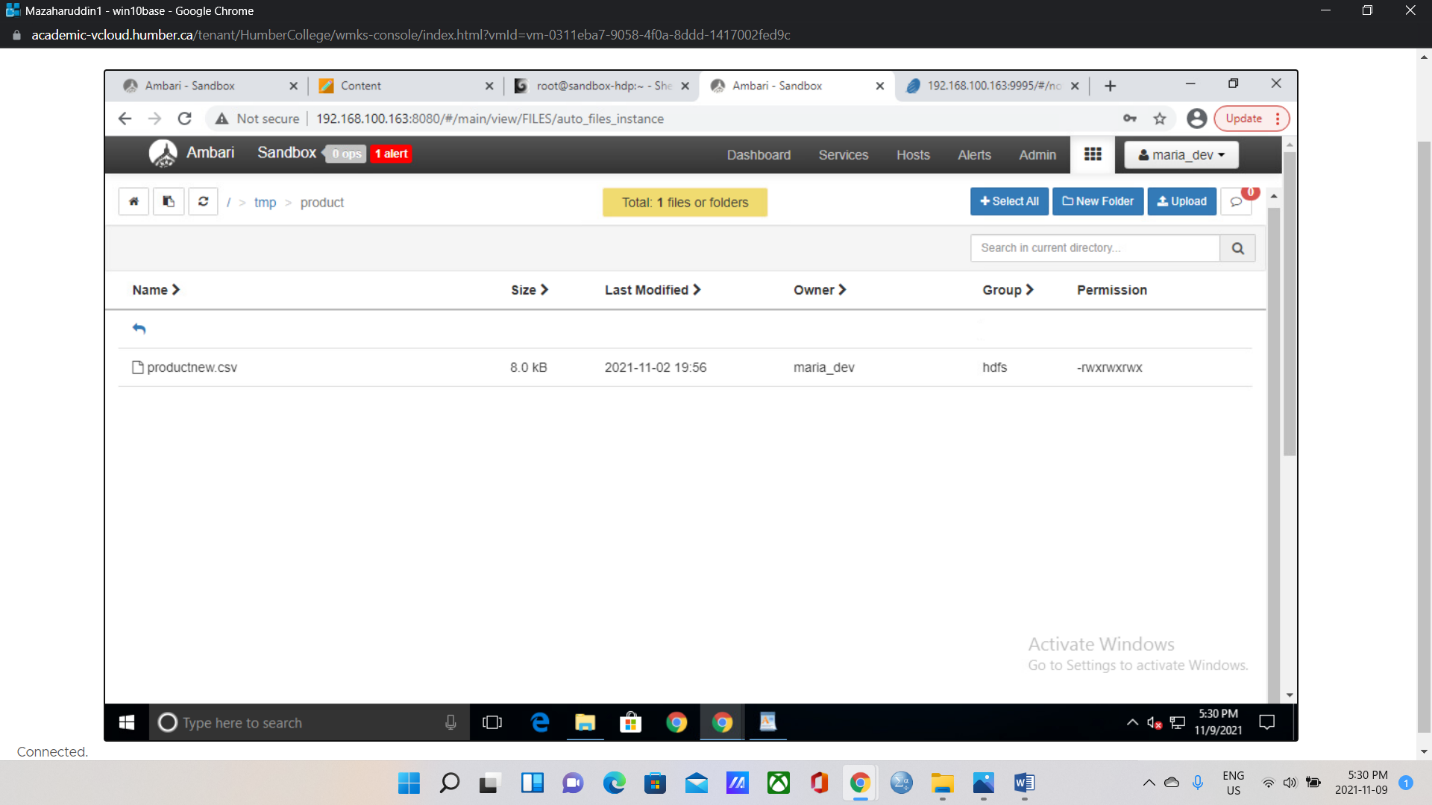
**HBase in class lab**

|  |  |
| --- | --- |
| **Student Name** | Mazaharuddin Mohammad |
| **Student Number** | N01464978 |
| **Subject Name** | [Big Data 1 - ITE-5200-IRA](https://learn.humber.ca/webapps/blackboard/execute/courseMain?course_id=_187043_1) |
| **Professor Name** | Orren Johnson |

1. **Load sales data**
   1. First load product. csv file into hadoop /tmp/product/ directory
   2. The product.csv file can be found in Course resources as Hbase product



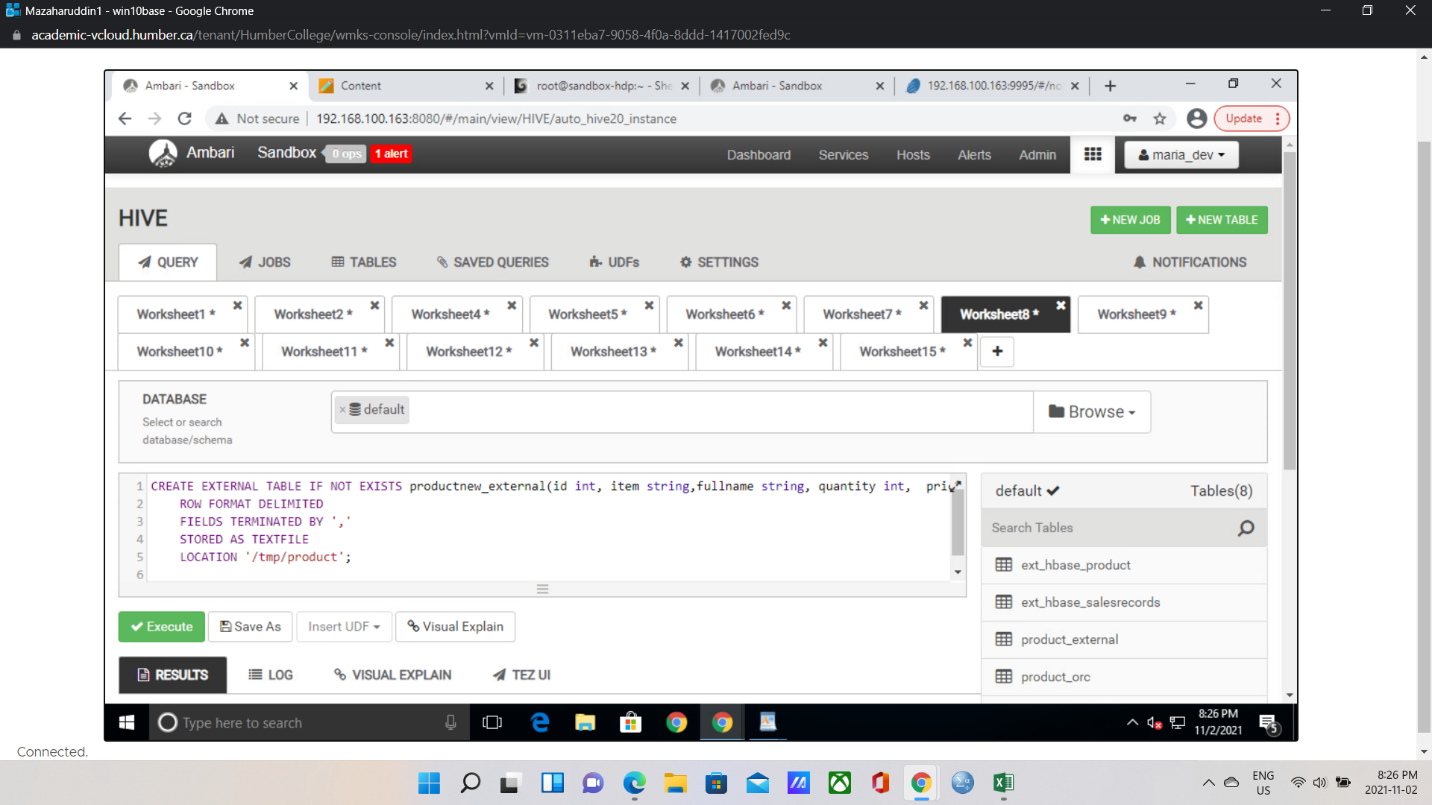
1. **then create a hive external table and load the csv file** 
   1. CREATE EXTERNAL TABLE IF NOT EXISTS product\_external(id int, item string,fullname string, quantity int, price int,item\_type string)

ROW FORMAT DELIMITED

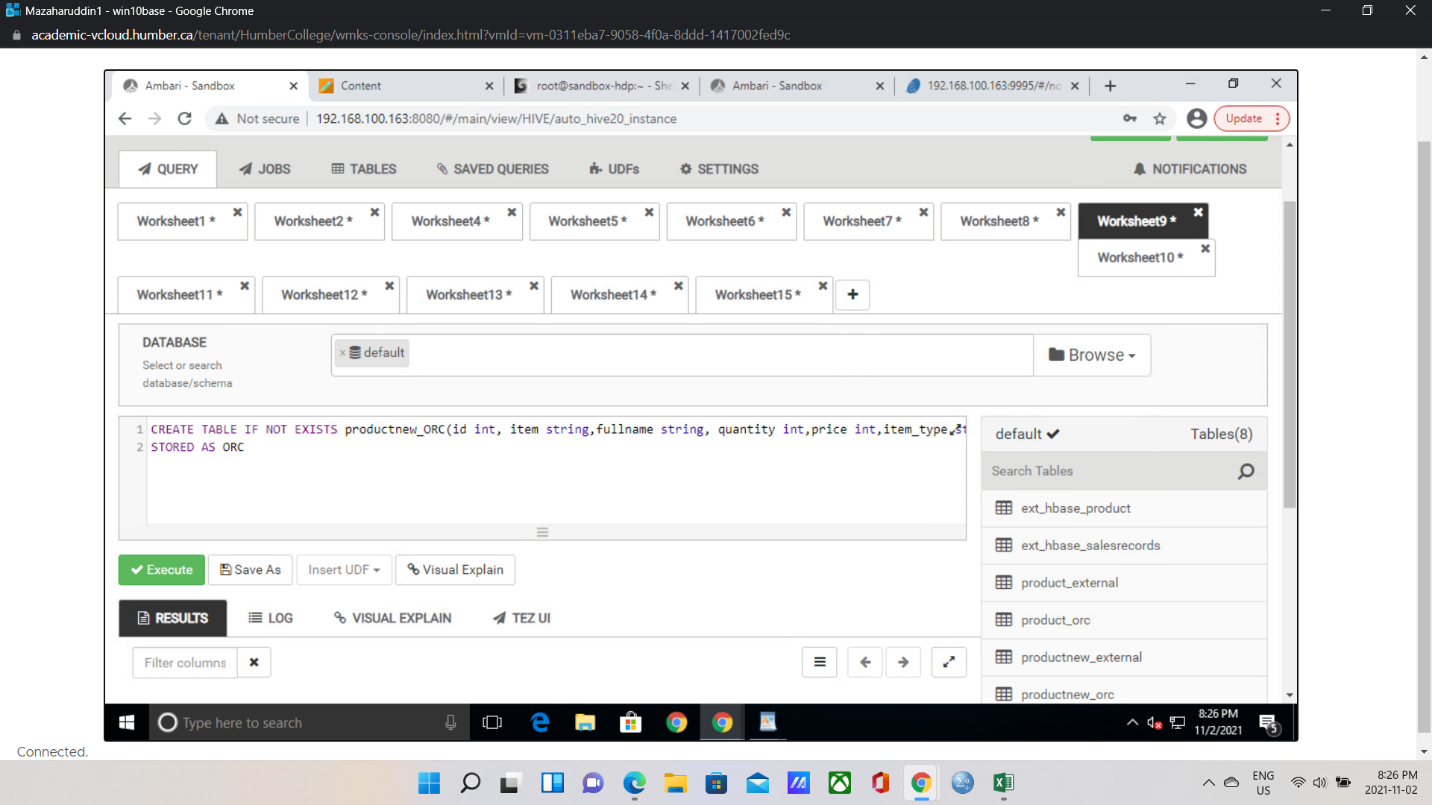
FIELDS TERMINATED BY ','

STORED AS TEXTFILE

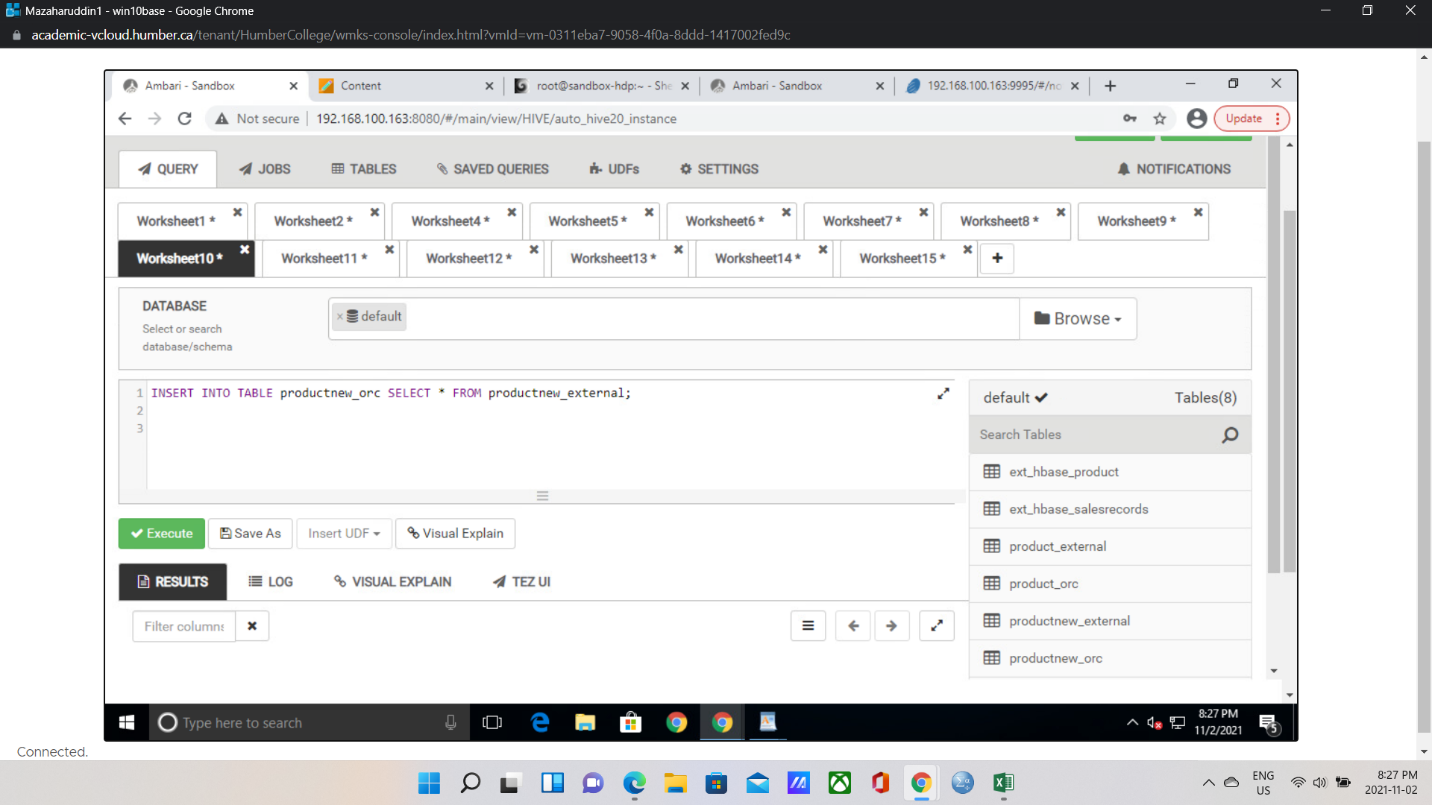
LOCATION '/tmp/product';



1. **Then create a hive internal table** 
   1. CREATE TABLE IF NOT EXISTS EXISTS product\_ORC(id int, item string,fullname string, quantity int, price int,item\_type string) STORED AS ORC;
   2. If you get an error – remove the redundant word 😊

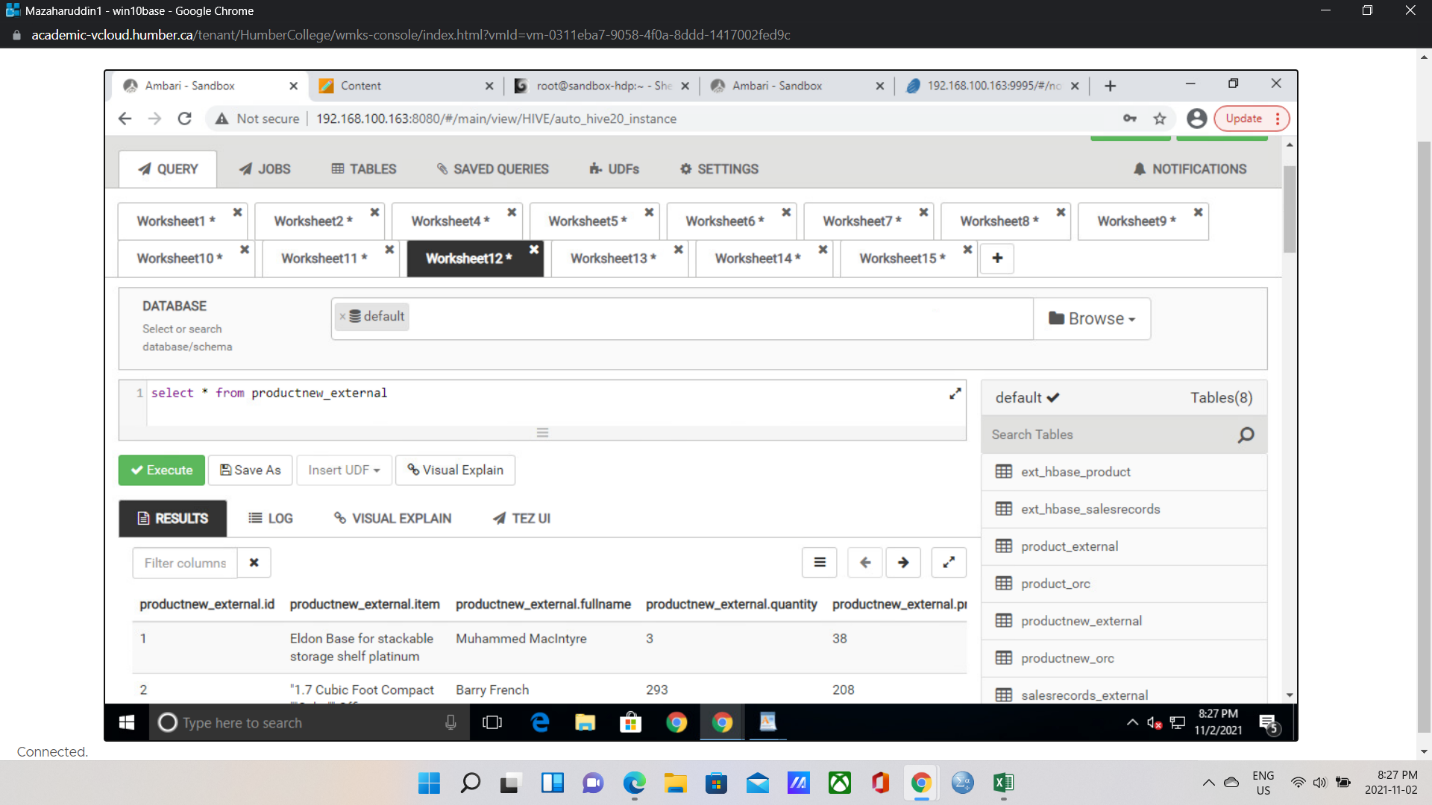


1. **Then load data from external to internal ORC** 
   1. INSERT INTO TABLE product\_orc SELECT \* FROM product\_external; - screen print results



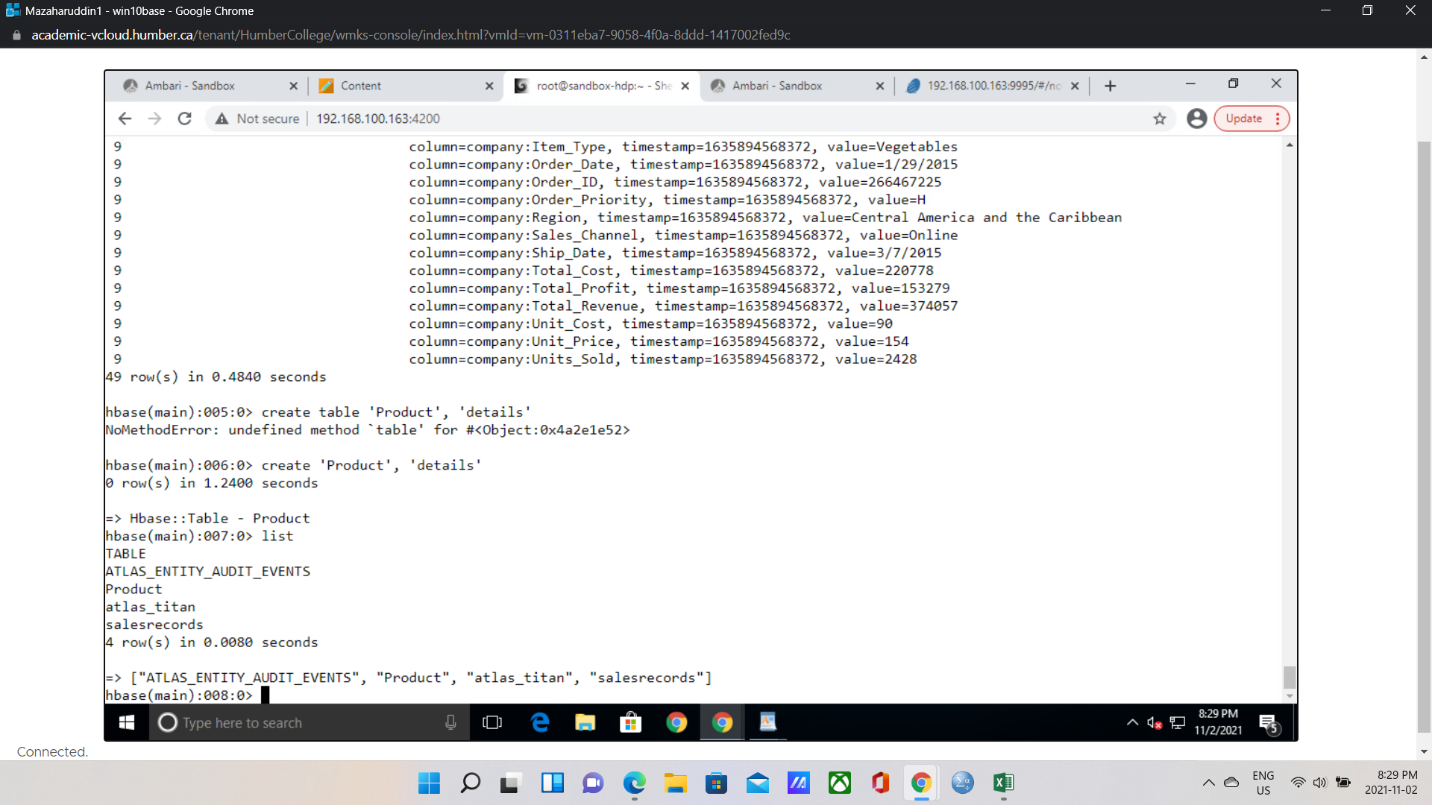


1. **select from both tables to see the data - screen print results**

****

****

1. **login to Hbase and create a hbase table**
   1. create table 'Product', 'details'



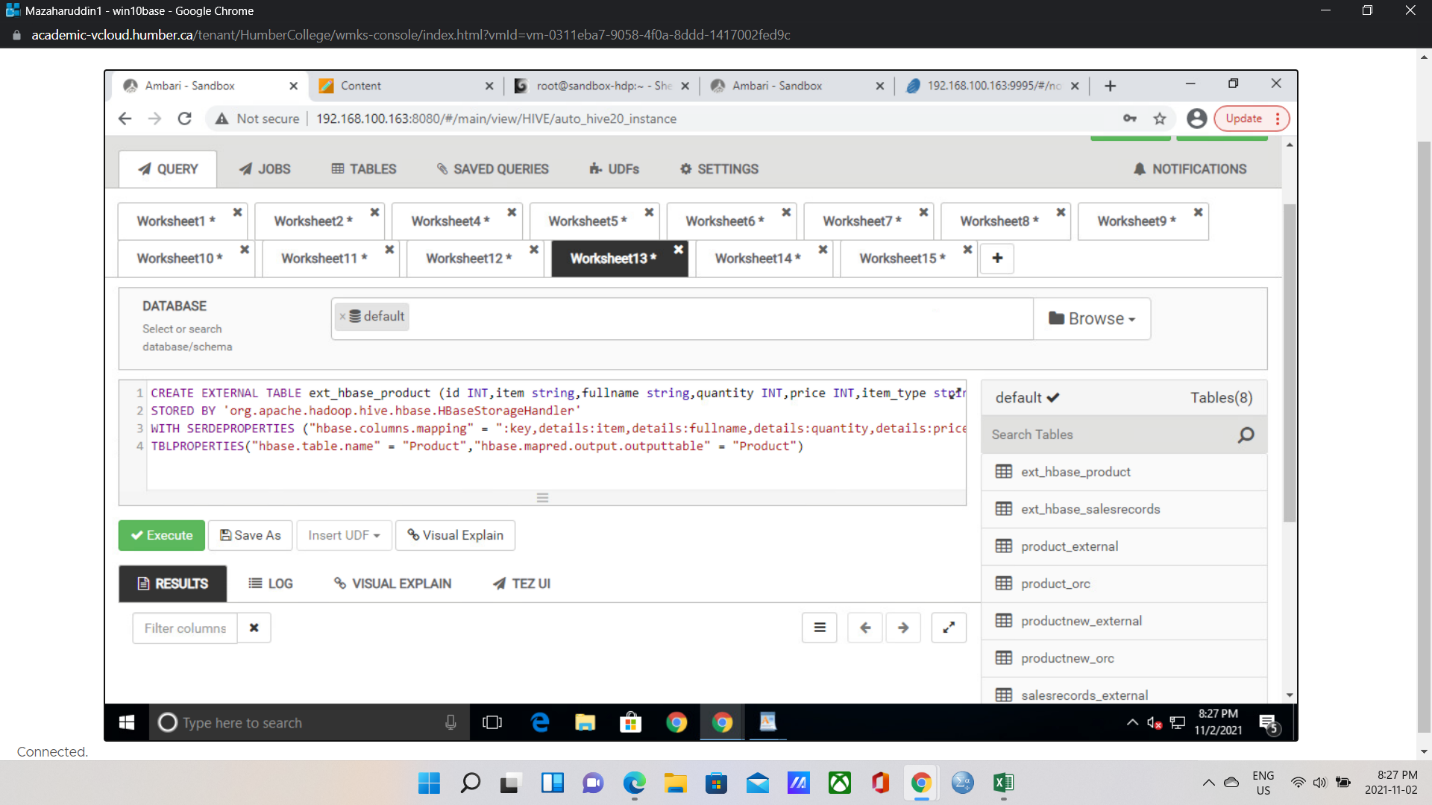
1. **create table in hive that maps directly to the hbase table** 
   1. Please review previous class notes to do this ? - **30%** of mark will be here once you create and print the definition use name ext\_hbase\_product

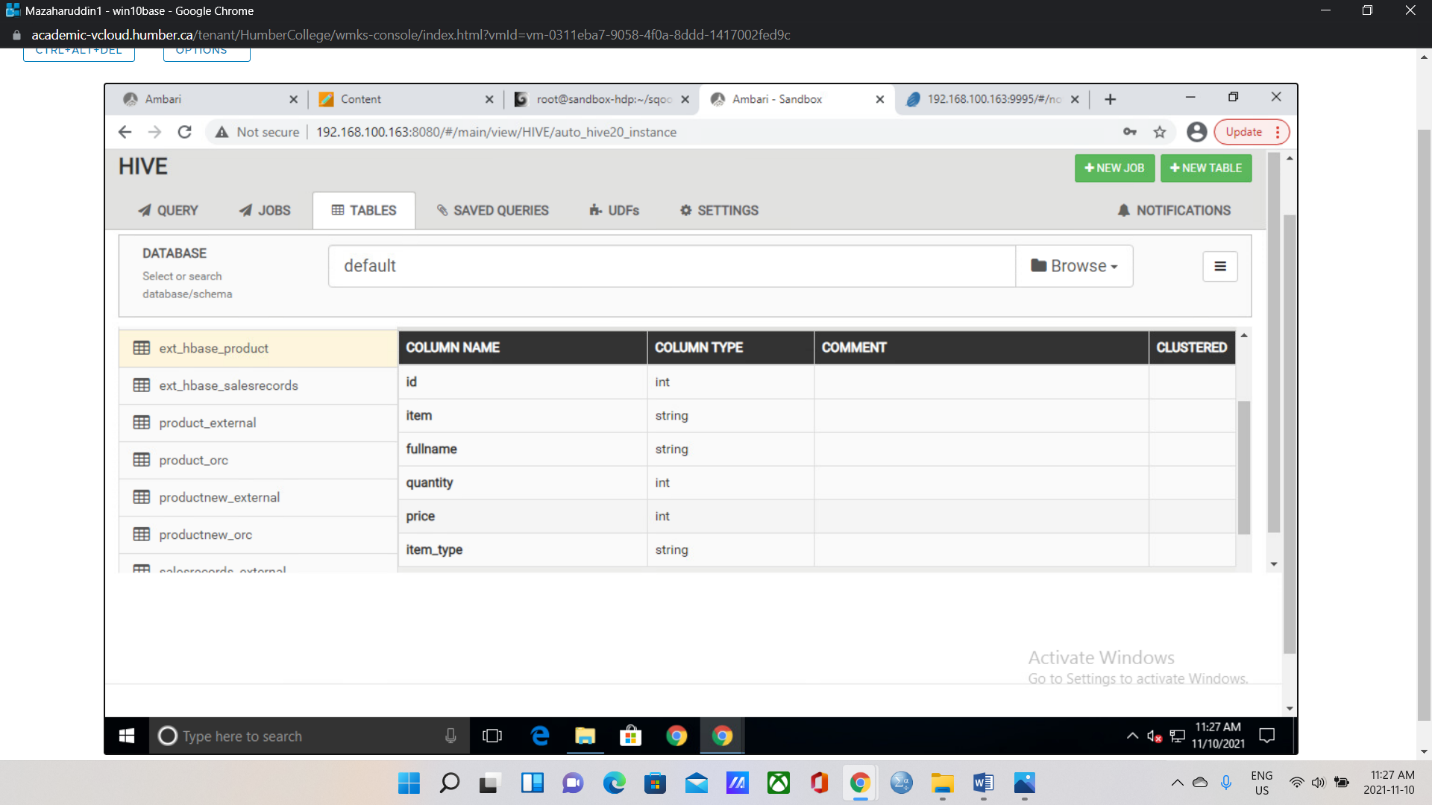
CREATE EXTERNAL TABLE ext\_hbase\_product (id INT,item string,fullname string,quantity INT,price INT,item\_type string)

STORED BY 'org.apache.hadoop.hive.hbase.HBaseStorageHandler'

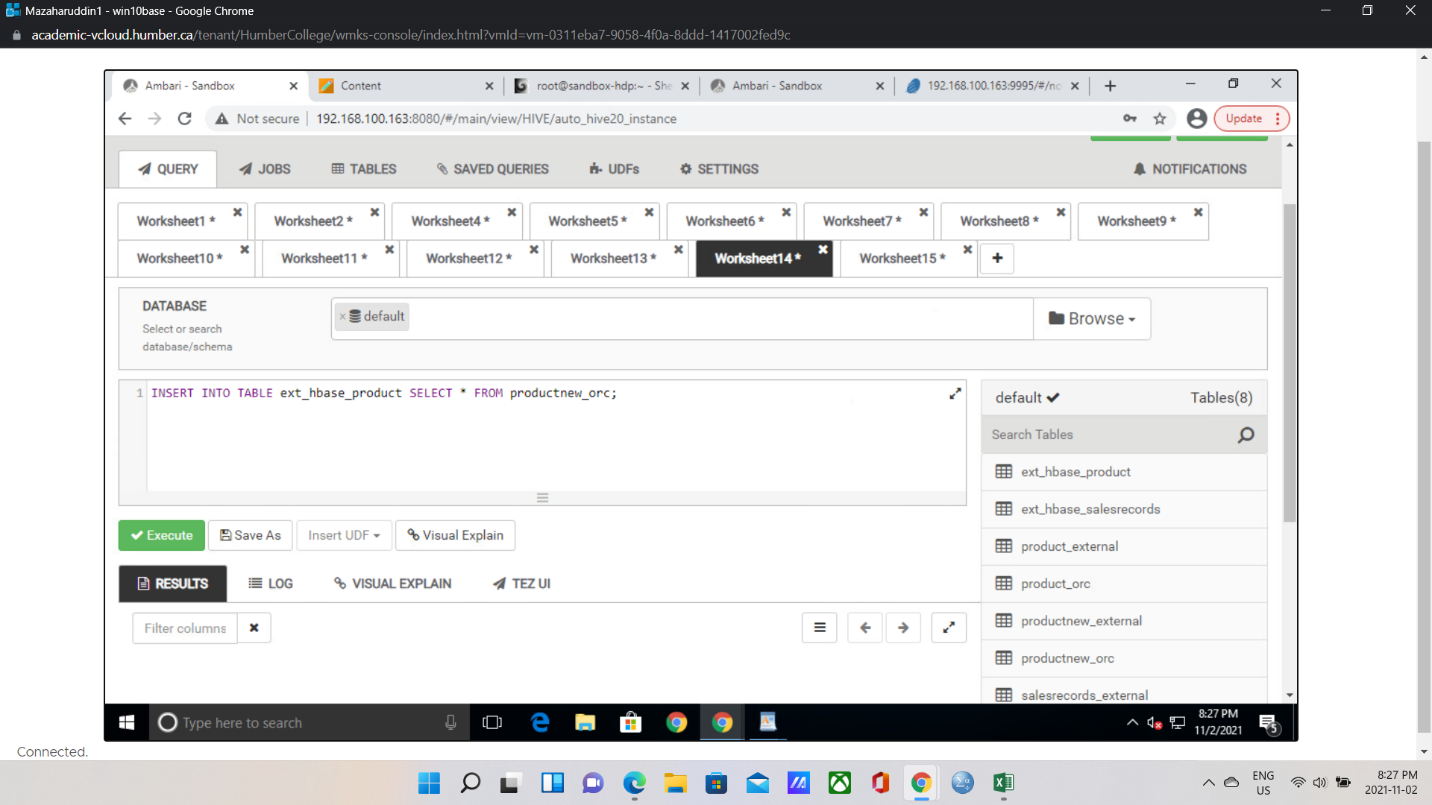
WITH SERDEPROPERTIES ("hbase.columns.mapping" = ":key,details:item,details:fullname,details:quantity,details:price,details:item\_type")

TBLPROPERTIES("hbase.table.name" = "Product","hbase.mapred.output.outputtable" = "Product")

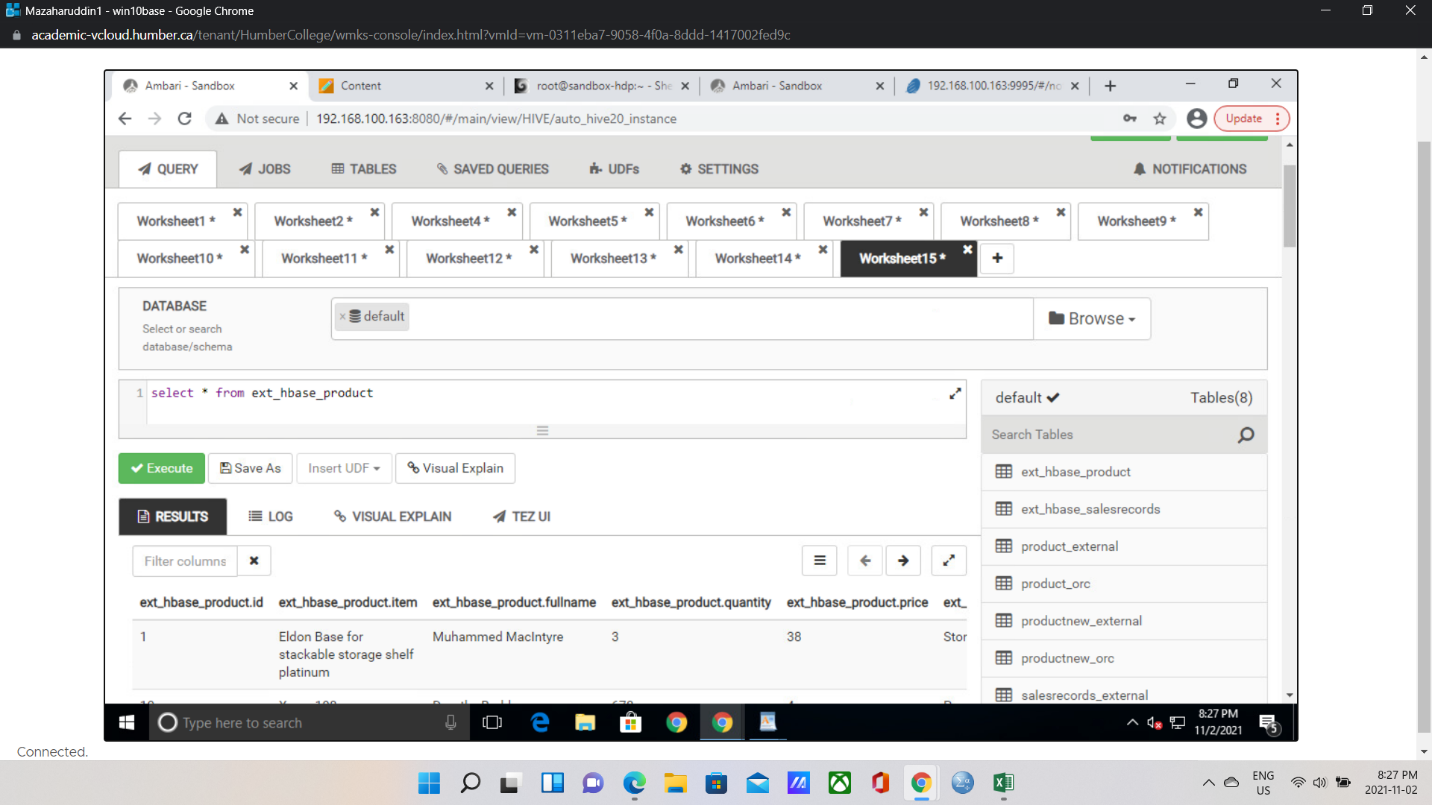




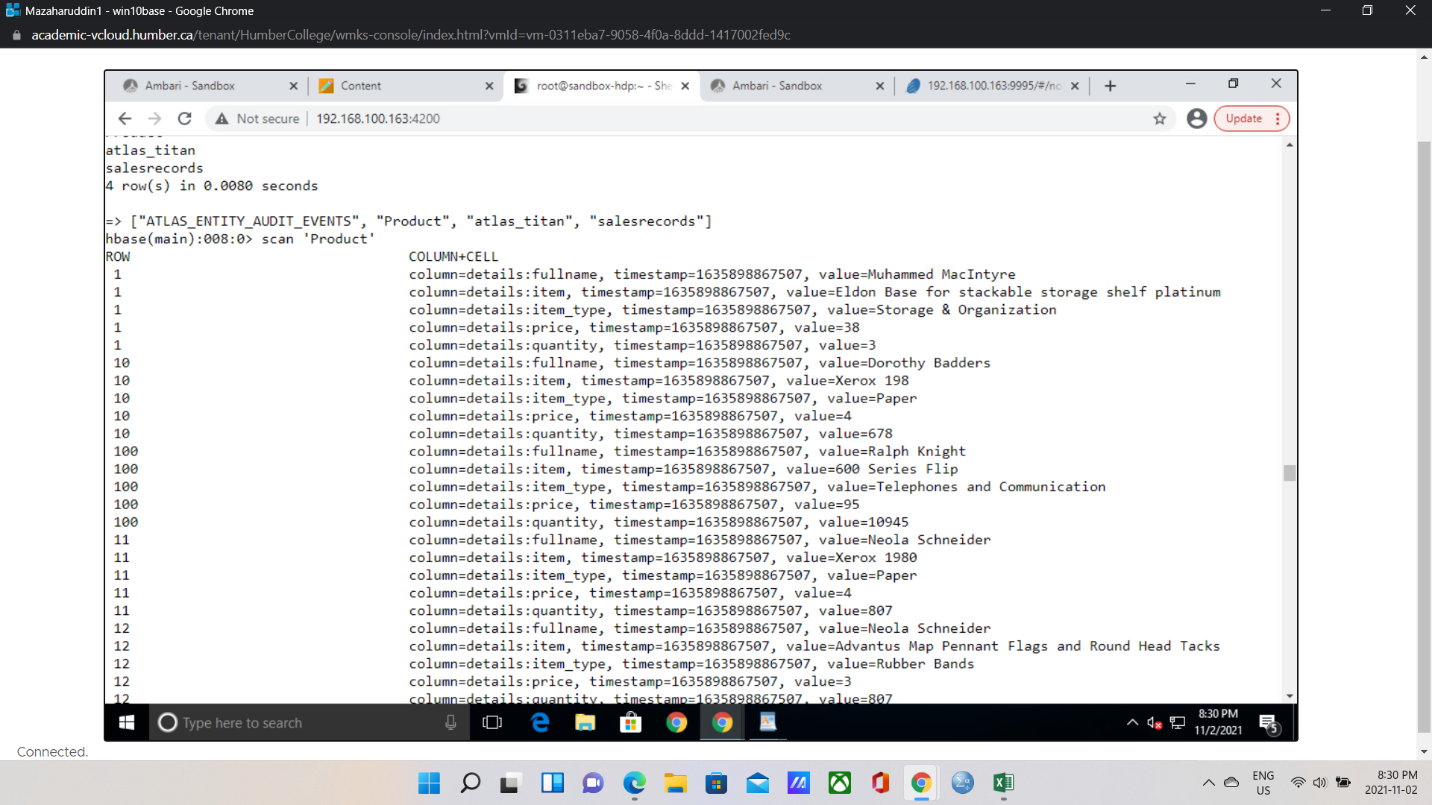
1. **From hive we insert records into Hbase table**
   1. INSERT INTO TABLE ext\_hbase\_product SELECT \* FROM product\_orc;

****

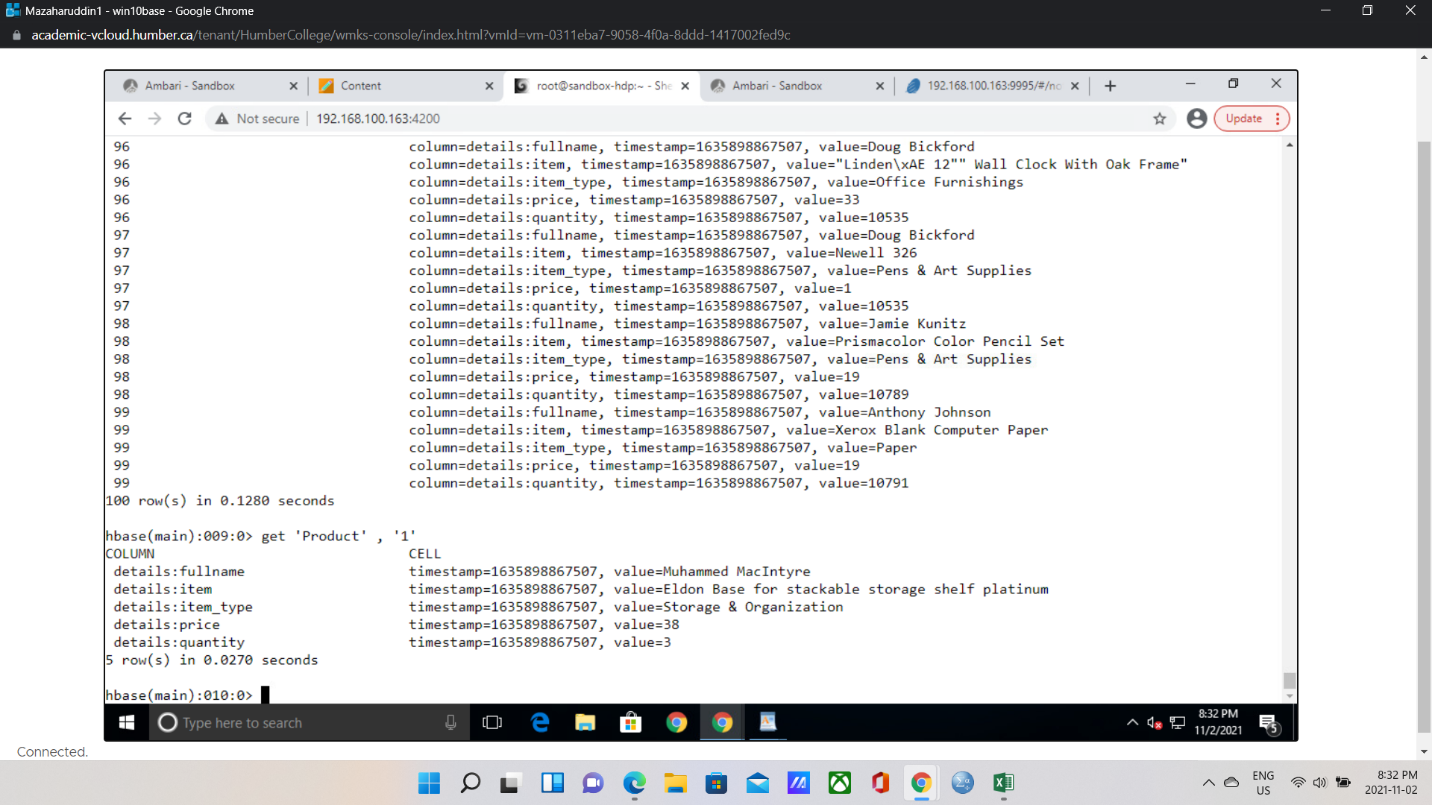
* 1. Select \* from ext\_hbase\_table - screen print first page only



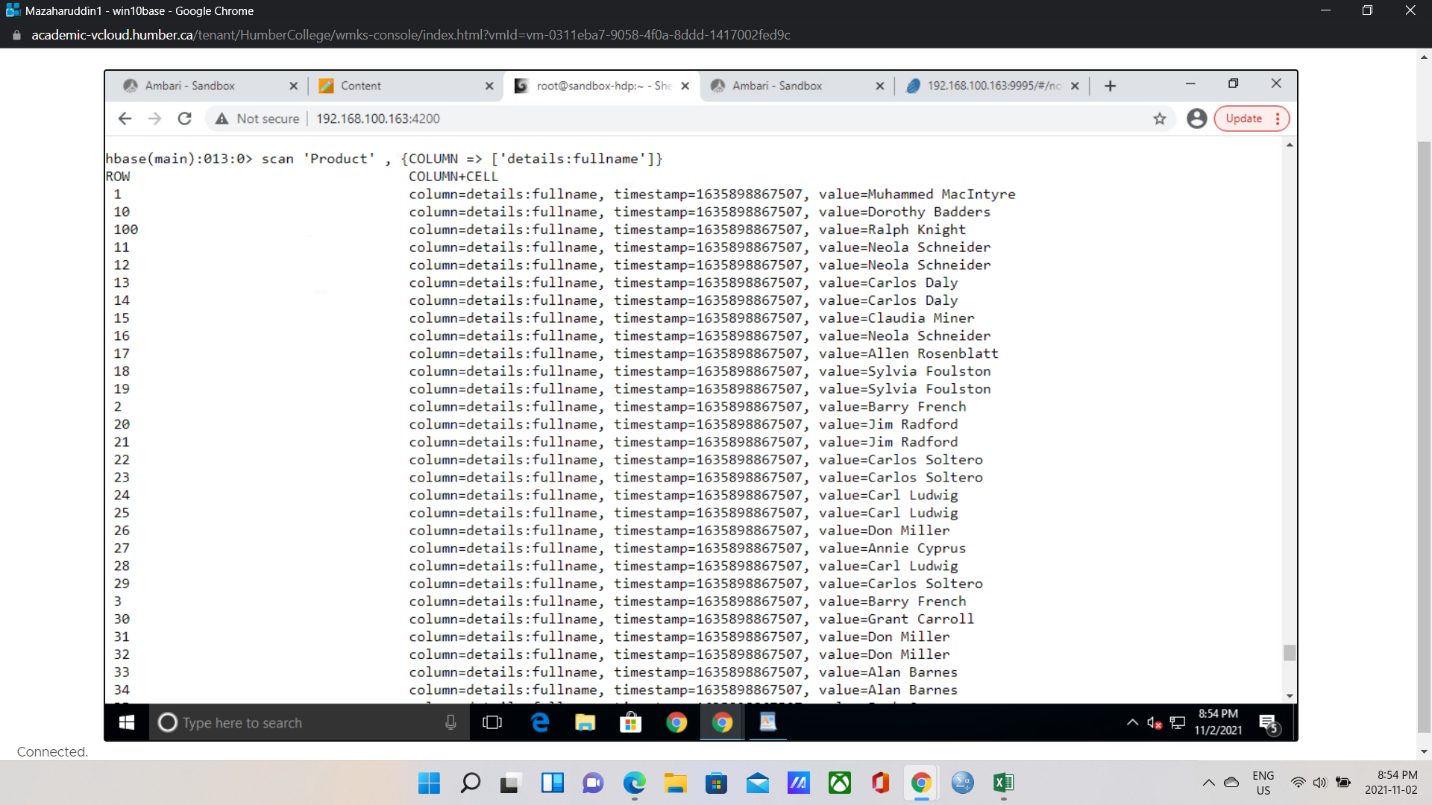
1. **From Hbase** 
   1. scan 'Product' - screen print the last page from hbase



* 1. get ‘Product’,’1’ – screen print

****

* 1. Write a filter command to show just the fullname



* 1. Write a filter command where item\_type = ‘paper’ screen print

