

## Split one column into multiple columns in hive

IN: HIVE



GCP: Google Cloud Platform: Data Engineer, Cloud Architect

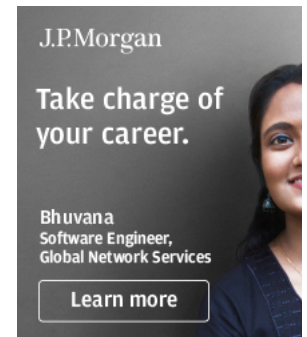
### SUGGESTED COURSE ON UDEMY

Apache Spark with Scala – Hands On with Big Data!

Spark and Python for Big Data with PySpark

Apache Kafka Series – Learn Apache Kafka for Beginners

Elasticsearch 7 and the Elastic Stack – In Depth & Hands On!



Azure Data Engineer Technologies for Beginners [DP-200, 201]

### HIVE TUTORIALS

Remove Header of CSV File in hive

Load CSV file in hive

Split one column into multiple columns in hive

Parse XML data in Hive

Partitioning in Hive

Load JSON Data in Hive non-partitioned table using Spark

Export hive data into file

Load CSV file into hive ORC table

String to Date conversion in hive

Pass variables from shell script to hive script

Load JSON Data into Hive Partitioned table using PySpark

Load Text file into Hive Table Using Spark

Load hive table into pig

Load multi character delimited file into hive

Calculate percentage in hive

Data migration from Hive to HBase

Find max value of a row in Hive

Load files into Hive Partitioned Table

Drop multiple partitions in Hive

\_\_hive\_default\_partition\_\_ in Hive

Exclude Column(s) From Select Query in Hive

Bucketing in Hive

Analytics Functions in Hive

Windowing Functions in Hive

Hive tips and shortcuts

Join in hive with example

Load CSV file into hive PARQUET table

Load CSV file into hive AVRO table

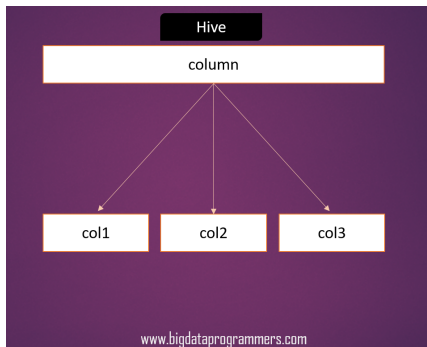
Export hive table to excel

external table in hive

Creation of hive schema

## Requirement

Suppose, you have one table in hive with one column and you want to split this column into multiple columns and then store the results into another Hive table.



## Solution

Assume the name of hive table is “transact\_tbl” and it has one column named as “connections”, and values in connections column are comma separated and total two commas are present in each value.

### Step 1: Create Hive Table

Create an input table transact\_tbl in bdp schema using below command.

- Show or hide the code

```
01. CREATE TABLE bdp.transact_tbl(connections STRING);
```

### Step 2: Insert Data in Hive

Insert data into connections columns, String should be comma separated. For e.g. “PSTP, POST, FRDE” is one value of connections column.

- Show or hide the code

```
01. INSERT INTO TABLE transact_tbl VALUES('PSTP,POST,FRDE'),
02. ('LEFD,HSYT,YDTH'),
03. ('AQSV,NESA,LRAJ'),
04. ('KSAI,SJAA,PSOA'),
05. ('SHAG,JSGS,AGAQ'),
06. ('AQDS,KAHY,AGQG'),
07. ('BAGW,AHQG,HAFY');
```

### Step 3: Show Hive data

Let’s see how data looks in a transact\_tbl table using below command:

- Show or hide the code



```
> SELECT connections FROM bdp.transact_tbl;
OK
connections
PSTP,POST,FRDE
LEFD,HSYT,YDTH
AQSV,NESA,LRAJ
KSAI,SJAA,PSOA
SHAG,JSGS,AGAQ
AQDS,KAHY,AQGS
BAGW,AHQG,HAFY
Time taken: 0.22 seconds, Fetched: 7 row(s)
```

Take charge of  
your career.

Bhuvana  
Software Engineer,  
Global Network Services

Learn more



Streaming Big Data with Spark Streaming  
& Scala – Hands On!

J.P.Morgan

Take charge of  
your career.

Bhuvana  
Software Engineer,  
Global Network Services

Learn more

#### RECENT POSTS

- [Create Delta Table with Partition in Databricks](#)
- [Create Delta Table from Path in Databricks](#)
- [Create Delta Table in Databricks](#)
- [Delta Table in Databricks](#)
- [Delta Lake in Databricks](#)
- [Read file from Azure Data Lake Gen2 using Python](#)
- [Read file from Azure Data Lake Gen2 using Spark](#)
- [Create Mount Point in Azure Databricks Using Service Principal and OAuth](#)
- [Create Service Principal in Azure](#)
- [filter a Map in spark using scala](#)

#### RECOMMENDATIONS

Certifications

Top Machine Learning Courses You Shouldn't Miss

Top courses for data engineers

Top Big Data Courses on Udemy You should Take

## Step 4: Create Output hive table

Create an output table where you want to store split values. Run below command in the hive:

- Show or hide the code

```
01. CREATE TABLE optrans_tbl(con1 STRING,con2 STRING,con3 STRING);
```

## Step 5: Use Hive function

There is a built-in function SPLIT in the hive which expects two arguments, the first argument is a string and the second argument is the pattern by which string should separate.

It will convert String into an array, and desired value can be fetched using the right index of an array.

## Data Science

Co-created with IBM and Designed by Industry Professionals  
to Get the Best Jobs in Top MNCs

intellipaat.com

OPEN

Use below query to store split records in the hive table:-

- Show or hide the code

```
01. INSERT INTO TABLE bdp.optrans_tbl
02. SELECT CONCTNS.splitted_cnctns[0] AS con1,
03. CONCTNS.splitted_cnctns[1] AS con2,
04. CONCTNS.splitted_cnctns[2] AS con3
05. FROM (SELECT split(connections,',') AS splitted_cnctns
06. FROM bdp.transact_tbl)CONCTNS;
```

In the above query, you can see that splitted\_cnctns is an array with three values in it, which can be extracted using the proper index as con1, con2, and con3.

Inner query is used to get the array of split values and the outer query is used to assign each value to a separate column.

## Step 6: Show output

Let's see how data looks like in an optrans\_tbl table using below command:



It should show output as below.-

```
>
> SELECT con1,con2,con3 FROM bdp.optrans_tbl;
OK
con1    con2    con3
PSTP    POST    PRDE
LEFD    HSYT    YDTH
AQSV    NESA    LRAJ
KSAI    SJAA    PSCA
SHAG    JSJS    AGAQ
AQDS    KAHY    AGQG
BAGW    AHQG    HAFY
Time taken: 0.333 seconds, Fetched: 7 row(s)
```



Keep Learning 😊

Don't miss the tutorial on [Top Big data courses on Udemy you should Buy](#)

Sharing is caring!

Share

Tweet

LinkedIn

Subscribe to our newsletter

Name\*

Email\*

SUBMIT



Know How Telegram Works

Ad Scaler

Load CSV file into hive ORC table

bigdataprogrammers.com

Business Analytics - 10 - Month Flagship Program

Ad docs.google.com

Export hive data into file

bigdataprogramn

Configure your server today - 3rd Gen. AMD EPYC 7003 Compare

Ad boston.co.uk

Partitioning in Hive

bigdataprogrammers.com

Load CSV file into hive PARQUET table

bigdataprogrammers.com

Load JSON Hive non-pa table using

bigdataprogramn

Previous Post: [Load CSV file in Pig](#)

Next Post: [Parse XML data in Hive](#)

## 2 COMMENTS

what if i want one more column other than the split String, say transact\_tbl table have one more column amount(Every connections have different





I have solved it, thank you 😊

KUMAR SOURABH | AUGUST 27, 2017 | [PERMALINK](#) | [LOG IN TO REPLY](#)

## LEAVE A REPLY

You must be [logged in](#) to post a comment.



# Buy Sell Signal Software

98% Accurate Trading Indicator Software

RichLiveTrade Software

## TAGS

[apache spark](#) [Azure](#) [Azure Databricks](#) [big data](#) [cloud](#) [csv](#) [csv file](#) [databricks](#) [dataframe](#) [Delta Table](#) [external table](#) [full join](#) [hadoop](#) [hbase](#) [hdfs](#) [hive](#) [hive interview](#) [import](#) [inner join](#) [IntelliJ](#) [interview](#) [qa](#) [interview questions](#) [join](#) [json](#) [key vault](#) [left join](#) [load](#) [MapReduce](#) [mysql](#) [partition](#) [percentage](#) [pig](#) [pyspark](#) [python](#) [quiz](#) [RDD](#) [right join](#) [sbt](#) [scala](#) [spark](#) [spark-shell](#) [spark dataframe](#) [spark sql](#) [sparksql](#) [sqoop](#)

[Recommended Courses For Data Engineers](#)

[Recommended Books to Become Data Engineer](#)

[Top Big Data Courses on Udemy You should Take](#)

[Top courses for data engineers](#)

[Certifications](#)

[Top Machine Learning Courses You Shouldn't Miss](#)

Subscribe to our newsletter

Name\*

Email\*

SUBMIT

[Big Data Mentorship](#)  
[Get Help From Big Data Expert](#)  
[Software Development](#)  
[QA Forum](#)  
[Quizzes](#)  
[Scenario Based Problems](#)  
[Interview Q&A](#)  
[BIG DATA Tutorials](#)  
[Video Tutorials](#)

[Spark-Scala Tutorials](#)  
[PySpark Tutorials](#)  
[Azure Tutorials](#)  
[Hive Tutorials](#)  
[HBase Tutorials](#)  
[MapReduce Tutorials](#)  
[Pig Tutorials](#)  
[Shell Script Tutorials](#)  
[Sqoop Tutorials](#)  
[Linux Tutorials](#)  
[Python Tutorials](#)  
[DataIngestion](#)  
[Getting Started](#)  
[Tech](#)

[About us](#)  
[Contact Us](#)  
[Cookie Policy](#)  
[Privacy Policy](#)  
[Terms of Service](#)

## DISCLAIMER

All trademarks and registered trademarks appearing on bigdataprogrammers.com are the property of their respective owners.

