

hive处理json和数组数据

```
-- 处理json数据
create table tmp.rca_t_module_his_1 as
select
get_json_object(moduleinputstr,'$.user_id') as user_id ,
get_json_object(moduleinputstr,'$.borrow_nid') as borrow_nid ,
get_json_object(moduleinputstr,'$.realname') as realname ,
get_json_object(moduleinputstr,'$.addproducttype') as addproducttype ,
get_json_object(engineoutputstr,'$.data.cy_d_middle_result') as scorer_id --多层嵌套
from es_rule.rca_t_module_his
where modulefieldname in
('cy_new_admit_1','cy_new_admit_2_1','cy_new_admit_2_2','cy_new_admit_2_3','cy_new_admit_3','cy_old_admit_3')
limit 10

-- 其实不是数组类型，用字符串处理 select instr('["1131","1183","1705"]','1131')
select scorer_id from tmp.rca_t_module_his_1 where instr(scorer_id,'1131')>0 limit 10 ;

-- 一行转多行 用 LATERAL VIEW explode
drop table tmp.rca_t_module_his_2
create table tmp.rca_t_module_his_2 as
select user_id, borrow_nid,addproducttype,realname ,sid
from tmp.rca_t_module_his_1
LATERAL VIEW explode(split(regexp_replace(scorer_id,"[^,0-9a-zA-Z]",","),',')) t as sid ;

select * from tmp.rca_t_module_his_2 limit 10
```