|  |  |
| --- | --- |
| CSV\_Name | Describition |
| Commits.csv | 包含提交的项目项目id，提交者id，作者id，创建时间等。 |
| Commit\_comments.csv | 用户提交项目时，其他用户对该提交的评论。包括评论内容已经与该评论相关的其他id |
| Commit\_parents.csv | 提交的父项目，也就是说是commit的原始项目。 |
| Followers.csv | Followers在GitHub的用户主页里，表示该用户的被关注者，当该用户提交任何仓库时，其followers都会收到提示。其包含关注者id、被关注者id以及关注时间 |
| Pull\_requests.csv | 在GitHub中其他用户对该项目提出的一些意见。类比于：在考试时，考试的试卷就是仓库（Repository），你的试卷会有很多错误（相当于Bug）。老师将你的试卷拿过来就当于fork。在你的卷子上做一些修改批注就类似于（git commit），最后把改好的试卷发给你（这就相当于pull\_request）而你对试卷进行改正就相当于（merge） |
| Issues.csv | 用户对该项目的一些评论，主要包含bug\_report和pull\_request |
| Issues\_commentss.csv | Issues里的相关评论。主要包括issue\_id,user\_id,comment\_id和创建时间 |
| Issue\_events.csv | Issue里的具体事件，主要是其要执行的动作（action） |
| Repo\_labels.csv | 仓库标签，由开发人员给仓库打的标签，主要用于检索。 |
| Issues\_labels.csv | Issues的标签，其主要有bug、duplicate、enhancement、good first issue、help wanted、invalid、question、wontfix |
| Organization\_members.csv | 组织成员，主要是GitHub社区的成员 |
| Project\_members.csv | 工程成员，主要是指指定工程的成员 |
| Project\_languages.csv | 工程语言，主要是编写该工程所使用的语言 |
| Project\_topics.csv | 工程topic，也就是我们所需要的tag，它主要由project\_id以及topic\_name组成，project\_id可以准确定位到指定工程，topic\_name为topic的名字 |
| Pull\_request\_comments,csv | Pull\_request的相关评论，其他用户对这个pull\_request的看法。主要包括评论内容，以及其相关id |
| Pull\_request\_commits.csv | 主要是一个连接表，对pull\_request和commits进行自然连接的表 |
| Pull\_request\_history.csv | 表示已经提交的过的pull\_request，主要包括history\_id、pull\_request\_id、创建时间以及 |
| Repo\_milestones.csv | 仓库的里程碑，开发者可以创建一个里程碑来记录自己的工程的整个生命周期 |
| Schem\_info.csv | 工程项目的版本信息，包含属性version |
| Users.csv | 用户，GitHub上的所有用户信息 |
| Watchers.csv | 查看者，在GitHub上查看过该项目的用户，其主要包含repo\_id、user\_id以及时间 |