

```
In [1]: from pyspark.sql import SparkSession

if __name__ == '__main__':
    #Create the spark session
    spark = SparkSession.builder.appName("filter transformation").getOrCreate()

    #Dataset
    data = spark.sparkContext.range(1,5)

    #Show the dataset
    print('Dataset')
    print(data.collect())
    print('-----')

    #Use the map function
    rdd = data.map(lambda x: (x, x*x, x*x*x))

    #Show the new dataset after the map function
    print('New Dataset')
    print(rdd.collect())

22/10/10 17:18:08 WARN Utils: Your hostname, Vaishalis-MacBook-Pro.local resolves to a loopback address: 127.0.0.1; using 192.168.0.105 instead (on interface en0)
22/10/10 17:18:08 WARN Utils: Set SPARK_LOCAL_IP if you need to bind to another address
Setting default log level to "WARN".
To adjust logging level use sc.setLogLevel(newLevel). For SparkR, use setLogLevel(newLevel).
22/10/10 17:18:09 WARN NativeCodeLoader: Unable to load native-hadoop library for your platform... using builtin-java classes where applicable
22/10/10 17:18:10 WARN Utils: Service 'SparkUI' could not bind on port 4040. Attempting port 4041.
22/10/10 17:18:10 WARN Utils: Service 'SparkUI' could not bind on port 4041. Attempting port 4042.
Dataset
[1, 2, 3, 4]
-----
New Dataset
[(1, 1, 1), (2, 4, 8), (3, 9, 27), (4, 16, 64)]
```

```
In [2]: columns = ["number","squared","cubed"]

#Create DataFrame
df = spark.createDataFrame(data = rdd, schema = columns)

#show() displays the contents of the DataFrame in a Table Row and Column Format
df.show()
```

number	squared	cubed
1	1	1
2	4	8
3	9	27
4	16	64

```
In [3]: #Applying Filter Transformation to result in a new DataFrame when column 1 is not 3

df1 = df.filter(df.number != 3).show(truncate = False)
```

number	squared	cubed
1	1	1
2	4	8
4	16	64

```
In [4]: # In[4]:

#Applying Filter Transformation to result in a new DataFrame when column 1 is not 2 and when column 3 is greater than 3
rdd3 = rdd.filter(lambda x: (x[2] > 5) & (x[0] != 2))
rdd3.toDF(["number","squared","cubed"]).show()
```

number	squared	cubed
3	9	27
4	16	64