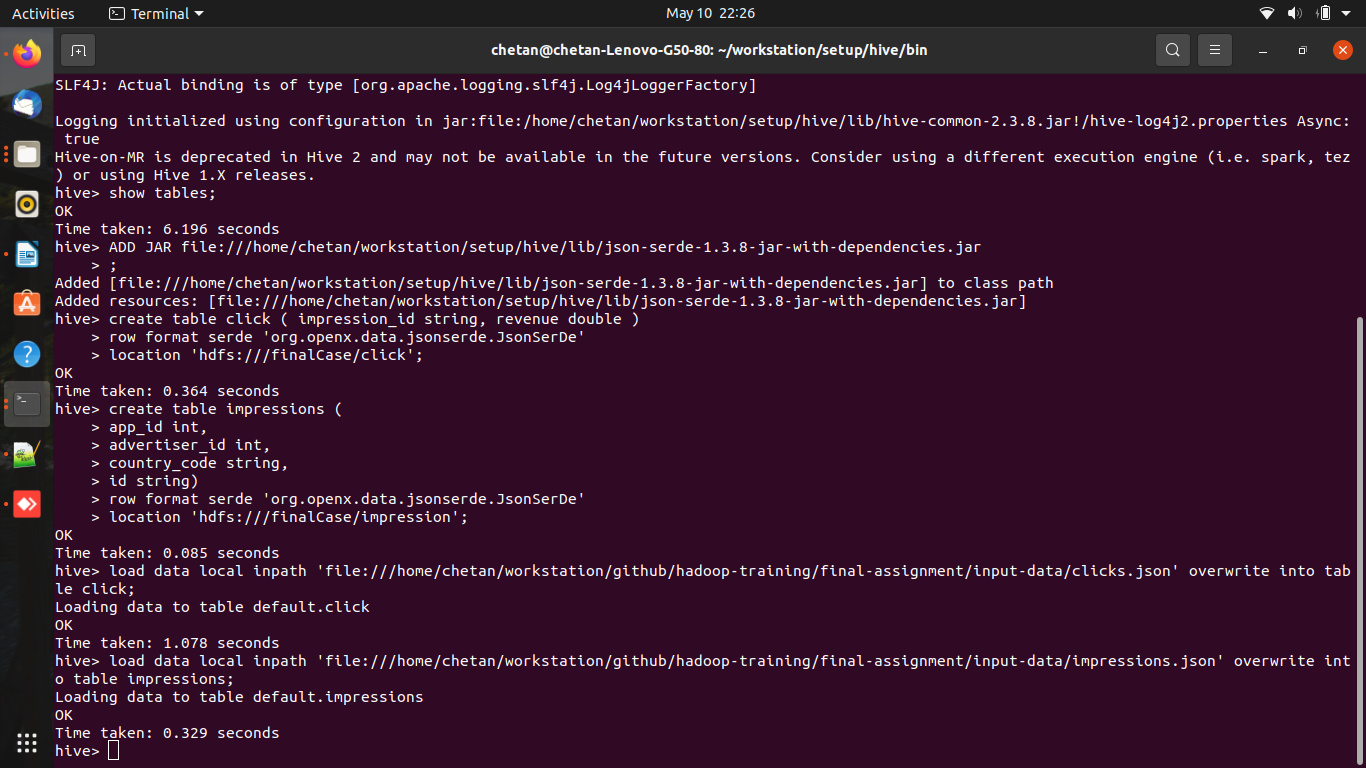
**1. Read events stored in JSON files**

**commands run -**

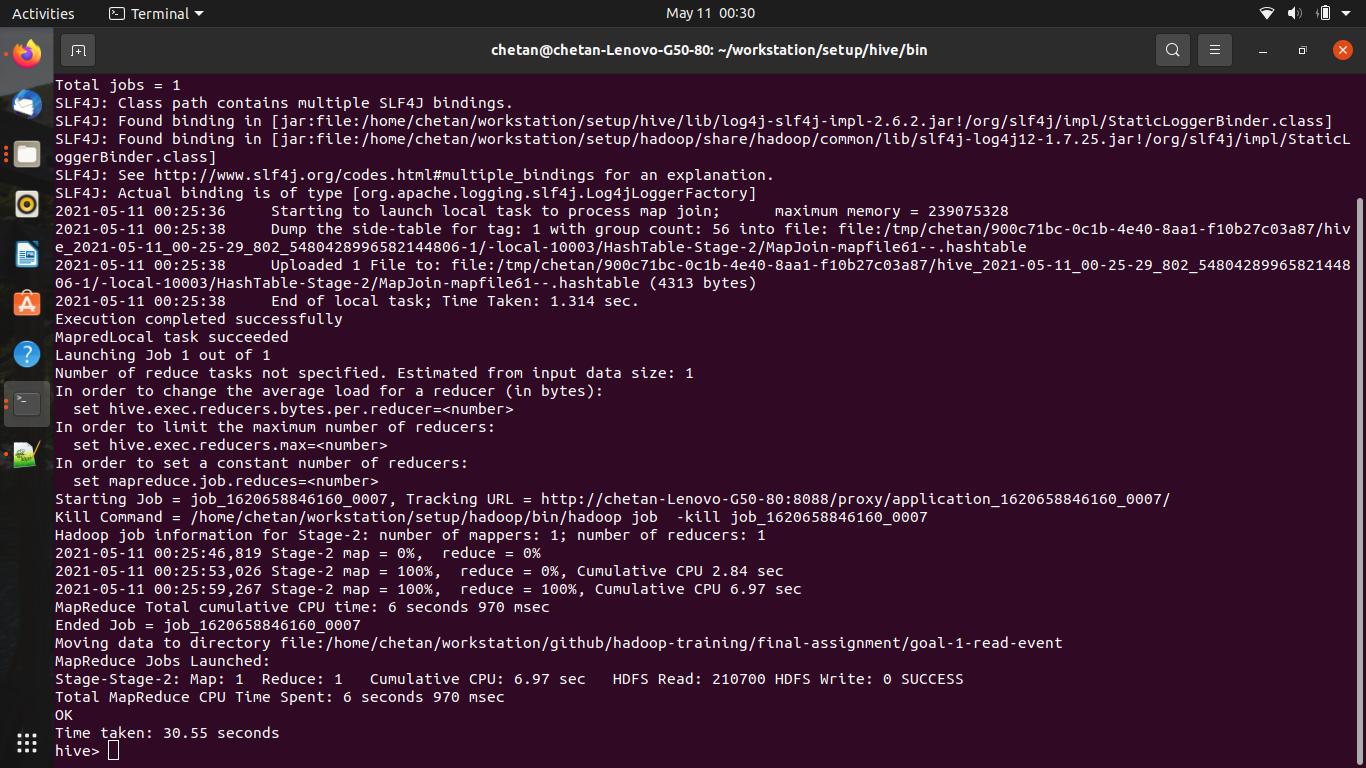
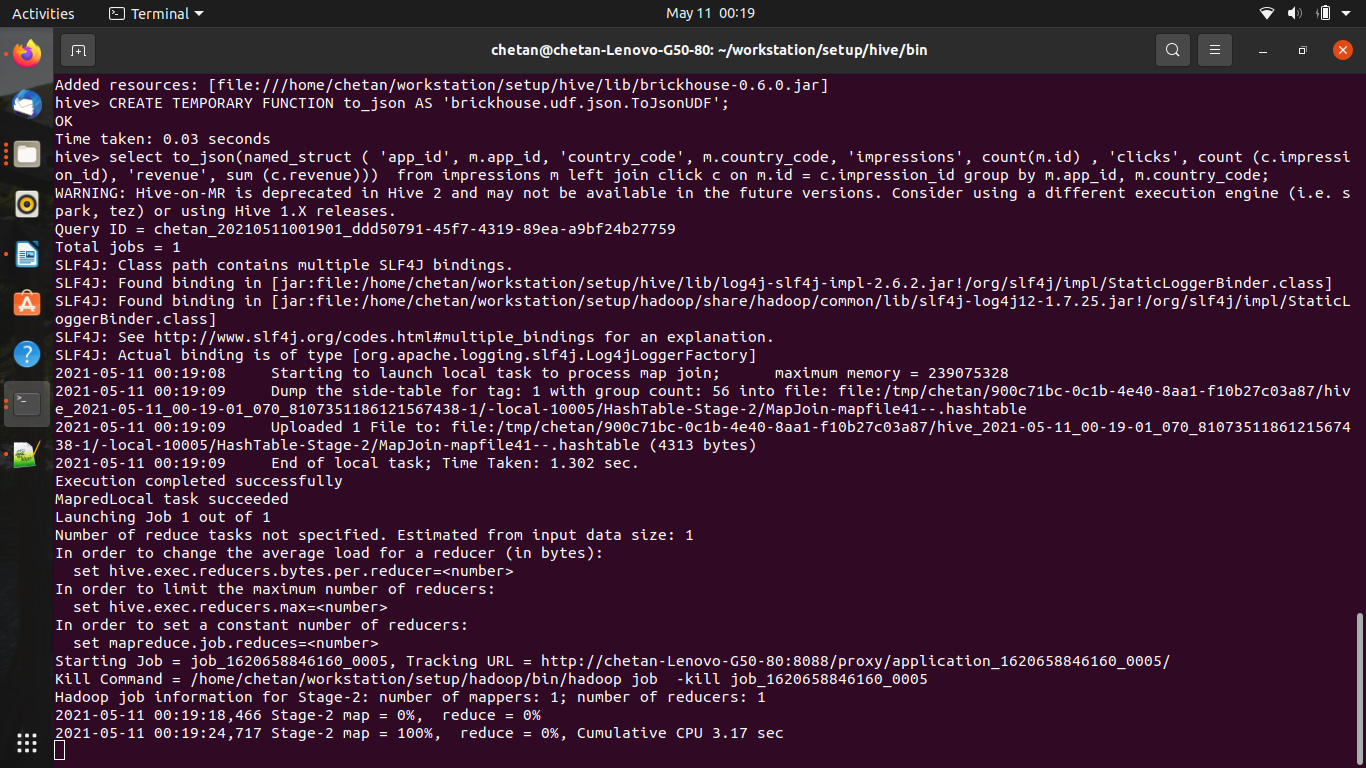
* create table click ( impression\_id string, revenue double ) row format serde 'org.openx.data.jsonserde.JsonSerDe' location 'hdfs:///finalCase/click';
* load data local inpath 'file:///home/chetan/workstation/github/hadoop-training/final-assignment/input-data/min-clicks.json' overwrite into table click;
* create table impressions ( app\_id int, advertiser\_id int, country\_code string, id string) row format serde 'org.openx.data.jsonserde.JsonSerDe' location 'hdfs:///finalCase/impression';
* load data local inpath 'file:///home/chetan/workstation/github/hadoop-training/final-assignment/input-data/min-impressions.json' overwrite into table impressions;



### 2. Calculate metrics for some dimensions

**commands run -**

* ADD JAR [file:///home/chetan/workstation/setup/hive/lib/brickhouse-0.6.0.jar](../workstation/setup/hive/lib/brickhouse-0.6.0.jar);
* CREATE TEMPORARY FUNCTION to\_json as 'brickhouse.udf.json.ToJsonUDF';
* INSERT OVERWRITE DIRECTORY 'file:///home/chetan/workstation/github/hadoop-training/final-assignment/goal-1-read-event' stored as textfile select to\_json(named\_struct ( 'app\_id', m.app\_id, 'country\_code', m.country\_code, 'impressions', count(m.id) , 'clicks', count (c.impression\_id), 'revenue', sum (c.revenue))) from impressions m left join click c on m.id = c.impression\_id group by m.app\_id, m.country\_code;



### **3. Make a recommendation for the top 5 advertiser\_ids to display for each app and country combination.**

**commands run -**

* INSERT OVERWRITE DIRECTORY 'file:///home/chetan/workstation/github/hadoop-training/final-assignment/goal-3-top-recommendation' ROW FORMAT DELIMITED FIELDS TERMINATED BY '\t' stored as textfile select m.app\_id, m.country\_code, m.advertiser\_id, sum (c.revenue) as t\_revenue from impressions m left join click c on m.id = c.impression\_id group by m.app\_id, m.country\_code, m.advertiser\_id order by m.app\_id, m.country\_code , t\_revenue desc;

