**必知必会的窗口函数二：课后习题窗口函数**

这次的文章我们来看一看，《SQL进阶教程》第二章节的课后习题。在学习的过程中我们将对某一些案例使用Pandas来重写和复现。这样就能对窗口函数有一个更为深度的认识。

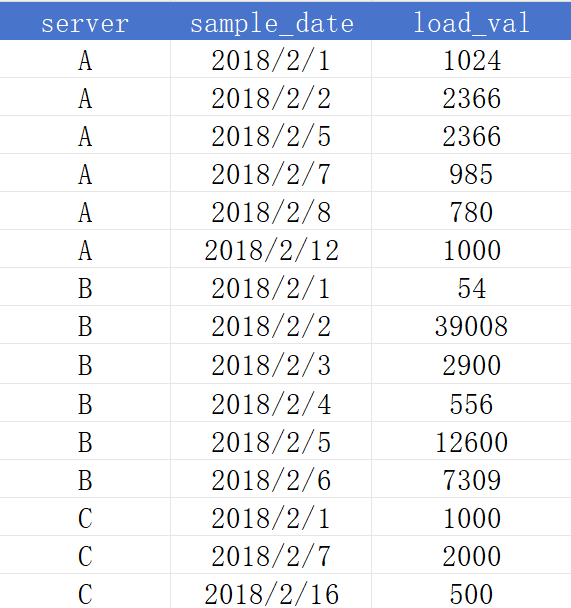
在上次的文章中，我们了解了窗口函数的基本知识，以及如何进行数据“移动”。具体文章可以点击以下链接：

**必知必会的窗口函数一\_基础知识**

本次的文章我们来学习一下课后的练习题。

**一、练习题1-2-1：窗口函数的结果预测（1）**

现在假设有表ServerLoadSample的数据如下所示，其数据如下：



创建该表的语句如下所示：

**CREATE** **TABLE** ServerLoadSample

(

**server** **varchar**(10),

sample\_date **date** **NOT** **NULL**,

load\_val **INTEGER** **NOT** **NULL**

);

**INSERT** **INTO** ServerLoadSample **VALUES**('A', **date**'2018-02-01',1024);

**INSERT** **INTO** ServerLoadSample **VALUES**('A', **date**'2018-02-02',2366);

**INSERT** **INTO** ServerLoadSample **VALUES**('A', **date**'2018-02-05',2366);

**INSERT** **INTO** ServerLoadSample **VALUES**('A', **date**'2018-02-07',985);

**INSERT** **INTO** ServerLoadSample **VALUES**('A', **date**'2018-02-08',780);

**INSERT** **INTO** ServerLoadSample **VALUES**('A', **date**'2018-02-12',1000);

**INSERT** **INTO** ServerLoadSample **VALUES**('B', **date**'2018-02-01',54);

**INSERT** **INTO** ServerLoadSample **VALUES**('B', **date**'2018-02-02',39008);

**INSERT** **INTO** ServerLoadSample **VALUES**('B', **date**'2018-02-03',2900);

**INSERT** **INTO** ServerLoadSample **VALUES**('B', **date**'2018-02-04',556);

**INSERT** **INTO** ServerLoadSample **VALUES**('B', **date**'2018-02-05',12600);

**INSERT** **INTO** ServerLoadSample **VALUES**('B', **date**'2018-02-06',7309);

**INSERT** **INTO** ServerLoadSample **VALUES**('C', **date**'2018-02-01',1000);

**INSERT** **INTO** ServerLoadSample **VALUES**('C', **date**'2018-02-07',2000);

**INSERT** **INTO** ServerLoadSample **VALUES**('C', **date**'2018-02-16',500);

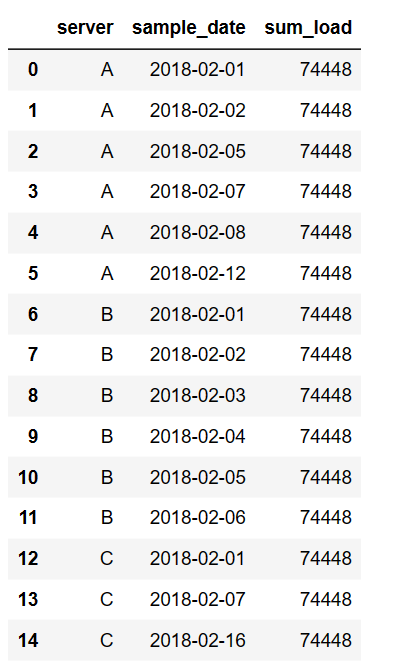
那么对于语句：

**select** **server**,sample\_date,

**sum**(load\_val) **over**() **as** sum\_load

**from** serverloadsample

其输出是怎样的呢。经过运行其输出为：



可以看出，结果集中server和sample\_date的数据样式是没有变化的。只不过sum\_load都是一样的数据，且我们可以发现sum\_load的值正好是load\_val列所有值的累加。可以看出在没有使用PARTITION BY和ORDER BY以及帧字句的时候，**此时OVER()作用的数据范围是整个表**，因此我们会看到预期的结果。

**二、练习题1-2-1：窗口函数的结果预测（2）**

对于表ServerLoadSample继续执行下述语句：

**select** **server**,sample\_date,

**sum**(load\_val) **over**(**partition** **by** **server**) **as** sum\_load

**from** serverloadsample

结果输出又是怎样的呢？输出如下：



上述语句就是一个正常的包含PARITION BY的语句的操作了，就是按照sever字段进行分组，可以知道分为A，B，C三组，然后分别对这三组中的load\_val进行求和，最后得到预期的输出

上述两个案例还是比较简单的，这里就不在进行扩张了，这里也不在使用Pandas实现了，大家自己来实现哈。

**三、参考文献**

1. MICK-《SQL进阶教程》第2版
2. Joe Celko-《SQL权威指南》