

Exploring Weather Trends

*In this project, the local (San Francisco City) and global temperature data are analyzed and compared to each other.

1. Data Extraction & Preparation & Description

- Data Extraction (SQL)
 - Download the cvs file of city_list and I found that San Francisco in the list is the city I am living.

```
SELECT *  
From city_list;
```

- Extract the information including year, average temperature, and etc about San Francisco from city_data and export to a cvs file.

```
SELECT *  
FROM city_data  
WHERE city LIKE 'San Francisco';
```

- Download the cvs file of global_data.

```
SELECT *  
From global_data;
```

- Data Preparation (Excel)
In order to produce a smoother output value than the actual data value, the 20-year moving average was calculated by taking the average of the first twenty years of the weather data and applying it to the rest of the years for the global data and the local (San Francisco) data.
- Data Description (Python: pandas)
 - a. global data set
It has total 266 records of average annual temperature from the year of 1750 to 2015. The maximal temperature is 9.49°C in 2015 and the minimum temperature is 7.62 °C in 1826.
After applying the method of moving average, it has 247 values ranging from the year of 1769 to 2015 and is used as the primary variable in the following analysis. The moving average temperature has a mean of 8.33 °C with a standard deviation of 0.39 °C.

```
df1=pd.read_csv(r'C:\Users\Yan\Desktop\udacity\Project 1\CSV\global_data.csv')  
print(df1)  
df1.loc[:, ['avg_temp', 'avg_temp_in_20_years']].describe()
```

	avg_temp	avg_temp_in_20_years
count	266.000000	247.000000
mean	8.369474	8.336518
std	0.584747	0.393123
min	5.780000	7.620000
25%	8.082500	8.085000
50%	8.375000	8.240000
75%	8.707500	8.640000
max	9.830000	9.490000

b. local data (San Francisco)

It has total 165 records of average annual temperature from the year of 1849 to 2013. The maximal temperature is 15.07 °C in 2013 and the minimum temperature is 14.07 °C in 1924 and 1925.

After applying the method of moving average, it has 146 values ranging from the year of 1849 to 2013 and will be used as the primary variable to analyze. The moving average temperature has a mean of 14.43 °C with a standard deviation of 0.26 °C.

```
df2=pd.read_csv(r'C:\Users\Yan\Desktop\udacity\Project 1\CSV\san_francisco.csv')
print(df2)
df2.loc[:, ['avg_temp', 'avg_temp_in_20_years']].describe()
```

	avg_temp	avg_temp_in_20_years
count	165.000000	146.0000
mean	14.450788	14.4250
std	0.504885	0.2558
min	13.220000	14.0700
25%	14.120000	14.2125
50%	14.390000	14.3800
75%	14.760000	14.5100
max	16.230000	15.0700

2. Visual Exploration and Analysis (Python: pandas, numpy, matplotlib.pyplot)

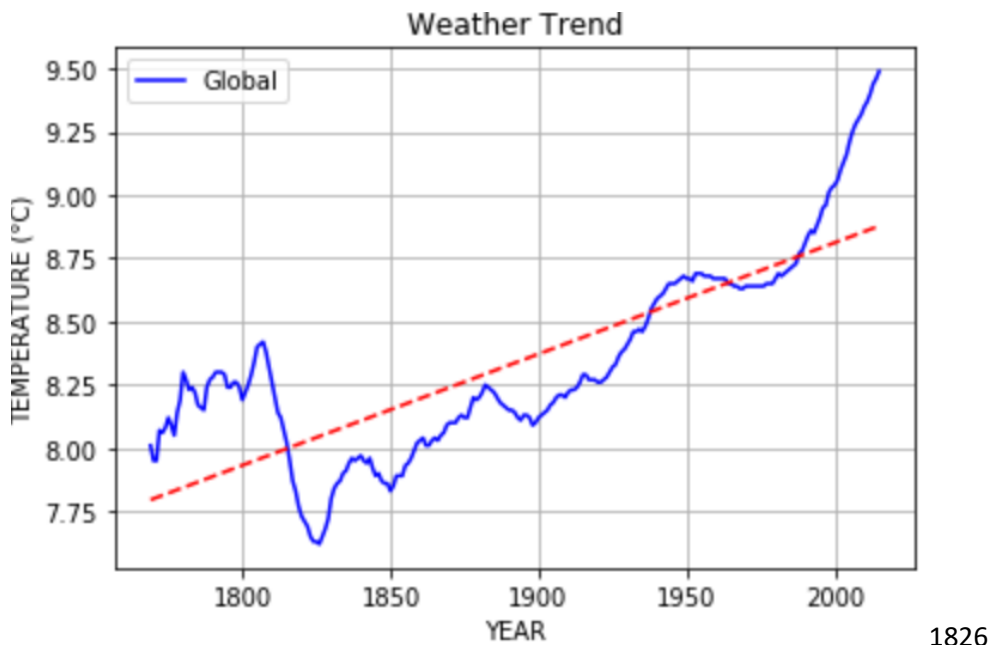
- Change of the global temperature (1750 - 2015)

- The below graph shows that the global temperature varies from 7.62 °C to 9.49 °C (1750 – 2015) and the temperature trend is increasing with years.
- The temperature curve has a steep drop between 1810 and 1826, and then increases with a fast growth after 1985.

```
import pandas as pd
import numpy as np
import matplotlib.pyplot as plt
import csv
year_of_x1=[]
avg_temp_y1=[]

with open(r'C:\Users\Yan\Desktop\udacity\Project 1\CSV\global_data.csv') as global_data:
    plots = csv.reader(global_data,delimiter=',')
    next(plots) # skip header
    for row in plots:
        if row[0].strip() and row[2].strip(): # skip missing values
            year_of_x1.append(int(row[0].strip()))
            avg_temp_y1.append(float(row[2].strip()))

plt.plot(year_of_x1,avg_temp_y1, label='Global', color="blue")
plt.grid(True)
plt.xlabel('YEAR')
plt.ylabel('TEMPERATURE (°C)')
x1=year_of_x1
y1=avg_temp_y1
z=np.polyfit(x1,y1,1)
p=np.poly1d(z)
plt.plot(x1,p(x1),"r--")
plt.title('Weather Trend')
plt.legend()
plt.show()
```



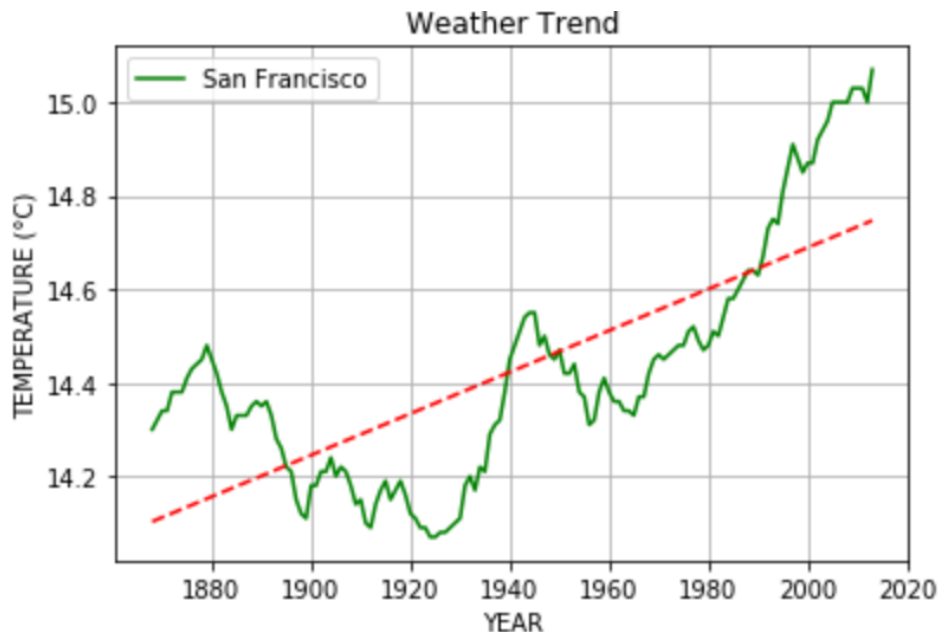
- Change of the San Francisco temperature (1849 - 2013)

- The below graph illustrates that the San Francisco's temperature is pretty constant varying from 14.07 °C and 15.07 °C (1849 – 2013) and the temperature trend is increasing with years.
- The temperature curve has two steep drops in about 1880 and 1945, and a **stable and rapid growth** after about 1960.

```
year_of_x2=[]
avg_temp_y2=[]

with open(r'C:\Users\Yan\Desktop\udacity\Project 1\CSV\san_francisco.csv') as san_francisco:
    plots = csv.reader(san_francisco, delimiter=',')
    next(plots) # skip header
    for row in plots:
        if row[0].strip() and row[4].strip(): # skip missing values
            year_of_x2.append(int(row[0].strip()))
            avg_temp_y2.append(float(row[4].strip()))

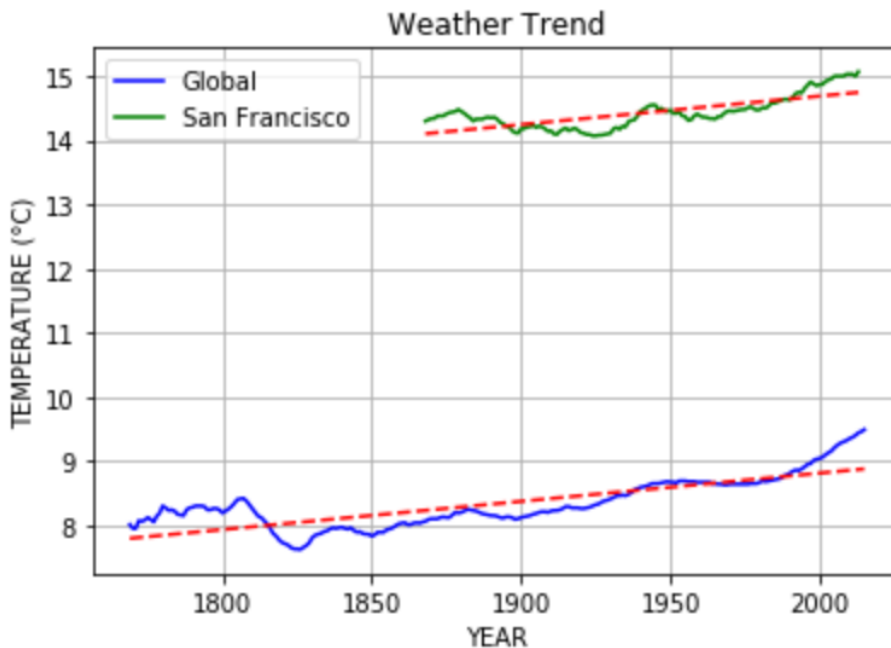
plt.plot(year_of_x2, avg_temp_y2, label='San Francisco', color="green", ms=2)
plt.grid(True)
plt.xlabel('YEAR')
plt.ylabel('TEMPERATURE (°C)')
x2=year_of_x2
y2=avg_temp_y2
z=np.polyfit(x2,y2,1)
p=np.poly1d(z)
plt.plot(x2,p(x2),"r--")
plt.title('Weather Trend')
plt.legend()
plt.show()
```



- Comparison of the weather trend between the global and the San Francisco

- The weather trending for both the global and San Francisco are growing with years and it seems that the two trending lines have pretty similar slope.
- The moving average temperatures for the two data have a swift growth after the late 20th century.
- It is much warmer in San Francisco with a mean of moving average temperature at 14.43 °C than that of the global temperature at 8.33 °C.
- Compared to the global weather, San Francisco has a stable average temperature. The fluctuations of temperature (1849 – 2013) is about 1 °C change vs. 2 °C for that of the global.

```
plt.plot(year_of_x1, avg_temp_y1, label='Global', color='blue')
plt.plot(year_of_x2, avg_temp_y2, label='San Francisco', color='green')
z1=np.polyfit(x1,y1,1)
p=np.poly1d(z1)
plt.plot(x1,p(x1), "r--")
z2=np.polyfit(x2,y2,1)
p=np.poly1d(z2)
plt.plot(x2,p(x2), "r--")
plt.grid(True)
plt.xlabel('YEAR')
plt.ylabel('TEMPERATURE (°C)')
plt.title('Weather Trend')
plt.legend()
plt.show()
```



Conclusion

Basing on the current analysis, whether the global or the local temperature is rising in the long term, which implies the climate change and global warming.