# Assignment 2: Data Transform Activity Vithushan Esan

**Student number: 100855316** 

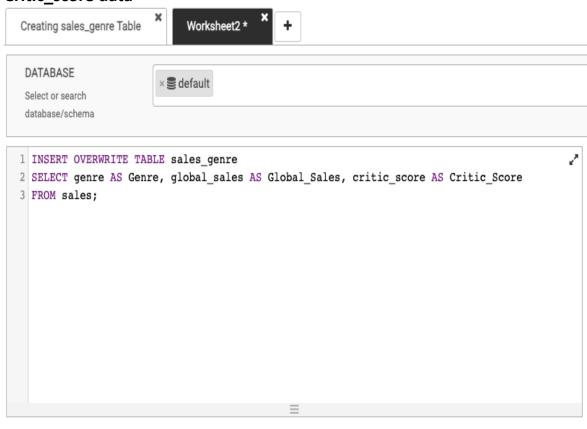
#### 1.) a) Create External Table

```
DATABASE
                             × ≥ default
 Select or search database/schema
1 CREATE EXTERNAL TABLE sales (
2 Name string,
3 Platform string,
   Year_of_Release int,
5 Genre string,
6 Publisher string,
7 NA_Sales float,
8 EU Sales float,
   JP_Sales float,
10 Other_Sales float,
11 Global Sales float,
12 Critic_Score int,
13 Critic_Count int,
14 User_Score float,
15 User_Count int,
16
   Developer string,
17 Rating string
18)
19 ROW FORMAT DELIMITED FIELDS TERMINATED BY ',' LINES TERMINATED BY '\n'
20 STORED AS TEXTFILE LOCATION '/Sales' TBLPROPERTIES("skip.header.line.count"="1");
```

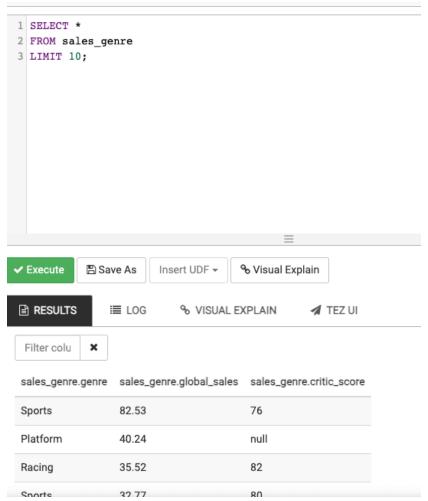
#### Create new table for sales\_genre

```
1 CREATE TABLE sales_genre
2 (
3    genre string,
4    global_sales float,
5    critic_score int
6  )
7
```

# Overwrote sales\_genre table, to include genre, global\_sales and Critic\_score data



#### Wrote Query to test if data is shown



#### 2.) Rounding data for Global sales column

```
1 INSERT OVERWRITE TABLE sales_genre
2 SELECT genre, ROUND(global_sales) AS global_sales, critic_score
3 FROM sales_genre;
```

#### Tested to see if it was rounded:

sales_genre.genre	sales_genre.global_sales	sales_genre.critic_score
Sports	83.0	76
Platform	40.0	null
Racing	36.0	82
Sports	33.0	80
Role-Playing	31.0	null
Puzzle	30.0	null
Platform	30.0	89
Misc	29.0	58
Platform	28.0	87
Shooter	28.0	null
Simulation	25.0	null
Racing	23.0	91
Role-Playing	23.0	null
Sports	23.0	80
Misc	22.0	61
Sports	22.0	80
Action	21.0	97
Action	21.0	95
Platform	21.0	null

# 3.) Filtered to look at items in Critic\_Score column that are greater than 0

```
1 INSERT OVERWRITE TABLE sales_genre
```

<sup>2</sup> SELECT genre, global\_sales, critic\_score

<sup>3</sup> FROM sales\_genre

<sup>4</sup> WHERE critic\_score > 0;

#### **Tested below:**

sales_genre.genre	sales_genre.global_sales	sales_genre.critic_score
Sports	83.0	76
Racing	36.0	82
Sports	33.0	80
Platform	30.0	89
Misc	29.0	58
Platform	28.0	87
Racing	23.0	91
Sports	23.0	80
Misc	22.0	61
Sports	22.0	80
Action	21.0	97
Action	21.0	95
Misc	20.0	77
Action	16.0	97
Action	16.0	95
Puzzle	15.0	77
Racing	15.0	95

# 4.) Ordered the data in Critic\_Score from highest to lowest

```
1 INSERT OVERWRITE TABLE sales_genre
```

<sup>2</sup> SELECT genre, global\_sales, critic\_score

<sup>3</sup> FROM sales\_genre

<sup>4</sup> ORDER BY critic\_score DESC;

### **Final Results:**

sales_genre.genre	sales_genre.global_sales	sales_genre.critic_score
Action	11.0	98
Action	11.0	98
Sports	5.0	98
Fighting	0.0	98
Action	5.0	97
Sports	4.0	97
Action	21.0	97
Shooter	3.0	97
Platform	11.0	97
Platform	8.0	97
Action	16.0	97
Shooter	6.0	97