### 1 行转列

**1）相关函数说明**

CONCAT(string A/col, string B/col…)：返回输入字符串连接后的结果，支持任意个输入字符串;

CONCAT\_WS(separator, str1, str2,...)：它是一个特殊形式的 CONCAT()。第一个参数剩余参数间的分隔符。分隔符可以是与剩余参数一样的字符串。如果分隔符是 NULL，返回值也将为 NULL。这个函数会跳过分隔符参数后的任何 NULL 和空字符串。分隔符将被加到被连接的字符串之间;

注意: CONCAT\_WS must be "string or array<string>

COLLECT\_SET(col)：函数只接受基本数据类型，它的主要作用是将某字段的值进行去重汇总，产生array类型字段。

**2）数据准备**

|  |  |  |
| --- | --- | --- |
| name | constellation | blood\_type |
| 孙悟空 | 白羊座 | A |
| 小张 | 射手座 | A |
| Sam | 白羊座 | B |
| 猪八戒 | 白羊座 | A |
| 凤姐 | 射手座 | A |
| 苍老师 | 白羊座 | B |

**3）需求**

把星座和血型一样的人归类到一起。结果如下：

射手座,A 小张|凤姐

白羊座,A 孙悟空|猪八戒

白羊座,B Sam|苍老师

**4）创建本地constellation.txt，导入数据**

[root@hadoop102 datas]$ vim person\_info.txt

孙悟空 白羊座 A

小张 射手座 A

Sam 白羊座 B

猪八戒 白羊座 A

凤姐 射手座 A

苍老师 白羊座 B

**5）创建hive表并导入数据**

create table person\_info(

name string,

constellation string,

blood\_type string)

row format delimited fields terminated by "\t";

load data local inpath "/opt/module/hive/datas/person\_info.txt" into table person\_info;

**6）按需求查询数据**

SELECT t1.c\_b , CONCAT\_WS("|",collect\_set(t1.name))

FROM (

SELECT NAME ,CONCAT\_WS(',',constellation,blood\_type) c\_b

FROM person\_info

)t1

GROUP BY t1.c\_b

### 2 列转行

**1）函数说明**

EXPLODE(col)：将hive一列中复杂的array或者map结构拆分成多行。

LATERAL VIEW

用法：LATERAL VIEW udtf(expression) tableAlias AS columnAlias

解释：用于和split, explode等UDTF一起使用，它能够将一列数据拆成多行数据，在此基础上可以对拆分后的数据进行聚合。

**2）数据准备**

表6-7 数据准备

|  |  |
| --- | --- |
| movie | category |
| 《疑犯追踪》 | 悬疑,动作,科幻,剧情 |
| 《Lie to me》 | 悬疑,警匪,动作,心理,剧情 |
| 《战狼2》 | 战争,动作,灾难 |

**3）需求**

将电影分类中的数组数据展开。结果如下：

《疑犯追踪》 悬疑

《疑犯追踪》 动作

《疑犯追踪》 科幻

《疑犯追踪》 剧情

《Lie to me》 悬疑

《Lie to me》 警匪

《Lie to me》 动作

《Lie to me》 心理

《Lie to me》 剧情

《战狼2》 战争

《战狼2》 动作

《战狼2》 灾难

**4）创建本地movie.txt，导入数据**

[root@hadoop102 datas]$ vi movie\_info.txt

《疑犯追踪》 悬疑,动作,科幻,剧情

《Lie to me》 悬疑,警匪,动作,心理,剧情

《战狼2》 战争,动作,灾难

**5）创建hive表并导入数据**

create table movie\_info(

movie string,

category string)

row format delimited fields terminated by "\t";

load data local inpath "/opt/module/hive/datas/movie\_info.txt" into table movie\_info;

**6）按需求查询数据**

SELECT movie,category\_name

FROM movie\_info

lateral VIEW

explode(split(category,",")) movie\_info\_tmp AS category\_name ;