

Name: Rohan Vinayak Chaudhari

Batch: Data Engineering

Date:23/01/2024

Topic: MYSQL(Data cleaning,rank,stored procedure)

Solution:

1. MYSQL:

ISNULL:

The screenshot shows a SQL IDE window with a query editor and a results grid. The query is as follows:

```
69
70 /*isnull*/
71 • SELECT DISTINCT * FROM animals a
72   LEFT JOIN adoptions ad ON ad.animal_id = a.id
73   WHERE ad.date IS NULL;
74
```

The results grid displays 12 rows of data. The columns are: id, name, breed, color, gender, status, species, shelter, animal_id, name, contact, and date. The data is as follows:

id	name	breed	color	gender	status	species	shelter	animal_id	name	contact	date
1	Bellyflop	Beagle	Brown	Male	0	Dog	1	NULL	NULL	NULL	NULL
2	Bellyflop	Beagle	Brown	Male	0	Dog	1	NULL	NULL	NULL	NULL
3	Snowy	Husky	White	Female	0	Dog	1	NULL	NULL	NULL	NULL
4	Princess	Pomeranian	Brown	Female	0	Dog	1	NULL	NULL	NULL	NULL
5	Princess	Poodle	Brown	Female	0	Dog	1	NULL	NULL	NULL	NULL
6	Meowmix	Munchkin	Yellow	Female	0	Cat	1	NULL	NULL	NULL	NULL
7	Ash	Persian	Gray	Female	0	Cat	1	NULL	NULL	NULL	NULL
8	Tiger	Bengal	Brown	Male	0	Cat	1	NULL	NULL	NULL	NULL
9	Snoops	Beagle	Brown	Male	0	Dog	2	NULL	NULL	NULL	NULL
10	Salt	Turkish Angora	White	Female	0	Cat	2	NULL	NULL	NULL	NULL
11											
12											

The output section shows the following message:

```
1 | 22:36:53 | SELECT DISTINCT * FROM animals a LEFT JOIN adoptions ad ON ad.animal_id = a.id WHERE ad.date IS NULL LIMIT ... | 10 row(s) returned
```

COALESCE:

SQL Fiddle

Limit to 1000 rows

```
75 /*coalesce*/
76 • SELECT *,COALESCE(date, 'Unknown') FROM animals a
77 LEFT JOIN adoptions ad ON ad.animal_id = a.id
78 WHERE ad.date IS NULL;
79
80
```

Result Grid

	id	name	breed	color	gender	status	species	shelter	animal_id	name	contact	date	COALESCE(date, 'Unknown')
▶	1	Bellyflo	Beagle	Brown	Male	0	Dog	1	NULL	NULL	NULL	NULL	Unknown
	2	Bellyflo	Beagle	Brown	Male	0	Dog	1	NULL	NULL	NULL	NULL	Unknown
	3	Snowy	Husky	White	Female	0	Dog	1	NULL	NULL	NULL	NULL	Unknown
	4	Princess	Pomeranian	Brown	Female	0	Dog	1	NULL	NULL	NULL	NULL	Unknown
	6	Princess	Poodle	Brown	Female	0	Dog	1	NULL	NULL	NULL	NULL	Unknown
	8	Meowmix	Munchkin	Yellow	Female	0	Cat	1	NULL	NULL	NULL	NULL	Unknown
	9	Ash	Persian	Gray	Female	0	Cat	1	NULL	NULL	NULL	NULL	Unknown
	10	Tiger	Bengal	Brown	Male	0	Cat	1	NULL	NULL	NULL	NULL	Unknown
	11	Snoops	Beagle	Brown	Male	0	Dog	2	NULL	NULL	NULL	NULL	Unknown
	12	Salt	Turkish Angora	White	Female	0	Cat	2	NULL	NULL	NULL	NULL	Unknown

Result 7

Output

Action Output

#	Time	Action	Message
1	22:37:25	SELECT *,COALESCE(date, 'Unknown') FROM animals a LEFT JOIN adoptions ad ON ad.animal_id = a.id WHERE ad.date IS NULL;	10 row(s) returned

STORED PROCEDURE:

SQL File 2

Limit to 1000 rows

```
281  /*stored procedure*/
282  DELIMITER //
283  • CREATE PROCEDURE GetAllanimals()
284  BEGIN
285      SELECT * FROM animals;
286  END //
287
288  DELIMITER ;
289  • CALL GetAllanimals();
```

Result Grid

	id	name	breed	color	gender	status	species	shelter
▶	1	Bellyflop	Beagle	Brown	Male	0	Dog	1
	2	Bellyflop	Beagle	Brown	Male	0	Dog	1
	3	Snowy	Husky	White	Female	0	Dog	1
	4	Princess	Pomeranian	Brown	Female	0	Dog	1
	5	Cricket	Chihuahua	Brown	Male	1	Dog	1
	6	Princess	Poodle	Brown	Female	0	Dog	1
	7	Spot	Dalmation	Black and White	Male	1	Dog	1
	8	Meowmix	Munchkin	Yellow	Female	0	Cat	1
	9	Ash	Persian	Gray	Female	0	Cat	1
	10	Tiger	Bengal	Brown	Male	0	Cat	1
	11	Snoops	Beagle	Brown	Male	0	Dog	2

Result 8

Output

Action Output

#	Time	Action	Message
✓ 1	22:38:14	CALL GetAllanimals()	12 row(s) returned

DENSE RANK:

SQL File 2

Limit to 1000 rows

```
90
91 /*dense rank*/
92 • SELECT id,name,breed, dense_rank()
93 OVER ( partition by name order by id desc )
94 AS 'dense_rank' FROM animals;
95
```

Result Grid

	id	name	breed	dense_rank
▶	9	Ash	Persian	1
	2	Bellyflop	Beagle	1
	1	Bellyflop	Beagle	2
	5	Crocket	Chihuahua	1
	8	Meowmix	Munchkin	1
	6	Princess	Poodle	1
	4	Princess	Pomeranian	2
	12	Salt	Turkish Angora	1
	11	Snoops	Beagle	1
	3	Snowy	Husky	1
	7	Spot	Dalmation	1

Result 9 x

Output

Action Output

#	Time	Action	Message
✓ 1	22:40:03	SELECT id,name,breed, dense_rank() OVER (partition by name order by id desc) AS 'dense_rank' FROM animals	12 row(s) returned

RANK:

```
36  /*rank*/
37  SELECT *,
38  RANK() OVER (ORDER BY id) my_rank
39  FROM animals;
40
41  /*PERCENT RANK*/
```

Result Grid									
Filter Rows:									
Exports: Wrap Cell Contents									
	id	name	breed	color	gender	status	species	shelter	my_rank
▶	1	Bellyflop	Beagle	Brown	Male	0	Dog	1	1
	2	Bellyflop	Beagle	Brown	Male	0	Dog	1	2
	3	Snowy	Husky	White	Female	0	Dog	1	3
	4	Princess	Pomeranian	Brown	Female	0	Dog	1	4
	5	Cricket	Chihuahua	Brown	Male	1	Dog	1	5
	6	Princess	Poodle	Brown	Female	0	Dog	1	6
	7	Spot	Dalmatian	Black and White	Male	1	Dog	1	7
	8	Meowmix	Munchkin	Yellow	Female	0	Cat	1	8
	9	Ash	Persian	Gray	Female	0	Cat	1	9
	10	Tiger	Bengal	Brown	Male	0	Cat	1	10
	11	Snoops	Beagle	Brown	Male	0	Dog	2	11

Result 10 x

Output

Action Output

#	Time	Action	Message
✓ 1	22:45:55	SELECT *, RANK() OVER (ORDER BY id) my_rank FROM animals	12 row(s) returned

PERCENT RANK:

The screenshot shows a SQL IDE interface. The top pane contains a SQL query: `/*PERCENT RANK*/
SELECT id,name,breed,
PERCENT_RANK() OVER (PARTITION BY gender ORDER BY id) my_rank
FROM animals;` The bottom pane displays the results of the query in a table with columns: id, name, breed, and my_rank. The results are sorted by id. The bottom-most pane shows the execution log with a single entry: `1 [22:41:36] SELECT id,name,breed, PERCENT_RANK() OVER (PARTITION BY gender ORDER BY id) my_rank FROM animals [12 rows returned]`

```
/*PERCENT RANK*/
SELECT id,name,breed,
PERCENT_RANK() OVER (PARTITION BY gender ORDER BY id) my_rank
FROM animals;
```

id	name	breed	my_rank
3	Snowy	Husky	0
4	Princess	Pomeranian	0.2
6	Princess	Poodle	0.4
8	Meowmix	Munchkin	0.6
9	Ash	Persian	0.8
12	Salt	Turkish Angora	1
1	Bellyflopp	Beagle	0
2	Bellyflopp	Beagle	0.2
5	Cindrat	Chihuahua	0.4
7	Spot	Dalmatian	0.6
10	Tiger	Bengal	0.8

Result 11

Output

Action Output

#	Time	Action	Message
1	[22:41:36]	SELECT id,name,breed, PERCENT_RANK() OVER (PARTITION BY gender ORDER BY id) my_rank FROM animals	[12 rows returned]