Exploring Weather Trends

Xueyun Zhou

1 An Outline

In this project, SQL and Excel were used. SQL was used for extracting the data, while Excel was used for doing calculation and creating charts.

The commands for extracting the data by SQL were as follows:

```
SELECT *
FROM city_list;
SELECT *
FROM city_data
WHERE city = 'Hangzhou';
SELECT *
FROM global_data;
```

I calculated the moving average in Excel by the formula "=AVERAGE()".

My key considerations when deciding how to visualize the trends were basically as follows:

1) How to compare two sets of data after the data were downloaded as two files?
2) How to deal with the missing data between the two sets of data?
3) How many years could be taken to calculate the average?

2 Line Chart

It would be easier to compare the two sets of data in a tab. So I consolidated the data by the formula VLOOKUP as below.
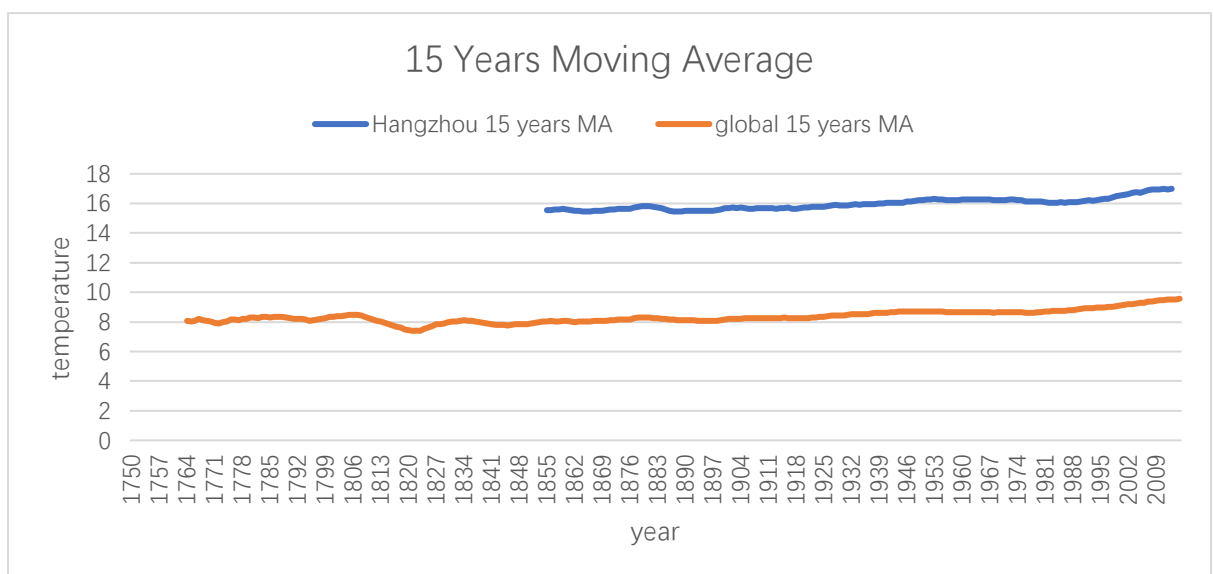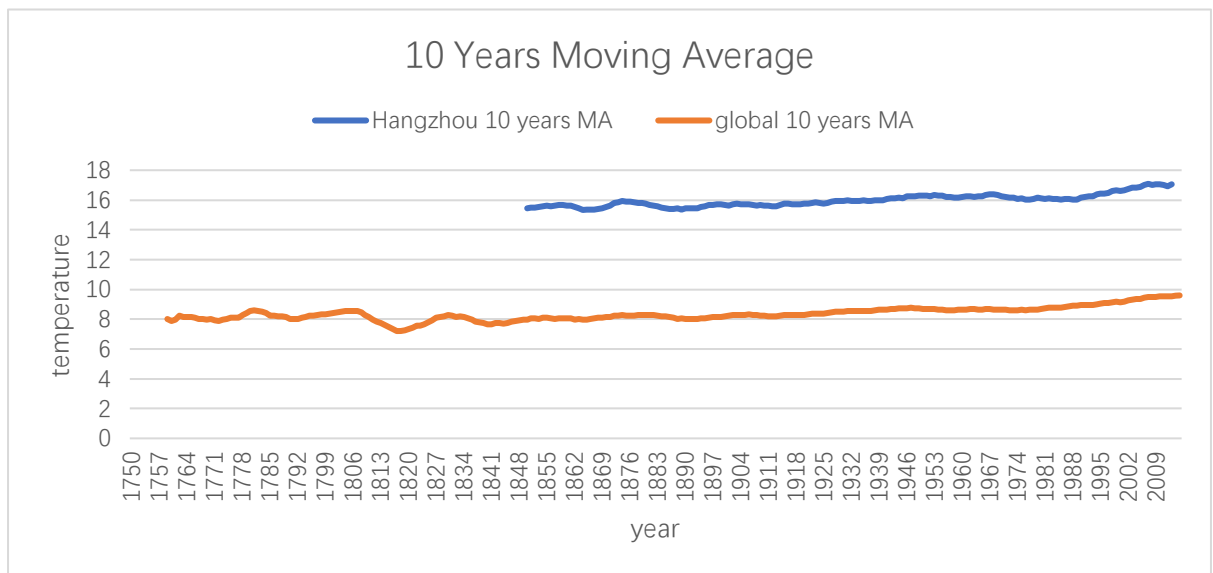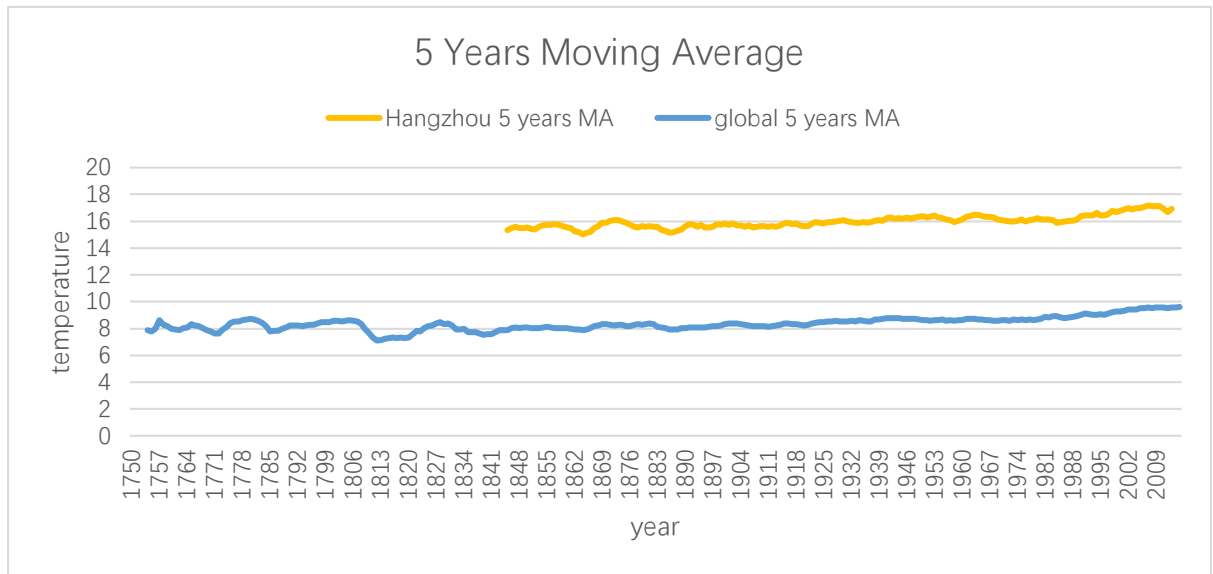
B93 | =VLOOKUP(Consolidated!A93,Hangzhou!$A$2:$D$174,4,0)

	A year	B Hangzhou yearly	C global yearly	D Hangzhou 5 years MA	E global 5 years MA	F Hangzhou 10 years MA	G global 10 years MA	H Hangzhou 15 years MA
91 1839		=VLOOKUP(A91,global!$A		=AVERAGE(C87:C91)		=AVERAGE(C82:C91)	
92 1840		=VLOOKUP(A92,global!$A		=AVERAGE(C88:C92)		=AVERAGE(C83:C92)	
93 1841	=VLOOKUP(Consolidated!	=VLOOKUP(A93,global!$A		=AVERAGE(C89:C93)		=AVERAGE(C84:C93)	
94 1842	=VLOOKUP(Consolidated!/	=VLOOKUP(A94,global!$A		=AVERAGE(C90:C94)		=AVERAGE(C85:C94)	
95 1843	=VLOOKUP(Consolidated!/	=VLOOKUP(A95,global!$A		=AVERAGE(C91:C95)		=AVERAGE(C86:C95)	
96 1844	=VLOOKUP(Consolidated!/	=VLOOKUP(A96,global!$A		=AVERAGE(C92:C96)		=AVERAGE(C87:C96)	
97 1845	=VLOOKUP(Consolidated!/	=VLOOKUP(A97,global!$A	=AVERAGE(B93:B97)	=AVERAGE(C93:C97)		=AVERAGE(C88:C97)	
98 1846	=VLOOKUP(Consolidated!/	=VLOOKUP(A98,global!$A	=AVERAGE(B94:B98)	=AVERAGE(C94:C98)		=AVERAGE(C89:C98)	
99 1847	=VLOOKUP(Consolidated!/	=VLOOKUP(A99,global!$A	=AVERAGE(B95:B99)	=AVERAGE(C95:C99)		=AVERAGE(C90:C99)	
100 1848	=VLOOKUP(Consolidated!/	=VLOOKUP(A100,global!$/	=AVERAGE(B96:B100)	=AVERAGE(C96:C100)		=AVERAGE(C91:C100)	
101 1849	=VLOOKUP(Consolidated!/	=VLOOKUP(A101,global!$/	=AVERAGE(B97:B101)	=AVERAGE(C97:C101)		=AVERAGE(C92:C101)	
102 1850	=VLOOKUP(Consolidated!/	=VLOOKUP(A102,global!$/	=AVERAGE(B98:B102)	=AVERAGE(C98:C102)	=AVERAGE(B93:B102)	=AVERAGE(C93:C102)	
103 1851	=VLOOKUP(Consolidated!/	=VLOOKUP(A103,global!$/	=AVERAGE(B99:B103)	=AVERAGE(C99:C103)	=AVERAGE(B94:B103)	=AVERAGE(C94:C103)	
104 1852	=VLOOKUP(Consolidated!/	=VLOOKUP(A104,global!$/	=AVERAGE(B100:B104)	=AVERAGE(C100:C104)	=AVERAGE(B95:B104)	=AVERAGE(C95:C104)	
105 1853	=VLOOKUP(Consolidated!/	=VLOOKUP(A105,global!$/	=AVERAGE(B101:B105)	=AVERAGE(C101:C105)	=AVERAGE(B96:B105)	=AVERAGE(C96:C105)	
106 1854	=VLOOKUP(Consolidated!/	=VLOOKUP(A106,global!$/	=AVERAGE(B102:B106)	=AVERAGE(C102:C106)	=AVERAGE(B97:B106)	=AVERAGE(C97:C106)	
107 1855	=VLOOKUP(Consolidated!/	=VLOOKUP(A107,global!$/	=AVERAGE(B103:B107)	=AVERAGE(C103:C107)	=AVERAGE(B98:B107)	=AVERAGE(C98:C107)	=AVERAGE(B93:B107)
108 1856	=VLOOKUP(Consolidated!/	=VLOOKUP(A108,global!$/	=AVERAGE(B104:B108)	=AVERAGE(C104:C108)	=AVERAGE(B99:B108)	=AVERAGE(C99:C108)	=AVERAGE(B94:B108)
109 1857	=VLOOKUP(Consolidated!/	=VLOOKUP(A109,global!$/	=AVERAGE(B105:B109)	=AVERAGE(C105:C109)	=AVERAGE(B100:B109)	=AVERAGE(C100:C109)	=AVERAGE(B95:B109)
110 1858	=VLOOKUP(Consolidated!/	=VLOOKUP(A110,global!$/	=AVERAGE(B106:B110)	=AVERAGE(C106:C110)	=AVERAGE(B101:B110)	=AVERAGE(C101:C110)	=AVERAGE(B96:B110)
111 1859	=VLOOKUP(Consolidated!/	=VLOOKUP(A111,global!$/	=AVERAGE(B107:B111)	=AVERAGE(C107:C111)	=AVERAGE(B102:B111)	=AVERAGE(C102:C111)	=AVERAGE(B97:B111)
112 1860	=VLOOKUP(Consolidated!/	=VLOOKUP(A112,global!$/	=AVERAGE(B108:B112)	=AVERAGE(C108:C112)	=AVERAGE(B103:B112)	=AVERAGE(C103:C112)	=AVERAGE(B98:B112)
113 1861	=VLOOKUP(Consolidated!/	=VLOOKUP(A113,global!$/	=AVERAGE(B109:B113)	=AVERAGE(C109:C113)	=AVERAGE(B104:B113)	=AVERAGE(C104:C113)	=AVERAGE(B99:B113)

Hangzhou | global | **Consolidated** | +

Firstly, I plotted the yearly temperatures to see the general trends and decide how many years could be taken to calculate moving averages. For the missing data, I plotted them as gaps. As the missing data only showed up prior and after the time period of the Hangzhou data, the visualization was not influenced.

The yearly average line chart is as below.
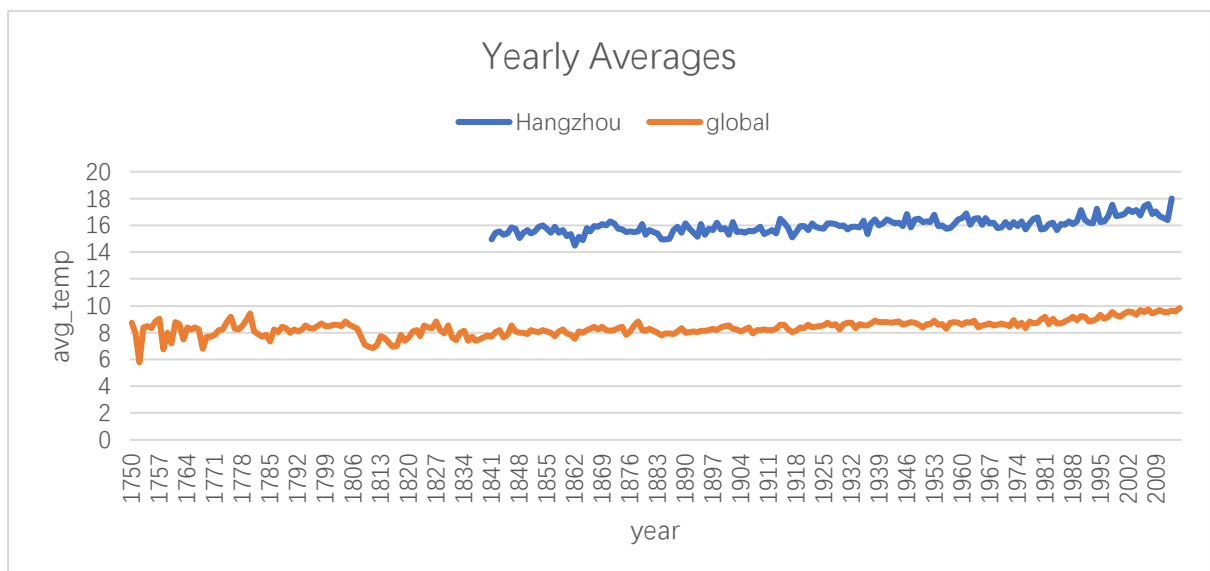
From the chart it can be seen the yearly average temperatures for both the city level data and the global data fluctuated wildly. To smooth the fluctuation and make trends more observable, 5 years, 10 years and 15 years were taken to calculate the average. The charts for each scenario are as follows.

5 Years Moving Average

— Hangzhou 5 years MA   — global 5 years MA

10 Years Moving Average

— Hangzhou 10 years MA   — global 10 years MA

15 Years Moving Average

— Hangzhou 15 years MA   — global 15 years MA

3 Four Observations

1) Overall, the temperature of the city I selected – Hangzhou was all time hotter on the yearly average than the global average. This difference has been consistent over time from 1841 to 2013.

2) From the 10 Years Moving Averages chart, it can be clearly seen that the global average fluctuated before around 1850, and steadily increased since then. Similarly, the Hangzhou average temperature fluctuated before around 1990, and steadily increased since then. There was a downward or stable trend during around 1965 to 1990 for both data. Afterwards, the average temperatures for both the city level and the global went up quickly.

3) From an overall view, the world was getting hotter. The trend has not been consistent over the 19th century and the 20th century, while in the 18th the temperatures moved up and down drastically.

4) The upward trend was more evident for the Hangzhou data than the global average since 1990. It can be supposed that the rapid urbanization and industrialization in Hangzhou since 1990 accelerated the rise of the temperature.