1.Command:

# Update system packages

sudo apt update

# Install required packages

sudo apt install apt-transport-https ca-certificates curl software-properties-common -y

# Add Docker’s GPG key

curl -fsSL https://download.docker.com/linux/ubuntu/gpg | \

sudo gpg --dearmor -o /usr/share/keyrings/docker-archive-keyring.gpg

# Add Docker repository

echo "deb [arch=$(dpkg --print-architecture) signed-by=/usr/share/keyrings/docker-archive-keyring.gpg] \

https://download.docker.com/linux/ubuntu $(lsb\_release -cs) stable" | \

sudo tee /etc/apt/sources.list.d/docker.list > /dev/null

# Update the package index

sudo apt update

# Install Docker

sudo apt install docker-ce -y

# Start Docker service and enable on boot

sudo systemctl start docker

sudo systemctl enable docker

# Optional: Add your user to the Docker group (log out and log back in after this)

sudo usermod -aG docker $USER

#2.Pull the Impala + Kudu Docker Image

docker pull josemyd/all-in-one-impala-kudu

#3- Run the Container

docker run -it --name impala-kudu-container josemyd/all-in-one-impala-kudu

# This command will start the container and bring you inside it.

#4. Re-enter Container Later

If you exited the container and want to access Impala again:

docker exec -it impala-kudu-container impala-shell

Code : Runn one by one shell and copy it to text file first

SHOW DATABASES;

CREATE DATABASE sl3\_impala;

USE sl3\_impala;

CREATE TABLE employees (

id INT,

name STRING,

department STRING,

salary DOUBLE

)

STORED AS PARQUET;

INSERT INTO employees VALUES

(1, 'Aarav', 'Engineering', 95000),

(2, 'Priya', 'Marketing', 72000),

(3, 'Rohan', 'Finance', 85000),

(4, 'Saanvi', 'Engineering', 98000),

(5, 'Karthik', 'Engineering', 87000),

(6, 'Neha', 'Marketing', 74000);

SELECT \* FROM employees WHERE department = 'Engineering';

SELECT \* FROM employees WHERE salary > 90000;

SELECT department, AVG(salary) AS avg\_salary FROM employees GROUP BY department;

SELECT \* FROM employees ORDER BY salary DESC LIMIT 3;

SELECT e.name, e.department, e.salary FROM employees e

JOIN ( SELECT department, AVG(salary) AS avg\_salary FROM employees

GROUP BY department

) dept\_avg

ON e.department = dept\_avg.department

WHERE e.salary < dept\_avg.avg\_salary;



