Problem 1132: Reported Posts II

Problem Information

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

Problem Description

Table:

Actions

+---------------+---------+ | Column Name | Type | +---------------+---------+ | user_id | int | | post_id | int | | action_date | date | | action | enum | | extra | varchar | +---------------+---------+ This table may have duplicate rows. The action column is an ENUM (category) type of ('view', 'like', 'reaction', 'comment', 'report', 'share'). The extra column has optional information about the action, such as a reason for the report or a type of reaction.

Table:

Removals

+---------------+---------+ | Column Name | Type | +---------------+---------+ | post_id | int | | remove_date | date | +---------------+---------+ post_id is the primary key (column with unique values) of this table. Each row in this table indicates that some post was removed due to being reported or as a result of an admin review.

Write a solution to find the average daily percentage of posts that got removed after being reported as spam,

rounded to 2 decimal places

.

The result format is in the following example.

Example 1:

Input:

Actions table: +---------+--------+------------+--------+--------+ | user_id | post_id | action_date | action | extra | +---------+--------+------------+--------+--------+ | 1 | 1 | 2019-07-01 | view | null | | 1 | 1 | 2019-07-01 | like | null | | 1 | 1 | 2019-07-01 | share | null | | 2 | 2 | 2019-07-04 | view | null | | 2 | 2 | 2019-07-04 | report | spam | | 3 | 4 | 2019-07-04 | view | null | | 3 | 4 | 2019-07-04 | report | spam | | 4 | 3 | 2019-07-02 | view | null | | 4 | 3 | 2019-07-02 | report | spam | | 5 | 2 | 2019-07-03 | view | null | | 5 | 2 | 2019-07-03 | report | racism | | 5 | 5 | 2019-07-03 | view | null | | 5 | 5 | 2019-07-03 | report | racism | +---------+--------+------------+--------+--------+ Removals table: +---------+-------------+ | post_id | remove_date | +---------+-------------+ | 2 | 2019-07-20 | | 3 | 2019-07-18 | +---------+-------------+

Output:

+----------------------+ | average_daily_percent | +----------------------+ | 75.00 | +----------------------+

Explanation:

The percentage for 2019-07-04 is 50% because only one post of two spam reported posts were removed. The percentage for 2019-07-02 is 100% because one post was reported as spam and it was removed. The other days had no spam reports so the average is (50 + 100) / 2 = 75% Note that the output is only one number and that we do not care about the remove dates.

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 reported_posts(actions: pd.DataFrame, removals: 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 reported_posts(actions: pd.DataFrame, removals: pd.DataFrame) ->
pd.DataFrame:
```