

I am trying to set the empty values in a csv file to zero in hive. But this code doesn't seem to work. What changes should I make?

Asked 4 years, 2 months ago    Active 1 year, 5 months ago    Viewed 1k times

**This is the input .csv file**

- 1"1","","Animation"
- "2","Jumanji",""
- "","Grumpier Old Men","Comedy"



**Hive Code**

```
CREATE TABLE IF NOT EXISTS movies(movie_id int, movie_name string,genre string)

ROW FORMAT SERDE 'org.apache.hadoop.hive.serde2.OpenCSVSerde'

WITH SERDEPROPERTIES (

"separatorChar" = ",",

"quoteChar" = "\"",

"serialization.null.format" = '0'

);
```

**Output**

1	Animation
2	Jumanji
	Grumpier Old Men      Comedy

csv    hadoop    hiveql    create-table    opencsv

Share   Improve this question   Follow

edited Sep 22 '19 at 7:36

 **leftjoin**

26.5k   6   43   80

asked Dec 19 '16 at 4:53

 **Srinivas Anant**

19   3

2

Empty strings in csv are interpreted as empty strings, not NULLs. To represent NULL inside a delimited text file you should use "\N". Also Hive provides you a table property "serialization.null.format" which can be used to treat a character of your choice as null in Hive SQL. In your case it should be empty string "". To convert NULLs to zeroes use `NVL(col, 0)` or `COALESCE(col, 0)` function depending on your hive version (COALESCE should work for all).



Share Improve this answer Follow

edited Dec 20 '16 at 9:48

answered Dec 20 '16 at 9:42



leftjoin

26.5k

6

43

80