1. Pyspark filter() function is used to filter the rows from DataFrame based on the given condition or SQL expression.
2. We can use where() also instead of filter().

from pyspark.sql import \*  
from pyspark.sql.functions import \*  
  
spark = SparkSession.builder.appName('alias').getOrCreate()  
  
myData = [(1,'Harsha','male',2000),  
 (2,'Mokshit','male',4000),  
 (3,'Harika','female',3000)]  
  
mySchema = ['id','name','gender','salary']  
  
df = spark.createDataFrame(myData,mySchema)  
  
*# single condition with filter*df.filter(df.gender=='male').show()  
  
*# single condition with where*df.where(df.salary > 3000).show()  
  
*# multiple condition's with filter*df.filter((df.gender=='male') & (df.salary >= 3000)).show()

**Output:**

+---+-------+------+------+

| id| name|gender|salary|

+---+-------+------+------+

| 1| Harsha| male| 2000|

| 2|Mokshit| male| 4000|

+---+-------+------+------+

+---+-------+------+------+

| id| name|gender|salary|

+---+-------+------+------+

| 2|Mokshit| male| 4000|

+---+-------+------+------+

+---+-------+------+------+

| id| name|gender|salary|

+---+-------+------+------+

| 2|Mokshit| male| 4000|

+---+-------+------+------+