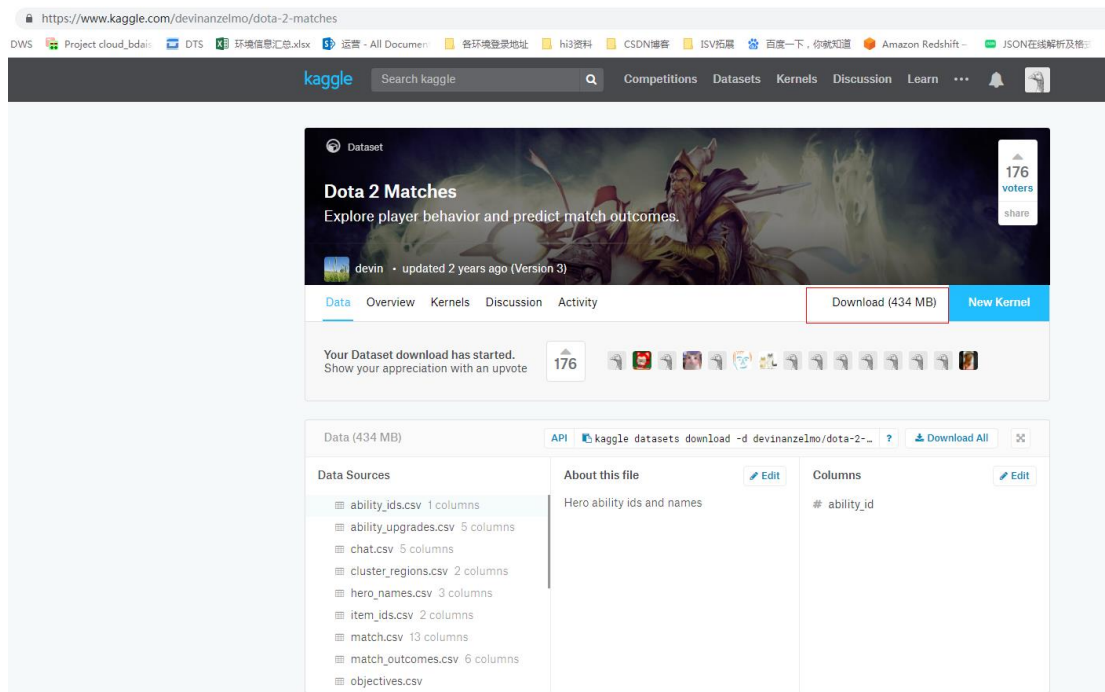


数据仓库服务DWS使用---OBS连接

一. 根据下面提供的资料与提示, 进行Dota2比赛分析

(1) 下载数据 (<https://dws.obs.cn-north-1.myhuaweicloud.com/dota-2-matches.zip>)



(2) (https://support.huaweicloud.com/devg-dws/import_from_obs_0001.html) 根据导入指导, 将数据导入以创建好的数据仓库

(3) 用户可根据自己的爱好, 对dota2的比赛数据进行分析, 步骤可以参照之前的操作文档。不强制要求将所有的表上传并分析, 用户可根据自己的实际情况, 上传几张表即可。

下面是样例, 请参考:

1、创建OBS桶:

区域

华北-北京一

不同区域的云服务产品之间网互不相通，请选择靠近您业务的区域，可减少网络时延，提高访问速度。桶创建成功后不支持变更区域，请谨慎选择。

桶名称

obs-demo

命名规则

- 需全局唯一，不能与已有的任何桶名称重复。
- 长度范围为3到63个字符，支持小写字母、数字、中划线（-）、英文句号（.）。
- 禁止两个英文句号（.）或英文句号（.）和中划线（-）相邻，禁止以英文句号（.）和中划线（-）开头或结尾。
- 禁止使用IP地址。
- 如果名称中包含英文句号（.），访问桶或对象时可能会进行安全证书校验。

存储类别

标准存储

低频访问存储

归档存储

适用于有大量热点文件或小文件，且需要频繁访问（平均一个月多次）并快速获取数据的业务场景。

桶策略

私有

公共读

公共读写

任何用户都可以对桶内对象进行读/写/删除操作。

多AZ

所选区域不支持配置多AZ。

2、导入数据到OBS桶中：

对象 已删除对象 碎片

对象是数据的基本单位，在OBS中文件和文件夹都是对象。您可以上传任何类型（文本、图片、视频等）的文件，并在桶中对这些文件进行管理。 [了解更多](#)

上传文件

新建文件夹

恢复

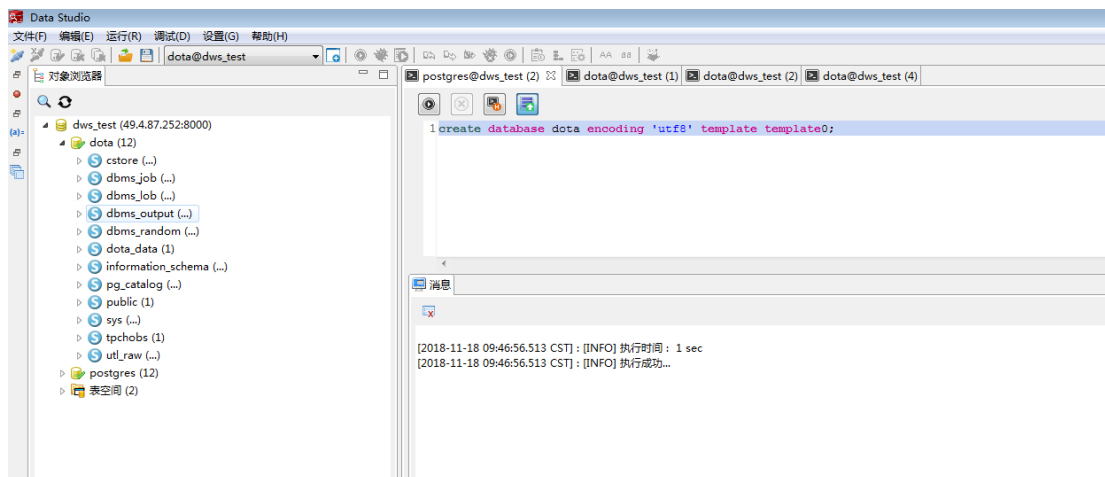
删除

输入对象名前缀搜索

<input type="checkbox"/>	名称	存储类别	大小	加密状态	恢复状态	最后修改时间	操作
← 返回上一级							
<input type="checkbox"/>	ability.csv	标准存储	18.807 KB	未加密	-	2018/11/18 10:31:21 GMT+08:00	下载 分享 更多
<input type="checkbox"/>	ability_ids.csv	标准存储	18.807 KB	未加密	-	2018/11/18 10:30:39 GMT+08:00	下载 分享 更多

3、在DWS集群中创建数据库：

```
create database dota encoding 'utf8' template template0;
```



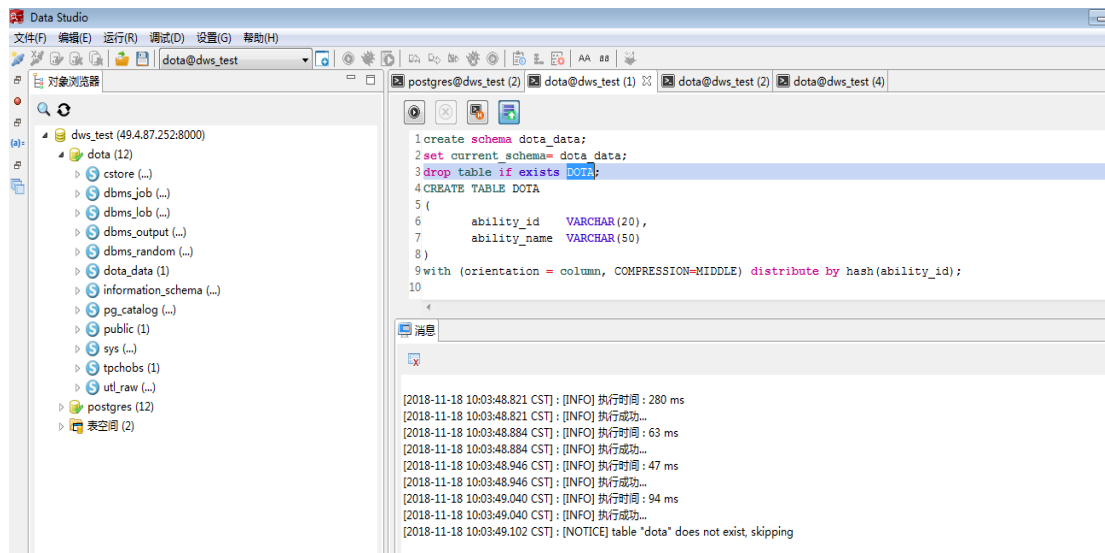
4、数据库中创建表：

```
drop table if exists DOTA;
CREATE TABLE DOTA
(
    ability_id VARCHAR(20),
```

```

        ability_name VARCHAR(50)
    )
    with (orientation = column, COMPRESSION=MIDDLE) distribute by hash(ability_id );

```



5、创建外表;

```

drop FOREIGN table if exists DOTA _OBS;
CREATE FOREIGN TABLE DOTA _OBS
(
    like dota_data.DOTA
)SERVER gsmpp_server
OPTIONS
(
    encoding 'utf8',
    location 'obs://obs-demo/ability_ids/ability.csv',
    format 'text',
    delimiter ',',
    access_key 'AAAAAAAAAAAAAAAAAAAA',
    secret_access_key 'BBBBBBBBBBBBBBBBBB',
    chunksize '64',
    IGNORE_EXTRA_DATA 'on'
);

```

Location: 文件在OBS桶的路径, 根据实际情况修改;

Access_key/secret_access_key:相应用户的ak/sk.

6、查询外表是否有数据, 并向内表插入数据;

```
Select * from DOTA_OBS limit 1;
```

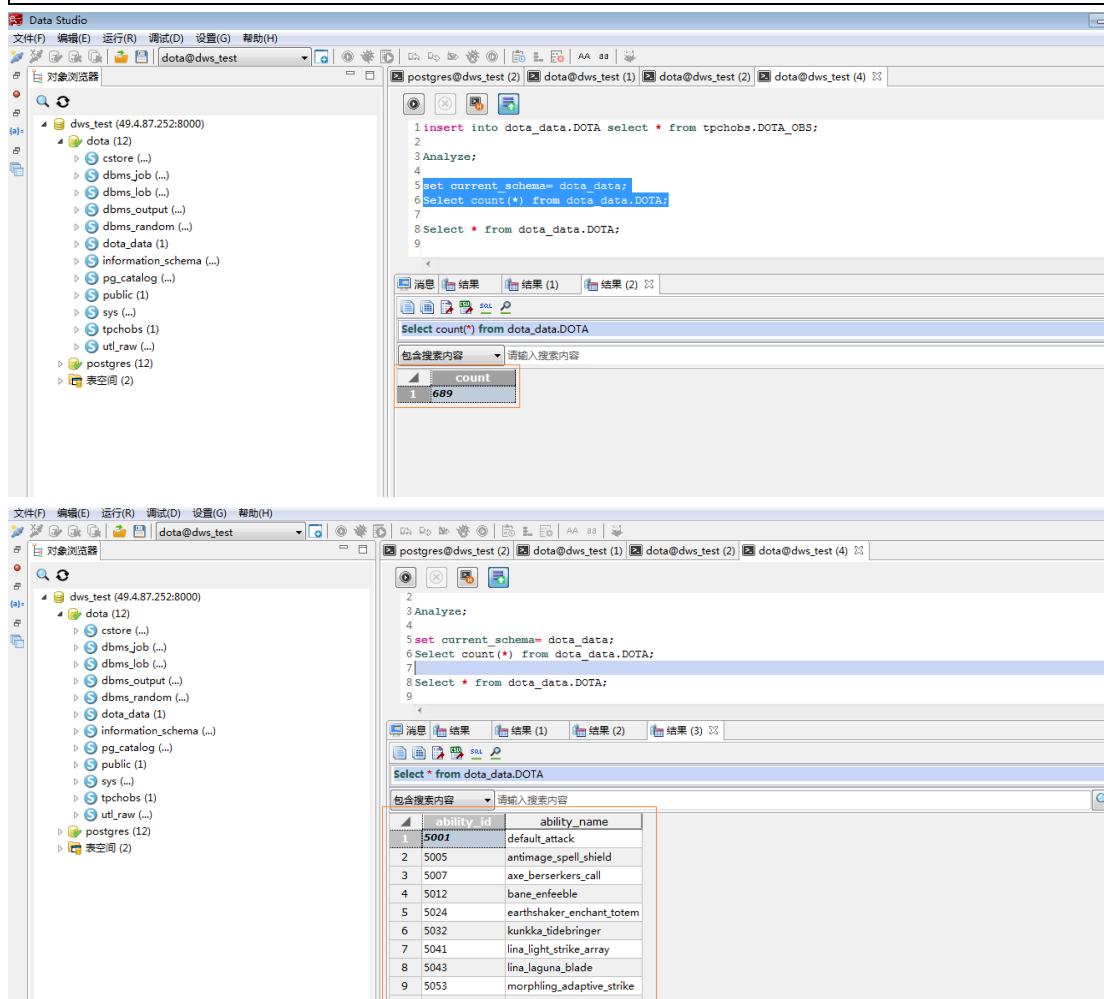
如果返回一条数据，表明外表中有数据，可以执行下面的语句进行插入操作；如果报错，或者没有返回数据，说明外表创建有误，需要检查外表配置，修改后重新创建外表。

```
insert into DOTA select * from DOTA_OBS;
```

7、查询数据样例：

```
analyze;
```

```
select count(*) from DOTA;
```



若上面流程可以走通，可根据自己的喜好，相同的步骤导入其他几张表到DWS集群。

7、使用Tableau工具连接DWS数据仓库：



连接

到文件

- Microsoft Excel
- 文本文件
- JSON 文件
- Microsoft Access
- PDF 文件
- 空间文件
- 统计文件
- 更多...

到服务器

- Tableau Server
- Microsoft SQL Server
- MySQL
- Oracle
- Amazon Redshift
- 更多...

已保存数据源

- Sample - Superstore

搜索

- MarkLogic
- MemSQL
- Microsoft Analysis Services
- Microsoft PowerPivot
- Microsoft SQL Server
- MonetDB
- MongoDB BI 连接器
- MySQL
- OData
- OneDrive
- Oracle
- Oracle Eloqua
- Oracle Essbase
- Pivotal Greenplum Database
- PostgreSQL
- Presto
- Progress OpenEdge
- Salesforce
- SAP HANA
- SAP NetWeaver Business Warehouse
- SAP Sybase ASE

PostgreSQL

服务器(V): 49.4.87.252 端口(P): 8000

数据库(D): postgres

输入数据库登录信息:

身份验证: 用户名和密码

用户名(U): dbadmin

密码(P): *****

☐ 需要 SSL

初始 SQL(I)...

登录

连接

49.4.87.252 PostgreSQL

数据库 postgres

表

- all_source (sys.all_source)
- all_tab_columns ...all_tab_columns
- all_tables (sys.all_tables)
- all_views (sys.all_views)
- dba_data_files (sys.dba_data_files)
- dba_ind_partitions ...dba_ind_partitions
- dba_indexes (sys.dba_indexes)
- dba_objects (sys.dba_objects)
- dba_part_indexes ...dba_part_indexes
- dba_part_tables ...dba_part_tables
- dba_procedures ...dba_procedures
- dba_sequences ...dba_sequences

新建自定义 SQL

新建并集

数据源

工作表 1

all_objects

all_all_tables

dba_data_files

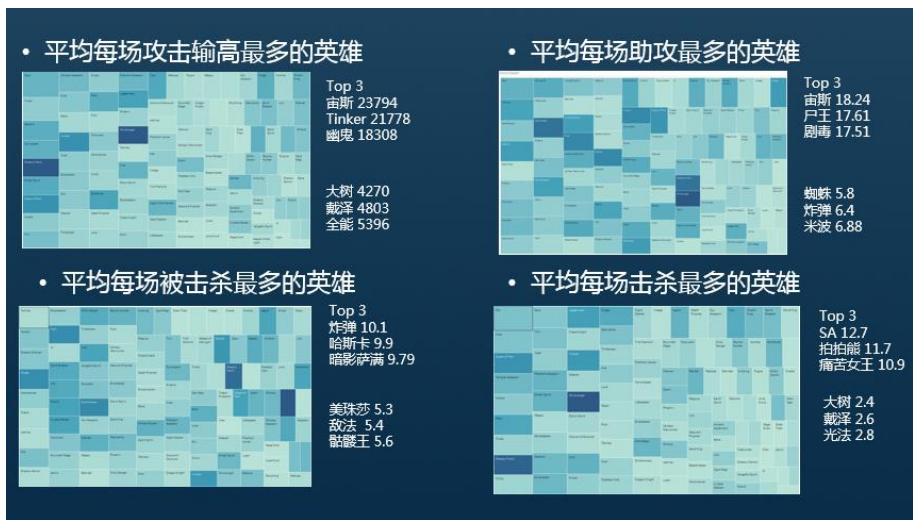
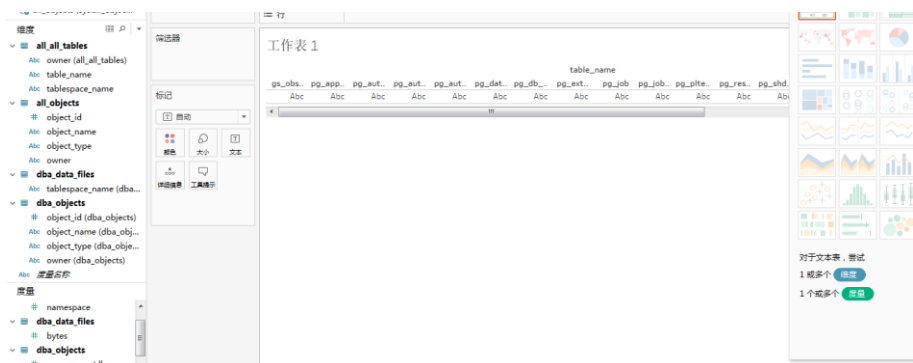
dba_objects

排序字段 数据源顺序

显示别名 显示隐藏字段 1,000

owner (all_all_tables)	table_name	tablespace_name	owner	object_name	object_id	object_type	namespace
Ruby	pg_db_role_setting	pg_global	Ruby	bit_or	2237	procedure	
Ruby	gs_obscaninfo	pg_global	Ruby	bit_or	2237	procedure	
Ruby	pg_extension_data_s...	pg_global	Ruby	bit_or	2237	procedure	
Ruby	pg_job_proc	pg_global	Ruby	bit_or	2237	procedure	
Ruby	pg_job	pg_global	Ruby	bit_or	2237	procedure	
Ruby	pg_thseclabel	pg_global	Ruby	bit_or	2237	procedure	
Ruby	pg_auth_history	pg_global	Ruby	bit_or	2237	procedure	
Ruby	pg_user_status	pg_global	Ruby	bit_or	2237	procedure	

8、使用Tableau分析数据：

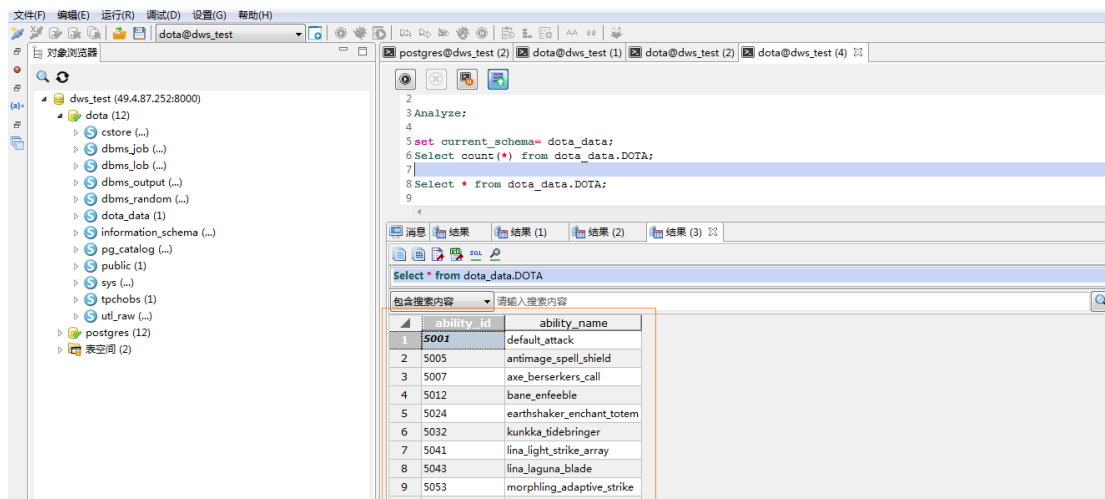


注：

使用Tableau连接DWS集群时要保证网络的连通性，否则会发生连接断开的现象。

打卡任务

完成上述的步骤，可以对分析结果进行截图。可以对DataStudio的结果截图，类似于下图：



也可对Tableau的分析结果截图，类似于下图：

