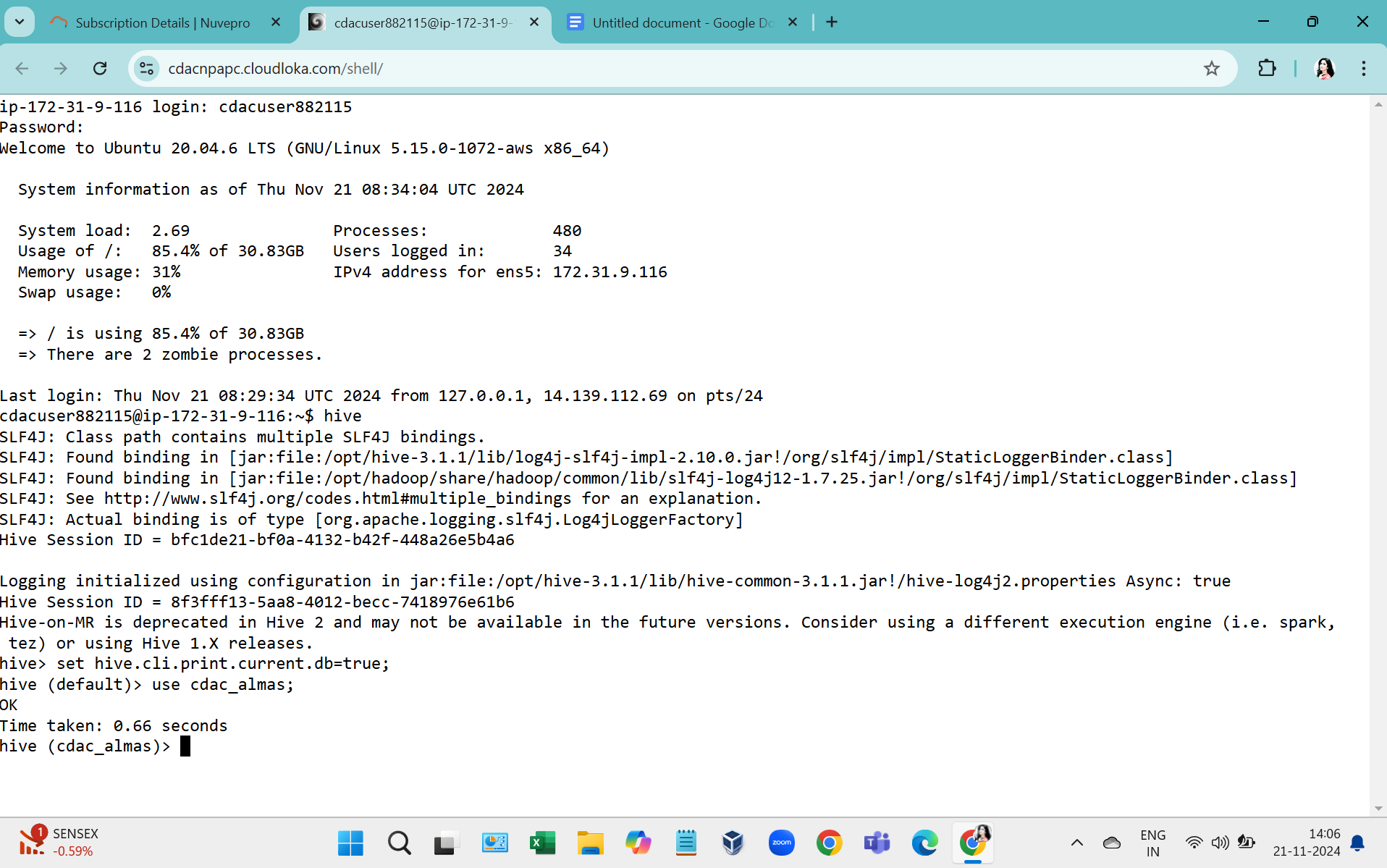
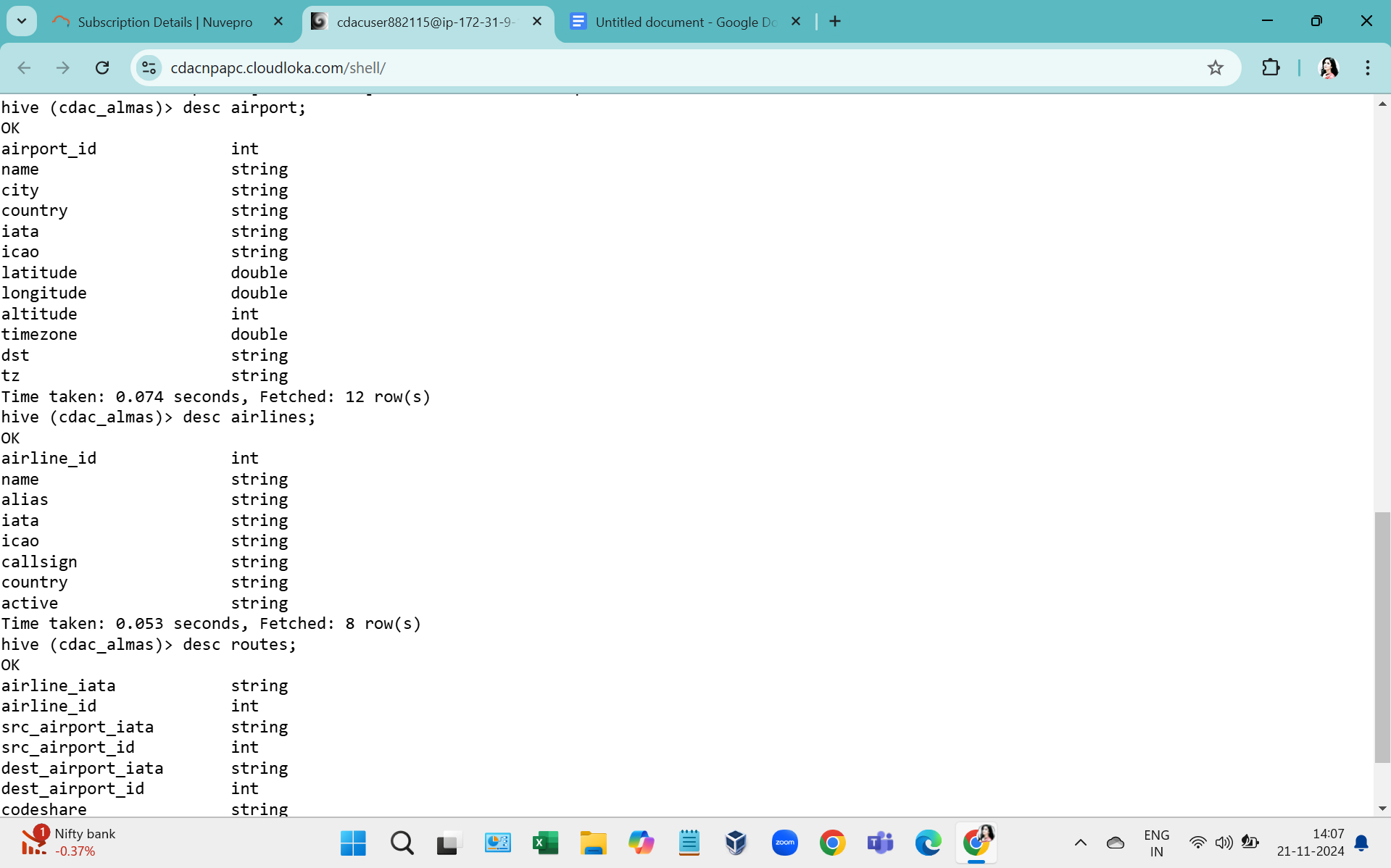
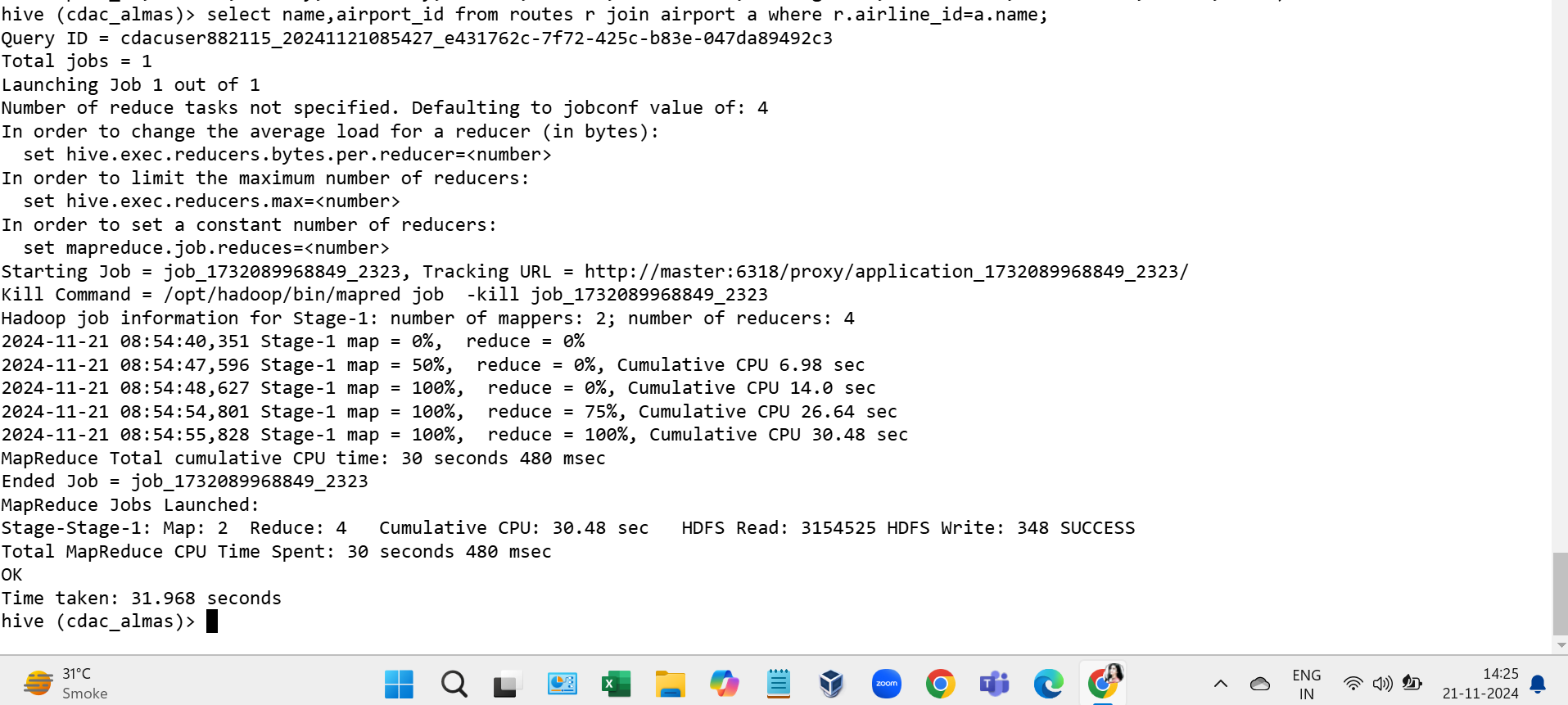
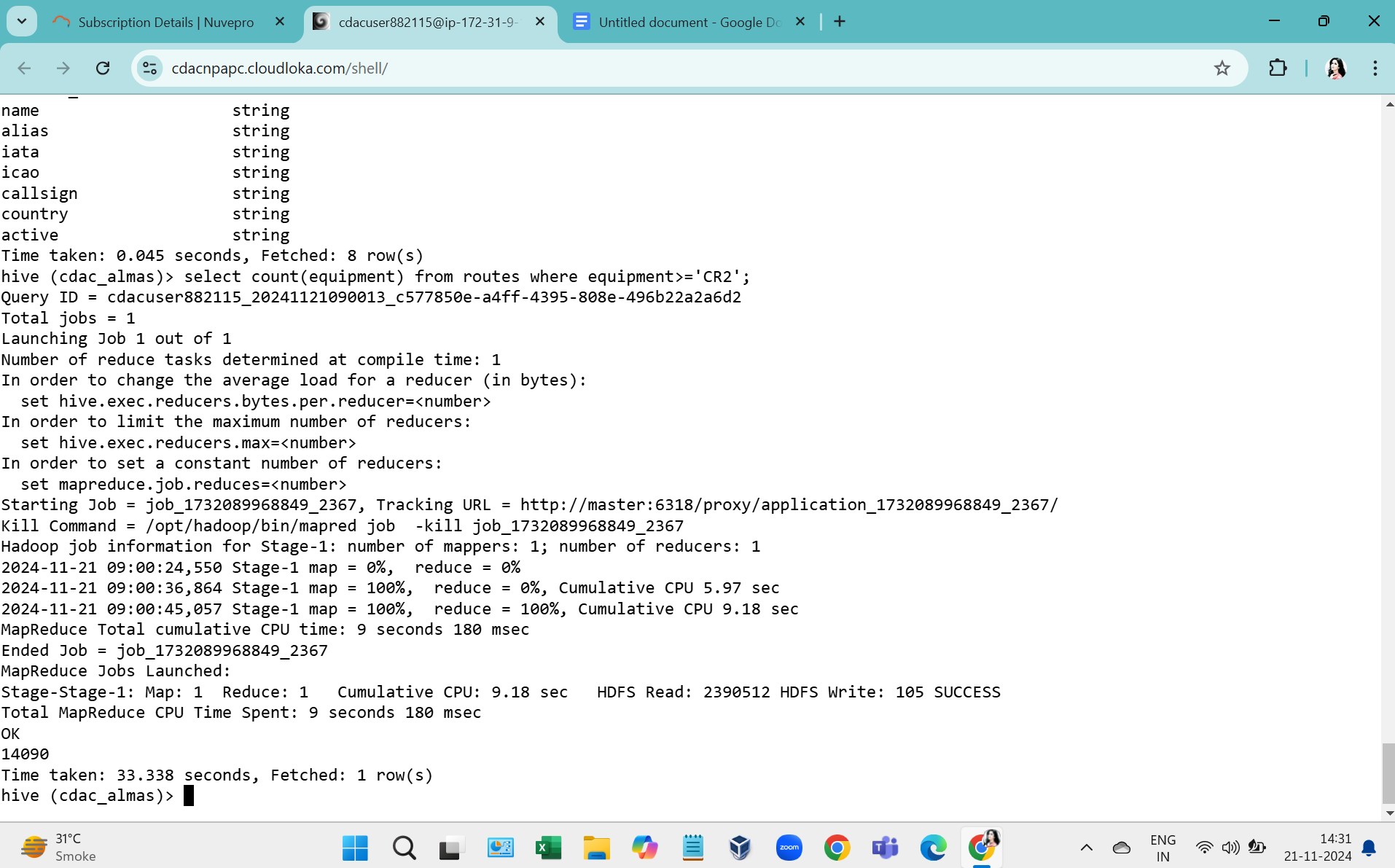
Q.1



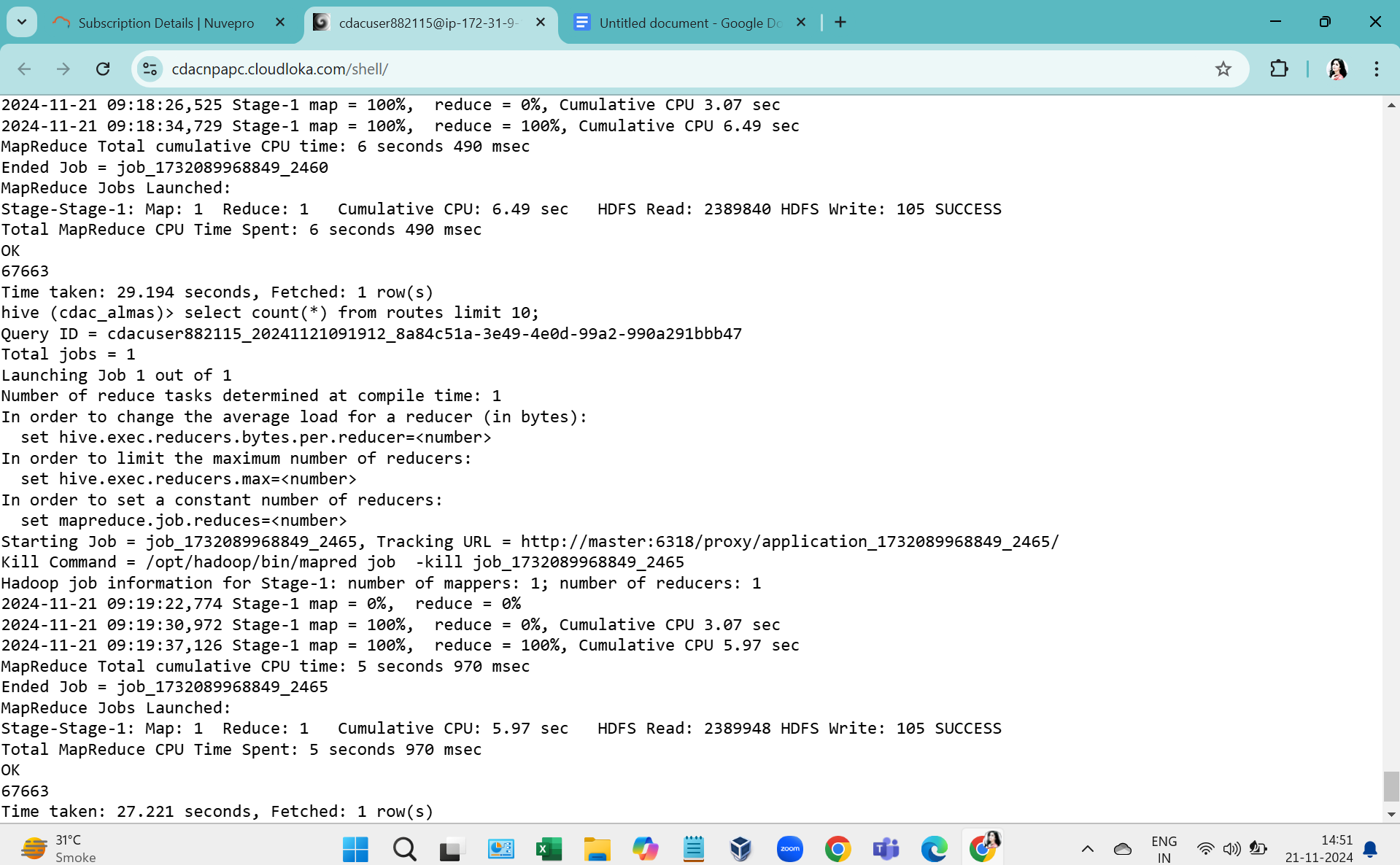


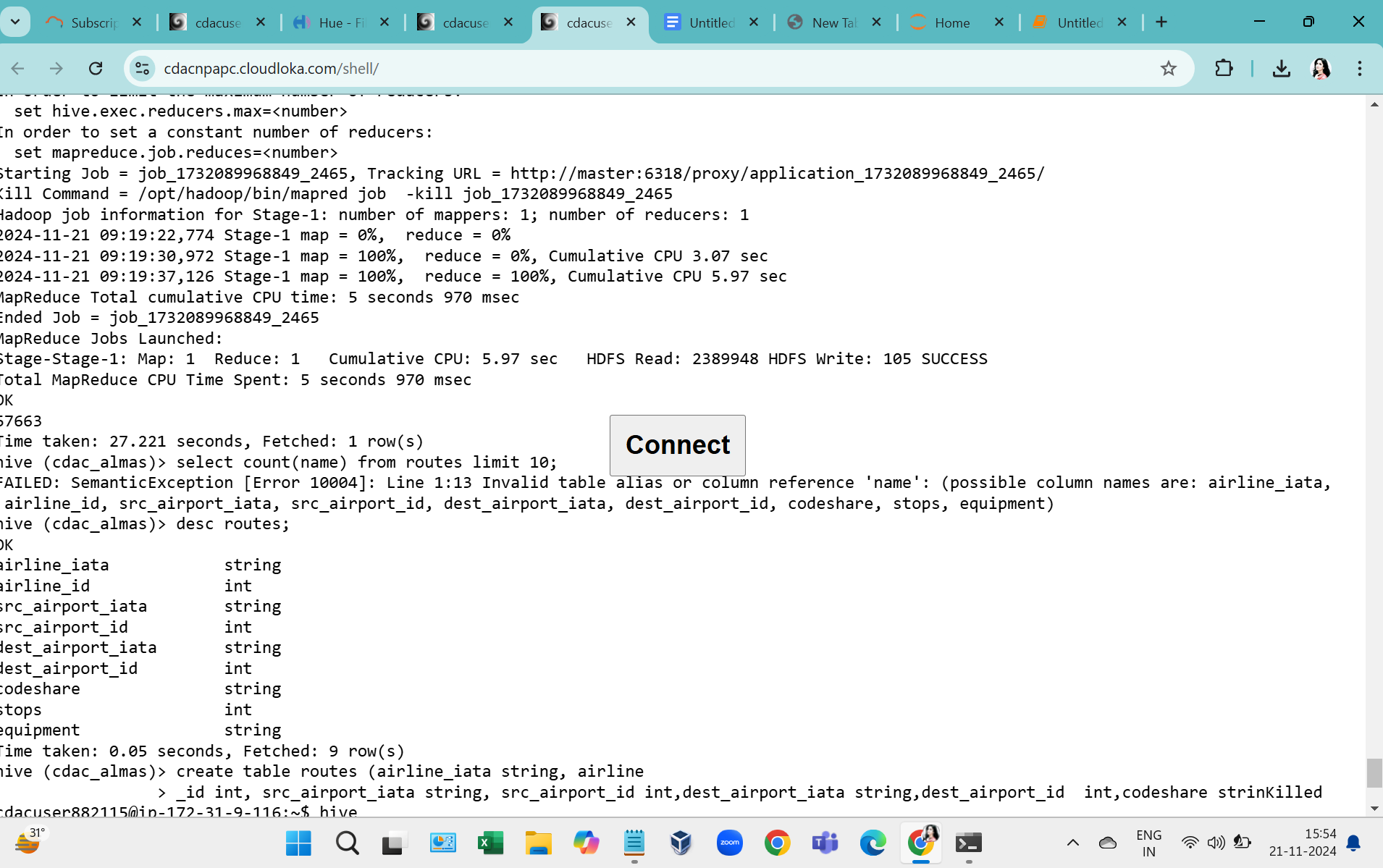


Q.2



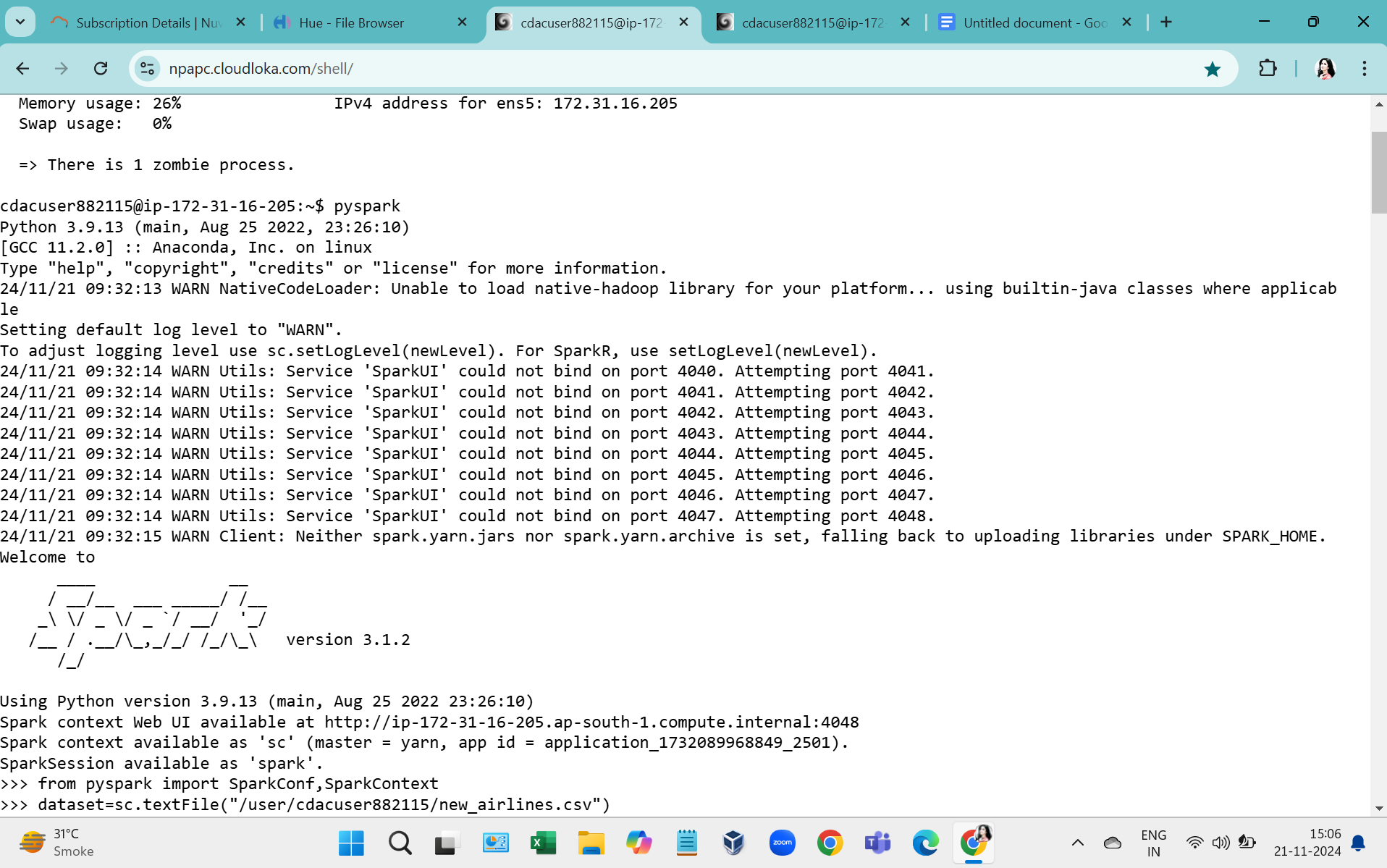
Q.3

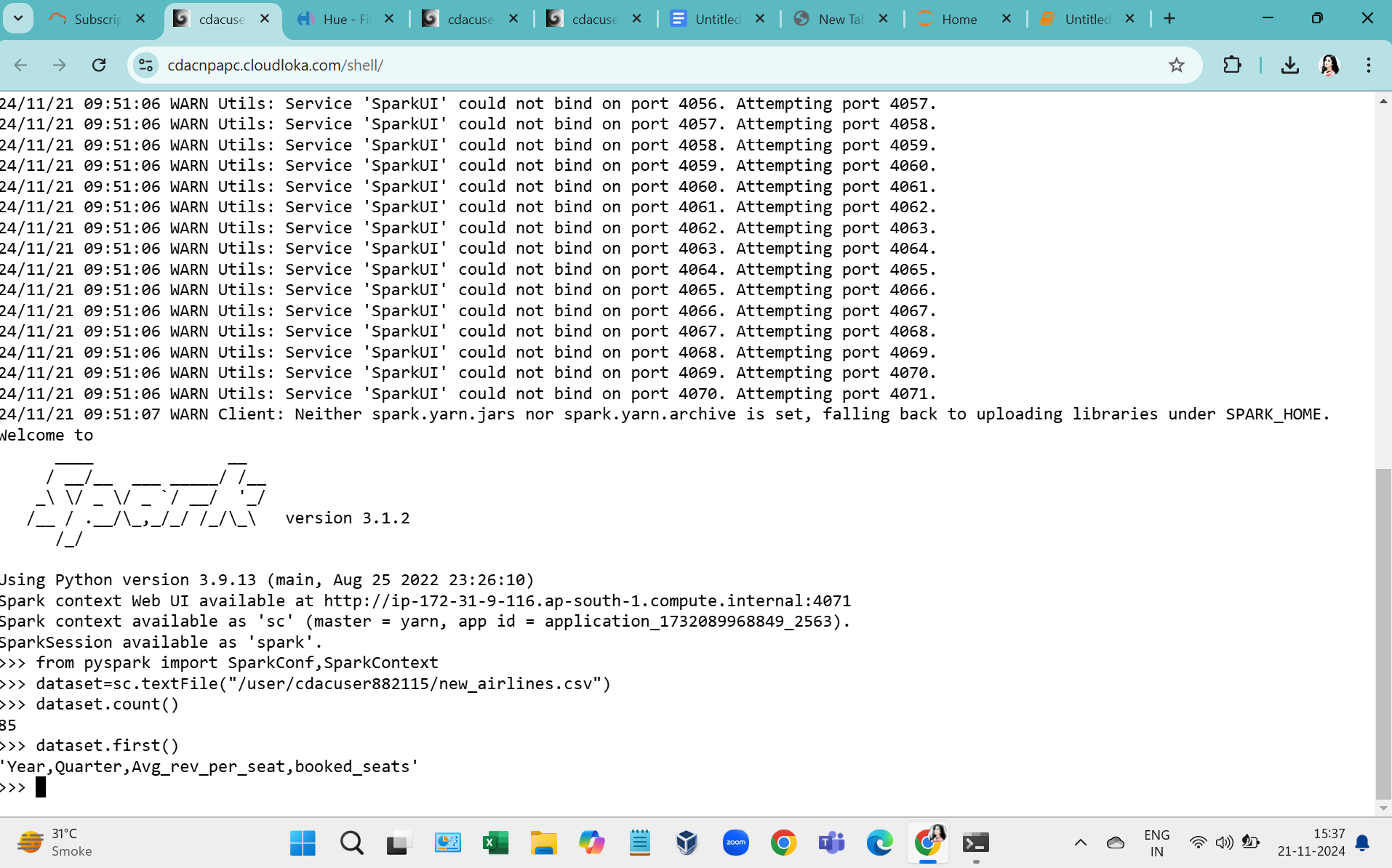


Question 2: 

Spark:

1.





split=dataset.map(lambda a : (a.split(",")[0],a.split(",")[1],a.split(",")[2],a.split(",")[3]))

combine1=dataset.map(lambda a: ((int(a[0]), int(a[1])), float(a[2])))

>>> for line in combine1:

... print(line)

Q.2

avg=dataset.filter(lambda a: -a[1])

>>> for line in avg.take(1):

... print(line)

Q.3 rdd=dataset.map(lambda a,b: a+b)

>>> for line in rdd.take(1):

... print(line)

Q.4