**NAME- PUJAN RASTOGI**

**EMAIL- rastogipujan59@gmail.com**

**ASSESSMENT 5**

**What Is Data Cleaning?**

The process of data cleaning (also called data cleansing) involves identifying any inaccuracies in a dataset and then fixing them. It’s the first step in any analysis and it includes deleting data, updating data, and finding inconsistencies or things that just don’t make sense.

**Data Cleaning Techniques**

* Deleting Duplicate Data
* Removing NULL Values
* Fix structural errors
* Handle missing data
* Normalize data
* Validate data

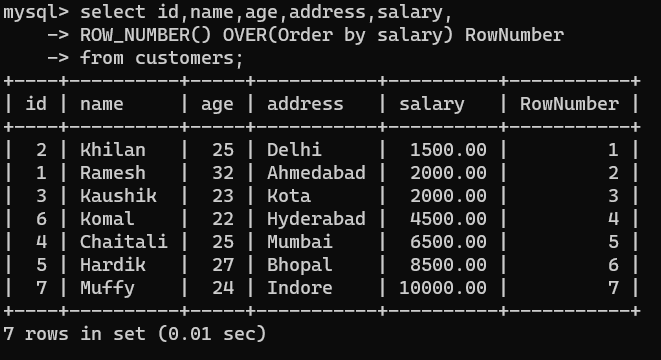
**RANK FUNCTIONS**

**ROW\_NUMBER()- It assigns the sequential rank number to each unique record.**

*select id,name,age,address,salary,*

*ROW\_NUMBER() OVER(Order by salary) RowNumber*

*from customers;*



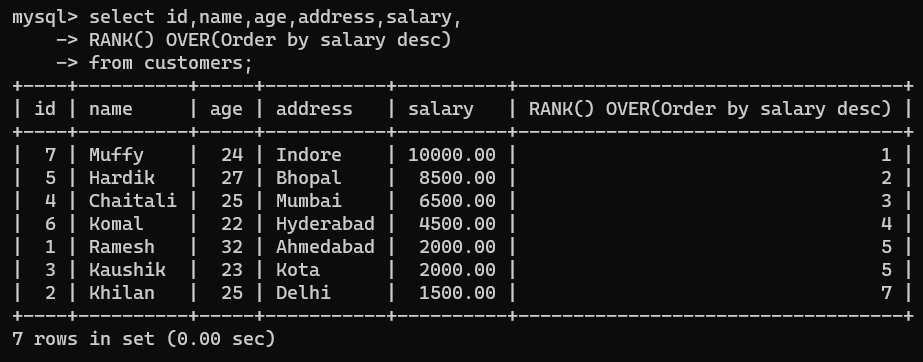
**RANK()-**

**It assigns the rank number to each row in a partition. It skips the number for similar values**

*select id,name,age,address,salary,*

*RANK() OVER(Order by salary desc)*

*from customers;*



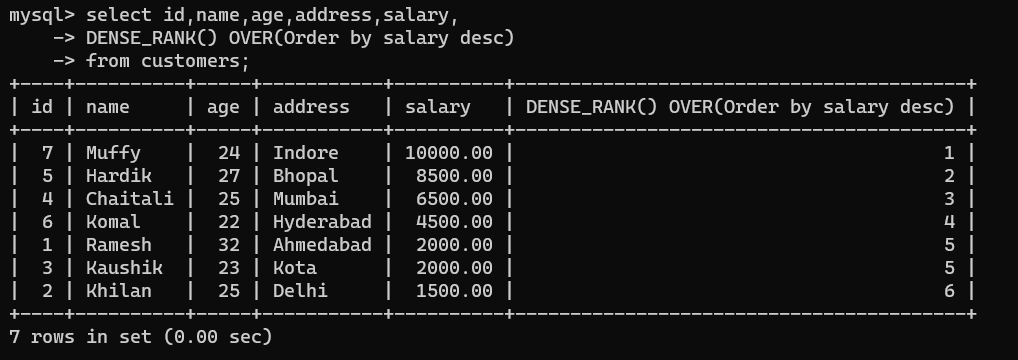
**DENSE\_RANK()**

**It assigns the rank number to each row in a partition. It does not skip the number for similar values.**

*select id,name,age,address,salary,*

*DENSE\_RANK() OVER(Order by salary desc)*

*from customers;*

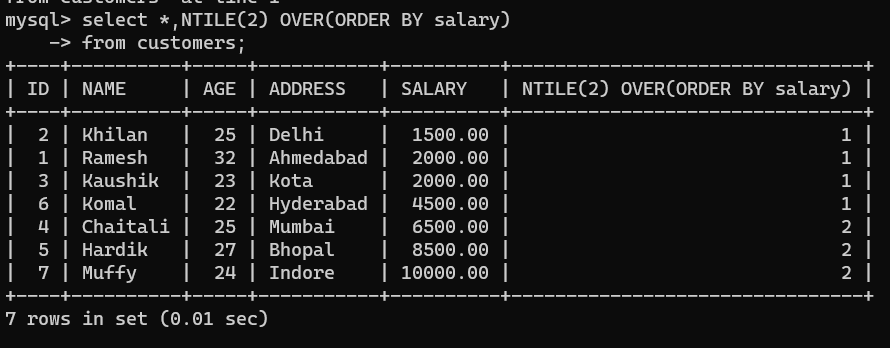


**NTILE()**

**It divides the number of rows as per specified partition and assigns unique value in the partition.**

*select \*, NTILE(2) OVER(ORDER BY salary)*

*from customers;*



**STORED PROCEDURE**

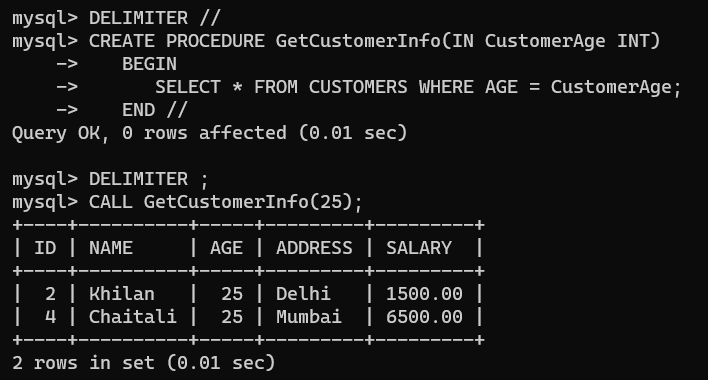
*DELIMITER //*

*CREATE PROCEDURE GetCustomerInfo(IN CustomerAge INT)*

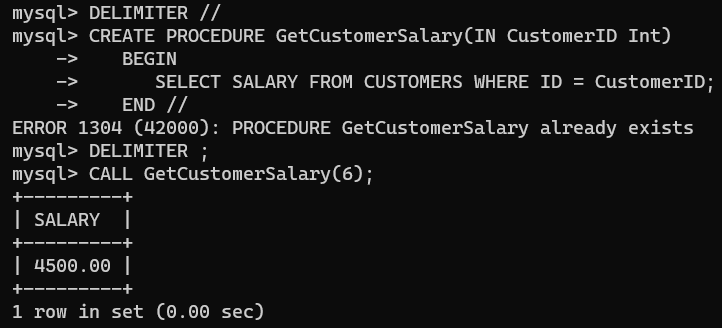
*BEGIN*

*SELECT \* FROM CUSTOMERS WHERE AGE = CustomerAge;*

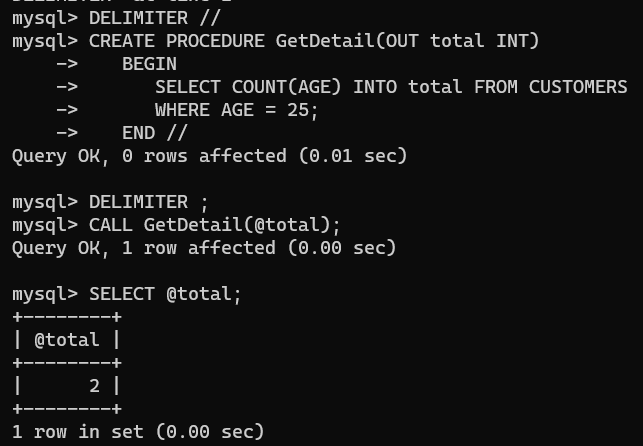
*END //*



Procedure with IN parameter



Procedure with OUT parameter



Procedure with INOUT parameter

