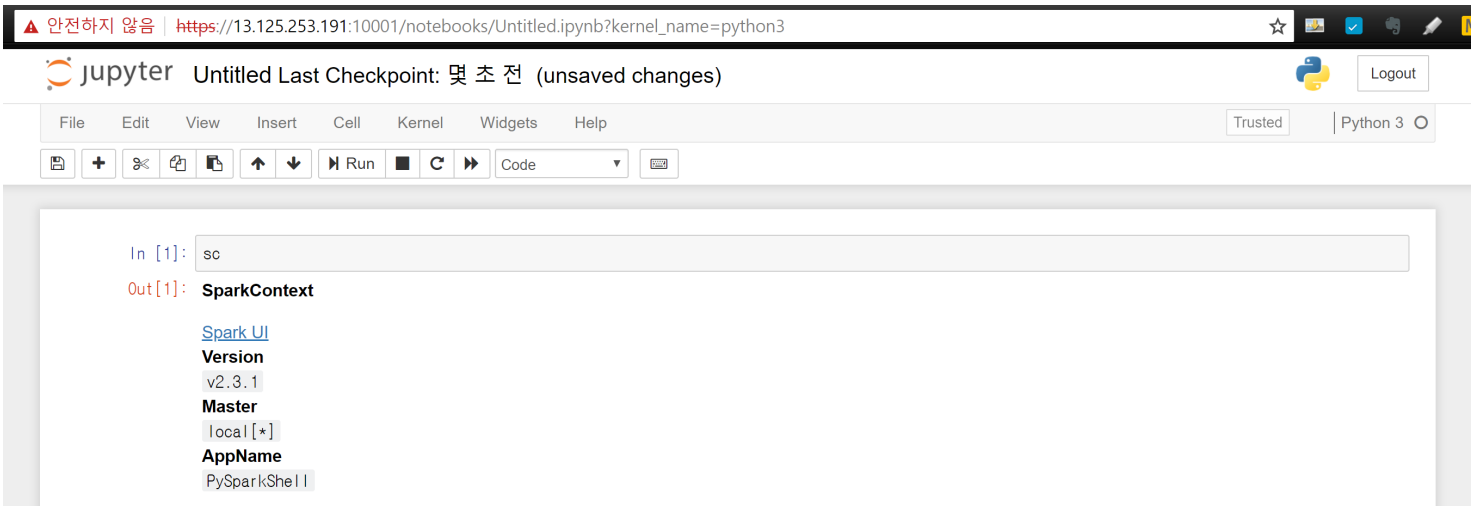
```
ubuntu@ip-172-31-26-63:~$ docker run -it --name yura_container -p 10001:10001 -v ~/workspace:/root/workspace ubuntu:16.04 /bin/bash
root@fdc4901b0427:/# ls
bin  boot  dev  etc  home  lib  lib64  media  mnt  opt  proc  root  run  sbin  srv  sys  tmp  usr  var
ubuntu@ip-172-31-26-63:~$ docker ps -a
CONTAINER ID        IMAGE               COMMAND             CREATED             STATUS                          PORTS               NAMES
fdc4901b0427        ubuntu:16.04        "/bin/bash"         2 minutes ago       Exited (0) 7 seconds ago                            yura_container
3979e4276f9e        hello-world         "/hello"            2 hours ago         Exited (0) 33 minutes ago                           serene_dubinsky
```

```
ubuntu@ip-172-31-26-63:~$ git clone https://github.com/xodhx4/ybigta_pyspark_docker.git
Cloning into 'ybigta_pyspark_docker'...
remote: Counting objects: 107, done.
remote: Compressing objects: 100% (48/48), done.
remote: Total 107 (delta 59), reused 106 (delta 59), pack-reused 0
Receiving objects: 100% (107/107), 26.47 KiB | 0 bytes/s, done.
Resolving deltas: 100% (59/59), done.
Checking connectivity... done.
```

```
ubuntu@ip-172-31-26-63:~/ybigta_pyspark_docker$ ls
codealike.json  Dockerfile     hdfs-site.xml  hive_init.sh   jupyter_init.sh            mapred-site.xml  spark-env.sh  start.sh
core-site.xml   hadoop-env.sh  hive-env.sh    hive-site.xml  jupyter_notebook_config.py  README.md        ssh_init.sh   yarn-site.xml
```

```
ubuntu@ip-172-31-26-63:~/ybigta_pyspark_docker$ docker build --tag yura_spark:1.0 .
Sending build context to Docker daemon  139.8kB
Step 1/38 : FROM ubuntu:16.04
 ---> e13f3d529b1a
Step 2/38 : MAINTAINER Taeoh Kim <kimtaeoh95@gmail.com>
 ---> Running in 5937dbef4efa
Removing intermediate container 5937dbef4efa
 ---> fbcb1ce575a9
Step 3/38 : RUN apt-get update && apt-get install -yqq     wget     bzip2     git &&     wget https://repo.anaconda.com/archive/Anaconda3-5.2.0-Linu
x-x86_64.sh -O ~/anaconda.sh &&     /bin/bash ~/anaconda.sh -b -p /opt/conda &&     rm ~/anaconda.sh
 ---> Running in 687579ea0a6a
Get:1 http://security.ubuntu.com/ubuntu xenial-security InRelease [107 kB]
Get:2 http://archive.ubuntu.com/ubuntu xenial InRelease [247 kB]
Get:3 http://security.ubuntu.com/ubuntu xenial-security/universe Sources [84.1 kB]
Get:4 http://archive.ubuntu.com/ubuntu xenial-updates InRelease [109 kB]
Get:5 http://security.ubuntu.com/ubuntu xenial-security/main amd64 Packages [668 kB]
Get:6 http://archive.ubuntu.com/ubuntu xenial-backports InRelease [107 kB]
Get:7 http://archive.ubuntu.com/ubuntu xenial/universe Sources [9802 kB]
Get:8 http://security.ubuntu.com/ubuntu xenial-security/restricted amd64 Packages [12.7 kB]
```

```
ubuntu@ip-172-31-26-63:~/ybigta_pyspark_docker$ docker run -it --name spark_container2 -p 10001:10001 -v ~/workspace:/root/workspace yura_spark:1.0
 * Starting OpenBSD Secure Shell server sshd                                                                                                  [ OK ]
Starting namenodes on [localhost]
localhost: starting namenode, logging to /root/hadoop-2.9.0/logs/hadoop-root-namenode-7fdf2dcdc31b.out
localhost: starting datanode, logging to /root/hadoop-2.9.0/logs/hadoop-root-datanode-7fdf2dcdc31b.out
Starting secondary namenodes [localhost]
localhost: starting secondarynamenode, logging to /root/hadoop-2.9.0/logs/hadoop-root-secondarynamenode-7fdf2dcdc31b.out
starting yarn daemons
starting resourcemanager, logging to /root/hadoop-2.9.0/logs/yarn--resourcemanager-7fdf2dcdc31b.out
localhost: starting nodemanager, logging to /root/hadoop-2.9.0/logs/yarn-root-nodemanager-7fdf2dcdc31b.out
2018-07-21 04:10:25: Starting Hive Metastore Server
SLF4J: Class path contains multiple SLF4J bindings.
SLF4J: Found binding in [jar:file:/root/apache-hive-2.3.3-bin/lib/log4j-slf4j-impl-2.6.2.jar!/org/slf4j/impl/StaticLoggerBinder.class]
SLF4J: Found binding in [jar:file:/root/hadoop-2.9.0/share/hadoop/common/lib/slf4j-log4j12-1.7.25.jar!/org/slf4j/impl/StaticLoggerBinder.class]
SLF4J: See http://www.slf4j.org/codes.html#multiple_bindings for an explanation.
SLF4J: Actual binding is of type [org.apache.logging.slf4j.Log4jLoggerFactory]
root@7fdf2dcdc31b:~# [I 04:10:28.967 NotebookApp] Writing notebook server cookie secret to /root/.local/share/jupyter/runtime/notebook_cookie_secret
[I 04:10:29.280 NotebookApp] JupyterLab beta preview extension loaded from /opt/conda/lib/python3.6/site-packages/jupyterlab
[I 04:10:29.280 NotebookApp] JupyterLab application directory is /opt/conda/share/jupyter/lab
[I 04:10:29.288 NotebookApp] Serving notebooks from local directory: /root
[I 04:10:29.289 NotebookApp] 0 active kernels
[I 04:10:29.289 NotebookApp] The Jupyter Notebook is running at:
[I 04:10:29.289 NotebookApp] https://7fdf2dcdc31b:10001/
[I 04:10:29.289 NotebookApp] Use Control-C to stop this server and shut down all kernels (twice to skip confirmation).
```



---

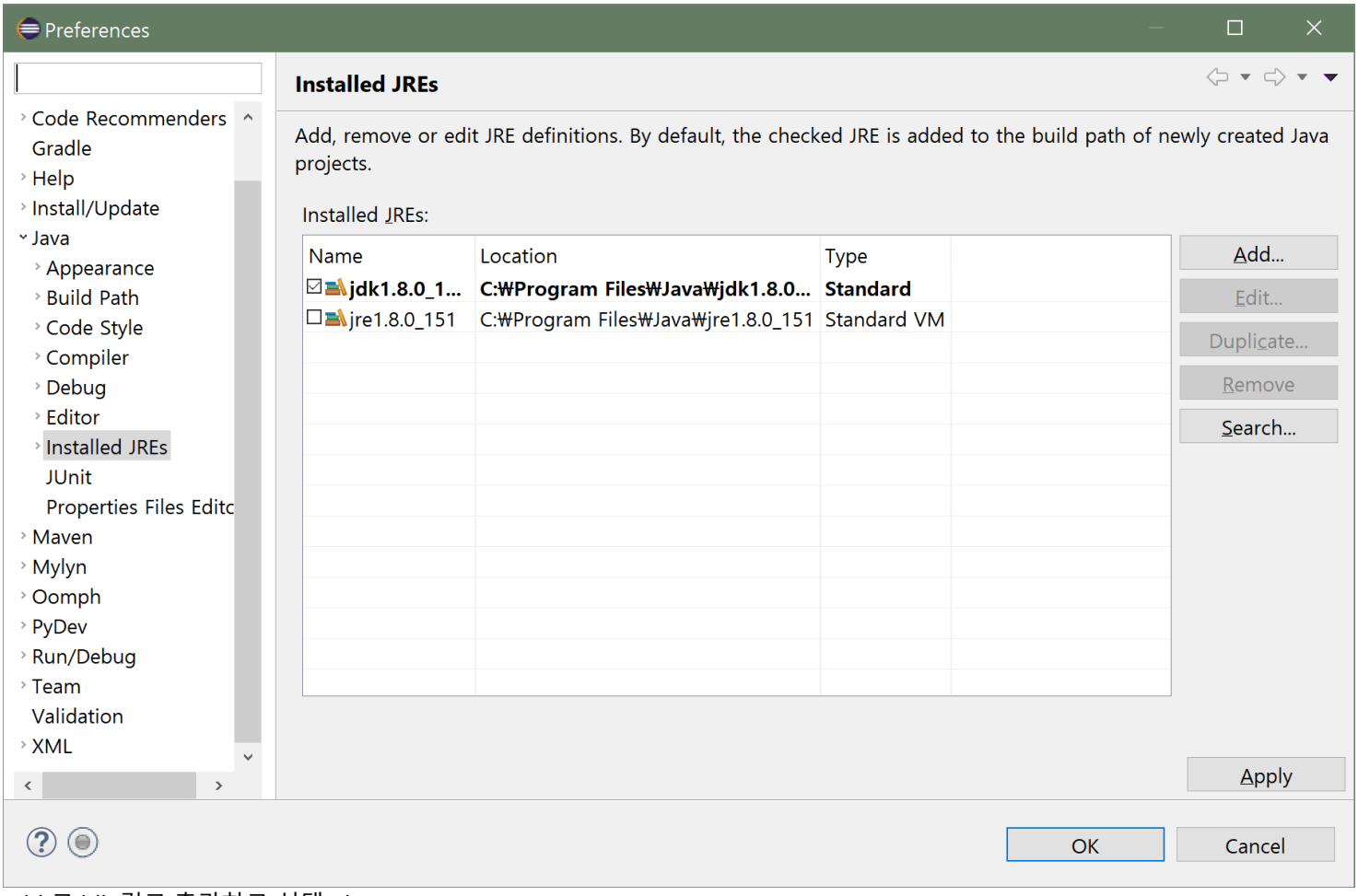
window->perspective->open perspective->other
git clone repository
url 붙이기
next,next
finish
하면 폴더가 생성됨

우측 키-> import project->finish
우클릭 -> run as -> maven build

안나오면
window->perspective->open perspective->other
java(default)로 바꾸기

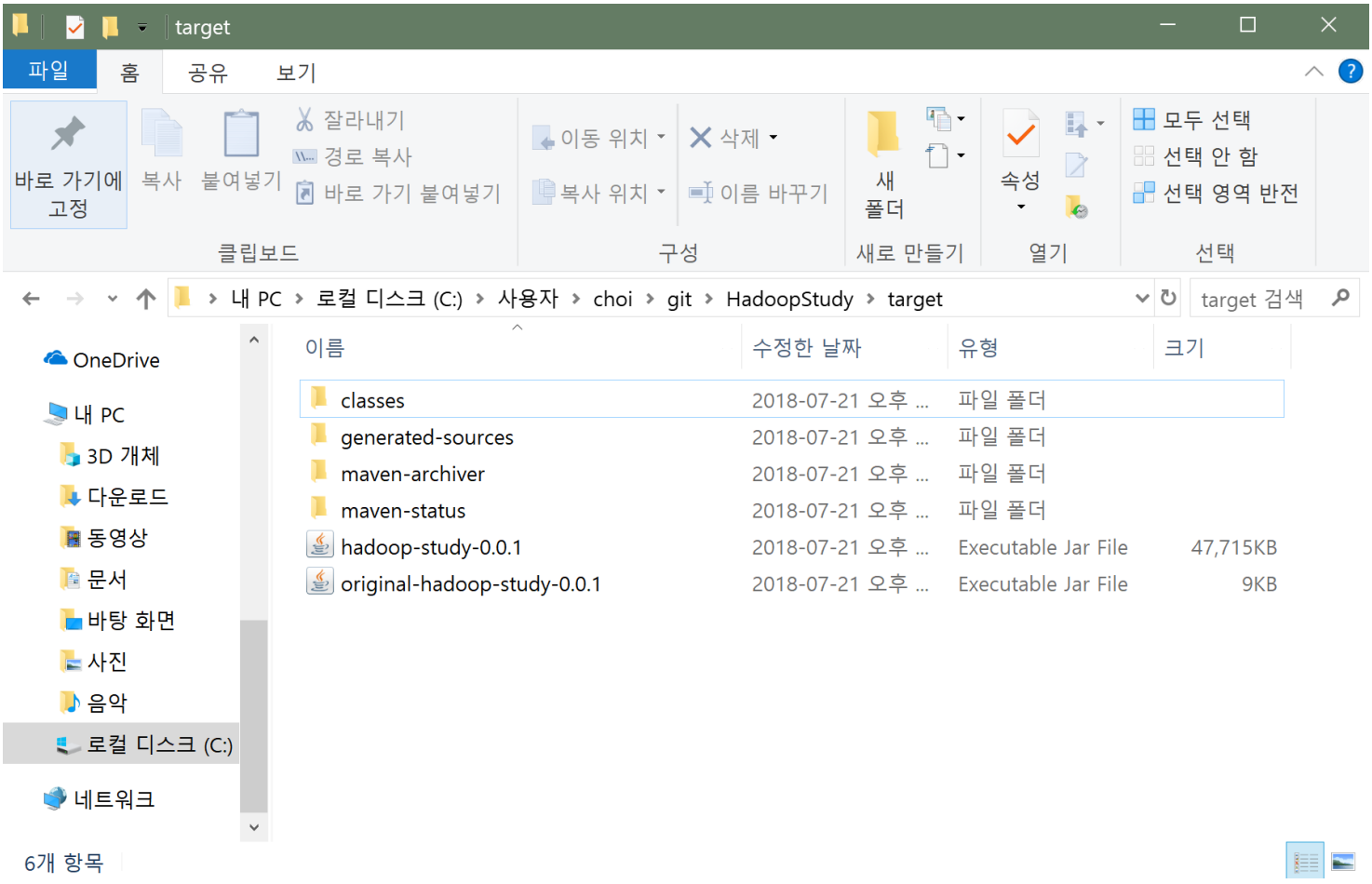
다시 우클릭 run as -> maven build -> goal에 clean install

에러 날 때:
window->preference ->java-> installed JRE



add 로 jdk 경로 추가하고 선택 ok
후 다시 run as -> maven build 끝

그럼 jar파일 생성됨



생성된 jar 파일을 local로 옮겨서 shakespeare파일 워드카운트 해본다
hdfs에 /root/yura/week3/ 폴더를 생성하고 싶을때

```
~/hadoop$ hdfs dfs -mkdir /root/
~/hadoop$ hdfs dfs -mkdir /root/yura/
~/hadoop$ hdfs dfs -mkdir /root/yura/week3/
```

/root/yura/week3/ 는 워드카운트 할 데이터가 있는 hdfs 내 폴더,
/root/yura/output_1/ 는 결과를 저장할 새로운 폴더(기존에 존재하지 않는 폴더)

```
ubuntu@ip-172-31-27-161:~/hadoop/bin$ ./hadoop jar /home/ubuntu/hadoop-study-0.0.1.jar com.ybigt
a.example.wordcount.WordCountDriver /root/yura/week3/ /root/yura/week3/output_1/
18/07/21 05:45:51 INFO client.RMProxy: Connecting to ResourceManager at localhost/127.0.0.1:8032
18/07/21 05:45:52 INFO input.FileInputFormat: Total input files to process : 6
18/07/21 05:45:52 INFO mapreduce.JobSubmitter: number of splits:6
18/07/21 05:45:52 INFO Configuration.deprecation: yarn.resourcemanager.system-metrics-publisher.
enabled is deprecated. Instead, use yarn.system-metrics-publisher.enabled
18/07/21 05:45:52 INFO mapreduce.JobSubmitter: Submitting tokens for job: job_1532146724932_0002
18/07/21 05:45:52 INFO impl.YarnClientImpl: Submitted application application_1532146724932_0002
18/07/21 05:45:52 INFO mapreduce.Job: The url to track the job: http://ip-172-31-27-161.ap-north
east-2.compute.internal:8088/proxy/application_1532146724932_0002/
18/07/21 05:45:52 INFO mapreduce.Job: Running job: job_1532146724932_0002
18/07/21 05:45:59 INFO mapreduce.Job: Job job_1532146724932_0002 running in uber mode : false
18/07/21 05:45:59 INFO mapreduce.Job:  map 0% reduce 0%
18/07/21 05:46:11 INFO mapreduce.Job:  map 17% reduce 0%
18/07/21 05:46:12 INFO mapreduce.Job:  map 50% reduce 0%
18/07/21 05:46:13 INFO mapreduce.Job:  map 100% reduce 0%
18/07/21 05:46:16 INFO mapreduce.Job:  map 100% reduce 100%
```

```
ubuntu@ip-172-31-27-161:~$ hdfs dfs -ls /root/yura/week3/
Found 7 items
-rw-r--r--   1 ubuntu supergroup    1784616 2018-07-21 05:37 /root/yura/week3/comedies
-rw-r--r--   1 ubuntu supergroup      58976 2018-07-21 05:37 /root/yura/week3/glossary
-rw-r--r--   1 ubuntu supergroup    1479035 2018-07-21 05:37 /root/yura/week3/histories
-rw-r--r--   1 ubuntu supergroup     238027 2018-07-21 05:37 /root/yura/week3/input.txt
drwxr-xr-x   - ubuntu supergroup          0 2018-07-21 05:46 /root/yura/week3/output_1
-rw-r--r--   1 ubuntu supergroup     268140 2018-07-21 05:37 /root/yura/week3/poems
-rw-r--r--   1 ubuntu supergroup    1752440 2018-07-21 05:37 /root/yura/week3/tragedies
ubuntu@ip-172-31-27-161:~$ hdfs dfs -ls /root/yura/week3/output_1
Found 2 items
-rw-r--r--   1 ubuntu supergroup          0 2018-07-21 05:46 /root/yura/week3/output_1/_SUCCESS
-rw-r--r--   1 ubuntu supergroup     872402 2018-07-21 05:46 /root/yura/week3/output_1/part-r-00
000
```