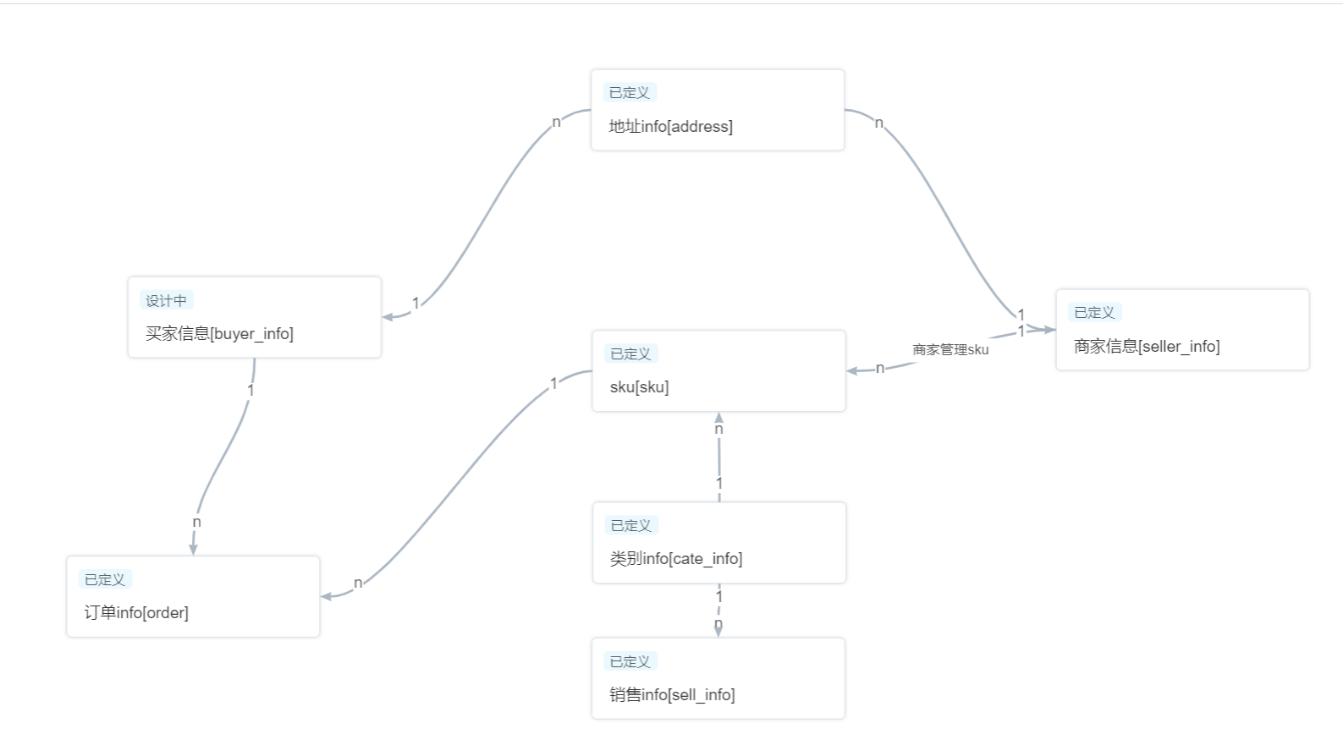
本周的作业主要内容包括了电商ER图创建和数据库建表、插入数据、执行sql查询。之后，也通过bingo平台对数据进行了数据检测。

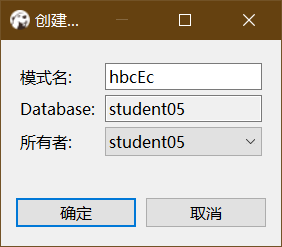
相关sql语句已经在文件夹中保存，可通过执行sql语句来分别创建数据库，插入数据和sql查询。

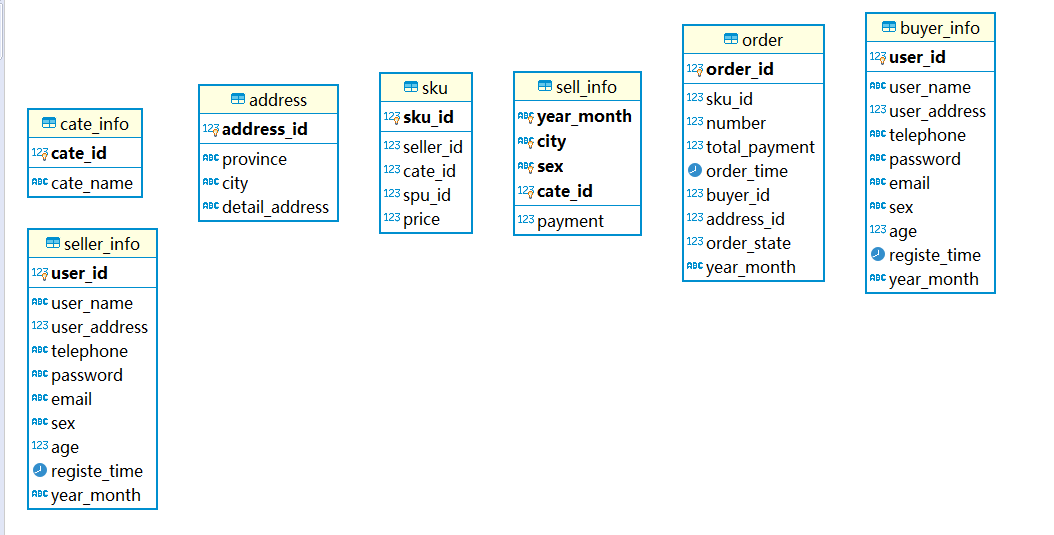
1. **物理模型创建**



我使用了bingo的物理建模功能，尝试在平台上进行ER图设计，并设计字段。之后连接我创建的数据源，进行物化。

创建数据源时，参数需要填写schema，而这个schema需要手动创建，否则用实操的schema的话，表会混杂在一起。

（通过DBeaver创建schema）

（具体字段）

1. **执行sql查询**

执行查询前，表内得有数据才能测试是否查询正确，于是我创建了几条数据插入进数据库。

--插入一些数据进去

--地址

insert into "hbcEc2".address  values (0,'广东省','广州市','番禺');

insert into "hbcEc2".address  values (1,'广东省','广州市','黄埔');

insert into "hbcEc2".address  values (2,'广东省','东莞市','麻涌');

insert into "hbcEc2".address  values (3,'广东省','深圳市','鹅城');

--买家

insert into "hbcEc2".buyer\_info values (0,'userA',0,'13712345678','password1','abcdefg@qq.com','1',21,to\_date('20210702','yyyymmdd'),'2021-07');

insert into "hbcEc2".buyer\_info values (1,'userB',1,'13712345678','password1','abcdefg@qq.com','0',21,to\_date('20210702','yyyymmdd'),'2021-07');

insert into "hbcEc2".buyer\_info values (2,'userC',2,'13712345678','password1','abcdefg@qq.com','1',21,to\_date('20210602','yyyymmdd'),'2021-06');

insert into "hbcEc2".buyer\_info values (3,'userD',3,'13712345678','password1','abcdefg@qq.com','1',21,to\_date('20210502','yyyymmdd'),'2021-05');

--商家

insert into "hbcEc2".seller\_info values (0,'seller0',3,'13712345678','password1','abcdefg@qq.com','1',21,to\_date('20210502','yyyymmdd'),'2021-05');

--商品类别

insert  into "hbcEc2".cate\_info values(0,'手机');

insert  into "hbcEc2".cate\_info values(1,'电脑');

insert  into "hbcEc2".cate\_info values(2,'鞋子');

--商品

insert  into "hbcEc2".sku values(0,0,0,0,100);

insert  into "hbcEc2".sku values(1,0,0,1,200);

insert  into "hbcEc2".sku values(2,0,1,0,150);

insert  into "hbcEc2".sku values(3,0,2,1,100);

--订单

insert  into "hbcEc2"."order"  values(0,0,1,100,to\_date('20210102','yyyymmdd'),0,0,1,'2021-01');

insert  into "hbcEc2"."order"  values(1,1,1,100,to\_date('20210102','yyyymmdd'),0,0,1,'2021-01');

insert  into "hbcEc2"."order"  values(2,1,1,100,to\_date('20210202','yyyymmdd'),1,1,1,'2021-02');

insert  into "hbcEc2"."order"  values(3,0,1,100,to\_date('20210302','yyyymmdd'),3,3,1,'2021-03');

insert  into "hbcEc2"."order"  values(4,3,1,100,to\_date('20210402','yyyymmdd'),2,2,1,'2021-04');

insert  into "hbcEc2"."order"  values(5,3,1,100,to\_date('20210102','yyyymmdd'),0,0,1,'2021-01');

insert  into "hbcEc2"."order"  values(6,0,1,100,to\_date('20210302','yyyymmdd'),3,1,1,'2021-03');

insert  into "hbcEc2"."order"  values(7,0,1,100,to\_date('20210102','yyyymmdd'),1,1,1,'2021-01');

插入完成后，进行sql查询。有以下几个查询：

1. 按月查看总体销售额，新增注册人数：

Sql语句：

--按月份统计新注册用户人数

select year\_month ,sum("人数") "注册人数" from(

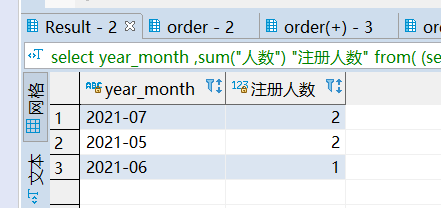
    (select year\_month , count(year\_month ) "人数" from "hbcEc2".buyer\_info bi group by year\_month )

    union all

    (select year\_month , count(year\_month ) "人数" from "hbcEc2".seller\_info si group by year\_month)

) a group  by year\_month ;

执行结果：



Sql语句：

--按月份查看总体销售额

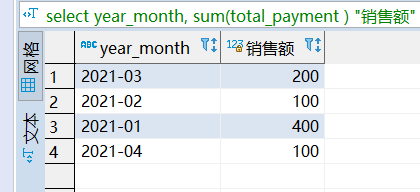
select year\_month, sum(total\_payment ) "销售额"

from "hbcEc2"."order" o

where order\_state = 1

group by year\_month;

执行结果：



1. 按月查看城市的销售额（收货城市）

Sql语句：

--按月和城市看销售额

select year\_month , a.city , sum(total\_payment ) "销售额"

from "hbcEc2"."order" o , "hbcEc2".address a

where o.address\_id = a.address\_id

group by year\_month ,a.city;

执行结果：



1. 按月查看城市、商品类别的销量额

Sql语句：

--按月查看城市、商品类别的销量额

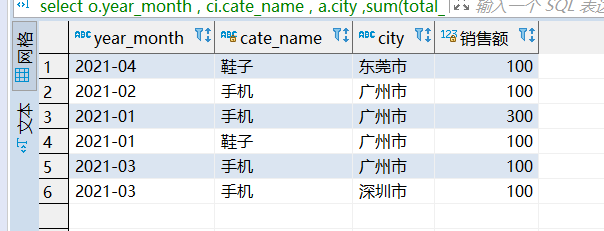
select  o.year\_month , ci.cate\_name , a.city ,sum(total\_payment ) "销售额"

from  "hbcEc2"."order" o , "hbcEc2".sku s , "hbcEc2".cate\_info ci, "hbcEc2".address a

where o.sku\_id =s.sku\_id and s.cate\_id = ci.cate\_id  and o.address\_id = a.address\_id

group by year\_month , ci.cate\_name , a.city;

执行结果：



1. 按月查看性别、商品类别的销售额

Sql语句：

--按月查看性别、商品类别的销售额

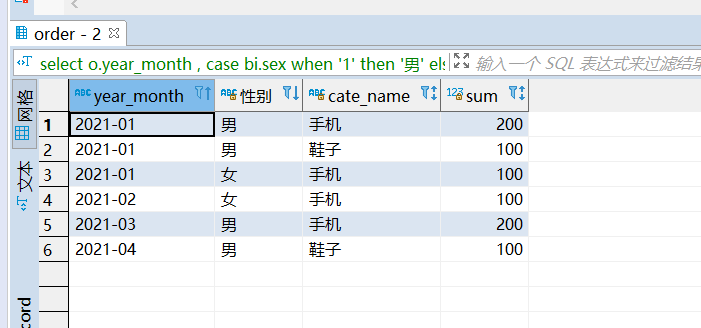
select o.year\_month , case bi.sex when '1' then '男'  else '女' end as 性别, ci.cate\_name , sum(total\_payment )

from "hbcEc2"."order" o , "hbcEc2".buyer\_info bi ,"hbcEc2".sku s ,"hbcEc2".cate\_info ci

where o.buyer\_id  = bi.user\_id and o.sku\_id =s.sku\_id  and s.cate\_id =ci.cate\_id

group by o.year\_month , bi.sex ,ci.cate\_name

执行结果：

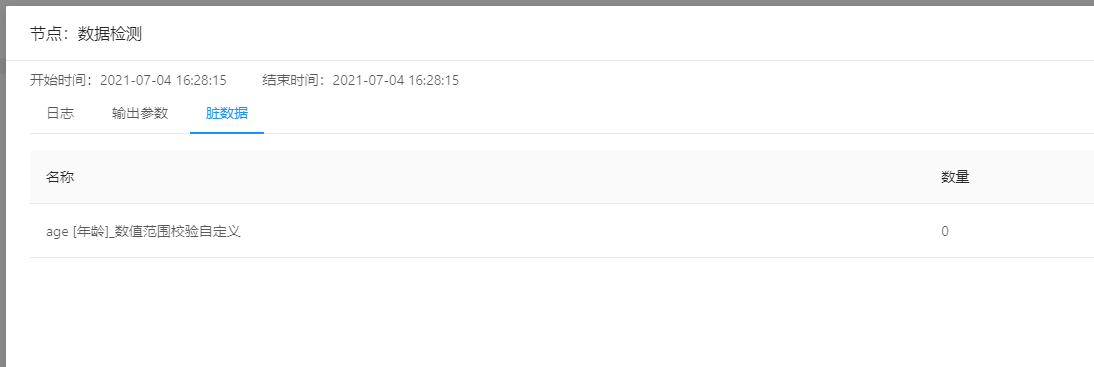


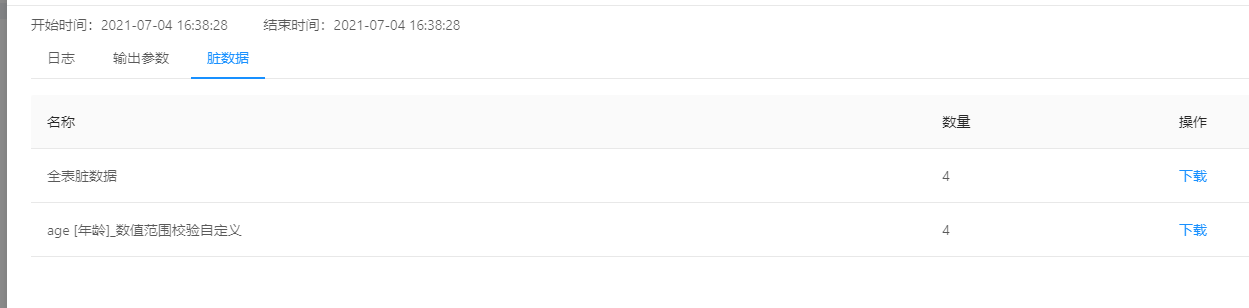
1. **数据质量检测任务**

我尝试创建检测任务，但是创建不出来，只能在数据工厂创建作业，流程里编辑为检测任务。

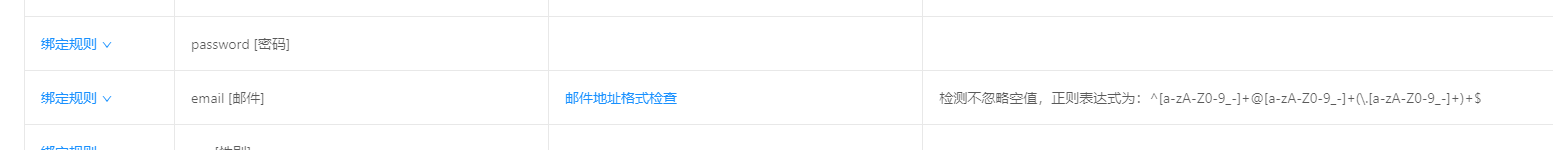


结果不是没有脏数据，因为年龄都为21，合格。



将年龄设置为>=22，结果如下

出现了脏数据。





同理，设置为检测邮箱或检测手机号。