06. 建表、数据导入、创建视图

一、建表

1）建临时表

建表

```
DROP TABLE IF EXISTS dev_dkx.xxx;
CREATE
        TABLE dev_dkx.xxx STORED AS ORC tblproperties
        (
                'orc.compress' = 'SNAPPY'
        ) AS

SELECT ...  FROM  ....
```
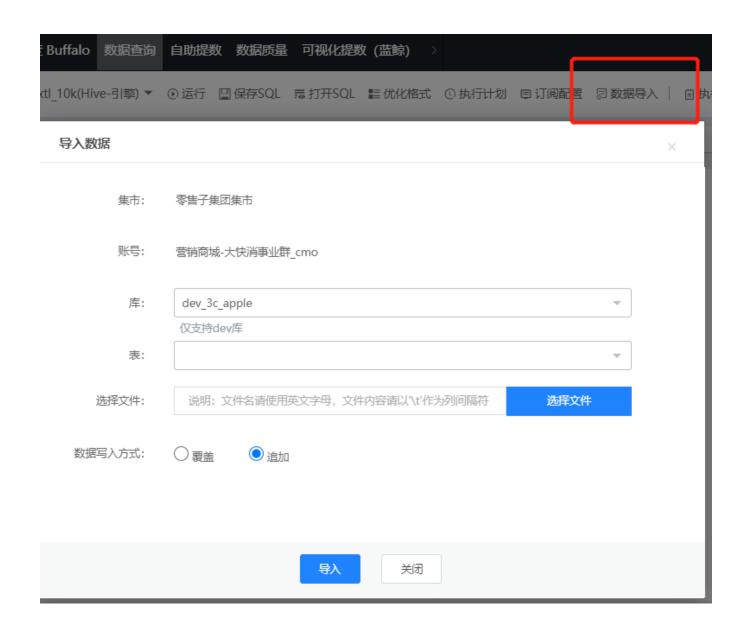
2）创建表导入数据： 先运行代码创建表，在数据查询里点击数据导入，选择库和已创建的表，再选择本地文件（txt格式）覆盖或追加。
等系统显示导入完成再关闭窗口！否则不能导入成功！！

创建表导入数据

```
CREATE
 TABLE dev_dkx.temp_bangong_sku_dic_2021
 (
  class string COMMENT '',
  brand_cn string COMMENT '',
  sku_id string COMMENT 'sku_id',
  sku_name string COMMENT ''
 )
 COMMENT 'temp_bangong_sku_dic_2021' row format delimited fields terminated BY '\t' stored AS textfile;
```

xtl_10k(Hive-引擎) ▼ ⊙运行 ▤保存SQL ▤打开SQL ▤优化格式 ⊙执行计划 ▤订阅配置 ▤数据导入 │ ▤执

导入数据 ✕

集市: 零售子集团集市

账号: 营销商城-大快消事业群_cmo

库: dev_3c_apple ▼

仅支持dev库

表: ▼

选择文件: 说明：文件名请使用英文字母，文件内容请以'\t'作为列间隔符 选择文件

数据写入方式: ◯ 覆盖 ⦿ 追加

导入 关闭

3）创建分区表

创建分区表

```
CREATE
        EXTERNAL TABLE IF NOT EXISTS app.app_jdmkt_visit_dept_sku_pin_order_d_mid
        (

        sale_ord_dt string comment '',
    YEAR_month string  comment'',
        item_sku_id string comment  'skuid',
        dept_id_1 string comment  'id',
        dept_name_1 string comment '',
        dept_id_2 string comment 'id',
        dept_name_2 string comment '',
        user_log_acct string comment '',
        parent_sale_ord_id string comment '',
        after_prefr_amount_1 string comment ''
        )
        COMMENT '_visit' PARTITIONED BY
        (
            dt string
        )
        ROW FORMAT DELIMITED FIELDS TERMINATED BY '\t' COLLECTION ITEMS TERMINATED BY '\n' STORED AS ORC
TBLPROPERTIES
        (
            'orc.compress' = 'SNAPPY'
        )
```
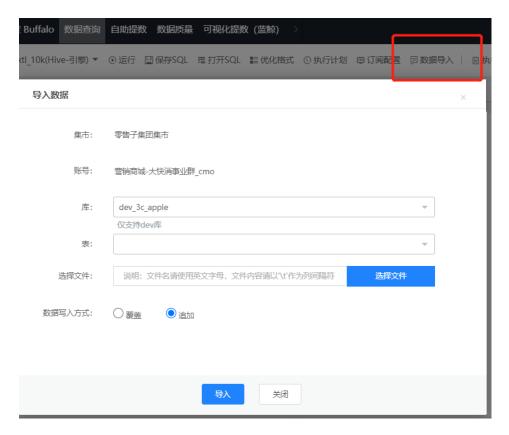
4）补充分区表插入数据

分区表插入数据

```
set mapred.min.split.size.per.node = 256000000;
set mapred.min.split.size.per.rack = 256000000;
set hive.input.format = org.apache.hadoop.hive.ql.io.CombineHiveInputFormat;
set hive.hadoop.supports.splittable.combineinputformat = true;
set mapreduce.input.fileinputformat.split.maxsize = 256000000;
set mapreduce.input.fileinputformat.split.minsize.per.node = 256000000;
set mapreduce.input.fileinputformat.split.minsize.per.rack = 256000000;
SET mapred.output.compress = true;
SET hive.exec.compress.output = true;
SET hive.default.fileformat = Orc;
SET mapred.output.compression.codec = org.apache.hadoop.io.compress.SnappyCodec;
SET hive.auto.convert.join = true;
SET mapred.job.priority = VERY_HIGH;
SET spark.sql.hive.mergeFiles = true;
SET hive.exec.dynamic.partition.mode = nonstrict;
SET hive.exec.dynamic.partition = true;
SET hive.exec.max.dynamic.partitions = 100000;
SET hive.exec.max.dynamic.partitions.pernode = 100000;
SET mapred.min.split.size = 1000000000;
SET mapred.max.split.size = 3000000000;
INSERT overwrite TABLE app.app_jdmkt_visit_dept_sku_pin_order_d_mid partition --- overwriteinto
        (dt
        )
SELECT ...
FROM....
```

二、数据导入

导入本地数据

可上传txt文件（utf-8格式）

三、创建视图：

对于比较敏感的大盘表，事业部只能申请视图使用本部门的数据

创建视图流程说明：https://cf.jd.com/pages/viewpage.action?pageId=493459650 （有问题找 曹玥）

创建视图参考代码：

```
create view if not exists app.v_adm_d04_trade_ord_det_sku_snapshot_fs
as
select a.* from
adm.adm_d04_trade_ord_det_sku_snapshot a
left semi join
(SELECT
        item_sku_id
     FROM
     gdm.gdm_m03_sold_item_sku_da
     WHERE
     dt = sysdate( - 1)
     AND
     dept_id_1 = 6368)
b
on a.item_sku_id = b.item_sku_id
```