

How to Rename Columns in PySpark: A Simple Guide

Authored by
stats writer

February 6, 2026

RECOMMENDED CITATION

stats writer (2026). *How to Rename Columns in PySpark: A Simple Guide*.

PSYCHOLOGICAL SCALES. Retrieved from <https://scales.arabpsychology.com/?p=129596>

PySpark is a powerful tool for data manipulation and transformation. One common task in data processing is to rename columns in a dataset. This can be easily achieved in PySpark using the "withColumnRenamed" function. This function takes two arguments - the current column name and the new column name. It then creates a new dataframe with the renamed column.

For example, if we have a dataframe with columns "id", "name", and "age", and we want to rename the "age" column to "years", we can use the following code:

```
df = df.withColumnRenamed("age", "years")
```

This will create a new dataframe with columns "id", "name", and "years". We can also rename multiple columns at once by chaining multiple "withColumnRenamed" functions.

In summary, renaming columns in PySpark is a simple and efficient process that can be achieved using the "withColumnRenamed" function. It allows for easy manipulation and organization of data in a dataframe.

Rename Columns in PySpark (With Examples)

You can use the following methods to rename columns in a PySpark DataFrame:

Method 1: Rename One Column

```
#rename 'conference' column to 'conf'
```

```
df = df.withColumnRenamed('conference', 'conf')
```

Method 2: Rename Multiple Columns

```
#rename 'conference' and 'team' columns
```

```
df = df.withColumnRenamed('conference', 'conf')
```

```
.withColumnRenamed('team', 'team_name')
```

Method 3: Rename All Columns

#specify new column names to use

col_names =

#rename all column names with new names

df = df.toDF(*col_names)

The following examples show how to use each of these methods in practice with the following PySpark DataFrame:

from pyspark.sql import SparkSession

spark = SparkSession.builder.getOrCreate()

#define data

data = ,

,

,

,

,

]

#define column names

columns =

```
#create dataframe using data and column names
```

```
df = spark.createDataFrame(data, columns)
```

```
#view dataframe
```

```
df.show()
```

```
+----+-----+-----+-----+  
|team|conference|points|assists|  
+----+-----+-----+-----+  
| A| East| 11| 4|  
| A| East| 8| 9|  
| A| East| 10| 3|  
| B| West| 6| 12|  
| B| West| 6| 4|  
| C| East| 5| 2|  
+----+-----+-----+-----+
```

Example 1: Rename One Column in PySpark

We can use the following syntax to rename just the conference column in the DataFrame:

```
#rename 'conference' column to 'conf'
```

```
df = df.withColumnRenamed('conference', 'conf')
```

```
#view updated DataFrame
```

df.show()

```
+----+----+-----+-----+
|team|conf|points|assists|
+----+----+-----+-----+
| A|East| 11| 4|
| A|East| 8| 9|
| A|East| 10| 3|
| B|West| 6| 12|
| B|West| 6| 4|
| C|East| 5| 2|
+----+----+-----+-----+
```

Notice that only the conference column has been renamed.

Example 2: Rename Multiple Columns in PySpark

We can use the following syntax to rename the conference and team columns in the DataFrame:

```
#rename 'conference' and 'team' columns
df = df.withColumnRenamed('conference', 'conf')
.withColumnRenamed('team', 'team_name')
```

#view updated DataFrame

df.show()

```
+-----+----+-----+-----+
|team_name|conf|points|assists|
+-----+----+-----+-----+
| A|East| 11| 4|
| A|East| 8| 9|
| A|East| 10| 3|
| B|West| 6| 12|
| B|West| 6| 4|
| C|East| 5| 2|
+-----+----+-----+-----+
```

Notice that the conference and team columns have been renamed while all other column names have remained the same.

Example 3: Rename All Columns in PySpark

We can use the following syntax to rename all columns in the DataFrame:

```
#specify new column names to use
```

```
col_names =
```

```
#rename all column names with new names
```

```
df = df.toDF(*col_names)
```

```
#view updated DataFrame
```

```
df.show()
```

```
+-----+-----+-----+-----+
|the_team|the_conf|points_scored|total_assists|
+-----+-----+-----+-----+
| A| East| 11| 4|
| A| East| 8| 9|
| A| East| 10| 3|
| B| West| 6| 12|
| B| West| 6| 4|
| C| East| 5| 2|
+-----+-----+-----+-----+
```

Notice that all of the column names have been renamed based on the new names that we specified.

The following tutorials explain how to perform other common tasks in PySpark: