**校园卡一卡通数据集介绍**

校园卡一卡通数据集由上海交通大学一卡通的消费记录产生，消费包括了食堂餐饮、洗浴、超市购物等，是大学校园生活的另一个角度的体现。数据集由交大网络信息中心和 OMNILab 联合提供,在隐私处理的基础上,将校园卡一卡通数据以比赛数据集的形式进行公开,参赛队伍在通过授权后方可使用该数据集参加本次技术创新大赛。

校园卡一卡通数据集包含3个数据，分别是用户信息（account）、交易记录表（trade）、商户信息表（merchant）。三个数据集均以txt文件的形式发布，总共336.3MB，各个字段之前以tab符分割。

**1.用户信息（account）**

用户信息中包含了2014.09.01~2015.01.31上海交通大学闵行校区使用过校园卡消费的学生信息。包扩卡号（匿名化）、学号（匿名化）、性别、年龄、入学年、学生类型。

|  |  |  |
| --- | --- | --- |
| **字段名称** | **字段含义** | **例子** |
| account | 学生卡号（匿名化） | A16100 |
| studentcode | 学生学号（匿名化） | 76118305 |
| gender | 性别 | 男 |
| yearofbirth | 出生年 | 1989 |
| grade | 入学年 | 2012 |
| type | 类型（本科/硕士/博士） | 硕士 |

account字段与交易记录（trade）中fromaccount字段关联；

studentcode与上海交通大学校园网WIFI流量数据集中net\_users.dat中的user\_id相关联。

**2.交易记录（trade）**

交易记录中包含了2014.09.01~2015.01.31，上海交通大学闵行校区的校园卡消费信息。每条交易记录包括消费者卡号、商户号、商户所属系统、交易时间、交易金额。

|  |  |  |
| --- | --- | --- |
| **字段名称** | **字段含义** | **例子** |
| fromaccount | 学生卡号（匿名化） | B62997 |
| toaccount | 商户代码 | 1000029 |
| syscode | 商户所属系统代码 | 34 |
| timestamp | 交易时间 | 2014-10-01 09:50:36 |
| amount | 交易金额（单位分） | 150 |

fromaccount字段与用户信息（account）中account字段关联；

toaccount字段与商户信息（merchant）中toaccount字段关联；

syscode字段与商户信息（merchant）中syscode字段关联。

**3.商户信息（merchant）**

商户信息中包含了上海交通大学大学闵行校区各个商户的信息。包括商户所属系统代码、商户所属系统名称、商户代码、商户名称、商户地点（有缺失）、商户成立时间。

|  |  |  |
| --- | --- | --- |
| **字段名称** | **字段含义** | **例子** |
| syscode | 商户系统代码 | 34 |
| codename | 商户系统名称 | 闵行第二食堂 |
| toaccount | 商户代码 | 1000031 |
| accountname | 商户名称 | 川味点心 |
| address | 商户地点（有缺失） |  |
| opendata | 商户成立时间 | 2007-07-04 |

toaccount字段与交易记录（trade）中toaccount字段关联；

syscode字段与交易记录（trade）中syscode字段关联。