

FINAL BDA Project - Sarah Khalid 27258

Hadoop Container

```
Administrator: Command Prompt - docker exec -it namenode /bin/bash
Microsoft Windows [Version 10.0.22621.1778]
(c) Microsoft Corporation. All rights reserved.

C:\Windows\System32>

D:\>cd big.data

D:\big.data>git clone https://github.com/big-data-europe/docker-hadoop.git
fatal: destination path 'docker-hadoop' already exists and is not an empty directory.

D:\big.data>cd docker-hadoop

D:\big.data\docker-hadoop>docker-compose up -d
[+] Running 21/28
   - resourcemanager 2 layers [=====] 0B/0B Pulling 1714.2s
   - b04764123f3e Download complete 1324.0s
   - b04394ddb35d Download complete 1325.4s
   - historyserver 3 layers [=====] 0B/0B Pulling 1714.2s
   - b2dc88cebe05 Download complete 1376.6s
   - 13b908760168 Download complete 1386.5s
   - 0991d53828a1 Download complete 1395.0s
   - namenode 13 layers [=====] 0B/0B Pulling 1714.2s
   - 3192219af404 Pull complete 880.5s
   - 7127a1d8cccd Downloading 1787.3s
   - 881a89599900 Download complete 9.1s
   - 77920a3e82af Download complete 30.3s
   - 92329e81ae4c Downloading 1707.3s
   - f373218fec59 Download complete 909.8s
   - ae53513fe997 Download complete 947.9s
   - 8b1800105b98 Download complete 1002.7s
   - c3a84a3e49c8 Download complete 1041.0s
   - ae5640e6da76 Download complete 1116.8s
   - facffb3afde3 Download complete 1119.8s
   - c71a6df73788 Download complete 1122.0s
   - 73b8c8ccb707 Download complete 1125.3s
   - nodemanager1 2 layers [=====] 0B/0B Pulling 1714.2s
   - beaa171f32f6 Download complete 1367.5s
   - 50ddad4de8a9 Download complete 1371.4s
   - datanode 3 layers [=====] 0B/0B Pulling 1714.2s
   - 3ca2ec07878c Download complete 1136.1s
   - 26c2d45430e Download complete 1141.0s
   - 13c9c87a46cb Download complete 1309.6s
Dialling registry-1.docker.io:443 with direct connection: resolving host registry-1.docker.io: lookup registry-1.docker.io: no such host

D:\big.data\docker-hadoop>docker-compose up -d
[+] Running 26/28
   - historyserver 13 layers [=====] 0B/0B Pulled 205.8s
   - 3192219af404 Pull complete 50.1s
   - 7127a1d8cccd Pull complete 109.5s
   - 881a89599900 Pull complete 109.6s
   - 77920a3e82af Pull complete 109.7s
   - 92329e81ae4c Pull complete 201.1s
   - f373218fec59 Pull complete 201.2s
   - ae53513fe997 Pull complete 201.2s
   - 8b1800105b98 Pull complete 201.3s
   - c3a84a3e49c8 Pull complete 201.4s
   - ae5640e6da76 Pull complete 201.4s
   - b2dc88cebe05 Pull complete 201.6s
   - 13b908760168 Pull complete 201.7s
   - 0991d53828a1 Pull complete 201.9s
   - datanode 3 layers [=====] 0B/0B Pulled 205.8s
   - 3ca2ec07878c Pull complete 201.6s
   - 26c2d45430e Pull complete 201.7s
   - 13c9c87a46cb Pull complete 201.9s
   - namenode 3 layers [=====] 0B/0B Pulled 205.7s
   - facffb3afde3 Pull complete 201.5s
   - c71a6df73788 Pull complete 201.7s
   - 73b8c8ccb707 Pull complete 201.8s
   - resourcemanager 2 layers [=====] 0B/0B Pulled 205.6s
   - b04764123f3e Pull complete 201.5s
   - b04394ddb35d Pull complete 201.7s
   - nodemanager1 2 layers [=====] 0B/0B Pulled 205.6s
   - beaa171f32f6 Pull complete 201.5s
   - 50ddad4de8a9 Pull complete 201.7s
[+] Running 9/9
   - Network docker-hadoop_default Created 0.7s
   - Volume "docker-hadoop_historyserver" Created 0.0s
   - Volume "docker-hadoop_hadoop_namenode" Created 0.0s
   - Volume "docker-hadoop_hadoop_datanode" Created 0.0s
   - Container datanode Started 3.9s
   - Container historyserver Started 5.4s
   - Container resourcemanager Started 5.7s
   - Container namenode Started 5.4s
   - Container nodemanager Started 4.1s

D:\big.data\docker-hadoop>docker ps
CONTAINER ID        IMAGE                                     COMMAND                  CREATED          STATUS          PORTS                               NAMES
2c6571458631       bde2020/hadoop-historyserver:2.0.0-hadoop3.2.1-java8  "/entrypoint.sh /run..." About a minute ago Up About a minute (healthy) 0.0.0.0:8188->8188/tcp historyserver
889454844c5d       bde2020/hadoop-nodemanager:2.0.0-hadoop3.2.1-java8    "/entrypoint.sh /run..." About a minute ago Up About a minute (healthy) 0.0.0.0:8042->8042/tcp nodemanager
3e567b4b21c0       bde2020/hadoop-namenode:2.0.0-hadoop3.2.1-java8      "/entrypoint.sh /run..." About a minute ago Up About a minute (healthy) 0.0.0.0:9800->9800/tcp, 0.0.0.0:9870->9870/tcp namenode
84d7bea1091c       bde2020/hadoop-resourceanager:2.0.0-hadoop3.2.1-java8 "/entrypoint.sh /run..." About a minute ago Up About a minute (healthy) 0.0.0.0:8088->8088/tcp resourcemanager
759783e2fcf4       bde2020/hadoop-datanode:2.0.0-hadoop3.2.1-java8      "/entrypoint.sh /run..." About a minute ago Up About a minute (healthy) 0.0.0.0:9864->9864/tcp datanode
```

```
Administrator: Command Prompt - docker exec -it namenode /bin/bash
[+] Running 26/28
   - historyserver 13 layers [=====] 0B/0B Pulled 205.8s
   - 3192219af404 Pull complete 50.1s
   - 7127a1d8cccd Pull complete 109.5s
   - 881a89599900 Pull complete 109.6s
   - 77920a3e82af Pull complete 109.7s
   - 92329e81ae4c Pull complete 201.1s
   - f373218fec59 Pull complete 201.2s
   - ae53513fe997 Pull complete 201.2s
   - 8b1800105b98 Pull complete 201.3s
   - c3a84a3e49c8 Pull complete 201.4s
   - ae5640e6da76 Pull complete 201.4s
   - b2dc88cebe05 Pull complete 201.6s
   - 13b908760168 Pull complete 201.7s
   - 0991d53828a1 Pull complete 201.9s
   - datanode 3 layers [=====] 0B/0B Pulled 205.8s
   - 3ca2ec07878c Pull complete 201.6s
   - 26c2d45430e Pull complete 201.7s
   - 13c9c87a46cb Pull complete 201.9s
   - namenode 3 layers [=====] 0B/0B Pulled 205.7s
   - facffb3afde3 Pull complete 201.5s
   - c71a6df73788 Pull complete 201.7s
   - 73b8c8ccb707 Pull complete 201.8s
   - resourcemanager 2 layers [=====] 0B/0B Pulled 205.6s
   - b04764123f3e Pull complete 201.5s
   - b04394ddb35d Pull complete 201.7s
   - nodemanager1 2 layers [=====] 0B/0B Pulled 205.6s
   - beaa171f32f6 Pull complete 201.5s
   - 50ddad4de8a9 Pull complete 201.7s
[+] Running 9/9
   - Network docker-hadoop_default Created 0.7s
   - Volume "docker-hadoop_historyserver" Created 0.0s
   - Volume "docker-hadoop_hadoop_namenode" Created 0.0s
   - Volume "docker-hadoop_hadoop_datanode" Created 0.0s
   - Container datanode Started 3.9s
   - Container historyserver Started 5.4s
   - Container resourcemanager Started 5.7s
   - Container namenode Started 5.4s
   - Container nodemanager Started 4.1s

D:\big.data\docker-hadoop>docker ps
CONTAINER ID        IMAGE                                     COMMAND                  CREATED          STATUS          PORTS                               NAMES
2c6571458631       bde2020/hadoop-historyserver:2.0.0-hadoop3.2.1-java8  "/entrypoint.sh /run..." About a minute ago Up About a minute (healthy) 0.0.0.0:8188->8188/tcp historyserver
889454844c5d       bde2020/hadoop-nodemanager:2.0.0-hadoop3.2.1-java8    "/entrypoint.sh /run..." About a minute ago Up About a minute (healthy) 0.0.0.0:8042->8042/tcp nodemanager
3e567b4b21c0       bde2020/hadoop-namenode:2.0.0-hadoop3.2.1-java8      "/entrypoint.sh /run..." About a minute ago Up About a minute (healthy) 0.0.0.0:9800->9800/tcp, 0.0.0.0:9870->9870/tcp namenode
84d7bea1091c       bde2020/hadoop-resourceanager:2.0.0-hadoop3.2.1-java8 "/entrypoint.sh /run..." About a minute ago Up About a minute (healthy) 0.0.0.0:8088->8088/tcp resourcemanager
759783e2fcf4       bde2020/hadoop-datanode:2.0.0-hadoop3.2.1-java8      "/entrypoint.sh /run..." About a minute ago Up About a minute (healthy) 0.0.0.0:9864->9864/tcp datanode
```

```
Administrator: Command Prompt - docker exec -it namenode /bin/bash
operable program or batch file.

D:\big.data\docker-hadoop>
'.' is not recognized as an internal or external command,
operable program or batch file.

D:\big.data\docker-hadoop>cd..
D:\big.data\docker-hive>
'docker-hive' is not recognized as an internal or external command,
operable program or batch file.

D:\big.data>cd docker hive
The system cannot find the path specified.

D:\big.data>cd docker-hive
D:\big.data\docker-hive>cd..
D:\big.data>cd docker-hadoop

D:\big.data\docker-hadoop>docker cp C:/Users/HP/Desktop/Test BDA/data/MDI_TEST.csv namenode:/tmp/
'docker cp' requires exactly 2 arguments.
See 'docker cp --help'.

Usage: docker cp [OPTIONS] CONTAINER:SRC_PATH DEST_PATH|-
       docker cp [OPTIONS] SRC_PATH|- CONTAINER:DEST_PATH

Copy files/folders between a container and the local filesystem

D:\big.data\docker-hadoop>docker cp C:/Users/HP/Desktop/TestBDA/data/MDI_TEST.csv namenode:/tmp/
Successfully copied 2.05KB to namenode:/tmp/

D:\big.data\docker-hadoop>docker exec -it namenode /bin/bash
root@3e567b4b21c0:/#
```

```
Administrator: Command Prompt - docker exec -it namenode /bin/bash

[-stat [format] <path> ...]
[-tail [-f] [-s <sleep interval>] <file>]
[-test [-d] [-s] [-w] <path>]
[-text [-ignoreCrc] <src> ...]
[-touch [-a] [-m] [-t TIME Stamp] [-c] <path> ...]
[-touchz <path> ...]
[-truncate [-w] <length> <path> ...]
[-usage [cmd ...]]

Generic options supported are:
-conf <configuration file>      specify an application configuration file
-D <property=value>             define a value for a given property
-fs <file:///hdfs://namenode:port> specify default filesystem URL to use, overrides 'fs.defaultFS' property from configurations.
-it <localResourceManager:port> specify a ResourceManager
-files <file1,...>              specify a comma-separated list of files to be copied to the map reduce cluster
-libjars <jar1,...>             specify a comma-separated list of jar files to be included in the classpath
-archives <archive1,...>       specify a comma-separated list of archives to be unarchived on the compute machines

The general command line syntax is:
command [genericOptions] [commandOptions]

Usage: hadoop fs [generic options] .mkdir [-p] <path> ...
root@3e567b4b21c0:/# hdfs dfs -mkdir -p /user/root/input
root@3e567b4b21c0:/# hdfs dfs -copyFromLocal /tmp/MDI_TEST.csv /user/root/input
2023-06-02 11:47:00,653 INFO sasl.SaslDataTransferClient: SASL encryption trust check: localhostTrusted = false, remoteHostTrusted = false
root@3e567b4b21c0:/# hdfs dfs -ls
Found 1 items
drwxr-xr-x    0 2023-06-02 11:47 input
root@3e567b4b21c0:/# hdfs dfs -ls /user/root/input
Found 1 items
drwxr-xr-x    3 root supergroup    273 2023-06-02 11:47 /user/root/input/MDI_TEST.csv
root@3e567b4b21c0:/#
```

```
Administrator: Command Prompt - docker exec -it namenode /bin/bash

Generic options supported are:
-conf <configuration file>      specify an application configuration file
-D <property=value>             define a value for a given property
-fs <file:///hdfs://namenode:port> specify default filesystem URL to use, overrides 'fs.defaultFS' property from configurations.
-jt <local|resource|manager:port> specify a ResourceManager
-files <file1,...>              specify a comma-separated list of files to be copied to the map reduce cluster
-libjars <jar1,...>             specify a comma-separated list of jar files to be included in the classpath
-archives <archive1,...>       specify a comma-separated list of archives to be unarchived on the compute machines

The general command line syntax is:
command [genericOptions] [commandOptions]

Usage: hadoop fs [generic options] -mkdir [-p] <path> ...
root@3e567b4b21c0:/# hdfs dfs -mkdir -p /user/root/input
root@3e567b4b21c0:/# hdfs dfs -copyFromLocal /tmp/MDI_TEST.csv /user/root/input
2023-06-02 11:47:00,653 INFO sasl.SaslDataTransferClient: SASL encryption trust check: localhostTrusted = false, remoteHostTrusted = false
root@3e567b4b21c0:/# hdfs dfs -ls
Found 1 items
drwxr-xr-x - root supergroup          0 2023-06-02 11:47 input
root@3e567b4b21c0:/# hdfs dfs -ls /user/root/input
Found 1 items
-rw-r--r-- 3 root supergroup          273 2023-06-02 11:47 /user/root/input/MDI_TEST.csv
root@3e567b4b21c0:/# exit
exit

D:\big.data\docker-hadoop>docker ps
CONTAINER ID   IMAGE                                COMMAND                  CREATED        STATUS        PORTS                                NAMES
2c8571458631   bde2020/hadoop-historyserver:2.0.0-hadoop3.2.1-java8  "/entrypoint.sh /run..." 19 minutes ago Up 19 minutes (healthy) 0.0.0.0:8188->8188/tcp   historyserver
89845484c5d    bde2020/hadoop-nodemanager:2.0.0-hadoop3.2.1-java8    "/entrypoint.sh /run..." 19 minutes ago Up 19 minutes (healthy) 0.0.0.0:8042->8042/tcp   nodemanager
1e567b4b21c0   bde2020/hadoop-namenode:2.0.0-hadoop3.2.1-java8      "/entrypoint.sh /run..." 19 minutes ago Up 19 minutes (healthy) 0.0.0.0:9800->9800/tcp, 0.0.0.0:9870->9870/tcp namenode
84d7bea1091c   bde2020/hadoop-resourcemanager:2.0.0-hadoop3.2.1-java8 "/entrypoint.sh /run..." 19 minutes ago Up 19 minutes (healthy) 0.0.0.0:8088->8088/tcp   resourcemanager
759782efcf4    bde2020/hadoop-datanode:2.0.0-hadoop3.2.1-java8      "/entrypoint.sh /run..." 19 minutes ago Up 19 minutes (healthy) 0.0.0.0:9864->9864/tcp   datanode

D:\big.data\docker-hadoop>docker exec -it namenode /bin/bash
root@3e567b4b21c0:/# hdfs dfs -ls /input
ls: '/input': No such file or directory
root@3e567b4b21c0:/# hdfs dfs -ls /user/root/input
Found 1 items
-rw-r--r-- 3 root supergroup          273 2023-06-02 11:47 /user/root/input/MDI_TEST.csv
root@3e567b4b21c0:/# hdfs dfs -mkdir -p /user/root/BDA
root@3e567b4b21c0:/# hdfs dfs -copyFromLocal /tmp/MDI_TEST.csv /user/root/BDA
2023-06-02 11:55:15,467 INFO sasl.SaslDataTransferClient: SASL encryption trust check: localhostTrusted = false, remoteHostTrusted = false
root@3e567b4b21c0:/# hdfs dfs -ls /user/root/BDA
ls: '/user/root/BDA': No such file or directory
root@3e567b4b21c0:/# hdfs dfs -ls /user/root/BDA
Found 1 items
-rw-r--r-- 3 root supergroup          273 2023-06-02 11:55 /user/root/BDA/MDI_TEST.csv
root@3e567b4b21c0:/#
```

There is a folder called BDA inside the HDFS root directory

There is a folder called input inside the HDFS root directory

Hive Container

```
Administrator: Command Prompt - docker-compose exec hive-server bash

Microsoft Windows [Version 10.0.22621.1778]
(c) Microsoft Corporation. All rights reserved.

C:\Windows\System32>D:
D:\>cd big.data
D:\big.data>cd docker-hive
D:\big.data\docker-hive>docker-compose up -d
[+] Running 6/0
  ✓ Container docker-hive-hive-server-1          Started      4.5s
  ✓ Container docker-hive-presto-coordinator-1    Started      3.8s
  ✓ Container docker-hive-hive-metastore-postgresql-1 Started      2.8s
  ✓ Container docker-hive-datanode-1             Started      3.7s
  ✓ Container docker-hive-hive-metastore-1        Started      3.0s
  ✓ Container docker-hive-namenode-1             Started      4.5s

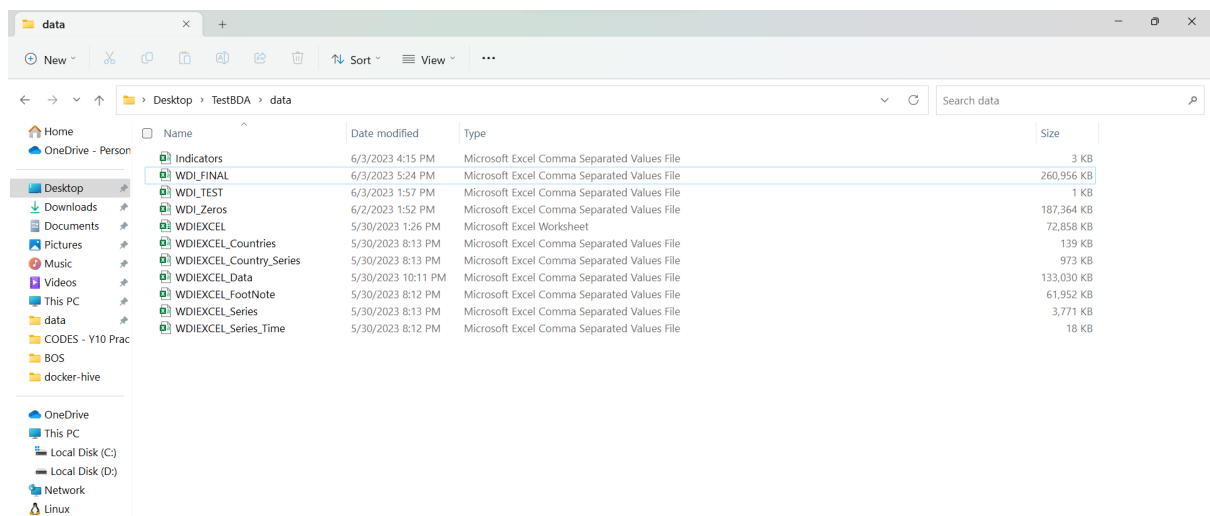
D:\big.data\docker-hive>docker cp C:/Users/HP/Desktop/TestBDA/data/MDI_TEST.csv namenode:/tmp/
Successfully copied 2,05kB to namenode:/tmp/

D:\big.data\docker-hive>docker exec -it namenode /bin/bash
root@3e567b4b21c0:/# hdfs dfs -mkdir -p /user/root/MDI_TEST_UR003
root@3e567b4b21c0:/# hdfs dfs -copyFromLocal /tmp/MDI_TEST.csv /user/root/MDI_TEST_UR003
2023-06-02 12:12:40,701 INFO sasl.SaslDataTransferClient: SASL encryption trust check: localhostTrusted = false, remoteHostTrusted = false
root@3e567b4b21c0:/# hdfs dfs -ls
Found 3 items
drwxr-xr-x - root supergroup          0 2023-06-02 11:55 BDA
drwxr-xr-x - root supergroup          0 2023-06-02 12:12 MDI_TEST_UR003
drwxr-xr-x - root supergroup          0 2023-06-02 11:47 input
root@3e567b4b21c0:/# hdfs dfs -ls /user/root/MDI_TEST_UR003
Found 1 items
-rw-r--r-- 3 root supergroup          273 2023-06-02 12:12 /user/root/MDI_TEST_UR003/MDI_TEST.csv
root@3e567b4b21c0:/# exit
exit

D:\big.data\docker-hive>docker-compose exec hive-server bash
root@2321ca988cd:/opt# /opt/hive/bin/beamline -u jdbc:hive://localhost:10000
SLF4J: Class path contains multiple SLF4J bindings.
SLF4J: Found binding in [jar:file:/opt/hive/lib/log4j-slf4j-impl-2.6.2.jar!/org/slf4j/impl/StaticLoggerBinder.class]
SLF4J: Found binding in [jar:file:/opt/hadoop-2.7.4/share/hadoop/common/lib/slf4j-log4j12-1.7.10.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]
scan complete in 1ms
scan complete in 756ms
No known driver to handle "jdbc:hive://localhost:10000"
beamline version 2.3.2 by Apache Hive
beamline: SHOW DATABASES;
No current connection
```

```
Administrator: Command Prompt - docker-compose exec hive-server bash
root@3e567b4b21c0:/# exit
exit
D:\big_data\docker-hive>docker-compose exec hive-server bash
root@2321c2a988c9:/opt# /opt/hive/bin/beeline -u jdbc:hive2://localhost:10000
SLF4J: Class path contains multiple SLF4J bindings.
SLF4J: Found binding in [jar:file:/opt/hive/lib/log4j-slf4j-impl-2.6.2.jar!/org/slf4j/impl/StaticLoggerBinder.class]
SLF4J: Found binding in [jar:file:/opt/hadoop-2.7.4/share/hadoop/common/lib/slf4j-log4j12-1.7.10.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]
scan complete in 1ms
scan complete in 75ms
No known driver to handle "jdbc:hive2://localhost:10000"
Beeline version 2.3.2 by Apache Hive
beeline> SHOW DATABASES;
No current connection
beeline> /opt/hive/bin/beeline -u jdbc:hive2://localhost:10000
beeline> .> show databases;
No current connection
beeline> !quit
root@2321c2a988c9:/opt# /opt/hive/bin/beeline -u jdbc:hive2://localhost:10000
SLF4J: Class path contains multiple SLF4J bindings.
SLF4J: Found binding in [jar:file:/opt/hive/lib/log4j-slf4j-impl-2.6.2.jar!/org/slf4j/impl/StaticLoggerBinder.class]
SLF4J: Found binding in [jar:file:/opt/hadoop-2.7.4/share/hadoop/common/lib/slf4j-log4j12-1.7.10.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]
Connecting to jdbc:hive2://localhost:10000
Connected to: Apache Hive (version 2.3.2)
Driver: Hive JDBC (version 2.3.2)
Transaction isolation: TRANSACTION_REPEATABLE_READ
Beeline version 2.3.2 by Apache Hive
0: jdbc:hive2://localhost:10000> _
```

Datasets



Name	Date modified	Type	Size
Indicators	6/3/2023 4:15 PM	Microsoft Excel Comma Separated Values File	3 KB
WDI_FINAL	6/3/2023 5:24 PM	Microsoft Excel Comma Separated Values File	260,956 KB
WDI_TEST	6/3/2023 1:57 PM	Microsoft Excel Comma Separated Values File	1 KB
WDI_Zeros	6/2/2023 1:52 PM	Microsoft Excel Comma Separated Values File	187,364 KB
WDIEXCEL	5/30/2023 1:26 PM	Microsoft Excel Worksheet	72,858 KB
WDIEXCEL_Countries	5/30/2023 8:13 PM	Microsoft Excel Comma Separated Values File	139 KB
WDIEXCEL_Country_Series	5/30/2023 8:13 PM	Microsoft Excel Comma Separated Values File	973 KB
WDIEXCEL_Data	5/30/2023 10:11 PM	Microsoft Excel Comma Separated Values File	133,030 KB
WDIEXCEL_FootNote	5/30/2023 8:12 PM	Microsoft Excel Comma Separated Values File	61,952 KB
WDIEXCEL_Series	5/30/2023 8:13 PM	Microsoft Excel Comma Separated Values File	3,771 KB
WDIEXCEL_Series_Time	5/30/2023 8:12 PM	Microsoft Excel Comma Separated Values File	18 KB

COPY CSV INTO NAMENODE TMP DIRECTORY

```
>docker cp C:/Users/HP/Desktop/TestBDA/data/WDI_TEST.csv
e00a03940e2d995f6a8d5805e500f5522155b8d80acb8189912eb28f28ef11e3:/tmp/
```

BASH INTO NAMENODE

```
docker exec -it  
e00a03940e2d995f6a8d5805e500f5522155b8d80acb8189912eb28f28ef11e3  
/bin/bash
```

CREATE A DIRECTORY IN HDFS

```
hdfs dfs -mkdir -p /user/root/input
```

CHECK ALL DIRECTORIES IN HDFS

```
hdfs dfs -ls
```

GO TO THE DIRECTORY IN HDFS

```
hdfs dfs -mkdir -p /user/root/NEW_TEST_BDA
```

COPY FROM NAMENODE TMP DIRECTORY TO THE DIRECTORY IN HDFS

```
hdfs dfs -copyFromLocal /tmp/WDI_TEST.csv /user/root/NEW_TEST_BDA
```

CHECK IF THE FILE WAS COPIED

```
hdfs dfs -ls
```

```
hdfs dfs -ls /user/root/NEW_TEST_BDA
```

BASH INTO THE HIVE SERVER

```
docker-compose exec hive-server bash
```

CONNECT TO BEELINE

```
/opt/hive/bin/beeline -u jdbc:hive2://localhost:10000
```

CREATE DATABASE

USE DATABASE

CREATE TESTING TABLES

```
create table test2( country_name string, indicator string, amount decimal)ROW
FORMAT DELIMITED
```

```
.....> FIELDS TERMINATED BY ','
.....> STORED AS TEXTFILE
.....> LOCATION '/user/root/NEW_TEST_BDA';
```

```
create table WDI_DATA( Country_name string, Country_Code string, Indicator
String, Indicator_code string , Year BIGINT, value decimal(40,10))
```

```
.....> ROW FORMAT DELIMITED
.....> FIELDS TERMINATED BY ','
.....> STORED AS TEXTFILE
.....> LOCATION '/user/root/NEW_TEST_BDA';
```

COPY THE CSV FROM THE HDFS DIRECTORY TO THE TEST TABLE

```
LOAD DATA INPATH '/user/root/NEW_TEST_BDA/WDI_TEST.csv' INTO TABLE
TEST;
```

CHECK IF THE RECORDS WERE INSERTED

```
SELECT * FROM TEST;
```

FINAL TABLES

```
create table WDI_FINAL(Country_Code string,Indicator_code string , Year BIGINT,
value decimal(38,10))
```

```
.....> ROW FORMAT DELIMITED
.....> FIELDS TERMINATED BY ','
.....> STORED AS TEXTFILE
.....> LOCATION '/user/root/NEW_TEST_BDA';
```

```
create table Countries(Country_code string, Country_Name string, )
```

```
.....> ROW FORMAT DELIMITED
.....> FIELDS TERMINATED BY ','
.....> STORED AS TEXTFILE
.....> LOCATION '/user/root/NEW_TEST_BDA';
```

```

create table Indicators(Indicator_code string, Indicator_name string)
.....> ROW FORMAT DELIMITED
.....> FIELDS TERMINATED BY ','
.....> STORED AS TEXTFILE
.....> LOCATION '/user/root/NEW_TEST_BDA';

```

```

create table Regions(Country_code string, Region string)
.....> ROW FORMAT DELIMITED
.....> FIELDS TERMINATED BY ','
.....> STORED AS TEXTFILE
.....> LOCATION '/user/root/NEW_TEST_BDA';

```

```
set hive.cli.print.header=200;
```

Databases

WDI - TABLE = TEST (3 ROWS) - HDFS DIRECTORY = NEW_TEST_BDA

WDI - TABLE = WDI_DATA (1.8 GB) - HDFS DIRECTORY = NEW_TEST_BDA

```

Administrator: Command Prompt
.....> ROW FORMAT DELIMITED
.....> FIELDS TERMINATED BY ','
.....> STORED AS TEXTFILE
.....> LOCATION '/user/root/NEW_TEST_BDA';
No rows affected (0.124 seconds)
0: jdbc:hive2://localhost:10000> Show tables;
+-----+
| tab_name |
+-----+
| test     |
| wdi_data |
+-----+
2 rows selected (0.036 seconds)
0: jdbc:hive2://localhost:10000>
D:\big.data\docker-hive> docker cp C:/Users/HP/Desktop/TestBDA/data/WDI_FINAL_TRANSFORMED.csv afc8bf2e48ae124169b241afd4d5f6408a5dad03a2ff23a6cb83d9790eaa54bb:/tmp/
Successfully copied 1.77GB to afc8bf2e48ae124169b241afd4d5f6408a5dad03a2ff23a6cb83d9790eaa54bb:/tmp/
D:\big.data\docker-hive>

```

RESULTS OF QUERIES

Query : Population, total SP.POP.TOTL

```

8 rows selected (1.325 seconds)
0: jdbc:hive2://localhost:10000> SELECT
. . . . .> wdi_final.country_code,
. . . . .> wdi_final.indicator_code,
. . . . .> SUM(wdi_final.value / 1000000) AS sum_value
. . . . .> FROM
. . . . .> wdi_final
. . . . .> WHERE
. . . . .> (wdi_final.year >= 2000 AND wdi_final.year <= 2022)
. . . . .> AND wdi_final.country_code IN ('USA', 'CHN', 'IND', 'PAK', 'BGD', 'KOR', 'JPN', 'ARE', 'RUS')
. . . . .> AND wdi_final.indicator_code = 'SP.POP.TOTL'
. . . . .> GROUP BY
. . . . .> wdi_final.country_code,
. . . . .> wdi_final.indicator_code
. . . . .> ORDER BY
. . . . .> sum_value DESC;
WARNING: Hive-on-MR is deprecated in Hive 2 and may not be available in the future versions. Consider using a different execution engine
+-----+-----+-----+
| wdi_final.country_code | wdi_final.indicator_code | sum_value |
+-----+-----+-----+
| CHN | SP.POP.TOTL | 29557.770000 |
| IND | SP.POP.TOTL | 27380.542051 |
| USA | SP.POP.TOTL | 6800.393419 |
| PAK | SP.POP.TOTL | 4276.147235 |
| BGD | SP.POP.TOTL | 3296.172196 |
| RUS | SP.POP.TOTL | 3167.178372 |
| JPN | SP.POP.TOTL | 2801.482593 |
| ARE | SP.POP.TOTL | 155.497329 |
+-----+-----+-----+

```

Conclusion:

Query : Population, female

SP.POP.TOTL.FE.IN

```

ERROR: ERROR WHILE COMPILING STATEMENT: FAILED: ParseException line 1:0 cannot recognize input near 'in' in 'this' modified (state=42000)
0: jdbc:hive2://localhost:10000> SELECT
. . . . .> wdi_final.country_code,
. . . . .> wdi_final.indicator_code,
. . . . .> SUM(wdi_final.value / 1000000) AS sum_value
. . . . .> FROM
. . . . .> wdi_final
. . . . .> WHERE
. . . . .> (wdi_final.year >= 2000 AND wdi_final.year <= 2022)
. . . . .> AND wdi_final.country_code IN ('USA', 'CHN', 'IND', 'PAK', 'BGD', 'KOR', 'JPN', 'ARE', 'RUS')
. . . . .> AND wdi_final.indicator_code = 'SP.POP.TOTL.FE.IN'
. . . . .> GROUP BY
. . . . .> wdi_final.country_code,
. . . . .> wdi_final.indicator_code
. . . . .> ORDER BY
. . . . .> sum_value DESC;
WARNING: Hive-on-MR is deprecated in Hive 2 and may not be available in the future versions. Consider using a different execution engine
+-----+-----+-----+
| wdi_final.country_code | wdi_final.indicator_code | sum_value |
+-----+-----+-----+
| CHN | SP.POP.TOTL.FE.IN | 14437.625341 |
| IND | SP.POP.TOTL.FE.IN | 13225.484117 |
| USA | SP.POP.TOTL.FE.IN | 3443.730773 |
| PAK | SP.POP.TOTL.FE.IN | 2085.632046 |
| RUS | SP.POP.TOTL.FE.IN | 1697.647000 |
| BGD | SP.POP.TOTL.FE.IN | 1644.151003 |
| JPN | SP.POP.TOTL.FE.IN | 1435.809662 |
| ARE | SP.POP.TOTL.FE.IN | 45.457716 |
+-----+-----+-----+
8 rows selected (2.444 seconds)
0: jdbc:hive2://localhost:10000>

```

Conclusion:

Query : Population, male

SP.POP.TOTL.MA.IN


```

8 rows selected (2.444 seconds)
0: jdbc:hive2://localhost:10000> SELECT
. . . . .> wdi_final.country_code,
. . . . .> wdi_final.indicator_code,
. . . . .> SUM(wdi_final.value / 1000000) AS sum_value
. . . . .> FROM
. . . . .> wdi_final
. . . . .> WHERE
. . . . .> (wdi_final.year >= 2000 AND wdi_final.year <= 2022)
. . . . .> AND wdi_final.country_code IN ('USA', 'CHN', 'IND', 'PAK', 'BGD', 'KOR', 'JPN', 'ARE', 'RUS')
. . . . .> AND wdi_final.indicator_code = 'SP.POP.TOTL.MA.IN'
. . . . .> GROUP BY
. . . . .> wdi_final.country_code,
. . . . .> wdi_final.indicator_code
. . . . .> ORDER BY
. . . . .> sum_value DESC;
WARNING: Hive-on-MR is deprecated in Hive 2 and may not be available in the future versions. Consider using a different execution engine
+-----+-----+-----+
| wdi_final.country_code | wdi_final.indicator_code | sum_value |
+-----+-----+-----+
| CHN                    | SP.POP.TOTL.MA.IN       | 15120.144663 |
| IND                    | SP.POP.TOTL.MA.IN       | 14155.057930 |
| USA                    | SP.POP.TOTL.MA.IN       | 3356.662648 |
| PAK                    | SP.POP.TOTL.MA.IN       | 2190.515186 |
| BGD                    | SP.POP.TOTL.MA.IN       | 1652.021194 |
| RUS                    | SP.POP.TOTL.MA.IN       | 1469.531369 |
| JPN                    | SP.POP.TOTL.MA.IN       | 1365.672933 |
| ARE                    | SP.POP.TOTL.MA.IN       | 110.039613 |
+-----+-----+-----+
8 rows selected (2.431 seconds)

```

Conclusion:

Query : Birth rate, crude (per 1,000 people) SP.DYN.CBRT.IN

Conclusion:

```

24 rows selected (1.257 seconds)
0: jdbc:hive2://localhost:10000> SELECT
. . . . .> wdi_final.country_code,
. . . . .> wdi_final.indicator_code,
. . . . .> SUM(wdi_final.value / 1000000) AS sum_value
. . . . .> FROM
. . . . .> wdi_final
. . . . .> WHERE
. . . . .> (wdi_final.year >= 2000 AND wdi_final.year <= 2022)
. . . . .> AND wdi_final.country_code IN ('USA', 'CHN', 'IND', 'PAK', 'BGD', 'KOR', 'JPN', 'ARE', 'RUS')
. . . . .> AND wdi_final.indicator_code = 'SP.DYN.CBRT.IN'
. . . . .> GROUP BY
. . . . .> wdi_final.country_code,
. . . . .> wdi_final.indicator_code
. . . . .> ORDER BY
. . . . .> sum_value DESC;
WARNING: Hive-on-MR is deprecated in Hive 2 and may not be available in the future versions. Consider using a different execution engine
+-----+-----+-----+
| wdi_final.country_code | wdi_final.indicator_code | sum_value |
+-----+-----+-----+
| PAK                    | SP.DYN.CBRT.IN          | 0.000694 |
| BGD                    | SP.DYN.CBRT.IN          | 0.000489 |
| IND                    | SP.DYN.CBRT.IN          | 0.000469 |
| ARE                    | SP.DYN.CBRT.IN          | 0.000291 |
| USA                    | SP.DYN.CBRT.IN          | 0.000285 |
| CHN                    | SP.DYN.CBRT.IN          | 0.000268 |
| RUS                    | SP.DYN.CBRT.IN          | 0.000247 |
| JPN                    | SP.DYN.CBRT.IN          | 0.000183 |
+-----+-----+-----+

```

Query : Death rate, crude (per 1,000 people) SP.DYN.CDRT.IN

Conclusion:

```
0: jdbc:hive2://localhost:10000> SELECT
. . . . .> wdi_final.country_code,
. . . . .> wdi_final.indicator_code,
. . . . .> SUM(wdi_final.value / 1000000) AS sum_value
. . . . .> FROM
. . . . .> wdi_final
. . . . .> WHERE
. . . . .> (wdi_final.year >= 2000 AND wdi_final.year <= 2022)
. . . . .> AND wdi_final.country_code IN ('USA', 'CHN', 'IND', 'PAK', 'BGD', 'KOR', 'JPN', 'ARE', 'RUS')
. . . . .> AND wdi_final.indicator_code = 'SP.DYN.CDRT.IN'
. . . . .> GROUP BY
. . . . .> wdi_final.country_code,
. . . . .> wdi_final.indicator_code
. . . . .> ORDER BY
. . . . .> sum_value DESC;
WARNING: Hive-on-MR is deprecated in Hive 2 and may not be available in the future versions. Consider using a different execution engi
```

wdi_final.country_code	wdi_final.indicator_code	sum_value
RUS	SP.DYN.CDRT.IN	0.000316
JPN	SP.DYN.CDRT.IN	0.000212
USA	SP.DYN.CDRT.IN	0.000186
PAK	SP.DYN.CDRT.IN	0.000168
IND	SP.DYN.CDRT.IN	0.000167
CHN	SP.DYN.CDRT.IN	0.000149
BGD	SP.DYN.CDRT.IN	0.000136
ARE	SP.DYN.CDRT.IN	0.000030

Query : Life expectancy at birth, total (years) SP.DYN.LE00.IN

Conclusion:

```
8 rows selected (2.683 seconds)
0: jdbc:hive2://localhost:10000> SELECT
. . . . .> wdi_final.country_code,
. . . . .> wdi_final.indicator_code,
. . . . .> SUM(wdi_final.value) AS sum_value
. . . . .> FROM
. . . . .> wdi_final
. . . . .> WHERE
. . . . .> (wdi_final.year >= 2000 AND wdi_final.year <= 2022)
. . . . .> AND wdi_final.country_code IN ('USA', 'CHN', 'IND', 'PAK', 'BGD', 'KOR', 'JPN', 'ARE', 'RUS')
. . . . .> AND wdi_final.indicator_code = 'SP.DYN.LE00.IN'
. . . . .> GROUP BY
. . . . .> wdi_final.country_code,
. . . . .> wdi_final.indicator_code
. . . . .> ORDER BY
. . . . .> sum_value DESC;
WARNING: Hive-on-MR is deprecated in Hive 2 and may not be available in the future versions. Consider using a different execution
```

wdi_final.country_code	wdi_final.indicator_code	sum_value
JPN	SP.DYN.LE00.IN	1825.17779
USA	SP.DYN.LE00.IN	1714.53898
ARE	SP.DYN.LE00.IN	1711.33900
CHN	SP.DYN.LE00.IN	1662.48900
BGD	SP.DYN.LE00.IN	1517.88800
RUS	SP.DYN.LE00.IN	1514.62439
IND	SP.DYN.LE00.IN	1475.77100
PAK	SP.DYN.LE00.IN	1419.04800

Query : Population growth (annual %) SP.POP.GROW

Conclusion:

```

8 rows selected (2.659 seconds)
0: jdbc:hive2://localhost:10000> SELECT
. . . . .> wdi_final.country_code,
. . . . .> wdi_final.indicator_code,
. . . . .> SUM(wdi_final.value) AS sum_value
. . . . .> FROM
. . . . .> wdi_final
. . . . .> WHERE
. . . . .> (wdi_final.year >= 2000 AND wdi_final.year <= 2022)
. . . . .> AND wdi_final.country_code IN ('USA', 'CHN', 'IND', 'PAK', 'BGD', 'KOR', 'JPN', 'ARE', 'RUS')
. . . . .> AND wdi_final.indicator_code = 'SP.POP.GROW'
. . . . .> GROUP BY
. . . . .> wdi_final.country_code,
. . . . .> wdi_final.indicator_code
. . . . .> ORDER BY
. . . . .> sum_value DESC;
WARNING: Hive-on-MR is deprecated in Hive 2 and may not be available in the future versions. Consider using a different execution engine
+-----+-----+-----+
| wdi_final.country_code | wdi_final.indicator_code | sum_value |
+-----+-----+-----+
| ARE | SP.POP.GROW | 110.63791 |
| PAK | SP.POP.GROW | 43.55606 |
| IND | SP.POP.GROW | 30.21591 |
| BGD | SP.POP.GROW | 28.97498 |
| USA | SP.POP.GROW | 17.34598 |
| CHN | SP.POP.GROW | 11.99329 |
| JPN | SP.POP.GROW | -0.75254 |
| RUS | SP.POP.GROW | -2.59108 |
+-----+-----+-----+
8 rows selected (2.659 seconds)

```

Query : CPIA gender equality rating (1=low to 6=high) IQ.CPA.GNDR.XQ

```

8 rows selected (2.659 seconds)
0: jdbc:hive2://localhost:10000> SELECT
. . . . .> wdi_final.country_code,
. . . . .> wdi_final.indicator_code,
. . . . .> SUM(wdi_final.value) AS sum_value
. . . . .> FROM
. . . . .> wdi_final
. . . . .> WHERE
. . . . .> (wdi_final.year >= 2000 AND wdi_final.year <= 2022)
. . . . .> AND wdi_final.country_code IN ('USA', 'CHN', 'IND', 'PAK', 'BGD', 'KOR', 'JPN', 'ARE', 'RUS')
. . . . .> AND wdi_final.indicator_code = 'IQ.CPA.GNDR.XQ'
. . . . .> GROUP BY
. . . . .> wdi_final.country_code,
. . . . .> wdi_final.indicator_code
. . . . .> ORDER BY
. . . . .> sum_value DESC;
WARNING: Hive-on-MR is deprecated in Hive 2 and may not be available in the future versions. Consider using a different execution engine
+-----+-----+-----+
| wdi_final.country_code | wdi_final.indicator_code | sum_value |
+-----+-----+-----+
| USA | IQ.CPA.GNDR.XQ | 65203565949323.21190 |
| RUS | IQ.CPA.GNDR.XQ | 65203565949323.21190 |
| JPN | IQ.CPA.GNDR.XQ | 65203565949323.21190 |
| CHN | IQ.CPA.GNDR.XQ | 65203565949323.21190 |
| ARE | IQ.CPA.GNDR.XQ | 65203565949323.21190 |
| IND | IQ.CPA.GNDR.XQ | 46239437514208.72510 |
| BGD | IQ.CPA.GNDR.XQ | 5645750368614.35990 |
| PAK | IQ.CPA.GNDR.XQ | 5645750368593.85990 |
+-----+-----+-----+
8 rows selected (2.538 seconds)
0: jdbc:hive2://localhost:10000>

```

Conclusion:

Query : Current health expenditure (% of GDP) SH.XPD.CHEX.GD.ZS

```
0: jdbc:hive2://localhost:10000> SELECT
. . . . .> wdi_final.country_code,
. . . . .> wdi_final.indicator_code,
. . . . .> SUM(wdi_final.value) AS sum_value
. . . . .> FROM
. . . . .> wdi_final
. . . . .> WHERE
. . . . .> (wdi_final.year >= 2000 AND wdi_final.year <= 2022)
. . . . .> AND wdi_final.country_code IN ('USA', 'CHN', 'IND', 'PAK', 'BGD', 'KOR', 'JPN', 'ARE', 'RUS')
. . . . .> AND wdi_final.indicator_code = 'SH.XPD.CHEX.GD.ZS'
. . . . .> GROUP BY
. . . . .> wdi_final.country_code,
. . . . .> wdi_final.indicator_code
. . . . .> ORDER BY
. . . . .> sum_value DESC;
WARNING: Hive-on-MR is deprecated in Hive 2 and may not be available in the future versions. Consider using a different execution engine

+-----+-----+-----+
| wdi_final.country_code | wdi_final.indicator_code | sum_value |
+-----+-----+-----+
| USA | SH.XPD.CHEX.GD.ZS | 9594000591930.60251 |
| JPN | SH.XPD.CHEX.GD.ZS | 9594000591796.42296 |
| RUS | SH.XPD.CHEX.GD.ZS | 9594000591713.56243 |
| CHN | SH.XPD.CHEX.GD.ZS | 9594000591699.14112 |
| IND | SH.XPD.CHEX.GD.ZS | 9594000591677.97126 |
| ARE | SH.XPD.CHEX.GD.ZS | 9594000591673.80759 |
| BGD | SH.XPD.CHEX.GD.ZS | 9594000591657.67769 |
| PAK | SH.XPD.CHEX.GD.ZS | 9594000591657.63435 |
+-----+-----+-----+
```

Conclusion:

Query : Employment to population ratio, ages 15-24, total (%) (national estimate)
SL.EMP.1524.SP.NE.ZS

```
8 rows selected (2.665 seconds)
0: jdbc:hive2://localhost:10000> SELECT
. . . . .> wdi_final.country_code,
. . . . .> wdi_final.indicator_code,
. . . . .> SUM(wdi_final.value) AS sum_value
. . . . .> FROM
. . . . .> wdi_final
. . . . .> WHERE
. . . . .> (wdi_final.year >= 2000 AND wdi_final.year <= 2022)
. . . . .> AND wdi_final.country_code IN ('USA', 'CHN', 'IND', 'PAK', 'BGD', 'KOR', 'JPN', 'ARE', 'RUS')
. . . . .> AND wdi_final.indicator_code = 'SL.EMP.1524.SP.NE.ZS'
. . . . .> GROUP BY
. . . . .> wdi_final.country_code,
. . . . .> wdi_final.indicator_code
. . . . .> ORDER BY
. . . . .> sum_value DESC;
WARNING: Hive-on-MR is deprecated in Hive 2 and may not be available in the future versions. Consider using a different execution engine

+-----+-----+-----+
| wdi_final.country_code | wdi_final.indicator_code | sum_value |
+-----+-----+-----+
| CHN | SL.EMP.1524.SP.NE.ZS | 65203565949323.21190 |
| BGD | SL.EMP.1524.SP.NE.ZS | 51045850705256.66430 |
| ARE | SL.EMP.1524.SP.NE.ZS | 34537423824301.60520 |
| IND | SL.EMP.1524.SP.NE.ZS | 31244240417998.51830 |
| PAK | SL.EMP.1524.SP.NE.ZS | 22869154866385.98120 |
| USA | SL.EMP.1524.SP.NE.ZS | 1116.78000 |
| JPN | SL.EMP.1524.SP.NE.ZS | 916.15000 |
| RUS | SL.EMP.1524.SP.NE.ZS | 709.89000 |
+-----+-----+-----+
8 rows selected (2.516 seconds)
0: jdbc:hive2://localhost:10000>
```

Conclusion:

Query : Literacy rate, adult total (% of people ages 15 and above)
SE.ADT.LITR.ZS

```

0: jdbc:hive2://localhost:10000> SELECT
. . . . .> wdi_final.country_code,
. . . . .> wdi_final.indicator_code,
. . . . .> SUM(wdi_final.value) AS sum_value
. . . . .> FROM
. . . . .> wdi_final
. . . . .> WHERE
. . . . .> (wdi_final.year >= 2000 AND wdi_final.year <= 2022)
. . . . .> AND wdi_final.country_code IN ('USA', 'CHN', 'IND', 'PAK', 'BGD', 'KOR', 'JPN', 'ARE', 'RUS')
. . . . .> AND wdi_final.indicator_code = 'SE.ADT.LITR.ZS'
. . . . .> GROUP BY
. . . . .> wdi_final.country_code,
. . . . .> wdi_final.indicator_code
. . . . .> ORDER BY
. . . . .> sum_value DESC;
WARNING: Hive-on-MR is deprecated in Hive 2 and may not be available in the future versions. Consider using a different execution engine (i.e.
+-----+-----+-----+
| wdi_final.country_code | wdi_final.indicator_code | sum_value |
+-----+-----+-----+
| USA | SE.ADT.LITR.ZS | 65203565949323.21190 |
| JPN | SE.ADT.LITR.ZS | 65203565949323.21190 |
| IND | SE.ADT.LITR.ZS | 55805407590522.09718 |
| CHN | SE.ADT.LITR.ZS | 51950835734446.75848 |
| RUS | SE.ADT.LITR.ZS | 51677202351565.19760 |
| ARE | SE.ADT.LITR.ZS | 47145817084815.41999 |
| PAK | SE.ADT.LITR.ZS | 27816622914549.69876 |
| BGD | SE.ADT.LITR.ZS | 23226820150746.98275 |
+-----+-----+-----+
8 rows selected (2.636 seconds)

```

Conclusion:

Query : Access to electricity (% of population)

EG.ELC.ACCS.ZS

```

0: jdbc:hive2://localhost:10000> SELECT
. . . . .> wdi_final.country_code,
. . . . .> wdi_final.indicator_code,
. . . . .> SUM(wdi_final.value) AS sum_value
. . . . .> FROM
. . . . .> wdi_final
. . . . .> WHERE
. . . . .> (wdi_final.year >= 2000 AND wdi_final.year <= 2022)
. . . . .> AND wdi_final.country_code IN ('USA', 'CHN', 'IND', 'PAK', 'BGD', 'KOR', 'JPN', 'ARE', 'RUS')
. . . . .> AND wdi_final.indicator_code = 'EG.ELC.ACCS.ZS'
. . . . .> GROUP BY
. . . . .> wdi_final.country_code,
. . . . .> wdi_final.indicator_code
. . . . .> ORDER BY
. . . . .> sum_value DESC;
WARNING: Hive-on-MR is deprecated in Hive 2 and may not be available in the future versions. Consider using a different execution eng
+-----+-----+-----+
| wdi_final.country_code | wdi_final.indicator_code | sum_value |
+-----+-----+-----+
| USA | EG.ELC.ACCS.ZS | 9594000593703.51000 |
| JPN | EG.ELC.ACCS.ZS | 9594000593703.51000 |
| ARE | EG.ELC.ACCS.ZS | 9594000593703.51000 |
| RUS | EG.ELC.ACCS.ZS | 9594000593699.01000 |
| CHN | EG.ELC.ACCS.ZS | 9594000593684.11909 |
| IND | EG.ELC.ACCS.ZS | 9594000593214.30087 |
| PAK | EG.ELC.ACCS.ZS | 9594000593099.39778 |
| BGD | EG.ELC.ACCS.ZS | 9594000592867.15676 |
+-----+-----+-----+
8 rows selected (2.488 seconds)

```

Conclusion:

Query : Unemployment, total (% of total labor force) (national estimate)

SL.UEM.TOTL.NE.ZS

```
0: jdbc:hive2://localhost:10000> SELECT
. . . . .> wdi_final.country_code,
. . . . .> wdi_final.indicator_code,
. . . . .> SUM(wdi_final.value) AS sum_value
. . . . .> FROM
. . . . .> wdi_final
. . . . .> WHERE
. . . . .> (wdi_final.year >= 2000 AND wdi_final.year <= 2022)
. . . . .> AND wdi_final.country_code IN ('USA', 'CHN', 'IND', 'PAK', 'BGD', 'KOR', 'JPN', 'ARE', 'RUS')
. . . . .> AND wdi_final.indicator_code = 'SL.UEM.TOTL.NE.ZS'
. . . . .> GROUP BY
. . . . .> wdi_final.country_code,
. . . . .> wdi_final.indicator_code
. . . . .> ORDER BY
. . . . .> sum_value DESC;
WARNING: Hive-on-MR is deprecated in Hive 2 and may not be available in the future versions. Consider using a different execution engine (i.e.
+-----+-----+-----+
| wdi_final.country_code | wdi_final.indicator_code | sum_value |
+-----+-----+-----+
| BGD | SL.UEM.TOTL.NE.ZS | 45606979217361.95130 |
| IND | SL.UEM.TOTL.NE.ZS | 31244240417818.45830 |
| ARE | SL.UEM.TOTL.NE.ZS | 28681017481662.35860 |
| PAK | SL.UEM.TOTL.NE.ZS | 15874415470369.47670 |
| CHN | SL.UEM.TOTL.NE.ZS | 6798114323204.98200 |
| RUS | SL.UEM.TOTL.NE.ZS | 144.07000 |
| USA | SL.UEM.TOTL.NE.ZS | 131.04000 |
| JPN | SL.UEM.TOTL.NE.ZS | 88.18000 |
+-----+-----+-----+
8 rows selected (2.545 seconds)
0: jdbc:hive2://localhost:10000>
```

Conclusion:

Query : Unemployment, youth total (% of total labor force ages 15-24) (national estimate) SL.UEM.1524.NE.ZS

```
8 rows selected (2.545 seconds)
0: jdbc:hive2://localhost:10000> SELECT
. . . . .> wdi_final.country_code,
. . . . .> wdi_final.indicator_code,
. . . . .> SUM(wdi_final.value) AS sum_value
. . . . .> FROM
. . . . .> wdi_final
. . . . .> WHERE
. . . . .> (wdi_final.year >= 2000 AND wdi_final.year <= 2022)
. . . . .> AND wdi_final.country_code IN ('USA', 'CHN', 'IND', 'PAK', 'BGD', 'KOR', 'JPN', 'ARE', 'RUS')
. . . . .> AND wdi_final.indicator_code = 'SL.UEM.1524.NE.ZS'
. . . . .> GROUP BY
. . . . .> wdi_final.country_code,
. . . . .> wdi_final.indicator_code
. . . . .> ORDER BY
. . . . .> sum_value DESC;
WARNING: Hive-on-MR is deprecated in Hive 2 and may not be available in the future versions. Consider using a different execution engine (i.e.
+-----+-----+-----+
| wdi_final.country_code | wdi_final.indicator_code | sum_value |
+-----+-----+-----+
| BGD | SL.UEM.1524.NE.ZS | 47607285268225.05570 |
| CHN | SL.UEM.1524.NE.ZS | 38399942916186.23090 |
| ARE | SL.UEM.1524.NE.ZS | 33188434797074.49060 |
| IND | SL.UEM.1524.NE.ZS | 31244240417908.96830 |
| PAK | SL.UEM.1524.NE.ZS | 19427460292256.15010 |
| RUS | SL.UEM.1524.NE.ZS | 357.17000 |
| USA | SL.UEM.1524.NE.ZS | 272.82000 |
| JPN | SL.UEM.1524.NE.ZS | 160.91000 |
+-----+-----+-----+
8 rows selected (2.557 seconds)
0: jdbc:hive2://localhost:10000>
```

Conclusion:

Query : Adjusted net national income (annual % growth) NY.ADJ.NNTY.KD.ZG

```

8 rows selected (2.557 seconds)
0: jdbc:hive2://localhost:10000> SELECT
. . . . .> wdi_final.country_code,
. . . . .> wdi_final.indicator_code,
. . . . .> SUM(wdi_final.value) AS sum_value
. . . . .> FROM
. . . . .> wdi_final
. . . . .> WHERE
. . . . .> (wdi_final.year >= 2000 AND wdi_final.year <= 2022)
. . . . .> AND wdi_final.country_code IN ('USA', 'CHN', 'IND', 'PAK', 'BGD', 'KOR', 'JPN', 'ARE', 'RUS')
. . . . .> AND wdi_final.indicator_code = 'NY.ADJ.NNTY.KD.ZG'
. . . . .> GROUP BY
. . . . .> wdi_final.country_code,
. . . . .> wdi_final.indicator_code
. . . . .> ORDER BY
. . . . .> sum_value DESC;
WARNING: Hive-on-MR is deprecated in Hive 2 and may not be available in the future versions. Consider using a different execution engine

+-----+-----+-----+
| wdi_final.country_code | wdi_final.indicator_code | sum_value |
+-----+-----+-----+
| ARE | NY.ADJ.NNTY.KD.ZG | 11554730015007.08638 |
| CHN | NY.ADJ.NNTY.KD.ZG | 9594000591790.11317 |
| IND | NY.ADJ.NNTY.KD.ZG | 9594000591733.29884 |
| BGD | NY.ADJ.NNTY.KD.ZG | 9594000591721.73077 |
| RUS | NY.ADJ.NNTY.KD.ZG | 9594000591699.94028 |
| PAK | NY.ADJ.NNTY.KD.ZG | 9594000591672.64126 |
| USA | NY.ADJ.NNTY.KD.ZG | 9594000591641.62392 |
| JPN | NY.ADJ.NNTY.KD.ZG | 9594000591613.98828 |
+-----+-----+-----+

```

Conclusion:

Query : Broad money (% of GDP)

FM.LBL.BMNY.GD.ZS

```

8 rows selected (2.716 seconds)
0: jdbc:hive2://localhost:10000> SELECT
. . . . .> wdi_final.country_code,
. . . . .> wdi_final.indicator_code,
. . . . .> SUM(wdi_final.value/1000) AS sum_value
. . . . .> FROM
. . . . .> wdi_final
. . . . .> WHERE
. . . . .> (wdi_final.year >= 2000 AND wdi_final.year <= 2022)
. . . . .> AND wdi_final.country_code IN ('USA', 'CHN', 'IND', 'PAK', 'BGD', 'KOR', 'JPN', 'ARE', 'RUS')
. . . . .> AND wdi_final.indicator_code = 'FM.LBL.BMNY.GD.ZS'
. . . . .> GROUP BY
. . . . .> wdi_final.country_code,
. . . . .> wdi_final.indicator_code
. . . . .> ORDER BY
. . . . .> sum_value DESC;
WARNING: Hive-on-MR is deprecated in Hive 2 and may not be available in the future versions. Consider using a different execution engine

+-----+-----+-----+
| wdi_final.country_code | wdi_final.indicator_code | sum_value |
+-----+-----+-----+
| RUS | FM.LBL.BMNY.GD.ZS | 10517125611.727054 |
| JPN | FM.LBL.BMNY.GD.ZS | 4.926013 |
| CHN | FM.LBL.BMNY.GD.ZS | 3.837134 |
| USA | FM.LBL.BMNY.GD.ZS | 1.875283 |
| IND | FM.LBL.BMNY.GD.ZS | 1.602345 |
| ARE | FM.LBL.BMNY.GD.ZS | 1.496495 |
| BGD | FM.LBL.BMNY.GD.ZS | 1.160877 |
| PAK | FM.LBL.BMNY.GD.ZS | 1.086921 |
+-----+-----+-----+
8 rows selected (2.596 seconds)

```

Conclusion:

Query : Central government debt, total (% of GDP)

GC.DOD.TOTL.GD.ZS

```

8 rows selected (2.596 seconds)
0: jdbc:hive2://localhost:10000> SELECT
. . . . .> wdi_final.country_code,
. . . . .> wdi_final.indicator_code,
. . . . .> SUM(wdi_final.value/1000) AS sum_value
. . . . .> FROM
. . . . .> wdi_final
. . . . .> WHERE
. . . . .> (wdi_final.year >= 2000 AND wdi_final.year <= 2022)
. . . . .> AND wdi_final.country_code IN ('USA', 'CHN', 'IND', 'PAK', 'BGD', 'KOR', 'JPN', 'ARE', 'RUS')
. . . . .> AND wdi_final.indicator_code = 'GC.DOD.TOTL.GD.ZS'
. . . . .> GROUP BY
. . . . .> wdi_final.country_code,
. . . . .> wdi_final.indicator_code
. . . . .> ORDER BY
. . . . .> sum_value DESC;
WARNING: Hive-on-MR is deprecated in Hive 2 and may not be available in the future versions. Consider using a different execution engine

wdi_final.country_code | wdi_final.indicator_code | sum_value |
+-----+-----+-----+
| CHN | GC.DOD.TOTL.GD.ZS | 65203565949.323214 |
| PAK | GC.DOD.TOTL.GD.ZS | 64280440930.201145 |
| ARE | GC.DOD.TOTL.GD.ZS | 62225945603.470754 |
| BGD | GC.DOD.TOTL.GD.ZS | 61802751752.523122 |
| IND | GC.DOD.TOTL.GD.ZS | 22487517469.320833 |
| RUS | GC.DOD.TOTL.GD.ZS | 2488262543.496102 |
| JPN | GC.DOD.TOTL.GD.ZS | 3.559619 |
| USA | GC.DOD.TOTL.GD.ZS | 1.758366 |
+-----+-----+-----+

```

Conclusion:

Query : Central government debt, total (current LCU) GC.DOD.TOTL.CN

```

8 rows selected (2.699 seconds)
0: jdbc:hive2://localhost:10000> SELECT
. . . . .> wdi_final.country_code,
. . . . .> wdi_final.indicator_code,
. . . . .> SUM(wdi_final.value/1000) AS sum_value
. . . . .> FROM
. . . . .> wdi_final
. . . . .> WHERE
. . . . .> (wdi_final.year >= 2000 AND wdi_final.year <= 2022)
. . . . .> AND wdi_final.country_code IN ('USA', 'CHN', 'IND', 'PAK', 'BGD', 'KOR', 'JPN', 'ARE', 'RUS')
. . . . .> AND wdi_final.indicator_code = 'GC.DOD.TOTL.CN'
. . . . .> GROUP BY
. . . . .> wdi_final.country_code,
. . . . .> wdi_final.indicator_code
. . . . .> ORDER BY
. . . . .> sum_value DESC;
WARNING: Hive-on-MR is deprecated in Hive 2 and may not be available in the future versions. Consider using a different execution engine (i.e. spark)

wdi_final.country_code | wdi_final.indicator_code | sum_value |
+-----+-----+-----+
| JPN | GC.DOD.TOTL.CN | 18849000000000.000000 |
| IND | GC.DOD.TOTL.CN | 808687517468.281789 |
| USA | GC.DOD.TOTL.CN | 298230000000.000000 |
| RUS | GC.DOD.TOTL.CN | 188578262543.125929 |
| PAK | GC.DOD.TOTL.CN | 67110440930.134392 |
| CHN | GC.DOD.TOTL.CN | 65203565949.323214 |
| BGD | GC.DOD.TOTL.CN | 64800751752.428592 |
| ARE | GC.DOD.TOTL.CN | 62253073530.684908 |
+-----+-----+-----+
8 rows selected (2.544 seconds)

```

Conclusion:

Query : Current account balance (% of GDP) BN.CAB.XOKA.GD.ZS


```

8 rows selected (2.544 seconds)
0: jdbc:hive2://localhost:10000> SELECT
      wdi_final.country_code,
      wdi_final.indicator_code,
      SUM(wdi_final.value/1000) AS sum_value
    FROM
      wdi_final
    WHERE
      (wdi_final.year >= 2000 AND wdi_final.year <= 2022)
      AND wdi_final.country_code IN ('USA', 'CHN', 'IND', 'PAK', 'BGD', 'KOR', 'JPN', 'ARE', 'RUS')
      AND wdi_final.indicator_code = 'BN.CAB.XOKA.GD.ZS'
    GROUP BY
      wdi_final.country_code,
      wdi_final.indicator_code
    ORDER BY
      sum_value DESC;
WARNING: Hive-on-MR is deprecated in Hive 2 and may not be available in the future versions. Consider using a different execution engine

```

wdi_final.country_code	wdi_final.indicator_code	sum_value
ARE	BN.CAB.XOKA.GD.ZS	65203565949.323214
RUS	BN.CAB.XOKA.GD.ZS	0.135484
CHN	BN.CAB.XOKA.GD.ZS	0.071904
JPN	BN.CAB.XOKA.GD.ZS	0.065017
BGD	BN.CAB.XOKA.GD.ZS	0.005110
IND	BN.CAB.XOKA.GD.ZS	-0.027219
PAK	BN.CAB.XOKA.GD.ZS	-0.040775
USA	BN.CAB.XOKA.GD.ZS	-0.075168

```

8 rows selected (2.53 seconds)

```

Conclusion:

Query : Expense (% of GDP)

GC.XPN.TOTL.GD.ZS

```

0: jdbc:hive2://localhost:10000> SELECT
      wdi_final.country_code,
      wdi_final.indicator_code,
      SUM(wdi_final.value/1000) AS sum_value
    FROM
      wdi_final
    WHERE
      (wdi_final.year >= 2000 AND wdi_final.year <= 2022)
      AND wdi_final.country_code IN ('USA', 'CHN', 'IND', 'PAK', 'BGD', 'KOR', 'JPN', 'ARE', 'RUS')
      AND wdi_final.indicator_code = 'GC.XPN.TOTL.GD.ZS'
    GROUP BY
      wdi_final.country_code,
      wdi_final.indicator_code
    ORDER BY
      sum_value DESC;
WARNING: Hive-on-MR is deprecated in Hive 2 and may not be available in the future versions. Consider using a different execution engine

```

wdi_final.country_code	wdi_final.indicator_code	sum_value
CHN	GC.XPN.TOTL.GD.ZS	65203565949.323214
PAK	GC.XPN.TOTL.GD.ZS	64280440930.149889
IND	GC.XPN.TOTL.GD.ZS	22487517468.579489
ARE	GC.XPN.TOTL.GD.ZS	16437505125.935863
BGD	GC.XPN.TOTL.GD.ZS	923125019.368952
RUS	GC.XPN.TOTL.GD.ZS	0.571152
USA	GC.XPN.TOTL.GD.ZS	0.506122
JPN	GC.XPN.TOTL.GD.ZS	0.375335

```

8 rows selected (2.584 seconds)
0: jdbc:hive2://localhost:10000>

```

Conclusion:

Query : Expense (current LCU)

GC.XPN.TOTL.CN

```

8 rows selected (2.584 seconds)
0: jdbc:hive2://localhost:10000> SELECT
    .>     wdi_final.country_code,
    .>     wdi_final.indicator_code,
    .>     SUM(wdi_final.value) AS sum_value
    .> FROM
    .>     wdi_final
    .> WHERE
    .>     (wdi_final.year >= 2000 AND wdi_final.year <= 2022)
    .>     AND wdi_final.country_code IN ('USA', 'CHN', 'IND', 'PAK', 'BGD', 'KOR', 'JPN', 'ARE', 'RUS')
    .>     AND wdi_final.indicator_code = 'GC.XPN.TOTL.CN'
    .> GROUP BY
    .>     wdi_final.country_code,
    .>     wdi_final.indicator_code
    .> ORDER BY
    .>     sum_value DESC;
WARNING: Hive-on-MR is deprecated in Hive 2 and may not be available in the future versions. Consider using a different execution eng
+-----+-----+-----+
| wdi_final.country_code | wdi_final.indicator_code | sum_value |
+-----+-----+-----+
| JPN                    | GC.XPN.TOTL.CN          | 198310000000000.00000 |
| RUS                    | GC.XPN.TOTL.CN          | 355840000000000.00000 |
| IND                    | GC.XPN.TOTL.CN          | 261397517468281.78900 |
| USA                    | GC.XPN.TOTL.CN          | 824200000000000.00000 |
| CHN                    | GC.XPN.TOTL.CN          | 65203565949323.21190 |
| PAK                    | GC.XPN.TOTL.CN          | 64938440930134.38990 |
| BGD                    | GC.XPN.TOTL.CN          | 24280125019188.82200 |
| ARE                    | GC.XPN.TOTL.CN          | 17097223051949.64070 |
+-----+-----+-----+
8 rows selected (2.654 seconds)

```

Conclusion:

Query : Fossil fuel energy consumption (% of total) EG.USE.COMM.FO.ZS

```

8 rows selected (2.654 seconds)
0: jdbc:hive2://localhost:10000> SELECT
    .>     wdi_final.country_code,
    .>     wdi_final.indicator_code,
    .>     SUM(wdi_final.value) AS sum_value
    .> FROM
    .>     wdi_final
    .> WHERE
    .>     (wdi_final.year >= 2000 AND wdi_final.year <= 2022)
    .>     AND wdi_final.country_code IN ('USA', 'CHN', 'IND', 'PAK', 'BGD', 'KOR', 'JPN', 'ARE', 'RUS')
    .>     AND wdi_final.indicator_code = 'EG.USE.COMM.FO.ZS'
    .> GROUP BY
    .>     wdi_final.country_code,
    .>     wdi_final.indicator_code
    .> ORDER BY
    .>     sum_value DESC;
WARNING: Hive-on-MR is deprecated in Hive 2 and may not be available in the future versions. Consider using a different execution eng
+-----+-----+-----+
| wdi_final.country_code | wdi_final.indicator_code | sum_value |
+-----+-----+-----+
| ARE                    | EG.USE.COMM.FO.ZS       | 41699304501057.25450 |
| RUS                    | EG.USE.COMM.FO.ZS       | 37464256449722.51834 |
| CHN                    | EG.USE.COMM.FO.ZS       | 37464256449646.93512 |
| IND                    | EG.USE.COMM.FO.ZS       | 37464256449385.01171 |
| BGD                    | EG.USE.COMM.FO.ZS       | 37464256449359.41728 |
| PAK                    | EG.USE.COMM.FO.ZS       | 37464256449259.88615 |
| JPN                    | EG.USE.COMM.FO.ZS       | 34250434422800.43702 |
| USA                    | EG.USE.COMM.FO.ZS       | 34250434422788.97668 |
+-----+-----+-----+
8 rows selected (2.477 seconds)

```

Conclusion:

Query : GDP growth (annual %) NY.GDP.MKTP.KD.ZG

```

8 rows selected (2.477 seconds)
0: jdbc:hive2://localhost:10000> SELECT
. . . . .> wdi_final.country_code,
. . . . .> wdi_final.indicator_code,
. . . . .> SUM(wdi_final.value) AS sum_value
. . . . .> FROM
. . . . .> wdi_final
. . . . .> WHERE
. . . . .> (wdi_final.year >= 2000 AND wdi_final.year <= 2022)
. . . . .> AND wdi_final.country_code IN ('USA', 'CHN', 'IND', 'PAK', 'BGD', 'KOR', 'JPN', 'ARE', 'RUS')
. . . . .> AND wdi_final.indicator_code = 'NY.GDP.MKTP.KD.ZG'
. . . . .> GROUP BY
. . . . .> wdi_final.country_code,
. . . . .> wdi_final.indicator_code
. . . . .> ORDER BY
. . . . .> sum_value DESC;
WARNING: Hive-on-MR is deprecated in Hive 2 and may not be available in the future versions. Consider using a different execution engine
+-----+-----+-----+
| wdi_final.country_code | wdi_final.indicator_code | sum_value |
+-----+-----+-----+
| CHN | NY.GDP.MKTP.KD.ZG | 190.63632 |
| BGD | NY.GDP.MKTP.KD.ZG | 131.98485 |
| IND | NY.GDP.MKTP.KD.ZG | 131.16232 |
| PAK | NY.GDP.MKTP.KD.ZG | 90.91655 |
| ARE | NY.GDP.MKTP.KD.ZG | 85.87208 |
| RUS | NY.GDP.MKTP.KD.ZG | 77.25373 |
| USA | NY.GDP.MKTP.KD.ZG | 44.87740 |
| JPN | NY.GDP.MKTP.KD.ZG | 13.93812 |
+-----+-----+-----+
8 rows selected (2.504 seconds)
0: jdbc:hive2://localhost:10000>

```

Conclusion:

Query : Tax revenue (% of GDP)

GC.TAX.TOTL.GD.ZS

```

8 rows selected (2.504 seconds)
0: jdbc:hive2://localhost:10000> SELECT
. . . . .> wdi_final.country_code,
. . . . .> wdi_final.indicator_code,
. . . . .> SUM(wdi_final.value) AS sum_value
. . . . .> FROM
. . . . .> wdi_final
. . . . .> WHERE
. . . . .> (wdi_final.year >= 2000 AND wdi_final.year <= 2022)
. . . . .> AND wdi_final.country_code IN ('USA', 'CHN', 'IND', 'PAK', 'BGD', 'KOR', 'JPN', 'ARE', 'RUS')
. . . . .> AND wdi_final.indicator_code = 'GC.TAX.TOTL.GD.ZS'
. . . . .> GROUP BY
. . . . .> wdi_final.country_code,
. . . . .> wdi_final.indicator_code
. . . . .> ORDER BY
. . . . .> sum_value DESC;
WARNING: Hive-on-MR is deprecated in Hive 2 and may not be available in the future versions. Consider using a different execution engine (
+-----+-----+-----+
| wdi_final.country_code | wdi_final.indicator_code | sum_value |
+-----+-----+-----+
| JPN | GC.TAX.TOTL.GD.ZS | 65203565949323.21190 |
| PAK | GC.TAX.TOTL.GD.ZS | 64280440930143.48677 |
| IND | GC.TAX.TOTL.GD.ZS | 22487517468477.65008 |
| ARE | GC.TAX.TOTL.GD.ZS | 16437505125893.43575 |
| CHN | GC.TAX.TOTL.GD.ZS | 15239750960309.60197 |
| BGD | GC.TAX.TOTL.GD.ZS | 923125019348.35624 |
| RUS | GC.TAX.TOTL.GD.ZS | 290.52510 |
| USA | GC.TAX.TOTL.GD.ZS | 229.42049 |
+-----+-----+-----+
8 rows selected (2.519 seconds)
0: jdbc:hive2://localhost:10000>

```

Conclusion:

Query : Gross savings (% of GDP)

NY.GNS.ICTR.ZS

```
0: jdbc:hive2://localhost:10000> SELECT
. . . . .> wdi_final.country_code,
. . . . .> wdi_final.indicator_code,
. . . . .> SUM(wdi_final.value/1000) AS sum_value
. . . . .> FROM
. . . . .> wdi_final
. . . . .> WHERE
. . . . .> (wdi_final.year >= 2000 AND wdi_final.year <= 2022)
. . . . .> AND wdi_final.country_code IN ('USA', 'CHN', 'IND', 'PAK', 'BGD', 'KOR', 'JPN', 'ARE', 'RUS')
. . . . .> AND wdi_final.indicator_code = 'NY.GNS.ICTR.ZS'
. . . . .> GROUP BY
. . . . .> wdi_final.country_code,
. . . . .> wdi_final.indicator_code
. . . . .> ORDER BY
. . . . .> sum_value DESC;
WARNING: Hive-on-MR is deprecated in Hive 2 and may not be available in the future versions. Consider using a different execution engine

+-----+-----+-----+
| wdi_final.country_code | wdi_final.indicator_code | sum_value |
+-----+-----+-----+
| ARE | NY.GNS.ICTR.ZS | 65203565949.323214 |
| CHN | NY.GNS.ICTR.ZS | 1.002627 |
| BGD | NY.GNS.ICTR.ZS | 0.778399 |
| IND | NY.GNS.ICTR.ZS | 0.713278 |
| JPN | NY.GNS.ICTR.ZS | 0.619713 |
| RUS | NY.GNS.ICTR.ZS | 0.618715 |
| USA | NY.GNS.ICTR.ZS | 0.402552 |
| PAK | NY.GNS.ICTR.ZS | 0.357279 |
+-----+-----+-----+
8 rows selected (2.512 seconds)
```

Conclusion:

Query : Inflation, consumer prices (annual %) FP.CPI.TOTL.ZG

```
8 rows selected (2.512 seconds)
0: jdbc:hive2://localhost:10000> SELECT
. . . . .> wdi_final.country_code,
. . . . .> wdi_final.indicator_code,
. . . . .> SUM(wdi_final.value/1000) AS sum_value
. . . . .> FROM
. . . . .> wdi_final
. . . . .> WHERE
. . . . .> (wdi_final.year >= 2000 AND wdi_final.year <= 2022)
. . . . .> AND wdi_final.country_code IN ('USA', 'CHN', 'IND', 'PAK', 'BGD', 'KOR', 'JPN', 'ARE', 'RUS')
. . . . .> AND wdi_final.indicator_code = 'FP.CPI.TOTL.ZG'
. . . . .> GROUP BY
. . . . .> wdi_final.country_code,
. . . . .> wdi_final.indicator_code
. . . . .> ORDER BY
. . . . .> sum_value DESC;
WARNING: Hive-on-MR is deprecated in Hive 2 and may not be available in the future versions. Consider using a different execution engine

+-----+-----+-----+
| wdi_final.country_code | wdi_final.indicator_code | sum_value |
+-----+-----+-----+
| ARE | FP.CPI.TOTL.ZG | 10260362068.564531 |
| RUS | FP.CPI.TOTL.ZG | 0.218360 |
| PAK | FP.CPI.TOTL.ZG | 0.174370 |
| BGD | FP.CPI.TOTL.ZG | 0.138094 |
| IND | FP.CPI.TOTL.ZG | 0.134975 |
| USA | FP.CPI.TOTL.ZG | 0.049342 |
| CHN | FP.CPI.TOTL.ZG | 0.048023 |
| JPN | FP.CPI.TOTL.ZG | 0.001849 |
+-----+-----+-----+
8 rows selected (2.512 seconds)
```

Conclusion:

Query : Net primary income (Net income from abroad) (current US\$)

NY.GSR.NFCY.CD

```

8 rows selected (4.068 seconds)
0: jdbc:hive2://localhost:10000> SELECT
. . . . .> wdi_final.country_code,
. . . . .> wdi_final.indicator_code,
. . . . .> SUM(wdi_final.value/1000) AS sum_value
. . . . .> FROM
. . . . .> wdi_final
. . . . .> WHERE
. . . . .> (wdi_final.year >= 2000 AND wdi_final.year <= 2022)
. . . . .> AND wdi_final.country_code IN ('USA', 'CHN', 'IND', 'PAK', 'BGD', 'KOR', 'JPN', 'ARE', 'RUS')
. . . . .> AND wdi_final.indicator_code = 'NY.GSR.NFCY.CD'
. . . . .> GROUP BY
. . . . .> wdi_final.country_code,
. . . . .> wdi_final.indicator_code
. . . . .> ORDER BY
. . . . .> sum_value DESC;
WARNING: Hive-on-MR is deprecated in Hive 2 and may not be available in the future versions. Consider using a different execution engine
+-----+-----+-----+
| wdi_final.country_code | wdi_final.indicator_code | sum_value |
+-----+-----+-----+
| JPN | NY.GSR.NFCY.CD | 12577417134.626510 |
| ARE | NY.GSR.NFCY.CD | 9639267392.149190 |
| USA | NY.GSR.NFCY.CD | 3606772000.000000 |
| BGD | NY.GSR.NFCY.CD | 192464589.446000 |
| PAK | NY.GSR.NFCY.CD | -80914010.000000 |
| IND | NY.GSR.NFCY.CD | -399775284.433000 |
| CHN | NY.GSR.NFCY.CD | -743756094.483000 |
| RUS | NY.GSR.NFCY.CD | -815843049.995000 |
+-----+-----+-----+
8 rows selected (2.423 seconds)
0: jdbc:hive2://localhost:10000>

```

Conclusion:

Query : Adjusted savings: education expenditure (current US\$)
 NY.ADJ.AEDU.CD

```

8 rows selected (2.423 seconds)
0: jdbc:hive2://localhost:10000> SELECT
. . . . .> wdi_final.country_code,
. . . . .> wdi_final.indicator_code,
. . . . .> SUM(wdi_final.value/1000) AS sum_value
. . . . .> FROM
. . . . .> wdi_final
. . . . .> WHERE
. . . . .> (wdi_final.year >= 2000 AND wdi_final.year <= 2022)
. . . . .> AND wdi_final.country_code IN ('USA', 'CHN', 'IND', 'PAK', 'BGD', 'KOR', 'JPN', 'ARE', 'RUS')
. . . . .> AND wdi_final.indicator_code = 'NY.ADJ.AEDU.CD'
. . . . .> GROUP BY
. . . . .> wdi_final.country_code,
. . . . .> wdi_final.indicator_code
. . . . .> ORDER BY
. . . . .> sum_value DESC;
WARNING: Hive-on-MR is deprecated in Hive 2 and may not be available in the future versions. Consider using a different execution engine
+-----+-----+-----+
| wdi_final.country_code | wdi_final.indicator_code | sum_value |
+-----+-----+-----+
| USA | NY.ADJ.AEDU.CD | 16608000000.000000 |
| ARE | NY.ADJ.AEDU.CD | 9765247582.398510 |
| JPN | NY.ADJ.AEDU.CD | 3290000000.000000 |
| CHN | NY.ADJ.AEDU.CD | 2910966799.810000 |
| IND | NY.ADJ.AEDU.CD | 1111542163.280000 |
| RUS | NY.ADJ.AEDU.CD | 1072430941.531000 |
| PAK | NY.ADJ.AEDU.CD | 84574080.587000 |
| BGD | NY.ADJ.AEDU.CD | 60592219.910600 |
+-----+-----+-----+
8 rows selected (2.427 seconds)
0: jdbc:hive2://localhost:10000>

```

Conclusion:

Query : Adolescents out of school (% of lower secondary school age)
 SE.SEC.UNER.LO.ZS

```
0: jdbc:hive2://localhost:10000> SELECT
. . . . .> wdi_final.country_code,
. . . . .> wdi_final.indicator_code,
. . . . .> SUM(wdi_final.value/1000) AS sum_value
. . . . .> FROM
. . . . .> wdi_final
. . . . .> WHERE
. . . . .> (wdi_final.year >= 2000 AND wdi_final.year <= 2022)
. . . . .> AND wdi_final.country_code IN ('USA', 'CHN', 'IND', 'PAK', 'BGD', 'KOR', 'JPN', 'ARE', 'RUS')
. . . . .> AND wdi_final.indicator_code = 'SE.SEC.UNER.LO.ZS'
. . . . .> GROUP BY
. . . . .> wdi_final.country_code,
. . . . .> wdi_final.indicator_code
. . . . .> ORDER BY
. . . . .> sum_value DESC;
WARNING: Hive-on-MR is deprecated in Hive 2 and may not be available in the future versions. Consider using a different execution engine
```

wdi_final.country_code	wdi_final.indicator_code	sum_value
PAK	SE.SEC.UNER.LO.ZS	65203565949.323214
CHN	SE.SEC.UNER.LO.ZS	65203565949.323214
USA	SE.SEC.UNER.LO.ZS	59424808318.771562
IND	SE.SEC.UNER.LO.ZS	46853187381.357503
JPN	SE.SEC.UNER.LO.ZS	46325889750.049459
ARE	SE.SEC.UNER.LO.ZS	44863237684.404675
RUS	SE.SEC.UNER.LO.ZS	37005016680.006385
BGD	SE.SEC.UNER.LO.ZS	27874910212.146773

Conclusion:

Query : Children in employment, total (% of children ages 7-14)
SL.TLF.0714.ZS

```
8 rows selected (2.506 seconds)
0: jdbc:hive2://localhost:10000> SELECT
. . . . .> wdi_final.country_code,
. . . . .> wdi_final.indicator_code,
. . . . .> SUM(wdi_final.value/1000) AS sum_value
. . . . .> FROM
. . . . .> wdi_final
. . . . .> WHERE
. . . . .> (wdi_final.year >= 2000 AND wdi_final.year <= 2022)
. . . . .> AND wdi_final.country_code IN ('USA', 'CHN', 'IND', 'PAK', 'BGD', 'KOR', 'JPN', 'ARE', 'RUS')
. . . . .> AND wdi_final.indicator_code = 'SL.TLF.0714.ZS'
. . . . .> GROUP BY
. . . . .> wdi_final.country_code,
. . . . .> wdi_final.indicator_code
. . . . .> ORDER BY
. . . . .> sum_value DESC;
WARNING: Hive-on-MR is deprecated in Hive 2 and may not be available in the future versions. Consider using a different execution engine
```

wdi_final.country_code	wdi_final.indicator_code	sum_value
USA	SL.TLF.0714.ZS	65203565949.323214
RUS	SL.TLF.0714.ZS	65203565949.323214
JPN	SL.TLF.0714.ZS	65203565949.323214
CHN	SL.TLF.0714.ZS	65203565949.323214
ARE	SL.TLF.0714.ZS	65203565949.323214
PAK	SL.TLF.0714.ZS	60715564062.983718
BGD	SL.TLF.0714.ZS	59560952703.293001
IND	SL.TLF.0714.ZS	56549321941.429615

Conclusion:

Query : Children out of school (% of primary school age)
SE.PRM.UNER.ZS

```
0: jdbc:hive2://localhost:10000> SELECT
. . . . .> wdi_final.country_code,
. . . . .> wdi_final.indicator_code,
. . . . .> SUM(wdi_final.value/1000) AS sum_value
. . . . .> FROM
. . . . .> wdi_final
. . . . .> WHERE
. . . . .> (wdi_final.year >= 2000 AND wdi_final.year <= 2022)
. . . . .> AND wdi_final.country_code IN ('USA', 'CHN', 'IND', 'PAK', 'BGD', 'KOR', 'JPN', 'ARE', 'RUS')
. . . . .> AND wdi_final.indicator_code = 'SE.PRM.UNER.ZS'
. . . . .> GROUP BY
. . . . .> wdi_final.country_code,
. . . . .> wdi_final.indicator_code
. . . . .> ORDER BY
. . . . .> sum_value DESC;
WARNING: Hive-on-MR is deprecated in Hive 2 and may not be available in the future versions. Consider using a different execution engine
```

wdi_final.country_code	wdi_final.indicator_code	sum_value
PAK	SE.PRM.UNER.ZS	65203565949.323214
CHN	SE.PRM.UNER.ZS	65203565949.323214
BGD	SE.PRM.UNER.ZS	45454360817.960845
JPN	SE.PRM.UNER.ZS	37005016680.012748
USA	SE.PRM.UNER.ZS	33006700719.659233
ARE	SE.PRM.UNER.ZS	31126436999.511741
RUS	SE.PRM.UNER.ZS	24602239618.403715
IND	SE.PRM.UNER.ZS	10512594413.414713

Conclusion:

Query : Compulsory education, duration (years)
SE.COM.DURS

```
8 rows selected (2.386 seconds)
0: jdbc:hive2://localhost:10000> SELECT
. . . . .> wdi_final.country_code,
. . . . .> wdi_final.indicator_code,
. . . . .> SUM(wdi_final.value/1000) AS sum_value
. . . . .> FROM
. . . . .> wdi_final
. . . . .> WHERE
. . . . .> (wdi_final.year >= 2000 AND wdi_final.year <= 2022)
. . . . .> AND wdi_final.country_code IN ('USA', 'CHN', 'IND', 'PAK', 'BGD', 'KOR', 'JPN', 'ARE', 'RUS')
. . . . .> AND wdi_final.indicator_code = 'SE.COM.DURS'
. . . . .> GROUP BY
. . . . .> wdi_final.country_code,
. . . . .> wdi_final.indicator_code
. . . . .> ORDER BY
. . . . .> sum_value DESC;
WARNING: Hive-on-MR is deprecated in Hive 2 and may not be available in the future versions. Consider using a different execution engine
```

wdi_final.country_code	wdi_final.indicator_code	sum_value
IND	SE.COM.DURS	12202457074.558608
JPN	SE.COM.DURS	9594000591.792510
USA	SE.COM.DURS	0.264000
PAK	SE.COM.DURS	0.264000
RUS	SE.COM.DURS	0.212000
CHN	SE.COM.DURS	0.198000
ARE	SE.COM.DURS	0.192000
BGD	SE.COM.DURS	0.110000

Conclusion:

Query : Educational attainment, at least Bachelor's or equivalent, population 25+,
total (%) (cumulative)
SE.TER.CUAT.BA.ZS

```

8 rows selected (2.379 seconds)
0: jdbc:hive2://localhost:10000> SELECT
. . . . .> wdi_final.country_code,
. . . . .> wdi_final.indicator_code,
. . . . .> SUM(wdi_final.value/1000) AS sum_value
. . . . .> FROM
. . . . .> wdi_final
. . . . .> WHERE
. . . . .> (wdi_final.year >= 2000 AND wdi_final.year <= 2022)
. . . . .> AND wdi_final.country_code IN ('USA', 'CHN', 'IND', 'PAK', 'BGD', 'KOR', 'JPN', 'ARE', 'RUS')
. . . . .> AND wdi_final.indicator_code = 'SE.TER.CUAT.BA.ZS'
. . . . .> GROUP BY
. . . . .> wdi_final.country_code,
. . . . .> wdi_final.indicator_code
. . . . .> ORDER BY
. . . . .> sum_value DESC;
WARNING: Hive-on-MR is deprecated in Hive 2 and may not be available in the future versions. Consider using a different execution engine

+-----+-----+-----+
| wdi_final.country_code | wdi_final.indicator_code | sum_value |
+-----+-----+-----+
| PAK | SE.TER.CUAT.BA.ZS | 65203565949.323214 |
| RUS | SE.TER.CUAT.BA.ZS | 62968823948.772905 |
| JPN | SE.TER.CUAT.BA.ZS | 62968823948.733557 |
| CHN | SE.TER.CUAT.BA.ZS | 62968823948.717222 |
| IND | SE.TER.CUAT.BA.ZS | 52562795873.445309 |
| USA | SE.TER.CUAT.BA.ZS | 42203537387.829906 |
| ARE | SE.TER.CUAT.BA.ZS | 36731889158.662438 |
| BGD | SE.TER.CUAT.BA.ZS | 26031505717.575019 |
+-----+-----+-----+

```

Conclusion:

Query : Government expenditure on education, total (% of GDP)

SE.XPD.TOTL.GD.ZS

```

8 rows selected (2.411 seconds)
0: jdbc:hive2://localhost:10000> SELECT
. . . . .> wdi_final.country_code,
. . . . .> wdi_final.indicator_code,
. . . . .> SUM(wdi_final.value/1000) AS sum_value
. . . . .> FROM
. . . . .> wdi_final
. . . . .> WHERE
. . . . .> (wdi_final.year >= 2000 AND wdi_final.year <= 2022)
. . . . .> AND wdi_final.country_code IN ('USA', 'CHN', 'IND', 'PAK', 'BGD', 'KOR', 'JPN', 'ARE', 'RUS')
. . . . .> AND wdi_final.indicator_code = 'SE.XPD.TOTL.GD.ZS'
. . . . .> GROUP BY
. . . . .> wdi_final.country_code,
. . . . .> wdi_final.indicator_code
. . . . .> ORDER BY
. . . . .> sum_value DESC;
WARNING: Hive-on-MR is deprecated in Hive 2 and may not be available in the future versions. Consider using a different execution engine

+-----+-----+-----+
| wdi_final.country_code | wdi_final.indicator_code | sum_value |
+-----+-----+-----+
| ARE | SE.XPD.TOTL.GD.ZS | 26031505717.513408 |
| USA | SE.XPD.TOTL.GD.ZS | 23796763716.946299 |
| IND | SE.XPD.TOTL.GD.ZS | 18666970264.447070 |
| CHN | SE.XPD.TOTL.GD.ZS | 15239750960.213648 |
| RUS | SE.XPD.TOTL.GD.ZS | 13361387805.864835 |
| JPN | SE.XPD.TOTL.GD.ZS | 11594306642.496511 |
| BGD | SE.XPD.TOTL.GD.ZS | 3583731027.712378 |
| PAK | SE.XPD.TOTL.GD.ZS | 3400814196.940652 |
+-----+-----+-----+
8 rows selected (2.465 seconds)

```

Conclusion: