**[root@sandbox-hdp ~]#** $HADOOP\_HOME/bin/hadoop jar /usr/hdp/current/hadoop-mapreduce-client/hadoop-streaming-3.1.1.3.0.1.0-187.jar \

> -input /kalyan/kalyan.txt \

> -output /kalyan/wcountoutput \

> -mapper mapper.py \

> -reducer reducer.py \

> -file /root/mapper.py \

> -file /root/reducer.py

**Input: kalyan.txt**

Oh whoa

Oh whoa

Oh whoa

You know you love me, I know you care

Just shout whenever and I'll be there

You are my love, you are my heart

And we will never, ever, ever be apart

**Mapper:**

Oh 1

Whoa 1

Oh 1

Whoa 1

Oh 1

Whoa 1

You 1

Know 1

You 1

Love 1

me, 1

I 1

Know 1

You 1

Care 1

Just 1

Shout 1

Whenever 1

And 1

I'll 1

Be 1

There 1

You 1

Are 1

My 1

Love, 1

You 1

Are 1

my 1

heart 1

And 1

We 1

Will 1

Never 1

Ever 1

Ever 1

Be 1

Apart 1

**Shuffle sort :**

Oh [1,1,1]

Whoa [1,1,1]

You [1,1]

Know [1,1]

you [1,1,1,]

love 1

me, 1

I 1

care 1

Just 1

shout 1

Whenever 1

and 1

I'll 1

be [1,1]

there 1

are [1,1]

my [1,1]

love, 1

heart 1

we 1

will 1

never 1

ever 1

ever, 1

apart 1

**Result(reduce -sum()):**

[root@sandbox-hdp ~]# hdfs dfs -cat /kalyan/wcountoutput/part-00000

And 1

I 1

I'll 1

Just 1

Oh 3

You 2

and 1

apart 1

are 2

be 2

care 1

ever 1

ever, 1

heart 1

know 2

love 1

love, 1

me, 1

my 2

never, 1

shout 1

there 1

we 1

whenever 1

whoa 3

will 1

you 3