sql读书笔记

——康珣

目录

[第九章 1](#_Toc459711951)

[9.1时间的本质 2](#_Toc459711952)

[9.2半开放时间模型 2](#_Toc459711953)

[9.2.1 用null作占位符 3](#_Toc459711954)

[9.2.2 单时间戳表 3](#_Toc459711955)

[9.2.3重叠的时间间隔 4](#_Toc459711956)

[9.4例题 5](#_Toc459711957)

# 第九章

Sql是具有时间类型的编程语言。然而每个sql产品都有自己独特的时间类型，这可能引发问题。标准sql认为数据库都是基于UTC（世界标准时间），然后转换为本地时间来实现。

# 9.1时间的本质

在数据处理中，大多数数据是离散的。但是时间却是个连续量。

时间子是一个量子概念，最早出现于20世界80年代的量子力学中，是亚原子物理学的准确定义。为了将时间这个连续性变量转变为离散变量，数据库中时间类型的研究人员，将这个术语用做时间的基本单位，构建以时间子为步长的时间模型。

以时间子构建时间模型会出现三个问题：

1. 一个时间子大约为2X10-23秒，大多数计算机硬件无法处理这么精细的值。
2. 数据不符合第一范式。
3. 离散数据必须有顺序，且不能有空隙，{d1、d2、d3、d4};但集合是无序的，

且集合自身的子集又是一个集合。

例子：x轴表示交货时间，y轴表示零件组。每个零件组都有交货日期。若用时间子作为基本单位表示的话，因为一天可以用n倍的时间子表示。如：零件1在5月1号交付，而坐标轴上的5月1号就是一段阴影区。会出现一个问题，5月1号内任何一个时间点都是5月1号，那么零件1究竟在哪个时间点交付呢？

5月1号

交货时间（/时间子）

零件组

# 9.2半开放时间模型

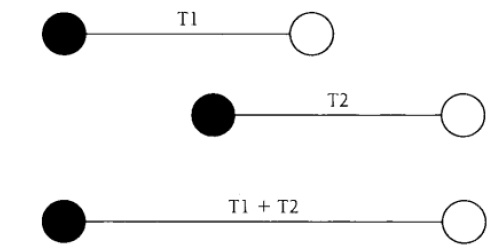
时间显示格式很重要，但是时间段基础模型更重要。

如：2006-12-31就是一个2006-12-31 00:00:00到2006-12-31 23:59:591999的时间段缩写。

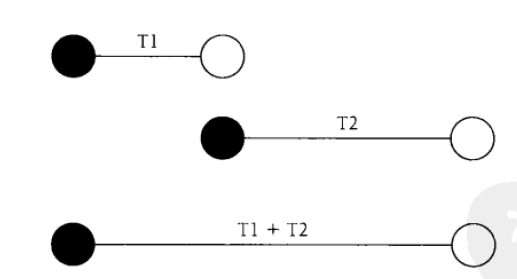
半开放时间模型就是左闭右开区间，实心点表示某点在时间线上，空心点表示某点不在时间线上。

不论对半开放时间模型进行怎样操作，仍然会得到半开放时间模型。原因如下：

1. 两个重叠的半开放时间间隔形成一个半开放时间间隔。

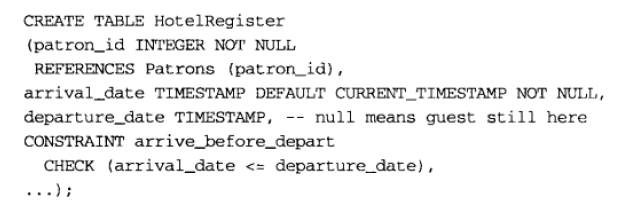


1. 如果从一个半开放时间间隔中删除另一个半开放时间间隔，仍会得到开放时间间隔（一个或两个）
2. 两个连续的半开放时间间隔形成一个半开放时间间隔。（开放端不会恰好连到下一个时间间隔的起点，约定当时间间隔差距小于某个△值，可以认为表示时间点相同，即前一个结束时间接上了下一个开始时间）。



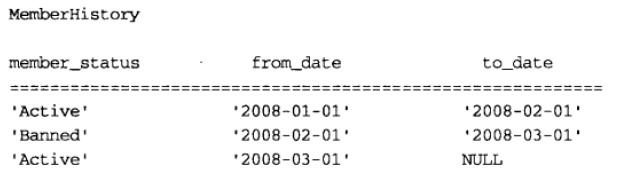
## 9.2.1 用null作占位符

业务：有人登记一家宾馆时，我们可以知道到达时间，却无法知道客户的离开时间。但是为了标记时间段，数据库建表中会有到达和离开时间，2个属性。我们可以先将离开时间标记为null，而当我们获得客户实际离开日期时，将null替换为离开日期。



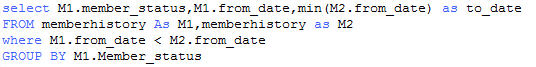
## 9.2.2 单时间戳表

业务：一个健身房的成员，总处于以下的状态之一，（active、dormant、cancelled、banned）。当他加入健身房(1月1号)为active，（2月1号）当天损坏健身设备，该人被健身房标记dormant，（3月1号）恢复active。

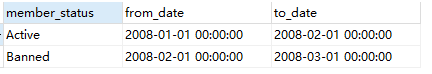


如果去除to\_date属性，若想查找此人2008年1月4日是什么状态，需要将表做自连接才能得到结果。

Sql语句如下：



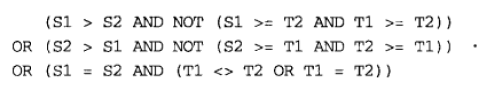
Result：



因此，需要数据对，即时间对方式表示时间段会更方便，表达方式更简单。

## 9.2.3重叠的时间间隔

OverLaps（）表示时间段比较，若重叠返回TRUE，否则返回FALSE。备注：S开始时间，T结束时间。



分别对应

T1

S2

S1

T2

1. 2.

T2

S1

S2

T1

T1

S1

S1

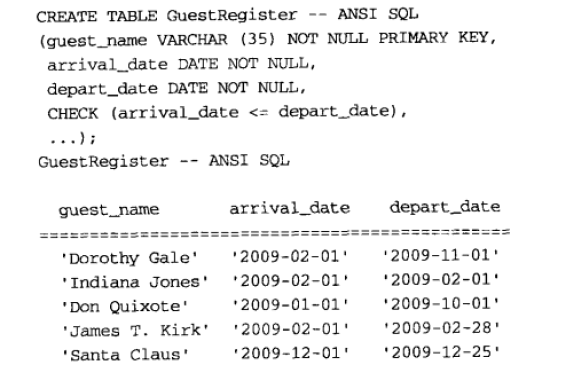
T2

制定标准规则：

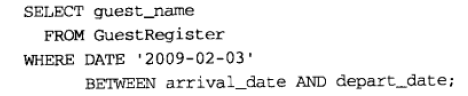
1. 时间段为半开放模型
2. 如果时间不是瞬时（时间点），而是时间段，S1-T1 与 S2与T2公共部分就会重叠。
3. 如果时间点与时间段出现重叠，那么时间点一定包含在时间段中。
4. 若第一项和第二项都为时间点，出现重叠的条件为两者相等。
5. 如果开始时间为null，结束时间为点时间，结束时间就是开始时间。如果开始时间为null，已知经过一个时间段（如2天）才结束，因为开始时间未知，虽然知道时间段，结束时间也未知，标记为null。

## 9.4例题

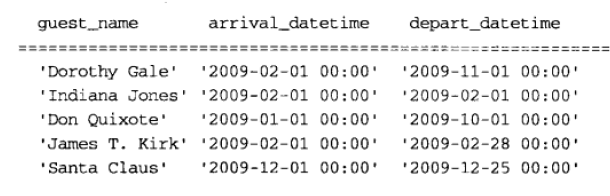
业务：旅店客人表，记录着客人居住时间，时间间隔为天。表如下。



1. 查询某个时间在旅馆的所有客人（使用between and）

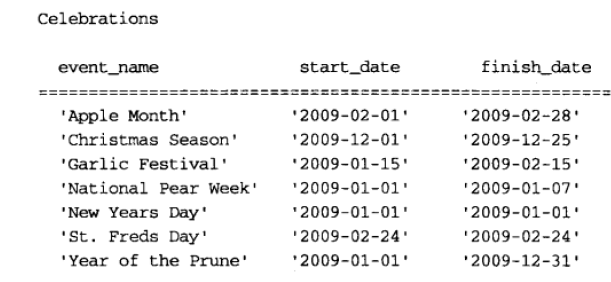


注意：如果数据库为sql server或sybase，时间类型只能为timestamp，原来2009-02-01会添加上时分表示，变为2009-02-01 00:00

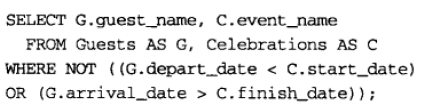


若客人结账时间为2009-02-01 13:00，我们在表中使用between and会失败，没有找出结果，因为原来2009-02-01符合第二行，但是因为timestamp原因，第二行时分变为00:00时分，不再符合between条件。

1. 若想知道旅客在节假日的居住情况。业务中给出了Celebration表。



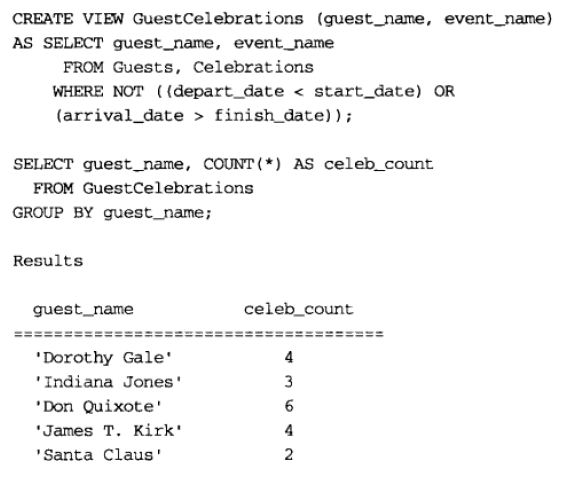
需要将Celebration表和Guests表做笛卡尔积。约束条件为客户居住时间和节假日时间重叠时，sql语句where条件分别表示（客人离开后节假日才开始）（节假日结束后客人才入住）这两种情况的否定。



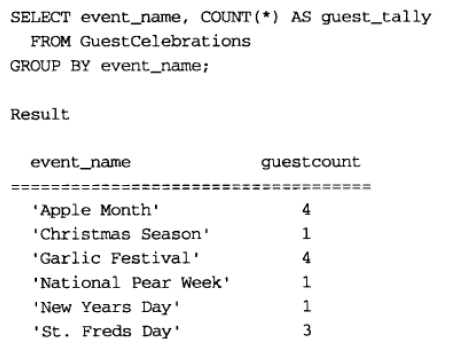
1. 为了查询每个客人观看多少庆祝活动，或者每个庆祝活动中各有多少客

人。可以创建视图（客户在节假日居住）来查询。

（1）



（2）



1. 查询庆祝活动中销售的房间宿数。（每个庆祝活动期间，所有客户居住时

间的总和）

节假日时间长度：Event.finish\_date – Event.start\_date + Interval

‘1’day

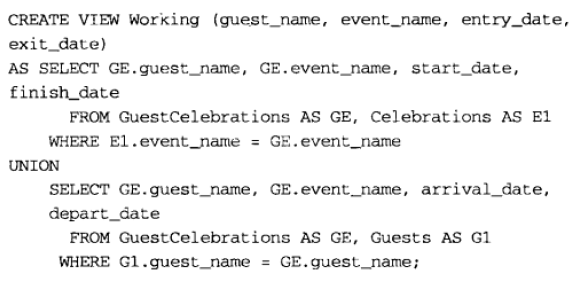
客人居住时间长度：Guest.depart\_date – Guest.arrival\_date

+Interval’1’Day

备注：Interval表示时间间隔。Interval’1’Day表示时间间隔为一天。

1. 将客户在节假日居住情况表分别与Guest表和Celebration表做连

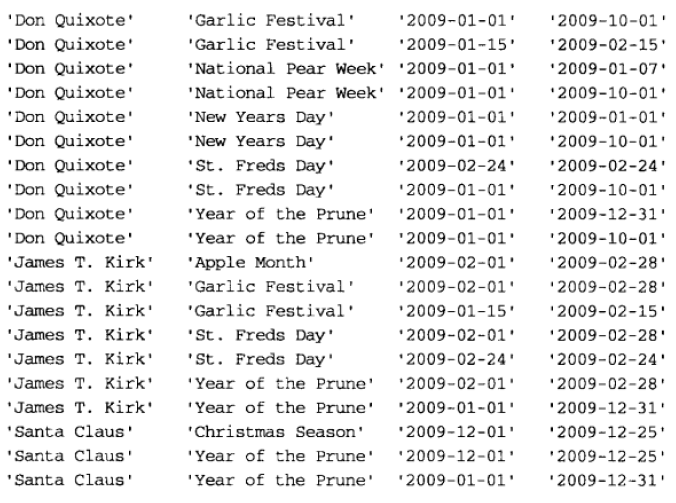
接，从中选出满足1、节假日期间居住客人的居住时间、离开时间2、有客人居住的节假日的开始时间、结束时间。



Result：

看Dorothy，共有8行记录。第1、3、5、7行表示的是该人在四个节假日期间的居住时间。第2、4、6、8行表示四个节假日的开始时间与结束时间。

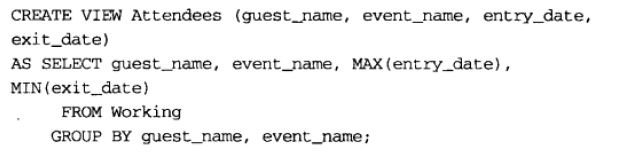




（2）为了计算客人在节假日期间的居住时间。如：8月1号入住，节假日2号才开始；8月20号退房，8月10号节假日结束。因此节假日期间的居住时间为8月2号—8月10号。

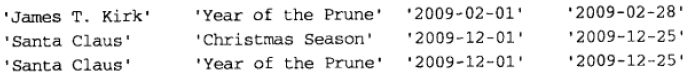
C:\Users\sherl\AppData\Roaming\Tencent\Users\1070746603\QQ\WinTemp\RichOle\T78%ZHZ2}SH%NN%AZXZ2]YV.png

计算公式如上

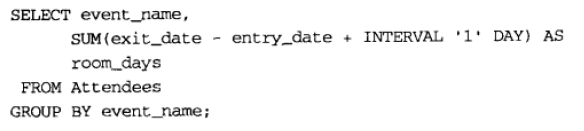


Result：





（3）每个节假日期间出售的房间宿数。（只需要将节假日分组，查询每个节假日所有客户的居住时间总和即可）



5．查询一月销售销售的房间宿数。

Where条件为了找出客户居住时间中与（1月1日至1月31日）存在重叠的居住时间。Case是将小于1月1日部分，大于1月31日部分做处理。

如：2008-12-30 – 2009-01-20，与（1月1日至1月31日）存在重叠，因为计算的是1月销售额，2008-12-30改为1月1日开始，2009-01-20不变。一月该客户居住时间为1月1日到1月20日。

