Explore Weather Trends using Python's Pandas

- By Jacob Barrow

Extract the Data

- Write a SQL query to extract the city level data.
  - select year, avg_temp from city_data where city='San Jose';
- Write a SQL query to extract the global data.
  - select * from global_data;

In [1]:
```
import pandas as pd
```

Open up the CSV with pandas

In [2]:
```
globalDF = pd.read_csv("Global Data.csv")
SanJoseDF = pd.read_csv("San Jose Data.csv")
```

Create a line chart comparing your city's temperatures with the global temperatures.

- Make sure to plot the moving average rather than the yearly averages in order to smooth out the lines, making trends more observable

In [3]:
```
print(globalDF.head(), " .... length is: ", len(globalDF))
```

```
   year  avg_temp
0  1750      8.72
1  1751      7.98
2  1752      5.78
3  1753      8.39
4  1754      8.47  .... length is:  266
```

In [4]:
```
print(SanJoseDF.head(), " .... length is: ",  len(SanJoseDF))
```

```
   year  avg_temp
0  1849     14.12
1  1850     13.80
2  1851     14.39
3  1852     13.81
4  1853     14.40  .... length is:  165
```

In [25]:
```
differenceOfYears = len(globalDF['year']) - len(set([x for x in globalDF['year']]).inte
print(differenceOfYears)
```

```
101
```

Custom 7 year moving average calculator

In [6]:
```
def calculate_7_year_moving_avg(arr):
    if len(arr) < 7:
        return "Not enough"

    mvAvg = [0 for _ in range(len(arr)-6)]

    for i in range(len(arr)-6):
        mvAvg[i] = sum(arr[i:i+7])/7

    return mvAvg
```

- Moving average graph for all years within globalDF: 7 year moving averages
- Used the following resource:
  https://matplotlib.org/stable/gallery/lines_bars_and_markers/simple_plot.html

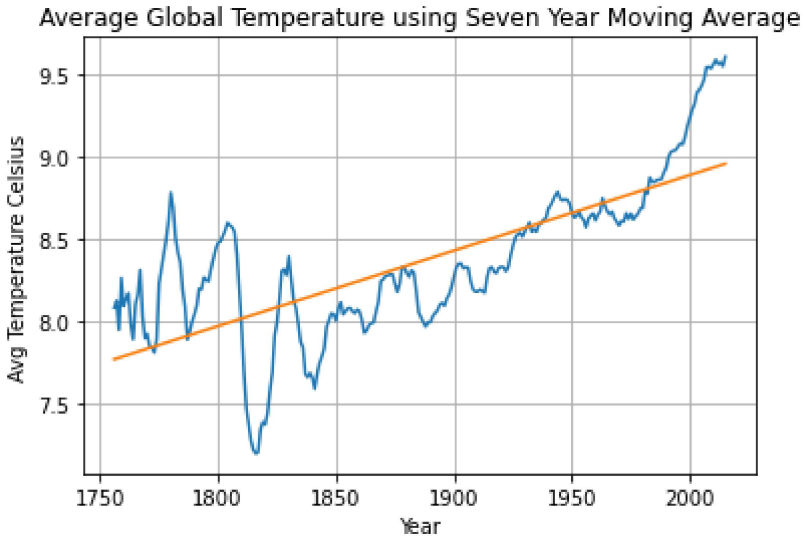
In [7]:
```
import matplotlib
import matplotlib.pyplot as plt
import numpy as np

# Data for plotting
s = calculate_7_year_moving_avg(globalDF['avg_temp'])
t = globalDF['year'][6:]

fig, ax = plt.subplots()
ax.plot(t, s)
m, b = np.polyfit(t, s, 1)
plt.plot(t, m*t + b)


ax.set(xlabel='Year', ylabel='Avg Temperature Celsius',
       title='Average Global Temperature using Seven Year Moving Average')
ax.grid()

fig.savefig("test.png")
plt.show()
```



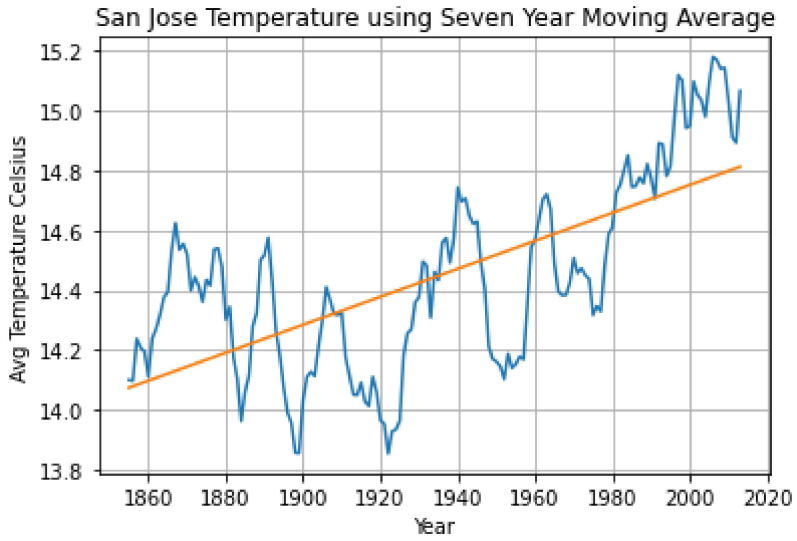
In [31]:
```
# Data for plotting
```

```
sanjose_s = calculate_7_year_moving_avg(SanJoseDF['avg_temp'])
sanjose_t = SanJoseDF['year'][6:]

fig, ax = plt.subplots()
ax.plot(sanjose_t, sanjose_s)
m, b = np.polyfit(sanjose_t, sanjose_s, 1)
plt.plot(sanjose_t, m*sanjose_t + b)

ax.set(xlabel='Year', ylabel='Avg Temperature Celsius',
       title='San Jose Temperature using Seven Year Moving Average')
ax.grid()

fig.savefig("test.png")
plt.show()
```

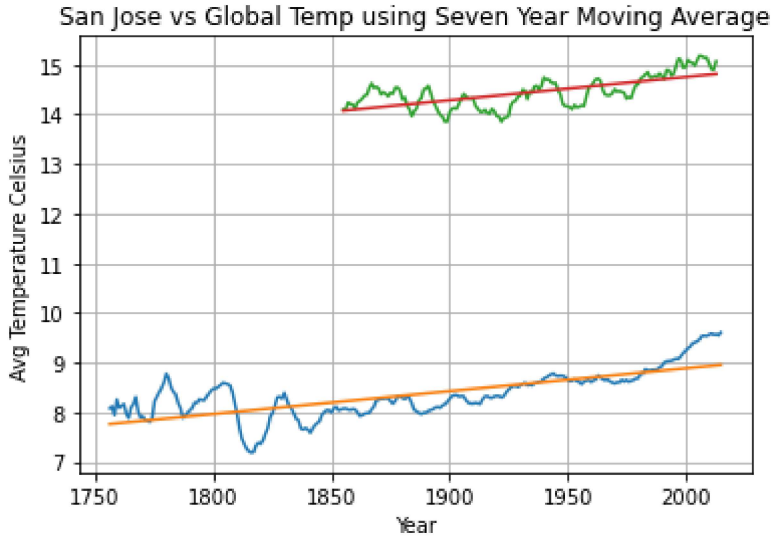


In [32]:
```
# Two plots together with red being San Jose and blue being Global
```

In [33]:
```
fig, ax = plt.subplots()
ax.plot(t, s)
m, b = np.polyfit(t, s, 1)
plt.plot(t, m*t + b)

ax.plot(sanjose_t, sanjose_s)
m, b = np.polyfit(sanjose_t, sanjose_s, 1)
plt.plot(sanjose_t, m*sanjose_t + b)
ax.set(xlabel='Year', ylabel='Avg Temperature Celsius',
       title='San Jose vs Global Temp using Seven Year Moving Average')

ax.grid()
fig.savefig("test.png")
plt.show()
```



Obsevations

- Is your city hotter or cooler on average compared to the global average? Has the difference been consistent over time?
  - For as long as this data has been collected, San Jose is hotter than the global average.
- "How do the changes in your city's temperatures over time compare to the changes in the global average?"
  - Both San Jose and the global average temperature appear to be increasing.
- What does the overall trend look like? Is the world getting hotter or cooler? Has the trend been consistent over the last few hundred years?
  - The overall for both San Jose and planet Earth point to the fact that the we are increasing both local and global temperatures. Historically, the world goes through warming and cooling periods as it ages. It is common to see spikes in temperatures, naturally corrected by freezing periods.
- Does planet Earth observe cyclic patterns of average global temperatures?
  - Based on the previous statement, yes planet Earth is known to go through changes based on numerous factors. The question hints out greenhouse gases, but I believe the more accurate description is an acceleration of an already predefined pattern. Just like physics and radar, the periodicity of the cycle may end up changing per time sample, but not necessarily the extreme points itself.