QUIZ 2
RESOURCES FOR TRENDING VIDEO

Arranged By:

Rajendra Rakha Arya Prabaswara

(1941720080/20)

PROGRAM STUDI D-IV TEKNIK INFORMATIKA
JURUSAN TEKNOLOGI INFORMASI
POLITEKNIK NEGERI MALANG

1. Move CAvideos.csv to server using scp

```
PS G:\Kuis2_BigData> scp D:\CAvideos.csv hadoopuser@192.168.0.10:Desktop
hadoopuser@192.168.0.10's password:
CAvideos.csv                                                   100%   61MB  99.7MB/s   00:00
PS G:\Kuis2_BigData>
```
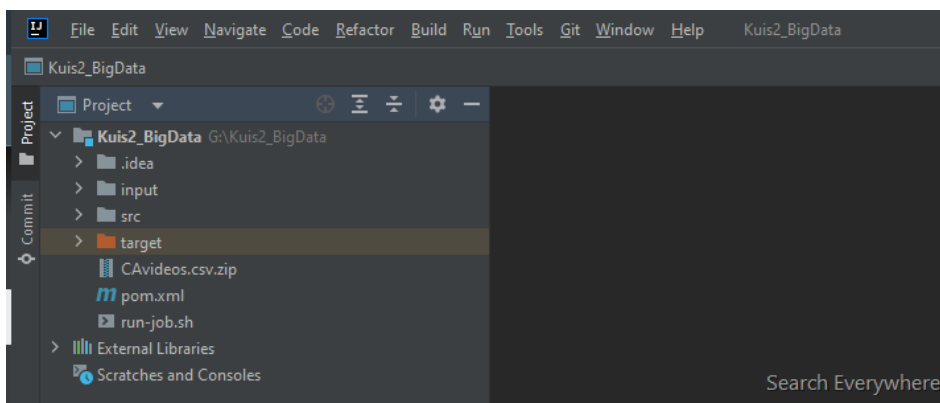
2. Make Input Folder For Hadoop

```
hadoopuser@hadoop-dataname:~/Desktop$ hadoop fs -mkdir -p / hadoopuser/input
hadoopuser@hadoop-dataname:~/Desktop$ hadoop fs -ls / hadoopuser
Found 4 items
drwxr-xr-x   - hadoopuser supergroup          0 2022-04-27 05:27 /hadoopuser/Input
drwxr-xr-x   - hadoopuser supergroup          0 2022-04-27 09:40 /hadoopuser/Ouput
drwxr-xr-x   - hadoopuser supergroup          0 2022-04-27 07:12 /hadoopuser/Output
drwxr-xr-x   - hadoopuser supergroup          0 2022-05-18 04:09 /hadoopuser/input
```
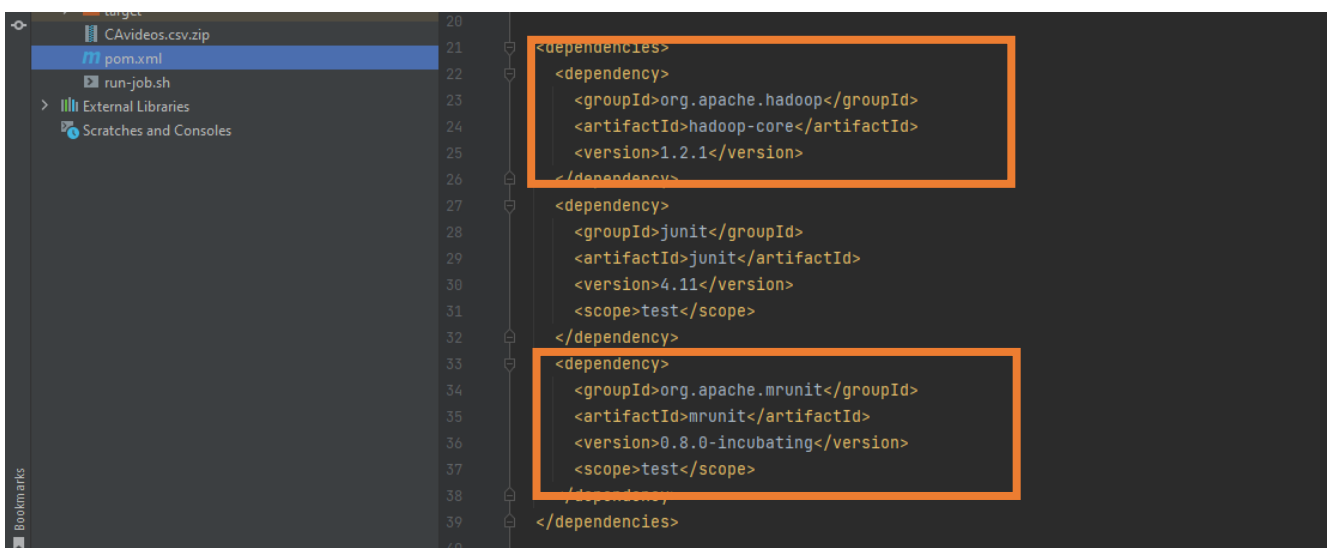
3. Move dataset to /hadoopuser/input

```
hadoopuser@hadoop-dataname:~/Desktop$ hadoop fs -put CAvideos.csv /hadoopuser/input
hadoopuser@hadoop-dataname:~/Desktop$ hadoop fs -ls /hadoopuser/input
Found 1 items
-rw-r--r--   2 hadoopuser supergroup   64067991 2022-05-18 04:23 /hadoopuser/input/CAvideos.csv
```

4. Make A new Project in Intellij Idea

5. Add Dependency to Pom.xml

6. Code For Mapper & Reducer

```java
package org.example;

import
org.apache.hadoop.io.LongWritable;
import org.apache.hadoop.io.Text;
import
org.apache.hadoop.mapreduce.Mapper
;

import java.io.IOException;
import java.util.ArrayList;
import java.util.List;
import java.util.Scanner;

import static
java.lang.String.format;

public final class MapStatistics
extends Mapper<LongWritable, Text,
Text, Text>
{
    private static final String
COMMA_DELIMITER = ",";
    public static final int
CHANNEL_ID_INDEX = 3;
    public static final int
VIEW_INDEX = 7;
    public static final int
LIKES_INDEX = 8;
    public static final int
DISLIKES_INDEX = 9;
    public static final int
COMMENT_COUNT_INDEX = 10;
    private Text
channelIdViewsLikesDislikesComment
Count = new Text();

    public void map(LongWritable
key, Text value, Context context)
throws IOException,
InterruptedException
    {
        List<String>
csvRecordsOfLine =
getRecordsFromLine(value.toString(
));
        try
        {
            String channelId =
csvRecordsOfLine.get(CHANNEL_ID_IN
DEX);
            String views =
csvRecordsOfLine.get(VIEW_INDEX);
            String likes =
csvRecordsOfLine.get(LIKES_INDEX);
            String dislikes =
csvRecordsOfLine.get(DISLIKES_INDE
X);
            String commentCount =
csvRecordsOfLine.get(COMMENT_COUNT
_INDEX);

channelIdViewsLikesDislikesComment
Count.set(format("%s-%s-%s-%s",
views, likes, dislikes,
commentCount));
            context.write(new
Text(channelId),
channelIdViewsLikesDislikesComment
Count);
        }
        catch
(IndexOutOfBoundsException ex)
        {
            // Do nothing
        }
    }

    private List<String>
getRecordsFromLine(String line)
    {
        List<String> values = new
ArrayList<>();
        try (Scanner rowScanner =
new Scanner(line))
        {

rowScanner.useDelimiter(COMMA_DELI
MITER);
            while
(rowScanner.hasNext())
            {

values.add(rowScanner.next());
            }
        }
        return values;
    }
}
```

```java
package org.example;

import org.apache.hadoop.io.NullWritable;
import org.apache.hadoop.io.Text;
import org.apache.hadoop.mapreduce.Reducer;

import java.io.IOException;

import static java.lang.String.format;

public final class ReduceStatistics extends Reducer<Text, Text, Text, NullWritable>
{
    private static final String HYPHEN_DELIMITER = "-";
    public static final int VIEW_INDEX = 0;
    public static final int LIKES_INDEX = 1;
    public static final int DISLIKES_INDEX = 2;
    public static final int COMMENT_COUNT_INDEX = 3;

    public void reduce(Text key, Iterable<Text> values, Context context) throws IOException, InterruptedException
    {
        long viewsSum = 0L;
        long likesSum = 0L;
        long dislikesSum = 0L;
        long commentSum = 0L;
        for (Text x : values)
        {
            String value = x.toString();
            String[] splitScores = value.split(HYPHEN_DELIMITER);
            try
            {
                long viewEntry = Long.parseLong(splitScores[VIEW_INDEX]);
                long likesEntry = Long.parseLong(splitScores[LIKES_INDEX]);
                long dislikesEntry = Long.parseLong(splitScores[DISLIKES_INDEX]);
                long commentCountEntry = Long.parseLong(splitScores[COMMENT_COUNT_INDEX]);
                viewsSum += viewEntry;
                likesSum += likesEntry;
                dislikesSum += dislikesEntry;
                commentSum += commentCountEntry;
            }
            catch (NumberFormatException ex)
            {
                // Do nothing
            }
        }
        context.write(new Text(format("%s,%s,%s,%s,%s",key, viewsSum, likesSum, dislikesSum, commentSum)), null);
    }
}
```

7. Code App.java

```java
package org.example;

import org.example.MapStatistics;
import
org.example.ReduceStatistics;
import org.apache.hadoop.fs.Path;
import
org.apache.hadoop.io.NullWritable;
import org.apache.hadoop.io.Text;
import
org.apache.hadoop.mapreduce.Job;
import
org.apache.hadoop.mapreduce.lib.in
put.FileInputFormat;
import
org.apache.hadoop.mapreduce.lib.ou
tput.FileOutputFormat;

public class App
{
    public static void
main(String[] args) throws
Exception
    {
        Job job =
Job.getInstance();


job.setJarByClass(App.class);

        job.setJobName("Trending
Statistics");


FileInputFormat.setInputPaths(job,
new Path(args[0]));
```

```java
FileOutputFormat.setOutputPath(job
, new Path(args[1]));


job.setMapperClass(MapStatistics.c
lass);

job.setReducerClass(ReduceStatisti
cs.class);


job.setMapOutputKeyClass(Text.clas
s);

job.setMapOutputValueClass(Text.cl
ass);


job.setOutputKeyClass(Text.class);

job.setOutputValueClass(NullWritab
le.class);

        boolean success =
job.waitForCompletion(true);
        System.exit(success ? 0 :
1);
    }

}
```

8. Code file Run-job.sh

```bash
# Variabel-variabel

FILE_JAR=target/kuis2-bigdata-1.0-
SNAPSHOT.jar

HADOOP_USERNAME=hadoopuser

IP_NAMENODE=192.168.0.10

HOME_DIR=/home/hadoopuser/

NAMA_JAR_TUJUAN=kuis2.jar

PACKAGE_ID=org.example.App
```

```bash
INPUT_FOLDER=/hadoopuser/input

OUTPUT_FOLDER=/hadoopuser/output-
kuis


# Bersihkan console

clear


# Jalankan SCP
```

```
SCP_ARG="${HADOOP_USERNAME}@${IP_N
AMENODE}:${HOME_DIR}${NAMA_JAR_TUJ
UAN}"

echo "Running SCP.."

echo "${SCP_ARG}"

scp $FILE_JAR $SCP_ARG



# SSH ke Name Node dan jalankan
MapReduce job

echo "Connecting to Namenode and
execute MapReduce Job.."
```

```
HADOOP_JAR_COMMAND="hadoop jar
${NAMA_JAR_TUJUAN} ${PACKAGE_ID}
${INPUT_FOLDER} ${OUTPUT_FOLDER}"

LS_OUTPUT_COMMAND="hadoop fs -ls
${OUTPUT_FOLDER}"

CAT_OUTPUT_COMMAND="hadoop fs -cat
${OUTPUT_FOLDER}/part-r-00000"

DELETE_OUTPUT_COMMAND="hadoop fs -
rm -r ${OUTPUT_FOLDER}"

ssh
"${HADOOP_USERNAME}@${IP_NAMENODE}
" "${HADOOP_JAR_COMMAND};
${LS_OUTPUT_COMMAND};
${CAT_OUTPUT_COMMAND};
${DELETE_OUTPUT_COMMAND}; exit"

echo "Selesai."
```
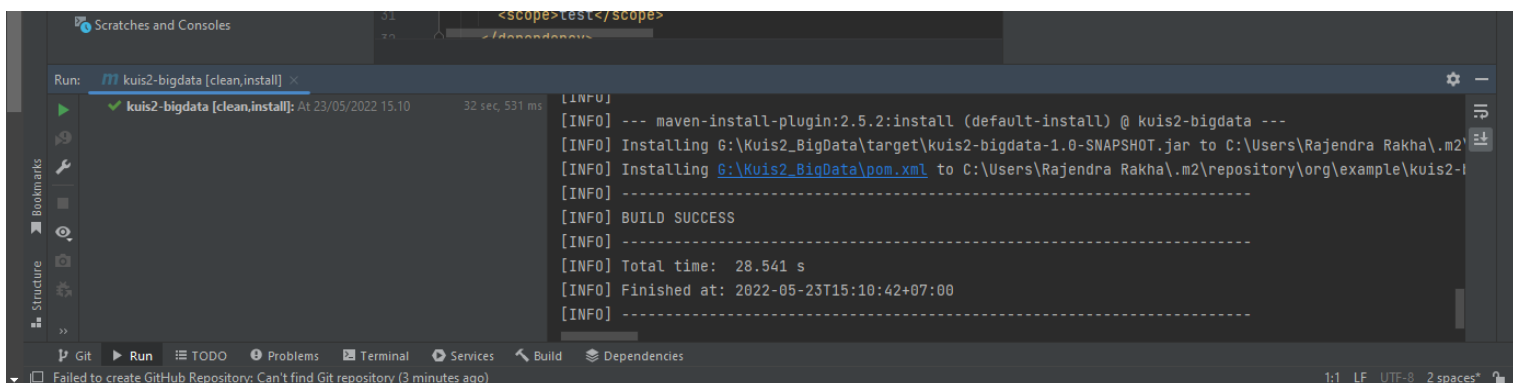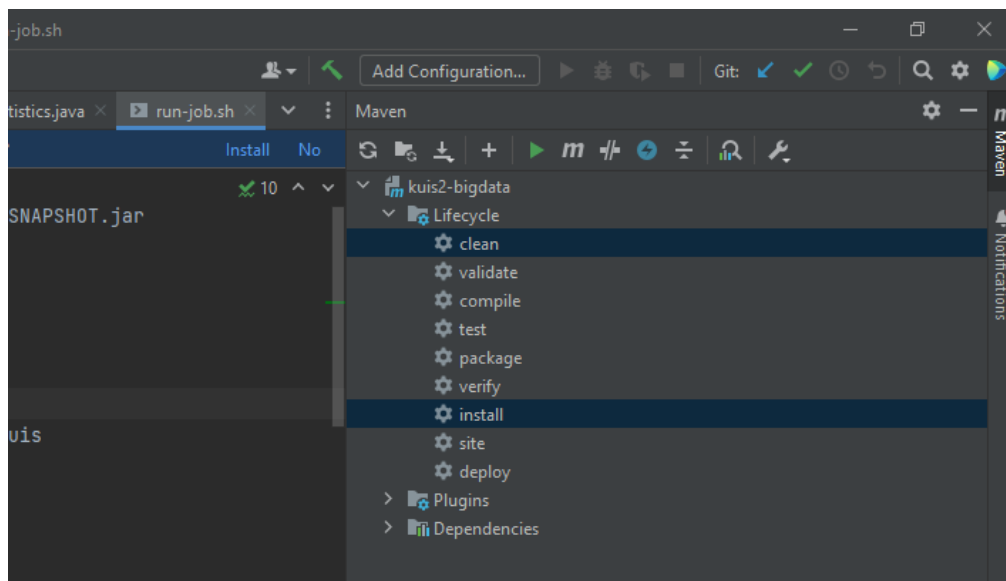
9. Select Clean and Install Then Run

10. After Succeed Compile , execute **run-job.sh**

```
Terminal:  Local ×  +  ∨

Windows PowerShell
Copyright (C) Microsoft Corporation. All rights reserved.

Install the latest PowerShell for new features and improvements! https://aka.ms/PSWindows

PS G:\Kuis2_BigData> sh "G:\Kuis2_BigData\run-job.sh"
```

11. The output results display CHANNEL NAME, NUMBER OF VIEWS, LIKE, DISLIKE, COMMENT

```
Terminal:  run-job.sh ×   run-job.sh ×  +  ∨
"Pero Like",383120,18338,166,2634
"Peter Brown",1465900,79612,810,14877
"Peter McKinnon",1748999,133578,1386,14919
"Peter Wong",120275,1233,99,67
"Petscop",321899,25841,132,11752
"PewDiePie",427916294,31436442,820325,3585585
"PhantomStrider",160110,9516,139,2705
"Phil Jones",2378,178,6,12
"PhilCrownVEVO",2206751,25611,4249,1491
"Philip DeFranco",22445950,1407841,55306,224963
"Phillip Rogers",877113,6759,676,1810
"Philly D",6826574,546375,6800,39135
"Phoenix Pictures International",1115328,7069,740,909
"Phú Quý Jr",33581,14,10,4
"Phạm Phạm",753055,4087,318,227
"Pick Up Limes",253621,25262,161,1870
"Pics Tube",3295146,19972,4831,3315
"Piemations",1776637,200105,1504,24844
```

Git   ▶ Run   ☰ TODO   ⊕ Problems   ⏩ Terminal   ⏺ Services   ⚒ Build   ⧉ Dependencies   ☰ Messages

Pushed 2 commits to origin/master (22 minutes ago)

```
Terminal:  run-job.sh ×   run-job.sh ×  +  ∨
' Man Says To Teen Sister",0,0,0,0
' Says Dad Accused Of Child Neglect",0,0,0,0
' Says Dr. Phil To Sisters In A Child Custody Dispute",0,0,0,0
' Says Ex-Wife Of Convicted Serial Killer Shawn Grate",0,0,0,0
' Says Father Of Munchausen By Proxy Target",0,0,0,0
' Says Woman Whose Parents Claim She's Desperate To Escape",0,0,0,0
" Asserts He Is "A King" | Kasie DC | MSNBC",0,0,0,0
"",0,0,0,0
刘昊然,0,0,0,0
或遭巨额罚款真的"非死不可"??小哥谈政事03月22日",0,0,0,0
為嫁豪門退出娛樂圈,0,0,0,0
竇驍,0,0,0,0
終於有一天公公忍不住了",0,0,0,0
黃景瑜 Johnny,0,0,0,0
黃景瑜 Johnny）【未刪減版】",0,0,0,0
```

FULL CODE LINK GITHUB

https://github.com/Rjndrkha/Kuis2_BigData