

Report

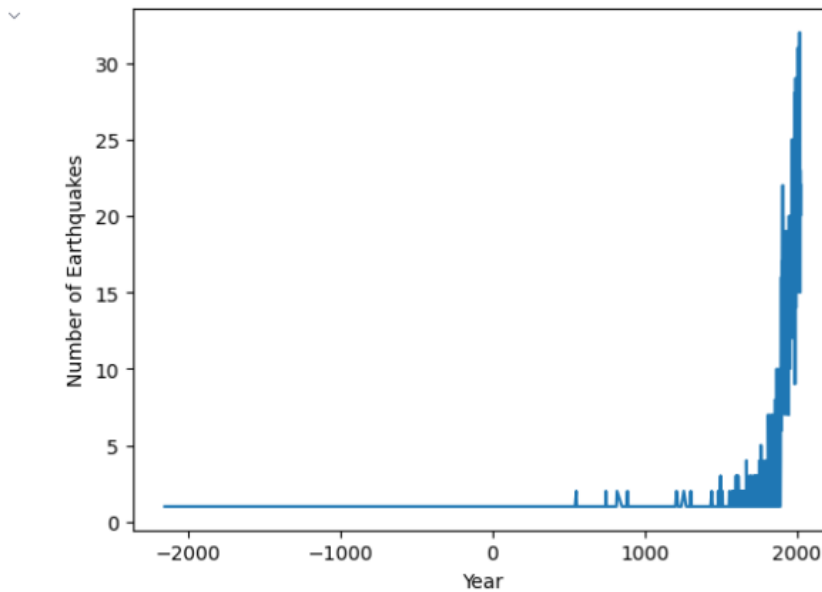
1

1.1 先将文件从网站里下载下来导入，取出公元前 2150 年以来的数据，在计算每个国家死亡的总人数，用 `sum()`，根据死亡人数降序排列，用 `sort_value(ascending=False).head(10)` 函数，

1.2

过滤震级大于 6.0 的地震数据，再提取每年地震的数量，通过 `count` 计算，然后直接绘制时间序列图，从时间序列图中可以看出 500 年以来，地震总数急剧升高，最高的年地震数达到了 30 次

name: mag, length: 556, dtype: int64



1.3

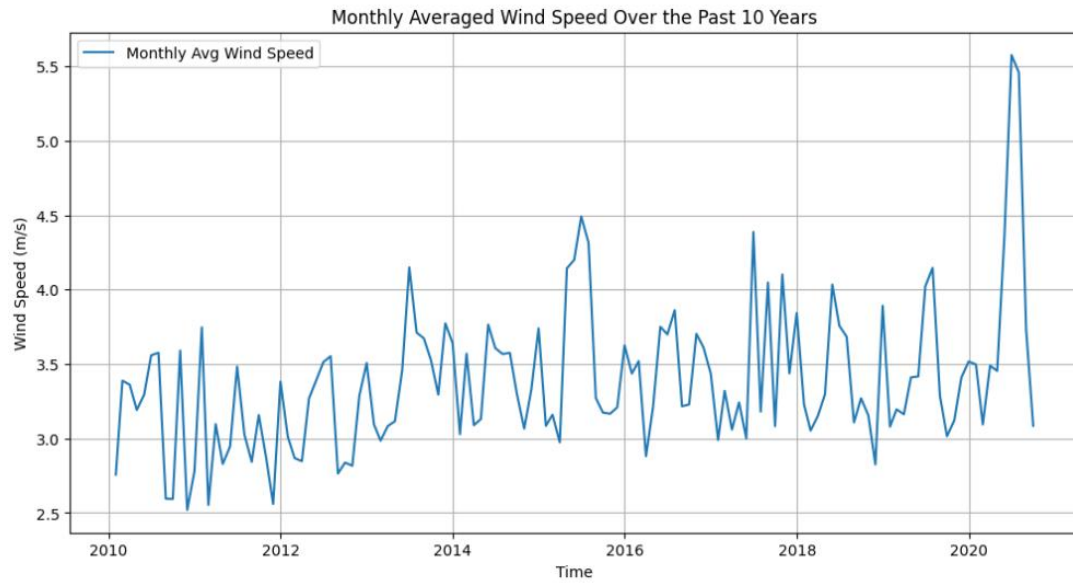
写出函数，以国家为索引，提取特定国家的地震数据，然后直接得出特定国家的地震总数，再通过 `max()` 求出该国家的最大震级，在输出某个国家的最大震级对应时间和震级。通过 `unique()` 获取唯一国家名，通过 `for` 循环以及构建的函数，得出每个国家的最大震级对应的日期和震级。

```
4872 1993.0 5.5
国家: BURUNDI, 总地震数: 1, 最大地震日期:      Year  Mag
5274 2004.0 4.7
国家: CZECH REPUBLIC, 总地震数: 1, 最大地震日期:      Year  Mag
5599 2008.0 4.1
国家: MADAGASCAR, 总地震数: 1, 最大地震日期:      Year  Mag
6054 2017.0 5.5
国家: ZAMBIA, 总地震数: 1, 最大地震日期:      Year  Mag
6061 2017.0 5.9
国家: COMOROS, 总地震数: 1, 最大地震日期:      Year  Mag
6146 2018.0 5.9
国家: nan, 总地震数: 0, 最大地震日期: Empty DataFrame
Columns: [Year, Mag]
Index: []
```

2、

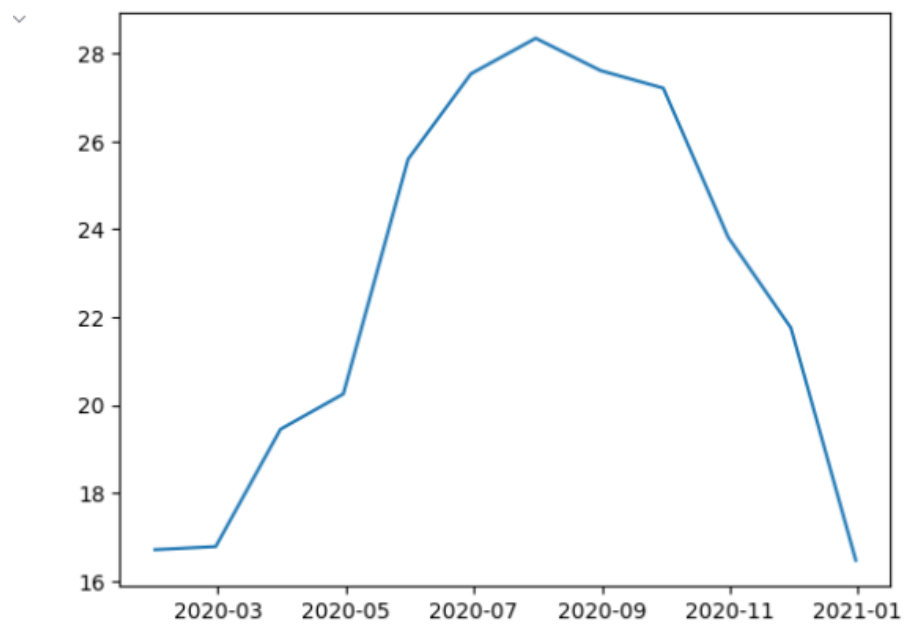
读取文件后，提取时间和风速这两列，可以看出风速这一列不常规，根据说明文件，取出字符串 8-12 个，也就是风速，再通过 `monthly_avg_wind_speed`

`=wind_data['WND'].resample('M').mean()` 计算月平均风速，除以放大因子 10，然后做出图像



3、

数据导入以后，筛选出不存在的值 NAN 值，或者是 (-999) 的值，绘制出了月平均温度



通过对数据进行统计检查，发现如不筛选出缺失值，则会导致结果偏高或偏低，甚至温度平均下来的结果为负值。