**Mock Test 2 - Big Data Question 3**

**Question 3: Write a Spark program to count the number of occurrences of each word in a given text file**

**Ans:**

from pyspark.sql import SparkSession

from pyspark.sql.functions import explode, split

spark = SparkSession.builder.appName("WordCount").getOrCreate()

# Read the text file into a DataFrame

lines = spark.read.text("<input\_file.txt>").rdd.map(lambda r: r[0])

# Split each line into words and explode the resulting array of words

words = lines.flatMap(lambda line: line.split(" "))

# Count the occurrences of each word

word\_counts = words.countByValue()

# Print the word count results

for word, count in word\_counts.items():

print(f"{word}: {count}")

# Stop the Spark session

spark.stop()