南华大学

计算机学院

实验报告

（ 2022 ~2023 学年度 第一学期 ）

|  |  |
| --- | --- |
| 课程名称 | 大数据平台及编程实践 |
| **实验名称** | 熟悉HIVE的基本操作 |

|  |  |  |  |
| --- | --- | --- | --- |
| 姓名 | 刘纯龙 | 学号 | 20184351149 |

|  |  |  |  |
| --- | --- | --- | --- |
| 专业 | 软件工程 | 班级 | 软卓01班 |

|  |  |  |  |
| --- | --- | --- | --- |
| 地点 |  | 教师 | 聂沛 |

1. **实验目的与要求**

实验目的：理解Hive作为数据仓库在Hadoop体系结构中的角色

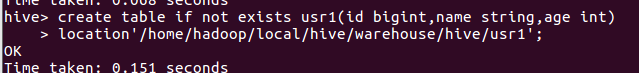
熟练使用常用的HiveQL语句

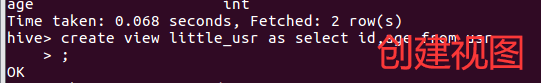
熟悉JAVA api访问hive

实验要求：配置好环境，对hive进行操作

1. **实验内容与步骤**
2. hive的基本操作
   * + 1. create 命令，创建数据库和表(指定路径)



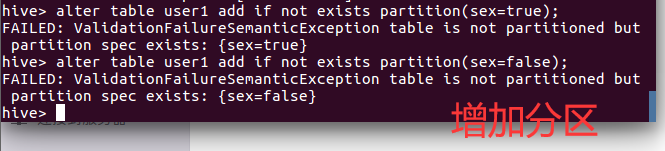
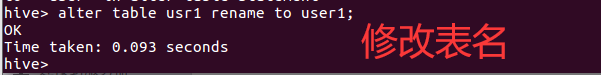




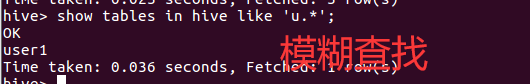
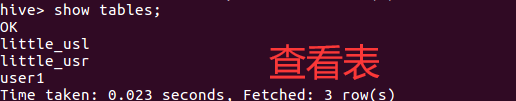
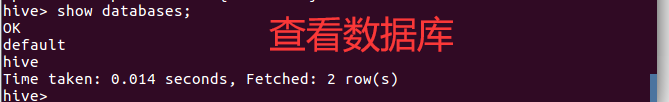
* + - 1. drop命令



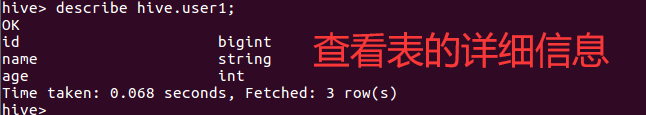
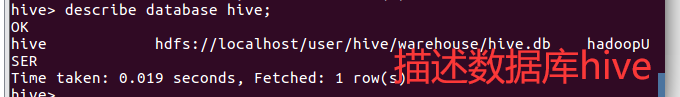
* + - 1. alter命令



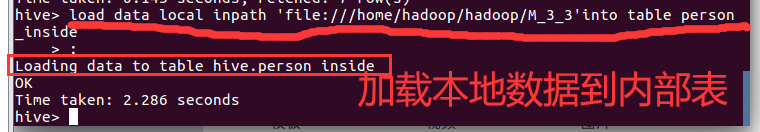
* + - 1. show命令

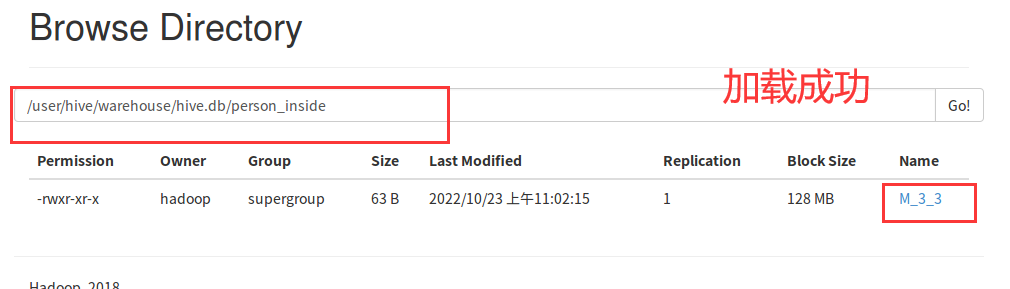


* + - 1. describe命令

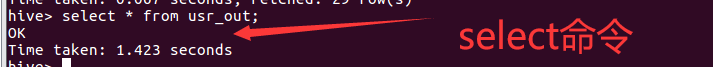


* + - 1. load命令加载文件





* + - 1. select命令

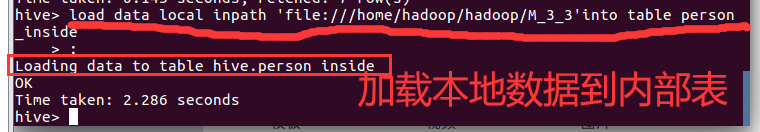
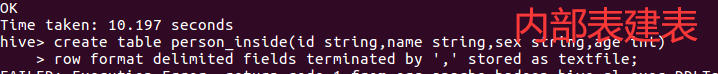


* + - 1. insert命令

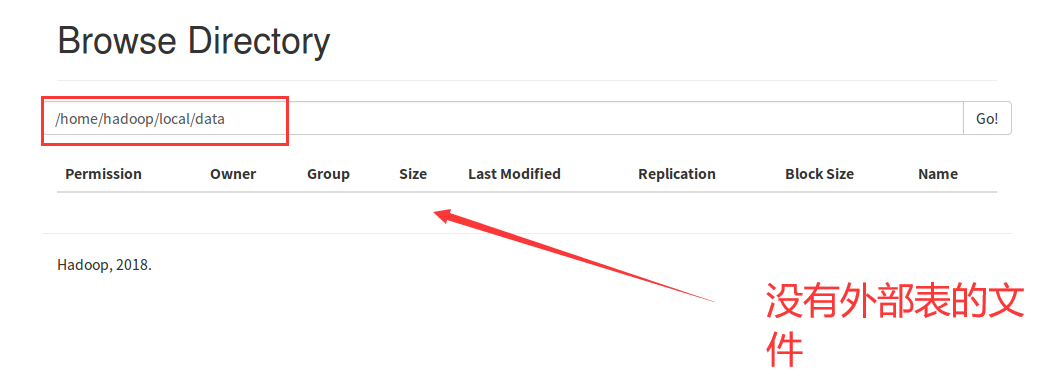
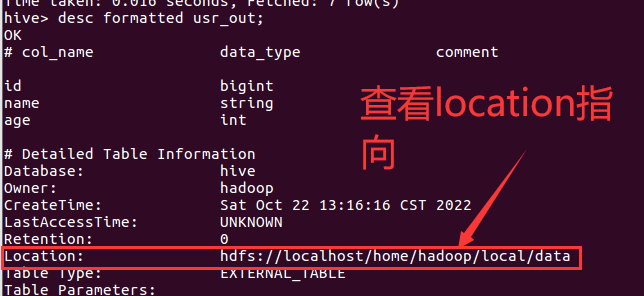
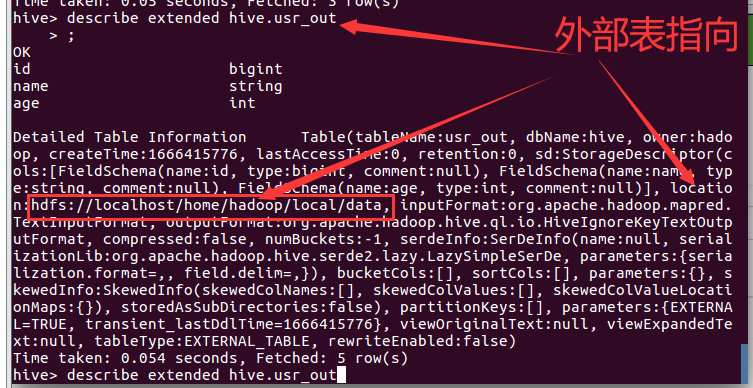
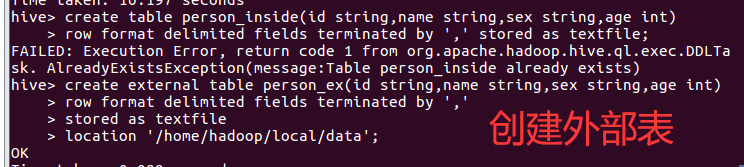


1. 内部表，外部表，分区，分桶

1.内部表



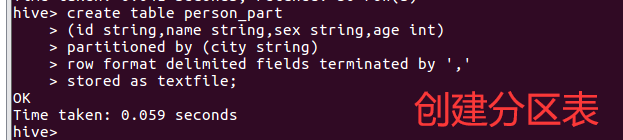
2.外部表

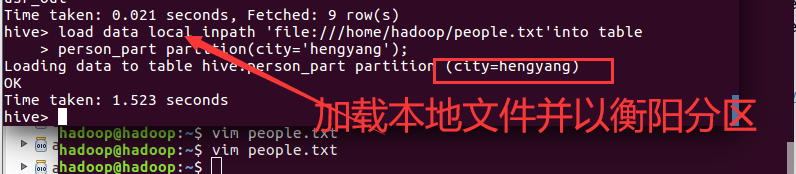


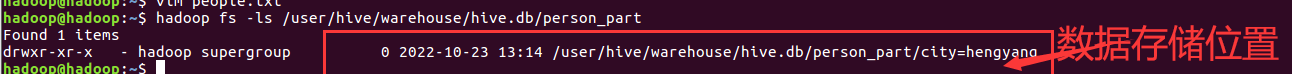


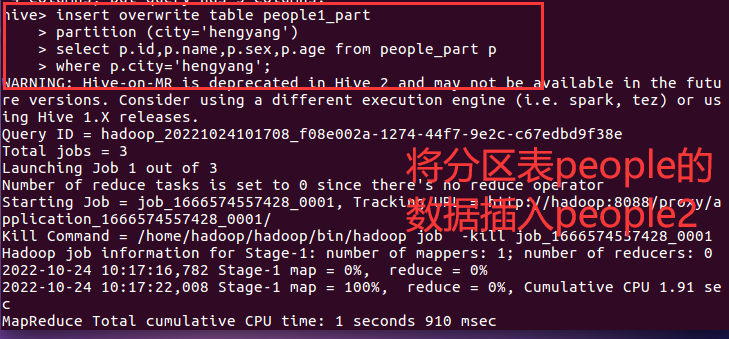
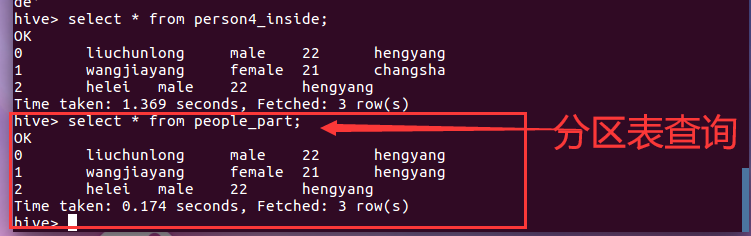


3.分区表

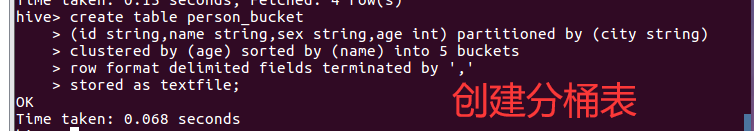


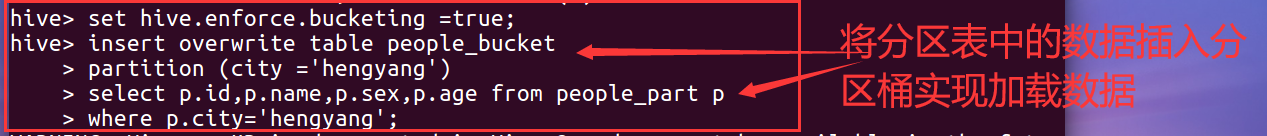


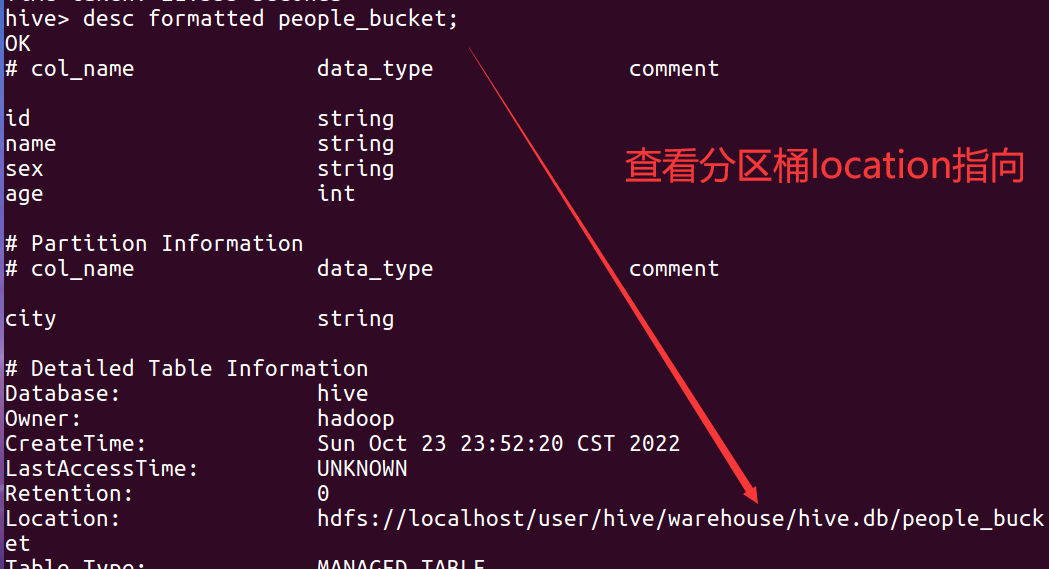


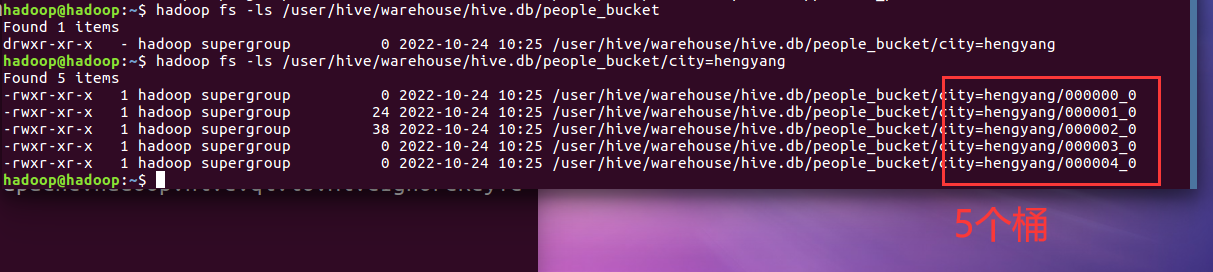


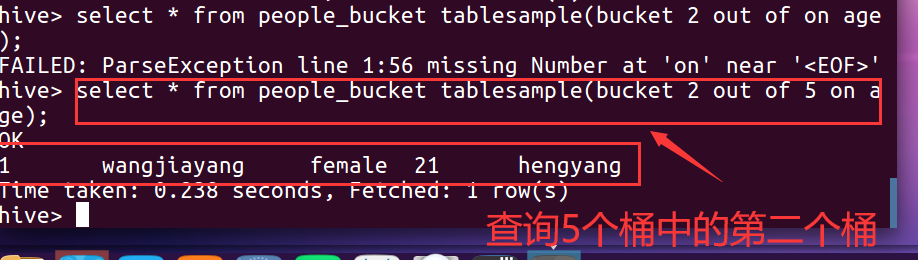
4.分桶表



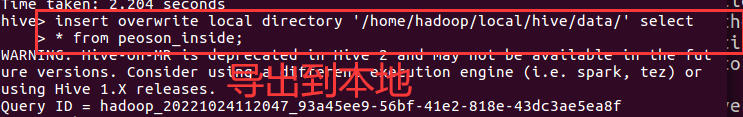


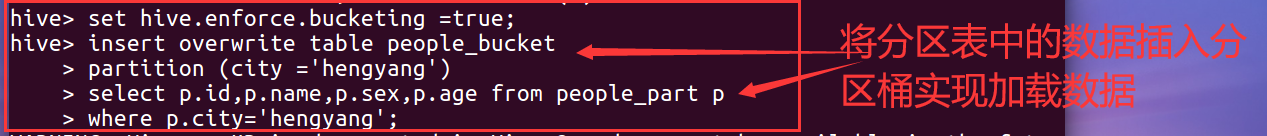


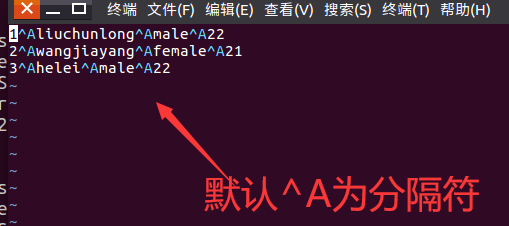


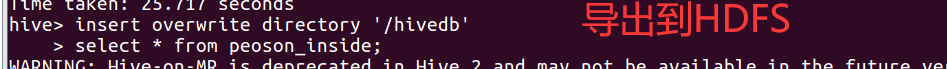


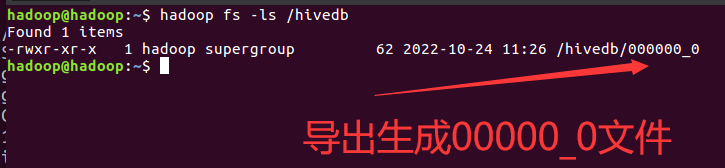
5.导出表



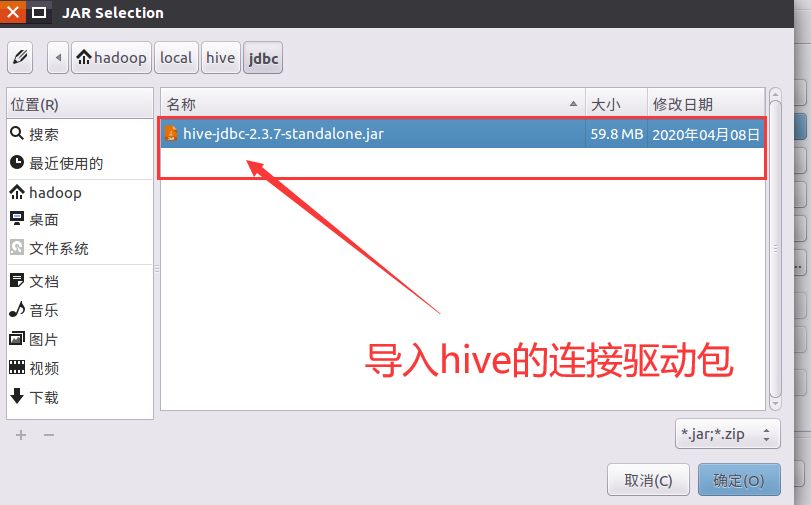






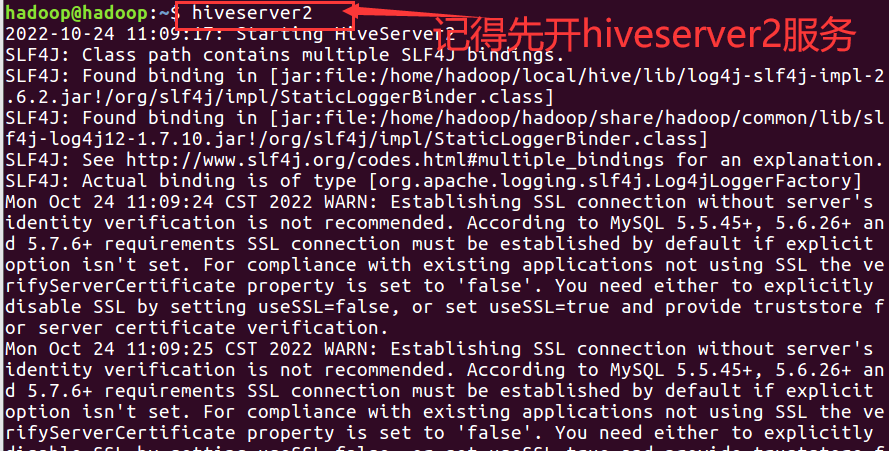


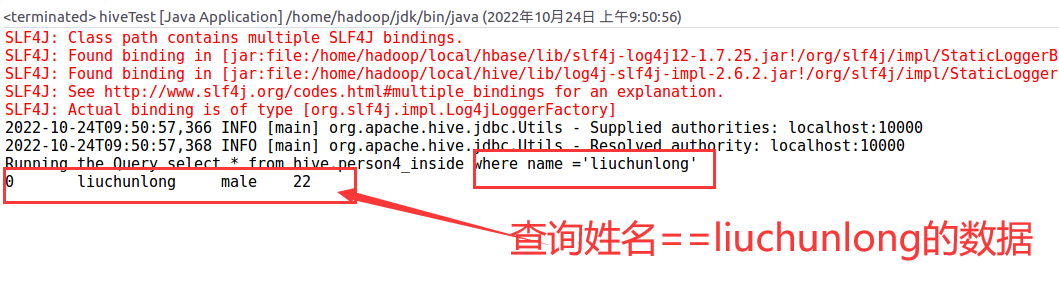
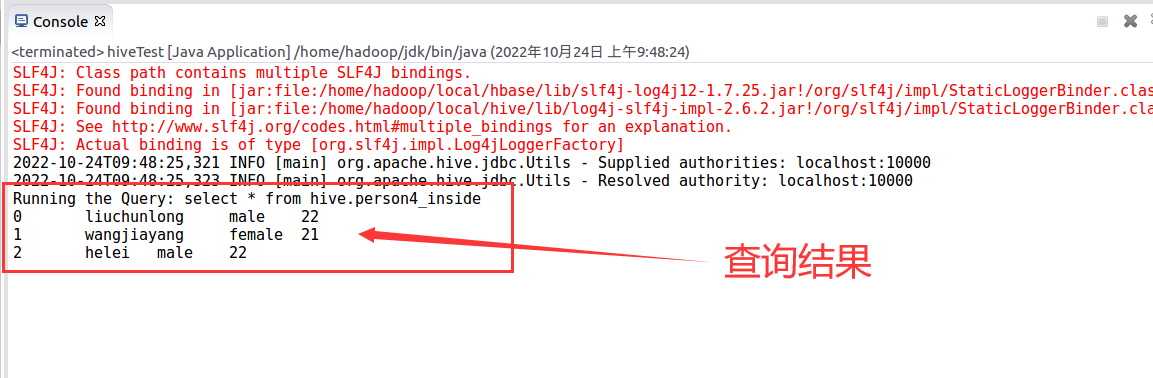
1. java api 访问hive

**1运行代码**



**2.运行结果截图**





1. **实验总结**

1.首先通过hive的基本操作和命令，对hiveQL查询语句有了一定的了解，比如create,drop,select,load,show,describe等等这些命令，以便于对hive进行数据的操作。

2．通过对内部表，外部表，分区表，分桶表的创建，加载数据，查看详细信息等，加深了对每种表的认识，以及他们的特性和注意事项。

3.通过JAVA api 访问hive,对hive中的数据进行查询等，过程一直报错连接不了访问不了hive，经排查才发现是没有开hiveserver2服务。