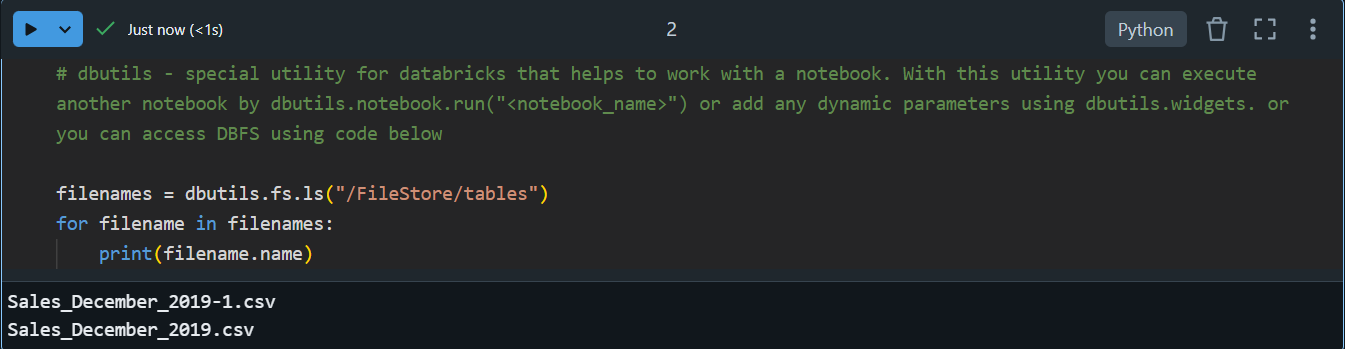
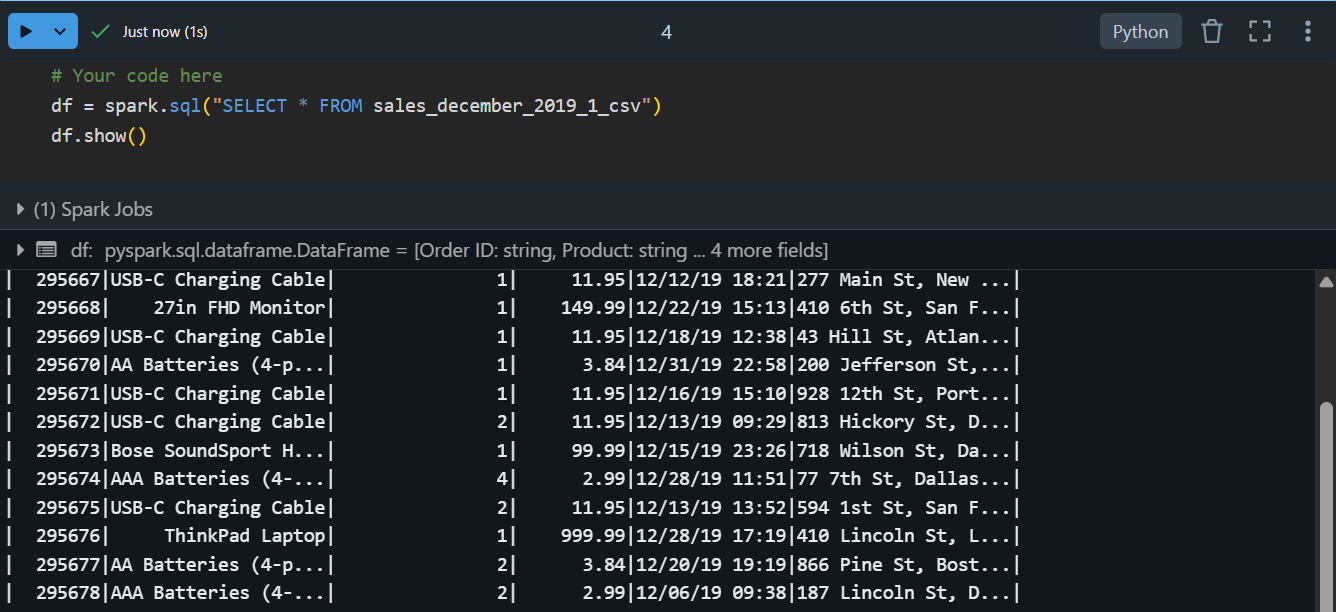
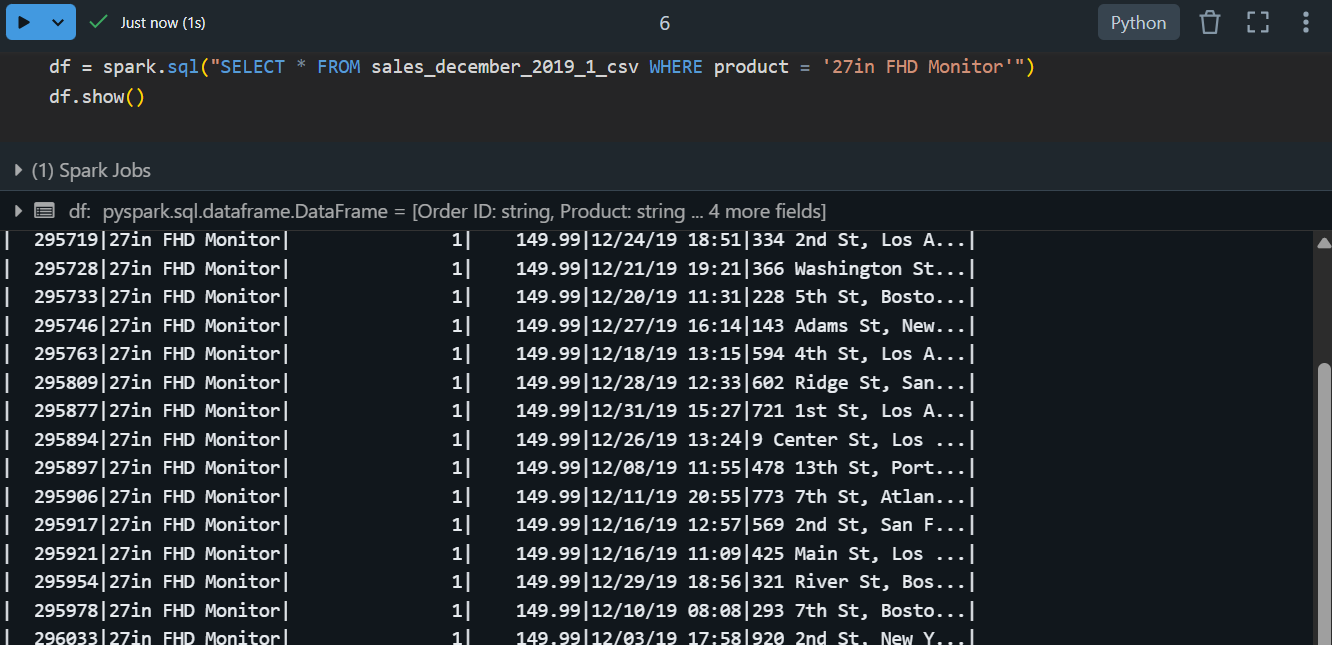
These are the files that the code returned when I ran it:



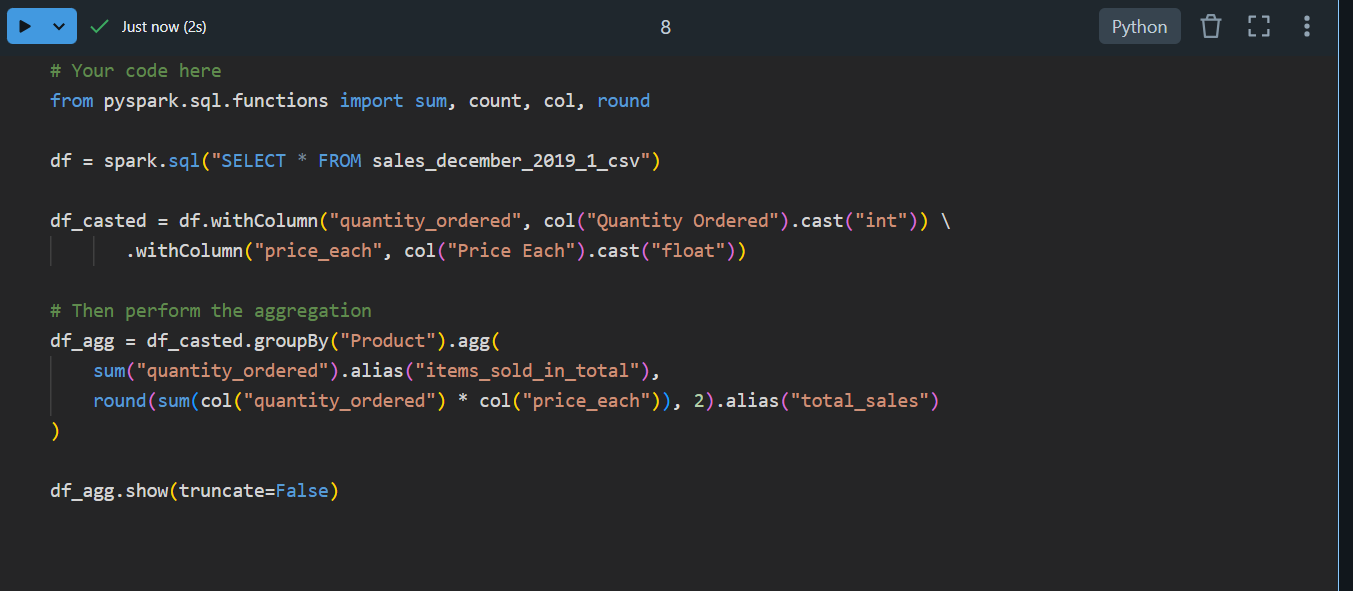
Task 1-Write data from table that you created into the dataframe using PySpark with SQL query. Show data in the dataframe



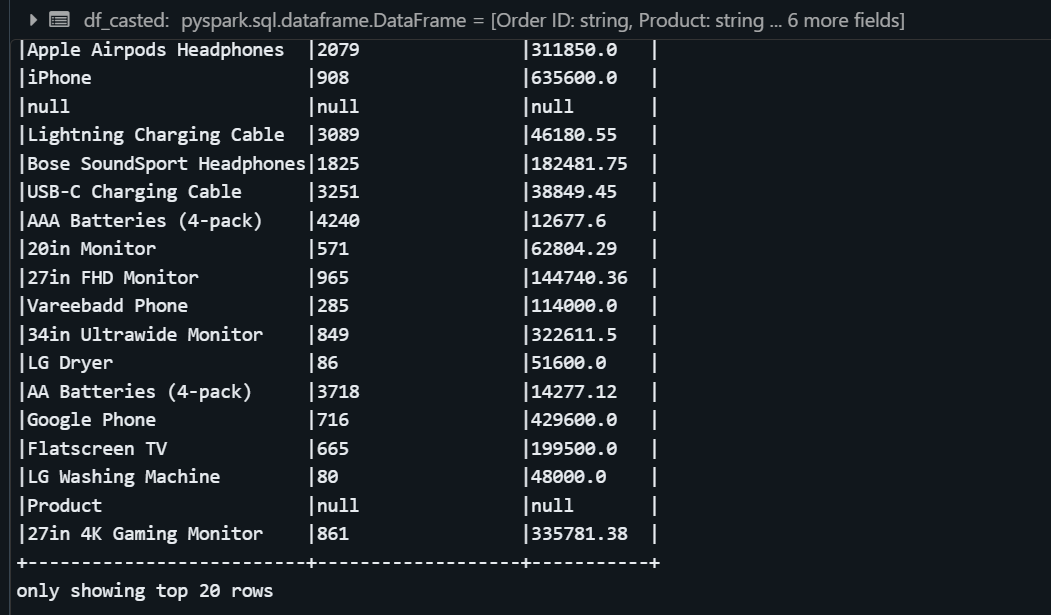
Task 2-Select data where product = "product\_name" from dataframe using PySpark



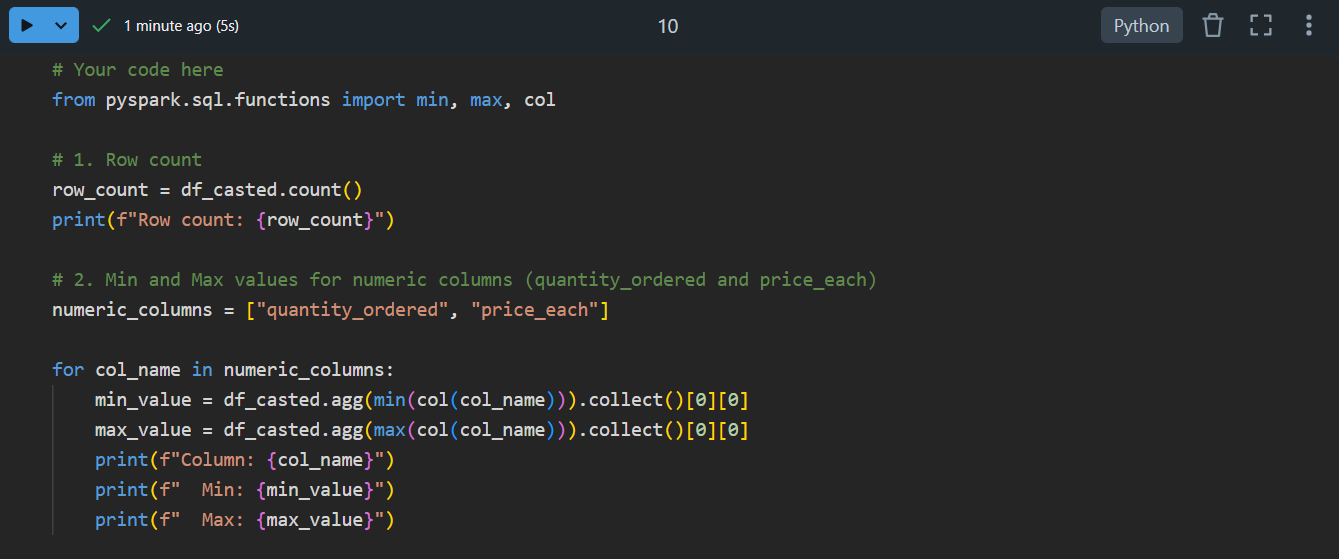
Task 3-Calculate the sales for each product, including the number of products sold

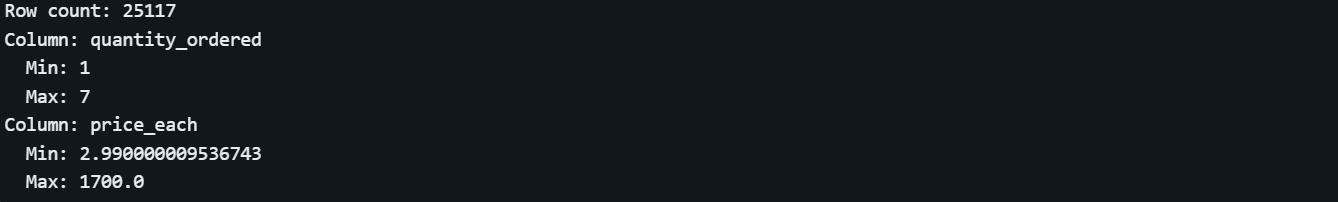


And this is the output of this code:



Task 4-Show data profiles output for the new dataframe of table sales\_december\_2019\_csv: row count, min and max value for each column

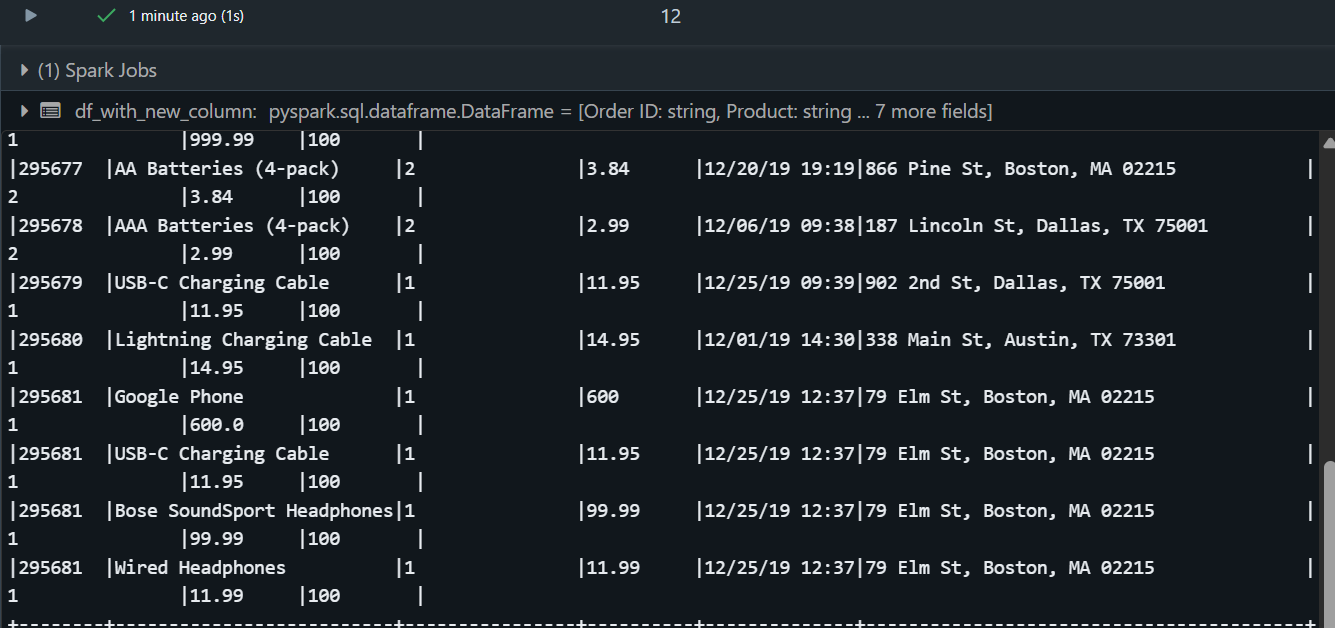


Output is this: 

Task 5-Add new column to the dataframe from previous task with any default value that you want

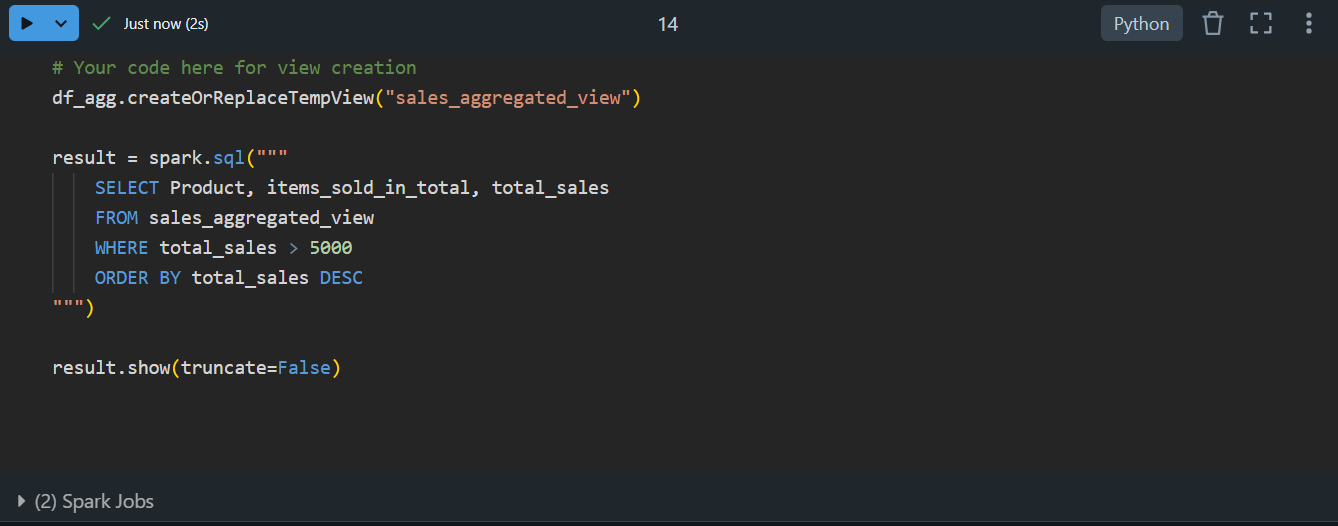


And this is the output:



Task 6-Create temporary view from task 4 dataframe using PySpark and perform any select using SQL

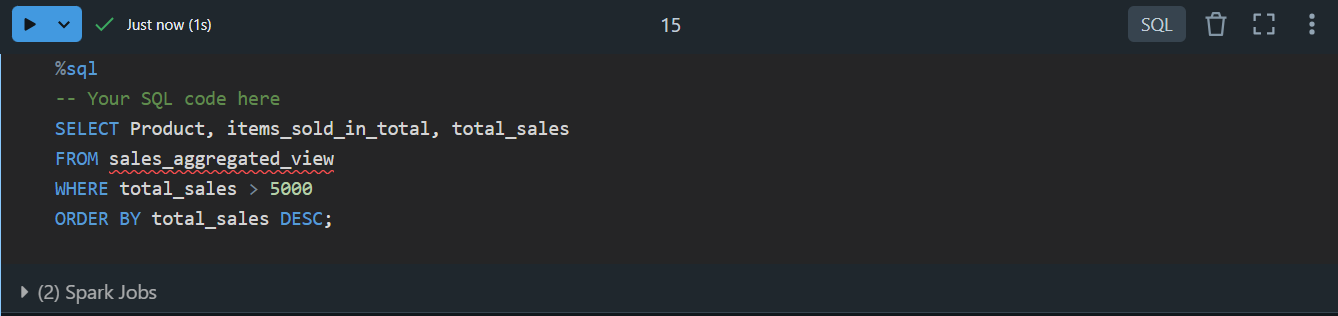
Here is the code:



And this is the output:



And I rechecked it with sql code also:



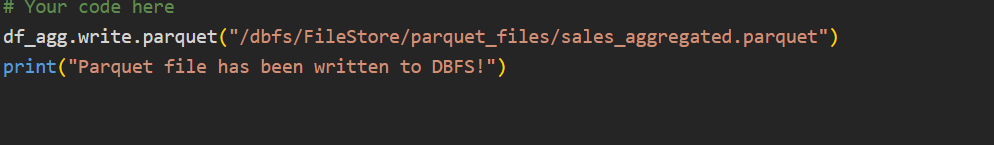
Result:



But this is the only small part

Task 7-Export dataframe from task 5 as .parquet file to the DBFS and show that databricks creates folder with .parquet files

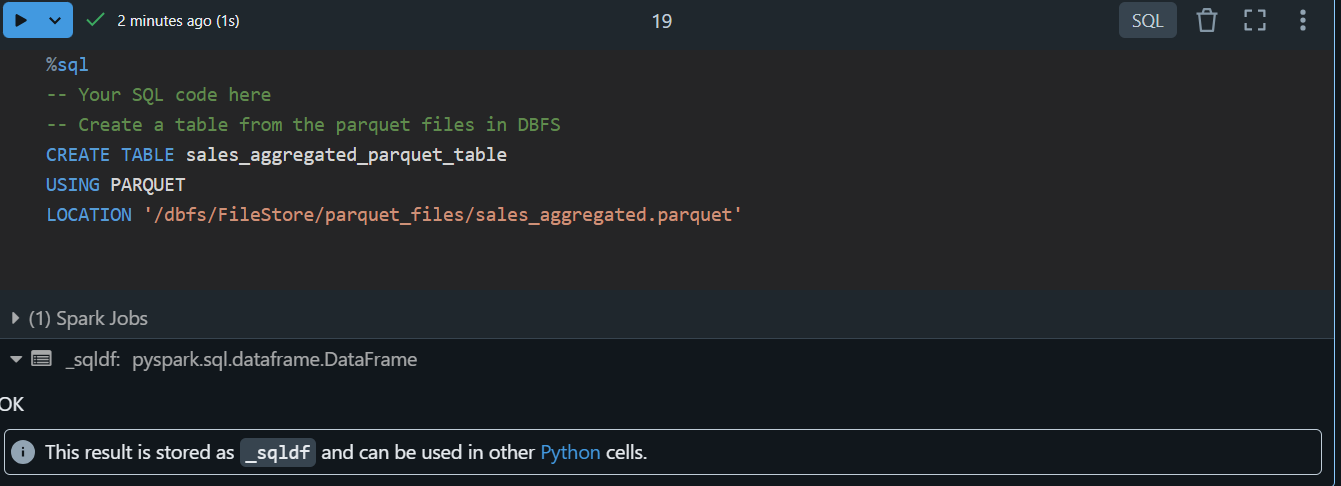
Code part:



Output:

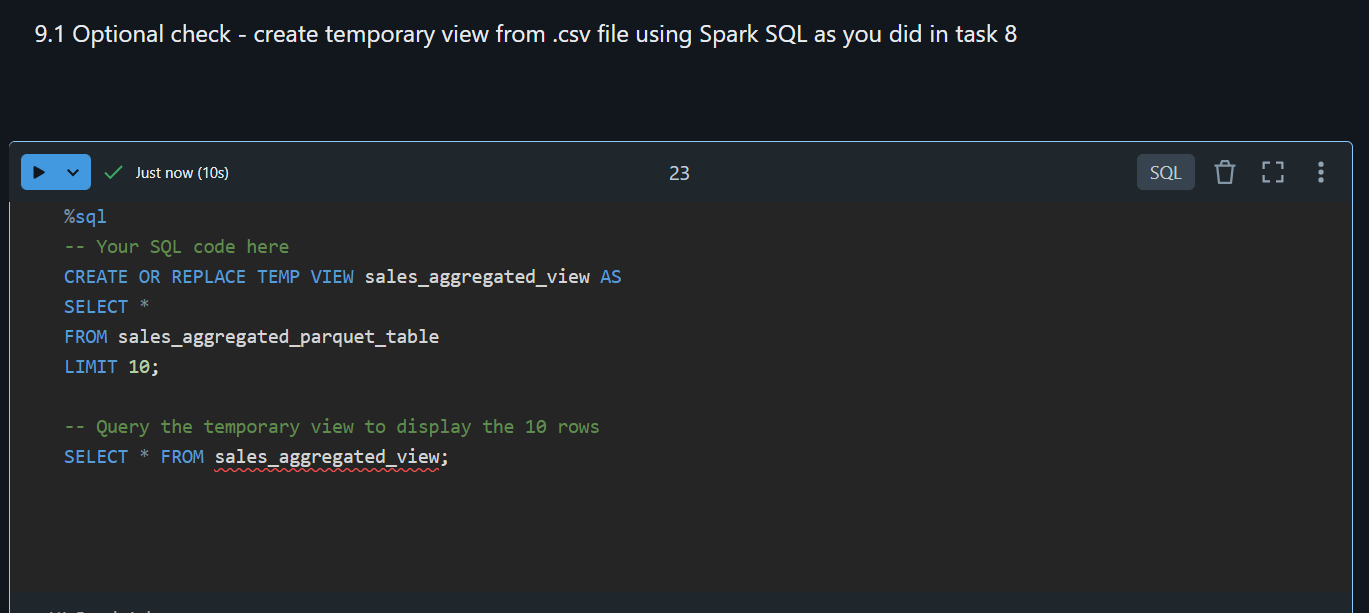


Task 8-Create table from .parquet files that was created in the previous cell using pure SQL

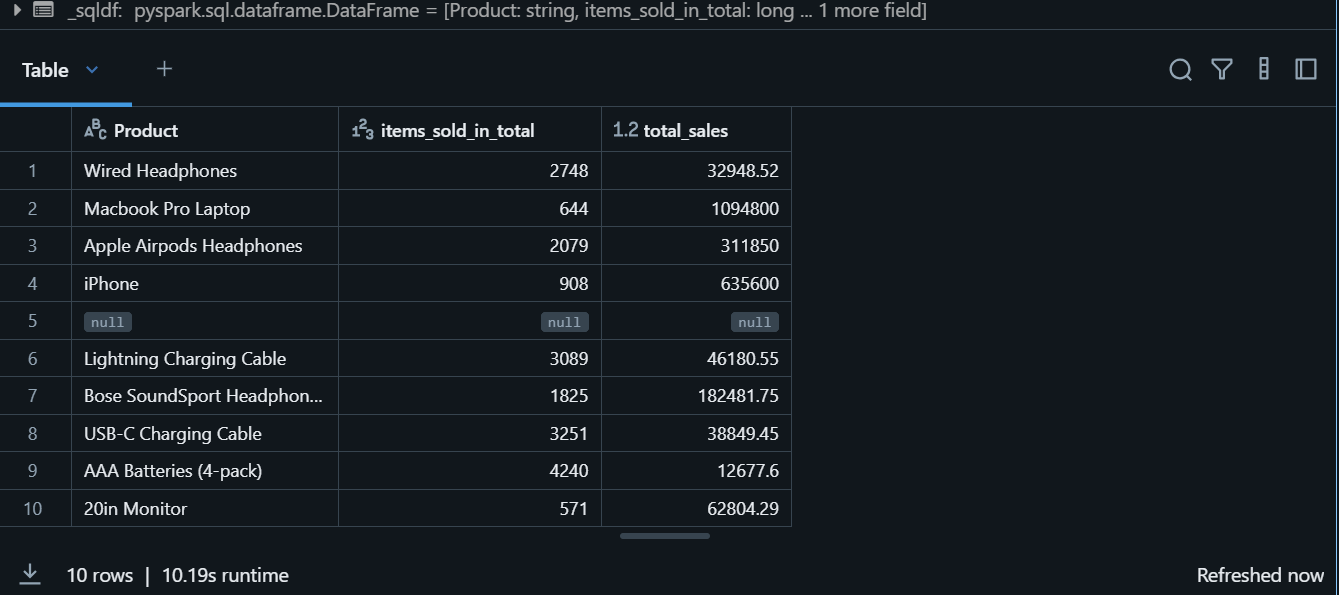


Task 9-

Task 9.1:



Output:



I don’t know if you are able to visit this link but for any case I will upload the link to this directory:

[https://community.cloud.databricks.com/editor/notebooks/1141659624815398?o=4469874836556395#command/1141659624815421](https://community.cloud.databricks.com/editor/notebooks/1141659624815398?o=4469874836556395%23command/1141659624815421)