**Pyspark基本语法（同盾DC）—使用udf编写sql语言**

————20190812花名：韵方

首先我们先进行导包：



Lit包是干嘛的我们先不管，udf是我们这次要用的包

然后我们来随便定义一个函数：



函数基本功能就是传进一个参数，把它转换成string形式的再在最后加上很多加号，现在我们要做的就是将这个函数通过udf使用到sql语言中去

首先我们来讲第一种写法，使此函数可以直接写入sql语言当中，就相当于我们自己定义了一种sql中的函数，直接上代码吧：



首先是类似java中方法实例化的一步，直接add1方法实例化命名为add\_spk，然后我们就可以在sql语言中使用它了：



这里我们在这个表后新加了名为product\_id\_c的一列，值为product\_id的值加上很多加号，现在我们将他打印出来，这里随便选了一列：



现在我们来看第二种写法，此写法同样也需要先将方法实例化，不过这种实例化不能使函数直接在sql语言中编写：



然后我们就不能直接在sql语言中写入此方法了，我们需要调用udf的方法：



这是一条查找语句，基本功能就是在上述表中再加名为produce\_id\_d的一列，内容与product\_id\_c的一列相同，都是调用add1的函数，我们将它打印出来

