```
-- Student Number: 10241485
-- Name: Jeslyn Ho Ka Yan
-- Date: 13 Oct 2024
-- Assignment 2, Task 4
HADOOP TERMINAL
# Start Hadoop services, but first, permissions issue prevents execution
bigdata@bigdata-VirtualBox:~$ ./startHadoop.sh
Starting namenodes on [localhost]
localhost: starting namenode, logging to /usr/share/hadoop-2.7.3/logs/hadoop-bigdata-
namenode-bigdata-VirtualBox.out
localhost: starting datanode, logging to /usr/share/hadoop-2.7.3/logs/hadoop-bigdata-
datanode-bigdata-VirtualBox.out
Starting secondary namenodes [0.0.0.0]
0.0.0.0: starting secondarynamenode, logging to /usr/share/hadoop-2.7.3/logs/hadoop-
bigdata-secondarynamenode-bigdata-VirtualBox.out
starting yarn daemons
starting resourcemanager, logging to /usr/share/hadoop-2.7.3/logs/yarn-bigdata-
resourcemanager-bigdata-VirtualBox.out
localhost: starting nodemanager, logging to /usr/share/hadoop-2.7.3/logs/yarn-
bigdata-nodemanager-bigdata-VirtualBox.out
starting historyserver, logging to /usr/share/hadoop-2.7.3/logs/mapred-bigdata-
historyserver-bigdata-VirtualBox.out
3536 Jps
2523 NameNode
3499 JobHistoryServer
2924 SecondaryNameNode
3261 NodeManager
3103 ResourceManager
2703 DataNode
# Creating the authors directory in the a2 database
bigdata@bigdata-VirtualBox:~$ $HADOOP_HOME/bin/hadoop fs -mkdir
/user/hive/warehouse/a2.db/authors
# Creating the items directory in the a2 database
bigdata@bigdata-VirtualBox:~$ $HADOOP_HOME/bin/hadoop fs -mkdir
/user/hive/warehouse/a2.db/items
# Listing contents of a2 database to confirm both authors and items directories are
bigdata@bigdata-VirtualBox:~$ $HADOOP_HOME/bin/hadoop fs -ls
/user/hive/warehouse/a2.db/
Found 2 items
            - bigdata supergroup
                                         0 2024-10-30 02:17
drwxr-xr-x
/user/hive/warehouse/a2.db/authors
           - bigdata supergroup
                                           0 2024-10-30 02:17
drwxr-xr-x
/user/hive/warehouse/a2.db/items
# Upload the author.tbl file to the authors directory in Hive warehouse
bigdata@bigdata-VirtualBox:~$ $HADOOP_HOME/bin/hadoop fs -put
~/ISIT312/A2/Task4/author.tbl /user/hive/warehouse/a2.db/authors/
# Upload the item.tbl file to the items directory in Hive warehouse
bigdata@bigdata-VirtualBox:~$ $HADOOP_HOME/bin/hadoop fs -put
~/ISIT312/A2/Task4/item.tbl /user/hive/warehouse/a2.db/items/
```

```
BEELINE TERMINAL
-- Start Beeline: Open the Beeline command line tool for Hive.
bigdata@bigdata-VirtualBox:~$ $HIVE_HOME/bin/beeline
Beeline version 2.1.1 by Apache Hive
-- Connect to Hive: Connect to the Hive server running on your local machine
beeline> !connect jdbc:hive2://localhost:10000
SLF4J: Class path contains multiple SLF4J bindings.
SLF4J: Found binding in [jar:file:/usr/share/apache-hive-2.1.1-bin/lib/log4j-slf4j-
impl-2.4.1.jar!/org/slf4j/impl/StaticLoggerBinder.class]
SLF4J: Found binding in [jar:file:/usr/share/hadoop-
2.7.3/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
Enter username for jdbc:hive2://localhost:10000:
Enter password for jdbc:hive2://localhost:10000:
Connected to: Apache Hive (version 2.1.1)
Driver: Hive JDBC (version 2.1.1)
24/10/30 02:16:57 [main]: WARN jdbc.HiveConnection: Request to set autoCommit to
false; Hive does not support autoCommit=false.
Transaction isolation: TRANSACTION_REPEATABLE_READ
-- Create Database: Make a new database called 'a2'.
0: idbc:hive2://localhost:10000> create database a2;
No rows affected (0.135 seconds)
-- Show Databases: List all existing databases to check if `a2` was created.
0: jdbc:hive2://localhost:10000> show databases;
| database_name |
l a2
| default |
+----+
2 rows selected (0.07 seconds)
-- Use Database: Switch to using the 'a2' database.
0: jdbc:hive2://localhost:10000> use a2;
No rows affected (0.019 seconds)
-- This table will use the data from the author.tbl file in HDFS
0: jdbc:hive2://localhost:10000> CREATE EXTERNAL TABLE IF NOT EXISTS author (
-- Unique author ID
                             a_id INT,
a_fname STRING,
                                               -- First name of author
-- Middle name of author
a_mname STRING,
a_lname STRING,
                                               -- Last name of author
a_dob DATE,
                                               -- Date of birth of author
a_bio STRING
                                               -- Biography of the author
```

. ROW FORMAT DELIMITED

```
. . . . . . . . . . . . . . . . . . FIELDS TERMINATED BY '|'
    . . . . . . . . . . . . . . . STORED AS TEXTFILE
  . . . . . . > LOCATION '/user/hive/warehouse/a2.db/authors/';
No rows affected (0.077 seconds)
-- This table will use the data from the item.tbl file in HDFS
0: jdbc:hive2://localhost:10000> CREATE EXTERNAL TABLE IF NOT EXISTS item (
 i_id INT,
                             -- Unique ID of item
                  i_title STRING,
                            -- Title of the item
                  i_a_id INT,
                             -- Author ID (foreign key
referencing AUTHOR table)
i_pub_date DATE,
                             -- Publication date of the
item
-- Publisher of the item
                  i_publisher STRING,
i_subject STRING,
                             -- Subject category of the
item
i_desc STRING,
                             -- Description of the item
i_related1 INT,
                             -- Related item ID 1
i_related2 INT,
                             -- Related item ID 2
i_related3 INT,
                             -- Related item ID 3
i_related4 INT,
                             -- Related item ID 4
i_related5 INT,
                             -- Related item ID 5
i_thumbnail STRING, -- Path to thumbnail image
 of the item
i_image STRING, -- Path to main image of
the item
i_srp DECIMAL(15,2), -- Suggested retail price
i_cost DECIMAL(15,2), -- Cost of the item
i_avail DATE,
                            -- Availability date of
the item
i_stock INT,
                             -- Quantity in stock
i_isbn STRING,
                             -- ISBN of the item
i_page INT,
                             -- Number of pages (for
   i_backing STRING, -- Book binding type
(e.g., paperback, hardback)
```

| i_dimensions STRING Physical dimensions of |
|--|
| the item |
| |
| |
| |
| |
| |
| LOCATION '/user/hive/warehouse/a2.db/items/'; No rows affected (0.108 seconds) |
| Query to count the total rows in the AUTHOR table 0: jdbc:hive2://localhost:10000> SELECT COUNT(*) AS author_row_count FROM author; 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, tez) or using Hive 1.X releases. |
| + |
| ++ |
| 2500 |
| 1 row selected (16.874 seconds) |
| 0: jdbc:hive2://localhost:10000> SELECT COUNT(*) AS item_row_count FROM item; 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, tez) or using Hive 1.X releases. ++ |
| item_row_count ++ |
| 10000 |
| ++ 1 row selected (20.8 seconds) |
| 1 Fow Selected (20.8 Seconds) |
| This query sums up the row counts from both tables to get the total row count across tables |
| 0: jdbc:hive2://localhost:10000> SELECT SUM(count) AS total_row_count > FROM (|
| > SELECT COUNT(*) AS count EPOM author |
| |
|) AS combined_counts; |
| 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, tez) or using Hive |
| 1.X releases. |
| + |
| ++ |
| 12500 ++ |
| 1 row selected (63.419 seconds) |
| Query to retrieve the first 3 rows from the AUTHOR table 0: jdbc:hive2://localhost:10000> SELECT * FROM author LIMIT 3; + |

| author.a_id author.a_fname author.a_mname author.a_dob author.a_bio | author.a_lname |
|--|---|
| + | |
| | |
| | |
| | |
| · | |
| Minutes search in the young, labour years. Individuals we continue in the following changes: Actions at the big, you clear, african ideas. Great, large pupils do pick. Necessa years carry long, external circumstances. Questions might might rest customers: Social, other players identify. Systride. Overseas, federal goods interfere. | ng aspects shall sleep ry eyes to the large, small carry. Lovely, actual talks ems could play. Standards ABABABABABABAALWH NULL the numbers realise ies forget to the members. dogs avoid. Sales to the as wear about. Names could between the mutual, prime ABABABABABABABARI NULL |
| | |
| | ays should move |
| imited, personal matters will have to make in the eyes. W | ays should move |
| Limited, personal matters will have to make in the eyes. W | ays should move |
| limited, personal matters will have to make in the eyes. W | ays should move |
| Limited, personal matters will have to make in the eyes. W | ays should move |
| limited, personal matters will have to make in the eyes. W | ays should move |
| Limited, personal matters will have to make in the eyes. W | ays should move |
| Limited, personal matters will have to make in the eyes. W | ays should move |
| imited, personal matters will have to make in the eyes. W | ays should move |
| Limited, personal matters will have to make in the eyes. W | ays should move |
| Limited, personal matters will have to make in the eyes. W | ays should move |
| Limited, personal matters will have to make in the eyes. W | ays should move |
| limited, personal matters will have to make in the eyes. W | ays should move |
| limited, personal matters will have to make in the eyes. W | ays should move |
| limited, personal matters will have to make in the eyes. W ++ 3 rows selected (0.118 seconds) Query to retrieve the first 3 rows from the ITEM table 0: jdbc:hive2://localhost:10000> SELECT * FROM item LIMIT ++ | ays should move |
| limited, personal matters will have to make in the eyes. W | ays should move |
| limited, personal matters will have to make in the eyes. W | ays should move |

| <pre>item.i_avail item.i_stock item.i_isbn item.i_page item.i_backing </pre> |
|---|
| item.i_dimensions |
| ++++ |
| _+ |
| |
| |
| |
| |
| |
| · |
| +++++ |
| |
| + |
| 1 Open, BABABABABABAOG wrong 1 |
| NULL Common occasions will have to ROMANCE |
| General reasons during the long, living subjects might take with the goods. Cells |
| continue. Small, new schools shift. Nations to the offences might show by the |
| subsequent, simple homes. Important, elderly years encourage. Dogs from the types |
| shall have to pull organic, top sources. Materials go to the big, french friends. |
| Social, nuclear examples should ask. Patients strengthen for the other, northern |
| 9272 9550 707 2253 5398 |
| c:\ |
| JQyQ5MIc91ANC |
| 2 Children with BABABABABABAL the other, necessary 2 |
| NULL Activities may forget. Blue birds need? COOKING |
| Similar, southern goods know to the centres. Round, domestic months must improve |
| probably on the thanks. Attitudes shall have to land. Full, russian walls could |
| control. Useful, economic hundreds make at the nuclear, great colours. Cultural, |
| immediate ministers give. Other, ordinary pictures make. Likely, great days satisfy |
| before the wrong, small companies. Economic, specific women lose. Other, |
| 8661 5159 7081 5701 3667 |
| |
| |
| |
| 3 Years will break BABABABABARI pleasant, free terms 3 |
| NULL |
| Hands on the british names come christian, other minutes. Members do make; Large |
| others can appear. Bodies suffer. Wages should know by the obvious minutes. Policies |
| ensure on the communities. Objects can cost in the interesting, early years. Active, |
| inner points meet. Departments in the particular, human years might evaluate such as |
| the weeks. Growing, left theories to the legal, old 4546 |
| 7449 |
| c:\ |
| hIs86uiW15pA6 9255 PAPERBACK 45.30x52.68x72.80 |
| ++++ |
| -+ |
| |
| |
| |
| |
| + |
| + |
| |
| + |
| 3 rows selected (0.15 seconds) |
| |

```
bigdata@bigdata-VirtualBox:~$ ./startHadoop.sh
Starting namenodes on [localhost]
localhost: starting namenode, logging to /usr/share/hadoop-2.7.3/logs/hadoop-bigdata-
namenode-bigdata-VirtualBox.out
localhost: starting datanode, logging to /usr/share/hadoop-2.7.3/logs/hadoop-bigdata-
datanode-bigdata-VirtualBox.out
Starting secondary namenodes [0.0.0.0]
0.0.0.0: starting secondarynamenode, logging to /usr/share/hadoop-2.7.3/logs/hadoop-
bigdata-secondarynamenode-bigdata-VirtualBox.out
starting yarn daemons
starting resourcemanager, logging to /usr/share/hadoop-2.7.3/logs/yarn-bigdata-
resourcemanager-bigdata-VirtualBox.out
localhost: starting nodemanager, logging to /usr/share/hadoop-2.7.3/logs/yarn-
bigdata-nodemanager-bigdata-VirtualBox.out
starting historyserver, logging to /usr/share/hadoop-2.7.3/logs/mapred-bigdata-
historyserver-bigdata-VirtualBox.out
3536 Jps
2523 NameNode
3499 JobHistoryServer
2924 SecondaryNameNode
3261 NodeManager
3103 ResourceManager
2703 DataNode
# Creating the authors directory in the a2 database
bigdata@bigdata-VirtualBox:~$ $HADOOP_HOME/bin/hadoop fs -mkdir
/user/hive/warehouse/a2.db/authors
# Creating the items directory in the a2 database
bigdata@bigdata-VirtualBox:~$ $HADOOP_HOME/bin/hadoop fs -mkdir
/user/hive/warehouse/a2.db/items
# Listing contents of a2 database to confirm both authors and items directories are
bigdata@bigdata-VirtualBox:~$ $HADOOP_HOME/bin/hadoop fs -ls
/user/hive/warehouse/a2.db/
Found 2 items
drwxr-xr-x
            - bigdata supergroup
                                           0 2024-10-30 02:17
/user/hive/warehouse/a2.db/authors
drwxr-xr-x - bigdata supergroup
                                           0 2024-10-30 02:17
/user/hive/warehouse/a2.db/items
# Upload the author.tbl file to the authors directory in Hive warehouse
bigdata@bigdata-VirtualBox:~$ $HADOOP_HOME/bin/hadoop fs -put
~/ISIT312/A2/Task4/author.tbl /user/hive/warehouse/a2.db/authors/
# Upload the item.tbl file to the items directory in Hive warehouse
bigdata@bigdata-VirtualBox:~$ $HADOOP_HOME/bin/hadoop fs -put
~/ISIT312/A2/Task4/item.tbl /user/hive/warehouse/a2.db/items/
```

Start Hadoop services, but first, permissions issue prevents execution

Screenshot

HADOOP TERMINAL

BEELINE TERMINAL

```
bigdata@bigdata-VirtualBox:~$ $HIVE_HOME/bin/beeline
Beeline version 2.1.1 by Apache Hive
beeline> !connect jdbc:hive2://localhost:10000
SLF4J: Class path contains multiple SLF4J bindings.
SLF4J: Found binding in [jar:file:/usr/share/apache-hive-2.1.1-bin/lib/log4j-slf4j-impl-2.4.1.jar
!/org/slf4j/impl/StaticLoggerBinder.class]
SĹF4Ĵ: Found binding in [jar:file:/usr/share/hadoop-2.7.3/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
Enter username for jdbc:hive2://localhost:10000:
Enter password for jdbc:hive2://localhost:10000:
Connected to: Apache Hive (version 2.1.1)
Driver: Hive JDBC (version 2.1.1)
24/10/30 02:16:57 [main]: WARN jdbc.HiveConnection: Request to set autoCommit to false; Hive does
not support autoCommit=false.
Transaction isolation: TRANSACTION_REPEATABLE_READ
0: jdbc:hive2://localhost:10000> create database a2;
No rows affected (0.135 seconds)
0: jdbc:hive2://localhost:10000> show databases;
| database_name |
I a2
I default
2 rows selected (0.07 seconds)
0: jdbc:hive2://localhost:10000> use a2;
No rows affected (0.019 seconds)
```

```
0: jdbc:hive2://localhost:10000> CREATE EXTERNAL TABLE IF NOT EXISTS author (
a id INT.
                                           -- Unique author ID
                          a_fname STRING,
                                           -- First name of author
a mname STRING,
                                           -- Middle name of author
                           a_lname STRING,
                                           -- Last name of author
                           a_dob DATE,
                                           -- Date of birth of author
                           a bio STRING
                                           -- Biography of the author
. . . . . . . . . . . . . . . . ROW FORMAT DELIMITED
. . . . . . . . . . . . . . . . FIELDS TERMINATED BY '|'
. . . . . . . . . . . . . STORED AS TEXTFILE
. . . . . . > LOCATION '/user/hive/warehouse/a2.db/authors/';
No rows affected (0.077 seconds)
```

```
0: idbc:hive2://localhost:10000> CREATE EXTERNAL TABLE IF NOT EXISTS item (
                              i_id INT,
                                                -- Unique ID of item
                              i_title STRING,
                                               -- Title of the item
 i_a_id INT,
                                                -- Author ID (foreign key referencing AUTHOR table)
   i pub date DATE,
                                                -- Publication date of the item
 i_publisher STRING,
                                                -- Publisher of the item
                              i_subject STRING,
                                                -- Subject category of the item
                                                -- Description of the item
                              i desc STRING.
                              i_related1 INT,
                                                -- Related item ID 1
                              i_related2 INT,
                                                -- Related item ID 2
                              i_related3 INT,
                                                -- Related item ID 3
                              i_related4 INT,
                                                -- Related item ID 4
 i_related5 INT,
                                                -- Related item ID 5
                              i_thumbnail STRING,
                                                -- Path to thumbnail image of the item
                                                -- Path to main image of the item
                              i_image STRING,
                              i_srp DECIMAL(15,2), -- Suggested retail price
 i_cost DECIMAL(15,2), -- Cost of the item
                              i_avail DATE,
                                                -- Availability date of the item
                              i_stock INT,
                                                -- Quantity in stock
 i_isbn STRING,
                                                -- ISBN of the item
                              i_page INT,
                                                -- Number of pages (for books)
                              i_backing STRING,
                                                -- Book binding type (e.g., paperback, hardback)
                              i\_dimensions \ STRING \ -- \ Physical \ dimensions \ of \ the \ item
```

```
AND FORMAT DELIMITED

PIELDS TERMINATED BY '|'

STORED AS TEXTFILE

LOCATION '/user/hive/warehouse/a2.db/ttens/';

No rows affected (e.108 seconds)

LOCATION '/user/hive/warehouse/a2.db/ttens/';

NORNING: MIVE on-MB its deprecated in Hive 2 and may not be available in the future versions. Consider using a different execution engine (i.e. spark, tez) or using Hive 1.X releases.

LOCATION '/ User in the future versions of the future
```

| a. idha.hiwaz.//laszlhast.18888- SE | LECT + EDOM itom LIMIT | | | | | | | | | |
|---|------------------------|------------------|-------------------|-----------------|-----------------|--------------------|--------------|--------------------|--------------------|-----------------------|
| 0: jdbc:hive2://localhost:10000> SELECT * FROM item LIMIT 3; | | | | | | | | | | |
| | | | | | ., | | | | | |
| | | | | | | | | | | |
| | -+ | | + | | -+ | | | + | | ++ |
| | | | | | | | | | | |
| item.i_id | item.i_title | 11 | item.i_a_id i | tem.i_pub_date | 1 | item.i_publisher | 1 | item.i_subject | ı | |
| | - | | ' | | | | | item.i | desc | |
| | | | | | | | | | | item.i_related2 |
| item.i_related3 item.i_related4 | item.i_related5 | item.i_thumbnail | item.i_image | item.i_srp | item.i_cost | item.i_avail | item.i_stock | item.i_isbn | item.i_page | item.i_backing i |
| tem.i_dimensions | | | | | | | | | | |
| + | | +- | | | -+ | | + | | | |
| | | | | | | | | | | |
| | | | | | | | | + | | ++- |
| | -+ | | + | + | -+ | + | -+ | + | + | ++ |
| | | | | | | | | | | |
| 1 Open, BABABABABABAGG ing subjects might take with the go | | | | | | sions will have to | | | | during the long, liv |
| shall have to pull organic, top so | | | | | | | | | | 1 9550 I l |
| 707 2253 | | c:\ | c:\ | 1727.42 | 821.00 | I NULL | 1 22 | JOVO5MIc91ANC | | PAPERBACK 8 |
| .49x69.77x55.64 | 1 2236 | c. (| 1 6.1 | 1 1/2/.42 | 021.00 | INOLL | 1 22 | 1 adadauteatwice | 1 42/3 | FAFERDACK 0 |
| | BABABAAL the other, ne | cessarv | 2 NI | 11.1 | I Activities n | may forget. Blue | hirds need? | COOKING | Similar souther | rn goods know to the |
| centres. Round, domestic months mus | | | | | | | | | | |
| immediate ministers give. Other, o | | | | | | | | | 3661 | l 5159 |
| 7081 5701 | | c:\ | c:\ | 8893.32 | 1 3909.30 | I NULL | 1 25 | I dFSe5WCeX3GJz | | AUDIO 19 |
| 3.76x84.33x1.71 | 1 | | 1 | , | , | 1 | 1 | , | 1 | , |
| I 3 I Years will break BAB | ABABABABARI pleasant. | free terms I | 3 N | JLL | I Simple, clea | ar feet represent | . Functions | SCIENCE-NATURE | Hands on the bri | itish names come chri |
| stian, other minutes. Members do ma | ke; Large others can a | ppear. Bodies su | ffer. Wages shou | ld know by the | | | | ities. Objects car | n cost in the inte | eresting, early years |
| . Active, inner points meet. Depart | ments in the particula | r, human years m | ight evaluate suc | ch as the weeks | . Growing, left | t theories to the | legal, old | 1.4 | 1546 | 7449 |
| 7146 9690 | 69 | c:\ | _ c:\ | 8345.49 | 3981.86 | NULL | 15 | hIs86uiW15pA6 | 9255 | PAPERBACK 4 |
| 5.30x52.68x72.80 | | | | | | | | | | |
| + | | +- | + | | -+ | | + | | + | |
| | | | | | | | | | | |
| | | | | | | | | | | ++- |
| | -+ | | + | + | -+ | + | -+ | + | + | ++ |
| | | | | | | | | | | |
| 3 rows selected (0.15 seconds) | | | | | | | | | | |
| 0: jdbc:hive2://localhost:10000> | | | | | | | | | | |