

# 2019

A low-angle, upward-looking shot of a modern glass skyscraper with a grid-like facade, set against a clear blue sky. The building's structure is composed of dark metal frames and large glass panels that reflect the sky.

## 企业级数据仓库实战

A stylized world map rendered in a light gray, dotted pattern, centered on the Atlantic Ocean. The map is positioned behind the main title text. The overall background features a blue gradient with geometric shapes and a dark blue wavy line at the bottom.



# Hive DDL操作介绍



按照动作类型分：

1、创建

CREATE DATABASE/SCHEMA, TABLE, VIEW, FUNCTION, INDEX

2、删除

DROP DATABASE/SCHEMA, TABLE, VIEW, INDEX

3、清空

TRUNCATE TABLE

4、更新

ALTER DATABASE/SCHEMA, TABLE, VIEW

5、查询

SHOW DATABASES/SCHEMAS, TABLES, TBLPROPERTIES, VIEWS, PARTITIONS, FUNCTIONS, INDEX[ES], COLUMNS, CREATE TABLE



按照操作主体分：

1、数据库

CREATE/DROP/ALTER/USER DATABASE

2、表

CREATE/DROP/TRUNCATE TABLE

3、视图

CREATE/DROP/ALTER VIEW

4、索引

CREATE/DROP/ALTER INDEX

5、函数

CREATE/DROP/RELOAD FUNCTION



### 1、创建数据库

```
CREATE (DATABASE|SCHEMA) [IF NOT EXISTS] database_name  
[COMMENT database_comment]  
[LOCATION hdfs_path]  
[WITH DBPROPERTIES (property_name=property_value, ...)];
```

### 2、删除数据库

```
DROP (DATABASE|SCHEMA) [IF EXISTS] database_name [RESTRICT|CASCADE];
```

### 3、更新数据库

```
ALTER (DATABASE|SCHEMA) database_name SET LOCATION hdfs_path;
```

### 4、切换数据库

```
USE DATABASE;
```





```
CREATE [TEMPORARY] [EXTERNAL] TABLE [IF NOT EXISTS] [db_name.]table_name
[(col_name data_type [column_constraint_specification] [COMMENT col_comment], ... [constraint_specification])]
[COMMENT table_comment]
[PARTITIONED BY (col_name data_type [COMMENT col_comment], ...)]
[CLUSTERED BY (col_name, col_name, ...) [SORTED BY (col_name [ASC|DESC], ...)] INTO num_buckets BUCKETS]
[SKEWED BY (col_name, col_name, ...)]
  ON ((col_value, col_value, ...), (col_value, col_value, ...), ...)
  [STORED AS DIRECTORIES]
[
  [ROW FORMAT row_format]
  [STORED AS file_format]
  | STORED BY 'storage.handler.class.name' [WITH SERDEPROPERTIES (...)]
]
[LOCATION hdfs_path]
[TBLPROPERTIES (property_name=property_value, ...)]
[AS select_statement];
```



```
CREATE [TEMPORARY] [EXTERNAL] TABLE [IF NOT EXISTS] [db_name.]table_name
[(col_name data_type [column_constraint_specification] [COMMENT col_comment], ... [constraint_specification])]
[COMMENT table_comment]
[PARTITIONED BY (col_name data_type [COMMENT col_comment], ...)]
[CLUSTERED BY (col_name, col_name, ...) [SORTED BY (col_name [ASC|DESC], ...)] INTO num_buckets BUCKETS]
[SKEWED BY (col_name, col_name, ...)]
  ON ((col_value, col_value, ...), (col_value, col_value, ...), ...)
  [STORED AS DIRECTORIES]
[
  [ROW FORMAT row_format]
  [STORED AS file_format]
  | STORED BY 'storage.handler.class.name' [WITH SERDEPROPERTIES (...)]
]
[LOCATION hdfs_path]
[TBLPROPERTIES (property_name=property_value, ...)]
[AS select_statement];
```



## DDL之表创建操作-标准样式

```
CREATE [TEMPORARY] [EXTERNAL] TABLE [IF NOT EXISTS] [db_name.]table_name
(
    col1 STRING COMMENT '字段注释1'
    ,col2 STRING COMMENT '字段注释2'
    ,col3 STRING COMMENT '字段注释3'
)
COMMENT '表注释'
PARTITIONED BY (pt STRING)
ROW FORMAT FIELDS TERMINATED BY '|' COLLECTION ITEMS TERMINATED BY ',' MAP KEYS TERMINATED BY ':'
STORED AS ORC/TEXTFILE/PARQUET
LOCATION '/home/user/hive/db_name/table_name/';
```

### 注意事项:

- 1、一般正式表需要使用外部表EXTERNAL字段进行标识，避免异常删除导致文件丢失
- 2、一般正式表需要有分区字段，避免全表扫描
- 3、一般正式表存储格式为TEXTFILE(文本格式)、PARQUET或者ORC(列存储+压缩)
- 4、一般正式表需要指定对应的LOCATION





## DDL之表创建操作-demo

```
CREATE EXTERNAL TABLE IF NOT EXISTS DW2.score
```

```
(  
    name string  
    , score map<string,int>  
)
```

```
COMMENT '建表测试'
```

```
PARTITIONED BY (pt STRING,hh STRING)
```

```
ROW FORMAT DELIMITED
```

```
FIELDS TERMINATED BY '|'
```

```
COLLECTION ITEMS TERMINATED BY ','
```

```
MAP KEYS TERMINATED BY ':'
```

```
STORED AS ORC
```

```
LOCATION 'hdfs://localhost:8020/dw2/score';
```



### 更新字段

```
ALTER TABLE table_name  
[PARTITION partition_spec]  
ADD|REPLACE COLUMNS (col_name data_type [COMMENT col_comment], ...)  
[CASCADE|RESTRICT]
```

### 更新分区

```
ALTER TABLE table_name DROP [IF EXISTS] PARTITION partition_spec[, PARTITION partition_spec, ...]  
ALTER TABLE table_name ADD [IF NOT EXISTS] PARTITION partition_spec [LOCATION 'location'][, PARTITION partition_spec [LOCATION 'location'], ...];
```

### 注意事项:

- 1、更新字段时记得添加CASCADE关键字，添加后所有的分区元数据都会更新到最新状态，否则只更新表元数据
- 2、一般在写入表分区数据时，先进行分区DROP，在进行分区挂载，可以减少因上一步中提到的问题造成数据访问异常



### 新建视图

```
CREATE VIEW [IF NOT EXISTS] [db_name.]view_name [(column_name [COMMENT column_comment], ...) ]  
[COMMENT view_comment]  
[TBLPROPERTIES (property_name = property_value, ...)]  
AS SELECT ...;
```

### 删除视图

```
DROP VIEW [IF EXISTS] [db_name.]view_name;
```

### 注意事项:

- 1、视图本质是一段SQL，其本身不存储数据
- 2、通过视图可以屏蔽底层数据，提供给相应的业务方使用



### 新建函数

```
CREATE FUNCTION [db_name.]function_name AS class_name  
[USING JAR|FILE|ARCHIVE 'file_uri' [, JAR|FILE|ARCHIVE 'file_uri' ]];
```

### 删除函数

```
DROP FUNCTION [IF EXISTS] function_name;
```

THANK YOU FOR YOUR GUIDANCE.

谢谢