

**Name: Jeslyn Ho Ka Yan**  
**Date: 9 Nov 2024**  
**Assignment 3, Task 2**

---

// This command Starts Hbase

```
bigdata@bigdata-VirtualBox:~$ $HBASE_HOME/bin/hbase shell
SLF4J: Class path contains multiple SLF4J bindings.
SLF4J: Found binding in [jar:file:/usr/share/hbase-1.2.6/lib/slf4j-log4j12-1.7.5.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.slf4j.impl.Log4jLoggerFactory]
HBase Shell; enter 'help<RETURN>' for list of supported commands.
Type "exit<RETURN>" to leave the HBase Shell
Version 1.2.6, rUnknown, Mon May 29 02:25:32 CDT 2017
```

// This command runs the HBase shell script solution2.hb, located at /home/bigdata/ISIT312/A3/, which executes all the HBase commands specified in the script.

```
hbase(main):001:0> source ('/home/bigdata/ISIT312/A3/solution2.hb')
solution2.hb(main):002:0> # Name: Jeslyn ho
solution2.hb(main):003:0* # SIM ID:10241485
solution2.hb(main):004:0* # Assingment 3 Task 2
solution2.hb(main):005:0* #
solution2.hb(main):006:0*
solution2.hb(main):007:0* #Query 1:
solution2.hb(main):008:0* # Retrieve all available information for the
product with product number 101, listing each version in a separate cell.
solution2.hb(main):009:0* get 'task2', 'product101', {COLUMN => 'PRODUCT',
VERSIONS => 4}
COLUMN                                CELL
0 row(s) in 0.6790 seconds

solution2.hb(main):010:0>
solution2.hb(main):011:0* #Query 2:
solution2.hb(main):012:0* # Retrieve all available information about sales of
product 101 made by customer 007, listing each version in a separate cell.
solution2.hb(main):013:0* scan 'task2', {FILTER => "PrefixFilter('sale101-
customer007')", VERSIONS => 4}
ROW                                    COLUMN+CELL
0 row(s) in 0.0680 seconds

solution2.hb(main):014:0>
solution2.hb(main):015:0*
solution2.hb(main):016:0* #Query 3:
solution2.hb(main):017:0* # Retrieve the name, location (address), and age
group of all customers, listing each customer in a separate cell.
solution2.hb(main):018:0* scan 'task2', {COLUMNS => ['CUSTOMER:customerName',
'CUSTOMER:location', 'CUSTOMER:ageGroup']}
ROW                                    COLUMN+CELL
```

```

customer|007
column=CUSTOMER:ageGroup, timestamp=1731231590179, value=15 - 64
customer|007
column=CUSTOMER:customerName, timestamp=1731231590164, value=James Bond
customer|007
column=CUSTOMER:location, timestamp=1731231590172, value=England
customer|666
column=CUSTOMER:ageGroup, timestamp=1731231590208, value=<15
customer|666
column=CUSTOMER:customerName, timestamp=1731231590193, value=Harry Potter
customer|666
column=CUSTOMER:location, timestamp=1731231590200, value=England
customer|888
column=CUSTOMER:ageGroup, timestamp=1731231590239, value=>64
customer|888
column=CUSTOMER:customerName, timestamp=1731231590216, value=Wukong
customer|888
column=CUSTOMER:location, timestamp=1731231590221, value=China
3 row(s) in 0.0940 seconds

```

```

solution2.hb(main):019:0>
solution2.hb(main):020:0*
solution2.hb(main):021:0* #Query 4:
solution2.hb(main):022:0* # Retrieve all information about products under the
brand 'Samsung', listing each product in a separate cell.
solution2.hb(main):023:0* scan 'task2', {FILTER =>
"SingleColumnValueFilter('PRODUCT', 'brand', =, 'binary:SAMSUNG')"}
ROW                                COLUMN+CELL
customer|007
column=CUSTOMER:ageGroup, timestamp=1731231590179, value=15 - 64
customer|007
column=CUSTOMER:customerId, timestamp=1731231590139, value=007
customer|007
column=CUSTOMER:customerName, timestamp=1731231590164, value=James Bond
customer|007
column=CUSTOMER:location, timestamp=1731231590172, value=England
customer|666
column=CUSTOMER:ageGroup, timestamp=1731231590208, value=<15
customer|666
column=CUSTOMER:customerId, timestamp=1731231590184, value=666
customer|666
column=CUSTOMER:customerName, timestamp=1731231590193, value=Harry Potter
customer|666
column=CUSTOMER:location, timestamp=1731231590200, value=England
customer|888
column=CUSTOMER:ageGroup, timestamp=1731231590239, value=>64
customer|888
column=CUSTOMER:customerId, timestamp=1731231590212, value=888
customer|888
column=CUSTOMER:customerName, timestamp=1731231590216, value=Wukong
customer|888
column=CUSTOMER:location, timestamp=1731231590221, value=China
date|20240110                        column=DATE:day,
timestamp=1731231590412, value=10

```

date 20240110	column=DATE:dayOfWeek,
timestamp=1731231590420, value=Wednesday	
date 20240110	column=DATE:month,
timestamp=1731231590408, value=01	
date 20240110	column=DATE:quarter,
timestamp=1731231590416, value=01	
date 20240110	column=DATE:year,
timestamp=1731231590403, value=2024	
date 20240128	column=DATE:day,
timestamp=1731231590435, value=28	
date 20240128	column=DATE:dayOfWeek,
timestamp=1731231590497, value=Sunday	
date 20240128	column=DATE:month,
timestamp=1731231590431, value=01	
date 20240128	column=DATE:quarter,
timestamp=1731231590446, value=01	
date 20240128	column=DATE:year,
timestamp=1731231590425, value=2024	
date 20240130	column=DATE:day,
timestamp=1731231590558, value=30	
date 20240130	column=DATE:dayOfWeek,
timestamp=1731231590567, value=Tuesday	
date 20240130	column=DATE:month,
timestamp=1731231590523, value=01	
date 20240130	column=DATE:quarter,
timestamp=1731231590563, value=01	
date 20240130	column=DATE:year,
timestamp=1731231590512, value=2024	
date 20240206	column=DATE:day,
timestamp=1731231590576, value=06	
date 20240206	column=DATE:dayOfWeek,
timestamp=1731231590583, value=Tuesday	
date 20240206	column=DATE:month,
timestamp=1731231590573, value=02	
date 20240206	column=DATE:quarter,
timestamp=1731231590580, value=01	
date 20240206	column=DATE:year,
timestamp=1731231590570, value=2024	
date 20240217	column=DATE:day,
timestamp=1731231590592, value=17	
date 20240217	column=DATE:dayOfWeek,
timestamp=1731231590598, value=Saturday	
date 20240217	column=DATE:month,
timestamp=1731231590589, value=02	
date 20240217	column=DATE:quarter,
timestamp=1731231590595, value=01	
date 20240217	column=DATE:year,
timestamp=1731231590587, value=2024	
date 20240222	column=DATE:day,
timestamp=1731231590614, value=22	
date 20240222	column=DATE:dayOfWeek,
timestamp=1731231590627, value=Thursday	
date 20240222	column=DATE:month,
timestamp=1731231590610, value=02	

date 20240222	column=DATE:quarter,
timestamp=1731231590622, value=01	
date 20240222	column=DATE:year,
timestamp=1731231590606, value=2024	
date 20240304	column=DATE:day,
timestamp=1731231590643, value=04	
date 20240304	column=DATE:dayOfWeek,
timestamp=1731231590654, value=Wednesday	
date 20240304	column=DATE:month,
timestamp=1731231590638, value=03	
date 20240304	column=DATE:quarter,
timestamp=1731231590649, value=01	
date 20240304	column=DATE:year,
timestamp=1731231590632, value=2024	
date 20240306	column=DATE:day,
timestamp=1731231590664, value=06	
date 20240306	column=DATE:dayOfWeek,
timestamp=1731231590674, value=Wednesday	
date 20240306	column=DATE:month,
timestamp=1731231590661, value=03	
date 20240306	column=DATE:quarter,
timestamp=1731231590669, value=01	
date 20240306	column=DATE:year,
timestamp=1731231590657, value=2024	
date 20240330	column=DATE:day,
timestamp=1731231590687, value=30	
date 20240330	column=DATE:dayOfWeek,
timestamp=1731231590697, value=Saturday	
date 20240330	column=DATE:month,
timestamp=1731231590683, value=03	
date 20240330	column=DATE:quarter,
timestamp=1731231590692, value=01	
date 20240330	column=DATE:year,
timestamp=1731231590679, value=2024	
date 20240416	column=DATE:day,
timestamp=1731231590709, value=16	
date 20240416	column=DATE:dayOfWeek,
timestamp=1731231590720, value=Tuesday	
date 20240416	column=DATE:month,
timestamp=1731231590705, value=04	
date 20240416	column=DATE:quarter,
timestamp=1731231590715, value=01	
date 20240416	column=DATE:year,
timestamp=1731231590701, value=2024	
date 20240427	column=DATE:day,
timestamp=1731231590733, value=27	
date 20240427	column=DATE:dayOfWeek,
timestamp=1731231590745, value=Saturday	
date 20240427	column=DATE:month,
timestamp=1731231590728, value=04	
date 20240427	column=DATE:quarter,
timestamp=1731231590737, value=01	
date 20240427	column=DATE:year,
timestamp=1731231590724, value=2024	

date 20240428	column=DATE:day,
timestamp=1731231590764, value=28	
date 20240428	column=DATE:dayOfWeek,
timestamp=1731231590776, value=Sunday	
date 20240428	column=DATE:month,
timestamp=1731231590759, value=04	
date 20240428	column=DATE:quarter,
timestamp=1731231590771, value=01	
date 20240428	column=DATE:year,
timestamp=1731231590754, value=2024	
sale001	column=SALES-
TRX:quantitySold, timestamp=1731232342281, value=10	
sale001	column=SALES-TRX:seller,
timestamp=1731232342267, value=seller1	
sale001	column=SALES-
TRX:totalSales, timestamp=1731232342293, value=200	
sale002	column=SALES-
TRX:quantitySold, timestamp=1731232342341, value=5	
sale002	column=SALES-TRX:seller,
timestamp=1731232342328, value=seller2	
sale002	column=SALES-
TRX:totalSales, timestamp=1731232342359, value=150	
salestrx 007:101:30012024	
column=CUSTOMER:customerId, timestamp=1731231590839, value=007	
salestrx 007:101:30012024	column=DATE:sDay,
timestamp=1731231590833, value=30	
salestrx 007:101:30012024	column=DATE:sMonth,
timestamp=1731231590829, value=01	
salestrx 007:101:30012024	column=DATE:sYear,
timestamp=1731231590824, value=2024	
salestrx 007:101:30012024	
column=PRODUCT:productId, timestamp=1731231590845, value=101	
salestrx 007:101:30012024	column=SALES-
TRX:discount, timestamp=1731231590858, value=0.2	
salestrx 007:101:30012024	column=SALES-
TRX:quantitySold, timestamp=1731231590849, value=2	
salestrx 007:101:30012024	column=SALES-
TRX:salesId, timestamp=1731231590820, value=20240102	
salestrx 007:101:30012024	column=SALES-
TRX:totalSale, timestamp=1731231590853, value=238	
salestrx 007:202:28012024	
column=CUSTOMER:customerId, timestamp=1731231590795, value=007	
salestrx 007:202:28012024	column=DATE:sDay,
timestamp=1731231590791, value=28	
salestrx 007:202:28012024	column=DATE:sMonth,
timestamp=1731231590787, value=01	
salestrx 007:202:28012024	column=DATE:sYear,
timestamp=1731231590784, value=2024	
salestrx 007:202:28012024	
column=PRODUCT:productId, timestamp=1731231590799, value=202	
salestrx 007:202:28012024	column=SALES-
TRX:discount, timestamp=1731231590816, value=0.1	
salestrx 007:202:28012024	column=SALES-
TRX:quantitySold, timestamp=1731231590802, value=1	

salestrx 007:202:28012024	column=SALES-
TRX:salesId, timestamp=1731231590780, value=20240101	
salestrx 007:202:28012024	column=SALES-
TRX:totalSale, timestamp=1731231590811, value=2100	
salestrx 666:102:22022024	
column=CUSTOMER:customerId, timestamp=1731231590938, value=666	
salestrx 666:102:22022024	column=DATE:sDay,
timestamp=1731231590935, value=22	
salestrx 666:102:22022024	column=DATE:sMonth,
timestamp=1731231590931, value=02	
salestrx 666:102:22022024	column=DATE:sYear,
timestamp=1731231590927, value=2024	
salestrx 666:102:22022024	
column=PRODUCT:productId, timestamp=1731231590943, value=102	
salestrx 666:102:22022024	column=SALES-
TRX:discount, timestamp=1731231590966, value=0.2	
salestrx 666:102:22022024	column=SALES-
TRX:quantitySold, timestamp=1731231590947, value=1	
salestrx 666:102:22022024	column=SALES-
TRX:salesId, timestamp=1731231590923, value=20240201	
salestrx 666:102:22022024	column=SALES-
TRX:totalSale, timestamp=1731231590951, value=1600	
salestrx 888:302:06032024	
column=CUSTOMER:customerId, timestamp=1731231590897, value=888	
salestrx 888:302:06032024	column=DATE:sDay,
timestamp=1731231590893, value=06	
salestrx 888:302:06032024	column=DATE:sMonth,
timestamp=1731231590885, value=03	
salestrx 888:302:06032024	column=DATE:sYear,
timestamp=1731231590872, value=2024	
salestrx 888:302:06032024	
column=PRODUCT:productId, timestamp=1731231590902, value=302	
salestrx 888:302:06032024	column=SALES-
TRX:discount, timestamp=1731231590918, value=0.3	
salestrx 888:302:06032024	column=SALES-
TRX:quantitySold, timestamp=1731231590908, value=2	
salestrx 888:302:06032024	column=SALES-
TRX:salesId, timestamp=1731231590868, value=20240301	
salestrx 888:302:06032024	column=SALES-
TRX:totalSale, timestamp=1731231590912, value=800	
salestrx 888:303:28042024	
column=CUSTOMER:customerId, timestamp=1731231590984, value=888	
salestrx 888:303:28042024	column=DATE:sDay,
timestamp=1731231590980, value=28	
salestrx 888:303:28042024	column=DATE:sMonth,
timestamp=1731231590977, value=04	
salestrx 888:303:28042024	column=DATE:sYear,
timestamp=1731231590974, value=2024	
salestrx 888:303:28042024	
column=PRODUCT:productId, timestamp=1731231590988, value=303	
salestrx 888:303:28042024	column=SALES-
TRX:discount, timestamp=1731231591002, value=0.4	
salestrx 888:303:28042024	column=SALES-
TRX:quantitySold, timestamp=1731231590992, value=5	

```

    salestrx|888:303:28042024                                column=SALES-
TRX:salesId, timestamp=1731231590970, value=20240401
    salestrx|888:303:28042024                                column=SALES-
TRX:totalSale, timestamp=1731231590996, value=500
    seller1                                                    column=SELLER:address,
timestamp=1731232342055, value=123 Seller St
    seller1                                                    column=SELLER:contact,
timestamp=1731232342086, value=John Doe
    seller1                                                    column=SELLER:name,
timestamp=1731232342046, value=Seller A
    seller2                                                    column=SELLER:address,
timestamp=1731232342131, value=456 Market Rd
    seller2                                                    column=SELLER:contact,
timestamp=1731232342143, value=Jane Smith
    seller2                                                    column=SELLER:name,
timestamp=1731232342119, value=Seller Two
24 row(s) in 0.4090 seconds

solution2.hb(main):024:0>
solution2.hb(main):025:0*
solution2.hb(main):026:0* #Task 1: Add a new column family named SELLER that
will store seller information (name, address, and contact person) and insert
data for at least two sellers.
solution2.hb(main):027:0* alter 'task2', NAME => 'SELLER'
Updating all regions with the new schema...
0/1 regions updated.
1/1 regions updated.
Done.
0 row(s) in 3.0790 seconds

solution2.hb(main):028:0> put 'task2', 'seller1', 'SELLER:name', 'Seller A'
0 row(s) in 0.1530 seconds

solution2.hb(main):029:0> put 'task2', 'seller1', 'SELLER:address', '123
Seller St'
0 row(s) in 0.0120 seconds

solution2.hb(main):030:0> put 'task2', 'seller1', 'SELLER:contact', 'John
Doe'
0 row(s) in 0.0070 seconds

solution2.hb(main):031:0> put 'task2', 'seller2', 'SELLER:name', 'Seller Two'
0 row(s) in 0.0220 seconds

solution2.hb(main):032:0> put 'task2', 'seller2', 'SELLER:address', '456
Market Rd'
0 row(s) in 0.0260 seconds

solution2.hb(main):033:0> put 'task2', 'seller2', 'SELLER:contact', 'Jane
Smith'
0 row(s) in 0.0250 seconds

solution2.hb(main):034:0>
solution2.hb(main):035:0* # Task 1 - Scan to verify seller information
solution2.hb(main):036:0* scan 'task2', {COLUMNS => ['SELLER']}

```

ROW	COLUMN+CELL
seller1	column=SELLER:address,
timestamp=1731232409017, value=123 Seller St	
seller1	column=SELLER:contact,
timestamp=1731232409040, value=John Doe	
seller1	column=SELLER:name,
timestamp=1731232408975, value=Seller A	
seller2	column=SELLER:address,
timestamp=1731232409168, value=456 Market Rd	
seller2	column=SELLER:contact,
timestamp=1731232409185, value=Jane Smith	
seller2	column=SELLER:name,
timestamp=1731232409135, value=Seller Two	
2 row(s) in 0.0830 seconds	
solution2.hb(main):037:0>	
solution2.hb(main):038:0*	
solution2.hb(main):039:0* #Task 2: Add information for two sales	
transactions, each by a different seller, with reasonable values.	
solution2.hb(main):040:0* put 'task2', 'sale001', 'SALES-TRX:seller',	
'seller1'	
0 row(s) in 0.0520 seconds	
solution2.hb(main):041:0> put 'task2', 'sale001', 'SALES-TRX:quantitySold',	
'10'	
0 row(s) in 0.0080 seconds	
solution2.hb(main):042:0> put 'task2', 'sale001', 'SALES-TRX:totalSales',	
'200'	
0 row(s) in 0.0160 seconds	
solution2.hb(main):043:0>	
solution2.hb(main):044:0* put 'task2', 'sale002', 'SALES-TRX:seller',	
'seller2'	
0 row(s) in 0.0120 seconds	
solution2.hb(main):045:0> put 'task2', 'sale002', 'SALES-TRX:quantitySold',	
'5'	
0 row(s) in 0.0100 seconds	
solution2.hb(main):046:0> put 'task2', 'sale002', 'SALES-TRX:totalSales',	
'150'	
0 row(s) in 0.0050 seconds	
solution2.hb(main):047:0>	
solution2.hb(main):048:0* # Task 2 - Scan to verify sales transaction data	
solution2.hb(main):049:0* scan 'task2', {FILTER => "PrefixFilter('sale')",	
COLUMNS => ['SALES-TRX']}	
ROW	COLUMN+CELL
sale001	column=SALES-
TRX:quantitySold, timestamp=1731232409389, value=10	
sale001	column=SALES-TRX:seller,
timestamp=1731232409350, value=seller1	
sale001	column=SALES-
TRX:totalSales, timestamp=1731232409401, value=200	



sale002	column=SALES-
TRX:quantitySold, timestamp=1731232409464, value=5	
sale002	column=SALES-TRX:seller,
timestamp=1731232409424, value=seller2	
sale002	column=SALES-
TRX:totalSales, timestamp=1731232409480, value=150	
salestrx 007:101:30012024	column=SALES-
TRX:discount, timestamp=1731231590858, value=0.2	
salestrx 007:101:30012024	column=SALES-
TRX:quantitySold, timestamp=1731231590849, value=2	
salestrx 007:101:30012024	column=SALES-
TRX:salesId, timestamp=1731231590820, value=20240102	
salestrx 007:101:30012024	column=SALES-
TRX:totalSale, timestamp=1731231590853, value=238	
salestrx 007:202:28012024	column=SALES-
TRX:discount, timestamp=1731231590816, value=0.1	
salestrx 007:202:28012024	column=SALES-
TRX:quantitySold, timestamp=1731231590802, value=1	
salestrx 007:202:28012024	column=SALES-
TRX:salesId, timestamp=1731231590780, value=20240101	
salestrx 007:202:28012024	column=SALES-
TRX:totalSale, timestamp=1731231590811, value=2100	
salestrx 666:102:22022024	column=SALES-
TRX:discount, timestamp=1731231590966, value=0.2	
salestrx 666:102:22022024	column=SALES-
TRX:quantitySold, timestamp=1731231590947, value=1	
salestrx 666:102:22022024	column=SALES-
TRX:salesId, timestamp=1731231590923, value=20240201	
salestrx 666:102:22022024	column=SALES-
TRX:totalSale, timestamp=1731231590951, value=1600	
salestrx 888:302:06032024	column=SALES-
TRX:discount, timestamp=1731231590918, value=0.3	
salestrx 888:302:06032024	column=SALES-
TRX:quantitySold, timestamp=1731231590908, value=2	
salestrx 888:302:06032024	column=SALES-
TRX:salesId, timestamp=1731231590868, value=20240301	
salestrx 888:302:06032024	column=SALES-
TRX:totalSale, timestamp=1731231590912, value=800	
salestrx 888:303:28042024	column=SALES-
TRX:discount, timestamp=1731231591002, value=0.4	
salestrx 888:303:28042024	column=SALES-
TRX:quantitySold, timestamp=1731231590992, value=5	
salestrx 888:303:28042024	column=SALES-
TRX:salesId, timestamp=1731231590970, value=20240401	
salestrx 888:303:28042024	column=SALES-
TRX:totalSale, timestamp=1731231590996, value=500	

7 row(s) in 0.0550 seconds

```

solution2.hb(main):050:0>
solution2.hb(main):051:0*
solution2.hb(main):052:0* #Task 3: Increase the total number of versions in
each cell of the PRODUCT column family to 4.
solution2.hb(main):053:0* alter 'task2', {NAME => 'PRODUCT', VERSIONS => 4}
Updating all regions with the new schema...
0/1 regions updated.

```

1/1 regions updated.  
Done.  
0 row(s) in 3.4560 seconds