**CS209A fall Final Project Report**

12012004 周兴雨

12012008 孙开轩

**选择的github repo：**

<https://github.com/carp-lang/Carp>

<https://github.com/alibaba/fastjson2>

**体系结构：**

对于每一项要进行分析处理的数据，都有对应的model以进行存储，即有commit, developer, issue, open\_issue, release五个类。

对于每一个model，都有对应的repository用于对数据进行查找，即有CommitRepository, DeveloperRepository, IssueRepository, OpenIssueRepository, ReleaseRepository五个类。

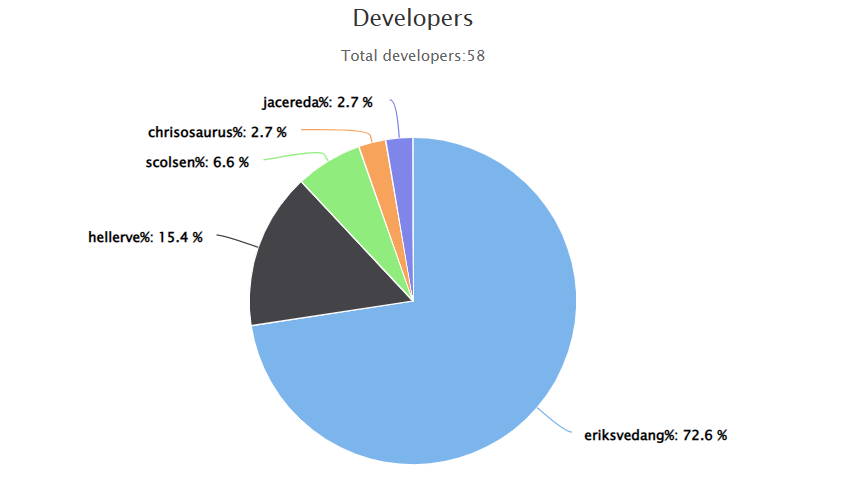
有三个controller类，ReleaseController, DeveloperController, IssueController，用于查找和处理数据，满足前端对三种数据的要求。

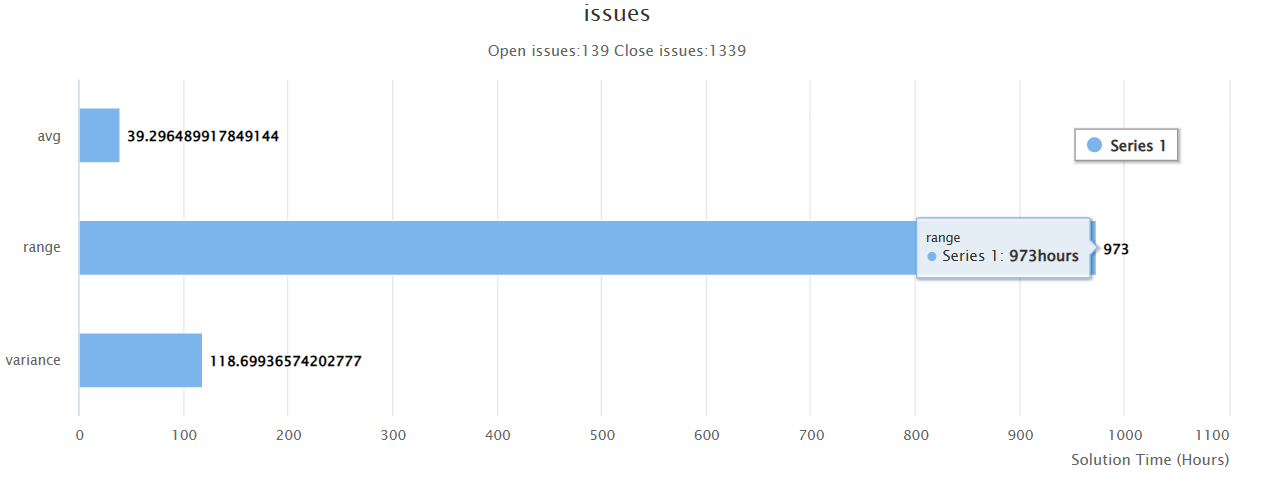
有一个service类，RepoService，用于对数据进行爬取。爬取数据使用了java的HttpRequest和HttpResponse，每次爬取数据时使用HttpRequest发送get请求，然后使用HttpResponse获取get到的数据，然后使用fastjson对获取的json格式信息进行解析，再将解析的结果存入数据库中。

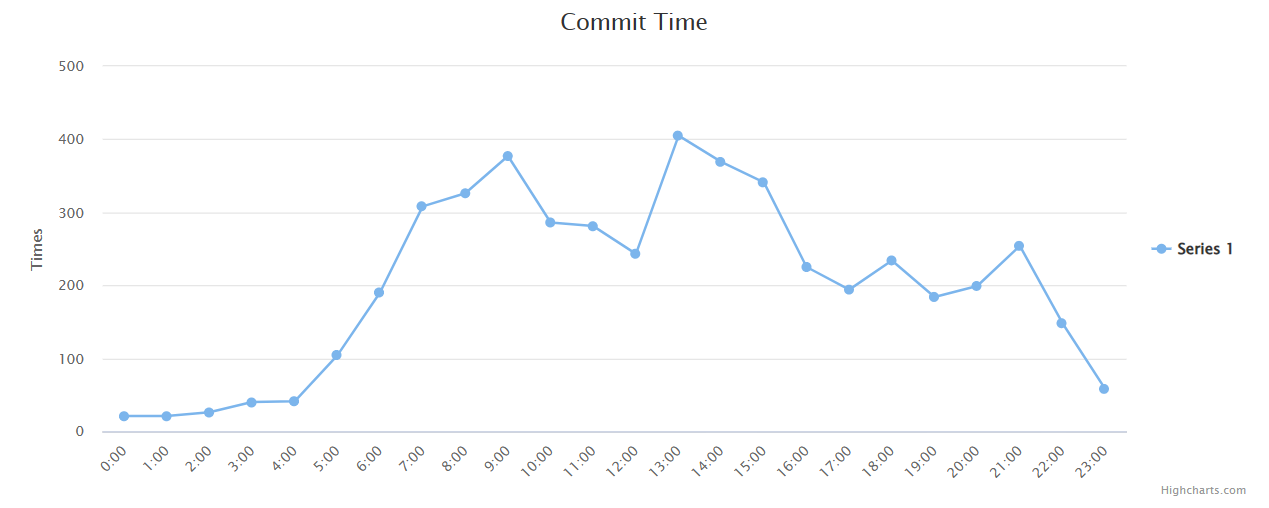
有一个前端页面index.html，用于从后端获取数据，并将数据直观地展示出来。

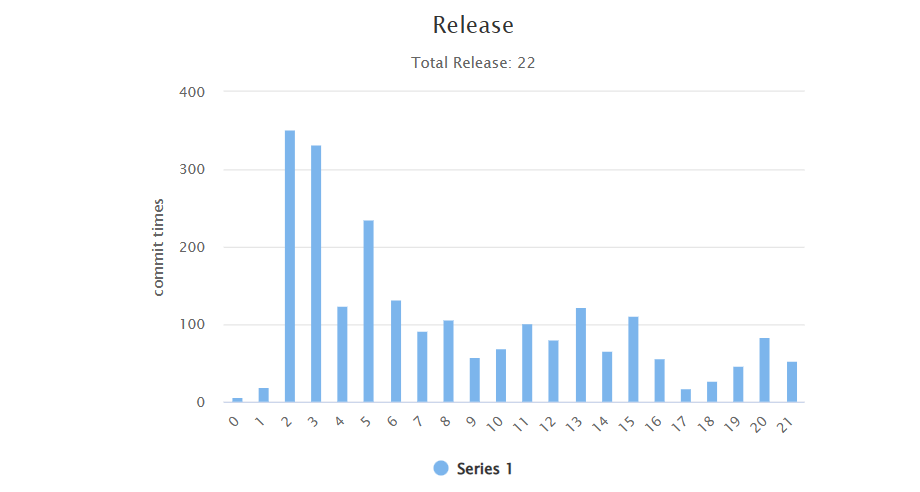
**前端效果：**

前端页面使用JavaScript，并通过HighCharts绘制数据统计图，效果如下：









**Insight：**

在我们数据处理的过程中，我们发现选取的两个repo有一个共同的特点，就是commit大部分是来自一个developer，可见一个项目中大部分工作都是由小部分成员完成的。

通过对commit时间的分析，我们发现这两个repo的commit时间均为北京时间的白天，可见这两个项目都是于工作时间完成的工作项目。

通过对issue的分析，我们发现这两个repo的issue问题解决时间的平均数都小于48小时，可见问题大部分回复的比较及时，但是极差较大，说明存在问题回复不及时的情况，可以改进。

通过对每一个release间commit的分析，我们发现在fastjson2中，release间的commit有很多，可见这个项目每个版本更新的内容较多。而Carp的release间commit较少，可见这个项目更新幅度较小。