

Report_PS2

Problem1

1.1 采用 read.delim 函数读取 txt 文件中的数据。

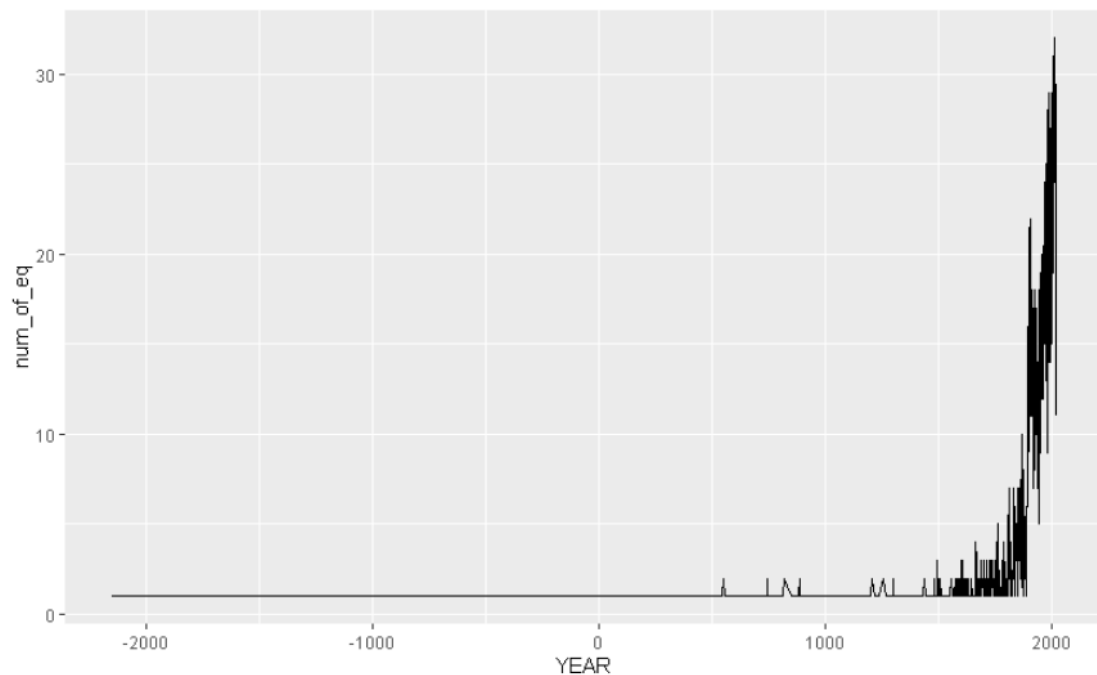
1.2 因地震死亡总人数列表：

```
# A tibble: 155 x 2
  COUNTRY      TOT_DEATHS
  <chr>      <int>
1 AFGHANISTAN 13069
2 ALBANIA      3092
3 ALGERIA     39308
4 ANTARCTICA    0
5 ANTIGUA AND BARBUDA 0
6 ARGENTINA   22520
7 ARMENIA    191890
8 ATLANTIC OCEAN 0
9 AUSTRALIA    12
10 AUSTRIA     5040
# ... with 145 more rows
```

死亡人数前十名的国家：

```
# A tibble: 10 x 2
  COUNTRY      TOT_DEATHS
  <chr>      <int>
1 CHINA    2074881
2 TURKEY   1074651
3 IRAN     1036676
4 SYRIA     439224
5 ITALY     434865
6 HAITI     321224
7 AZERBAIJAN 317219
8 JAPAN     278137
9 ARMENIA    191890
10 PAKISTAN  148692
```

1.3 地震总数的时间序列：



从图中可以看到大型地震发生的频数从 1500 年之后明显上升。我认为几千年地壳状态的变化不足以解释地震次数迅速增加的情况，产生这样趋势的原因是 1500 年之前科技水平有限，导致很多地震并未被记录，是历史记载的原因，而非地球本身的原因。

1.4 函数应用举例：以中国和日本为例：

```
> CountEq_LargestEq("CHINA")
# A tibble: 1 x 5
  Country eq_of_country year month   day
  <chr>      <int> <int> <int> <int>
1 CHINA      606  1668     7    25
> CountEq_LargestEq("JAPAN")
# A tibble: 1 x 5
  Country eq_of_country year month   day
  <chr>      <int> <int> <int> <int>
1 JAPAN      406  2011     3    11
```

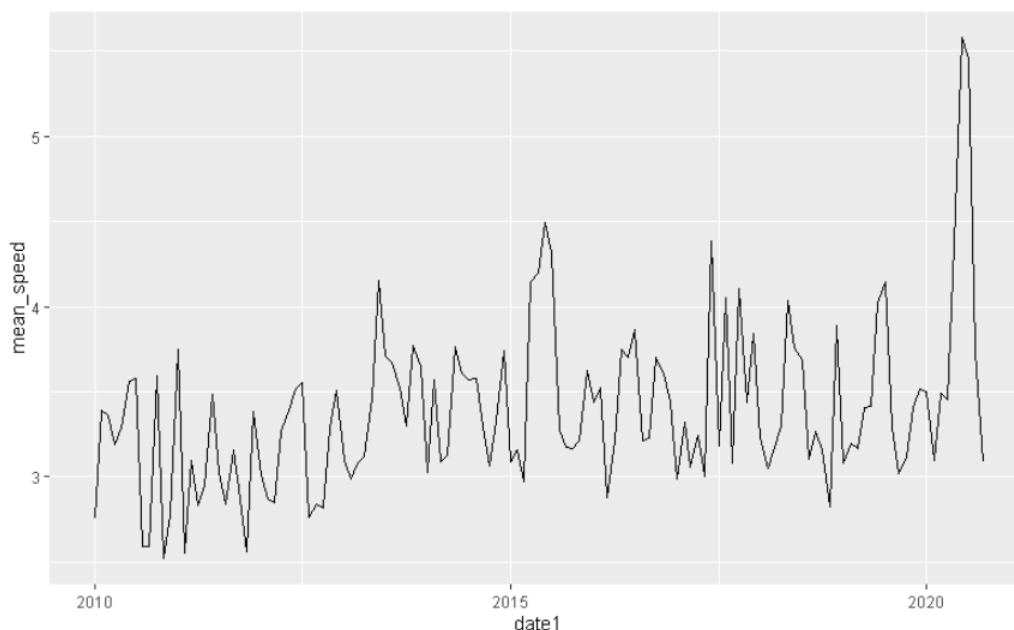
按降序排列每个国家：

```
Country eq_of_country year month   day
  <chr>      <int> <int> <int> <int>
1 CHINA      606  1668     7    25
2 JAPAN      406  2011     3    11
3 INDONESIA  394  2004    12    26
4 IRAN       380   856    12    22
5 TURKEY     329  1912     8     9
6 TURKEY     329  1916     1    24
7 ITALY      326  1915     1    13
8 USA        268  1964     3    28
9 GREECE     264   365     7    21
10 GREECE     264  1303     8     8
# ... with 152 more rows
```

Problem2

对风速数据进行筛选：由于研究的是风速大小变化，因此忽略掉前面对风向的探测数据。对于风速的探测数据，首先去除风速大小为 9999 的缺失值，然后再筛选出质量码为 1 的数据即可。

深圳十年来月平均风速变化图：



从图中可以看到深圳 10 年来月平均风速有上升的趋势，但不明显。

Problem3

选取变量为水面叶绿素浓度。将水中和水底的数据去除，再数据类型转换过程中已经将无效值自动以 NA 代替，因此不需要做额外的数据整理。得到水面叶绿素浓度的时间序列：

