**Aminer团队公开的新浪微博数据说明文文档**

**公开地址：**

链接：https://pan.quark.cn/s/37c29307ba85?pwd=Pwmg

提取码：Pwmg

**用户表：uidlist**

user\_id: 用户真实ID（可在新浪微博中查询）

map\_id：用户映射ID（本数据库多数表采用此ID）

**用户画像表：user\_profile**

user\_id: 用户的user\_id

**bi\_followers\_count：双向关注的用户数**

city：城市代码

verified：认证码

**followers\_count：粉丝数(关注我的）**

location：位置

province：省份

**friends\_count：关注数（我关注的）**

name：昵称

gender：性别

created\_at：创建时间

verified\_type：认证类型

**statuses\_count：微博数**

description：用户描述

**用户关系网络：weibo\_network**

user\_id: 用户的map\_id

friends\_count:用户关注数

friends：具体关注的用户的map\_id，以#隔开

bi\_followers\_count：双向关注的用户数

bi\_followers：具体双向关注的用户的map\_id，以#隔开

**原创微博表：root\_content**

original\_mid：原创微博ID

content：原创微博内容

original\_uid：发布微博的用户map\_id

original\_time：微博发布时间

retweet\_num：在新浪微博中实际转发的数量

re\_with\_num& re\_without\_num：本数据中统计的转发数量

原址中将转发微博分两种形式存放，retweetwithcontent和retweetwithoutcontent，前者统计了转发微博的全部信息，后者去除了转发微博的内容（缩小数据量）

**无内容转发微博表：retweetwithoutcontent**

id：主键ID

retweet\_uid：转发微博的用户map\_id

retweet\_time：转发微博时间

original\_mid：原创微博ID

**用户动态关注关系表：graph\_170w\_1month**

user\_id: 用户的map\_id

friends\_id：用户关注的用户map\_id，以#隔开

follows\_count：关注数量

timestamp：关注时间