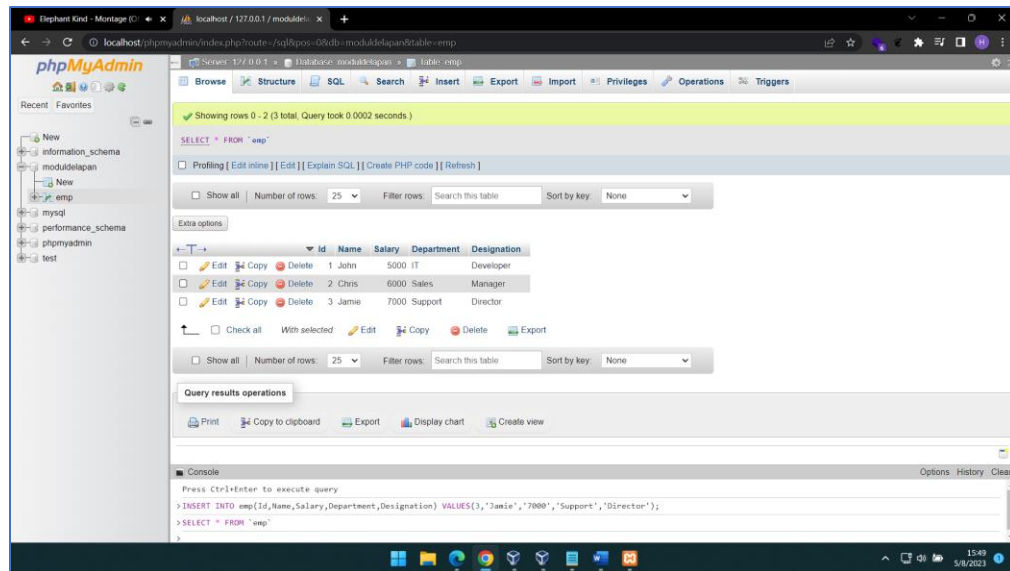


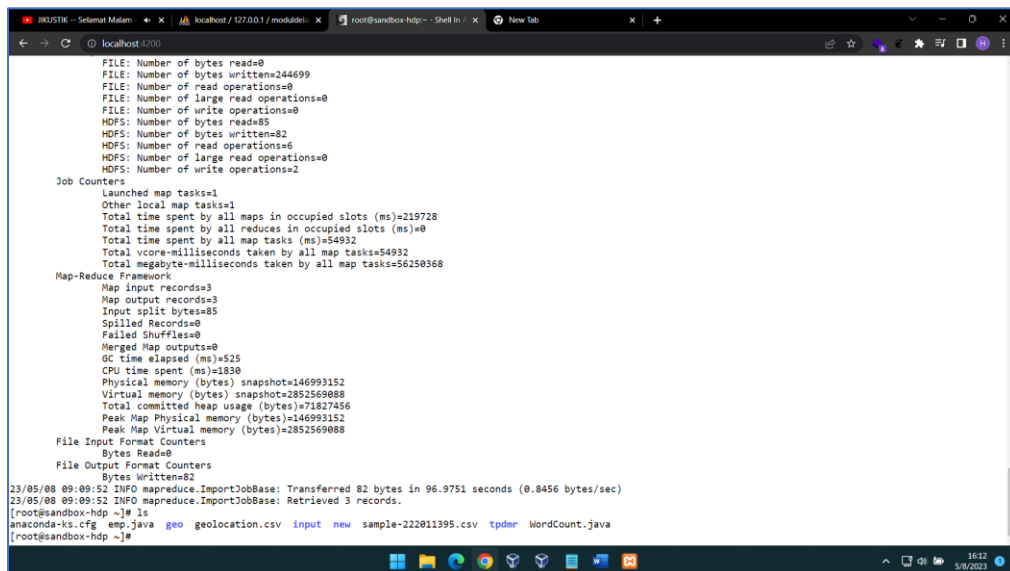
Praktikum Teknologi Perekayasaan Data

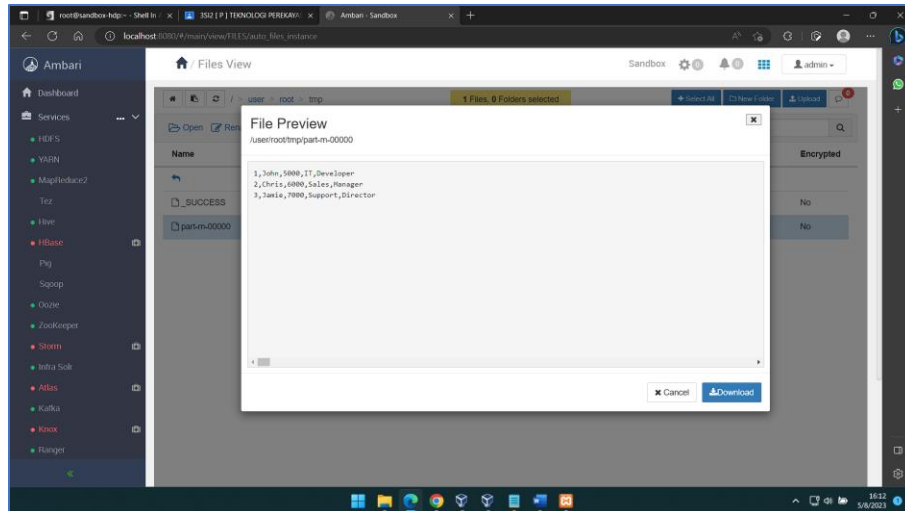
Tanggal Praktikum : Monday, May 8, 2023

1. Database (mysql from xampp)

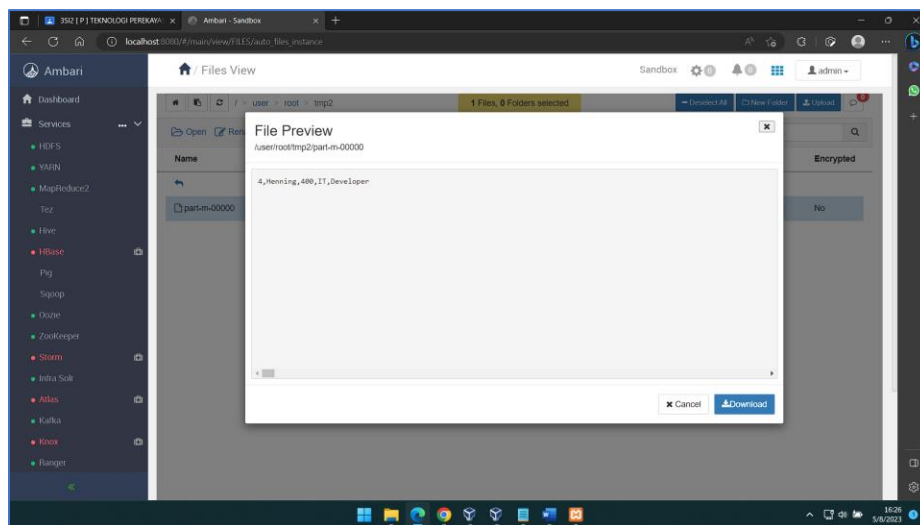


2. Import table

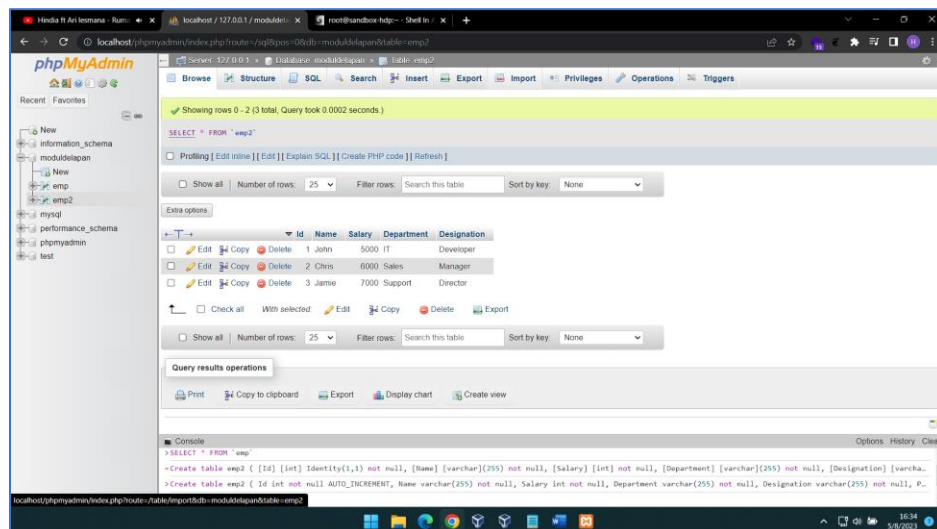




3. Import Incremental



4. Export Data

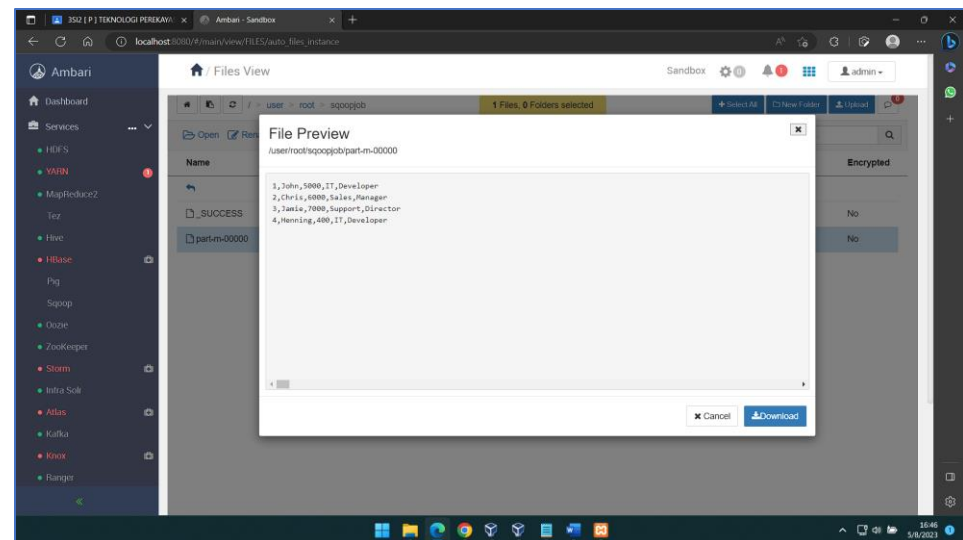


5. Sqoop Job

```
23/05/08 09:40:06 INFO sqoop.Sqoop: Running Sqoop version: 1.4.7.3.0.1.0-187
23/05/08 09:40:06 ERROR tool.BaseSqoopTool: Error parsing arguments for job:
23/05/08 09:40:06 ERROR tool.BaseSqoopTool: Unrecognized argument: --import
23/05/08 09:40:06 ERROR tool.BaseSqoopTool: Unrecognized argument: --connect
23/05/08 09:40:06 ERROR tool.BaseSqoopTool: Unrecognized argument: jdbc:mysql://10.0.2.2:3306/moduldelapan
23/05/08 09:40:06 ERROR tool.BaseSqoopTool: Unrecognized argument: --username
23/05/08 09:40:06 ERROR tool.BaseSqoopTool: Unrecognized argument: root
23/05/08 09:40:06 ERROR tool.BaseSqoopTool: Unrecognized argument: -P
23/05/08 09:40:06 ERROR tool.BaseSqoopTool: Unrecognized argument: --table
23/05/08 09:40:06 ERROR tool.BaseSqoopTool: Unrecognized argument: emp
23/05/08 09:40:06 ERROR tool.BaseSqoopTool: Unrecognized argument: --target-dir
23/05/08 09:40:06 ERROR tool.BaseSqoopTool: Unrecognized argument: sqoopjob
23/05/08 09:40:06 ERROR tool.BaseSqoopTool: Unrecognized argument: --m
23/05/08 09:40:06 ERROR tool.BaseSqoopTool: Unrecognized argument: 1

Try --help for usage instructions.
[root@sandbox-hdp ~]# sqoop job --create sqoopjob --import --connect "jdbc:mysql://10.0.2.2:3306/moduldelapan" --username root -P --table emp --target-dir sqoopjob --m
1
Warning: /usr/hdp/3.0.1.0-187/accumulo does not exist! Accumulo imports will fail.
Please set $ACCUMULO_HOME to the root of your Accumulo installation.
SLF4J: Class path contains multiple SLF4J bindings.
SLF4J: Found binding in [jar:file:/usr/hdp/3.0.1.0-187/hadoop/lib/slf4j-log4j12-1.7.25.jar!/org/slf4j/impl/StaticLoggerBinder.class]
SLF4J: Found binding in [jar:file:/usr/hdp/3.0.1.0-187/hive/lib/log4j-slf4j-impl-2.10.0.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.slf4j.impl.Log4jLoggerFactory]
23/05/08 09:41:18 INFO sqoop.Sqoop: Running Sqoop version: 1.4.7.3.0.1.0-187
23/05/08 09:41:19 INFO manager.SqlManager: Using default fetchSize of 1000
Enter password:
[root@sandbox-hdp ~]# sqoop job --list
Warning: /usr/hdp/3.0.1.0-187/accumulo does not exist! Accumulo imports will fail.
Please set $ACCUMULO_HOME to the root of your Accumulo installation.
SLF4J: Class path contains multiple SLF4J bindings.
SLF4J: Found binding in [jar:file:/usr/hdp/3.0.1.0-187/hadoop/lib/slf4j-log4j12-1.7.25.jar!/org/slf4j/impl/StaticLoggerBinder.class]
SLF4J: Found binding in [jar:file:/usr/hdp/3.0.1.0-187/hive/lib/log4j-slf4j-impl-2.10.0.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.slf4j.impl.Log4jLoggerFactory]
23/05/08 09:42:12 INFO sqoop.Sqoop: Running Sqoop version: 1.4.7.3.0.1.0-187
23/05/08 09:42:12 INFO manager.SqlManager: Using default fetchSize of 1000
Available jobs:
sqoopjob
[root@sandbox-hdp ~]#
```

```
23/05/08 09:45:18 INFO mapreduce.Job: Counters: 32
File System Counters
  FILE: Number of bytes read=0
  FILE: Number of bytes written=245131
  FILE: Number of read operations=0
  FILE: Number of large read operations=0
  FILE: Number of write operations=0
  HDFS: Number of bytes read=85
  HDFS: Number of bytes written=109
  HDFS: Number of read operations=6
  HDFS: Number of large read operations=0
  HDFS: Number of write operations=2
Job Counters
  Launched map tasks=1
  Other local map tasks=1
  Total time spent by all maps in occupied slots (ms)=47672
  Total time spent by all reduces in occupied slots (ms)=0
  Total time spent by all map tasks (ms)=11918
  Total time spent by all map tasks (ms)=11918
  Total megabyte-milliseconds taken by all map tasks=12204032
Map-Reduce Framework
  Map input records=4
  Map output records=4
  Input split bytes=85
  Spilled Records=0
  Failed Shuffles=0
  Merged Map outputs=0
  GC time elapsed (ms)=62
  CPU time spent (ms)=1650
  Physical memory (bytes) snapshot=189456384
  Virtual memory (bytes) snapshot=2851680256
  Total committed heap usage (bytes)=103731968
  Peak Map Physical memory (bytes)=189456384
  Peak Map Virtual memory (bytes)=2851680256
File Input Format Counters
  Bytes Read=0
File Output Format Counters
  Bytes Written=109
23/05/08 09:45:18 INFO mapreduce.ImportJobBase: Transferred 109 bytes in 42.7428 seconds (2.5501 bytes/sec)
23/05/08 09:45:18 INFO mapreduce.ImportJobBase: Retrieved 4 records.
[root@sandbox-hdp ~]#
```



6. Comman Sqoop Lainnya

a. View All Table

```
Launched map tasks=1
Other local map tasks=1
Total time spent by all maps in occupied slots (ms)=47672
Total time spent by all reduces in occupied slots (ms)=0
Total time spent by all map tasks (ms)=11918
Total vcore-millisecods taken by all map tasks=11918
Total megabyte-millisecods taken by all map tasks=12204032

Map-Reduce Framework
  Map input records=4
  Map output records=4
  Input split bytes=85
  Spilled Records=0
  Failed Shuffles=0
  Merged Map outputs=0
  GC time elapsed (ms)=62
  CPU time spent (ms)=1650
  Physical memory (bytes) snapshot=189456384
  Virtual memory (bytes) snapshot=2851680256
  Total committed heap usage (bytes)=123731968
  Peak Map Physical memory (bytes)=189456384
  Peak Map Virtual memory (bytes)=2851680256

File Input Format Counters
  Bytes Read=0
  File Output Format Counters
    Bytes Written=109
23/05/08 09:45:18 INFO mapreduce.ImportJobBase: Transferred 109 bytes in 42.7428 seconds (2.5501 bytes/sec)
23/05/08 09:45:18 INFO mapreduce.ImportJobBase: Retrieved 4 records.
[root@sandbox-hdp ~]# sqoop list-tables --connect 'jdbc:mysql://10.0.2.2:3306/modulelapan' --username root -P
Warning: /usr/hdp/3.0.1-0-187/accumulo does not exist! Accumulo imports will fail.
Please set $ACCUMULO_HOME to the root of your Accumulo installation.
SLF4J: Class path contains multiple SLF4J bindings.
SLF4J: Found binding in [jar:file:/usr/hdp/3.0.1-0-187/hadoop/lib/slf4j-log4j12-1.7.25.jar!/org/slf4j/impl/StaticLoggerBinder.class]
SLF4J: Found binding in [jar:file:/usr/hdp/3.0.1-0-187/hive/lib/log4j-slf4j-impl-2.10.0.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.slf4j.impl.Log4jLoggerFactory]
23/05/08 09:48:47 INFO sqoop.Sqoop: Running Sqoop version: 1.4.7.3.0.1-0-187
Enter password:
23/05/08 09:48:49 INFO manager.MySQLManager: Preparing to use a MySQL streaming resultset.
emp
emp2
[root@sandbox-hdp ~]#
```

b. Read Table

```
23/05/08 09:52:34 INFO sqoop.Sqoop: Running Sqoop version: 1.4.7.3.0.1-0-187
Enter password:
23/05/08 09:52:36 ERROR tool.BaseSqoopTool: Error parsing arguments for eval:
23/05/08 09:52:36 ERROR tool.BaseSqoopTool: Unrecognized argument: --query
23/05/08 09:52:36 ERROR tool.BaseSqoopTool: Unrecognized argument: select * from emp2

Try --help for usage instructions.
[root@sandbox-hdp ~]# sqoop eval --connect 'jdbc:mysql://10.0.2.2:3306/modulelapan' --username root -P --query "SELECT * FROM emp2"
Warning: /usr/hdp/3.0.1-0-187/accumulo does not exist! Accumulo imports will fail.
Please set $ACCUMULO_HOME to the root of your Accumulo installation.
SLF4J: Class path contains multiple SLF4J bindings.
SLF4J: Found binding in [jar:file:/usr/hdp/3.0.1-0-187/hadoop/lib/slf4j-log4j12-1.7.25.jar!/org/slf4j/impl/StaticLoggerBinder.class]
SLF4J: Found binding in [jar:file:/usr/hdp/3.0.1-0-187/hive/lib/log4j-slf4j-impl-2.10.0.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.slf4j.impl.Log4jLoggerFactory]
23/05/08 09:53:39 INFO sqoop.Sqoop: Running Sqoop version: 1.4.7.3.0.1-0-187
Enter password:
23/05/08 09:53:40 ERROR tool.BaseSqoopTool: Error parsing arguments for eval:
23/05/08 09:53:40 ERROR tool.BaseSqoopTool: Unrecognized argument: --query
23/05/08 09:53:40 ERROR tool.BaseSqoopTool: Unrecognized argument: SELECT * FROM emp2

Try --help for usage instructions.
[root@sandbox-hdp ~]# sqoop eval --connect 'jdbc:mysql://10.0.2.2:3306/modulelapan' --username root -P --query "select * from emp2"
Warning: /usr/hdp/3.0.1-0-187/accumulo does not exist! Accumulo imports will fail.
Please set $ACCUMULO_HOME to the root of your Accumulo installation.
SLF4J: Class path contains multiple SLF4J bindings.
SLF4J: Found binding in [jar:file:/usr/hdp/3.0.1-0-187/hadoop/lib/slf4j-log4j12-1.7.25.jar!/org/slf4j/impl/StaticLoggerBinder.class]
SLF4J: Found binding in [jar:file:/usr/hdp/3.0.1-0-187/hive/lib/log4j-slf4j-impl-2.10.0.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.slf4j.impl.Log4jLoggerFactory]
23/05/08 09:54:50 INFO sqoop.Sqoop: Running Sqoop version: 1.4.7.3.0.1-0-187
Enter password:
23/05/08 09:54:52 INFO manager.MySQLManager: Preparing to use a MySQL streaming resultset.
+----+-----+-----+-----+-----+
| Id | Name | Salary | Department | Designation |
+----+-----+-----+-----+-----+
| 1 | John | 5000 | IT | Developer |
| 2 | Chris | 6000 | Sales | Manager |
| 3 | Jamie | 7000 | Support | Director |
+----+-----+-----+-----+-----+
[root@sandbox-hdp ~]#
```

c. Insert row in table

```
phpMyAdmin
Database: modulelapan
Table: emp2

Showing rows 0 - 3 (4 total, Query took 0.0002 seconds)

SELECT * FROM `emp2`

+----+-----+-----+-----+-----+
| Id | Name | Salary | Department | Designation |
+----+-----+-----+-----+-----+
| 1 | John | 5000 | IT | Developer |
| 2 | Chris | 6000 | Sales | Manager |
| 3 | Jamie | 7000 | Support | Director |
+----+-----+-----+-----+-----+

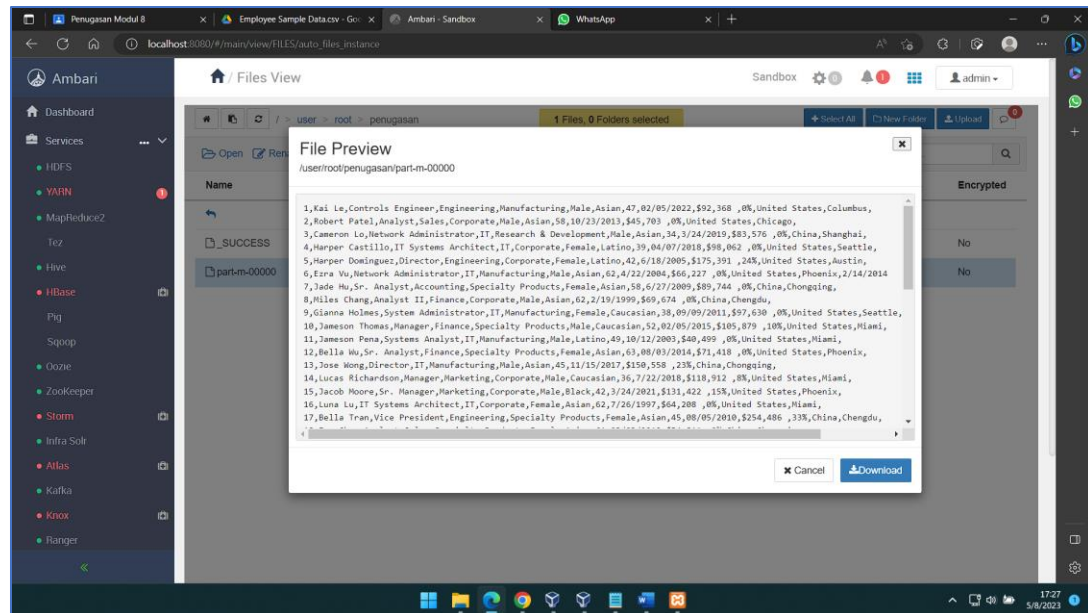
Query results operations
Print Copy to clipboard Export Display chart Create view

Console
> SELECT * FROM `emp2`
> Create table emp2 ( [Id] [int] Identity(1,1) not null, [Name] [varchar](255) not null, [Salary] [int] not null, [Department] [varchar](255) not null, [Designation] [varc
> Create table emp2 ( [Id] int not null AUTO_INCREMENT, Name varchar(255) not null, Salary int not null, Department varchar(255) not null, Designation varchar(255) not null,
> SELECT * FROM `emp2`
```

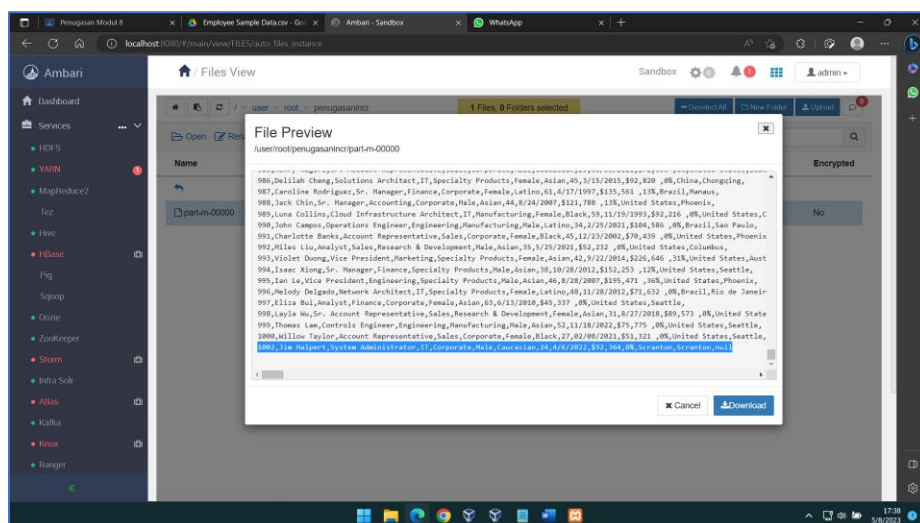
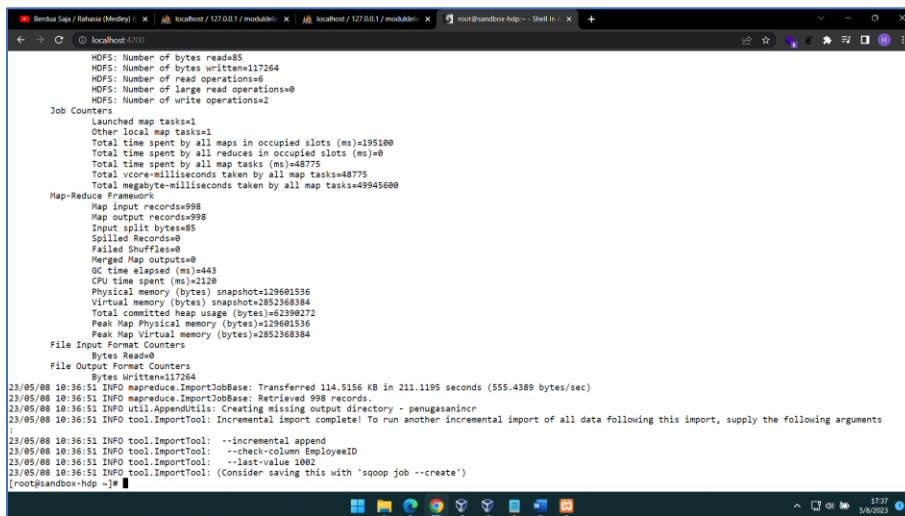
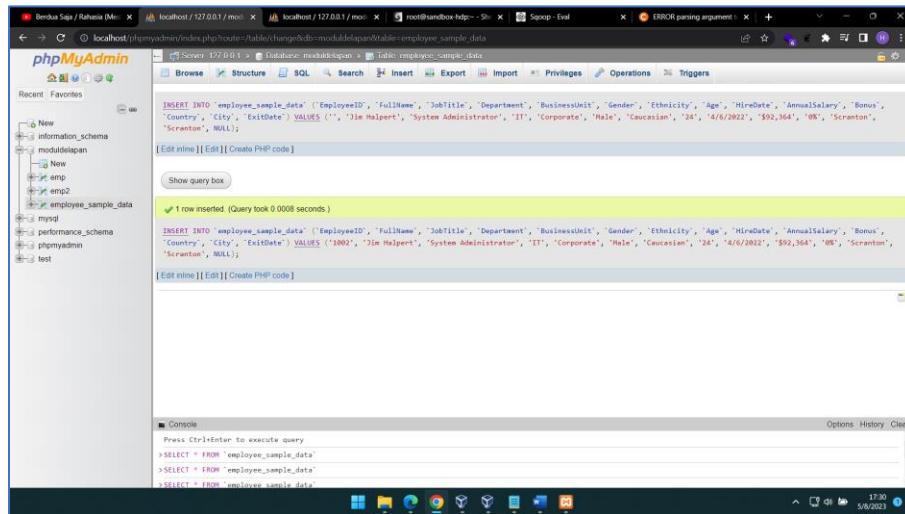
Penugasan

1. Proses dan hasil impor data ke HDFS.

```
23/05/08 10:06:44 INFO mapreduce.Job: Counters: 32
  File System Counters
    FILE: Number of bytes read=0
    FILE: Number of bytes written=244868
    FILE: Number of read operations=0
    FILE: Number of large read operations=0
    FILE: Number of write operations=0
    HDFS: Number of bytes read=85
    HDFS: Number of bytes written=117478
    HDFS: Number of read operations=6
    HDFS: Number of large read operations=0
    HDFS: Number of write operations=2
  Job Counters
    Launched map tasks=1
    Other local map tasks=1
    Total time spent by all maps in occupied slots (ms)=288876
    Total time spent by all reduces in occupied slots (ms)=0
    Total time spent by all map tasks (ms)=72219
    Total vcore-milliseconds taken by all map tasks=72219
    Total megabyte-milliseconds taken by all map tasks=73952256
  Map-Reduce Framework
    Map input records=1000
    Map output records=1000
    Input split bytes=85
    Spilled Records=0
    Failed Shuffles=0
    Merged Map outputs=0
    GC time elapsed (ms)=421
    CPU time spent (ms)=1790
    Physical memory (bytes) snapshot=142917632
    Virtual memory (bytes) snapshot=2852265984
    Total committed heap usage (bytes)=76621760
    Peak Map Physical memory (bytes)=142979880
    Peak Map Virtual memory (bytes)=2852265984
  File Input Format Counters
    Bytes Read=0
  File Output Format Counters
    Bytes Written=117478
23/05/08 10:06:44 INFO mapreduce.ImportJobBase: Transferred 114.7246 KB in 110.3185 seconds (1.0399 KB/sec)
23/05/08 10:06:44 INFO mapreduce.ImportJobBase: Retrieved 1000 records.
[root@sandbox-hdp ~]#
```

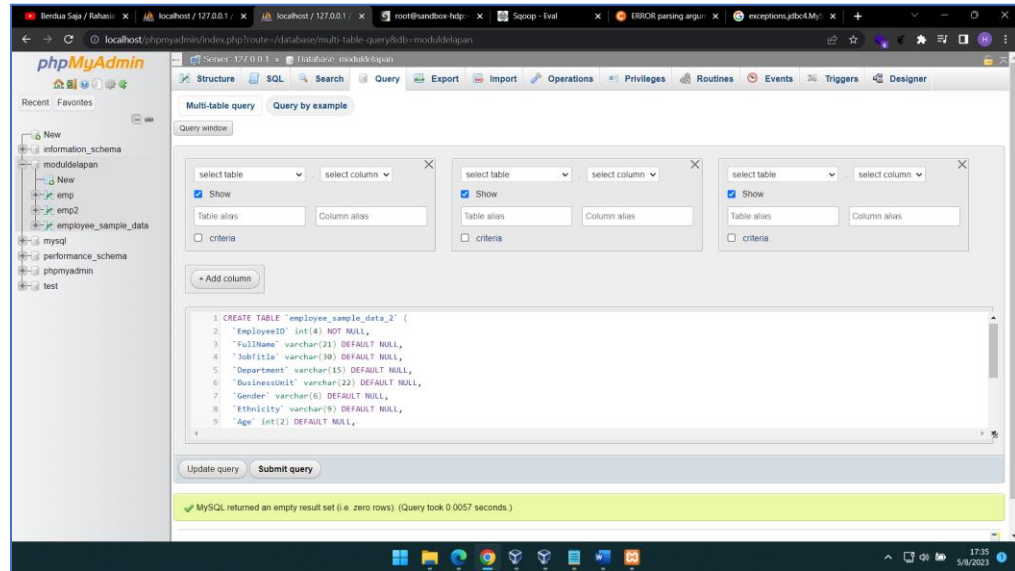


2. Proses incremental load untuk data berikut sekaligus hasil incremental load-nya (untuk record yang akan di-load dapat dibuat sendiri)

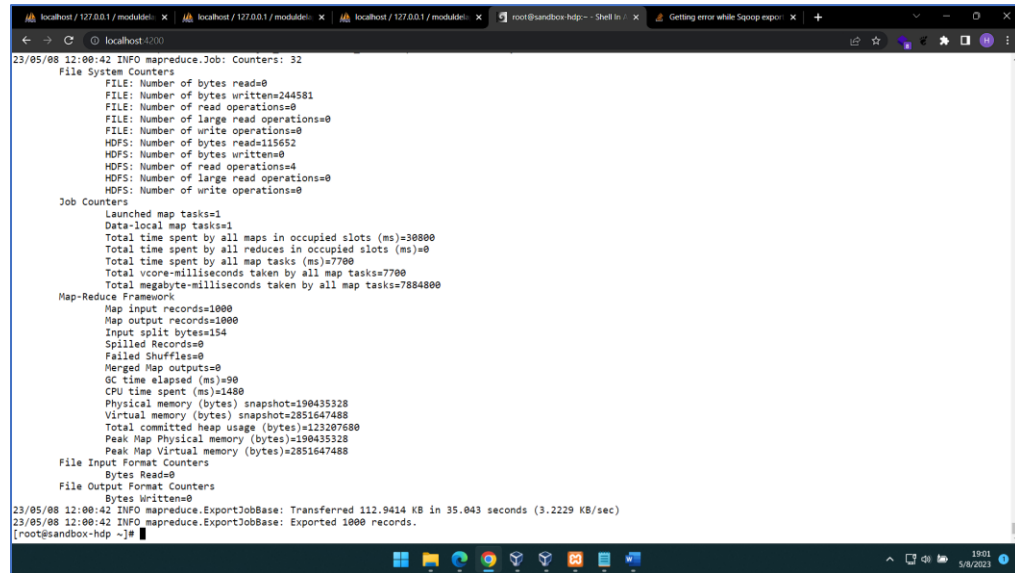


3. Proses ekspor data:

- Create table



- Proses ekspor data dari HDFS ke RDBMS



- Hasil ekspor data

Showing rows 0 - 24 (1000 total. Query took 0.0002 seconds)

```
SELECT * FROM 'employee_sample_data2'
```

EmployeeID	FullName	JobTitle	Department	BusinessUnit	Gender	Ethnicity	Age	HireDate	AnnualSalary	Bonus	Country	City
1	Kai Le	Controls Engineer	Engineering	Manufacturing	Male	Asian	47	02/05/2022	\$92	368	0%	United States
2	Robert Patel	Analyst	Sales	Corporate	Male	Asian	58	10/23/2013	\$45	703	0%	United States
3	Cameron Lo	Network Administrator	IT	Research & Development	Male	Asian	34	3/24/2019	\$83	576	0%	China
4	Harper Castillo	IT Systems Architect	IT	Corporate	Female	Latino	39	04/07/2018	\$98	962	0%	United States
5	Harper Dominguez	Director	Engineering	Corporate	Female	Latino	42	6/18/2005	\$175	391	24%	United States
6	Ezra Vu	Network Administrator	IT	Manufacturing	Male	Asian	62	4/22/2004	\$66	227	0%	United States
7	Jade Hu	Sr Analyst	Accounting	Specialty Products	Female	Asian	58	6/27/2009	\$89	744	0%	China
8	Miles Chang	Analyst II	Finance	Corporate	Male	Asian	62	2/19/1999	\$60	674	0%	China

4. Proses insert data ke RDBMS dengan menggunakan sqoop eval.

```
Enter password:
23/05/08 10:56:27 INFO manager.MySQLManager: Preparing to use a MySQL streaming resultset.
23/05/08 10:56:27 WARN tool.EvalSqlTool: SQL exception executing statement: com.mysql.jdbc.exceptions.jdbc4.MySQLSyntaxErrorException: You have an error in your SQL syntax; check the manual that corresponds to your MariaDB server version for the right syntax to use near ''employee_sample_data2'('EmployeeID', 'FullName', 'JobTitle', 'De
partment', 'B...' at line 1
at sun.reflect.NativeConstructorAccessorImpl.newInstance0(Native Method)
at sun.reflect.NativeConstructorAccessorImpl.newInstance(NativeConstructorAccessorImpl.java:62)
at sun.reflect.DelegatingConstructorAccessorImpl.newInstance(DelegatingConstructorAccessorImpl.java:45)
at java.lang.reflect.Constructor.newInstance(Constructor.java:423)
at com.mysql.jdbc.Util.handleNewInstance(Util.java:425)
at com.mysql.jdbc.Util.getInstance(Util.java:488)
at com.mysql.jdbc.SQLException.createSQLException(SQLException.java:944)
at com.mysql.jdbc.MySQLIO.checkErrorPacket(MySQLIO.java:3973)
at com.mysql.jdbc.MySQLIO.checkErrorPacket(MySQLIO.java:3989)
at com.mysql.jdbc.MySQLIO.sendCommand(MySQLIO.java:2527)
at com.mysql.jdbc.MySQLIO.sqlQueryDirect(MySQLIO.java:2698)
at com.mysql.jdbc.ConnectionImpl.execSQL(ConnectionImpl.java:2484)
at com.mysql.jdbc.PreparedStatement.executeInternal(PreparedStatement.java:1858)
at com.mysql.jdbc.PreparedStatement.execute(PreparedStatement.java:1197)
at org.apache.sqoop.tool.EvalSqlTool.run(EvalSqlTool.java:68)
at org.apache.sqoop.Sqoop.run(Sqoop.java:158)
at org.apache.hadoop.util.ToolRunner.run(ToolRunner.java:76)
at org.apache.sqoop.Sqoop.runSqoop(Sqoop.java:186)
at org.apache.sqoop.Sqoop.runTool(Sqoop.java:248)
at org.apache.sqoop.Sqoop.runTool(Sqoop.java:249)
at org.apache.sqoop.Sqoop.main(Sqoop.java:258)

[root@sandbox-hdp ~]# sqoop eval --connect 'jdbc:mysql://10.0.2.2:3306/moduldelapan' --username root -P --query "INSERT INTO employee_sample_data VALUES ('1004','Pamela
Beasley','Receptionist','Temp','Sales','Female','MGA','32','3/5/2022','$212','3%','West Virginia','Virginia','3/3/2023')"
```