## 示例数据说明

本示例数据包含以下几部分：

1. 约80W条行程数据 trips.csv
2. 约12W条用户数据 users.csv
3. 站点名称数据 station.csv
4. 2020年全年节假日数据 workdays.csv

本示例数据格式为 csv文件。

* csv：逗号分隔表，可以表示一个二维表，文件中的每一行对应表格一行数据，行内各个列使用逗号分隔。

### trips

格式：csv，乘车记录。

|  |  |  |
| --- | --- | --- |
| 列序号 | 说明 | 类型说明 |
| 1 | 用户id | 与user.csv 表中用户id进行关联，可以获取用户对应的行程 |
| 2 | 进站名 |  |
| 3 | 进站时间 | YYYY-MM-DD hh:mm:ss |
| 4 | 出站名 |  |
| 5 | 出站时间 | YYYY-MM-DD hh:mm:ss |
| 6 | 渠道类型 | 本例子中不使用 |
| 8 | 票价 | 单位分 |

### stations

格式：csv，站点列表。所有线路所有站点的列表。

|  |  |  |
| --- | --- | --- |
| 列序号 | 说明 | 类型说明 |
| 1 | 站点id | 无业务意义 |
| 2 | 站点名 |  |
| 3 | 线路名 |  |
| 4 | 站点所在区 |  |

### workdays

格式：csv，日期的工作日节假日属性，由于节假日和调休的时间是不规律的，这个表中会定义每一天的日期属性。1为工作日（包括周末调休变成的工作日）；2为正常周末；3为节假日，包括法定节假日，以及与法定节假日相连的周末。

|  |  |  |  |
| --- | --- | --- | --- |
| 列序号 | 说明 | 类型 | 类型说明 |
| 1 | 日期 | string | YYYYMMDD |
| 2 | 节假日属性 | 枚举 | 1，2，或3 |

### users

格式：csv，

|  |  |  |
| --- | --- | --- |
| 列序号 | 说明 | 类型说明 |
| 1 | 用户id | 和trip.csv里面的用户ID关联可获取用户对应的行程 |
| 2 | 省市 |  |
| 3 | 出生年份 |  |
| 4 | 性别 | 0 男，1 女 |