

## #1

### #1.1

首先按 Country 分类，再计算 Total Deaths 的总数，之后按降序排序，取前十个即为拥有死亡人数最多的 ten countries。

### #1.2

首先将 Mag>6.0 的筛选出来，按照 Year 分类后再计算地震数量并画图。

趋势：根据画出的图，我发现年份越大，地震数量越多，我认为是因为随着时间的推进，科技发展越来越好，对地震的观测技术更好了，所以地震的数量更多了。

### #1.3（参考了 chaptgpt 的解题思路）

定义函数 CountEq\_LargestEq(country)：筛选出国家的地震数据后计算该国家的总地震数量（total\_eqs），然后找出该国家 Mag 的最大值（Mag\_max），并获取与最大震级和该国家相对应的日期（data\_largest），最后返回国家名称、总地震数量和最大地震日期。循环遍历每个国家并调用 CountEq\_LargestEq(country)函数，并将每个国家的统计结果添加到 results 列表中。创建统计结果的 DataFrame，包含三列：'Country'（国家名称）、'Total Earthquakes'（总地震数量）和'Date of Largest Earthquake'（最大地震日期），按照总地震数量降序排序，最后输出结果。

## #2

读取 2281305.csv 文件，对 WND 列进行字符串分割，提取其中的第四个元素，并存储为 wind\_speed 变量中。将 wind\_speed 中的数据转换为数值型，并存储在 WS 列中。将 DATE 列中的数据转换为日期时间类型，提取出年份存储在 YEAR 列中，提取出月份存储在 MONTH 列中。根据条件筛选，将 'WS' 列中数值不等于 9999 的行提取出来，并将结果存储在 data\_new 列中。按照年份和月份进行分组，再计算每个组中 WS 列的均值，并绘图。

趋势：总体风速随时间增大而增大。

## #3

### #3.1

读取文件，然后再遍历每一行并删除空值

### #3.2

根据 X 轴为 DATE，y 列为变量 DP01 画图。

### #3.3

使用 describe()计算 CDS D 的描述统计信息，并将结果存储在 variable1\_stats 中。之后分别计算 CDS D 的均值、标准差、最小值、最大值和取值范围，最后输出这些结果。