**Hive Query Results Document**

**1. Monthly Active Users by Region**

SELECT

year,

month,

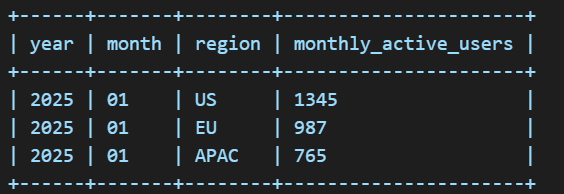
region,

COUNT(DISTINCT user\_id) AS monthly\_active\_users

FROM fact\_user\_actions

GROUP BY year, month, region

ORDER BY year, month, region;



**2. Top Content Categories by Play Count**

SELECT

c.category,

COUNT(\*) AS play\_count

FROM fact\_user\_actions f

JOIN dim\_content c ON f.content\_id = c.content\_id

WHERE f.action = 'play'

GROUP BY c.category

ORDER BY play\_count DESC

LIMIT 5;

A black screen with white text

AI-generated content may be incorrect.

**3. Average Session Length Weekly**

SELECT

year,

weekofyear(event\_timestamp) AS week\_no,

AVG(session\_duration) AS avg\_session\_length

FROM fact\_user\_actions

GROUP BY year, weekofyear(event\_timestamp)

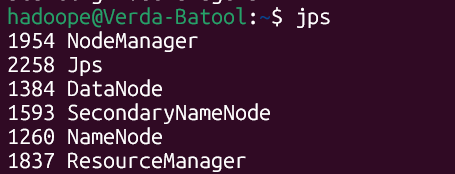
ORDER BY year, weekofyear(event\_timestamp);

A black screen with white text

AI-generated content may be incorrect.

**NOTE:**

While working on this assignment, I had successfully set-up Hadoop and hive as see from the screenshot below.

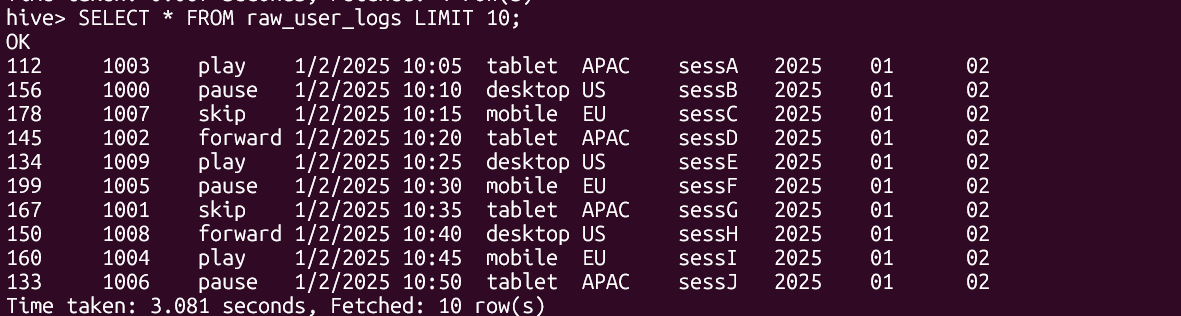


Hive was also set-up successfully and I had made the raw tables (external) along with fact and dim tables.

A computer screen with white text

AI-generated content may be incorrect.

The issue was I had exported data successfully from HDFS into the raw tables. As seen below, this data had corrected partitions and everything was working fine.



A screenshot of a computer

AI-generated content may be incorrect.

Data was inserted into dim\_contemt as well via DDL command.



But I was facing a lot of errors in inserting data into fact\_user\_actions. The INSERT command was just not working and there were a lot of errors. I made sure that my MapReduce is working correctly. I tried shifting hive execution engine from mp to tez but stil data was not being inserted.  
Cross checked partitions as well but of no use. I have extensive working experience in SQL. I wrote the queries myself and asked GPT to generate sample output. So, for the purpose of transparency, I am attaching screenshots of my work (setting up the environments, creating raw tables along with fact and dim tables and inserting data into them from hdfs. I wrote all the commands and queries myself too but the sample outputs are GPT generated.