# 项目四 Hive数据定义



《Hive数据仓库应用教程》

# 目 录 / CONTENTS


1 任务4-1 HiveQL操作数据库

2 任务4-2 HiveQL操作数据表

3 任务4-3 HiveQL操作分区表

4 任务4-4 HiveQL操作分桶表

5 任务4-5 HiveQL操作视图

6 任务4-6 HiveQL相关操作

# 任务 4-1 HiveQL操作数据库

# 4.1.1 任务描述

为了便于数据的存储和管理，在Hive数据仓库中会根据数据的分类将数据存储在数据库的指定数据表中，可以理解为数据库是数据存储与管理的基础。

本次任务将详细讲解如何在Hive中操作数据库，主要包括对数据库的创建、查询、查看、切换、修改、删除等基本操作。

# 4.1.2 任务准备知识

## 1.创建数据库

Hive中创建数据库的语法格式如下：

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

说明：

- CREATE (DATABASE|SCHEMA)：表示创建数据库的语句；
- IF NOT EXISTS：可选，用于判断创建的数据库是否已经存在，若不存在则创建，反之不创建；
- database_name：表示创建的数据库名称；
- COMMENT database_comment：可选，表示数据库的相关描述；
- LOCATION hdfs_path：可选，用于指定数据库在HDFS上的存储位置；
- WITH DBPROPERTIES (property_name=property_value, ...)：可选，用于设置数据库属性。

## 注意事项

- 蓝色字体是建库语法的关键字，用于指定某些功能。
- [ ]中括号的语法表示可选。
- | 表示使用的时候，左右语法二选一。
- **建库语句中的语法顺序要和语法树中顺序保持一致。**

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

## 2．查询数据库

Hive中查询数据库的语法格式如下：

SHOW (DATABASES|SCHEMAS) [LIKE 'identifier_with_wildcards’ ];

说明：

- SHOW (DATABASES|SCHEMAS)：表示查询数据库的语句；
- LIKE 'identifier_with_wildcards'：可选，LIKE子句用于模糊查询。

## 3．查看数据库

Hive中查看数据库信息的语法格式如下：

DESCRIBE|DESC (DATABASES|SCHEMAS) [EXTENDED] db_name;

说明：

- DESCRIBE|DESC (DATABASES|SCHEMAS)：表示查询数据库信息的语句；
- EXTENDED：可选，在查询数据库的信息中显示属性；
- db_name：用于指定查询的数据库名称。

## 4．切换数据库

切换数据库的语法格式：

```
USE db_name;
```

说明：

- USE：表示切换数据库的语句；
- db_name：指定数据库名称。

## 5．修改数据库

Hive中可以修改数据库信息中的属性和所有者，修改数据库信息的语法格式如下：

```
/*修改数据库属性*/
ALTER (DATABASE|SCHEMA) database_name SET DBPROPERTIES
 (property_name=property_value, ...);
/*修改数据库所有者*/
ALTER (DATABASE|SCHEMA) database_name SET OWNER  [USER|ROLE] user_or_role;
```

说明：

- ALTER (DATABASE|SCHEMA)：表示修改数据库信息的语句；
- database_name：指定数据库名称；
- SET DBPROPERTIES (property_name=property_value, ...)：指定修改数据库信息中的属性；
- SET OWNER [USER|ROLE] user_or_role：指定修改数据库信息中的所有者。

## 6．删除数据库

Hive中删除数据库的语法格式如下：

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

说明：

- DROP (DATABASE|SCHEMA)：表示删除数据库的语句；
- IF EXISTS：可选，用于判断数据库是否存在；
- database_name：用于指定数据库名称；
- [RESTRICT|CASCADE]：可选，表示数据库中存在表时是否可以删除数据库。

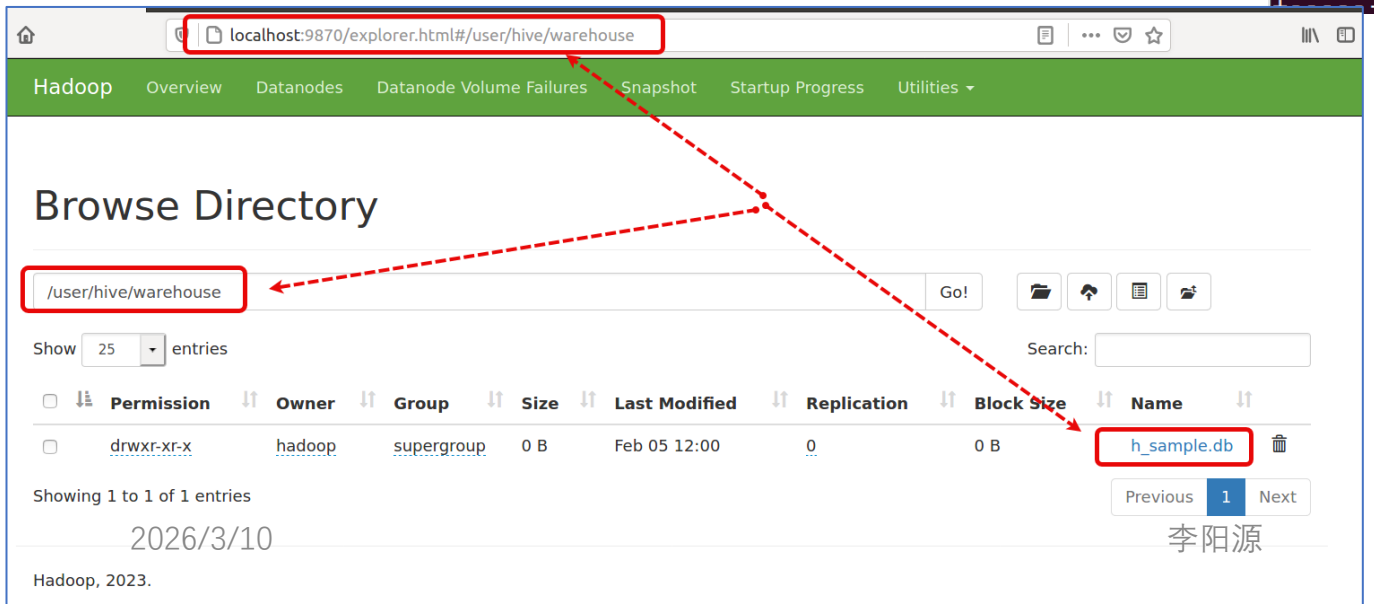
# 4.1.3 任务实施

## 1．创建数据库

完成数据库h_sample的创建，在Hive客户端工具Beeline中创建数据库h_sample，并指定数据库文件存放在HDFS的/hive_db/create_db/目录中，语句如下：

```
CREATE DATABASE if not exists h_sample
COMMENT "This is h_sample database"
LOCATION '/user/hive/warehouse/h_sample.db'
with dbproperties ("date"="2026-02-05");
```

```
0: jdbc:hive2://localhost:10000> CREATE DATABASE if not exists h_sample
0: jdbc:hive2://localhost:10000> COMMENT "This is h_sample database"
0: jdbc:hive2://localhost:10000> LOCATION '/user/hive/warehouse/h_sample.db'
0: jdbc:hive2://localhost:10000> with dbproperties ("date"="2026-02-05");
0: jdbc:hive2://localhost:10000> show databases;
+----------------+
| database_name  |
+----------------+
| default        |
| h_sample       |
+----------------+
```



## 2．查询已有的数据库

完成数据库的查询，查询Hive中所有数据库和查询Hive中数据库名称的首字母是“h”的数据库。

首先，查询Hive中所有数据库，语句如下：

SHOW DATABASES;

然后，查询Hive中数据库名称首字母是“h”的数据库，语句如下：

SHOW DATABASES LIKE 'h*';

```
show databases;

show databases like 'h*';
```

```
0: jdbc:hive2://localhost:10000> show databases;
+----------------+
| database_name  |
+----------------+
| default        |
| h_sample       |
+----------------+
0: jdbc:hive2://localhost:10000> show databases like 'h*';
+----------------+
| database_name  |
+----------------+
| h_sample       |
+----------------+
```

## 3．查看数据库信息

查询Hive数据库h_sample的信息，语句如下：

```
DESC DATABASE EXTENDED h_sample;
```

```
0: jdbc:hive2://localhost:10000> desc database extended h_sample;
+-----------+----------------------------+----------------------------------------------------+-------------+-------------+---------------------+
| db_name   |          comment           |                      location                      | owner_name  | owner_type  |     parameters      |
+-----------+----------------------------+----------------------------------------------------+-------------+-------------+---------------------+
| h_sample  | This is h_sample database  | hdfs://hadoop01:9000/user/hive/warehouse/h_sample.db | hadoop      | USER        | {date=2026-02-05}   |
+-----------+----------------------------+----------------------------------------------------+-------------+-------------+---------------------+
```

## 4．切换数据库

将当前使用的数据库切换至数据库h_sample，语句如下：

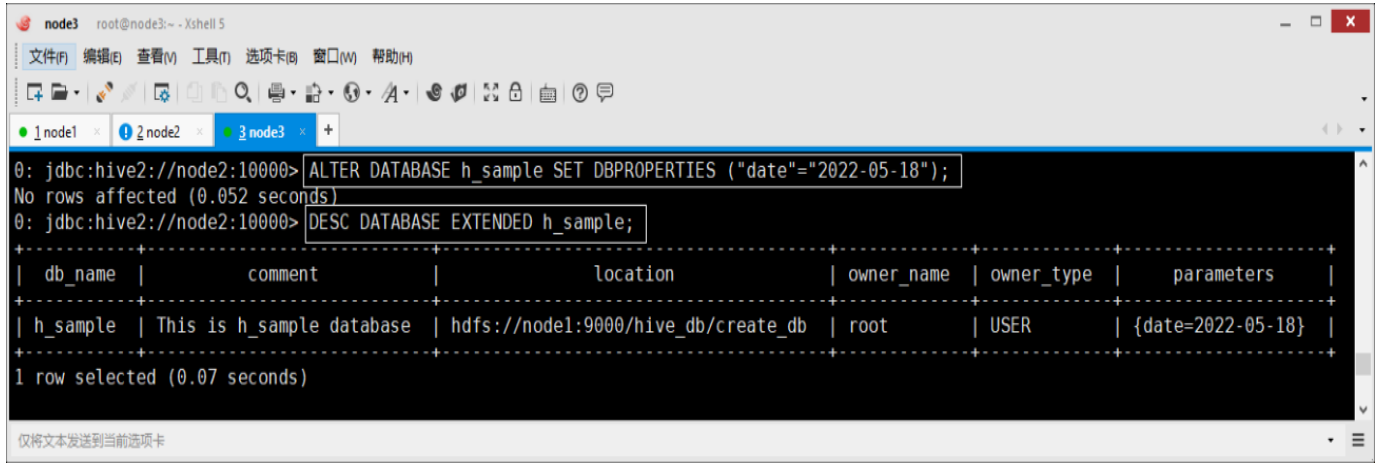
```
USE h_sample;
```

## 5．修改数据库

修改数据库h_sample中的属性，语句如下：

```
ALTER DATABASE h_sample SET DBPROPERTIES
 ("date"="2022-05-18");
```

执行结果如图4-4所示：




## 6．删除数据库

删除数据库h_sample，语句如下：

```
DROP DATABASE IF EXISTS h_sample;
```

# 感谢您的观看



《Hive数据仓库应用教程》