

Hive数据类型转换

写在前面的话

，学Hive这么久了，发现目前国内还没有一本完整的介绍Hive的书籍，而且互联网上面的资料很乱，于是我决定写一些关于[《Hive的那些事》](#)序列文章，分享给大家。我会在接下来的时间整理有关Hive的资料，如果对Hive的东西感兴趣，请关注本博客。/archives/tag/hive的那些事

在[《Hive内置数据类型》](#)

文章中，我们提到了Hive内置数据类型由基本数据类型和复杂数据类型组成。今天的话题是Hive数据类型之间的转换。同Java语言一样，Hive也包括隐式转换（implicit conversions）和显式转换（explicitly conversions）。

Hive在需要的时候将会对numeric类型的数据进行隐式转换。比如我们对两个不同数据类型的数字进行比较，假如一个数据类型是INT型，另一个是SMALLINT类型，那么SMALLINT类型的数据将会被隐式转换地转换为INT类型，这个到底和Java中的一样；但是我们不能隐式地将一个INT类型的数据转换成SMALLINT或TINYINT类型的数据，这将会返回错误，除非你使用了CAST操作。

任何整数类型都可以隐式地转换成一个范围更大的类型。TINYINT,SMALLINT,INT,BIGINT,FLOAT和STRING都可以隐式地转换成DOUBLE；是的你没看出，STRING也可以隐式地转换成DOUBLE！但是你要记住，BOOLEAN类型不能转换为其他任何数据类型！

下表列出了Hive内置的数据类型之间是否可以进行隐式的转换操作：

	bl	tinyint	si	int	bigint	float	double	dm	string	vc	ts	date	ba
boolean	true	false	false	false	false	false	false	false	false	false	false	false	false
tinyint	false	true	true	true	true	true	true	true	true	true	false	false	false
smallint	false	false	true	true	true	true	true	true	true	true	false	false	false
int	false	false	false	true	true	true	true	true	true	true	false	false	false
bigint	false	false	false	false	true	true	true	true	true	true	false	false	false
float	false	false	false	false	false	true	true	true	true	true	false	false	false
double	false	false	false	false	false	false	true	true	true	true	false	false	false

ble													
decim al	false	false	false	false	false	false	false	true	true	true	false	false	false
strin g	false	false	false	false	false	false	true	true	true	true	false	false	false
varc har	false	false	false	false	false	false	true	true	true	true	false	false	false
ts	false	false	false	false	false	false	false	false	true	true	true	false	false
date	false	false	false	false	false	false	false	false	true	true	false	true	false
bina ry	false	false	false	false	false	false	false	false	false	false	false	false	true

注：由于表格比较大，这里对一些比较长的字符串进行缩写，ts是timestamp的缩写,bl是boolean的缩写,sl是smallint的缩写,dm是decimal的缩写,vc是varchar的缩写,ba是binary的缩写。

我们可以用CAST来显式的将一个类型的数据转换成另一个数据类型。如何使用？CAST的语法为cast(value AS TYPE)。举个例子：假如我们一个员工表employees，其中有name、salary等字段；salary是字符串类型的。有如下的查询：

```
SELECT name, salary FROM employees
WHERE cast(salary AS FLOAT) < 100000.0;
```

这样salary将会显示的转换成float。如果salary是不能转换成float，这时候cast将会返回NULL！对cast有以下几点需要说明的：

- 如果将浮点型的数据转换成int类型的，内部操作是通过round()或者floor()函数来实现的，而不是通过cast实现！
- 对于BINARY类型的数据，只能将BINARY类型的数据转换成STRING类型。如果你确信BINARY类型数据是一个数字类型(a number)，这时候你可以利用嵌套的cast操作，比如a是一个BINARY，且它是一个数字类型，那么你可以用下面的查询：

```
SELECT (cast(cast(a as string) as double)) from src;
```

我们也可以将一个String类型的数据转换成BINARY类型。

- 对于Date类型的数据，只能在Date、Timestamp以及String之间进行转换。下表将进行详

细的说明：

有效的转换	结果
cast(date as date)	返回date类型
cast(timestamp as date)	timestamp中的年/月/日的值是依赖与当地的时区，结果返回date类型
cast(string as date)	如果string是YYYY-MM-DD格式的，则相应的年/月/日的date类型的数据将会返回；但如果string不是YYYY-MM-DD格式的，结果则会返回NULL。
cast(date as timestamp)	基于当地的时区，生成一个对应date的年/月/日的时间戳值
cast(date as string)	date所代表的年/月/日时间将会转换成YYYY-MM-DD的字符串。

本博客文章除特别声明，全部都是原创！
原创文章版权归过往记忆大数据（[过往记忆](#)）所有，未经许可不得转载。
本文链接：[【】（）](#)