Project 1: Explore Weather Trends

1. Extra data from a database using SQL

(1) Extra local temperature data

```
SELECT *
FROM city_data
WHERE country LIKE ('Singapore') and avg_temp IS NOT NULL
ORDER BY year
```

(2) Extra my hometown temperature data

```
SELECT *
FROM city_data
WHERE city LIKE ('Fuzhou') and avg_temp IS NOT NULL
ORDER BY year
```

(3) Extra global temperature data

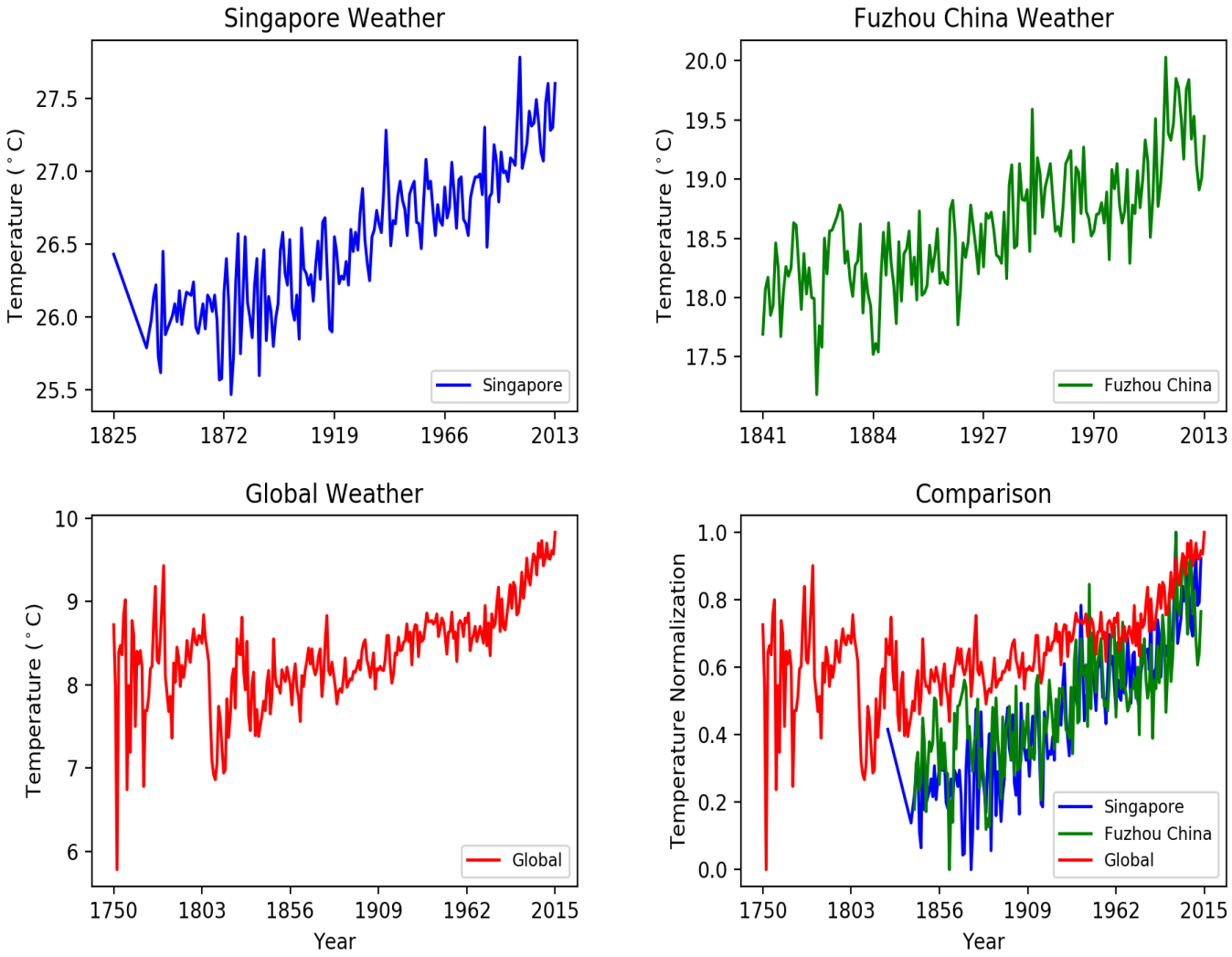
```
SELECT *
FROM global_data
```

2. Create a clear data visualization using Python

```
from pylab import plot, show
import pandas as pd
import numpy as np
import matplotlib.pyplot as plt
import matplotlib.gridspec as gridspec
# Read data from csv files
df1 = pd.read_csv('city_data_sg.csv')
df2 = pd.read_csv('global_data.csv')
df3 = pd.read_csv('city_data_fuzhou.csv')
year_sg, avg_temp_sg = df1['year'], df1['avg_temp']
nor1=(avg_temp_sg-min(avg_temp_sg))/(max(avg_temp_sg)-min(avg_temp_sg))
year_fz, avg_temp_fz = df3['year'], df3['avg_temp']
nor2=(avg_temp_fz-min(avg_temp_fz))/(max(avg_temp_fz)-min(avg_temp_fz))
year_glo, avg_temp_glo = df2['year'], df2['avg_temp']
nor3=(avg_temp_glo-min(avg_temp_glo))/(max(avg_temp_glo)-min(avg_temp_glo))
# Plot data
fig, axs = plt.subplots(2, 2)
axs[0, 0].plot(year_sg,avg_temp_sg,'blue',label='Singapore')
axs[0, 0].set_title('Singapore Weather')
axs[0, 0].set_xticks(np.linspace(min(year_sg),max(year_sg),5, endpoint=True))
axs[0, 0].legend(loc=4, fontsize='small')
axs[0, 0].set_ylabel('Temperature ($^\circ$C)')
axs[0, 1].plot(year_fz,avg_temp_fz,'green',label='Fuzhou China')
axs[0, 1].set_title('Fuzhou China Weather')
axs[0, 1].legend(loc=4, fontsize='small')
axs[0, 1].set_ylabel('Temperature ($^\circ$C)')
axs[0, 1].set_xticks(np.linspace(min(year_fz),max(year_fz),5, endpoint=True))
axs[1, 0].plot(year_glo,avg_temp_glo,'red',label='Global')
axs[1, 0].set_title('Global Weather')
axs[1, 0].legend(loc=4, fontsize='small')
axs[1, 0].set_xticks(np.linspace(min(year_glo),max(year_glo),6, endpoint=True))
axs[1, 0].set_xlabel('Year')
axs[1, 0].set_ylabel('Temperature ($^\circ$C)')
axs[1, 1].set_title('Compare')
axs[1, 1].legend(loc=4, fontsize='small')
axs[1, 1].set_xticks(np.linspace(min(year_glo),max(year_glo),6, endpoint=True))
axs[1, 1].set_xlabel('Year')
axs[1, 1].set_ylabel('Temperature Normalization')
axs[1, 1].plot(year_sg,nor1,'blue',label='Singapore')
axs[1, 1].plot(year_fz,nor2,'green',label='Fuzhou China')
axs[1, 1].plot(year_glo,nor3,'red',label='Global')
plt.subplot_tool()
plt.legend(loc=4, fontsize='small')
plt.show()
```

3. Interpret a data visualization

I use python to plot the data where I extracted using SQL. And the results are shown below:



In Figure1, it illustrates the change in Singapore temperature relative to 1825-2013 average temperatures. And the temperature range is around **25.5-27.5°C**, which is a bit higher than Fuzhou city where the temperature is around **17.5-20°C** in figure2, and it is much higher than the global average temperature if we look at the third figure where the range is only around **6-10°C**.

Since the temperature range of Singapore, Fuzhou city and global is different, and in order to better compare three of them easier. I do a normalization calculation. And the result is shown on figure4. We can see that they all show a waring trend neither in Singapore, Fuzhou city nor in global cities. The temperature keeps increasing, which point to the conclusion that ***our planet is warming***.