

Problem Statement:

A **ride-sharing company** wants to analyze the consistency of its drivers. For each driver, calculate the **variance** in trip earnings over the past 6 months. Also, determine how many times the driver's daily earnings deviated by more than **1.5 times** their average daily earnings within this period.

Table: TRIPS

- TRIP_ID (INT) – Unique identifier for each trip
- DRIVER_ID (INT) – ID of the driver
- TRIP_DATE (DATE) – Date of the trip
- EARNINGS (DECIMAL) – Earnings from the trip

Output:

- DRIVER_ID
- EARNINGS_VARIANCE (Variance of daily earnings over 6 months)
- HIGH_DEVIATION_DAYS (Count of days where daily earnings > 1.5 × average daily earnings)

Variance Formula:

The variance of daily earnings is calculated as:

$$\sigma^2 = \frac{\sum (x_i - \bar{x})^2}{n}$$

where:

- x_i = daily earnings
- \bar{x} = average daily earnings
- n = number of days

**QUERY:

```
WITH TOTAL_DAILY AS(
SELECT
    DRIVER_ID,
    TRIP_DATE,
    SUM(EARNINGS) AS DAILY_EARNING
FROM TRIPS
WHERE TRIP_DATE >= CURRENT_DATE - INTERVAL 180 DAY
GROUP BY 1,2
)
```

```
,AVG_DAILY AS(  
SELECT  
    DRIVER_ID,  
    AVG(DAILY_EARNING) AS AVG_EARNING,  
    COUNT(TRIP_DATE) AS N_DAYS  
FROM TOTAL_DAILY  
GROUP BY 1  
)
```

```
SELECT  
    A.DRIVER_ID,  
    SUM(CASE WHEN A.DAILY_EARNING > (1.5*B.AVG_EARNING) THEN 1 ELSE 0) AS  
HIGH_DEVIATION_DAYS,  
    SUM(POWER((A.DAILY_EARNING - B.AVG_EARNING),2))/B.N_DAYS AS EARNINGS_VARIANCE  
FROM  
    TOTAL_DAILY AS A  
INNER JOIN  
    AVG_DAILY AS B  
ON A.DRIVER_ID = B.DRIVER_ID  
GROUP BY 1  
;
```