Problem 3055: Top Percentile Fraud

Problem Information

Difficulty: Medium
Acceptance Rate: 0.00%
Paid Only: No

Problem Description

Table:

Fraud

+-------------+---------+ | Column Name | Type | +-------------+---------+ | policy_id | int | | state | varchar | | fraud_score | int | +-------------+---------+ policy_id is column of unique values for this table. This table contains policy id, state, and fraud score.

The Leetcode Insurance Corp has developed an ML-driven

predictive model

to detect the

likelihood

of fraudulent claims. Consequently, they allocate their most seasoned claim adjusters to address the top

5%

of

claims

flagged

by this model.

Write a solution to find the top

5

percentile

of claims from

each state

.

Return

the result table ordered by

state

in

ascending

order,

fraud_score

in

descending

order, and

policy_id

in

ascending

order.

The result format is in the following example.

Example 1:

Input:

Fraud table: +-----------+------------+-------------+ | policy_id | state | fraud_score | +----------+------------+-------------+ | 1 | California | 0.92 | | 2 | California | 0.68 | | 3 | California | 0.17 | | 4 | New York | 0.94 | | 5 | New York | 0.81 | | 6 | New York | 0.77 | | 7 | Texas | 0.98 | | 8 | Texas | 0.97 | | 9 | Texas | 0.96 | | 10 | Florida | 0.97 | | 11 | Florida | 0.98 | | 12 | Florida | 0.78 | | 13 | Florida | 0.88 | | 14 | Florida | 0.66 | +-----------+------------+-------------+

Output:

+-----------+------------+-------------+ | policy_id | state | fraud_score | +----------+------------+-------------+ | 1 | California | 0.92 | | 11 | Florida | 0.98 | | 4 | New York | 0.94 | | 7 | Texas | 0.98 | +-----------+------------+-------------+

Explanation

- For the state of California, only policy ID 1, with a fraud score of 0.92, falls within the top 5 percentile for this state. - For the state of Florida, only policy ID 11, with a fraud score of 0.98, falls within the top 5 percentile for this state. - For the state of New York, only policy ID 4, with a fraud score of 0.94, falls within the top 5 percentile for this state. - For the state of Texas, only policy ID 7, with a fraud score of 0.98, falls within the top 5 percentile for this state. Output table is ordered by state in ascending order, fraud score in descending order, and policy ID in ascending order.

Code Snippets

MySQL:

```
# Write your MySQL query statement below
```

MS SQL Server:

```
/* Write your T-SQL query statement below */
```

PostgreSQL:

```
-- Write your PostgreSQL query statement below
```

Oracle:

```
/* Write your PL/SQL query statement below */
```

Pandas:

```
import pandas as pd

def top_percentile_fraud(fraud: pd.DataFrame) -> pd.DataFrame:
```

Solutions

MySQL Solution:

```
# Write your MySQL query statement below
```

MS SQL Server Solution:

```
/* Write your T-SQL query statement below */
```

PostgreSQL Solution:

```
-- Write your PostgreSQL query statement below
```

Oracle Solution:

```
/* Write your PL/SQL query statement below */
```

Pandas Solution:

```
import pandas as pd

def top_percentile_fraud(fraud: pd.DataFrame) -> pd.DataFrame:
```