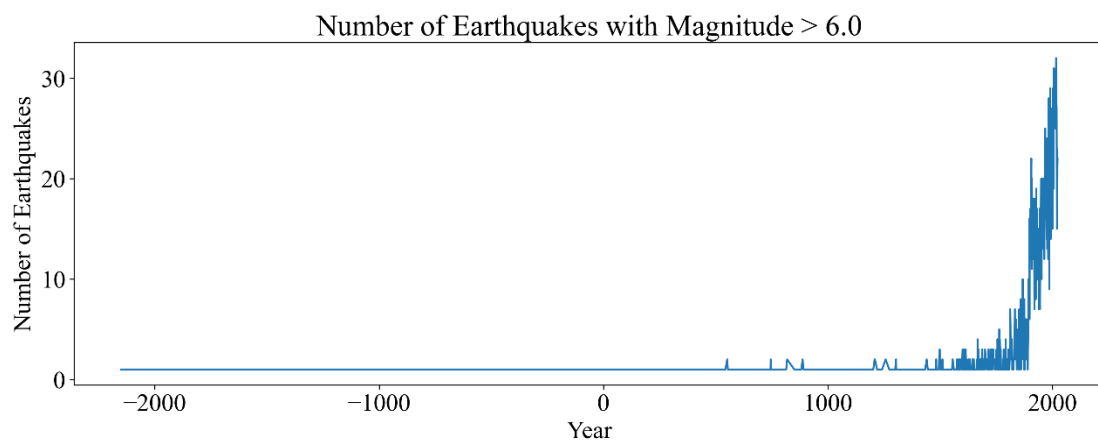


# 1. Significant earthquakes since 2150 B.C.

**1.1 [5 points]** Compute the total number of deaths caused by earthquakes since 2150 B.C. in each country, and then print the top ten countries along with the total number of deaths.  
排名前十的国家以及死亡总人数：

Country	
CHINA	2075045.0
TURKEY	1188881.0
IRAN	1011449.0
ITALY	498478.0
SYRIA	439224.0
HAITI	323478.0
AZERBAIJAN	317219.0
JAPAN	279085.0
ARMENIA	191890.0
PAKISTAN	145083.0

**1.2 [10 points]** Compute the total number of earthquakes with magnitude larger than 6.0 (use column Mag as the magnitude) worldwide each year, and then plot the time series. Do you observe any trend? Explain why or why not?



从图中可以观察到以下特点：

从公元前 2000 年到公元 0 年，地震的记录非常稀疏。从公元 1000 年左右开始，尤其是近 200 年，地震的记录数量呈现出急剧上升的趋势。

**古代记录的稀疏性：**从公元前 2000 年到公元 0 年，地震的记录非常稀疏。这很可能是因为古代没有先进的仪器和设备来记录地震，同时古代的文字记录也较少，导致很多地震事件没有被记录下来。

**近现代的显著增长：**从公元 1000 年左右开始，尤其是近 200 年，地震的记录数量呈现出急剧上升的趋势。这可能有以下几个原因：

**技术进步：**近现代，地震监测技术得到了迅猛的发展，全球各地建立了大量的地震监测站，因此能够记录到更多的地震事件。

**人口增长与城市化：**随着人口的增长和城市的扩张，一些之前未被注意的地震区域也开始被密切关注，因此地震记录也相应增加。

**数据共享：**随着全球化和信息化的发展，各国之间的地震数据共享变得更加方便，导致地震记录更加完整。

**近年的波动：**尽管近 200 年地震次数整体上升，但中间也有一些波动，这可能与地球内部的地壳活动有关。

**自然地震活动的周期性：**地球的地壳活动具有一定的周期性，某些时期地震活动可能会更加活跃，而某些时期则相对稳定。

**结论：**图反映了地震记录技术的进步、人类社会的发展和地球内部地壳活动的变化。近现代地震记录的急剧增长既与技术和社会因素有关，也可能与地球内部的自然变化有关。

**1.3 [10 points]** Write a function `CountEq_LargestEq` that returns both (1) the total number of earthquakes since 2150 B.C. in a given country AND (2) the date of the largest earthquake ever happened in this country. Apply `CountEq_LargestEq` to every country in the file, report your results in a descending order.

CHINA: 总地震次数 = 589, 最大地震发生日期 = 1668.0-7.0-25.0

JAPAN: 总地震次数 = 351, 最大地震发生日期 = 2011.0-3.0-11.0

INDONESIA: 总地震次数 = 331, 最大地震发生日期 = 2004.0-12.0-26.0

IRAN: 总地震次数 = 259, 最大地震发生日期 = 856.0-12.0-22.0

USA: 总地震次数 = 223, 最大地震发生日期 = 1964.0-3.0-28.0

TURKEY: 总地震次数 = 214, 最大地震发生日期 = 1939.0-12.0-26.0

GREECE: 总地震次数 = 155, 最大地震发生日期 = 365.0-7.0-21.0

PERU: 总地震次数 = 153, 最大地震发生日期 = 1716.0-2.0-6.0

CHILE: 总地震次数 = 146, 最大地震发生日期 = 1960.0-5.0-22.0

RUSSIA: 总地震次数 = 142, 最大地震发生日期 = 1952.0-11.0-4.0

PHILIPPINES: 总地震次数 = 136, 最大地震发生日期 = 1897.0-9.0-21.0

MEXICO: 总地震次数 = 125, 最大地震发生日期 = 1899.0-1.0-24.0

ITALY: 总地震次数 = 102, 最大地震发生日期 = 1915.0-1.0-13.0

TAIWAN: 总地震次数 = 97, 最大地震发生日期 = 1920.0-6.0-5.0

PAPUA NEW GUINEA: 总地震次数 = 93, 最大地震发生日期 = 1919.0-5.0-6.0

INDIA: 总地震次数 = 84, 最大地震发生日期 = 1950.0-8.0-15.0

COLOMBIA: 总地震次数 = 66, 最大地震发生日期 = 1826.0-6.0-18.0

NEW ZEALAND: 总地震次数 = 65, 最大地震发生日期 = 1826.0-nan-nan

SOLOMON ISLANDS: 总地震次数 = 61, 最大地震发生日期 = 1977.0-4.0-21.0

AFGHANISTAN: 总地震次数 = 60, 最大地震发生日期 = 1909.0-7.0-7.0

ECUADOR: 总地震次数 = 57, 最大地震发生日期 = 1906.0-1.0-31.0

VANUATU: 总地震次数 = 48, 最大地震发生日期 = 1913.0-10.0-14.0

PAKISTAN: 总地震次数 = 44, 最大地震发生日期 = 1945.0-11.0-27.0

ALGERIA: 总地震次数 = 39, 最大地震发生日期 = 1980.0-10.0-10.0

ALBANIA: 总地震次数 = 35, 最大地震发生日期 = 1893.0-6.0-14.0

VENEZUELA: 总地震次数 = 30, 最大地震发生日期 = 1894.0-4.0-29.0

GUATEMALA: 总地震次数 = 29, 最大地震发生日期 = 1942.0-8.0-6.0

NICARAGUA: 总地震次数 = 27, 最大地震发生日期 = 1898.0-4.0-29.0

TAJIKISTAN: 总地震次数 = 27, 最大地震发生日期 = 1907.0-10.0-21.0

COSTA RICA: 总地震次数 = 26, 最大地震发生日期 = 1950.0-10.0-5.0

MYANMAR (BURMA): 总地震次数 = 25, 最大地震发生日期 = 1912.0-5.0-23.0

NEW CALEDONIA: 总地震次数 = 25, 最大地震发生日期 = 1875.0-3.0-28.0

USA TERRITORY: 总地震次数 = 23, 最大地震发生日期 = 1902.0-9.0-22.0

AUSTRALIA: 总地震次数 = 22, 最大地震发生日期 = 1989.0-5.0-23.0

EL SALVADOR: 总地震次数 = 21, 最大地震发生日期 = 1915.0-9.0-7.0

KERMADEC ISLANDS (NEW ZEALAND): 总地震次数 = 21, 最大地震发生日期 = 1986.0-10.0-20.0

SOUTH KOREA: 总地震次数 = 20, 最大地震发生日期 = 1643.0-7.0-25.0

PANAMA: 总地震次数 = 20, 最大地震发生日期 = 1882.0-9.0-7.0

TONGA: 总地震次数 = 20, 最大地震发生日期 = 1919.0-4.0-30.0

NEPAL: 总地震次数 = 17, 最大地震发生日期 = 1505.0-6.0-6.0

ARGENTINA: 总地震次数 = 17, 最大地震发生日期 = 1944.0-1.0-15.0

FIJI: 总地震次数 = 17, 最大地震发生日期 = 1919.0-1.0-1.0

GEORGIA: 总地震次数 = 15, 最大地震发生日期 = 1905.0-10.0-21.0

FRANCE: 总地震次数 = 15, 最大地震发生日期 = 1817.0-3.0-11.0

CANADA: 总地震次数 = 15, 最大地震发生日期 = 1949.0-8.0-22.0

BANGLADESH: 总地震次数 = 15, 最大地震发生日期 = 1918.0-7.0-8.0

KYRGYZSTAN: 总地震次数 = 14, 最大地震发生日期 = 1946.0-11.0-2.0

AZERBAIJAN: 总地震次数 = 14, 最大地震发生日期 = 1667.0-11.0-nan

SLOVENIA: 总地震次数 = 14, 最大地震发生日期 = 1511.0-3.0-26.0

SPAIN: 总地震次数 = 13, 最大地震发生日期 = 881.0-5.0-26.0

ROMANIA: 总地震次数 = 13, 最大地震发生日期 = 1977.0-3.0-4.0

SOUTH AFRICA: 总地震次数 = 13, 最大地震发生日期 = 1942.0-11.0-10.0

BULGARIA: 总地震次数 = 12, 最大地震发生日期 = 1904.0-4.0-4.0

UKRAINE: 总地震次数 = 12, 最大地震发生日期 = 103.0-nan-nan

UZBEKISTAN: 总地震次数 = 12, 最大地震发生日期 = 1976.0-4.0-8.0

CROATIA: 总地震次数 = 12, 最大地震发生日期 = 1667.0-4.0-6.0

TURKMENISTAN: 总地震次数 = 10, 最大地震发生日期 = 1895.0-7.0-8.0

KAZAKHSTAN: 总地震次数 = 10, 最大地震发生日期 = 1889.0-7.0-11.0

SERBIA: 总地震次数 = 10, 最大地震发生日期 = 1922.0-3.0-24.0

AZORES (PORTUGAL): 总地震次数 = 10, 最大地震发生日期 = 1968.0-2.0-28.0

HAITI: 总地震次数 = 10, 最大地震发生日期 = 1842.0-5.0-7.0

ISRAEL: 总地震次数 = 9, 最大地震发生日期 = -31.0-9.0-2.0

BOSNIA-HERZEGOVINA: 总地震次数 = 9, 最大地震发生日期 = 1969.0-10.0-27.0

HONDURAS: 总地震次数 = 9, 最大地震发生日期 = 1856.0-8.0-4.0

PORTUGAL: 总地震次数 = 8, 最大地震发生日期 = -60.0-nan-nan

ARMENIA: 总地震次数 = 8, 最大地震发生日期 = 1988.0-12.0-7.0

DOMINICAN REPUBLIC: 总地震次数 = 8, 最大地震发生日期 = 1946.0-8.0-4.0

TANZANIA: 总地震次数 = 8, 最大地震发生日期 = 1910.0-12.0-13.0

SAMOA: 总地震次数 = 8, 最大地震发生日期 = 1917.0-6.0-26.0

EGYPT: 总地震次数 = 7, 最大地震发生日期 = 1995.0-11.0-22.0

MACEDONIA: 总地震次数 = 7, 最大地震发生日期 = 1979.0-5.0-24.0

UK: 总地震次数 = 7, 最大地震发生日期 = 1580.0-4.0-6.0

ETHIOPIA: 总地震次数 = 7, 最大地震发生日期 = 1906.0-8.0-25.0

CONGO: 总地震次数 = 7, 最大地震发生日期 = 1992.0-9.0-11.0

SOUTH GEORGIA AND THE SOUTH SANDWICH ISLANDS: 总地震次数 = 7, 最大地震发生日期 = 1929.0-6.0-27.0

POLAND: 总地震次数 = 7, 最大地震发生日期 = 2004.0-9.0-21.0

NORTH KOREA: 总地震次数 = 6, 最大地震发生日期 = 1518.0-7.0-2.0

ICELAND: 总地震次数 = 6, 最大地震发生日期 = 1912.0-5.0-6.0

CUBA: 总地震次数 = 6, 最大地震发生日期 = 2020.0-1.0-28.0

TRINIDAD AND TOBAGO: 总地震次数 = 6, 最大地震发生日期 = 1888.0-1.0-10.0

BOLIVIA: 总地震次数 = 6, 最大地震发生日期 = 1994.0-6.0-9.0

MONGOLIA: 总地震次数 = 6, 最大地震发生日期 = 1905.0-7.0-9.0

BRAZIL: 总地震次数 = 6, 最大地震发生日期 = 1963.0-11.0-9.0

JORDAN: 总地震次数 = 5, 最大地震发生日期 = -2150.0-nan-nan

GERMANY: 总地震次数 = 5, 最大地震发生日期 = 1978.0-9.0-3.0

MONTENEGRO: 总地震次数 = 5, 最大地震发生日期 = 1979.0-4.0-15.0

GHANA: 总地震次数 = 5, 最大地震发生日期 = 1862.0-7.0-10.0

ATLANTIC OCEAN: 总地震次数 = 5, 最大地震发生日期 = 1941.0-11.0-25.0

HUNGARY: 总地震次数 = 5, 最大地震发生日期 = 1834.0-10.0-15.0

GUADELOUPE: 总地震次数 = 5, 最大地震发生日期 = 1843.0-2.0-8.0

VIETNAM: 总地震次数 = 5, 最大地震发生日期 = 1935.0-11.0-1.0

ANTARCTICA: 总地震次数 = 5, 最大地震发生日期 = 1998.0-3.0-25.0

RWANDA: 总地震次数 = 5, 最大地震发生日期 = 2015.0-8.0-7.0

BHUTAN: 总地震次数 = 5, 最大地震发生日期 = 2009.0-9.0-21.0

SWITZERLAND: 总地震次数 = 4, 最大地震发生日期 = 1601.0-9.0-18.0

MOROCCO: 总地震次数 = 4, 最大地震发生日期 = 2023.0-9.0-8.0

ERITREA: 总地震次数 = 4, 最大地震发生日期 = 1875.0-11.0-2.0

UGANDA: 总地震次数 = 4, 最大地震发生日期 = 1912.0-7.0-9.0

MALAWI: 总地震次数 = 4, 最大地震发生日期 = 1989.0-3.0-10.0

SYRIA: 总地震次数 = 3, 最大地震发生日期 = 1202.0-5.0-20.0

LEBANON: 总地震次数 = 3, 最大地震发生日期 = 551.0-7.0-9.0

CYPRUS: 总地震次数 = 3, 最大地震发生日期 = 1953.0-9.0-10.0

THAILAND: 总地震次数 = 3, 最大地震发生日期 = 2014.0-5.0-5.0

IRAQ: 总地震次数 = 3, 最大地震发生日期 = 1864.0-12.0-2.0

YEMEN: 总地震次数 = 3, 最大地震发生日期 = 1982.0-12.0-13.0

JAMAICA: 总地震次数 = 3, 最大地震发生日期 = 1899.0-6.0-14.0

MARTINIQUE: 总地震次数 = 3, 最大地震发生日期 = 1906.0-12.0-3.0

MICRONESIA, FED. STATES OF: 总地震次数 = 3, 最大地震发生日期 = 1911.0-8.0-16.0

INDIAN OCEAN: 总地震次数 = 3, 最大地震发生日期 = 1928.0-3.0-9.0

MALAYSIA: 总地震次数 = 3, 最大地震发生日期 = 1976.0-7.0-26.0

NETHERLANDS: 总地震次数 = 3, 最大地震发生日期 = 1992.0-4.0-13.0

SAUDI ARABIA: 总地震次数 = 3, 最大地震发生日期 = 2009.0-5.0-19.0

MOZAMBIQUE: 总地震次数 = 3, 最大地震发生日期 = 2006.0-2.0-22.0

TUNISIA: 总地震次数 = 2, 最大地震发生日期 = 1957.0-2.0-20.0

AUSTRIA: 总地震次数 = 2, 最大地震发生日期 = 1590.0-9.0-15.0

ANTIGUA AND BARBUDA: 总地震次数 = 2, 最大地震发生日期 = 1690.0-4.0-16.0

SOUTH SUDAN: 总地震次数 = 2, 最大地震发生日期 = 1990.0-5.0-20.0

COTE D'IVOIRE: 总地震次数 = 2, 最大地震发生日期 = 1879.0-2.0-11.0

SOLOMON SEA: 总地震次数 = 2, 最大地震发生日期 = 1895.0-3.0-6.0

KENYA: 总地震次数 = 2, 最大地震发生日期 = 1928.0-1.0-6.0

PACIFIC OCEAN: 总地震次数 = 2, 最大地震发生日期 = 1932.0-11.0-2.0

LAOS: 总地震次数 = 2, 最大地震发生日期 = 2007.0-5.0-16.0

SLOVAKIA: 总地震次数 = 1, 最大地震发生日期 = 2004.0-1.0-10.0

FRENCH GUIANA: 总地震次数 = 1, 最大地震发生日期 = 1885.0-8.0-4.0

TOGO: 总地震次数 = 1, 最大地震发生日期 = 1788.0-nan-nan

SIERRA LEONE: 总地震次数 = 1, 最大地震发生日期 = 1795.0-5.0-20.0

NORWAY: 总地震次数 = 1, 最大地震发生日期 = 1819.0-8.0-31.0

UK TERRITORY: 总地震次数 = 1, 最大地震发生日期 = 1983.0-11.0-30.0

FRENCH POLYNESIA: 总地震次数 = 1, 最大地震发生日期 = 1848.0-7.0-12.0

KIRIBATI: 总地震次数 = 1, 最大地震发生日期 = 1905.0-6.0-30.0

CAMEROON: 总地震次数 = 1, 最大地震发生日期 = 1945.0-9.0-12.0

PALAU: 总地震次数 = 1, 最大地震发生日期 = 1914.0-10.0-23.0

CENTRAL AFRICAN REPUBLIC: 总地震次数 = 1, 最大地震发生日期 = 1921.0-9.0-16.0

LIBYA: 总地震次数 = 1, 最大地震发生日期 = 1963.0-2.0-21.0

GABON: 总地震次数 = 1, 最大地震发生日期 = 1974.0-9.0-23.0

BELGIUM: 总地震次数 = 1, 最大地震发生日期 = 1983.0-11.0-8.0

GUINEA: 总地震次数 = 1, 最大地震发生日期 = 1983.0-12.0-22.0

DJIBOUTI: 总地震次数 = 1, 最大地震发生日期 = 1989.0-8.0-20.0

BERING SEA: 总地震次数 = 1, 最大地震发生日期 = 1991.0-2.0-21.0

WALLIS AND FUTUNA (FRENCH TERRITORY): 总地震次数 = 1, 最大地震发生日期 = 1993.0-3.0-12.0

SUDAN: 总地震次数 = 1, 最大地震发生日期 = 1993.0-8.0-1.0

BURUNDI: 总地震次数 = 1, 最大地震发生日期 = 2004.0-2.0-24.0

CZECH REPUBLIC: 总地震次数 = 1, 最大地震发生日期 = 2008.0-11.0-22.0

MADAGASCAR: 总地震次数 = 1, 最大地震发生日期 = 2017.0-1.0-11.0

ZAMBIA: 总地震次数 = 1, 最大地震发生日期 = 2017.0-2.0-24.0  
COMOROS: 总地震次数 = 1, 最大地震发生日期 = 2018.0-5.0-15.0  
nan: 总地震次数 = 0, 最大地震发生日期 = No Data  
IRELAND: 总地震次数 = 0, 最大地震发生日期 = No Data  
SAINT LUCIA: 总地震次数 = 0, 最大地震发生日期 = No Data  
CANARY ISLANDS: 总地震次数 = 0, 最大地震发生日期 = No Data  
GRENADA: 总地震次数 = 0, 最大地震发生日期 = No Data  
BARBADOS: 总地震次数 = 0, 最大地震发生日期 = No Data  
SAINT VINCENT AND THE GRENADINES: 总地震次数 = 0, 最大地震发生日期 =  
No Data  
BRITISH VIRGIN ISLANDS: 总地震次数 = 0, 最大地震发生日期 = No Data  
SRI LANKA: 总地震次数 = 0, 最大地震发生日期 = No Data  
URUGUAY: 总地震次数 = 0, 最大地震发生日期 = No Data  
MONTserrat: 总地震次数 = 0, 最大地震发生日期 = No Data

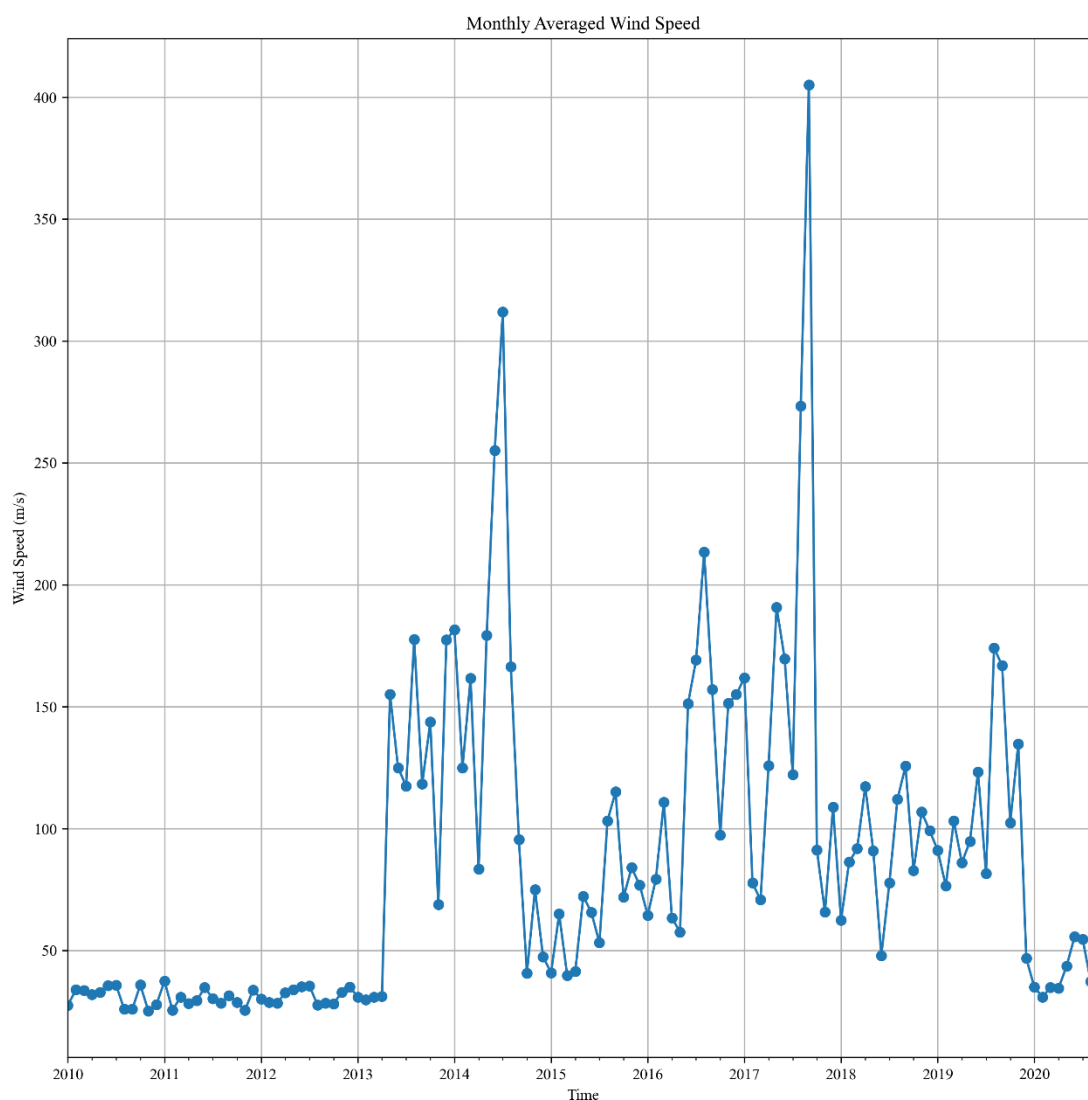
## 2. Wind speed in Shenzhen during the past 10 years

In this problem set, we will examine how wind speed changes in Shenzhen during the past 10 years, we will take a look at the hourly weather data measured at the BaoAn International Airport. The data set is from [NOAA Integrated Surface Dataset](#). Download the file [2281305.zip](#), where the number 2281305 is the site ID. Extract the zip file, you should see a file named 2281305.csv. Save the .csv file to your working directory.

Read page 8-9 (POS 65-69 and POS 70-70) of the comprehensive [user guide](#) for the detailed format of the wind data. Explain how you filter the data in your report.

**[10 points]** Plot monthly averaged wind speed as a function of the observation time. Is there a trend in monthly averaged wind speed within the past 10 years?





以观察到以下关于深圳近 10 年风速的信息：

**稳定的低风速段：**在 2010 年到 2020 年期间，有一个相对稳定的低风速段，风速大约在 50mg 以下。这可能表示深圳在这 10 年中的大部分时间都有稳定的低风速。

**显著的风速峰值：**图中有几个显著的风速峰值，尤其是在 2014 年、2017 年和 2019 年。这些峰值可能代表了某些特定的风暴或台风事件。

**风速的波动：**尽管大部分时间风速相对较低，但图中还是展示了一些风速的波动。这可能是由于季节性因素或其他大气条件的变化。

**2019 年后的下降趋势：**从 2019 年末开始，风速呈现了一个下降的趋势，直到 2020 年。这可能表示在这段时间内深圳的风速较为稳定或减弱。

### 3. Explore a data set

Browse the [CASEarth](#), [National Centers for Environmental Information \(NCEI\)](#), or [Advanced Global Atmospheric Gases Experiment \(AGAGE\)](#) website. Search and download a data set you are interested in. You are also welcome to use data from your group in this problem set. But the data set should be in `csv`, `XLS`, or `XLSX` format, and have temporal information.

**3.1 [5 points]** Load the `csv`, `XLS`, or `XLSX` file, and clean possible data points with missing values or bad quality.

```
beijingdata=pd.read_csv('data/Beijing.csv',header=0,sep
=',')

beijingdata=beijingdata.drop(labels=['SNOW','SNWD
(Inches)'],axis=1)

beijingdata.dropna(axis=0,how='any',inplace=True)

beijingdata['Date']=pd.to_datetime(beijingdata['Date'])

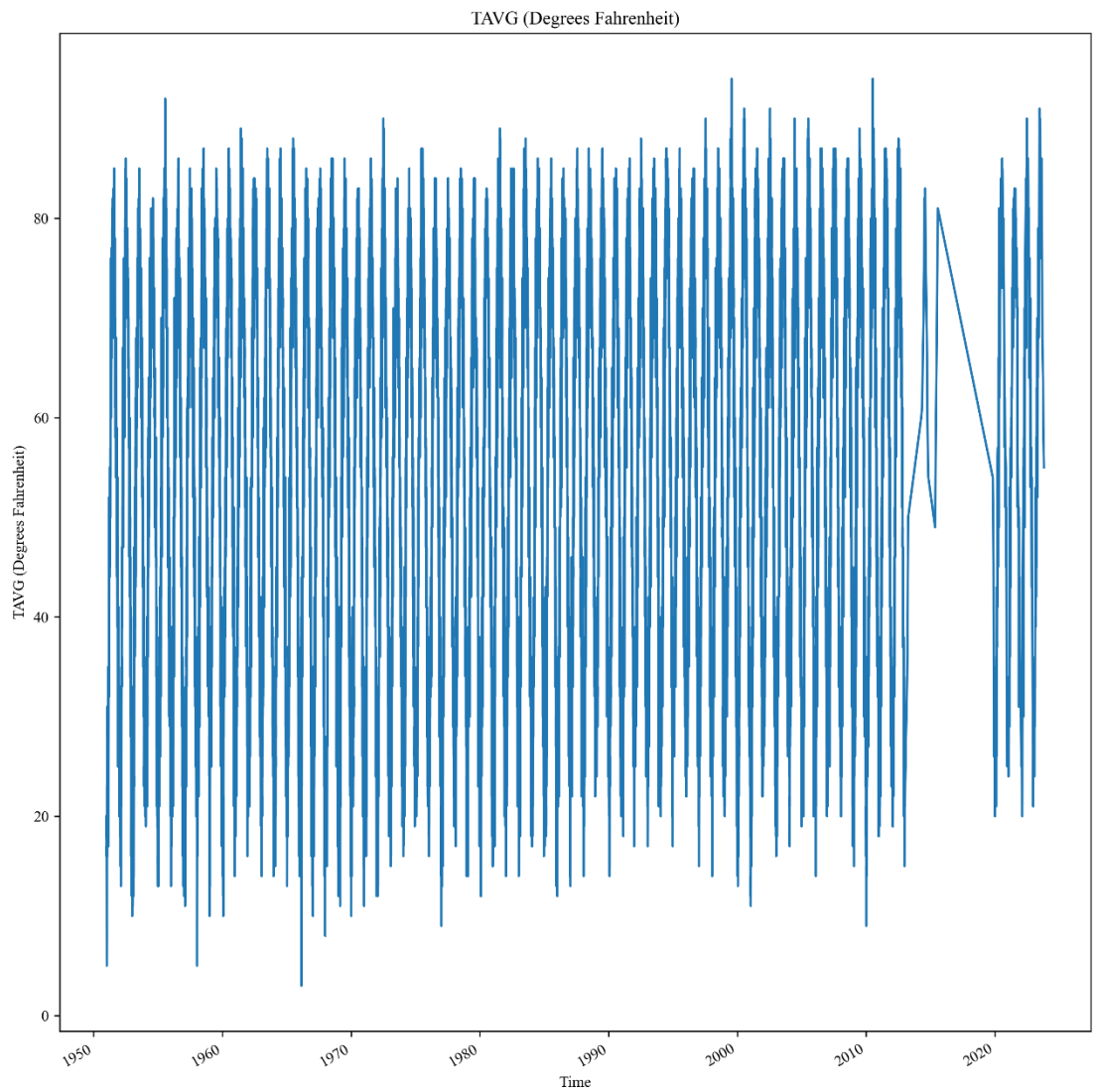
beijingdata.set_index('Date',inplace=True)
```

先删除最后两列，因为基本全是 `nan` 值，然后删除存在 `nan` 的行，最后留下：

	TAVG (Degrees Fahrenheit)	TMAX (Degrees Fahrenheit)	TMIN (Degrees Fahrenheit)	PRCP (Inches)
Date				
1951-01-02	20.0	30.0	14.0	0.02
1951-01-03	18.0	25.0	11.0	0.00
1951-01-04	19.0	23.0	14.0	0.05
1951-01-05	16.0	20.0	13.0	0.18
1951-01-06	17.0	20.0	14.0	0.00
...	...	...	...	...
2023-09-09	68.0	82.0	67.0	2.24
2023-09-28	64.0	83.0	51.0	0.00
2023-10-01	62.0	79.0	48.0	0.00
2023-10-05	59.0	75.0	44.0	0.00
2023-10-23	55.0	74.0	43.0	0.00

22988 rows × 4 columns

**3.2 [5 points]** Plot the time series of a certain variable.



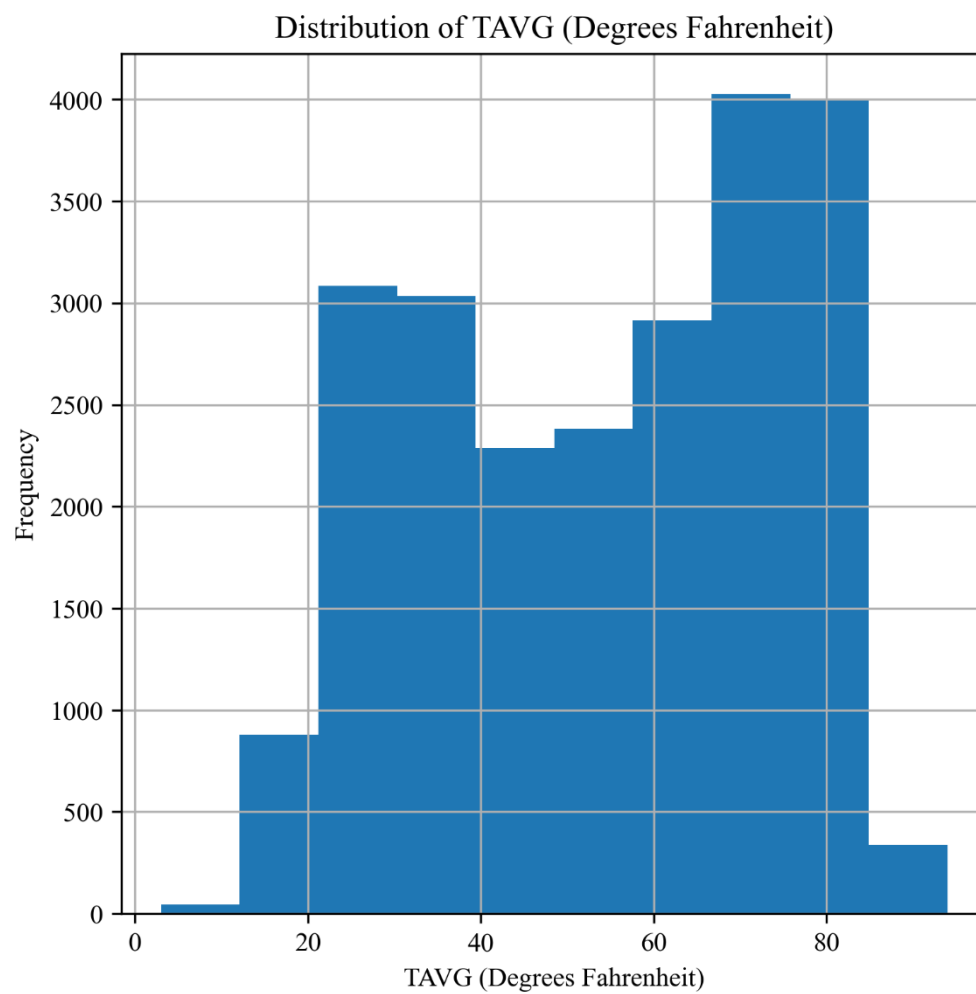
**3.3 [5 points]** Conduct at least 5 simple statistical checks with the variable, and report your findings.

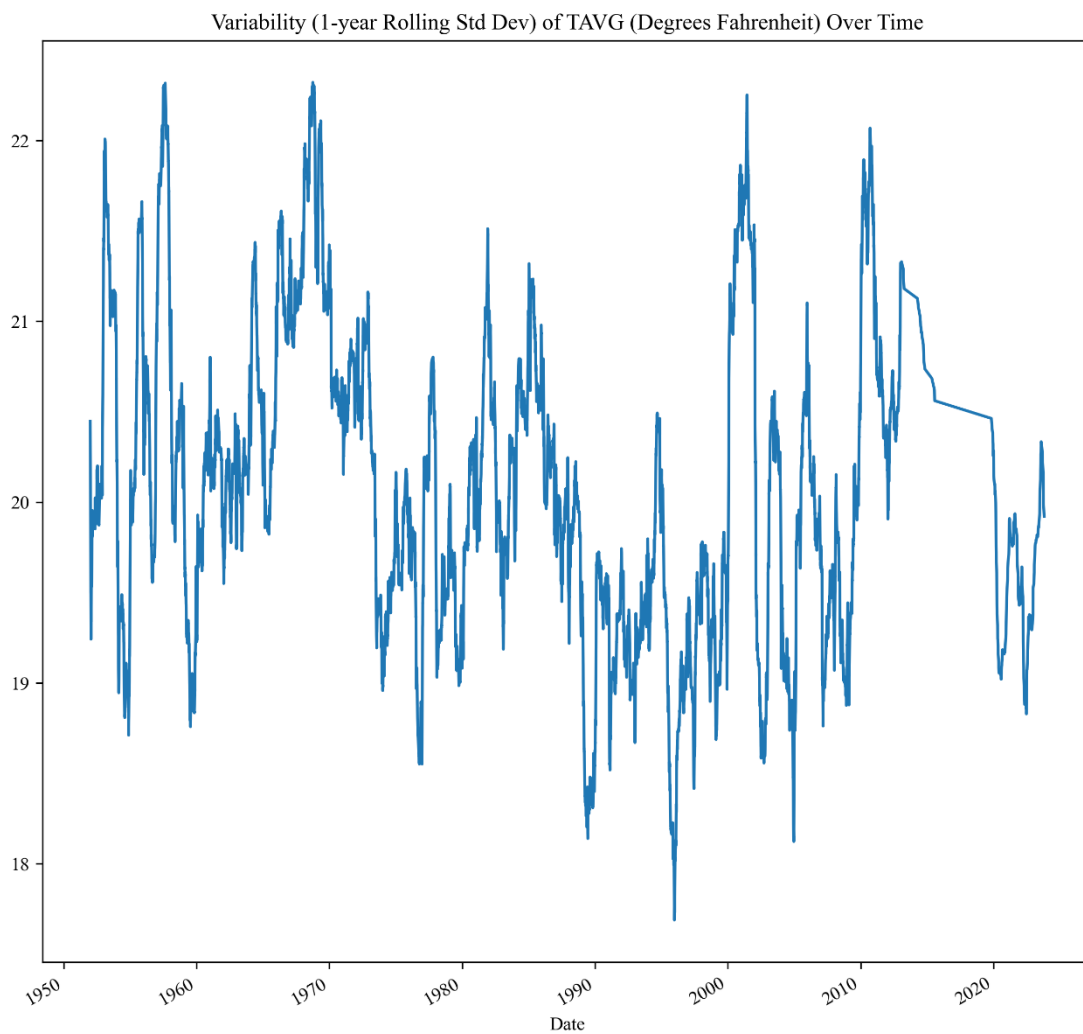
1、TAVG (Degrees Fahrenheit)	TMAX (Degrees Fahrenheit)	\
count	22988.000000	22988.000000
mean	54.126327	64.385070
std	20.166218	20.559877
min	3.000000	9.000000
25%	35.000000	46.000000
50%	57.000000	68.000000
75%	73.000000	83.000000

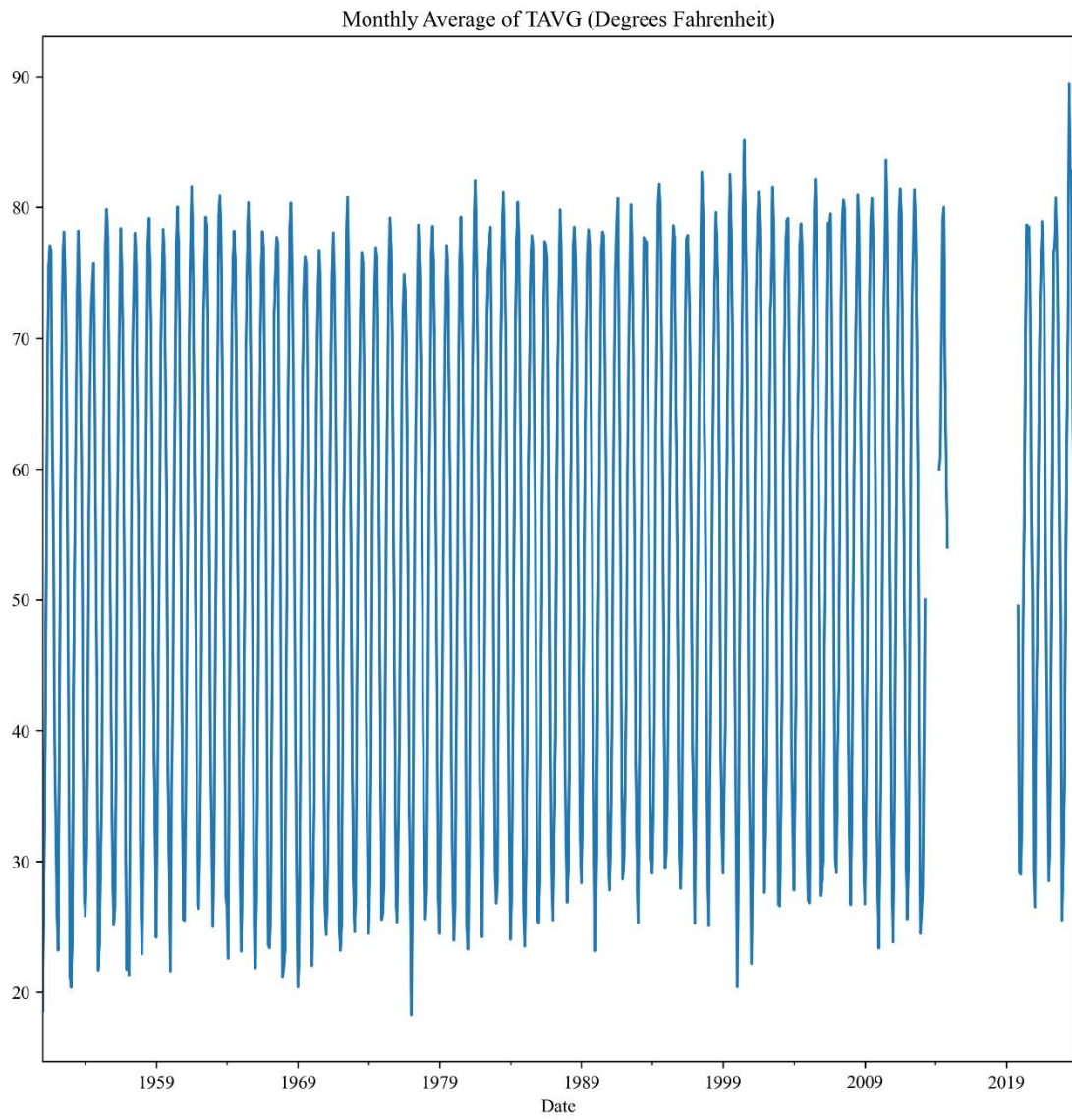
max	94.000000	107.000000
-----	-----------	------------

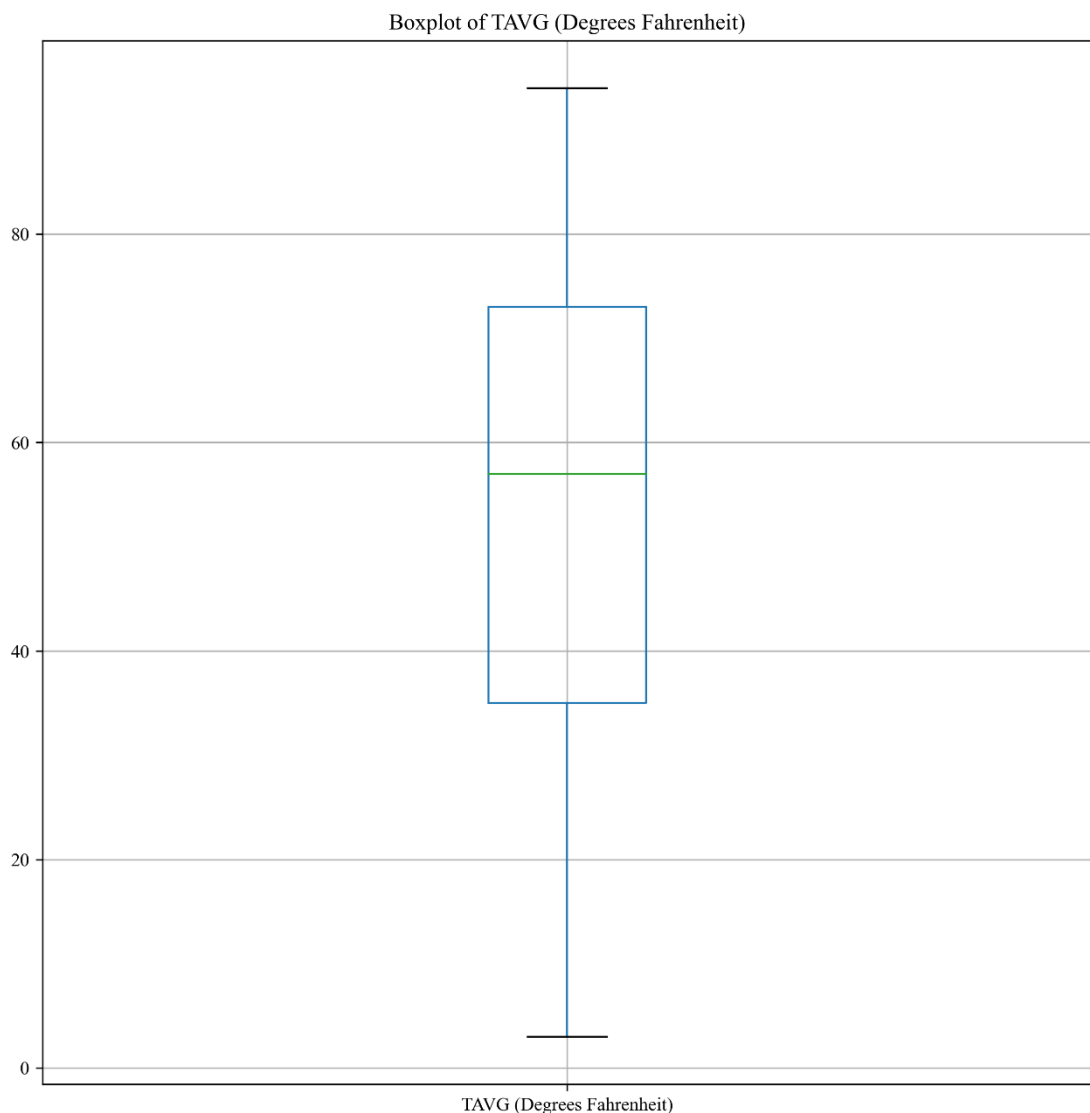
	TMIN (Degrees Fahrenheit)	PRCP (Inches)
count	22988.000000	22988.000000
mean	44.940230	0.06466
std	20.106953	0.30949
min	-17.000000	0.000000
25%	27.000000	0.000000
50%	46.000000	0.000000
75%	63.000000	0.000000
max	85.000000	9.61000

2、









**温度分布：**可以看到大多数的温度分布集中在 20 到 70 度范围内，尤其是在 40 到 60 度范围。这可能意味着这个范围内的温度是北京市最常见的平均温度。

**温度变化趋势：**可以看出从 1950 年到 2020 年间，北京市的温度波动是有周期性的。但总体上，似乎在近些年来平均温度的变动有所增大，这可能与全球气候变化有关。

**变暖趋势：**尽管温度在各个月份之间有所不同，但从每年的起始和结束点来看，近年来北京市的平均温度呈上升趋势。这与全球变暖的观点是一致的。

**温度的可变性：**通过计算 1 年滚动标准差来得出的。从图中可以看出，虽然温度的可变性在某些年份有所上升，但在最近的几年中呈下降趋势。这可能意味着近年来北京的气候变得相对稳定。

结合这些图表，可以得出结论：北京市近年来的平均温度呈现上升趋势，与

全球变暖趋势一致，同时温度的可变性在最近几年中有所下降，表明气候变得相对稳定。