软件描述文档

Version1.0

2018.12.27

Tiki-Taka Visual足球联赛数据可视化平台

郭智溢

1简介

11文档目的

本软件的设计文档提供了Tiki-Taka Visual完整的设计细节与需求分析。此文档的目的是为了规整整个软件的功能，并为以后的改进提供范本。

12设计概览

Tiki-Taka Visual软件允许用户查询2010-2017年欧洲五大联赛的球队、单场球赛数据，同时提供截至2018年的最新球员数据联网查询功能。

整个软件的基本结构为pyqt5编写的客户端配合python+sqlite3数据库来进行的数据处理以及pyecharts可视化工具来共同构建，而可视化的视图为Html文件，通过pyqt webengine来加载。

13文档概览

本文确定了Tiki-Taka Visual的架构、功能、数据、组件以及用户接口设计。整个系统为Layered Architecture结构，有完整的UI界面，并在底层有数据库查询、联网查询的功能，细节将在后文进行描述。

14参考文献

Python 《Beginning python 3》

D3数据可视化开源库文档

PyQT 5.0 技术文档

《数据库系统概念第六版》

Scrapy文档 <https://github.com/scrapy/scrapy>

Echarts文档 [**http://echarts.baidu.com/feature.html**](http://echarts.baidu.com/feature.html)

2系统概览

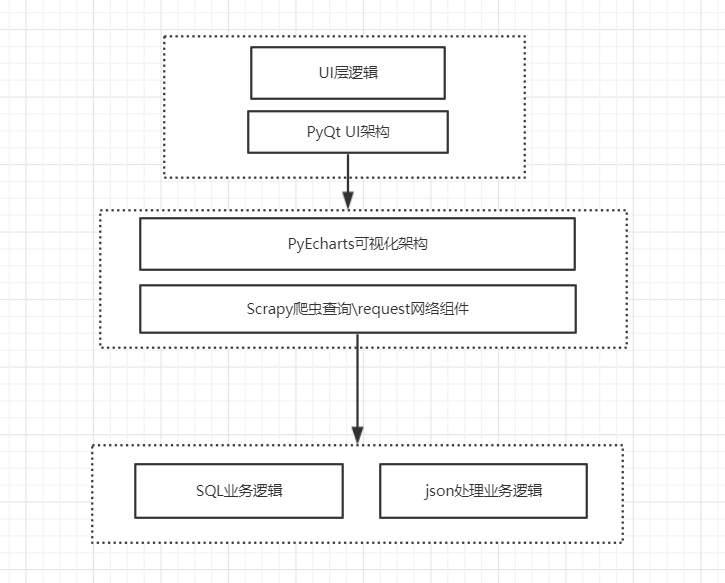
Tiki-Taka Visual软件允许使用者根据自定义的筛选信息，来筛选单场比赛或者球队单赛季的数据，并有可视化视图来显示这些数据。软件的功能包括查询五大联赛2010-2017的积分榜、射手榜、助攻榜，并且能够以联赛、年份、轮数为筛选条件查看单场比赛的数据统计，并提供可交互的可视化视图，同时提供2010-2017年在五大联赛出战的所有球队的赛季数据统计，并提供可交互的可视化视图，最后为联网查询截至2018年的球员数据以及技术指标统计，数据来源于FIFA，该功能可以联网获取球员的头像、国家、效力的俱乐部的标志图片，以及技术特色的数据。

本软件需要在联网的情况下于PC端python环境下使用，否则无法完全使用所有的功能。

3系统架构

31架构设计

采用分层模式Layered Architecture，每一层的功能简述如下图



32 系统分解描述

321 数据处理模块data\_process

包含所有的数据处理代码、脚本，并有配套的sql查询结果转为json的流程工具（部分配置代码于数据使用json可以做到轻量化、快速化）

322 数据可视化模块chars

该模块包含整个软件所需要的pyecharts模块的调用，并根据业务需求做了相当一部分改进

4数据设计

41数据描述

整个软件使用的数据分为零时数据和永久数据。

永久数据方面，后台存放的数据有单场赛事数据、积分榜（射手榜、助攻榜）数据、球员信息数据，其中积分榜（射手榜、助攻榜）数据使用了自己编写的定制爬虫来爬取。同时还有单场赛事、球队赛季信息的查询索引信息的config文件，这部分文件用于快速形成数据库索引语句，同时用于部分ui界面信息加载的配置文件。

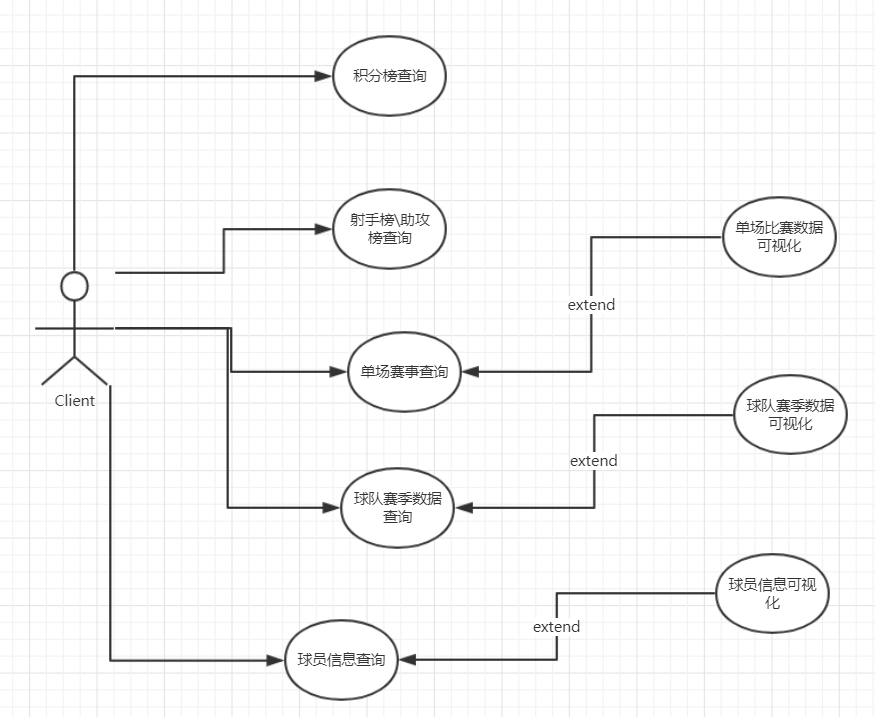
临时数据则是通过爬虫和球员信息数据库中的部分信息来从网络上临时下载的球员照片、球员国家旗帜和效力俱乐部的logo，以及部分球员信息。

42数据库信息

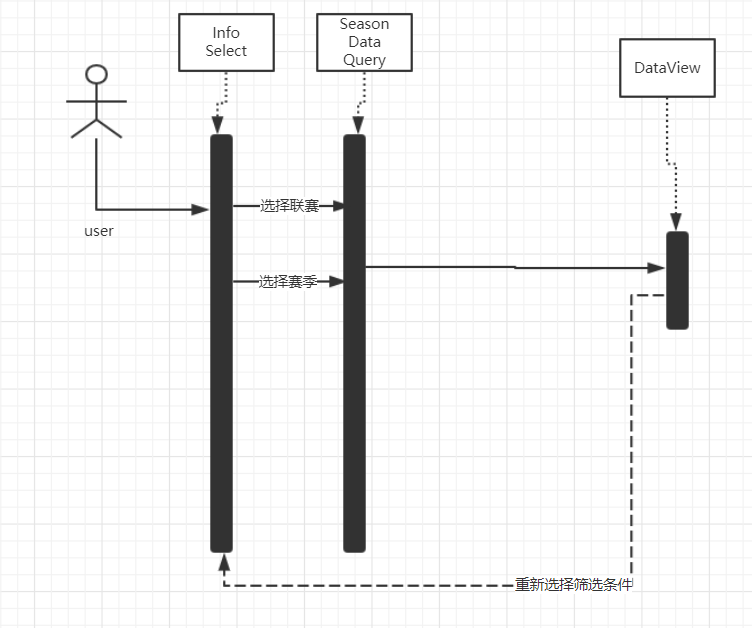
|  |  |
| --- | --- |
| 数据库名称 | 数据库描述 |
| game.db | 单场赛事数据，其中有5个table分别代表5大联赛，每个table存放的每个联赛2010-2017的所有赛事数据 |
| league\_table.db | 五大联赛积分榜以及球队赛季信息统计 |
| player\_data.db | 包含部分不易获取的球员数据，作为本地索引 |

5模块设计

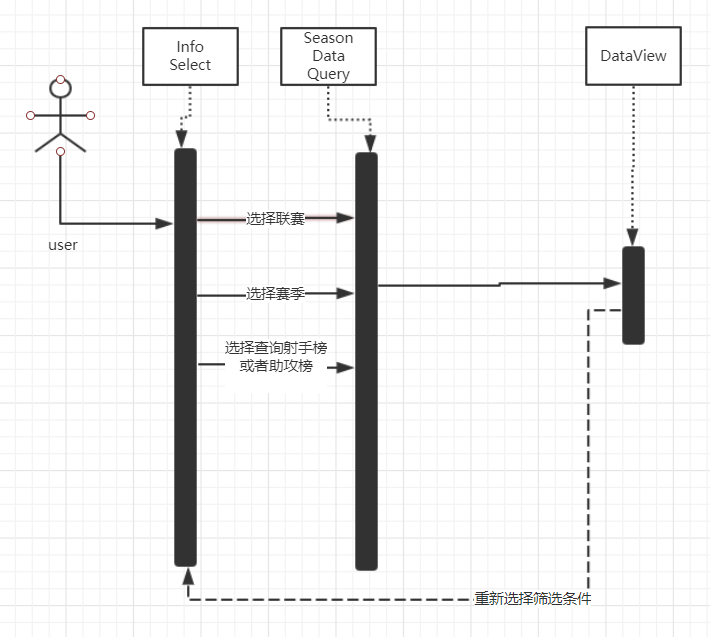
51系统设计用例图如图所示



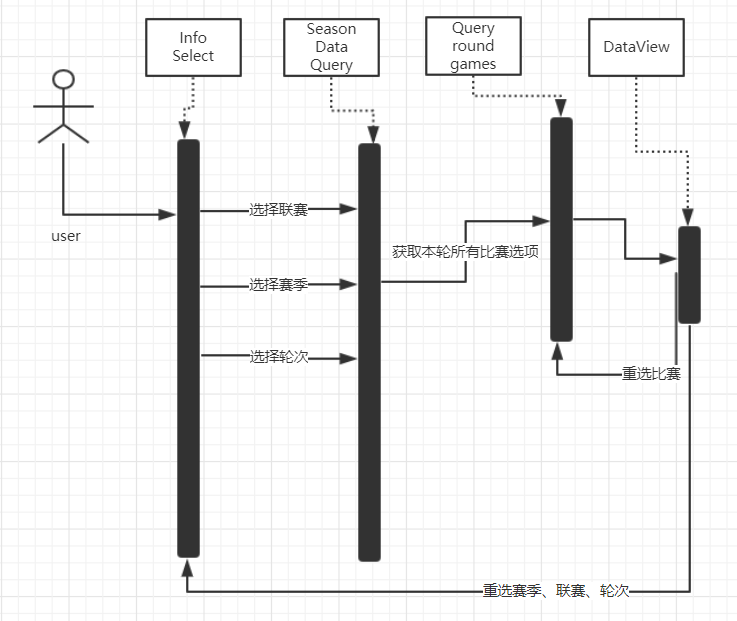
521积分榜查询



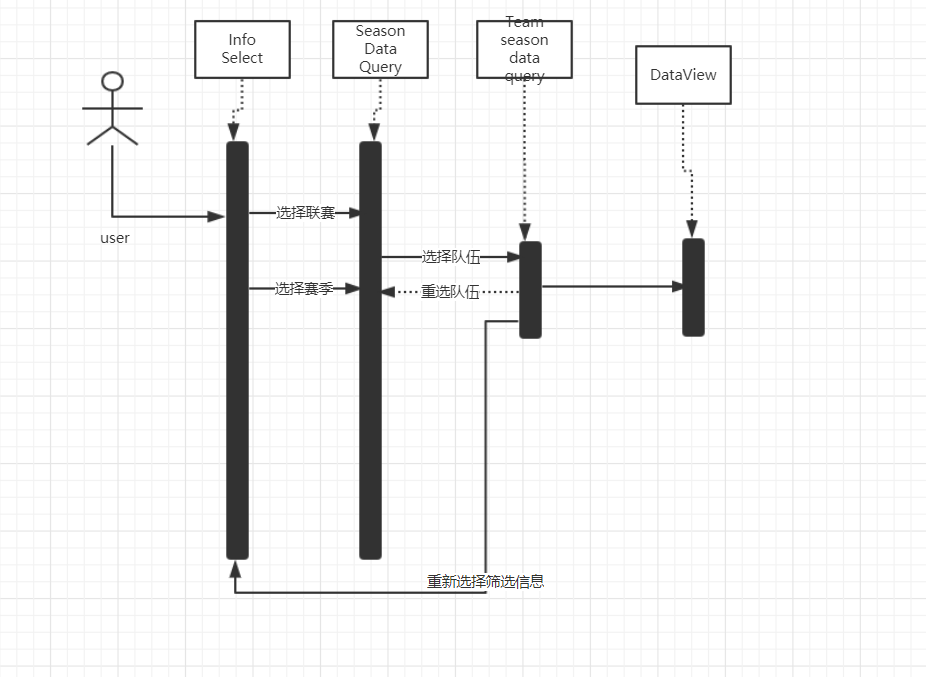
522射手助攻榜



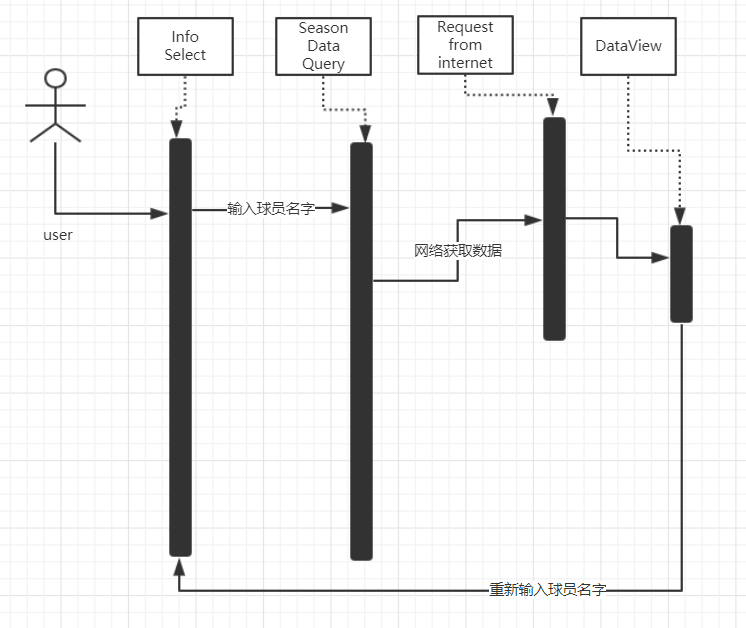
523单场赛事查询



524球队赛季数据查询



525球员信息查询



52功能说明

521积分榜查询

|  |  |
| --- | --- |
| 功能编号 | 01 |
| 功能名称 | 积分榜查询 |
| 参与者 | 用户 |
| 输入 | 选择联赛门类与联赛年份 |
| 输出 | 格式化输出根据筛选条件从数据库中调取的联赛积分榜数据 |
| 基本事件流 | 1. 选择联赛 2. 选择赛季 3. 指定区域展示积分榜 |

522射手榜/助攻榜查询

|  |  |
| --- | --- |
| 功能编号 | 02 |
| 功能名称 | 射手榜/助攻榜查询 |
| 参与者 | 用户 |
| 输入 | 选中的联赛、年份、显示射手榜或者助攻榜的标签 |
| 输出 | 格式化输出根据筛选条件从数据库中调取射手榜或者助攻榜 |
| 基本事件流 | 1. 选择联赛 2. 选择年份 3. 选择显示积分榜或者助攻榜 |

523单场赛事查询与可视化视图生成

|  |  |
| --- | --- |
| 功能编号 | 03 |
| 功能名称 | 单场比赛数据可视化查询 |
| 参与者 | 用户 |
| 输入 | 选中的联赛、年份、轮次、单场比赛 |
| 输出 | 单场比赛的基本数据，以及比赛射门、犯规等数据的可视化视图 |
| 基本事件流 | 1. 选择联赛 2. 选择年份 3. 选择轮次 4. 选择比赛 5. 操作生成的可视化视图 |

524球队2010/2011-2017/2018赛季数据查询与可视化视图生成

|  |  |
| --- | --- |
| 功能编号 | 04 |
| 功能名称 | 球队赛季数据查询与可视化视图分析 |
| 参与者 | 用户 |
| 输入 | 选中的联赛、选中的球队 |
| 输出 | 球队名称与球队所在国，球队赛季数据统计可视化视图 |
| 基本事件流 | 1. 选择联赛 2. 选择球队 3. 操作生成的可视化视图 |

525球员数据查询与可视化视图生成

|  |  |
| --- | --- |
| 功能编号 | 05 |
| 功能名称 | 球员数据查询与可视化视图的生成 |
| 参与者 | 用户 |
| 输入 | 球员名字 |
| 输出 | 球员数据可基本信息、网络加载的球员头像、球员国家旗帜logo、球员效力的俱乐部logo |
| 基本事件流 | 1. 输入球员名称 2. 加载网络数据 3. 显示可视化视图 |

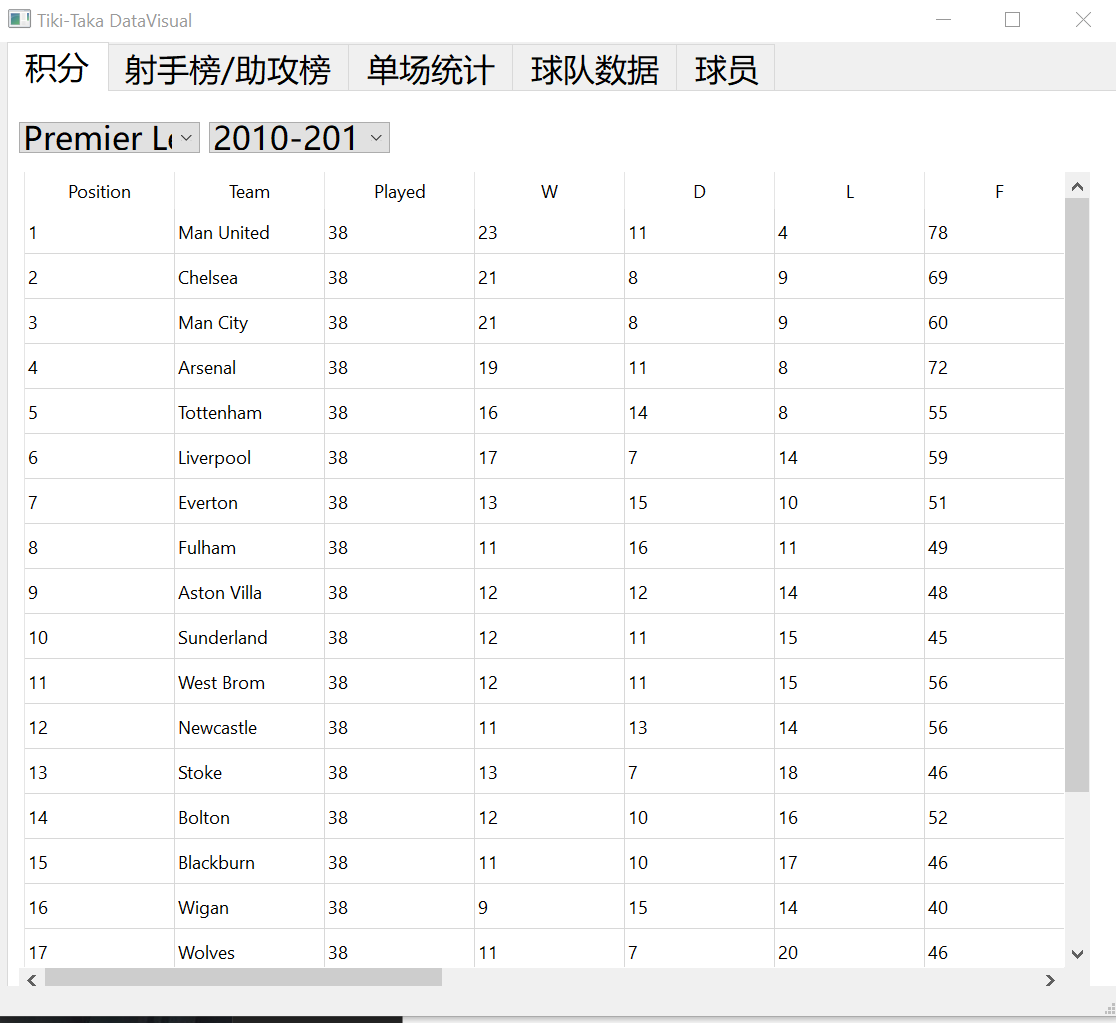
6UI设计概览

61分界面简介

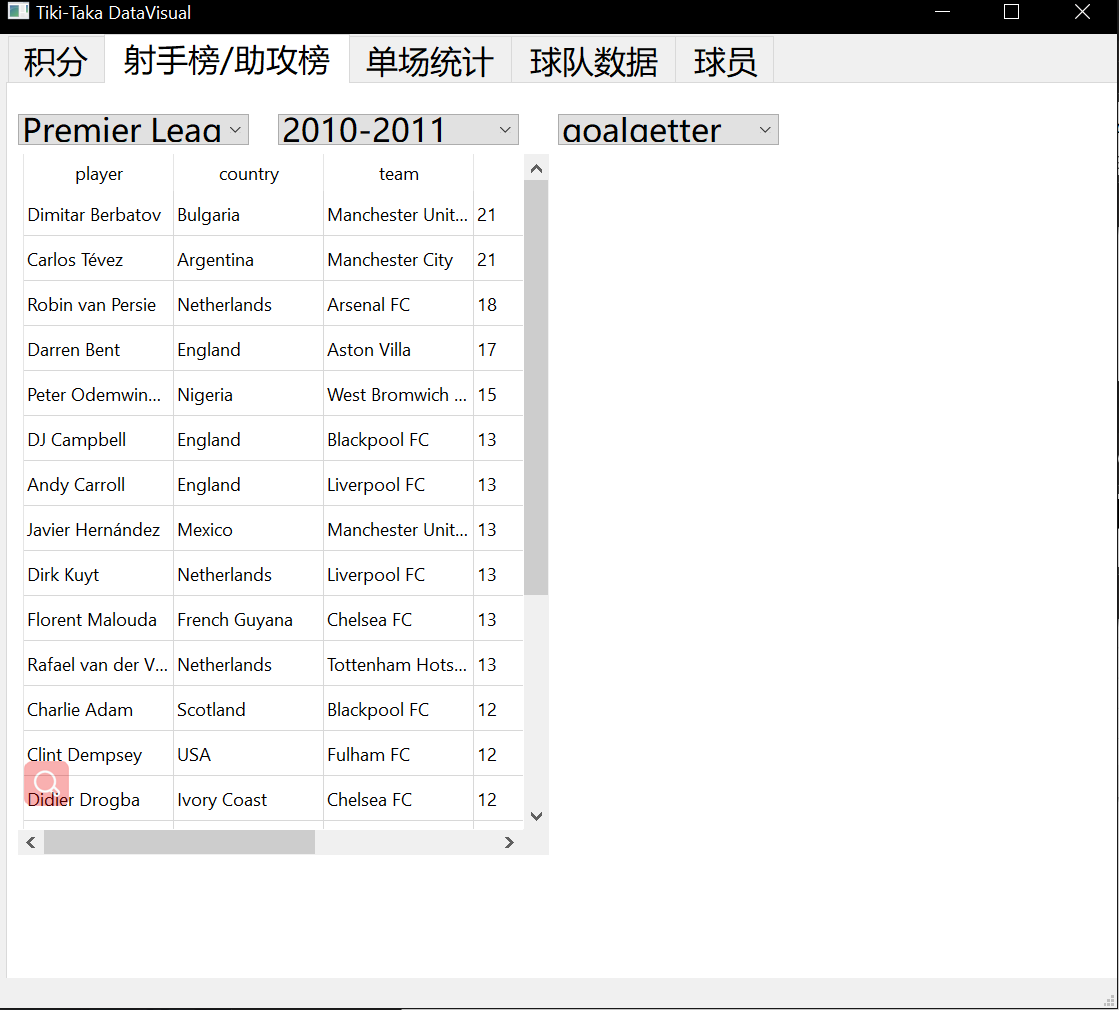
用户界面主要分为以下几个：积分榜界面、射手榜\助攻榜界面、单场比赛数据统计界面、球队赛季数据界面、球员数据界面

62界面样式

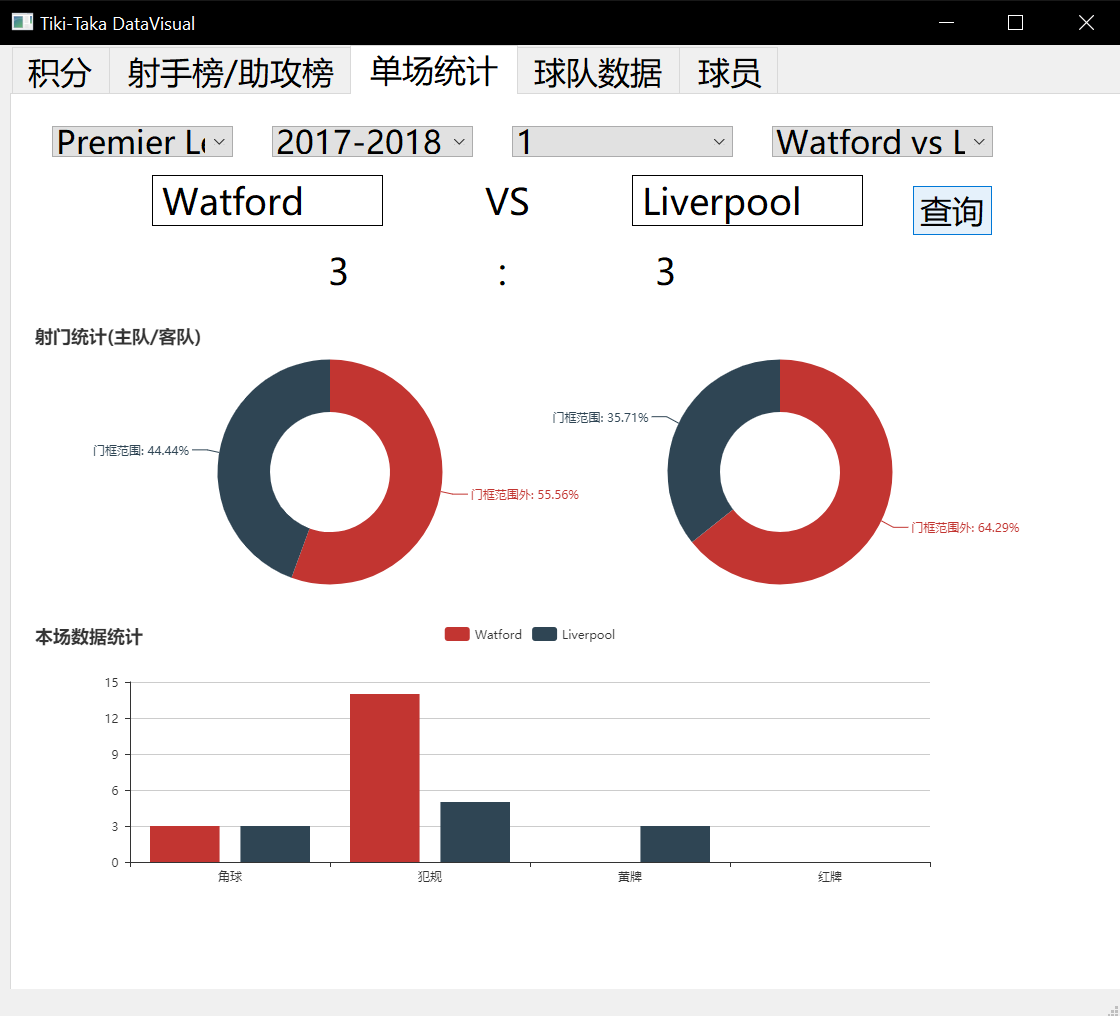
621积分榜界面



622射手榜/助攻榜界面



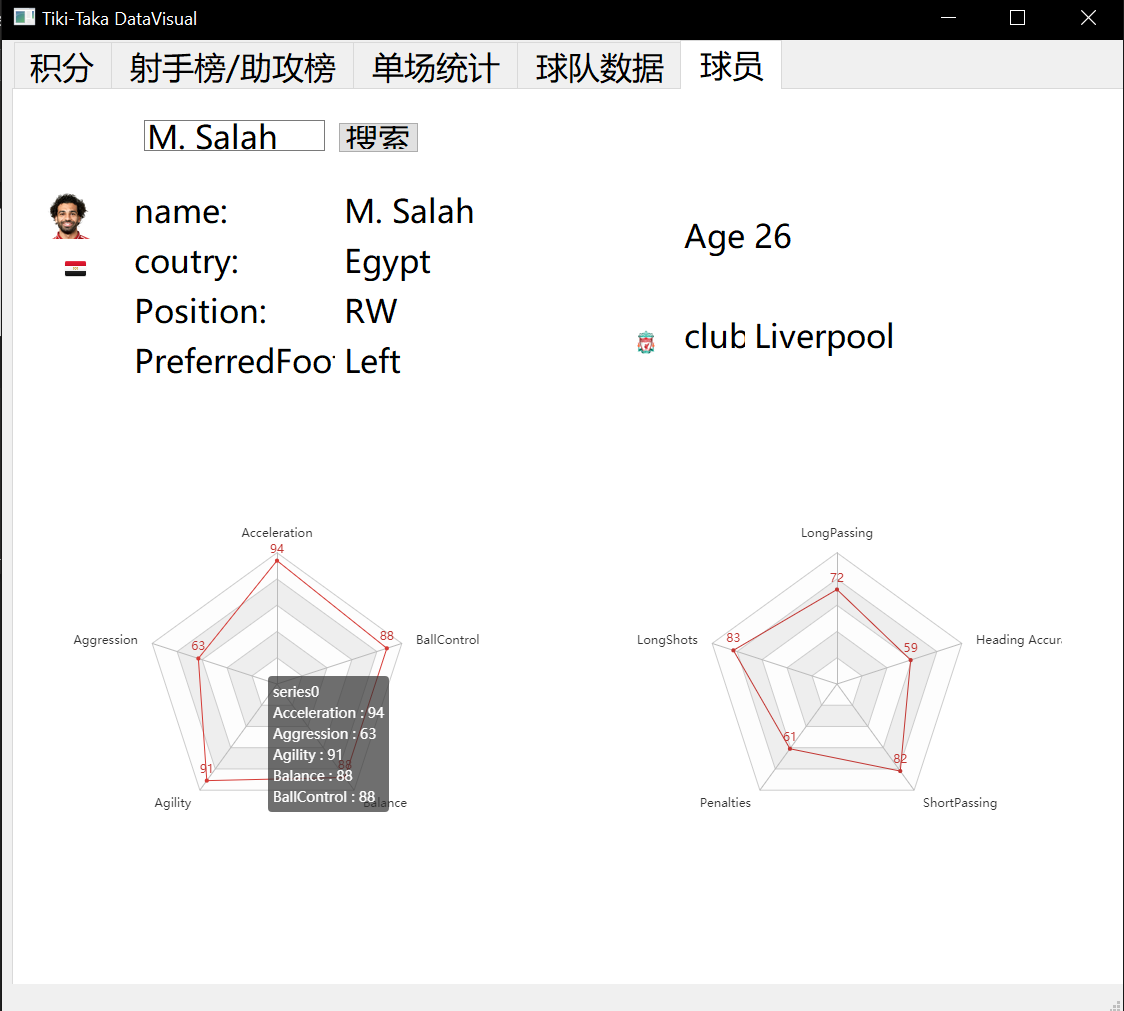
623单场统计界面



624球队赛季数据界面



625球员信息界面



7项目开发流程简介

71项目开发平台

整个项目基于python3.5，使用的IDE为pycharm community 2018，使用github来对整个开发做跟踪。

711开发环境依赖

dependencies:

- blas=1.0=mkl

- certifi=2016.2.28=py35\_0

- chardet=3.0.4=py35\_0

- mkl=2017.0.3=0

- numpy=1.13.1=py35\_0

- pandas=0.20.3=py35\_0

- pip=9.0.1=py35\_1

- python=3.5.4=0

- python-dateutil=2.6.1=py35\_0

- pytz=2017.2=py35\_0

- requests=2.14.2=py35\_0

- setuptools=36.4.0=py35\_1

- six=1.10.0=py35\_1

- vc=14=0

- vs2015\_runtime=14.0.25420=0

- wheel=0.29.0=py35\_0

- wincertstore=0.2=py35\_0

- pip:

- click==7.0

- dukpy==0.2.2

- future==0.17.1

- javascripthon==0.10

- jinja2==2.10

- jupyter-echarts-pypkg==0.1.2

- lml==0.0.2

- macropy3==1.1.0b2

- markupsafe==1.1.0

- pillow==5.3.0

- prettytable==0.7.2

- pyecharts==0.5.11

- pyecharts-javascripthon==0.0.6

- pyecharts-jupyter-installer==0.0.3

- pyecharts-snapshot==0.1.9

- pyqt5==5.11.3

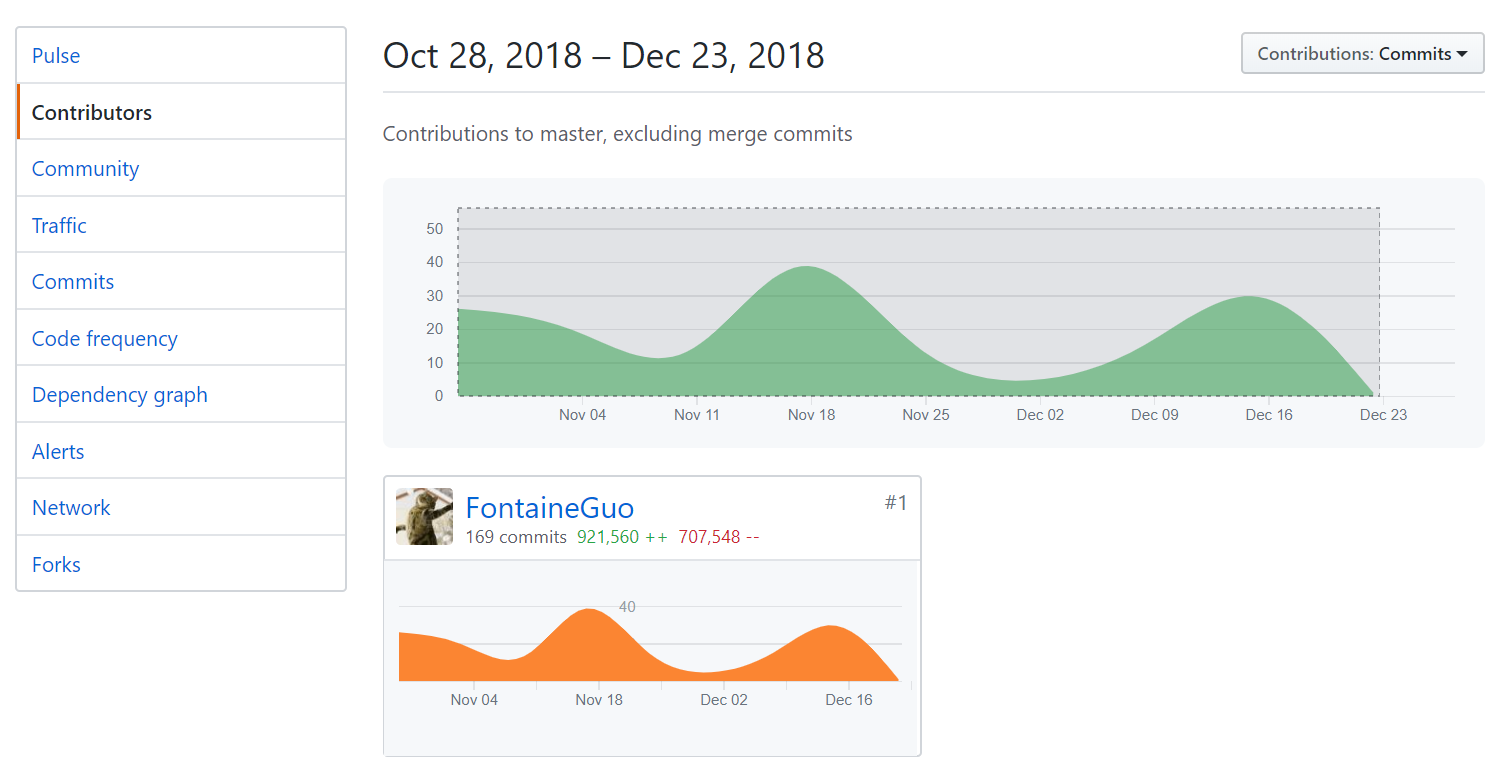
- pyqt5-sip==4.19.13

- pyqt5-tools==5.11.2.1.3

- python-dotenv==0.10.1

72开发进度简述

开发时间2018.10.28-2018.12.23，共计169次提交。Github统计报告如图



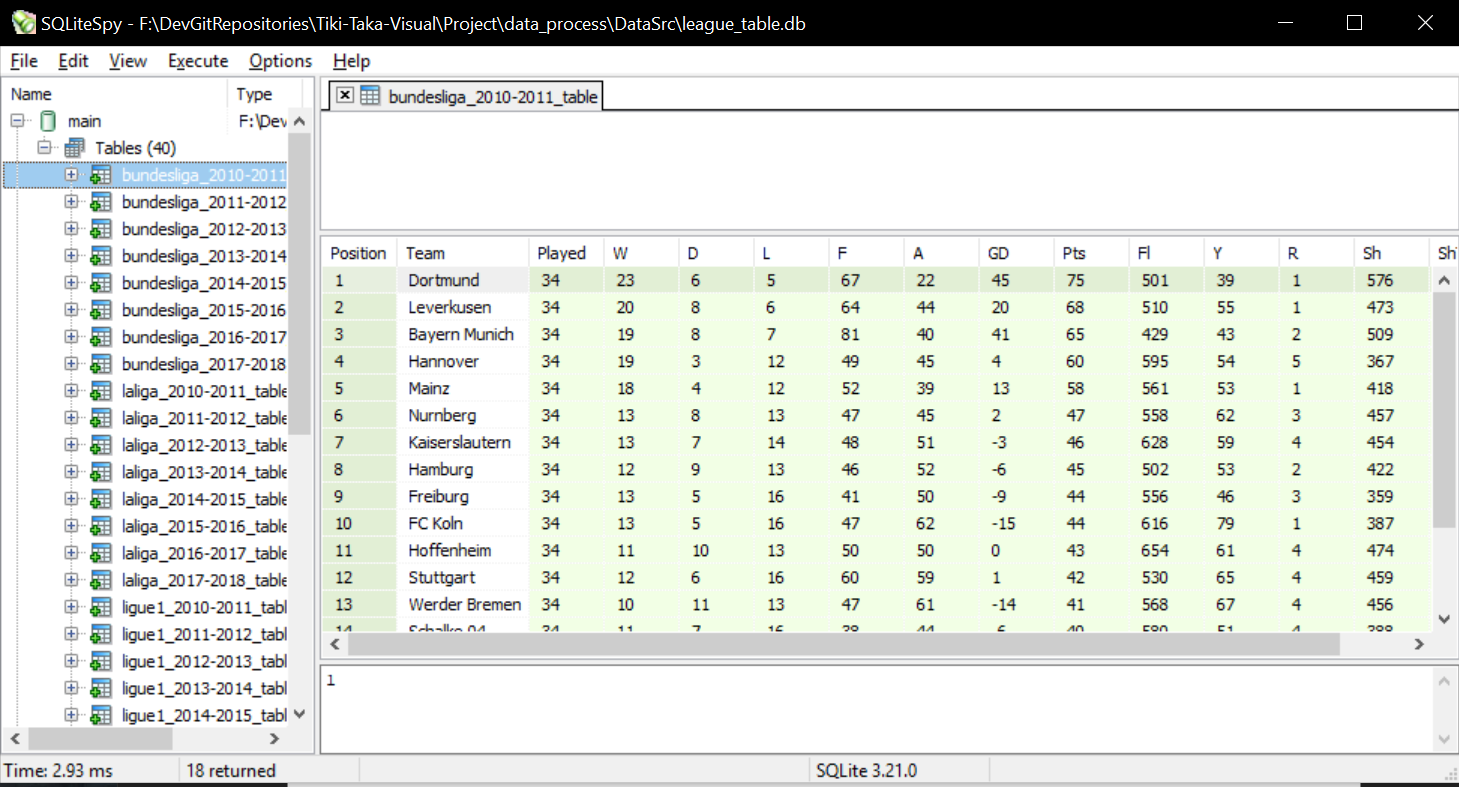
8额外的工作

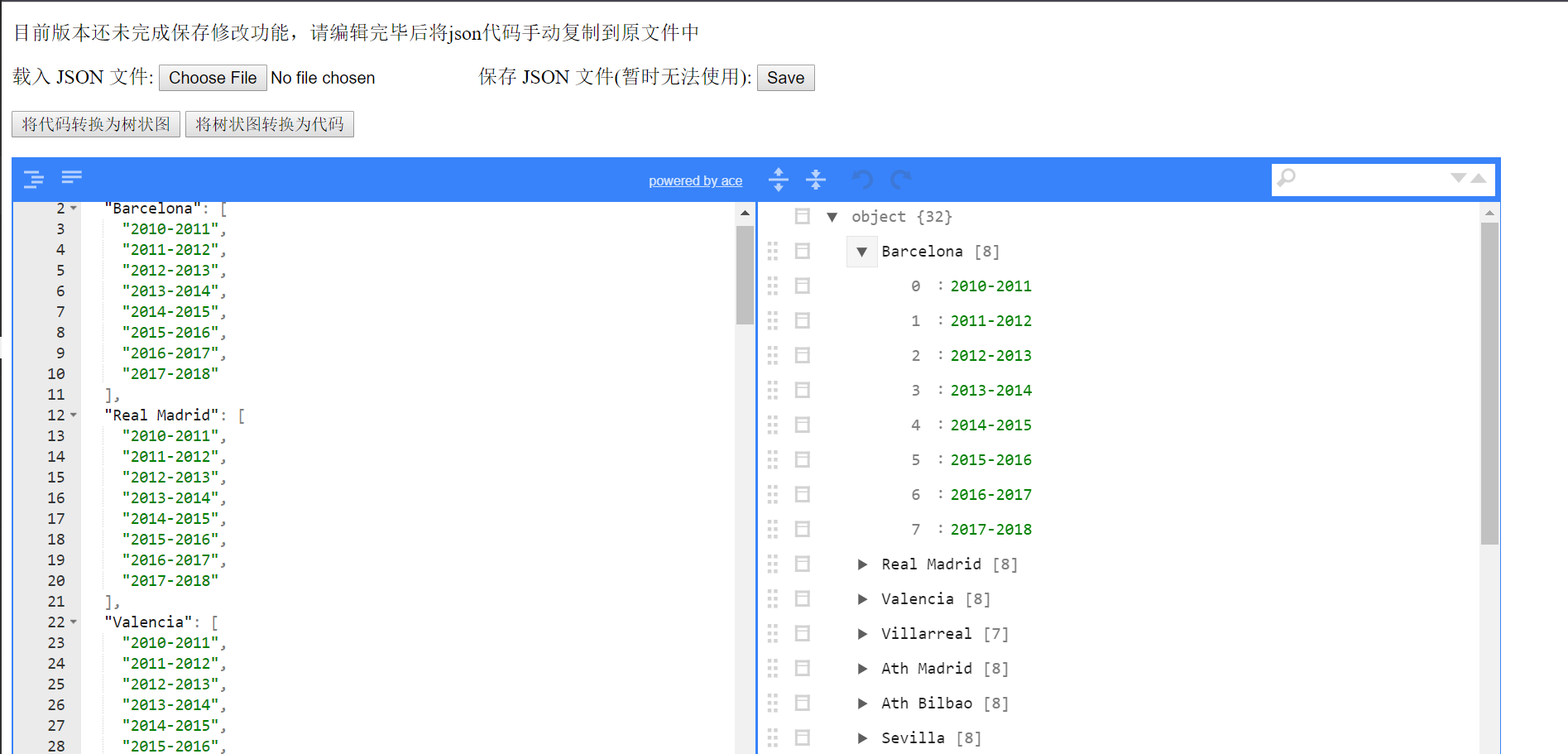
81数据爬取

基于scrapy开源爬虫架构，为项目定制了专用的爬虫用于爬取积分榜、助攻榜、射手榜的数据。

82数据导入与清洗

大部分爬取或者公开的数据均基于csv文件格式，通过数据处理脚本将部分数据固化在了基于sqlite3的sql数据库中，同时需要大量查询时间的部分数据被提取、清洗后使用脚本将其转化为了json配置文件，同时为了格式化json，使用开源的插件定制了专用的json查看工具jsonEditor。





83专业数据对照

很多爬取的数据表格都是英文的标注，且大部分都是缩写，因此做了一些额外的工作去查询这些缩写代号的具体含义，并做单独的收集整理。

