

**BDT WEEK 8**

**NAME:**  **DATE:** 20/11/2024

**ROLL NUMBER:** **BRANCH:** CSD-B

**PROBLEM STATEMENT :**

1)Writing and running Spark program

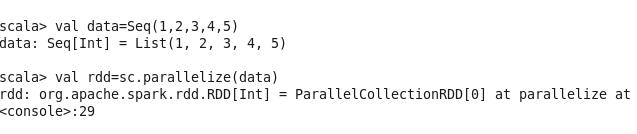
**CODE:**

**1)Opening spark:** open the terminal and type ‘spark-shell’

**2)Basic spark command:**



**3)Creating RDDs:**



**4)Actions on RDDs:**

**Collect the elements of an RDD:**



**Count the elements in an RDD:**

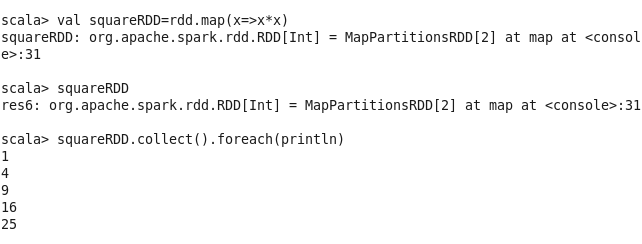


**Take the first n elements from an RDD:**

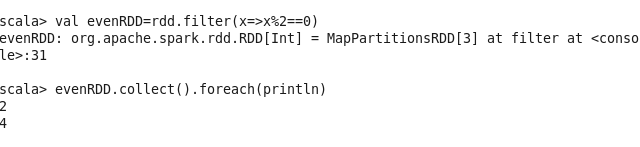


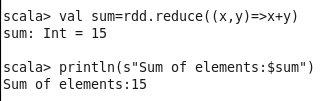
**5)Transformations on RDDs:**

**Map Transformation (Square each element):**

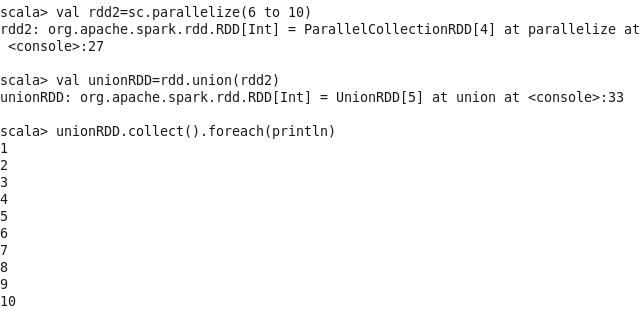


**Filter Transformation (Filter even numbers):**

**Reduce Transformation (Sum all elements):**



**Union Transformation (Combine two RDDs):**



**Distinct Transformation (Remove duplicate elements):**

