# 下载指定位置的数据

下载附件中的数据文件【任务1清洗后的数据.zip】，解压，使用SFTP相关工具将数据传至slave2服务器的/home目录下

# 使用Hadoopshell命令创建目录/data/

Hadoop fs –mkdir /data

# 上传数据至HDFS的/data目录下

使用命令：hadoop fs –put /home/任务1清洗后的数据 /data/

# Hive创建数据库

Create database db\_phone\_raw\_1;

# Hive创建表

## 原始数据表

Create table tbl\_phone\_data\_1(fld\_Brand\_Name string, fld\_Phone\_color string, fld\_Phone\_size string) row format delimited fields terminated by '|';

## 手机品牌销量查询暂存表

Create table tbl\_brand\_count(fld\_Brand\_Name string, fld\_sale\_count int);

## 手机颜色销量查询暂存表

Create table tbl\_color\_count(fld\_phone\_color string, fld\_sale\_count int);

## 手机屏幕尺寸销量查询暂存表

Create table tbl\_size\_count(fld\_phone\_size string, fld\_sale\_count int);

# 数据加载命令

Load data inpath '/data/\*' into table tbl\_phone\_data\_1;

# 查询分析

## 手机品牌销量查询

insert overwrite table tbl\_brand\_count select fld\_Brand\_Name,count(1) as fld\_sale\_count from tbl\_phone\_data\_1 group by fld\_Brand\_Name;

## 手机颜色销量查询

insert overwrite table tbl\_color\_count select fld\_phone\_color,count(1) as fld\_sale\_count from tbl\_phone\_data\_1 group by fld\_phone\_color;

## 手机屏幕尺寸销量查询

insert overwrite table tbl\_size\_count select fld\_phone\_size,count(1) as fld\_sale\_count from tbl\_phone\_data\_1 group by fld\_phone\_size;