PS2\_1

1. 设置工作环境并且载入程辑包，读入数据文件并将其转化为tbl 型，结果发现程序报错如下：

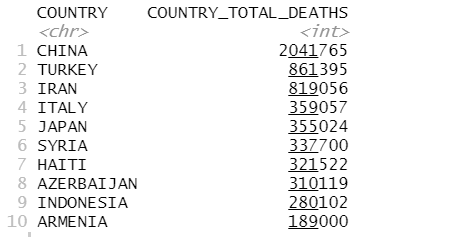


“关于报错和quote的用法，求助了CSDN论坛，得到了解决办法”

检查发现转化为tbl型后，NA值都以“”呈现在数据集中，所以在read.table函数中添加条件quote=“”，禁用引号。报错解除。

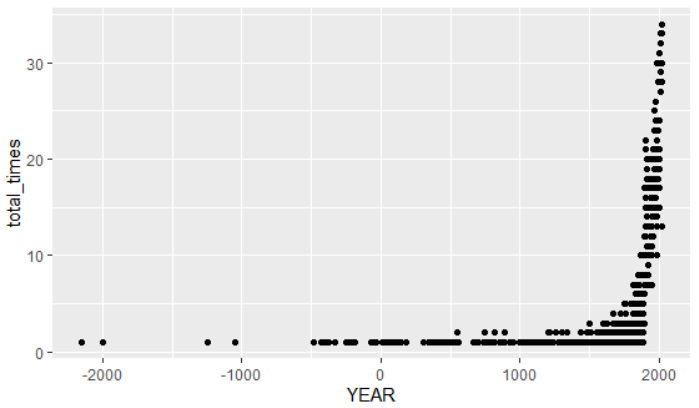
1. 首先剔除死亡人数未知的地震，不计算在死亡人数之和里，以国家聚类，求地震死亡人数之和。再按死亡人数进行反向排序，输出前10行。

输出结果：



1. 首先筛选出地震等级大于6.0的地震，再生成一列用来计数，如果地震级数大于6则计数为一次，否则为0次。选取年份，地震级数，级数三列，以年份聚类，统计每年的地震次数，以年份为横坐标，地震数为纵坐标做图。

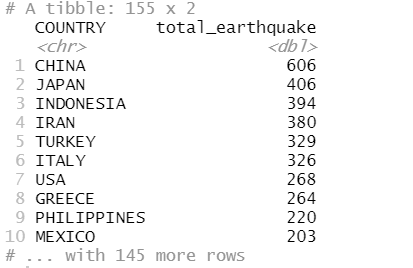
输出结果：



1. 首先将NA值赋值为0，便于之后的加和。

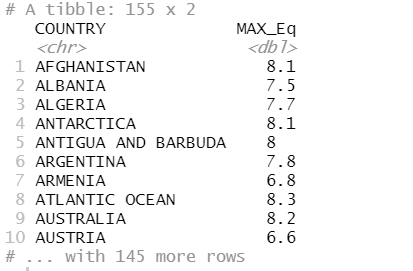
为便于后续函数输出结果，首先将每个国家的地震总数统计出来并按照地震总数大小排序，储存在变量total\_times\_country中。

结果为：

：

再将每个国家的最大地震等级统计出来，储存在变量maxEq\_country中。

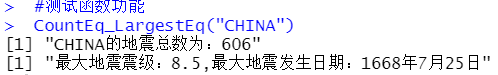
结果为：



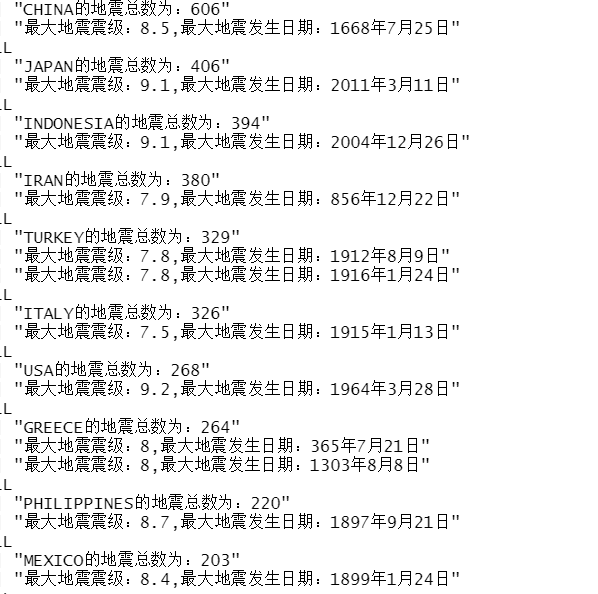
将函数中需要用到的列存在 all\_data 变量中，便于后续调取。

接着编写函数，首先循环是为了遍历所有国家，接着进行判断，如果国家名字匹配，则输出之前得到的对应的地震总数。接着判断，如果国家匹配，再通过循环遍历all\_data里的各场地震，当国家名能对应且地震级数等于最大地震级数时，输出对应的日期。

以“CHINA”为例，检验函数功能，结果如下：



题目还要求我们按照地震总数降序输出各国家的地震信息，在这里我用循环语句调用total\_times\_country里的国家名，并将其逐个代入函数。由于国家名已经按照地震总数排序，所以输出的结果也是按照地震总数排序的。由于国家太多，只展示前10个结果：



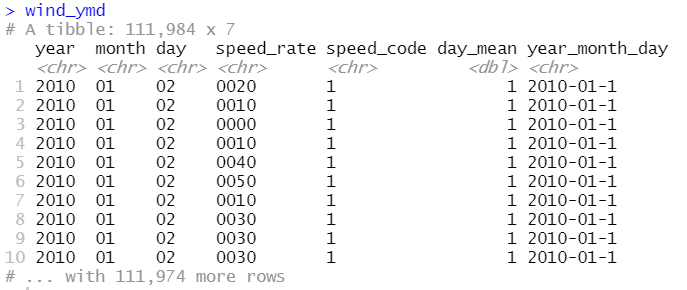
“与王超同学交流，他给了我首先按照地震总数对国家进行排序的思路，使得最后一问可以按照地震总数进行各个国家地震信息的输出”

PS2\_2

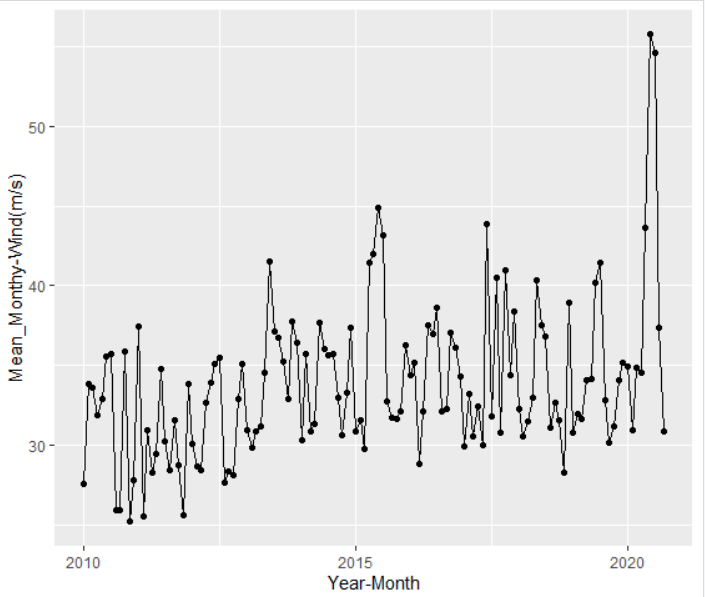
首先设置工作环境，载入程辑包，然后读取数据文件，转换为tbl型。

根据数据说明文件，对DATE和WIND这两列进行拆分处理，得到所需的年 月 日 风速 风速标识等信息，数据单独成列。

然后挑选 年 月 风速 风速标识等信息储存在一个变量中，方便后续使用，为求月平均风速，需要模糊天数信息，故生成一列新的天数信息，将其赋值为1。接着，再将年月日拼接，方便转化为日期格式。此时变量所含的信息：



最后，筛选风速标识为“1”或“0”或“4”或“5”的数据，将其视为合适的好数据，然后以年-月-日聚类，统计月平均风速（需注意要将风速先转换为数值型），最终以日期形式的年月日为横坐标，和以月平均风速为纵坐标画图。结果如下：



“起初在将拼接的年-月转化为日期格式时，出现了报错，日期格式转化需要有年-月-日三项，因为求月平均风速，所以将日期模糊处理，统一定为1号，此思路是与王超同学讨论的结果。”

PS2\_3

首先设置工作环境，载入程辑包，然后读取数据文件，转换为tbl型。

根据数据说明文件， 以DATE为横坐标，PRCP（降雨量）为纵坐标画图。

