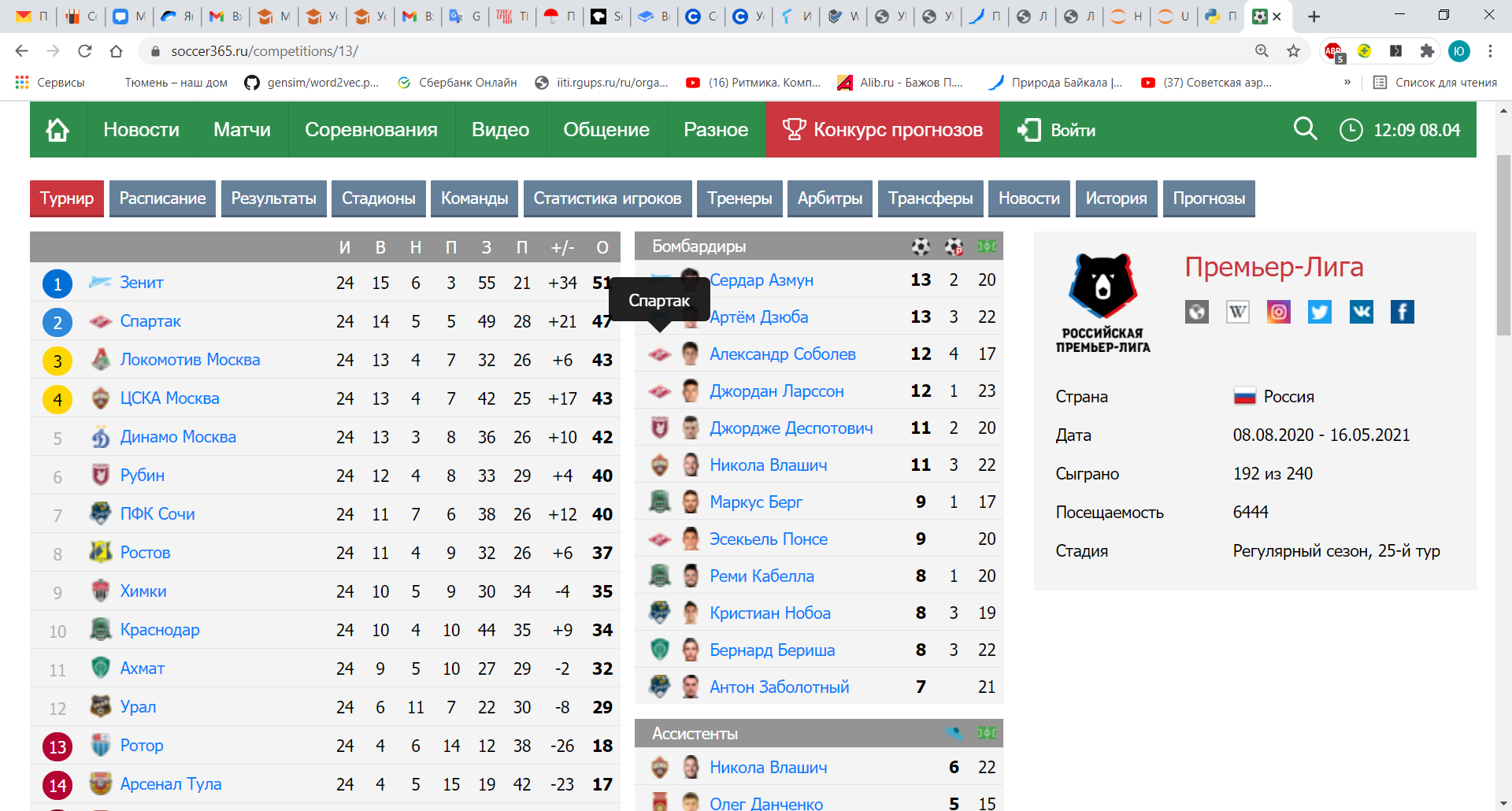
**Python: web-скрапинг, извлечение данных**

**Задание**

Имеются таблицы с данными о футбольных матчах Премьер-лиги России, опубликованные на сайте по адресу: <https://soccer365.ru/competitions/13/>



Требуется написать приложение, реализующее следующий функционал:

1. Извлечь данные по таблицам Бомбардиры, Ассистенты, Штрафники.  
   Внимание! Один и тот же игрок может присутствовать в нескольких таблицах.
2. Собрать из этих данных датафрейм со следующими столбцами:

|  |  |  |
| --- | --- | --- |
| **Имя поля** | **Значения, тип** | **Источник** |
| Команда | Строка (Зенит, Спартак и т.д.) | Иконка рядом с аватаром игрока в таблицах Бомбардиры, Ассистенты, Штрафники. |
| ФИ игрока | Строка (Артём Дзюба и и.д.) | Таблицы Бомбардиры, Ассистенты, Штрафники |
| Роль | Категория: бомбардир, ассистент, не определено | -\\- |
| Голы | Integer | Таблица Бомбардиры |
| Пенальти | Integer | Таблица Бомбардиры |
| Пасы | Integer | Таблица Ассистенты |
| Матчи | Integer | Таблицы Бомбардиры, Ассистенты, Штрафники |
| Штрафные | Булево: 1 – имеются штрафные, 0 – не имеются. | Таблица Штрафники |
| Fair play | Integer | Таблица Штрафники |
| Желтые карточки | Integer | Таблица Штрафники |
| 2ЖК | Integer | Таблица Штрафники |
| Красные карточки | Integer | Таблица Штрафники |

1. Получить по датафрейму следующую аналитику:
2. Первая тройка команд по числу забитых голов с выводом их числа.
3. Первая тройка команд по числу желтых карточек.
4. Список игроков, которые участвовали не во всех играх своей команды. Число игр команды определить по максимальному числу матчей ее игроков.
5. Доля пенальти по отношению к числу голов для каждой команды.
6. Корреляция числа голов с количеством очков команды. Очки взять из первой таблицы на странице по рейтингу команд.

Материал для справки:

[https://python-scripts.com/beautifulsoup-html-parsing#method-select](https://python-scripts.com/beautifulsoup-html-parsing" \l "method-select)

Пример парсинга отзывов с Banki.ru: файл *пример\_BS4.py*

Пример парсинга таблиц: *файл парсинг\_таблиц.html*