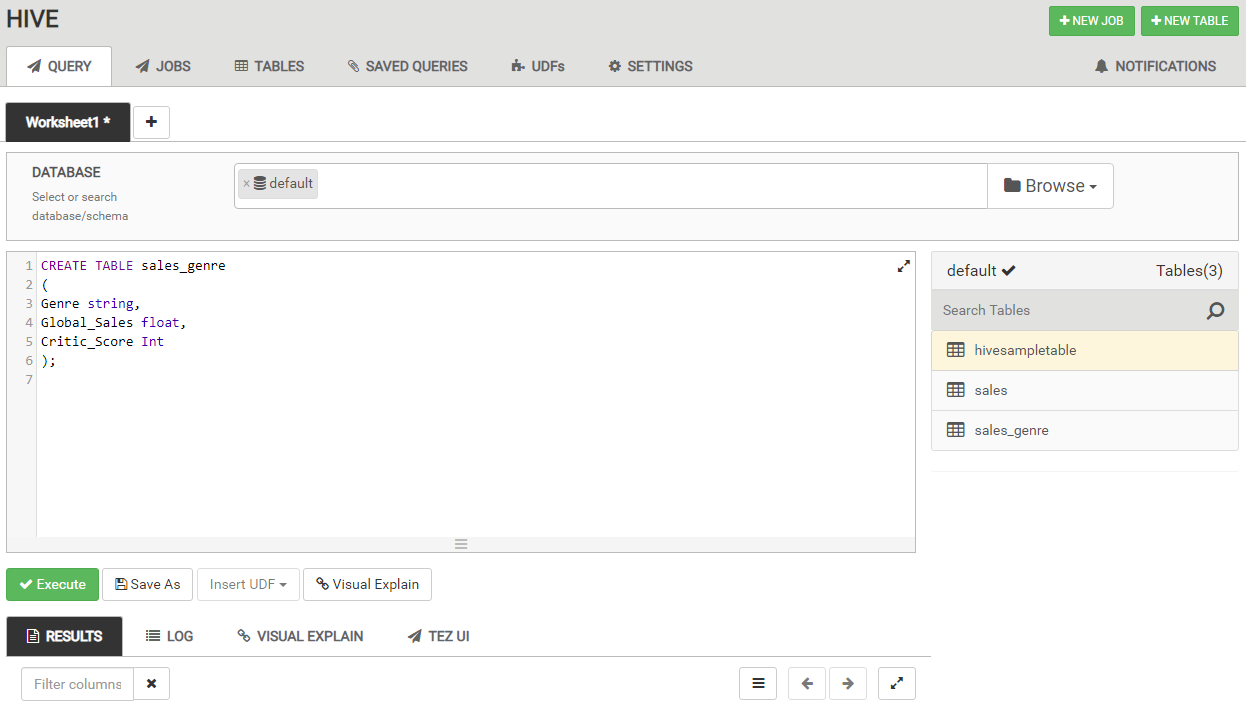
**Transforming your Data**

The below transformations were performed using the Microsoft Azure platform using Ambari Views, specifically Hive View 2.0. Before the following transformations could be completed an HDInsight’s Cluster was created and the Video games sales data was loaded into a table called “sales”. The below transformations take into consideration that the previously mentioned steps have already been completed.

**Step 1 (Create a new table for your analysis called “sales\_genre”):**

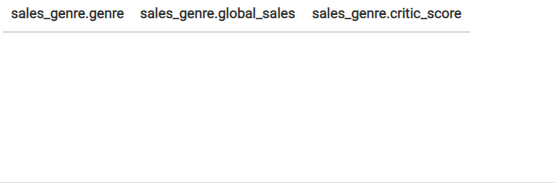
Use the code below to create the new table that will include Genre, Global\_Sales and Critic\_Score.

CREATE TABLE sales\_genre  
(  
Genre string,  
Global\_Sales float,  
Critic\_Score Int  
);



Perform the following code to ensure that your table has been created properly:

SELECT \* FROM sales\_genre LIMIT 10;



You should see a result similar to the empty table above.

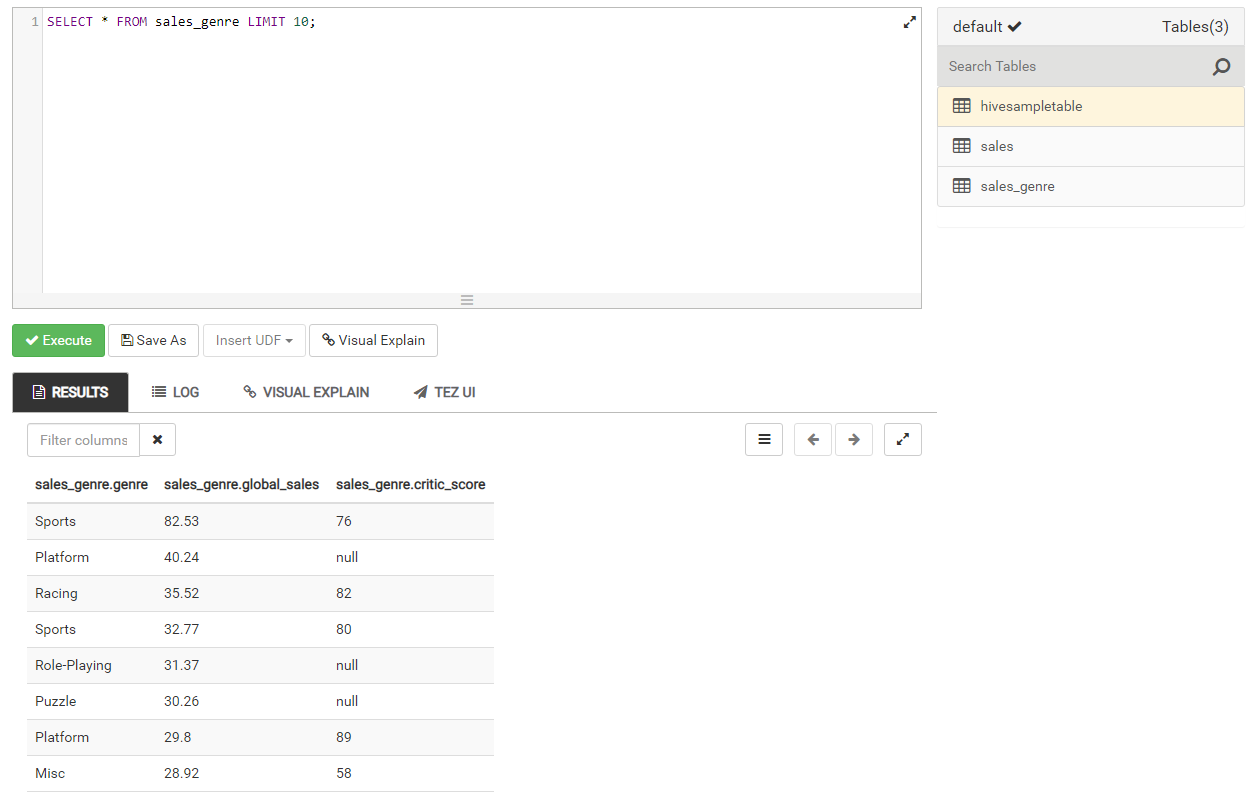
**Step 2 (Load the “sales” table data into your new “sales\_genre” table):**

Use the code below to load the data from the “sales” table into your new “sales\_genre” table:

INSERT OVERWRITE TABLE sales\_genre  
SELECT Genre, Global\_Sales, Critic\_Score  
FROM sales;

Perform the following code to ensure that your table has been created properly:

SELECT \* FROM sales\_genre LIMIT 10;



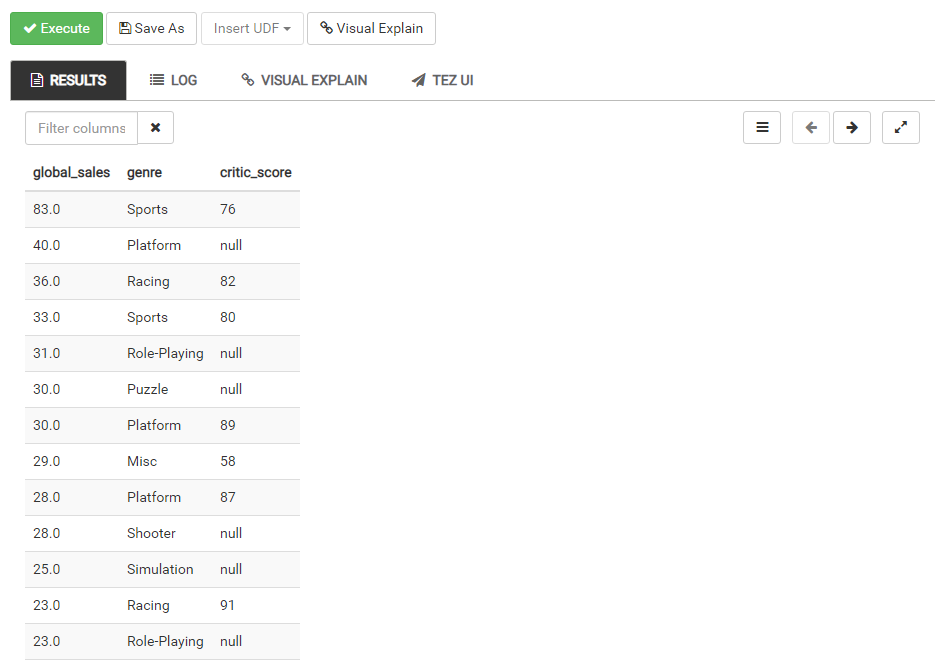
You should see a result similar to the table above.

**Step 3 (Round the data found in the “Global\_Sales” column):**

Use the code below to round the data in the “Global\_Sales” column:

SELECT ROUND(Global\_Sales) AS Global\_Sales, Genre, Critic\_Score

FROM sales\_genre;



You should see a result similar to the table above.

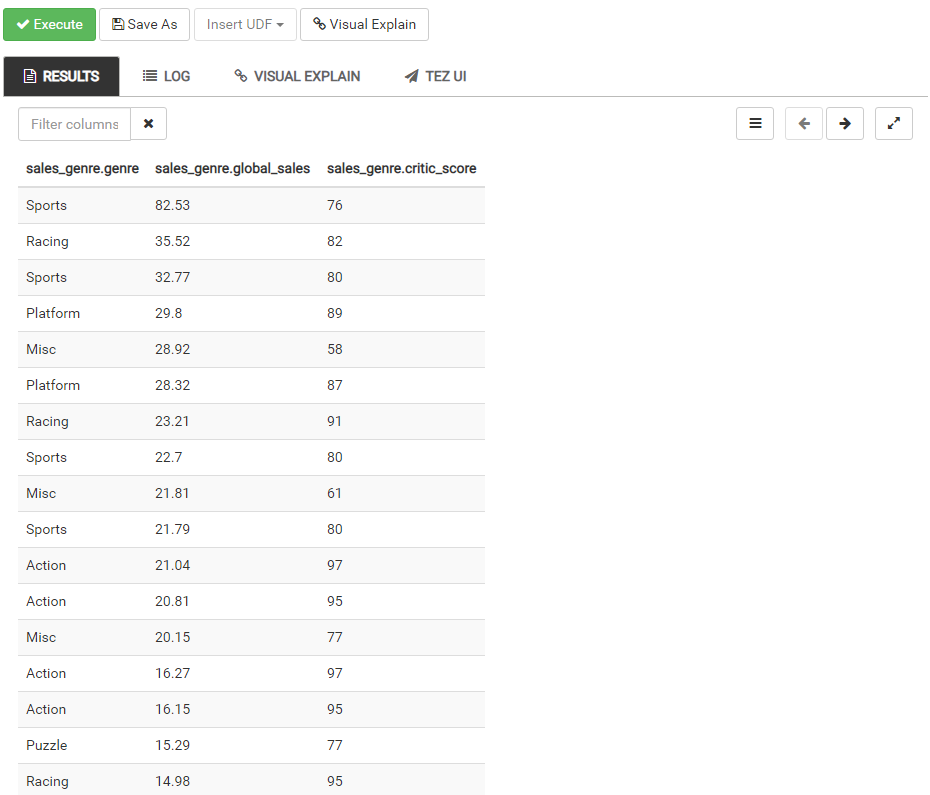
**Step 4: (Filter the data to only look at those items in the “Critic\_Score” column that are greater than 0):**

Use the code below to view all items in the “Critic\_Score” column that are greater than 0:

SELECT \*

FROM sales\_genre

WHERE critic\_score > 0;



You should see a result similar to the table above.

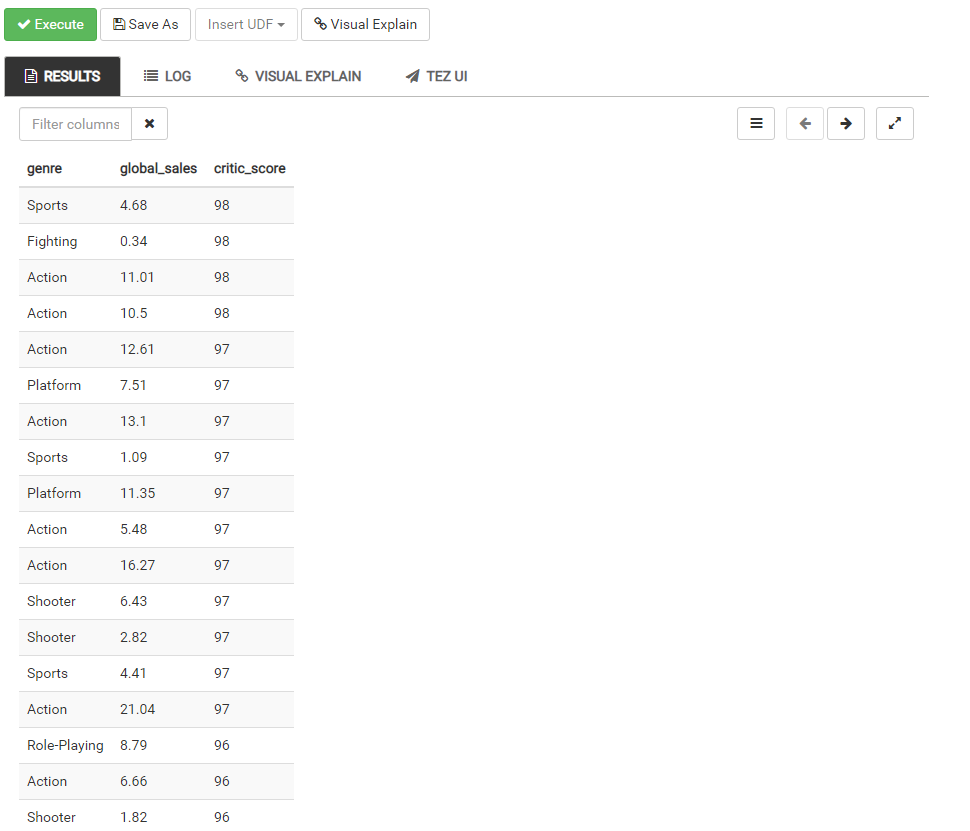
**Step 5 (Order the data in the “Critic\_Score” column from highest to lowest):**

Use the code below to order the data in the Critic\_Score column from highest to lowest:

SELECT Genre, Global\_Sales, Critic\_Score

FROM sales\_genre

ORDER BY Critic\_Score DESC;



You should see a result similar to the table above.

Congratulations you have now completed all of your data transformations!