

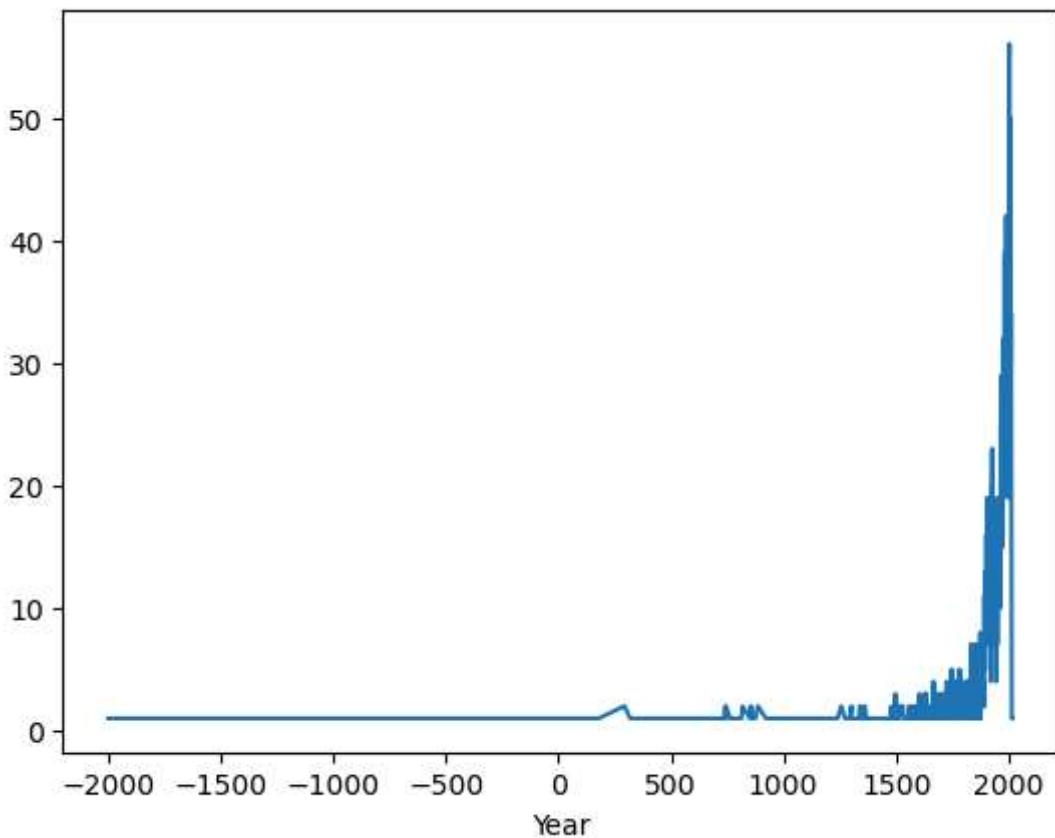
```
In [13]: import pandas as pd
import numpy as np
from matplotlib import pyplot as plt
#使用sep='\t', 可以指定制表符为分隔符
Sig_Eqs = pd.read_csv('earthquakes-2024-10-29_10-37-56_+0800.tsv', sep='\t')
```

```
In [25]: #1.1
#过滤出公元前2150年以来的数据
year_data = Sig_Eqs[Sig_Eqs['Year'] >= -2150]
#计算不同国家的死亡总数并分组
Sig_Eqs1 = year_data.groupby(['Country']).sum()['Deaths']
#将分组以降序排序并取前20个
Sig_Eqs1.sort_values(ascending = False).head(20)
```

```
Out[25]: Country
CHINA          2075947.0
TURKEY         1188881.0
IRAN           1011453.0
ITALY           498418.0
SYRIA           439224.0
HAITI            323478.0
AZERBAIJAN      317219.0
JAPAN            279607.0
ARMENIA          191890.0
PAKISTAN         145083.0
IRAQ             136200.0
ECUADOR          135496.0
TURKMENISTAN     117412.0
PERU              102169.0
ISRAEL            90388.0
PORTUGAL          83572.0
GREECE            80378.0
CHILE             64277.0
INDIA              63507.0
TAIWAN            57153.0
Name: Deaths, dtype: float64
```

```
In [23]: #1.2
#I got a inspiration from Yi Yang, 可以用作计数器, 表示每一行（即每一次地震记录）都
Sig_Eqs['Number']=1
#将每年震级大于3的Number加起来
Sig_Eqs2 = Sig_Eqs.loc[Sig_Eqs['Ms']>3].groupby(['Year']).sum()['Number']
#画图
Sig_Eqs2.plot.line()
#观察到从公元1500年开始, 地震频率越来越高, 像指数型增长, 有可能是全球进入了地震活
```

```
Out[23]: <Axes: xlabel='Year'>
```



```
In [63]: #1.3
#本题向我的师姐龙师倩寻求了帮助
#检查输入的是否为DataFrame，并返回地震数量
def get_num_eq(largest_eq):
    if isinstance(largest_eq,pd.DataFrame):
        return len(largest_eq)
    else:
        return 0

def CountEq_LargestEq(country):
    country_data = Sig_Eqs[Sig_Eqs['Country'] == country]
    #如果有地震数量
    if len(country_data) > 0:
        #根据Ms降序排列，找出最大的地震
        largest_eq = country_data.sort_values('Ms', ascending = False)
        #获得地震的年份，月份与日期，拼接成字符串
        largest_da = f'{largest_eq.iloc[0]["Year"]}-{largest_eq.iloc[0]["Mo"]}-{largest_eq.iloc[0]["Dy"]}'
        #获取地震发生的地方
        largest_lo = largest_eq.iloc[0]['Location Name']
        #如果没有地震，返回相关信息
    else:
        largest_eq = 'No earthquakes'
        largest_da = 'No earthquakes'
        largest_lo = 'No earthquakes'
    return largest_eq, largest_da, largest_lo
#创建一个空列表储存每个国家的结果
results = []
for country in Sig_Eqs['Country'].unique():
    #调用函数获取地震信息
    largest_eq, largest_da, largest_lo = CountEq_LargestEq(country)
    #获取地震数量
    num_eq = get_num_eq(largest_eq)
    #将各个元素添加到results列表中
```

```
results.append((country, num_eq, largest_da, largest_lo))
#对地震数量进行排序(降序)
results.sort(key=lambda x: x[1], reverse=True)
#打印每个国家的地震信息
for country, num_eq, largest_da, largest_lo in results:
    print(f"Country:{country} Earthquakes:{num_eq} Date:{largest_da} Location:{largest_lo}")
```

Country:CHINA Earthquakes:623 Date:1920.0-12.0-16.0 Location:CHINA: GANSU PROVINCE, SHANXI PROVINCE  
Country:JAPAN Earthquakes:419 Date:869.0-7.0-13.0 Location:JAPAN: SANRIKU  
Country:INDONESIA Earthquakes:412 Date:2004.0-12.0-26.0 Location:INDONESIA: SUMATRA: ACEH: OFF WEST COAST  
Country:IRAN Earthquakes:386 Date:856.0-12.0-22.0 Location:IRAN: DAMGHAN, QUMIS  
Country:TURKEY Earthquakes:337 Date:1999.0-8.0-17.0 Location:TURKEY: ISTANBUL, KOCaeli, SAKARYA  
Country:ITALY Earthquakes:333 Date:1915.0-1.0-13.0 Location:ITALY: MARSICA, AVEZzano, ABRUZZI  
Country:USA Earthquakes:278 Date:1957.0-3.0-9.0 Location:ALASKA  
Country:GREECE Earthquakes:273 Date:1303.0-8.0-8.0 Location:GREECE: CRETE  
Country:PHILIPPINES Earthquakes:226 Date:1897.0-9.0-21.0 Location:PHILIPPINES: MINDANAO, ZAMBOANGA, SULU, ISABELA  
Country:MEXICO Earthquakes:212 Date:1787.0-3.0-28.0 Location:MEXICO: SAN MARCOS, OAXACA  
Country:CHILE Earthquakes:199 Date:1730.0-7.0-8.0 Location:CHILE: VALPARAISO  
Country:PERU Earthquakes:190 Date:1619.0-2.0-14.0 Location:PERU: TRUJILLO, PIURA, SANTA  
Country:RUSSIA Earthquakes:155 Date:1952.0-11.0-4.0 Location:RUSSIA: KAMCHATKA PENINSULA  
Country:PAPUA NEW GUINEA Earthquakes:103 Date:2000.0-11.0-16.0 Location:PAPUA NEW GUINEA: NEW IRELAND, DUKE OF YORK  
Country:INDIA Earthquakes:101 Date:1897.0-6.0-12.0 Location:INDIA: ASSAM; BANGLADESH  
Country:TAIWAN Earthquakes:101 Date:1920.0-6.0-5.0 Location:TAIWAN  
Country:COLOMBIA Earthquakes:80 Date:1904.0-1.0-20.0 Location:PANAMA-COLOMBIA  
Country:NEW ZEALAND Earthquakes:71 Date:1826.0-nan-nan Location:NEW ZEALAND: FJORDLAND  
Country:ECUADOR Earthquakes:68 Date:1906.0-1.0-31.0 Location:ECUADOR: OFF COAST  
Country:AFGHANISTAN Earthquakes:67 Date:1909.0-7.0-7.0 Location:AFGHANISTAN: HINDUKUSH  
Country:VENEZUELA Earthquakes:66 Date:1900.0-10.0-29.0 Location:VENEZUELA: MACUTO  
Country:SOLOMON ISLANDS Earthquakes:62 Date:1977.0-4.0-21.0 Location:SOLOMON ISLANDS  
Country:ALGERIA Earthquakes:57 Date:1980.0-10.0-10.0 Location:ALGERIA: NORTHERN  
Country:ALBANIA Earthquakes:56 Date:1893.0-6.0-14.0 Location:ALBANIA: HIMARA, DHERMI, KUC, KUDHESI, VLORE, KANINA  
Country:VANUATU Earthquakes:56 Date:1910.0-6.0-16.0 Location:VANUATU ISLANDS  
Country:PAKISTAN Earthquakes:53 Date:1945.0-11.0-27.0 Location:PAKISTAN: MAKRAN COAST  
Country:CROATIA Earthquakes:53 Date:1667.0-4.0-6.0 Location:BALKANS NW: CROATIA: DUBROVNIK: RAGUSA  
Country:FRANCE Earthquakes:43 Date:1909.0-6.0-11.0 Location:FRANCE: VERNEGUES, CHARLEVAL, LA ROQUE D'ANTHERON  
Country:USA TERRITORY Earthquakes:40 Date:1902.0-9.0-22.0 Location:GUAM: AGANA  
Country:GUATEMALA Earthquakes:39 Date:1942.0-8.0-6.0 Location:GUATEMALA: NEAR S COAST  
Country:NICARAGUA Earthquakes:39 Date:1898.0-4.0-29.0 Location:NICARAGUA: LEON, CHINANDEGA, MANAGUA  
Country:MYANMAR (BURMA) Earthquakes:38 Date:1912.0-5.0-23.0 Location:MYANMAR (BURMA): MANDALAY, MOGOK, MAYMYO  
Country:EL SALVADOR Earthquakes:38 Date:2001.0-1.0-13.0 Location:EL SALVADOR; GUATEMALA  
Country:COSTA RICA Earthquakes:36 Date:2012.0-9.0-5.0 Location:COSTA RICA: NICOA  
Country:SPAIN Earthquakes:34 Date:1884.0-12.0-25.0 Location:SPAIN: ARENAS DEL RAY, VEGA, ALHAMA, MALAGA  
Country:SYRIA Earthquakes:33 Date:1202.0-5.0-20.0 Location:SYRIA: SOUTHWESTERN

Country:SWITZERLAND Earthquakes:31 Date:1021.0-5.0-12.0 Location:SWITZERLAND  
Country:PORTUGAL Earthquakes:28 Date:1761.0-3.0-30.0 Location:PORTUGAL: LISBON, PORTO  
Country:AZORES (PORTUGAL) Earthquakes:27 Date:1980.0-1.0-1.0 Location:AZORES: T  
ERCEIRA, ANGRA DO HEROISMO  
Country:TAJIKISTAN Earthquakes:27 Date:1907.0-10.0-21.0 Location:TAJIKISTAN: KARATAG  
Country:NEW CALEDONIA Earthquakes:25 Date:1875.0-3.0-28.0 Location:NEW CALEDONIA:  
LOYALTY ISLANDS: LIFOU ISLAND  
Country:ISRAEL Earthquakes:24 Date:1759.0-10.0-30.0 Location:ISRAEL: ZEFAT (SAFE  
D)  
Country:IRAQ Earthquakes:24 Date:1864.0-12.0-2.0 Location:IRAQ: ZURBATIYAH, BADR  
AH, TURSAQ, BAGHDAD  
Country:TONGA Earthquakes:24 Date:1919.0-4.0-30.0 Location:TONGA ISLANDS  
Country:AUSTRALIA Earthquakes:24 Date:1989.0-5.0-23.0 Location:MACQUARIE ISLAND:  
MACQUARIE STATION; NEW ZEALAND  
Country:PANAMA Earthquakes:23 Date:1882.0-9.0-7.0 Location:PANAMA: SAN BLAS ARCHI  
PELAGO  
Country:SLOVENIA Earthquakes:22 Date:1511.0-3.0-26.0 Location:BALKANS NW: SLOVENIA:  
IDRIJA, SKOFJA LOKA  
Country:CANADA Earthquakes:22 Date:1949.0-8.0-22.0 Location:CANADA: QUEEN CHARLOTTE ISLANDS  
Country:SOUTH KOREA Earthquakes:21 Date:1643.0-7.0-25.0 Location:SEA OF JAPAN  
Country:MOROCCO Earthquakes:21 Date:2004.0-2.0-24.0 Location:MOROCCO: AL HOCEIMA,  
IMZOURENE, BENI ABDALLAH  
Country:ARGENTINA Earthquakes:21 Date:1894.0-10.0-27.0 Location:ARGENTINA: LA RIOJA,  
SAN JUAN, MENDOZA  
Country:KERMADEC ISLANDS (NEW ZEALAND) Earthquakes:21 Date:1986.0-10.0-20.0 Location:  
KERMADEC ISLANDS: RAOUL  
Country:NEPAL Earthquakes:20 Date:1505.0-6.0-6.0 Location:NEPAL-INDIA  
Country:HAITI Earthquakes:20 Date:1842.0-5.0-7.0 Location:HAITI: CAP-HAITIEN  
Country:JAMAICA Earthquakes:19 Date:1899.0-6.0-14.0 Location:JAMAICA  
Country:FIJI Earthquakes:19 Date:1919.0-1.0-1.0 Location:FIJI ISLANDS  
Country:BULGARIA Earthquakes:18 Date:1904.0-4.0-4.0 Location:BULGARIA: STRUMA  
Country:DOMINICAN REPUBLIC Earthquakes:18 Date:1946.0-8.0-4.0 Location:DOMINICAN  
REPUBLIC: NORTHEASTERN COAST  
Country:ICELAND Earthquakes:17 Date:2000.0-6.0-17.0 Location:ICELAND: VESTMANNAEYJAR, HELLA  
Country:BANGLADESH Earthquakes:17 Date:1918.0-7.0-8.0 Location:BANGLADESH: SRIMANGAL  
Country:AZERBAIJAN Earthquakes:16 Date:1667.0-11.0-nan Location:AZERBAIJAN: SHEMAKHA (SAMAXI)  
Country:KYRGYZSTAN Earthquakes:15 Date:1911.0-1.0-3.0 Location:KAZAKHSTAN: ALMA-ATA,  
TURKESTAN; AFGHANISTAN  
Country:EGYPT Earthquakes:15 Date:1995.0-11.0-22.0 Location:EGYPT: NUWAYBI; SAUDI ARABIA;  
ISRAEL; JORDAN  
Country:GEORGIA Earthquakes:15 Date:1905.0-10.0-21.0 Location:GEORGIA: CAUCASUS  
Country:SERBIA Earthquakes:15 Date:1980.0-5.0-18.0 Location:BALKANS NW: SERBIA  
Country:ROMANIA Earthquakes:15 Date:1986.0-8.0-30.0 Location:ROMANIA: BUCHAREST,  
KISHINEV-KAGUL  
Country:LEBANON Earthquakes:14 Date:-590.0-nan-nan Location:LEBANON: SUR (TYRE)  
Country:UK Earthquakes:14 Date:1931.0-6.0-7.0 Location:UK: SCARBOROUGH, GRIMSBY  
Country:UZBEKISTAN Earthquakes:14 Date:1976.0-4.0-8.0 Location:UZBEKISTAN: GAZLI  
Country:CUBA Earthquakes:14 Date:1992.0-5.0-25.0 Location:CUBA: PILON, MANZANILLO  
Country:SOUTH AFRICA Earthquakes:14 Date:1942.0-11.0-10.0 Location:SOUTH AFRICA:  
PRINCE EDWARD ISLAND  
Country:ARMENIA Earthquakes:13 Date:1988.0-12.0-7.0 Location:ARMENIA: LENINAKAN,  
SPITAK, KIROVAKAN  
Country:HONDURAS Earthquakes:13 Date:1856.0-8.0-4.0 Location:HONDURAS: COAST

Country:UKRAINE Earthquakes:12 Date:1927.0-9.0-11.0 Location:UKRAINE: CRIMEA: S EBASTOPOL

Country:MACEDONIA Earthquakes:12 Date:1979.0-5.0-24.0 Location:BALKANS NW: MACEDONIA: DEBAR

Country:TURKMENISTAN Earthquakes:11 Date:1895.0-7.0-8.0 Location:TURKMENISTAN: UZUN-ADA

Country:BOSNIA-HERZEGOVINA Earthquakes:11 Date:1990.0-11.0-27.0 Location:BALKANS NW: BOSNIA-HERZEGOVINA: TITOGRAD

Country:YEMEN Earthquakes:10 Date:1982.0-12.0-13.0 Location:YEMEN: DHAMAR

Country:MONTENEGRO Earthquakes:10 Date:1979.0-4.0-15.0 Location:BALKANS NW: MONTENEGRÖ

Country:MARTINIQUE Earthquakes:10 Date:1906.0-12.0-3.0 Location:MARTINIQUE

Country:TUNISIA Earthquakes:9 Date:1957.0-2.0-20.0 Location:TUNISIA: SIDI ABID, SIDI TOUIL (LA MEDJA), CAILLOUX

Country:KAZAKHSTAN Earthquakes:9 Date:1978.0-3.0-24.0 Location:KAZAKHSTAN: ALMA-ATA

Country:GERMANY Earthquakes:9 Date:1978.0-9.0-3.0 Location:GERMANY

Country:ETHIOPIA Earthquakes:9 Date:1961.0-6.0-1.0 Location:ETHIOPIA: KARAKORE

Country:GUADELOUPE Earthquakes:9 Date:1969.0-12.0-25.0 Location:GUADELOUPE: GRAND BOURG

Country:TRINIDAD AND TOBAGO Earthquakes:8 Date:1831.0-12.0-3.0 Location:TRINIDAD & ST. CHRISTOPHER

Country:TANZANIA Earthquakes:8 Date:1910.0-12.0-13.0 Location:TANZANIA: RUKWA

Country:SAMOA Earthquakes:8 Date:1917.0-6.0-26.0 Location:SAMOA ISLANDS

Country:BRAZIL Earthquakes:8 Date:1986.0-11.0-30.0 Location:BRAZIL: JOAO CAMARA, NATAL

Country:CYPRUS Earthquakes:7 Date:1953.0-9.0-10.0 Location:CYPRUS: PAPHOS

Country:AUSTRIA Earthquakes:7 Date:1998.0-4.0-12.0 Location:AUSTRIA: ARNOLDSTEIN; SLOVENIA: BOVEC, KOBARID

Country:CONGO Earthquakes:7 Date:2005.0-12.0-5.0 Location:CONGO: KALEMIE

Country:BOLIVIA Earthquakes:7 Date:1916.0-8.0-25.0 Location:BOLIVIA-NORTHERN CHILE

Country:SOUTH GEORGIA AND THE SOUTH SANDWICH ISLANDS Earthquakes:7 Date:1929.0-6.0-27.0 Location:SOUTH SANDWICH ISLANDS

Country:POLAND Earthquakes:7 Date:2019.0-7.0-1.0 Location:POLAND: KATOWICE

Country:NORTH KOREA Earthquakes:6 Date:1668.0-7.0-31.0 Location:NORTH KOREA: YELLOW SEA

Country:ERITREA Earthquakes:6 Date:1915.0-9.0-23.0 Location:ERITREA: ASMARA

Country:ATLANTIC OCEAN Earthquakes:6 Date:1941.0-11.0-25.0 Location:ATLANTIC OCEAN: NORTHERN

Country:MONGOLIA Earthquakes:6 Date:1905.0-7.0-23.0 Location:MONGOLIA

Country:JORDAN Earthquakes:5 Date:362.0-5.0-24.0 Location:JORDAN: AL-KARAK

Country:GHANA Earthquakes:5 Date:1939.0-6.0-22.0 Location:GHANA: ACCRA

Country:HUNGARY Earthquakes:5 Date:1763.0-6.0-28.0 Location:HUNGARY: GYOR; CZECH OSLOVAKIA: KOMAROM

Country:VIETNAM Earthquakes:5 Date:1983.0-6.0-24.0 Location:VIETNAM

Country:ANTARCTICA Earthquakes:5 Date:1998.0-3.0-25.0 Location:BALLENY ISLANDS

Country:RWANDA Earthquakes:5 Date:2008.0-2.0-14.0 Location:RWANDA: GISENYI

Country:BHUTAN Earthquakes:5 Date:2009.0-9.0-21.0 Location:BHUTAN: TASHIGANG

Country:THAILAND Earthquakes:4 Date:1983.0-4.0-22.0 Location:THAILAND: BANGKOK

Country:MICRONESIA, FED. STATES OF Earthquakes:4 Date:1911.0-8.0-16.0 Location:MICRONESIA, FED. STATES OF: CAROLINE ISLANDS

Country:UGANDA Earthquakes:4 Date:1966.0-3.0-20.0 Location:UGANDA: KICHWAMBA, BONDIBOGYO; TANZANIA; DR CONGO

Country:MALAWI Earthquakes:4 Date:1989.0-3.0-10.0 Location:MALAWI: SALIMA, DEDZA, MOHINJI

Country:SLOVAKIA Earthquakes:3 Date:1443.0-6.0-5.0 Location:SLOVAKIA: BANSKA STI AVNICA

Country:ANTIGUA AND BARBUDA Earthquakes:3 Date:1690.0-4.0-16.0 Location:ANTIGUA; SAINT KITTS AND NEVIS

Country:SOUTH SUDAN Earthquakes:3 Date:1990.0-5.0-20.0 Location:SOUTH SUDAN: JUBA, MAYA; UGANDA: NAKURA

Country:KENYA Earthquakes:3 Date:1928.0-1.0-6.0 Location:KENYA: SUBUKIA

Country:INDIAN OCEAN Earthquakes:3 Date:1928.0-3.0-9.0 Location:INDIAN OCEAN: S

Country:MALAYSIA Earthquakes:3 Date:1976.0-7.0-26.0 Location:MALAYSIA: SABAH: L AHAD,DATU,KANAK

Country:NETHERLANDS Earthquakes:3 Date:1992.0-4.0-13.0 Location:THE NETHERLANDS: ROERMOND; GERMANY: BONN, HEINSBERG

Country:SAUDI ARABIA Earthquakes:3 Date:2009.0-5.0-19.0 Location:SAUDI ARABIA: WE STERN

Country:MOZAMBIQUE Earthquakes:3 Date:2006.0-2.0-22.0 Location:MOZAMBIQUE

Country:OMAN Earthquakes:2 Date:1497.0-nan-nan Location:OMAN: QALHAT

Country:FRENCH GUIANA Earthquakes:2 Date:1774.0-nan-nan Location:FRENCH GUIANA: CAYENNE

Country:SAINT LUCIA Earthquakes:2 Date:1788.0-10.0-12.0 Location:SAINT LUCIA

Country:TOGO Earthquakes:2 Date:1788.0-nan-nan Location:TOGO: AGUNAH, KANA; BENI N: ABOMEY

Country:CANARY ISLANDS Earthquakes:2 Date:1800.0-nan-nan Location:CANARY ISLANDS: PENON DE LA GOMERA

Country:UK TERRITORY Earthquakes:2 Date:1983.0-11.0-30.0 Location:INDIAN OCEAN: CHAGOS ARCHIPELAGO: DIEGO GARCIA

Country:COTE D'IVOIRE Earthquakes:2 Date:1879.0-2.0-11.0 Location:COTE D'IVOIRE: ABIDJAN

Country:SOLOMON SEA Earthquakes:2 Date:1895.0-3.0-6.0 Location:W. SOLOMON SEA

Country:CAMEROON Earthquakes:2 Date:1909.0-4.0-26.0 Location:CAMEROON: BUEA

Country:PACIFIC OCEAN Earthquakes:2 Date:1958.0-11.0-4.0 Location:EAST PACIFIC RIDGE

Country:LAOS Earthquakes:2 Date:2007.0-5.0-16.0 Location:LAOS: BOKEO; THAILAND: CHIANG RAI, CHIAN SAEN

Country:IRELAND Earthquakes:1 Date:1490.0-nan-nan Location:IRELAND: SLIGO, MAYO

Country:SIERRA LEONE Earthquakes:1 Date:1795.0-5.0-20.0 Location:SIERRA LEONE

Country:NORWAY Earthquakes:1 Date:1819.0-8.0-31.0 Location:NORWAY: RANA REGION: LUROY

Country:GRENADA Earthquakes:1 Date:1822.0-12.0-1.0 Location:GRENADA

Country:BARBADOS Earthquakes:1 Date:1831.0-8.0-11.0 Location:BARBADOS, SAINT VINCENT, DOMINICA, ANTIGUA

Country:SAINT VINCENT AND THE GRENADINES Earthquakes:1 Date:1844.0-8.0-30.0 Location:SAINT VINCENT: KINGSTOWN

Country:FRENCH POLYNESIA Earthquakes:1 Date:1848.0-7.0-12.0 Location:FRENCH POLynesia: TAHITI

Country:BRITISH VIRGIN ISLANDS Earthquakes:1 Date:1871.0-9.0-nan Location:BRITISH VIRGIN ISLANDS: TORTOLA ISLAND

Country:SRI LANKA Earthquakes:1 Date:1882.0-1.0-nan Location:SRI LANKA: TRINCOMALEE

Country:URUGUAY Earthquakes:1 Date:1888.0-6.0-5.0 Location:URUGUAY: COLOGNE

Country:MONTSERRAT Earthquakes:1 Date:1897.0-4.0-25.0 Location:MONTSERRAT

Country:KIRIBATI Earthquakes:1 Date:1905.0-6.0-30.0 Location:KIRIBATI: PHOENIX ISLANDS

Country:PALAU Earthquakes:1 Date:1914.0-10.0-23.0 Location:MICRONESIA, FED. STATE OF: CAROLINE ISLANDS

Country:CENTRAL AFRICAN REPUBLIC Earthquakes:1 Date:1921.0-9.0-16.0 Location:CENTRAL AFRICAN REPUBLIC: NOLA

Country:LIBYA Earthquakes:1 Date:1963.0-2.0-21.0 Location:LIBYA: BARCE (AL MARJ)

Country:GABON Earthquakes:1 Date:1974.0-9.0-23.0 Location:GABON

Country:BELGIUM Earthquakes:1 Date:1983.0-11.0-8.0 Location:BELGIUM

Country:GUINEA Earthquakes:1 Date:1983.0-12.0-22.0 Location:GUINEA: GAOUAL-KOUMBA

Country:DJIBOUTI Earthquakes:1 Date:1989.0-8.0-20.0 Location:DJIBOUTI: GALAFI, YABAHI; ETHIOPIA

Country:BERING SEA Earthquakes:1 Date:1991.0-2.0-21.0 Location:BERING SEA

```
Country:WALLIS AND FUTUNA (FRENCH TERRITORY) Earthquakes:1 Date:1993.0-3.0-12.0 Location:FUTUNA ISLAND
Country:SUDAN Earthquakes:1 Date:1993.0-8.0-1.0 Location:SUDAN: KHARTOUM
Country:BURUNDI Earthquakes:1 Date:2004.0-2.0-24.0 Location:BURUNDI: RUYAGA
Country:CZECH REPUBLIC Earthquakes:1 Date:2008.0-11.0-22.0 Location:CZECH REPUBLIC: KARVINA
Country:MADAGASCAR Earthquakes:1 Date:2017.0-1.0-11.0 Location:MADAGASCAR: ANTSIRABE
Country:ZAMBIA Earthquakes:1 Date:2017.0-2.0-24.0 Location:ZAMBIA: KAPUTA
Country:COMOROS Earthquakes:1 Date:2018.0-5.0-15.0 Location:COMOROS: MAYOTTE
Country:nan Earthquakes:0 Date:No earthquakes Location:No earthquakes
```

In [141...]

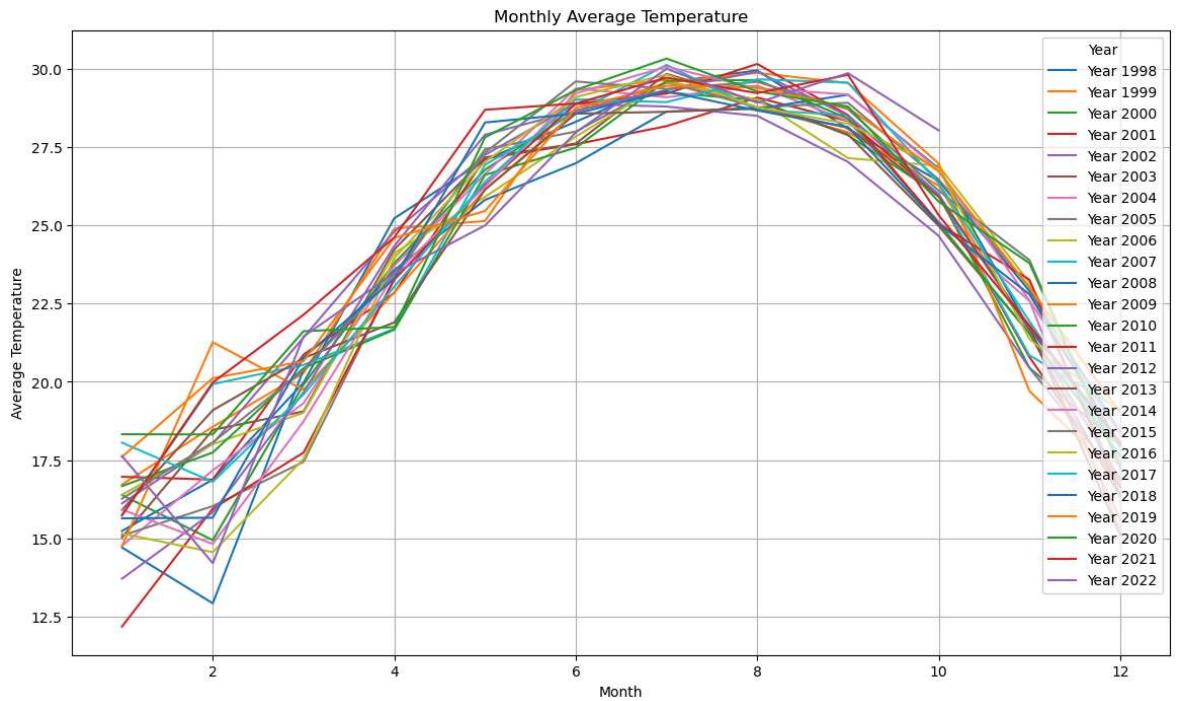
```
#2
#难点是读取文件时有错误，某些列包含混合数据类型，low_memory=False是为了减少混合类型
Weather = pd.read_csv('Baoan_Weather_1998_2022.csv', low_memory = False)
#查看TMP数据发现为(+xxxx,x),我认为+xxxx为10倍的温度，逗号后面的是判断类型的ASCII
#首先我将TMP中的+号去除
Weather['TMP'] = Weather['TMP'].astype(str).str.replace('+','')
#将TMP分为Temperature和QC_CODE两部分
Weather[['Temperature', 'QC_CODE']] = Weather['TMP'].str.split(',', expand=True)
#转换温度为浮点数
Weather['Temperature'] = Weather['Temperature'].astype(float)
#转换QC_CODE为浮点数
Weather['QC_CODE'] = Weather['QC_CODE'].astype(float)
#算真实温度
Weather['Temperature'] = Weather['Temperature'] / 10
#筛选不要的值（即+999.9与ASCII中除了0,1,4,5的值）
new_Weather = Weather[(Weather['Temperature'] != 999.9) & (Weather['QC_CODE'].is
#将日期转换成datetime
new_Weather['DATE'] = pd.to_datetime(new_Weather['DATE'], errors='coerce')
#提取年份
new_Weather['YEAR'] = new_Weather['DATE'].dt.year
#提取月份
new_Weather['MONTH'] = new_Weather['DATE'].dt.month
#按照年和月计算均值
mon_tmp = new_Weather.groupby(['YEAR', 'MONTH'])['Temperature'].mean(numeric_only=True)
#创建透视表
draw_TMP = mon_tmp.pivot_table(index= 'MONTH', columns='YEAR', values='Temperature')
#检验数据
print(draw_TMP)
#创建图形大小为14*8
plt.figure(figsize = (14,8))
#绘制每一年的每月平均温度数据，并为每条曲线添加标签
for year in draw_TMP.columns:
    plt.plot(draw_TMP.index, draw_TMP[year], label=f'Year {year}')
#设置图例
plt.legend(title='Year')

plt.title('Monthly Average Temperature')
plt.xlabel('Month')
plt.ylabel('Average Temperature')
#添加网格线
plt.grid(True)
plt.show()
#过去的25年中，夏天的平均气温差异较小，冬天的平均气温相差较大
```

```
C:\Users\Administrator\AppData\Local\Temp\ipykernel_13880\2976577178.py:12: SettingWithCopyWarning:  
  A value is trying to be set on a copy of a slice from a DataFrame.  
  Try using .loc[row_indexer,col_indexer] = value instead  
  
See the caveats in the documentation: https://pandas.pydata.org/pandas-docs/stable/user\_guide/indexing.html#returning-a-view-versus-a-copy  
    new_Weather['DATE'] = pd.to_datetime(new_Weather['DATE'],errors='coerce')  
C:\Users\Administrator\AppData\Local\Temp\ipykernel_13880\2976577178.py:13: SettingWithCopyWarning:  
  A value is trying to be set on a copy of a slice from a DataFrame.  
  Try using .loc[row_indexer,col_indexer] = value instead  
  
See the caveats in the documentation: https://pandas.pydata.org/pandas-docs/stable/user\_guide/indexing.html#returning-a-view-versus-a-copy  
    new_Weather['YEAR'] = new_Weather['DATE'].dt.year  
C:\Users\Administrator\AppData\Local\Temp\ipykernel_13880\2976577178.py:14: SettingWithCopyWarning:  
  A value is trying to be set on a copy of a slice from a DataFrame.  
  Try using .loc[row_indexer,col_indexer] = value instead  
  
See the caveats in the documentation: https://pandas.pydata.org/pandas-docs/stable/user\_guide/indexing.html#returning-a-view-versus-a-copy  
    new_Weather['MONTH'] = new_Weather['DATE'].dt.month
```

YEAR MONTH	1998	1999	2000	2001	2002	2003	\	
1	15.233447	16.718875	16.379526	16.963957	16.112745	15.025641		
2	16.875304	18.562635	14.937097	16.874513	18.025229	18.463608		
3	19.971246	20.279426	19.911480	20.869220	21.417967	19.055263		
4	25.228365	24.900331	23.794485	22.842232	24.803134	24.231215		
5	27.098454	25.140681	26.779589	27.137987	27.257064	27.174689		
6	28.313107	29.326210	28.672650	27.594037	28.886527	27.617192		
7	29.564120	29.653387	29.250413	28.174542	28.799858	29.846122		
8	29.954622	28.780511	29.005964	29.087143	28.502216	28.970902		
9	28.089919	27.983466	27.882812	28.335639	27.030360	27.907604		
10	26.418671	26.268303	25.975580	25.943776	24.664663	24.982497		
11	22.900497	21.620893	20.447923	20.789146	20.443140	21.555899		
12	18.395948	15.611723	17.961399	16.645134	16.831429	16.333056		
YEAR MONTH	2004	2005	2006	2007	...	2013	2014	\
1	14.758056	15.084211	16.389222	15.723848	...	15.907133	15.917713	
2	17.172687	16.028018	18.007623	19.927136	...	19.092262	14.818253	
3	19.316138	17.441463	19.012534	20.541217	...	20.759623	18.733602	
4	23.715113	23.309179	24.113924	21.697753	...	21.899164	23.437200	
5	26.643284	27.435081	25.911413	26.940621	...	26.146472	26.244975	
6	29.380114	28.000698	27.839331	28.619776	...	28.572851	29.219476	
7	29.095597	29.574290	29.597177	30.132561	...	28.631282	30.075770	
8	29.469241	28.771968	29.020565	28.691817	...	28.759192	29.376224	
9	28.337339	28.920723	27.152222	28.524401	...	28.120438	29.186192	
10	25.114384	26.467429	26.905299	25.995658	...	25.915672	26.776822	
11	22.585112	23.028973	23.174270	20.850485	...	21.735000	22.598955	
12	18.341905	16.646849	17.855177	19.110512	...	15.097270	15.209899	
YEAR MONTH	2015	2016	2017	2018	2019	2020	\	
1	16.264315	15.147379	18.055295	15.635685	17.600706	18.324855		
2	18.067039	14.555028	16.805134	15.656040	20.115196	18.318742		
3	20.372553	17.531276	19.597984	20.772077	20.657625	21.618182		
4	23.444885	23.966387	23.044896	23.304651	24.601877	21.751240		
5	27.325633	27.034511	26.394556	28.290000	25.461492	27.815952		
6	29.602601	29.102607	29.029158	28.573514	28.759583	29.328651		
7	29.366095	29.825605	28.941935	29.287935	29.464819	30.334409		
8	29.588687	28.773104	29.671719	28.673790	29.418567	29.284820		
9	28.713793	28.250939	29.564546	28.145179	28.759158	28.796159		
10	26.107770	26.819556	26.342886	25.047658	26.724289	25.758530		
11	23.890417	21.368980	21.937605	22.797083	23.014127	23.794606		
12	17.223679	19.029062	17.615237	17.962500	19.013705	17.322386		
YEAR MONTH	2021	2022						
1	15.740576	17.626970						
2	19.979144	14.202688						
3	22.142344	21.451882						
4	24.618056	23.473186						
5	28.695461	24.998679						
6	28.898785	27.971803						
7	29.723333	30.005550						
8	29.229172	28.919057						
9	29.814251	29.866772						
10	25.324395	28.035110						
11	21.801146	NaN						
12	18.047225	NaN						

[12 rows x 25 columns]



In [143... pip install netCDF4

```
Collecting netCDF4
  Note: you may need to restart the kernel to use updated package
  s.
```

```
    Downloading netCDF4-1.7.2-cp312-cp312-win_amd64.whl.metadata (1.8 kB)
Collecting cftime (from netCDF4)
    Downloading cftime-1.6.4.post1-cp312-cp312-win_amd64.whl.metadata (8.9 kB)
Requirement already satisfied: certifi in d:\anaconda\lib\site-packages (from net
CDF4) (2024.8.30)
Requirement already satisfied: numpy in d:\anaconda\lib\site-packages (from netCD
F4) (1.26.4)
    Downloading netCDF4-1.7.2-cp312-cp312-win_amd64.whl (7.0 MB)
       ----- 0.0/7.0 MB ? eta :---:-
       ----- 0.0/7.0 MB ? eta :---:-
       ----- 0.0/7.0 MB 445.2 kB/s eta 0:00:16
       ----- 0.0/7.0 MB 281.8 kB/s eta 0:00:25
       ----- 0.1/7.0 MB 393.8 kB/s eta 0:00:18
       ----- 0.1/7.0 MB 492.8 kB/s eta 0:00:14
       ----- 0.1/7.0 MB 532.5 kB/s eta 0:00:13
       ----- 0.2/7.0 MB 726.5 kB/s eta 0:00:10
       ----- 0.3/7.0 MB 842.9 kB/s eta 0:00:08
       ----- 0.4/7.0 MB 997.7 kB/s eta 0:00:07
       ----- 0.5/7.0 MB 1.2 MB/s eta 0:00:06
       ----- 0.7/7.0 MB 1.4 MB/s eta 0:00:05
       ----- 1.0/7.0 MB 1.8 MB/s eta 0:00:04
       ----- 1.2/7.0 MB 2.1 MB/s eta 0:00:03
       ----- 1.5/7.0 MB 2.4 MB/s eta 0:00:03
       ----- 1.9/7.0 MB 2.9 MB/s eta 0:00:02
       ----- 2.6/7.0 MB 3.7 MB/s eta 0:00:02
       ----- 3.3/7.0 MB 4.3 MB/s eta 0:00:01
       ----- 4.2/7.0 MB 5.2 MB/s eta 0:00:01
       ----- 5.8/7.0 MB 6.7 MB/s eta 0:00:01
       ----- 7.0/7.0 MB 7.8 MB/s eta 0:00:01
       ----- 7.0/7.0 MB 7.6 MB/s eta 0:00:00
    Downloading cftime-1.6.4.post1-cp312-cp312-win_amd64.whl (178 kB)
       ----- 0.0/178.9 kB ? eta :---:-
       ----- 178.9/178.9 kB 10.5 MB/s eta 0:00:00
Installing collected packages: cftime, netCDF4
Successfully installed cftime-1.6.4.post1 netCDF4-1.7.2
```

```
In [145...]: pip install xarray
```

```
Requirement already satisfied: xarray in d:\anaconda\lib\site-packages (2023.6.0)
Requirement already satisfied: numpy>=1.21 in d:\anaconda\lib\site-packages (from
xarray) (1.26.4)
Requirement already satisfied: pandas>=1.4 in d:\anaconda\lib\site-packages (from
xarray) (2.2.2)
Requirement already satisfied: packaging>=21.3 in d:\anaconda\lib\site-packages
(from xarray) (23.2)
Requirement already satisfied: python-dateutil>=2.8.2 in d:\anaconda\lib\site-pac
kages (from pandas>=1.4->xarray) (2.9.0.post0)
Requirement already satisfied: pytz>=2020.1 in d:\anaconda\lib\site-packages (fro
m pandas>=1.4->xarray) (2024.1)
Requirement already satisfied: tzdata>=2022.7 in d:\anaconda\lib\site-packages (f
rom pandas>=1.4->xarray) (2023.3)
Requirement already satisfied: six>=1.5 in d:\anaconda\lib\site-packages (from py
thon-dateutil>=2.8.2->pandas>=1.4->xarray) (1.16.0)
Note: you may need to restart the kernel to use updated packages.
```

```
In [147...]: pip install nc-time-axis
```

```
Collecting nc-time-axis
  Downloading nc_time_axis-1.4.1-py3-none-any.whl.metadata (4.7 kB)
Requirement already satisfied: cftime>=1.5 in d:\anaconda\lib\site-packages (from nc-time-axis) (1.6.4.post1)
Requirement already satisfied: matplotlib in d:\anaconda\lib\site-packages (from nc-time-axis) (3.8.4)
Requirement already satisfied: numpy in d:\anaconda\lib\site-packages (from nc-time-axis) (1.26.4)
Requirement already satisfied: contourpy>=1.0.1 in d:\anaconda\lib\site-packages (from matplotlib->nc-time-axis) (1.2.0)
Requirement already satisfied: cycler>=0.10 in d:\anaconda\lib\site-packages (from matplotlib->nc-time-axis) (0.11.0)
Requirement already satisfied: fonttools>=4.22.0 in d:\anaconda\lib\site-packages (from matplotlib->nc-time-axis) (4.51.0)
Requirement already satisfied: kiwisolver>=1.3.1 in d:\anaconda\lib\site-packages (from matplotlib->nc-time-axis) (1.4.4)
Requirement already satisfied: packaging>=20.0 in d:\anaconda\lib\site-packages (from matplotlib->nc-time-axis) (23.2)
Requirement already satisfied: pillow>=8 in d:\anaconda\lib\site-packages (from matplotlib->nc-time-axis) (10.3.0)
Requirement already satisfied: pyparsing>=2.3.1 in d:\anaconda\lib\site-packages (from matplotlib->nc-time-axis) (3.0.9)
Requirement already satisfied: python-dateutil>=2.7 in d:\anaconda\lib\site-packages (from matplotlib->nc-time-axis) (2.9.0.post0)
Requirement already satisfied: six>=1.5 in d:\anaconda\lib\site-packages (from python-dateutil>=2.7->matplotlib->nc-time-axis) (1.16.0)
Downloading nc_time_axis-1.4.1-py3-none-any.whl (17 kB)
Installing collected packages: nc-time-axis
Successfully installed nc-time-axis-1.4.1
Note: you may need to restart the kernel to use updated packages.
```

In [233...]

```
#3
#导入文件
df = pd.read_csv('ibtracs.ALL.list.v04r00.csv',
                  usecols=range(17),
                  #跳过文件中的第一行和第二行
                  skiprows=[1, 2],
                  #将ISO_TIME解析成日期时间格式
                  parse_dates=['ISO_TIME'],
                  #将NOT_NAMED,NAME视为缺失值
                  na_values=['NOT_NAMED', 'NAME'])
df.head()
```

```
C:\Users\Administrator\AppData\Local\Temp\ipykernel_13880\1120637988.py:3: DtypeWarning: Columns (5) have mixed types. Specify dtype option on import or set low_memory=False.
```

```
df = pd.read_csv('ibtracs.ALL.list.v04r00.csv',
```

Out[233...]

	SID	SEASON	NUMBER	BASIN	SUBBASIN	NAME	ISO_TIME	NAT
0	1842298N11080	1842	1	NI	BB	NaN	1842-10-25 06:00:00	
1	1842298N11080	1842	1	NI	BB	NaN	1842-10-25 09:00:00	
2	1842298N11080	1842	1	NI	BB	NaN	1842-10-25 12:00:00	
3	1842298N11080	1842	1	NI	BB	NaN	1842-10-25 15:00:00	
4	1842298N11080	1842	1	NI	AS	NaN	1842-10-25 18:00:00	

In [237...]

## #3.1

```
#由于之前写的代码并没有将WMO_WIND转换成数值类型，所以在处理过程中无法将转换的值（）
df['WMO_WIND'] = pd.to_numeric(df['WMO_WIND'], errors='coerce')
#根据SID进行分组，并找到这个风暴的最大风速，将结果重置成DataFrame格式
WIND_STORM = df.groupby(['SID', 'NAME'])['WMO_WIND'].max().reset_index()
#按照最大风速进行降序排序，选择前10大风暴
largest_hurricanes = WIND_STORM.sort_values('WMO_WIND', ascending = False).head(1)
print(largest_hurricanes)
```

	SID	NAME	WMO_WIND
5027	2015293N13266	PATRICIA	185.0
2114	1980214N11330	ALLEN	165.0
2813	1988253N12306	GILBERT	160.0
4230	2005289N18282	WILMA	160.0
3603	1997253N12255	LINDA	160.0
5346	2019236N10314	DORIAN	160.0
4546	2009288N07267	RICK	155.0
5171	2017242N16333	IRMA	155.0
4220	2005261N21290	RITA	155.0
3688	1998295N12284	MITCH	155.0

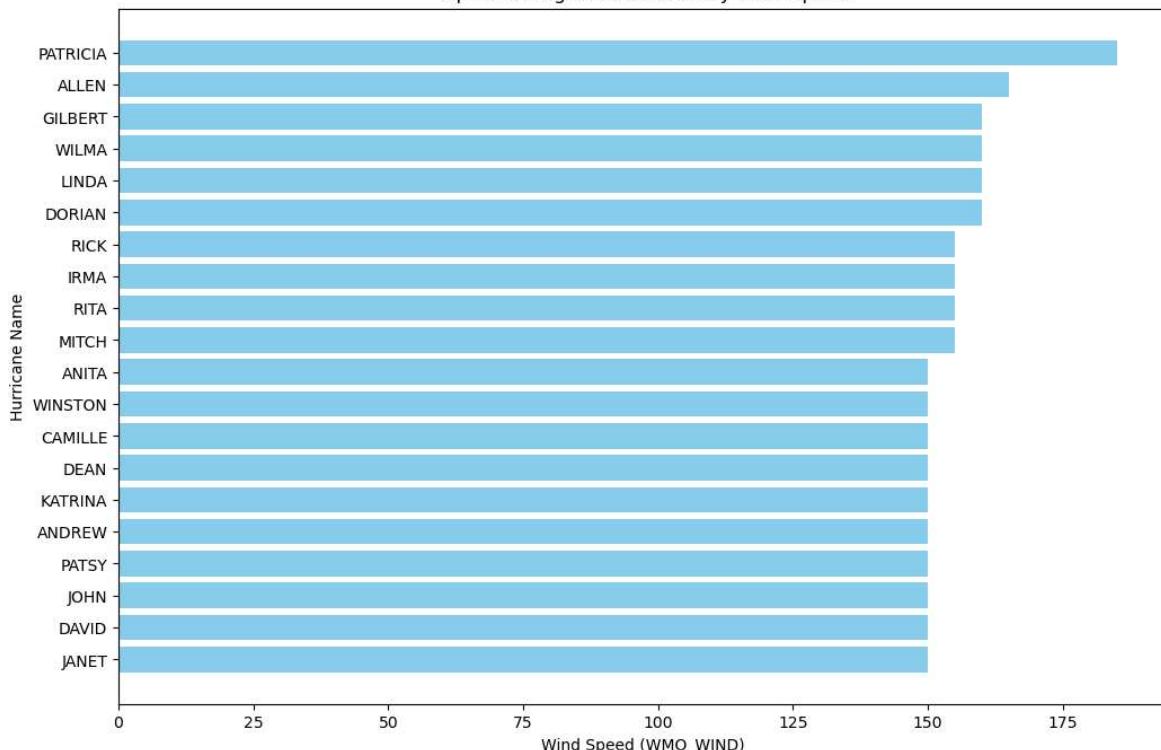
In [241...]

## #3.2

```
large_20_hurricanes = WIND_STORM.sort_values('WMO_WIND', ascending = False).head(20)
print(large_20_hurricanes)
plt.figure(figsize=(12, 8))
plt.barh(large_20_hurricanes['NAME'], large_20_hurricanes['WMO_WIND'], color='sk'
plt.xlabel('Wind Speed (WMO_WIND)')
plt.ylabel('Hurricane Name')
plt.title('Top 20 Strongest Hurricanes by Wind Speed')
# 反转Y轴，使得风速最高的在最上面
plt.gca().invert_yaxis()
plt.show()
```

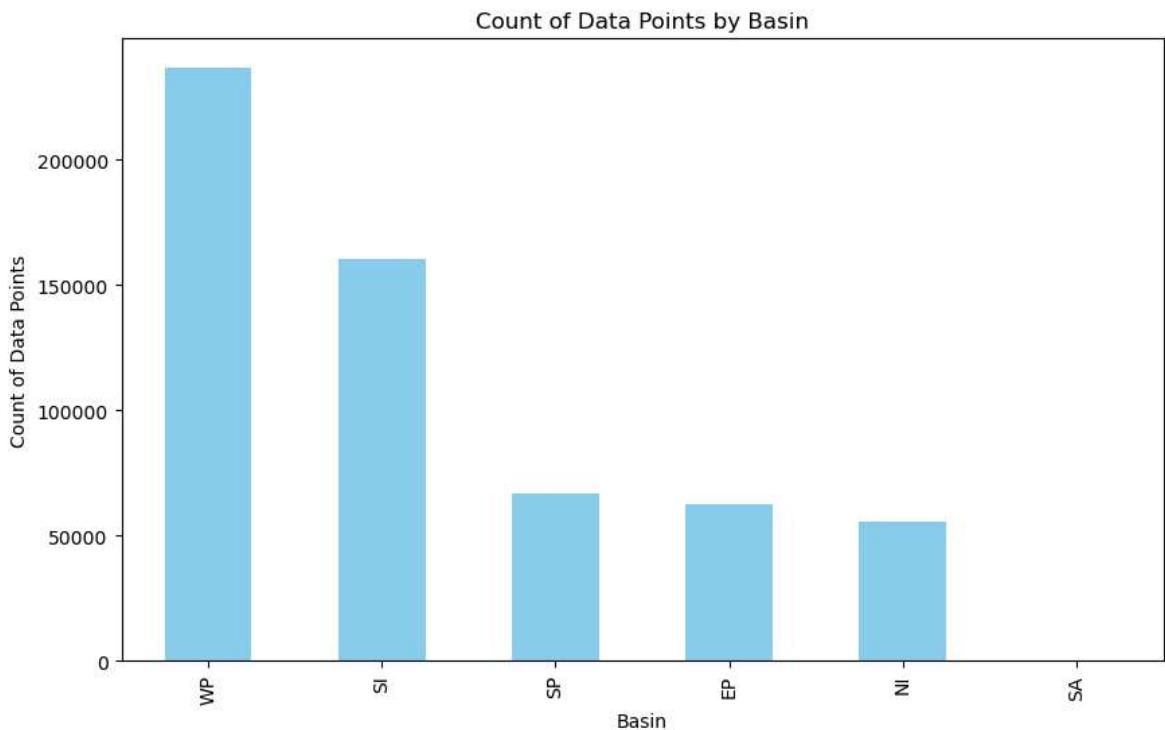
	SID	NAME	WMO_WIND
5027	2015293N13266	PATRICIA	185.0
2114	1980214N11330	ALLEN	165.0
2813	1988253N12306	GILBERT	160.0
4230	2005289N18282	WILMA	160.0
3603	1997253N12255	LINDA	160.0
5346	2019236N10314	DORIAN	160.0
4546	2009288N07267	RICK	155.0
5171	2017242N16333	IRMA	155.0
4220	2005261N21290	RITA	155.0
3688	1998295N12284	MITCH	155.0
1860	1977242N27272	ANITA	150.0
5046	2016041S14170	WINSTON	150.0
1091	1969226N18280	CAMILLE	150.0
4354	2007225N12331	DEAN	150.0
4205	2005236N23285	KATRINA	150.0
3157	1992230N11325	ANDREW	150.0
407	1959249N19181	PATSY	150.0
3333	1994222N11267	JOHN	150.0
2042	1979238N12324	DAVID	150.0
283	1955265N13307	JANET	150.0

Top 20 Strongest Hurricanes by Wind Speed



In [245...]

```
#3.3
#计算每个BASIN数据点的数量
basin_counts = df['BASIN'].value_counts()
#绘制条形图
plt.figure(figsize=(10, 6))
basin_counts.plot(kind='bar', color='skyblue')
plt.xlabel('Basin')
plt.ylabel('Count of Data Points')
plt.title('Count of Data Points by Basin')
plt.show()
```



In [249...]

```
#3.4
plt.figure(figsize=(10, 6))
#绘制六边形图, mincnt=1是为了确保至少有一个数据点的六边形会被显示
plt.hexbin(df['LON'], df['LAT'], gridsize=50, cmap='Blues', mincnt=1)
plt.colorbar(label='Count of Data Points')
plt.xlabel('Longitude')
plt.ylabel('Latitude')
plt.title('Hexbin Plot of Data Points Location')
plt.show()
```



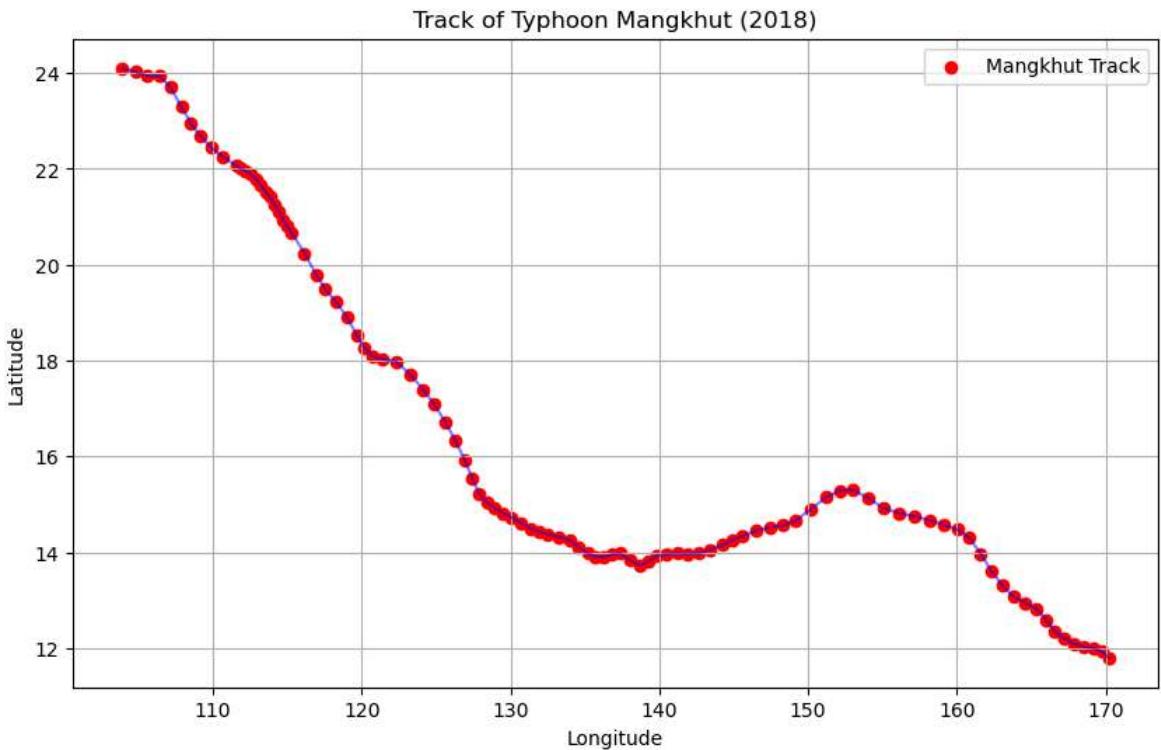
In [253...]

```
#3.5
#寻找名字为MANGKHUT, 年份为2018年的数据
mangkhut = df[(df['NAME'] == 'MANGKHUT') & (df['ISO_TIME'].dt.year == 2018)]
```

```

# 绘制散点图
plt.figure(figsize=(10, 6))
#绘制散点图
plt.scatter(mangkhut['LON'], mangkhut['LAT'], c='red', marker='o', label='Mangkhut')
#连接轨迹点
plt.plot(mangkhut['LON'], mangkhut['LAT'], linestyle='-', color='blue', alpha=0.5)
plt.title('Track of Typhoon Mangkhut (2018)')
plt.xlabel('Longitude')
plt.ylabel('Latitude')
#添加图例
plt.legend()
plt.grid()
plt.show()

```



In [261]: #3.6  
`filtered_df = df[(df['ISO_TIME'].dt.year >= 1970) & (df['BASIN'].isin(['WP', 'EP']))]`  
`filtered_df`

Out[261...]

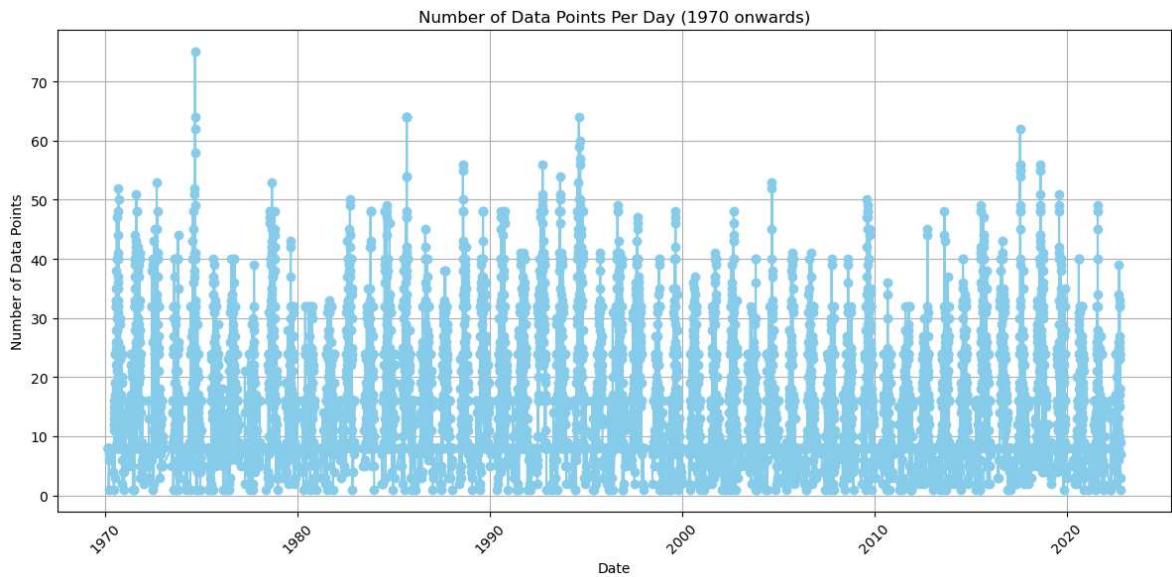
	SID	SEASON	NUMBER	BASIN	SUBBASIN	NAME	ISO_TIME
350393	1970050N07151	1970	22	WP	MM	NANCY	1970-02-10 00:00:00
350394	1970050N07151	1970	22	WP	MM	NANCY	1970-02-10 03:00:00
350395	1970050N07151	1970	22	WP	MM	NANCY	1970-02-10 06:00:00
350396	1970050N07151	1970	22	WP	MM	NANCY	1970-02-10 09:00:00
350397	1970050N07151	1970	22	WP	MM	NANCY	1970-02-10 12:00:00
...	...	...	...	...	...	...	.
707084	2022275N10316	2022	76	EP	MM	JULIA	2022-10-11 15:00:00
707085	2022275N10316	2022	76	EP	MM	JULIA	2022-10-11 18:00:00
707173	2022286N15151	2022	80	WP	MM	NaN	2022-10-12 12:00:00
707174	2022286N15151	2022	80	WP	MM	NaN	2022-10-12 15:00:00
707175	2022286N15151	2022	80	WP	MM	NaN	2022-10-12 18:00:00

176352 rows × 17 columns



In [263...]

```
#3.7
#按日期计算数据点的数量
daily_counts = filtered_df['ISO_TIME'].dt.date.value_counts().sort_index()
# 绘制折线图(绘制柱状图的时候x轴标签重叠在一起了, 因此使用折线图)
plt.figure(figsize=(12, 6))
daily_counts.plot(kind='line', color='skyblue', marker='o')
plt.xlabel('Date')
plt.ylabel('Number of Data Points')
plt.title('Number of Data Points Per Day (1970 onwards)')
# 旋转x轴标签以便于阅读
plt.xticks(rotation=45)
plt.grid()
# 自动调整布局以避免标签重叠
plt.tight_layout()
plt.show()
```



In [389]:

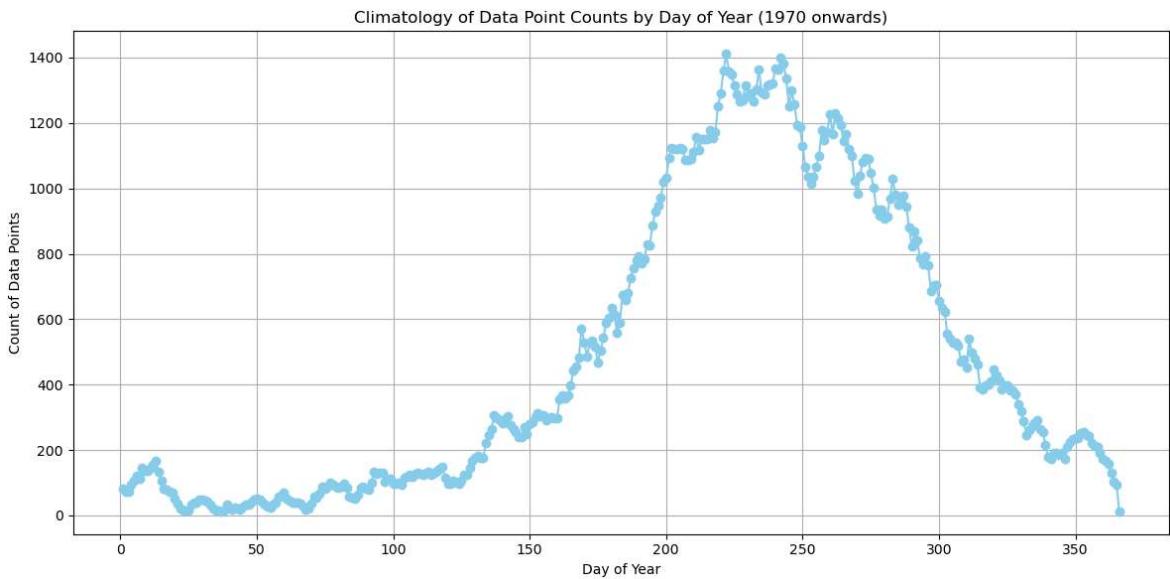
```
#3.8
#将时间转化成DOY
filtered_df['day_of_year'] = filtered_df['ISO_TIME'].dt.dayofyear

# 计算每一天的数据点数量
climatology_counts = filtered_df['day_of_year'].value_counts().sort_index()

# 绘制气候学图表
plt.figure(figsize=(12, 6))
climatology_counts.plot(kind='line', color='skyblue', marker='o')
plt.xlabel('Day of Year')
plt.ylabel('Count of Data Points')
plt.title('Climatology of Data Point Counts by Day of Year (1970 onwards)')
plt.grid()
plt.tight_layout()
plt.show()
```

```
C:\Users\Administrator\AppData\Local\Temp\ipykernel_13880\4235628547.py:2: SettingWithCopyWarning:
A value is trying to be set on a slice from a DataFrame.
Try using .loc[row_indexer,col_indexer] = value instead
```

```
See the caveats in the documentation: https://pandas.pydata.org/pandas-docs/stable/user_guide/indexing.html#returning-a-view-versus-a-copy
filtered_df['day_of_year'] = filtered_df['ISO_TIME'].dt.dayofyear
```

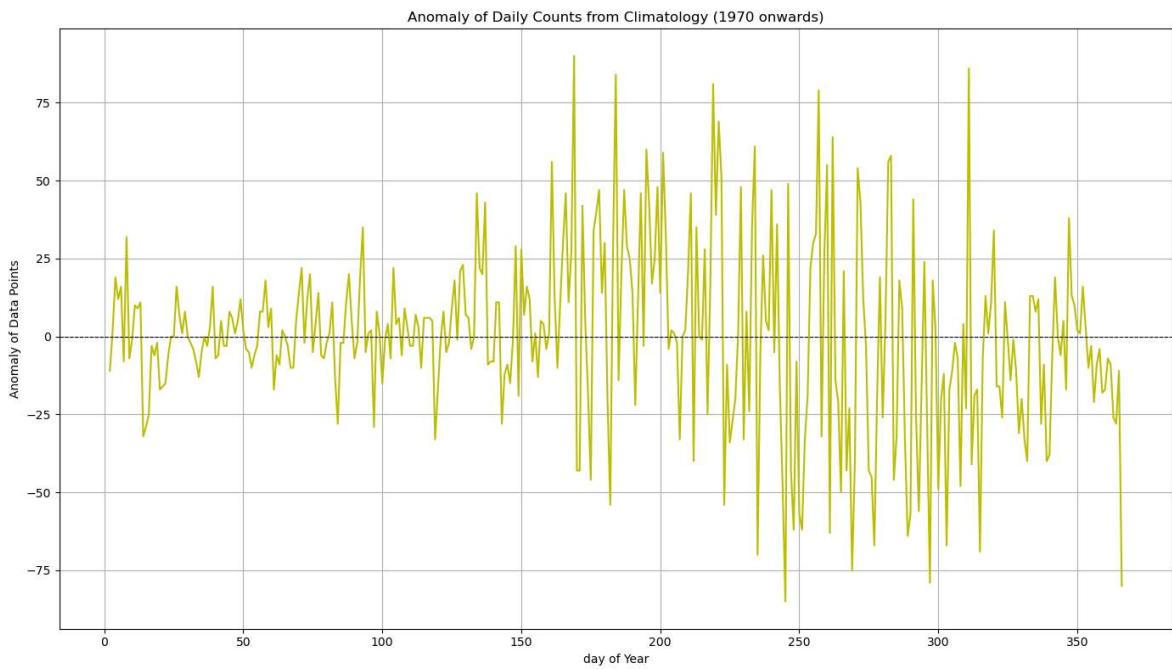


In [415...]

```
#3.9
#我从杨诣那里得到了灵感
#计算气候学平均值
climatology_mean = filtered_df.groupby('day_of_year').size().reset_index(name='Count')
#计算实际计数
daily_counts = filtered_df.groupby('day_of_year').size().reset_index(name='Actual_Count')
#计算异常值
daily_counts['Anomaly']=daily_counts['Actual_Count']- climatology_mean.set_index('day_of_year')['Count']
#显示结果
print(daily_counts)
plt.figure(figsize=(14, 8))
plt.plot(daily_counts['day_of_year'], daily_counts['Anomaly'], color='y')
#添加水平线表示平均值
plt.axhline(0, color='black', linestyle='--', linewidth=0.8)
plt.xlabel('day of Year')
plt.ylabel('Anomaly of Data Points')
plt.title('Anomaly of Daily Counts from Climatology (1970 onwards)')
plt.grid(True)
plt.tight_layout()
plt.show()
```

	day_of_year	Actual_Count	Anomaly
0	1	83	NaN
1	2	72	-11.0
2	3	74	2.0
3	4	93	19.0
4	5	105	12.0
..	...	...	...
361	362	158	-9.0
362	363	132	-26.0
363	364	104	-28.0
364	365	93	-11.0
365	366	13	-80.0

[366 rows x 3 columns]



In [567...]

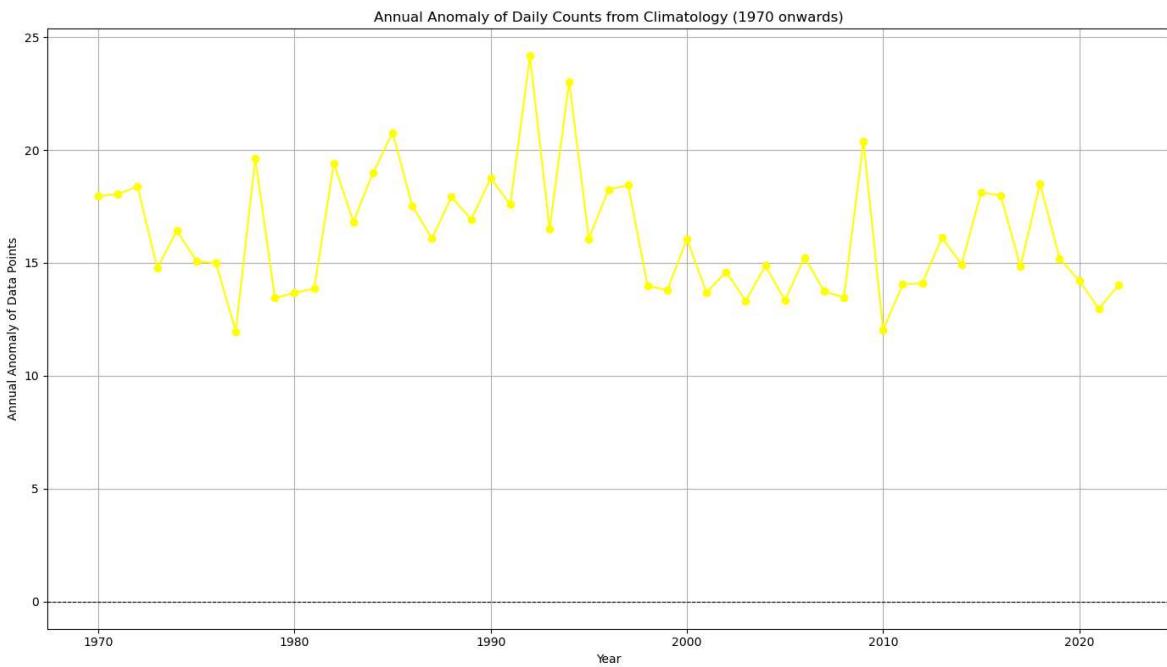
```
#3.10
# 将异常值设置为时间序列
daily_counts.index = pd.to_datetime(daily_counts.index)

# 以年分辨率进行重采样并计算均值
annual_counts = daily_counts.resample('Y').mean()

# 绘制年均异常值图表
plt.figure(figsize=(14, 8))
plt.plot(annual_counts.index.year, annual_counts, color='yellow', marker='o', li
plt.axhline(0, color='black', linestyle='--', linewidth=0.8) # 添加水平线表示平
plt.xlabel('Year')
plt.ylabel('Annual Anomaly of Data Points')
plt.title('Annual Anomaly of Daily Counts from Climatology (1970 onwards)')
plt.grid(True)
plt.tight_layout()
plt.show()
#1985, 1992, 1994, 2009年的飓风活动异常突出
```

C:\Users\Administrator\AppData\Local\Temp\ipykernel\_13880\1546102801.py:6: Future Warning: 'Y' is deprecated and will be removed in a future version, please use 'Y E' instead.

```
annual_counts = daily_counts.resample('Y').mean()
```



In [555...]

```
#4.1 选择了布法罗尼亚加拉国际机场的信息进行导入
inf = pd.read_csv('72528014733.csv', low_memory = False)
inf.head()
#将+9999,9的值替换成NaN
inf = inf.replace(9999.9, np.nan)
#删除为全部NaN的列
inf=inf.dropna(axis=1,how="all")
#使用前一个非NaN值来填充NaN值, 即进行向后填充
inf=inf.fillna(axis=0,method='bfill')
```

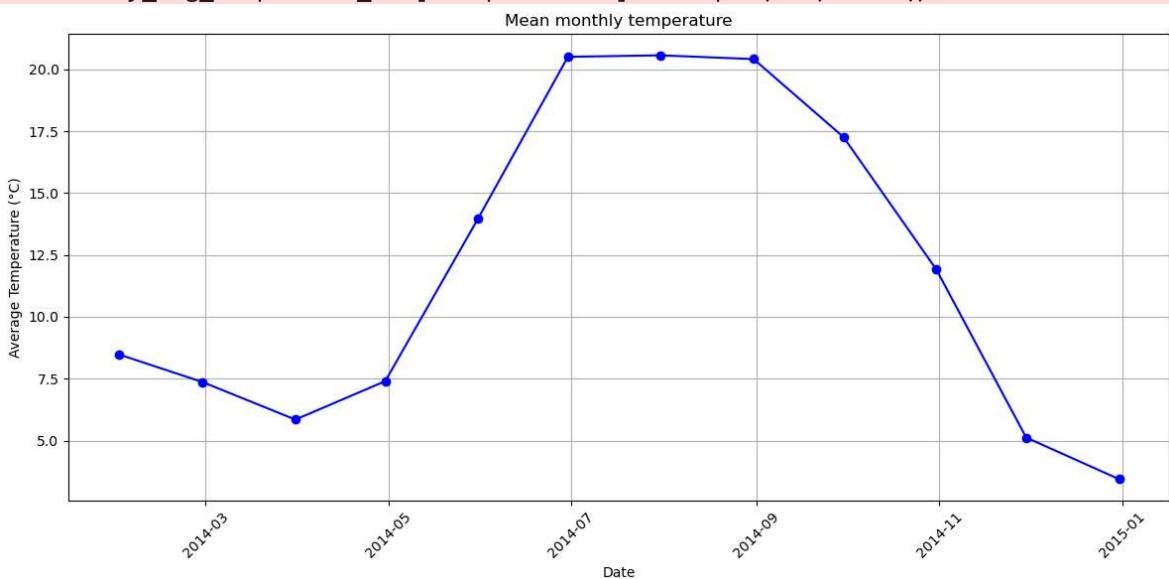
C:\Users\Administrator\AppData\Local\Temp\ipykernel\_13880\4287247817.py:9: Future Warning: DataFrame.fillna with 'method' is deprecated and will raise in a future version. Use obj.ffill() or obj.bfill() instead.  
inf=inf.fillna(axis=0,method='bfill')

In [557...]

```
#4.2
inf['DATE'] = pd.to_datetime(inf['DATE'], errors='coerce')
inf.set_index('DATE', inplace=True)
inf['TMP'] = inf['TMP'].astype(str).str.replace('+',' ',regex=False).str.replace(
#print(inf['TMP'].head())
inf[['temperature', 'qc_code']] = inf['TMP'].str.split(',',expand=True)
inf['temperature'] = inf['temperature'].astype(float)
inf['qc_code'] = inf['qc_code'].astype(str)
inf['temperature'] = inf['temperature'] / 10
valid_qc_codes = ['0', '1', '4', '5'] # 合格的质量控制代码
new_inf = inf[inf['qc_code'].isin(valid_qc_codes)]
monthly_avg_temp = new_inf['temperature'].resample('M').mean()
plt.figure(figsize=(12, 6))
plt.plot(monthly_avg_temp.index, monthly_avg_temp, marker='o', linestyle='-', color='blue')
plt.title('Mean monthly temperature')
plt.xlabel('Date')
plt.ylabel('Average Temperature (°C)')
plt.grid(True)
plt.xticks(rotation=45) # 旋转 x 轴标签以便更好地显示
plt.tight_layout()
plt.show()
```

```
C:\Users\Administrator\AppData\Local\Temp\ipykernel_13880\1937861391.py:12: FutureWarning: 'M' is deprecated and will be removed in a future version, please use 'ME' instead.
```

```
monthly_avg_temp = new_inf['temperature'].resample('M').mean()
```



```
In [569...]
```

#### #4.3

```
# (1) 对均值、中位数、标准差、最小值和最大值进行统计检查
summary_stats = new_inf['temperature'].describe()
print(summary_stats)
#平均温度相对较低
```

```
count    14374.000000
mean     11.471428
std      7.615531
min      0.000000
25%     4.400000
50%     11.000000
75%     17.800000
max     31.100000
Name: temperature, dtype: float64
```

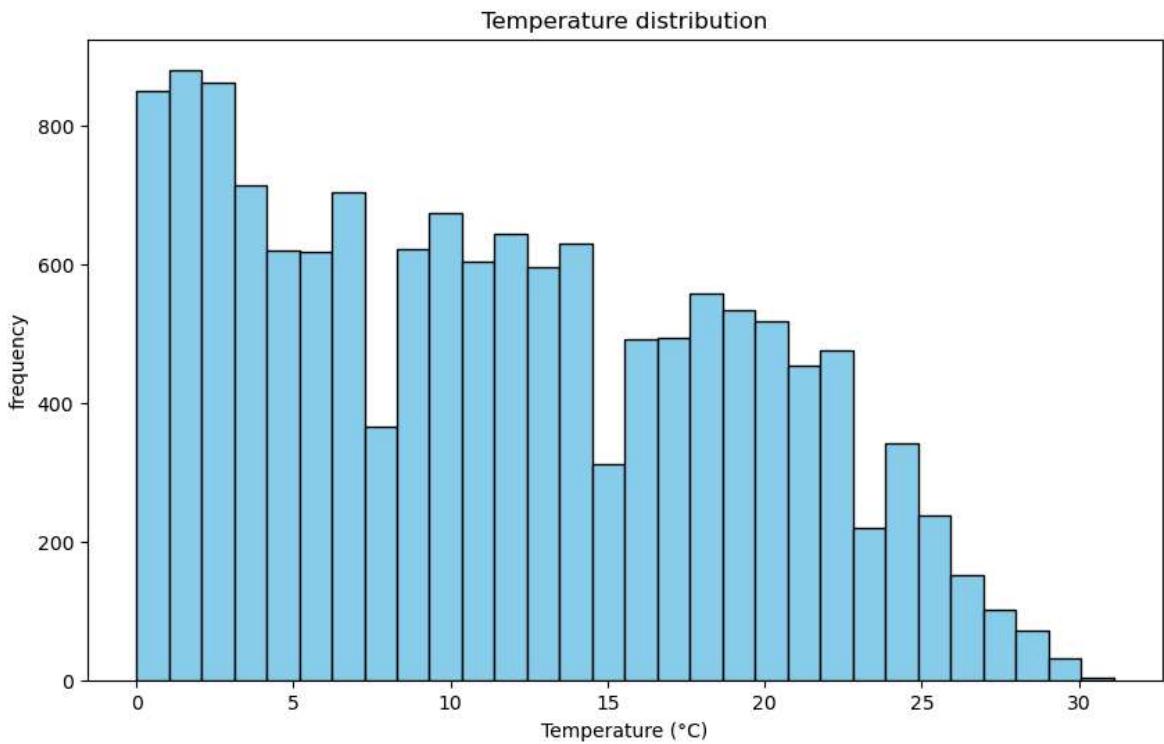
```
In [571...]
```

```
# (2) 对缺失值进行检查
missing_values = new_inf['temperature'].isnull().sum()
print(f'温度的缺失值数量: {missing_values}')
```

温度的缺失值数量: 0

```
In [575...]
```

```
# (3) 使用直方图检查温度值的分布
plt.figure(figsize=(10, 6))
new_inf['temperature'].hist(bins=30, color='skyblue', edgecolor='black')
plt.title('Temperature distribution')
plt.xlabel('Temperature (°C)')
plt.ylabel('frequency')
plt.grid()
plt.show()
#温度在0-5区间的较多，高温较少
```



In [577...]

```
#(4) 使用四分位数法识别潜在的异常值
Q1 = new_inf['temperature'].quantile(0.25)
Q3 = new_inf['temperature'].quantile(0.75)
IQR = Q3 - Q1
lower_bound = Q1 - 1.5 * IQR
upper_bound = Q3 + 1.5 * IQR

outliers = new_inf[(new_inf['temperature'] < lower_bound) | (new_inf['temperature'] > upper_bound)]
print(f'异常值数量: {len(outliers)})'
```

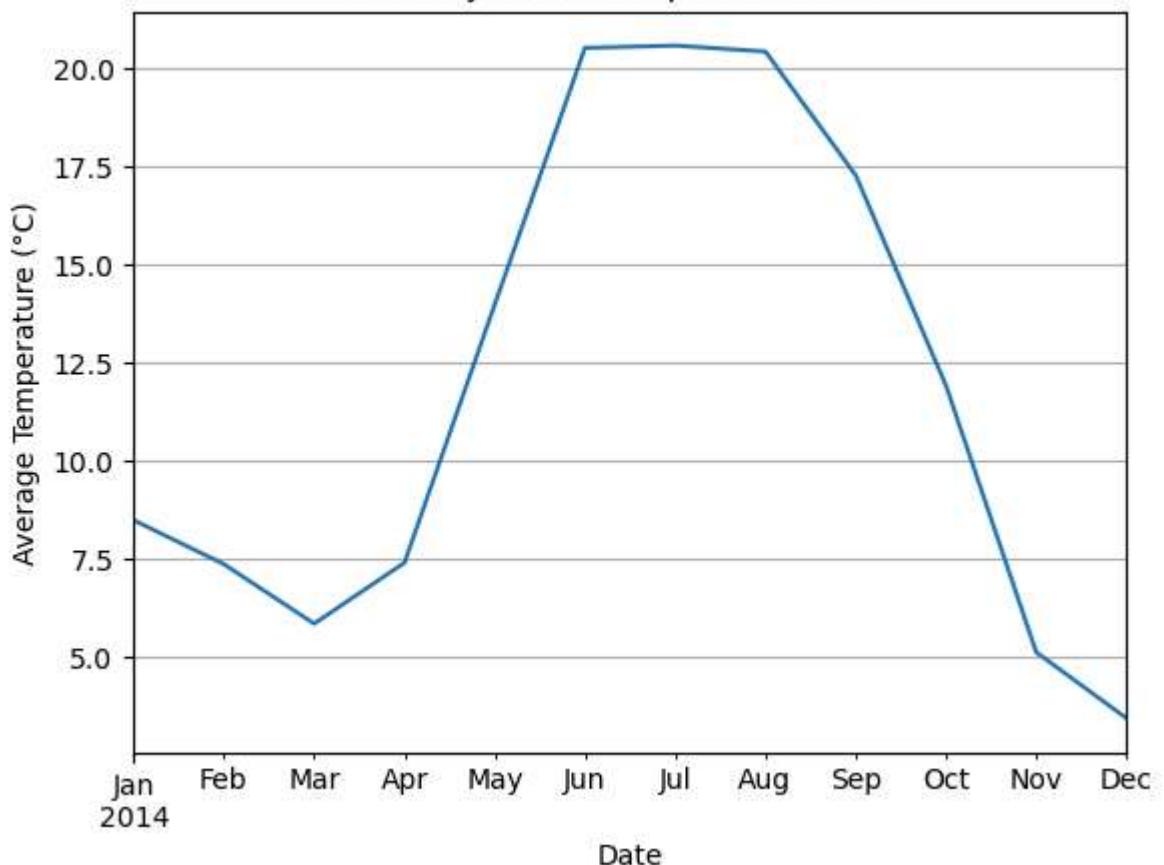
异常值数量: 0

In [579...]

```
#(5) 平均月温度随时间的变化趋势
monthly_avg_temp.plot()
plt.title('Monthly mean temperature trend')
plt.xlabel('Date')
plt.ylabel('Average Temperature (°C)')
plt.grid()
plt.show()

# 计算趋势线的斜率
slope = np.polyfit(range(len(monthly_avg_temp)), monthly_avg_temp, 1)[0]
print(f'趋势斜率: {slope}')
```

Monthly mean temperature trend



趋势斜率: 0.1241233345598398