

期末大作业：分析新冠疫情数据。

找一个有全球新冠病毒数据的网站，爬取其中的数据（禁止使用数据接口直接获取数据）。

要求爬取从 2020 年 12 月 1 日开始的连续 15 天的数据，国家数不少于 100 个。

1. 标明你的数据来源：包括网址和首页截图
2. 数据分析和展示应包括：
 - a) 15 天中，全球新冠疫情的总体变化趋势；
 - b) 累计确诊数排名前 20 的国家名称及其数量；
 - c) 15 天中，每日新增确诊数累计排名前 10 个国家的每日新增确诊数据的曲线图；
 - d) 累计确诊人数占国家总人口比例最高的 10 个国家；
 - e) 死亡率（累计死亡人数/累计确诊人数）最低的 10 个国家；
 - f) 用饼图展示各个国家的累计确诊人数的比例（你爬取的所有国家，数据较小的国家可以合并处理）；
 - g) 展示全球各个国家累计确诊人数的箱型图，要有平均值
 - h) 其它你希望分析和展示的数据。

以上图形应包括完整的坐标、刻度、标签、图例等，如有必要请配上说明文字，对图中的内容进行解释。

3. 根据以上数据，列出全世界应对新冠疫情最好的 10 个国家，并说明你的理由。
4. 针对全球累计确诊数，利用前 10 天采集到的数据做后 5 天的预测，并与实际数据进行对比。说明你预测的方法，并分析与实际数据的差距和原因。
5. 请贴上爬虫程序的核心代码、数据处理及数据展示的核心代码（应有足够的注释）；
6. 以 pdf 格式提交到爱课堂平台上，文件名为学号，总页数不超过 30 页。
7. 截止时间 12 月 27 日 24 点。