**Problem Statement**

Read a stream of Strings, fetch the words which can be converted

to numbers. Filter out the rows, where the sum of numbers in that

line is odd.

Provide the sum of all the remaining numbers in that batch.

Assignment 20 Code:

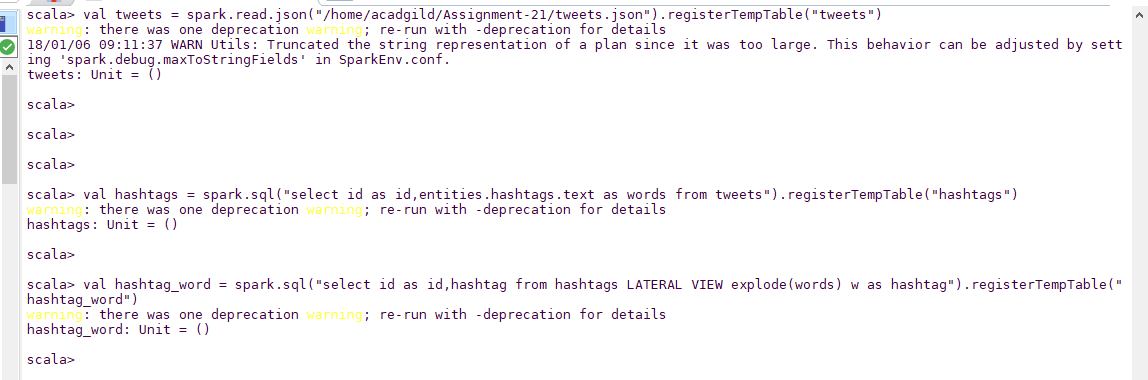
val tweets = spark.read.json("/home/acadgild/Assignment-21/tweets.json").registerTempTable("tweets")

val hashtags = spark.sql("select id as id,entities.hashtags.text as words from tweets").registerTempTable("hashtags")

val hashtag\_word = spark.sql("select id as id,hashtag from hashtags LATERAL VIEW explode(words) w as hashtag").registerTempTable("hashtag\_word")

val popular\_hashtags = spark.sql("select hashtag, count(hashtag) as cnt from hashtag\_word group by hashtag order by cnt desc").show()

Screen-Shot





Submitted By

Shishir KumarJha