Create logic for linkedIn profile views in Apache Spark

Insert the below records

id,from\_member,to\_member

1,Manasa,Sumit

2,Deepa,Sumit

3,Sumit,Manasa

4,Manasa,Deepa

5,Deepa,Manasa

6,Shilpy,Manasa

from pyspark.sql import SparkSession

import getpass

username = getpass.getuser()

spark = SparkSession. \

builder. \

config('spark.ui.port','0'). \

config("spark.sql.warehouse.dir", f"/user/{username}/warehouse"). \

enableHiveSupport(). \

master('yarn'). \

getOrCreate()

rdd1 = spark.sparkContext.textFile("/user/itv017244/data/input/linkedin\_views.csv")

rdd2 = rdd1.map(lambda x : x.split(",")[2])

rdd3 = rdd2.map(lambda x: (x,1))

rdd4 = rdd3.reduceByKey(lambda x,y : (x+y))

rdd4.collect()