



### Problem Statement 01 :

Create and RDD[String] as shown below

C Programming	C++	Java	Scala	Python
---------------	-----	------	-------	--------

1. Print the elements which are greater than length 5

Expected output :

```
C Programming
Python
```

2. Print the elements in reverse order.

Expected output

```
Python
C Programming
```

```
!pip install pyspark py4j

Collecting pyspark
  Downloading pyspark-3.4.1.tar.gz (310.8 MB)
    Preparing metadata (setup.py) ... done
Requirement already satisfied: py4j in /usr/local/lib/python3.10/dist-packages (0.10.9.7)
Building wheels for collected packages: pyspark
  Building wheel for pyspark (setup.py) ... done
  Created wheel for pyspark: filename=pyspark-3.4.1-py2.py3-none-any.whl size=311285388 sha256=8c76b3253f468d029c20d6d39518c167a9078a873dc0ca473090706eeb0955
  Stored in directory: /root/.cache/pip/wheels/0d/77/a3/ff2f74cc9ab41f8f594dabf0579c2a7c6de920d584206e0834
Successfully built pyspark
Installing collected packages: pyspark
Successfully installed pyspark-3.4.1
```

```
from pyspark import SparkContext, SparkConf

# Spark configuration
conf = SparkConf().setAppName("RDDExample")
sc = SparkContext(conf=conf)

# Create an RDD from a list of strings
languages_rdd = sc.parallelize(["C Programming", "C++", "Java", "Scala", "Python"])

# Elements with length greater than 5
filtered_languages_rdd = languages_rdd.filter(lambda lang: len(lang) > 5)

# Collect and print the filtered elements
filtered_languages = filtered_languages_rdd.collect()
for lang in filtered_languages:
    print(lang)

# Stop the Spark context
sc.stop()
```

C Programming  
Python

```
RDD.ipynb
File Edit View Insert Runtime Tools Help All changes saved
+ Code + Text
[19] from pyspark import SparkContext, SparkConf

# Spark configuration
conf = SparkConf().setAppName("RDEExample")
sc = SparkContext(conf=conf)

# Create an RDD from a list of strings
languages_rdd = sc.parallelize(["C Programming", "C++", "Java", "Scala", "Python"])

# Elements with length greater than 5
filtered_languages_rdd = languages_rdd.filter(lambda lang: len(lang) > 5)

# Collect the filtered elements and reverse the list
filtered_languages = filtered_languages_rdd.collect()[::-1]

# Print the elements in reverse order
for lang in filtered_languages:
    print(lang)

# Stop the Spark context
sc.stop()

Python
C Programming
```



## Problem Statement 02 :

Generate an RDD[Int] which is multiples of 1024 lets say the numbers are in kilobytes and we want to convert each element to its respective MB and perform the following tasks.

1. Convert each element of RDD to its respective MB.
2. Filter elements which are greater than 10 MB
3. Apply transformation to convert MB to GB

```
RDD.ipynb
File Edit View Insert Runtime Tools Help All changes saved
+ Code + Text
from pyspark import SparkConf, SparkContext

# Spark configuration and context
conf = SparkConf().setAppName("RDDTransformations").setMaster("local")
sc = SparkContext(conf=conf)

# RDD of multiples of 1024 (kilobytes)
rdd = sc.parallelize(range(1, 10001)).map(lambda x: x * 1024)

# 1: Convert each element of RDD to its respective MB
mb_rdd = rdd.map(lambda x: x / 1024.0)

# 2: Filter elements which are greater than 10 MB
filtered_rdd = mb_rdd.filter(lambda x: x > 10.0)

# 3: Apply transformation to convert MB to GB
gb_rdd = filtered_rdd.map(lambda x: x / 1024.0)

# Print the results
print("Original RDD:")
print(rdd.collect())

print("Elements in MB greater than 10:")
print(filtered_rdd.collect())

print("Elements in GB:")
print(gb_rdd.collect())

# Stop the Spark context
sc.stop()
```

```
RDD.ipynb
File Edit View Insert Runtime Tools Help All changes saved
+ Code + Text
[20] print(rdd.collect())

print("Elements in MB greater than 10:")
print(filtered_rdd.collect())

print("Elements in GB:")
print(gb_rdd.collect())

# Stop the Spark context
sc.stop()

Original RDD:
[1024, 2048, 3072, 4096, 5120, 6144, 7168, 8192, 9216, 10240, 11264, 12288, 13312, 14336, 15360, 16384, 17408, 18432, 19456, 20480, 21504, 22528, 23552, 24576, 25600, 26624, 27648, 28672, 29696, 30720, 31744]
Elements in MB greater than 10:
[11.0, 12.0, 13.0, 14.0, 15.0, 16.0, 17.0, 18.0, 19.0, 20.0, 21.0, 22.0, 23.0, 24.0, 25.0, 26.0, 27.0, 28.0, 29.0, 30.0, 31.0, 32.0, 33.0, 34.0, 35.0, 36.0, 37.0, 38.0, 39.0, 40.0, 41.0, 42.0, 43.0, 44.0, 45.0, 46.0, 47.0, 48.0, 49.0, 50.0, 51.0, 52.0, 53.0, 54.0, 55.0, 56.0, 57.0, 58.0, 59.0, 60.0, 61.0, 62.0, 63.0, 64.0, 65.0, 66.0, 67.0, 68.0, 69.0, 70.0, 71.0, 72.0, 73.0, 74.0, 75.0, 76.0, 77.0, 78.0, 79.0, 80.0, 81.0, 82.0, 83.0, 84.0, 85.0, 86.0, 87.0, 88.0, 89.0, 90.0, 91.0, 92.0, 93.0, 94.0, 95.0, 96.0, 97.0, 98.0, 99.0]
Elements in GB:
[0.0107421875, 0.011171875, 0.0116015625, 0.01203125, 0.0124609375, 0.012890625, 0.0133203125, 0.01375, 0.0141796875, 0.014609375, 0.0150390625, 0.01546875, 0.0158984375, 0.016328125, 0.0167578125, 0.0171875, 0.0176171875, 0.018046875, 0.0184765625, 0.01890625, 0.0193359375, 0.019765625, 0.0201953125, 0.020625, 0.0210546875, 0.021484375, 0.0219140625, 0.02234375, 0.0227734375, 0.023203125, 0.0236328125, 0.0240625, 0.0244921875, 0.024921875, 0.0253515625, 0.02578125, 0.0262109375, 0.026640625, 0.0270703125, 0.0275, 0.0279296875, 0.028359375, 0.0287890625, 0.02921875, 0.0296484375, 0.030078125, 0.0305078125, 0.0309375, 0.0313671875, 0.031796875, 0.0322265625, 0.03265625, 0.0330859375, 0.033515625, 0.0339453125, 0.034375, 0.0348046875, 0.035234375, 0.0356640625, 0.03609375, 0.0365234375, 0.036953125, 0.0373828125, 0.0378125, 0.0382421875, 0.038671875, 0.0391015625, 0.03953125, 0.0399609375, 0.040390625, 0.0408203125, 0.04125, 0.0416796875, 0.042109375, 0.0425390625, 0.04296875, 0.0433984375, 0.043828125, 0.0442578125, 0.0446875, 0.0451171875, 0.045546875, 0.0459765625, 0.04640625, 0.0468359375, 0.047265625, 0.0476953125, 0.048125, 0.0485546875, 0.048984375, 0.0494140625, 0.04984375, 0.0502734375, 0.050703125, 0.0511328125, 0.0515625, 0.0519921875, 0.052421875, 0.0528515625, 0.05328125, 0.0537109375, 0.054140625, 0.0545703125, 0.055, 0.0554296875, 0.055859375, 0.0562890625, 0.05671875, 0.0571484375, 0.057578125, 0.0580078125, 0.0584375, 0.0588671875, 0.059296875, 0.0597265625, 0.06015625, 0.0605859375, 0.061015625, 0.0614453125, 0.061875, 0.0623046875, 0.062734375, 0.0631640625, 0.06359375, 0.0640234375, 0.064453125, 0.0648828125, 0.0653125, 0.0657421875, 0.066171875, 0.0666015625, 0.06703125, 0.0674609375, 0.067890625, 0.0683203125, 0.06875, 0.0691796875, 0.069609375, 0.0700390625, 0.07046875, 0.0708984375, 0.071328125, 0.0717578125, 0.0721875, 0.0726171875, 0.073046875, 0.0734765625, 0.07390625, 0.0743359375, 0.074765625, 0.0751953125, 0.075625, 0.0760546875, 0.076484375, 0.0769140625, 0.07734375, 0.0777734375, 0.078203125, 0.0786328125, 0.0790625, 0.0794921875, 0.079921875, 0.0803515625, 0.08078125, 0.0812109375, 0.081640625, 0.0820703125, 0.0825, 0.0829296875, 0.083359375, 0.0837890625, 0.08421875, 0.0846484375, 0.085078125, 0.0855078125, 0.0859375, 0.0863671875, 0.086796875, 0.0872265625, 0.08765625, 0.0880859375, 0.088515625, 0.0889453125, 0.089375, 0.0898046875, 0.090234375, 0.0906640625, 0.09109375, 0.0915234375, 0.091953125, 0.0923828125, 0.0928125, 0.0932421875, 0.093671875, 0.0941015625, 0.09453125, 0.0949609375, 0.095390625, 0.0958203125, 0.09625, 0.0966796875, 0.097109375, 0.0975390625, 0.09796875, 0.0983984375, 0.098828125, 0.0992578125, 0.0996875, 1.0, 1.00296875, 1.0059375, 1.00890625, 1.011875, 1.01484375, 1.0178125, 1.02078125, 1.02375, 1.02671875, 1.0296875, 1.03265625, 1.035625, 1.03859375, 1.0415625, 1.04453125, 1.0475, 1.05046875, 1.0534375, 1.05640625, 1.059375, 1.06234375, 1.0653125, 1.06828125, 1.07125, 1.07421875, 1.0771875, 1.08015625, 1.083125, 1.08609375, 1.0890625, 1.09203125, 1.095, 1.09796875, 1.1009375, 1.10390625, 1.106875, 1.10984375, 1.1128125, 1.11578125, 1.11875, 1.12171875, 1.1246875, 1.12765625, 1.130625, 1.13359375, 1.1365625, 1.13953125, 1.1425, 1.14546875, 1.1484375, 1.15140625, 1.154375, 1.15734375, 1.1603125, 1.16328125, 1.16625, 1.16921875, 1.1721875, 1.17515625, 1.178125, 1.18109375, 1.1840625, 1.18703125, 1.19, 1.19296875, 1.1959375, 1.19890625, 1.201875, 1.20484375, 1.2078125, 1.21078125, 1.21375, 1.21671875, 1.2196875, 1.22265625, 1.225625, 1.22859375, 1.2315625, 1.23453125, 1.2375, 1.24046875, 1.2434375, 1.24640625, 1.249375, 1.25234375, 1.2553125, 1.25828125, 1.26125, 1.26421875, 1.2671875, 1.27015625, 1.273125, 1.27609375, 1.2790625, 1.28203125, 1.285, 1.28796875, 1.2909375, 1.29390625, 1.296875, 1.29984375, 1.3028125, 1.30578125, 1.30875, 1.31171875, 1.3146875, 1.31765625, 1.320625, 1.32359375, 1.3265625, 1.32953125, 1.3325, 1.33546875, 1.3384375, 1.34140625, 1.344375, 1.34734375, 1.3503125, 1.35328125, 1.35625, 1.35921875, 1.3621875, 1.36515625, 1.368125, 1.37109375, 1.3740625, 1.37703125, 1.38, 1.38296875, 1.3859375, 1.38890625, 1.391875, 1.39484375, 1.3978125, 1.40078125, 1.40375, 1.40671875, 1.4096875, 1.41265625, 1.415625, 1.41859375, 1.4215625, 1.42453125, 1.4275, 1.43046875, 1.4334375, 1.43640625, 1.439375, 1.44234375, 1.4453125, 1.44828125, 1.45125, 1.45421875, 1.4571875, 1.46015625, 1.463125, 1.46609375, 1.4690625, 1.47203125, 1.475, 1.47796875, 1.4809375, 1.48390625, 1.486875, 1.48984375, 1.4928125, 1.49578125, 1.49875, 1.50171875, 1.5046875, 1.50765625, 1.510625, 1.51359375, 1.5165625, 1.51953125, 1.5225, 1.52546875, 1.5284375, 1.53140625, 1.534375, 1.53734375, 1.5403125, 1.54328125, 1.54625, 1.54921875, 1.5521875, 1.55515625, 1.558125, 1.56109375, 1.5640625, 1.56703125, 1.57, 1.57296875, 1.5759375, 1.57890625, 1.581875, 1.58484375, 1.5878125, 1.59078125, 1.59375, 1.59671875, 1.5996875, 1.60265625, 1.605625, 1.60859375, 1.6115625, 1.61453125, 1.6175, 1.62046875, 1.6234375, 1.62640625, 1.629375, 1.63234375, 1.6353125, 1.63828125, 1.64125, 1.64421875, 1.6471875, 1.65015625, 1.653125, 1.65609375, 1.6590625, 1.66203125, 1.665, 1.66796875, 1.6709375, 1.67390625, 1.676875, 1.67984375, 1.6828125, 1.68578125, 1.68875, 1.69171875, 1.6946875, 1.69765625, 1.700625, 1.70359375, 1.7065625, 1.70953125, 1.7125, 1.71546875, 1.7184375, 1.72140625, 1.724375, 1.72734375, 1.7303125, 1.73328125, 1.73625, 1.73921875, 1.7421875, 1.74515625, 1.748125, 1.75109375, 1.7540625, 1.75703125, 1.76, 1.76296875, 1.7659375, 1.76890625, 1.771875, 1.77484375, 1.7778125, 1.78078125, 1.78375, 1.78671875, 1.7896875, 1.79265625, 1.795625, 1.79859375, 1.8015625, 1.80453125, 1.8075, 1.81046875, 1.8134375, 1.81640625, 1.819375, 1.82234375, 1.8253125, 1.82828125, 1.83125, 1.83421875, 1.8371875, 1.84015625, 1.843125, 1.84609375, 1.8490625, 1.85203125, 1.855, 1.85796875, 1.8609375, 1.86390625, 1.866875, 1.86984375, 1.8728125, 1.87578125, 1.87875, 1.88171875, 1.8846875, 1.88765625, 1.890625, 1.89359375, 1.8965625, 1.89953125, 1.9025, 1.90546875, 1.9084375, 1.91140625, 1.914375, 1.91734375, 1.9203125, 1.92328125, 1.92625, 1.92921875, 1.9321875, 1.93515625, 1.938125, 1.94109375, 1.9440625, 1.94703125, 1.95, 1.95296875, 1.9559375, 1.95890625, 1.961875, 1.96484375, 1.9678125, 1.97078125, 1.97375, 1.97671875, 1.9796875, 1.98265625, 1.985625, 1.98859375, 1.9915625, 1.99453125, 1.9975, 2.0, 2.00296875, 2.0059375, 2.00890625, 2.011875, 2.01484375, 2.0178125, 2.02078125, 2.02375, 2.02671875, 2.0296875, 2.03265625, 2.035625, 2.03859375, 2.0415625, 2.04453125, 2.0475, 2.05046875, 2.0534375, 2.05640625, 2.059375, 2.06234375, 2.0653125, 2.06828125, 2.07125, 2.07421875, 2.0771875, 2.08015625, 2.083125, 2.08609375, 2.0890625, 2.09203125, 2.095, 2.09796875, 2.1009375, 2.10390625, 2.106875, 2.10984375, 2.1128125, 2.11578125, 2.11875, 2.12171875, 2.1246875, 2.12765625, 2.130625, 2.13359375, 2.1365625, 2.13953125, 2.1425, 2.14546875, 2.1484375, 2.15140625, 2.154375, 2.15734375, 2.1603125, 2.16328125, 2.16625, 2.16921875, 2.1721875, 2.17515625, 2.178125, 2.18109375, 2.1840625, 2.18703125, 2.19, 2.19296875, 2.1959375, 2.19890625, 2.201875, 2.20484375, 2.2078125, 2.21078125, 2.21375, 2.21671875, 2.2196875, 2.22265625, 2.225625, 2.22859375, 2.2315625, 2.23453125, 2.2375, 2.24046875, 2.2434375, 2.24640625, 2.249375, 2.25234375, 2.2553125, 2.25828125, 2.26125, 2.26421875, 2.2671875, 2.27015625, 2.273125, 2.27609375, 2.2790625, 2.28203125, 2.285, 2.28796875, 2.2909375, 2.29390625, 2.296875, 2.29984375, 2.3028125, 2.30578125, 2.30875, 2.31171875, 2.3146875, 2.31765625, 2.320625, 2.32359375, 2.3265625, 2.32953125, 2.3325, 2.33546875, 2.3384375, 2.34140625, 2.344375, 2.34734375, 2.3503125, 2.35328125, 2.35625, 2.35921875, 2.3621875, 2.36515625, 2.368125, 2.37109375, 2.3740625, 2.37703125, 2.38, 2.38296875, 2.3859375, 2.38890625, 2.391875, 2.39484375, 2.3978125, 2.40078125, 2.40375, 2.40671875, 2.4096875, 2.41265625, 2.415625, 2.41859375, 2.4215625, 2.42453125, 2.4275, 2.43046875, 2.4334375, 2.43640625, 2.439375, 2.44234375, 2.4453125, 2.44828125, 2.45125, 2.45421875, 2.4571875, 2.46015625, 2.463125, 2.46609375, 2.4690625, 2.47203125, 2.475, 2.47796875, 2.4809375, 2.48390625, 2.486875, 2.48984375, 2.4928125, 2.49578125, 2.49875, 2.50171875, 2.5046875, 2.50765625, 2.510625, 2.51359375, 2.5165625, 2.51953125, 2.5225, 2.52546875, 2.5284375, 2.53140625, 2.534375, 2.53734375, 2.5403125, 2.54328125, 2.54625, 2.54921875, 2.5521875, 2.55515625, 2.558125, 2.56109375, 2.5640625, 2.56703125, 2.57, 2.57296875, 2.5759375, 2.57890625, 2.581875, 2.58484375, 2.5878125, 2.59078125, 2.59375, 2.59671875, 2.5996875, 2.60265625, 2.605625, 2.60859375, 2.6115625, 2.61453125, 2.6175, 2.62046875, 2.6234375, 2.62640625, 2.629375, 2.63234375, 2.6353125, 2.63828125, 2.64125, 2.64421875, 2.6471875, 2.65015625, 2.653125, 2.65609375, 2.6590625, 2.66203125, 2.665, 2.66796875, 2.6709375, 2.67390625, 2.676875, 2.67984375, 2.6828125, 2.68578125, 2.68875, 2.69171875, 2.6946875, 2.69765625, 2.700625, 2.70359375, 2.7065625, 2.70953125, 2.7125, 2.71546875, 2.7184375, 2.72140625, 2.724375, 2.72734375, 2.7303125, 2.73328125, 2.73625, 2.73921875, 2.7421875, 2.74515625, 2.748125, 2.75109375, 2.7540625, 2.75703125, 2.76, 2.76296875, 2.7659375, 2.76890625, 2.771875, 2.77484375, 2.7778125, 2.78078125, 2.78375, 2.78671875, 2.7896875, 2.79265625, 2.795625, 2.79859375, 2.8015625, 2.80453125, 2.8075, 2.81046875, 2.8134375, 2.81640625, 2.819375, 2.82234375, 2.8253125, 2.82828125, 2.83125, 2.83421875, 2.8371875, 2.84015625, 2.843125, 2.84609375, 2.8490625, 2.85203125, 2.855, 2.85796875, 2.8609375, 2.86390625, 2.866875, 2.86984375, 2.8728125, 2.87578125, 2.87875, 2.88171875, 2.8846875, 2.88765625, 2.890625, 2.89359375, 2.8965625, 2.89953125, 2.9025, 2.90546875, 2.9084375, 2.91140625, 2.914375, 2.91734375, 2.9203125, 2.92328125, 2.92625, 2.92921875, 2.9321875, 2.93515625, 2.938125, 2.94109375, 2.9440625, 2.94703125, 2.95, 2.95296875, 2.9559375, 2.95890625, 2.961875, 2.96484375, 2.9678125, 2.97078125, 2.97375, 2.97671875, 2.9796875, 2.98265625, 2.985625, 2.98859375, 2.9915625, 2.99453125, 2.9975, 3.0, 3.00296875, 3.0059375, 3.00890625, 3.011875, 3.01484375, 3.0178125, 3.02078125, 3.02375, 3.02671875, 3.0296875, 3.03265625, 3.035625, 3.03859375, 3.0415625, 3.04453125, 3.0475, 3.05046875, 3.0534375, 3.05640625, 3.059375, 3.06234375, 3.0653125, 3.06828125, 3.07125, 3.07421875, 3.0771875, 3.08015625, 3.083125, 3.08609375, 3.0890625, 3.09203125, 3.095, 3.09796875, 3.1009375, 3.10390625, 3.106875, 3.10984375, 3.1128125, 3.11578125, 3.11875, 3.12171875, 3.1246875, 3.12765625, 3.130625, 3.13359375, 3.1365625, 3.13953125, 3.1425, 3.14546875, 3.1484375, 3.15140625, 3.154375, 3.15734375, 3.1603125, 3.16328125, 3.16625, 3.16921
```