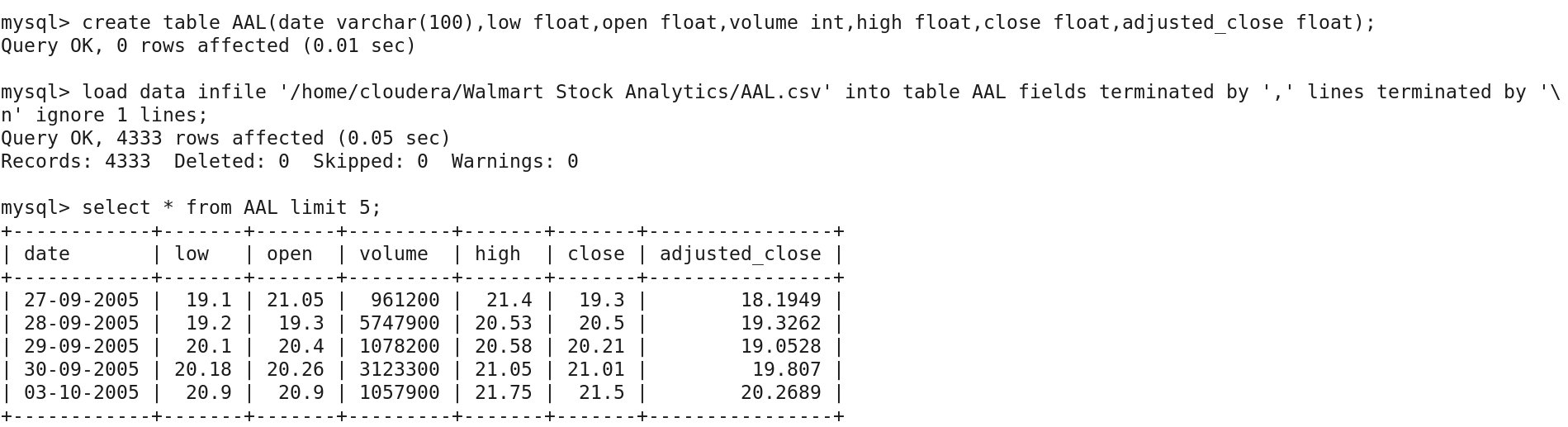
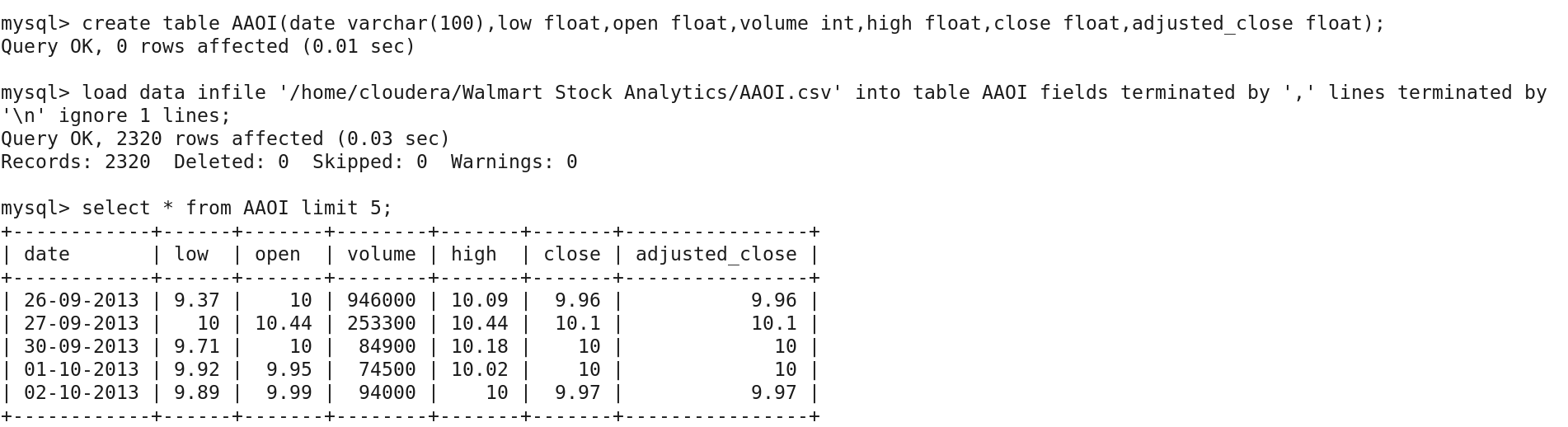
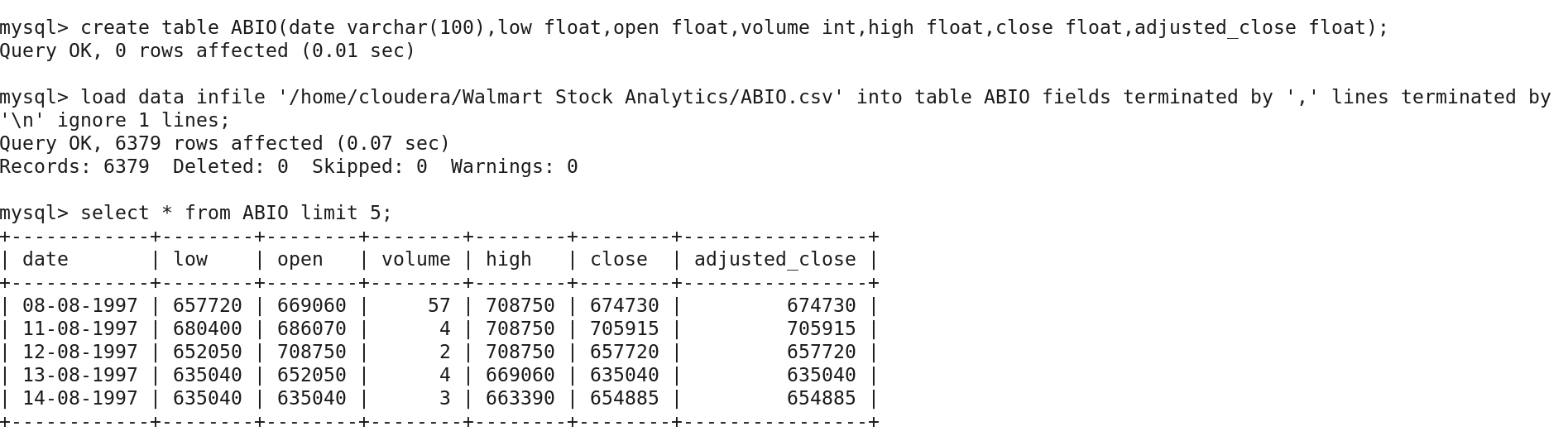
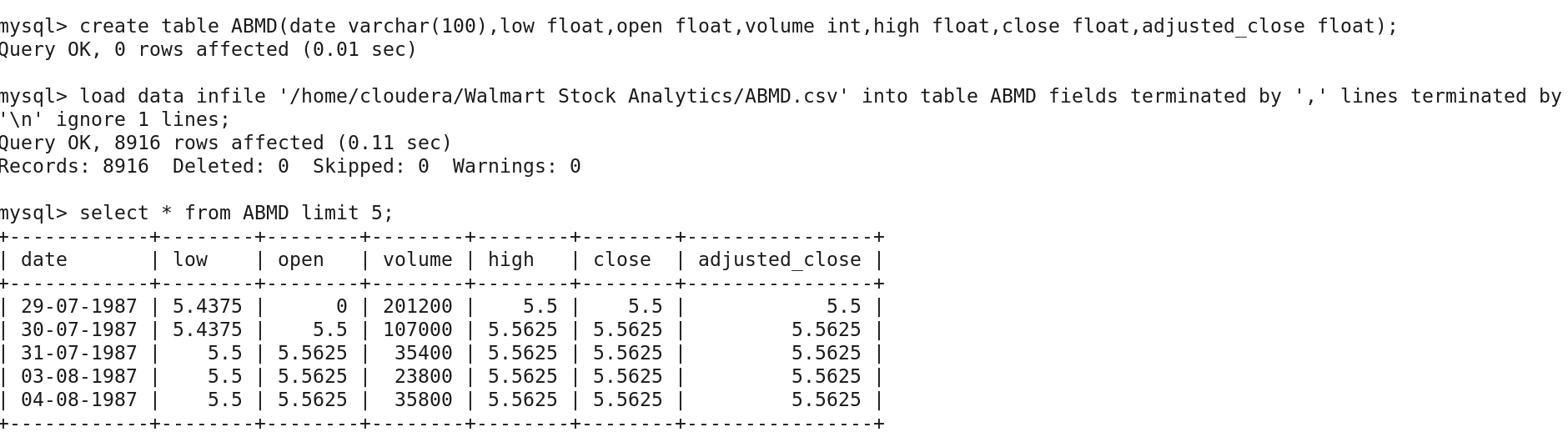
Creating tables

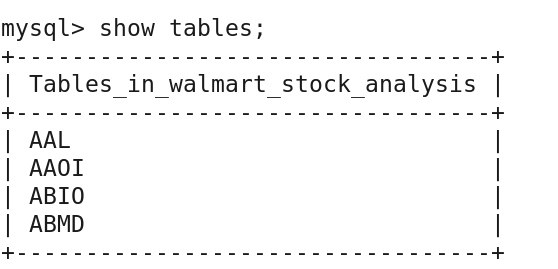








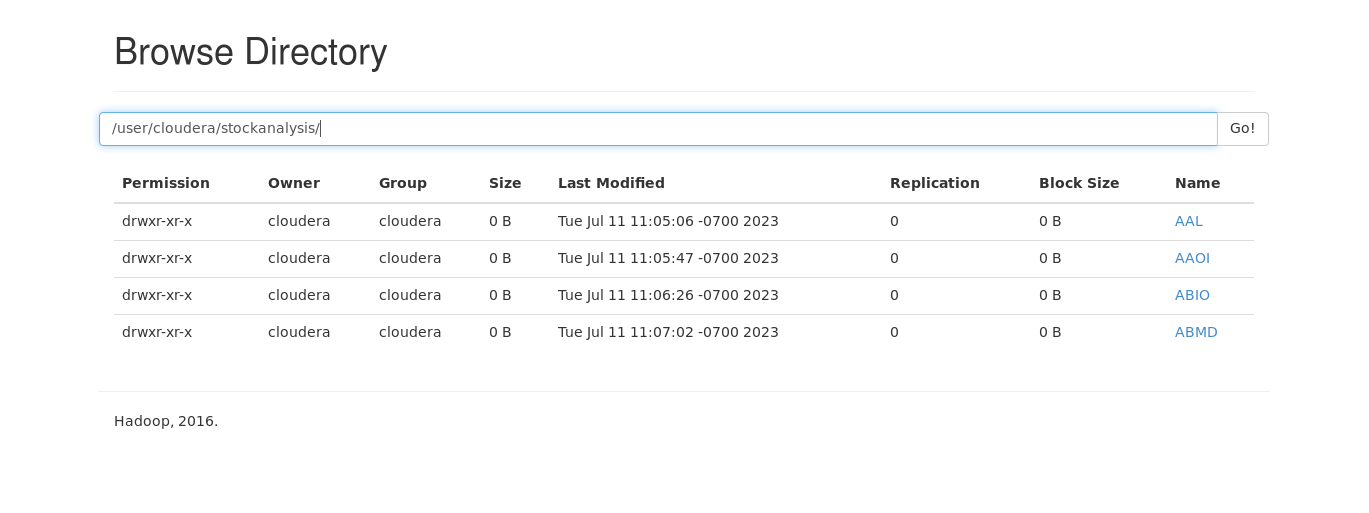
Tables in the Data Base



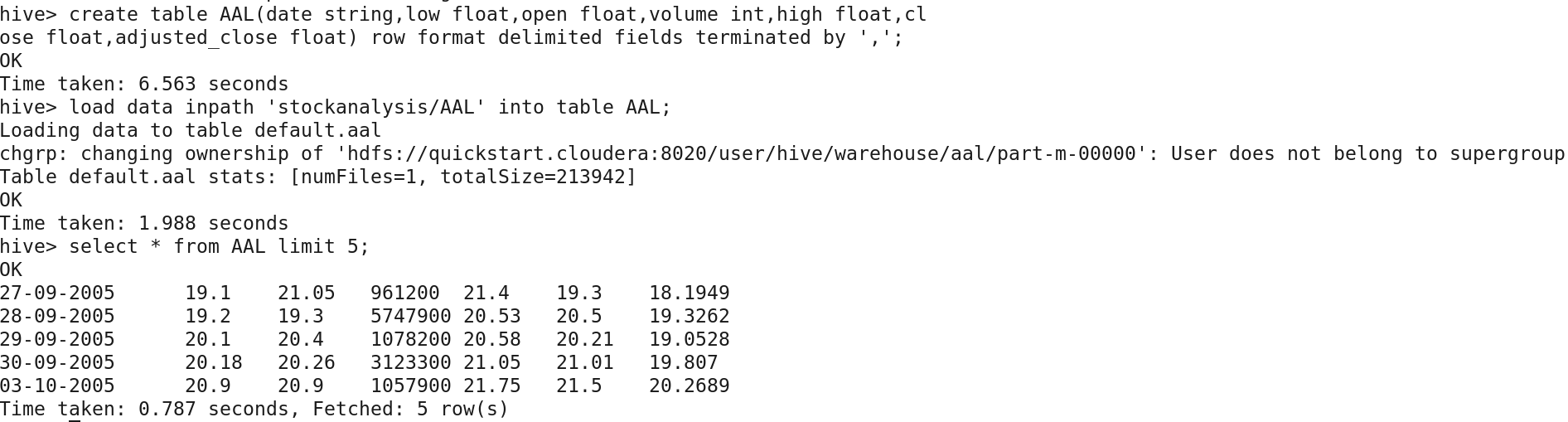
IMPORTING DATA FROM RDBMS TO HDFS USING SQOOP

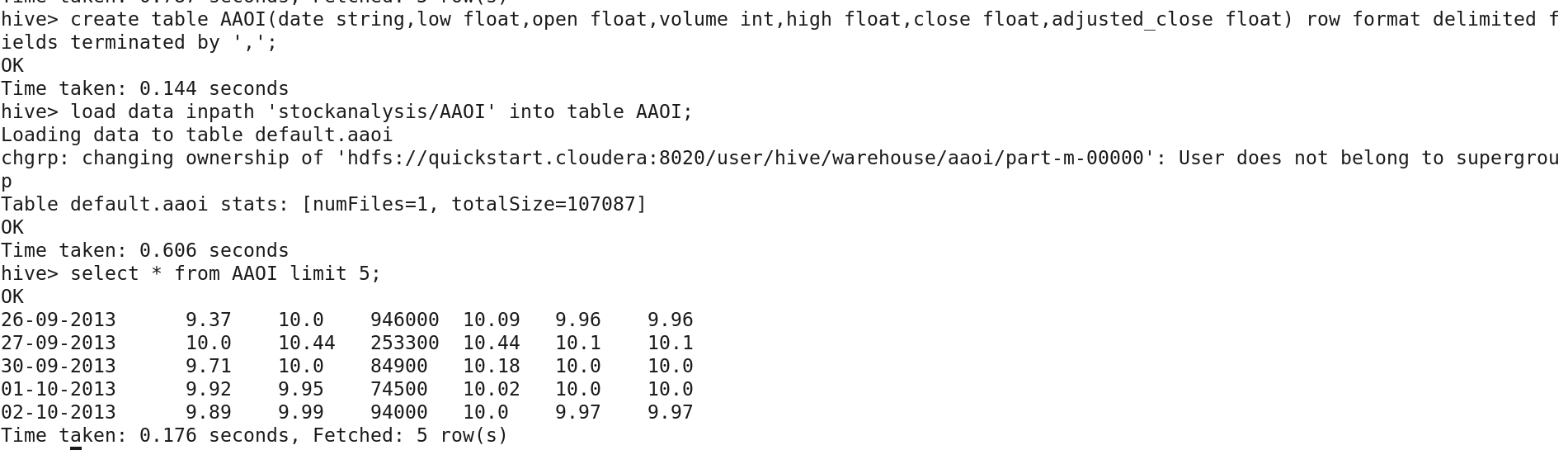


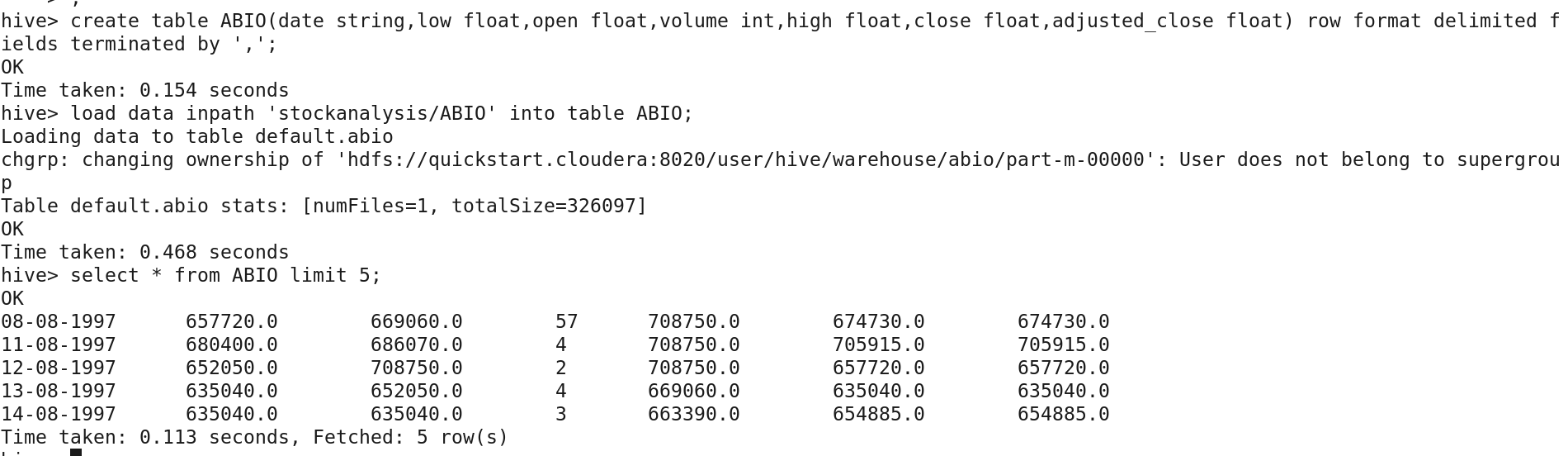
HDFS Directory :

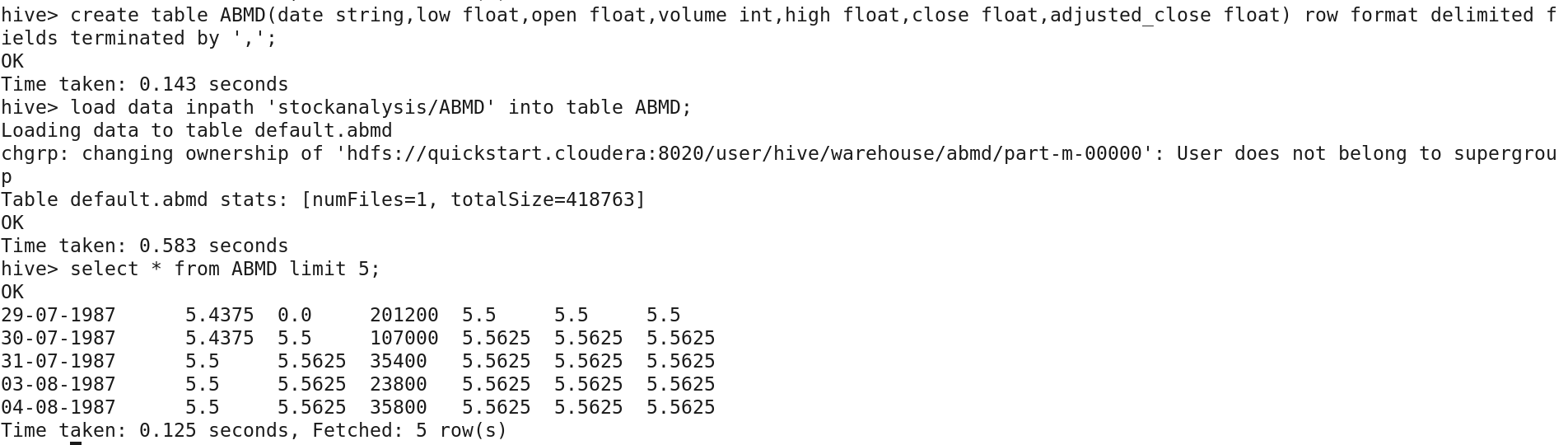


HDFS to HIVE WAREHOUSE :





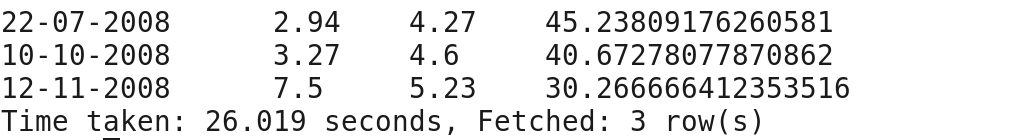




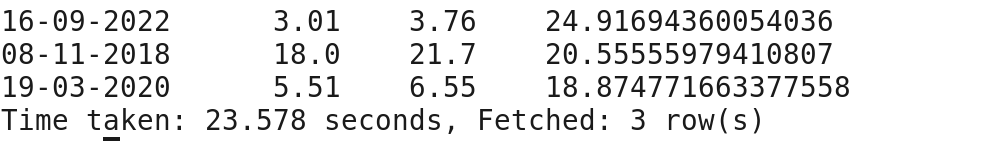
Scenario Solutions :

**PS 1 : Write a Hive query to identify the top three dates that experienced the largest percentage change in stock price (from open to close) for every stock.**

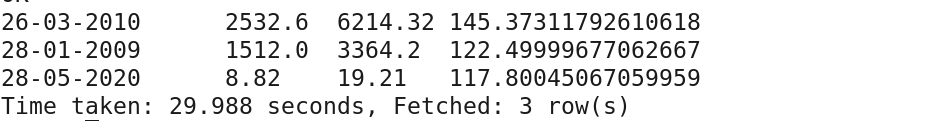




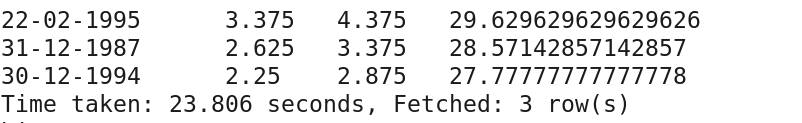








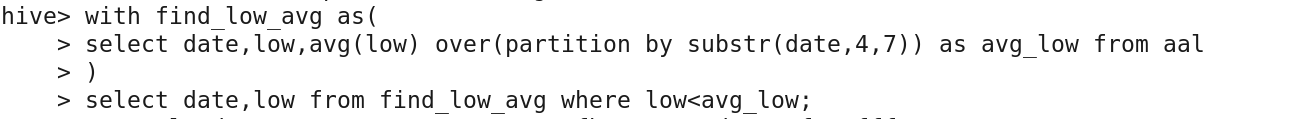




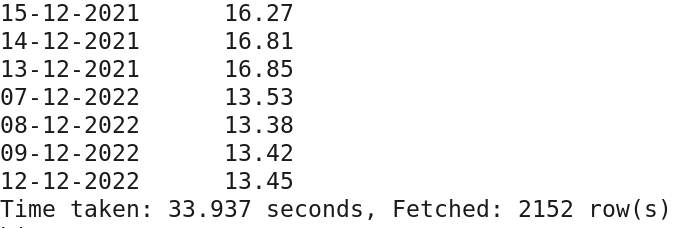
**PS2 :**

**write a Hive query to identify the dates where Low is less than average month low for every stock.**

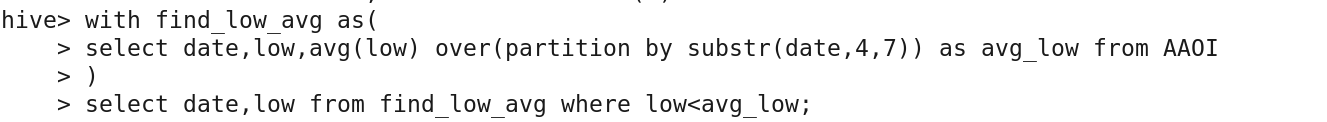
**AAL**



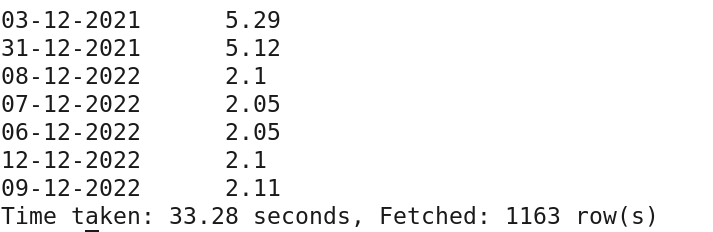
**SAMPLE OUTPUT :**



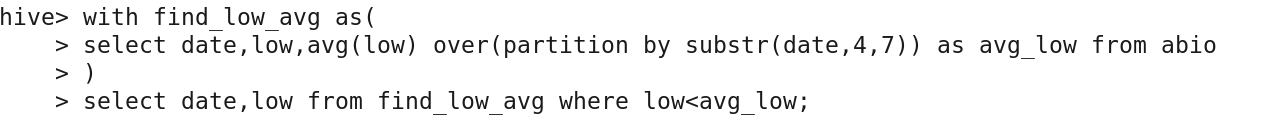
**AAOI**

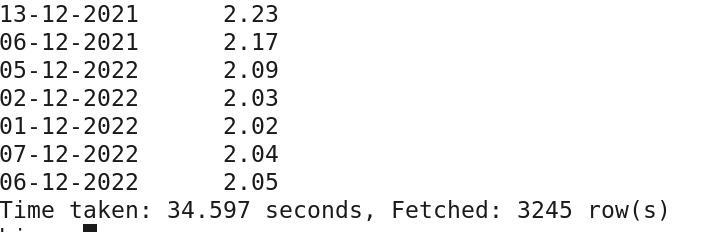


**SAMPLE OUTPUT :**

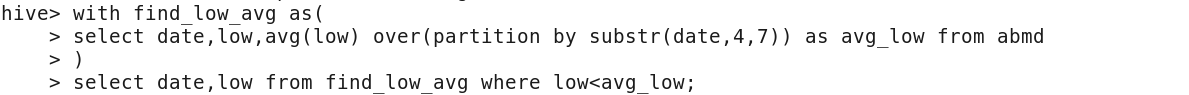


**ABIO**

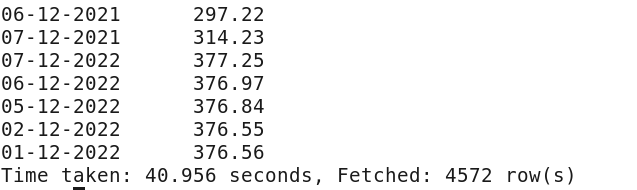




**ABMD**

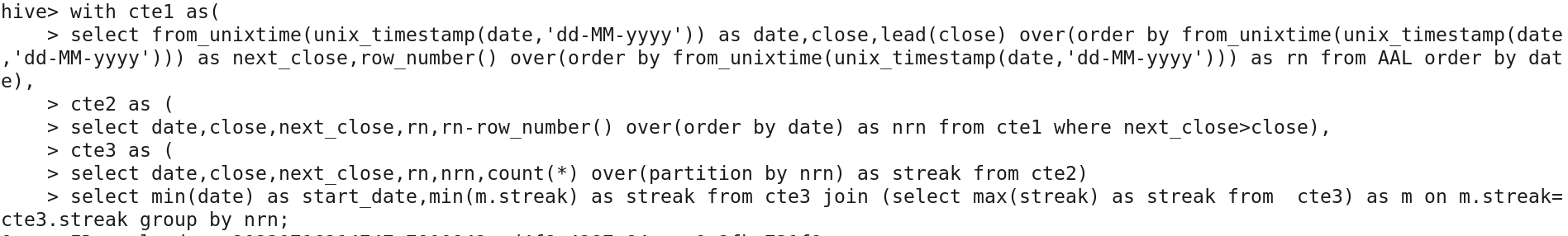


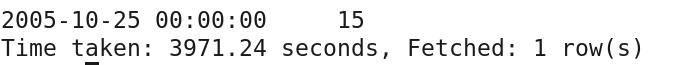
**SAMPLE OUTPUT :**



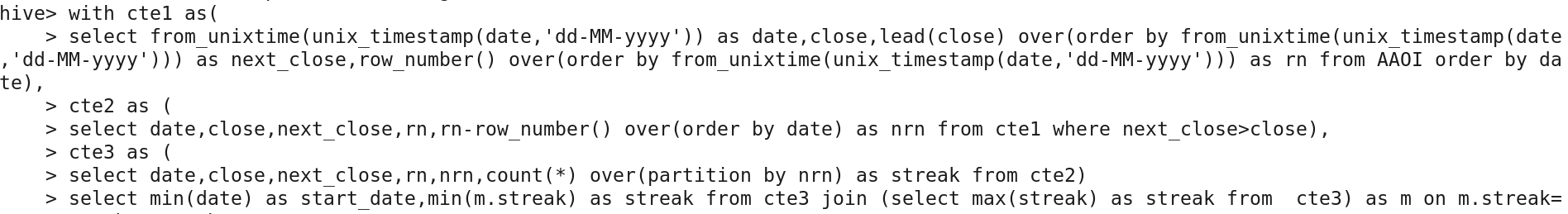
**PS 3 : Write a Hive query to find the date with the longest consecutive streak of increasing closing prices for every stock.**

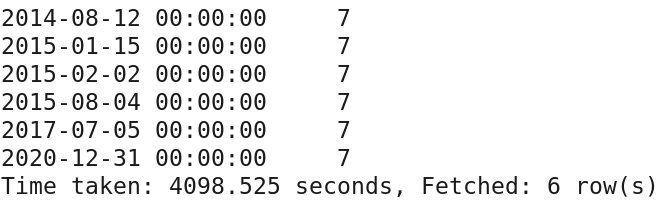
**AAL**

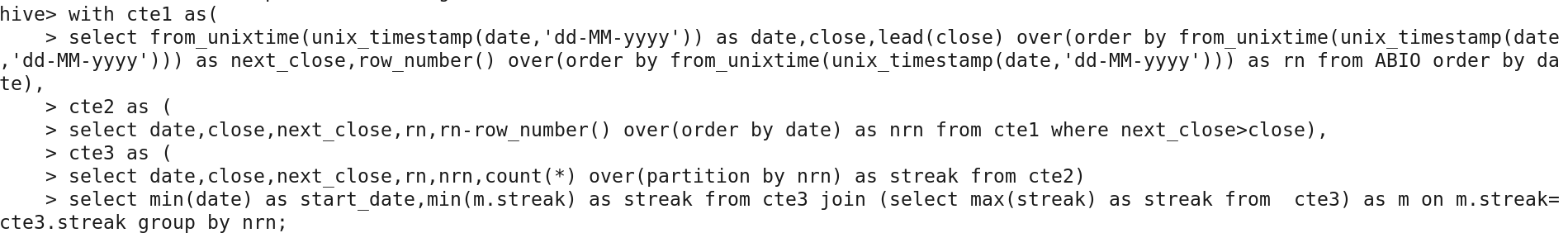


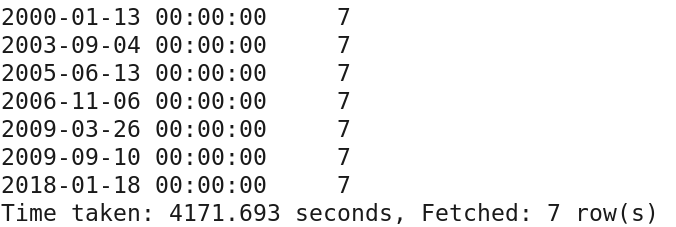


**AAOI**

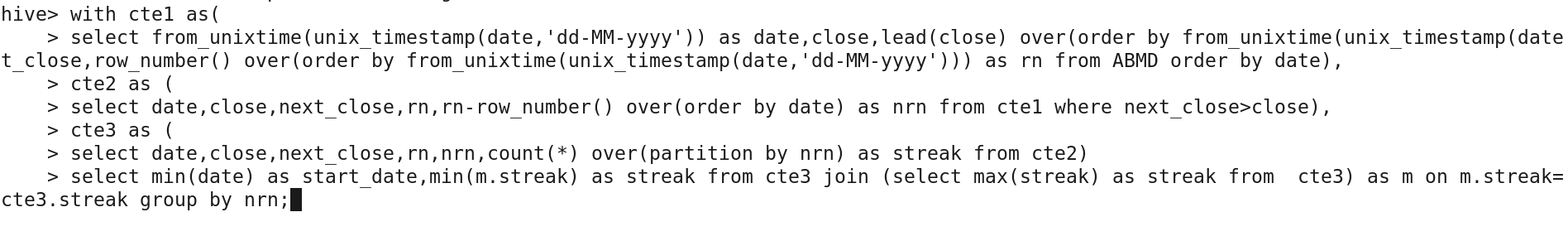


****

**ABIO**

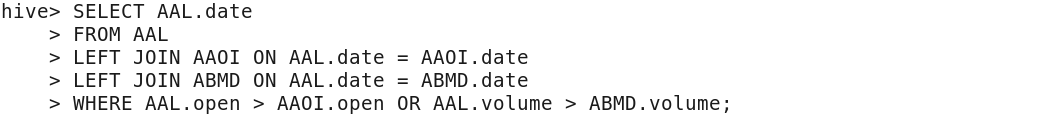


**ABMD**

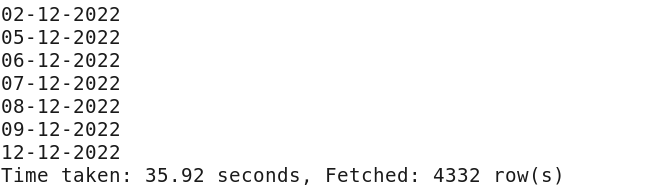




**PS 4 : write a Hive query to find the dates where AAL open price is higher than AAOI open price OR AAL volume greater than AMBD (write your query in an optimised way).**

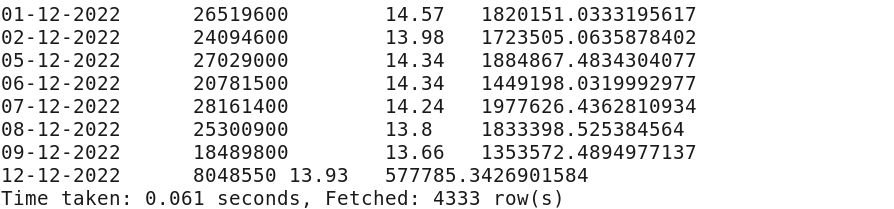
****

**SAMPLE OUTPUT :**

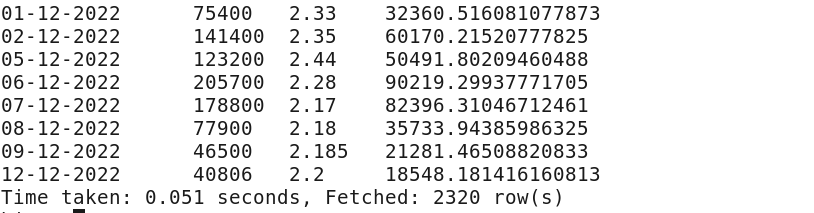


**PS 5: write a Hive query to calculate VH ratio(volume to hive ratio).**

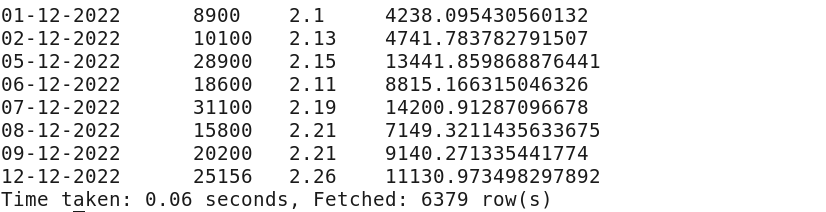




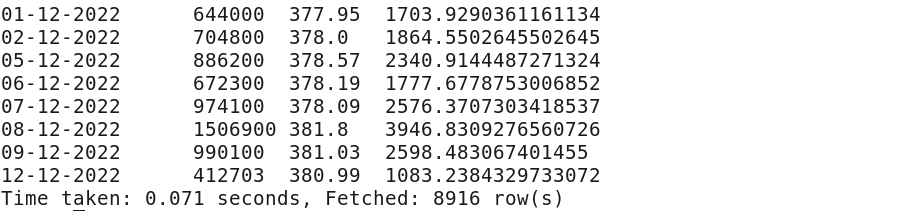










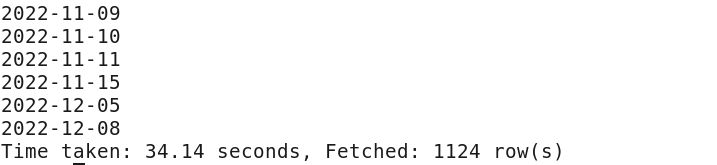


**PS 6 : 6. Write a Hive query to find the dates where previous day close and current day open difference is greater than 0 for each stock.**

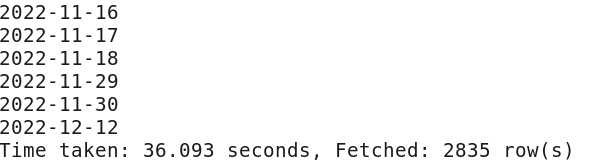
****

****

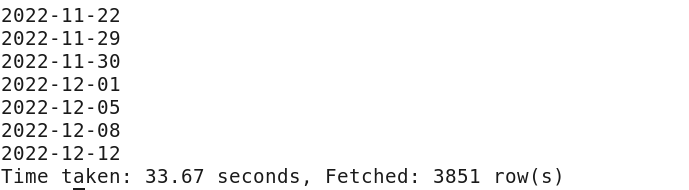
****

****

****

****

****

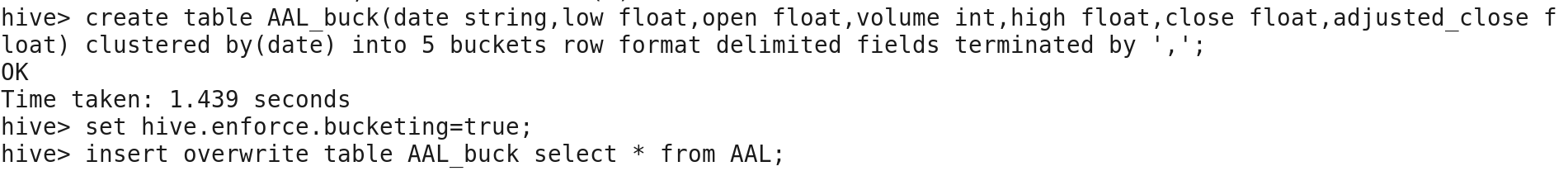
****

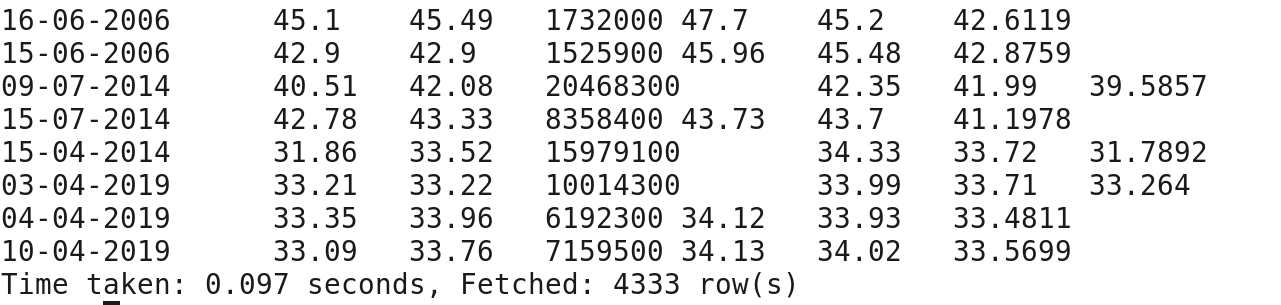
**PS 7 :** **Find median of volume for ABIO.**

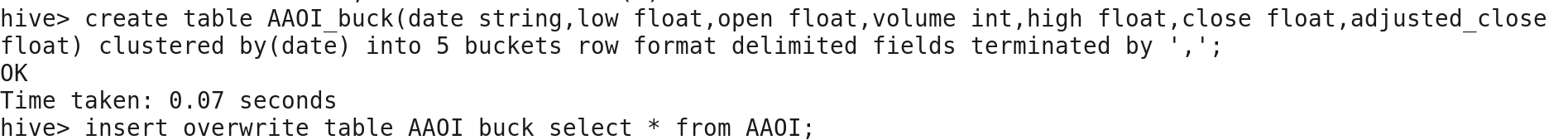
****

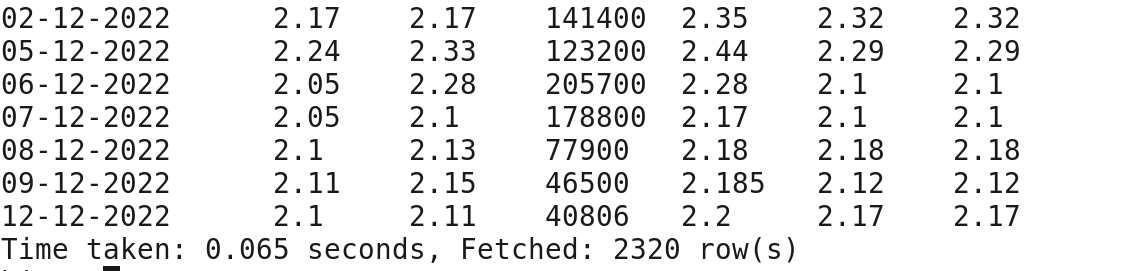


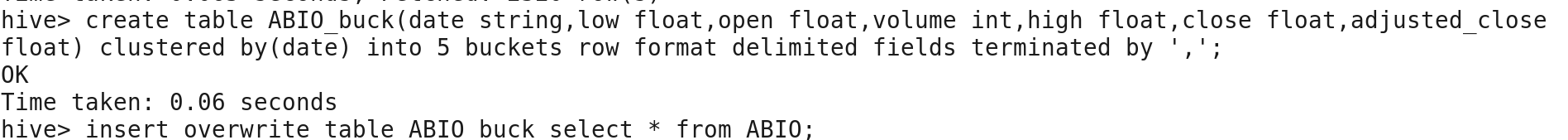
**CREATING BUCKING TABLES AND INSERTING DATA INTO IT**

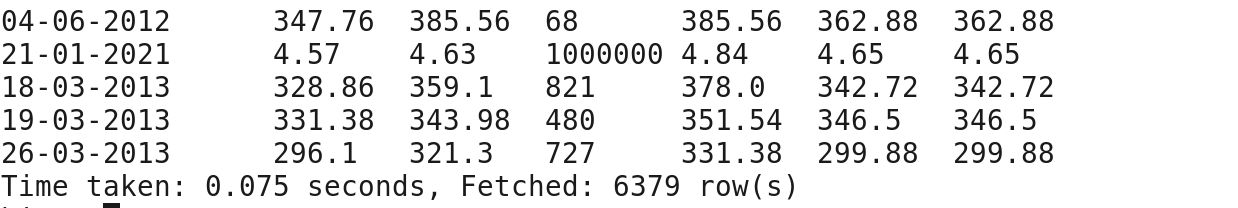


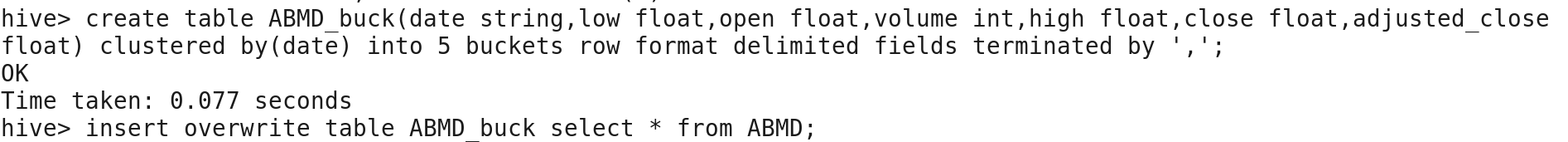
****

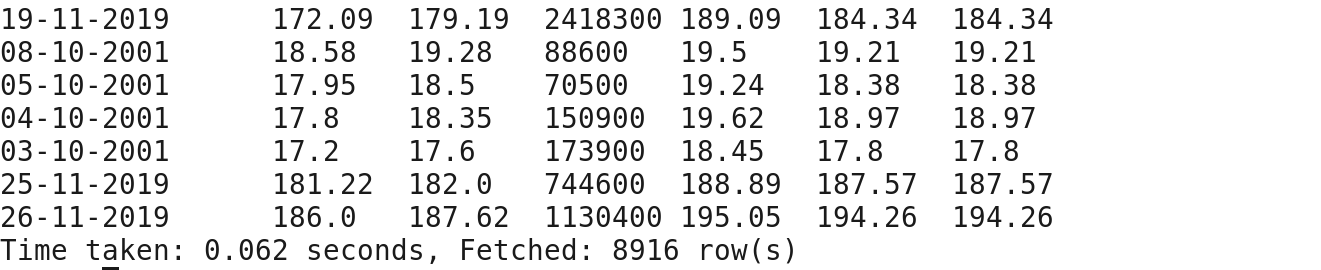
****

****

****

****

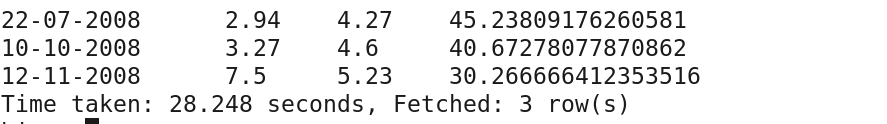
****

****

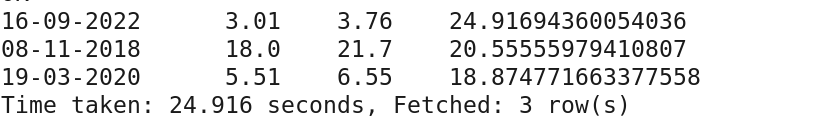
**ANALYZING DATA ON BUCKETED TABLES**

**PS 1 : Write a Hive query to identify the top three dates that experienced the largest percentage change in stock price (from open to close) for every stock.**

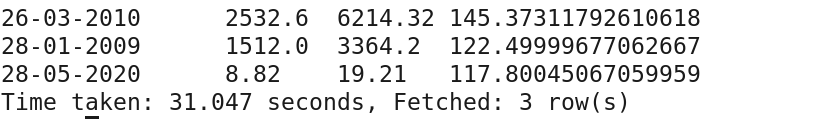
****

****

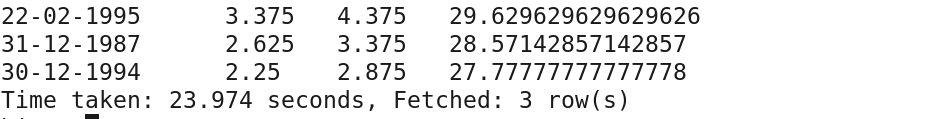
****

****



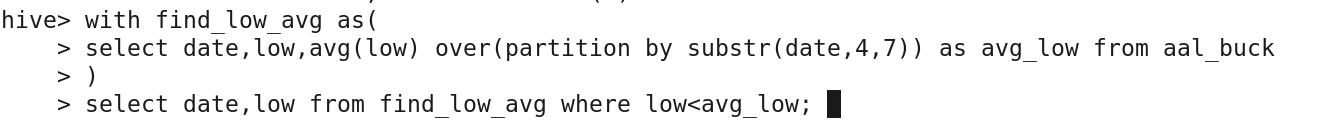


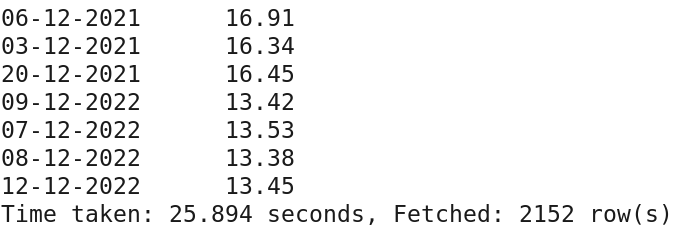
****

****

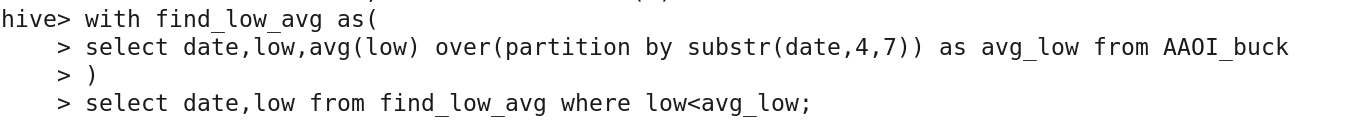
**PS2 : write a Hive query to identify the dates where Low is less than average month low for every stock.**

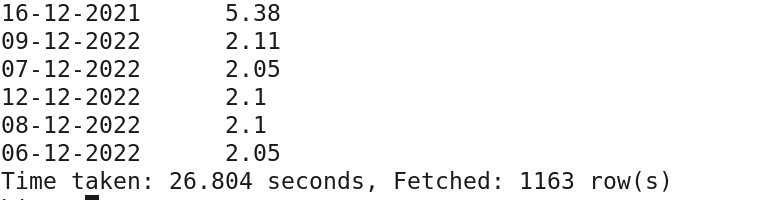
**AAL**

****

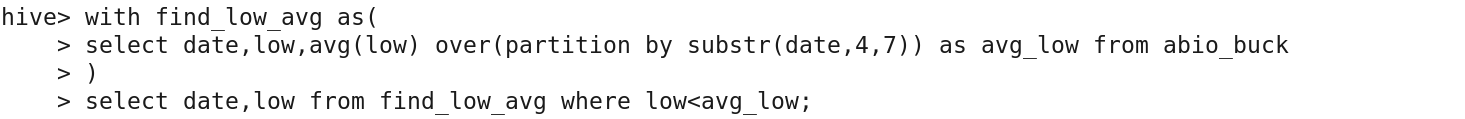
****

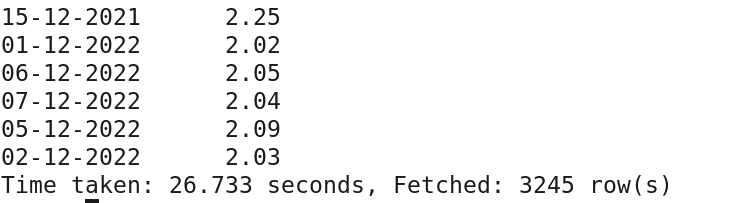
**AAOI**

****

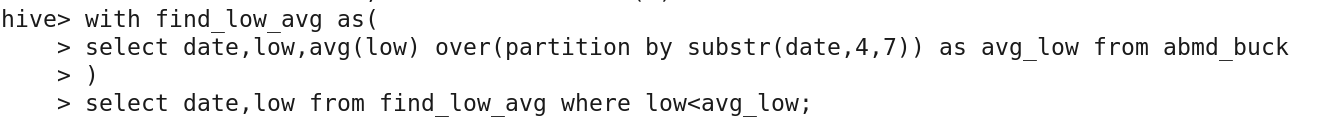
****

**ABIO**

****

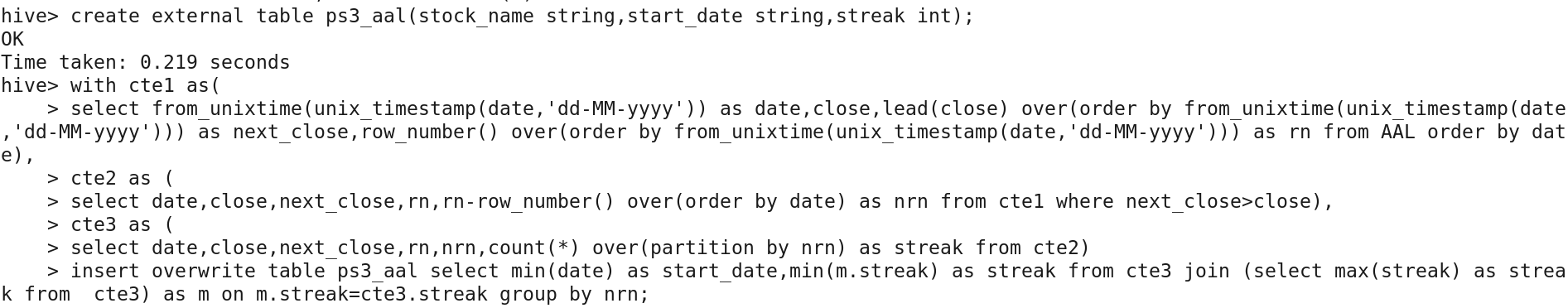
****

**ABMD**

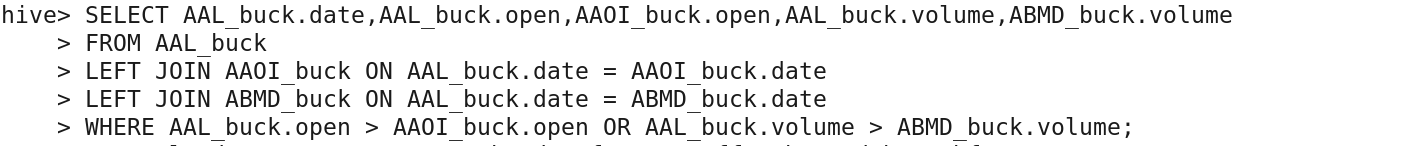
****

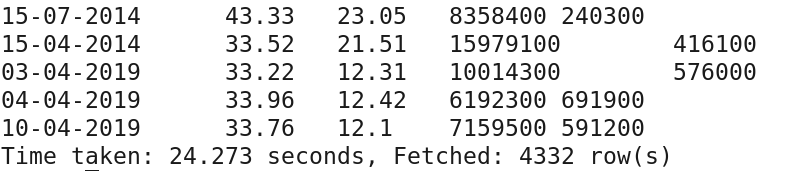
****

**PS 3 : Write a Hive query to find the date with the longest consecutive streak of increasing closing prices for every stock.**

****

**PS 4 : write a Hive query to find the dates where AAL open price is higher than AAOI open price OR AAL volume greater than AMBD (write your query in an optimised way).**

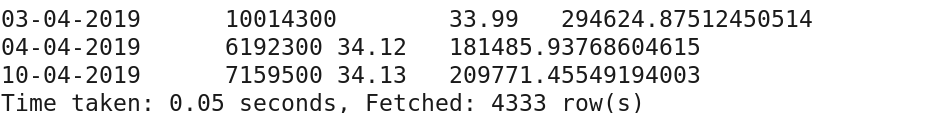
****

****

**PS 5: write a Hive query to calculate VH ratio(volume to hive ratio).**

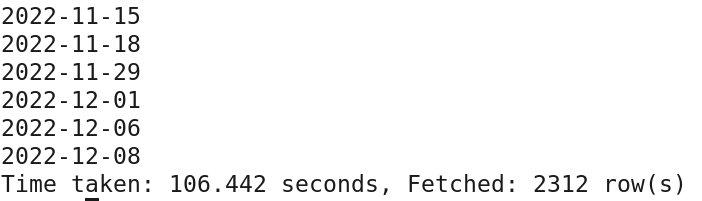
**AAL**

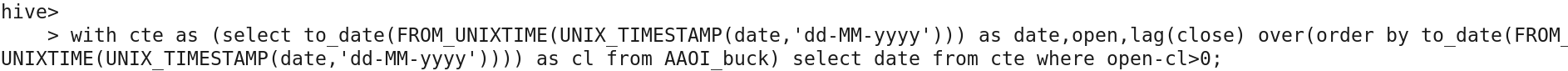
****

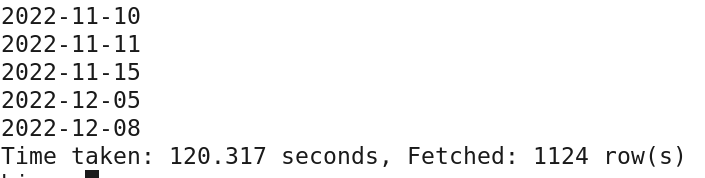
****

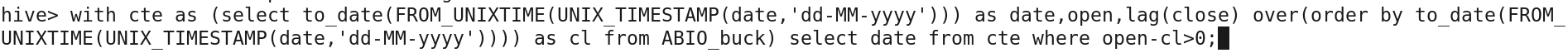
**PS 6 : 6. Write a Hive query to find the dates where previous day close and current day open difference is greater than 0 for each stock.**

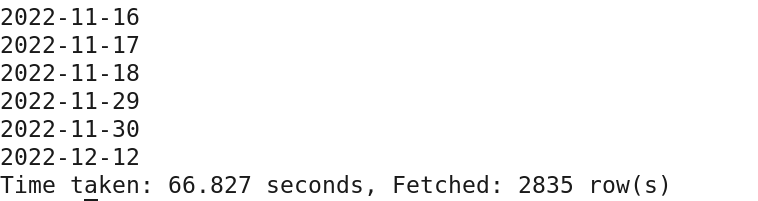
****

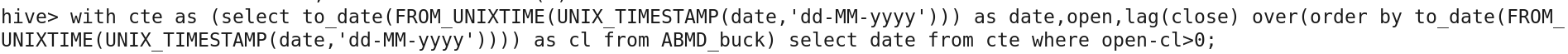
****

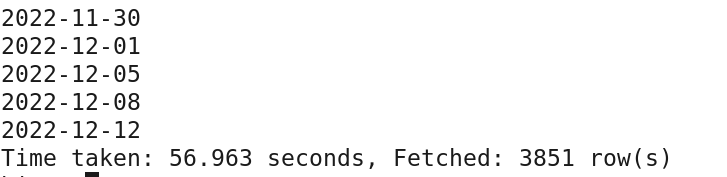
****

****

****

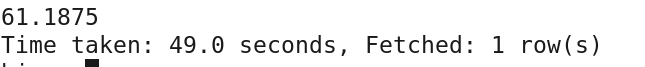
****

****

****

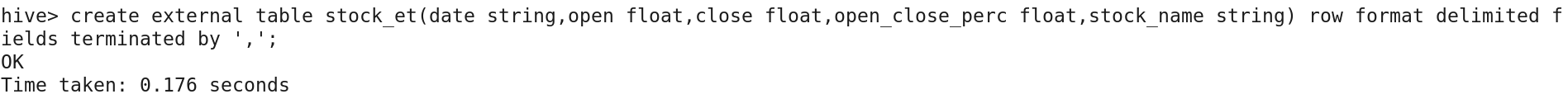
**PS 7 :** **Find median of volume for ABIO.**

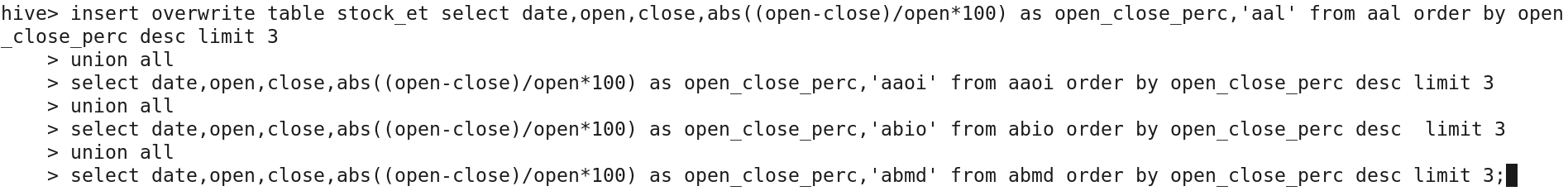
****

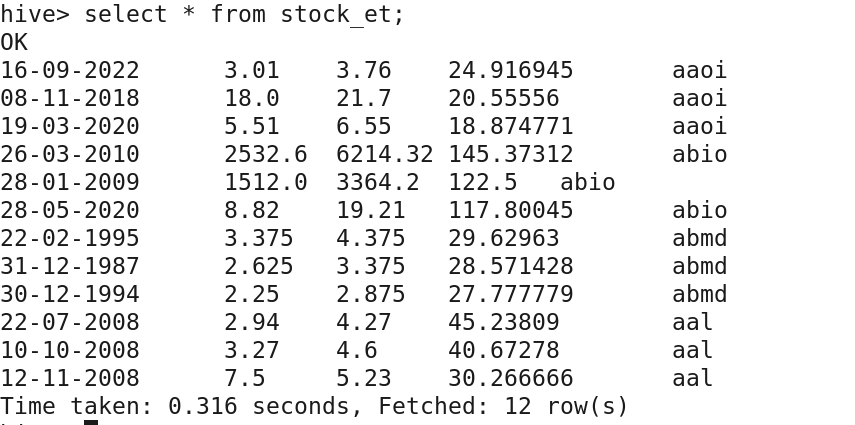
****

**CREATING EXTERNAL TABLES AND INSERTING OUTPUT DATA :**

**PS 1 :**

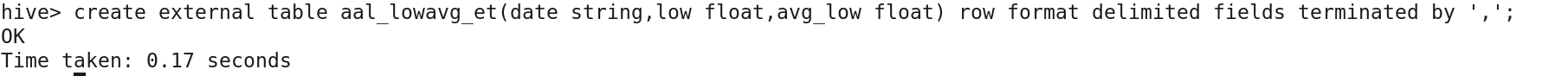
****

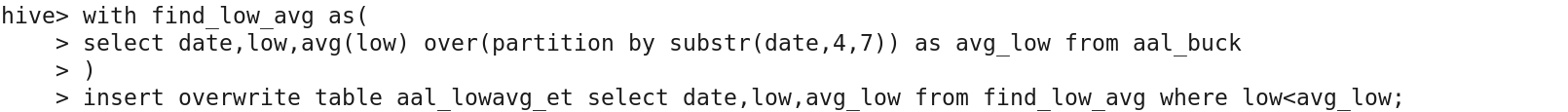
****

****

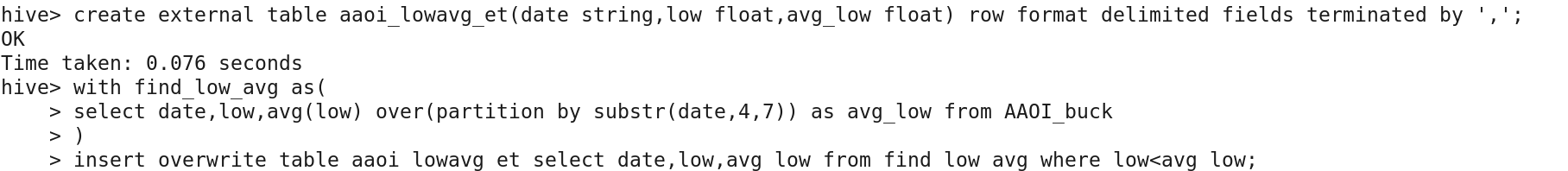
**PS 2:**

**AAL**

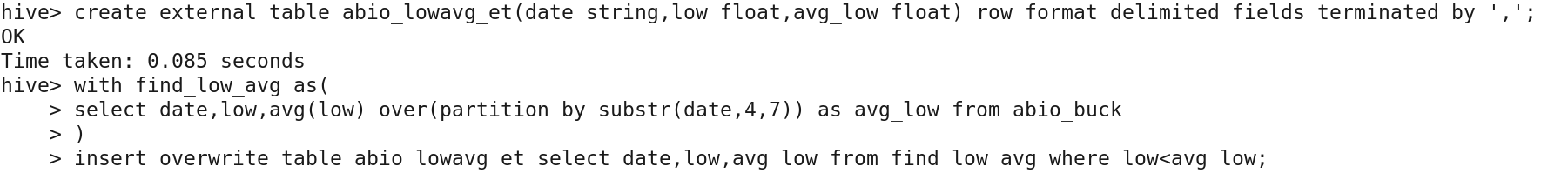
****

****

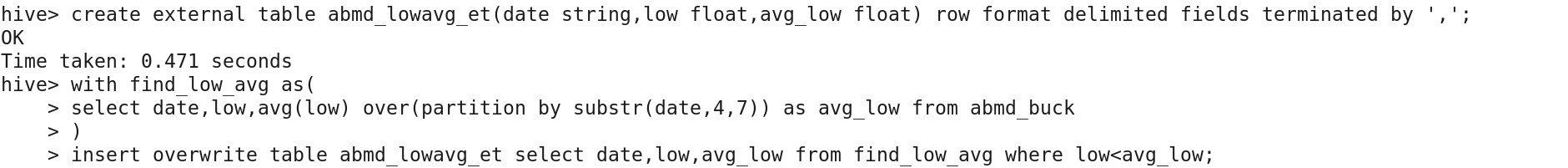
**AAOI**

****

**ABIO**

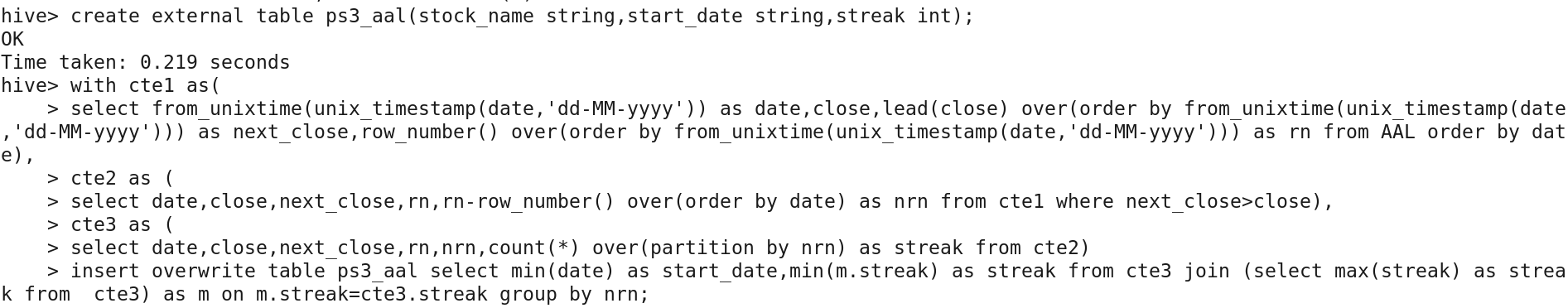
****

**ABMD**

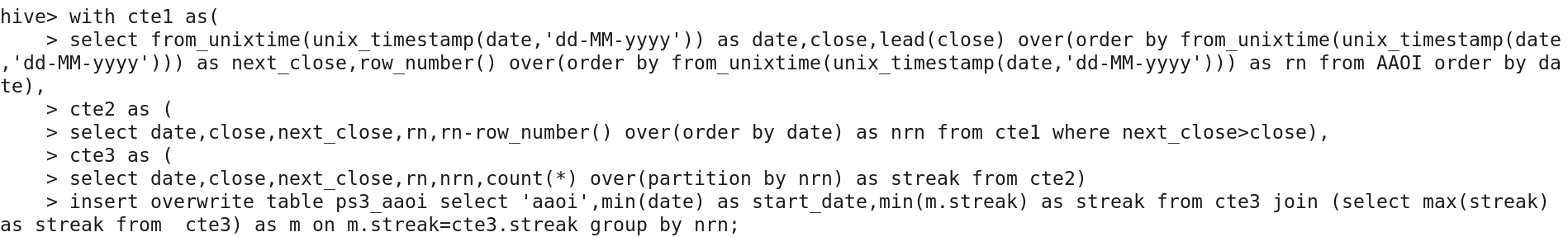
****

**PS 3:**

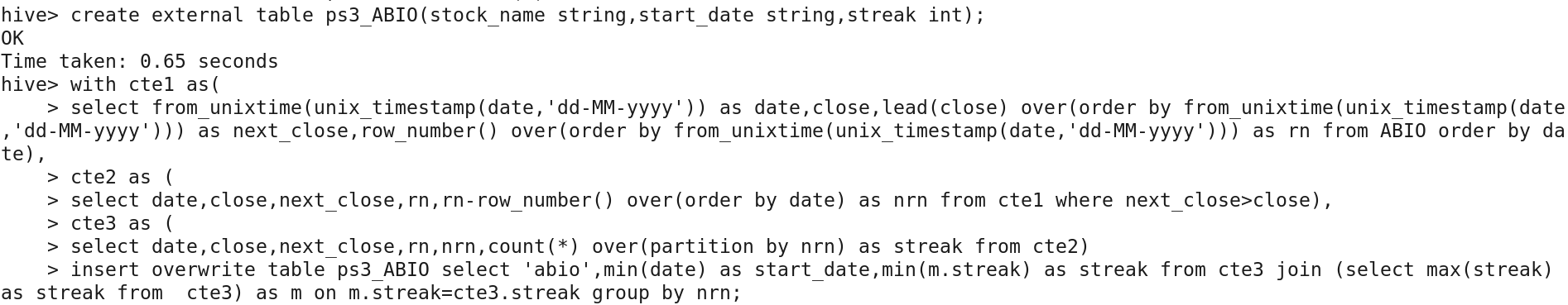
**AAL**

****

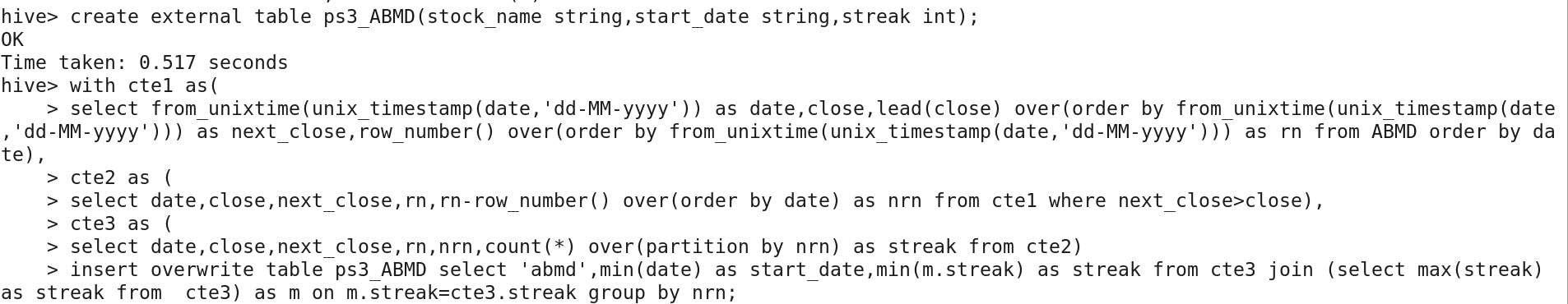
**AAOI**

****

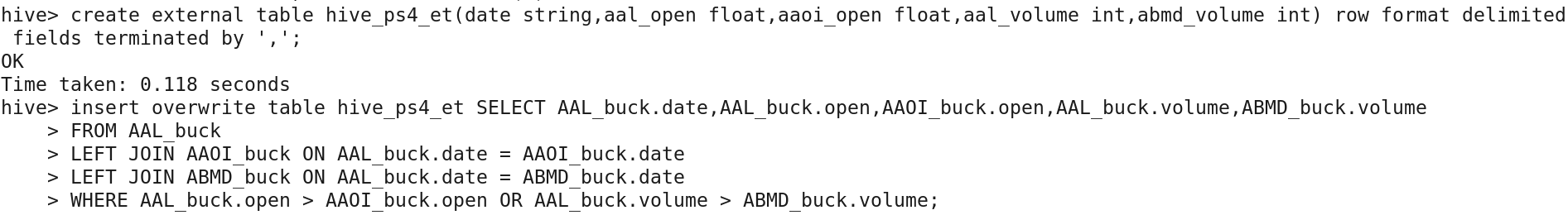
**ABOI**

****

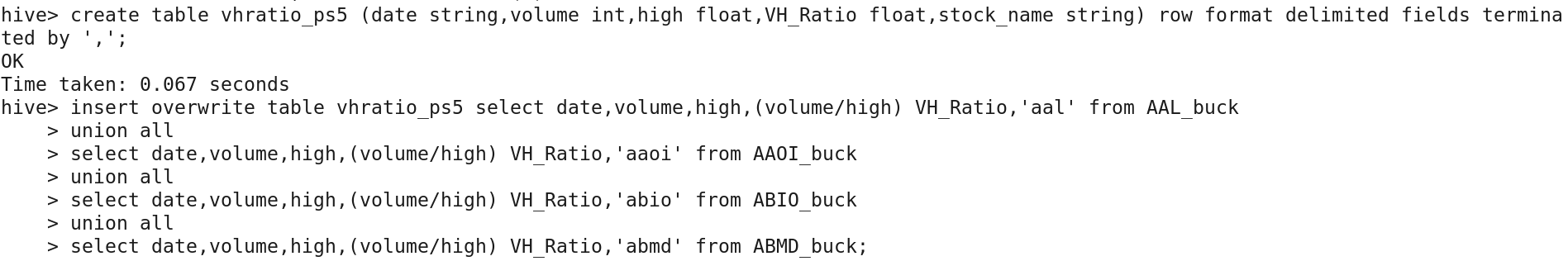
**ABMD**

****

**PS 4:**

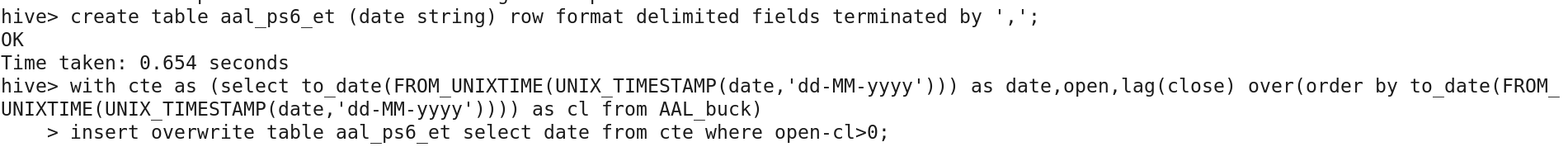
****

**PS 5:**

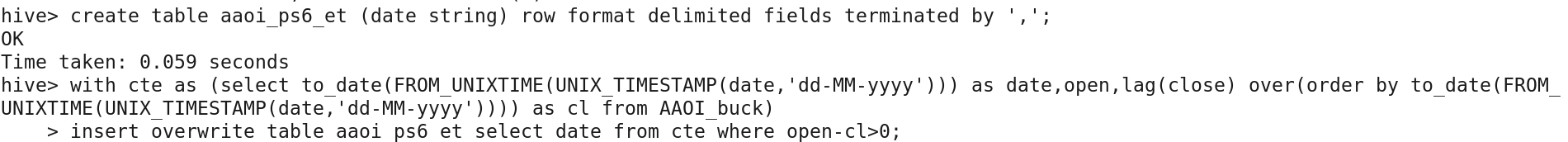
****

**PS 6:**

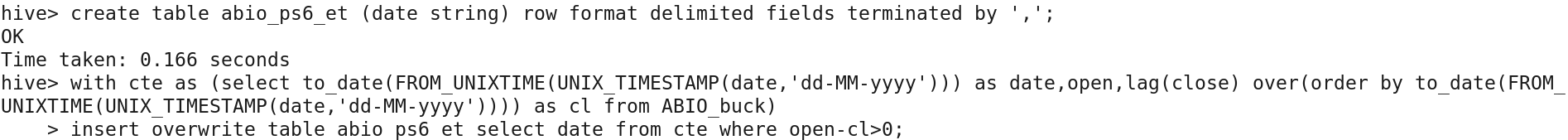
**AAL**

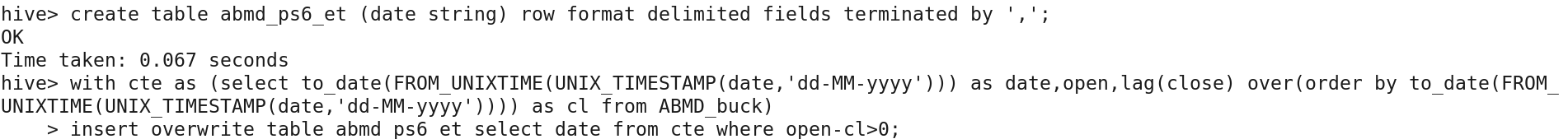
****

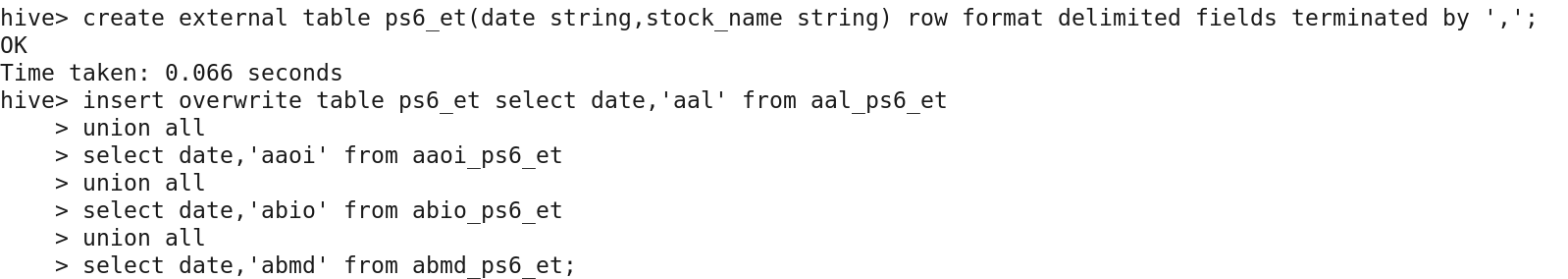
**AAOI**

****

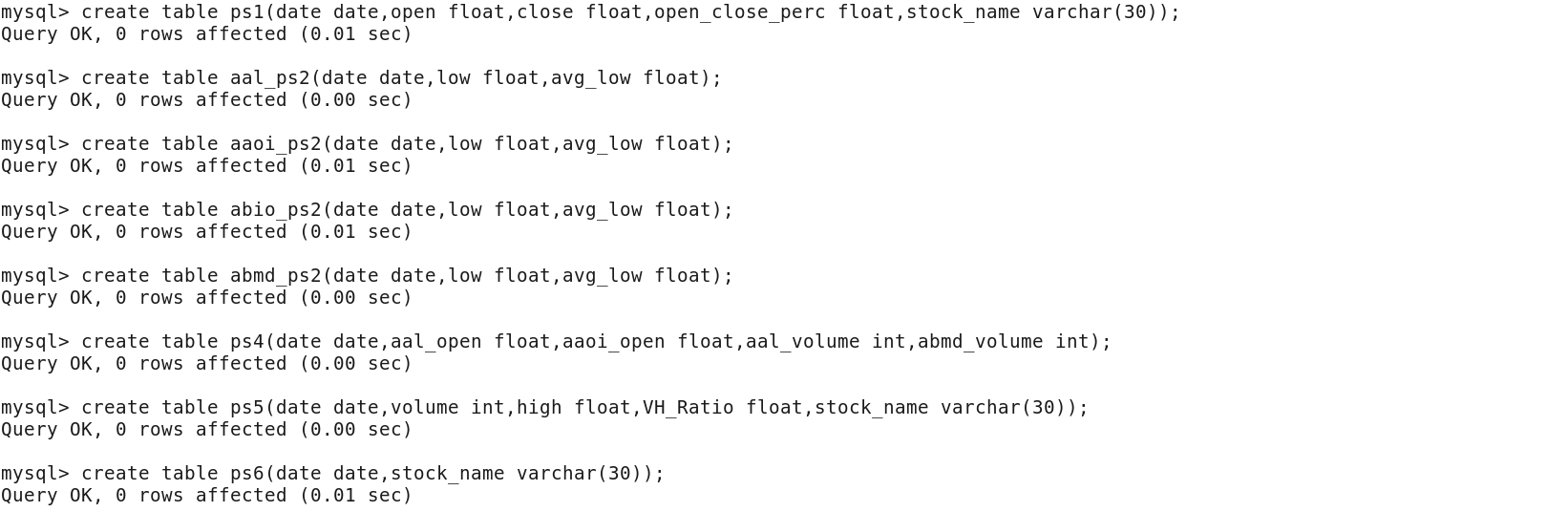
**ABIO**

****

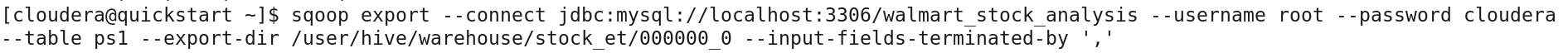
**ABMD  
**



**CREATING TABLES IN MYSQL DATABASE FOR EXTERNAL TABLES**



**SQOOP EXPORT FROM EXTERNAL TABLES TO MYSQL**

****

sqoop export --connect jdbc:mysql://localhost:3306/walmart\_stock\_analysis --username root --password cloudera --table aal\_ps2 --export-dir /user/hive/warehouse/aal\_lowavg\_et/\* -m 1

sqoop export --connect jdbc:mysql://localhost:3306/walmart\_stock\_analysis --username root --password cloudera --table aaoi\_ps2 --export-dir /user/hive/warehouse/aaoi\_lowavg\_et/\* -m 1

sqoop export --connect jdbc:mysql://localhost:3306/walmart\_stock\_analysis --username root --password cloudera --table abio\_ps2 --export-dir /user/hive/warehouse/abio\_lowavg\_et/\* -m 1

sqoop export --connect jdbc:mysql://localhost:3306/walmart\_stock\_analysis --username root --password cloudera --table abmd\_ps2 --export-dir /user/hive/warehouse/abmd\_lowavg\_et/\* -m 1

****

sqoop export --connect jdbc:mysql://localhost:3306/walmart\_stock\_analysis --username root --password cloudera --table ps4 --export-dir /user/hive/warehouse/hive\_ps4\_et/\* -m 1 --input-null-non-string '\\N' --input-null-string '\\N'

sqoop export --connect jdbc:mysql://localhost:3306/walmart\_stock\_analysis --username root --password cloudera --table ps5 --export-dir /user/hive/warehouse/vhratio\_ps5 -m 1

sqoop export --connect jdbc:mysql://localhost:3306/walmart\_stock\_analysis --username root --password cloudera --table ps6 --export-dir /user/hive/warehouse/ps6\_et/\* -m 1