Home (/) About (https://sparkbyexamples.com/about-sparkbyexamples/) | { One stop for all Spark Examples }

Spark by {Examples} (https://sparkbyexamples.com/)

Spark (https://sparkbyexamples.com/)
PySpark (https://sparkbyexamples.com/pyspark-tutorial/)
Hive (https://sparkbyexamples.com/apache-hive-tutorial/)
HBase (https://sparkbyexamples.com/apache-hbase-tutorial/)
Kafka (https://sparkbyexamples.com/apache-kafka-tutorials-with-examples/)
FAQ's (https://sparkbyexamples.com/spark-questions/)
More (https://sparkbyexamples.com/)

Spark Tutorial

Spark – Installation on Windows (https://sparkbyexamples.com/spark/apache-spark-installation-on-windows/)

Spark – Installation on Linux | Ubuntu (https://sparkbyexamples.com/spark/spark-installation-on-linux-ubuntu/)

Spark – Cluster Setup with Hadoop Yarn (https://sparkbyexamples.com/spark/spark-setup-on-hadoop-yarn/)

Spark – Web/Application UI (https://sparkbyexamples.com/spark/spark-web-ui-understanding/)

Spark – Setup with Scala and IntelliJ (https://sparkbyexamples.com/spark/spark-setup-run-with-scala-intellij/)

Spark – How to Run Examples From this Site on IntelliJ IDEA (https://sparkbyexamples.com/spark/how-to-run-spark-examples-from-intellij/)

Spark – SparkSession (https://sparkbyexamples.com/spark/sparksession-explained-with-examples/)

Spark – SparkContext (https://sparkbyexamples.com/spark/spark-sparkcontext/)

Spark RDD Tutorial

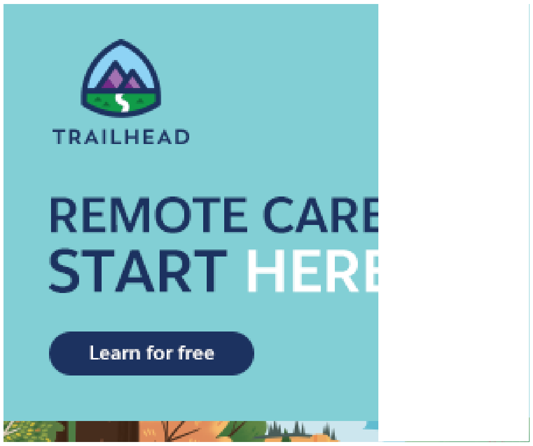
Spark RDD – Parallelize (https://sparkbyexamples.com/apache-spark-rdd/how-to-create-an-rdd-using-parallelize/)

Spark explode array and map columns to rows

NNK (https://sparkbyexamples.com/author/admin/) - Apache Spark (https://sparkbyexamples.com/category/spark/) / Spark SQL Functions (https://sparkbyexamples.com/category/spark/spark-sql-functions/)

In this article, I will explain how to explode array or list and map DataFrame columns to rows using different Spark explode functions (explode, explore_outer, posexplode, posexplode_outer) with Scala example.

- Spark RDD – Read text file (https://sparkbyexamples.com/apache-spark-rdd/spark-read-multiple-text-files-into-a-single-rdd/)
- Spark RDD – Read CSV (https://sparkbyexamples.com/apache-spark-rdd/spark-load-csv-file-into-rdd/)
- Spark RDD – Create RDD (https://sparkbyexamples.com/apache-spark-rdd/different-ways-to-create-spark-rdd/)
- Spark RDD – Create Empty RDD (https://sparkbyexamples.com/apache-spark-rdd/spark-how-to-create-an-empty-rdd/)
- Spark RDD – Transformations (https://sparkbyexamples.com/apache-spark-rdd/spark-rdd-transformations/)
- Spark RDD – Actions (https://sparkbyexamples.com/apache-spark-rdd/spark-rdd-actions/)
- Spark RDD – Pair Functions (https://sparkbyexamples.com/apache-spark-rdd/spark-pair-rdd-functions/)
- Spark RDD – Repartition and Coalesce (https://sparkbyexamples.com/spark/spark-repartition-vs-coalesce/)
- Spark RDD – Shuffle Partitions (https://sparkbyexamples.com/spark/spark-shuffle-partitions/)
- Spark RDD – Cache vs Persist (https://sparkbyexamples.com/spark/spark-difference-between-cache-and-persist/)
- Spark RDD – Persistance Storage Levels (https://sparkbyexamples.com/spark/spark-persistence-storage-levels/)



While working with structured files like JSON (https://sparkbyexamples.com/spark/spark-read-and-write-json-file/), Parquet (https://sparkbyexamples.com/spark/spark-read-write-dataframe-parquet-example/), Avro (https://sparkbyexamples.com/spark/read-write-avro-file-spark-dataframe/), and XML (https://sparkbyexamples.com/spark/processing-xml-files-in-spark/) we often get data in collections like arrays, lists, and maps, In such cases, these explode functions are useful to convert collection columns to rows in order to process in Spark effectively.

Related:

- How to flatten nested Struct column (https://sparkbyexamples.com/spark/spark-flatten-nested-struct-column/)
- How to flatten nested array column (https://sparkbyexamples.com/spark/spark-flatten-nested-array-column-to-single-column/)

Though I've explained here with Scala, a similar method could be used to explode array and map columns to rows with PySpark and if time permits I will cover it in the future. If you are looking for PySpark, I would still recommend reading through this article as it would give you an Idea on Spark explode functions and usage.

[Spark RDD – Broadcast Variables (https://sparkbyexamples.com/spark/spark-broadcast-variables/)](https://sparkbyexamples.com/spark/spark-broadcast-variables/)

[Spark RDD – Accumulator Variables (https://sparkbyexamples.com/spark/spark-accumulators/)](https://sparkbyexamples.com/spark/spark-accumulators/)

[Spark RDD – Convert RDD to DataFrame (https://sparkbyexamples.com/apache-spark-rdd/convert-spark-rdd-to-dataframe-dataset/)](https://sparkbyexamples.com/apache-spark-rdd/convert-spark-rdd-to-dataframe-dataset/)

Spark SQL Tutorial

[Spark SQL – Create DataFrame (https://sparkbyexamples.com/spark/different-ways-to-create-a-spark-dataframe/)](https://sparkbyexamples.com/spark/different-ways-to-create-a-spark-dataframe/)

[Spark SQL – Select Columns (https://sparkbyexamples.com/spark/spark-select-columns-from-dataframe/)](https://sparkbyexamples.com/spark/spark-select-columns-from-dataframe/)

[Spark SQL – Add and Update Column (withColumn) (https://sparkbyexamples.com/spark/spark-dataframe-withcolumn/)](https://sparkbyexamples.com/spark/spark-dataframe-withcolumn/)

[Spark SQL – Rename Nested Column (https://sparkbyexamples.com/spark/rename-a-column-on-spark-dataframes/)](https://sparkbyexamples.com/spark/rename-a-column-on-spark-dataframes/)

[Spark SQL – Drop column (https://sparkbyexamples.com/spark/spark-drop-column-from-dataframe-dataset/)](https://sparkbyexamples.com/spark/spark-drop-column-from-dataframe-dataset/)

[Spark SQL – Where | Filter (https://sparkbyexamples.com/spark/spark-dataframe-where-filter/)](https://sparkbyexamples.com/spark/spark-dataframe-where-filter/)

[Spark SQL – When Otherwise (https://sparkbyexamples.com/spark/spark-case-when-otherwise-example/)](https://sparkbyexamples.com/spark/spark-case-when-otherwise-example/)

Before we start, let's create a DataFrame with array and map fields, below snippet, creates a DF with columns "name" as StringType, "knownLanguage" as [ArrayType (https://sparkbyexamples.com/spark/spark-array-arraytype-dataframe-column/)](https://sparkbyexamples.com/spark/spark-array-arraytype-dataframe-column/)and "properties" as [MapType (https://sparkbyexamples.com/spark/spark-dataframe-map-maptype-column/)](https://sparkbyexamples.com/spark/spark-dataframe-map-maptype-column/).

And, from below code, "spark" is an instance of SparkSession, please refer to complete code at the end to see how to create SparkSession object.

```
import spark.implicits._

val arrayData = Seq(
  Row("James",List("Java","S
Row("Michael",List("Spark","
Row("Robert",List("CSharp","
Row("Washington",null,null),
Row("Jefferson",List(),Map()
)

val arraySchema = new Struct
  .add("name",StringType)
  .add("knownLanguages", Arr
  .add("properties", MapType

val df = spark.createDataFra
df.printSchema()
df.show(false)
```

[Spark SQL – Collect data to Driver](https://sparkbyexamples.com/spark/spark-dataframe-collect/) (https://sparkbyexamples.com/spark/spark-dataframe-collect/)

[Spark SQL – Distinct](https://sparkbyexamples.com/spark/spark-remove-duplicate-rows/) (https://sparkbyexamples.com/spark/spark-remove-duplicate-rows/)

[Spark SQL- Pivot Table DataFrame](https://sparkbyexamples.com/spark/how-to-pivot-table-and-unpivot-a-spark-dataframe/) (https://sparkbyexamples.com/spark/how-to-pivot-table-and-unpivot-a-spark-dataframe/)

[Spark SQL – Data Types](https://sparkbyexamples.com/spark/spark-sql-dataframe-data-types/) (https://sparkbyexamples.com/spark/spark-sql-dataframe-data-types/)

[Spark SQL – StructType | StructField](https://sparkbyexamples.com/spark/spark-sql-structtype-on-dataframe/) (https://sparkbyexamples.com/spark/spark-sql-structtype-on-dataframe/)

[Spark SQL – Schema](https://sparkbyexamples.com/spark/spark-schema-explained-with-examples/) (https://sparkbyexamples.com/spark/spark-schema-explained-with-examples/)

[Spark SQL – Groupby](https://sparkbyexamples.com/spark/using-groupby-on-dataframe/) (https://sparkbyexamples.com/spark/using-groupby-on-dataframe/)

[Spark SQL – Sort DataFrame](https://sparkbyexamples.com/spark/spark-how-to-sort-dataframe-column-explained/) (https://sparkbyexamples.com/spark/spark-how-to-sort-dataframe-column-explained/)

[Spark SQL – Join Types](https://sparkbyexamples.com/spark/spark-sql-dataframe-join/) (https://sparkbyexamples.com/spark/spark-sql-dataframe-join/)

[Spark SQL – Union and UnionAll](https://sparkbyexamples.com/spark/spark-dataframe-union-and-union-all/) (https://sparkbyexamples.com/spark/spark-dataframe-union-and-union-all/)

[Spark SQL – map() vs mapPartitions()](https://sparkbyexamples.com/spark/spark-map-vs-mappartitions-transformation/) (https://sparkbyexamples.com/spark/spark-map-vs-mappartitions-transformation/)

explode – spark explode array or map column to rows

Spark function `explode(e: Column)` is used to explode or create array or map columns to rows. When an array is passed to this function, it creates a new default column “col1” and it contains all array elements. When a map is passed, it creates two new columns one for key and one for value and each element in map split into the row.

This will ignore elements that have null or empty. from the above example, Washington and Jefferson have null or empty values in array and map, hence the following snippet out does not contain these rows.

explode – array column example

```
  df.select($"name",explode($"
    .show(false)
```

Outputs:

```
+-------+------+
|name   |col   |
+-------+------+
|James  |Java  |
|James  |Scala |
|Michael|Spark |
|Michael|Java  |
|Michael|null  |
|Robert |CSharp|
|Robert |      |
+-------+------+
```

[Spark SQL – foreach() vs foreachPartition() (https://sparkbyexamples.com/spark/spark-foreachpartition-vs-foreach-explained/)]

[Spark SQL – map() vs flatMap() (https://sparkbyexamples.com/spark/spark-map-vs-flatmap-with-examples/)]

[Spark SQL – Persist and Cache (https://sparkbyexamples.com/spark/spark-dataframe-cache-and-persist-explained/)]

[Spark SQL – UDF (User Defined Functions) (https://sparkbyexamples.com/spark/spark-sql-udf/)]

[Spark SQL – Array (ArrayType) Column (https://sparkbyexamples.com/spark/spark-array-arraytype-dataframe-column/)]

[Spark SQL – Map (MapType) column (https://sparkbyexamples.com/spark/spark-dataframe-map-maptype-column/)]

[Spark SQL – Flatten Nested Struct Column (https://sparkbyexamples.com/spark/spark-flatten-nested-struct-column/)]

[Spark SQL – Flatten Nested Array Column (https://sparkbyexamples.com/spark/spark-flatten-nested-array-column-to-single-column/)]

[Spark SQL – Explode Array & Map Columns (https://sparkbyexamples.com/spark/explode-spark-array-and-map-dataframe-column/)]

[Spark SQL – Sampling (https://sparkbyexamples.com/spark/spark-sampling-with-examples/)]

explode – map column example

```
    df.select($"name",explode($"
      .show(false)
```

Outputs:

```
+-------+----+-----+
|name   |key |value|
+-------+----+-----+
|James  |hair|black|
|James  |eye |brown|
|Michael|hair|brown|
|Michael|eye |null |
|Robert |hair|red  |
|Robert |eye |     |
+-------+----+-----+
```

explode_outer – Create rows for each element in an array or map.

Spark SQL `explode_outer(e: Column)` function is used to create a row for each element in the array or map column. Unlike explode, if the array or map is null or empty, explode_outer returns null.

[Spark SQL – Partitioning (https://sparkbyexamples.com/spark/spark-partitioning-understanding/)](https://sparkbyexamples.com/spark/spark-partitioning-understanding/)

Spark SQL Functions

[Spark SQL String Functions (https://sparkbyexamples.com/spark/usage-of-spark-sql-string-functions/)](https://sparkbyexamples.com/spark/usage-of-spark-sql-string-functions/)

[Spark SQL Date and Timestamp Functions (https://sparkbyexamples.com/spark/spark-sql-date-and-time-functions/)](https://sparkbyexamples.com/spark/spark-sql-date-and-time-functions/)

[Spark SQL Array Functions (https://sparkbyexamples.com/spark/spark-sql-array-functions/)](https://sparkbyexamples.com/spark/spark-sql-array-functions/)

[Spark SQL Map Functions (https://sparkbyexamples.com/spark/spark-sql-map-functions/)](https://sparkbyexamples.com/spark/spark-sql-map-functions/)

[Spark SQL Sort Functions (https://sparkbyexamples.com/spark/spark-sql-sort-functions/)](https://sparkbyexamples.com/spark/spark-sql-sort-functions/)

[Spark SQL Aggregate Functions (https://sparkbyexamples.com/spark/spark-sql-aggregate-functions/)](https://sparkbyexamples.com/spark/spark-sql-aggregate-functions/)

[Spark SQL Window Functions (https://sparkbyexamples.com/spark/spark-sql-window-functions/)](https://sparkbyexamples.com/spark/spark-sql-window-functions/)

[Spark SQL JSON Functions (https://sparkbyexamples.com/spark/spark-most-used-json-functions-with-examples/)](https://sparkbyexamples.com/spark/spark-most-used-json-functions-with-examples/)

Spark Data Source API

[Spark – Read & Write CSV file (https://sparkbyexamples.com/spark/spark-read-csv-file-into-dataframe/)](https://sparkbyexamples.com/spark/spark-read-csv-file-into-dataframe/)

Spark – Read and Write JSON file (https://sparkbyexamples.com/

explode_outer – array example

```
    df.select($"name",explode_ou
      .show(false)
```

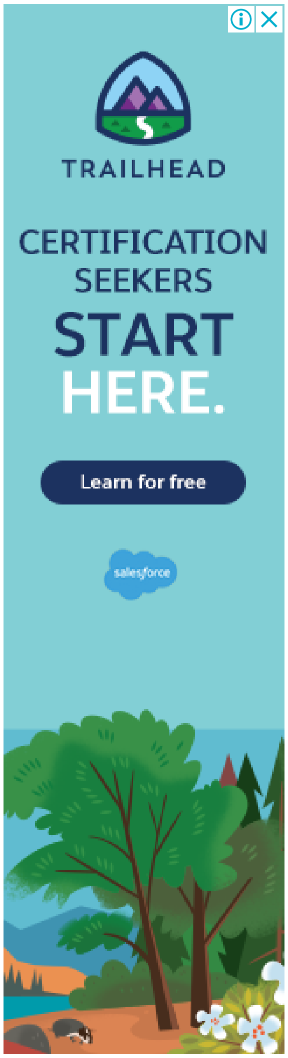
Outputs:

```
+----------+------+
|name      |col   |
+----------+------+
|James     |Java  |
|James     |Scala |
|Michael   |Spark |
|Michael   |Java  |
|Michael   |null  |
|Robert    |CSharp|
|Robert    |      |
|Washington|null  |
|Jeferson  |null  |
+----------+------+
```

explode_outer – map example

```
    df.select($"name",explode_ou
      .show(false)
```

Outputs:



spark/spark-read-and-write-json-file/)

Spark – Read & Write Parquet file (https://sparkbyexamples.com/spark/spark-read-write-dataframe-parquet-example/)

Spark – Read & Write XML file (https://sparkbyexamples.com/spark/spark-read-write-xml/)

Spark – Read & Write Avro files (https://sparkbyexamples.com/spark/read-write-avro-file-spark-dataframe/)

Spark – Read & Write Avro files (Spark version 2.3.x or earlier) (https://sparkbyexamples.com/spark/using-avro-data-files-from-spark-sql-2-3-x/)

Spark – Read & Write HBase using "hbase-spark" Connector (https://sparkbyexamples.com/spark/spark-read-write-using-hbase-spark-connector/)

Spark – Read & Write from HBase using Hortonworks (https://sparkbyexamples.com/spark/create-spark-dataframe-from-hbase-using-hortonworks/)

Spark – Read & Write ORC file (https://sparkbyexamples.com/spark/spark-read-orc-file-into-dataframe/)

Spark – Read Binary File (https://sparkbyexamples.com/spark/spark-read-binary-file-into-dataframe/)

Spark Streaming & Kafka

Spark Streaming – OutputModes (https://sparkbyexamples.com/spark/spark-streaming-outputmode/)

```
+----------+----+-----+
|name      |key |value|
+----------+----+-----+
|James     |hair|black|
|James     |eye |brown|
|Michael   |hair|brown|
|Michael   |eye |null |
|Robert    |hair|red  |
|Robert    |eye |     |
|Washington|null|null |
|Jeferson  |null|null |
+----------+----+-----+
```

posexplode – explode array or map elements to rows

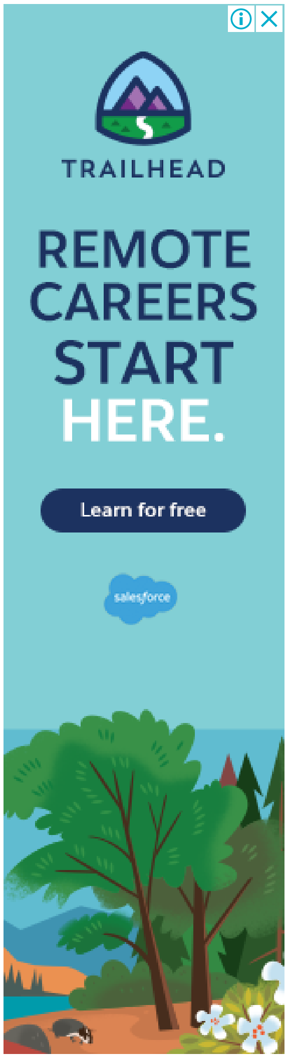
`posexplode(e: Column)` creates a row for each element in the array and creates two columns "pos' to hold the position of the array element and the 'col' to hold the actual array value. And when the input column is a map, posexplode function creates 3 columns "pos" to hold the position of the map element, "key" and "value" columns.

This will ignore elements that have null or empty. Since the Washington and Jefferson have null or empty values in array and map, the following snippet out does not contain these.

posexplode – array example

```
    df.select($"name",posexplode
      .show(false)
```

Outputs:



[Spark Streaming – Reading Files From Directory (https://sparkbyexamples.com/spark/spark-streaming-read-json-files-from-directory/)](https://sparkbyexamples.com/spark/spark-streaming-read-json-files-from-directory/)

[Spark Streaming – Reading Data From TCP Socket (https://sparkbyexamples.com/spark/spark-streaming-from-tcp-socket/)](https://sparkbyexamples.com/spark/spark-streaming-from-tcp-socket/)

[Spark Streaming – Processing Kafka Messages in JSON Format (https://sparkbyexamples.com/spark/spark-streaming-with-kafka/)](https://sparkbyexamples.com/spark/spark-streaming-with-kafka/)

[Spark Streaming – Processing Kafka messages in AVRO Format (https://sparkbyexamples.com/spark/spark-streaming-consume-and-produce-kafka-messages-in-avro-format/)](https://sparkbyexamples.com/spark/spark-streaming-consume-and-produce-kafka-messages-in-avro-format/)

[Spark SQL Batch – Consume & Produce Kafka Message (https://sparkbyexamples.com/spark/spark-batch-processing-produce-consume-kafka-topic/)](https://sparkbyexamples.com/spark/spark-batch-processing-produce-consume-kafka-topic/)

```
+-------+---+------+
|name   |pos|col   |
+-------+---+------+
|James  |0  |Java  |
|James  |1  |Scala |
|Michael|0  |Spark |
|Michael|1  |Java  |
|Michael|2  |null  |
|Robert |0  |CSharp|
|Robert |1  |      |
+-------+---+------+
```

posexplode – map example

```
    df.select($"name",posexplode
      .show(false)
```

Outputs:

```
+-------+---+----+-----+
|name   |pos|key |value|
+-------+---+----+-----+
|James  |0  |hair|black|
|James  |1  |eye |brown|
|Michael|0  |hair|brown|
|Michael|1  |eye |null |
|Robert |0  |hair|red  |
|Robert |1  |eye |     |
+-------+---+----+-----+
```

posexplode_outer – explode array or map columns to rows.

Spark `posexplode_outer(e: Column)` creates a row for each element in the array and creates two columns “pos’ to hold the position of the array element and the ‘col’ to hold the actual array value. Unlike posexplode, if the array or map is null or empty, posexplode_outer function returns null, null for pos and col columns. Similarly for the map, it returns rows with nulls.

posexplode_outer – array example

```
    df.select($"name",posexplode
      .show(false)
```

Outputs:

```
+----------+----+------+
|name      |pos |col   |
+----------+----+------+
|James     |0   |Java  |
|James     |1   |Scala |
|Michael   |0   |Spark |
|Michael   |1   |Java  |
|Michael   |2   |null  |
|Robert    |0   |CSharp|
|Robert    |1   |      |
|Washington|null|null  |
|Jeferson  |null|null  |
+----------+----+------+
```

posexplode_outer – map example

```
    df.select($"name",posexplode
      .show(false)
```

Outputs:

```
+----------+----+----+-----+
|name      |pos |key |value|
+----------+----+----+-----+
|James     |0   |hair|black|
|James     |1   |eye |brown|
|Michael   |0   |hair|brown|
|Michael   |1   |eye |null |
|Robert    |0   |hair|red  |
|Robert    |1   |eye |     |
|Washington|null|null|null |
|Jeferson  |null|null|null |
+----------+----+----+-----+
```

The complete example of exploding array or maps to rows

```
package com.sparkbyexamples.spar

import com.sparkbyexamples.spark
import org.apache.spark.sql.{Row
import org.apache.spark.sql.func
import org.apache.spark.sql.type

object ExplodeArrayAndMap{

  def main(args:Array[String]) :

    val spark = SparkSession.bui
      .master("local[1]")
      .getOrCreate()

     // create DataFrame

    val arrayData = Seq(
      Row("James",List("Java","S
      Row("Michael",List("Spark"
      Row("Robert",List("CSharp"
      Row("Washington",null,null
      Row("Jeferson",List(),Map(
    )

    val arraySchema = new Struct
      .add("name",StringType)
      .add("knownLanguages", Arr
      .add("properties", MapType

    val df = spark.createDataFra
    df.printSchema()
    df.show()

    import spark.implicits._
    // Below are Array examples
    //explode
    df.select($"name",explode($"
      .show()

    //explode_outer
    df.select($"name",explode_ou
      .show()

    //posexplode
    df.select($"name",posexplode
      .show()
    //posexplode_outer
    df.select($"name",posexplode
      .show()

    // Below are Map examples

    //explode
    df.select($"name",explode($"
      .show()
    //explode_outer
    df.select($"name",explode_ou
      .show()
    //posexplode
    df.select($"name",posexplode
      .show()

    //posexplode_outer
    df.select($"name",posexplode
      .show()
  }
}
```

Some common faq’s of explode functions

What is explode function
Spark SQL explode function is used to create or split an array or map DataFrame columns to rows. Spark defines several flavors of this function; explode_outer – to handle nulls and empty, posexplode – which explodes with a position of element and posexplode_outer – to handle nulls.

Difference between explode vs explode_outer
explode – creates a row for each element in the array or map column by ignoring null or empty values in array. whereas explode_outer returns all values in array or map including null or empty.

Difference between explode vs posexplode

explode – creates a row for each element in the array or map column. whereas posexplode creates a row for each element in the array and creates two columns 'pos' to hold the position of the array element and the 'col' to hold the actual array value. And, for the map, it creates 3 columns 'pos', 'key' and 'value'

Conclusion

In this article, you have learned how to how to explode or convert array or map DataFrame columns to rows using explode and posexplode SQL functions and their's respective outer functions and also learned differences between these functions.

Share this:

(https://sparkbyexamples.com/spark/explode-spark-array-and-map-dataframe-column/?share=facebook&nb=1)

(https://sparkbyexamples.com/spark/explode-spark-array-and-map-dataframe-column/?share=reddit&nb=1)

(https://sparkbyexamples.com/spark/explode-spark-array-and-map-dataframe-column/?share=pinterest&nb=1)

(https://sparkbyexamples.com/spark/explode-spark-array-and-map-dataframe-column/?share=tumblr&nb=1)

(https://sparkbyexamples.com/spark/explode-spark-array-and-map-dataframe-column/?share=pocket&nb=1)

(https://sparkbyexamples.com/spark/explode-spark-array-and-map-dataframe-column/?share=linkedin&nb=1)

(https://sparkbyexamples.com/spark/explode-spark-array-and-map-dataframe-column/?share=twitter&nb=1)

TAGS: EXPLODE (HTTPS://SPARKBYEXAMPLES.COM/TAG/EXPLODE/) , EXPLODE_OUTER (HTTPS://SPARKBYEXAMPLES.COM/TAG/EXPLODE_

OUTER/), POSEXPLODE (HTTPS://SPARKBYEXAMPLES.COM/TAG/POSEXPLODE/), POSEXPLODE_OUTER (HTTPS://SPARKBYEXAMPLES.COM/TAG/POSEXPLODE_OUTER/)

(https://sparkbyexamples.com/author/admin/)

NNK (Https://Sparkbyexamples.Com/Author/Admin/)

SparkByExamples.com is a Big Data and Spark examples community page, all examples are simple and easy to understand and well tested in our development environment Read more .. (https://sparkbyexamples.com/about-sparkbyexamples/)

Leave a Reply

Categories

Apache Hadoop (https://sparkbyexamples.com/category/hadoop/)

Apache Spark (https://sparkbyexamples.com/category/spark/)

Apache Spark Streaming (https://sparkbyexamples.com/category/spark/apache-spark-streaming/)

Apache Kafka (https://sparkbyexamples.com/category/kafka/)

Apache HBase (https://sparkbyexamples.com/category/hbase/)

Apache Cassandra (https://sparkbyexamples.com/category/cassandra/)

Snowflake Database (https://sparkbyexamples.com/category/snowflake/)

H2O Sparkling Water (https://sparkbyexamples.com/category/h2o-sparkling-water/)

PySpark (https://sparkbyexamples.com/category/pyspark/)

Recent Posts

Spark regexp_replace() – Replace String Value (https://sparkbyexamples.com/spark/spark-regexp_replace-replace-string-value/)

How to Run a PySpark Script from Python? (https://sparkbyexamples.com/pyspark/run-pyspark-script-from-python-subprocess/)

Spark SQL like() Using Wildcard Example (https://sparkbyexamples.com/spark/spark-sql-like-using-wildcard-example/)

Spark isin() & IS NOT IN Operator Example (https://sparkbyexamples.com/spark/spark-isin-is-not-in-operator-example/)

Spark – Get Size/Length of Array & Map Column (https://sparkbyexamples.com/spark/spark-get-size-length-of-array-map-column/)

Spark Using Length/Size Of a DataFrame Column (https://sparkbyexamples.com/spark/spark-using-length-size-of-a-dataframe-column/)

Spark rlike() Working with Regex Matching Examples (https://sparkbyexamples.com/spark/spark-rlike-regex-matching-examples/)

Spark Check String Column Has Numeric Values (https://sparkbyexamples.com/spark/spark-check-string-column-has-numeric-values/)

About SparkByExamples.Com

SparkByExamples.com is a Big Data and Spark examples community page, all examples are simple and easy to understand, and well tested in our development environment Read more .. (https://sparkbyexamples.com/about-sparkbyexamples/)

Follow Us

(https://twitter.com/sparkbyexamples)

(https://www.facebook.com/sparkbyexamples/)

(https://www.linkedin.com/in/n-nk-b860a8193/)

(https://github.com/spark-examples/)

Spark Check Column Data Type is Integer or String (https://sparkbyexamples.com/spark/spark-check-column-data-type-is-integer-or-string/)