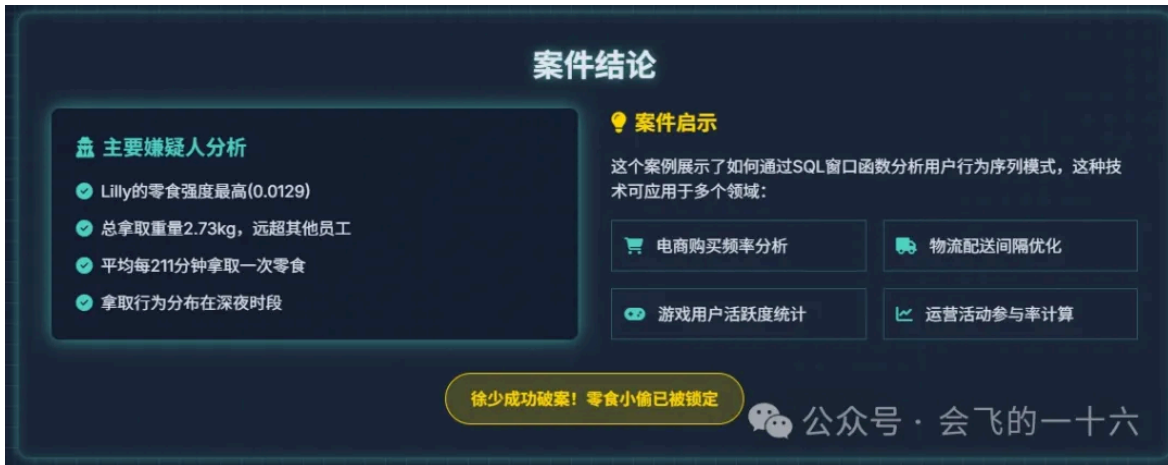
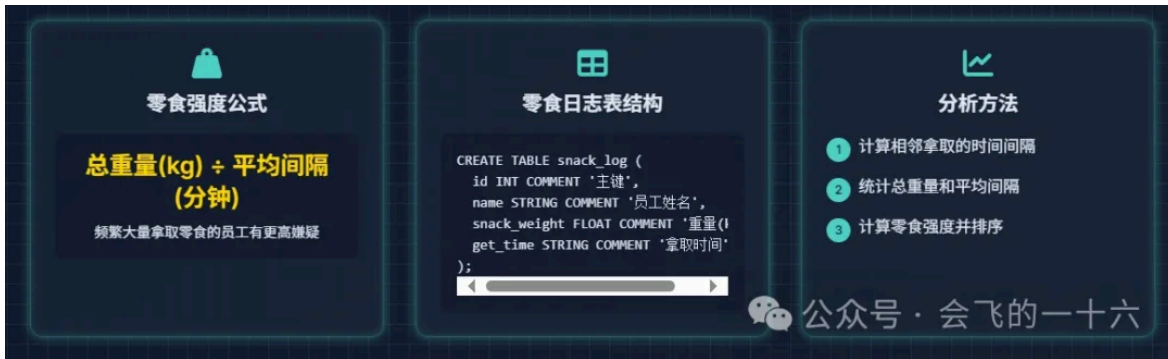


有趣的SQL题 | 零食大侦探：谁偷吃了办公室零食？

原创 江月明203 会飞的一十六 2025年08月07日 08:09 重庆



一、问题背景

徐少去拿零食发现并没有零食，众所周知，徐少的父亲是公司的保卫科科长，今日便是要替父出征，查出到底谁在偷吃零食。聪明的徐少发明了“零食强度”公式，通过该指标来找出最高的嫌疑人。已知零食日志表 `snack_log` 记录了员工拿取零食的时间与重量，需计算每位员工的零食强度（kg/分钟），定义为：

零食强度 = 总拿取重量(kg) / 平均两次拿取间隔时间(分钟)

📖 案件背景

徐少发现办公室零食神秘消失，作为保卫科科长之子，他决心查明真相。通过分析零食日志数据，他发明了“零食强度”指标公式：

$$\text{零食强度} = \Sigma \text{拿取重量(kg)} / \text{平均间隔时间(分钟)}$$

通过对每个员工的零食拿取行为进行分析，找出最可疑的“零食大盗”。

📁 数据表结构 (snack_log)

```
id INT 主键
name STRING 员工姓名
snack_weight FLOAT 拿取重量(kg)
get_time STRING 拿取时间
```

👤 公众号 · 会飞的一十六

二、问题分析

1. 核心指标定义

徐少发明的“零食强度”公式是：

定义：零食强度 = 总拿取重量 (kg) / 相邻拿取的平均间隔 (分钟)

$$\text{零食强度} = \frac{\sum (\text{snack_weight})}{\text{AVG}(\text{interval_mins})}$$

👤 公众号 · 会飞的一十六

零食强度 = 总拿取重量 (kg) / 相邻拿取的平均间隔 (分钟)



总重量越大： 拿得越多，嫌疑越大。



平均间隔越短： 拿得越频繁，嫌疑越大。

👤 公众号 · 会飞的一十六

指标逻辑很直观：总重量越大、间隔越短，越可能是偷吃者（比如频繁拿且拿得多的人）。

2. 问题拆解

要计算“零食强度”，需要解决三个子问题：

1. 计算每个员工相邻两次拿取零食的时间间隔（分钟）；
2. 统计每个员工的总拿取重量和平均间隔时间；
3. 用“总重量÷平均间隔”计算强度，按强度降序排序（找出Top1）。

- 1

计算时间间隔

相邻两次拿取零食的时间差（分钟）
- 2

统计核心数据

每位员工的总拿取重量和平均间隔时间
- 3

计算强度排序

用"总重量÷平均间隔"计算强度，按强度降序找出Top1
- 公众号 · 会飞的一十六

三、表结构与示例数据

员工拿取零食的日志存在 `snack_log` 表中，结构如下（Hive DDL）：

```
CREATETABLE IF NOT EXISTS snack_log (  
  id INT COMMENT '主键，自增',  
  name STRING COMMENT '员工姓名',  
  snack_weight FLOAT COMMENT '拿取重量 (kg) ',  
  get_time STRING COMMENT '拿取时间 (YYYY-MM-DD HH:MM:SS) '  
)  
COMMENT '员工零食拿取日志表'  
ROW FORMAT DELIMITED  
FIELDS TERMINATED BY'\t'  
STORED AS TEXTFILE;
```

字段名	类型	描述
id	INT	主键，自增
name	STRING	员工姓名
snack_weight	FLOAT	拿取重量(kg)
get_time	STRING	拿取时间

兑换合集后可阅读剩余71%

#SQL面试精要

已完结 共28个

1. SQL面试提问：如何找出支付金额在前 20% 的用户？
2. SQL高频面试题 | 如何查询用户的累计消费金额及VIP等级？

合集详情

3. SQL面试提问：如何找出每月GMV环比下降超20%的城市？

[SQL面试精要 · 目录](#) ≡

[< 上一篇 · 面试提问：数据开发中如何通过指标拆解来编写正确的SQL？（附拆解模板）](#)

个人观点，仅供参考