**Step 6: Dynamic Data Masking**

Vaishnavi Pravin Apsingkar

NN: 142

**Part A: Implementing Dynamic Data Masking**

**To mask the “Name” and “Salary”:**

use KOT142;

ALTER TABLE Driver

ALTER COLUMN Name ADD MASKED WITH (FUNCTION = 'default()');

ALTER TABLE Driver

ALTER COLUMN Salary ADD MASKED WITH (FUNCTION = 'default()');

**Result:**

Commands completed successfully.

**Part B: Creating user and granting permission**

**To create new user:**

use KOT142;

CREATE USER user142 WITHOUT LOGIN;

**Result:**

Commands completed successfully.

**To grant permissions to the user:**

use KOT142;

GRANT SELECT ON Driver TO user142;

**Result:**

Commands completed successfully.

**For Testing:**

**As admin:**

SELECT Name, Salary FROM Driver WHERE DriverID = 'D107';

**Result:**

Name Salary

-------------------- ---------------------

Dave White 59100.00

(1 row affected)

This shows original data to the admin.

**As user142 (for whom data is masked) to test the mask:**

EXECUTE AS USER = 'user142';

SELECT Name, Salary FROM Driver WHERE DriverID = 'D107';

REVERT;

**Result:**

Name Salary

-------------------- ---------------------

xxxx 0.00

(1 row affected)

This shows masked data for user142.

**As user142 trying to access another table (who has no access to it):**

EXECUTE AS USER = 'user142';

SELECT \* FROM Truck;

REVERT;

**Result:**

Msg 229, Level 14, State 5, Line 2

The SELECT permission was denied on the object 'Truck', database 'KOT142', schema 'dbo'.

This shows that it denies access to the data for user142 for ‘Truck’ table as no access has been given to it.

**Process to find good estimate of salary range for driver with driverID ‘D107’:**

**Step 1:**

Estimate the broad range of salary. Then, check data for salary ranges in 0 to 100000 for driver ID D107.

**Step 2:**

a. Divide the range into half. Continue checking in first half range.

b. If data available, continue step 2.

c. If data not available in first half, check in next half range. Continue step 2.

**Step 3:**

Stop when you get good and close estimate of salary ranges within range of 0 to 500.

**Let’s implement above steps:**

**As an admin to see actual data for taking broad range of salaries:**

SELECT \* FROM Driver;

**Result:**

DriverID Name Salary StartDate Status

-------- -------------------- --------------------- ---------- ----------

D101 Bill Green 65200.00 2021-11-12 Full

D102 Bob Jones 66100.00 2021-10-30 Full

D107 Dave White 59100.00 2022-01-12 Full

D110 Ann Black 65200.00 2021-06-22 Part

D142 Vaishnavi Apsingkar 60000.00 2023-01-06 Part

D225 John Dough 58000.00 2021-05-15 Full

D280 Gail Brown 58000.00 2022-01-12 Full

D301 Mark White 59200.00 2022-03-05 Part

D302 Bob Jones 59000.00 2022-03-06 Part

(9 rows affected)

This shows the original data of Driver. So, we can continue estimating within this broad range of salaries from $0 to $100000.

**As user142, finding good estimate for the salary of driver with ID ‘D107’:**

* **Checking in range from 0 to 100000:**

EXECUTE AS USER = 'user142';

SELECT \* FROM Driver WHERE Salary>0

AND Salary<100000;

**Result:**

DriverID Name Salary StartDate Status

-------- -------------------- --------------------- ---------- ----------

D101 xxxx 0.00 2021-11-12 Full

D102 xxxx 0.00 2021-10-30 Full

D107 xxxx 0.00 2022-01-12 Full

D110 xxxx 0.00 2021-06-22 Part

D142 xxxx 0.00 2023-01-06 Part

D225 xxxx 0.00 2021-05-15 Full

D280 xxxx 0.00 2022-01-12 Full

D301 xxxx 0.00 2022-03-05 Part

D302 xxxx 0.00 2022-03-06 Part

(9 rows affected)

The data with D107 is available in result, means this driver has salary is range from $0 to $100000.

* **Checking in range from 0 to 50000:**

EXECUTE AS USER = 'user142';

SELECT \* FROM Driver WHERE Salary>0

AND Salary<50000;

**Result:**

DriverID Name Salary StartDate Status

-------- -------------------- --------------------- ---------- ----------

(0 rows affected)

Here, the data for driver ID ‘D107’ is not available in result. Thus, salary of this driver is in range from $50000 to $100000.

* **Checking in range of 50000 to 100000:**

EXECUTE AS USER = 'user142';

SELECT \* FROM Driver WHERE Salary>50000

AND Salary<100000;

**Result:**

DriverID Name Salary StartDate Status

-------- -------------------- --------------------- ---------- ----------

D101 xxxx 0.00 2021-11-12 Full

D102 xxxx 0.00 2021-10-30 Full

D107 xxxx 0.00 2022-01-12 Full

D110 xxxx 0.00 2021-06-22 Part

D142 xxxx 0.00 2023-01-06 Part

D225 xxxx 0.00 2021-05-15 Full

D280 xxxx 0.00 2022-01-12 Full

D301 xxxx 0.00 2022-03-05 Part

D302 xxxx 0.00 2022-03-06 Part

(9 rows affected)

The data for DriverID D107 is there. This shows, salary of this driver is in range of $50000 to $100000.

* **Checking in range of 50000 to 75000:**

EXECUTE AS USER = 'user142';

SELECT \* FROM Driver WHERE Salary>50000

AND Salary<75000;

**Result:**

DriverID Name Salary StartDate Status

-------- -------------------- --------------------- ---------- ----------

D101 xxxx 0.00 2021-11-12 Full

D102 xxxx 0.00 2021-10-30 Full

D107 xxxx 0.00 2022-01-12 Full

D110 xxxx 0.00 2021-06-22 Part

D142 xxxx 0.00 2023-01-06 Part

D225 xxxx 0.00 2021-05-15 Full

D280 xxxx 0.00 2022-01-12 Full

D301 xxxx 0.00 2022-03-05 Part

D302 xxxx 0.00 2022-03-06 Part

(9 rows affected)

The data for D107 driverID is available in result. This shows the salary of this driver is in range of $50000 to $75000.

* **Checking in range of 50000 to 62000:**

EXECUTE AS USER = 'user142';

SELECT \* FROM Driver WHERE Salary>50000

AND Salary<62000;

**Result:**

DriverID Name Salary StartDate Status

-------- -------------------- --------------------- ---------- ----------

D107 xxxx 0.00 2022-01-12 Full

D142 xxxx 0.00 2023-01-06 Part

D225 xxxx 0.00 2021-05-15 Full

D280 xxxx 0.00 2022-01-12 Full

D301 xxxx 0.00 2022-03-05 Part

D302 xxxx 0.00 2022-03-06 Part

(6 rows affected)

Still, the data for D107 driver ID can be seen in result. This shows that the salary of this driver is in range of $50000 to $62000.

* **Checking in range of 50000 to 56000:**

EXECUTE AS USER = 'user142';

SELECT \* FROM Driver WHERE Salary>50000

AND Salary<56000;

**Result:**

DriverID Name Salary StartDate Status

-------- -------------------- --------------------- ---------- ----------

(0 rows affected)

The data is not available for D107. This shows that salary of driver is not in the range of $50000 to $56000. Thus, it must be in the range of $56000 to $62000.

* **Checking in the range of 56000 to 62000:**

EXECUTE AS USER = 'user142';

SELECT \* FROM Driver WHERE Salary>56000

AND Salary<62000;

**Result:**

DriverID Name Salary StartDate Status

-------- -------------------- --------------------- ---------- ----------

D107 xxxx 0.00 2022-01-12 Full

D142 xxxx 0.00 2023-01-06 Part

D225 xxxx 0.00 2021-05-15 Full

D280 xxxx 0.00 2022-01-12 Full

D301 xxxx 0.00 2022-03-05 Part

D302 xxxx 0.00 2022-03-06 Part

(6 rows affected)

The data is available for D107 driverID. This shows that salary of this driver is in range of $56000 to $62000.

* **Checking in the range of 56000 to 59000:**

EXECUTE AS USER = 'user142';

SELECT \* FROM Driver WHERE Salary>56000

AND Salary<59000;

**Result:**

DriverID Name Salary StartDate Status

-------- -------------------- --------------------- ---------- ----------

D225 xxxx 0.00 2021-05-15 Full

D280 xxxx 0.00 2022-01-12 Full

(2 rows affected)

There is no data for driverID D107. This means, salary of driver is not in range of $56000 to $59000. So, it is in range from 59000 to 62000.

* **Checking in range of 59000 to 62000:**

EXECUTE AS USER = 'user142';

SELECT \* FROM Driver WHERE Salary>59000

AND Salary<62000;

**Result:**

DriverID Name Salary StartDate Status

-------- -------------------- --------------------- ---------- ----------

D107 xxxx 0.00 2022-01-12 Full

D142 xxxx 0.00 2023-01-06 Part

D301 xxxx 0.00 2022-03-05 Part

(3 rows affected)

The data for driverID D107 is available. Thus, salary is in range of $59000 to $62000.

* **Checking in range of 59000 to 60500:**

EXECUTE AS USER = 'user142';

SELECT \* FROM Driver WHERE Salary>59000

AND Salary<60500;

**Result:**

DriverID Name Salary StartDate Status

-------- -------------------- --------------------- ---------- ----------

D107 xxxx 0.00 2022-01-12 Full

D142 xxxx 0.00 2023-01-06 Part

D301 xxxx 0.00 2022-03-05 Part

(3 rows affected)

Data is available. This shows salary for driver with driverID D107 is in range of 59000 to 60500.

* **Checking in range of 59000 to 59750:**

EXECUTE AS USER = 'user142';

SELECT \* FROM Driver WHERE Salary>59000

AND Salary<59750;

**Result:**

DriverID Name Salary StartDate Status

-------- -------------------- --------------------- ---------- ----------

D107 xxxx 0.00 2022-01-12 Full

D301 xxxx 0.00 2022-03-05 Part

(2 rows affected)

Thus salary of driver with driver ID D107 is in range of $59000 to $59750.

**Now,**

* **Checking in range of 59000 to 59375:**

EXECUTE AS USER = 'user142';

SELECT \* FROM Driver WHERE Salary>59000

AND Salary<59375;

**Result:**

DriverID Name Salary StartDate Status

-------- -------------------- --------------------- ---------- ----------

D107 xxxx 0.00 2022-01-12 Full

D301 xxxx 0.00 2022-03-05 Part

(2 rows affected)

Data is available for Driver ID D107. So, salary of driver is in range of $59000 to $59375.

This gives very close and good estimate of salary for driver with driverID D107.

In this way, we can use binary search technique to find the good estimate of data range in case of masked data.