

# How to Count Values in a PySpark Column Based on a Condition

Authored by  
**stats writer**

January 20, 2026

## RECOMMENDED CITATION

stats writer (2026). *How to Count Values in a PySpark Column Based on a Condition*. PSYCHOLOGICAL SCALES. Retrieved from <https://scales.arabpsychology.com/?p=126726>

You can use the following methods to count the number of values in a column of a PySpark DataFrame that meet a specific condition:

### Method 1: Count Values that Meet One Condition

```
#count values in 'team' column that are equal to 'C'  
df.filter(df.team == 'C').count()
```

### Method 2: Count Values that Meet One of Several Conditions

```
from pyspark.sql.functions import col  
  
#count values in 'team' column that are equal to 'A' or 'D'  
df.filter(col('team').isin()).count()
```

The following examples show how to use each method in practice with the following PySpark DataFrame that contains information about various basketball players:

```
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|
+---+-----+-----+
| A| East| 11|
| A| East| 8|
| A| East| 10|
| B| West| 6|
| B| West| 6|
| C| East| 5|
| C| East| 15|
| C| West| 31|
| D| West| 24|
+---+-----+-----+
```

### Example 1: Count Values that Meet One Condition

We can use the following syntax to count the number of values in the **team** column that are equal to C:

```
#count values in 'team' column that are equal to 'C'
df.filter(df.team == 'C').count()
```

3

We can see that a total of **3** values in the **team** column are equal to C.

### Example 2: Count Values that Meet One of Several Conditions

We can use the following syntax to count the number of values in the **team** column that are equal to either A or D:

```
from pyspark.sql.functions import col
```

```
#count values in 'team' column that are equal to 'A' or 'D'
df.filter(col('team').isin()).count()
```

4

We can see that a total of **4** values in the **team** column are equal to either A or D.

**Note:** You can find the complete documentation for the PySpark **filter** function .

ARABPSYCHOLOGY.COM