

```
from pyspark import SparkContext
```

```
# Create a SparkContext
sc = SparkContext("local", "RDD Exploration")
```

⊕ **ValueError:** Cannot run multiple SparkContexts at once; existing SparkContext(app=Databricks Shell, master=local[8]) created by \_\_init\_\_ at /databricks/python\_shell/dbruntime/spark\_connection.py:127

```
# Create an RDD from a list
data = [1, 2, 3, 4, 5]
rdd = sc.parallelize(data)
```

```
# Perform some basic operations on the RDD
# 1. Count the number of elements
count = rdd.count()
print("Number of elements:", count)
```

Number of elements: 5

```
# 2. Sum all elements
total_sum = rdd.sum()
print("Sum of all elements:", total_sum)
```

Sum of all elements: 15

```
# 3. Calculate the mean
mean = total_sum / count
print("Mean of elements:", mean)
```

Mean of elements: 3.0