

在Hive中，用户可以自定义一些函数，用于扩展HiveQL的功能。Hive 自定义函数主要包含以下三种：

- UDF (user-defined function) 单独处理一行，输出也是以行输出。许多Hive内置字符串，数学函数，时间函数都是这种类型。大多数情况下编写对应功能的处理函数就能满足需求。如：concat, split, length, rand等。这种UDF主要有两种写法：继承实现UDF类和继承GenericUDF类（通用UDF）。
- UDAF (user-defined aggregate function) 用于处理多行数据并形成累加结果。一般配合group by使用。主要用于累加操作，常见的函数有max，min，count，sum，collect_set等。这种UDF主要有两种写法：继承实现UDAF类和继承实现AbstractGenericUDAFResolver类。
- UDTF (user-defined table function) 处理一行数据产生多行数据或者将一行打成多列。如explode, 通常配合Lateral View使用，实现列转行的功能。parse_url_tuple将一行转为多列。

Hive的UDF机制是需要用户实现: Resolver和Evaluator，其中Resolver就用来处理输入，调用Evaluator，Evaluator就是具体功能的实现。