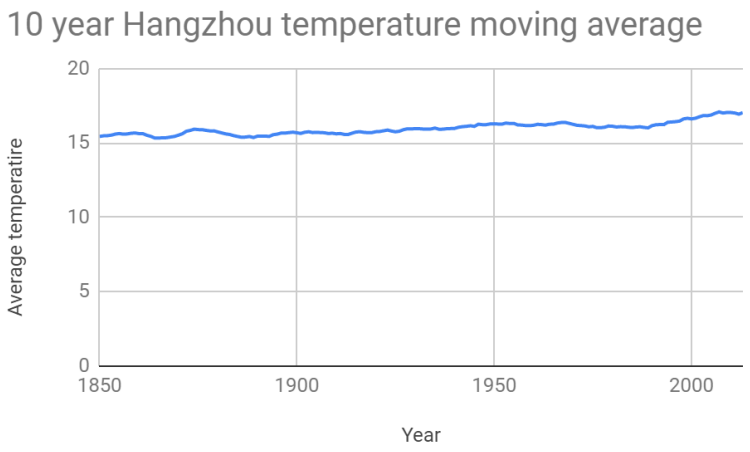
Data Preparation

- Tools used for each step:
    - Extract the data: SQL is used with queries to extract needed data.
        - SQL query for city_data:
          *SELECT * FROM city_data WHERE city = 'Hangzhou'*
        - SQL query for global_data:
          *SELECT * FROM global_data*
    - CSV Data storage and data analysis: Google Sheets is used.
- Calculate the moving average:
    - Both 5 year and 10 year moving averages are used to analyze the data
- Data visualization:
    - First plot the raw data of average temperature by the year, we can observe some clear fluctuations and maybe the overall trend as well.
    - To smooth the data, by the distance between fluctuations we can assume that 5 or 10 years moving average can be good intervals.
    - Plot 5 year and 10 year moving averages for both local and global temperatures, use *CORREL()* in google sheet to calculate the correlation coefficient.
    - check similarities and differences in trends.

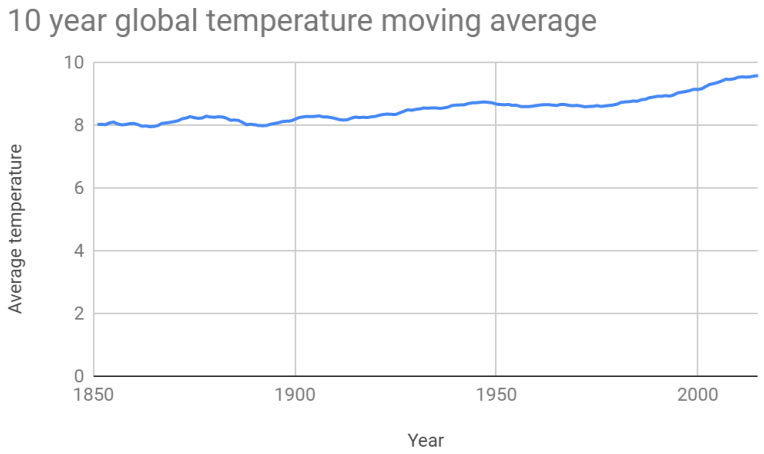
Results Graphs

City moving average:



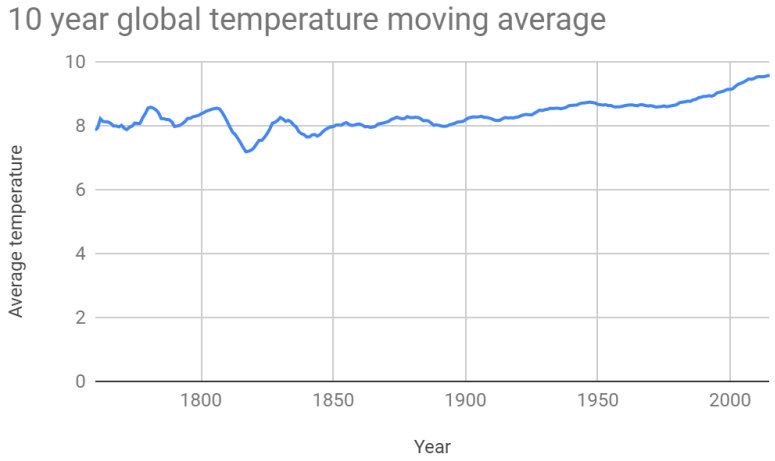
correlation coefficient: 0.879

Global moving average:



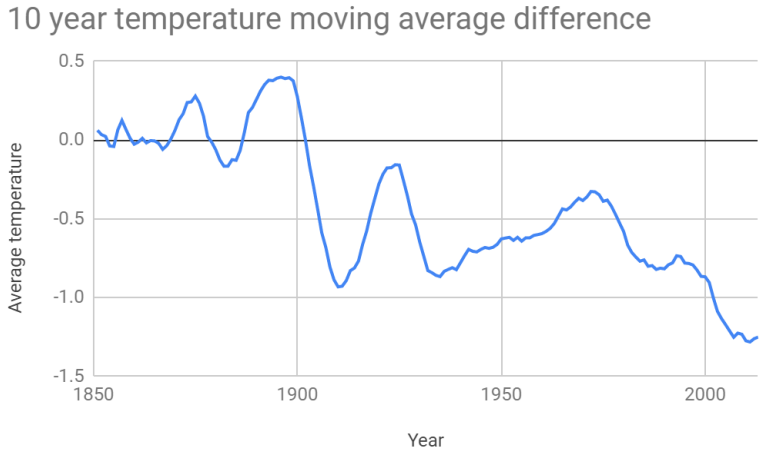
correlation coefficient: 0.926

Global moving average from year 1750:



correlation coefficient: 0.751

Difference in temperature moving averages between Hangzhou and global:



correlation coefficient: -0.795

Results observations:

- The trends of both global and Hangzhou's temperature are getting hotter, the temperature increasing speed of global is faster than Hangzhou according to the positive correlation coefficients.
- The temperature increase speeds of both global and Hangzhou are getting faster around 1990 because curves are getting steeper.
- Compared to the global temperature, the increase of temperature of Hangzhou is slower, this can be seen both from the curves and the correlation coefficients.
- The getting hotter overall trend is not consistent over the last 200 years, as we can see from the two graphs of global moving averages, there are fluctuations on the same horizontal line of 8 degree before 1900, after that the temperature is going overall higher. This can also be seen from the higher correlation coefficient after 1850.
- It is difficult to estimate the average temperature in Hangzhou based on the average global temperature, because differences in the moving averages are quite inconsistent, although the overall trend is smaller. More research or other investigations need to be done to make the estimation.