

# Exploring Weather Trends

## Extract the data

Selecting the global data is easy:

```
SELECT * FROM global_data;
```

Find the exact spelling of my neighboring city Xi'an and extract its yearly temperatures:

```
SELECT * FROM city_list  
WHERE country LIKE 'China';
```

```
SELECT * FROM city_data  
WHERE city = 'Xian';
```

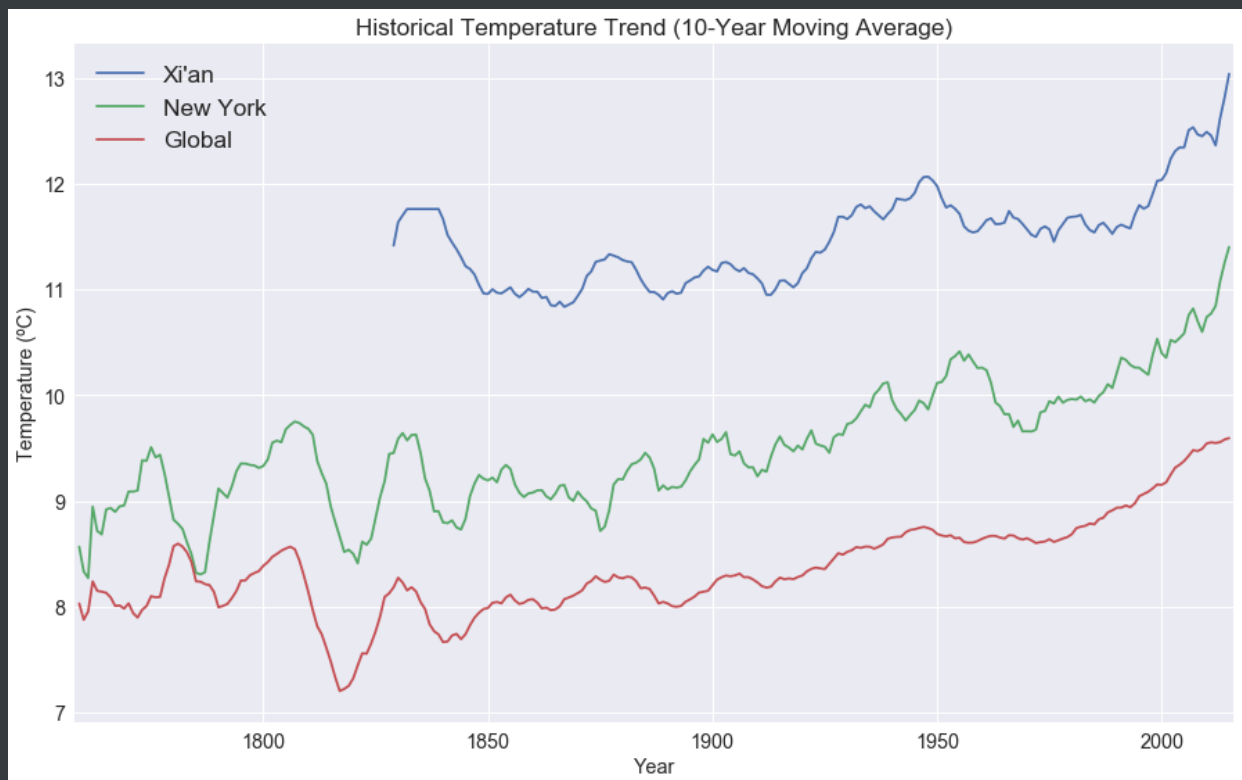
The same for New York as comparison:

```
SELECT * FROM city_list  
WHERE city LIKE '%York%';
```

```
SELECT * FROM city_data  
WHERE city = 'New York';
```

## Open up the CSV and Create a line chart

After importing and cleaning the data a bit in `DAND-P0.ipynb` with pandas, I've created the following line chart:



## Make observations

- All three lines fluctuate more or less in the same steps (after removing potential outliers), with Xi'an warmer than New York, which in turn is warmer than global averages.
- Since 1850, all three lines have been in a general rising trend.
- Temperatures of Xi'an and of the whole world have a Pearson correlation coefficient of 0.78, which means they are strongly correlated.
- Since temperatures of Xi'an are 2.96 °C higher than global averages on average, we could probably predict the former by add 2.96 °C to the latter.