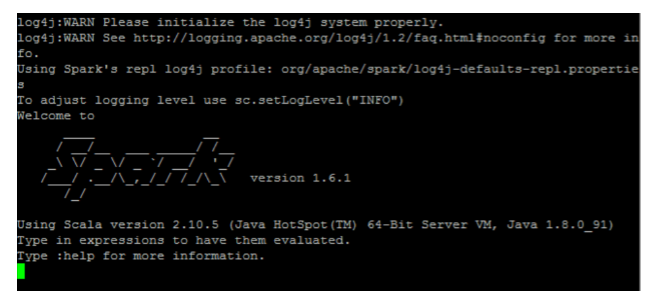
**ASSESSMENT-18 Vinutha S**

**23/12/2023**

**1. To start the Spark shell**

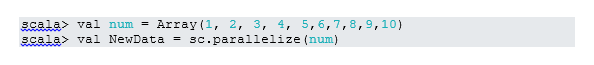
Spark shell



**2. Read a file from the local system:**

local system

**3. Create RDD through parallelizing**

**4. Count Items in RDD**RDD 2

**6. Read the first 3 Items from RDD**

RDD 3

**7. Save output/processed data into the text file**

RDD 4

**Intermediate Spark Commands**

**1. Filter on RDD**

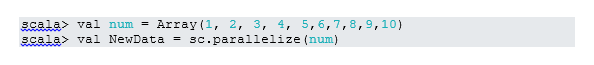
Let’s create a new RDD for items that contain “yes.”

Filter on RDD

**2. Chain Operation**

Chain Operation

**3. Create RDD through parallelizing**



.

**6. Read the first 3 Items from RDD**

RDD 3

**7. Save output/processed data into the text file**

RDD 4