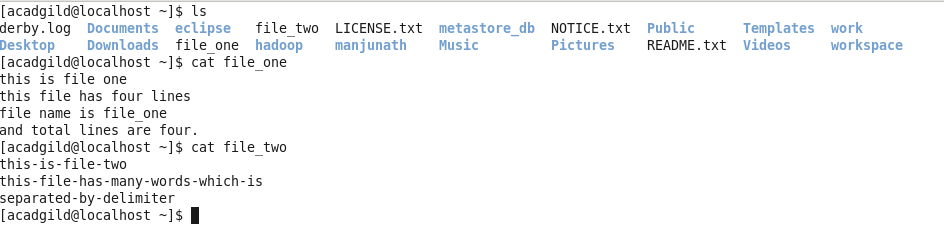
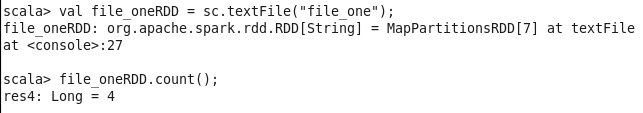
**Assignment 17.1:**

Create 2 text files and add data in it.



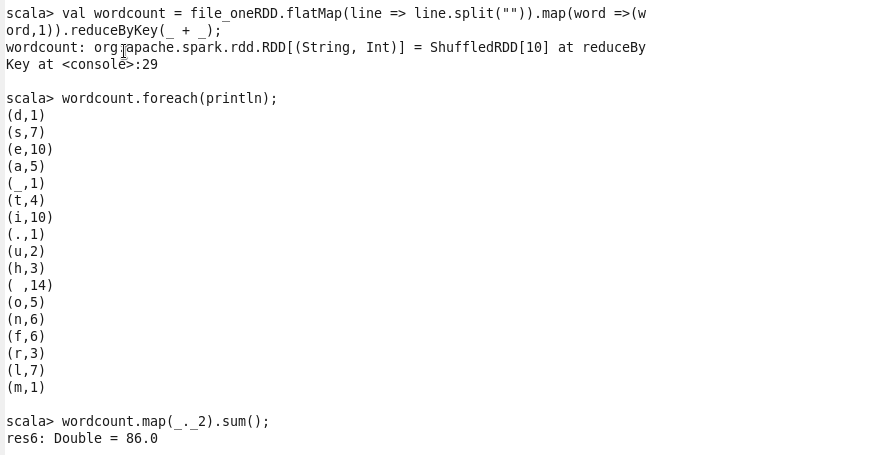
1. **Write a program to read a text file and print the number of rows of data in the document.**

* Load the file1 to RDD using spark context.
* Count the number of words present in that file.



1. **Write a program to read a text file and print the number of words in the document.**

* As RDD is already created we are using the same rdd for file\_one.txt to read and print the no of words in the document.
* Split the line with space and count the each words per lines.



1. **We have a document where the word separator is -, instead of space. Write a spark code, to obtain the count of the total number of words present in the document.**

* Load the second file which had word separator – to RDD.
* Split the line by space and count the words in the file with -



So file 2 has 14 words.