

1 BASIC LEVEL

Task No.	Description	Hive Command	Example / Notes
1	Create database	create database employee_db;	Creates DB
	Switch DB	use employee_db;	—
2	Create simple employee table	create table employee_basic(emp_id int, emp_name string, age int);	3-column table
3	Insert sample data	insert into employee_basic values (1,'Amit',25),(2,'Riya',29);	Adds 2 rows
4	Simple select	select * from employee_basic;	View records

2 INTERMEDIATE LEVEL

Task No.	Description	Hive Command	Example / Notes
5	Create detailed employee table with delimiter	create table employee_details(emp_id int, emp_name string, age int, salary float, city string, department string) row format delimited fields terminated by ',';	CSV-friendly table
6	Load data from HDFS	load data inpath '/data/emp_details.csv' into table employee_details;	Loads CSV
7	Filter employees by city	select * from employee_details where city='Bangalore';	Filtering example
8	Max salary	select max(salary) from employee_details;	Aggregation
9	Count by city	select city, count(*) from employee_details group by city;	Grouping

3 ADVANCED LEVEL

Task No.	Description	Hive Command	Example / Notes
10	Create partitioned salary table	create table employee_salary(emp_id int, salary float) partitioned by (year int, month int);	Partitioned by year/month
11	Insert into partition	insert into table employee_salary partition(year=2025, month=1) select emp_id, salary from employee_details;	Partition load
12	Create department table	create table department(dept_id int, dept_name string);	—
	Insert dept values	insert into department values (1,'IT'),(2,'HR'),(3,'Finance');	—
13	Join employee + department	select e.emp_name, e.salary, d.dept_name from employee_details e join department d on e.department=d.dept_name;	Inner join
14	Avg salary by department	select department, avg(salary) from employee_details group by department;	Aggregation

4 EXPERT LEVEL

Tas k No.	Descriptio n	Hive Command	Example / Notes
15	Create bucketed table	set hive.enforce.bucketing=true; create table employee_bucketed(emp_id int, emp_name string, salary float) clustered by(emp_id) into 4 buckets stored as orc;	Bucketed + ORC
16	Insert data into bucketed table	insert into table employee_bucketed select emp_id, emp_name, salary from employee_details cluster by(emp_id);	Ensures bucket distribution
17	Enable ACID	set hive.txn.manager=org.apache.hadoop.hive.ql.lockmgr.DbTxnManager; set hive.support.concurrency=true;	For update/delete
18	Update employee salary	update employee_bucketed set salary = salary + 5000 where emp_id = 1;	Transactional update
19	Top 3 highest-paid per department	select emp_name, department, salary from (select emp_name, department, salary, dense_rank() over (partition by department order by salary desc) as rn from employee_details) t where rn <= 3;	Window function
20	Most experienced employee per city (custom question)	select emp_name, city, experience from (select emp_name, city, experience, row_number() over (partition by city order by experience desc) as rn from employee_details) x where rn=1;	Advanced analytics