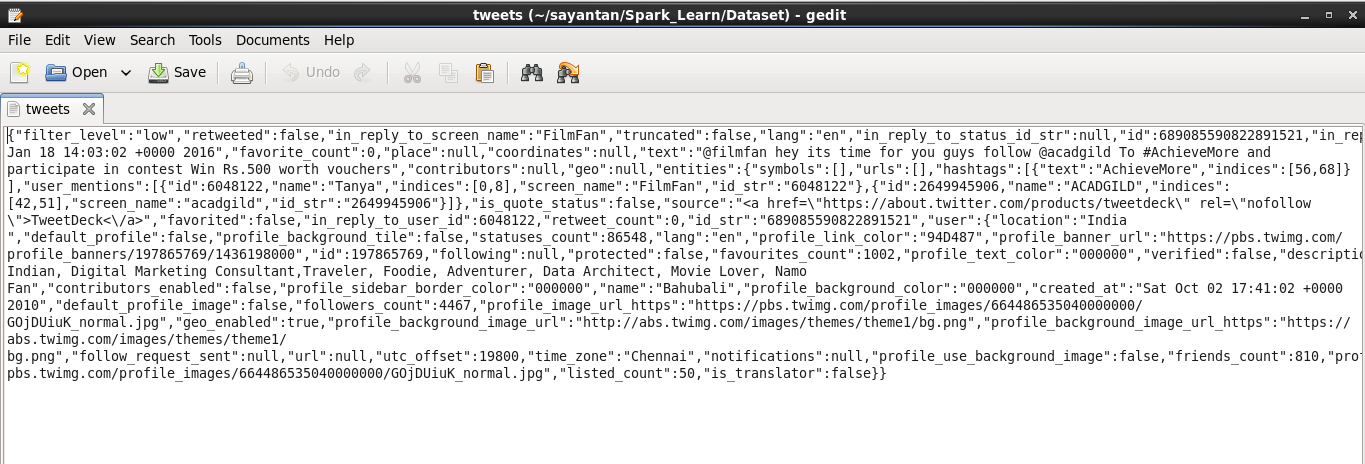
**Assignment23.1**

**Problem Statement:**

**Counting popular hashtags using Spark sql**

**The sample JSON file is as follows**



we will first read this json file using sqlcontext

val tweets = sqlContext.jsonFile("file:///home/cloudera/sayantan/Spark\_Learn/Dataset/tweets").registerTempTable("tweets")

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

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

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

**the output is as follows:**

