1.) table structure below:

emp\_table:

emp\_id, emp\_name, dept\_id,salary

dept\_table:

dept\_id,dept\_name

find 5th higest salary in each department:

output:

emp\_id,emp\_name,salary,department.

**SELECT emp\_name,dept\_name,salary,**

**DENSE\_RANK() OVER (PARTITION BY d. dept\_name ORDER BY salary DESC) AS salary\_rank**

**FROM dept\_table  JOIN emp\_table ON dept\_table.dept\_id = emp\_table.dept\_id;**

2.) table structure below:

custome\_table:

customer\_id,name,email.

order\_table:

order\_id,product\_id,customer\_id,quantity

product\_table:

product\_id,product\_name,price

write a query to find below output:

1.)customer\_id,name,product\_name,quantity.

**select customer\_id,name,product\_name,quantity from customer\_table**

**join order\_table on customer\_table.customer\_id = order\_table .customer\_id**

**join product\_table on order\_table. product\_id= product\_table .product\_id**

2.) customer\_id, total\_purchase\_amount.

**select order\_table.customer\_id, (order\_table.quantity\* product\_table .price) as total\_purchase\_amount from order\_table**

**join product\_table on order\_table.product\_id = product\_table .product\_id**

3.)highest sold product based on quantity.

**select product\_id, max(quantity) from order\_table**

**group by product\_id**

3.) table structure below:

emp\_table:

id,name,gender,city,dept\_id

dept\_table:

dept\_id,dept\_name

write a query to find below output::

1. write query to convert male to female and female to male.

**UPDATE emp\_table SET gender =**

**(CASE**

**WHEN gender ='male' THEN 'female' else 'male'**

**END)**

1. write a query to find emp from same city.

**select distinct employee\_name**

**from emp\_table as T**

**where employee.city = T.city;**

1. write query to find dept\_id not in department table.

**SELECT dept\_id FROM emp\_table  
RIGHT JOIN dept\_table  
ON emp\_table. dept\_id = dept\_table. dept\_id;**

4.) write a python program to change date format.

1-11-2022 to 11-1-2022

mm-dd-yyyy to dd-mm-yyyy

**Ans : datetime.datetime.strptime("1-11-2022", '%m-%d-%Y').strftime('%d/%m/%Y')**

5.) write a pyspark code to read a csv,parquet,json files.

**Ans : df = spark.read.format(“text”/ “csv”/”parquet”/”json”).load("file path")**

**print(df)**