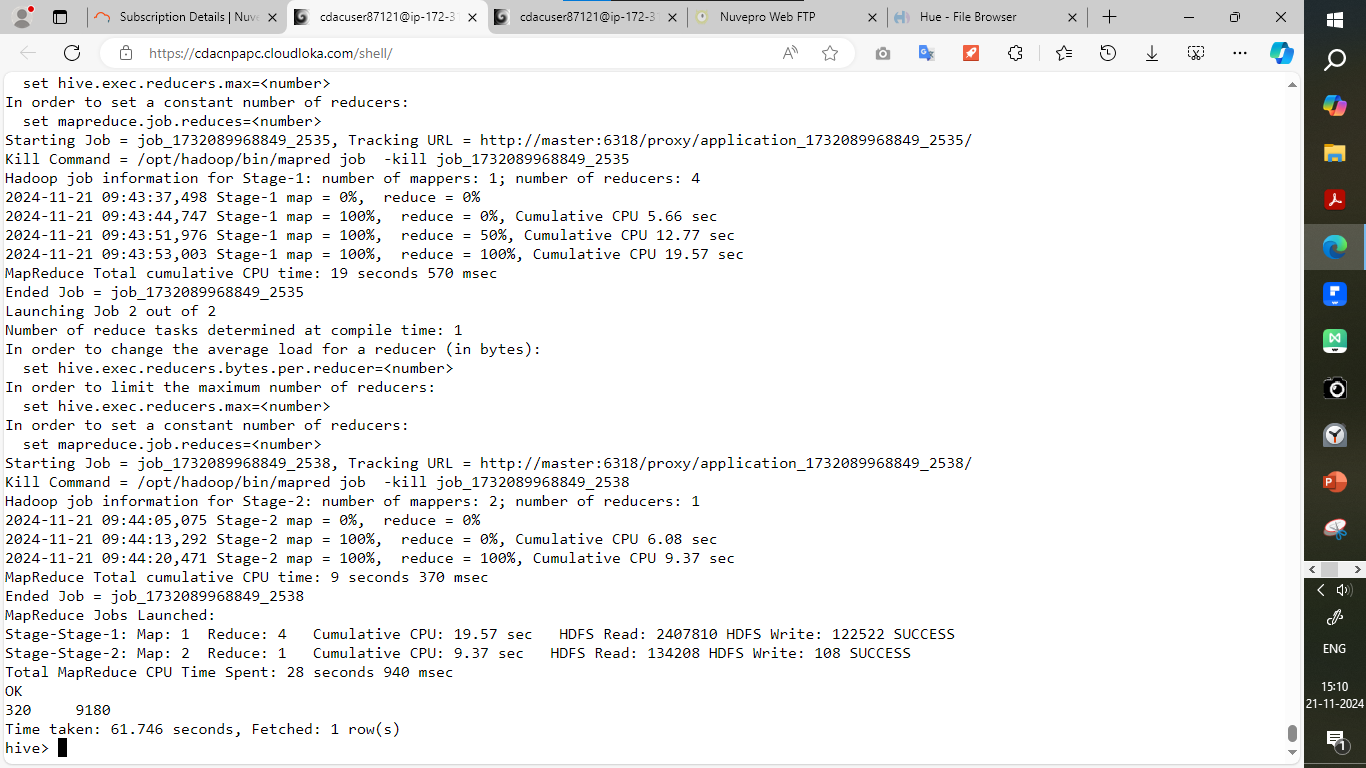
Hive

Determine the aircraft type (Equipment) that is used on the highest number of route.

* Code:

select equipment, count(\*) as ct from routes group by equipment order by ct desc limit 1;



Spark

Find the minimum, maximum and average of the average revenue per seat.

* Average:
* Code:

df.agg(avg("Avg\_rev\_per\_seat").alias("Average revenue")).show()



Maximum:

* Code:

df.select("Avg\_rev\_per\_seat").orderBy("Avg\_rev\_per\_seat",ascending = False).limit(1).show()

Minimum:

* Code:

df.select("Avg\_rev\_per\_seat").orderBy("Avg\_rev\_per\_seat",ascending = True).limit(1).show()



Find the total number of booked seats for all quarters combined.

Code:

df.groupBy("Year").agg(sum("booked\_seats").alias("Total\_booked\_seats")).orderBy("Year").show()

