

PS2_1

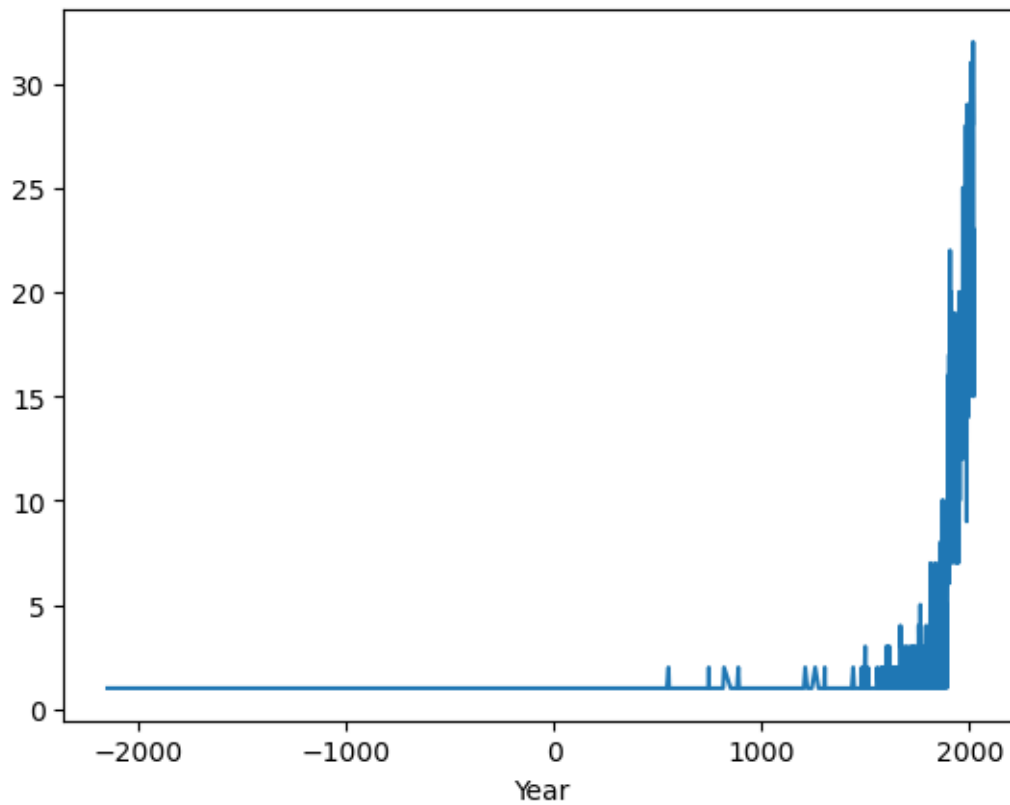
1.1

The top ten countries along with the total number of deaths are:

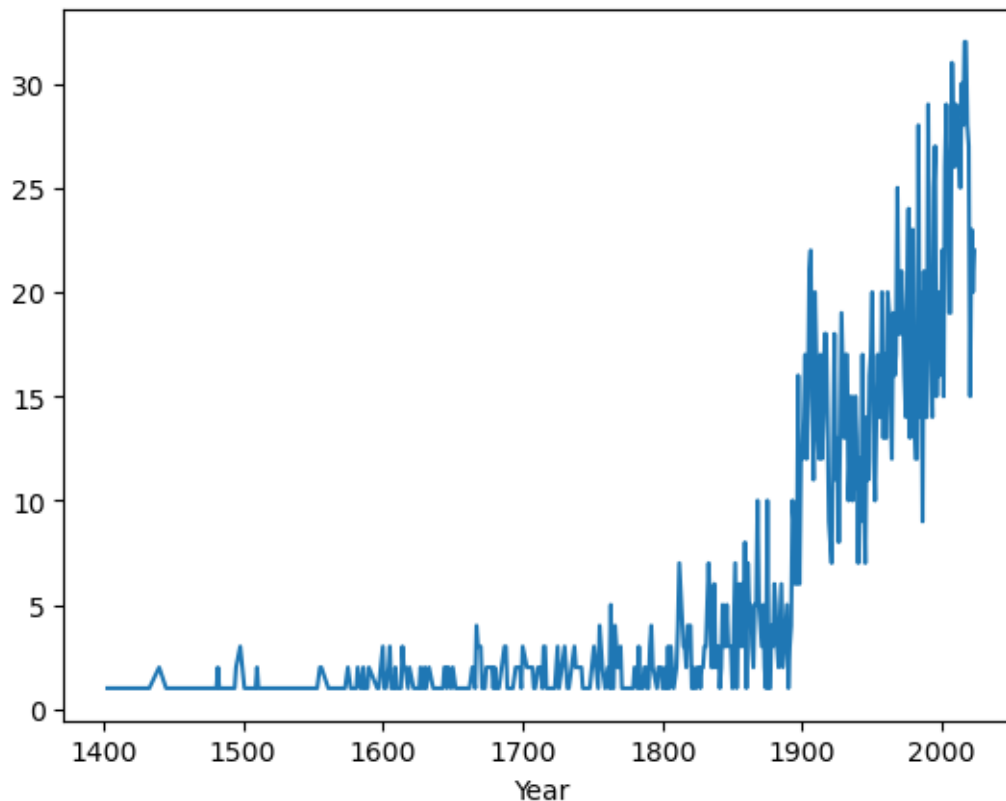
CHINA	2041929.0
TURKEY	995648.0
IRAN	758650.0
SYRIA	437700.0
ITALY	422679.0
JAPAN	356083.0
HAITI	323776.0
AZERBAIJAN	310119.0
INDONESIA	282819.0
ARMENIA	189000.0

1.2

The time series is as follows:



Plot the time series from 15C to now to make the line clearer:



We can find a trend that earthquakes with magnitude larger than 6.0 have become much more than before since 1900. The possible reasons are the development of the measurements for earthquakes and more sites monitoring earthquakes, rather than really more large earthquakes.

1.3

I only wrote the function successfully (I referred to some blogs when writing the function), but couldn't apply it to every country in the file...

```
# 1.3
def CountEq_LargestEq(country):
    Eqs_TotalNumber = Sig_Eqs.groupby(Sig_Eqs['Country']).count()['Id']
    Eqs_TotalNumber[country]

    reordered = Sig_Eqs.sort_values(['Country', 'Mag'], ascending=[1, 0])
    grouped = reordered.groupby(['Country'])
    largest = grouped.apply(lambda x: x.head(1))
    sort_country = largest.loc[largest['Country'] == country]
    sort_country['Date'] = sort_country['Year'].astype(str) + '/' + sort_country['Mo'].astype(int).astype(str) + '/' + sort_country['Dy']

    return Eqs_TotalNumber[country], sort_country['Date']

CountEq_LargestEq('CHINA')

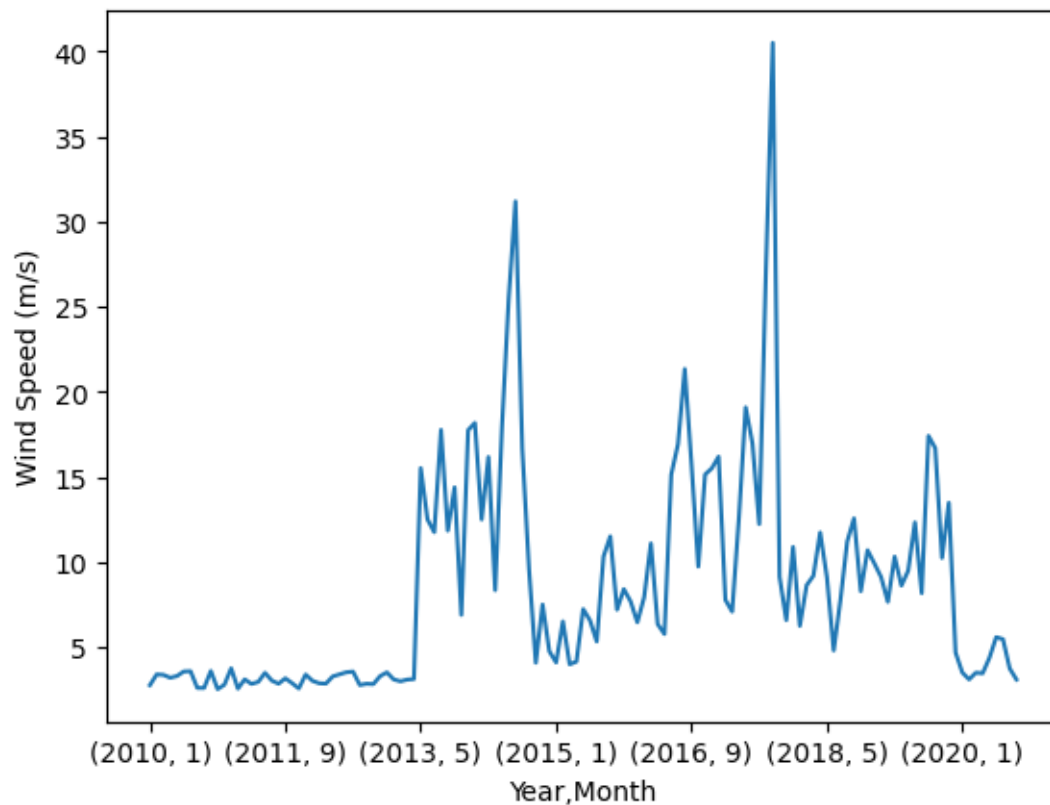
# for EachCountry in largest['Country']:
#     try:
#         CountEq_LargestEq(EachCountry)
#     except:
#         pass
```

PS2_2

How to filter the data:

The last column (WND) contains 'N' or 'C' (Normal or Calm), and at the same time the last string is '1' (Passed all quality control checks). Then choose the wind speed not equal to 9999.

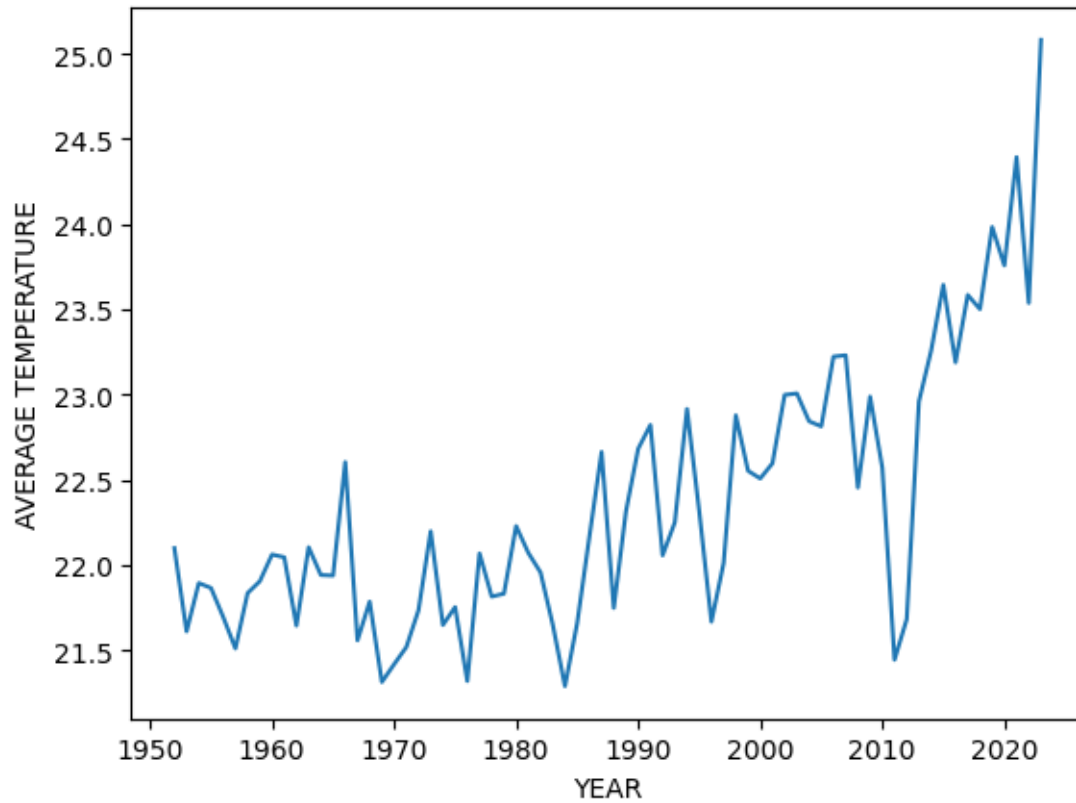
Plot monthly averaged wind speed as follows:



The result seems quite strange for some large anomalies and can't see the trends from the plot...

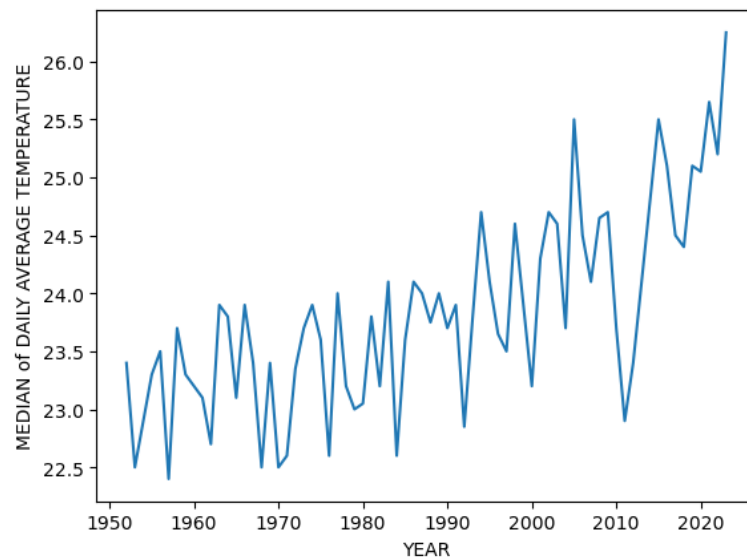
PS2_3

1. I downloaded a dataset of historical climatology in Guangzhou from NCEI. For average temperature, I eliminated NaN value.
2. Plot annual average temperature from 1952 to now (The reason why choosing 1952 as beginning is that there are lots of missing data before 1952).

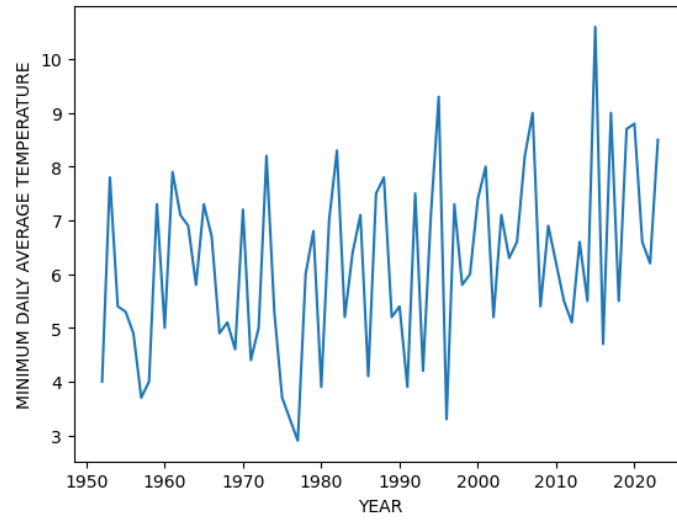


(a) Annual average temperature

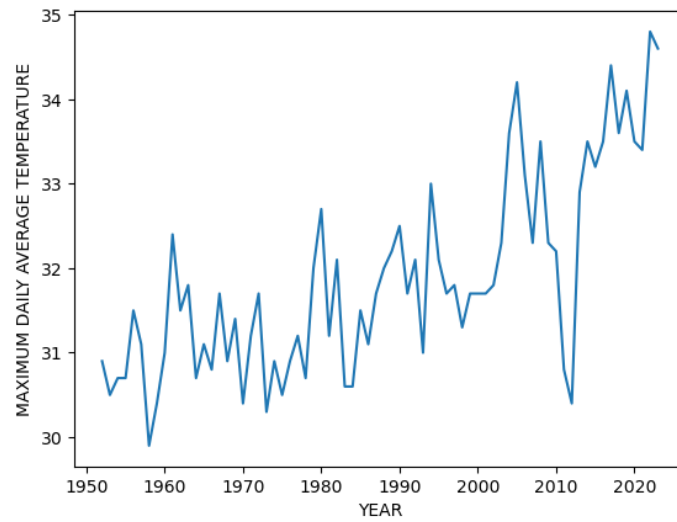
3. Plot more statistical characteristic curves as follows:



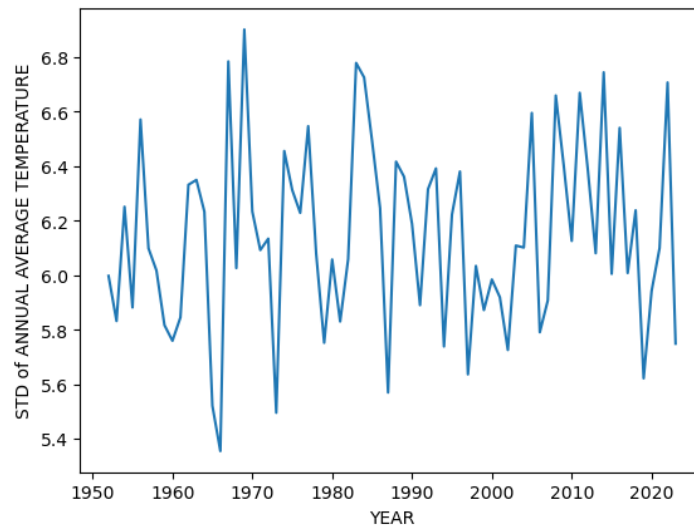
(b) Median of daily average temperature



(c) Minimum daily average temperature



(d) Maximum daily average temperature



(e) Standard deviation of annual average temperature

Findings:

- (1) The annual average temperature in Guangzhou has been rising since 1950s, and the warming has been accelerating since 2010.
- (2) Median of daily average temperature shows similar trend.
- (3) Minimum average temperature in each year shows a fluctuant trend, while maximum average temperature in each year shows an apparent rising trend.
- (4) Standard deviation of average temperature fluctuates from 5.3 to 7 and doesn't show any trend, which means the variation of average temperature changes little.