Assignment #2:

1. Create "weather" external table under /user/training/weather

CREATE EXTERNAL TABLE IF NOT EXISTS weather (WBANNO int, LST\_DATE DATE, CRX\_VN FLOAT, LONGITUDE FLOAT,LATITUDE FLOAT,T\_DAILY\_MAX FLOAT,T\_DAILY\_MIN FLOAT,T\_DAILY\_MEAN FLOAT, T\_DAILY\_AVG FLOAT,P\_DAILY\_CALC FLOAT, SOLARAD\_DAILY String,SUR\_TEMP\_DAILY\_TYPE FLOAT, SUR\_TEMP\_DAILY\_MAX FLOAT,

SUR\_TEMP\_DAILY\_MIN FLOAT,SUR\_TEMP\_DAILY\_AVG FLOAT,RH\_DAILY\_MAX FLOAT, RH\_DAILY\_MIN FLOAT,RH\_DAILY\_AVG FLOAT,SOIL\_MOISTURE\_5\_DAILY FLOAT,

SOIL\_MOISTURE\_10\_DAILY FLOAT, SOIL\_MOISTURE\_20\_DAILY FLOAT, SOIL\_MOISTURE\_50\_DAILY FLOAT, SOIL\_MOISTURE\_100\_DAILY FLOAT,

SOIL\_TEMP\_5\_DAILY FLOAT, SOIL\_TEMP\_10\_DAILY FLOAT, SOIL\_TEMP\_20\_DAILY FLOAT, SOIL\_TEMP\_50\_DAILY FLOAT, SOIL\_TEMP\_100\_DAILY FLOAT)

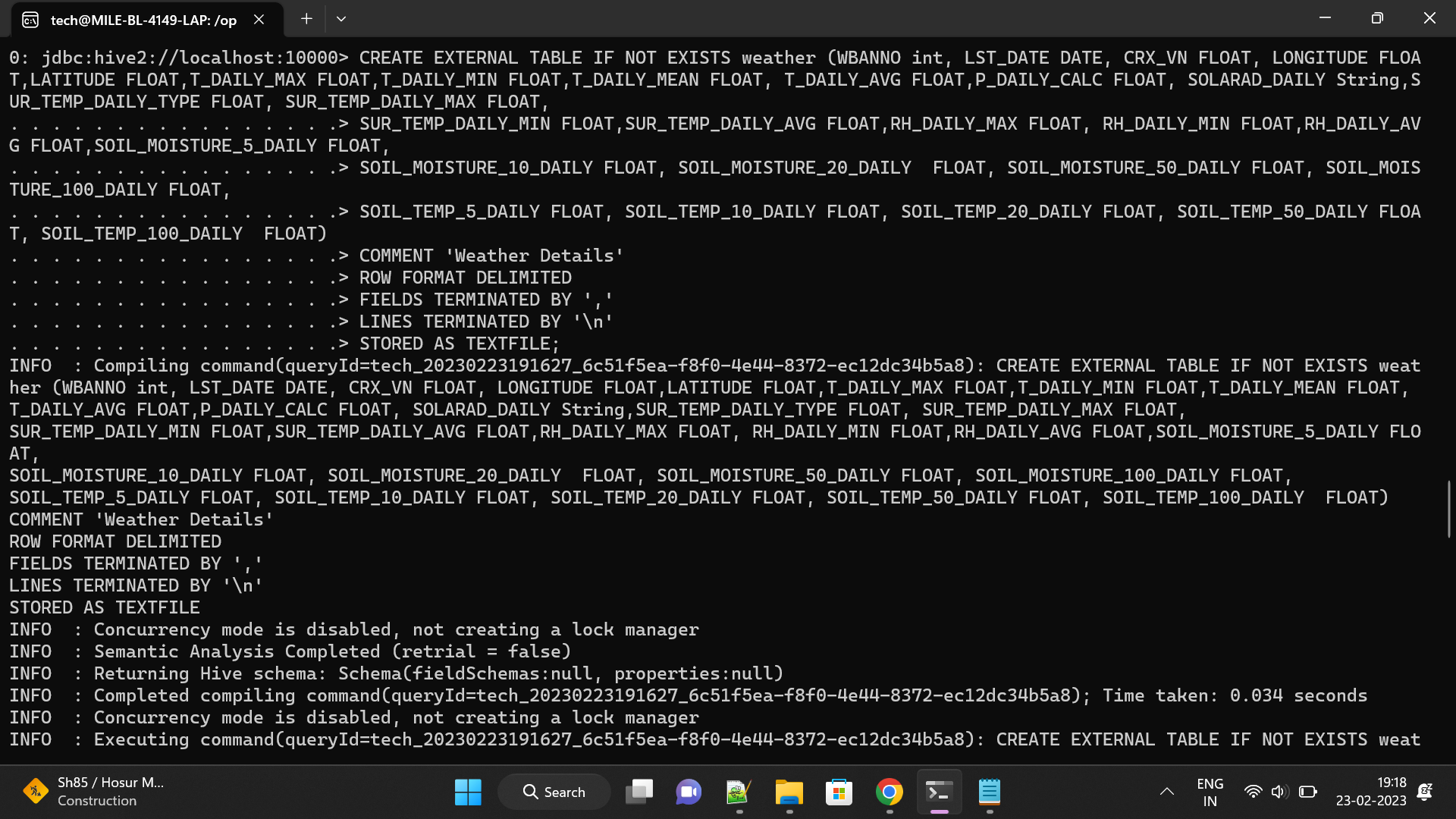
COMMENT 'Weather Details'

ROW FORMAT DELIMITED

FIELDS TERMINATED BY ','

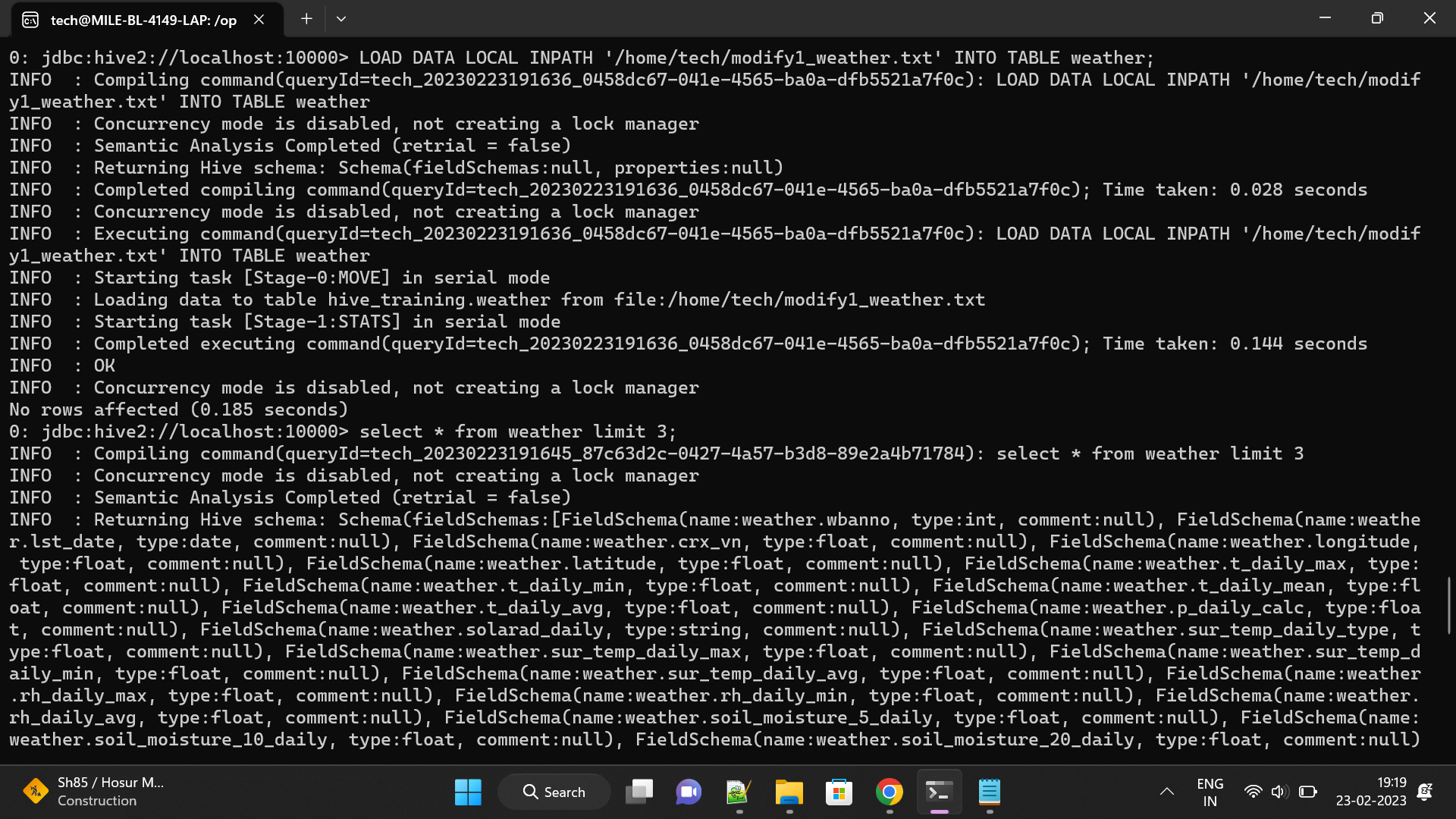
LINES TERMINATED BY '\n'

STORED AS TEXTFILE;



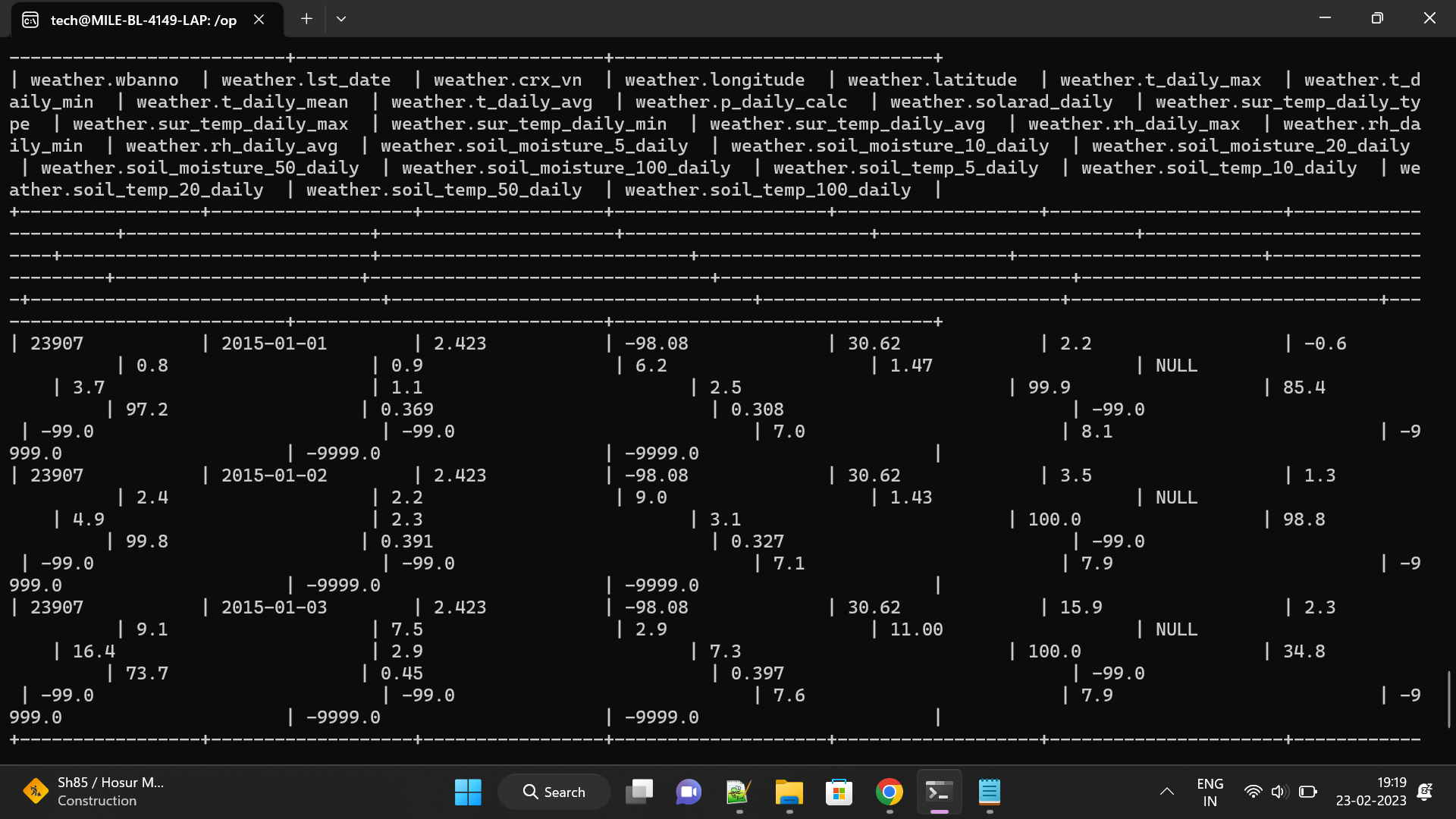
2. Load the data from ~/futurence\_hadoop-pyspark/labs/dataset/weather to /user/training/weather

LOAD DATA LOCAL INPATH '/home/tech/modify1\_weather.txt' INTO TABLE weather;



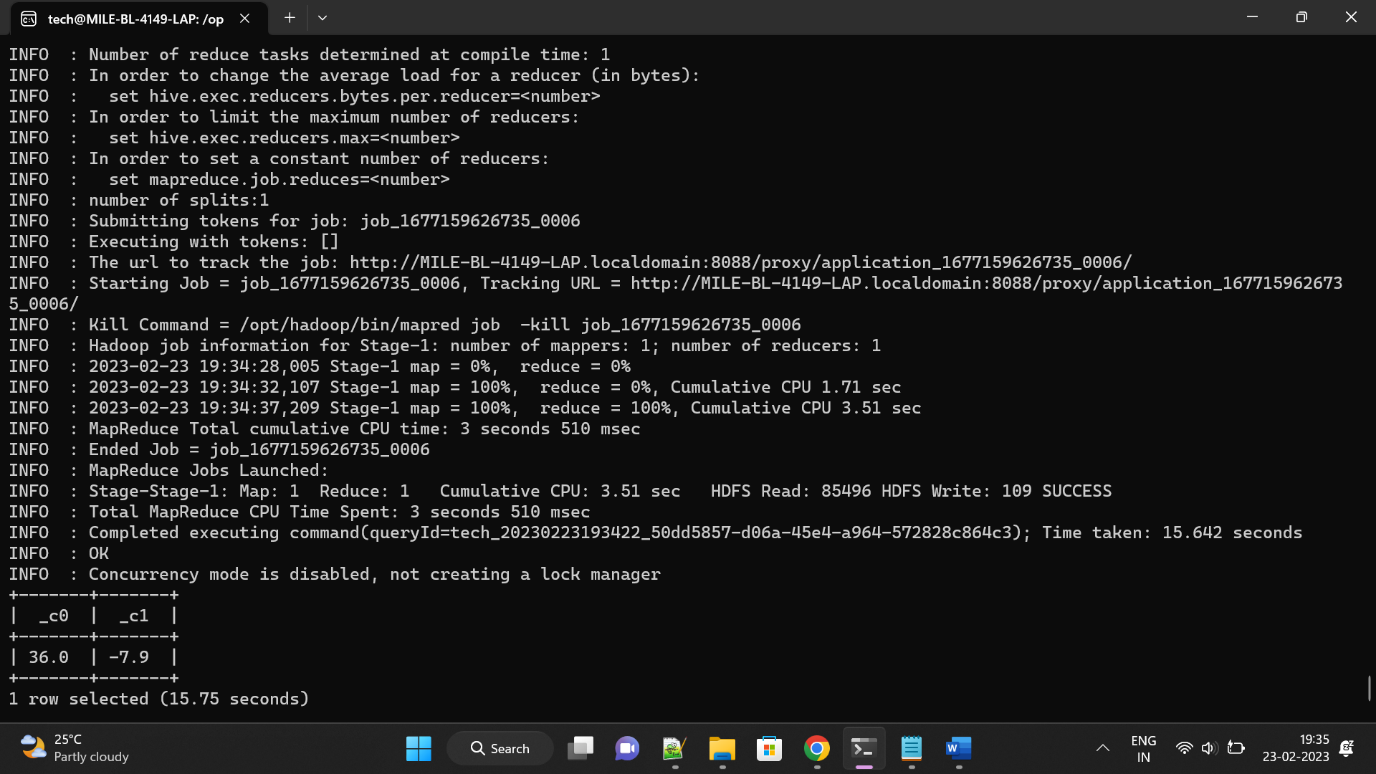
3. Display the weather

select \* from weather;



4. Display Max, Min weather

select max(t\_daily\_max),min(t\_daily\_min) from weather;



5. Display month wise Max and Min weather

select month(lst\_date),max(t\_daily\_max) from weather group by month(lst\_date);

