地震数据可视化项目技术报告

1. 项目目的

本项目目的是根据中国地震信息网上的地震历史数据，利用python做一个地震小动画，让人直观的看到地震时间、地震位置以及震级。

1. 项目设计

本项目主要分两个阶段来进行。第一阶段从中国地震信息网上获取数据信息，因为该网站未提供直接下载数据表格的链接，因此需要使用python网络爬虫来爬取地震信息。第二阶段是根据爬取的信息进行可视化处理，通过文件加载，再使用pygame中的图形绘制以及屏幕刷新机制来实现图像的变动。

1. 项目实现

为了获取信息，先要引入requests库来实现网络连接，主要使用的函数是r=requests.get(url)，作用是返回一个包含服务器资源的Response对象，其中包含从服务器返回的所有的相关资源，再通过r.text将相关资源全部转换为字符串，下面给出获取字符串的函数gethtmltext(url)：

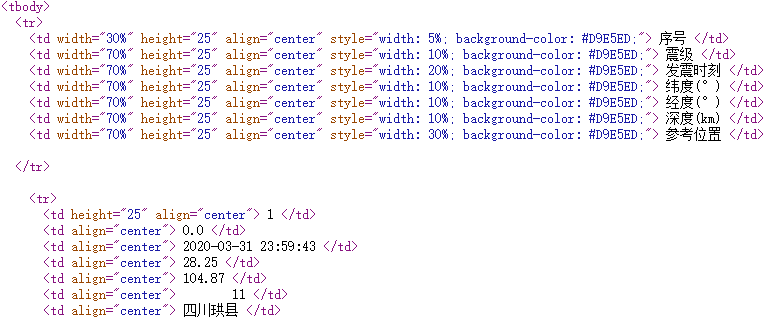
def gethtmltext(url):  
 try:  
 r = requests.get(url, timeout=30)  
 r.raise\_for\_status()  
 r.encoding = r.apparent\_encoding  
 return r.text  
 except:  
 return "产生异常"

这里通过try…except处理了网络连接的两种异常情况：超时和连接失败。通过r.apparent\_encoding将资源编码转换为合理形式，再转换为文本字符串作为返回值。

接下来引入beautifulsoup库来对返回的HTML格式的文本字符串做进一步处理。BeautifulSoup函数将复杂HTML字符串转换成一个复杂的树形结构，每个节点都是Python对象。

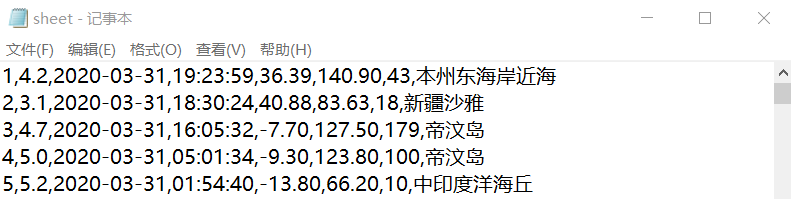
sp = BeautifulSoup(txt, "html.parser")

通过查看网页的源代码可以得知，包含地震信息的标签<tag>是<tbody>:



for child in sp.tbody.find\_all()

直接找到<tbody>标签的所有子标签，遍历所有子标签，把每个子标签组成一个列表，再把所有列表合成一个大列表，通过去头、去空格、去换行、加逗号等一系列处理，写入txt文件，得到如下CSV格式的数据：



遍历网页数据库的所有页数，对每一页进行相同处理，最终获得完整的地震数据信息txt文件。接下来，引入pygame库，载入txt文件，通过readlines函数将文件分行形成列表：

line = f.readlines()

利用pygame库中的pygame.image.load函数加载地图：

earthmap = pygame.image.load("earthmap.jpg")

遍历line列表，利用pygame库中的pygame.draw.circle函数绘制表示地震的红色透明圆圈，坐标、直径直接运用公式得到：

pygame.draw.circle(screen1, (255, 0, 0, 128), (  
int((eval(line[i].split(",")[5]) + 180) \* 10 / 3 + 700) % 1200,  
int(-3 \* eval(line[i].split(",")[4]) + 270)), int(eval(line[i].split(",")[1])\*\*3//10))

当遍历列表时，如果出现前一个地震数据的日期与后一个不同，则跳出循环，在屏幕上打印出日期：

f1.render\_to(screen, (0, 545), current[2]+" Magnitude >= 3", fgcolor=(255, 0, 0), size=50)

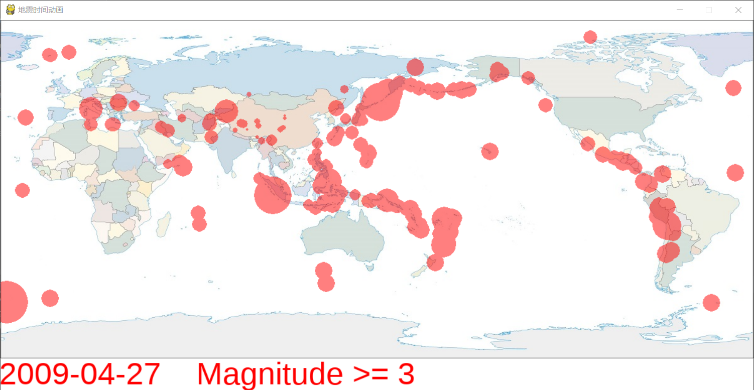
动画中屏幕刷新频率定为fps=200，因为日期跳动很快，所以红圈存在时间很短，因此引入延迟处理，当天除了显示当天的地震外，还另外显示该天地震前的300个地震，使用循环遍历画红圈：

for i in range(flag+1, flag+301)

利用blit函数不断更新打印地图来实现红圈的动态刷新：

screen.blit(earthmap, (0, 0))

1. 项目结果

 如图所示，该动画能够比较直观的表示地震的时间，震级，位置等信息，且运行未出现错误。但可以发现，除中国以外，基本没怎么出现小型地震，主要原因应该是本数据来源是中国的网站，因此对于国外的小型地震，可能未纳入统计范围。不过只要将找到的数据同样转化为上述CSV格式的txt文件，并载入程序就可以实现相同的功效。