

中级SQL

约瑟夫·哈利特

2023 年 2 月 13 日



上次…

我们引入 SQL 作为查询数据库的语言

- ▶ 如何创建表格
- ▶ 如何添加和删除数据
- ▶ 如何运行基本查询

这次…

更高级的SQL! ▶ 更多特性,更

多功能 ▶ 其他连接 ▶ NULL

NULL 就是什么都没有

SQL中有一个特殊的值来表示缺失的数据:NULL。▶但它们几乎总是
一个坏主意▶比较它们的逻辑非常古怪

空=空？

假设我们有一个包含下表的数据库：

人	果
约瑟夫·莱姆	
马特	苹果
帕塔	

让我们找到每个我们知道他们最喜欢的水果是什么的人！

```
SELECT * FROM 水果 WHERE 水果 <> NULL;
```

呃……，我们试试相反的方法？

```
SELECT * FROM 水果 WHERE 水果 = NULL;
```

呃什么？

从水果中选择*，其中水果像“%”；

人	果
约瑟夫·莱姆	
马特	苹果

所以…

从水果中选择*，其中水果不喜欢“%”；

NULL 很奇怪...

因为 NULL 意味着属性缺失... ►与它比

较的结果只是简单的愚蠢有点出乎意料简单的解决方案是将所有内容声明为 NOT NULL

►使用更高范式 (5NF)那么你会发现它们几乎完全消失

否则你必须记住一堆愚蠢的特殊比较器

SELECT * FROM 水果 WHERE 水果为空;

人果
帕塔

从水果中选择*,其中水果不为空;

人果
约瑟夫·莱姆
马特 苹果

棘手的加入

当 NULL 出现问题时,明确测试相等性。

►那么当您想要将两个表中包含 NULL 的表连接在一起时会发生什么

人果

约瑟夫·莱姆

马特

苹果

帕塔

水果

盘子

苹果 苹果崩溃了

香蕉 香蕉分割

樱桃

酸橙

代基里酒

我最喜欢的食物是什么？

那么对于你们的每位讲师来说,什么才是一道美味佳肴呢？

►（自然连接类似于常规连接,但假设相同命名的列应该相等）。

人水果盘

约瑟夫·莱姆代基里酒

马特

苹果 苹果崩溃了

但可怜的帕塔呢?我们如何让他出现在我们的表中？

左连接和右连接

在执行之前的 JOIN 操作时,我们只需要匹配的行.....

►技术上称为 INNER JOIN...

有时,如果我们想保留无法进行连接的列,数据库可以将 NULL 保留在其中.....

从水果中选择人、水果.水果、菜肴 LEFT JOIN 食

谱 ON 水果.水

果 = 食谱.水果;

人水果盘

约瑟夫·莱姆代基里酒

马特

苹果 苹果崩溃了

帕塔

右连接

右连接类似于左连接,但反过来.....

从水果中选择水果.水果,菜肴,人,右键加入水果
上的食谱。水果
=食谱.水果;

水果盘
青柠代基里酒
苹果苹果崩溃马特
香蕉船

人
约瑟夫

香蕉去哪儿了?!

从水果中选择食谱.水果,菜肴,人,右键加入水果上的
食谱.水果=食
谱.水果;

水果 盘子
酸橙 代基里酒
苹果苹果崩溃马特
香蕉 香蕉分割
樱桃

人
约瑟夫

(或者只是自然连接,它通常会处理它.....)

选择水果、菜肴、人物
来自水果
正确的自然加入食谱;

水果 盘子
酸橙 代基里酒
苹果 苹果碎马特
香蕉 香蕉分割
樱桃

人
约瑟夫

又一个加入!

如果我们想同时进行 LEFT 和 RIGHT JOIN 怎么办?

选择 *

来自水果

完整的外部自然连接食谱;

人果

盘子

约瑟夫·莱姆

代基里酒

马特

苹果 苹果崩溃了

帕塔

香蕉 香蕉分割

樱桃

统计函数呢？

在上一讲中,我们介绍了 COUNT 作为一种计算存在多少事物的方法？

►外连接表中可能有哪些不同的水果？

选择 *

来自水果

完整的外部自然连接食谱；

人果

约瑟夫·莱姆

马特

帕塔

苹果

香蕉 香蕉分割

樱桃

盘子

代基里酒

苹果弄碎

选择数量 (水果)

来自水果

完整的外部自然连接食谱

计数 (水果)4

…所以看起来 COUNT 忽略了 NULL

其他统计数据...

来给水果排名吧！

水果	星星
苹果	0
香蕉	4
樱桃	
酸橙	5

```
SELECT AVG(stars) AS Average FROM 排名;
```

平均3.0

```
选择 SUM(星星)/COUNT(水果) AS 平均值  
来自排名;
```

平均 2

记住电脑很糟糕

►将计数乘以 1.0 来“修复”？ ►另外,星星
的数量是序数数据,所以无论如何都不应该使用平均值.....

标准差呢？

标准差是指某事物平均偏离平均值的程度。

选择 SQRT(AVG(偏差)) AS STDDEV

从 (

选择水果、星星、平均值、

(星星-平均值)* (星星-平均值)AS 偏差

FROM排名加入 (

选择 AVG (星星)作为平均值

来自排名

)

WHERE 星星不为空

);

标准差

2.16024689946929

您可以将查询嵌套在另一个查询中（子查询！）▶这是让您的 SQL变慢的

秘诀▶也许只使用 SQL 进行数据检索,而将复杂的统计数据

留给统计编程

语言？

这就是 SQL!

使用技巧? ▶不

要把事情搞得太复杂! ▶范式

让事情变得更简单! ▶像躲避瘟疫一
样避免 NULL