Sentiment\_analysis\_on\_demonetization

val tweets = sc.textFile("/home/acadgild/hadoop/demonetization-tweets.csv").map(x => x.split(",")).filter(x=>x.length>=2).map(x => (x(0).replaceAll("\"",""),x(1).replaceAll("\"","").toLowerCase)).map(x => (x.\_1,x.\_2.split(" "))).toDF("id","words")

tweets.registerTempTable("tweets")

val explode = spark.sql("select id as id,explode(words) as word from tweets").registerTempTable("tweet\_word")

val afinn = sc.textFile("/home/acadgild/hadoop/AFINN.txt").map(x => x.split("\t")).map(x => (x(0),x(1))).toDF("word","rating").registerTempTable("afinn")

val join = spark.sql("select t.id,AVG(a.rating) as rating from tweet\_word t join afinn a on t.word=a.word group by t.id order by rating desc").show

