数据源方面讨论汇总：

最后想要进行页面汇总展示的内容：

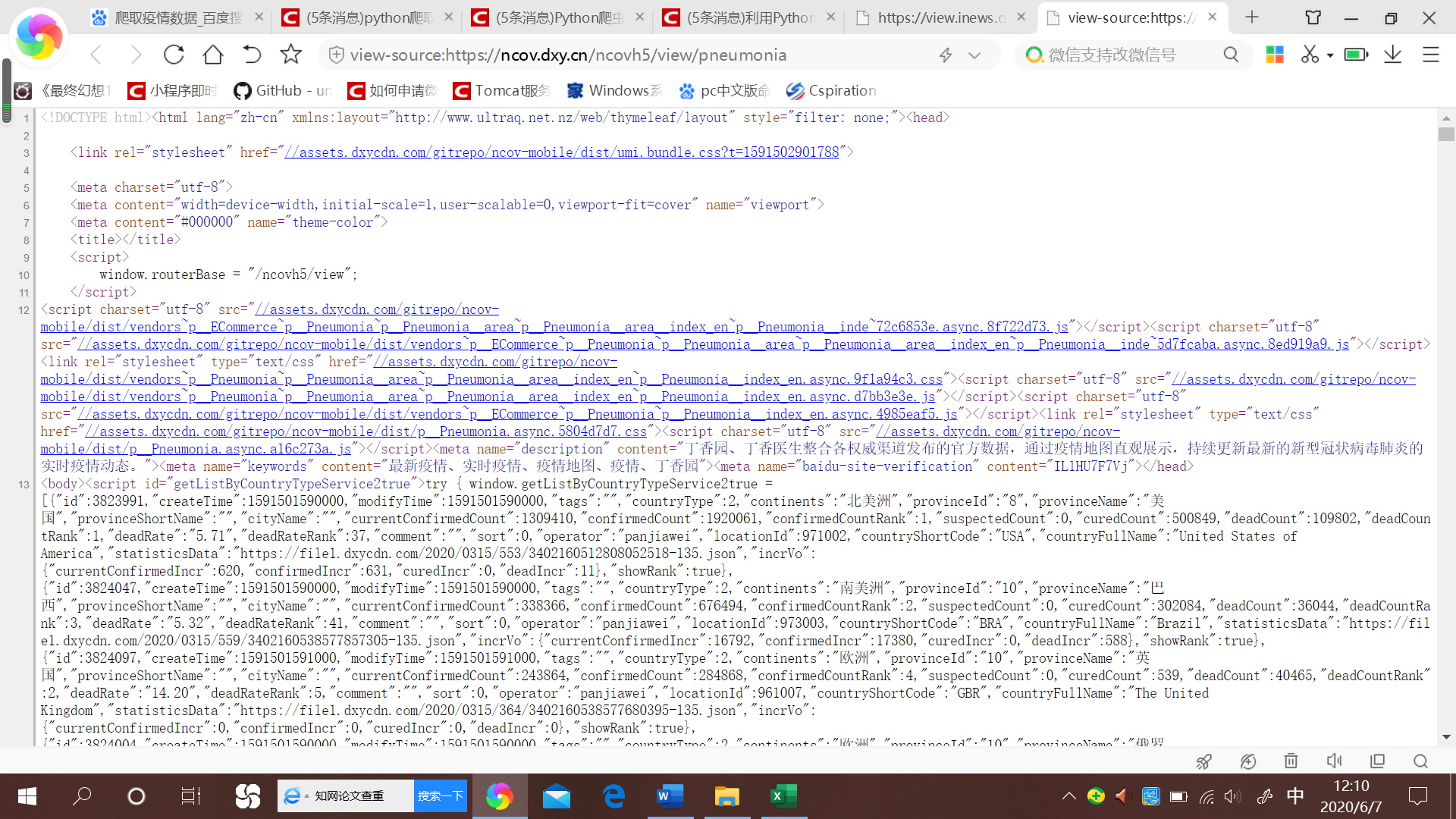
* 全球方面疫情确诊情况（精确到天，以国家为划分）——可以考虑以柱状图或热力图的方式来进行展示，人力观察全球疫情各地区数据来确定不同颜色（或柱状图刻度）划分区间
* 国内疫情确诊情况（精确到天，以省份作为划分）——与以上采取相同展示形式
* 全球各国病死率和治愈率分析（精确到周）——根据汇总进行前十位排序，柱状图中以不同颜色分别代表病死率和治愈率
* 前十位国家累计死亡人数（精确到天）——既要展示数据量又要展示死亡数在全球死亡数中所占比例，所以采用玫瑰图形式，半径长短区分数值大小，角度大小表示占比
* 前五位国家某月累计确诊人数按天展示——可采用动态图形式，或者各国分别折线图形式展示趋势

支付宝疫情数据：

在提出的几个数据源中最完整最全面，在划分上最细致，地区上可以到定位点所在小区疫情方面信息。但是安全防护措施很高，除去登录之外还有换IP的手机验证码验证流程等保护用户信息的措施，且本身支付宝原则上不允许爬虫爬取应用内数据。另外，它的疫情数据显示是根据所使用设备的定位来实时刷新数据显示的，爬取时从页面抓取数据，需要一次性获取全球范围的所有地域细分数据较为困难。所以将这条先否决。

丁香园疫情数据：https://ncov.dxy.cn/ncovh5/view/pneumonia

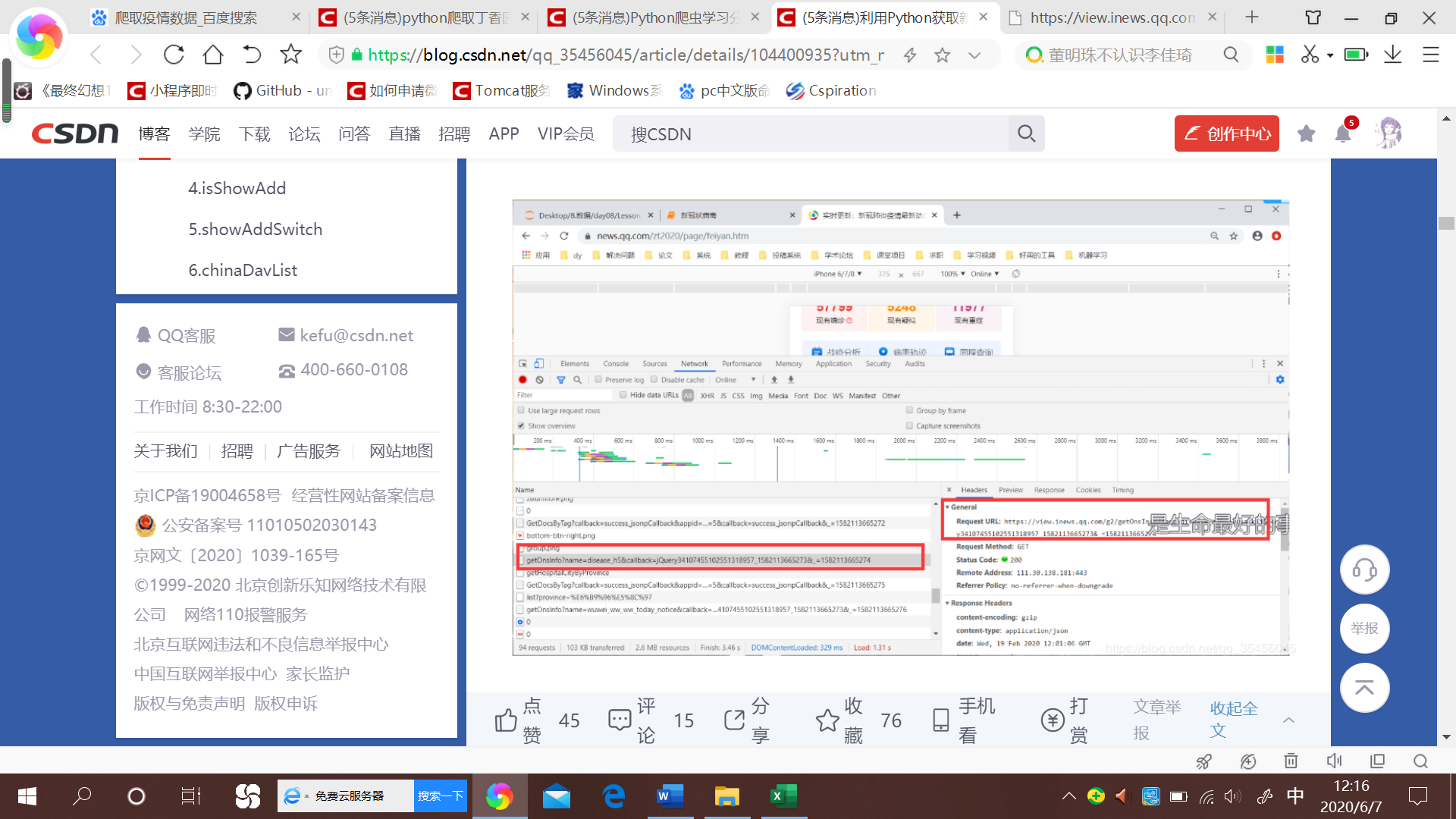
丁香园的数据按天进行更新，也都是汇总类型数据，在需求上与以上我们想要做的展示数据基本是符合的，它的疫情数据可以直接页面右键显示源代码来进行分析爬取，方便快捷，以时间流的方式进行更新。



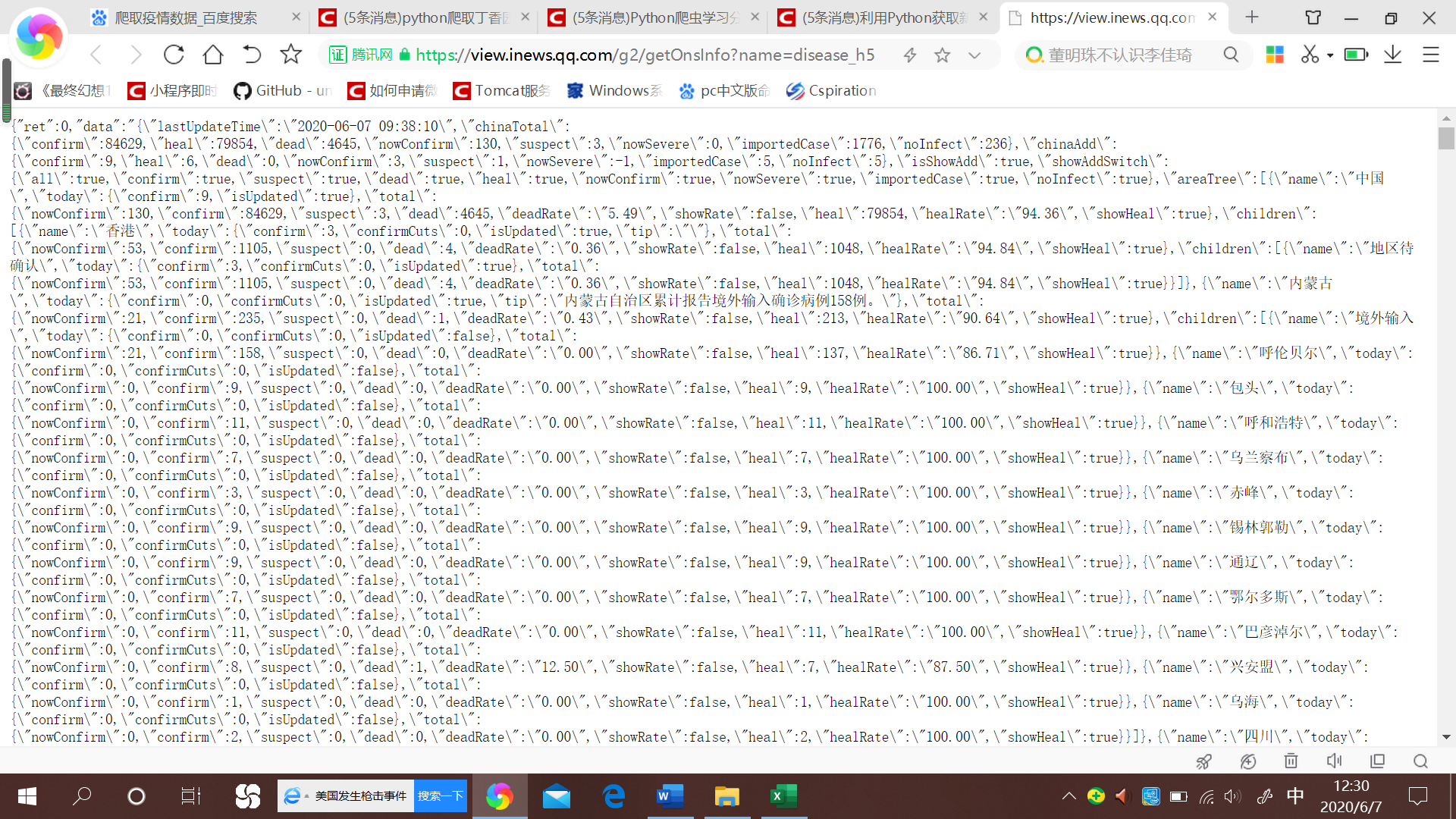
但它的数据主要为统计汇总型数据，要查找实时数据采用这类数据源很方便，但想要看之前某段时间内的数据的话，这类数据源就很难满足要求，因为它不会保留之前某天的具体新增数据，而是全部加在一起的累计数据。

腾讯和网易官方维护的疫情数据：

这两个数据源的获取方式很类似，数据源类型也几乎相同，所以放在一起说。它们的疫情数据是通过网页请求其他url，返回json格式的数据再根据用户需求渲染到网页上显示。检查开发者工具中的Network项可以得到



从这个部分的urlRequest中，获得数据源，大致如下：



不过腾讯的数据源也需要处理一个用户登录操作，试验上来看有时需要验证码有时不用，而网易是直接将数据源接口给出来的，不需要登录也不检验ua，考虑到数据类型更新频率都相似所以优先选取网易。

初步结果：

我们目前认为可以综合使用网易和丁香园的数据，丁香园部分数据在展示统计类型可视化时更方便，而网易的更细致。我们在接下来的几天里会把这两个网站的数据源分别先爬取少量下来，用这少量数据先做出几张可视化html文件给前端页面编写的同学，方便她们先用这个开始前端页面的编写工作。后台数据爬取和可视化的同学会同时对这两个网页进行大量数据的爬取，并尝试按照我们提出的需求进行可视化操作，看看这两种哪种更合适，或者是两种类型综合一起使用。