

Plagiarism Scan Report

Summary

Report Generated Date	23 Jul, 2018
Plagiarism Status	100% Unique
Total Words	184
Total Characters	1603
Any Ignore Url Used	

Content Checked For Plagiarism:

package bigdatasummertraining

```
import org.apache.spark._
import org.apache.spark.SparkContext._
import org.apache.log4j._
```

```
/** Count up how many of each star rating exists in the MovieLens 100K data set. */
object RatingsCounter {
  /** Our main function where the action happens */
  def main(args: Array[String]) {
    // Set the log level to only print errors
    Logger.getLogger("org").setLevel(Level.ERROR)
    // Create a SparkContext using every core of the local machine, named RatingsCounter
    val sc = new SparkContext("local[*]", "RatingsCounter")
    // Loading the dataset of 100k
    val lines = sc.textFile("../MovieRatings/ml-100k/u.data")
    // Convert each line to a string, split it out by tabs, and extract the third field.
    // (The file format is lines.map(x => x.toString().split("\t")(2))
    // Count up how many times each value (rating) occurs
    val results = ratings.countByValue()
    // Sort the resulting map of (rating, count) tuples
    val sortedResults = results.toSeq.sortBy(_._1)
    println("\t\t*****\n")
    println("\t\t Welcome to 100K Movies Rating BigData Problem \n")
    println("\t\t*****\n\n")
    println("\t\t Solved By: Nikhil Kumar(11606029) K1642 \n")
    println("\t\t*****\n\n")
    // Print each result on its own line.
    // Print each result on its own line.
    sortedResults.sorted.foreach(x => println("\t\t-----
\n"+"\t\tMovie Star(*) = " + x._1+" And Number of Movies = "+x._2))
    println("\t\t-----\n")
  }
}
```