

数据解析概述

在上一章中, 我们基本上掌握了抓取整个网页的基本技能. 但是呢, 大多数情况下, 我们并不需要整个网页的内容, 只是需要那么一小部分. 怎么办呢? 这就涉及到了数据提取的问题.

本课程中, 提供三种解析方式:

1. re解析
2. bs4解析
3. xpath解析

这三种方式可以混合进行使用, 完全以结果做导向, 只要能拿到你想要的. 用什么方案并不重要. 当你掌握了这些之后, 再考虑性能的问题.