Project 4 - Data Analytics with Hadoop

Project ini merupakan tugas individu dan ditujukan untuk melatih dan mengukur pemahaman terhadap konsep pada materi Data Analytics with Hadoop .

Soal:

- Analisis code project
- Jalankan project dilocal machine masing-masing
- Tambahkan code mapreduce untuk mendapatkan total pendapatan transaksi setiap bulannya
- Kumpulkan tugas dalam bentuk link Github.

Upload project yang sudah dibahas dikelas ke dalam Google Classroom paling lambat Hari Rabu, Jam 23.59 WIB.

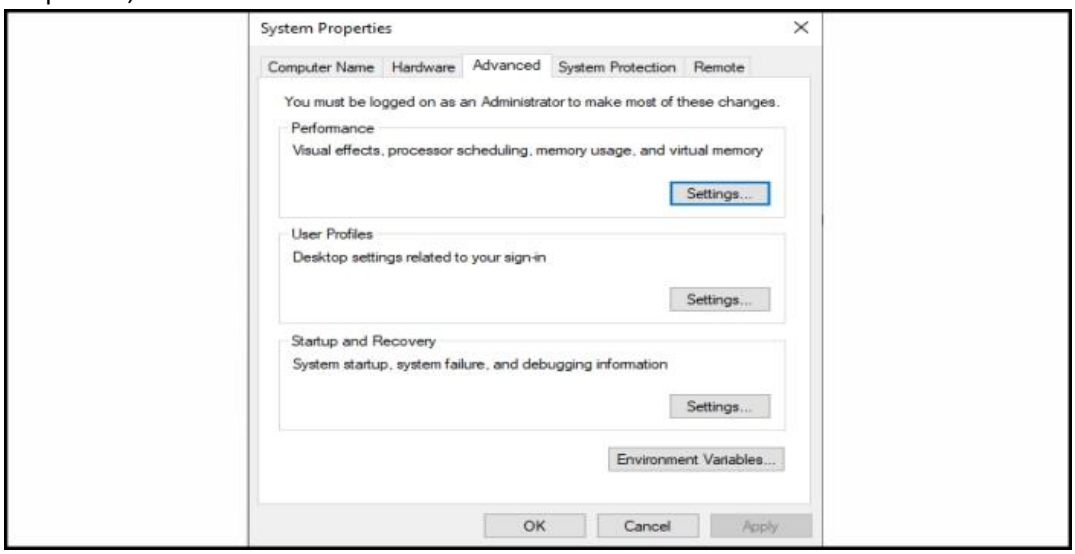
<u>Jawaban</u>

Panduan Install Apache Hadoop 3.2.1 pada OS Windows

Prasyarat sebelum melakukan instalasi Hadoop versi 3.2.1 pada windows 10 adalah melakukan instalasi Java. Semua versi hadoop hanya support pada Java versi 8. Berikut langkah-langkahnya:

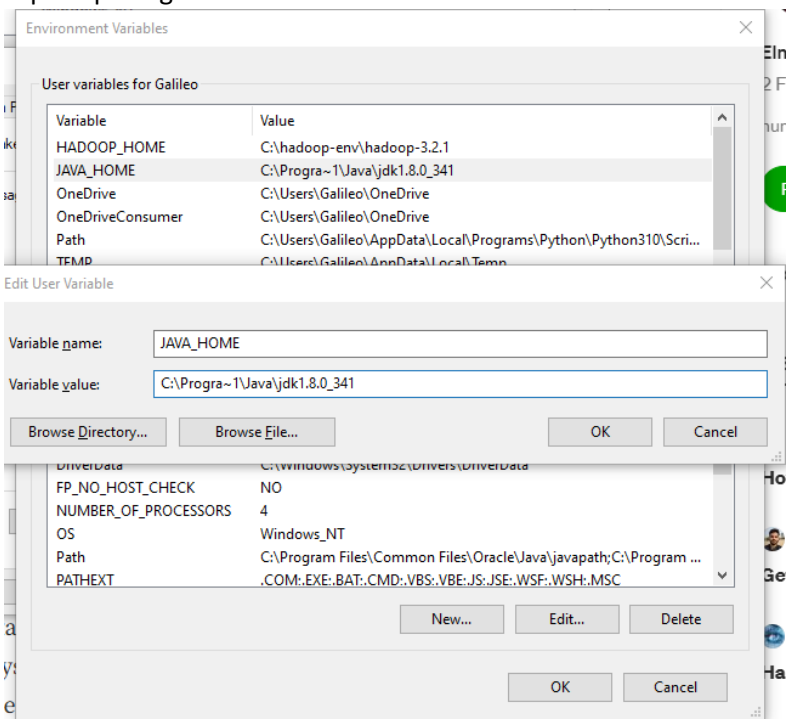
1. Masuk ke website oracle berikut https://www.oracle.com/java/technologies/javase/javase-jdk8-downloads.html
2. Download Java Development Kit 8 (JDK 8) windows x64
3. Instalasi Java
   a. Buka file jdk .exe yang telah download, kemudian ikuti proses instalasi sampai selesai

b. Lakukan Settings Environment Variables Java. Pertama, buka Control Panel – System and Security – System – Advanced System Settings. Kemudian akan muncul dialog box System Properties, lalu klik Environment Variables.

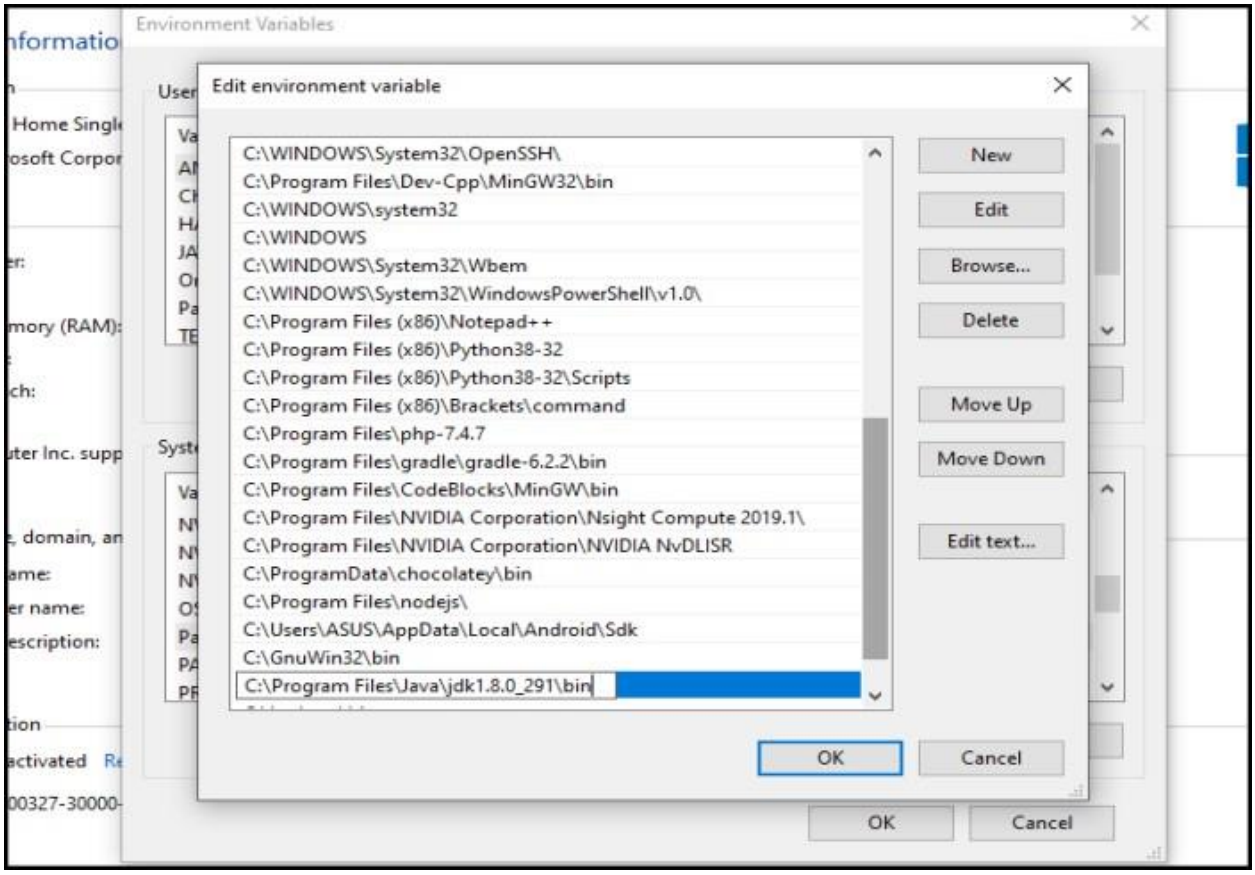


c. Atur Home Java. Pada User variables, klik New kemudian isi Variable name dengan JAWA_HOME dan Variable value dengan direktori C:\Program Files\Java\ jdk1.8.0_341. Kemudian klik OK.

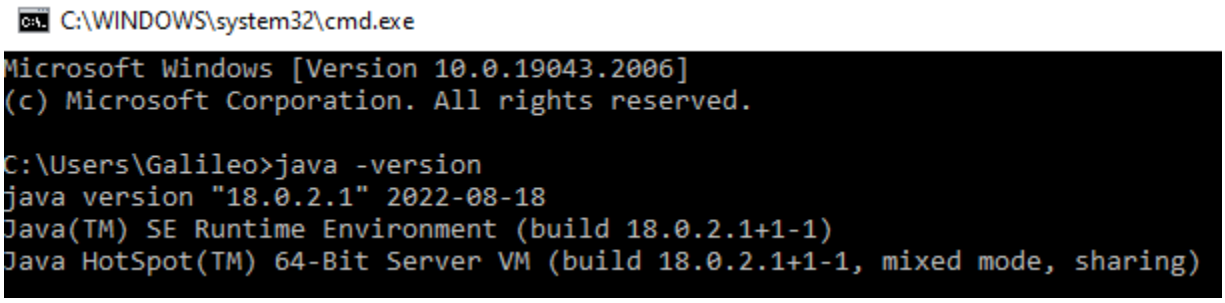
Di tahap ini, perlu ada sedikit modifikasi dimana penulisan Variable value tidak diperkenankan ada spasi (Program Files) karena akan menyebabkan perintah tidak berjalan. Maka lakukan sedikit editing, ubah Program Files menjadi Progra~1. Atau jika file jdk kamu berada di Program Files(x86) maka ubah menjadi Progra~2. Sehingga tampilannya akan seperti pada gambar di bawah.



d. Atur path Java. Pada System variables, klik Path. Kemudian klik New dan isi dengan direktori jdk\bin. Lalu klik OK pada Edit environment variable, Environment Variables dan System Properties.

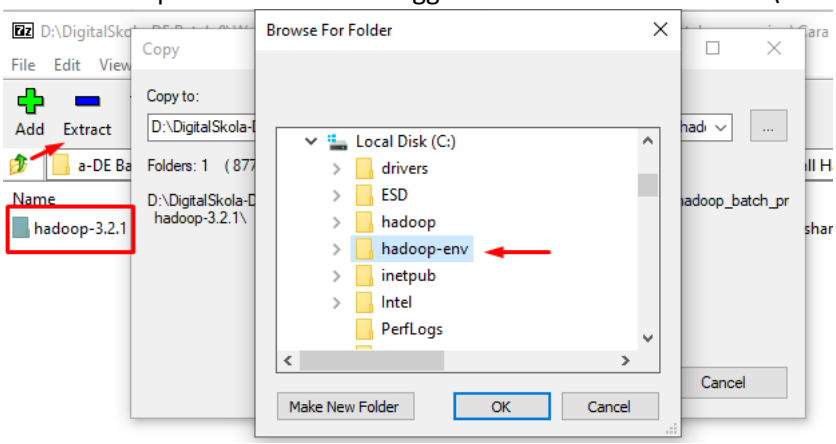


e. Buka cmd, cek versi java dengan perintah **java –version**. Apabila muncul versi java yang diinstall, maka proses instalasi berhasil.

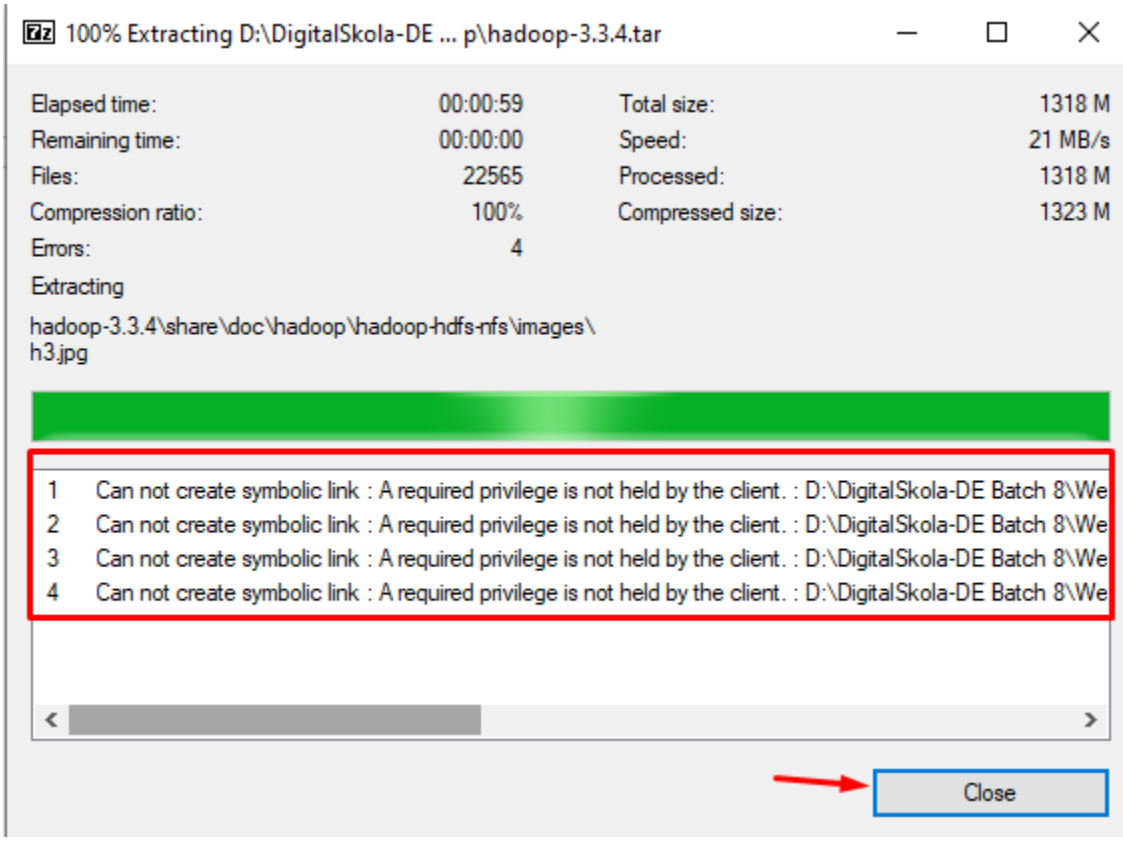


Proses Instalasi Hadoop versi 3.2.1

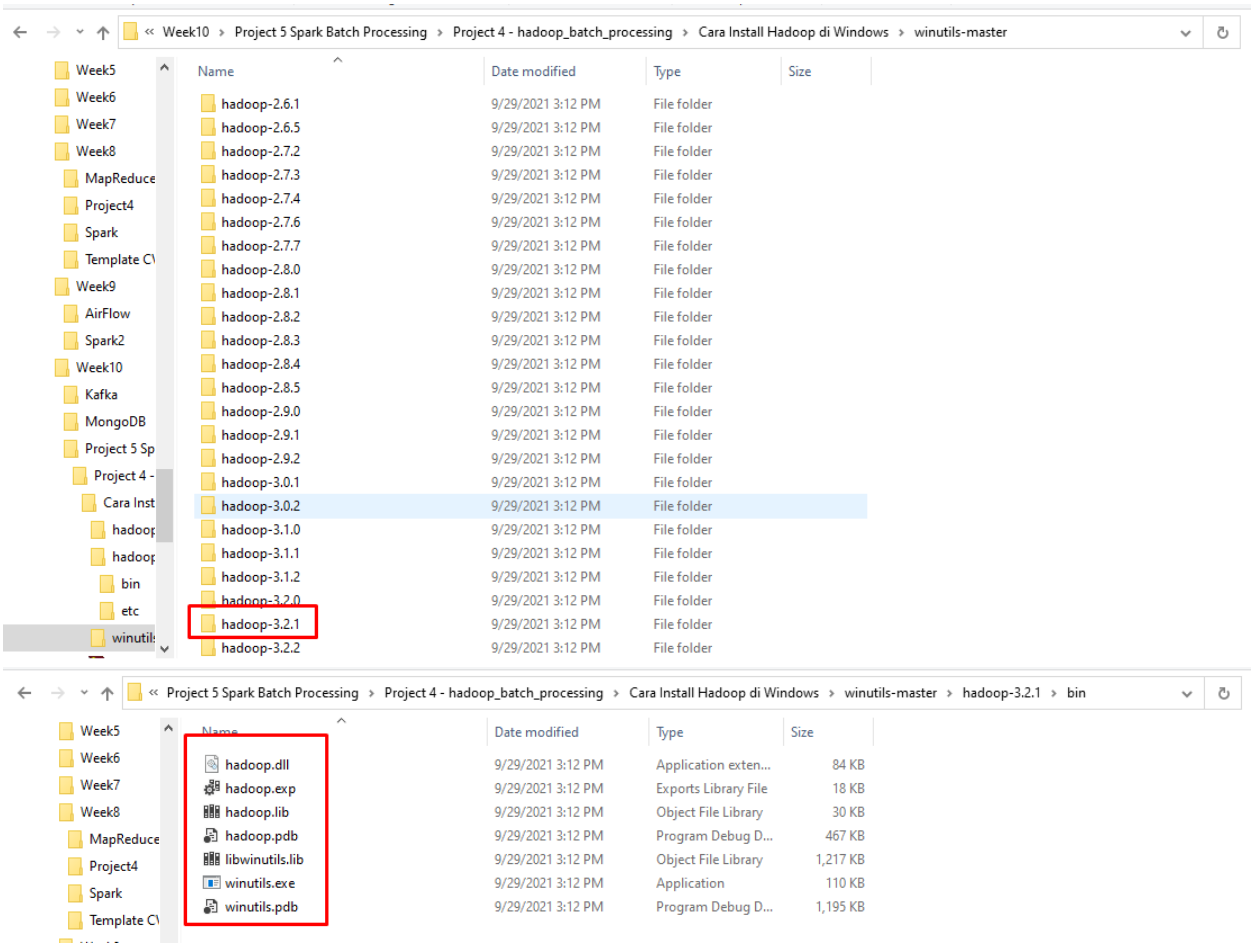
1. Download Hadoop versi 3.2.1 melalui link https://archive.apache.org/dist/hadoop/common/hadoop-3.2.1/hadoop-3.2.1.tar.gz
2. Copy file hadoop ke drive C, buat folder hadoop-env kemudian ekstrak file hadoop ke dalam folder hadoop-env tersebut. Sehingga file hasil ektraksi ada di C:\hadoop-env



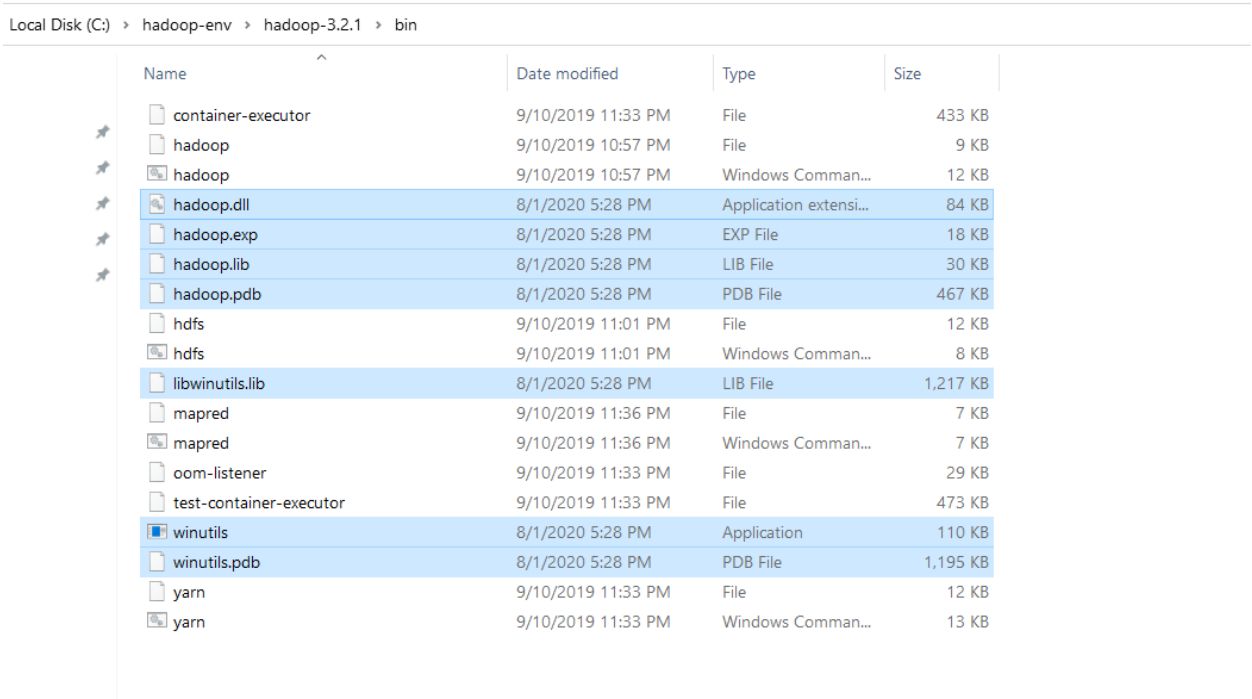
3. Ketika selesai dan menemukan peringatan seperti dibawah ini, maka klik close.



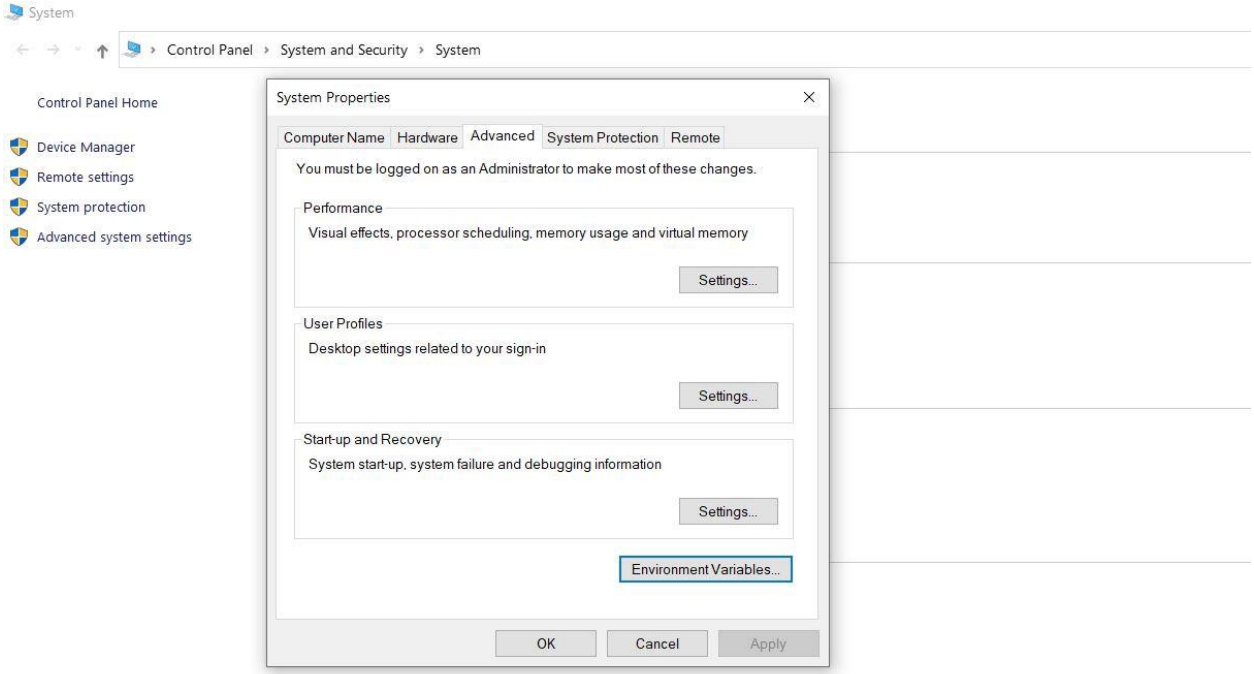
4. Agar Hadoop dapat dioperasikan pada OS Windows, diperlukan patch file Hadoop yang spesifik untuk Windows. Download file zip **bin** yang ada di link ini (https://github.com/cdarlint/winutils/archive/refs/heads/master.zip), kemudian extract, dan hasilnya seperti dalam gambar di bawah.



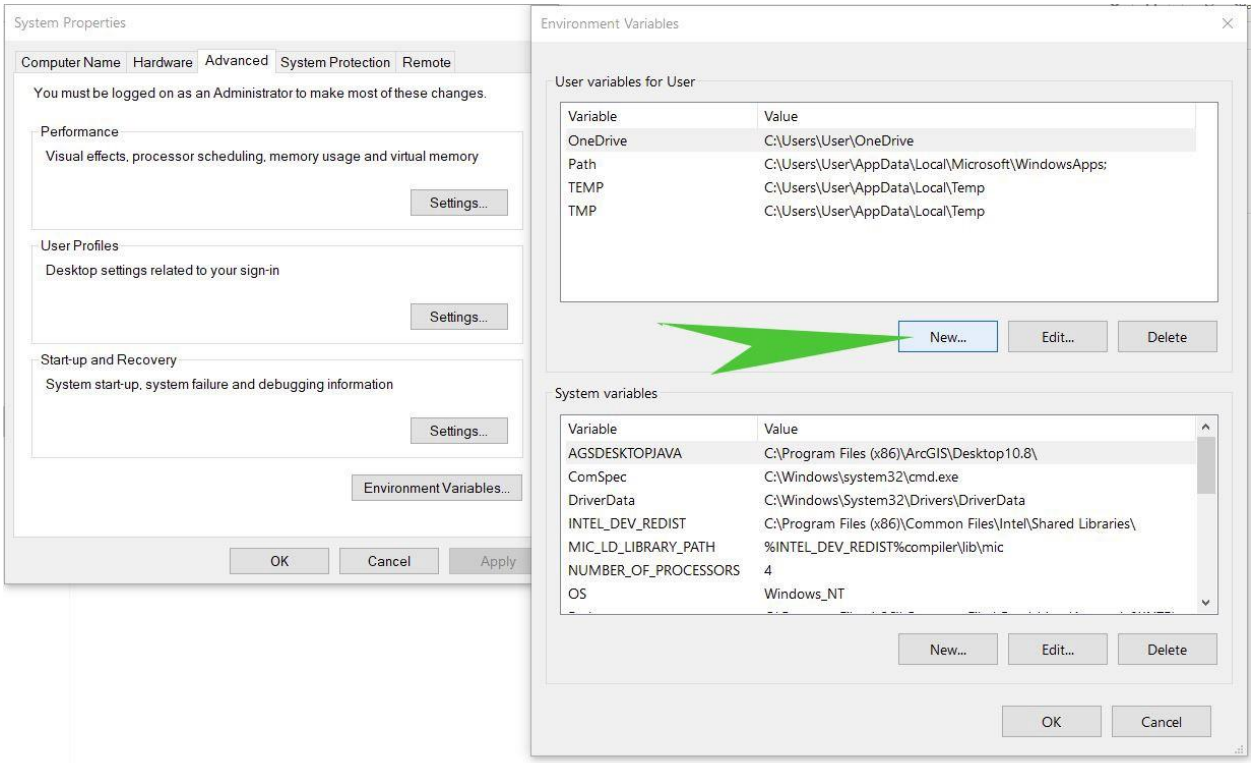
5. Letakkan semua file-file tersebut di folder C:\hadoop-env\hadoop-3.2.1\bin, seperti pada file-file yang tersorot biru pada gambar di bawah.



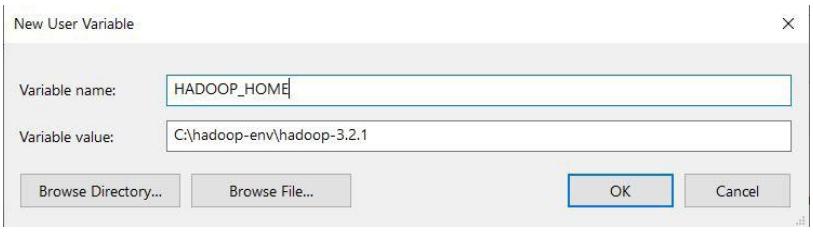
6. Setting pada environment variables Klik Control Panel > System and Security > System > Advanced system settings (lihat gambar di bawah)



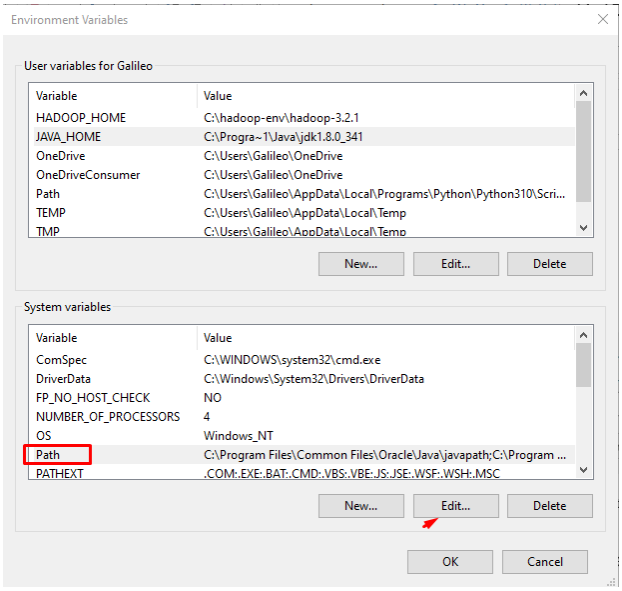
7. Lalu, klik Environment Variables. Dan lanjutkan klik New pada user varibles for User (lihat gambar di bawah)



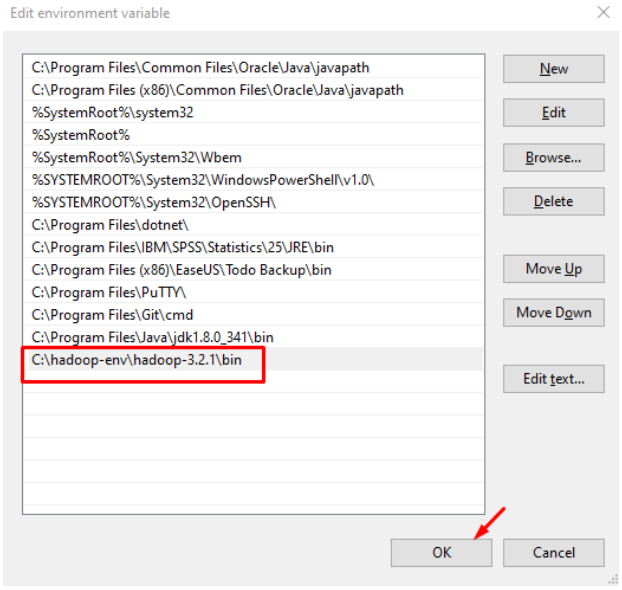
8. Setelah klik New, akan muncul popup Kemudian ketikkan HADOOP_HOME pada form Variable name, dan C:\hadoop-env\hadoop-3.2.1 pada form Variable value, seperti pada gambar di bawah. Lalu klik OK.



9. Setelah klik OK di atas, akan muncul tampilan seperti pada gambar di bawah. Pada bagian System variables, scroll sehingga muncul Path (lihat gambar di bawah), select kemudian klik Edit.



10. Klik New dan isi dengan direktori C:\hadoop-env\hadoop-3.2.1\bin. Lalu klik OK pada Edit environment variable, Environment Variables dan System Properties.



11. Sampai dengan tahap ini, semoga berjalan lancar. Lakukan testing untuk melihat apakah setting java dan hadoop di environment variables telah berhasil. Caranya: buka cmd, kemudian ketik hadoop -version. Hasilnya seperti pada gambar di bawah:

```
C:\Users\Galileo>hadoop -version
java version "1.8.0_341"
Java(TM) SE Runtime Environment (build 1.8.0_341-b10)
Java HotSpot(TM) 64-Bit Server VM (build 25.341-b10, mixed mode)
```

12. Kemudian Kita akan melakukan konfigurasi terhadap 4 file hadoop. Buka folder C:\hadoop-env\hadoop-3.2.1\etc\hadoop. Kemudian buka file hdfs-site.xml, core-site.xml, mapred-site.xml, yarn-site.xml di text editor.
13. Tambahkan code berikut pada file core-site.xml

```
<property>
<name>fs.default.name</name>
<value>hdfs://localhost:9820</value>
</property>
```

core-site.xml - Notepad

File Edit Format View Help

```
    http://www.apache.org/licenses/LICENSE-2.0

  Unless required by applicable law or agreed to in writing, software
  distributed under the License is distributed on an "AS IS" BASIS,
  WITHOUT WARRANTIES OR CONDITIONS OF ANY KIND, either express or implied.
  See the License for the specific language governing permissions and
  limitations under the License. See accompanying LICENSE file.
-->

<!-- Put site-specific property overrides in this file. -->

<configuration>

<property>
        <name>fs.default.name</name>
        <value>hdfs://localhost:9820</value>
</property>

</configuration>
```

14. Tambahkan code berikut pada file mapred-site.xml

```
<property>
<name>mapreduce.framework.name</name>
<value>yarn</value>
<description>MapReduce framework name</description>
</property>
```

mapred-site.xml - Notepad

File Edit Format View Help

```
  Unless required by applicable law or agreed to in writing, software
  distributed under the License is distributed on an "AS IS" BASIS,
  WITHOUT WARRANTIES OR CONDITIONS OF ANY KIND, either express or implied.
  See the License for the specific language governing permissions and
  limitations under the License. See accompanying LICENSE file.
-->

<!-- Put site-specific property overrides in this file. -->

<configuration>

<property>
        <name>mapreduce.framework.name</name>
        <value>yarn</value>
        <description>MapReduce framework name</description>
</property>

</configuration>
```

15. Tambahkan code berikut pada file yarn-site.xml

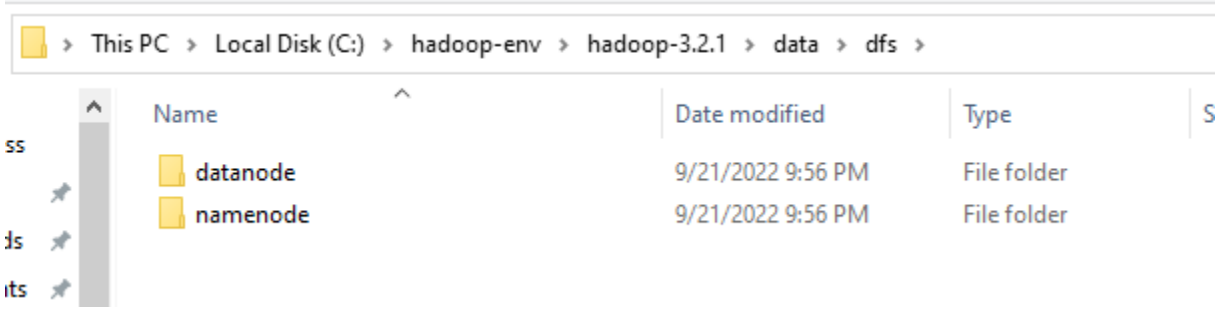
```
<property>
<name>yarn.nodemanager.aux-
services</name><value>mapreduce_shuffle</value>
<description>Yarn Node Manager Aux Service</description>
</property>
```

yarn-site.xml - Notepad

File Edit Format View Help

```
    http://www.apache.org/licenses/LICENSE-2.0

  Unless required by applicable law or agreed to in writing, software
  distributed under the License is distributed on an "AS IS" BASIS,
  WITHOUT WARRANTIES OR CONDITIONS OF ANY KIND, either express or implied.
  See the License for the specific language governing permissions and
  limitations under the License. See accompanying LICENSE file.
-->
<configuration>

<!-- Site specific YARN configuration properties -->

<property>
        <name>yarn.nodemanager.aux-services</name><value>mapreduce_shuffle</value>
        <description>Yarn Node Manager Aux Service</description>
</property>

</configuration>
```

16. Pada direktori hadoop buatlah folder baru dengan nama **data**, didalam folder **data** buat folder baru dengan nama **dfs**. Pada folder **dfs** tersebut buat 2 folder baru dengan nama **datanode** dan **namenode**.



17. Tambahkan code berikut pada file hdfs-site.xml. Untuk tag value disesuaikan dengan direktori dimana folder namenode dan datanode dibuat.

```
<property>
<name>dfs.replication</name>
<value>1</value>
</property>
<property>
<name>dfs.namenode.name.dir</name>
<value>file:///C:/hadoop-env/hadoop-3.2.1/data/dfs/namenode</value>
</property>
<property>
<name>dfs.datanode.data.dir</name>
<value>file:///C:/hadoop-env/hadoop-3.2.1/data/dfs/datanode</value>
</property>
```

hdfs-site.xml - Notepad

File Edit Format View Help

```
<!-- Put site-specific property overrides in this file. -->

<configuration>

<property>
        <name>dfs.replication</name>
        <value>1</value>
</property>
<property>
        <name>dfs.namenode.name.dir</name>
        <value>file:///C:/hadoop-env/hadoop-3.2.1/data/dfs/namenode</value>
</property>
<property>
        <name>dfs.datanode.data.dir</name>
        <value>file:///C:/hadoop-env/hadoop-3.2.1/data/dfs/datanode</value>
</property>

</configuration>
```

Ln 1, Col 1 | 100% | Unix (LF) | UTF-8

18. **[OPTIONAL]** Pada file hadoop-env.cmd, sesuaikan direktori JAVA_HOME dengan direktori java jdk.

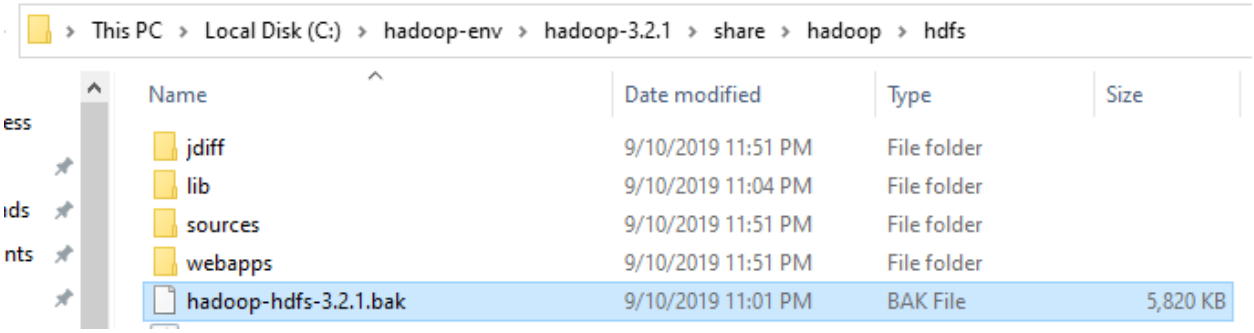
File Edit Search View Encoding Language Settings Tools Macro Run Plugins Window ?

core-site.xml | hadoop-env.cmd | hdfs-site.xml | yarn-site.xml | mapred-site.xml.template

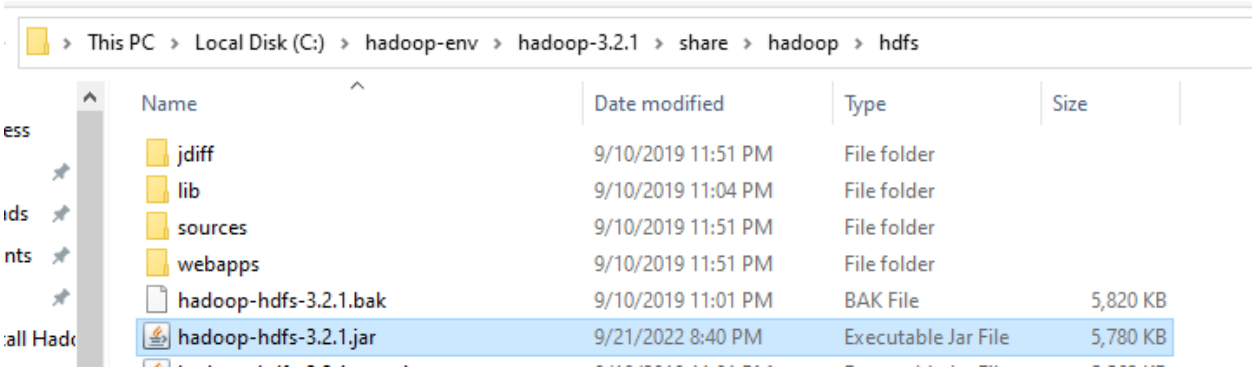
```
1  @echo off
2  @rem Licensed to the Apache Software Foundation (ASF) under one or more
3  @rem contributor license agreements.  See the NOTICE file distributed with
4  @rem this work for additional information regarding copyright ownership.
5  @rem The ASF licenses this file to You under the Apache License, Version 2.0
6  @rem (the "License"); you may not use this file except in compliance with
7  @rem the License.  You may obtain a copy of the License at
8  @rem
9  @rem     http://www.apache.org/licenses/LICENSE-2.0
10 @rem
11 @rem Unless required by applicable law or agreed to in writing, software
12 @rem distributed under the License is distributed on an "AS IS" BASIS,
13 @rem WITHOUT WARRANTIES OR CONDITIONS OF ANY KIND, either express or implied.
14 @rem See the License for the specific language governing permissions and
15 @rem limitations under the License.
16
17 @rem Set Hadoop-specific environment variables here.
18
19 @rem The only required environment variable is JAVA_HOME.  All others are
20 @rem optional.  When running a distributed configuration it is best to
21 @rem set JAVA_HOME in this file, so that it is correctly defined on
22 @rem remote nodes.
23
24 @rem The java implementation to use.  Required.
25 set JAVA_HOME=C:\Program Files\Java\jdk1.8.0_181
26
27 @rem The jsvc implementation to use. Jsvc is required to run secure datanodes.
28 @rem set JSVC_HOME=%JSVC_HOME%
29
30 @rem set HADOOP_CONF_DIR=
31
32 @rem Extra Java CLASSPATH elements.  Automatically insert capacity-scheduler.
33 if exist %HADOOP_HOME%\contrib\capacity-scheduler (
34   if not defined HADOOP_CLASSPATH (
35     set HADOOP_CLASSPATH=%HADOOP_HOME%\contrib\capacity-scheduler\*.jar
36   ) else (
37     set HADOOP_CLASSPATH=%HADOOP_CLASSPATH%;%HADOOP_HOME%\contrib\capacity-sch
```

Penulisan direktori JAVA_HOME sebenarnya tidak boleh terdapat folder yang mengandung spasi, namun apabila ingin tetap terdapat folder dengan spasi, maka ditambahkan "…." agar tidak terjadi error.

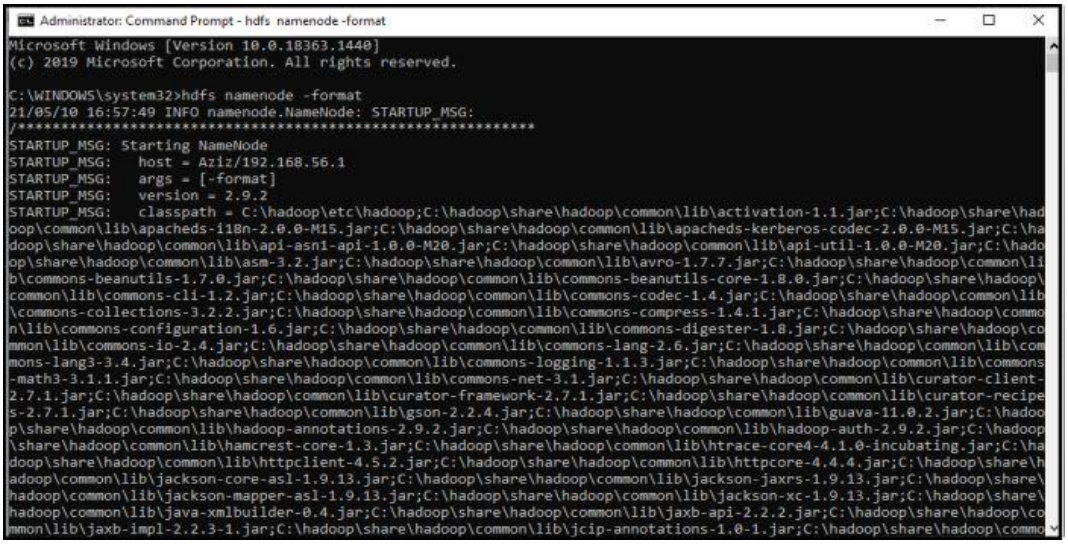
19. Melakukan format pada Name node. Di tahap ini cukup **krusial** karena jika tidak dilakukan akan ada error saat menjalankan formatting. Sebelum melakukan format, lakukan 3 tahapan ini:
    a. Download file hadoop-hdfs-3.2.1.jar pada link https://github.com/FahaoTang/big-data/raw/master/hadoop-hdfs-3.2.1.jar
    b. Buka folder C:\hadoop-env\hadoop-3.2.1\share\hadoop\hdfs kemudian Rename nama file hadoop-hdfs-3.2.1.jar menjadi hadoop-hdfs-3.2.1.bak



    c. Copy file hadoop-hdfs-3.2.1.jar yang sebelumnya telah di download pada langkah 9.a. ke dalam folder C:\hadoop-env\hadoop-3.2.1\share\hadoop\hdfs



20. Buka Command prompt dan ketikkan perintah **hdfs namenode -format**. maka lihatlah hasilnya, pada Startup_msg akan menampilkan Starting NameNode.



```
Select Command Prompt
2022-09-21 20:21:40,254 INFO util.GSet: Computing capacity for map cachedBlocks
2022-09-21 20:21:40,254 INFO util.GSet: VM type       = 64-bit
2022-09-21 20:21:40,255 INFO util.GSet: 0.25% max memory 889 MB = 2.2 MB
2022-09-21 20:21:40,256 INFO util.GSet: capacity      = 2^18 = 262144 entries
2022-09-21 20:21:40,280 INFO metrics.TopMetrics: NNTop conf: dfs.namenode.top.window.num.buckets = 10
2022-09-21 20:21:40,280 INFO metrics.TopMetrics: NNTop conf: dfs.namenode.top.num.users = 10
2022-09-21 20:21:40,282 INFO metrics.TopMetrics: NNTop conf: dfs.namenode.top.windows.minutes = 1,5,25
2022-09-21 20:21:40,292 INFO namenode.FSNamesystem: Retry cache on namenode is enabled
2022-09-21 20:21:40,293 INFO namenode.FSNamesystem: Retry cache will use 0.03 of total heap and retry cache entry expiry
 time is 600000 millis
2022-09-21 20:21:40,300 INFO util.GSet: Computing capacity for map NameNodeRetryCache
2022-09-21 20:21:40,300 INFO util.GSet: VM type       = 64-bit
2022-09-21 20:21:40,302 INFO util.GSet: 0.029999999329447746% max memory 889 MB = 273.1 KB
2022-09-21 20:21:40,302 INFO util.GSet: capacity      = 2^15 = 32768 entries
2022-09-21 20:21:40,380 INFO namenode.FSImage: Allocated new BlockPoolId: BP-513676949-192.168.88.102-1663766580359
2022-09-21 20:21:40,444 INFO common.Storage: Storage directory C:\hadoop\data\namenode has been successfully formatted.
2022-09-21 20:21:40,504 INFO namenode.FSImageFormatProtobuf: Saving image file C:\hadoop\data\namenode\current\fsimage.c
kpt_0000000000000000000 using no compression
2022-09-21 20:21:40,779 INFO namenode.FSImageFormatProtobuf: Image file C:\hadoop\data\namenode\current\fsimage.ckpt_000
0000000000000000 of size 402 bytes saved in 0 seconds .
2022-09-21 20:21:40,808 INFO namenode.NNStorageRetentionManager: Going to retain 1 images with txid >= 0
2022-09-21 20:21:40,852 INFO namenode.FSNamesystem: Stopping services started for active state
2022-09-21 20:21:40,853 INFO namenode.FSNamesystem: Stopping services started for standby state
2022-09-21 20:21:40,863 INFO namenode.FSImage: FSImageSaver clean checkpoint: txid=0 when meet shutdown.
2022-09-21 20:21:40,864 INFO namenode.NameNode: SHUTDOWN_MSG:
/************************************************************
SHUTDOWN_MSG: Shutting down NameNode at User/192.168.88.102
************************************************************/

C:\hadoop\sbin>
```

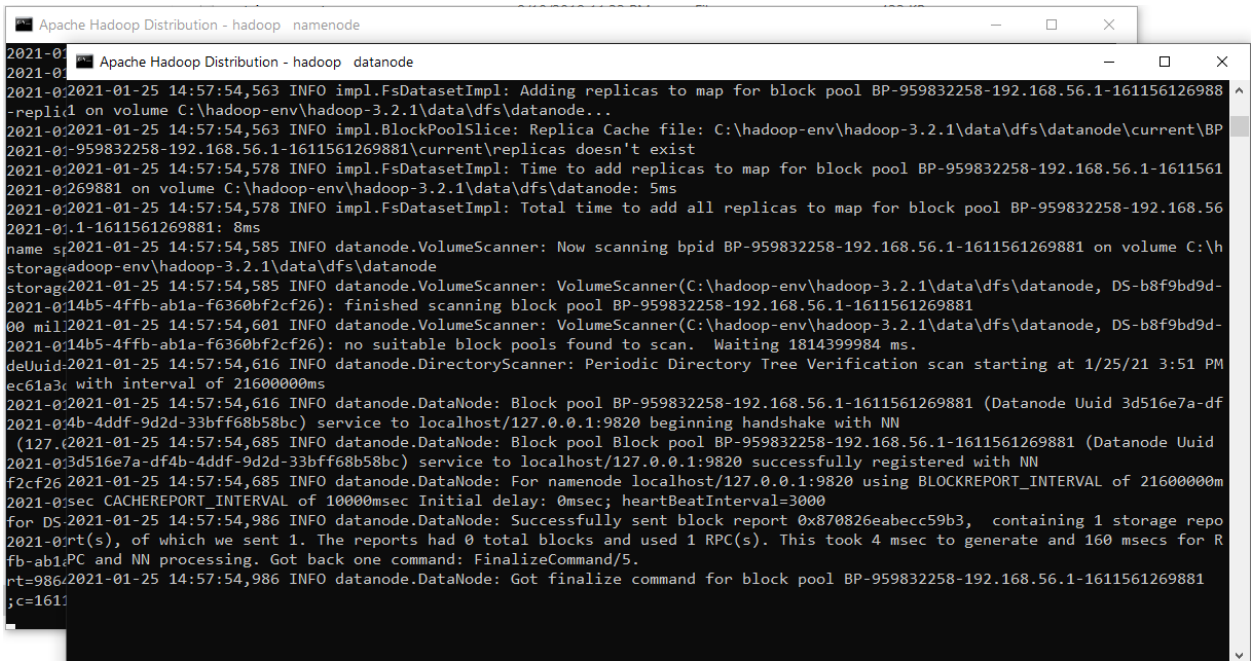
21. Lalu aktifkan Hadoop Services dengan buka kembali command prompt, dan arahkan ke C:\hadoop-env\hadoop-3.2.1\sbin.

```
Command Prompt
2022-09-21 21:55:42,711 INFO namenode.FSNamesystem: Retry cache will use 0.03 of total heap and retry cache entry expiry
 time is 600000 millis
2022-09-21 21:55:42,719 INFO util.GSet: Computing capacity for map NameNodeRetryCache
2022-09-21 21:55:42,719 INFO util.GSet: VM type       = 64-bit
2022-09-21 21:55:42,723 INFO util.GSet: 0.029999999329447746% max memory 889 MB = 273.1 KB
2022-09-21 21:55:42,723 INFO util.GSet: capacity      = 2^15 = 32768 entries
2022-09-21 21:55:42,827 INFO namenode.FSImage: Allocated new BlockPoolId: BP-1217908123-192.168.88.102-1663772142804
2022-09-21 21:55:42,898 INFO common.Storage: Storage directory C:\hadoop-env\hadoop-3.2.1\data\dfs\namenode has been suc
cessfully formatted.
2022-09-21 21:55:42,984 INFO namenode.FSImageFormatProtobuf: Saving image file C:\hadoop-env\hadoop-3.2.1\data\dfs\namen
ode\current\fsimage.ckpt_0000000000000000000 using no compression
2022-09-21 21:55:43,251 INFO namenode.FSImageFormatProtobuf: Image file C:\hadoop-env\hadoop-3.2.1\data\dfs\namenode\cur
rent\fsimage.ckpt_0000000000000000000 of size 402 bytes saved in 0 seconds .
2022-09-21 21:55:43,280 INFO namenode.NNStorageRetentionManager: Going to retain 1 images with txid >= 0
2022-09-21 21:55:43,296 INFO namenode.FSImage: FSImageSaver clean checkpoint: txid=0 when meet shutdown.
2022-09-21 21:55:43,297 INFO namenode.NameNode: SHUTDOWN_MSG:
/************************************************************
SHUTDOWN_MSG: Shutting down NameNode at User/192.168.88.102
************************************************************/

C:\Users\Galileo>cd C:\hadoop-env\hadoop-3.2.1\sbin
```

22. Jalankan command **.\start-dfs.cmd**.

```
C:\hadoop-env\hadoop-3.2.1\sbin>.\start-dfs.cmd
```

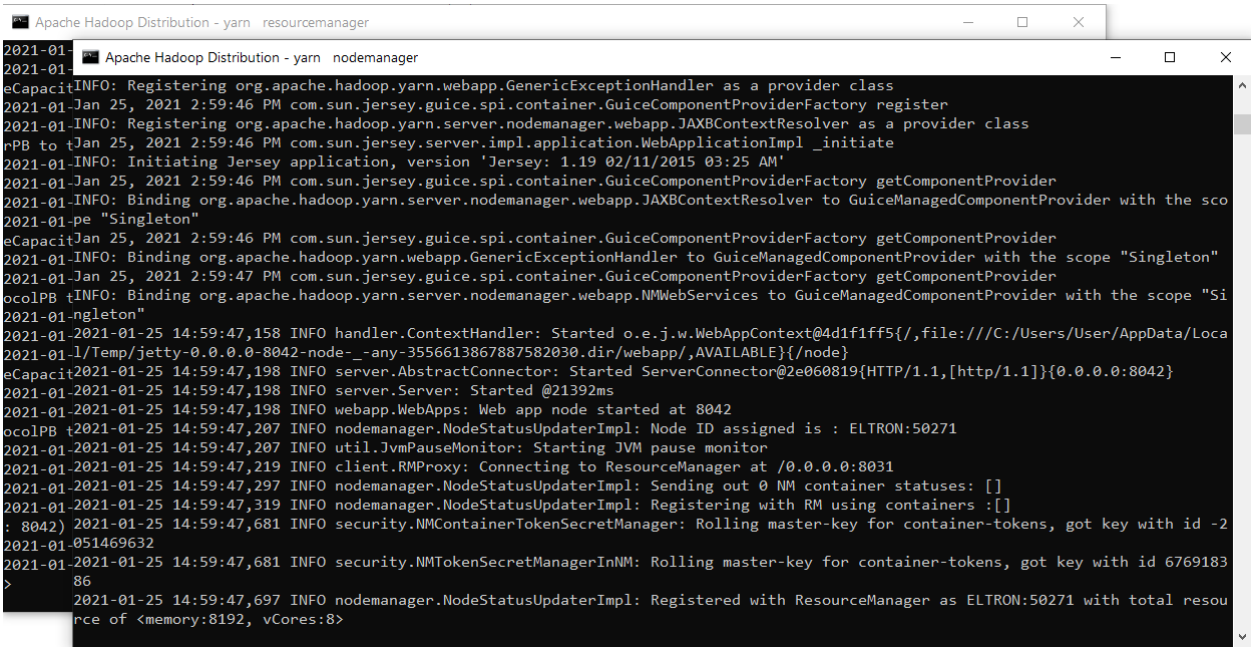
Dan muncullah seperti pada gambar di bawah, yang menandakan service Hadoop dfs telah aktif.



23. Mengaktifkan service Hadoop yarn. Ketik instruksi **.\start-yarn.cmd**, enter:

```
C:\hadoop-env\hadoop-3.2.1\sbin>.\start-yarn.cmd
starting yarn daemons
```

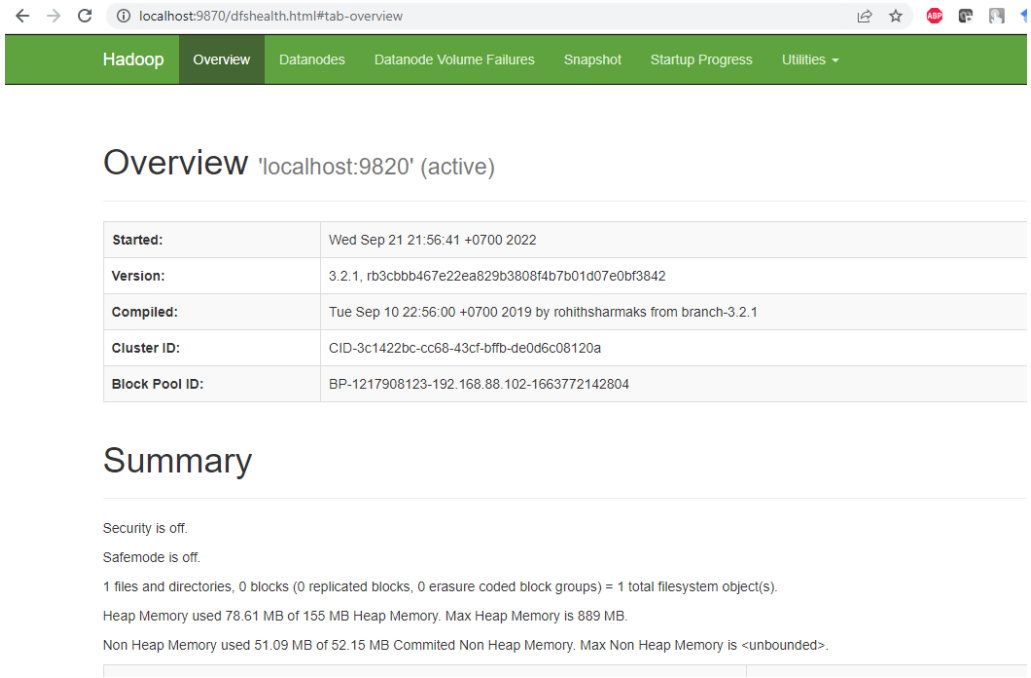
Dan muncullah seperti pada gambar di bawah, yang menandakan service Hadoop yarn telah aktif.



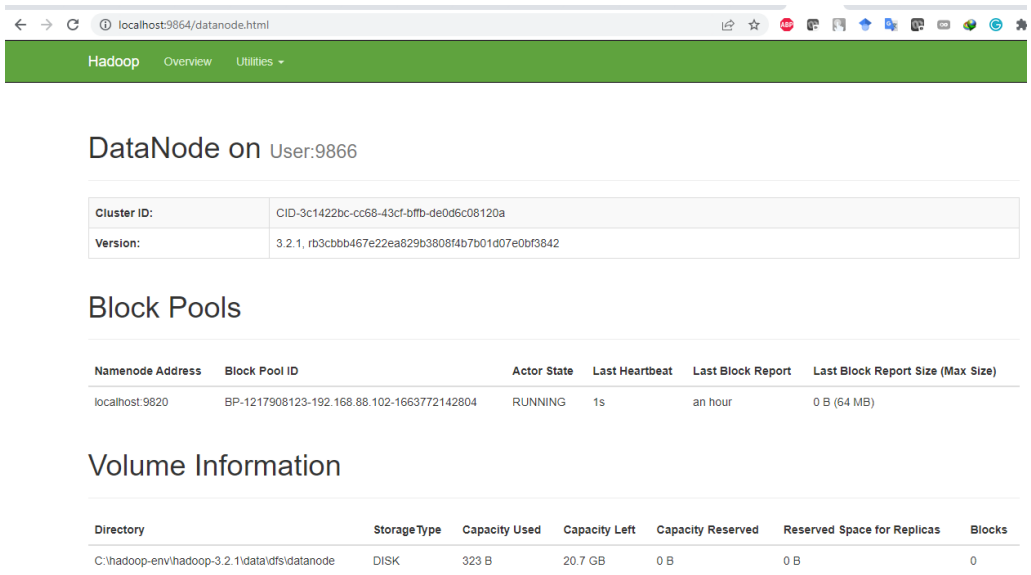
24. Memastikan semua services telah aktif, ketikkan perintah **jps**

```
C:\hadoop-env\hadoop-3.2.1\sbin>jps
15616 NameNode
4992 Jps
14500 ResourceManager
10040 DataNode
5404 NodeManager
```

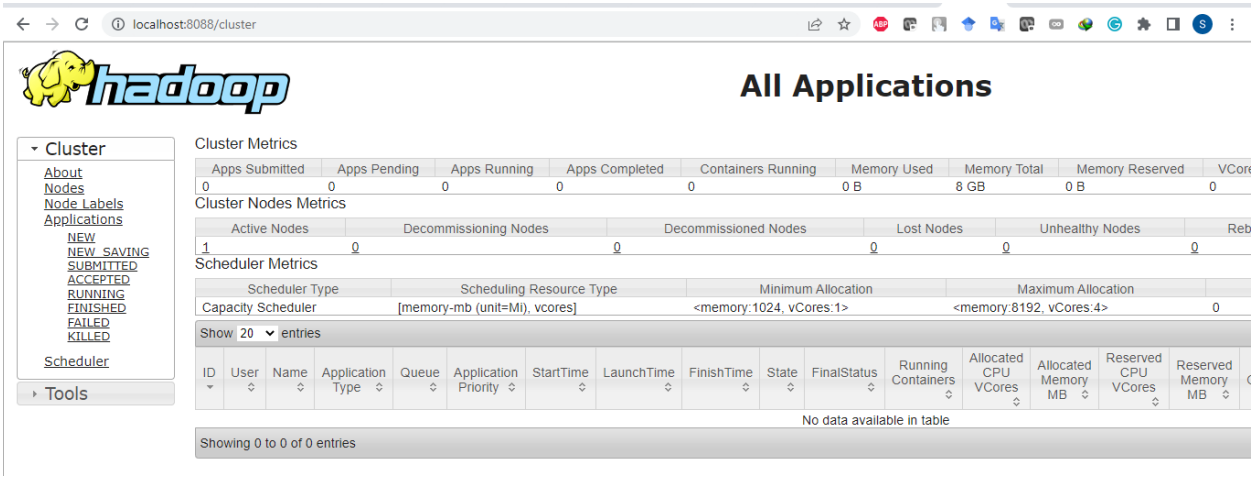
25. Membuka Hadoop Web UI dengan cara jalankan browser kemudian ketikan alamat URL Name node: http://localhost:9870/dfshealth.html



Data node: http://localhost:9864/datanode.html

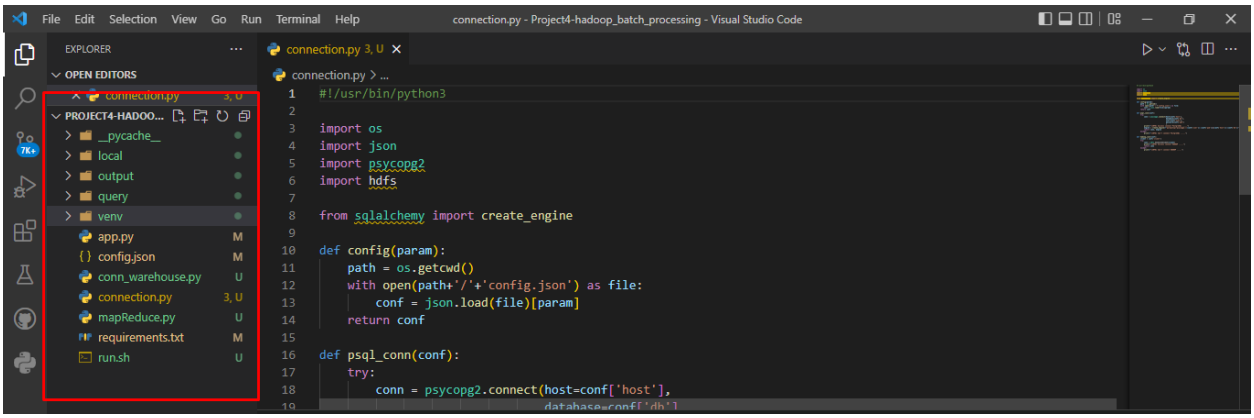


Yarn: http://localhost:8088/cluster

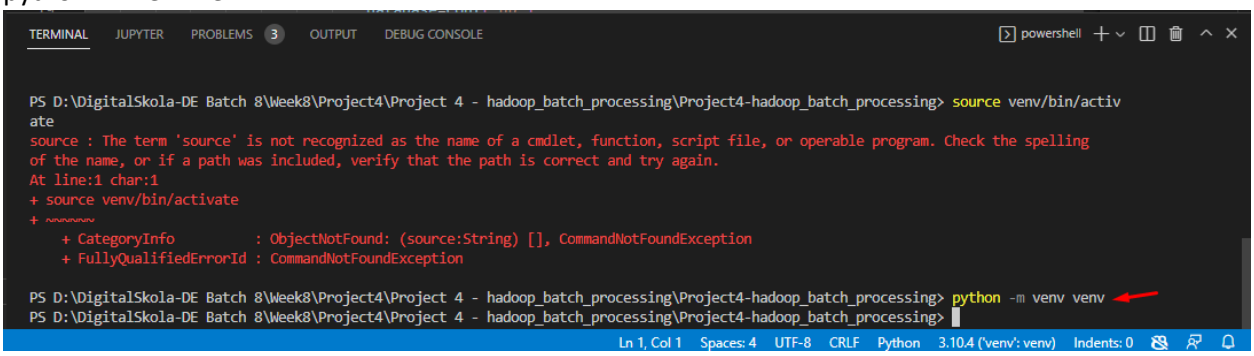


Data Analytic dengan Hadoop

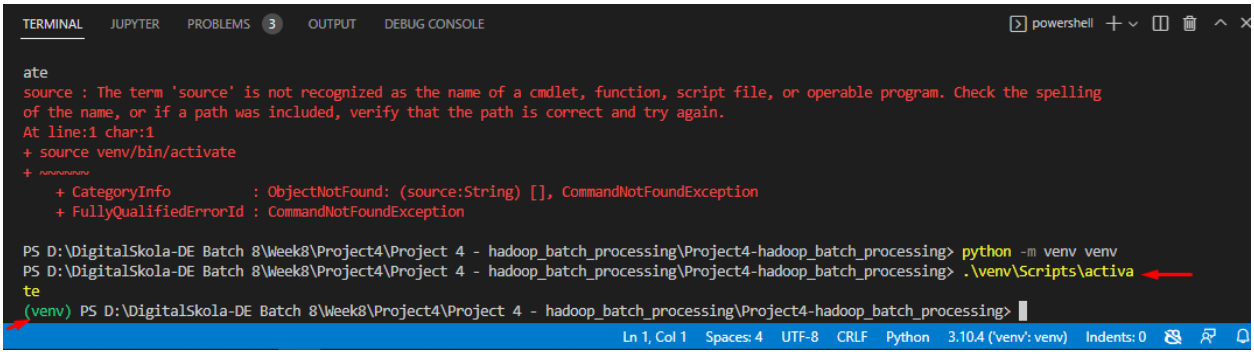
1. Buka visual studio code. Open folder dan arah ke folder Project4-hadoop_batch_processing



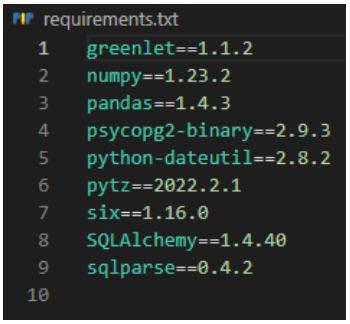
2. Pada Visual Studio Code aktifkan terminal kemudian buat virtual environment dengan perintah python -m venv venv



3. Aktifkan virtual environment tersebut dengan perintah .\venv\Scripts\activate
   Pastikan terdapat tulisan (venv) di sebelah kiri



4. Install library yang dibutuhkan sesuai list pada file requirements.txt



dengan perintah pip install -r .\requirements.txt

```
TERMINAL   JUPYTER   PROBLEMS 3   OUTPUT   DEBUG CONSOLE                                   pip

(venv) PS D:\DigitalSkola-DE Batch 8\Week8\Project4\Project 4 - hadoop_batch_processing\Project4-hadoop_batch_processing> pip install -r
 .\requirements.txt
Collecting greenlet==1.1.2
  Using cached greenlet-1.1.2-cp310-cp310-win_amd64.whl (101 kB)
Collecting numpy==1.23.2
  Using cached numpy-1.23.2-cp310-cp310-win_amd64.whl (14.6 MB)
Collecting pandas==1.4.3
  Using cached pandas-1.4.3-cp310-cp310-win_amd64.whl (10.5 MB)
Collecting psycopg2-binary==2.9.3
  Using cached psycopg2_binary-2.9.3-cp310-cp310-win_amd64.whl (1.2 MB)
Collecting python-dateutil==2.8.2
  Using cached python_dateutil-2.8.2-py2.py3-none-any.whl (247 kB)
Ln 1, Col 1   Spaces: 4   UTF-8   CRLF   Python   3.10.4 ('venv': venv)   Indents: 0
```

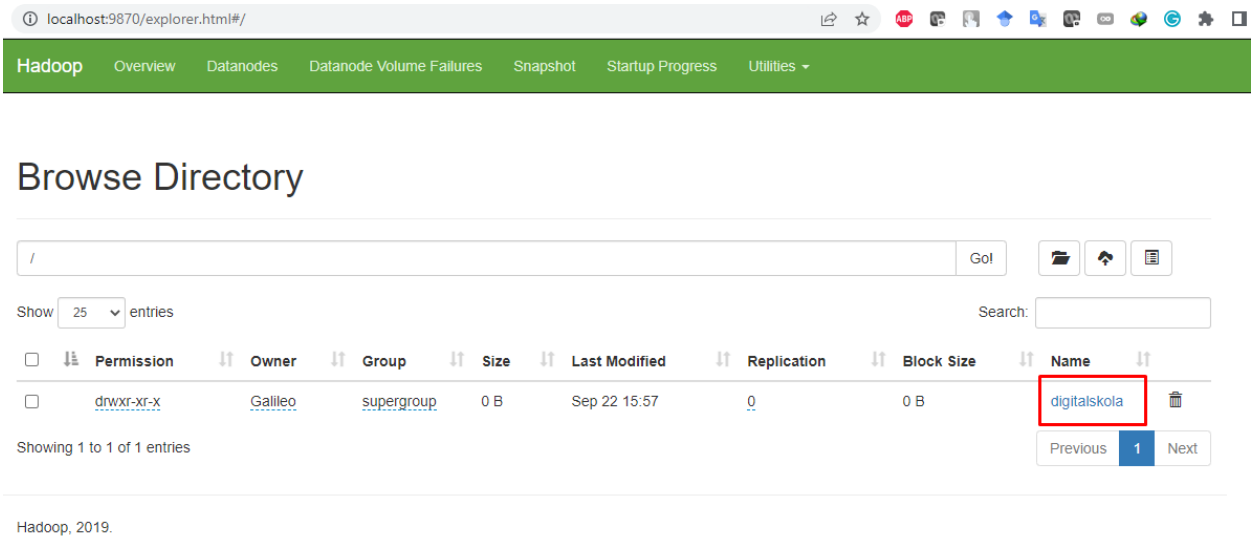
Tunggu sampai proses instalasi library selesai

```
TERMINAL   JUPYTER   PROBLEMS 1   OUTPUT   DEBUG CONSOLE                                   powershell

Collecting six==1.16.0
  Using cached six-1.16.0-py2.py3-none-any.whl (11 kB)
Collecting SQLAlchemy==1.4.40
  Using cached SQLAlchemy-1.4.40-cp310-cp310-win_amd64.whl (1.6 MB)
Collecting sqlparse==0.4.2
  Using cached sqlparse-0.4.2-py3-none-any.whl (42 kB)
Installing collected packages: pytz, sqlparse, six, psycopg2-binary, numpy, greenlet, SQLAlchemy, python-dateutil, pandas
Successfully installed SQLAlchemy-1.4.40 greenlet-1.1.2 numpy-1.23.2 pandas-1.4.3 psycopg2-binary-2.9.3 python-dateutil-2.8.2 pytz-2022.
2.1 six-1.16.0 sqlparse-0.4.2
WARNING: You are using pip version 22.0.4; however, version 22.2.2 is available.
You should consider upgrading via the 'D:\DigitalSkola-DE Batch 8\Week8\Project4\Project 4 - hadoop_batch_processing\Project4-hadoop_bat
ch_processing\venv\Scripts\python.exe -m pip install --upgrade pip' command.
(venv) PS D:\DigitalSkola-DE Batch 8\Week8\Project4\Project 4 - hadoop_batch_processing\Project4-hadoop_batch_processing>
```

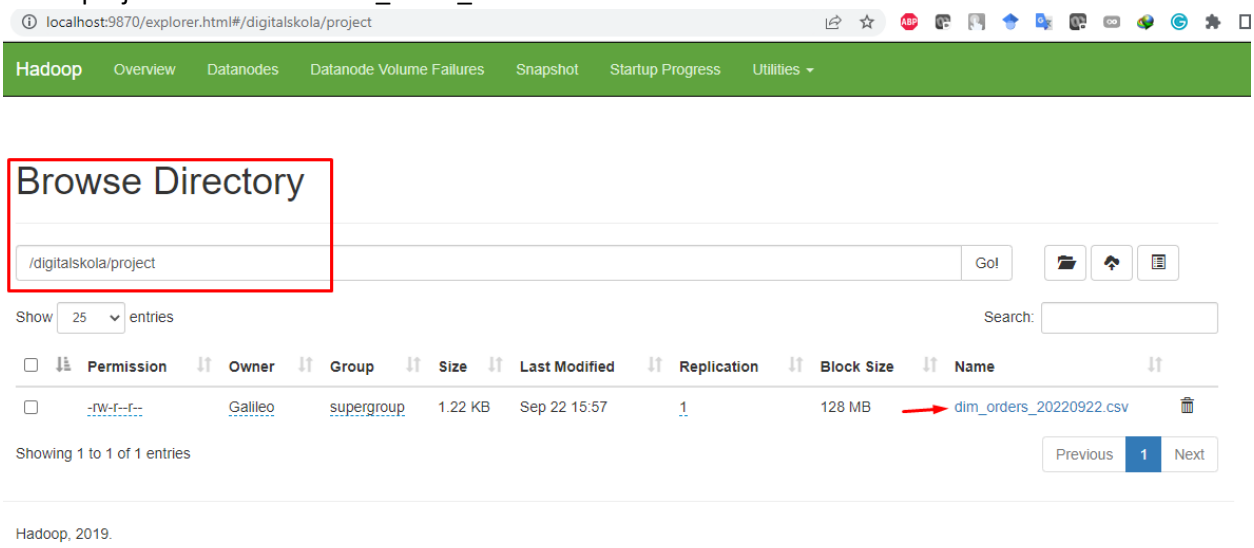
5. Jalankan program proses python hadoop batch processing dengan perintah **python app.py**

```
(venv) PS D:\DigitalSkola-DE Batch 8\Week8\Project4\Project 4 - hadoop_batch_processing\Project4-hadoop_batch_processing> python .\app.py
[INFO] Service ETL is Starting .....
[INFO] Success connect Warehouse .....
[INFO] Success connect PostgreSQL .....
[INFO] Success connect HADOOP .....
[INFO] Service ETL is Running .....
[INFO] Upload Data in HADOOP Success .....
[INFO] Upload Data in LOCAL Success .....
[INFO] Update WDH Success .....
[INFO] Service ETL is Success .....
(venv) PS D:\DigitalSkola-DE Batch 8\Week8\Project4\Project 4 - hadoop_batch_processing\Project4-hadoop_batch_processing>
```

6. Bila program berjalan dengan baik tanpa error, maka di hadoop akan muncul folder bernama digitalskola



7. Bila di-explore isi dari folder digitalskola, maka akan ada folder bernama project dan di dalam folder project akan ada file dim_ordes_20220922.csv



8. Setelahnya jalankan program python mapreduce untuk mengolah data dim_orders_20220922 yang telah ada di hadoop dengan perintah **python .\mapReduce.py -r hadoop hdfs:///digitalskola/project/dim_orders_20220922.csv**

   Dan tunggu sampai proses mapreduce selesai running

```
(venv) PS D:\DigitalSkola-DE Batch 8\Week8\Project4\Project 4 - hadoop_batch_processing\Project4-hadoop_batch_processing> python .\mapReduce.py -r ha
doop hdfs:///digitalskola/project/dim_orders_20220922.csv
No configs found; falling back on auto-configuration
No configs specified for hadoop runner
Looking for hadoop binary in C:\hadoop-env\hadoop-3.2.1\bin...
Found hadoop binary: C:\hadoop-env\hadoop-3.2.1\bin\hadoop.CMD
Using Hadoop version 3.2.1
Looking for Hadoop streaming jar in C:\hadoop-env\hadoop-3.2.1...
Looking for Hadoop streaming jar in D:\DigitalSkola-DE Batch 8\Week8\Project4\Project 4 - hadoop_batch_processing\Project4-hadoop_batch_processing...
Looking for Hadoop streaming jar in /home/hadoop/contrib...
Looking for Hadoop streaming jar in /usr/lib/hadoop-mapreduce...
Hadoop streaming jar not found. Use --hadoop-streaming-jar
Creating temp directory C:\Users\Galileo\AppData\Local\Temp\mapReduce.Galileo.20220926.142237.165822
uploading working dir files to hdfs:///user/Galileo/tmp/mrjob/mapReduce.Galileo.20220926.142237.165822/files/wd...
Copying other local files to hdfs:///user/Galileo/tmp/mrjob/mapReduce.Galileo.20220926.142237.165822/files/
```

9. Hasil output mapreduce yang dihasilkan adalah untuk total pendapatan (counting) transaksi setiap bulannya di tahun 2022.

A	B	C	D	E	F	G	H	I	J	K
order_id	order_date	user_id	payment_name	shipper_name	order_price	order_discount	voucher_name	voucher_price	order_total	rating_status
1110001	1/20/2022	100101	Debit	JNE Express	250000	15000	New User	5000	230000	Medium Impact
1110002	1/29/2022	100102	Debit	JNE Express	620000	40000	New User	5000	575000	Medium Impact
1110003	2/13/2022	100103	Credit	JNE Express	6000000	1000000	New User	5000	4995000	Very Low Impact
1110004	3/6/2022	100102	Wallet	JNE Express	3150000	45000			3105000	High Impact
1110005	4/28/2022	100105	Debit	Sicepat Express	4000000	1000000	New User	5000	2995000	Very Low Impact
1110006	5/9/2022	100103	Debit	Sicepat Express	4500000	1030000			3470000	Medium High Impact
1110007	5/21/2022	100106	Debit	JNE Express	870000	25000			845000	High Impact
1110008	6/2/2022	100108	Credit	Sicepat Express	2000000	0	New User	5000	1995000	Medium High Impact
1110009	6/23/2022	100103	Credit	Lazada Express	2000000	0			2000000	High Impact
1110010	7/1/2022	100102	Credit	Lazada Express	1050000	45000			1005000	Low Impact
1110011	7/21/2022	100110	Wallet	Sicepat Express	550000	15000			535000	High Impact
1110012	7/30/2022	100110	Debit	JNE Express	490000	35000	Body Soap Promo	10000	445000	Medium High Impact

Data Source

dim_orders_20220922

ordercount_output_local_map.txt - Notepad

File Edit Format View Help

```
"2022-01"	2
"2022-02"	1
"2022-03"	1
"2022-04"	1
"2022-05"	2
"2022-06"	2
"2022-07"	3
```

Hasil