In [1]:

```
#1.1
import pandas as pd
import numpy as np
from matplotlib import pyplot as plt
# 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.
data = pd.read_csv(’C:/Users/ASUS/Desktop/ESS5023 Assigment/Sig_Eqs.csv’)
data1 = data[[’Country’,’Deaths’]]
print(data1.groupby([’Country’]).sum().sort_values(’Deaths’,ascending=False).iloc[0:10])
```

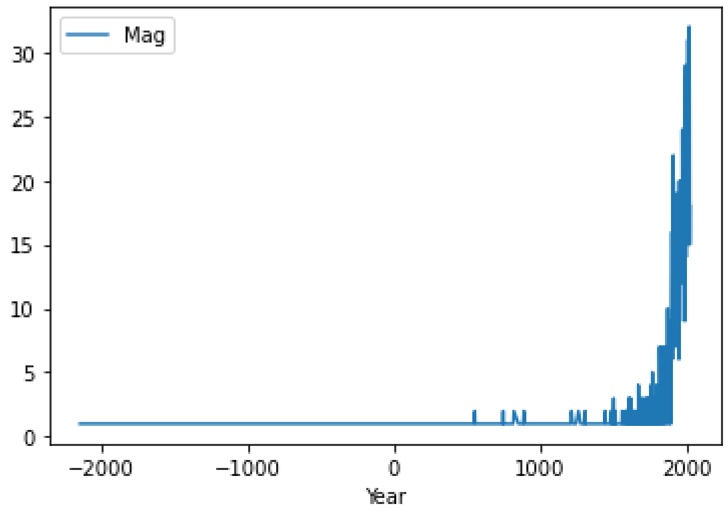
```
              Deaths
Country
CHINA      2074900.0
TURKEY     1074769.0
IRAN       1011437.0
SYRIA       439224.0
ITALY       434863.0
HAITI       323472.0
AZERBAIJAN  317219.0
JAPAN       278138.0
ARMENIA     191890.0
PAKISTAN    148783.0
```

In [2]:

```
#1.2
# 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?
data2 = data[[’Year’,’Mag’]].loc[data[’Mag’]>6.0]
data2.groupby(’Year’).count()[’Mag’].plot()
plt.legend()
```

Out[2]:

```
<matplotlib.legend.Legend at 0x25f77059ac0>
```



In [3]:

```
#1.3
# Write a function CountEq_LargestEq that returns both (1) the total number of earthquakes since 215
# 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.
def CountEq_LargestEq(a):
    data3 = data.loc[data['Country']==a]
    a = len(data3)
    b= data['Year'].iloc[data3['Mag'].idxmax()]
    print(a,int(b))

CountEq_LargestEq('CHINA')
```

```
610 1668
```

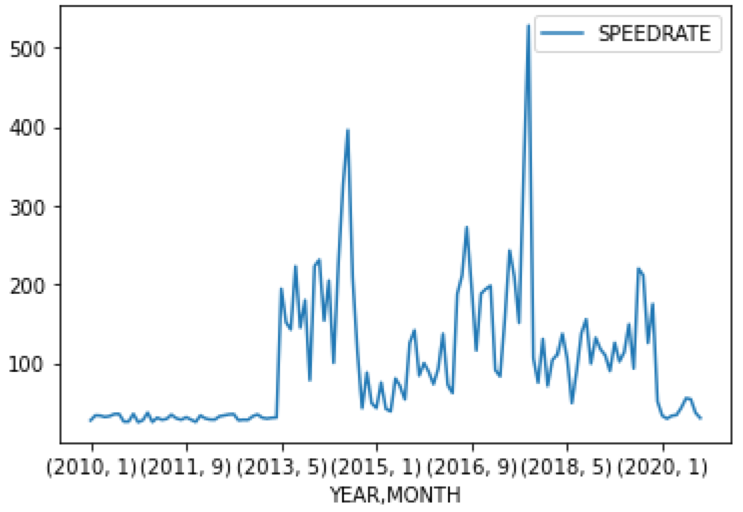
In [4]:

```
#2
# -*- coding: utf-8 -*-

# 1.direction angle : 001~360 Angular Degrees 999 = Missing. If type code (below) = V, then 999 indi
# 2.direction quality code :
    #0 = Passed gross limits check
    #1 = Passed all quality control checks
    #2 = Suspect
    #3 = Erroneous
    #4 = Passed gross limits check, data originate from an NCEI data source
    #5 = Passed all quality control checks, data originate from an NCEI data source
    #6 = Suspect, data originate from an NCEI data source
    #7 = Erroneous, data originate from an NCEI data source
    #9 = Passed gross limits check if element is present
# 3.type code :
    #A = Abridged Beaufort
    #B = Beaufort
    #C = Calm
    #H = 5-Minute Average Speed
    #N = Normal
    #R = 60-Minute Average Speed
    #Q = Squall
    #T = 180 Minute Average Speed
    #V = Variable
    #9 = Missing
# 4.speed rate : 0~900 m/s 9999 = Missing
# 5.speed quality code :
    #0 = Passed gross limits check
    #1 = Passed all quality control checks
    #2 = Suspect
    #3 = Erroneous
    #4 = Passed gross limits check, data originate from an NCEI data source
    #5 = Passed all quality control checks, data originate from an NCEI data source
    #6 = Suspect, data originate from an NCEI data source
    #7 = Erroneous, data originate from an NCEI data source
    #9 = Passed gross limits check if element is present
import pandas as pd
import numpy as np
from matplotlib import pyplot as plt

data = pd.read_csv('C:/Users/ASUS/Desktop/ESS5023 Assigment/2281305.csv')
data = data[['DATE','WND']]
data['MONTH'] = pd.to_datetime(data['DATE']).dt.month
data['YEAR'] = pd.to_datetime(data['DATE']).dt.year
a = data['WND'].str.split(',', expand = True)
a['DATE'] = data['DATE']
data = pd.merge(data,a,on='DATE')
data = data[['YEAR','MONTH',3]]
data.rename(columns={3:'SPEEDRATE'},inplace=True)
data['SPEEDRATE'] = data['SPEEDRATE'].apply(int)
data.groupby(['YEAR','MONTH']).mean().plot()
```

```
C:\Users\ASUS\anaconda3\lib\site-packages\IPython\core\interactiveshell.py:3165: Dty
peWarning: Columns (4,8,9,12,15,21,22,24,26,31,33,34) have mixed types.Specify dtype
option on import or set low_memory=False.
  has_raised = await self.run_ast_nodes(code_ast.body, cell_name,
```

Out[4]:

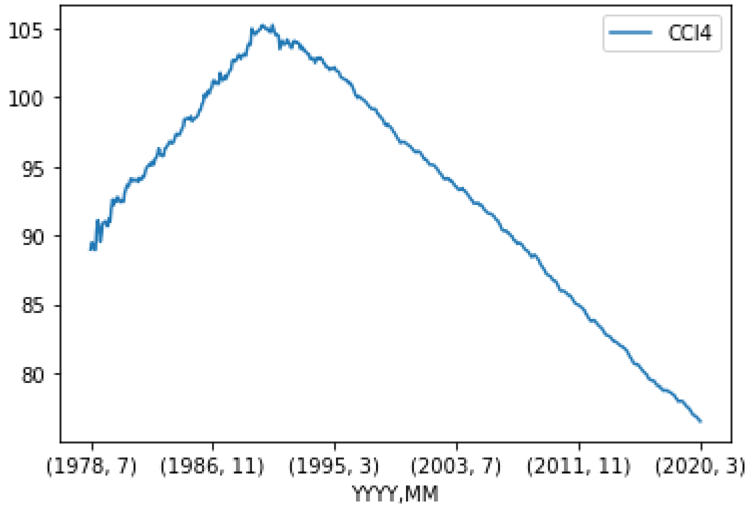
```
<AxesSubplot:xlabel=’YEAR,MONTH’>
```



In [5]:

```
#3
# -*- coding: utf-8 -*-
import pandas as pd
import numpy as np
from matplotlib import pyplot as plt

data = pd.read_csv(’C:/Users/ASUS/Desktop/ESS5023 Assigment/global_mean_md.csv’)
data.groupby([’YYYY’,’MM’]).mean()[’CCl4’].plot()
plt.legend([’CCl4’])
print(data[’CCl4’].describe())
```

```
count     501.00000
mean       93.18991
std         8.09018
min        76.54100
25%        87.14000
50%        94.12000
75%       100.05100
max       105.23300
Name: CCl4, dtype: float64
```



In [ ]: